CD-ROM INCLUDED

NAB Broadcast Engineering Conference Proceedings 2007

Conferences: April 14-19, 2007 Exhibits: April 16-19, 2007
Las Vegas Convention Center, Las Vegas, Nevada USA

NAB

THE WORLD'S LARGEST ELECTRONIC MEDIA SHOW

www.nabshow.com

A recap of the
61st Broadcast Engineering
Conference at NAB2007

National Association of
NAB®
BROADCASTERS

2007 NAB BROADCAST ENGINEERING CONFERENCE COMMITTEE

Mr. Lewis Zager (Chairman)
Director
PBS DTV Strategic Services Group
2100 Crystal Drive
Arlington, VA 22314

Mr. Talmage Ball
VP, Engineering
Bonneville International
55 North 300 West
Salt Lake City, UT 84180

Mr. Dave Converse
VP & Director of Engineering
ABC/Disney Television Stations Group
500 S. Buena Vista Street
Burbank, CA 91521

Mr. Steve Fluker
Director of Engineering
Cox Radio, Orlando
4192 N. John Young Parkway
Orlando, FL 32825

Mr. Al Grossniklaus
Director of Engineering & Operations
WTHR NBC 13
1000 North Moridian Street
Indianapolis, IND 46204

Mr. Gary Kline
Vice President of Engineering & IT
Cumulus Media
3535 Piedmont Road
Building 14 - Suite 1400
Atlanta, GA 30339

Mr. Wayne Kube
Technology Manager, Broadcast Media
Belo Corporation
400 S. Record Street
Dallas, TX 75202

Mr. John Merrill
Director of Engineering
CBS 5 KPHO-TV
4016 N Black Canyon Highway
Phoenix, AZ 85017

Mr. Paul Shulins
Director of Technical Operations
Greater Media, Boston
55 Morrissey Blvd.
Boston, MA 02125

Mr. Jeffrey Smith CEA, CBNT
President
JRS Broadcast Engineering
11 Hampshire Pl
Monroe Twp., NJ 00831

SBE Representatives:
Mr. John L. Poray
Executive Director, SBE
9102 North Meridian Street, Suite 150
Indianapolis, IN 46260

Mr. Frederick Baumgartner, CPBE
Dir Broadcast Engineering
Qualcomm MediaFLO
E-148 D, 5775 Morehouse Dr
San Diego, Ca 92121-1714

TABLE OF CONTENTS

NAB Broadcast Engineering Conference Opening Keynote

Sunday, April 15, 2007
9:00 AM – 9:30 AM

Chairperson: Lynn Claudy
NAB, Washington, DC

***Destination in Broadcast Technology -- from Mobile to Ultra-HDTV**
Hirokazu Nishiyama, Managing Director, NHK, Tokyo, Japan

*Paper not available at the time of publication

Multicasting for Radio

Sunday, April 15, 2007
9:30 AM – 12:00 PM

Chairperson: Jeffrey Smith
JRS Broadcast Engineering, Monroe Twp, NJ

***HD Radio Advanced Services - A Market Reality**
Joseph D'Angelo, iBiquity Digital, Warren, NJ

***Implications of Advanced HD Radio Services on Station Workflow**
Raymond Miklius, Broadcast Electronics, Inc., Quincy, IL

Considerations for Automation & Playout in Multi-Stream Applications
Eugene Novacek, ENCO, Southfield, MI

HD Radio® Conditional Access: What It Is, How It Fits in the Broadcast Station; and Why It Can Provide Outstanding Return on Investment
Thomas Rucktenwald, NDS, Costa Mesa, CA

***Radio Engineering - Managing Multiple Formats Made Easy**
Richard Darr, RCS, White Plains, NY

---

*Paper not available at the time of publication

Considerations for Automation and Playout in Multi-stream Applications

Eugene Novacek P.E.
David Turner
Neil S. Price

ENCO Systems Inc.
Southfield, Michigan

Background and History of PC based Station Automation

Even with the accepted use of digital audio in master recording studios and the adoption of the digital CD, broadcast automation began as a largely electro-mechanical analogue process in the 80's. Early Station Automation sometimes referred to a large reel of tape playing on an auto-reversing tape deck. This automation arrangement could provide most of a day's programming with minimal operator intervention, but the audio quality was mediocre due to the slow tape speed, and it required production personnel to invest a great deal of time assembling the contents of the tapes. Once the tapes were assembled, it was very difficult to make any changes, especially if the tape was playing on-air. A more flexible arrangement used a stack of cart machines with the secondary cue (SEC) output of one wired to the start input of the next, and so on with the last machine connected back to the first. This allowed the operator to load several songs and commercials to play automatically in sequence, freeing up short amounts of time for him (or her) to do other important operations. This crude form of automation was followed by numerous variations of mechanical beasts designed to hold and cycle dozens and sometimes hundreds of carts through multiple playback decks. These electro-mechanical marvels were capable of automatically sequencing several days worth of programming. These machines were incredible achievements in station automation and faithfully served the broadcast industry for many years. However, their dependency on extensive mechanical transports made them high maintenance devices and limited their flexibility. Even the audio quality was hard to maintain due to the number of capstans and tape heads involved. Broadcasters needed newer, higher quality, easier to maintain, more versatile automation systems. As CD players and DAT tape machines became common place in the home and in the studio, the audience demand for CD Quality broadcast audio also grew. Along with the improved audio this new generation of equipment brought improved control features that allowed broadcasters to elevate automation to a new level. In addition to the standard START and STOP functions, a much more complete set of instructions including shuttling and indexing were now available, usually through a serial data protocol. This enabled a single electronic controller to "talk" to multiple devices, directing them to cue up then play individual tracks from CDs and tapes containing multiple tracks. With CD Jukebox devices available that could hold and play tracks from over 300 CDs and DAT tape machines that could cue and play hours worth of programming, a station could now automate for days at a time, if the controller was programmed properly. While several manufacturers developed automation systems for their own equipment, these were mostly proprietary hardware and not able to communicate with devices from other manufacturers. These units often employed special keypads for data entry and ran ROM based programs that were difficult to upgrade or modify. This is where the Personal Computer made its biggest mark on station automation. PCs were already being used in business and at home to run database applications, perfect for storing a log of scheduled events. PCs had excellent serial (and parallel) communications capability and could be configured/programmed to "speak" any protocol required. PCs used a standard user interface (keyboard and monitor). Programs were stored on disk and ran from RAM and were therefore relatively simple to modify and reload. The Personal Computer was an excellent platform for station automation controllers. A number of manufacturers created custom software packages using standard PC hardware to automate CD jukeboxes, DAT players and fire relays to play standard cart machines. These were quite effective and could provide a good level of walk away automation, but some of the stations most important material, e.g. commercials, IDs, and promos were left to the weakest link - the cart machine. It became increasingly obvious that a more sophisticated "CD quality" version of the cart machine was needed to play spots, IDs, jingles etc. As digital cart machines began appearing as direct hardware replacements for the existing analog units, some clever computer people were realizing that everything needed for a digital cart machine was already in a PC except the Analog to Digital and Digital to Analog conversion electronics. And this could easily be added by building a custom signal processing circuit board designed to plug into one of the existing PC expansion slots.

Since PCs were already driving the development of the hard disk storage technology, they had access to the capacity required to store many hours of digital audio. This could be divided into any number of any size pieces, allowing an extensive inventory; and unlike tape or floppy disk based systems, PC based systems could provide random access to any and all of the material they contained. Adding these powerful audio capabilities to the automation strengths of the PC has created the perfect platform for the continued evolution of station automation.

Low cost PC computing, coupled with high quality audio processing cards, created the opportunity for the typical radio broadcaster to fully automate and the market for PC based radio automation took off. Throughout the 90's radio stations found that they could fully automate most studio operations. The concept of a programmed playlist including all commercial content, controlled to precise timing opened the door to totally unmanned operation, ideal for overnight shifts and periods where minimal support or on-air personnel are available. Consolidation in the broadcast industry was driving need for minimizing the cost of operations and reducing staff requirements for on-air operation. While we, in a "new" broadcast automation industry, seized upon an opportunity to fully automate cumbersome electro-mechanical systems and brought huge manpower savings to the broadcaster, needs quickly broadened to include entirely new concepts that were not envisioned in the days of the CART machine. The broadcast industry was moving away from the automation of a basic analog process into that of a data based, IT driven structure. New clusters of stations wanted to share libraries and processes. Commercial traffic and scheduling systems needed to be integrated and on-air talent needed to operate with minimal engineering support. The pace of change continues to accelerate for the broadcaster and multicasting associated with HD Radio™ is but another challenge which will undoubtedly open the door for others. From an automation and playout developers' perspective, we must address three areas outside the basic functions within our systems. First, we must deal with and create data that describes audio content for other systems (metadata); second, we must build-in open connectivity to other systems using standard protocols as needed; and thirdly, we must build in the flexibility to allow our users to rapidly meet as yet undefined new requirements as they develop. These are the areas that set professional grade broadcast playout and automation tools apart from the plethora of media player type applications that now abound.

New Data Requirements Emerging

In order to facilitate many automation requirements, new information that describes what is in the audio file or program is required. The association of this "metadata" with audio files is generally referred to as "tagging". As various types of rich media moved to the digital world, Microsoft developed a standard called RIFF for Resource Interchange File Format. This included a representation for audio files called wave files denoted with the ".wav" extension. The wave format supports a number of audio formats, compressed and uncompressed. The European Broadcast Union extended the definition to better enable file exchange between production systems in the world of broadcasting and music production. Not surprisingly, this became known as the Broadcast Wave Format and included metadata specific to the needs of audio production systems. The audio formats are limited to linear PCM and MPEG coded compressed audio. Inspired by the goals of BWF, we were pleased to help author a specification targeted at the radio broadcaster called Cart Chunk. In 2002 Cart Chunk format was adopted as AES standard AES46-2002. Cart Chunk defined important metadata used by on-air applications such as automation and scheduling systems.

Over on the world wide web, use of rich media and audio was also skyrocketing. Users of the internet made use of highly compressed audio files, typically MP3, which also required metadata so that browsers could interpret and visually display information about the content. A simple tagging format called ID3v1 became the de facto standard for representing basic artist and title information in an MP3 file. This has since been expanded through several iterations. The current IBOC HD Radio™ specification calls for metadata (described as MPSD for Main Program Service Data and SPSD for Supplemental Program Service Data) to be formatted based on ID3v2.3.0.
The general design of all of these schemes is to provide an adaptable framework that can be changed, or added to, without affecting earlier applications. For example, an application reading a tagged file could interpret those "frames" of metadata associated with that application but ignore others, whether new or unused. The automation and playout system must adequately represent and store appropriate metadata to allow its use as needed within the application, as well as maintain it for use in downstream or ancillary applications. We must also have the ability to move between applications that may represent the same metadata in varying formats such as between Cart Chunk and the IBOC specification. The transfer of this information is addressed through standardized hardware and software tools that define the connectivity between systems.

The Need for Connectivity

An automation system rarely contains everything necessary for complete walk-away operation. For it to be truly effective, a system needs to be able to interface with various other broadcast devices, like audio switchers, audio consoles, satellite receivers, CD players, DAT players, station clocks, and more. These interfaces are created in two parts: the hardware interface, which provides proper electronic inter-connection, and the software interface that communicates specific commands to and from the peripheral. Systems based on personal computers can take advantage of a multitude of off the shelf hardware (plug-in I/O boards) that can provide any mix of RS-232, RS-422, RS- 485 and large numbers of opto-isolated contact closure inputs and dry contact outputs. Systems that are able to address these generic I/O products are more flexible than those that offer their own custom I/O hardware that may offer a mix of these features. In addition to these "traditional" communication interfaces, many new consoles and routing switchers feature network interface ports and are able to communicate via IP over standard Ethernet networks. This is a perfect match for the network hardware of a PC platform and can greatly simplify the physical interconnections for systems that are capable of communicating through this channel. This can allow PC based automation systems to control input routing, fader labeling, bus assignment and even automate mixes through a simple network connection. With the hardware connections becoming more standardized and common place, we see that the real strength of an automation system is in the software portion of this interface - how it uses the hardware to communicate with the various devices external to the PC. A wide variation of command syntax and order exists among the various peripheral devices an automation system might connect to. Command strings, string lengths and the use of carriage return and/or line feed vary from device to device. Even simple contact closure controlled devices can often require specific closure sequences to perform a given function. It is therefore very important that the automation system software support the creation and transmission of custom serial protocols, be capable of storing and running control macros that execute specific command sequences and provide these features through a user programmable mechanism so that each station's unique compliment of equipment can be administered and maintained. Many system manufacturers provide software modules (drivers) that perform the data translations necessary to allow the automation system to communicate with other popular devices. In many cases that is a simple and adequate way to provide interoperability, as long as the configuration never changes. But things will change. An automation system that is able to utilize industry standard methods of communication and provides a mechanism to "tweak" communication details is more flexible and able to adapt to future developments.

Data communication abilities of automation systems are now extending beyond peripheral control and are now being used to create a better user experience through ancillary data services. An automation system can send database information about its audio (often referred to as metadata) to external data channels that can provide relevant "Now Playing" information to the end user. Examples of this include RDS encoders feeding conventional FM transmitters to display title, artist and station ID text strings on user's receivers and roadside billboards. HD Radio™ MPSD and SPSD are also examples. And metadata over IP is now allowing broadcasters to tightly integrate their on-air product with their station web sites and streaming services. These services typically use standardized information exchange formats like eXtensible Mark-up Language (XML) which is a powerful and accepted format specification for the exchange of almost any type of data. XML is a data description language that is general-purpose, Internet protocol-friendly, and very easy to learn and author. XML is unlike custom data interchange formats which are generally proprietary, special-purpose, "binary" formats that cannot be easily shared by different software applications or across different computing platforms, much less authored and maintained in common text editors. Using XML, data can be moved between applications very easily, even as applications undergo independent change. At ENCO we have utilized XML extensively to carry playlist information, provide a user customization platform and fully integrate partner music and traffic scheduling applications. XML is also utilized in communicating with HD Radio™ Importers.

Flexibility : Empowering the Broadcast User

Applying computer technology to automate broadcast processes for handling audio, from storage and management through playout and on-air assist have become the standard worldwide. It would be difficult to find a broadcaster today that relies on manual methods to get their audio program to air. The demands are as great as the variety of broadcasters that range from the webcaster operating out of his room, to global networks. For some, this means having the feature set to support a one man operation to easily set up and play without interruption 24/7. For others it's the capabilities to manage multiple programs on multiple stations. And yet others look for on-air assist features to ensure everything that occurs live runs smoothly without error. To address these needs, the broadcast automation industry has continued to add increasing numbers of features supporting specific tasks to their systems. The list is long and continues to grow. However the real underlying requirement to address dynamically changing needs in modern broadcasting is flexibility. Flexibility can reduce a broadcaster's reliance on vendor development cycles in meeting rapidly changing needs, for example a change in a format specification in a new media application or the requirement to repurpose programs for a different venue.

A fundamental characteristic required for broadcaster flexibility is the concept of the "virtual machine". Playout and automation has always been designed to mimic the traditional devices used in broadcast. However, they are implemented in software, rather than hardware. Software implementations of virtual devices can be free from the encumbrances and limitations associated with physical devices. We need only be concerned with the physical connections and system performance. So an application designed to provide playout functionality for a broadcast stream can be designed to run in duplicate at the same time with the same or differing content, limited only by performance of the computer and enough physical connections to support the required number of outputs. We have found that todays' high performance PC compute engines are quite capable of supporting numerous program streams. And, broadcast level sound card interfaces can now easily support dozens of audio channels. Audio over IP interfaces essentially have no limitation. This virtual machine support makes multicasting for HD Radio™ just another variation of the process often used for remote station operation and streaming applications with the addition of formatting the program and data for the HD Importer or Exporter.

Figures 1 thru 3 show examples of various multiple stream configurations we have implemented for ENCO clients. Figure 1 shows how a broadcaster might use multiple streams and virtual playback machines to provide programs for multiple stations from a single studio workstation. The broadcaster is running six stations from one studio and in fact, from one automation workstation. In this case traffic and music are scheduled on separate workstations and the logs are

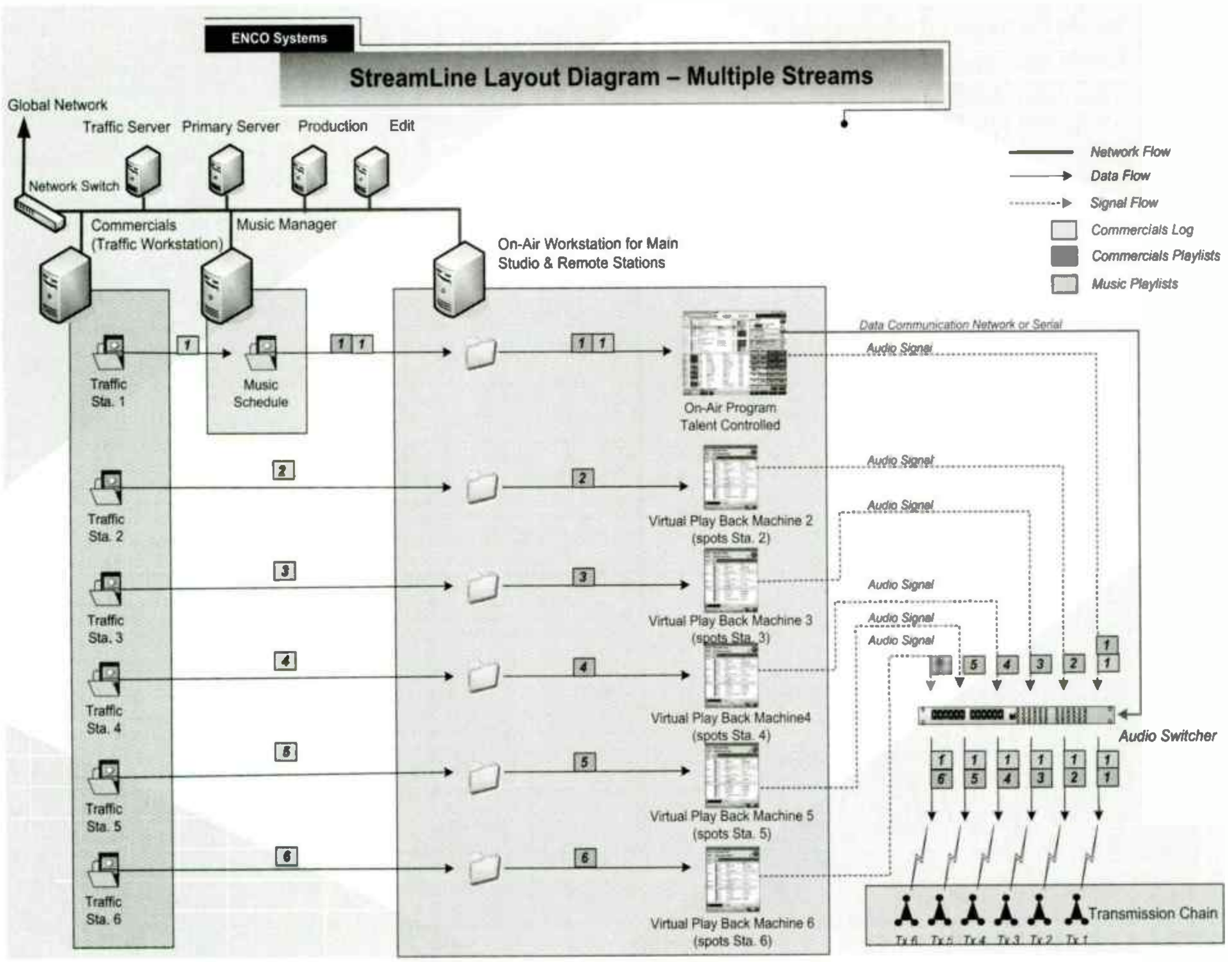


Figure 1

This configuration illustrates six radio stations being multicast from a single automation workstation. All share same music and live content. All have differing commercial spots. Automation command sequences in the playlist ensure that the appropriate commercial playback is routed to the correct station.

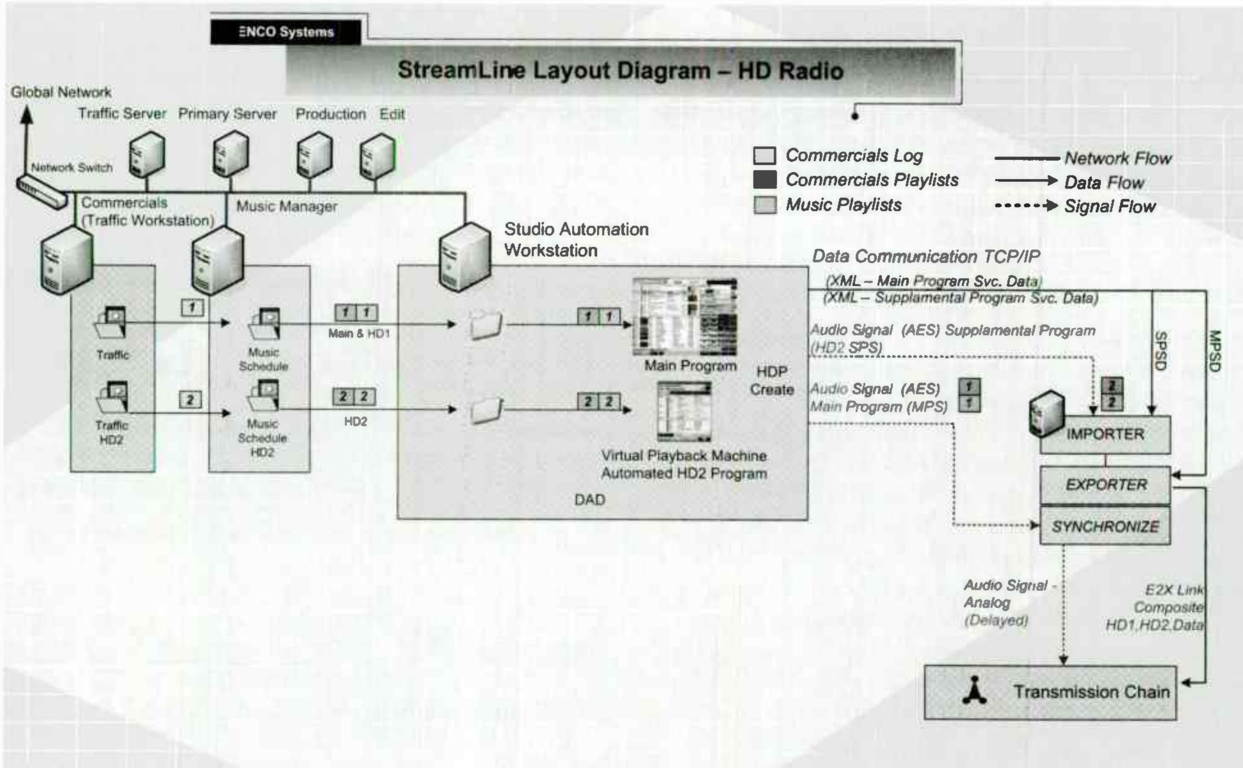


Figure 2
A StreamLine configuration supporting differing main program and HD2 program streams.

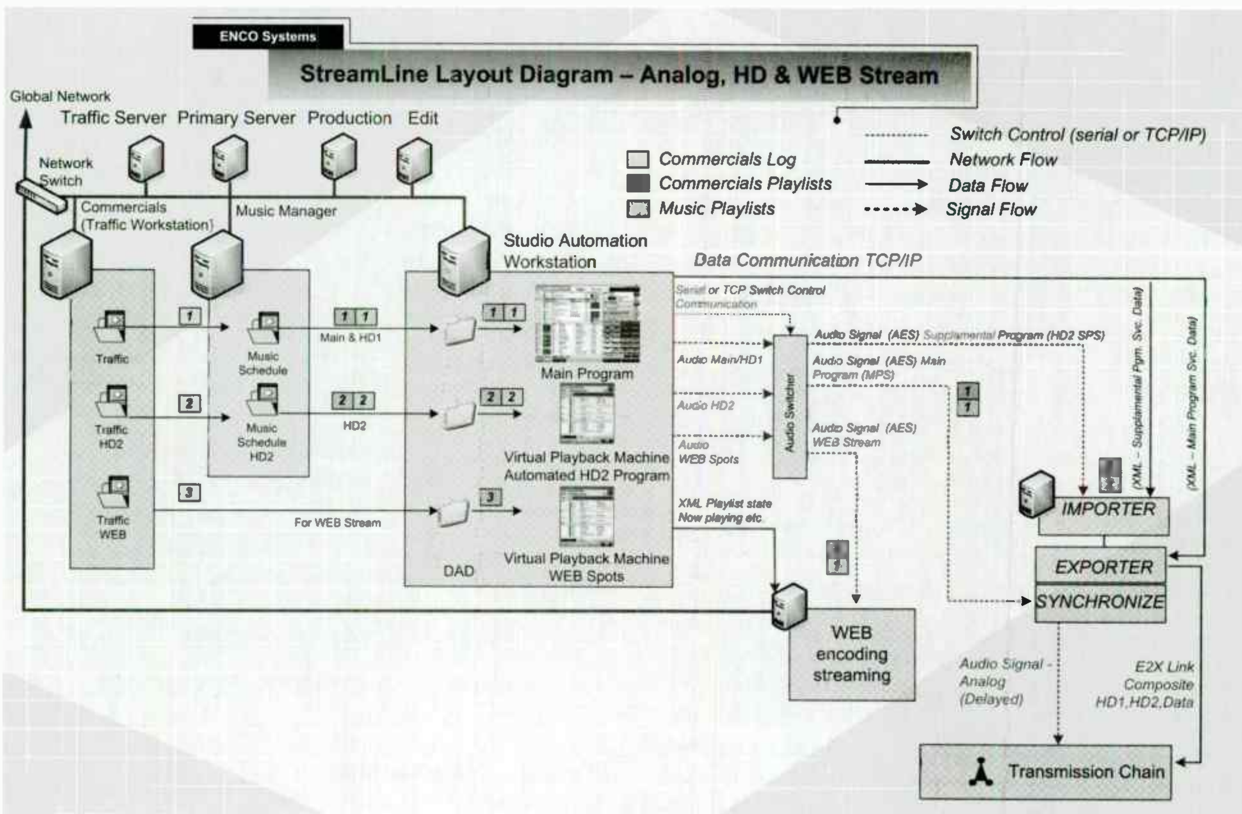
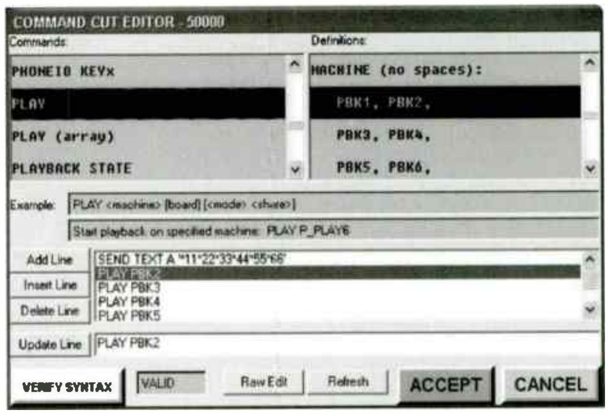


Figure 3
This configuration adds a WEB stream that contains the main program with different commercial spots from a virtual playback machine.

automatically ingested in the appropriate playlist. The music and live content need to be the same across all stations. However each station has unique commercial spots to air. When a commercial spot is initiated, either in the live studio from the on-air screen, or encountered in the main studio playlist, commands are automatically generated to the audio switcher routing station specific commercial spots from virtual playback machines to the appropriate station transmission. The combination of playlist entities doubling as "command cuts" enables a smooth segue transition of unique material for each station while the main program material is from a single source. Dealing with new HD supplemental programs can be handled in a similar fashion. Figure 2 illustrates a multicast configuration for HD Radio™. An audio switcher is not employed as the HD2 program stream is fully automated and routed directly to the exporter. Figure 3 shows the same configuration with an added WEB stream program containing different commercial spots. Similar to the first example, an audio switcher is used to rout just the WEB specific content to encoding and streaming machine or service.

Designing-in measures to ensure easy system adaptation to change is an important factor we must consider in building automation technology that will stand the test of time. One way to accomplish this is through a system architecture similar to that found in the process control industry. Process control runs factories producing everything from potato chips to nuclear power. So what does that have to do with broadcasting? Well you may be surprised! Managing a broadcast stream is in many ways similar; it has a sequence of events controlled in real time, requires interface of a large variety of devices and often needs to make automatic adjustments based on numerous criteria. Systems used to implement process control must be extremely flexible. After all, making potato chips and producing nuclear power both employ process control, but the processes are quite different. Yet the tools created to implement these dissimilar processes might in fact be very similar. Process focused systems are built around a philosophy that permits the user to implement their own features at a high level of abstraction. This means the user can implement powerful and far reaching changes to their system in easy to understand steps and common language. Systems must be designed to allow the engineer that understands the production process to create or modify control of the process in easy to understand terminology and commands, that do not necessitate hard coding of new features in low level computer programming languages. For example, many processes a broadcaster may wish to implement involve how they operate their real and virtual machines under automated sequences or based on external conditions like time of day, program information or other events. It makes sense to give the user control of these things in easy to understand broadcast terminology enabling them real control of their broadcast process. The broadcast engineer should not need to deal with program languages or cryptic command macros. The engineer need only understand his process sequence, not the intricacies of computer programming. Below is a simple example of ENCO's process language, called DCL, used to rout audio in the scenario described in Figure 1.



This example includes the following format:

SEND TEXT A '<serial string for switcher>'
PLAY PBK2
PLAY PBK3
PLAY PBK4
PLAY PBK5
PLAY PBK6

The serial text string will be formatted to match the particular audio switcher in use, but will match the audio inputs to the desired outputs during commercial breaks (in this case 1to1, 2to2 etc.). It also defines the communication port. This sequence may be initiated either manually from the on-air interface or may be automatic when a station specific break is reached in the main program. Upon execution of the switcher command, the virtual playback machines are started with their station specific content routed through the switch to the proper output. Upon completion of the break, the last command cut in each virtual playback stops the playback. The main program then returns the audio switch routing to all outputs matched to the main program input. Of course the duration of the break segments would need to match across all stations. This is a simple example, however we are constantly amazed at the variety of new scenarios and creative solutions we see implemented by customers to solve their unique problems.

Future Considerations

It is difficult to predict how HD Radio™ applications may impact automation systems in the future. Many aspects that have been discussed have the potential to require adaptation through the automation system.

Applications such as electronic program guides, file downloads to the listener, and listener controlled options all have the potential to require adjustments to automation. Broadcasters that implemented any automated actions based absolute real time (non-delayed) for live programming may need to consider how to address issues related to delay and synchronization of HD with analogue streams. Others that may have automated control of audio processing based on playing content may need to consider other differences.

Flexibility and customization has never been more important in the broadcast industry, particularly in light of new HD Radio™ acceptance. When introduced in the early 90's, automation systems were designed to simply automate a well defined, decades old process, and they do an outstanding job. But now, and for the foreseeable future, the process is changing in dramatic and rapid fashion. It is no longer just a radio broadcast stream. There are many streams, both audio and data, both analog and digital, to many locations, using many devices. And this is just the beginning. The future promises more innovation at an ever faster pace. It is clear that a broadcaster that must wait for "new features" or enabling technology to be hard coded into their system, risks falling seriously behind. They must now be able to react to their changing environment quickly, and on their own. The automation and playout system is the backbone that defines and delivers the broadcasters product. It must be designed with this flexibility at its heart.

References

Turner, David " Station Automation and Networking", N.A.B. Handbook, Section 3-6 1999

iBiquity Digital "HD Radio™ Air Interface Design Description - Program Service Data" March, 2005

iBiquity Digital "HD Radio™ Data Network Requirements" October, 2006

Chalmers, R. " The Broadcast Wave Format - An Introduction" EBU Technical Review, Winter 1997

Pierce, Dick and Steadman, Geoff "Introduction" CART CHUNK.ORG web site, 2007

ID3.ORG "What is ID3V2?" ID3.org web site 2006

European Broadcasting Union "BWF - A Format for Audio Data Files in Broadcasting" July 2001

Ballard, Phil "Absolute Beginners XML Tutorial and Resources" The Mouse Whisperer web site 2007

Price, Neil "DAD's Critical Advantage - Protecting Your Future" Letter from DAD Newsletter" Sep. 2006

HD RADIO® CONDITIONAL ACCESS: WHAT IT IS; HOW IT FITS IN THE BROADCAST STATION; AND WHY IT CAN PROVIDE OUTSTANDING RETURN ON INVESTMENT

THOMAS RUCKTENWALD
NDS
Costa Mesa, California

ABSTRACT

Conditional Access is the precise definition of the two words put together; access is based upon certain conditions. When applied to digital broadcasting, a consumer gets just what they want, no more and no less. This model works well for subscription pay services including pay TV.

So, how can this conditional access work for terrestrial HD Radio? How does such a system fit into a radio station environment and into radio station operations? Does terrestrial HD Radio have advantages over other transmission media? What programming works with conditional access?

This introductory paper examines, in understandable professional terms, the possible implementations of conditional access within HD Digital Radio. Conditional access equipment fits into the station, mates to other equipment that may already be within the station, operates within the station's work flow, and creates new operations that the station must perform. This paper's explanations will allow all to comprehend the impact of this new and emerging radio broadcast capability. It suggests some business possibilities that may make conditional access implementation profitable.

WHAT IS CONDITIONAL ACCESS?

Under a conditional access system, reception of transmitted content only occurs when the receiver is authorized to receive the transmission and has the appropriate capability to decrypt the broadcast content. The receiver meets the broadcast conditions and provides access to the content.

Two main concepts embody the technology of conditional access: entitlement and scrambling. The entitlement is an authorization; the scrambling is an encryption of the content.

Entitlement is a right, privilege, or claim. In order to get an entitlement, the recipient must make a contract. In a standard broadcast model, the consumer calls the provider or visits a website and supplies information. They either already have or they arrange for receiver equipment and they order content, programming, or channels. The system handles everything automatically from this point forward, with entitlements based upon the arrangement transmitted from the provider to the correct and appropriate receiver only. The reception equipment knows what is supposed to receive and it provides the desired service.

Scrambling is mixing up or jumbling. For conditional access, such scrambling must be done so that the original content can be reconstructed without error or significant delay. The transmission side performs the scrambling, often using a known and standard type of methodology like 3DES or AES and a standard number of bits, like 128. Information about how to descramble the transmission is sent along with the content. The scrambling, while using a standard methodology, is an unknown due to frequent key changes and will not be decipherable unless the receiver is qualified by an entitlement. The scrambling key is dynamic and changes often over time. These keys, which are used to unlock the scrambling, are never sent directly, but are themselves disguised or encoded so that only a true and entitled receiver may enjoy the programming or content.

Both transmission and reception are a part of this ecosystem. The transmitter and the receiver have complimentary provisioning. The broadcaster transmits entitlements; the receiver recognizes only its own entitlements. The broadcaster transmits scrambled or encrypted programming; if entitled, the receiver can descramble or decrypt the programming.

HOW CONDITIONAL ACCESS WORKS FOR HD RADIO

In many conditional access implementations, the service is offered by one single platform supplier. That supplier provides all the services, the content or channels, and is the source for all the entitlements. Examples include DirecTV for satellite TV, Cablevision for Cable TV, Qwest for IPTV, and Sirius for satellite radio. All content, all services, all entitlements, and all authorized equipment for that system originate from one source, the platform provider.

This is not and cannot be true for terrestrial digital radio. Content suppliers, radio stations, and station groups will continue to independently compete. A conditional access system for HD Radio must coordinate all of these entities, without business interference, so that any equipped receiver may possibly receive these broadcasts.

Every station or station group that deploys conditional access will have the same type of equipment and perform a similar scrambling process. At the same time, each station or station group must be uniquely distinct and recognizable by any equipped receiver. There can be no identity, channel identification, or programming ID duplication.

In order for the system to operate correctly, every radio must be uniquely identified so that its entitlements can be individually addressed to it. Consumers should receive only the programming that they desire and only the programming that is intended for them. Each radio or receiver must perform similarly with every broadcast source.

The best possible answer is to have something within the system that ties everything together. This portion of the system must provide unique station and conditional access service identification to each participating broadcaster. It must authenticate the broadcaster and then provide the information that differentiates the stations. This portion of the system also holds information about each unique equipped radio. It must be able to identify each unique radio and provide it information so that it will perform with the desired broadcast or content.

Such a system is shown in Figure 1. The system includes the existing HD Radio system Exciter, Exporter, and Importer. Only channels created through the Importer may be encrypted as it is anticipated that the main program will remain free-to-air for station license purposes. The Importer will contain a new capability called the Scrambler. The Scrambler will be inert unless the conditional access equipment is in the system. The Scrambler can be used to encrypt multiple channels or programs simultaneously.

The conditional access system contains several operational components; the two main functions are the Entitlement Management Message Generator (EMMG) and the Entitlement Control Message Generator (ECMG). The EMMG generates Entitlement Management Messages (EMMs) which qualify individual radios. The EMMs are transmitted through the Importer on a low bit-rate data channel. The ECMG generates information about the scrambling keys. The Scrambler uses the keys to encrypt the content; the Entitlement Control Messages (ECMs) are transmitted through the Importer and to the receivers. The receivers use this information to recreate the control words that will descramble the content.

The ECMG, because of timing and associative reasons, must be co-located with the Importer at the station. The EMMG and the control over the system operation can be located anywhere, including a Network Operations Center (NOC). A NOC can control multiple stations and route entitlements to either a single or many definable stations. In this fashion, a station group or an aggregator can operate and control numerous stations in a cost efficient manner.

The portion of this system that ties competitive stations and all radio receivers into a cohesive unit is called the National Resource Manager (NRM). The NRM verifies station authenticity, provides unique conditional access service identification, verifies and signs radio entitlements, and holds the database of all radios.

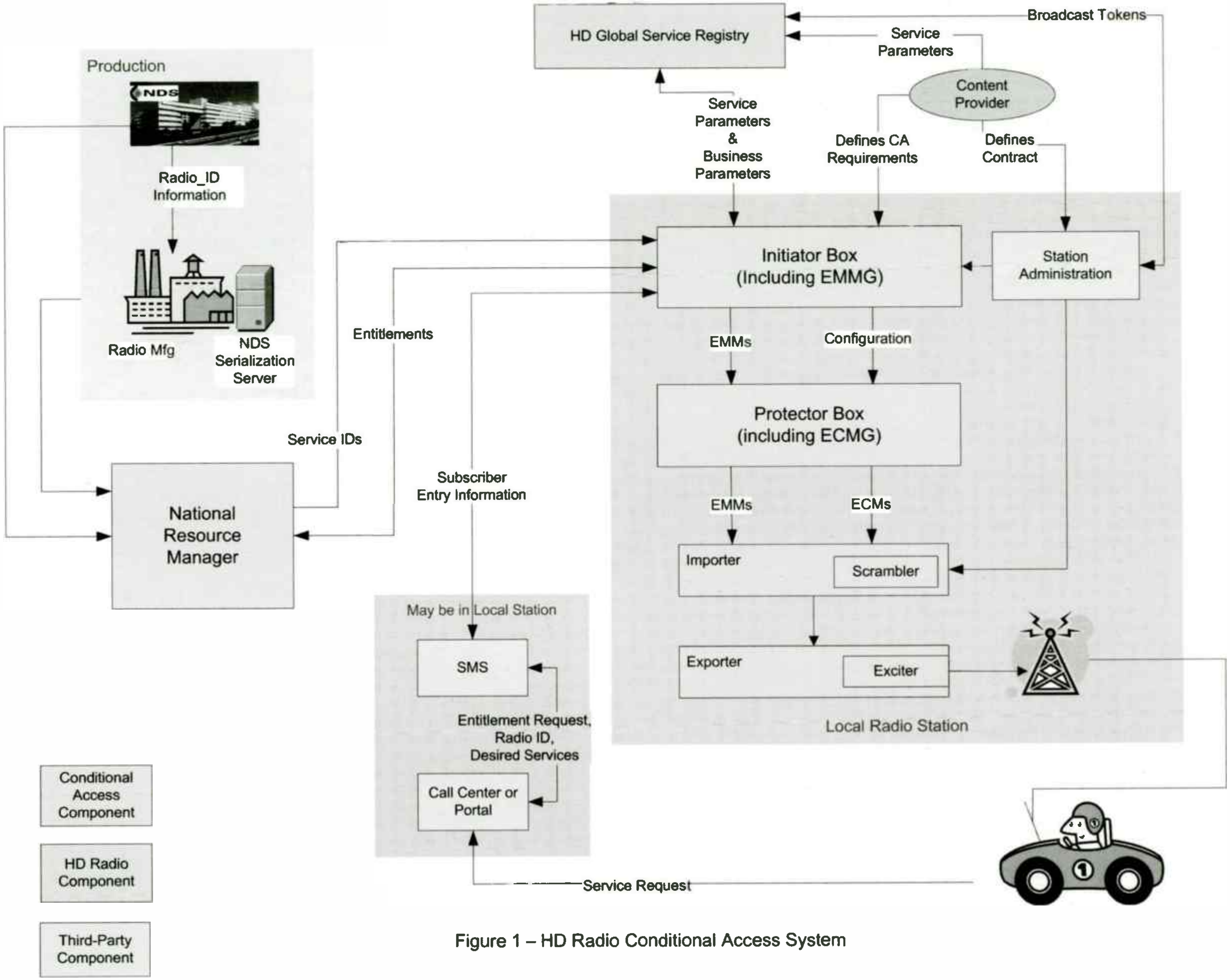


Figure 1 – HD Radio Conditional Access System

RADIOS AND RECEIVERS

This paper focuses on the broadcast implementation of conditional access. However, in order to understand the broadcast requirements, it is important to understand some basic radio requirements.

The radio must be able to decode or decrypt the scrambled content transmission in real time. In order to do that, the radio must know how the content was scrambled and it must already have the information it needs from the system to decrypt it. In a secure system, the information about descrambling and how the content was scrambled is only available to authorized receivers. The authorization comes from entitlements within the broadcast. Through an entitlement, the receiver knows that it is supposed to receive the scrambled signals and it knows how to obtain the descrambling information.

Addressing radios in a system that can receive from many broadcasting sources requires something special. Every radio must be unique to the system, even though the radios can come from many different manufacturers. The most efficient technology that makes every radio unique is serialization.

Each radio is uniquely serialized through the decoder chip. Each decoder chip contains some unique codes and with it, some embedded secrets. The chip/radio identification can be accessed through an activation sequence on the radio. When the consumer calls or registers via a website with the radio information, the system can identify the radio ID authenticity and individually address that radio.

The serialization information is provided to the decoder IC manufacturer by the conditional access manufacturer. The National Resource Manager also knows all the serialization information. Servers located at chip manufacturers and connected to the conditional access manufacturer will program the data that individualizes each HD Radio decoder chip. This process is done for other broadcast systems and is well known in integrated circuit manufacturing.

When the consumer wishes to register their radio and receive conditional access programming, they call the station or register with the online site. With the proper radio ID information, the receiver, within seconds, will obtain its entitlements and automatically turn on, as shown in Figure 2.

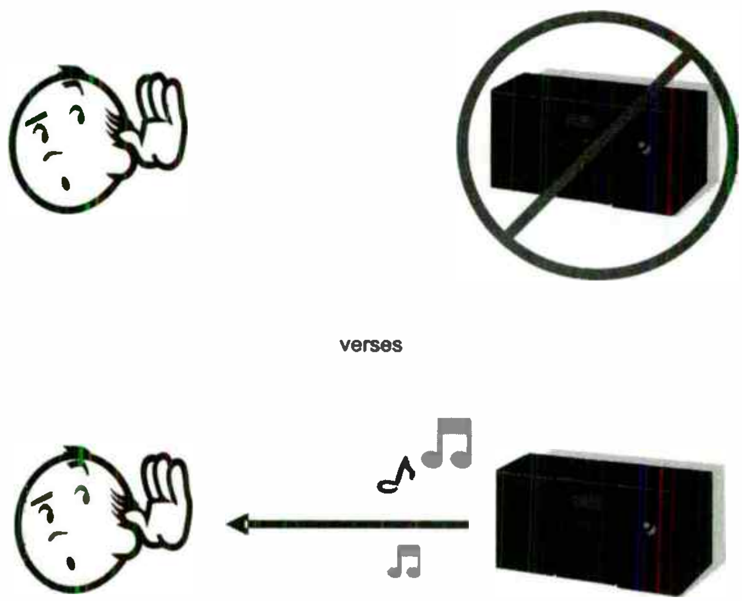


Figure 2 – Entitlements Allow the Encrypted Programming

A CONDITIONAL ACCESS SYSTEM WITHIN THE HD RADIO STATION ENVIRONMENT

Conditional access equipment must fit into the station, mate to other equipment that may already be within the station, operate within the station's work flow, and create the new operations that the station requires.

A conditional access system will connect to the HD Radio Importer version 3.0 or higher. V3.0 contains the HD Radio scrambler module; previous versions do not have this capability and cannot be used for conditional access. The scrambler will be inert or inactive without a conditional access system. Once activated, the scrambler will continue to operate as last instructed, even if disconnected from the CA system.

For an individual station, the conditional access system is embodied in two boxes.

The first box, called the Protector™, mates directly to and must be co-located with the HD Radio Importer. Within the Protector is the Entitlement Control Message Generator (ECMG), and the Entitlement Management Message Spooler, which is a buffering and smart carousel transmission of radio entitlements.

The second box, termed the Initiator™, mates to the Protector and to the National Resource Manager. The Initiator can exist anywhere in the station environment. It contains the EMMG, a User Interface, setup and control over both its own operation and the Protector, the ability to enter information about radios that are authorized to receive from the station, and connectivity to and from the NRM for authorization, verification, and unique radio receiver communication.

The Initiator is intended for single station operation. The type of operations that this box provides may be created on a much larger scale through a Network Operations Center. The NOC may then set up and control conditional access for many station Protector boxes, as shown in Figure 3. EMMs from the NOC EMMG are steered to the appropriate station or stations. The NOC connects to the NRM on behalf of all broadcast stations within its system.

If there is both an Initiator and a NOC, the NOC makes submissions to the Initiator. Content providers, such as Premier Radio or Westwood One, submit their content and conditional access parameters to the station group NOC or the station Initiator. This submission process provides control to the most local entity, preserving traditional US terrestrial broadcast radio localism.

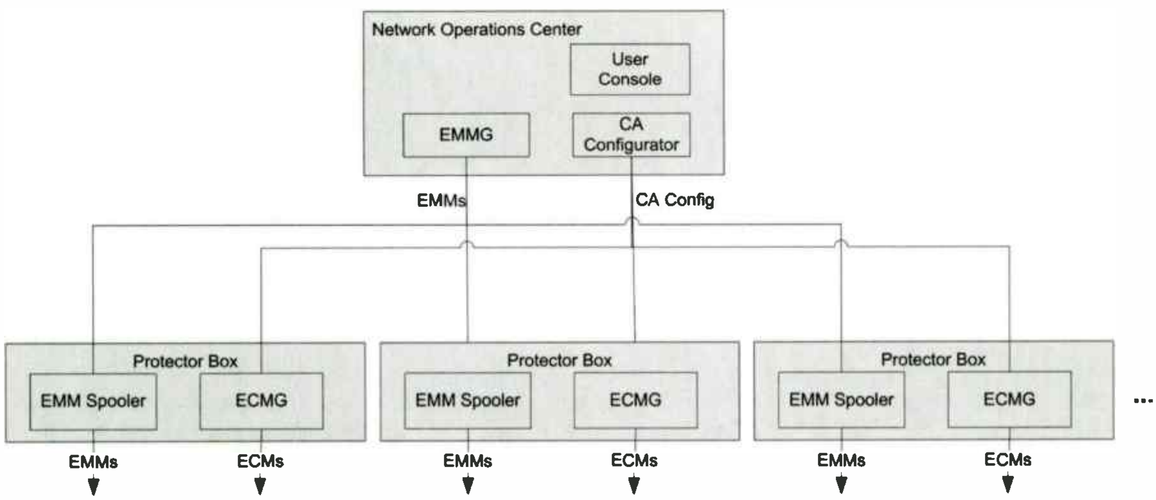


Figure 3 – A Network Operations Center Replaces the Initiator and Controls Multiple Station Protector Boxes

CONDITIONAL ACCESS OPERATIONS

Equipment setup for the conditional access system will be straightforward. A structured entry of station information will activate the equipment.

Application of conditional access can be as simple as turning it on for specific multicast channels and leaving that to run until the station decides otherwise.

The system will also support constant change. Conditional access changes will require operational personnel to access the user interface. CA can be applied to specific programs or channels and changed at the discretion of the station operators. One program may be encrypted but the next program may be free-to-air.

Future automation systems will mate to conditional access. Because of the new programming opportunities, automation systems will need to control multiple playlists and will also need to control additional station equipment. The automation system may provide one place to access and setup the entire station system.

Radio/receiver data entry is another operational requirement. When the consumer registers their radio, the radio ID information is the critical information for entitlement. However, this registration process is an important moment for the radio broadcaster. This is an opportunity to learn more about the consumer. That information will be extremely valuable for the station advertisers. Stations and station groups will have some free choice concerning what information is required for registration.

The standard conditional access equipment will provide capability to entitle radios for the beginning of the station business or for the life of a small station. If the registered radios exceed several thousand, the station or station group should consider either a separate subscriber management system (SMS) or some subscriber or membership software that can be integrated with the conditional access system. Integration with most well-known SMS systems should be an easy process.

Radio registration information can be accepted as direct data entry from operations personnel or a contracted service, entries from an SMS system, or from a web portal that allows the user to self-register. Internet self-registration, once successfully integrated, may provide the lowest cost but may also prove frustrating when the Internet presents a potential barrier, or for those that do not understand the required radio equipment information.

THE NATIONAL RESOURCE MANAGER

The NRM is outside the station environment and is a facility supplied by a third party, the conditional access provider. It is not part of the radio station or station group. Every new station installation is verified by the NRM. Every request for radio entitlement is verified by the NRM. It is a very large database and is the source information for and about the system.

The National Resource Manager will be an equipment-redundant installation. As the conditional access system proliferates, there will be site redundancy.

Under most operations, the radio station or station group will never know that the NRM is in the system. The only time that the NRM is evident is when there is a problem. The primary problem will be that the radio receiver that is being registered does not have appropriate serialization. The consumer is trying to register a radio that is not part of the system.

DIGITAL TERRESTRIAL HD RADIO ADVANTAGES OVER OTHER TRANSMISSION MEDIA

It may be tempting to complain that each radio station allocation represents a narrow bandwidth and because there are already signals occupying that allocation, there is relatively little space for great accomplishments. However, the HD Radio system is proving such complaints unfounded.

HD Radio has shown how a digital overlay can perform tremendous service and that the digital signals can be grouped to provide multiple offerings even though there is an analog signal occupying the associated bandwidth. While there may be less transmission bit space per radio station than TV broadcasts, cable, or telephony, the over 13,000 active US terrestrial radio stations represent tremendous throughput. There is a roadmap to an all-digital station, which will provide even more bit space and greater opportunities.

Our population is already habitually involved in terrestrial radio. 233 million people listen to terrestrial radio an average of 19.5 hours per week. There are 230 million registered cars with radios; 17 million new cars are sold each year, virtually all with terrestrial radio. New models will include HD Radio. Approximately 70 million new radios were made last year and there are over 800 million radios in the US.

HD Radio's creators have had an opportunity to examine the successes and failures of other transmission media. It is easy to see that WebTV failed and that digital video recorders are a hit, although the penetration of such DVR devices is still statistically low. The iPod is an audio experience; even Video iPod usage is 99% audio. An iPod device that can also receive content from a broadcast source may have significant value. The same companies that brought successful TV and Internet technology will bring successful HD Digital Radio technology.

Technology by itself does not guarantee success. Content is King. There will undoubtedly be some national offerings, particularly music from the larger music labels. Conditional access and digital rights management will be used for stored programming and will be particularly important for the music labels and their agent, the Recording Industry Association of America (RIAA).

The biggest advantage for terrestrial broadcasting, particularly radio, is localism. Even with higher penetration of national offerings and increased speed in communication, people within a local area still value their hometown above all else. Morals and values in Peoria, Illinois are different than New York City. People in Texas like Friday night high school football and they want programming that speaks to their lives and to their neighborhood. While programming success might be spelled somewhat differently from one local community to the next, HD Radio is well positioned to accommodate this.

HD RADIO PROGRAMMING THAT WORKS WITH CONDITIONAL ACCESS

Satellite radio is heavily focused on music programming with little to no commercials. Patrons pay for this service, with rates around $13 per month. The success of satellite subscription radio is debatable.

Because terrestrial radio is traditionally a free service, one might expect that HD Radio subscription services would be the last thing to develop. While that may be true for music programming, one might expect data subscription services would be successful from inception. Traffic services are a good example of a data service. Radio signals are hearty and data transmitted to specific applications would require little bandwidth while providing maximum satisfaction. Conditional access is required to secure these data services.

Public service may fuel early conditional access successes. Local Fire, Police, and Emergency Services may require a private emergency channel. The International Association of Audio Information Services (IAAIS) provides radio reading services for the blind. These readings include copyrighted books, newspapers, and magazines; under agreement, these must be offered only to those who are impaired. Conditional access is required to maintain public service privacy.

Pay-per-listen events have high value. Concerts, in particular, seem to extract a strong positive consumer response. Special events of any kind, those that occur on a one-time basis, can provide a new source of revenue for terrestrial radio.

Membership has its privileges. For example, NPR stations can provide additional programming that is entitled only for its members. During pledge periods, where members receive pledge-free programming, the general public receives donation requests.

The most promising offering, however, seems to be "opt-in" services. Opt-in means that the consumer wishes to receive these program services and they make the appropriate arrangements to obtain them. The consumer pays no fees or subscriptions because the service or subscription is advertiser supported. Like cable TV or satellite subscription radio, such opt-in services may be outside profanity restrictions. While a decision specifically on this has not been obtained from the FCC prior to this paper's publication, such a decision would be consistent with other rulings.

Protected opt-in services, available only to those that subscribe through registration, would then have certain artistic freedom and will open up new avenues and programming opportunities. This may affect talk shows, comedy channels, and music that may not be available on a free-to-air station, as well as medical and religious programming.

OUTSTANDING RETURN ON INVESTMENT

The radio station management might be tempted to ask "What is the Killer App?" and "What will be my return on investment in conditional access?"

The good news is that the conditional access system is flexible. It will support data transmission for applications as well as audio. It will support channels that are scrambled all the time as well as part-time and pay-per events.

The "killer app" is going to depend upon the offering, the local station, and the intended audience. For a blind person, the "killer app" is Radio Reading Services. For NPR, it may be membership benefits. For the hardcore commuter, it may be traffic information.

The offering that seems to extract the largest positive response from both broadcasting professionals and consumers involves opt-in. Opt-in continues the advertising-based model. Since the consumer is known, advertising is more valuable because it can be focused.

With opt-in, the radio station will have additional creative license and should be free to program offerings that compete and exceed satellite channels. For the consumer, receiving a free, fashionable service is extremely attractive.

While there are no predictive statistics for increased income or return on investment at this time, we believe that the broadcaster should, for a modest additional equipment investment in their present HD Radio installation, expect to double their present income with HD Radio conditional access protected services.

Pick an application or programming that you feel suits your station, community, and expected target audience.

CONCLUSIONS

The conditional access product for HD Radio is a well-conceived offering that easily fits and integrates with compatible standard HD Radio equipment. One conditional access box provides the data for scrambling and data streams that will be used by the complimentary and entitled radio receiver. Either a second box or a Network Operations Center registers users, enables entitlements, and provides setup and control over the conditional access system. The system should be easy to install and set up. Operational requirements vary depending upon the desires of the station or station group, spanning from set up once to integrating with a subscriber management system or an automation system.

Conditional access also provides for new offerings and opportunities. This technology coupled with content or programming that may appeal to the local broadcast audience should deliver tremendous commercial success.

ACKNOWLEDGEMENTS

HD Radio is a registered trademark of iBiquity Digital Corporation

Protector and Initiator are trademarks of NDS Ltd.

The Consumer Electronic Association for published radio statistics

Gershon Nachshon, Ron Katz, and Patti Daino of NDS for their contributions to this paper

February 2009

Sunday, April 15, 2007
9:30 AM – 1:30 PM

Chairperson: Wayne Kube
Belo Corporation, Dallas, TX

***Analog RF Systems-Uses in a Digital World, Retuning and Reusing**
James York, Dielectric Communications, Raymond, ME

Redundancy in Metadata Generation Systems
Srinath V. Ramaswamy, Triveni Digital Inc, Princeton Junction, NJ

***Planning the Final DTV Channel Change for your Facility**
Dennis Wallace, Meintel, Sgrignoli, & Wallace, Waldorf, MD

The Use of Analog Equipment for DTV Transmission
Myron Fanton, Electronics Research, Inc., Chandler, IN

***An Experimental ATSC-DTV Single Frequency Network in Ottawa Canada**
Khalil Salehian, Communications Research Centre Canada (CRC), Ottawa, Ontario, Canada

A New Generation Smart Antenna for DTV Service
Aldo Cugnini, AGC Systems, LLC, Long Valley, NJ

---

*Paper not available at the time of publication

REDUNDANCY IN METADATA GENERATION SYSTEMS

Srinath V Ramaswamy
Rich Chernock
Triveni Digital Inc
Princeton Junction, NJ

ABSTRACT

Metadata generation systems have become mission critical components in today's DTV plants. Metadata is used in the DTV transmission to facilitate tuning and provide information about the shows and schedule. Metadata is also used to enable many of the enhanced features available with DTV: support for multiple languages, advanced captioning and even scheduling of DVR recordings. In the US, the FCC has mandated that all DTV transmissions must contain correct PSIP (metadata) information. Besides violating the FCC mandate, loss of PSIP information in the broadcast stream can disrupt viewability. Outside of the US, the DVB System Information (SI) [1] serves the same purpose, with similar problems if disrupted. For IPTV, the program guide (based on metadata distributed from the Central Office) plays the same pivotal role.

It has become necessary to ensure that the flow of metadata into the broadcast stream is uninterrupted, best achieved through redundant configurations. Redundant metadata system architectures need to ensure that both internal and external failures are detected and appropriate actions taken for continued operation. For example, internal failures could include database failure, application errors, etc. and external failures could include failures in the link between the metadata generator and downstream devices (such as Multiplexers/Encoders). Redundant architectures also need to support automatic switchover so that the backup generation system is brought into place seamlessly.

This paper details the importance of redundancy in metadata generation systems and also discusses several possible architectures to achieve uninterrupted operation.

INTRODUCTION

Metadata generators play a vital role in providing information to consumers that would allow them to experience DTV to its fullest. For instance metadata in the form of Program and System Information Protocol (PSIP, an ATSC standard [2]) contains tuning information and schedule data that allows the user to tune to his/her channel of interest and view programming information. Besides provide tuning capabilities and a program guide, PSIP preserves station branding. Similarly, outside the US in the DVB arena, the Service Information (SI) standard specifies metadata that allows DVB users to discover services and view programming information. Similarly in IPTV systems programming data in proprietary format and discovery information are provided to consumers IPTV set top boxes. The discussion below will focus on PSIP metadata, but mention will be made of considerations for DVB SI and IPTV metadata as appropriate.

WHY DO WE NEED REDUNDANCY?

In 2005, to protect the consumer, the FCC mandated that PSIP (as prescribed in the ATSC A65/B standard) be carried in all DTV signals emitted by terrestrial broadcasters. This mandate resulted in PSIP generators being widely deployed in terrestrial broadcast stations. Another result of this mandate was that the reliability requirements for PSIP generators became equivalent to other equipment in DTV stations, requiring them to be fully operational 24/7. Loss of PSIP information or incorrect PSIP would not only violate FCC requirements but could disrupt viewability and also push consumers into receiving their television content via other medium. Inaccurate PSIP could also create issues in DTV pass thru devices and DVR enabled receivers.

FCC mandated PSIP requirements

The FCC mandate places the following important requirements on the PSIP present in the broadcast emission:

- PSIP base tables MGT, STT, TVCT, RRT (except region 1) shall be present and carry accurate information. The STT shall be accurate +- 1 second
- EIT's 0-3 carrying correct event information for up to the next 12hrs shall be present
- The DTV major channel number used will be the NTSC RF channel number[1]
- Closed captioning service descriptions (caption_service_descriptor) shall be present when a program carries closed captioning data. These descriptions shall be in present in both the EITs and PMTs for that program
- When a program has a content advisory the parental advisory V-Chip description (content_advisory_descriptor) shall be present

[1] A small number of exceptions to this requirement are documented in A/65

- If audio is present then the audio stream description (audio_stream_descriptor) shall be present

PSIP METADATA GENERATOR SYSTEM ARCHITECTURE

A PSIP metadata generator typically consists of several distributed software components and hardware devices that it relies on to gather schedule information, resolve and generate metadata in the appropriate format and send the metadata to downstream devices in the appropriate format. Figure 1, below, depicts a typical metadata generation system and its sources of schedule information and its metadata destination points.

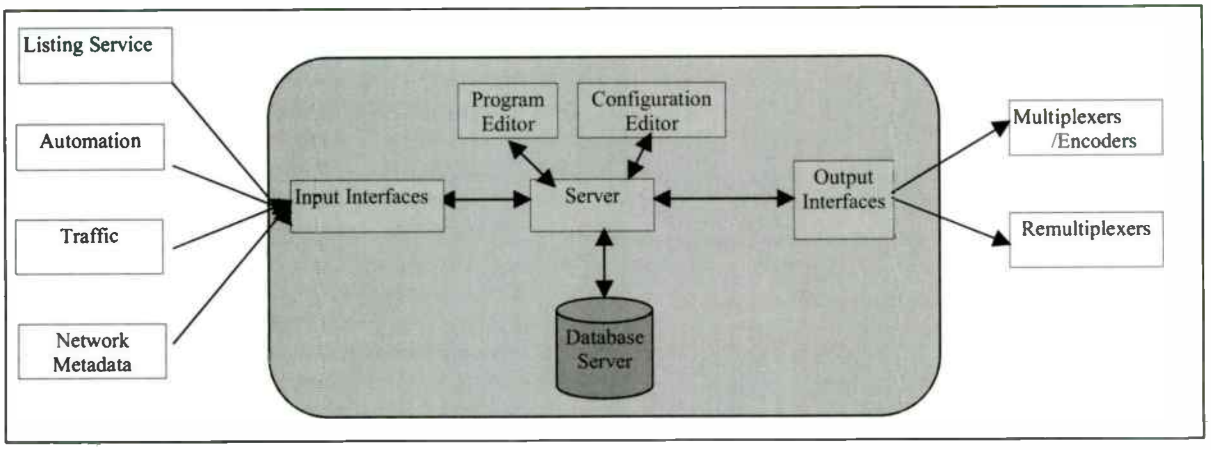


Figure 1. Metadata (PSIP) Generator system

In the above system:

- The database server holds the schedule and generator configuration information
- The Program Editor allows the user to modify/add/remove schedule data and provides a view of programming information that is currently being transmitted as well as allowing the user to make last minute changes.
- The Configuration Editor allows the user to create and maintain the station's PSIP configuration. It also allows one to set up input and output interfaces in order to specify the metadata flow.
- Input interfaces gather schedule data from one or more sources - such as Listing, Traffic, Automation, network PSIP ingest, multiplexers, encoders etc…
- Output interfaces supply the PSIP metadata to external multiplexers, encoders and remultiplexers. In many cases the output interfaces might tightly integrate with external MPEG-2 stream sources to generate the correct PSIP/PSI/SI tables. These connections typically use ASI or IP.

Coordination is necessary between PSIP and PSI. Per the FCC mandate on PSIP, the caption_service descriptor needs to present in both the EIT and PMT for events that are captioned (both carrying the same information). If present, the Redistribution Control Descriptor (Broadcast Flag) must also be present in both locations. This need for coordination often results in the generation of PSI by the Metadata generator. When the PSIP generator is tightly integrated with the multiplexer, it becomes easier to have the metadata generator generate PSI consistent with PSIP based on the current active transport stream configuration setup on the multiplexer/encoder. In such configurations it becomes even more necessary to deploy a redundant system since without proper PSI and PSIP, DTV tuners are guaranteed not to operate.

In many Terrestrial DTV deployments the output signal is remuxed into cable and satellite feeds. At the local remux locations, cable/satellite/iptv equipment may rely on the PSIP/PSI metadata to discover services and provide programming information to their consumers. Failure of correct metadata would result in disruption to carry through of these terrestrial DTV services and affect important revenue source.

TYPICAL METADATA GENERATOR CONFIGURATIONS

Single Station

For a typical single station, a single PSIP generator system will provide PSIP for the single DTV emission signal.

Dual/Triple Station

For this configuration, a single PSIP generator system can provide the PSIP needed for two or three stations that operate within a state or in close proximity. One physical PSIP generator with multiple outputs will be suitable for this type of configuration.

PSIP Central Casting

PSIP Central Casting systems generate PSIP for many more than 3 stations, which typically are geographically widespread. A similar situation exists for the DVB SI scenario - providing SI data for many streams spanning many networks. One or more servers that are centrally located will gather schedule information for all of the stations and generate the appropriate metadata. Once encoded, the metadata is transferred across a WAN to each of the emission stations, where it is injected into the emission multiplex. Redundant metadata systems are very important for the central casting configuration since one system provides PSIP/SI data for many DTV streams.

POTENTIAL REASONS FOR PSIP GENERATOR FAILURES

Listed below are potential causes of metadata generation malfunctions:

- Software errors
    - Operating system issues
    - Driver issues
    - Database corruption
    - Viruses
    - Metadata Generator software errors
    - Network connectivity issues

- Hardware errors
    - Disk failures
    - Time code card failure
    - (ASI) Output card failure
        - Missing tables
        - Table timing errors

PSIP REDUNDANCY SOLUTIONS

Some simple hardware failures may be protected against by relatively simple hardware solutions, For example, a metadata system may be protected against hard disk failures by incorporating RAID arrays or power supply issues by incorporating redundant power supply units. However, the majority of the failure modes listed above require more complex redundancy solutions as discussed below.

What about carouselling?

A common configuration for PSIP generation involves the PSIP generator pushing encoded PSIP tables to the multiplexer, where they are played out at rates defined by the PSIP generator (carouselling as opposed to streaming). While there is a degree of resiliency due to this design (when communications fail between the PSIP generator and the multiplexer, PSIP continues to be present in the emission stream), the need for redundancy is not obviated for the following reasons:

- PSIP/PSI could be inaccurate after a channel line up change. As an example, the viewer may not be able to tune to a channel that was previously declared inactive but is active now.
- Changes that occur during the time the PSIP generator is down are no longer available. For instance the rating for a program may have changed, or there might be an event overrun whose information is no longer available.
- Event information is no longer available after the current set of information expires.

Redundant Muliplexers/Encoders

While not a metadata redundancy system, in some cases the operator feels that his system is more at risk of a failure in the multiplexer or encoder than in the PSIP generator. In this case a single PSIP generation system would meet the metadata needs of both the primary and redundant multiplexers/encoders. The metadata generation system would provide a similar output to both the primary and redundant multiplexers/encoders. The redundancy is not required in the metadata generation system itself and is missing in such configurations. While this configuration protects against failures in the multiplexing and encoding equipment, there is no protection against PSIP generator issues.

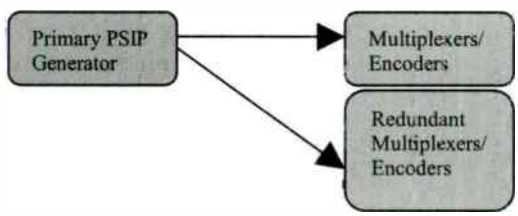


Figure 2. Redundant multiplexers/Encoders

Cold Spare Redundant PSIP Generation System

For this configuration, the redundant or spare PSIP generation system is not activated full time - it is activated only when the primary system fails. This setup requires user intervention when a failure occurs and does not offer the benefits of more automated approaches, such as automatic switchover, error detection from active redundancy with monitoring, database mirroring etc. This configuration does not guarantee uninterrupted delivery of metadata to downstream devices.

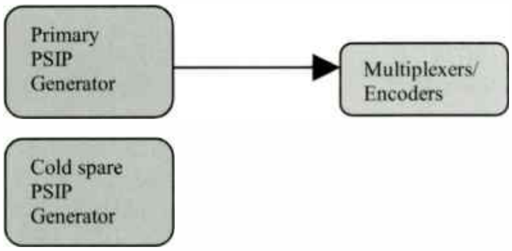


Figure 3. Cold Spare Redundant System

1+1 Metadata Generator Redundancy

In this configuration, there are two separate physical metadata generators (the Primary and Redundant system) that are tightly coupled. The Redundant system monitors the Primary's output for errors, and the Primary's health using SNMP and heartbeat mechanisms. Upon detection of metadata generation failure, the redundant system can automatically take over metadata generation.

Figure 4, below, depicts the interactions between the Primary and Redundant system internal components in facilitating redundancy.

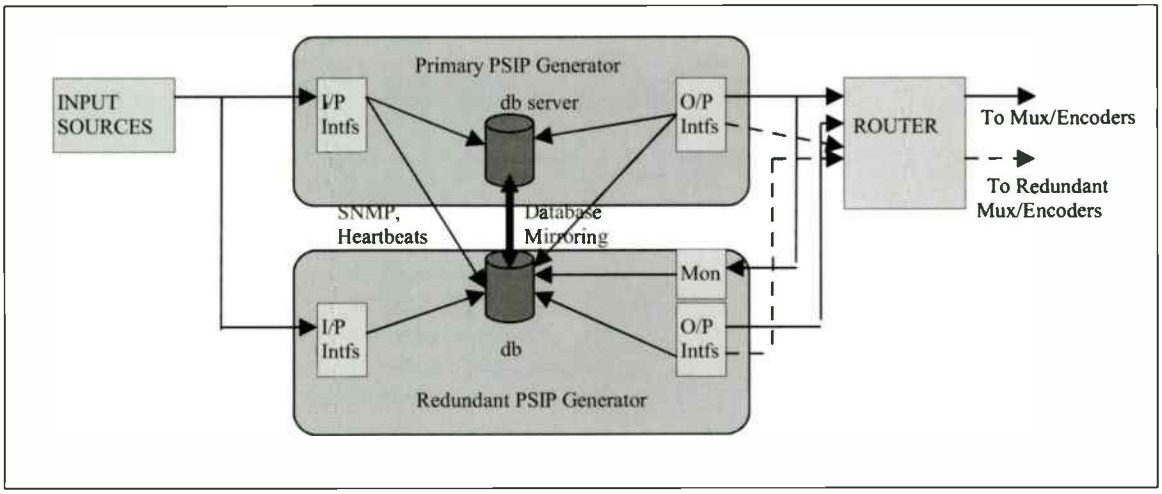


Figure 4. 1+1 redundancy solution

The Primary PSIP generator operates as normal, dynamically generating metadata by acquiring schedule data from various input sources and creating PSIP/SI/PSI tables that are then fed into multiplexers/encoders via a router. The router might be an ASI or IP router, depending upon the input connection into the multiplexer. In this configuration the Primary's database is periodically replicated into the redundant database server. This allows the redundant system to not only automatically create a copy of the database (which will include any manually scheduled events) but also to take over metadata generation with most current configuration setup on Primary with minimal user intervention.

Besides feeding the multiplexer, the output of the primary PSIP generator is also routed to the monitoring module ("Mon" in Figure 4), where the PSIP is checked for validity. Errors in PSIP that would be representative of a failure can cause a switchover in PSIP source.

Based on a user configurable switching criteria (described later) the redundant system could be activated automatically and it could request the router to route its generated metadata to the downstream multiplexer/encoder and block the metadata from the primary reaching the downstream devices. Depending upon the connection needs, the router could be a typical ASI router or an IP router whose routing functionality can be controlled from the redundant system.

To facilitate redundancy the Primary system monitors its internal component's health and provides heartbeats and SNMP data to the redundant system. The primary system could also request the redundant system to take over metadata generation when it detects alarms from various internal failures, one of which might be its inability to communicate with cards that generate ASI streams or hard disk failures.

In the redundant mode the redundant system does the following,

- Mirror Primary's database

- Monitor the validity of PSIP on the Primary's IP/ASI output
- Play the role of watchdog, listening in to heartbeats from the Primary's internal components.
- Monitor SNMP parameters and traps to detect Primary internal components failures.
- Monitor alarms from tightly integrated multiplexers/encoders to detect metadata errors.

The monitoring module described above has the following features:

- Handles both ASI/IP input errors
- Detects MPEG-2 Transport packet errors
- Detects PSIP errors, including improper schedule data
- Table timing errors
- PSIP errors per ATSC A/78 [3]
- SI errors per ETR 101 290 [4]

The 1+1 redundant solution would meet the metadata redundancy needs of the majority of DTV stations configured in the single station or dual/triple station configuration, which requiring a single Metadata generator. This system would also provide the metadata needs of redundant downstream devices, but the switchover criteria for the redundant system may be less stringent compared to the primary multiplexer's feed, There could be more than one router in such configurations. The redundant system could be on a low cost server and could also function as a minimal monitoring solution.

N+M Redundancy solution

For complex DVB SI metadata generation or PSIP Central Casting configurations there typically might be more than one metadata generator providing SI/PSIP to many networks and stations. In such scenarios, for N metadata generators, there could be M redundant units -where M<N and M often is one. Figure 5, below, shows the architecture for such a configuration.

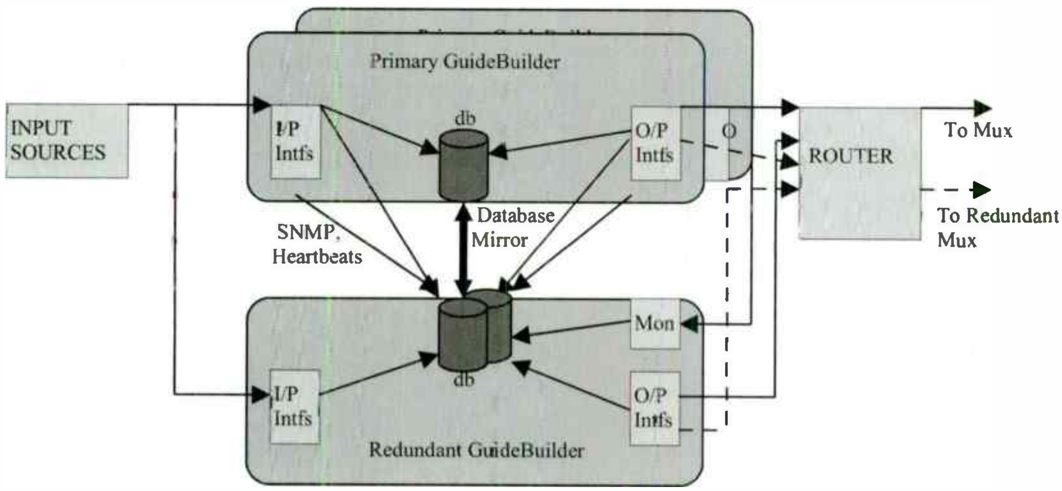


Figure 5. N+M Redundancy configuration

For this configuration, each redundant unit would be the backup for a subset of the primary unit(s) with the load assignment made during initial configuration. The redundant unit would mirror the databases of the primary units that it is assigned to. Similar to the 1+1 configuration there would be SNMP and heartbeat based monitoring and also the capability to automatically take over the operation of the primary upon detecting a failure. This is a much more complex system than those described above, especially when automated switching is involved as the redundant system would need to correctly configure itself to match the failed Metadata generator's configuration.

Criteria for Automatic Switchover

The following conditions would cause the redundant metadata generator to automatically take over metadata generation (alternatively, in the manual mode these could trigger an alarm to the operator indicating that attention is needed):

- Emission MPEG-2 Transport packets are not MPEG-2 compliant (too many Continuity Counter errors, CRC errors, section errors etc...)

- PSI data incorrect or not present (this is chosen if the user has activated PSI generation on the metadata generator)
- PSIP data is not A/65 compliant
- DVB SI data is not compliant
- No output stream detected
- Heartbeats are missing from primary
- Unable to communicate with end devices
- Metadata generator internal errors
- Errors detected in multiplexers and encoders tightly integrated with the Metadata generator

The above automatic switching parameters would be user configurable such that the user can determine the appropriate switching criteria to be applied to match the station's operational requirements.

External Communications

In all of the above redundancy configurations the PSIP generator system should be capable of providing alarms via email, SNMP traps etc… to external monitoring equipments or users.

CONCLUSION

Metadata (PSIP/PSI/SI) plays a vital role in today's DTV systems. There is a requirement for metadata generation systems to be safeguarded against potential impairments through the placement of redundant metadata generation systems. The FCC mandate requiring the inclusion of PSIP in broadcast DTV streams has made it even more compelling to implement redundancy in PSIP generators. The fact that viewers may not be able to tune at all without proper PSIP/PSI/SI information makes it very important to have the metadata generator operational 24/7.

Feasible redundancy solutions are described above. By choosing the appropriate redundancy solution a DTV station can guarantee reliable metadata generation thus ensuring viewership and quality of service.

REFERENCES

1. ETSI EN 300 468 V1.7.1 "Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems"

2. ATSC Standard A/65C "Program and System Information Protocol for Terrestrial Broadcast and Cable, Revision C, with Amendment No. 1"

3. ATSC Recommended Practice A/78 "Transport Stream Verification"

4. ETSI TR 101 290 V.1.2.1 "Digital Video Broadcasting (DVB); Measurement Guidelines for DVB Systems"

THE USE OF ANALOG EQUIPMENT FOR DTV TRANSMISSION

Myron D. Fanton, PE
Electronics Research, Inc.

INTRODUCTION

At the termination of DTV simulcasting, many broadcasters are assigned their original NTSC analog channel for their ATSC digital transmission. Faced with this channel assignment the use of the original transmission equipment is a real option. An engineering analysis of RF transmission equipment informs the decision and defines the limits in using existing equipment designed for analog operation.

POWER CAPACITY

Key to the analysis of equipment designed for analog TV service is the specification of power capacity. The average power capacity denotes the management of heat within the RF component, primarily affecting product lifetime. Peak power capacity governs the maximum sustainable voltage gradient without ionization of the dielectric material.

Peak power, called at times peak envelope power, is specified as the average power in a steady-state, continuous wave (CW) that produces the desired instantaneous peak voltage [1]. In this manner the power of NTSC transmissions are specified as the peak power of the visual synch pulse. Digital power is specified as average power, the ratio of peak to average power varying with the modulation scheme.

The computation of instantaneous peak power values should be avoided as impractical. Though the semantics may be clumsy, peak power quantities are the RMS average of sine waves that achieve specified peaks. Instantaneous voltage peaks have both computational and practical value.

Equating the total, instantaneous peak voltage produced by NTSC and ATSC television transmission yields the following.

$$\frac{P_{DTV}}{P_{NTSC}} = \frac{\left(1+\sqrt{x}\right)^2}{w} \qquad (1)$$

The fraction, x, is the aural-to-visual ratio of the NTSC transmission, typically assumed to be 0.10 but can be larger in practice. The peak-to-average power ratio of the DTV transmission is the factor w, typically assumed to be 5 for 8-VSB modulation, but also can be larger in practice.

$$\frac{P_{DTV}}{P_{NTSC}} = (x+y) \qquad (2)$$

The total Average Power of the combined visual and aural NTSC transmission is typically assumed to be 70% of the specified visual power. The average power of the visual transmission, a fraction y of the specified peak power of the visual transmission, is 0.60 for the worst case, black screen condition.

Therefore, 7kW of DTV power produces the same voltage peaks as 20kW of NTSC power and the same average, heating power as 10kW of NTSC power. With a 20% reduction in specified power between the two systems, RF component peak power capacity has a 70% safety factor and average power 350%.

Peak Power in Coax

The power handling capabilities of coaxial lines are based primarily on two factors: the maximum peak power (or maximum voltage gradient that can safely be present) and the maximum average power, which is determined by the allowable temperature rise of the inner conductor.

Deriving the peak power capacity of coax begins with the computation of a required DC test voltage and relates this to an RF voltage and peak power.

The DC test voltage is derived from the following equation that includes the air density derating factor:

$$E_p = \left(3.17\times10^4\right)(d\delta)\left(\log\left(\frac{d}{D}\right)\right)\left(1+\frac{0.273}{\sqrt{d\delta}}\right) \qquad (3)$$

where
E_p = production test voltage, V
δ = air density factor = 3.92 B/T
B = absolute pressure, cm of mercury
T = temperature, K
(δ = 1 for B = 76 cm and T = 23°C = 296 K)

The production test voltage is converted to RF RMS voltage,

$$E_{RF} = (0.7)\left(\frac{1}{SF\sqrt{2}}\right)E_p \quad (4)$$

where

E_{RF} = maximum RF RMS operating voltage with no derating for VSWR or modulation, but includes a safety factor, *SF*.

$1/\sqrt{2}$ = RMS factor

0.7 = DC to RF factor [2]

SF = safety factor for voltage (typically 1.4 for coaxial cables and 2 for rigid coax)

The peak power rating in watts, P_{pk}, can now be calculated:

$$P_{pk} = \frac{E_{rf}^2}{Z_o} = \frac{\left(\frac{(0.7)E_p}{SF\sqrt{2}}\right)}{Z_o} \quad (5)$$

On the first order, derating for the presence of standing waves may be performed by simply dividing by the VSWR. The nature of reflections originating from the RF system has been studied [3]. The Channel Reflected Energy (CRE) is defined from measured reflection data (i.e., VSWR data) and is related to SNR degradation in the DTV receiver. The CRE is the RF system performance parameter relating directly to DTV reception and analog system reflections should present negligible degradation to DTV system performance.

For most installations, the peak power ratings will not be a significant factor as they are typically much higher than a single transmitter system can generate. The primary concern will be for multiple channel installations where two or more TV signals are combined into the same transmission line. If the peak voltages from two or more signals of equal power add together in phase, the equivalent peak power rises as the square of the number of carriers. In this situation, voltage levels can become the primary concern in specifying the transmission line type.

Average Power in Coax

The average power rating is determined by the amount of heat created due to line losses. The temperature rise is primarily limited by the safe, lifetime performance of the dielectric material used to support the inner conductor. Since the temperature rise on the inner conductor is greater than the outer conductor, the maximum allowable temperature is normally specified based on inner conductor temperature at the rated power level. Typical industry conditions have been to allow the inner conductor to reach a temperature of 100°C with an ambient temperature of 40°C. This means the inner conductor temperature is allowed to rise 60°C above the ambient. The average power rating can then be calculated using the following equation:

$$P_{avg} = \frac{(16380)\sigma D}{\alpha M_\alpha} \quad (6)$$

where

P_{avg} = average power rating for 60°C rise of inner conductor temperature

D = outer conductor outside diameter, in

σ = heat transfer coefficient of outer conductor, watts/in2

M_α = correction factor for attenuation (relative to 20°C)

α = attenuation constant, dB/100 ft at 20°C

Standard heat transfer coefficients are listed below in Table 1 for rigid coaxial line types.

Line Size	Zo	σ
7/8	50	0.1280
1-5/8	50	0.1200
3-1/8	50	0.1070
4-1/16	50	0.1035
6-1/8	50	0.0970
6-1/8	75	0.0770
7-3/16	75	0.0760
8-3/16	75	0.0740
9-3/16	50	0.0900
9-3/16	75	0.0660

Table 1: Heat Transfer Parameters for Rigid Coax

The average power rating is based on the temperature rise on the inner conductor and this in turn affects the lifetime performance of the dielectric material. Therefore, operation at higher temperatures will result in a reduction in the life expectancy and reliability of the line relative to the lower temperature performance.

Barring improper installation or damage, the typical failure mode of coaxial lines is damage to the connection points as a result of excessive heating over time. Based on this observation, long-term operation of coaxial lines at elevated temperatures is not recommended.

Power Capacity of Antennas

Like coaxial transmission line, radiating elements reach peak power limits when the dielectric surrounding them begins to break-down. In a gas media this occurs as the molecules ionize, causing arcing, scintillation, or corona. In coax, this would typically occur at the inner conductor, where the voltage gradient (E-field) peaks, and an antenna with coaxial conductors will exhibit this same limitation at high input power levels.

Pressurized radome enclosures eliminate the effects of the environment on the internal components of the antenna in the same way a supply of pressurized, dry gas protects transmission lines. This reduces the maintenance involved and increases peak voltage capacity [4].

Slot antenna elements possess large voltage gradients between the walls of the slots and between coupling devices and nearby conductors. Dipole elements have high fields near the feed-points and at the tips of the conductors. The presence of this high field limits the peak power capacity and highlights the need for surrounding the element with pressurized, dry air.

The average power capacity of an antenna is constrained by the management of energy dissipated in components. Many antenna designs have short-circuits to tune the input and terminate elements. This facilitates heat flow from the inner conductor. Radomes and other plastic components must be designed to dissipate a minimum amount of heat, keeping operating temperatures low.

In any case, an antenna originally designed for analog operation may be operated at equivalent peak and average powers. Antenna equipment that has been operating in a pressurized, dry environment will have increased safety margins and extended useful life.

Creating Enclosures

Waveguide and coaxial transmission lines contain many joints sealed with o-rings around the flanges. Terrestrial and satellite microwave transmission lines are similarly sealed and pressurized [5]. All such transmission lines contain critical junctions that would corrode or otherwise degrade in the presence of contaminants, and for this reason that almost all high-power broadcast transmission facilities have pressurization equipment supplying dry gas to transmission lines.

These systems contain o-rings and high-pressure flanges providing a seal, and they are pressurized as well. With a sealed volume of air, or any other gas, a pressure vessel exists. Such volumes are created in an effort to eliminate ingress of water or other contaminants. A small amount of gas is leaked in every seal, and over time a pressure differential will not exist between the sealed volume of gas and the environment surrounding it (Figure 1).

Efforts to create volumes sealed from environments that include the passage of weather systems are often thwarted by the pressure changes. A situation where the outside pressure rises above the pressure of the sealed environment is inevitable and accompanied by moisture (Figure 2). The moisture and other contaminants are ultimately pumped into the sealed volume. Without equipment to constantly provide greater pressure in the enclosure, water and other contaminants would be pumped into the sealed line every time a front passes. Without a supply of dry air, condensation forms on the inside of the enclosure.

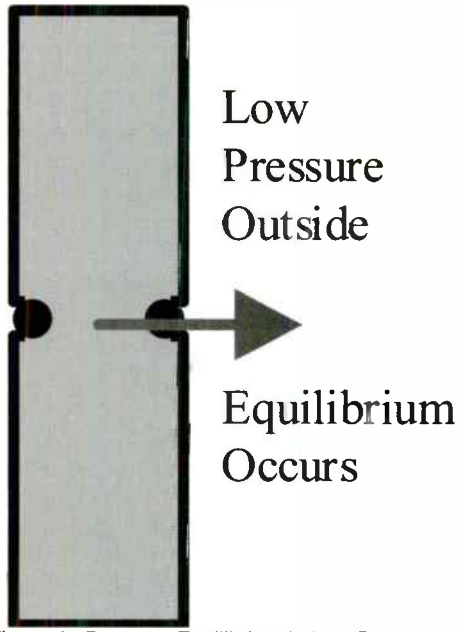


Figure 1: Pressure Equilibrium in Low-Pressure Storm

Ingress Protection Ratings

The IEC Test Standard EN 60529 [6] outlines an international classification system for the sealing effectiveness of enclosures of electrical equipment against the intrusion into the equipment of foreign objects (i.e. tools, dust, fingers) and moisture. This classification system uses the letters "IP" ("Ingress Protection") followed by two digits.

The first digit of the IP code indicates the degree that persons are protected against contact with moving parts (other than smooth rotating shafts, etc.) and the degree that equipment is protected against solid foreign bodies intruding into an enclosure. The second digit indicates the degree of protection of the equipment inside the enclosure against the harmful entry of various forms of moisture (e.g. dripping, spraying, submersion, etc.)

Pressurized Antennas

Broadcast television antennas have critical components that would be impaired greatly in the presence of water or other contaminants. Many antennas have flange junctions nearly identical to those critical junctions found in the transmission lines mentioned above. In addition, the performance of radiating elements, coupling devices, and power dividing components suffers if subjected to corrosive or contaminated environments. Without the use of pressurization equipment to eliminate the ingress of contaminants, periodic maintenance is required to clean and repair the damaged surfaces.

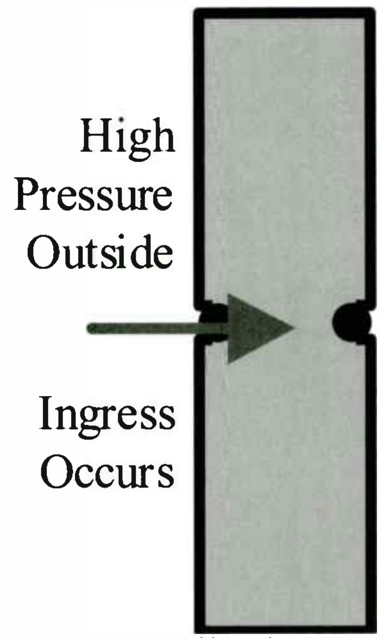


Figure 2: High Pressure System Moves In

Ingress protected antenna equipment, with IP67 or better ratings, requires no periodic maintenance. No dust or debris from the harsh external environment will contaminate critical conducting surfaces and points of conductor contact. No water will be pumped into the antenna by passing storms and pressure changes. After 30 years of service, critical components will still be shiny and undamaged.

COVERAGE AND INTERFERENCE

The FCC rules for De Minimus Interference have strategic application to enhance a DTV station coverage performance. The original analog antenna pattern may not be optimal for DTV coverage, especially if the original azimuth pattern is omni-directional. The advantage of directional radiation patterns is that they simultaneously reduce interference and cover important population centers at the same time.

The planning factors and analysis tools involved in the design of the original analog station are substantially different for DTV planning. Numerical models of propagation and interference account for terrain and population parameters. These changes may make the current analog RF system undesirable for DTV transmission.

PRODUCT RELIABILITY AND LIFETIME

The components in an RF system that limit lifetime are the items made of plastic. By way of example, the average power of a cable is designed to limit the inner conductor temperature to a level than allows a satisfactory life of 20 years. Clearly, running a cable hot enough to cause melting, or even softening, of the dielectric is unacceptable. While the polyethylene materials used in coaxial cables have melting points around 130°C, the operating temperature needs to be significantly lower than this to limit the long term degradation of the dielectric by oxidation.

Levels of Protection	1st Digit	2nd Digit
0	No special protection	
1	Objects greater than 50mm in diameter.	Protection from dripping water.
2	Object not greater than 80mm in length and 12mm in diameter.	Protection from vertically dripping water.
3	Tools, wire, etc., of thickness greater than 1.0mm.	Protection from sprayed water.
4	Any object with a diameter or thickness greater than 1.0mm	Protection from splashed water.
5	Volume of dust that would interfere with operation	Protection from water projected from a nozzle
6	Dust tight.	Protection against heavy seas, or powerful jets of water.
7	NA	Protection against immersion.
8	NA	Protection against complete, continuous submersion in water.

Figure 3: Ingress Protection (IP) Ratings

Even with less than 1% oxygen in the atmosphere, dielectric materials react with it at elevated temperatures, ultimately degrading their electrical and mechanical properties. Dielectric loss increases and the plastic turns brittle; this can lead to movement of the inner conductor and line failure. Like all chemical reactions, the reaction rate increases substantially with quite modest temperature increases. Typically, the rate of oxidative degradation increases, and lifetime decreases, by a factor of 2 for each 10°C rise in temperature.

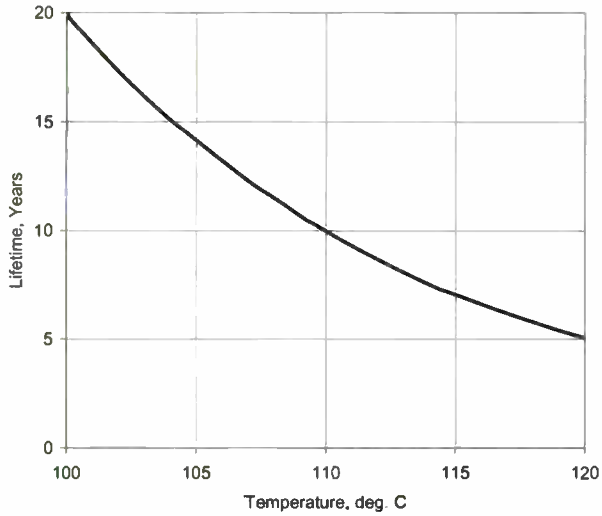


Figure 5: Lifetime reduction and conductor temperature

Therefore, and increase in inner conductor temperature will accelerate oxidation of the dielectric and shorten usable life. You can estimate this effect for a polyethylene dielectric cable from Figure 5. For a maximum inner conductor temperature of 100° C. oxidation proceeds slowly enough to assure a cable life of approximately 20 years. However, with the inner conductor temperature raised only to 115° C, cable life is cut to 7 years.

CONCLUSIONS

Applying DTV to analog RF equipment is permissible if care is taken to equate peak power. This creates additional margin for average power, allowing the analog equipment to reach its original design lifetime. However, the antenna pattern and consequently the coverage and interference performance of the transmission system may not be optimized for DTV.

REFERENCES

[1] Sgrignoli G., "Measuring Peak/Average Power Ratio of the Zenith/AT&T DSC-HDTV Signal with a Vector Signal Analyzer," *IEEE Trans. On Broadcasting*, vol. 39, pp. 255–264, Jun. 1993.

[2] Cooper R., "Experiments on the Electric Strength of Air at Centimetre Wavelengths," *Jour. I.E.E. (London)*, vol. 94, pp. 315-324, Sep. 1947.

[3] Fanton M, "Analysis and Measurement of RF System Reflections and DTV Transmission," *www.broadcastpapers.com/tvtran/tvtran01.htm*, May 2003.

[4] Silver S., *Microwave Antenna, Theory and Design*, V12 of MIT Radiation Laboratory Series, McGraw Hill, New York, 1949.

[5] Wittaker, Jerry, ed., *NAB Engineering Handbook, 9th Edition,* Washingtion, DC, 1999

[6] IEC 60529, Edition 2.1, *Degrees of protection provided by enclosures (IP Code),* Geneva, Switzerland, 2001.

A NEW GENERATION SMART ANTENNA FOR DTV SERVICE

Akira Hayashi
Funai Corporation
Rutherford, NJ

Eiji Shibuya
DX Antenna Co., Ltd.
Kobe, Japan

Aldo Cugnini
AGC Systems LLC
Long Valley, NJ

ABSTRACT

With the arrival of exclusively digital terrestrial broadcasting in the United States now less than two years away, an increasing number of viewers will be relying on indoor antennas for free over-the-air programming. Because many viewers will be situated in DMAs that are served by multiple transmitting locations, a fixed antenna will not provide optimum reception across the available channels. In addition, even community antennas using the same transmitting site may cause receivers to experience different multipath reception conditions across different channels. To aid these situations, electronically-steerable Smart Antennas have been developed that automatically adjust the preferred signal direction for each particular broadcast emission. This optimization can take into account various factors, such as signal strength, multipath strength, and/or bit error rate. One such antenna has been developed by DX Antenna, a subsidiary company of Funai. We will discuss the performance and other characteristics of Smart Antenna systems that include a controlling circuit in a typical DTV set-top box.

MOTIVATION

The U.S. Congress (and by enforcement, the FCC) have stipulated that, as of February 18, 2009, all full-power TV stations nationwide must switch off their analog broadcasts. In addition, so that analog TV viewers are not disenfranchised from broadcast service, Congress has mandated the implementation of a program through which households in the United States may obtain coupons that can be applied toward the purchase of digital-to-analog converter boxes (DTAs). In response to the creation of this program, to be administered by the NTIA, cross-industry submissions by broadcasters and consumer electronics manufacturers have validated the need for high-quality, easy-to-use, low-cost digital converter boxes, so that consumers and program service providers are given the best opportunity to maintain quality TV services.

Digital Signal Reception

As is well known, terrestrial television reception – both analog and digital – is subject to many transmission path impairments, including that of multipath, where multiple "echoes" of the transmitted signal arrive at the receiving location. Whereas this impairment results in "ghosts" upon display of an analog television broadcast, the digital consequence is an increase in the received error rate of the signal. This can, in moderate-to-severe multipath situations, lead to an impaired signal – with compromised video and audio presentation – or, in the worst case, no reception at all. While this situation can be improved by careful physical aiming of the receiving antenna, this adjustment will not be ideal for all received stations, due to their different transmission powers, frequencies, and locations, and can therefore antenna adjustment can be problematic under poor reception conditions.

Furthermore, in a report on challenges that could arise in connection with the DTV transition, the GAO cited reports that some people may need new antennas or adjustment of existing antennas in order to receive digital over-the-air signals, and that improved antenna technology may be needed for some households.

ANTENNA SYSTEM DESIGN

In order to improve reception under varying conditions, smart antennas have already been implemented at cellular telephone base stations. It is now practical to utilize the same technology for consumer digital television reception. By providing an automatic mechanism to adjust the antenna, the direction and gain (amplification) of the antenna can be electronically changed, with no need for user intervention or physical adjustment of the antenna.

One such antenna functions by changing the relative gain and phase (delay) of the internal elements. While offering a high degree of optimization for both signal capture and interference rejection, this kind of adaptive antenna is somewhat complex, and hence expensive, to implement. Phased arrays have been in use for quiet some time by the military.

Outdoor Smart Antenna

An alternate type of smart antenna is the so-called switched beam antenna system. In this system, multiple fixed elements within the antenna are selectively utilized, so that a primary receiving direction is favored. At the same time, strong sources of multipath can be negated; an optimization algorithm can perform a tradeoff between the two factors. The user simply plugs in this antenna – to a suitably-equipped DTV receiver – and the receiver will automatically adjust the antenna for optimal reception of each DTV station.

Indoor Smart Antenna

Selecting an antenna direction and gain setting for optimum signal reception involves an assessment of the signal quality over the operating extent of the antenna. Various parameters of the received signal can be evaluated and weighed, including signal strength, mean squared error of the channel equalizer, spectral flatness, and unwanted interference. Depending on the system architecture, this optimization process can be tightly integrated with the demodulator, or can be implemented separately.

The combination of direction and gain can also be used in a more sophisticated algorithm that anticipates 3rd-order intermodulation interference from strong UHF taboo channels, or from the n ± 1, 2 channel pairs where tuner RF selectivity may be minimal.

CEA-909 SMART ANTENNA INTERFACE

While a Smart Antenna can be an option to the consumer, it will only function if the appropriate interface is available at the receiver. Such an interface has been developed and standardized by the Consumer Electronics Association in CEA-909, entitled "Antenna Control Interface." This standard allows any compliant receiver to operate with any compliant antenna, regardless of manufacturer. The standard also defines the data algorithms used, connection standards, and other requirements. The antenna configuration is neither specified nor implied, leaving specific design considerations to the manufacturer.

At the time of this writing, the CEA R4-WG4 Working Group was defining a control protocol that works over the antenna coax, with a target of two options for the CEA-909 Standard: the existing control protocol that uses a separate connector and a new one that shares the RF signal connector/coax.

PROTOTYPE

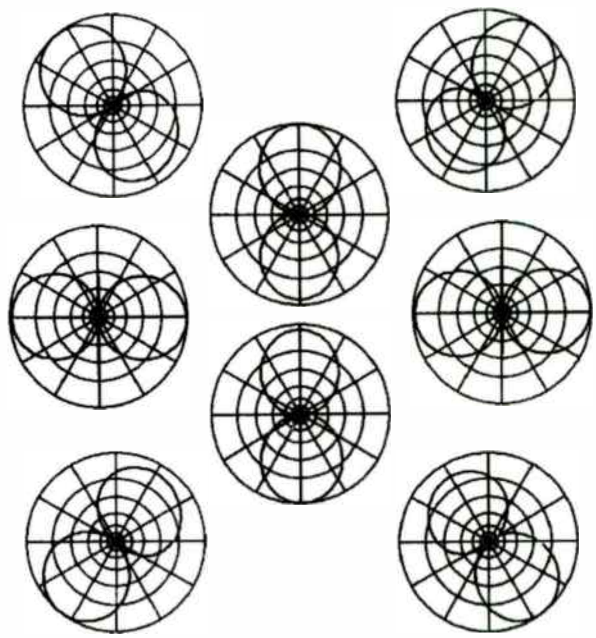
While an optimum antenna for outdoor use can be somewhat large, indoor use necessitates a smaller, more elegant design. While this will inherently place limits on the performance of such a device, the unit will generally be used in areas of high signal strength, and can thus be smaller than an outdoor unit designed for fringe area reception.

Frequency (MHz)	VHF: 54–216 (ch2–13)
	UHF: 470–806 (ch14–69)
Antenna Segments	8
Polarization	horizontal
Gain over monopole	15–20 dB; includes selectable gain amp and FM trap
Noise Figure	<4.0 dB
VSWR	< 2.5:1
Impedance	75Ω
Interface	Type F, CEA-909
Dimensions (mm, in.)*	216(W) × 216(L) × 140(H) (8.5 × 8.5 × 5.5)
Weight	2.0 kg (4.4lbs.)
* without VHF rod antenna	

Prototype Parameters

Based on a successful larger design, DX Antenna, a subsidiary of Funai, has developed a small prototype smart antenna with the preliminary parameters shown here. Some variants on this are possible; e.g., collapsible rods can be added if needed for VHF service.

In addition to selecting one of eight different azimuth directions, the unit can operate with two different levels of RF amplification. This is useful in areas of high signal strength to avoid direct overload of the receiver front end.



Prototype Smart Antenna Reception Patterns

FIELD EXPERIENCE

The prototype has been tested in a variety of field locations with promising results. The antenna was setup indoors on various levels of urban buildings. Using a proprietary setup algorithm, a DTV receiver proceeded to optimize the antenna at each location for the each of the received DTV signals. Once setup, there was no further human intervention for the smart antenna. A representative sample of the actual received station count data is shown below.

	Floor	Smart Antenna	Mini Yagi	Amplified Loop
Torrance	2nd	0	2+0	0
	4th	21	17+4	15+2
Santa Clara	1st	19	12+4	13+3
	2nd	6	3+2	4+1
San Francisco	24th	12	7+0	6+3*
Indianapolis	1st	5	3+2	3+2
NY Metro	3rd	10	7+2	9+1
* not amplified				

Typical Received Station Count

At each location, two other indoor antennas were also used for reference: a popular "mini yagi," with no amplification, and a combination set-top UHF loop/VHF rod antenna, with built-in amplification. All of the indoor antennas were tested at the same location, position and approximate time. With the reference antennas, the direction of the antennas was first established to receive at least one station, and then was optimized manually for each station, in order to maximize the number of receivable channels.

As can be seen, the smart antenna performs at least as well as the hand-optimized yagi—but requires no user intervention. In a very small number of locations, the smart antenna was unable to receive a signal—in those few cases, it was seen that the yagi required meticulous adjustment, apparently due to a very high level of multipath.

SUMMARY

A new smart antenna has been developed for DTV service that can automatically optimize DTV reception for viewers relying on indoor antennas. With this ease-of-use, consumers can now enjoy trouble-free DTV reception in their digital home. The many stakeholders in the DTV transition have emphasized that consumers should not incur a hardship in making the transition in 2009. This technical development helps to ensure that both integrated televisions as well as digital-to-analog converter boxes provide consumers with the ability to continue to receive free over-the air TV content without degradation in picture, sound quality, and viewing experience after the DTV transition.

ACKNOWLEDGEMENTS

Our thanks to Bob Cannistraro of Funai and Art Allison of NAB for their useful comments and assistance in preparing this paper.

Some of the schemes described in this paper may be covered by issued or pending U.S. and international patents. Appearance of antennas is subject to change without notice.

REFERENCES

Digital Television Transition and Public Safety Act of 2005, Title III of the Deficit Reduction Act of 2005, Pub. L. 109–171, 120 Stat. 4, 21, February 8, 2006.

Implementation and Administration of a Coupon Program for Digital-to-Analog Converter Boxes, Department of Commerce, National Telecommunications and Information Administration, Docket Number: 060512129–6129–01, July 25, 2006.

Comments of the Association for Maximum Service Television, Consumer Electronics Association, and National Association of Broadcasters, In the Matter of Implementation and Administration of a Coupon Program for Digital-to-Analog Converter Boxes, September 25, 2006

Digital Broadcast Television Transition: Several Challenges Could Arise in Administering a Subsidy Program for DTV Equipment, GAO–05–623T, May 26, 2005.

The Multipath to Clarity – Receiving HDTV over the air takes luck and lots of patience, Philip Yam, Scientific American, June 2005.

Smart Antenna Systems, International Engineering Consortium, September 2000.

Antenna Control Interface, (Standard), CEA-909, June 1, 2002.

Method and apparatus for controlling a smart antenna using metrics derived from a single carrier digital signal, European Patent EP1484847, published August 12, 2004.

Anticipating Signal Behavior, Charles W. Rhodes, TV Technology, October 20, 2004.

RF Implementation for HD Radio

Sunday, April 15, 2007
1:00 PM – 5:30 PM

Chairperson: Steve Fluker
Cox Radio/Orlando, Orlando, FL

***HD Radio Implementation Case Studies: Dual-Polarized Master-Antenna**
Myron Fanton, Electronics Research, Inc., Chandler, IN,

Isolation in FM IBOC Multi-Channel Systems-St. Louis, Real World Data
Henry Downs, Dielectric Communications, Raymond, ME

***Pre-correction Techniques for Radio Broadcast Transmitters**
Richard Hinkle, Broadcast Electronics, Inc., Quincy, IL

***Low Loss FM Filters Employ Evanescent Modes**
Derek Small, Myat Inc., Portland, ME

***FM Digital Radio Implementation Update**
Kerry Cozad, Dielectric Communications, Raymond, ME

***Location Availability Coverage Analysis of AM HD Radio**
Charles Cooper, du Treil, Lundin & Rackley, Inc., Sarasota, FL

FM- HD-Radio Field Trial Results under European Frequency Planning Conditions
Markus Ruoss, Ruoss AG, Rotkreuz, Switzerland

A Precision, Low-Cost GPS-Based Synchronization Scheme for Improved AM Reception
Stephen Smith, Oak Ridge National Laboratory, Oak Ridge, TN

***HD Radio Technology Implementation in Brazil**
Marco Tulio, Radio Globo Brazil
Marcelo Cacheiro, Brazil-Harris Broadcast
Perry Priestley, iBiquity Digital, Columbia, MD

*Paper not available at the time of publication

Isolation in FM IBOC Multi-Channel Systems-St Louis, Real World Data.

Henry Downs
Dielectric Communications
Raymond, Maine.

Abstract

There are a number of important parameters to be considered when building an FM IBOC system. One, which has proved to be most crucial, is the aspect of combiner isolation between the transmitters. Due to the 20 dB difference between the analog and digital power levels, much of the consideration up to this point has revolved around the Isolation from the High Power analog transmitter to the Low Power digital transmitter. The new solid state digital transmitters have been shown to perform adequately with isolation as low as 40 dB. However, when considering the existing tube type analog transmitters, it has to be recognized that both the digital and analog signals are at the **"same frequency"**. As such there is no "turn around loss" as typically experienced in multi-station combiner systems where the transmitter to transmitter combiner isolation has a typical minimum goal of 55 dB with an additional turn around loss of approximately 20 dB. For IBOC systems with 40 dB of combiner isolation coupled with a 20 dB power level differential and zero turn around loss from the tube type transmitter, the possibility of spectral re-growth becomes very real. This paper addresses the system factors which have to be taken into account for every installation and real world data from a multi-station system is presented.

Introduction

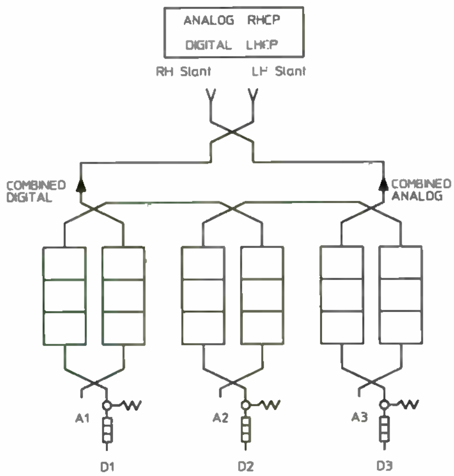
Isolation is fast becoming the IBOC deployment "war-cry". As more and more stations go live with IBOC, it has become readily apparent that, although the numerous components in an IBOC system provide excellent individual performance with respect to isolation, their collective performance, when combined as a system, is less than optimum. This phenomenon is found both in single station systems and also in the large multi-channel systems. This paper discusses primarily a large multi-channel system, but the concepts described may be applied equally well to a single station installation. When considering a multi-station IBOC system, it is necessary to examine many different facets of operation. Typically, such systems, will fall into one of two categories.

- Pre-combined IBOC signals will be fed into a multichannel combiner and subsequently fed to a wideband antenna.
- Analog and digital signals will be combined through the use of a multi-channel combiner and then fed to the two inputs of a dual input antenna or, in some cases, two separate antennas.

In either set-up, the parameter which invariably has to be studied exhaustively is the system isolation. The importance of this particular system parameter cannot be expressed enough. During the early IBOC development years, it was routinely stated that the isolation requirement between the analog and digital transmitters had to be of the order of 35 dB. This number has ebbed and flowed +/- 5 dB or so but, generally speaking, 35 dB has been the "minimum norm". The various antenna and system FM manufacturers have been continually striving for optimum isolation from each IBOC component and many critical milestones have been reached at the component level. Isolation numbers in excess of 40 dB have been recorded for many combining components on the ground. Most dual input IBOC antennas offer a minimum of 20 dB isolation between the inputs, although the Dielectric HDFMVee and the reverse polarized interleaved antenna have demonstrated isolation numbers greater than 35 dB over most of the FM band (1,2,3,4). In order to demonstrate the interoperability of the various components used in a typical system the results of a recent 10 channel system upgrade will be presented.

St Louis 10 Channel Expansion/IBOC Upgrade

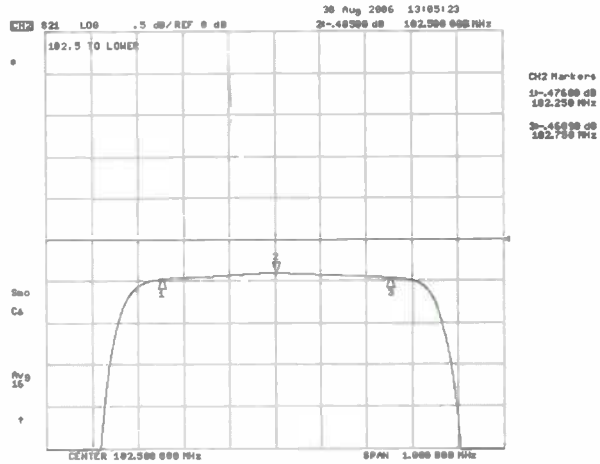
A recent upgrade was completed in St Louis, which entailed the addition of 2 channels to an existing 8 channel system. In addition, the complete system was also upgraded to be completely IBOC ready. On the ground, the existing system was comprised of a number of Constant impedance filter (CIF) Combiner modules daisy chained together to facilitate the combination of all of the channels onto a common output line. The combined signal was then split via a patch panel/matched Tee set-up to allow two equal amplitude signals to be fed up dual line runs to a dual input top half/bottom half antenna. Such a patch panel system also allows for a reduced power mode whereby the combined signal may be directed to either the top half or bottom half of the antenna in the event of a partial antenna failure. Customer requirements dictated that the CIF combiner be used to facilitate the combination of the analog signals and the digital signals onto two separate output lines.



Such systems, which have been built in the past, offer an efficient and cost effective method to Implement IBOC.

Existing Channel Analysis

The initial step, as in any IBOC conversion/upgrade was to perform a complete system analysis on the existing system to determine what, if any, components could be re-used. This entailed an examination of the combiner all the way through to the antenna to ensure adequate system bandwidth characteristics. As shown here the existing filters in the combiner had sufficient bandwidth such that IBOC implementation could be accomplished. However, it was determined that the input port to port isolation would be marginal and new superior performance hybrids for the inputs were recommended.

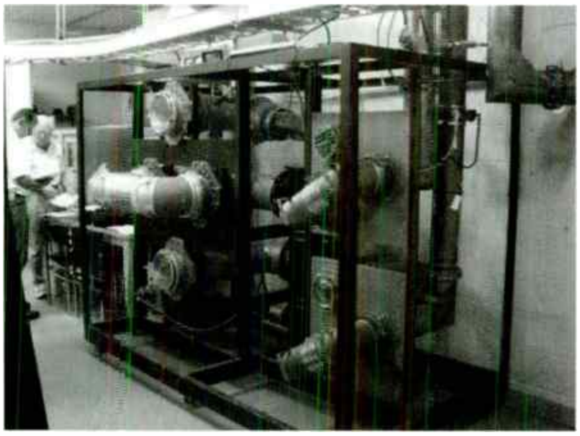


The original system incorporated patch panels on both the input and output of each CIF module in order to effect a by-pass capability for emergency use.

In order to utilize the combiner modules to combine both analog and digital signals onto separate output lines, it would be necessary to remove these patch panels.

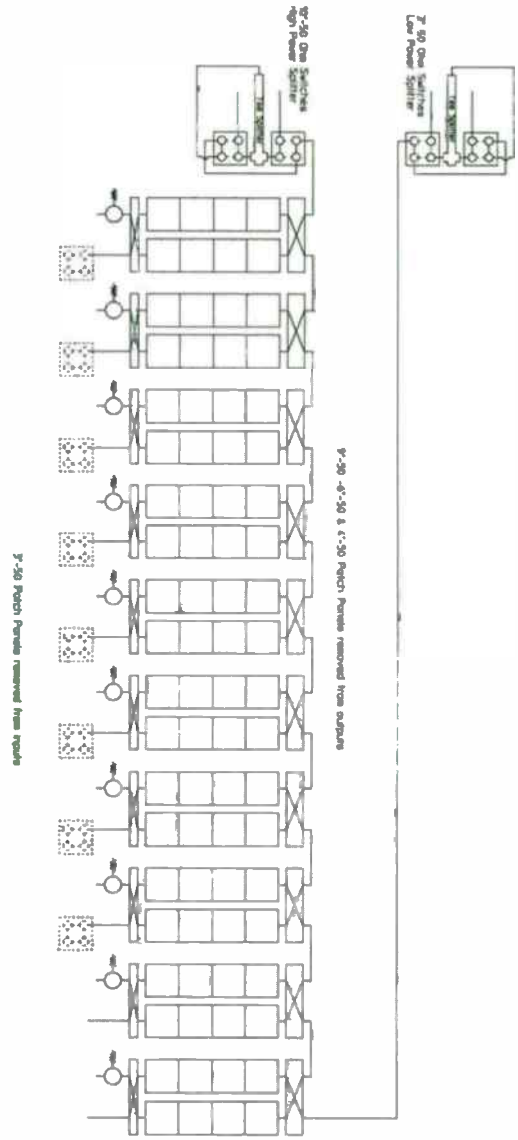
Patch Panel/Splitter Configuration

The existing configuration allowed the option of splitting the combiner output signal equally in order to feed it to both halves of the Dielectric CBR antenna. It also allowed a "reduced power" combined output signal to be routed to either half of the antenna for use in emergency mode. It was determined that this system could stay in place and work satisfactorily if it had sufficient power handling capability.



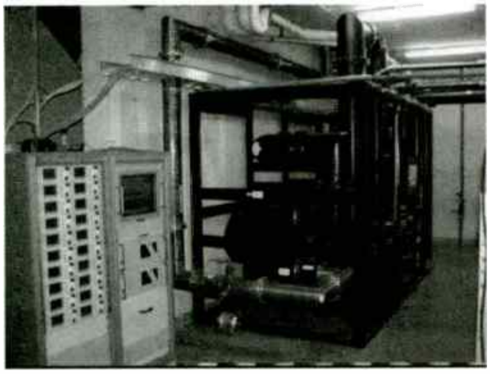
Dual Line Runs and Dual Input Antenna

The existing antenna system had been in place for 20 years and was at the upper limit of it's
power capacity. The addition of IBOC along with two more stations was determined to push it beyond reasonable operating limits. In addition, previous investigations carried out had demonstrated that pre- IBOC dual input antennas had evolved over time to favor a single input by making the individual cross dipoles of each bay of the antenna asymmetrical in order to maximize the broadband VSWR response. .(5) It has been demonstrated however, that such asymmetries result in an extremely poor antenna input to input isolation. Based upon the original St Louis geometry, this isolation number was estimated to be as low as 13 dB in some areas of the FM band.
With the system analysis complete and the customer desires taken into account, the following system was proposed.



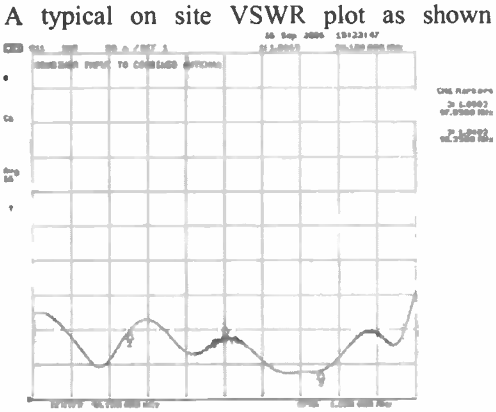
System Performance

The new system installation was completed in July/August of 2006 with no install issues and to-date, has performed flawlessly.

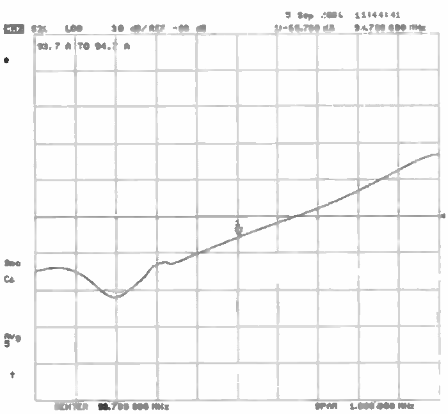


Typical system data is presented here

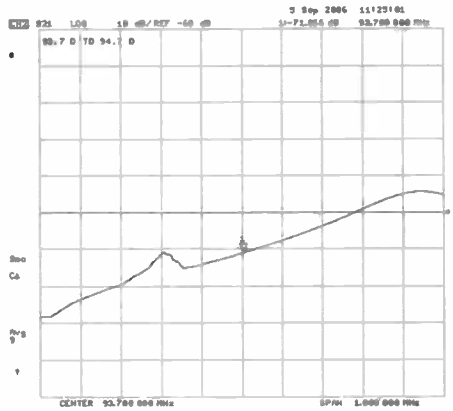
A typical on site VSWR plot as shown here



is well within the system specifications.
The isolation between like ports, namely Analog to Analog

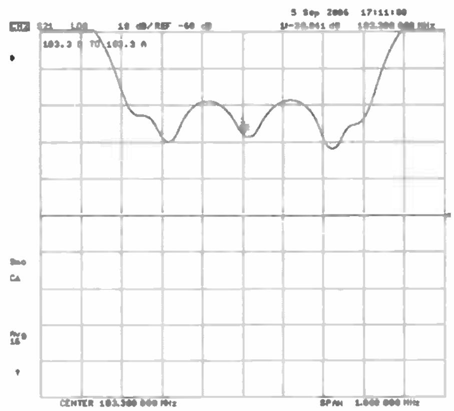


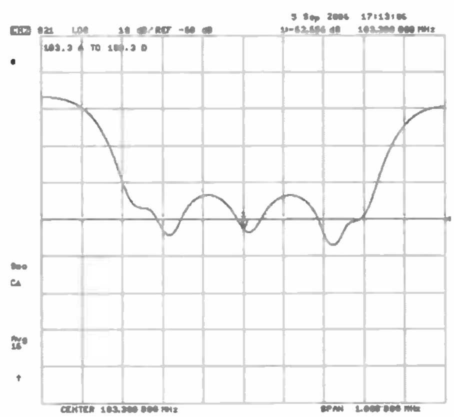
and Digital to Digital as shown here for closely spaced channels is, as expected, greater than the 60 dB minimum design goal.



Analog to Digital Same Channel Isolation

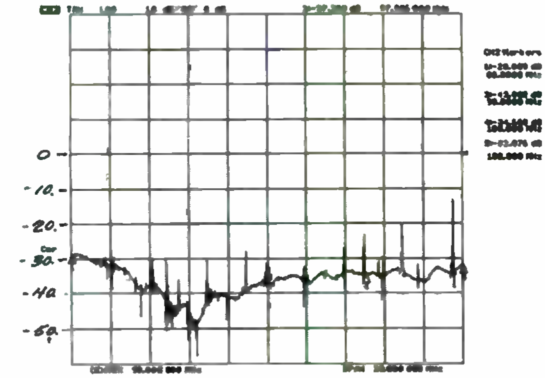
One of the key isolation figures is between the analog and digital transmitters for the same channel. As is evident from the two plots here





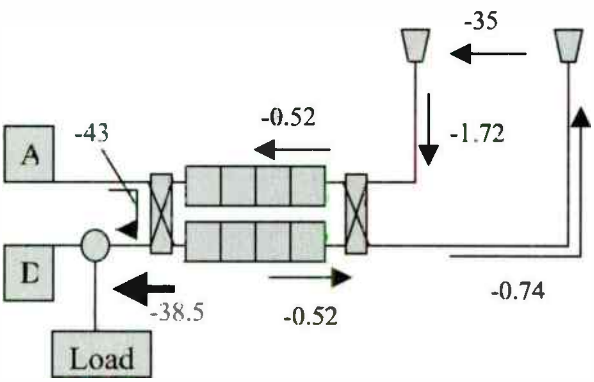
the traces are essentially identical, differing by the effective Circulator isolation of approximately 25 dB. It is also noted that the analog to digital isolation is somewhat less than would be expected when considering the isolation numbers of the individual components in the system.

The system isolation plots shown are for 103.3 MHz where the antenna isolation alone is of the order of 35 dB



for a system isolation number, added to this 35 dB for the antenna are the line losses and combiner losses in both directions. When added up the resultant "antenna isolation is approximately 38.5 dB
The input of the combiner module yielded an isolation of 43dB.

The old adage that a system is only as good as it's weakest link certainly does not appear to apply to the isolation in an IBOC system. One has to consider all of the possibilities. In this case, consider the following :



In the above diagram, the invocation of "Murphy's Law" states that the path lengths through the system will invariably cause the signals due to Antenna Isolation and Combiner Isolation to be combined in some manner, such that the true analog to digital isolation of the system is less than that of either the combiner or the antenna. In this case, the antenna system isolation including combiner and line losses is of the order of 38.5dB while that of the combiner input was optimized for 43 dB. The resulting analog to digital isolation is 38 dB, as depicted in the earlier digital to analog isolation plot. In this instance, this exceeds the accepted minimum of 35 dB. However, in order to safeguard the digital transmitter it was decided to implement a circulator on each IBOC input port. It should be noted here that the circulator has to be sized appropriately for the system in order to handle both forward and reflected power levels. One has to continually bear in mind that the analog power level is 20 dB higher than that of the digital.

Digital to Analog Isolation issues

The system configuration is symmetrical with respect to the analog and digital input characteristics. As such the digital to analog isolation is also 38 dB. Taking into account the difference in signal levels this isolation results in the analog transmitter being subjected an incoming digital signal level of -58 dB with respect to the analog output. The burning question is, "Is this enough?" As far as the existing analog transmitter is concerned this is seen as VSWR and will not cause the analog transmitter to fold back. However, because the digital side bands are effectively in band for the analog transmitter, one has to consider the inherent lack of turn around loss of the transmitter. At best this turn around loss is of the order of 6 dB. As such the digital side bands will be reflected out of the analog transmitter only 6 dB down and could possibly cause the system to be "out of mask".

In addition, the possibility of intermod products caused by the mixing of the side bands with the analog signal are very real. A total Intermod sweep of the system has been completed with no evidence of spurious emissions detected.

Conclusion

It is critical to perform a complete system analysis prior to recommending any equipment for an IBOC upgrade. The parameter which continually comes up is the system isolation and a very rigorous examination is required every time. Any change in system parameters at any time can cause the isolation parameters to change. Such changes could be caused by antenna icing affecting VSWR, antenna input Isolation, or both. Since there are many different ways to construct an IBOC system, it is critical to analyze every possible scenario. In an ongoing system operation the detection of changing isolation parameters may well be the key to success. The more system points that can be monitored, the better the possibility that events which could lead to a catastrophic system failure, can be pin-pointed prior to disaster.

References

(1) "IBOC Considerations for Multi-Channel Installations". by Henry Downs, NAB Proceedings, 2005

(2) "An Innovative Approach to Common Radiator Implementation of Digital Radio". by John Schadler, Radio World Engineering Extra, October, 2006

(3) "Maximizing HDRadio Signals by Maximizing Isolation". by Henry Downs, Radio Business Report, August 2006.

(4) "Dielectric's HDR Plus Antenna Raises the Bar", by Henry Downs, Radio Magazine, April 2005

(5) "A New Master FM Antenna Design", by William Kerkhoff, NAB Proceedings, 2002

FM- HD- Radio Field Trial Results under European Frequency Planning Conditions

M.A.Ruoss
Ruoss AG, Switzerland

ABSTRACT

FM-Radio in Europe uses different Frequency planning and some transmitting parameters are not the same as in the US. Field- and Lab-Testing by Ruoss AG/Radio Sunshine Lucerne Switzerland shows in an interims Report to the Swiss OFCOM Dec 06, already after the first part of the Work very promising perspectives to use FM-HD-Radio in Europe as easy as in the US, and to go on Air operationally as early as second half of 2009. The Project, funded by the Radio-Industry and supported by the OFCOM will continue during 2007 ,to demonstrate more of the HD-Radio Capabilities with Translators/ Boosters ,RDS-AF Functionality, extensive Indoor testing and a European Set of HD-Radio planning Parameters. Many European Countries started HD-Radio Trials as well, and even Trials in Germany are expected in the first half of 2007.

HISTORY

In respect to digital Radio, Europe is known as (commercially questionable!), successful 100% DAB-Eureka 147 Area and the „image" of IBOC in Europe is very poor. Almost "unproven" Arguments are distributed like "technically not feasible in Europe because of many differences in the FM-System, including mix-up with am-HD-Radio Questions.
Also based on the big DAB-Investments (and Commitments of the BIG-Players and Authorities) , and an European digital FM-System on its Way (DRM+) to be finalized, comprehensible-wise no one is really interested in a "foreign" FM-Digital Alternative.

One should also not overlook, that not even one European Country has a commercially profitable DAB-Operation very soon, with the eventual Exception of UK where there is the Chance to become profitable within some Years from now. (CGap, the biggest UK-Operator predicts "digital break even" for 2010). More so, some thousands of local and regional (one Program-) Broadcasters are not so sure anymore that a Multiplex-Technology is the proper Solution for them, and that maybe more feasible and economically viable systems with a slow evolution path to digital should be found.

That's why the Swiss OFCOM and the Association of Private Broadcasters (VSP), started to support our Initiative to do some HD-radio Field Trials and to follow any other kind of FM-digital System who could be an alternative.

Nowadays this is on the Way to change dramatically, thanks to several testing sites all over Europe; especially in Switzerland where the OFCOM granted the first European HD-radio Test License for 2 Years to Ruoss AG/Radio Sunshine.

Of course, there is no doubt about the fact, which especially DAB+ (new audio codec and new error correction), when introduced very fast, is a very good solution for multiprogramming-Broadcasters and for large area coverage (SFN's).

FIELD AND LAB TEST- SET UP

For the Field Testing the Main Transmitter of Radio Sunshine (88.00Mc, located at around 400 Meter HAAT , with Mountains in approx 10 , 20 and 40Km behind the coverage Area) is used in a High Level Combined Standard Hybrid-Mode with 3500 Watt analog and 35 Watt digital Power. The Main Coverage Area is the Town of Zug and Vicinity. (Radio Sunshine has a total of 14 Transmitters to cover Central Switzerland) See Figure one: Simplified Excerpt of critical Field Test Area.

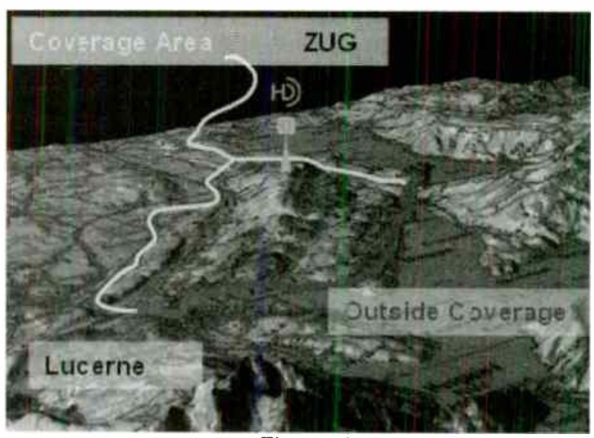


Figure 1

88.00Mc TX ROBE TX-Site Equipment HD-Radio Test-Vehicle LAB-Test SET UP Rotkreuz

The Content transmitted in 24/7 Multicast Mode is:

- Radio Sunshine analog FM
- Radio Sunshine digital (HD 1) 48kbs
- Radio Energy Zurich digital (HD 2) 32kbs
- Service digital (HD-3) 15kbps, Voice: Traffic, Weather, Sport, Events…
- PAD HD-display 1kbs

A second Transmitting Equipment together with other FM- and Test- Equipment Is used for Lab-Simulations and testing.

For the first Part of the Trial a Number of JVC, SANYO and Kenwood Car-Receivers are used together with some Table Tops from POLK and Boston Acoustics. ADA Receivers are used for Reference. 12 FM-Receivers where used for LAB testing, 3 Walk-Radio/Cell phone, 3 Car, 4 portable, 2 Compact HiFi.

DEVIATION AND MPX-POWER ISSUES

The FM-System in the US uses 75uS Pre-Emphasis, a maximum Deviation of 50Kc and MPX Power above 0dBr. Europe uses 50uS Pre-Emphasis, a 75Kc max. Deviation and max. MPX Power Levels ranging from 0dBr to "some dBr" depending on Country and "best practice". Some European Countries only allow a max. of 0dbr MPX Audio Power Other Countries like Switzerland allow +3dBr , and some Countries using even higher MPX-Levels . See Figure 2 for Typical used Peak Deviation and MPX-Power for all HD-Radio (LAB- and Field)-testing.

All our testing so far, shows clearly that „the European Values" of Peak-Deviation/Pre-Emphasis, and MPX-Power Levels (up to 6dBr), do not give any relevant Limitations for the use of the HD-Radio System

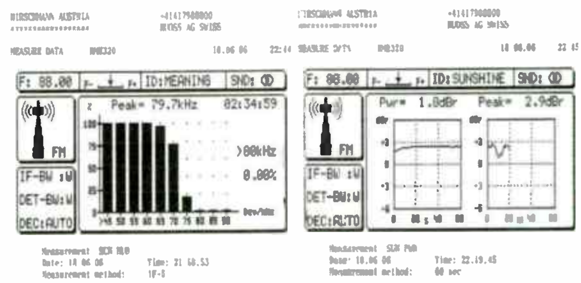


Figure 2

The simple Explanation why HD-Radio works "easy" with European higher Deviation, is, that the HD-System was designed to work with higher SCA carriers like 76 and 96Kc in the us , and that there is a FAC (First Adjacent Canceller) in the System.

HOST COMPATIBILITY

Is the European Receiver Universe different from the US? Not really, but in most Countries in Europe, portable and INDOOR-Listening of FM-Radio is much more important than in the US. Actual Figures from Switzerland shows that approx. 2/3 of all Radio Listening is Indoor /Portable! This means that Portable Receivers has to be tested as well as Car-Receivers. Home- Stereo's are a bit less important nowadays because they are mostly connected to Cable and not to an Aerial (anymore).

Earlier Receiver Studies from Nozema (Netherlands, 1998) and the Swiss OFCOM in the Year 2002 indicated that a greater part of the portable Receivers even are not able to meet the min SNR Requirement and the min RF-Selection Criteria's as it is recommended by the ITU-R BS-412-9, BS.415-2 and ITU-R BS.641 Rules suggest.

Figure 3 shows the SNR Performance (with and without HD-Signal) of a non Representative Sample of typical European FM-Receivers.

The Good News is, that in the last some few Years, the FM-Receivers in the Cell-Phones ,PDA's , PMP's and in the Low-Cost Portables and Walk Radios get very good Receivers, in Respect to SNR and Selectivity built in.

SNR (in dBA) of typical European FM Receivers:

	CAR I	CAR II	Home-Stereo I	Home Stereo II	New Walk-Radio FM/DAB	New Cell phone	Old Mini-Portable	Old Walkradio Portable
FM- Only	**56**	**56**	**69**	**68**	**54**	**53**	**47**	**47**
HD-ON	**56**	**55**	**67**	**62**	**52**	**51**	**46**	**46**

Figure 3

Not even one Complaint from a Listener (> 50 000 FM-Receivers involved on 88.00Mc) about SNR – Reduction from HD-Radio Sidebands, nor a Complaint about RDS-AF Function , what is very important in the Test-Area (more than 4 different Frequencies needed for FM-analog Coverage around Lucerne/Zug)

The SNR- Influence of the 2 digital HD-Radio Side-Bands on Europe's FM-Receiver Universe at the Time of operational Introduction of HD-Radio will be very minimal, and the commonly used planning Criteria on minimum SNR Performance can be met. No any Influence on RDS-AF Function was detected so far.

Extended Hybrid and Multicast

Simple LAB- Measurements and Comparisons with Extended Hybrid Operation did not give relevant differences to the Measurements carried in the US. However, to get best possible SNR-Protection for some of the existing High Performance FM-Receivers, I would actually recommend waiting with the Full-Extended Hybrid Operation till after initial operational Introduction in European Radio- Markets. Looking from marketing perspectives, of course Multicast is a Must right from beginning!

FIELD TEST RESULT

Digital Coverage in the FM-Coverage Area

Within the FM-Coverage Area the digital Coverage for (all digital Content) is perfect and 100% stable with Car Receivers. Examples see in Figure 5. As also can be seen that, coverage Extension happens in Areas where FM analog is unusable (black /blue and green in the Figure 4) because of strong Multi-Path Reception, but with sufficient Field strength and an off Spectrum protection. Figure 4, shows the FM analog Quality (ITU, 5 Grade Scale) of almost the complete Reception Area of the 88.00 Mc Transmitter. Almost all of the Green/Blue and Black Parts of that Picture will have digital Reception with some blending to analog. The Coverage of the Additional Content is smaller than for the Host-Program, because of no blending Feature. This Results are not yet "validated" for complete Indoor Reception in all that Area where Car-reception is good. See also "Indoor portable Reception "

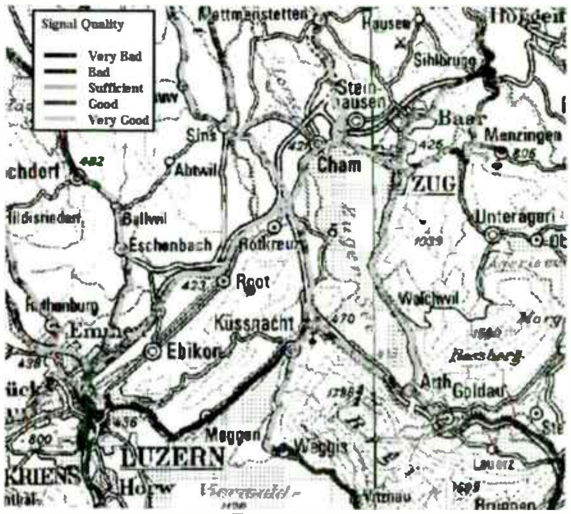


Figure 4

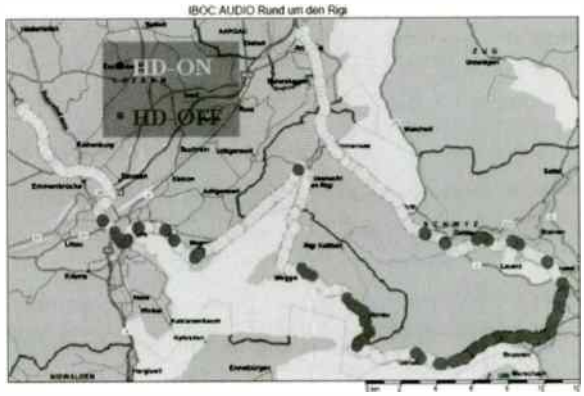
RDS- AF Functionality / Host compatibility

The Main-Transmitter of Radio Sunshine (88.00Mc) is used by more than 50 000 FM-Receivers. In the first 9 Month of Operation with the HD-Radio Signal, no one Complaint from Listeners (on 88.00, or close other Frequencies) was received about additional Noise or distortions. Also there is no any Complaint about the RDS-AF Functionality.

This is a clear Indication that HD-Radio can be introduced on existing FM- Transmitters in a European Frequency Planning Environment.

Indoor portable Reception

Untill now, very little Indoor Reception Experiences are made. Extensive Indoor testing with 40 or more Portable Receivers distributed to Listeners in the fringe Area of the 88.00Mc Transmitter will take place in Part II of the Trial. We already know that compared to the Street- Field strength in the critical Areas, we have around 15 -20dB "Headroom" for digital Reception. Based on the common used Number of 14 dB for FM Building Penetration, we are very confident that the



Indoor-Results will be very promising, with the Exception that we cannot yet quantify the influence of today's typical indoor noise environment on digital reception in the FM-Band. To achieve deep Indoor Reception, approx. 60dBuV/Meter at 2 Meter (commonly used in Switzerland for actual FM planning) will be necessary.

HD-Robustness at High Speeds

Because 88.00Mc is a relatively "well protected" (in Respect to ITU 412-9 Recommendation), additional "artificial" Interferers on both sides of 88.00Mc where needed to simulate digital Reception robustness in fast moving vehicles. Compared to actual Lab-Results some Degradation could be recognized, but at maximum allowed ITU-412 interfering Levels, the digital Reception still is very robust. More "quantifiable" work will be done on this Issue in Part II of the Trial.

LAB-TESTING RESULTS

For LAB-testing the ITU-R-BS-412-9 and ITU-R- BS 641 Recommendation for FM Planning and Measurement is taken as basic Reference. Because FM-HD-Radio Hybrid Operation is a Mixture between an analog and a digital System and as example, because the 2 HD-Sidebands are redundant, some of the recommended Procedures are not anymore useful for HD-Radio Considerations or does not make much Sense. Also today's Station's Sound design produces somewhat other Spectral Densities as at the Time the Recommendations were made. For judging the HD-Radio System we must have different Set of Protection-Ratios for the following Interfering Situations:

D to A: (Digital to Analog)
An HD-Station interferes with an analog only FM-Station
A to D:
An analog only Station's interferes with a HD-Station
D to D:
Interference between Digital Hybrid Stations

For A to D and D to D one will also need different Protection Ratios for stationary/portable and Reception in fast moving vehicles. To get realistic Numbers, both Sidebands have to be interfered at the same Time.

Some subjective evaluation of the Difference between a FM modulated signal interfering an analog FM-Signal and an OFDM HD-Signal interfering an analog FM-Signal at the typical HD-Signal Differences has to take place.
Later there will be also a need for Full Digital Mode Protection Ratios

For the D to A and D to D Measurements, as for the Off Air Signal, Figure 6 shows the typical used RF-Spectrum for Hybrid and Extended Hybrid Operation.

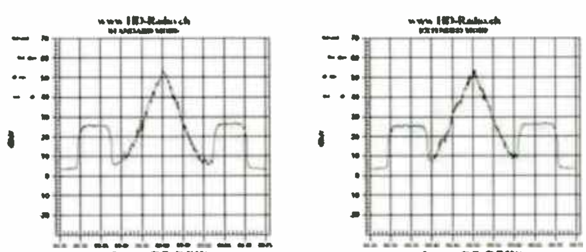


Figure 6

Critical Spectrum Issues?

For the D to A Measurements the same Receivers as for the Host compatibility were used.

For A to D and D to D- Measurements, HD- Radios from Sanyo, JVC and Kenwood was used.

CO-Channel and 400Kc Interfering Situations or not critical at ITU- Limits and the same as in US, so no special additional attention is necessary for this Situations.

Figure 7 shows European 100Kc Interferer at ITU 412-Limit. This Constellation is not critical for D to A, A to D and D to D Interference. The analog Interference is always stronger as Interference from digital Sidebands. Depending on the FM Receiver, more than 10dB "Headroom" was measured on the ITU Limit.

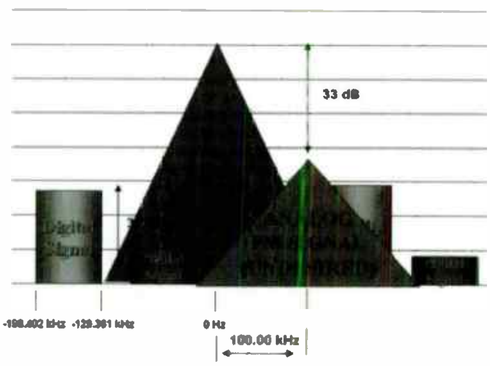


Figure 7

Figure 8 shows the 300Kc Interference Situation at ITU-412-9 Limit. This Constellation is not critical for D to A, A to D and D to D Interference. The "Headroom" to ITU-412-Limits is only "some dB's", depending on Receiver, and symmetrical Interferers, both strong at same Time, can be critical in fast moving Vehicles.

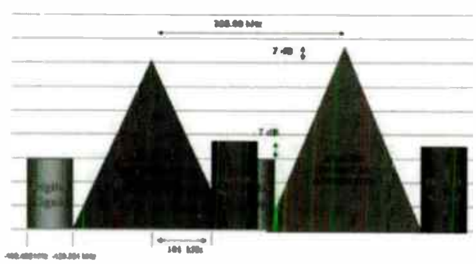


Figure 8

Figure 9 shows the worst possible Constellation for Europe with both Sidebands Interfered with a HD Signal at ITU Limits. This is the most critical Situation in the Field, (same as in the US), and is fortunately not standard practice to happen in the Coverage Area in most European countries. It is not common as well, that if one strong +200Kc Interferer exists that at same Time and Place there is also a strong – 200Kc Interferer.

D to A: as long as the Listener does not use a HD-Radio, the max Stereo- SNR can be reduced well below the recommended Value. In practice, almost all today's CAR-Receivers will blend to MONO (produces approx. 20 dB better SNR) during this kind of strong Interference, and the average Listener will not notice.

Home-Stereo's with individual Arial (does nowadays exists only in rare Cases) will have degraded noise Performance, and "old" Receivers with poor RF-Selectivity (and missing Low Pass after Demod) will have analog Interference before they notice the additional Noise from the OFDM Signal.

Therefore it is recommended (to protect the existing FM Receivers) that special Attention has to be taken for 200Kc Interferer Situations, especially when both sidebands are affected.

A to D and D to D Interference at ITU-Limits will still produce stable HD-Radio reception for the wanted Signal.

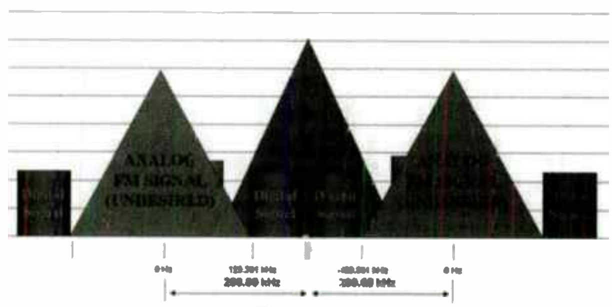


Figure 9

New FM-HD-Radio Planning Rules?

If the ITU-R- 412-9 Recommendations are respected in a way as it was "common practice" for the earlier FM-Planning in Europe, and the estimated near future average FM- Receiver-Performance is used, no new Rules are necessary to implement HD-Radio in Europe! Following Exceptions may apply:

- avoid symmetrical, (and same Time /same Location) 200Kc Interferers in the Coverage Area
- try some Headroom at 200 Khz Frequency Difference (average 7dB would be great)
- Use 60dBuV /Meter at 2 Meter at Edge of Coverage Area
- Correct eventual earlier "non conforming 200Kc Interferers" in the Coverage Area

In Europe, the same Interferers (+/- 200Kc) are the HD-Radio System-Limit as in the US! And not as told around for Years 100 and 200Kc !

Figure 10 shows typical LAB-Test Configurations as they were used for preliminary Protection ratio measurements.

Receiver Sensitivity

Coming as no big surprise, all tested Car HD-Radio Receivers has very good Sensitivity and very good FM-analog Performance as well. HD-Radio uses same proven FM-Radio RF-Front-end as it is known for very long Time. The Car Receivers we measured so far works all in digital Mode a bit below 30dBuV at the Receivers Antenna Input. The first Generation Tabletop's are some dB less sensitive

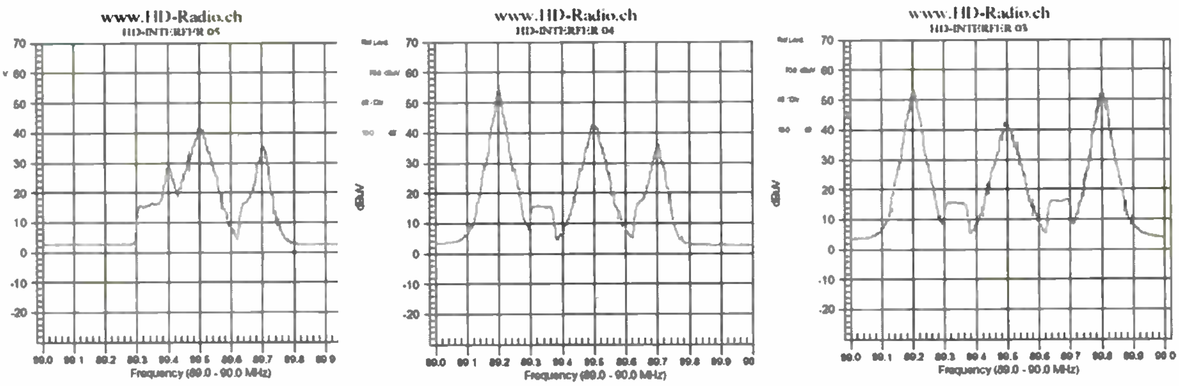


Figure 10

How selective are today's FM-Receivers?

Unfortunately not all Receivers in the actually existing FM-Radio Universe have the same performance when it comes to selectivity. As Figure 11 shows, there are dramatically big differences between Receivers.

As Example at 200Kc Frequency Difference, a relevant part of the Receiver Universe (BAKOM Study 2003) is more than 20dB worse than the ITU Recommendation (portable /mobile), but another Group of Receivers (especially Car-Receivers) is up to more than 50dB better than ITU-R is looking for.

The good News is, that with the Exception of 100Kc, the actual FM-Receiver-Chips built in, in cheap portable Receivers and Cell phones are much better as some Years ago, and are reaching ITU 412-9 min. Performance.

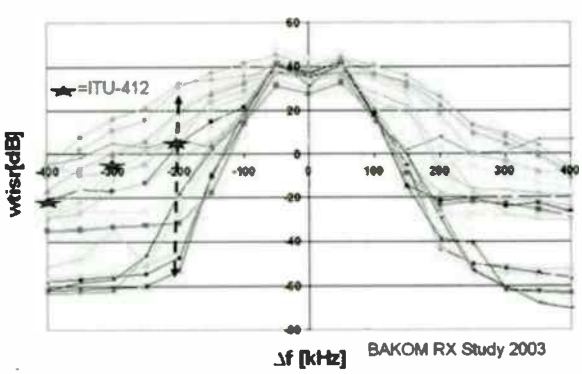


Figure 11

HD-Radio an (ETSI) Standard?

The initial Steps have already been taking to make HD-Radio a European ETSI (European Technical Standard Institute) Standard. An Important Part of the is Process will be the inclusion of the European RDS Standard with the AF-Functionality (Alternate Frequencies),because in contrary to the US , an average Private Broadcaster has a high number of Translators and not just one Main Transmitter.

OTHER FM-HD-Radio TRIALS IN EUROPE

HD-Radio's growing popularity in Europe has led to some "HD-Radio Tourism" to our Location in Central Switzerland for a Test-Drive with one of our Test-Vehicles and a Laboratory Demonstration of all critical Protection Ratio Issues. It is a Fact, that almost all "critical Visitors" (only had "heard of", and "read about" HD-Radio in US) becomes real "Fan's" of the Digital FM Idea after the Demonstration in the Field and in the LAB.

Some of the leading OEM's for the German Automotive Industry have started to use central Switzerland is the Test-Place for there HD-Radio Equipment.

Even after the Announcement by the World DAB-Organization about the soon arrival of the far more efficient new DAB+ (with AAC+ Audio Codec, witch is very similar to HDC used by HD-Radio, and the

better Error Correction) the European Interest from Regulating Bodies and Private Broadcasters is still growing very strong.
In France (Towercast and NRJ-Group, backed from Sirti) doing intense testing since April 2006 in the Paris Area with 2 HD-Transmitter on-air, in a critical "European frequency Assignment". Poland and Ukraine are actually also running some HD-Radio Tests.
Actual new plans for HD-Radio Trials and Interest in HD-Radio (we know of at time of writing) in Europe includes:

- Austria (Krone-Hit-Vienna and others)
- Italy (RVR with some Broadcasters in north Italy, Aeranti Corallo-Radio Association)
- Germany (several independent projects ,)
- Romania
- Czech Republic
- Bosnia

THE KEY ADVANTAGES OF FM-HD-RADIO™

HD- Radio can become the "optimum" Path for European "one Program Broadcaster" to the digital Radio Age. The Advantages are obvious:

- Same frequency and same basic FM-infrastructure
- Additional services possible on the same frequency
- Better quality of service in the coverage area
- Backwards compatible with FM (simple communication!)
- Slow evolution without much cost for the broadcaster
- EMC Issues practically non-existent
- Good indoor and portable reception achievable
- Less Translators/Boosters as for analog only Operation

NEXT STEPS

Even with the very positive Results so far, more detailed Investigations in the Lab and in the field are necessary before Regulating Bodies can start issuing operational Licenses for HD-Radio. Following the major steps in our Trial Project till end of 2007:

- Comparison between FM and HD-Radio™ „indoor" coverage
- Support for OFCOM Lab-Measurement for HD-Radio Protection Ratio for all major operating Constellations and all major Receiver Segments.
- Consequences (if there are any) on FM-planning for FM-HD-Radio- implementation in Switzerland (and Europe), comparison with USA and ITU-412-9 recommendation
- Translators and SFN-Boosters for HD-Radio in the field (and RDS-AF Functionality test)
- Implementation and operation costs in Comparison to FM operation , and other digital Platforms
- Implementation Comparison with other digital FM-Systems like DRM +
- Support of timely creation of operational HD-Radio operational licensing Guidelines
- More general awareness in Europe for FM-HD-Radio , including at Receiver Manufacturer Level , for "Multiplatform Receivers"

The Presentation of the Progress and the Results will be presented during the second HD-Radio days in Lucerne, Switzerland on 4./5. October 2007. To stay informed on the "Work in Progress", see WWW.HD-radio.CH . The Final Report for the OFCOM is planned for December 2007. An operational start up of HD-Radio is estimated for approx. 3./4. Quarter 2007.

ACKNOWLEDGEMENTS

Many Thanks to all Sponsors and Supporters of the HD-Radio Field Trial in Switzerland and for the Help for this Paper. January 2007 M.A.Ruoss

Gold sponsors

Supported by the Federal Office of Communications OFCOM

Silver sponsors

Bronze sponsors

A Precision, Low-Cost GPS-Based Synchronization Scheme for Improved AM Reception

Stephen F. Smith and James A. Moore
Oak Ridge National Laboratory
Oak Ridge, Tennessee

David W. Allan
Allan Space-Time Solutions
Fountain Green, Utah

ABSTRACT

This paper describes a highly accurate carrier-frequency synchronization scheme for automatically locking multiple, remotely located AM broadcast transmitters to a common frequency/timing reference source such as GPS. The extremely tight frequency lock (to typically ≤ 1 part in 10^9) permits the effective elimination of audible and even sub-audible beats between the local (desired) station's carrier signal and the distant co-channel stations' carriers, *usually* received via skywave propagation during evening and nighttime hours but *also* present in daytime fringe areas. These carrier-beat components cause annoying low-frequency noises and modulations of the desired station's audio at the receiver and concurrent distortion of the distant station(s)' audio, often causing listeners to "tune out" due to the poor reception quality. Significant reduction or elimination of the beats and related effects can significantly enlarge the effective (interference-limited) listening area of the desired station (***from 4 to 10 times*** as indicated in our laboratory listening tests) and *simultaneously* reduce the corresponding interference of the local transmitter to the distant stations. In addition, HD reception will clearly benefit via the reduction in beats from analog AM signals. The automatic frequency-control hardware described is inexpensive (<$2K), reliable, adjustment-free, and easy to install; even older transmitters can be easily adapted.

INTRODUCTION

Generally the major sources of interference to broadcast stations are the co-channel signals from stations in adjacent areas of the country. In traditional analog (NTSC) television, the co-channel distant video carrier (which is amplitude-modulated as well) causes the generation of diagonal interference patterns ("herringbone") in the picture, plus often a visible, horizontally oriented "sync-bar" which very slowly drifts vertically in the received display. To minimize the severity of these effects when the two carriers were very close to the same frequency, the FCC long ago mandated that geographically adjacent co-channel TV stations offset their carrier frequencies from each other by roughly 10 kHz, to keep the magnitude of the resulting interference pattern to fringe-area viewers acceptably low. However, this scheme is practical only because of the considerable width (6 MHz) of the TV channels; no such separation or offset scheme is feasible for AM radio stations, whose channel allocations are only 20 kHz in total width. Traditionally, the FCC has required that AM broadcast stations operate within ±20 Hz of their assigned carrier frequencies, which theoretically permits difference-frequency beats between co-channel carriers of up to 40 Hz. Typically, the stations operate at much smaller frequency errors; beat frequencies are generally now below 5 Hz, and are sometimes even below 1 Hz. Unfortunately, in areas that are located far from the local transmitter (i.e., the desired station), distant-station carrier-beat components frequently cause fast fluttering-type modulations of, and/or large-amplitude swishing sounds in, the desired station's audio at the receiver and concurrent distortion of the audio modulation by the more distant station(s), even if they are sub-Hertz in nature, since the received desired-station's carrier's amplitude and phase are noticeably modulated by the distant station's carrier signals. Modern-day AM radio receivers are provided with automatic gain-control circuitry to level out signal-strength variations over time. However, the typical AM radio receiver's automatic gain-control (AGC) [also called "automatic volume control" (AVC)], to avoid bass distortion, usually will respond far too slowly to "average out" or suppress these inter-carrier beat modulations; thus, these highly annoying modulation effects are largely passed on intact to the listener.

There are two distinct AM interference scenarios to be considered — day and night. During daytime hours, the ionosphere is absorptive and the dominant interference is from adjacent co-channel groundwave signals. Generally, an AM radio listener during the evening and nighttime hours, and to a lesser extent in the early morning, receives undesired skywave signals from several distant stations as well as the desired local (groundwave) signal. If the carrier frequencies of *all* of these signals are maintained to within about 0.01-0.001 Hz of each other, any resulting carrier beats will be of such long periods that the beats can be effectively suppressed by the action of the receiver's AGC/AVC circuitry and will thus be unnoticeable to the listener. For best results, all stations on the channel in question (at least those with signals above the noise floor at the receiver) should be closely frequency-locked to a common precise reference as just described, or the beats will not be completely eliminated [1].

OVERALL SYSTEM IMPLEMENTATIONS

It is useful to examine how the phases/delays of the audio and RF components of the AM radio signals can affect reception quality in the field, particularly in signal-overlap regions. For instance, the RF signal delay is very roughly 1 millisecond for 186 miles (corresponding to the speed of the propagating signal). At a point equidistant from two omnidirectional, co-phased (synchronous) transmitters with equal power and modulation, propagating via groundwaves over land paths of identical RF conductivity, the two RF signals will arrive with equal amplitudes and delays (phases). Now if we assume that the RF carriers and the sideband audio signals are precisely in phase (matched in time) as they leave the two antennas, at the exact midpoint between the two transmitters the RF signals and the detected audio will also be in phase; the signals can be added algebraically to calculate the resultant. Now for points *not* equidistant from the two transmitters, the RF signals will vectorially add; in general, there will be augmentations and cancellations of the two waves occurring at spatial intervals of one-half wavelength, essentially the same as is the case for standing waves on a mismatched transmission line. Obviously, near the equal-signal points, the standing wave patterns will exhibit maximum variations; in fact, §73.182(t) of the FCC Rules defines the region of "satisfactory service" for synchronous stations as areas where the ratio of field strengths is ≥ 6 dB (i.e., ≥ 2:1). In contrast, the focus of this paper is ***not*** this classic "synchronous broadcasting" case, but the much more practical (*and economically significant*) scenarios involving signal quality near the fringes of a local station's coverage area, where the disparity of desired-vs.-undesired signal levels (e.g., 12 dB or more) will reduce this localized "standing-wave" modulation effect to negligible levels.

From the standard FCC propagation data, at the nominal fringe signal level of 0.5 mV/m (for all Classes of stations except A, defined as 0.1 mV/m), the daytime, groundwave co-channel signals [re 47CFR§73.182] must each be no more than $^1/_{20}$ the amplitude (–26 dB) at the stated field-strength contour [or 25μv/m, (5μv/m for A)]. The same corresponding nighttime values of acceptable co-channel interference levels (–26 dB) are specified for Class A, 0.5 mV/m (50% skywave) contours and the 2 mV/m contours for Class B (groundwave). No such protection constraints exist for Class C and D nighttime operation, where average interfering skywave levels may be as high as 1.25 mV/m! [2] Allowing for finite ground conductivities, it is evident that an improvement of 6 dB in effective co-channel levels will ***nearly double*** the interference-limited contours of the stations compared with the standard, non-synchronous case (*almost quadrupling the equivalent coverage area*). As will be described later, our simulations with real broadcast audio demonstrated that for some types of programming (i.e., with good audio masking properties) the effective improvement can even approach 10 dB, which could nearly triple the interference-limited coverage range!

With the beats eliminated, the background audio from the co-channel stations will be clean; often, the so-called "cocktail party" effect will reduce the apparent level of those signals to the listener even further, especially in high-background ambients such as automobiles. The net result of these effects will be universally evident but particularly beneficial to nighttime operations at local Class-C and Class-D stations, whose coverage areas are already acutely curtailed by heavy co-channel skywave interference. For these latter classes, the near-quadrupling of effective coverage at night should be a major benefit, particularly to listeners in outlying suburban areas.

TX SYNCHRONIZER HARDWARE SYSTEMS

The basic configuration of the GPS-disciplined oscillator is extremely simple, as shown in the accompanying diagram of Fig. 1. The basic oscillator may be a more-or-less conventional high-stability quartz crystal type, usually ovenized or otherwise temperature-compensated. To counter long-term drifts, aging effects, and loading-circuit changes, the basic oscillator is ***slightly*** adjusted via electronic (e.g., varactor diode) or mechanical means (e.g., turning a capacitor-adjustment shaft) to track a high-precision source of standard frequency obtained from a specialized GPS satellite receiver or other source, often at either 5.000 or 10.000 MHz. This very stable local reference frequency is then used as a clock for a standard digitally implemented frequency synthesizer, which is programmed to generate the specific (AM broadcast) transmitter carrier frequency desired. The stability of the disciplining source, typically ~ 1 part in 10^9 to 10^{11}, is thus transferred to the final AM transmitter carrier output frequency. In Fig. 1, the multi-channel GPS satellite receiver outputs a standard reference frequency, generally either 10.000 or 5.000 MHz (or a 1-Hz pulse) in commercially available units. The output at 10.000 MHz is applied to a synchronous digital divider chain to produce the desired phase-comparison frequency (and, thus, step size) of the overall phase-locked loop (PLL) type of frequency synthesizer illustrated to develop the precision AM transmitter carrier frequency source. In the example embodiment, three successive stages of divide-by-10 are cascaded to provide the overall ÷1000 frequency-division ratio to achieve the desired 10.000-kHz phase-comparison reference frequency. The dividers and phase comparator (detector) can be implemented with any suitably fast circuit technology (i.e., TTL, HCMOS, etc.). Likewise, the "programmable divider" block is optimally implemented with high-speed, bipolar-transistor emitter-coupled logic (ECL) or high-speed

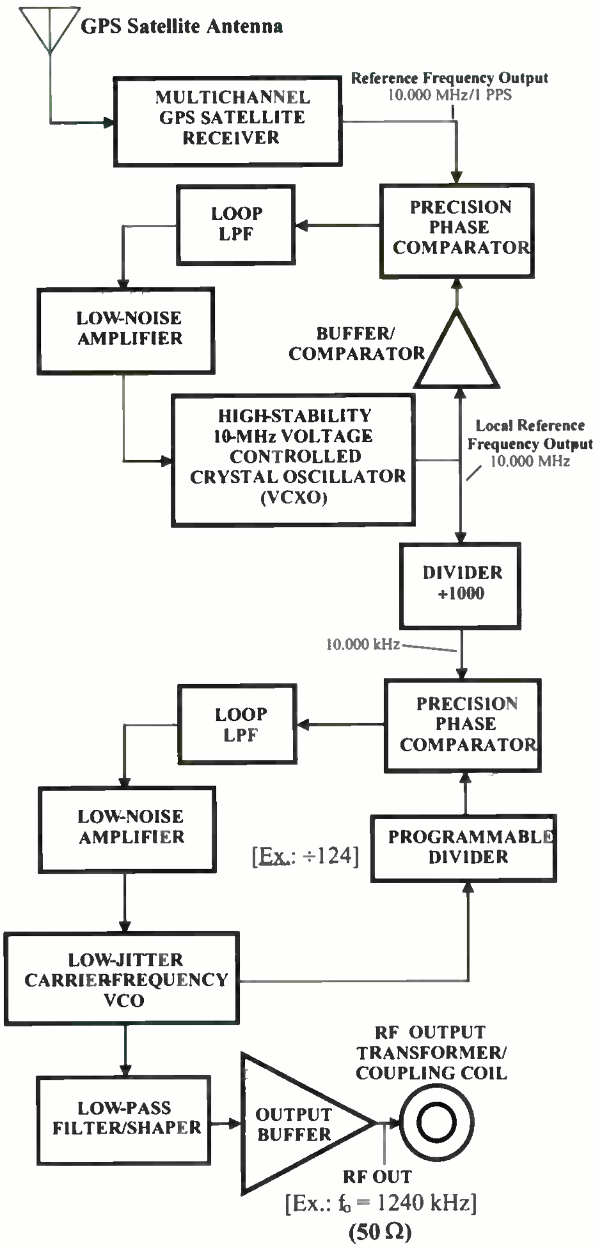


Fig. 1 — Basic GPS-disciplined oscillator system.

CMOS circuits, which are presently far more prevalent in the industry. As shown in the drawing of Fig. 1, this stage is selected to have the ratio needed to divide the specified carrier frequency (here, for example, 1240 kHz) down to the desired phase-comparison frequency of 10 kHz; thus, here the ratio is set to 124. [Obviously, for international AM broadcasting applications, where stations may be allocated at intervals of 5 kHz, 9 kHz, etc. (all multiples of 1 kHz), the phase comparator would probably operate at a frequency of 1 kHz to achieve this smaller frequency-step size and the two divider chains would each be modified to raise their overall ratios by a factor of ten. The loop low-pass-filter (LPF), low-noise amplifier, and high-stability voltage-controlled oscillator are selected to achieve low-noise, low-jitter performance of the locked loop, as would be expected from the high-accuracy, high-stability requirements of the application. The currently available versions of high-stability ovenized quartz VCXOs (used here as the local reference) rival the much more expensive rubidium- and even cesium-based units traditionally used for such applications. Closing the upper loop, a fast buffer/comparator stage squares up the usually sinusoidal output of the main oscillator and develops logic levels suitable for driving the logic circuitry. The ÷1000 fixed divider is used to divide the reference frequency (here, 10.000 MHz) down to the common 10-kHz step size, and the final transmitted carrier frequency is synthesized in a conventional integer-N PLL configuration via a low-jitter VCO, often realized via a standard square-wave oscillator, which drives a switched-integrator stage to produce a triangular wave at the same frequency, which in turn is soft-clipped (shaped) to a very nearly sinusoidal waveform to yield the final low-distortion output carrier frequency. The principal advantage to this approach is the elimination of the traditional custom-manufactured carrier-frequency crystal; as a result, the design is completely standardized. An internal programmable switch or logic device selects the specific final RF carrier-frequency value. A second variation of the overall precision-control scheme using a direct-digital synthesizer (DDS) device is shown in Fig. 2, where the GPS-derived reference frequency is *multiplied* to serve as the high-frequency clock (here, 200 MHz) for a fast digital numerically controlled oscillator (NCO) circuit, such as that included in the AD9854 chip from Analog Devices, Inc. (ADI) or in numerous other, similar chips from ADI and several other manufacturers (e.g., Intel, Intersil, and Philips). Here, the main NCO accumulates digital phase at its clocking rate; the phase data is presented continually to a phase-to-amplitude data converter [read-only memory] ("Sine-ROM") block, which outputs a series of amplitude values corresponding to a sine-wave at the programmed frequency. These data words, typically at resolutions of from 8 to 14 bits, are then fed to a high-speed digital-to-analog (D/A) converter, which generates the stepwise analog carrier output signal. A final downstream low-pass filter/wave-shaper stage smooths the output signal and reduces distortion and spurious components to negligible levels; the final RF output amplifier boosts the signal power to a level adequate to drive the associated RF transmitter circuitry. For ease and reliability in interfacing the reference signal into an (older) existing transmitter's oscillator stage (operating at 1×, 2×, or 4× the carrier frequency), an inductively coupled drive loop is typically employed around one of the leads to the existing crystal to injection-lock the transmitter crystal to the external reference. Then, if the reference signal fails for any reason, the transmitter will simply "slip" back to its original frequency, without any interruption in its

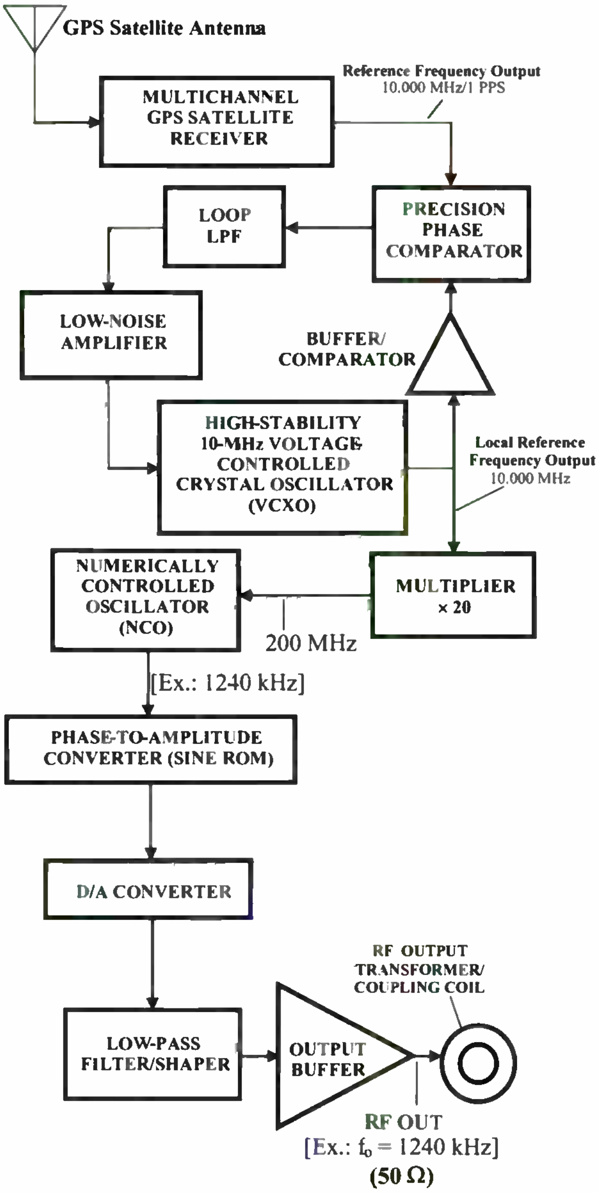


Fig. 2 — Basic GPS-locked DDS Oscillator System

broadcast transmission. Alternatively, the GPS-disciplined master reference frequency (at typically 5.000 or 10.000 MHz) could be used as an external reference for the frequency synthesizer generally employed in more modern AM transmitter designs. The stability of the disciplining source, typically much better than ~ 1 part in 10^9, is thus directly transferred to the final AM transmitter carrier output frequency.

The projected implementation for the synchronization hardware package consists of three basic components: the GPS timing receiver; its external antenna; and the rack-mounted synchronizer unit. The synchronizer chassis will probably provide at least one 10.000-MHz output for driving the external inputs of newer synthesizer-based transmitters, plus one or more selectable outputs to provide oscillator reference signals at 1×, 2×, or 4× the carrier frequency for most older rigs. Using modern low-cost GPS receiving hardware and efficient microcomputer/FPGA-based designs for the synchronizer unit, we anticipate that the final price for the AM broadcaster would be in the \$1K - \$2K range. The fundamental block diagram of the unit will most likely resemble that in Figs. 1 or 2.

The basic architecture of Fig. 1 incorporates a classic integer-N PLL-based synthesizer, but other types could be used, so long as the final RF carrier frequency is *precisely* (i.e., within at least 1 part in 10^{10} of) the ideal value. Most fractional-N and direct-digital synthesizer (DDS) type systems, as in Fig. 2, actually generate sets of discrete output frequencies, based on the input reference clock and the internal logic divide ratio. The effective frequency resolution (in Hz) is equal to the reference clock frequency (f_{ref}) (in Hz), divided by the binary exponent of the accumulator or frequency-control word width of the DDS device. For the overall synchronization system to perform optimally, we would like to keep all individual beat frequencies comfortably below 1 mHz (with corresponding beat periods exceeding 1000 seconds); thus each individual transmitter should exhibit a frequency error well below 10^{-3} Hz (ideally, closer to 10^{-4} Hz). This figure is consistent with earlier studies by NBS/NIST of ionosphere-induced frequency shifts in AM skywave signals for standard-frequency stations WWV and WWVH [3]. Now for a DDS with a 10.000000-MHz clock and a 32-bit accumulator, the minimum step size is $10^7 \div 2^{32} = 2.3283\ldots \times 10^{-3}$ Hz, which is greater than the desired local frequency error of $< 10^{-4}$ Hz. Some quick arithmetic shows that in fact an accumulator width of 36 bits is very marginal to maintain the desired accuracy; 40-48 bits would be much better. On the other hand, by choosing a DDS reference clock of 16.777216 MHz (precisely 2^{24} Hz), we can go as low as 24 bits in accumulator width and still achieve the *exact* desired 1-Hz step size (and, thus, zero frequency error). Obviously, there are many other possible implementations of the synchronizer unit, so long as the fundamental frequency-error constraints are met.

AM TX SYNCHRONIZATION LAB TESTING

Referring to Fig. 3, the effectiveness of the synchronization concept for AM broadcasting was demonstrated with a laboratory test setup. To receive local-area broadcast signals as convenient multiple peak-limited audio sources, an external rooftop FM antenna was connected to a 3-way splitter, which in turn fed three different FM broadcast tuners. Each of the tuners was coupled to an individual audio level-adjust amp and thence to an externally modulated frequency-synthesized AM signal generator. A low-frequency function generator was also employed to phase-modulate one of the signal generators. The "main" signal generator was directly coupled to a 3-

way combiner, while the other two "interfering" signal generators were each coupled to the 3-way combiner via a variable attenuator. The output of the 3-way combiner was fed to a high-quality AM tuner via a fixed attenuator; the tuner then fed the lab loudspeakers via a commercial component-grade high-power stereo audio receiver used as a monitor amplifier. Still referring to Fig. 3, the laboratory test setup

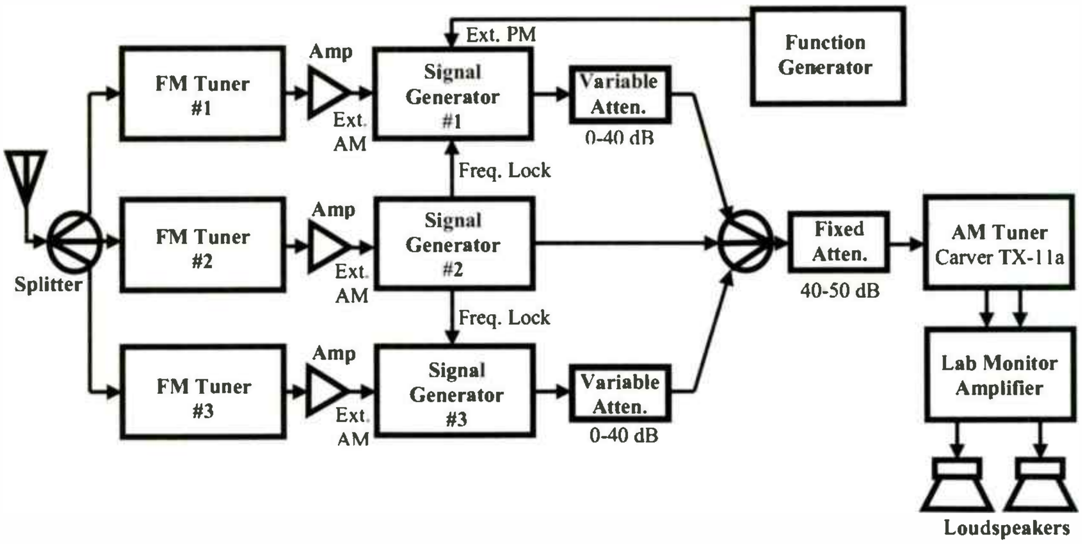


Fig. 3 — Laboratory AM Synchronization Testing Setup

employed three externally modulatable signal generators. The "B" channel of a high-performance synthesized two-channel signal generator, a Hewlett-Packard model 3326A, was used as the reference ("local") AM station; the "A" side was set up as one of two "interfering" co-channel stations. The other "interferer" was simulated by a separate single-channel synthesized H-P 3325B generator. All three generator units were commonly locked to a high-stability ovenized 10.000-MHz clock within the 3326A chassis. For the test, all units were nominally set to 1000 kHz, as was the high-quality Carver TX-11a audiophile AM Stereo/FM Stereo tuner which was used to demodulate the combined signals and feed an external lab monitoring amplifier/speaker system. The generator outputs were all set at a nominal 0-dBm output level (50 ohms) and fed to a three-input passive RF power combiner. The combiner's 50-Ω output was attenuated about 40-50 dB before being fed to the RF antenna input of the high-quality tuner, which itself was terminated and de-sensitized by a 50-Ω resistor.

Most of the interference tests were initially conducted with standard monaural audio modulation, although some subsequent trials included the "main" signal operated in CQUAM analog stereo to test the overall effect of synchronization on stereo reception. Three consumer-grade FM broadcast tuners were set to three different Knoxville/Oak Ridge, TN area stations to obtain reasonably clean, peak-limited audio sources to modulate the three AM generators via their external-AM audio inputs. The main generator was fed directly to the AM signal combiner, while each of the two "interferers" was connected through a pair of switchable step attenuators, one with 10-dB steps and the other with 1-dB increments. Many hours of careful subjective listening were conducted, with the two interfering units both precisely on-frequency with the main unit (synchronous operation) and with the two interferers at various frequency offsets, from below 1 Hz to above 10 Hz. The most audibly annoying beats were generally judged to be below roughly 2 Hz, so several tests were conducted with offsets of 0.7 and 1.7 Hz, respectively, which tend to more closely emulate current typical AM broadcast channel beat characteristics.

Referring to Fig. 4, subjective measurements to determine the familiar audible interference assessment criteria of "imperceptible", "perceptible", "annoying", and "objectionable" were made and documented. An extra sub-category within the "perceptible" bracket, denoted "long-term-listenable" ("slightly perceptible"), was added to assist in delineating a listening quality level adequate to avoid "tune-outs" caused by carrier beats and other background interference. For this test campaign, impulse noises such as from lightning were deemed out of the scope of this effort and were therefore not included. To help simulate significant nighttime ionospheric phase/propagation-delay shifts, in both synchronous and non-synchronous scenarios, an auxiliary low-frequency function generator was added to externally phase-modulate the larger interferer's signal at rates varying from 0.5 Hz to below 0.01 Hz (a 100-second period); the usual periods of fading were from 10 to 30 seconds, following past experience of the authors. Obviously, the interference in the standard (non-synchronous) case consists of both carrier beats and background co-channel audio modulations and beats, while in the synchronous setup the carrier beats are absent. (Asynchronous sideband beats, as mentioned by Kahn [4], are highly variable, at much lower levels, and are thus a *very* minor factor). At low signal-to-interference ratios (SIRs), i.e., below about 12 dB, the unsynchronized beats are highly annoying and cause gross intermodulation of the received audio program. At much higher SIRs, approaching 30 dB, even the non-synchronized beats become imperceptible to most listeners. At *intermediate* SIR levels, the effects are strongly dependent on the exact nature of the three audio modulation signals, due to masking of lower-level sounds by louder concurrent ones in the human auditory system. Slightly different results are therefore obtained for fast ("pop") to slow music [M], average voice [V], fast-paced voice [FV] (e.g., commercial advertisements) and slow-tempo voice [SV] (e.g., telephone-talk shows). Overall, the net advantage to the listener of synchronizing the AM carriers and thereby eliminating the beats is on average about 6 dB minimum and can often be as great as 10 dB; this is of major importance in evening and nighttime situations where the SIR due to incoming skywave signals can degrade to levels of 12 dB or worse. Thus for comparable audibility, with synchronization interfering signals from co-channel stations may be from 6 to 10 dB higher than in the current non-synchronized scenario. The net result is that a listener may be at a position where the SIR is 6-10 dB worse (i.e., at roughly *twice to three times* the distance from the desired station) before the interference becomes annoying, as compared with the present situation. Even at an 8-dB SIR, when the carrier beats from unsynchronized stations are totally objectionable, interference from well-synchronized remote signals can be masked surprisingly well by highly modulated audio from the stronger ("local") station and thus render a reasonably listenable signal. For those interested, an audio CD from these lab tests is available from the authors.

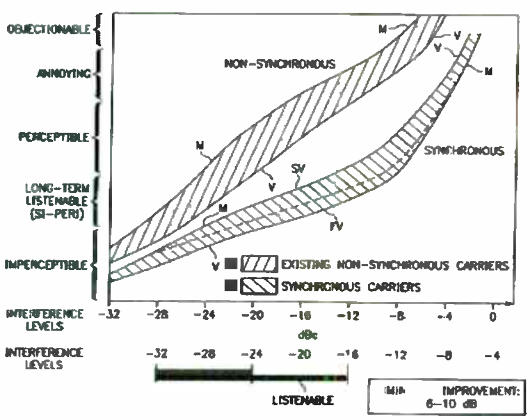


Fig. 4 — Audibility of Synchronous vs. Non-Synchronous Interference

MOBILE DOPPLER EFFECTS

Referring to Fig. 5, field contours of two overlapping synchronous (co-channel) AM transmitters are depicted with typical mobile-receiver trajectories. For a listener traveling in a path connecting two 1.0-MHz transmitters at a velocity of 30 m/s (about 67 mph), as shown on trajectory V_1 of Fig. 5, the signal minima will appear at an interval of (150 m) ÷ (30 m/s) = 5 s, or at a rate of 0.2 Hz. For the typical automobile AM radio receiver with an effective automatic gain-control (AGC) loop bandwidth of roughly 5 to 10 Hz, this 0.2-Hz variation will be easily leveled out by the AGC loop so long as there is sufficient residual gain in the receiver to amplify the resultant signal at these partial null-points. For velocity vectors at other angles from the transmitters, the standing-wave amplitude-oscillation intervals will be longer by a factor of the inverse of the cosine of the angle from the direct or radial path; for instance, if the path is at 45° from the radial, the oscillation period will be (5 s) ÷ (cos 45°) = 5/0.707 = 7.07 s. For example, on trajectory V_2 of Fig. 5, the receiver is moving tangentially with respect to Station 1 but radially toward Station 2. Thus there will be no relative shift in the receiver of signal 1, but the maximum Doppler-frequency shift (for that velocity) of +0.1 Hz will be observed for signal 2. For paths which are neither tangential or radial to either transmitter, such as V_3 of Fig. 5, both shifts will be present but will be of intermediate frequencies (i.e., $|f_{beat}| < 0.1$ Hz). For velocity paths essentially perpendicular to the radial (cos 90° = 0), as on path V_4 of Fig. 5, the oscillations will be of very long periods and of vanishing amplitudes. Although theoretically there will be cancellation nulls of great depth at many discrete points in the equal-strength areas of the two transmitters, in reality the nulls will only be partial, due to local variations in the amplitude and phase of the transmitted ground waves, scattering effects, and diffraction over hills and other large terrain features. If the receiver gain range is sufficient to compensate for a null depth of about 20-30 dB worst case, the audible effects will usually be modest, except for some distortion in the nearly-equal signal zones (reddish area).

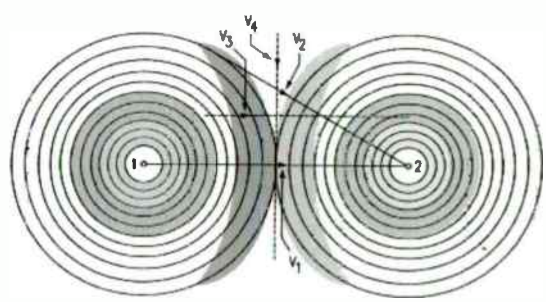


Fig. 5 — Plot Illustrating Doppler Effects in Mobile Receivers for Synchronous Stations 1 & 2

Another way of describing this effect is that of Doppler-induced carrier-frequency shifts in moving (vehicle) receivers. For the aforementioned scenario, where a vehicle is moving at 30 m/s on a linear path between two frequency-locked stations (path V_1) at 1.0 MHz, the apparent frequency of the transmitter being approached will be shifted *upward* by $(3 \times 10^1 \text{ m/s}) \div (3 \times 10^8 \text{ m/s}) \times (10^6 \text{ Hz}) = 0.1$ Hz, while the transmitter being receded from will exhibit an apparent *downward* shift of 0.1 Hz. The net result will be an observed offset of 0.2 Hz between the two carriers, with the attendant beat. The effective beat frequency with respect to each station will thus be a function of the product of vehicle speed, the cosine of the velocity-vector angle with the radial from the station, and the true transmitted RF carrier frequency. Equivalently, the Doppler shift from each station (**1** or **2**) is the product of the radial component of the relative vehicle velocity (referred to the specific transmitter) and the carrier frequency, divided by the speed of light (or RF) in the atmosphere. The total composite beat frequency is simply the sum of the two Doppler shifts (for both **1** and **2**). In general, the path on the common radial will produce the worst-case (highest) beat frequency; ideally, a path perpendicular to that line will produce no beats, since the Doppler shifts of each station will be equal in magnitude *and* sign and thus cancel at the moving receiver. The mathematical equations are thus:

$$f_{beat(total)} = \sum_n f_{beat(n)} \qquad \textbf{(1)}$$

$$f_{beat(n)} = (v_{Rn} \cos \theta_n)(f_0/c) \qquad \textbf{(2)}$$

where $f_{beat(n)}$ is the nth beat frequency in Hz, v_{Rn} is the velocity in m/s relative to station n, θ_n is the angle of the trajectory from the radial from station n, f_0 is the original carrier frequency in Hz, n is the number of received co-channel stations, and c is the speed of light in m/s. Thus the combined beat signal is merely the sum of the Doppler frequency components due to the relative radial velocities with respect to each station, times the inverse of the nominal RF wavelength.

A significant consequence of this Doppler effect is the *very* low-frequency beat-modulation of the audio envelope in mobile receivers (none of this occurs for fixed radios), but several factors ameliorate the situation in real vehicular listening environments. First, the apparent modulation from near 0 to 0.3 Hz (typically less than 0.2 Hz) is largely suppressed by the action of the radio's internal feedback AGC circuitry, which rapidly and effectively levels these slow amplitude variations to maintain a fairly constant detected carrier magnitude. Second, the presence of relatively high levels of ambient "road noise" in the vehicle at higher speeds, particularly in the low-frequency region of the audible spectrum, serves to mask these cyclic but low-level variations. Third, local RF field irregularities also cause overall level variations which "dither" (randomly modulate) these cyclic field variations; these variations also tend to mask the beats. When the vehicle slows and thus produces less road noise to mask the beats, their frequencies drop to negligible values and generally fall below audibility. Finally, the dynamic nature of most types of music and voice broadcast programming also inherently tends to aurally mask these very low-frequency components. Obviously, the magnitudes of the beats will be dependent on the relative amplitudes of the two co-channel signals being received; for most areas, where the signals are at least 10 dB different in level, the resultant beats are very weak. ***Thus, the bottom line is that these Doppler effects are overall very minor.***

Compared with the standard static-receiver synchronous AM reception case discussed previously, the presence of these sub-Hertz Doppler beats in *mobile* listening environments typically causes a degradation (i.e., increase) in the overall beat audibility of only about 1-2 dB compared with the curves in Fig. 4. *It is important to understand that virtually all of the major benefits of synchronous co-channel station operation are still retained even for the mobile listener.*

SUMMARY AND CONCLUSIONS

A low-cost synchronization scheme to minimize beat-related interference among co-channel AM stations has been discussed. Although the general technique of *local* synchronization has been known and studied in the past, including for AM stations, only recently has the feasibility of assembling an *economical* wide-area (continental to worldwide) synchronization system for AM broadcasting emerged, largely due to the availability of very low-cost (< $100) GPS timing receivers and inexpensive electronic devices such as microprocessors and FPGAs. We anticipate that for a cost between $1-2K, a GPS-referenced frequency-synchronizer unit can be deployed at transmitter sites. The system would be capable of holding both modern and older transmitters to well within 0.5-1 ppb of the assigned frequency, thus essentially eliminating carrier beat interference among co-channel AM stations. According to our lab tests, the net result would be a near doubling of the existing ***interference-limited*** coverage radius of stations (and nearly a quadrupling of the serviceable reception area). Unlike previous studies of close-spaced synchronous stations [5] which indicated significant distortion in the overlap areas with comparable signal levels {also see 47 CFR 73.182(t) [6]}, the proposed synchronization scheme is largely for improving the reception *in the fringe areas* of existing stations which are via allocations well-separated from their co-channel interferers (Fig.6).

For a hypothetical station in a high-conductivity area with interference-limited contours (i.e., on Class A or B channels), a 6-dB reduction in interference due to beats can provide a corresponding almost 2:1 gain in listenable range (the gold and blue dotted annular regions). Obviously, other factors such as power-line noise may well intervene, ***but both daytime and nighttime listenable areas will improve significantly.*** Although essentially all stations will benefit during early-morning and near-sunset (critical) hours, this is of paramount importance to local Class-C stations, which often can no longer cover their population centers at night due to suburban sprawl, as well as Class-D stations with very limited nighttime power levels.

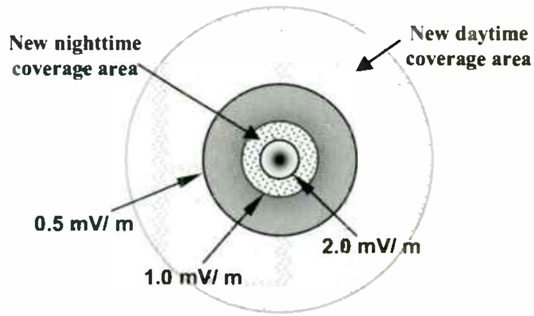


Fig. 6 — Effective Synchronous Day/Night Interference-Limited Coverage Improvements

The broad use of wide-area AM synchronization technology will also significantly improve long-distance skywave reception by minimizing the carrier beat effects from co-channel interfering signals; new HD installations will also benefit from a reduction in beats from AM analog signals. The major deployment issue is not really technical or even economic, but political — to make synchronization work optimally, all broadcasters must join in. For the majority of stations to rapidly benefit, we believe the FCC must mandate the technology. From our perspective, the low cost ($1-2K) should not constitute an economic hardship for any station, particularly when weighed against the major coverage gains to be realized. Certainly, stations near the Canadian and Mexican borders will need the additional cooperation of our neighboring nations' broadcasters, but the same benefits will apply to them as well.

In conclusion, although synchronization is a solid step in improving the AM radio medium, many more issues remain to be solved, including reduction of RFI from electrical utilities, poor receiver quality (especially regarding bandwidth, sensitivity, and noise-limiting), and transmitting antenna and network maintenance problems. Hopefully, the growth of HD radio will help address the latter issues, but strong FCC enforcement action on behalf of the broadcasters will still be needed to properly address the increasing levels of power-line induced noise and other common-carrier interference to broadcast AM radio reception.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the assistance, encouragement, and valuable technical insights offered by Trevor F. Swoyer, WETB; Tom King, Kintronic Labs; George DeVault, Holston Valley Broadcasting; Ben Dawson, Hatfield & Dawson; Paul Wisniewski, Bliley Technologies; David Layer, NAB; Bob Weirather and Geoff Mendenhall of Harris Broadcast; and Peter Conlon and Tim Hardy of Nautel, Ltd.

REFERENCES

[1] Smith, S. F. and Moore, J. A., *Carrier-Frequency Synchronization System for Improved Amplitude Modulation and Television Broadcast Reception*, U. S. Patent 6,563,893, May 13, 2003.

[2]Benjamin Dawson, private communication, January 25, 2007.

[3]Allan, David W., *Frequency and Time Coordination, Comparison, and Dissemination*, **Precision Frequency Control**, Vol. 2, Ch.13, pp. 233-273, January 1, 1985; also available online from NIST, in Bin 619 at http://tf.nist.gov/timefreq/general/publications.htm

[4]Kahn, Leonard R., *POWER-side™*, **IEEE Transactions on Broadcasting**, Vol. 34, No. 3, September 1988, pp.407-420.

[5]Whitaker, George, Sr., *Case History: Synchronous Broadcasting*. **SBE Proceedings**, 1991, pp.101-107.

[6] Part 73 FCC Rules for AM, FM, and TV, §73.182, http://www.fcc.gov/mb/audio/bickel/amfmrule.html

Transition to HD News Production

Sunday, April 15, 2007
1:30 PM – 5:30 PM

Chairperson: Dave Converse
ABC/Disney Television Stations Group, Burbank, CA

DoD Uplinks Status Report
Dane Ericksen, Hammett & Edison, Inc., San Francisco, CA

***2 GHz Relocation Project Update**
Michael Degitz, Sprint Nextel, Reston, VA

Improving Digital News Backhaul Efficiency with AVC H.264 Compression
Bob Titus, Modulus Video, Inc., Sunnyvale, CA

Affordable Wireless Broadband for HDTV Production
Robert Ehlers, HauteSpot Networks Corporation, San Luis Obispo, CA

***Technologies that Improve Digital Content for Broadcast and Film**
Jeff Brown, NVIDIA, Santa Clara, CA

Format-Transparent News Operations
Fred Schultz, Harris Corporation, Burbank, CA

High-performance HD News Production and Broadcasting Linking Plural Servers
Katsumasa Hirose, NHK, Tokyo, Japan

News Editing and Transfer System in a Distant Region
Seong-Gyu Jeon, Broadcast Technical Research Institute Korean, Broadcasting System, Seoul, Republic of Korea

*Paper not available at the time of publication

DoD Uplinks Status Report

Dane E. Ericksen, P.E., CSRTE, 8-VSB, CBNT
Chairman, SBE FCC Liaison Committee
Chairman, ATSC TSG S3 Specialist Group on Digital ENG
Senior Engineer, Hammett & Edison, Inc., Consulting Engineers
San Francisco, California

Abstract

In the ET Docket 00-258 Fourth MO&O, the FCC affirmed that DoD uplinks could use both the 1.8 GHz federal government SGLS band and the refarmed 2 GHz TV BAS band, albeit subject to successful frequency coordination with incumbent BAS, CARS and LTTS licensees. As a result of that decision, SBE has been in contact with DoD regarding how those uplinks will protect 2 GHz TV BAS operations. This paper reports the results of the first two SBE meetings with DoD representatives.

Background

In the July 7, 2003, ET Docket Fourth Notice of Proposed Rulemaking (NPRM) the FCC proposed authorizing up to eleven Department of Defense (DoD) tracking, telemetry and commanding (TT&C) satellite uplink stations that now use 1,761–1,862 MHz federal government Space Ground Link Subsystem (SGLS) band to be able to also use the refarmed 2,025–2,110 MHz TV Broadcast Auxiliary Services (BAS) band. This move was triggered by the reallocation of the 1,710–1,755 MHz federal government band to the commercial sector, for still more Advanced Wireless Services (AWS) spectrum.

This, in turn, supposedly meant finding additional capacity for DoD uplinks now operating in that band due to the relocation of other Federal users from 1710-1755 MHz into the SGLS band. DoD, IRAC (Interdepartmental Radio Advisory Committee), NTIA (National Telecommunications and Information Administration (NTIA), and, of course, the FCC, decided that the 2 GHz electronic news gathering (ENG) band would be a good place for these relocated DoD uplinks. The DoD uplinks are located in Camp Parks (CA, near San Francisco), New Boston (NH), Denver, Albuquerque, Cape Canaveral (FL), Colorado Springs, Lompoc (CA), Kaena Point, Oahu (HI), Prospect Harbor (ME) and Guam. The uplinks will have maximum equivalent isotropic radiated powers (EIRPs) of up to 104 dBm operating into a 13-meter uplink dish. The side lobe suppression of the uplink dish is only about 60 dB or so, meaning that the leakage toward the horizontal, and therefore toward ENG receive only (ENG-RO) sites, could be as high as 44 dBm EIRP. Since a typical ENG truck has an EIRP of around 65 dBm, this means that the DoD uplinks would have co-channel "leakage" at low elevation angles that could be only 20 dB less power than the main beam power of an ENG truck.

Despite the opposition comments of the Society of Broadcast Engineers, Inc. (SBE), the National Association of Broadcasters (NAB), the Association for Maximum Service Television (MSTV), and others, in the October 21, 2004, ET Docket 00-258 Seventh Report and Order (R&O) the FCC decided that DoD uplinks in the 2 GHz TV BAS band could be made to work.

SBE, NAB and MSTV then filed Petitions for Reconsideration of the Seventh R&O, but in an April 11, 2006, Memorandum Opinion and Order (MO&O) the Commission affirmed its Seventh R&O decision. Thus, the issue of the wisdom of creating a co-channel allotment with 2 GHz

TV BAS stations, and also Community Antenna Relay Service (CARS) TV Pickup stations and Local Television Transmission Service (LTTS) stations, is moot. The task facing broadcasters is now to make the sharing work. Perhaps the brightest spot of the Seventh R&O was the unequivocal decision that the newcomer DoD uplinks must protect incumbent TV BAS, CARS and LTTS operations. Further, in the Fourth MO&O, the Commission clarified that protection of 2 GHz TV BAS operations includes the protection of existing ENG-RO sites.

First DoD/SBE Meeting- El Segundo

The first DoD/SBE meeting occurred on November 8, 2006, in El Segundo, California, at The Aerospace Corporation. The Aerospace Corporation is a major government non-profit consulting firm that supports DoD in all of its space programs. Present for SBE were President Chriss Scherer, Immediate Past President Ray Benedict, Director Ralph Beaver (Chairman of the SBE Frequency Coordination Committee), Director Dane Ericksen (Chairman of the SBE FCC Liaison Committee), SBE General Council Chris Imlay, and SBE Executive Director John Poray. There were nine DoD representatives, three of which were active duty USAF officers, plus four civilian DoD and two Aerospace Corporation persons.

Decisions from the first meeting included:

1. While no agreements have been reached formally, DoD considers that the protection criteria for ENG-RO sites would be a noise threshold degradation of no greater than 0.5 dB has merit (this was the protection criteria cited in the Seventh R&O).

2. DoD indicated that they will make every effort to avoid operation in the 2,025-2,025.5 MHz lower Data Return Link (DRL) band, and in the 2,109.5-2,110 MHz upper DRL band. This is in recognition of the important role that the forty 25-kHz wide DRL channels are likely to play in future ENG operations.

3. While the DoD/SBE meetings are intended to provide guidance to both DoD and broadcasters regarding protection of ENG operations, the actual frequency coordination for each DoD uplink operating in the 2 GHz TV BAS band will have to include all local TV BAS, CARS and LTTS licensees. Meetings with these local licensees will be held in a venue near the to-be-converted uplink, and will include the local above-1 GHz BAS frequency coordinator.

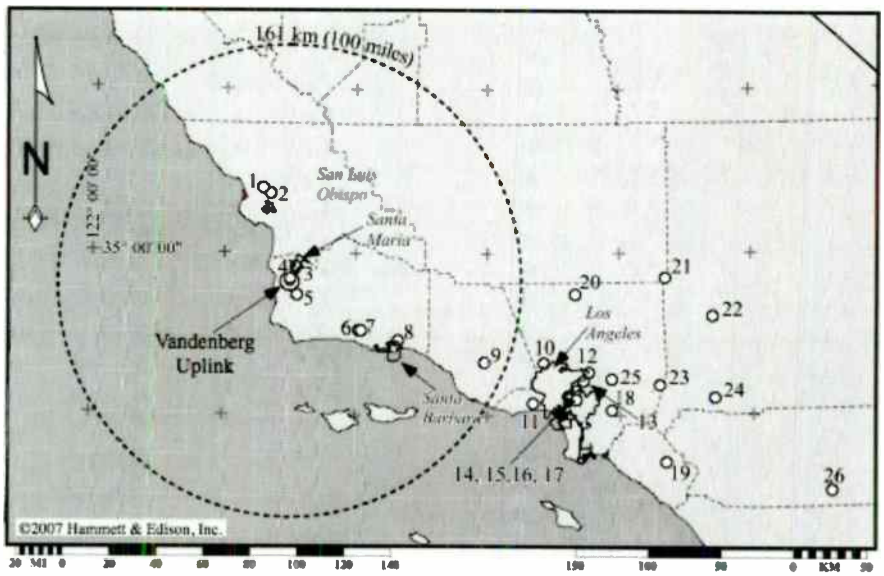
It should be noted that a 0.5 dB noise threshold degradation criteria constitutes a *frequency re-use* criteria, which is an important element of an overarching *frequency-sharing* criteria. Under frequency re-use, the two operations are engineered so that both co-channel transmissions can occur in the same area and at the same time, without interference being caused to either party. This is commonly done for fixed, point-to-point terrestrial microwave links. However, those links have the considerable benefit of using highly directive, large-diameter parabolic dish transmitting and receiving antennas, and typically being cross-polarized with a nearby co-channel link. Additionally, the path geometries between fixed links in known in advance, and does not change. Unfortunately, these conditions do not apply for ENG operations, which, by their very nature, are itinerant or mobile in nature, and are often not scheduled in advance (although some ENG uses, such as the coverage of national political conventions or sports coverage, can benefit from advance frequency coordination).

Other technical details learned during the first meeting were than the uplink dishes would be 13-meter antennas rather than the 10-meter antennas estimated by SBE in its Fourth NPRM comments. Also that the maximum Unified S-Band (USB) EIRP would be 104 dBm, as opposed to the 115 dBm maximum EIRP for SGLS operations (USB is the term the NASA and DoD uses for the 2 GHz TV BAS band). Further, it was learned that the 104 dBm EIRP an emergency, spacecraft-in-trouble, power level. The EIRP typically used for normal, 99% of the time communications would be a 14 dB lower 90 dBm EIRP. Finally, it was learned that the occupied bandwidth of a typical USB uplink signal would be about 2 MHz.

Second DoD/SBE Meeting- Las Vegas

The second SBE/DoD meeting was held on January 11, 2007, in Las Vegas, Nevada (this coincided with the Consumer Electronics Show and the winter SBE Executive Committee meeting). Preliminary interference calculations for the first DoD uplink scheduled to operate in the 2 GHz TV BAS band, at Vandenberg AFB near Lompoc, California, were reviewed at this meeting.

The validity of the SBE-proposed 161 km (100 mile) inclusion radius around a DoD uplink was tested. As shown by Figures 1, 2 and 3 below, a total of twenty-six Southern California known or possible ENG-RO receive sites were tested.

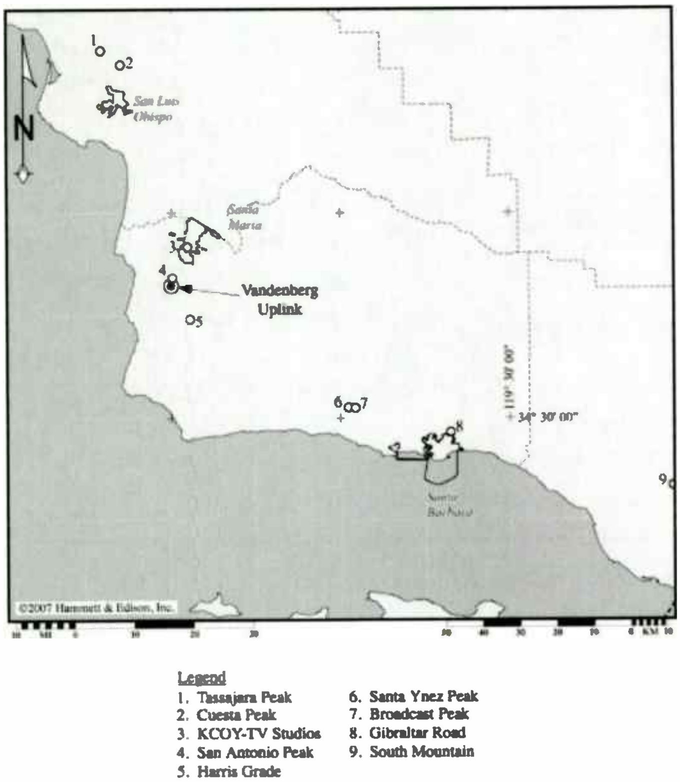


Legend

1. Tassajara Peak	10. Oat Mountain	19. Santiago Peak
2. Cuesta Peak	11. Saddle Peak	20. Palos Verdes
3. KCOY-TV Studios	12. Mt. Lukens	21. Sierra Peak
4. San Antonio Peak	13. Verdugo Peak	22. Quartzite Mountain
5. Harris Grade	14. KNBC Studios	23. Sunset Ridge
6. Santa Ynez Peak	15. KYSR	24. Keller Peak
7. Broadcast Peak	16. Hollywood Hills	25. Mt. Wilson
8. Gibraltar Road	17. CNN	26. Toro Peak
9. South Mountain	18. South El Monte	

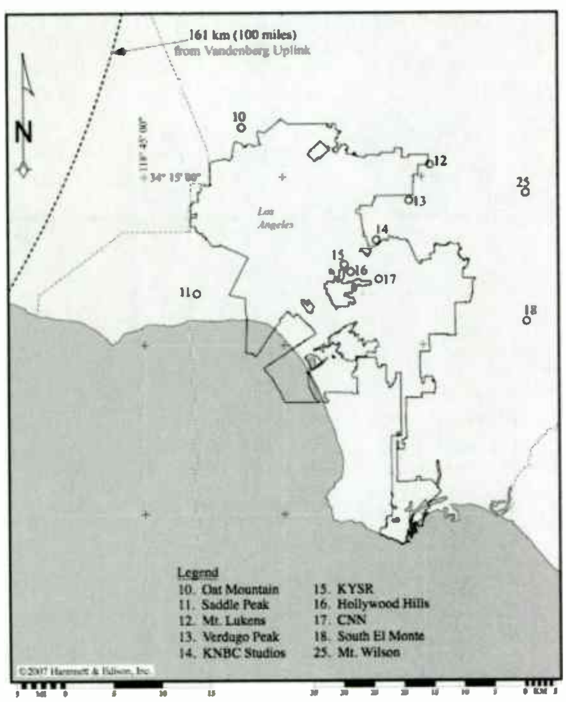
Albers equal area map projection. Map data taken from Sectional Aeronautical Charts, published by the National Ocean Survey. Geographic coordinate marks shown at 60-minute increments. City limits shown taken from 2000 U.S. Census Bureau TIGER data.

Figure 1. Possible and actual Southern California ENG-RO sites tested for main-beam illumination by the Vandenberg USB uplink dish.



Lambert conformal map projection. Map data taken from Sectional Aeronautical Charts, published by the National Ocean Survey. Geographic coordinate marks shown at 30-minute increments. City limits shown taken from 2000 U.S. Census Bureau TIGER data.

Figure 2 (top): Vandenberg uplink area sites.



Lambert conformal conic map projection. Map data taken from Sectional Aeronautical Charts, published by the National Ocean Survey. Geographic coordinate marks shown at 15-minute increments. City limits shown taken from U.S. Census Bureau TIGER/Line 2000 data.

Figure 3 (bottom): Los Angeles area sites.

Shadowgraph studies were then made. These studies indicated that all of the possible/actual ENG-RO sites more than 161 km from the Vandenberg uplink lacked line-of-sight to a presumed 9.1-meter (30-feet) AGL receiving antenna at the studied sites. While greater AGL heights for mountain-top ENG receiving antennas are certainly possible, these shadowgraph maps indicate that at least for the Vandenberg uplink that ENG-RO sites more than 161 km distant are unlikely to be affected.

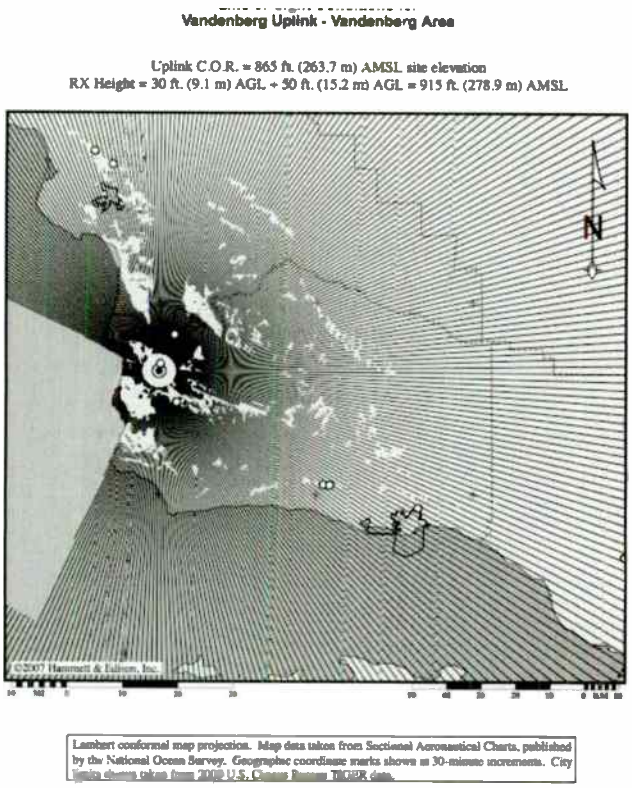


Figure 4. Shadowgraph showing line-of-sight conditions for the Vandenberg uplink antenna in the San Luis Obispo/Santa Maria/Santa Barbara area.

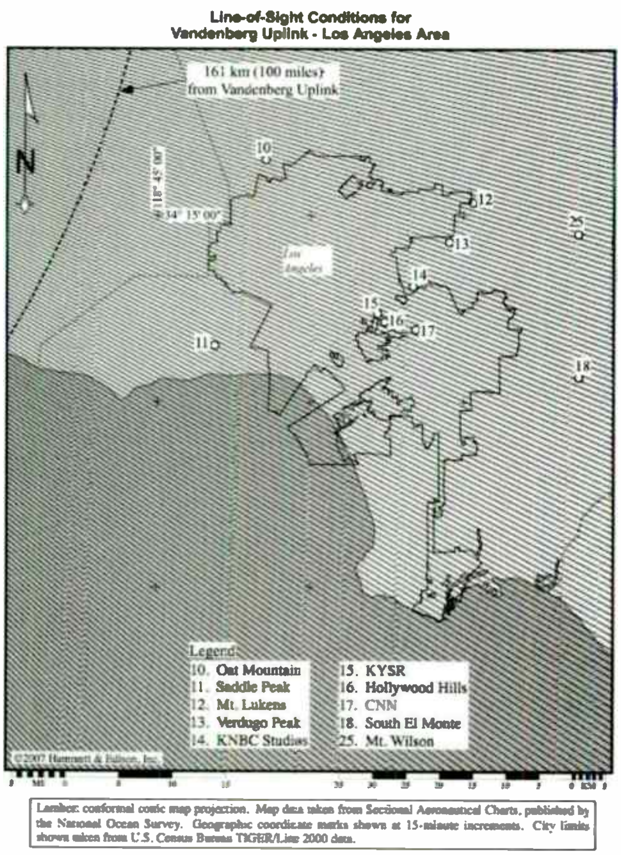
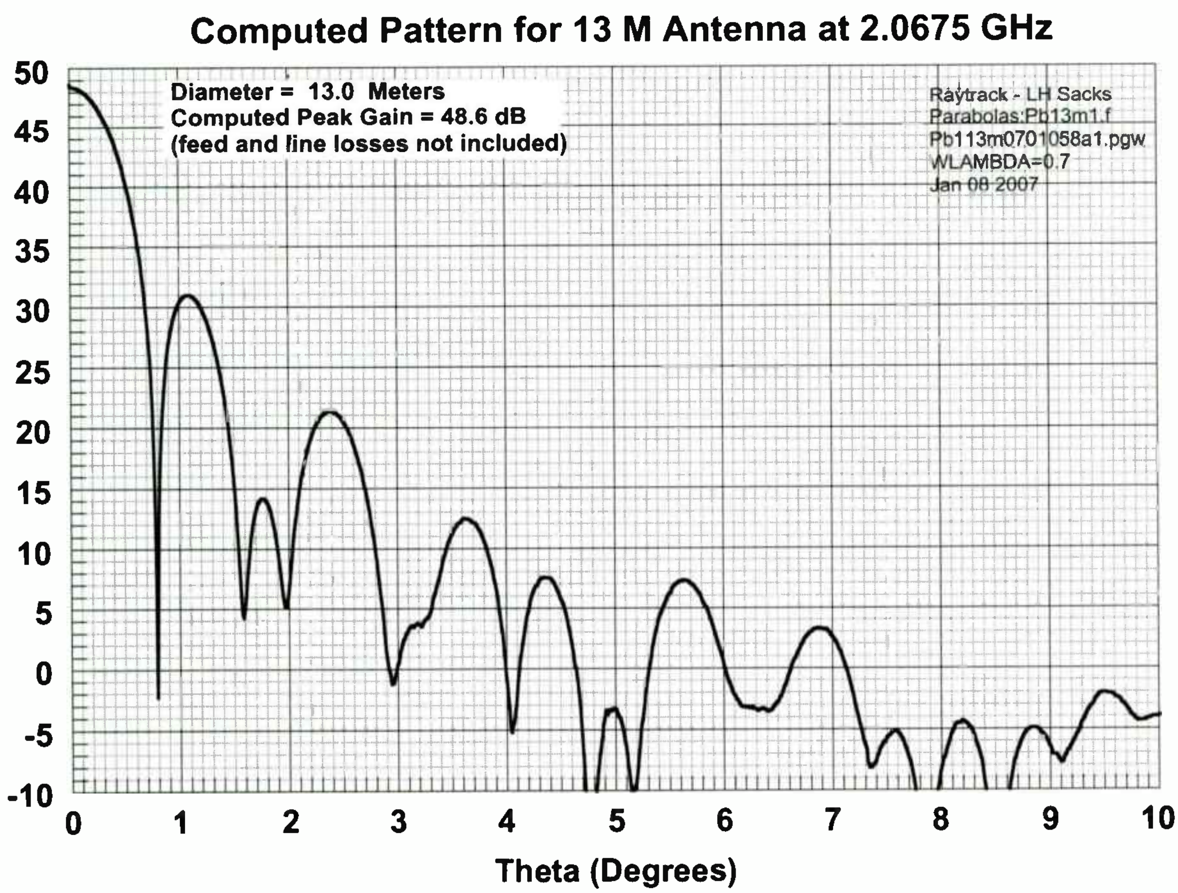


Figure 5. Shadowgraph showing line-of-sight conditions for the Vandenberg uplink antenna in the Los Angeles area.

The radiation pattern envelope (RPE) of the 13 meter USB uplink antenna that DoD will use was also discussed. Figure 6 shows the RPE from zero to 10° off axis, and Figure 7 shows the RPE from zero to 180° off axis (these are calculated, rather than measured, RPEs). Note that Figure 7 shows the additional side lobe suppression that might be possible by applying a 1.5-meter (5-foot) metal shroud around the circumference of the uplink antenna. These tentative shroud calculations were done by The Aerospace Corporation engineers; they do not include further suppression of side lobes that might be possible if the interior of the shroud is lined with microwave absorbing material. although whether this will prove to be a practical retrofit is still being studied. For example, and as shown by Figures 8 and 9, when a terrestrial point-to-point microwave dish is shrouded and lined with microwave absorbing material, improvements in the off-axis suppression of almost 30 dB are possible. It should be noted that any DoD antenna modification must be studied in detail and may not be supportable when all system factors are considered.

Figure 6 (below). Calculated RPE of 13-meter DoD USB uplink dish, zero to ±10° off axis. Courtesy of The Aerospace Corporation (Preliminary).



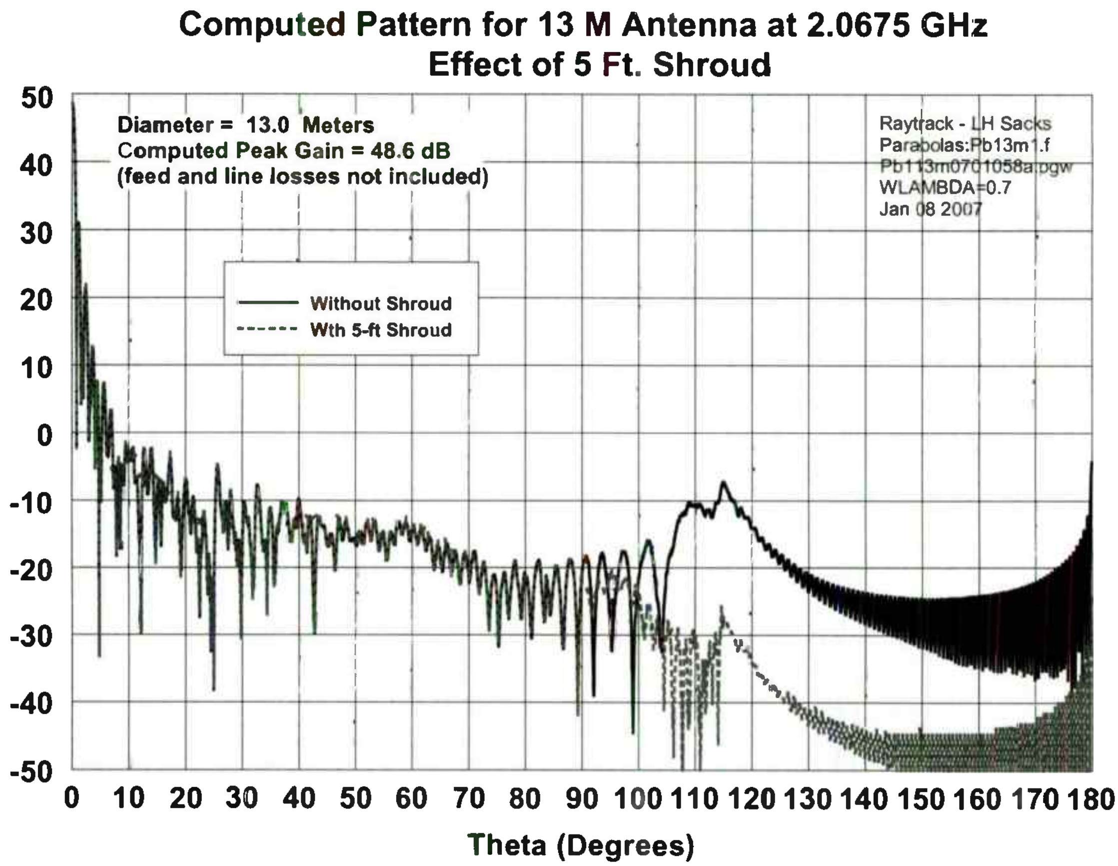
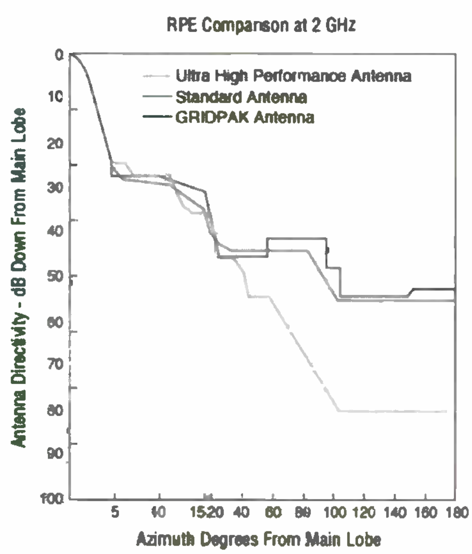
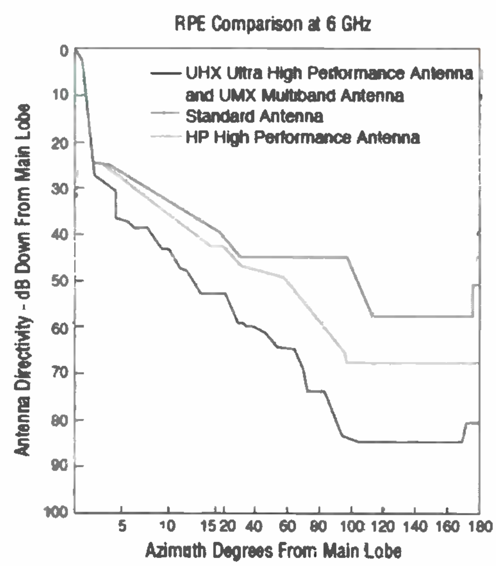


Figure 7 (above). Calculated RPE of 13-m USB uplink dish, zero to ±180° off axis. Courtesy of The Aerospace Corporation (Preliminary).





Figures 8 (left) and 9 (right). RPEs for Andrew Corporation standard performance, high performance and ultra high performance parabolic antennas at 2 GHz and at 6 GHz. These figures are courtesy of Andrew Corporation.

Next Tasks

Look Angle Histograms

DoD personnel indicated that the azimuths and elevation angles for a particular uplink as it tracks various military satellites (many in non-geosynchronous orbits) could be made, to see if these operational arcs include the azimuth(s) and elevation angle(s) to nearby ENG-RO sites that might not be terrain-shielded to the uplink location. For example, Figure 10 shows the terrain profile from the Vandenberg uplink to the ENG-RO site at Broadcast Peak, and Figure 11 shows the terrain profile to the ENG-RO site at Cuesta Peak. These are clearly unobstructed paths, where free space path loss (FSPL) conditions would apply.

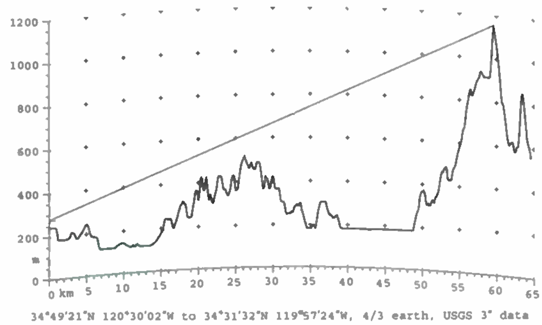


Figure 10. Vandenberg uplink to Broadcast Peak terrain profile. Path is 59.7 km at 123°T. Profile extends 5 km past the ENG-RO site.

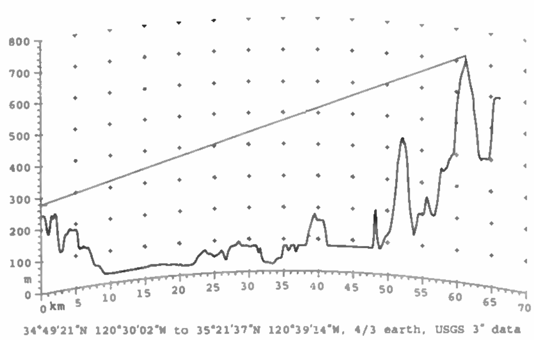


Figure 11. Vandenberg uplink to Cuesta Peak terrain profile. Path is 61.3 km bearing 347°T. Profile extends 5 km past the ENG-RO site.

If the Vandenberg uplink elevation angle is at least 10°, then for the Broadcast Peak ENG-RO site the predicted receive carrier level (RCL) at the ENG receiver input port can be estimated as +104 dBm EIRP worst case "emergency" power, -14 dB for routine, low power operation, -7.8 dB bandwidth factor (*i.e.*, 10log(2 MHz/12 MHz = -7.8 dB), -50 dB side lobe suppression, -134.3 dB FSPL, + 20 dBi assumed receive antenna gain = -82.1 dBm RCL. This misses the 0.5 dB noise threshold criteria RCL limit of -104.1 dBm by some 22 dB (*i.e.*, a co-channel undesired signal of -104.1 dBm added to a -95 dBm at-threshold coded orthogonal frequency division multiplex (COFDM) desired ENG signal would degrade the noise threshold to -94.5 dBm, or a 0.5-dB degradation).

However, if the uplink azimuth never got within ±100° of the 123°T direction from the Vandenberg uplink to Broadcast Peak, the off-axis suppression would improve by another 24 dB, reducing the RCL of the undesired DoD signal from -82.1 dBm to -106.1 dBm. This would then not only meet, but surpass, the 0.5 dB noise threshold degradation criteria. So a statistical look-arc analysis may demonstrate that a DoD uplink dish is so unlikely to be aimed in the direction of an ENG-RO site that protection of that ENG-RO can be effectively provided, even when line-of-sight conditions exist. Of course, if the look arc statistical analysis shows that the uplink antenna frequently sweeps through an arc that places less suppressed side lobes towards the ENG-RO site, then other mitigation measures will be required.

Possible D/U Interference Criteria in Lieu of Noise Threshold Protection Criteria in Certain Cases

Another area for planned future work is to derive the desired-to-undesired (D/U) signal ratio that a 12 or 6 MHz wide digital ENG signal needs in order to not be degraded by a 2 MHz wide USB signal. This could be accomplished in a laboratory using variable attenuators and directional couplers to introduce an undesired USB uplink signal into a digital ENG signal. The necessary D/U ratios for various levels of the "desired" ENG signal, and for various digital modulation schemes and combinations of forward error correction, could then be derived. Let's assume, for discussion purposes, that the most susceptible ENG signal is found to require a D/U ratio of 25 dB or better in order to not be affected. Let us also assume that an analysis of RCLs for a particular ENG-RO site show that 99% of the time the weakest incoming signal is -60 dBm (as opposed to the assumed receiver threshold value of -95 dBm). Then applying a D/U interference criteria instead of a noise threshold protection criteria may be reasonable. Doing so would then allow a RCL of -85 dBm for the undesired DoD uplink signal, as opposed to a limit of less than or equal to -104.1 dBm. Of course, it would be up to each TV Pickup, CARS, and LTTS licensee near a newcomer DoD uplink whether to accept a D/U ratio interference criteria rather than a noise threshold protection criteria. But in some situations that might be a reasonable approach.

Summary

Encouraging progress continues to be made. Especially if adding microwave absorbing material to an uplink dish shroud can reduce the side lobe leakage by another 20 dB or so (and this must be examined further and approached cautiously), then at least for the Vandenberg uplink the goal of frequency re-use has utility as a critical element of achieving an equitable frequency sharing strategy, at least to established ENG-RO sites.

For 2 GHz TV BAS receivers used by ENG relay trucks or satellite ENG (S-ENG) trucks that are themselves equipped with ENG receivers, frequency sharing, and real-time frequency coordination, will still be necessary. This is because, unlike fixed ENG-RO sites, the location of an ENG relay truck or an S-ENG truck will often not be known in advance. However, this should prove to be a much more tolerable burden than having to frequency coordinate regularly used ENG-RO sites.

Acknowledgements

The author wishes to acknowledge the cooperation of Mr. Fred Moorefield of the USAF Frequency Management Agency. Also the cooperation and assistance of Dr. Albert "Buzz" Merrill of The Aerospace Corporation, as well as all of the other DoD representatives.

The work of SBE President Chriss Scherer, SBE Immediate Past President Ray Benedict, SBE Frequency Coordination Director Ralph Beaver, SBE General Counsel Chris Imlay, and SBE Executive Director John Poray, has also been instrumental in this effort. It should be noted that SBE officers and directors volunteer their time, and are generally *not* reimbursed for travel expenses.

Finally, I wish to thank Mr. Howard Fine of the Southern California Frequency Coordinating Committee (SCFCC) for his ongoing and immeasurably helpful assistance.

Improving Digital News Backhaul Efficiency With AVC H.264 Compression

Bob Titus
Modulus Video, Inc.
Sunnyvale, California

Introduction

Despite growth in the availability and affordability of bandwidth around the world there never seems to be enough for broadcasters to use when they need it. This is particularly true of news backhaul circuits where many have had to resort to proprietary compression algorithm and lower than desired picture quality and frame rates.

AVC/H.264 encoding for point to point links offers compelling quality advantages for given bit rates compared to other commonly used compression standards, and is an excellent candidate for backhaul of DNG feeds, particularly as these migrate to HD in the future. Many organizations are struggling with how to handle HD feeds within their existing communications infrastructure. AVC can provide sufficient improvement in compression efficiency to allow carriage of HD feeds within the channel space currently used for SD links.

For SD backhaul the challenge is a little different. In this case broadcasters are looking to get data bandwidth requirements down to the point where signals can be reliably carried over public data circuits and/or over the emerging cellular data networks. AVC is equally applicable here and allows a standards-based approach to be retained while still preserving frame rate and image quality at very low bit rates.

The fact is, while much attention is paid to the increasing number of transmission technologies – wireless, fiber, free space optics, leased lines, microwave, WiMAX, millimeter wave radio, unlicensed bands and licensed microwave radio – none of these resolve the underlying bandwidth challenge. There are simply too many bits that need to be moved reliably through too many narrow pipes.

As a result, broadcasters find themselves faced with an increasingly frenetic juggling act, keeping bandwidth cost, technology expenditure, quality demands, and more in the air simultaneously.

This paper analyses the requirements and opportunities of these two segments and discusses some of the trade-offs inherent in any such compression system.

AVC Background

Advanced Video Coding (AVC) is the latest generation of video compression. AVC, also known as ITU-T H.264/AVC and ISO/IEC MPEG-4 part 10/AVC, was developed in the first few years of the 21st century in response to widespread calls for more efficient video compression. It succeeds MPEG-2 in most application areas and provides tools that enable a wide variety of new and improved applications.

AVC would be better described as a tool kit than as a monolithic CODEC (COmpressor-DECompressor). Compared to its popular predecessor, MPEG-2, AVC provides more video processing versatility. Both MPEG-2 and AVC are block-based transform-based CODECs, but in AVC the blocks are variable in size and the transforms are optimized for real world processors. AVC also incorporates very efficient prediction modes that don't exist in previous CODECS. Prediction is the way of recycling data that has already been sent to a decoder so that it doesn't have to be sent again, thus saving bandwidth.

Current generation professional AVC encoders implement Main Profile, which is a particular subset of the full AVC tool kit. AVC Main Profile encompasses both standard definition (SD) and high definition (HD) formats. The most advanced MPEG-4 AVC encoders embrace more of the full AVC standard, providing more inclusive profiles that provide tools for additional kinds of data transformations, data reduction, and color processing.

It was claimed early on that AVC would be about twice as good as MPEG-2 in terms of compression efficiency. In practice, commercially available top-tier AVC-based broadcast encoders not only meet the twice-as-good mark, but in some cases do even better. By way of comparison, many IPTV services are planning to launch HDTV services at just 8 Mbps per channel, a dramatic reduction from the 18 Mbps or more required to achieve equivalent quality with MPEG-2. Future improvements in AVC compression efficiency will come from incorporation of AVC High Profile tools into 2nd and 3rd generation professional encoders, and the use of more sophisticated proprietary video processing techniques such as noise filters that enhance the efficiency of the core AVC tools. One can already see AVC HDTV television services in development at 5 to 6 Mbps, maximum.

The Difference with News Backhaul

Backhaul takes many forms, some more bandwidth constrained than others. Major programming feeds and high profile sports, for example, typically have significant fixed infrastructure behind them including very large data pipelines.

News, however, is very different and while many broadcasters are fine for the moment, the writing is on the wall. The advent of high definition news, in particular will put tremendous strain on existing bandwidth.

The fundamental problem facing broadcast news is the need to move an ever increasing number and variety of feeds within the same bandwidth they are using today. When viewers turn on today's news, whether it is the biggest national news program or the smallest local news show, they expect the coverage to include reports from many localities, near and far; locations that often have no dedicated circuits upon which to rely. Increasingly, people expect a large percentage of that coverage to be live.

Some of the most compelling content to come out of the Iraq War was live shots of embedded journalists moving with troops on maneuver. Arresting as the coverage may have been, the image quality was horrendous. The ability to more than half the data rate using MPEG-4 AVC encoding through the same channel will dramatically improve this type of coverage.

Local stations are increasingly finding themselves under pressure from national networks to expand their coverage from remote locations. People are used to seeing video and reports from every point on the globe. If a story can be delivered from the mountains of Afghanistan, they quite reasonably expect to see the local high school basketball team and a current score delivered from across town, across the state, or across their region.

More Bandwidth vs. Greater Efficiency

As a result, broadcasters find themselves using ever more backhaul to meet that demand. Unlike many of the technologies used in broadcast that have benefited in reduced costs through commoditization, bandwidth remains a precious resource. Current rates for transponder rental are in the area of $50,000 per MByte pr year. If you want more feeds, broadcasters either need to pay for more bandwidth or find ways to make more efficient use of their existing bandwidth, wedging more signals into the same bandwidth.

Making better use of existing bandwidth is the clear choice for most broadcasters. That means they need to be smarter about how they compress their signals.

MPEG-4 AVC Opportunity

In existing standard definition video backhaul processes, the status quo is fine…for today. The problem is the increasing number of feeds that need to be sent, many coming from inconvenient places with no line of site. A variety of new delivery channels such as satellite phones, public internet and leased circuits hold great appeal for news applications, however they lack the bandwidth for broad use.

To staff and equip and ENG remote can cost any station thousands of dollars. The low bandwidth requirements enabled by MPEG 4 allow stations to use some of the more current wireless technologies and reducing the manpower requirements to staff the remote.

The challenge is that a T1 only offers about 1.5 Mbps, considerably less than the microwave and satellite links broadcasters are accustomed to. Today's only options are to use MPEG-2 or MPEG-1 utilizing half D1 resolution, which will provide an image at low bit rates, but at 1.5 Mbps the quality is unacceptable by most standards.

On the contrary, MPEG-4 AVC has been demonstrated to deliver high quality at an astounding 1.5 Mbps and less. The low bit rates offered by AVC make possible some interesting scenarios. At ~1.5 Mbps it is theoretically possible to do live transmission over a T-1 line, particularly when coupled with advances in forward error correction. The issues which must be overcome are video packet loss, packet reordering and jitter. Any one of these can cause and adverse affect on the video quality. To overcome these issues forward error correction must be used to ensure the integrity of the signal at the receiving end. One method of FEC which is currently being deployed is COP 3 or Pro_MPEG. There are other methods being discussed of which there are several which are hybrids, employing technologies shared by several companies. All have their pros and cons. When encoder manufacturers combine FEC with the use of high efficiency codecs, such as AVC, then it becomes possible to reliably move high quality images over some of these new low cost backhaul paths.

Enabling HD for News

By far the biggest bandwidth squeeze broadcasters face is in backhauling HD signals in scenarios where they do not have any more bandwidth than they had for SD. Helicopters, for example, confront a major challenge trying to deliver HD over their microwave links as higher bandwidth requirement significantly impact their range, in effect nullifying their value.

HD coverage, however, has been deemed critical in markets of all sizes.

Point to point microwave links will not typically suffer from a bandwidth limitation, however, a satellite backhaul, where the broadcaster must essential pay by the bit, there is much more to be gained by squeezing the video into smaller channels.

Historically there has always been a question about latency in any compression scheme. History has also shown, however, that station managers and financial decision makers have determined that the benefits of implementing improved compression technology are too good to ignore. When MPEG-2 was first deployed, no one was happy with its delays, but, people grew used to it. Over a relatively short period of time, the encoding grew faster and today we find latencies of only a few frames.

Having seen this process before, it is reasonably safe to assume that the dramatic increase in visual quality per bit combined with the need to backhaul high definition video news stories will lead to implementation of MPEG-4 AVC for news. Latency in today's top MPEG-4 AVC encoders is shrinking rapidly.

For satellite backhaul, where people expect a bit of delay anyway, the difference to the viewer is hardly noticeable, but the potential impact on cost savings, better pictures and greater flexibility is significant. Broadcasters will want to use AVC on their satellite links because the cost per bit is quite high. Right now, an MPEG-2 link from and SNG truck at about 12-15 Mbps delivers a good SD picture. The problem is that to achieve the same quality for HD video, a much higher bit rate will be required.

MPEG-4 AVC, however, already delivers stunning HD at the same 12 Mbps bit rate used for MPEG-2 SD, and in fact maintains this image quality at significantly lower bit rates. The bottom line is that AVC enables news organizations to deliver HD video within the infrastructure that it already uses. The compatibility between MPEG-2 and MPEG-4 AVC further means that you can readily move that data within existing links as both use an MPEG2 transport format.

During the transition from MPEG-2 to MPEG-4, it is important to be able to support both formats. Significant investment has been made in legacy devices that will only handle MPEG-2 and a tremendous quantity of MPEG-2 content will linger for years. At the same time deployments of IPTV over bandwidth-limited legacy infrastructure expands, their will be increasing demand at the distribution end to move to MPEG-4 AVC. Lastly, the need to support multiple formats, such as HD and SD, as well as to support an increasingly broad range of display technologies such as mobile devices, means that efficient transcoding will be very important. Developers including Modulus Video have done extensive transcoding testing with excellent results concatenating AVC and MPEG-2 compression. In fact, side by side demonstrations at recent public events have pitted original content, Modulus Video AVC content at sub 5 Mbps and content transcodes to MPEG-2 at 18 Mbps with excellent results.

Standards based approach

Of course, other technologies have been put forth with some success. Some of these offer reasonably good compression and final picture quality. The concern put forth by many broadcasters is that these other technologies tend to be based on proprietary schemes. These days, most broadcasters prefer to see standards based solutions in which equipment from the widest variety of manufacturers is interoperable.

MPEG-4 AVC fits that requirement, delivering very high compression rates, excellent picture quality, and an internationally established standards foundation that allows the broadcaster to choose best of breed hardware to fit their requirements.

Conclusion

Contribution operations such as news are a good choice for early MPEG-4 AVC adoption. Unlike other elements of the broadcast chain, the lower relative costs of contribution will enable a quicker return on the investment and put the infrastructure in place for later moves to broader MPEG-4 deployment.

With further advances on the way, including greater compression efficiency, lower latency, and higher signal resiliency using Forward Error Correction, MPEG-4 AVC offers broadcasters a compelling way forward for the future.

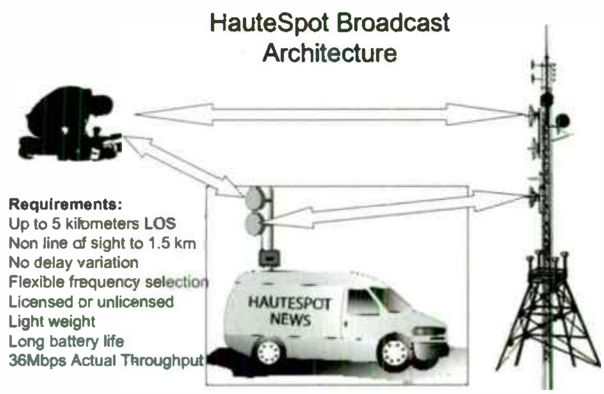
Affordable Wireless Broadband for HDTV Production

Bob Ehlers
HauteSpot Networks Corporation
San Luis Obispo, California

Abstract

High Definition Television presents many challenges to everyone involved in both the production and broadcast (distribution) of high quality video and audio. One of the biggest challenges is moving the signal around, either from the studio to the consumer's home (broadcast) or from the camera man to the switcher and then to the broadcasting network (production).

For the most part HD content begins life out of a camera as a HD-SDI (Serial Digital Interface SMPTE 292M or SMPTE 372M stream of between 1.485 and 2.970Gbps).



In most cases the HD content is delivered to consumers as MPEG2 or MPEG4 compression encoded digital streams that fill anywhere from 300kbps to 25Mbps "pipes". Television manufacturers build monitors which decode these streams and "fill in the gaps" to produce high quality images.

In between the original source and the end consumer, the content can take many forms, but for the sake of flexibility, most content producers would prefer to keep the content uncompressed or less compressed. It is for this reason that content producers prefer to do all of their work using the HD-SDI raw stream and then compress the final output to whatever formats are desired for the end consumer.

One of the biggest challenges with HD-SDI in production is that it is very difficult to move around. HD-SDI has a cable distance limitation of approximately 100 meters. Some manufacturers have created proprietary twinax solutions which can extend this distance out to 300 meters, but this also requires more rigid and expensive cable.

Ideally, HD production would use wireless technologies. However, in order to move the content in raw HD-SDI form, you need huge bandwidth. This has necessitated the development of extremely expensive, short range, high data rate proprietary microwave solutions. Beyond the requirement for raw bandwidth, wireless technologies also require support of low latency, no retransmission of dropped data, little interference, and low jitter (delay variation). Further, if more than one source is used, consideration needs to be given to source synchronization or common clocking.

What we will explore in this presentation is using a more thoughtful approach to combining high resolution compression technologies, new wireless network protocols, multicast streaming protocols, and digital microwave technologies to yield a more portable, standard based architecture for wireless HD production transport.

How Big A Pipe?

The SMPTE 259M standard for SD-SDI defines a serial data stream with a nominal rate of 270 Mbps for uncompressed standard definition video with up to 16 channels of audio.

For High Definition, the SMPTE 292M standard for HD-SDI defines a serial data stream with a nominal rate of 1.485 Gbps and even going as high as 2.97 Gbps which is required to support 1080p at a 59.94 Hz frame rate with 4:4:4 color sampling.

Raw HD video serial streams are very large indeed. In fact there are Ultra High Definition cameras that can produce streams that exceed 16Gbps. Obviously moving raw data around wirelessly is a great challenge.

In order to get the bandwidth necessary to move this amount of raw data complex modulation schemes such as QAM-128 or QAM-256 and very high frequencies are required. Frequencies over 10GHz are needed, with channel widths measured in hundred of MHz in order to carry these uncompressed signals. The obvious consequence of this is that applications are limited to being purely line of sight, short range, and subject to interference. Plus they are going to be very expensive. This pretty much defeats the reasons to look at wireless technologies for HD production to begin with.

In order to move video streams at lower frequencies, extending range and allowing for non-line of sight applications, the stream must be compressed. The question is: how much?

The answer lies in the end use of the stream. To better understand the bandwidth required, we need to understand the technologies used to compress and stream HD over packet networks like TCP/IP. Bear with me as I discuss some of the basics of MPEG, JPEG2000, muxing, streaming and transcoding.

Most HD content is delivered to consumers using either MPEG2 or MPEG4 encoding, with the industry moving quickly towards MPEG4 AVC H.264 as the standard. This allows broadcasters to compress down to a fairly compact 8 to 9Mbps. Set top boxes and HD receivers incorporate decoders which decompress the stream and fill in the gaps to create high resolution images.

Today, most satellite and cable operators are delivering HD content over MPEG2 with compression rates averaging about 15Mbps.

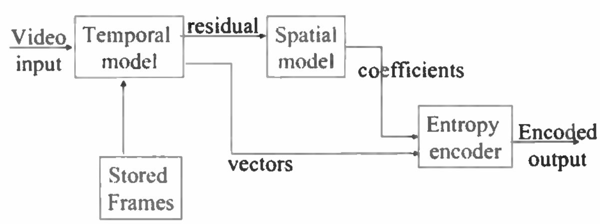
HD-DVD and Blu-Ray disks are capable of running at up to 36Mbps, however most content is delivered in streams of 19.2Mbps. Let's say that they have an upper capacity of 36Mbps and are the "bandwidth leader" among the consumer HD delivery media.

Mode	Rate	Note
Video Conferencing	300Kbps	
Typical IP Internet Video Streaming (YouTube)	50-700Kbps	
MPEG4/AVC SD Broadcast	1.5-2Mbps	Will be used for channel density on satellite and cable where video quality is less important
DVB, DSS MPEG2 SD Broadcast	3.75Mbps	Being phased out in favor of MPEG4
MPEG4/AVC HD Broadcast	8-10Mbps	Assumed data rate for the next 5 years to the home
MPEG2 SD Broadcast	13-18Mbps	Being phased out in favor of MPEG4
HD-DVD, Blu-Ray	19-36Mbps	Assumed hard media and best image quality for home
SMPTE 259M SD	270Mbps	Nominal rate over SDI serial interface
SMPTE 292M HD	1.485Gbps	Nominal rate over HD-SDI serial interface
SMPTE 372M or 424M	2.970Gbps	Nominal rate over two HD-SDI serial interfaces
Professional HD 1920x1080@180fps	10Gbps	Proprietary
Ultra HD 4096x9620@60fps	16Gbps	In development
3D Holographic 1600x2900@60fps	6.3Tbps	In development

MPEG 4 AVC H.264

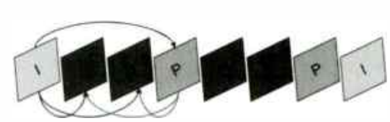
So the consumer will be receiving their HD as MPEG 4 AVC H.264. What does that mean?

MPEG4 Video Encoder Block Diagram



MPEG 4 encodes streaming content using a combination of vectoring and spatial modeling. The input signal is processed by the encoder comparing time sampled image frames and then calculating the differences between frames.

MPEG Frame Structure



I-Frames (keys) – These frames are coded using intra-coding. They are the key frames on which indexing occurs for restart or random access. Moderate compression

P-Frames – Predicted frames based on the previous I-Frame and used as one of the sources for prediction of the B-Frames. High compression

B-Frames – These are bi-directionally predicted frames based on the prior I-Frame and the next P-Frame. Highest compression.

Temporal modeling is based around making predictions from previous video frame to the next frame. These changes are due to motion and use estimates trajectory or optical flow and sub pixel motion compensation.

Calculated vectors from the temporal model are fed to the entropy encoder and the remaining content of the image is passed off to the spatial model for predictive coding calculated based on neighboring pixel data. The spatial model provides additional data to the entropy encoder.

Final encoding is made based on a combination of the vector data, variable block size motion compensation, and calculated spatial data.

The result is a video stream which can be compressed up to 83:1.

Frames within MPEG4 are keyed and the encoding works on both the key frames and the differences between the key frames (I-frames), creating predicted frames (P and B frames) in between the keys.

MPEG4 is not without its drawbacks. The primary area where it is weak is latency, where encoding can take up to 300ms and decoding could take the same amount of time. MPEG4 requires a full frame to be received before it can start to calculate its algorithms.

Another issue with MPEG4 is that if a key frame is dropped, an entire set of predicted frames may be lost, forcing the decoder at the receive end of the stream to wait for the next I-frame (key) and resync, which can be a very processor intensive and time sensitive process. Of course the majority of the data is stored in the I-frames (key frames), and the predicted frame information makes up a smaller proportion of total data, since the predicted frames only have to carry the differences from the key and not all of the data.

Motion JPEG 2000

Motion JPEG 2000 is another compression technology which uses wavelet encoding technology in place of the prediction model of MPEG-4. Motion JPEG 2000 does not use interframe prediction, rather every frame is indexed. The lack of use of interframe prediction results in a loss of compression capability, but eases video editing, since simple edits can be performed at any frame with no need to recreate predicted frames.

Using only intraframe coding technology also makes the degree of compression capability independent of the amount of motion in the scene, since temporal prediction is not being used.

Motion JPEG has the added advantage that each frame can start to be decoded immediately upon arrival of the first bits of the frame. You don't have to wait for the whole frame to be received, making latency lower than protocols requiring I-frames. It also consumes less CPU cycles to perform compression and decompression.

Motion JPEG Frame Structure

I-Frames (keys) – In every frame there is index information and encoding is accomplished using wavelet algorithms. Compression of each I-frame is greater than the compression of I-frames in MPEG, but since every frame carriers full image detail, the overall stream is larger than MPEG.

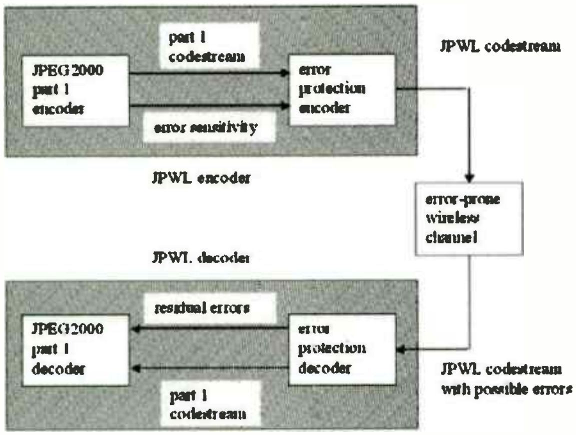
Finally, an important attribute for wireless transmission is that losing a frame does not affect any other frame. The frame can be dropped, but the stream can continue on, without skipping any of the P and B frames between the I-frames, since every frame is an I-frame. Where your network is not absolutely reliable (i.e. wireless) being able to drop one frame without it effecting any other frames is advantageous, particularly when looking at processor utilization and latency.

Ideally, for wireless multimedia applications, any encoding scheme should be natively more robust to media transmission errors. Motion JPEG 2000 has to be robust to transmission errors. The core coding system of Motion JPEG 2000 provides some error resilience to improve performance over noisy channels. However, these tools only detect where errors occur, conceal the erroneous data, and resynchronize the decoder. In other words, they do not correct transmission errors. Furthermore, these tools do not apply to the image header which is the most important part of the code stream. For real time streams, retransmission really doesn't make much sense, but for near real time, where buffering is acceptable and latency is not as important, this is greatly improve the quality of video transmitted over wireless.

The proposed Motion JPEG2000 wireless system supports three functionalities:

- protection of the code stream against transmission errors,
- description of the degree of sensitivity of different parts of the code stream to transmission errors,
- description of the locations of residual errors in the code stream.

The proposed Motion JPEG2000 system is illustrated in the figure below.



"The error protection encoder and decoder consist in techniques to protect a code stream against transmission errors. The error protection encoding process modifies the code stream to make it more resilient to errors, e.g. by adding redundancy or by partitioning and interleaving the data. In particular, provisions are made to specifically protect the image header, as it is the most critical part of the code stream. Techniques to protect the code stream include Forward Error Correcting (FEC) codes, data partitioning and interleaving, and unequal error protection. The error protection decoding process detects the occurrence of errors and corrects them whenever possible.

The error sensitivity descriptor describes the degree of sensitivity of different parts of the code stream to transmission errors. This information is typically generated when the image is encoded using a JPEG 2000 Part 1 encoder, but it can also be directly derived from a Part 1 code stream. This information can subsequently be used when protecting the image. More specifically, sensitive parts of the code stream can be more heavily protected than less sensitive parts (unequal error protection).

The residual errors descriptor specifies the locations of residual errors in the code stream. The residual errors are the errors which cannot be corrected by the error protection decoder. This information is typically generated during the error protection decoding process. This information can subsequently be used in the JPEG 2000 Part 1 decoder to prevent decoding corrupted portions of the stream." [2]

[2] The JPEG committee web site describing the JPWL working group.

The primary disadvantage of Motion JPEG2000 for wireless is that it is implemented through fairly proprietary codec and that the compression rates achieved using it are less than MPEG4. Motion JPEG2000 offers compression rates of approximately 40:1 with good quality.

Much debate exists between supporters of the MPEG4 standards and the Motion JPEG2000 implementations. Both are well suited to deliver compression rates that are good for wireless networks. Which codec are ultimately adopted by the broadcast industry for production remains to be seen.

Transcoding

"Transcoding is the direct digital-to-digital conversion from one (usually lossy) codec to another. It involves decoding/decompressing the original data to a raw intermediate format (i.e. PCM for audio or YUV for video), in a way that mimics standard playback of the lossy content, and then re-encoding this into the target format.

Transcoding can also refer to the encoding of files to a lower bitrates without changing video formats, a process that is also known as transrating.

Compression artifacts are cumulative; therefore transcoding between lossy codec causes a progressive loss of quality with each successive generation. For this reason, it is generally discouraged unless unavoidable." [1]

When we compress video streams for transport over wireless media for production, the key issue that we need to be concerned with is transcoding. Ideally, streams should be compressed once and sent to their destination without decoding or transcoding. In a perfect world this means using MPEG4 AVC H.264 or Motion JPEG2000 compression at the camera and send it directly to a decoder at the home or burning it to disk.

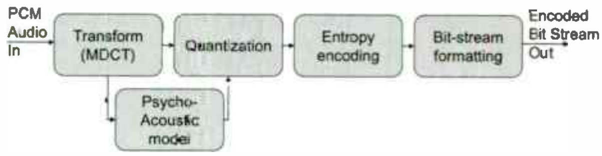
When transcoding is required, going from a Motion JPEG2000 stream to a MPEG4 stream should yield better results, as there is less prediction used in Motion JPEG, meaning that the stream will not be compressing elements that were predicted by the upstream codec.

[1] Wikipedia definition of Transcoding

Audio

Audio compression works in much the same way as video. The algorithms are different. In most actual implementations MPEG1 or MPEG2 are used in conjunction with a MPEG2 or MPEG4 video stream. Although MPEG4 has its own audio encoding scheme which yields very good compression without sacrificing quality. MPEG4 introduced AAC encoding which is very scalable and allows for many channels of audio to be multiplexed with the video stream, allowing such support as 5 channel surround sound to be delivered without more complex external multiplexing.

Audio Encoding



More recently MPEG4 amendments have added ALS lossless compression which can be delivered at 128kbps.

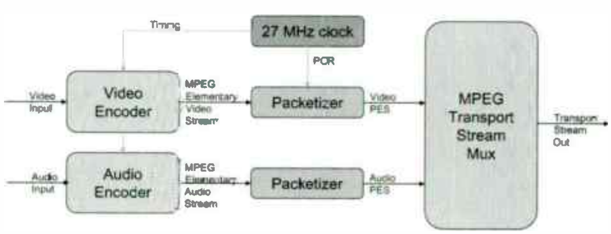
In the case of Motion JPEG2000 audio can be muxed into the stream at similar sampling rates.

The size of audio streams is pretty trivial compared to video, assuming that you are limiting yourself to 2 channels. HD-SDI supports up to 16 channels embedded, so if you use all of the available HD-SDI channels, you can have a significant amount of bandwidth required for audio. Typically allocating 128kbps is more than sufficient.

Muxing and Synchronization

Both MPEG4 and JPEG2000 define layers for video and audio muxing and synchronization. These layers operate by providing a common clock to both audio and video streams and then time stamping frames as the exit encoding.

MPEG Multiplexing



This stamping and synchronization allows streams to be placed into transport stream wrappers so that they can be sent across a network and reassembled in order at the receiving end.

Streaming Control

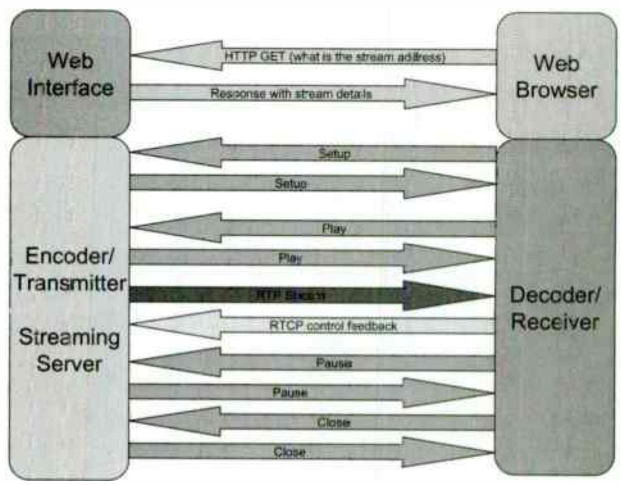
Finally, once the video and audio content have been compressed, encoded, muxed and clocked for synchronization, the stream needs to be transported. In proprietary schemes this may be nothing more than opening a simple data pipe over the wireless link. A one way transmission scheme is fairly simple to construction, but allows for no adaptation.

If a full bidirectional network is available, streaming media control is possible. Protocols such as RTP (Real Time Protocol), RTSP (Real Time Streaming Protocol) and RTCP (Real Time Control Protocol) allow for much more robust and higher quality streaming.

- RTP is a transport protocol for the delivery of real-time data, including streaming audio and video.
- RTSP is a control protocol for initiating and directing delivery of streaming multimedia from media servers, the "Internet VCR remote control protocol". RTSP does not deliver data, though the RTSP connection may be used to tunnel RTP traffic for ease of use with firewalls and other network devices.
- RTCP is a part of RTP and helps with lip synchronization and QOS management, among others.
- RTP and RTSP will likely be used together in many systems, but either protocol can be used without the other.

This diagram shows how RTSP can be used to get an RTP stream from the encoder/transmitter.

RTP, RTSP and RTCP



RTP, RTSP and RTCP combined add less than 5% total overhead to any streaming media, and in the case of large files such an HD stream, they are negligible.

Conclusion on Pipe Size

For now, the assumption I will make is that a typical consumer will expect delivery of a MPEG4 AVC H.264 encoded video stream at a data rate of 8 to 10Mbps and a hard media stream (HD-DVD or Blu-Ray) with an actual TCP/IP data rate of **36Mbps**. While there may be other stream requirements for venues such as stadiums, these are the exception and not the rule.

Now that we have established a basic premise as to the size of the data stream and the compression method, let's look at wireless system requirements.

Requirements for a Wireless HD Production Solution

There are many different sets of requirements for a wireless HD solution for broadcast production. In this case we are talking about a camera back or relatively compact solution that can be operated from battery, are light weight, are simple to configure and use, have good range and are reliable. Of course, they must also be able to transport the video streams reliably.

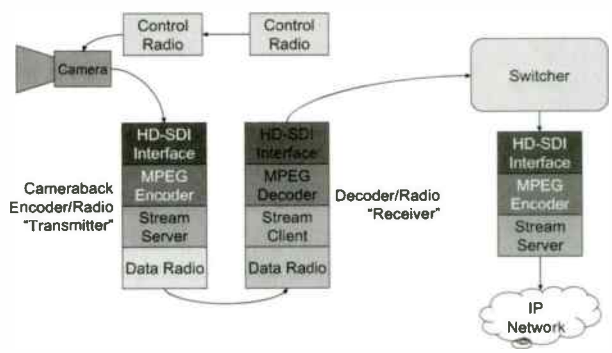
The following table gives a good example of a set of requirements for such a solution:

Requirement Title	Requirement Description
Modularity	User replaceable parts/components
User Interface	Simple to use, with keypad and LCD or browser based interface
Video Input Format	SD definition: NTSC format at 29.97fps: Full D1: 720x480. PAL format* at 25fps: Full D1: 720x576 HD definition formats: 1080/50i(1125) 1080/50i(1250) 1080/60i(1125) 720/60p(750) 576/50i(625) 576/50p(625) 480/60i(525) 480/60p(525)
Video Input Connectors	HD-SDI, SDI, HDMI, Component
Audio Input Format	1 stereo pair with 48kHz response
Audio Input Connectors	XLR
Radio Band	Unlicensed NII/ISM, BAS, and upcoming 1764-1803MHz
Range	Up to 5 km line of sight and 1.5 km with obstructions
Bandwidth	User adjustable 5, 10, 20 or 40MHz
Modulation	OFDM or COFDM
Channelization	Conforming to FCC regulations
FEC Coding Rate	1/2, 1/3, 1/4
Security	Support for industry standard cryptography and authentication, including the ability to meet FIPS140-2
Antenna Connectors	Standard such as SMA or N type
Antenna Options	Ability to attach small omni directional antennas up to high gain parabolics for really long range
Ethernet	Desired - Ability to use for camera control, IP devices or cameras, or other like back channel operations
Serial Port	Desired - Ability to use for camera control, IP devices or cameras, or other like back channel operations
USB Port	Desired - Ability to use for camera control, IP devices or cameras, or other like back channel operations
Visual Indicators	Power on, Status, Ethernet Link, Ethernet traffic, Alignment, Frequency/Channel set, Site Survey
Audible Indicators	Alignment for directional feeds
Network Clocking	Ability to set to GPS clock or network clock for synchronization
Input Power	12VDC "low power" – connect to Anton Bauer Batteries
Enclosure Material	Light Weight, extremely rugged
Operating Temperature	-40 °C to + 70 °C
Weight	Less than 3 lbs
Dimensions	Same as a camera battery
Compliance	FCC compliant and able to interoperate with other devices

There are several manufacturers currently delivering solutions that meet many of these requirements. Their products are, for the most part:

- Proprietary
- One way (data)
- Very Expensive

The following diagram illustrates the typical wireless HD solution.



At first glance they appear to provide raw HD-SDI to HD-SDI pipes. But on closer inspection, they use compression encoding in order to fit the video stream into their available bandwidth, then they decompress at the other end. While they may get fairly good results with their compression methods, any signal carried over these solutions, which is intended for rebroadcast (and therefore recompression again) will suffer from cumulative resolution loss due to transcoding.

Note that in most of the wireless HD solutions today, camera control is delivered completely out of band with a separate radio system. Why is this?

Do you Jitter?

Typical IP wireless networks are half duplex by design. Both ends of the wireless link are transceivers, capable of both receiving and transmitting data over the air. When a frame of data is sent, it is done so with some error checking in the header of the frame, often just parity. Once a frame is received, an acknowledgement frame is sent back to the sender, indicating that the frame was received without error.

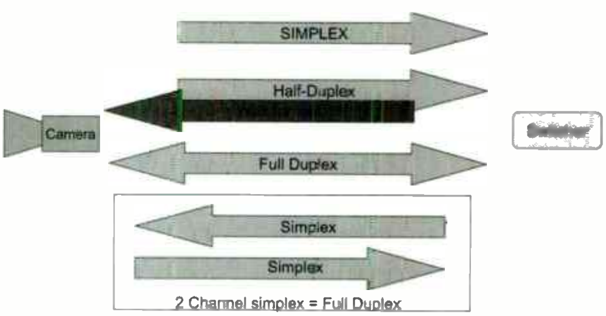
This acknowledgement scheme assures that the wireless connection is reliable and allows the sender the ability to resend any lost frames.

The problem is that the link is half duplex. While the receiver of data is acknowledging the frame receipt to the sender, the sender must wait. Making the sender wait causes the frames to arrive unevenly at the destination. Inconsistent delay variation between frames results in jitter.

Buffering at the send and receive points can reduce delay variation, but it still can present a problem, particularly as distances increase and it takes longer for frames to flow from source to destination.

The TCP/IP protocol goes one step further and can add packet reliability, but using the TCP (transmission control protocol) part of the suite to acknowledge receipt of packets and request resends of lost packets.

By using a simplex transmission scheme it eliminates delay variation, as there is no reverse traffic to the main data flow. However, this also means that there is no ability to use in-band management and the link can no longer be used to carry TCP/IP data.



Wouldn't it be better just to suppress acknowledgements of radio frames and IP packets, but have the ability to have in-band back channel capabilities available? In other words, keep the ability to use half duplex but only use it when you need to.

Alternatively, why not create a symmetric or asymmetric, full duplex link using two radio channels? This would yield the best of both worlds. We would have a forward channel that could support the video stream with acknowledgements and other management traffic being sent back on the back channel. The back channel could actually be narrower than the forward channel, as it would not need nearly the bandwidth. For instance a forward channel of 10Mhz could be used with a back channel of 5Mhz.

RTP and Acknowledgements

The RTP protocol supports two types of transmissions, a unicast method which allows for one client to connect and receive the transmission from the source, or there is a multicast method which allows multiple clients to connect to the server. Additionally, the multicast method does not require the source to know the destination IP address, it simply allows clients to subscribe to the stream using a multicast address and associated TCP/IP. The destination then "tunes" into the stream from the source, allowing point to multipoint architectures to be created. The unicast method requires the source to know the destinations address and push the content to the destination.

RTP is a protocol which essentially rides on top of a transport protocol such as UDP (User Datagram Protocol) or TCP (Transmission Control Protocol). An RTP application creates RTP data packets which are then wrapped with TCP or UDP headers to create the desired packet or datagram. TCP provides error checking and congestion control. The problem is that in a real-time setting, those checks and balances can be quite costly and be the source of delay and jitter. UDP makes absolutely no guarantee of time or control or sequencing or even an assurance that the datagram will arrive, and if the data is lost, no effort is made to resend it. Acknowledgement of the receipt of a UDP packet is pointless, as a lost packet will not be resent anyway.

By streaming content using RTP using UDP, a true half duplex link is possible, eliminating jitter and delay variation.

Why do I need bi-directional communications?

When we talk about bi-directional networks, we really are referring to a fully TCP/IP compliant network that is capable of transporting data using IP v4 or v6. By using TCP/IP and presenting an Ethernet connection at each end of the link, we are free to incorporate any of the millions of IP capable devices or protocols that exist.

Bi directional communications can be useful in many ways in HD production. Obviously you can replace the camera control functions that the out of band solutions provide. But there are many other less intuitive reasons to use a fully bi-directional network:

Adaptive control – Virtually all codec have the ability to vary their compression rates in order to take advantage of changing network conditions. If the decoder is able to provide real time feedback on network latency, congestion, delay variation, etc, the encoder and retune itself to accommodate these conditions.

For instance, if interference reduces available throughput of the wireless link, the encoding can increase compression, therefore reducing load and delay over the link, without dropping frames. Then, when the interference subsides the compression level can be reduced, increasing image quality.

Remote control – The entire encoder/transmitter application can be remotely administered over the link. Camera operation can be completely remote. Not only telemetry (P-T-Z) but also channel selection, rates, input source selection, Quality of Service, and more.

Sensoring – An intelligent encoder/transmitter can serve as a platform for much more than just HD stream transport. Addition of technologies such as RFID, motion sensing, environmental monitors, and the like could provide critical information needed to better manage your resources.

Audio or Video Conferencing – With the availability of a bi-directional network, Voice Over IP or back channel streaming are possible, allowing camera operators to receive audio direction from directors or even to watch a stream of post production video.

A Contentious Issue

Standard WLAN technology such as IEEE 802.11, or WiFi, is designed for data. Data can wait. Delay variation does not affect data like it does audio and video streams. The very nature of the protocols used in 802.11, 802.16 and other similar protocols allow for contention between network nodes for transmission time.

In 802.11 all stations associated with an access point vie for time using collision detection and back off scheme. Even with Quality of Service management in place, unless a high bandwidth streaming node is given 100% of the available bandwidth on the network, there will be delay variation resulting in buffering and packet delays or drops.

In 802.16, the protocol is a little smarter in that it uses time slot reservations, but even this scheme is not appropriate to high bandwidth streaming media. As more nodes enter or leave the network, the actual time that any node will receive their slot will vary from cycle to cycle, even with Quality of Service management.

Unfortunately, if the streaming media requires all of the available bandwidth of a channel, then contention cannot be tolerated…At least on the forward (data carrier) channel.

This means that each encoder/transmitter, or stream source, should have its own dedicated channel for data. That is not to say that each stream needs its own dedicated channel, just that each source needs its own channel.

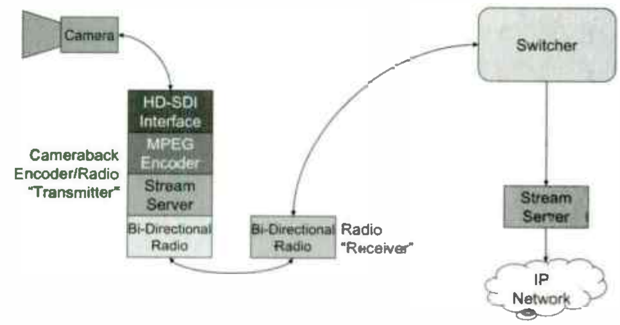
Consider this example: If we have an encoder/transmitter which also has the ability to encode three simultaneous streams from three different sources. Let's say that the radio link supports a total bandwidth of 30Mbps and each stream needs 10Mbps. Using RTP multicast, we can actually stream all three streams can be transported over the same wireless link with good performance. The RTP server can sequence the packets for each stream for consistent delay variation. As long as there are no acknowledgements, or the acknowledgements are sent back on a separate back channel, no interruptions in the stream will occur and jitter buffering can be kept to a minimum.

802.11 and 802.16 cannot effectively support high bandwidth video streaming, even with QoS, in a network with more than two nodes. This does not mean that there is no place for standards based networks in wireless HD production, only that the contention management schemes of these protocols is not ideal for video streaming.

Why "Double Compress"? – Stream Direct to Consumer

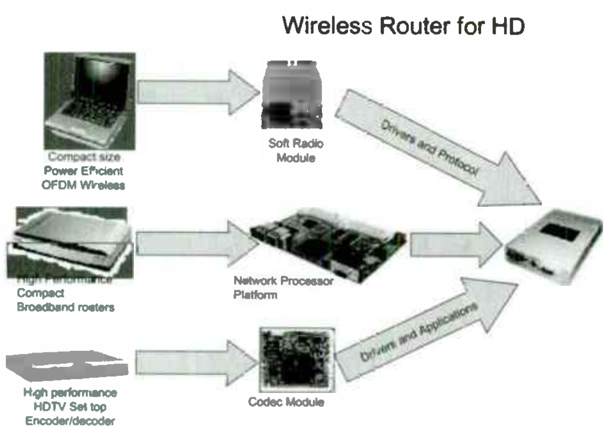
As mentioned earlier, most competing wireless HD production products use MPEG encoding to compress their HD-SDI stream before transmission, and then decompress it before converting back to HD-SDI at the receive site. Since this is done in proprietary equipment, it cannot be routed to the end user without recompression after it is mixed in post production.

If a native IP, bi-directional network is used for the wireless HD production link, and MPEG4 or Motion JPEG2000 standard encoding occurs at the source, then the stream can be routed, manipulated, and streamed directly to the end user without recompression, resulting in a higher quality image.



Soft Radios, Network Processing Platforms and Set Top Boxes

Just because the 802.11 and 802.16 protocols are less than ideal for wireless HD production, does not mean that the components that support these protocols cannot be used. Over the last several years there have been incredible innovations in the WiFi component world. Thanks to the volume economics of millions of laptop computers all looking for the nearest hotspot, manufacturers have created soft radios that incorporate bleeding edge features at commodity prices.



The soft radios of today have the ability to be repurposed for the needs of the wireless HD production market. These are OFDM components with great transmit power, incredible sensitivity, good noise rejection, and drivers which can be rewritten to support modifications to the 802.11 MAC which make the radios highly appropriate to this application.

For instance, the 802.11 protocol is designed for public service broadcast advertisement (beaconing), which allows clients looking for a hotspot to find the SSID of an access point and connect to it. In a wireless HD production application this feature is not required. A soft radio allows you to suppress this feature. The same is true with all sorts of management overhead like acknowledgements, polling, etc.

A noiseless, jitter free wireless environment can be built using these components and modifications to the protocols.

Compare this to custom built radios which need to be built in low volume for the broadcast industry. There is no volume application to support driving the cost of these parts down, so the unit price has to be high.

An added benefit of using these commodity parts is that there are a broad array of compatible components such as up/down converters, cavity filters, amplifiers, antennas, cables and other accessories that can be used with them.

Combine these soft radios with advances in components for embedded network processing, and you are able to deliver amazing performance in extremely small footprints. As an example, for the Cable Set Top Box market, the major CPU manufacturers have introduced network processing engines, based on ARM technology, that integrate 8 to 16 micro engines. Each micro engine is capable of processing network packets at extremely fast rates. The multi-threading design of these processors makes them ideal for applications such as video streaming.

Additionally, these network processors use the Linux operating system, which greatly accelerates application development, and provides a pool of open source applications and developers for the platform.

Finally, with the advent of HD Television sets and satellite set top boxes, low cost, high performance MPEG4 codec chips are coming to market. Not only are these high performance parts, but they are much smaller than the discrete DSP components of the past.

Combine the OFDM soft radio components, the network processors, the highly integrated video/audio encoding chips, extremely flexible Linux development environments and you have everything required to deliver low cost, high performance wireless HD video streaming for broadcast production.

Modulation, Distance and Noise

Rates and Schemes

As stated at the beginning of this paper, most wireless technologies for HD are focused on using high order constellation QAM modulation, in order to maximize data rates. Quadrature amplitude modulation (QAM) is a modulation scheme which conveys data by changing the amplitude of two carrier waves. These two waves, usually sinusoids, are out of phase with each other by 90° and are thus called quadrature carriers. For instance 128-QAM or 256-QAM are what are used for digital cable TV.

Unfortunately, if the mean energy of the constellation (EIRP of FCC regulations for band) is to remain the same (by way of making a fair comparison), the points must be closer together and are thus more susceptible to noise and other corruption; this results in a higher bit error rates and so higher-order QAM can deliver more data less reliably than lower-order QAM.

It is for this reason that at lower rates, where reliability is more important than throughput, 4-PSK types of modulation, or QPSK Quadrature phase-shift keying, is used. QPSK is much more tolerant of noise and corruption, as the points in the constellation can be spread further apart and, therefore, easier to discern from one another.

We have found that QPSK modulation actually is the best trade off between reliability and throughput.

Modern OFDM routers can use a variety of modulation schemes under different conditions and can even dynamically adjust between schemes as conditions vary.

A good wireless router will give you the option of specifying either automatic rate negotiation (modulation scheme selection) between the two nodes in your network, setting a cap on rate (locking down to more reliable, less complex, modulation schemes), or fixing the rate (modulation scheme) at a particular level. The dynamic negotiation of modulation scheme can be accomplished in a number of ways, but almost all use a combination of receive signal strength and bit error rates. After a set threshold is met in both criteria, the rate is adjusted upwards. If either the signal strength drops or the bit errors increase, then the rate will fall back.

In most cases dynamic rate negotiation is something that is not desirable. When a rate is escalated to and fails, then data will be lost, resulting in delay and jitter. It is always suggested to lock the rate or set a cap rate so that escalation to an impossible rate will never occur.

Channel Width

Another element of throughput is channel width. In OFDM (Orthogonal Frequency-Division Multiplexing) carriers are spread out from a center frequency to several sub-channels. The more sub-channels you have available, the more throughput is available.

802.11 and 802.16 define channel width as 20MHz. A "turbo" channel is the bonding of two standard channels into one 40MHz channel. A ½ channel is 10MHz wide and a ¼ channel is 5MHz wide.

The highest data rates in the 802.11 world, are achieved with a 40MHz channel width and 64-QAM modulation (theoretical rate of 108Mbps including overhead).

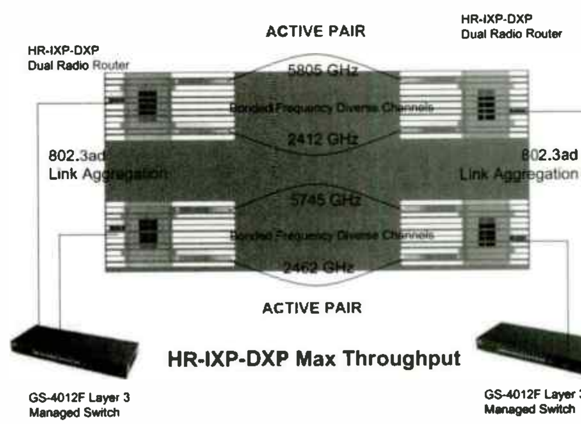
In the real world broadcasts would be carried on 5,10, or 12MHz channels for licensed frequencies and 20 or 40MHz in the unlicensed frequencies.

Most customers would target a 10MHz channel width. Using QPSK modulation, we could expect an actual throughput of approximately 10Mbps. This would be highly resilient. Using QAM-64 the actual throughput would be approximately 30Mbps, but would be less reliable and have a shorter range.

Link Aggregation/Interface Bonding

Modern wireless routers act on TCP/IP packets, allowing them to route these packets over one or many physical interfaces. This is one of the benefits of using a standards based, bidirectional, link for wireless HD.

Link Aggregation, or interface bonding, allows the router to abstract the underlying physical interfaces into one logical interface. This is accomplished using either IEEE 802.3ad or similar proprietary protocols at Layer 2, or by using IP filters and routing table modifications at Layer 3.



In either method, interfaces are grouped together into logical trucks. The router monitors the condition of each physical interface in the trunk. If the link is up and running, then traffic is forwarded over it and the interface remains active in the trunk. If the link is down, traffic is routed to alternate interfaces and the interface is removed from the trunk.

Latency, throughput, bit error rates, etc can all be used to monitor the status of each interface.

Above this is a load balancing algorithm that systematically shares traffic between all interfaces in the trunk. This is generally a round robin scheme rotating packet forwarding to each interface in the trunk sequentially.

By using link aggregation, the bandwidth of multiple interfaces can be combined.

The benefits of link aggregation are two fold: obviously when all interfaces are functioning, you have the combined throughput of all the interfaces; but you also have improved reliability, making the use of wireless possible under conditions that might not otherwise be possible.

For instance, it is possible to run two, three or four radio interface modules in the same router. Each interface could run on a separate band in any of 900MHz, 1.76GHz, 2.225GHz, 2.3-2.4GHz, 4.9GHz, or 5.1-6.1GHz spectrums.

Running this way would allow for frequency diversity.

- Inference on one band would not necessarily affect the other band.
- Roaming from line of sight to non line of sight could be supported by moving between a low frequency band like 900MHz and a higher band like 5GHz.
- Channel availability would be greatly increased, as each remote would have many more options available.
- Unlicensed spectrum would be more attractive, as interference issues could be mitigated.
- Reliability would be greatly increased.
- And scalability could become almost limitless.

Distances

Another consideration is distance and obstruction. Signal gain and receive sensitivity are the keys to overcoming distance and obstruction. Of course we are always bound by EIRP rules for any device. To achieve good communication, you need to look at link budget. In simplex systems you only need to consider the transmit gain of the transmitter and the receive sensitivity of the receiver. In bi-directional links, we have to consider both the transmit gain and the receive sensitivity of both ends.

Frequency in MHz	Total Link Distance in Miles	Total Link Distance in Feet	Total Link Distance in Meters	Obstruction Distance in Miles (if any)
5180	0.18	950.4	288	0.09

	From Radio Characteristics Worksheet		Amplifier		From Antenna Characteristics Worksheet				
	Transmit Power in dBm:	Receive Sensitivity in dBm:	Transmit Power in dBm:	Receive Sensitivity in dBm:	Cable Loss in dBi:	Antenna Gain in dBi:	Additional Loss for Obstructions	Transmit Power Budget:	Receive Sensitivity Budget:
Basestation:	26	-90	0	0	-1	12	-10	27	101
Remote:	26	-90	0	0	-1	5	-10	20	94

	Total Link Budget	Theoretical Margin	Free Space Loss	First Fresnel Zone Max Radius in Feet	Obstacle Free Radius in Feet	Earth Curvature Height Over Link in Feet
Outbound Link:	121	25.0479547	95.95205	6.720114019	4.032068412	1.022E-06
Inbound Link:	121	25.0479547	95.95205	6.720114019	4.032068412	1.022E-06

Frequency in MHz	Total Link Distance in Miles	Total Link Distance in Feet	Total Link Distance in Meters	Obstruction Distance in Miles (if any)
5180	0.6	3168	960	0.3

	From Radio Characteristics Worksheet		Amplifier		From Antenna Characteristics Worksheet				
	Transmit Power in dBm:	Receive Sensitivity in dBm:	Transmit Power in dBm:	Receive Sensitivity in dBm:	Cable Loss in dBi:	Antenna Gain in dBi:	Additional Loss for Obstructions	Transmit Power Budget:	Receive Sensitivity Budget:
Basestation:	26	-90	0	0	-1	12	0	37	101
Remote:	26	-90	0	0	-1	5	0	30	94

	Total Link Budget	Theoretical Margin	Free Space Loss	First Fresnel Zone Max Radius in	Obstacle Free Radius in Feet	Earth Curvature Height Over Link
Outbound Link:	131	24.5903798	106.4096	12.26919346	7.361516075	1.136E-05
Inbound Link:	131	24.5903798	106.4096	12.26919346	7.361516075	1.136E-05

A high power transmitter and high sensitivity receiver are preferred for mobile applications, so that a link can be configured with the smallest possible antennas. The downside of high sensitivity is that your radio may be overly sensitive to noise as well. Good noise rejection is also essential. A tunable noise-floor is the best option. The downside of high transmit power is over modulation and low battery life.

At 1.7 to 6GHz, a reasonable assumption of transmitter gain is 24-25dBm, receive sensitivity can range from -90 to -94dBm, and a typical antenna choice would be a 5dBi rubber duck for the transmitter (camera back) and a 12dBi omni for the receiver (base station).

With a pair of these types of devices, you could expect a line of sight range of about 960 meters and a non line of sight range, with a couple of walls, of 280 meters.

For longer distances, higher gain antennas are required. By using 12dBi antennas at either end range can be extended out to 2km and by using 3 ft 32dBi parabolic antennas range can exceed 40 km.

Of course, true non-line of sight requires lower frequencies. Broadband operations down to 900MHz

What's Available Today and What is Coming Tomorrow

Right now routers that provided bi-directional, IP based wireless solutions for HD broadcast production are capable of delivering 65Mbps actual throughput. They accept 100Mbps Ethernet input/output and transport it up to 3000 feet using small 5dBi gain antennas on the remote and a 12dBi gain antenna at the base station.

They weigh less than 1 lb and run off virtually any 12VDC power source, drawing less than 8 watts.

Having been derived from 802.11 radio components, the routers can support the 900MHz, 2.4GHz, and 5GHz bands. They also can support the 1.769-1.850GHz band when that band is approved by the FCC for commercial use. Work is underway to port these radios into the 2.025GHz BAS band as well.

A variety of video encoders have been tested with these routers, supporting both proprietary, MPEG4 and Motion JPEG2000, all with good results. HD-SDI to proprietary high quality encoding and decoding, which can be transported over

SD MPEG2 video with MPEG1 audio encoding is also already available as an add-in module to these routers.

By the end of the 2nd quarter of this year, HD-SDI input to an MPEG4 AVC H.264 encoder with full RTP/RTSP streaming support will be incorporated into the routers, in a single, compact device.

The devices of today have the ability to leverage OFDM radio modules supporting a range of frequencies. What is currently available for the market includes:

Frequency	**MAX TX gain**	**MAX RX sensitivity**	**MAX TCP/IP Throughput**	**Most likely TCP/IP Throughput[3]**
900MHz	700mW/28.5dBm	-93dBm	18Mbps	6Mbps
1.763-1.850GHz	500mW/27dBm	-99dBm	65Mbps	10Mbps
2.3-2.5GHz	400mW/26dBm	-94dBm	65Mbps	10Mbps
4.9GHz	400mW/26dBm	-94dBm	65Mbps	10Mbps
5.1-6.1GHz	400mW/26dBm	-94dBm	65Mbps	10Mbps

[3] Most Likely Throughput is based on an assumed channel width of 10MHz and QPSK modulation at a range of up to 1km using a 5dBi gain antenna at one end and 12dBi antenna at the other.

What should this cost?

Today proprietary solutions that are designed from the ground up for the low volume, broadcast market, cost anywhere from $40,000 to $100,000 per link. These solutions have no volume market to help drive down the cost of their components. Their quality is good and they can accomplish what they claim, but they are not economically advantaged.

We think that solutions built using components that are driving by volume applications should cost about ½ of what proprietary, custom built products.

Format-Transparent News Operations

Fred Schultz
Harris Corporation
Northridge, California

ABSTRACT

Until recently the market pressure on manufacturers of digital news technology has been to embed values from file-oriented workflow and IT technology into proprietary video server/editor offerings. Escalating consumer purchase of HD sets has shifted that focus to the need for viable HD production solutions which support mixed SD/HD viewership. This paper examines technologies and workflows for HD local news that offer efficiency, format-transparency and competitiveness while the viewing audience is in the process of migrating from SD to HD.

THE LONG HELLO

For both broadcasters of television news and the digital technology vendors that have traditionally supported them, the run-up to producing local news in HD has been underway a long, long time. Despite the fact that the industry is finally compliant with the requirements for digital broadcasting, few stations have seriously addressed creating local news in HD until very recently.

As a result the customer pressure felt by tech vendors had not been directed so much at end-to-end HD production solutions as on updating offerings to include the advances rapidly arriving from the IT sector, specifically the plummeting cost of digital storage, the rise in networking bandwidth, and the widespread availability of high-speed connectivity. The new generation of file-based server/editor products that resulted were crafted for SD but for the most part also embodied infrastructures sufficiently open and powerful to address the inevitable demands of HD.

THE BROAD SHOULDERS OF NEWS

Within broadcasting it is well known that news operations serve stations as both a financial backbone and the station's public face. A successful and well-run news department can and usually does provide the lion's share of a station's revenue, so that news is rightly regarded by ownership and management as key to present and future profitability. In prime time many of the commercial time slots for a given program are pre-filled by the networks. Slots in syndicated programming may be similarly pre-filled, or the show itself must otherwise be paid for. In contrast, during local news all commercial time belongs entirely to the station.

Of course, not all local newscasts can generate the same level of profits. In any given market the cost of running a news department might be roughly equal from station to station, but the ad rates that can be charged scales in proportion to the show's viewership. Being No. 1 (or a competitive No. 2) brings a disproportionately large reward, and a powerful motivation to hold that position.

Along with the revenue that news (potentially) provides, news programming also provides its station with a public face and brand identity. News consumers today are awash with attractive options ranging from print media through in-depth programming on public and satellite radio, to the ubiquitous and evolving Internet, with options in development now for cell phones and other follow-on mobile technologies. Television is just one of many players offering the public enlightenment about news events.

What television succeeds at far better than any of its news-vending competitors is the ability to extend a feeling of community into the lives of people who otherwise may share little but geography. Newspapers might competitively inform a region, but the humanizing face of television news can uniquely evoke a sense of belonging and membership across its community.

The competitive challenge each station therefore faces is to build and retain a unique identity that is sufficiently compelling to attract viewers and keep them away from alternatives. The process each station chooses impacts its decisions about when to start producing news in HD – as opposed to simply upconverting its SD feed – and what technology pathway will be most appropriate for their circumstances.

HIGH-DEF EXPECTATIONS, LEGACY REALITY

Because so much of a station's revenue is attached to the performance of its news, and because any errors or problems during a newscast are so visible and potentially costly, news operations view change very conservatively, holding closely to the maxim about not fixing anything that's not clearly broken. And while a substantial number of stations have in fact moved from tape to file-based SD news operations, a very large number of them have instead struggled to postpone the cost and pain of conversion until the need for local HD production forces an unavoidable rebuild.

Back when the number of HD receivers in a community was small, upconverting the SD news feed to HD was clearly an appropriate decision for broadcasters. But the appeal of HD as a consumer product has finally become well established. Consequently, for all stations, irrespective of whether their news is tape- or file-based, each new HD set that enters a home is a challenge to the status quo and an opportunity to seize or lose viewer loyalty. The risk for each station as it plans and implements its HD rollout is that by waiting too long, or by scrimping on resources, once-loyal viewers may discover that a competitor's newscast looks better on their new HD screen.

February 17, 2009, is locked in as the day when all over-the-air broadcasting in the U.S. must be digital. While this regulation does not mandate broadcasting news or any other content in HD, and while legacy televisions may continue to be used via cable boxes or other adapters, the HD penetration that is already underway will spike on the runup to that date. This effectively creates the deadline by which all stations that want to remain competitive must already have their HD news operations in place and up to speed.

The reality on the ground is that today the hardware used by stations for news ranges from color-under analog tape to HD-ready server/editor systems. In an industry that acknowledges how rarely anything is done exactly the same from station to station, clearly no single buildout pathway can fit circumstances everywhere.

Despite the wide variety of broadcaster's legacy technology and the wide range in their market size and capital budgets, they expect their key vendors to offer solutions that will support straightforward, format-transparent news operations. For years most manufacturers have been pushing their core product line in that general direction, and in the process have come to recognize two key benchmarks.

Operational Benchmark #1

> Operators want to create all news elements just once – in HD – using the best content available at the moment, without any added steps or effort.

The fact that HD is destined to become the viewer's standard for news underscores the need for HD to become the domain in which all news will be shot, produced and edited. The SD viewership will be well-served by this process since SD derived from HD masters can be as good as, and is often better than, content created from equivalent SD-only processes.

While shooting in HD will be the standard, there will always be a need to accommodate content of different formats and resolutions. Even material shot on a cell phone has found ready acceptance when it is the best content available. Simultaneously, new formats (and their underlying codecs) will always be arriving. More and more content is being acquired from a growing range of non-industry sources.

Broadcasters expect their technology to accommodate the growing range of content transparently. Ingest must be simple and editing must be straightforward, with clear and flexible options for handling scaling and aspect ratio. And they expect that neither the operator's workflow nor the system performance should be slowed to accommodate any kind of content.

Furthermore, broadcasters have made it known that news production systems must be open to technology built outside the traditional tight circle of broadcast vendors. Apple's Final Cut Pro is currently the most obvious example, but will certainly not be the last.

Operational Benchmark #2

> During on-air operations, the system will transparently output the best available content – principally but not always HD – and auto-create SD and HD feeds with appropriately templated graphics, resolution and aspect ratio.

All portions of the system must intelligently and flexibly accommodate the best available content. When the moment comes to air a story, the system needs to not only play the HD edited content, the story graphics, SD from the library, a download from the Internet, or a clip or still from a mobile phone. It must also format the output intelligently, flexibly and transparently. It must perform all up/down conversions along with automatic aspect ratio conversions, and do so without needing input from the operator. The system should also accommodate linked SD and HD graphics templates so the same text content will generate separate looks optimized for 4:3 and 16:9.

AREAS OF DEVELOPMENT

The progress that vendors are making towards meeting these two benchmarks is the cumulative result of efforts across a number of areas of development.

File-Based Ingest

Ingest of HD news content will initially follow processes developed with file-based SD. External news service feeds will continue being pushed into dedicated local feed servers, and must then be processed into the news server. Locally shot material will either arrive through the door on drives or memory cards, or will be transferred from the field into the server via network connectivity.

Field-Generated Metadata and Selection Support

HD cameras are going to make widely available a number of efficiencies that were initially developed for SD operations. They can automatically embed a growing range of metadata along with the video in their media files. The ability to capture focus, aperture, zoom and color balance makes it easier to match shots, which speeds the edit and improves the product. And while automatically logging the GPS coordinates of every shot may seem of little daily value, it would certainly simplify assigning re-shoots or follow-on stories to a different crew, and could supply crucial documentation in the legal face-offs that good news efforts occasionally generate. A separate feature that enables the camera operator to mark material of high value while rolling has already proven invaluable for speeding the editing process and shortening total time to air.

HD cameras are also enabling efficiencies by onboard generation of low-bitrate proxies. Thumbnail stills and low-bitrate video offer workflow options for shortening the time to air that range from selection (and sequencing) of shots while still in the field, to speeding transfer from the field media, through support of editing the proxy video straight from the field media so that only the frames that are actually used require transfer. These capabilities can be of great value for breaking news, but since stations generally want all new field content to be accessible to their full production staff, time and storage capacity for full transfers must be also be accommodated.

Assignment-Generated Metadata

Other than with breaking news, the identity and general shape of most news stories originates at the time of their assignment. Technology is increasingly allowing this information to be exposed not just on the newsroom computer desktops, but through the camera in the field, as well as in the editing environment. This capability can make the process of creating the story and linking all its media both faster and more foolproof, and offers the prospect of giving both the shooters and cutters access to all known relevant information.

File Exchange and MXF

HD news operations stand to be the beneficiary of considerable prior work spent developing Material eXchange Format (MXF) as a container format for digital video and audio media to support standards-based file exchange. Both Sony (XDCAM) and Panasonic (P2) use it to make their content and metadata available to all editing platforms in HD as well as SD. MXF also enables straightforward exchange of files for news operations which span across more than one type of server domain, *e.g.* between Thomson K2 and Harris NEXIO.

Bandwidth Accommodation

Manufacturers were able to successfully upgrade their early-generation video servers into today's flexible, high-performance offerings thanks in no small part to the commoditization of Gigabit Ethernet, with its order-of-magnitude increase in bandwidth. And of equal importance to the increase in bandwidth were improvements in Ethernet infrastructure, like managed switches that provided the robustness and fault-tolerance necessary to support on-air operations.

But even for SD-only operations the bandwidth boost provide by GigE was quickly challenged by customer demands for systems with more inputs, more editing and more storage. The first solution for working around bandwidth limitations was to create low-resolution proxies for viewing and editing. This process added a non-trivial overhead to accommodate creation, storage and management of the proxies and EDL conformance. But the benefits that browse enabled – increased availability and lowered cost of access – were overwhelming.

The bitrate needed for browse viewing alone is fairly low, but editing's need for frame accuracy and picture detail led to common adoption of 1.5 Mbps MPEG-1 to service both. Recent developments with MPEG-4 have begun opening the door for even lower bitrates that support both editing and higher-quality viewing. But whether MPEG-1 or MPEG-4, browse operations offer HD news operations a valuable way to view and edit without diminishing the bandwidth available to central system functions.

It is servicing the high data rate required by each HD operation that is the main challenge which must be addressed by all server/editor systems. The solutions they can create are shaped by their system architecture and storage infrastructure.

A solution offered for one platform incorporates outboard compression/decompression of HD to a mezzanine data rate equivalent to SD. This extends HD capability to some of their existing operations without obsoleting much equipment or introducing many changes to their workflow.

For customers of other vendors, accommodation of HD has been made possible by the rollout of 4 Gbps Fibre Channel storage technology. Some of the solution manufacturers who were previously instrumental in expanding Fibre Channel's inherent robust protection capabilities have developed further architectural innovations for Fibre Channel's faster new technology. This has allowed extending additional enhancements in robustness as well as the full new bandwidth to a record number of simultaneous operations by utilizing the flexibility and economy of GigE.

At some point the arrival of 10 GigE infrastructure will provide another rising tide that will again raise all boats. Until that happens, the bandwidth demands of HD give vendors the opportunity to demonstrate their technical prowess and workflow understanding by creating solutions to more fully exploit available storage technologies.

OPTIMAL WORKFLOW FOR THE NEAR TERM

As long as the largest viewing audience is still watching on SD, stations will need to optimize production for a 4:3 audience. An appropriate workflow would be as follows:

The assignment editor specifies a new story on the newsroom computer system, which creates and links a placeholder ID on the server to receive the future edited video. The reporter searches wires and previous story scripts, browses video archives, and then goes into the field with the photographer.

The photographer shoots the story in HD, always framing the content to fit within the 4:3 center of the full 16:9 picture. When great takes worth marking are obvious, a button is pressed to identify them. While riding back to the station the reporter reviews material, picks in- and out-points for desired shots, and saves it all as a sequence of selects on the camera media.

At the station the reporter goes to his desk and begins writing the story. On a desktop NLE the editor locates the assigned story within a list and clicks it, which opens a new linked bin and timeline. The camera media is mounted, displaying proxies of the selects which are dragged into the story bin. Editing is begun, and in the background the NLE pulls full-resolution HD from the camera media. Desired HD content on the server from an earlier story is dragged into the bin. Desired SD content is retrieved from the archive, restored to the server, and dragged into the bin. The edit is completed and saved as HD into the placeholder ID that was created earlier for it. The completed ID arriving on the server generates status updates for the director and producer informing them that the clip is ready to air.

At airtime, the clip plays through the switcher and into the HD feed. The switcher also splits off a separate stream that is downconverted and aspect ratio converted for SD. Downstream of the split, on-screen graphics are fed into both the SD and HD feeds so that templates which make the best creative use of the 4:3 and 16:9 frames can be used (such downstream templating also lends itself to providing news for cell phones and handhelds with graphics optimized for their small screens).

This workflow produces a fully-HD newscast that assures SD viewers of seeing the best-looking, most competitive product (4:3 is also the best framing for small screens in the mobile market). As penetration of HD rises, the value of this workflow will warrant re-evaluation, but for the mid-term shooting in HD and protecting the 4:3 center provides the best news programming for all audiences and the most transparent operational workflow.

THIS MOVE AND THE NEXT

Clearly no one HD technology – or workflow – can fit the full and diverse needs of news broadcasting. Since the viewing public is the ultimate arbiter for justifying all changes, plans must carefully track their desires, and more importantly, their behaviors. HD appears clearly on track to become the transmission format by which everything will soon be judged. Once most viewers start watching on HD, whether they will come to prefer news framed and shot in 16:9 or whether they will continue to favor the look of protected 4:3 will drive the next change in workflow.

High-performance HD News Production and Broadcasting Linking Plural Servers

K.HIROSE, Y.IMAZEKI, T.YAMADA, A.NARUMI, H.TABUCHI,
H.SATO, J.KIYOHARA, M.SHIBATA, S.MATSUDA
NHK (Japan Broadcasting Corp.)
Tokyo, JAPAN

Abstract

NHK is converting to tapeless broadcasting with the aim of configuring a new broadcasting system. Recently, NHK has developed and installed a high performance, HD (high definition) news production and broadcasting system (HD server system) linking multiple video servers at the BSNC (Broadcasting Satellite News Center) for satellite channels. This is a total system that enables recording, editing, and playing out, and features HD nonlinear editors with powerful video effects, and SD (standard definition) nonlinear editors (proxy editors) for quick editing. By combining an additional standard-resolution server and a low-resolution server for PC previews, NHK was able to create a server system more compact than an HD server alone. The system is also closely linked to NHK's NDACS[1] (News Data Acquisition and Control System), which is the main system for news. The system has such highly cooperative functions with NDACS as an automatic recording function based on the NDACS schedule and a script link function that can play out edited footage just by affixing a thumbnail to the e-script[2]. This system is capable of creating an extremely efficient and stable workflow for broadcasting by closely linking the meta-data[3] stored on the NDACS to the server footage.

[1] NDACS
A support system for creating and broadcasting news accurately and promptly. It supports all work involved in news reporting, such as coverage schedules, meta-data represented by manuscripts and rights information, and the news script. NDACS stands for News Data Acquisition and Control System.

[2] E-Scripts
A system that creates news scripts on the computer before broadcast and unfolds them during broadcast at the touch of a single button. It describes video switching effects and audio level, specified moving images, still images, and VCR and server startups.

[3] Meta-Data
Here, meta-data means text information that supplements video information. It also refers to data to be searched. The meta-data of filmed footage includes the title, filming date, filming location, etc. It can also include rights information such as a person's likeness, copyright, and conditions of use for each video. Meta-data comprises both data entered automatically and manually.

Turning to Tapeless System

NHK broadcasts approximately 17 hours of news per day for satellite channels from the BSNC. Many broadcasters play out footage with broadcast -standard VCR's tapes during news programs, but in recent years, broadcasting stations in some countries have introduced tapeless broadcasting systems for news production and broadcasting. Replacing broadcasting tapes with video servers has led to many benefits such as reduced preparation time and more efficient workflow. NHK is converting to HD-quality, not just SD-quality, tapeless broadcasting and is planning to upgrade the news center with a tapeless broadcasting system. So far, NHK has developed and installed an HD server system at the BSNC as the first step (Phase 1[4]).

Overall System Outline

An overall outline of the BSNC broadcasting facilities including the HD server system is shown in Fig. 1. BSNC broadcasting equipment upgraded the BS news controller, studio equipment, CG center equipment, and NDACS to HD in 2006. The HD server system started operating as a resource of the BS news controller this year. In the HD server system, NHK plays out footage at the BSNC using the flow described below.

First, NHK receives raw footage via various transmission routes. The raw footage is encoded into video files at three levels of resolution: high (material server), standard (proxy server), and low (preview server), and each is recorded on the server accordingly. The recording controller has an automatic recording function based on the NDACS schedule, a periodic recording function and an emergency manual recording function.

[4] Phase 1
NHK is promoting the conversion to HD-quality tapeless images and is making progress in Phase 2, a plan to equip the whole news center with a tapeless broadcasting system. In Phase 1, NHK installed the HD server system at the BSNC.
In Phase 2, NHK will verify the operability of the equipment and develop an HD server system at the entire NHK news center for terrestrial broadcasting.

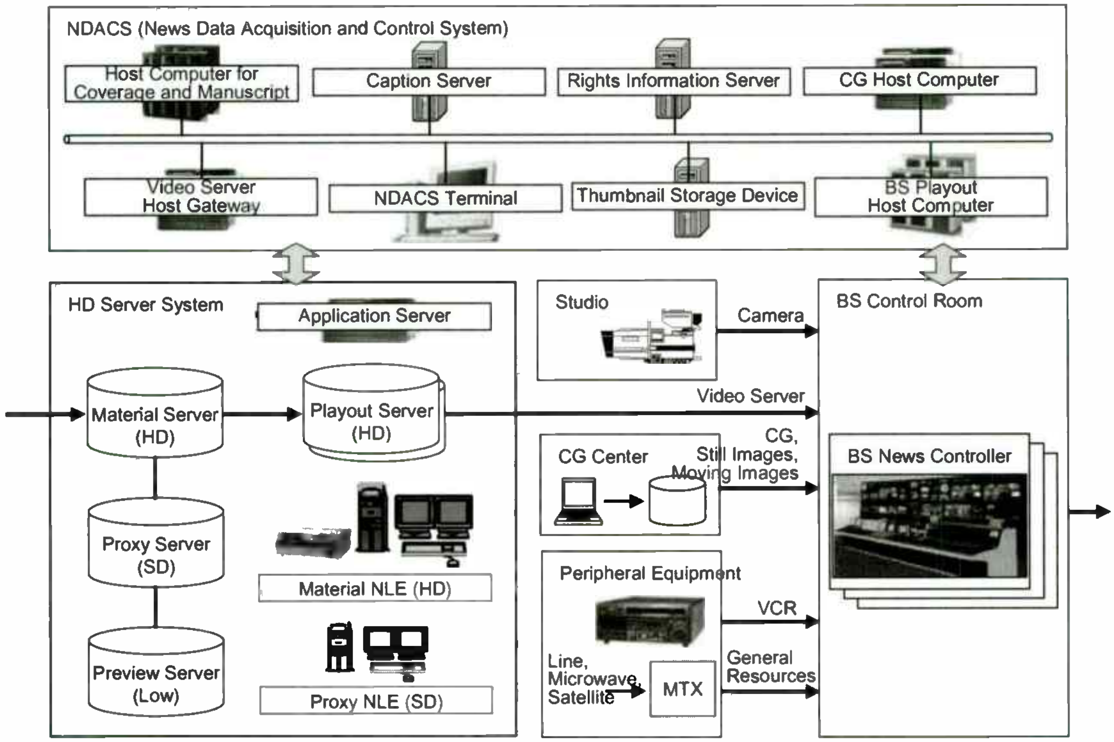


Fig. 1: Overall Outline of BS News Center

Next, the editing operators revise the raw footage recorded on the server (recorded footage) using the HD nonlinear editor and the proxy editor. They divide the use, depending on objective, between the HD nonlinear editor with high video effects, and the proxy editor for quick editing. They can use the proxy editor merely to occupy bandwidth narrower than the HD nonlinear editor, making the system more compact. After editing, the material server immediately transfers the raw footage to the HD server for playing out (playout server) as edited raw footage (edited footage). It also simultaneously transfers the edited footage to the preview server.

Supervisors and other people responsible for news programs create e-scripts such as news scripts with resources while previewing the edited footage on the preview server with an NDACS terminal. The e-script contains high-quality functions linked to the HD server system and the NDACS, such as an e-script link function. The e-script can play out specified edited footage merely by affixing its thumbnails. So, they can play out the edited footage at the touch of a single button along with the e-script on the BS news controller.

System Configuration

The HD server system is a total server system that enables recording, editing, and playing out. A diagram of the system's configuration is shown in Fig. 2. The HD server system is broadly classified into four parts: recording, editing, playout, and footage management.

1. Recording System

In addition to material, proxy, and preview servers, which record the raw footage, the system also configures server controllers, down-converters, encoders, transcoders, and control terminals. The specifications of the recording system servers are shown in Table 1.

Table 1: Recording System Server Specifications

No.	Configuration	Recording Format	Recording Rate	Available Memory
1	Material Server (HD)	MPEG2 I-only	100 Mbps	400 hours
2	Proxy Server (SD)	MPEG2 I-only	20 Mbps	400 hours
3	Preview Server	MPEG4/AVC	500 Kbps	450 hours

2. Editing System

In addition to the HD nonlinear editors and proxy editors, which edit the raw footage, the system also configures rendering devices and transfer devices. The various elements of the nonlinear editor are shown in Table 2. A photo of the nonlinear editor is shown in Fig. 3.

Table 2: Various Elements of the Nonlinear Editor

No.	Configuration	No. of Formats	Model	Software
1	Nonlinear Editor (HD)	3	HP xw9300	TGV Aurora Edit HDX
2	Proxy Editor (SD)	10	HP xw9300	TGV Aurora Edit SDR

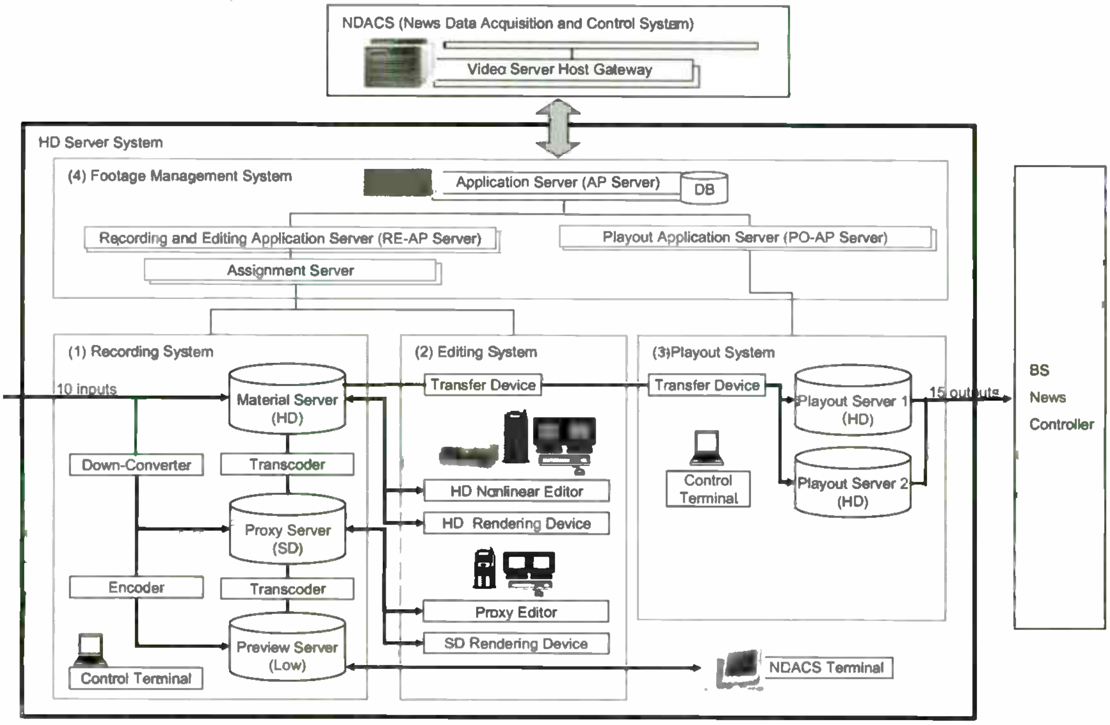


Fig. 2: Overall Outline of HD Server System

Fig. 3: Nonlinear Editor

3. Playout System

In addition to the playout servers, which play out the edited footage, the system also configures server controllers, transfer devices, and control terminals. The specifications of the playout system servers are shown in Table 3.

Table 3: Playout System Server Specifications

No.	Configuration	Recording Format	Recording Rate	Available Memory
1	Playout Server No. 1 (HD)	MPEG2 I-only	100 Mbps	50 hours
2	Playout Server No. 2 (HD)	MPEG2 I-only	100 Mbps	50 hours

4. Footage Management System

In addition to using the application server (AP server) to control all the footage. the system configures a database, a recording and editing AP server (RE-AP server), an assignment server, and a playout AP server (PO-AP server). The specifications of the footage management system server are shown in Table 4.

Characteristics

The HD server system has the following characteristics.

■ Compact design

By recording the raw footage to both the material and proxy servers. the system reduces the load on the material server, making it more compact than an HD server-only configuration.

■ High reliability

To ensure continuous operation, the system is configured so that each subsystem functions independently and can continue operating even when there is a system malfunction.

- By separating the material server and the playout server, the system is able to continue playing out the edited footage on the playout server in case of something trouble with the material server.
- By duplexing the playout server, the system is able to continue playing out the edited footage on one of the playout servers in case of something

Table 4: Footage Management System Server Specifications

No.	Configuration	No. of Formats	Application
1	AP Server /Database	Single-format duplex	Server combines and controls information throughout entire system. Also interfaces with NDACS.
2	RE-AP Server /Assignment Server	Single-format duplex	Server controls footage within material, proxy and preview servers.
3	PO-AP Server	Single-format duplex	Server controls footage within playout server.

trouble with another playout server. In emergencies we can play out footage on the material server even if, by some chance, both playout servers should malfunction.

- The system can play out SD-quality footage on the proxy server in emergencies, in case of something trouble with both the material and playout servers.
- The system can continue operating without greatly affecting the HD recording or playout systems in case of something trouble with the proxy and preview servers themselves.
- By using RAID6 [5] (Redundant Arrays of Inexpensive Disks, level 6) in the storage device, the system can continue operating in case of something trouble with two recording disks within the same blocks.
- By using the subsystem server, the RE-AP server and the PO-AP server, the system can continue operating without being greatly affected in case of something trouble with the AP server.

■ High usability

To enable compatibility with product upgrades to be implemented in the future, the nonlinear editor is not customized and conforms to standard product specifications.

■ Improved user interface

The control terminals are equipped with easy-to-use GUI (Graphical User Interface) so that there is no need for people operating the system to be highly specialized. Some monitors use touch panel screens.

■ Environment and security

NHK installed a special new rack room for the system, which enables it to function regardless of earthquakes, dust, or high temperatures. The rack room is locked for extra security.

System Functions

This section explains the HD server system's functions and outlines the workflow for each system.

1. Recording

So far, operators have manually recorded vast amounts of raw footage to VCR's tapes delivered from within Japan and abroad. When introducing this server system, we have considered various methods of uploading the raw footage more efficiently to the video server. This directly affects the efficiency of recording operations to the video server from tapes.

Normally, if supervisors and other people responsible for news programs want to record raw footage (either to tape or server), they create the meta-data at the NDACS, which is the main computer system for news. The meta-data includes the titles, the recording start and end times, the name of the recorded resource, and the recording medium (i.e., whether to record to tape or server). The recording schedule in the BSNC is determined based on the meta-data. The HD server system obtains the meta-data created by the NDACS via the AP server and assigns server recording time. A maximum of 10 pieces of footage can be recorded simultaneously in the servers, and the schedule is calculated and assigned so as to prevent duplication. As shown in the control terminal screen in Fig. 4, the assigned data is displayed graphically. When the start time is reached, the recording starts automatically. The raw footage is simultaneously recorded on the HD material server, the SD proxy server, and the preview server. The server recording repeats this operation without human assistance or intervention.

Fig. 4: Recording Schedule

The HD server system has a periodic recording function that records predetermined resources at a predetermined time daily. It stores these recording schedules to a meta-file and assigns the filed schedules daily. This function is very useful in recording daily set-time transmissions.

Further, the HD server system has a manual recording (emergency recording) function by which operators can record to the server by pressing the "*REC* button" on the control terminal. This function enables immediate recording without waiting for meta-data from the NDACS if the system switches to emergency coverage. In this way, the system can

---

[5] RAID6

RAID6 is based on RAID5 and generates a second parity information in all hard disk drives in the system, increasing the system reliability. RAID is a hard disk drive array that can be assembled with the following goals:

Mirroring: Automatically copies, thru hardware, all data sent to one hard disk drive to another hard drive. If the first hard drive fails, the second disk, which has a backup of all data stored in the first disk, enters in action automatically.

Data stripping: Increases the storage system performance by accessing two or more hard disk drives in parallel. The files aren't stored in just one hard disk drive; they are split and each part of the file is written on a different hard drive. Since each hard drive will store just a small part of the file and not the whole file, the storage is done faster.

always supply prompt recordings.

2. Editing

The footage recorded on the server can be edited using the HD nonlinear editor and the proxy editor. The HD nonlinear editor uses the HD recorded footage stored on the material server, and the proxy editor uses the SD footage stored on the proxy server. So, they can reduce the load on the material server, making the system more compact than the HD server alone. The nonlinear editor can also edit mixtures of footage recorded on the server and footage uploaded from the attached VCR.

Until now, nonlinear editors have been used in many news production sites, but there have been few editors this closely connected to the server. Therefore, operations linking HD nonlinear editors and proxy editors are groundbreaking technologies. This system enables strong links between multiple editors using editing information called EDL (Edit Decision List). This system achieves an effective workflow using multiple servers. A simple diagram of this editing operation is shown in Fig. 5.

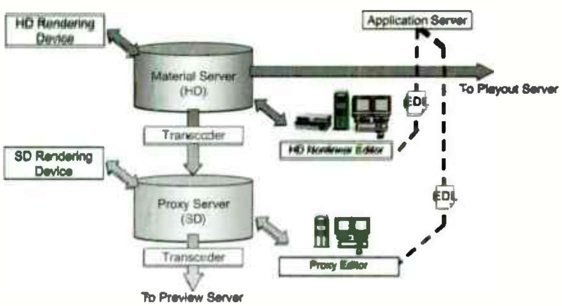


Fig. 5: Editing Workflow

The HD nonlinear editor features various excellent video effects such as blur and mosaic and audio effects such as voice disguise and equalizing. External rendering devices render the effects, and the footage created by this process is called rendered footage. To edit with an HD nonlinear editor, the EDL that the HD nonlinear editor created is sent to the proxy server. When an HD rendering device renders the effects, the SD rendered footage is created on the proxy server at the same time as the HD rendered footage is created on the material server. Depending on the effects, the SD rendered footage is transcoded from the material server or is created by the SD rendering device itself. After editing, when the HD nonlinear editor "publishes" the footage ("publish" means that a recorded footage is edited completely and announces the completion of the edited footage), the edited footage is transferred to the playout and preview servers and is simultaneously created on the material server. All the EDL created by the HD nonlinear editor is sent to the proxy server, and SD-quality edited footage is created on the proxy server as if the proxy editor had edited it.

The proxy editor, in contrast, is an editor that focuses on speed, and edits only simple transition effects such as wipe and dissolve. When the proxy server "publishes" after editing, the EDL is sent to the material server as the edited footage is being created on the proxy server. The material server, which receives the EDL, creates edited footage with the same content as the footage that the proxy server created and sends the edited footage to the playout and preview servers.

These editors can share the EDL even while editing is still being conducted, so the work can be divided such that the nonlinear editor can edit the parts that require effects, and the proxy editor can edit the other parts. In this way, the editing system can perform the special tasks of each editor, greatly contributing to efficiency.

3. Playing Out

The playout server can output a maximum of 15 pieces of footage simultaneously. Normally, the e-script system plays out edited footage on the playout server. When not using the e-script system, operators can play out the edited footage manually with control terminals.

Before broadcast, supervisors and others responsible for news programs prepare the e-script, identifying which edited footage is to be played out and in which order it is to be played. In creating e-script, they determine which news script and resources (camera, moving images, still images, etc.) including server footage to use with the NDACS terminal, and enter the meta-data necessary to broadcast the news program. The HD server system features a script link function. If the edited server footage in the e-script on the NDACS is scheduled to be played out, the label frames in the control terminal on the HD server system change color from blue to red as shown in Fig. 6. This indicates playout of the edited footage for a news program, and means that the edited footage specified in the server control terminal should be labeled. Operators label the edited footage by clicking and dragging the thumbnails for the edited footage specified. Then, supervisors check that the specified footage is registered and finish creating the e-script. In this way, they can proceed in order through "stand-by" and "play-out" of the registered footage according to the progress of the e-script for the news program.

Fig. 6: Script Link Function

4. Footage Management

The AP server on the HD server system manages recorded footage and edited footage by linking to the meta-data on the NDACS, which is a host computer. The AP server also manages the storage period of those pieces of footage after broadcasting. The storage period is relatively short (between two days and a week) due to server memory space. When the period has elapsed, the footage is completely deleted from the server. If operators don't want to lose the footage after the storage period has elapsed, they must copy it to other media such as tapes before it is deleted.

As shown in Fig. 7, the NDACS manages recorded and edited footage on the server by previewing footage with a PC. So operators can preview server footage with their own PC, they can easily manage meta-data like rights information. Rights information, which is particularly important information for the broadcaster, concerns restricted items according to the content of the footage (conditions of use and the right to the use of likenesses, etc., distributed to other broadcasters and overseas).

Fig. 7: Preview Screen on NDACS Terminal

Contrary to expectations, operators can manage cutting units of server footage on the NDACS. The cutting points in recorded footage are detected automatically when uploading to the server, and edited footage is cut by referring it to the EDL. Operators can change the cutting points and thumbnails for each cutting with PCs. In this way, they can manage meta-data such as rights information by units of footage and by each cutting point in the footage, so the NDACS makes meta-data an extremely powerful tool for managing footage.

Introduced Effects

By introducing the HD server system, NHK expects to benefit in following ways.

■ Faster reporting

In addition to changing the workflow for each function, the system dramatically reduces total production time due to "follow-up work". When recording to existing VCR's tapes, operators could not start editing until the footage had finished being recorded, as the tape could not be removed. But when recording to the server, operators can start editing with the nonlinear editor in less than a minute. In addition, when editing with existing tapes, operators need to transport the tape to the control room and to fill it for playing out. But when "publishing" with the nonlinear editor, operators can transfer the footage to the playout server immediately and can play it out at the news controller in seconds. In this way, "follow-up work" will greatly increase the speed of news reporting.

■ Effective use of footage

Because operators can use and store footage on the server simultaneously, they can edit the same recorded footage with multiple editors and can play out the same edited footage to multiple outlets. Editing operators and others can also simultaneously preview the edited footage with their own PCs. In this way, the server footage can be used more effectively.

■ Reduced tape costs

NHK is actively involved in the reusing of VCR's tapes. However, the cost of tapes for news has risen greatly over the years. Introducing the HD server system eliminates the need for multiple recordings or copies of the same footage, so the cost of tapes is expected to fall greatly as demand drops.

■ Combined management of footage and meta-data

The HD server system unifies footage and meta-data by closely linking to the NDACS, and can combine management of resources.

Conclusion

The HD server system makes possible a very reliable and compact system by combining multiple servers and an extremely effective and easy-to-use system by closely linking to the meta-data on the existing system. In using, it is necessary to proceed with thorough verification, and to further improve operability and controllability. In Phase 2, NHK is proceeding to develop and design a more advanced HD server system.

Finally, we would like to thank to makers (NEC, TGV, Grass Valley Japan Ltd.) and all those who worked hard to introduce this system.

News Editing and Transfer System in a Distant Region

Seong-Gyu Jeon, Gi-Sun Yang, Man-Kyu Lee, Hee-Jung Kim, and Sung-Choon Park
Broadcast Technical Research Institute, KBS
Republic of Korea

ABSTRACT

In this paper we introduce a newsgathering system that assists news reporters to easily edit source video in a distant site, and transfer the result to the TV station. We have incorporated IT (Information Technology) components to develop the system. Presently reporters are bothered to edit source material, compress video, and transfer it using several tools. We have integrated file transfer and video editing functions together for seamless workflow. For promptness our system is designed to make it possible one-stop news production outside the TV station. The system also reflects the demands of field workers, i.e., easy usage for beginners and features for speedy news production. One of the most focused is the reliability of the system in the process of transfer. We have implemented several error-resilient schemes in order to cope with unexpected network errors.

Our system consists of remote laptop editing terminals and a newsgathering server. DV-format based editing and analog-to-DV converting device are adopted to capture video from not only lightweight DV camcorders but also old-fashioned ENG cameras into computers. We expect that news reporters and cameramen can take terminals anywhere the Internet is available and both editing and sending will be done on the spot with our newsgathering system.

1. INTRODUCTION

1.1 Background

Since the past, many attempts toward more simplicity and promptness have been made in news production process. However it was not as dramatically as for the last decade when broadcast systems have started to incorporate more techniques and components based on IT. VCR editing and tapes are being replaced by non-linear editing and digital storages and analog video lines and control panels are being changed into network lines and computers. The alternative to satellite transmission, such as Internet file transfer, has been emerged. IT has enabled broadcasters to easily transfer program material over the Internet, not using an expensive dedicated network line. These days the availability of the Internet is propagating at tremendous speed. Most news reporters are carrying laptop computers connected to the Internet for newsgathering. IT infrastructure has already become a part of news production.

There are some existing tools that news reporters outside the station or abroad can utilize to edit video and transfer news materials to the station. They are, however, at a disadvantage in the sense of cost-effectiveness, user-friendliness, the easiness of customization, and security. Professional news delivery products are too expensive to be fully equipped all over the news production workflow. As for satellite transmission, due to high costs and the limitation of coverage where signals can reach, an alternative means, if any, will be a better option; for example, an cable modem Internet. In addition, commercial news editing tools are too complicated for most unskilled users and heavily featured for rough-cut editing.

In KBS it has not been long since deploying IT-based newsgathering components. However, it is easy to see that news reporters make use of nonlinear editing software and professional video delivery devices which automatically compress and transmit video files from a distant place. Recently, what is more, a try to deliver video files to the station by email has been made. In the meantime, we have experienced those tools, coming to wish a more efficient, low-cost, easy-to-use, and customized solution on our own. It is also required file transfer to be robust for network errors.

1.2 Workflow

Figure 1 shows the workflow of our newsgathering system. It is comprised of 5 major steps: video capture, voice recording, non-linear editing, video compressing and transfer, and video tape-out. In KBS it is usual that a newsgathering team consist of a reporters and a cameraman in pair. The reporter is responsible for writing a news article, and he records his voice while reading the article using a recording module. After that, the cameraman captures video materials, edits clips according to the reporter's voice. Once editing is done,

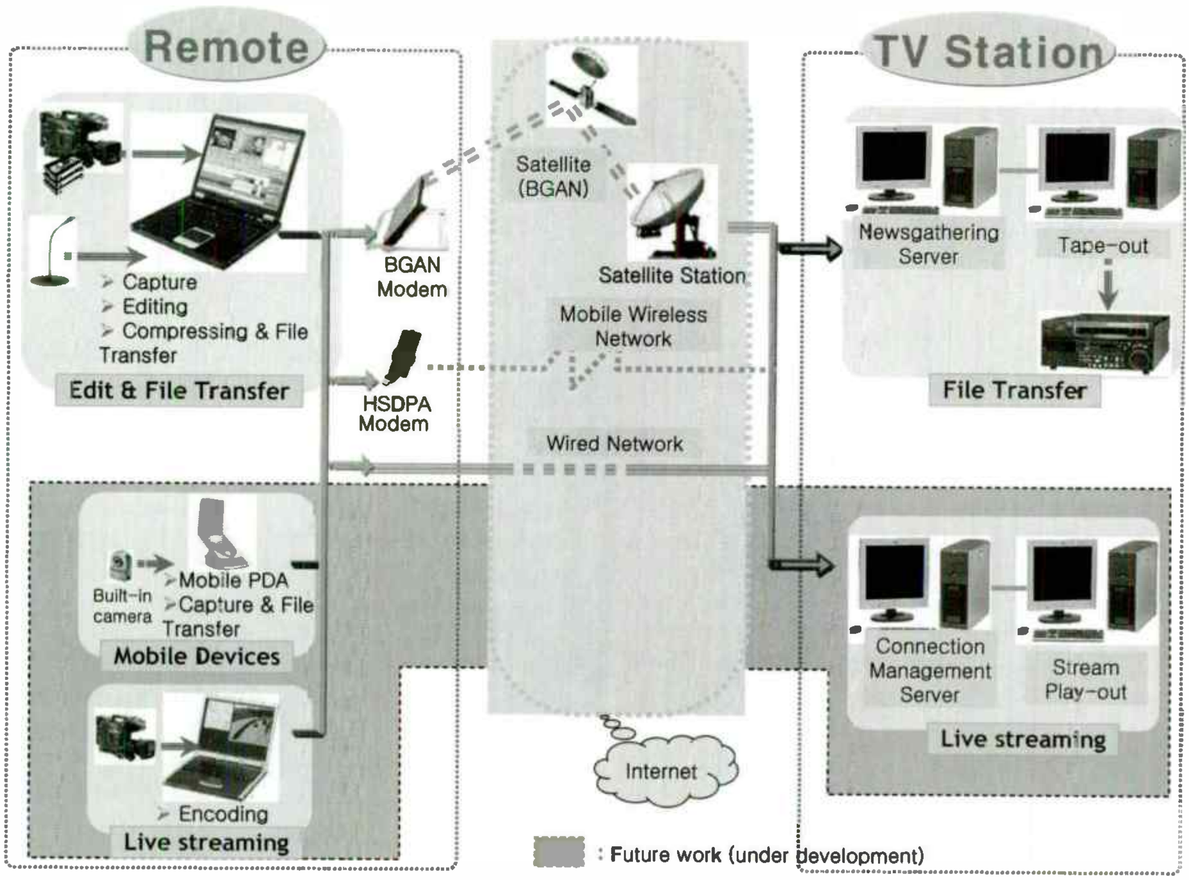


Figure - 1 the workflow of the system

the final video file is rendered through the compressing step. And then, the file is transferred to the domestic TV station using FTP (File Transfer Protocol) over the Internet. The newsgathering server in the station monitors and manages all the file transfer processes. After the transfer is completed, the received file passes through the tape-out step where the video is stored on a VCR tape.

The transfer process can be carried out on any network which supports IP (Internet Protocol). We expect that video transmission can be done at a hotel or anywhere the Internet is wired. We also employ BGAN (Broadband Global Area Network) which is a satellite-based uplink that provides IP network [1]. BGAN will be of great help in overseas countries where the Internet is not available.

In Figure 1, the grey-colored section means future work. We have a plan to extend our system to support live streaming and mobile devices this year. We suppose that mobile devices can be used for an audience participation newsgathering, and TV viewers can create news items about their neighborhood and send them to the TV station.

1.3 Design

It is well known that it is not easy to reflect users' needs when you develop a system. In particular it becomes much more difficult when you step in the implementation process of the system. Therefore we have tried our best to collect in-field requirements as many as possible in the design step. As a result, the basic design concepts are easy usage, one-stop operation, customization for news editing, and reliability. To satisfy these requirements, we have adopted the followings features:

- ✓ *Employment of easy-to-use buttons in UI (User Interface)*
- ✓ *Exclusion of unnecessary features*
- ✓ *3 steps of workflow: capture, editing, and transmission*
- ✓ *Audio-first editing*
- ✓ *3 tracks of editing: a video, a voice audio, and a field audio*
- ✓ *Automatic transmission retrial and file sending continuation when network errors*
- ✓ *Time estimation to render and send files*

Audio-first editing means that a user edits the reporter's voice audio track first and finishes video editing afterward. Detailed descriptions of some of the above features will be introduced later.

2. SYSTEM ARCHITECTURE AND FUNCTIONS

2.1 Overview

Our system is divided into two parts, remote terminal and newsgathering server as shown in Figure 2. The newsgathering server is responsible for the reception of the transfer. The main function of the remote terminal is to edit video material and transmit files. Remote terminals are, for the most part, installed in laptop computers for mobility.

Every work step is implemented as a module which is independent of others. Between modules, data are communicated only through the predefined interface. In this way, system errors are easily corrected when they are found.

As for the transmitting module, we have designed it to work ceaselessly even when network errors occur. Also, in case of low network bandwidth, time estimation feature is provided so that a user can predict total time for delivery. We hope that these features will save trouble for a user to supervise the overall file transfer process.

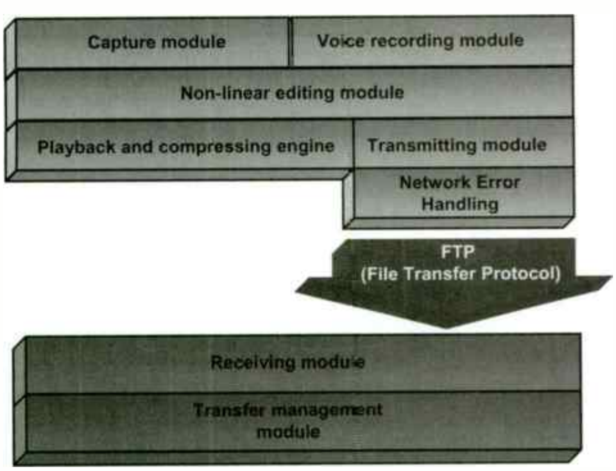


Figure - 2 the system architecture

2.2 Video Capture and Voice Recording

Capture is a step that converts video tapes to DV-format files. Figure 3 shows the video capture UI. The module supports three types of capture of DV-format stream: manual capture (Fig.3-a), automatic capture (Fig.3-b), and batch capture (Fig.3-c). Manual capture is used when you capture video from an analog camera using analog-to-DV converter. Automatic capture is used when you have a DV camcorder supporting DV device controls. Batch capture is used when you do multiple automatic captures at a time. When the capture is completed, you get clips which are element units of editing. The file format of the captured clip is DV-AVI. For the convenience of editing, the capture module has SCD (Scene Changes Detection) feature which divides a clip into sub-clips automatically according to scene changes during video capture.

Figure - 3 video capture user interface

Voice recording module records a reporter's voice for news. This module has 'start record', 'pause record', 'save record' and 'preview record' features to afford convenience to a user. In case that a reporter records wrong or wants to record again, it offers 'record cancel' and 'record continuation' functions at the desired position. Thus the reporter can re-record anytime and anywhere.

In addition, voice recording offers several basic features such as volume level meter, microphone volume control, waveform display, saved audio clip list, and status information (the free space of the storage, time code, etc.).

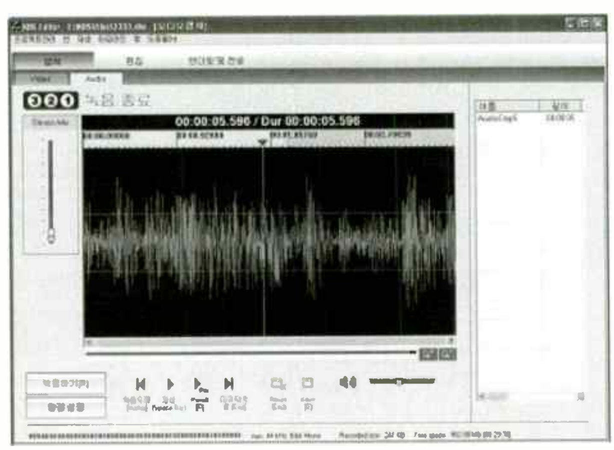


Figure - 4 voice recording user interface

2.3 Non-linear Editing

Non-linear editing is a video and audio editing that allows random access on the source material. In this step, a user cuts clips, merges clips, and gives a transition effect between two clips. All these processes are non-destructive, that is, the actual source files in storage are not lost or modified during editing.

As mentioned before, we employ user-friendly UI design all over the editing operations. Buttons are implemented as big as possible for readability and shows balloon descriptions that guide beginners to easily find what the button means and the short-cut key for that function. Additionally, we remove unnecessary functions which make it complicated to use the system. We provide basic cut operations such as trim editing (see Figure 6), ripple mode editing, overwrite mode editing, 3 point editing, etc. As well as those features, we also provide basic transitions, audio volume controls, shuttling, video still capture, and undo.

Figure - 5 non-linear editing user interface

The layout of editing UI is shown in Figure 5. The UI is comprised of 4 main windows: bin, clip monitor, preview monitor, and timeline. The bin window manages captured clip lists and effect lists. Bin means a folder where clips are stored. The clip monitor is for previewing clips in a bin window, and plays the role of a source VCR in tape-to-tape editing. Besides, it can be used as a preview window of news text to save the space of a computer screen. The timeline is a window where the actual editing operations are performed. In the timeline, there are 3 tracks; a track for video, two tracks for audio editing, which are enough for news editing. For audio editing, we provide a voice audio track and a field audio track, the first one is for editing reporter's voice and the second one is for field sound recorded on the spot. The preview monitor is for the playback of the timeline and used for confirming the result of editing.

Our system supports audio-first editing features, such as track-locking to protect the voice track from unintended modifications and audio track mapping which let the user select audio tracks where he wants to put clips. We classify clip folders, namely, bins, to interview bin, audio dubbing bin, and default bin, according to the audio track mapping properties. And we provide automatic clip trimming which determines the end position of an insert clip by a predefined length when a user just clicks on a start position. It accelerates clip insertion operation.

Figure - 6 trim editing user interface

2.4 Rendering and File Transfer

Rendering and transmission functions are integrated in a UI. The layout of the UI is shown in Figure 7. There are three sections: rendering (Fig.7-a), compression parameters information (Fig.7-b), and transmission (Fig.7-c).

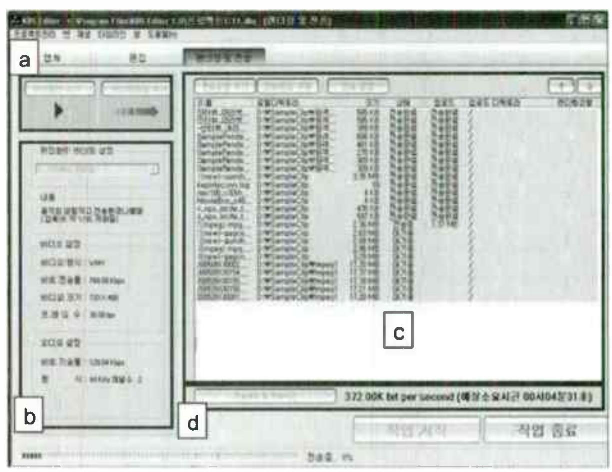


Figure - 7 rendering and transmission user interface

Rendering is a step that makes a file as the result of editing, which includes compressing process. We provide several predefined profiles of compressing parameters classified by video resolution, compressing bit rate, frame rate, etc. A user can easily configure video/audio compressing parameters by a mouse click, and don't have to know about complicated parameters.

Once rendering is completed, file transmission automatically starts. File transmission is to send files to the remote server using FTP. As introduced earlier, we have incorporated some features to cope with unexpected network errors, i.e., 'automatic transmission retrial', 'time estimation', and 'file sending continuation'.

Firstly, 'automatic transmission retrial' is to try to connect with the server automatically every few minutes and re-send the interrupted files when the server-side or client-side network is down. Secondly, 'time estimation' is to estimate total time required for rendering and file transfer (Fig.7-d). This enables a user to meet a deadline by selecting a proper compressing profile: for example, if network speed is good, a high quality profile can be applied, if bad, a low quality one can be used. Finally, 'file sending continuation' is to continue the file transfer from a break position when a transmission interruption occurs by a user or a network error.

Figure - 8 newsgathering server user interface

Newsgathering server receives files from a remote terminal and monitors receiving status. Figure 8 shows newsgathering server UI. The server supports functions such as: 'server log', 'security management', 'user management', 'environment setup', 'connection status display', 'statistical data display', and 'receiving status display'. 'Receiving status display' function shows an administrator transmitted file size, transmission speed, transmission time, received file path, etc.

When the transmission is completed, a user can record the received file on a tape in the VCR using a tape-out feature on the server.

3. IN-FIELD TEST AND RESULTS

We have tested mainly two parts, the quality of compressed video and file transfer speed. In the transfer test, we have performed wireless Internet transmission that uses a BGAN satellite Internet service and wired Internet transmission that uses a cable modem. In the domestic regions, the uplink speed of tested BGAN service is about 128 kbps, and the cable modem is about 320 kbps. We have also made tests in foreign countries (Tianjin in China, Moscow, London, and Berlin) over the wired Internet. In China, the upload speed is 50 kbps. In the other areas, the transmission speed are about 160~240 kbps. In Korea, a new mobile Internet service, namely HSDPA (High-Speed Downlink Packet Access) has been launched recently and the test speed is about 270 kbps. These test results depend on circumstances and we are planning more experiments this year.

We have also made video compressing test using VC-1 Codec [2]. The original file size to encode is 150 MB

Resolution	Bitrate (kbps)	Rendering Time	File Si e (MB)	Video Quality	Block Artifact Existence
720x480	64	4m30s	1.4	very bad	O
	128	4m31s	1.95	very bad	O
	256	4m34s	2.97	very bad	O
	512	4m43s	4.93	bad	O
	768	4m40s	6.67	not bad	△
	1000	4m47s	8.71	not bad	△
	1500	4m50s	12.07	good	X
	2000	4m52s	15.55	good	X
	5000	5m15s	26.67	very good	X
360x240	64	1m7s	1.8	very bad	O
	128	1m11s	2	bad	O
	256	1m22s	2.9	bad	O
	512	1m27s	4.74	not bad	△
	768	1m28s	6.2	good	X
	1000	1m30s	7.1	good	X
	1500	1m39s	8.68	very good	X
	2000	1m46s	10.24	very good	X
	5000	1m50s	17.6	very good	X

Table - 1 Video Compressing Test Results

(DV-format, a minute length). The audio encoding bit rate is fixed as 128 kbps. As shown in the test result (Table 1), when the compression bit rate is more than 1.5 Mbps in SD (Standard Definition) resolution, we achieve a result that satisfies picture quality to broadcast. We have tested with a laptop (CPU 1.86 GHz, RAM 1G) and picture quality test is subject to opinions of video experts.

In addition to the above tests, we are operating a test system in the field and several real-world problems have been introduced. In the first place, network QoS (Quality of Service) problem has been raised. Overseas branch offices have a dedicated network line connected to the station, thus network speed is guaranteed to some extent. However, if transmission is carried out somewhere else, worst cases should be taken into consideration. For instance, in China network status is not good as shown in the earlier test. As to BGAN satellite service, network speed is guaranteed from a remote site to the satellite base station in Netherlands, however it is not between the base station and a domestic region, which causes a bottleneck to overall network speed. As a matter of course, network line status is beyond our capability in sense that it is the responsibility of the Internet service provider. However, some actions can be taken to improve network speed, such as placing a newsgathering server at IDC (Internet Data Center) to provide fast access to the server.

In the second place, tape-out video quality problem has been raised. Even if the compressed video quality is good, the final output will be useless unless tape-out process is carried out properly. We are in testing several video out devices and trying to optimize the solution. The last problem is the compatibility of video file formats. Currently we supports a DV format, however there are some cases that other video file formats created by a third party, such as MPEG-2, MPEG-4, HDV (High Definition Video: a video format designed to record compressed HDTV video on standard DV media), should be imported to our system when it comes to the overseas regions. As to this problem, we are planning to provide a solution this year.

4. CONCLUSION AND FUTURE WORK

We have developed a system for editing and transferring news items easily and quickly in an exterior region. Our system places emphasis on the one-stop workflow, easy-to-use UI, in-field requirements, and system stability. We have tested the system both at overseas and domestic places and have deployed a test system, which is in use in the field. There are some remaining improvements of the system to be made this year.

As mentioned briefly before, we have a plan to extend our system. The main work will be to support live streaming. It means that a news reporter can broadcast live on the spot from a laptop computer. This will include state-of-art technologies such as H.264 video coding, and focus on a low-bandwidth network situation. Also, we will be developing a version working on a mobile platform.

While the current system only supports SD, we are planning to support HD (High Definition) cameras. Current laptop computers are not able to process HD video smoothly. Therefore we are now making an effort at supporting proxy video rather than HD sources themselves. We will support MPEG-4 proxy format, which is widely used in tapeless HD cameras. As to HDV, it does not have a proxy format supported and it should be down-converted to a DV format before being imported to our system.

KBS has a plan to upgrade the legacy news production line to a digital news room system in a few years. At that time, we expect, our system will be customized to operate together with the news room system. This job will include the implementation of handling metadata, such as cataloging information about news materials. Since the reason, we will define a first version of metadata template this year and support it in our system.

REFERENCES

1. http://broadband.inmarsat.com/
2. SMPTE 421M-2006 Television – "VC-1 Compressed Video Bitstream Format and Decoding Process"

ACKNOWLEDGEMENTS

This work is a joint project between KBS and Honest Technology, Daejeon, Korea

Audio Solutions for Television

Monday, April 16, 2007
10:30 AM – 12:00 PM

Chairperson: Bruce Jacobs
Twin Cities Public Television, Saint Paul, MN

Seeking the "Holy Grail" of Loudness Management: Effective Yet Unobtrusive Control

David Reaves, TransLanTech Sound, LLC, Recklinghausen, Germany

Audio-Video Timing

Paul Briscoe, Harris Corporation, North York, Ontario, Canada

Specifying Audio for HD

Thomas Lund, TC Electronic A/S, Brabrand, Denmark

SEEKING THE 'HOLY GRAIL' OF LOUDNESS MANAGEMENT: EFFECTIVE YET UNOBTRUSIVE CONTROL

David P. Reaves, III
TransLanTech Sound, LLC, Recklinghausen, Germany

INTRODUCTION

It's been over 40 years since loudness research became an industry focus, and it was not long after that automatic loudness control became a reality in broadcasting.

At the 1966 NAB convention held in Chicago, a paper was delivered by researchers from CBS Laboratories, updating their ongoing work in the field of broadcast loudness measurement and control.

While some of us in the industry look upon the CBS work as a benchmark in the serious application of psycho-acoustic research to meet real world needs, we've come a long way since then, with radical changes in technology, delivery and source quality.

With DSP techniques we are now able to perform complex signal control more easily and, if designed properly, more transparently than ever.

In this paper, I will outline some ideas that have been used in the design of minimally-audible level control along with thoughts about how such concepts might be applied to loudness control.

THE IDEAL LOUDNESS CONTROLLER

If we were to describe an ideal loudness controller, we might use a description like the following:

1) it would keep the bulk of audio elements within the 'comfort zone' of the typical audience member;
2) it would not draw attention to its operation;
3) it would be compatible with and serve to integrate multiple programming types;
4) if multiple units are used in series at different points in the program chain, only the first unit would make dramatic changes (if needed), while succeeding units would be able to automatically recognize the presence of previously processed audio and reduce or disable their control actions accordingly;
5) it would be open to some form of 'undoing,' by use of dynamic metadata, if desired.

Let's go over each of these features individually, because at first glance they seem a bit ambitious.

Point 1, Keep Levels in the Comfort Zone

Point One is a given. Keeping the levels 'comfortable' has to be the primary goal. But don't we need to define 'comfortable' levels and 'typical' audience members? I would say that those are subjective judgements best left to those making programming decisions. What we *can* do is to give our best effort to offer the tools to work effectively within those definitions, whatever they may be. Here's a better question: Do we want or need precise, tenth of a dB accurate control of loudness levels? Fortunately, probably not. Human hearing is rather forgiving, especially over time. The audio events which most need our corrective attention are abrupt, extreme variations and sounds of lengthy duration that, on the loud end, are energy-dense and on the soft end contain intelligence that is critical to the understanding and/or enjoyment of the program, such as softly spoken dialog.

A certain amount of homogeneity adds to the comfort level of the program, but too much can detract. Humans seem to like a bit of variety, so we're not looking for continuous RMS density that would result in a dynamically boring and relentless signal (we can listen to broadcast radio for that!), however the dynamic extremes *must* be controlled, ideally in an artful manner.

This we can do.

Point 2, System (In)Audibility

This is a serious requirement, and will be the focus of this presentation. Why would we want loudness control to be unobtrusive? Beacuse it's long been accepted that processing's artifacts can create *audience fatigue.*

I have a pet theory:

If we can avoid giving the audience any, even sub-conscious reason to reach for the remote, they have one less opportunity to change channels.

A great level controller, like a great edit, is never noticed. An aggressive, audible approach to processing cannot help but alienate certain segments of the audience, but with today's technology there is no reason for a processor to have a telltale 'sound' unless that sound is desired as an effect.

Subtle, almost imperceptible control is possible by using well-chosen sets of detection and control methods, which this paper will ultimately outline.

Point 3, Multiple Program Compatibility

Our idealized loudness controller needs to be able to transition artfully between multiple program sources of varying degrees of quality, spectral balance, level and dynamic range, and make it all one consistent program. All the while still conforming to Point Two.

Point 4, Stackability

The envisioned loudness controller will have a system that analyzes incoming audio, recognizing programming that is already within the comfort zone, and reducing or even shutting off the process for the duration of such programming. As will be seen, this feature is a logical outcome of our particular response to Points One and Two, with the further advantage that a system so designed will naturally shy away from adding more processing to material that is already processed.

Point 5, Undo-able

If a process is simple enough, its control can easily be described with a minimal number of bits of ongoing metadata, allowing an opposite process to occur in those receivers where it might be required or desired. But keep in mind that our ideal Loudness Controller would not sound 'processed' in the first place, so this requirement may be moot for many types of programming and the majority of audience members.

MAKING AN AGC'S ACTION INAUDIBLE

About a decade ago, I set out to design an AGC with minimally audible control artifacts. At that point in time, the product orientation of broadcast processor manufacturers seemed focused upon maximum peak level control and ultimate high levels of loudness (most particularly in radio). Almost without exception, the idea that an AGC could or should be unobtrusive seemed of little importance.

My thought was that by making the AGC nearly invisible, the following peak control could be used more consistently and judiciously. By carefully noting and correcting for the design flaws that produced various well-known processing artifacts, we eventually came upon a series of control algorithms that when used together met our 'no sound' criteria. The fact that this also created an environment where we met most of the five 'ideal loudness controller' points previously outlined turns out to be a pleasant bonus when it comes to the subject at hand.

The most egregious AGC artifacts are:

- Inappropriate signal detection and timing
- Spectral modulation ('pumping')
- Crossover anomalies
- An insistent, continuously varying control signal ('breathing')

We were able to find an elegant combination of methods to process level audio by exploiting a few features of human hearing.

RMS detection

The human ear/brain combination appears to integrate audio levels over time, with an exponential relationship over the first few hundred milliseconds. In other words, within a certain time range of less than about one-third of a second, the longer the duration of the audio at a given level, the louder it sounds (Plomp and Bouman, 1959; Zwislocki, 1969). Any detection scheme that fails to take this into account will have incongruities with the human hearing model, and traditional peak detection is particularly noted for such errors. By using RMS detection with appropriate time constants, we can easily approximate the required integration, such that our detection mechanism closely follows the ear's perception.

Additionally, by measuring not only the absolute level but also the *variation* over time of incoming RMS energy, it is possible to derive an additional signal that indicates the instantaneous dynamic range (IDR). This signal comes closer to zero the more consistent the audio energy level is over time, to correspond to the human ear's similar sensitivity. We will use this signal to prevent unnecessary control, a function which we will cover shortly.

Multiple bands

We discovered early on that one of the most audible level control artifacts is *spectral modulation*. You hear this effect, commonly referred to as 'pumping,' when one portion of the spectrum's audio dominates and (often rhythmically) controls the overall level when used in a single band ('wideband') AGC.

By using frequency crossover filters to create multiple frequency bands, each with their own control system, it is possible to nearly eliminate spectral intermodulation, but multi-band operation creates problems of its own when it re-shapes the audio in an unnatural spectral balance.

First-order Crossovers

In today's world of sophisticated digital audio processing, crossover designers lean very heavily upon sharp FIR filters. They are inherently phase linear, and thus sound very good. The temptation is to use as radical a rolloff as possible, with the thought that a clean, signal-free stop band is vital. In reality, as far as ear-friendly level correction is concerned, there is no such requirement. On the contrary, by using gentle, first-order IIR filters to separate the control bands, not only does the sonic spectrum more smoothly adhere to that of the original when program audio is spread over multiple bands of differing gains, when an instrument plays a scale that moves from one band into another the transition from band to band is inaudible. A first-order crossover system can easily be designed to have an overall phase deviation of less than 30 degrees under any operating condition, which means that phase anomalies are not a problem.

Spectral skewing

There's still, however, a serious problem to deal with: the unnatural, artificial-sounding spectral re-balancing called 'spectral skewing.'

By definition, good-sounding audio does not require any form of equalization. Historically, using multi-band systems for level control would mean that certain portions of the spectrum would be abnormally boosted or cut. Over a narrow dynamic control range of a few decibels, this may not be noticed, But in an environment where levels may need to be altered as much as 25 dB or more, spectral skewing becomes aggravating.

Equal energy distribution

Fortunately, since we are working with RMS energies in our processors, we can take advantage of this close relationship to loudness perception to use the spectral energy balance of what we consider to be well-balanced programming as a reference.

Using this reference spectral balance program material, we set the crossover frequencies to yield near identical energy levels in each band. By doing so, the typical 'good-sounding' programming maintains its spectral balance, as all the bands' control systems will naturally track their gains together over extremely wide dynamic variations. Programming that does not have the same spectral balance as our reference will be subjected to a slight amount of dynamic EQ. However, since the crossovers are first-order, even with several decibels' difference in gain in adjacent bands the transition between bands is gentle and any EQ effects are subtle and typically even useful for finely correcting material with a poor spectral balance.

The problem of spectral skewing is greatly diminished to the point where generally only those who can A/B monitor the input versus output will notice a difference.

'Window' Hysteresis with Release Gating

'Breathing,' is what we call the audible effect when the attack and release of automatic level control become noticeable, usually more obvious with gross level changes. Multiple bands won't reduce or eliminate breathing, but multiple time constants or a *Window Hysteresis* characteristic release algorithm will. Using intelligence to decide when and if to make a change is the key to eliminating breathing. This concept was first introduced with CBS Laboratories Audimax automatic level controller.

Here is where we use that second RMS signal we mentioned before, the IDR, which represents the variation of energy over time. The lower this number, the less processing is required to keep the level consistent. Our version of window hysteresis uses that signal to determine when to shut off ('freeze' or 'gate') the release system. We also shut off the release when the incoming level goes so low that bringing it up would cause an increase of noise or other annoyance.

In dynamic control, it's the transitions that create the audible artifacts. After all, if there is no change, there is no effect. So, by opportunistically stopping the processor from making changes, we reduce the amount of time that the processor is making level transitions.

It should be noted that when programming is itself changing level, one can adjust it almost with impunity. As a level goes down, a corresponding counter upward correction is not noticeable, so long as the compression ratio is not higher than about four to one. Similarly, as a level goes up, a downward level correction is also not readily audible. About the only time you can't change levels inaudibly is when they are extremely steady. A Hysteresis release system exploits this fact, because *it stops the processing exactly when it would be most audible!*

By looking for opportunities to stop or freeze the dynamic correction signal, we have been able to create a system that is exceedingly transparent to human hearing, all the while being able to control nearly any level it might be expected to receive.

This type of processing not only allows what I feel is the most faithful reproduction of the short-term dynamics in the program, the algorithm coincidentally allows previously processed or over-processed material to avoid further processing, and it has yet another bonus of also neatly avoiding the usual audible reminder that processing is going on, by opportunistically time-separating attack and release with periods of no gain change.

APPLYING THESE ALGORITHMS TO LOUDNESS CONTROL

A simple AGC *can* be used as a loudness control. But loudness control that is truly 'ear friendly' must have some kind of compensation for the peculiarities of human hearing. In particular it must take into account the spectral loudness response of the human ear, most often measured by what are called 'equal loudness curves,' the best known example being the Fletcher-Munson curves from 1933.

Those curves, and the numerous similar curves that have come since illustrate the reasons we can not expect a traditional AGC to do more than a passable job of loudness control.

Importantly, the ear is most sensitive in the midrange, with less sensitivity progressively towards the extremes of the spectrum. Also, this measured difference in frequency sensitivity changes with the absolute loudness, becoming more pronounced at lower levels.

But loudness is even more than that. It is a moving target, depending upon, in addition to intensity and spectral balance:

- Spectral density
- Duration of the sound
- The individual, and the individual's
    - Mood
    - Hearing ability
    - Subjective preferences
- Environment

Loudness is clearly a subjective experience. The sensations 'too loud' or 'too soft' can be experienced differently by two different people or even the same person on a different day in a different environment or even just in a different mood.

We can't hope to create a 'perfect' loudness controller, and a definitive study of all the possible sensations of loudness is clearly beyond the scope of a paper such as this. But fortunately, incredible amounts of research are readily available for us to refer to.

Which Loudness standard to use?

The familiar saying "the great thing about standards is that there are so many of them to choose from" has no greater validity than when referring to the study and measurement of loudness.

If we start with the well-known Fletcher-Munson report published in 1933, over the years there have been dozens if not hundreds of studies on the subject of human loudness sensitivity. There have been numerous methods developed attempting more or less successfully, to describe loudness objectively, certainly a daunting task. Some research results are more applicable to real-world broadcast than others.

The research I have paid most attention to is that which brought forth the original CBS loudness controllers and companion products, the Loudness Monitor and Loudness Indicator.

The CBS Approach

To the best of my knowledge, the aforementioned CBS studies included a number of 'firsts.' Most importantly, they were the first broadcast group to dedicate a serious budget to research in the effort to discover the mechanisms of loudness perception in broadcast programming

Equal loudness curves had been based upon single tones, in laboratory anechoic listening conditions. It would appear the CBS Laboratories research team (including Benjamin Bauer, Emil Torick and R. G. Allen) were the first to use pink noise filtered to 1/3 octave centered at 1 kHz as a reference (rather than a tone). They also were early in their extensive use of noise bursts for testing, along with employing a living room environment rather than anechoic laboratory for listening tests, and the use of actual programming for comparative samples. They were the first company to offer a useful loudness meter ("Loudness Level Monitor," 1967) and the first to offer a loudness controller.

CBS Labs went on to produce a second more sophisticated version of their CBS Loudness Controller and over the next few years (re-formed as the CBS Technology Center) continued to explore the realm of loudness control, introducing a revised loudness monitor, the "Loudness Indicator" in 1981.

Bauer, Torick et al, defined their loudness meter as having eight bands, each roughly corresponding to three of the 24 so-called 'critical bands.' In the analog 1960s and 70s, the design of narrow band filters was expensive and tedious, so they had elected to use fewer bands. It can be debated as to whether that shortcut made their instrument noticeably less accurate than it may have been, but given the variation in humans it was not really such a bad compromise, and was a clear improvement over existing art.

In today's world of DSP, however, not only is direct replication of the 24 critical bands possible, it is actually not difficult, once enough DSP firepower is given to it. A weighted sum of the RMS values of these critical band-matched filters would probably be a good starting place for a loudness reference, and we are pursuing such a project.

Now, by taking the CBS approach a step further, and building in the loudness contour derived from our 24 critical bands of RMS detection, we apply that loudness reference to the threshold of a smaller number of AGC bands whose crossover frequencies are chosen to result in an equal distribution of RMS energy per band with normal programming. Further, by implementing the aforementioned algorithm of hysteresis release with release gating which preserves entirely the momentary variations in the dynamic range, then a powerful loudness control might be effected in a way that is pleasant and nearly inaudible in action.

But that is just one possibility. Once you have derived the loudness signature, based upon whatever curve you believe is best and whatever summation of band energy best meets your specification, it can be applied similarly to the AGC control system. One might even include multiple detection schemes built into a single unit, where the user can best decide which standard they prefer.

SUMMATION

Preliminary listening tests are showing that we can consistently control loudness with minimal audible artifacts using adaptations of AGC algorithms that have been proven and built upon over the course of several decades and many products. We've described an ideal Loudness Controller for broadcast, and the methods that might be used to implement it. This system would incorporate a chosen loudness level detection reference signal, controlling an algorithm that is a combination of RMS detection with release gating and dynamic window-release hysteresis with an appropriate compression ratio and carefully selected timing, using feed forward control in a multiple band configuration featuring gentle crossovers and an equal-energy per band scheme.

Such a system promises to be powerful, transparent, undo-able and prior processing aware. In the ongoing quest for consistent or improved audience levels, a quality loudness controller is another important ingredient in the alchemy of audience maintenance.

"The Science of Sound, Second Edition" Thomas D. Rossing, Addison-Wesley Publishing Company, 1990 ISBN 0-201-15727-6

U. S. Patent #3187268; *"Gain Control Apparatus providing Constant Gain Interval"* Inventors: Benjamin B. Bauer; Arthur Kaiser; Assignee: Columbia Broadcasting Systems, Inc.; Publication Date: June 1, 1965

U. S. Patent #3260957; *"Compensated Platform Gain Control Apparatus"* Inventors: Arthur Kaiser; Emil Torick; Assignee: Columbia Broadcasting Systems, Inc.; Publication Date: July 12, 1966

U. S. Patent #3594506; *"Loudness Level Indicator"* Inventors: Bauer, Benjamin B.: Torick, Emil L.; Allen, Richard G.; Rosenheck, Allen J.; Assignee: Columbia Broadcasting Systems, Inc.; Publication Date: July 20, 1971

U. S. Patent #3582964; *"Automatic Loudness Controller"* Inventors: Torick, Emil L.; Allen; Richard G.; Bauer, Benjamin B.; Rosenheck, Allan J.; Owner/Assignee: Columbia Broadcasting System, Inc.; Publication Date: June 1, 1971

U. S. Patent #24406923; *"Automatic Loudness Controller"* Inventor: Burne, John C., III; Torick. Emil L.; Owner/Assignee: CBS Inc.; Publication Date: September 27, 1983

"Loudness, Its Definition, Measurement and Calculation," Harvey Fletcher and W. A. Munson, The Journal of Acoustical Society of America. vol. 5, pp. 81-108 (1933)

"A Re-determination of the Equal-Loudness Relations for Pure Tones." D. W. Robinson and R. S. Dadson, British Journal of Applied Physics, vol. 7, pp. 166-181, 1956

"Calculation of the Loudness of a Complex Noise," S. S. Stevens, The Journal of Acoustical Society of America vol. 28, pp. 807-832, 1956

"Relation Between Hearing Threshold and Duration for Tone Pulses," Plomp, R., Bouman, M. A., The Journal of Acoustical Society of America vol. 31, pp 749-758, 1959

"Procedure for Calculating Loudness," S. S. Stevens, The Journal of Acoustical Society of America, vol. 33, pp. 1577-1585, 1961

"Researches in Loudness Measurement," Bauer and Torick, IEEE Transactions on Audio and Electro-Acoustics, vol. Au-14, No. 3, pp. 141-151, 1966

"A Loudness-Level Monitor for Broadcasting," Benjamin B. Bauer, Emil L. Torick, Allan J. Rosenheck and Richard G. Allen, IEEE Transactions on Audio and Electroacoustics, Vol AU-15, No. 4, pp 177-182, December 1967

"The Measurement of Loudness Level," Benjamin B. Bauer, Emil L. Torick and Richard G. Allen, The Journal of Acoustical Society of America vol. 50, pp. 405-414, 1971

"Temporal Summation of Loudness: an Analysis," Zwislocki, J.J., The Journal of Acoustical Society of America vol. 46, pp 431-441, 1969

"A New Loudness Indicator For Use In Broadcasting," Bronwyn L. Jones, Emil L. Torick, CBS Technology Center; Audio Engineering Society, Presented at the 71stConvention 1982 March 2-5 Montreux; originally published in JSMPTE, Sept., 1981

"On levelling and loudness problems at television and radio broadcast studios," Siegfried Klar, Gerhard Spikofski; Institut für Rundfunktechnik GmbH; Audio Engineering Society Convention Paper 5538; Presented at the 112th Convention 2002 May 10–13 Munich, Germany

"Comparison of Objective Measures of Loudness Using Audio Program Material," Eric Benjamin; Audio Engineering Society Convention Paper 5703; Presented at the 113th Convention 2002 October 5–8 Los Angeles, California, USA

"Precise and Full-range Determination of Two-dimensional Equal Loudness Contours," Y. Suziki, et al, Research Institute of Electrical Communications, Tohoku Unversity: Japan, (2000-2003)

"Normal Equal-Loudness-Level Contours ISO 226:2003 Acoustics," Suzuki, et al: the International Organization for Standardization (ISO); revision of standard curves as defined in ISO 226.

"Recommendation ITU-R BS.1770," Algorithms to measure audio programme loudness and true-peak audio level (2006)

Audio-Video Timing

Paul Briscoe
Manager, Strategic Engineering
Harris Broadcast Communications
Toronto, Canada

Prepared for NAB2007 Conference – "Audio Solutions for Television"

Overview

Maintaining synchronization between audio and video is one of the oldest technical challenges in media. Because they are usually handled by different means throughout the production and presentation processes, many opportunities exist for audio and video to get out of sync. This paper discusses the history, context and nature of the problem and explores possible solutions.

A Bit Of History

The challenge of audio to video synchronization in media is not a new one, indeed when the first silent movies were projected to awestruck audiences, the piano player was challenged with matching his timing and changes to the content on-screen, and of course this became a very advanced art with the music even more dramatic than the dialogue it replaced. The real challenges began with the introduction of 'talkies'. Here we saw the joint usage of two different technologies and methods to record pictures and sound, requiring a means of establishing synchronization between them. Though many methods have been noted, including the clapping of hands in front of the camera, the one that stuck was invented in the 1920's, purportedly in Australia, which was called the clapperboard or slate. The latter name comes from the hand-held blackboard that is shot at the beginning of each take, containing the film, scene, take and other information (that we'd call metadata today), to this was added a 'clapper, consisting of a hinged wooden arm which could be raised and clapped back down onto the slate, creating simultaneous aural and visual events. Once the sound recorder and camera were running and up to speed (this is the '20's remember), the clapper is framed and the clapstick operated, from there the scene is shot. Any time the sound or picture recording are disrupted in any way, the clapper is again used to restart shooting.

Bringing the sound and pictures into the editing room, it was now possible for the editor to align them by finding the audible clap on the sound track and the visual cue of the impact of the clap stick, to the nearest frame. At this point, the sound and pictures could be rolled simultaneously, and plus or minus startup time differences and running speeds, alignment would be maintained, more or less. As the motion picture art evolved, sound was handled in many different ways in the editing room, but the clapper remained as the primary means of establishing and verifying alignment. In the earliest days of movie distribution, the sound was distributed on discs which were aligned with a short beep to the film countdown which the projectionist would do prior to pressing 'go'. If he did a good job, and his equipment was in top form, the movie would start and stay in sync, otherwise he'd have to modulate the speed of the turntable on an ongoing basis to keep things in alignment. As technology evolved, sound-on-film techniques made this alignment unnecessary, and of course further evolution brought discrete mag sound which again had to be aligned for projection, and so on, but at least the problem was constrained to one device or system and could be fixed readily.

With the advent of television, the problem of AV asynchrony was gone – TV was live, and the camera picked up and sent the picture in almost real time, and the sound was likewise picked up, amplified and delivered along with the picture to the transmitter. The signal from the transmitter went directly to the viewer's set, where it was demodulated and immediately displayed. Delay times down either the audio or video path were measured in microseconds, and there were no elements in the system that could introduce any significant delay anyway. The biggest challenge was a legacy one – when playing movies on television the same old requirements for maintaining AV sync came to bear, but with sound-on-film by now well established, this was a small problem. If the Telecine operators were on the ball, the broadcaster knew that their sound and pictures were in sync, and would stay that way simply because there was no way to make them otherwise. The biggest problems in this realm came in the sound effects room, where synchronizing the effects to the visuals are critical, and were done live.

Even with the introduction of videotape in the 1950's, the sound and picture were recorded simultaneously on tape, and here again the delays were inconsequential in the AV sync context, in the record path there were virtually none, and on the play side where video timebase correction techniques evolved which delayed the output video somewhat, the magnitude of these delays were small, primarily to accommodate the mechanically-induced timing variances in the off-tape video. In the 70's, however, when timebase correction further evolved to using frame buffers to accommodate unlocked VTRs and provide a genlocked output, we first begin to see the unraveling of our synchronized universe.

These devices delayed the video by up to a frame with respect to when it came off tape, but audio was not processed, simply taken from the VTR directly. We now have a situation where, under certain circumstances, the picture could arrive later than the sound by up to a frame, even if the content recorded on tape was perfectly aligned. If this were a dub, the recorded result would contain this error; and when played back would have the same effect added by that playback operation. The tip of the iceberg had been revealed.

The origin of the problem is subtle, as realtime baseband video is inherently time-critical on a scale of nanoseconds. One of the traditional headaches of system design has been timing the video to ensure that all pictures arrived in time at points of composition or switching to allow seamless operation. This was accomplished through many schemes before digital processing techniques were available, one of which was identifying the latest signal in a group that needed to be timed, and delaying the others using reels of cable. This technique, however, only yielded delays in the microsecond range, and were thus useful only for fixing H-timing-sized issues within a facility. Within this environment, it was easy to maintain AV synch, there was no opportunity for it to be otherwise.

VTRs were only the first of many AV sync offenders. As the '80's progressed, more and more advanced digital techniques were brought to bear in the industry, most of which were on the video side. Frame synchronization became the defacto way to handle video timing as it could deal equally with freerunning or locked sources, and didn't care what the source timing was, it'd simply add or remove delay including the dropping and repeating of full frames on an ongoing basis to keep the output video locked to the genlock reference. Of course digital techniques were expensive, so putting a parallel delay system into the associated audio path wasn't done, after all the delay was at most one frame, hardly noticeable, right? Frame buffers prevail in much of todays processing and production equipment, integration with computer technologies require having the entire frame in hand, thus such live equipment has inherent delay.

Nature of the problem

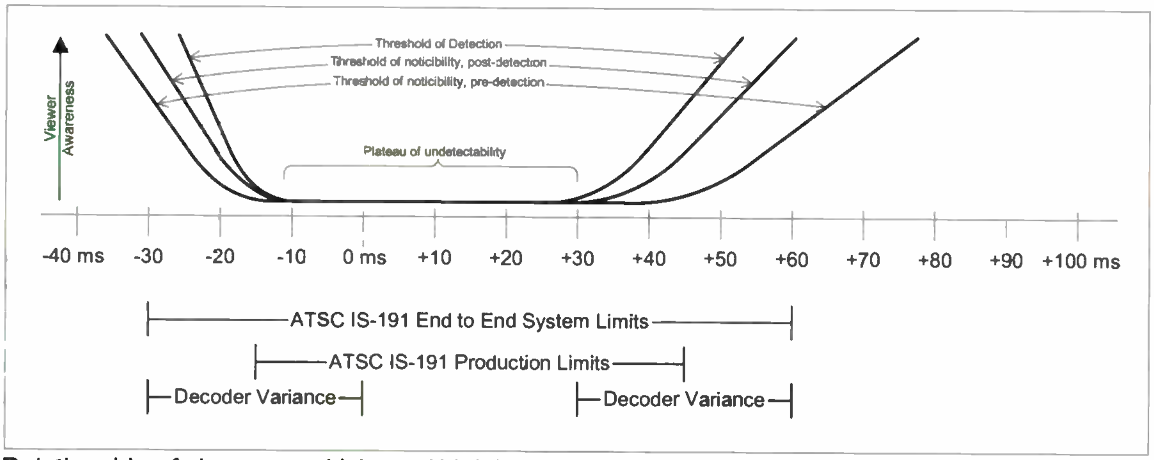
Our human perceptual systems are highly evolved tools which allow us to directly extract as well as infer vast amounts of information from our surroundings. These systems collaborate to give us a richer experience than the raw data provides – the screech of tires makes people brace automatically for the crash as the adrenaline flows, seeing lightning in the distance creates a natural expectation of thunder to follow. These systems evolved through our human experience interacting with the physical world, and are tuned to enhancing experience through learned effects. One of these learned effects is the natural lateness of sound compared to vision, such that our brains are quite comfortable with audio arriving after the image. By the same token, when this natural situation is reversed, putting the sound before the image, the result is quite disturbing and uncomfortable.

If one is watching the news and can read lips with the volume turned off, you have no issues, likewise if you don't watch, but only listen (also called radio). You get the full information without any distraction. This story changes however when you look and listen to material where the audio and video is out of sync. Your brain has reasonable expectations, two of which are that sound often comes after the picture due to the differences in physical propagation of sound and light, and that this latency is of a magnitude proportional to distance. When these are within reason, meaning the sound closely follows the visually–associated cue, nothing looks amiss, but if the sound lags the picture by an amount which is not obviously natural (such as a close-up shot), or simply lags by 'too much', we can quickly notice and become distracted by this. How much is too much? Studies from the early '80's yielded ITU BT1359-1, where it was recommended that audio should not lead video by more than 90 ms, and should not lag by more than 185 ms. This was based on reliable detection of the occurrence, and by today's standards is far too loose. Subsequently the ATSC developed IS-191 which tightened the times considerably, down to 15 ms maximum lead and 45 ms maximum lag; combined with a decoder tolerance of +/- 15 ms, end to end performance limits are 30 ms lead and 60 ms lag.

There are actually two thresholds to be considered, detectability and noticeability. Detectability is the point at which a viewer can observe AV asynchrony when they're looking for it; noticeability is the point at which the viewer notices it without trying. It's been noted that the noticeability threshold is reduced somewhat after the initial notice, such that it takes a significant reduction in error to become again unnoticed. This hysterisis means that should a brief excursion in AV sync occur that is spotted, previously unnoticed smaller errors will now be more readily spotted and it will take some period of compliance for the viewer to forget the trigger. This means that brief errors of large magnitude are far more dangerous than ongoing errors of lower magnitude, and their occurrence can predispose the viewer to spot the smaller ones.

The effect of the sound being later than the picture is natural and viewers are somewhat more tolerant of it, however the problem worsens when the opposite happens. In our natural human experience, we don't have a benchmark for the sound preceding the picture, so our brains find this experience particularly rattling. Not only is the threshold of detectability significantly lower, so is the threshold of noticeability. In addition, the hysterisis is greater than the audio-later case, so once noticed, it will remain so with tenacity. Finally, studies show that prior to being triggered, or even being able to detect it, viewers find that early audio nevertheless results in an uncomfortable viewing experience. Of course, once it's been detected and remains noticed, the result is a distracting, difficult to watch and unpleasant experience overall.

It's also been identified that audio-early content is received with much poorer intelligibility by the viewer, and it's message may be lost or it's legitimacy questioned as a result. Bad enough for programming, this is catastrophic for advertisers, where retention of the message is key. Even worse, consider that a critical political message is being delivered to a population by it's leader, using a studio where the audio precedes the video by enough to cause viewer discomfort, not noticible, but enough to trigger negative effects. Is their message less understandable or believable due to this? Is the trust of their public in any way influenced by it? What should be a minor annoyance could have implications beyond soap and soap operas.



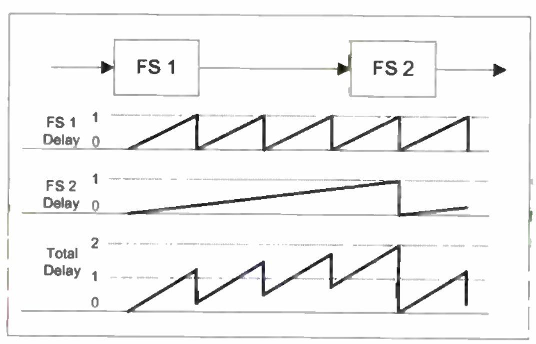
Relationship of viewer sensitivity to AV delay

Content plays a significant role in positioning these thresholds. Where there are few or no mutual events, there is no problem, and if the associations are soft or ambiguous (sports, where the distance to the event makes the AV sync error imperceptible, or shots where mutual events are few), the problem is of no consequence to a point, obviously a threshold will be reached where even the softest associations are obviously wrong. The problem becomes more marked with people talking, as there is a natural recognition and brain expectation of correlation between lip action and sound, although this is still a somewhat soft perception domain. The worst problem comes with sound effects, be they natural or Foley. Like the clapper board, percussive events with an equally striking visual provide a very crisp sampling of the AV alignment, and are often the trigger that causes ongoing detection in content that otherwise may not have triggered the viewer.

So we have the situation where sensitivity to AV asynchrony is asymmetrical with audio early being critical, and we use technologies which have traditionally (and continue) to delay the video. This means that the problem is lurking around many corners, ready to bite.

Sources of Asynchrony
If we had the benefit of inventing a whole new end-to-end broadcasting technology which used a uniform and robust scheme for content time stamping and alignment, we could get a handle on the problem. We have this in many places in today's systems, and where it works, it works well, but limited implementations, particularly in the consumer space, can cause it to fail. Throughout the program path from creation to viewer, we build systems of increasing capability and complexity, sometimes without full attention to the management of audio alignment.

Use of framesynchronizers to 'fix' timing has become commonplace, but if audio is not considered, asynchrony triggers can occur. A single framesync can generate only a maximum of a frame of error, but this already starts to encroach on the noticeability threshold. If this is the only delay in the video path, things may be OK, but consider having two of them cascaded in a path where audio is not handled – much of the time their net contribution is not worrisome, but there are occasions where the delays add to push the delay well over the threshold. Without handling audio to compensate for these delays, we have a system which can intermittently break otherwise good content.



If single stages of frame synchronization were the only source of the problem, it might not be as insidious, but the advancement of production technologies has continued to exacerbate the issue. Since the beginning of the analog / digital hybrid era, production switchers with DVEs could add and remove frames of delay on the fly, so in doing a live show, one could have perfect audio alignment with a straight bus mix, but cut to a DVE for an over the shoulder, for example, and suddenly the audio is a frame early. This can be limiting in terms of the visual composition of the program, as shots and mixes which suffer this fate must be avoided. The risk in this is that at one time or another, simply due to the visual composition, the content may trigger the user threshold of noticeability, leaving them predispositioned to detecting an ongoing (and intermittent) problem they wouldn't have otherwise detected.

Routing Switchers – how can this be? More and more manufacturers are providing format conversion natively on router inputs and outputs, as well as synchronization. The matrix delay is irrelevant, but framesyncing or otherwise processing the video first can be an unexpected source of delay.
Because they're internal to the device, video servers typically have robust AV synchronization mechanisms that guarantee alignment. While they may guarantee that the server is not contributing to the problem, it should be remembered that the server will happily capture and protect any incoming AV alignment, good or bad, and add to that whatever its internal characteristic might be. Something that looked OK during the ingest because it was inside the threshold may suddenly cross the line at playout due to the server's contribution.

CCD Cameras can introduce their own delays - depending upon the CCD block and downstream processing, there can be a field or more of delay between the optical image and the resultant output video. Many camcorders consider this when recording audio and video and compensate accordingly, but a standalone camera knows nothing of audio. When used for live coverage, the linked-back video us usually framesynced to time it into the facility, adding another variable delay factor. What looked good in the field has now broken badly.

Compression technologies provide mechanisms to manage and maintain AV sync, but many are not infallible. Depending upon the technology used and the specific implementation within that compression toolset, the robustness of the AV synch mechanisms can be quite variable. Some schemes are able to tolerate wandering of streams better than others, and how they recover can differ somewhat. It's desirable for the content to never be disrupted, and some systems provide compress / expand solutions for this, others simply burp the audio to fix the timing. If the compressed audio stream comes from an unlocked clock (or there are other system – level violations), the receiver may not cope well, achieving sync when the program is acquired and not doing anything more about it until it's so bad it requires a relock. This is doubly bad, as the user may have long begun noticing the error and are treated to a burp which fixes it, maybe only for a while. This sort of problem is most prevalent in the distribution chain to the end viewer, where set-top boxes, designed to be extremely cost-effective, have only the most necessary and often rudimentary alignment and recovery mechanisms. Also in the compressed domain we have processing like transrating, switching, Transport Stream grooming and local content insertion, all of which must obey the rules of synchronization and handle the audio correctly.

What about the end of the road? The viewing environment can equally be guilty of contributing to the problem and it's beyond our direct grasp as broadcasters. Consider the case where a viewing environment is set up utilizing 5.1 sound and HD video, say 720p. Coming out of the STB, the audio and video are in time, and the audio goes directly to amplifier channels and speakers, the video to a native resolution (1280x720) LCD panel. The video is likely delivered promptly to the screen pixels, requiring no temporal processing in the display and the audio and video will be fairly tightly timed, within the capabilities of the STB. Now we switch over to a 1080i program. The display now has to perform deinterlacing and scaling, requiring some amount of delay which is dependent upon the technology and implementation used. Now we have pixels arriving at the display somewhat later than they did before, without the upstream system being aware that audio compensation is necessary. Even worse, the STB might be set to an output format which requires it to pre-scale the content which the panel scales again, incurring even more delay. In the era of TV 'sets' with everything self-contained and built by one manufacturer (with one video format to deal with), this could be possibly managed, but with today's multiple formats and user component choice, even a correctly aligned program is not safe when it gets to the viewer. One almost longs for the days of the TV "set" where everything could be managed in one place.

Because the end to end system involves many stages, the accumulation of errors along the way can result in accumulation of the problem. A device which introduces a small amount of video delay but doesn't do likewise to the audio is probably OK on it's own, but in a long content path, it doesn't take many of these to accumulate enough delay to detect or notice, and a marginal system with a framesync or two along the way which periodically jump the video into and out of the region of noticeability will be annoying to troubleshoot when reported. Small amounts of video processing delay in each device will accumulate, often with complex results, and different paths will have vastly different behaviours.

Solutions

There is no silver bullet for AV asynchrony. The best defense, it would seem, is a good offence – building systems with AV sync integrity at every stage, so that no element in the path contributes any biasing error and such that at the input and output of major (and minor) system blocks, the AV sync is maintained. The golden rule should be to always build subsystems with proper audio handling, so that when cascaded, a number of small errors don't accumulate to the point of noticeability. This won't, for example, stop an incoming satellite feed from having AV sync errors, but it will preserve them accurately so that they can be corrected at one place in the system. Equipment selection and careful system design can prevent delay accumulation scenarios, at least within a facility.

The common use of embedded audio has certainly gone a long way to eliminating many sources of the problem, as most processing devices will de-embed the audio and delay it appropriately for the video path, then re-embed, but this can lead to a false sense that everything's OK. Embedding can make audio easy to forget about.

When framesyncing video, it's necessary to process the audio as well. At a minimum, the audio should be delayed by a fixed amount so that the maximum excursions of the synchronizer don't push the audio early enough to be noticed. It's of course preferable to have the audio processed in concert with the video, using audio sample rate conversion to minimize or eliminate the instantaneous AV sync errors as frames are dropped or repeated. Where a synchronizer is being used for a fixed delay, the audio should be delayed by the same amount. Framesynchs used on the ingest side of a facility may also convert domain (A to D) and/or format. The additional delay from some of these processes must also be considered and compensated for in the audio path. The goal should be to deliver AV synchronous feeds to the facility, so should incoming errors exist, the ingest subsystem is the place to correct them, most ingest hardware provides the ability to adjust the audio delay for this purpose.

In studio systems, other than crafting all shots and mixes to avoid the problem, one solution here is system design where audio synchronizers are aware of ongoing video delay and use sample rate conversion to quickly re-align audio on the fly based on advice from the video switcher while maintaining pitch. While not the best solution as it can disrupt tempo in musical content, this has a comparatively mild viewer impact compared with the risk of noticing AV sync errors. Another common option is the compromise route, where a static audio delay is installed at the studio output to provide enough delay to guarantee that the early audio noticeability threshold isn't reached, hopefully without the simplest video path resulting in noticeably late audio. It could be that for a total DVE contribution of 2 frames, a one-frame audio delay would suffice provided the 2-DVE shots didn't carry clear mutual events. This is indeed a compromise solution, but can and does yield good results.

To ensure that the routing switcher is not a contributory source of the problem, AV alignment of any input-processed source should be verified on a router output bus, ensuring that the full internal processing and routing is in the path. Depending upon manufacturer, the audio may be handled properly, or may require adjustment (where possible) on the router front end or upstream.

CCD-block and processing delay errors can be handled in a couple of ways. With discrete picture and sound acquisition, timecode can be used to guarantee a known relationship, but may still not consider the camera delay. Shooting a slate is still the best bet for a bottom-line confirmation when recording, for live ENG, system design (both ends) needs to ensure that the AV is in time when it arrives into the production environment.

In compressed systems, one is at the mercy of the methods, which are pretty good, and the real-world implementations, some of which are maybe less so. A set-top box may have a solid AV timestamp alignment scheme, but perhaps has a flaky relock algorithm that results in a viewer seeing several seconds of AV asynchrony after a channel change before it is fixed. This is annoying, but not the end of the world, *unless* the content which on it's own was fine is now is noticed as having problems, and of course once seen, the sensitized viewer keeps seeing it. If this happens with every channel change, we'll have a viewer who is well aware of and likely not too happy with the problem. With all compressed systems, it's crucial that equipment be characterized and verified for AV robustness, not just in quiescent operation, but through LOS / AOS, channel change, etc..

Test and Measurement Tools

Over the years, there have been a number of schemes which enable a user to measure and establish AV synchronization. Of course the clapperboard / slate has evolved and is still used as insurance, although timecode has pretty much straightened things out in the acquisition end, although due to CCD delay, etc., it's entirely possible that the beautiful timecode-aligned content is not *really* in sync, so the clapstick still serves a purpose.

This is referred to as 'offline' testing as it's separate and distinct from the program content and not intended for viewer consumption. The electronic equivalent to this would be something like Harris' A2V audio sync test system where synthesized picture and sound with mutual audio and video events is generated by one piece of equipment and received by another which can provide the user with a display of relative AV timing. A signal of this type is used outside of content carriage for system setup, and is sufficiently simple in nature that it can even be evaluated with ears and eyes if necessary. Such a scheme also allows downstream equipment to be auto-adjusting. The primary advantage of this type of testing is that it can be applied at one end of any system and its successful propagation is guaranteed regardless of the conversion and processing along the way.

Online testing uses the live program content to supply the mutual events and uses downstream receiving devices to perform extraction and analysis. These events can range from insertion of specific triggers in the form of invisible content marking (such as watermarking, spectral signatures, etc.), to analysis of program content to find and track mutual events and use them for measurement / correction. These systems are inherently complex and require specialized receiving devices to do their job. Invisible marking of the content can be less than robust and may not propagate through all of the processing, particularly deinterlacing / scaling and compression / decompression of a system path. Content recognition systems count on intelligent analysis of visual features and look for mutual audio events, but their effectiveness and accuracy is a function of their degree of advancement and availability of appropriate content.

With the proliferation of technologies used in the broadcast path and the many ways today's methods can conspire to break AV sync, there is more industry interest than ever in solving the challenge. Various standards bodies have been and continue to work to better quantify the problem, and consider some new issues as well: the increased spatial resolution of HDTV may yield lower viewer thresholds by virtue of them being able to better see the mutual event. With progressive framerates of 50/59/60 Hz, the opportunity to be triggered is greater – an event that is over a field but under a frame in the interlaced format now has become visible. With a growing number of disparate production and distribution technologies, it's even more important than ever to find solutions which work from acquisition to viewing.

Industry Activities

In an industry where there are literally dozens of manufacturers' products within an overall program path, it's been challenging to find a holistic solution. Various manufacturers have put forward products to resolve the issue within the scope of devices or subsystems, but to date, no system-wide automated means of maintaining AV sync has gelled. It's a difficult challenge for any single manufacturer to solve, not just for competitive reasons, but simply due to the ever-increasing methods and technologies used. Perhaps the only opportunity for a fix comes from standards bodies, but to date this has been a tough start. Various organizations have struck committees to look at the situation, and to date only timing guidelines have come forward. Even quantifying the problem has even been tough, with the original ITU constraints being tightened further by ATSC, and now realizing that this too is insufficient, tighter numbers are being considered. What's needed however is a solution, not specification. At this writing the SMPTE's latest Ad-Hoc group on Lip Sync is building momentum which could result in adequate guidelines, Recommended Practices, and perhaps Standards which help constrain the problem. This could further evolve to development of a system-wide solution space in which all manufacturers can play and the problem can be addressed and resolved transparently and dynamically at the system level.

Conclusion

Ever since the invention of the clapperboard slate, AV synchronization has been an ongoing and evolving challenge for the media industries, and in particular television broadcasting. The problems are biased against the nature of the technology (video later) and accumulate in unexpected ways, and from many diverse sources. Video system re-timing within the path can make the problem come and go with seeming randomness, and viewer triggering can make what's mostly an infrequent problem one that's seen a lot more. HD with it's resolution and –p rates just makes matters worse. Awareness is the first key to solving the puzzle, as it will lead to equipment selection and system architectural and operational decisions which consider the issue and minimize or prevent it by design. Managing AV delay within subsystems of an overall production / distribution system will ensure that it's repaired wherever it gets broken, and the accumulation of any remainder is understood within the end-to-end system and kept within acceptability. Still, given the diversity of technologies and methods in use today, this is not going to go away any time soon. A problem as old as sound and pictures, AV asynchrony today is more prevalent and difficult to cure than ever before, but with awareness and design consideration, it can be managed, albeit marginally. The path from actor to viewer remains a long and wild one with many opportunities to mess up the AV sync. With today's technologies we can keep a lid on it, and with tomorrow's, we can possibly put it to rest.

Then again, they probably thought they had it nailed with the invention of the clapperboard; we need solutions equally simple and long-lived.

The author wishes to thank the following individuals for their contributions to and review of this paper.

Leigh Whitcomb, HARRIS Corporation
Randy Conrod, HARRIS Corporation

The only type of standard level instrument that does not display some sort of peak level is the *VU meter*, [13]. Though developed for another era, this kind of meter is arguably better at presenting an audio segment's *center of gravity*. However, a VU meter is not perceptually optimized, or ideal for looking at audio with markedly different dynamic range signatures.

Instead of a peak meter which isn't really peak (a sample meter), a quasi-peak meter which is neither peak nor average, and an average meter which isn't optimum average (VU), the ITU set out to define an updated standard describing the issues one by one [14] and introducing the perceptual measure, loudness.

Unlike electrical level, *loudness* is subjective, and listeners weigh the most important factors differently:

- Sound pressure level
- Frequency contents
- Duration

In search of an "objective" loudness measure, a certain Between Listener Variability (BLV) and Within Listener Variability (WLV) must be accepted, meaning that even loudness assessments by the same person are only consistent to some extent, and depends on the time of day, her mood, the degree of attention etc. BLV adds further to the blur, when sex, culture, age etc. are introduced as variables.

Because of the variations, a generic loudness measure is only meaningful when it is based on large subjective reference tests and solid statistics. Like other technical definitions, the results also have to be repeatable.

Together with McGill University in Montreal, TC Electronic has undertaken extensive loudness model investigations and evaluations [8, 9, 10]. The results denounce a couple of Leq measures, namely A and M weighted, as generic loudness measures.

In fact, a quasi-peak meter showed better judgement of loudness than Leq(A) or Leq(M). Even used just for speech, Leq(A) is a poor pick [7, 9], and it performs worse on music and effects.

An appropriate choice for a low complexity, generic measurement algorithm has been labeled Leq(RLB) [17]. Though it better describes loudness than a quasi-peak meter [9], its performance against normal VU or slow VU has not yet been systematically tested.

It should be noted that the idea of a perceptually based level calculation is not new. An aging, but respectable measure such as "CBS Loudness", is still being used with success for automated level control [6]. This model has served as a de facto reference for objective loudness measurement, in the broadcast community for decades.

ITU BS.1770

In 2006, ITU-R Working Party 6J drafted a new loudness and peak level measure, BS.1770 [15].

It has been debated if the loudness part of the standard is robust enough. As a global loudness reference, it will obviously be exploited where possible. However, with homogenous mono material, Leq(RLB) has been verified in independent studies to be a relatively accurate measure, taking its simplicity into account [9, 10, 17].

It seems justified to use Leq(RLB) as a *baseline* measure for long-term loudness, as long as room for improvement is built into the standard.

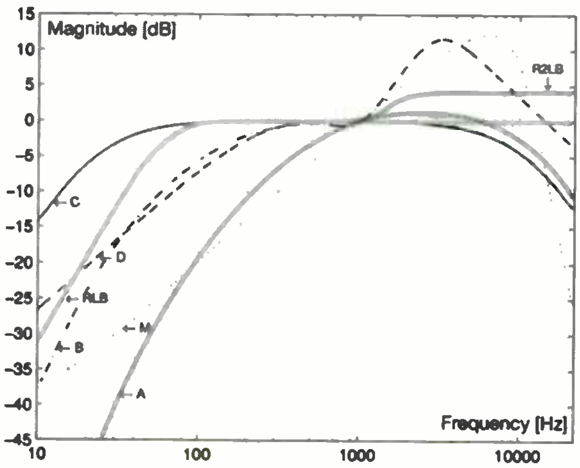


Fig 4. Weighting filters used in combination with Leq measures. A: Red, RLB: Green, R2LB: Blue (follows RLB below 500 Hz).

Note that a stereo and multichannel annex using a revised weighting filter, R2LB [18], is less verified. Especially the multichannel extension should be used with caution [10]. The different weighting curves are shown in *fig 4*.

The other aspect of BS.1770, the algorithm to measure *true-peak*, is built on more solid ground. Inconsistent peak meter readings, unexpected overloads, distortion in data reduced delivery and conversion etc. has been extensively described [1-5].

In liaison with AES SC-02-01, an over-sampled true-peak level measure has been specified. Depending on the measurement over-sample ratio, different under-read ratios can be expected. For instance, up to 0.7 dB at four times over-sampling, but better than the 3 dB uncertainty with a sample based measure. This improved peak level measure now being included in BS.1770 will hopefully make its way back to the music industry, and help put an end to the damaging of our musical heritage due to overly hot level.

In conclusion, BS.1770 is an honorable attempt at specifying loudness and peak level separately, instead of the simplistic (sample peak) and mixed up measures (quasi-peak) in use today.

Though combined peak and loudness measures already exist, the time is right for an international standardization. Analog level is not an issue anymore, so one of the historically difficult variables rooted in hardware has disappeared, and the dynamic range is wider. It is now more a question of how the average and maximum level is used across a variety of platforms.

BS.1770 COMPLIANT METERING

The BS.1770 measure may be presented to the user in a traditional, realtime way with certain rise and fall times to be specified, see *Fig 5*. The ITU-R BS.1771 draft is this type of meter [19].

Meter 1 uses negative LU numbers like a typical digital sample meter. NABA has suggested this method [20], at least until a Reference Loudness has been agreed upon. The other meters have a "0" Reference point like proposed by Australia [19].

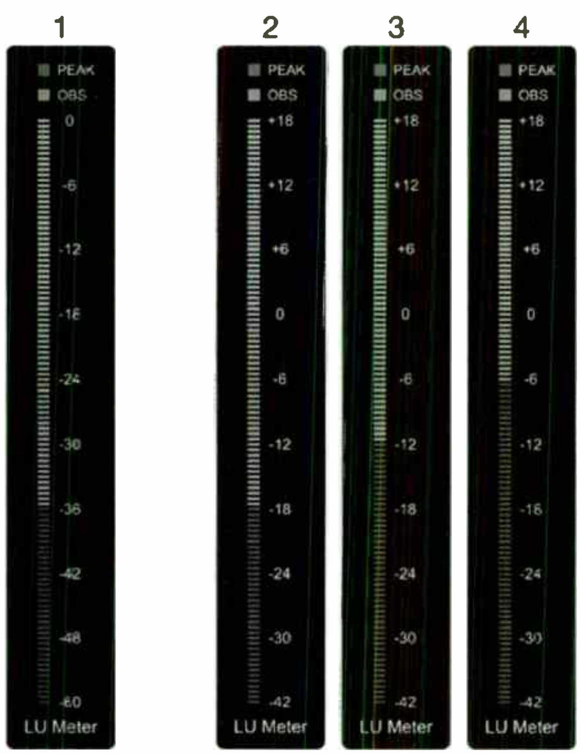


Fig 5. Loudness meters using a traditional, linear display, and color coded in accordance with Fig 2.
1: HDTV meter using negative only LU numbers.
2: HDTV meter with 0 LU at Reference Loudness.
3: SDTV meter with 0 LU at Reference Loudness.
4: iCast meter with 0 LU at Reference Loudness.

Note that the same LU meter display is expected to be valid across a number of audio formats. This raises some issues about mono, stereo and 5.1 summing, and reference tone displaying that were still open for discussion at the time of writing.

Informative combined loudness and peak level meters of course already exist, for instance the ones from Dorroughs. BS.1770 "just" offers a standardized way of measuring both numbers, and the means to tie short-term and long-term loudness consistently together.

BS.1770 AND LOUDNESS HISTORY

Loudness control is not just a matter of absolute limits. A change in loudness is by evolutionary default meant to grab our attention. Therefore, alerting auditory events don't have to last very long before we react.

TC and McGill University have conducted extensive listening tests to design a precise loudness model suited for both short-term and long-term measurements, speech, music and effects [8, 9].

We were concerned that describing the level variations of an entire program using just one number was an over-simplification, and would not provide enough information about its broadcast suitability, see *Fig 6*.

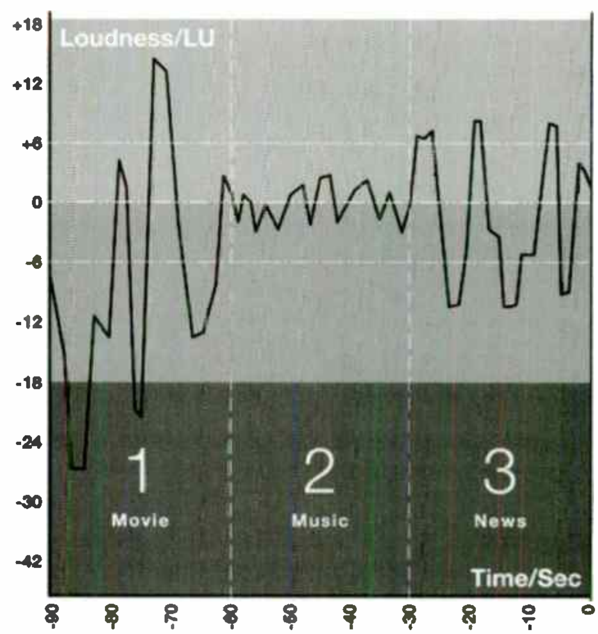


Fig 6. A one-number-per-program approach to Loudness is an over-simplification. Audio segment 1, 2 and 3 may be normalized and produce the same Standard Loudness number, but their profiles over time are clearly different.

To control loudness developments consistently over time, the most effective method is probably to have the loudness history visualized from production onwards. A mixing engineer or a journalist should be able to identify long-term as well as short-term loudness developments.

If only short-term loudness is displayed, a program's Standard Loudness Measure, its "center of gravity", is unknown, and therefore how well it fits with other programs, and across a variety of broadcast platforms.

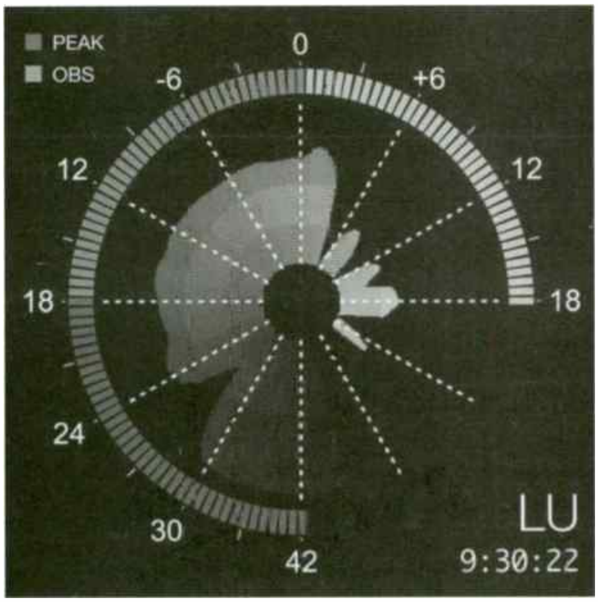


Fig 7. Loudness meter with loudness history in the "radar view" and color coded in accordance with the HDTV targets of Fig 2.

The loudness history aspect is not addressed in the BS.1771 draft meter, but an example with such virtues is shown in *Fig 7-8.* The loudness history can be set at, for instance, one revolution per minute.

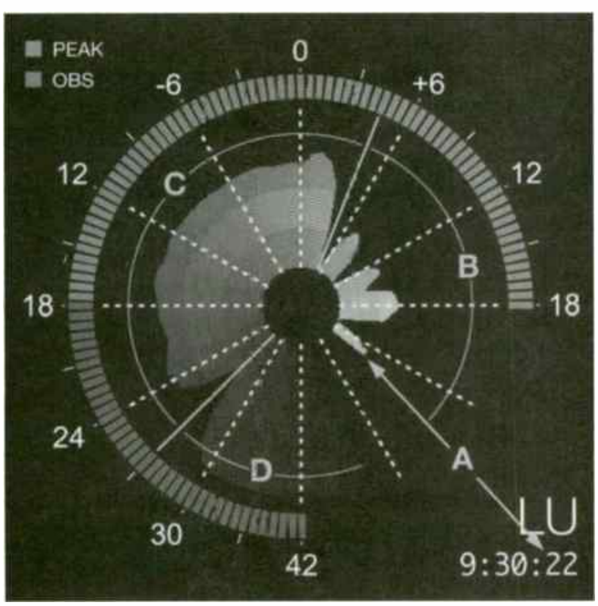


Fig 8. Loudness Meter explained. "Radar head", A, is moving clockwise, one revolution (e.g.) per minute. The outer ring shows current loudness at around –12 VU.
A: Current time, 22 seconds past the minute.
B: Current segment (Film): Low Consistency and soft.
C: Rock finished 17 sec ago. High Consistency and loud.
D: Pop finished 45 sec ago. High Consistency and very loud.

The round display distinguishes itself from a normal PPM or VU meter, making a point that the measure is also different. Its angular reading means that the numbers need not be visible, thereby enabling condensed views like the superimposed one shown in *Fig 9*. Reference loudness is at 12 o' clock, and can be identified from a small picture and at a distance.

SPEECH, MUSIC AND EFFECTS

For broadcast programming meant to be distributed over a number of platforms, it is fundamental to define its Standard Loudness, or "center of gravity". With this center point well defined, it is simple to transcode a given program to any platform.

It has been suggested to reference programming to the level of its dialog, which to some extent works for film. However, this has bad consequences in broadcast, where mixing esthetics between programs may vary significantly, where dialog not always take center stage, where any type of sound may be disturbing, and where the consumer Dynamic Range Tolerance is lower.

The sound of a phone ringing in a commercial, John Frusciante's guitar, or a fighting scene in Pirates of the Caribbean can all make some people grab the remote, and should naturally have an influence on the loudness of a program. Even if a station is news only, documentaries or drama, it will still have accompanying sounds that can be annoying.

Fig 9. Loudness meter superimposed onto a TV picture. The outer ring shows short-term loudness, the arrow shows long-term loudness. For both indicators,12 o' clock represents Reference loudness. The red center indicator would normally be off. It signifies notable channel differences, and electrical low or high level.

The most compelling reason to identify and treat speech differently from other sources would be to guarantee its intelligibility, not to anchor the entire program around it. Loudness is just one parameter when determining the intelligibility of dialog.

The right production techniques in combination with platform specific preconditioning, however, can ensure speech intelligibility, cross-platform fitness, and the elimination of unacceptable level fluctuations. All in accordance with BS.1770, and without additional time being spent at the station.

DELIVERY SPECS: STANDARD LOUDNESS

BS.1770 is an open standard for measuring peak level and loudness. It may be used to have level offsets (long-term loudness) correspond with realtime measuring and correction (short-term loudness) across program transitions, and across multiple broadcast platforms such as HD, SD, IP and iCast. It is necessary, though, to establish a consistency between off-line level offsets and on-line correction.

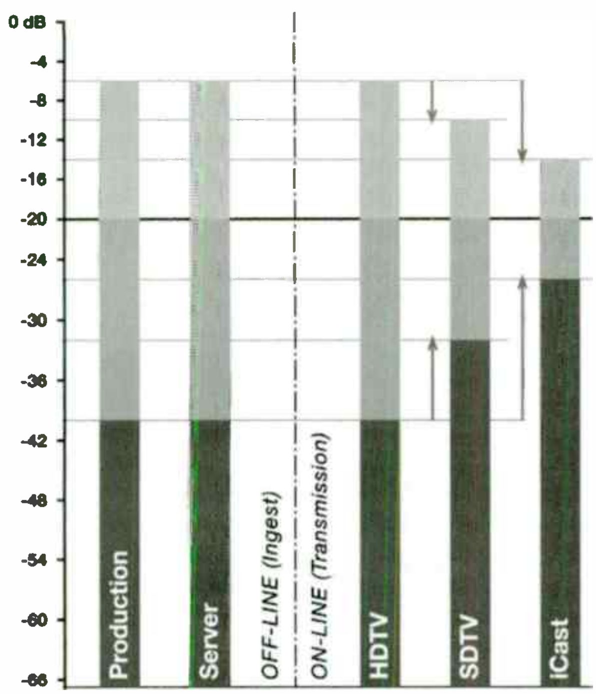


Fig 10. Level handling from Production to Broadcast.
All platforms use the same nominal SLM.
Red arrows: Dynamics processing.

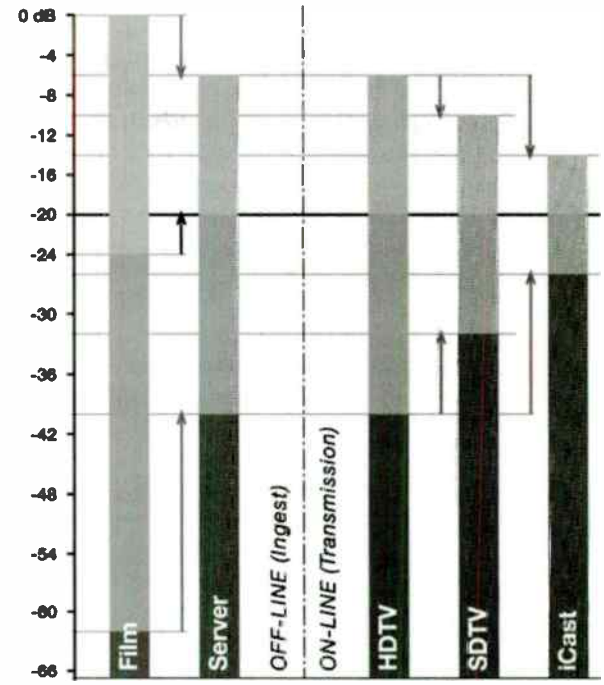


Fig 11. Level transcoding from Film to Broadcast.
First align SLM, then process as indicated.
Black arrows: Level offset.
Red arrows: Dynamics processing.

When all programming hits master control at the same Standard Loudness Measure (SLM), on-line correction for the various platforms can be centered around this value, and be as gentle, transparent and foreseeable as possible, see *Fig 10-13.* Content with different dynamic range signatures can be seamlessly mixed this way.

Production, live and external content should be aimed at the HDTV dynamic range signature, which is a little wider than what is used for today's analog TV delivery. The HDTV signature is automatically narrowed during transmission to fit other broadcast platforms.

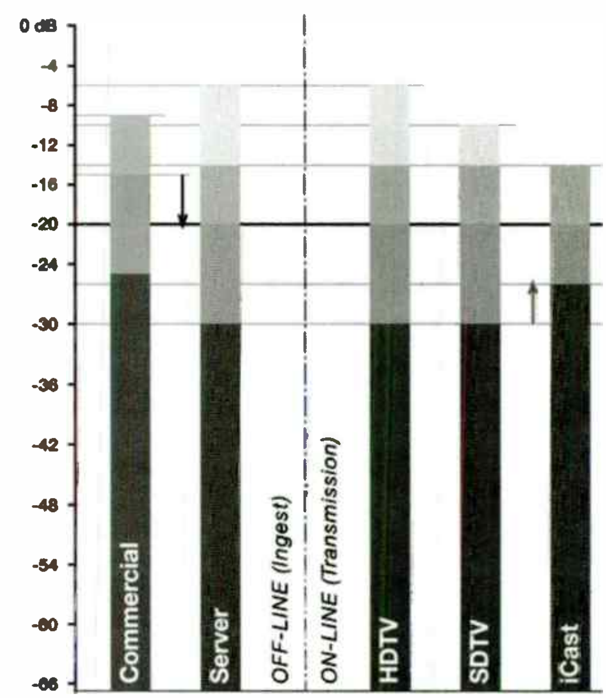


Fig 12. Level transcoding from Commercial to Broadcast.
First align SLM, then process as indicated.
Black arrows: Level offset.
Red arrows: Dynamics processing.
Gray marking: Unused headroom.

Despite the improved loudness consistency enabled through BS.1770 compliance, it is important not to aim at a wider dynamic range for HDTV than requested by most consumers, ref *Fig 1*. For some stations, it may even be advantageous to use the SDTV dynamic range signature for all HD delivery, with the possible exception of film.

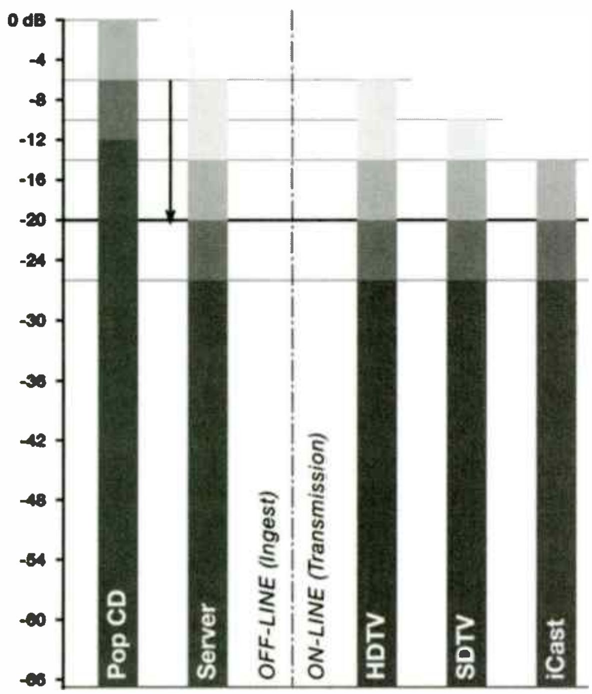


Fig 13. Level transcoding from CD Levels to Broadcast.
First align SLM. More processing is rarely required.
Black arrows: Level offset.
Gray marking: Unused headroom.

Long-term loudness normalization (level offsets) should therefore be taken care of during ingest or inside a file server. Under the same off-line conditions, relevant information besides from a program's SLM may also be derived.

DELIVERY SPECS: LOUDNESS CONSISTENCY

Studies of dialog from broadcast and film, music, commercials and effect sounds have led to the conclusion that at least one more telling parameter should be used for program delivery specification, namely the Consistency Loudness Measure (CLM).

CLM is a long-term statistical measure also rooted in BS.1770. It has been designed to indicate intrinsic loudness variations inside a program, and relate them to the HDTV range signature. A combination of the Standard Loudness Measure, SLM, and the Consistency Loudness Measure, CLM, is a superior broadcast suitability predictor to a single number such as, for instance, Dialnorm.

CLM	Description
Below -12	Very wide dynamic range, e.g. a movie. Significant processing to be expected for HD delivery.
-12 to -6	Wide dynamic range. Notable processing to be expected for HD delivery.
-6 to -2	Extended dynamic range. Moderate processing to be expected for HD delivery.
-2 to +2	On target for HD. Little or no processing to be expected for HD delivery.
+6 to +12	More uniform than needed. Little or no processing to be expected for SD delivery.
Above +12	Very narrow dynamic range, e.g. a pop CD. On target for iCast.

Table 1. The Consistency Loudness Measure, CLM.

In the example of *Fig 14*, the source is a hot pop track from CD, Madonna's Hung Up. The current loudness is at +14 LU (outer ring), the history is almost as loud (SLM=+13.5), and the consistency history shows very little variation (CLM=+14.8). The radar view reveals that the previous segment, a battle scene from the movie Pearl Harbor, ended 15 seconds ago. Pearl Harbor shows lower loudness and less consistency, and SLM/CLM descriptors at +6.4/+1.1 confirm this.

Fig 14. LU meter showing Realtime Loudness, Loudness History and Long-term Loudness SLM and CLM descriptors.

It should be noted how the SLM and CLM numbers are directly operational. In the example, Hung Up would be broadcast-fit if offset by -13.5 dB, while the scene in Pearl Harbor needs a correction of -6.4 dB. In both cases with positive CLM values, no further dynamics processing is indicated for HD. When all of Pearl Harbor is logged instead of just the battle scene, SLM shows -4.8 and CLM -11.5, suggesting an average offset of +4.8 dB plus substantial dynamics range processing before broadcast, like the curve in *Fig 15*.

METADATA

In DTV using Dolby AC3 data reduction, provision has been made to include extra information on top of the audio itself. Such information is known as "metadata", and added before transmission at the station. AC3 metadata allows three end-listener level control parameters to be set.

Dialnorm adjusts the receiver's level control in order to keep dialog at a constant level. The closer this setting gets to 0 dBFS, the lower the reproduction level. Line-mode DRC enables dynamic range restriction with wideband boost of low level, and compression of high level. RF-mode DRC does the same with additional level boost and limiting meant to be compatible with analog TV. The DRC settings specify a dynamic range reduction profile, with names such as "none", "speech", "music light", "film standard" etc.

The hope that AC3 decoders deployed inside consumer equipment would thus be able to restrict dynamic range appropriately at the listener has not been fulfilled, because AC3 is far from able to fill the gap between cinema and iPod. With a wideband design like this, pumping and other artefacts already become quite notable at boost or cut ratios of 6 dB [10], with much more regulation being indicated, see *Fig 1*.

Metadata only get used if they provide clear advantages without downsides. When benefits are not obvious, the extra work and equipment needed to create metadata, and the potential compatibility issues they may pose over time, naturally work against the concept. It's no wonder why broadcasters around the world are seeking more effective methods to control loudness than basing a station on part of a solution for one platform.

Fortunately, AC3 can work well without stations having to go through the trouble of using more of its metadata extension than changes between stereo and 5.1.

END-LISTENER DYNAMIC RANGE CONTROL

Using AC3 metadata as the main level and range control has other well known downsides than the ones described before. It is unpredictable how a consumer has her receiver set, and reproduction level becomes a mess when metadata is missing or wrong.

Acknowledging these problems, Dolby has introduced a chipped loudness control solution, Dolby Volume, to manufacturers of consumer equipment. Dolby Volume is single-ended, and doesn't require metadata to function. If its complexity is high enough, it may completely disregard metadata and not worry about if they are correct or not.

Single-ended consumer control of loudness has been a long time coming [4], but should be welcomed. Apple's relatively simple solution in iTunes was the first to offer a solution better than peak level normalization to the general public. With Dolby Volume, and other solutions to come, we can finally hope to rebuke the loudness war in music and film production.

With regard to broadcast, however, intangible consumer processing cannot be relied on. Metadata is one layer of extra unpredictability, single-ended consumer processing is another. Audio should therefore be adequately preconditioned at the station, and transmitted with fixed metadata to keep uncertainties at a minimum.

BEST PRACTICE

Based on experiences from broadcasters around the world, consistent audio and satisfied listeners are best assured when aiming HDTV transmission on nearly the same dynamic range signature as SDTV. For best results, the dynamic range should be only slightly wider, ref. *Fig 2*, with other platforms being fed and suitably processed off the HDTV stream.

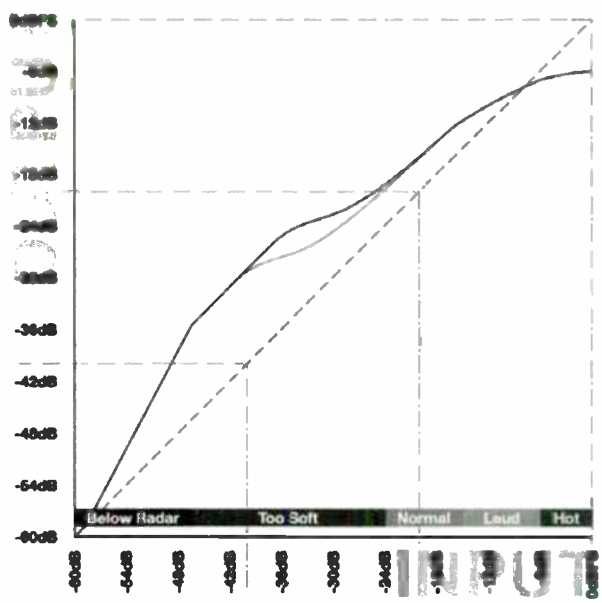


Fig 15. Example of dynamic range re-mapping of a 5.1 feature film to HD broadcast.
Black curve: Center channel.
Orange curve: L, R, Ls, Rs.

The widened dynamic range is made possible by centering all programming around a long-term loudness measure derived from ITU-R BS.1770 rather than the varying degrees of peak normalization used in broadcast today.

During *ingest* or inside the *server*, programming is normalized using the long-term Standard Loudness Measure. If the Loudness Consistency of ingested material is not high enough, dynamics processing is applied to comply with the HDTV dynamic range signature, example *Fig 15*. To precondition very wide dynamic range movies at ingest, it may be indicated to use a "telecine approach", transcoding sensitive scenes one at a time.

Access to a BS.1770 based loudness meter should also be provided in *production and editing*. A combination of a suitable realtime display, a history view and long-term descriptors has been shown in *Fig 14*. This type of display presents a target loudness and consistency measure during mixing, and informs about how much downstream processing will be applied for the production to fit various broadcast platforms. The new loudness measure also has the advantage of being understandable not only to audio experts, but to video editors, journalists and other non-specialists as well.

In *master control* and *transmission*, dynamic range conditioning for the different platforms takes place. HD has already been targeted upstream, so processing on this platform only plays a role when errors have been made at previous stages. Audio conditioning for other platforms is performed automatically, ref *Fig 10-13*.

For transmission where *metadata* is required, e.g. with Dolby AC3, the best practice is to keep Dialnorm, Line-mode DRC and RF-mode DRC fixed at certain values. With the gently widened dynamic range suggested here, Dialnorm should be set between -20 and -24 dBFS. A lower setting may generate more loudness at the end listener, but also more wideband processing taking place. Therefore, it's a sign of inadequate upstream level and/or processing, if the Dialnorm number has to be lowered to keep loudness aligned with other stations.

The DRC parameters too are set in a way not asking the impossible of the decoding processing. The best and most predictable audio results are obtained with Line-mode DRC disabled (setting "None"). RF-mode DRC should also be disabled, or set to one of the gentle profiles such as "Music Light" or "Film Light".

A content provider *delivery specification* describes the required SLM and max peak level. The peak level detection uses over-sampling, and is typically set 10 to 14 dB above the SLM. If only a sample peak measure is available, which is the rule today, digital ingest and files transfer may be louder and more distorted than expected [5]. It may also be helpful for a content provider to know a CLM target in order to understand what processing can be expected when the program is delivered to various platforms.

CONCLUSION

The paper has described how the open ITU-R BS.1770 standard can be used to streamline the flow of audio at the station by aligning content based on loudness, and implementing an automated trickle-down routine from HD to SD to IP to iCast. The "center of gravity" alignment also enables a station to mix content with different dynamic range signatures seamlessly, and without annoying level fluctuation.

Speech intelligibility and consistency of loudness are still the responsibilities of a broadcaster across all delivery platforms. Intangible consumer processing can not be relied on when dealing with such fundamental issues, and the Dynamic Range Tolerance of a consumer has to be respected.

A realtime loudness meter and powerful statistical descriptors have therefore been derived from BS.1770, and may be employed to automatically ensure loudness consistency within and between programs. The new SLM and CLM descriptors are also effective at specifying content, and make an AC3 based HD delivery system perform better by shedding light on its blind angles. Furthermore, less time needs being spent on audio issues at the station, and the tools proposed can be used effectively inside a file server, or by a person who is not an audio expert.

Over the next years, BS.1770 will have to prove itself against the currently used level measurement and control techniques at the broadcaster. The advantages of a separate peak and average measure will soon become clear, regardless if users in different parts of the world stick to their traditional reference levels or not.

In conclusion, BS.1770 is more about applying the same "center of gravity" level measure everywhere than necessarily having the most advanced loudness measure available from the start. A respectable and standardized average measure is better than any peak level normalization, and can help streamlining production, program interchange and transmission to various platforms with or without the use of metadata.

BS.1770 and the procedures described in this paper could also help put an end to the digital production loudness war, and maybe make the CD format its last casualty.

REFERENCES

[1] Julian Dunn: Measurement Techniques for Digital Audio, Digital Filter Overshoot and Headroom. Tech paper for Audio Precision. Beaverton, OR, 2001.

[2] Nielsen, S. & Lund, T.: Level Control in Digital Mastering. Proceedings of the AES 107 convention. New York, NY, 1999. Preprint 5019.

[3] Nielsen, S. & Lund, T.: Overload in Signal Conversion. Proceedings of the AES 23 conference. Copenhagen, 2003.

[4] Lund, T.: Distortion to The People. Proceedings of the Tonmeistertagung 23. Leipzig 2004. Paper A05.

[5] Lund, T: Stop Counting Samples. Proceedings of the AES 121 convention. San Francisco, CA, October 2006. Preprint 6972.

[6] Jones, B.L. & Torick, E.L.: A New Loudness Indicator for Use in Broadcasting. Proceedings of the AES 71 Convention. Montreux, 1982.

[7] Brian C. J. Moore, Brian R. Glasberg & Michael A. Stone: Why Are Commercials so Loud? – Perception and Modeling of Amplitude-Compressed Speech. JAES, no. 12, December 2003.

[8] Skovenborg, Quesnel & Nielsen: Loudness Assessment of Music and Speech. Proceedings of the AES 116 convention. Berlin, 2004. Preprint 6143.

[9] Skovenborg, Quesnel & Nielsen: Evaluation of Different Loudness Models with Music and Speech. Proceedings of the AES 117 convention. San Francisco, 2004. Preprint 6234.

[10] Lund, T: Control of Loudness in Digital TV. Proceedings of NAB BEC. Las Vegas, NV, April 2006.

[11] IRT: ARD Pflichtenheft 3/6, Aussteurungsmesser. Second edition. München, 1998.

[12] IEC: IEC 268-18, Peak program level meters - Digital audio peak level meter, First edition. Geneva, 1995.

[13] IEC: IEC 268-17, Standard volume indicators. First edition. Geneva, 1990.

[14] ITU-R, WP6P: Audio Metering Characteristics Suitable for the Use in Digital Sound Production. Geneva, 2000.

[15] ITU-R, Draft BS.1770: Algorithms to measure audio programme loudness and true-peak audio level. Geneva, April 2006.

[16] Kurozumi, K & Hoshi, H. Realtime Loudness Meter for Digital Broadcast. AES 11th regional Convention. Tokyo, 2003.

[17] Norcross, S & Soulodre, G: Objective Measures of Loudness. Proceedings of the AES 115 convention. New York, NY, 2003. Preprint 5896.

[18] Norcross, S & Soulodre, G: Stereo and Multichannel Loudness Perception and Metering. Proceedings of the AES 119 convention. New York, NY, 2005. Preprint 6618.

[19] ITU-R, Draft BS.1771: Requirements for loudness and true-peak indicating meters. Geneva, April 2006.

[20] NABA presenting North American viewpoints to the ITU on question R-2/6: Proposed specification for audio programme loudness. Geneva, August 2006.

STL Bandwidth Requirements for Radio

Monday, April 16, 2007
10:30 AM – 12:00 PM

Chairperson: Milford Smith
Greater Media, Inc., Lawrenceville, NJ

Audio Over IP – Can It Really Work?
Kevin Campbell, APT, Belfast, Northern Ireland

E2X Bandwidth and Bit Error Requirements for Ethernet Synchronization
Philipp Schmid, Nautel Ltd, Hackett's Cove, Canada

Networking and STL Issues When Implementing Multicasting for HD Radio
Richard Hinkle, Broadcast Electronics, Inc., Quincy, IL

AUDIO OVER IP – CAN IT REALLY WORK?

Kevin Campbell

APT, Waltham, MA

INTRODUCTION

In the past broadcasters have utilized a number of different telecommunication mediums to move audio content between locations. Balanced analog copper circuits, microwave systems both licensed and unlicensed and synchronous based systems such as V.35 or T1 have been at the heart of broadcast networks for many years. Now, however we are at the beginning of a sea change in the way broadcasters move content for applications such as Studio Transmitter Links, Transmitter Studio Links, Inter-Studio Networking and Remotes. We are at the beginning of the IP age.

This paper seeks to clarify the pitfalls and challenges faced by those looking to create an IP Broadcast Chain. In doing so, I have divided these challenges and pitfalls into two areas, Network related parameters and Equipment Related Parameters. We will also examine the advantages to implementing or migrating to an IP transport mechanism. Finally we will look at all the issues covered and answer the question posed...Audio Over IP Can It Really Work?

NETWORK RELATED PARAMETERS

Network Selection

It is of course of paramount importance that the network be fit for purpose. It must be of sufficient quality to support the uninterrupted flow of packets from point A to point B and it must be configurable with regards to the parameters which I will describe in detail later in this paper. In short, Audio Over IP for broadcast quality applications cannot be achieved on the Open Internet, contentious ADSL links or contentious WANs or LANs. If you cannot control the IP Network, then you cannot control the quality of the audio emanating from that network.

Protocol Selection

There is a common mistaken assumption when broadcasters first broach the subject of audio over IP that TCP will be the most appropriate protocol. TCP is a connection-oriented protocol. This means that a connection is created between sender and receiver and, as packets flow over this connection in the order they were sent, every packet must be acknowledged as received by the receiver to the sender. For the purposes of real time audio delivery, the fundamental flaw with this protocol comes when a packet is dropped. Having sent a packet over TCP/IP, the sender will repeatedly ask the receiver for an acknowledgment that the packet was received and, if the packet is dropped, this acknowledgement can never come. This produces the ugly side effect of data peaks on the IP link as bandwidth is swallowed up in a fruitless request for a receipt acknowledgement. The repeated acknowledgement request packets and the compressed audio payload are now in competition for the available bandwidth and the audio will begin to glitch.

UDP as a connectionless protocol employs a "send and forget" strategy with no acknowledgement required making it much more suited to real time audio applications over reasonable data bandwidths. RTP is now also being deployed along with UDP to aid the time stamping and sequencing of packets. Strategies for dealing with packet loss are discussed in further detail later in this paper.

Specifying Network Bandwidth and Packet Size

The choices made with regard to audio settings will define the data bandwidth you need to transport encoded audio over the IP network. For synchronous connections this actually equates to the bandwidth required to transport the compressed audio.

However, in IP we must add an overhead that is required to packetize the audio data. Encapsulation into an IP packet adds header bytes containing information related to the routing of the packet along the IP Network. The packet information is examined by the routers and switches which constitute the routing intelligence on the network and, based on the packet header information, decisions are made on how, when and where to route a packet. Included in the IP Packet Header are:

- 8 Bits - Containing the Type of Service or Quality of Service
- 16 Bits – Containing the length of the packet

- 16 Bits – Containing an identification tag to assist in reconstruction of fragmented packets
- 8 Bits that contain the protocol (TCP, UDP, RTP)
- 32 Bits that contain the source IP Address
- 32 Bits that contain the destination IP Address

A packet header containing this data (and more) must be included in every packet that originates from an IP device on an IP network. There is a correlation then to be found between packet size and bandwidth requirement. The table below details the IP Data Rate for audio transportation using the Enhanced apt-X algorithm as an example.

Table 1: IP Data Rates using Enhanced apt-X

Audio Algorithm	Frequency Response	Audio Data Rate (kbps)	Packet Size Bytes	IP Packet Size Bytes	IP Pkts/Sec	Packetization Delay mS	IP Data Rate (kbps)
Eapt-X 16-Bit	3.5kHz Stereo	64	128	194	62.5	16	97
Eapt-X 16-Bit	7kHz Mono	64	256	322	31.25	32	80.5
		64	512	578	15.625	64	72.3
		64	1280	1346	6.25	160	67.3
Eapt-X 16-Bit	7kHz Stereo	128	128	194	125	8	194
Eapt-X 16-Bit	15kHz Mono	128	256	322	62.5	16	161
		128	512	578	31.25	32	144.5
		128	1280	1346	12.5	80	134.6
Eapt-X 16-Bit	15kHz Stereo	256	128	194	250	4	388
Eapt-X 24-Bit	20kHz Mono	256	256	322	125	8	322
		256	512	578	62.5	16	289
		256	1280	1346	25	40	269.2
Eapt-X 16-Bit	20kHz Stereo	384	128	194	375	2.7	582
		384	256	322	187.5	5.3	483
		384	512	578	93.75	10.7	433.5
		384	1280	1346	37.5	26.7	403.8
Eapt-X 24-Bit	20kHz Stereo	576	128	194	562.5	1.8	873
		576	256	322	281.25	3.6	724.5
		576	512	578	140.625	7.1	650.3
		576	1280	1346	56.25	17.8	605.7

It is clear that choosing a larger packet size will reduce the bandwidth as the header bytes will represent a smaller proportion of the total bytes i.e. more payload, less overhead. However, with larger packets there is a downside. Firstly, latency will increase as the packetization delay is greater. Secondly, if a large packet is dropped there will be a correspondingly larger amount of payload dropped - in the real world this equates to dropped audio.

Choosing the optimum packet size will always be a balance between bandwidth efficiency, network performance (in terms of dropped packets) and audio quality.

Packet Loss – Correction, Concealment, SLAs & Performance Monitoring, Synchronous Backup

All packet-based systems are susceptible to dropped packets, resulting in dropped audio. As discussed, we can negate the problem by choosing smaller packet sizes but in so doing we incur greater packetization delays and bandwidth requirements. The other options for dealing with packet loss are concealment, correction or temporarily abandoning the packetized network in favor of an automated backup to a synchronous network.

Correction

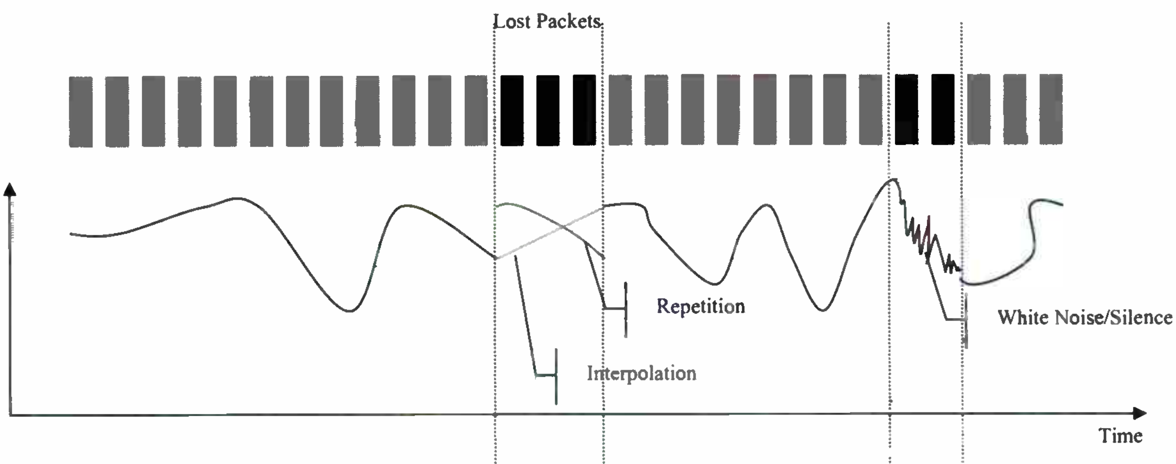
No forward error correction is currently used in the UDP packet. The disadvantage of this is that, as the integrity of the received data is not verified once received, packets are susceptible to undetected bit errors. On the flip side, bypassing the additional correction stage increases the delivery speed of packets. Forward Error Correction techniques have been implemented successfully in audio streaming applications, however the implications for real time audio delivery due to the processing and data overheads have yet to be reconciled. Other issues that limit the effectiveness of FEC in audio broadcast applications over IP are the dynamic nature of the packet losses on a network. For example, will the FEC implementation be able to handle bursty packet losses or a percentage of lost packets if the packets lost are non consecutive? Also, to maintain compatibility between IP codec vendors, the FEC Information should be sent on a different port to ensure that the decoding codec does not become confused if it cannot handle the sent implementation of an FEC scheme.

Concealment

With correction not a viable option for dealing with dropped packets the next viable option is to look at concealing those dropped packets.

There are a number of methods of concealment ranging from simple repetition of the last good packet received to silence/noise injection or interpolation and retransmission. All have an impact on the reproduced audio.

Figure 1 Packet Loss Concealment



In listening tests injection of silence predictably produced unacceptable breaks in the audio that led to a degree of incoherence. The injection of white noise improved intelligibility of the reproduced audio but was again noticeable. The use of repetition of the last known good frame produced favorable results. The use of interpolation/pattern matching/waveform substitution to conceal the loss of packets is also possible and again produced favorable results. However, when complexity is measured against the benefit to the listener, these last two techniques are delivering a diminishing return. Also the results of all these techniques will be governed by subjective improvements in audio quality and by the amount of packets being concealed.

SLAs & Performance Monitoring

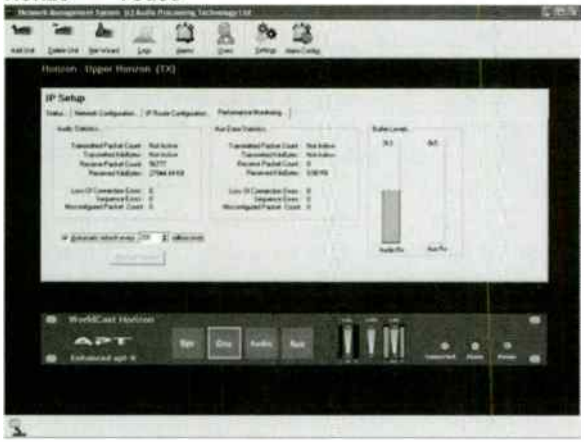
Although not a short term fix to packet loss the use of SLAs (Service Level Agreements) and Performance Monitoring can provide long term stability on an IP Broadcast link in guarding against packet loss. Any Telco or provider will issue a SLA against an IP link typically guaranteeing uptime in percentage terms. This percentage can be reconciled to lost packets and actual down time on the link. However, to truly measure that SLAs are being adhered to, the terminating codec on a link must run non-intrusive performance monitoring on that link.

Figure 2 Perceived audio quality

Silence	White noise	Repetition	Interpolation	Retransmission

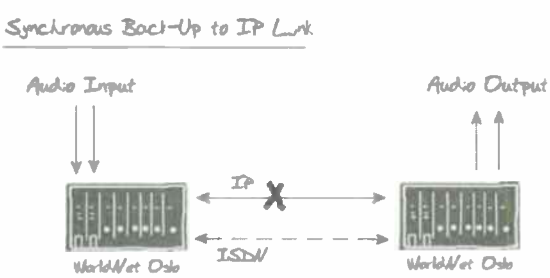
Complexity/Delay

Figure3: Performance Monitoring on APT's WorldCast Horizon IP Codec



Synchronous Backup

Even with all necessary due diligence applied in the selection of the IP Network and Service Provider the network is likely to suffer a major outage at some stage. This can equate to consecutive dropped packets over a sustained period of time but is more likely to be large consecutive bursts of dropped packets. In this instance, the broadcaster is effectively off-air unless they have a backup. With a synchronous link over IP or T1 a point-to-point backup can be created. A many to one backup is also possible, however this backup will only be effective if the Primary IP links are diversely routed.



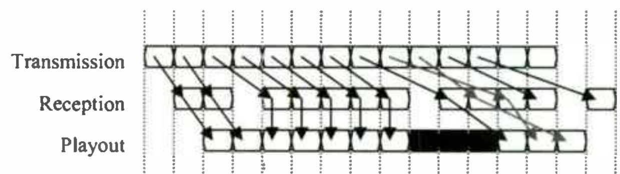
The backup from the primary IP link to the secondary synchronous link can be triggered by a number of different criteria on the APT WorldNet Oslo system. The trigger to switch to backup could be silence on the audio output of a specific audio module or a defined threshold in the Performance Monitoring log. Similarly, the automated restore back to the primary IP link could be defined in the Performance Monitoring log i.e. number of consecutive packets received without a single drop would equate to a restoration of the primary link.

EQUIPMENT RELATED PARAMETERS

Jitter

Jitter can be defined as packets that are received either side of the predicted arrival time. Given that any packet can take any route from transmit to receive, jitter is an inherent characteristic of packet switched networks. Packet jitter can be stipulated and defined under an SLA with the service provider, establishing an upper limit of the permitted jitter value. However it will always be present in an IP based network and so codec selection is vital in dealing with this. The buffer depth is critical in allowing the codec to provide enough time for the packet to be received and decoded before its play-out time. Reducing the buffer size reduces the jitter time mask available and increases the possibility of being forced to drop packets that have arrived beyond their play-out time.

Figure 3 Network Jitter Effects:



This diagram shows the effect of network jitter on the reception of audio and its subsequent play-out through an audio system. The buffer depth is usually set in milliseconds; in this case it is set to a two-packet buffer. For the purpose of this example this allows for up to a two-packet jitter delay in the system. Provided the network jitter is low the system is unaffected and plays out the packets received in sequence. However, should the jitter increase, there is the distinct possibility that the packets will arrive after the determined play-out time. In this example the packets have to be dropped, which results in the audio being corrupted.

Latency

All networks have transport latency due to the natural laws of physics. Transporting an electronic signal through whatever medium takes a finite amount of time that cannot be removed. In a switched network there is both the standard transmission delay and also the packetizing delay to contend with. This packetizing delay ensures that the delay on an IP network is always greater than on a synchronous network. Therefore special attention must be paid to latency if real-time audio delivery is to be achieved.

By definition a packet must be assembled and consists of a header plus payload. The size of that payload can be varied but ultimately it consists of an audio sample. Take for example the Enhanced apt-X algorithm, a system that uses a four to one compression algorithm and has a packet size of 128 bytes. That's 512 audio samples, equal to 666 µsec in Mono and 333 µsec in stereo. Then take the time it takes to propagate through the UDP stack after being assembled into a Real-time Transport Protocol packet. In real, live-unit tests this is approximately 20-30 ms using the most optimal codec equipment available on the market. It is important to realize that this natural latency increases as the sample frequency decreases. With the inherent latency in the protocol stack, the need for efficient management of the UDP stack and the packetizing process by the hardware codec is crucial for the maintenance of low latency broadcast circuits. Also crucial to countering latency is the implementation of QOS within the packet header. QOS can be set through the IP communicating equipment, for example an IP codec, and determines that the packet is given priority over other packets on the network that do not have a QOS value. The additional delay of the audio coding is also critical in real time audio applications and is my final consideration in equipment selection.

Algorithm Selection

Having thoroughly investigated the intricacies of IP as a method of moving program content from Point A to Points B, C, D and through to Point X, the next step is to look at the best method of layering in audio on top of the transport stream. In essence there are two options – PCM / linear or using compression to reduce bit rates. Within compressed there are two sub-options - perceptual or ADPCM.

PCM or linear audio is well defined in terms of the audio - what you get in should be what you get out, assuming there are no problems relating to analog-to-digital conversions, signal-to-noise ratios or quantization issues. The compelling reason not to choose linear is directly related to the data bandwidths required.

A stereo signal sampled at 44.1 kHz, with a word depth of 16 bit, will require a data rate of 1.411Mbit/s (plus 10 – 15% overhead and additional for FEC and synchronization algorithms). This data rate bandwidth will cause stress on the IT network passing the data. If the broadcaster adds in additional channels (5.1 or more Stereo signals), deepens the word depth to 24 bit and increases the sampling frequency to 96kHz (or even 192kHz for the small furry animals that happen to be listening), it soon becomes apparent that what was a benign solution has now turned into a network nightmare.

Making the decision to use compression opens up an interesting argument. Two options are available:

The perceptual based algorithms using psycho-acoustic based principles that can generally be described as "Lossy". Some examples are MPEG Layer II, MPEG Layer III (MP3) and AAC (including the myriad of derivatives). These algorithms are heavily processor hungry and remove content that is perceived to be irrelevant. As such, they result in content that vaguely resembles the original (especially after several passes) and has a long latency i.e. 50+ milliseconds.

The other option is to use the relatively non-destructive Enhanced apt-X algorithm, which is based on ADPCM principles. This algorithm offers a low delay of less than 2 milliseconds and has exceptional acoustic properties. These acoustic claims have been confirmed by independent listening tests. The most recent listening test undertaken was with a group (approximately 20) of Chief Engineers from the GWR group (now GCAP after they merged with the Capital Group). This was a double blind listening test with 10 audio samples (different genres (Classical, Pop), a cappella, spoken voice (Male / Female)). We tested Enhanced apt-X, MPEG and J.41. Enhanced apt-X was shown to be indistinguishable from the original PCM. The Enhanced apt-X algorithm can also offer word depths of 16, 20 & 24 bit, thus significantly improving the dynamic range to greater than 110dB.

Working on the assumption that the IP transport stream will naturally introduce a minimum delay of 20+ milliseconds, reducing the latency of the compression algorithm becomes an imperative when considering the design of a broadcast network. In essence, using a perceptual coder will render the solution unusable for any level of live event that requires off-air monitoring, whereas using Enhanced apt-X will offer broadcasters a viable alternative. Any overhead using apt-X relates to the RTP/UDP protocol used for the transmission of the compressed audio data and will be approximately 10-30% depending on packet size.

Along with the well-documented features of low latency and audio performance, Enhanced apt-X also has an embedded word pattern to aid connection and synchronization. AutoSync aids the ability to quickly synchronize i.e. 3 milliseconds on start up or in the event of a drop out. In addition the predictive nature of Enhanced apt-X allows for the masking of lost packets. As such, both features allied together act as a form of FEC.

On a more subjective issue, using multiple passes of a perceptual codec (for example, consider the final emission for HD Radio or DAB) will result in content heavy with artifacts. Ultimately these will cause "listener fatigue," swiftly followed by users tuning to another station that sounds better because it uses less destructive coding algorithms.

CONCLUSION

It is patently clear that today's broadcasters are convinced not only that Audio Over IP can work but are determined to ensure that it does. The benefits of packetized networks over existing synchronous networks for audio transport have proved too persuasive for them to ignore:

- Synchronous networks are not easily scalable in comparison to IP
- Synchronous networks can be inefficient i.e. only supplied as T1 (1.5Mbits) or E1 (2Mbits)
- Synchronous links tend to be more expensive
- Not easy to achieve flexible multipoint configurations over Synchronous
- Widespread viability of IP Networks
- Consolidation of Engineering skills

However, despite these overwhelming arguments in favour of IP audio transport, broadcasters must also ensure that the quality and reliability of their existing networks are not sacrificed in the quest for greater efficiency and cost-savings. Therefore, the following checkboxes must be ticked

- An IP Network which is fit for purpose running a Protocol suitable for real-time delivery of Audio
- A robust hardware codec solution with redundant options and powerful performance monitoring to help enforce SLAs
- A hardware codec solution that allows parameters to be adjusted to take account of network conditions, parameters such as Jitter, Packet Size and Latency
- A low delay algorithm such an Enhanced apt-X to counter the inherent delay associated with IP Networks

E2X Bandwidth and Bit Error Requirements for Ethernet Synchronization

Introducing a Reliable Real-Time Point-to-Multipoint E2X Transport Protocol

PHILIPP SCHMID
Nautel Limited
Bangor, Maine

INTRODUCTION

The introduction of the generation 3 digital radio systems broadcast architecture transforms the broadcast system from an audio based broadcast system to a general digital data broadcast system that carries audio content along with advanced application services (AAS). The transition to this architecture necessitates changing traditional audio based studio transmitter links (STLs) to generic IP based data links.

Whereas IP streams are ideal for transferring arbitrary digital information not limited to audio content, many of the challenges inherent to IP based streaming are now imposed on this architecture. The nature of this system places stringent requirements on the STL that can cause significant on air data outages.

The network and bandwidth requirements of carrying the data stream to the exciter has already been well analyzed and published[1]. However, to deploy a system that is successfully synchronized based on the data stream rather than external GPS based synchronization brings an additional set of challenges which are presented in this article.

An alternate transport protocol to UDP or TCP/IP is presented in this article, that drastically improves data loss across the STL by implementing an automatic repeat request (ARQ) scheme. This protocol considers the impact on studio transmitter synchronization due to retransmissions.

The protocol operation is detailed and illustrates its capability of addressing multiple exciters with a reliable data stream. The article demonstrates a main/standby configuration, a dual station topology, as well as, the protocol's applicability for satellite distribution of the E2X data stream.

GENERATION 3 IBOC BROADCAST SYSTEM

The 3rd generation IBOC radio broadcast system architecture differs from previous generations in that most of the IBOC functions and equipment are moved from the transmitter site to the studio site. Illustration 1 provides a basic overview of the architecture.

The only function in the IBOC system that remains at the transmitter site is the digital IBOC modulation performed inside the exciter engine (Exgine) modulator board. Keeping the modulation function at the transmitter site is necessary since the digital bit stream to be modulated can be transferred in fewer than 150 kbps. IBOC modulation turns this bit stream into a discrete time signal that requires the equivalent of 23 Mbps to be transferred across an STL. Though required to contain the complete FM baseband frequency spectrum, it is inefficient to carry this amount of data across bandwidth limited STLs.

The transmitter site receives two data feeds from the studio: an audio feed for traditional FM broadcast, which may be carried in any digital or analog industry standard, and an Ethernet based digital data stream for IBOC modulation. This digital data stream is not limited to carrying audio data and may contain program service data or other data services along with one or more additional compressed audio streams.

The major IBOC system components at the studio site are the Exporter, Importer and system synchronization unit. The Exporter's responsibilities include delaying the analog audio to match the digital on-air delay and passing the audio signal to the STL. On the digital side

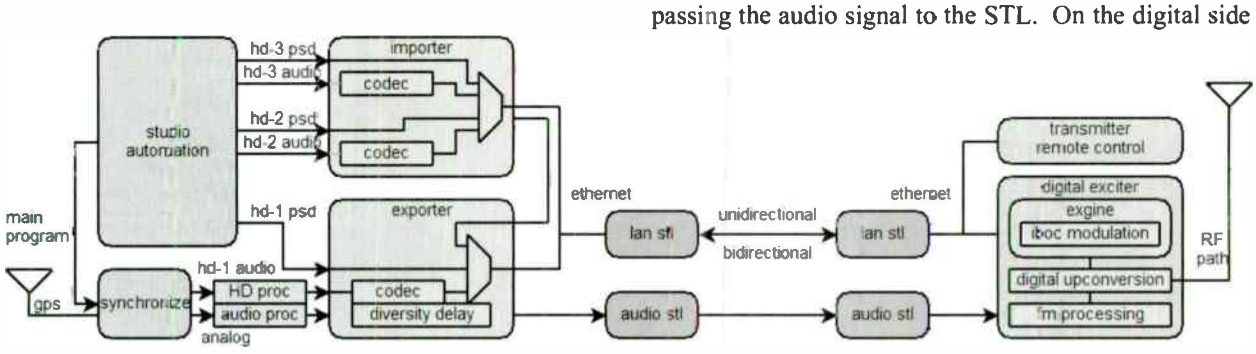


Illustration 1: Generation 3 IBOC FM Radio Broadcast Architecture

the Exporter performs the audio coding of the digital main program and multiplexing all digital content including program service data for the main program, as well as, all digital content originating from the Importer.

The Importer can capture two additional audio feeds for secondary program services which are encoded in the Importer and are passed to the Exporter along with their corresponding program service data.

The broadcast architecture is synchronous in nature in that a discrete number of audio samples captured at the studio translates into a discrete number of digital symbols produced at the exciter. The synchronization unit synchronizes the incoming audio feed to a GPS based timing reference, which in turn affects the timing of all operations in the system.

The STL link must now be able to carry Ethernet/IP based traffic along with traditional audio streams, requiring a digital STL. Many stations have already transitioned to digital STLs in favor of a digital audio feed over an analog feed for improved audio quality. Most STL vendors now provide upgrade paths to share the existing digital STL bandwidth between audio paths and Ethernet based paths. However, since most STLs have traditionally been unidirectional in nature, many STLs today also only provide a unidirectional Ethernet path. Even basic network functions, such as the address resolution protocol, often require a bidirectional data path between a sender and receiver on a network. It is important that the digital data stream protocol does not require a bidirectional data path, but it may make use of a bidirectional link if one is present.

Classifying the bandwidth requirements of the data portion of the digital data stream is relatively straight forward and iBiquity Digital Corporation has sufficiently detailed these requirements. The bandwidth requirements for the synchronization stream, on the other hand, are more difficult to determine.

SYNCHRONIZING THE IBOC DATA STREAM

The digital data stream across the STL is a generic bit stream rather than a discrete time signal. This means that this data stream cannot be sample rate converted, since each bit must remain unaltered. Consequently, both the Exporter at the studio and the Exgine at the transmitter must remain frequency locked in order to process the data at precisely the same rate. At the transmitter site the reception of the data stream is closely coupled to the digital modulation stage and, therefore, the data stream indirectly dictates the digital symbol rate. Disciplining the exciter with a second GPS receiver at the transmitter site does allow for effective frequency synchronization and relieves STL requirements somewhat. However, this approach increases the overall component count and does not control digital signal throughput delay without absolute frame alignment.

This synchronization requirement fundamentally differs from standard network communications, where data exchanges are often flow controlled by the receiver as in the case of TCP/IP. The synchronization unit cannot be placed at the transmitter site, as it does require a bidirectional STL in order to rate control the studio. Placing the synchronization unit at the studio does account for unidirectional STLs and allows for the extension of the architecture to synchronize multiple transmitter sites simultaneously.

Illustration 2 depicts a source synchronous clocking scheme operating across an asynchronous transport medium, such as an Ethernet or IP based link, and recovers the source processing rate in order to rate lock the digital modulation process. The system is disciplined by a high quality reference frequency, such as the 1 Hz signal generated by a GPS receiver, all other processes fall into place with respect to this reference signal. The audio feed is re-sampled with the GPS disciplined sample rate to ensure frequency stability rather than depending on the clock signal within the audio feed. The re-sampled audio feed is encoded and other digital processing is applied to create the data stream to traverse the link. After a fixed number of audio samples, clock packets are inserted into the data stream.

At the transmitter site, the exciter uses the clock packets to

a) recover the studio's processing rate by disciplining a frequency locked loop or a phase locked loop by aggressively rejecting high frequency clock packet jitter introduced by the link, and

b) maintain a constant throughput delay by estimating an appropriate starting time based on averaging an initial number of clock

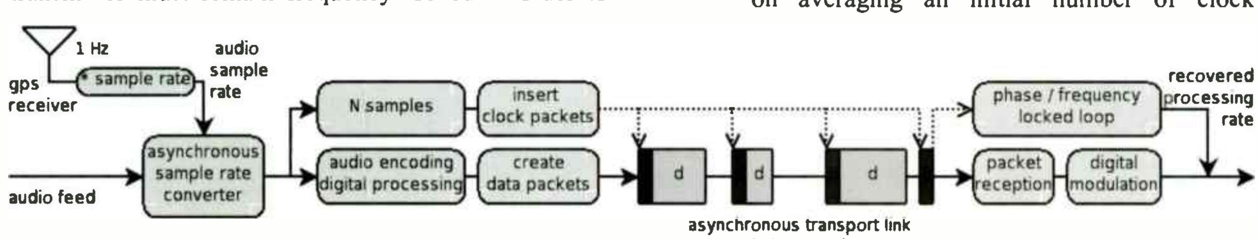


Illustration 2: Source Synchronous Clocking across an Asynchronous Data Link

packets. Clock packet jitter normally falls in the range of milliseconds, but in extraordinary circumstances can be as high as several seconds or more depending on the situation and type of STL. Therefore, it is not prudent to determine the initial digital modulation based on a single link dependent event. A phase locked loop can correct for this initial phase offset error, but a frequency locked loop will maintain this error indefinitely.

Exciter synchronization does not generally care about the link throughput delay. It will, however, react to long term changes in throughput delay, hence, STLs that automatically adjust modulation schemes dynamically based on transmit conditions are generally undesirable in this situation.

It does make sense to discipline exciter synchronization using clock packets. This minimizes the impact of various link characteristics by reducing the impact of packet serialization delays through a smaller packet structure. However, in order for clock packets to represent a true measure of throughput delay, the intermediate data packets must not impact clock packet timing, so we can determine a dedicated link bandwidth rate of

$$dedicated\ link\ rate = \frac{8 * max(data\ bytes)}{clock\ packet\ period}\ bps$$

Two synchronization levels are specified by the IBOC standard that define the maximum allowed symbol clock frequency error.

- Level I: must meet a maximum error of 0.01 ppm This requirement is difficult to attain given the typical amount of clock packet jitter even in optimal network deployments. In this case, a second GPS synchronizer unit is required at the transmitter site.
- Level II: must meet a maximum error of 1 ppm This requirement can be met using Ethernet based synchronization, provided the STL does not inject clock packet jitter greater than 15 ms and the STL bandwidth meets the minimum dedicated link rate outlined above.

The impact of data stream induced timing errors due to an insufficient bandwidth can be mitigated to some extent, however, the embedded timing information is degraded leading to less stable digital symbol rates and less predictable throughput delays.

EXPORTER TO EXCITER PROTOCOL

The data stream protocol between the Exporter and exciter components is termed the E2X protocol and is defined as part of the standard IBOC Radio System Broadcast Architecture. It implements a source synchronous clocking scheme similar to the one described in the previous section. The FM IBOC system transfers all its payload data in Layer 1 frame intervals of 1.48 seconds. This frame is broken into 16 data packets sent after a clock packet sent every 92.8 ms. However, the payload is not evenly distributed across all 16 data packets in the L1 frame as shown in Illustration 3.

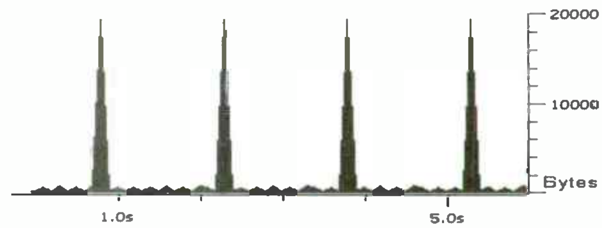


Illustration 3: Service Mode 3 Bandwidth Utilization

With a maximum data packet size ranging from 18788 Bytes to 19330 Bytes, the minimum dedicated link bandwidth is around 1.5 Mbps using the relationship developed in the preceding section. Even though these data packets are broken into smaller IP packets they are all delivered to the STL at the same time creating a temporary congestion condition. On a 256 kbps link, this congestion period can be around 600 ms affecting several subsequent clock periods. The exciter's receive buffer must be sufficiently deep to absorb this data imbalance; a receive buffer of at least one layer 1 frame of 1.48s is recommended, bringing the overall analog-to-digital delay to just under 9 seconds.

It should be noted that the E2X protocol can be successfully deployed with the recommended bandwidth requirements stated by iBiquity Digital Corporation for the E2X protocol[1], but unless a bandwidth greater than or equal to the ideal dedicated bandwidth is employed exciter synchronization is degraded using Ethernet based synchronization. GPS based synchronization can be used for STLs with lower bandwidth.

The E2X protocol follows the standard Open Systems Interconnection (OSI) networking model and limits itself to the application, presentation, and session layers. It assumes standard Ethernet technology for the physical and data link layers, but intermediate links do not strictly have to be Ethernet. The Internet protocol (IP) is employed for the network layer. E2X initially only supported the User Datagram Protocol (UDP) as a transport layer protocol, which is a fire-and-forget best effort protocol that does not guarantee data delivery. UDP can, however, address multiple destinations at once through the use of broadcast or multicast communications and also works on unidirectional STLs. The transmission control protocol (TCP) is in the process of being adopted by the E2X protocol, but at the time of writing is not yet fully supported. TCP does offer guaranteed ordered end-to-end data delivery, but is limited to point-to-point communications and requires a bidirectional link.

Since UDP is widely used in many E2X deployments, it is of interest to investigate the associated STL quality requirements. A single bit error on the STL can cause an entire data packet to be dropped, which affects an entire L1 frame in the IBOC data stream causing a 1.48 second audible HD outage. The main program may blend back to the analog audio transmission, but secondary programs and data services experience an irrecoverable interruption. Bit error rate is not an adequate measure for quantifying STL requirements in a networking environment, since many types of packet loss in a network are not directly related to bit errors. However, bit error rate is a convenient measure to assess protocol performance and provides a guideline for an appropriate signal to noise ratio on an RF based STL.

The effect of a single bit error on the STL can be contrasted between a comparable AES audio feed and the E2X data stream. With the exception of the impact on the control and status flow of the AES stream, a single bit error only extends to a single sample and even in this case errors can be concealed by interpolating neighboring samples. So it is unlikely that a single bit error is even perceivable to a listener. In the IBOC system an interruption of 1.48 seconds is very perceptible. If we define the delivered quality of service (QoS) as the ratio of flawless transmission over the total transmission time, then we can define the allowable mean time between failures (MTBF) as:

$$MTBF = \frac{error\,duration}{1 - QoS}$$

$$MTBF = \frac{1.48\,s}{1 - 0.99999} = 148000\,s$$

The required bit error rate to sustain this quality of service is then:

$$BER = \frac{1}{MTBF * bit\,rate}$$

$$BER = \frac{1}{148000\,s * 115503\,bps} = 5.85 * 10^{-11}$$

Correspondingly for AES at 44.1kHz, the same quality of service requires a bit error rate of:

$$MTBF = \frac{22.68\,us}{1 - 0.99999} = 2.268\,s$$

$$BER = \frac{1}{2.268\,s * 1.41\,Mbps} = 3.12 * 10^{-7}$$

For the digital performance this means an interruption of service roughly every 41.1 hours. The bit error rate can be converted into a packet loss rate of about $4.22*10^{-7}$. Ibiquity recommends a maximum packet loss rate of 10^{-5}, which may cause several interruptions every hour[1].

As shown, the performance requirements on the STL in a UDP type environment are very stringent. We can use either forward error correction (FEC) or automatic repeat requests (ARQ) to effective reduce the required bit error rate. FECs add redundant information into the data stream that allow the receiver to reconstruct the original message with corrupted or missing message parts. This technique adds additional bandwidth requirements to the data stream regardless of the actual error rate. Forward error correction works well on the physical link level where many bit errors can be present, but only provides diminishing returns in environments of lower bit error rates.

ARQ, on the other hand, relies on error detection at the receiver and issues a retransmission request back. This only requires additional bandwidth in the case of detected bit errors. This scheme, does require a feedback path in order to issue retransmission requests.

Even a single retransmission of a bit in error can greatly improve the effective bit error rate, as shown in the following example that provides a much better effective bit rate based on a channel bit rate of $7.65*10^{-6}$.

$$BER_{eff} = BER^2 = (7.65 * 10^{-6})^2 = 5.85 * 10^{-11}$$

Regardless of the type of error correction that is chosen, the impact of redundant data transmission or retransmissions can impact the exciter's synchronization ability, if it impacts clock packet transmission.

E2X TRANSPORT REQUIREMENTS

Overall the 3rd generation radio broadcast system architecture is required to propel digital radio to the next level. It is architecturally sound and the E2X protocol in itself works reasonably well. The protocol is not directly responsible for the previously illustrated challenges. Rather, the 3rd generation radio broadcast architecture is confronted with the challenges of high performance real time data streaming, a non-trivial challenge confronting the networking community. Good networking practices can ease the presented challenges, however, the problem is that there is no established transport protocol that adequately addresses all the requirements of the E2X protocol.

The E2X protocol poses the following transport requirements:

- Some packet types of the E2X protocol such as the initial control packet that carries critical system information, such as the configured service mode, require **guaranteed delivery reliability** in order to start digital modulation.
- E2X data packets only require **limited delivery reliability**, since data packets only convey meaningful information prior to the time that they

are intended to be on air. Retransmission requests beyond that time are futile and waste valuable bandwidth resources. A late delivery is essentially a packet discard.

- It is undesirable to retransmit any E2X clock packets, since any retransmission introduces a throughput delay measurement error. Depending on the implementation of the clock recovery circuit, most implementations can handle lost clock pulses better than incorrect clock pulses. **No delivery reliability** is, therefore, desired for clock packets.
- The protocol requires a **very high degree of reliability** in order to maximize the mean time between HD dropouts. If the transport protocol cannot provide an adequate degree of reliability, then the physical link layer is responsible for reliable data delivery instead.
- Clock packet delivery must not be impacted by
  - data packet transmissions, which can introduce a patterned throughput delay error that is difficult to recover from.
  - data packet retransmissions, which could make the throughput delay dependent on current link characteristics, such as environmental conditions.
  - aggressor traffic that is trying to share the link with the E2X protocol and may upset synchronization and even starve the digital modulator.
  - other E2X traffic sharing the link with the E2X protocol.
- **Unidirectional and bidirectional STLs** must be supported.
- **Multiple destinations** should be addressable to support multi exciter configurations, as well as, multi and single frequency networks.

A NEW E2X TRANSPORT PROTOCOL

Looking at this list of requirements it is apparent that neither TCP nor UDP adequately cover all items on the list. Nautel fully supports the existing E2X protocol design including the standard transport protocol choices. In addition Nautel has implemented a new transport layer protocol that addresses the E2X protocol transport requirements and allows multiple exciters to synchronize to the same Exporter synchronization stream while providing reliable data transfers to each destination.

The proposed E2X transport protocol implementation is layered on top of UDP, since UDP allows broadcast and multicast communication, however, it may be possible to port the protocol to the real time protocol (RTP) in the future. The protocol is best illustrated by looking at a main/standby exciter configuration across a generic STL as shown in Illustration 4.

The protocol is broken into a control plane and data plane. The data plane carries the encapsulated and segmented E2X protocol to both exciters. The Exporter can directly communicate with a single exciter using unicast IP, which directly maps a single IP address, such as 10.10.10.100 to a unique MAC address, such as 00:50:C2:59:70:8C. In order for the Exporter to use this mode of communications, a bidirectional STL is required to be able to use the address resolution protocol. Alternatively, the intermediate router can be programmed with a static MAC entry in the case of a unidirectional STL. To address multiple exciters simultaneously, the Exporter must resort to either broadcast mode (without an intermediate router), directed broadcast (through the router), or multicast IP to address only a subset of exciters.

The control plane communication from the Exporter to the exciters follows the same communication modes as the data plane, but the Exporter may choose to only directly address one of several exciters. For example, the Exporter may grant the ability to issue retransmission request only to a subset of listening exciters allowing for a highly scalable point to multipoint deployment of this protocol. For the time being this function is reserved but not implemented.

The control plane communication from the exciters back to the Exporter is always unicast directly to the Exporter. In the case of a unidirectional STL, this communication stream will not make it back to the

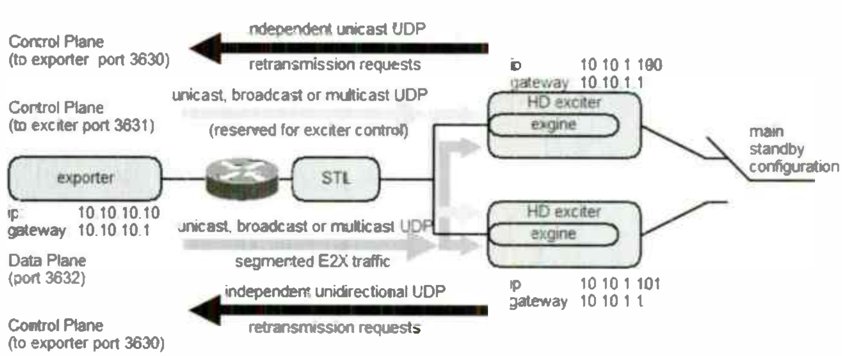


Illustration 4: E2X Transport Protocol for Main Standby Exciter Operation

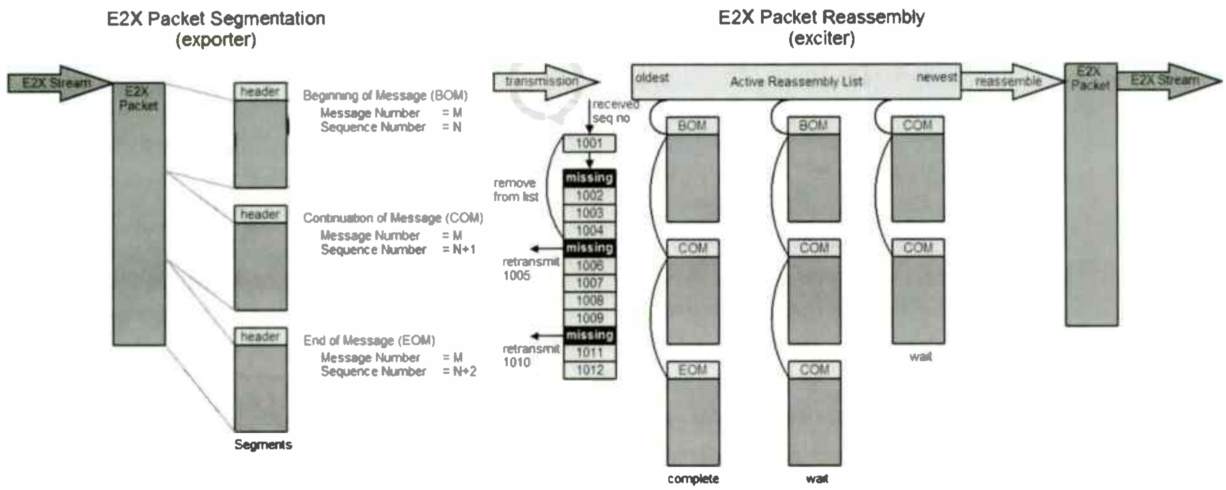


Illustration 5: E2X Packet Segmentation and Reassembly

Exporter, so the transport protocol cannot make the assumption that retransmission requests will in fact reach the Exporter. This will result in a loss of on-air quality due to increased data outages, but the E2X protocol continues to operate correctly.

Some packet types, such as the Exporter's control packet, require guaranteed delivery in order for the exciter to start modulation with correct operating parameters. Allowing exciters to start using locally stored operating parameters can cause a configuration mismatch between Exporter and exciter. Rather than relying on stored parameters, the E2X transport protocol maintains a copy of the most recent control packet and periodically retransmits the control packet to the exciter.

Using this configuration, both HD exciters are always ready to transmit HD and the digital modulator is not aware whether it is active or in hot-standby mode. The Exporter also is not aware which exciter is currently active. Both exciters independently issue retransmission requests back to the Exporter. If only a single exciter has lost a packet, it will issue a corresponding retransmission request. If both exciters lose a packet, both will issue a retransmission request but the Exporter will only retransmit the same data packet once provided the retransmission requests have been received close enough in time. Retransmitted packets are delivered to both exciters just like a regular data transmission. An exciter knows based on sequence numbers whether it has already received a packet and will discard duplicate data packets. Retransmission requests are made continuously until the data packet arrives or until the operation times out accounting for the fact that retransmission requests themselves may be lost.

This approach not only allows for redundancy in the exciters, it also allows the entire data path between Exporter and exciters to be split and made redundant. However, there is no guarantee that both digital modulators are always in the same state; IBOC sequence numbers and interleaving may be different and frame alignment may be offset. An exciter change-over normally causes an RF power cycle, so the exciter switch is not continuous in the first place. Receivers will detect the loss of digital coverage, blend back to analog and eventually re-acquire the digital signal.

E2X Segmentation and Reassembly

The main standby example demonstrates how the E2X transport protocol addresses multiple destinations and how it can support unidirectional and bidirectional links, it does not illustrate how the remaining requirements are met. The mechanics of the data transfer is shown in Illustration 5. In a real time data stream control over the PDU's payload size is crucial in order to guarantee a reasonable quality of service. Hence, the arbitrary length E2X packets are broken up into smaller segments with their own header information. Sequencing information in the header allows the segment to be related back to the original E2X packet.

The transmission path across the link may discard individual segments or reorder segments. Any segment that is received is queued in the active reassembly list in accordance with it's message number carried in the segment header. Once every segment in a reassembly list is sequential and the list consists of at least one beginning of message (BOM) and end of message (EOM) segment, the reassembly process is complete. The reassembled E2X packet is passed along to the digital modulator for processing.

In the event that the E2X packet to be segmented does not carry enough payload information to span more than one segment, the E2X packet is packaged into a single segment message, which then bypasses reassembly at the exciter and directly gets turned into an E2X packet. Clock packets fall into this category for faster processing.

Every sequence number that is received at the exciter gets entered into a sequence number list in ascending numerical order. After entering the sequence number,

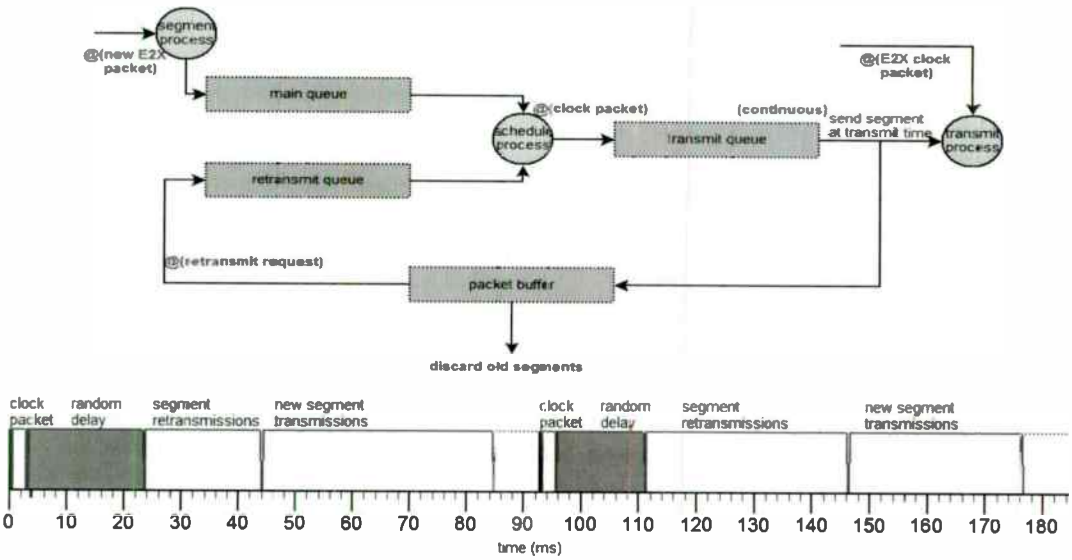


Illustration 6: Transport Protocol Scheduling at the Studio

the list is collapsed up to the next missing sequence number or to the end of the list. Now this list can be iterated across to determine what sequence numbers to retransmit. Multiple sequence numbers can be placed into a single retransmission request in order to minimize the number of retransmission requests.

While it is beneficial to decrease segment sizes in order to be able to control traffic flow more accurately, it does increase computational complexity, as well as, decrease bandwidth efficiency due to the added header structure. Larger segment sizes have a greater impact on the synchronization on another E2X stream on the same link. The segment size is configurable and not limited to Ethernet frame sizes; the underlying UDP protocol will break the segment into smaller parts automatically.

Segment Scheduling

Segmenting the E2X data stream in and of itself does not provide any benefits, since UDP does precisely perform the same function already. What complicates the situation is the fact that most STLs operate at a much lower network bandwidth compared to most LAN deployments. Consequently, all segmented packets will likely be placed on FIFO queues effectively negating the effect of segmentation unless the STL provides more sophisticated queuing. Illustration 6, however, presents a queuing structure that can effectively utilize the segments and compute packet dispatch times that match the reduced STL throughput rate. On the Exporter every new E2X packet is segmented right away and all resulting segments are placed on the main queue and each segment is treated independently from thereon.

The reception of a clock packet triggers the scheduling process that is controlled by two parameters:

1. The main bandwidth parameter sets the bandwidth required to sustain the regular E2X data stream without retransmissions. It determines how many segments may be transmitted in one clock interval and causes any instantaneous bandwidth variations to smooth out over time.
2. Overall STL bandwidth dedicated to this E2X stream dictates the rate at which segments are dispatched from the Exporter. It also governs how much additional bandwidth is available for retransmissions.

The scheduling process places segments from both the main queue and the retransmit queue on to the transmit queue as shown in Illustration 6. However, clock packets bypass the transmit queue and are directly sent to the exciter with a flag set to disabled retransmissions for clock packets. As each segment is queued on the transmit queue a dispatch time stamp is attached to it. The transmit process only dispatches segments after the given time stamp has elapsed.

The scheduler may first insert a random delay where no segments are scheduled. This minimizes the impact of one E2X stream onto another on the same STL, as it allows the two streams to interlock better while minimizing the synchronization impact. The scheduler first determines how much bandwidth is used up by new segments, but it schedules retransmissions prior to new segments, since retransmissions are more time critical.

All dispatched segments are placed on the packet buffer until they get too old to be of interest (i.e. greater than the exciter receiver buffer depth). A retransmission request from any destination searches the packet buffer for the given sequence number and places it on the retransmit queue. If two or more requests for the same sequence number from different exciters are received, only the first request is serviced, since the same same packet is delivered to both destinations.

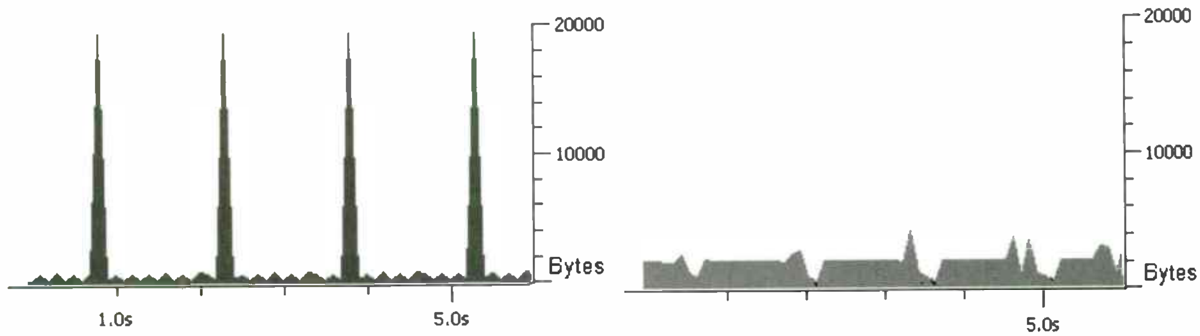


Illustration 7: E2X Instantaneous Bandwidth (left) and E2X Transport Protocol Bandwidth (right)

Illustration 7 shows a marked improvement in bandwidth utilization compared to standard E2X bandwidth utilization and demonstrates how multiple E2X streams can coexist on the same STL much more easily.

Bit and Burst Error Performance

In order to evaluate the effectiveness of this retransmission scheme, random bit errors are intentionally introduced at the segment level rather than causing discards. That way bandwidth resources are still consumed while not conveying any useful information. Illustration 8 contrasts the transport protocol's bit error performance with the E2X protocol in UDP mode based on the relationships discussed earlier. The exciter has a configurable receive buffer depth that allows the station operator to choose between increased protocol performance and total signal propagation delay. Two buffer levels of 16 packets (1.48 seconds) and 25 packets (2.32 seconds) are illustrated. Keep in mind that the E2X stream is spread almost over an entire L1 frame depending on how much bandwidth is allocated for it, reducing the effective buffer depth to a degree. This means that not every segment has the same chance of successful retransmission. With a buffer depth of 16 packets some packets may only be retransmitted once or twice; a buffer depth of 25 packets provides increased reliability due to a greater chance of successful retransmission.

The number of retransmissions depends on the throughput delay across the STL. The presented data takes this delay into account as it has been collected using a Moseley Starlink / Lanlink combination as presented in the application examples section.

A station engineer should analyze the effective bit error performance of his STL and select the desired quality of service. This identifies the operating point on this chart. If necessary the exciter's buffer depth can be adjusted. High quality RF/T1 based STLs do exist with bit error rates that are sufficient for good quality IBOC transmission. However, the STL itself is not the only source of data loss. Intermediate network equipment can contribute to packet loss due to congestion conditions or other circumstances; most network equipment is not designed to provide extremely high reliability, as in normal network situations, the end nodes provide reliable data delivery as part of their

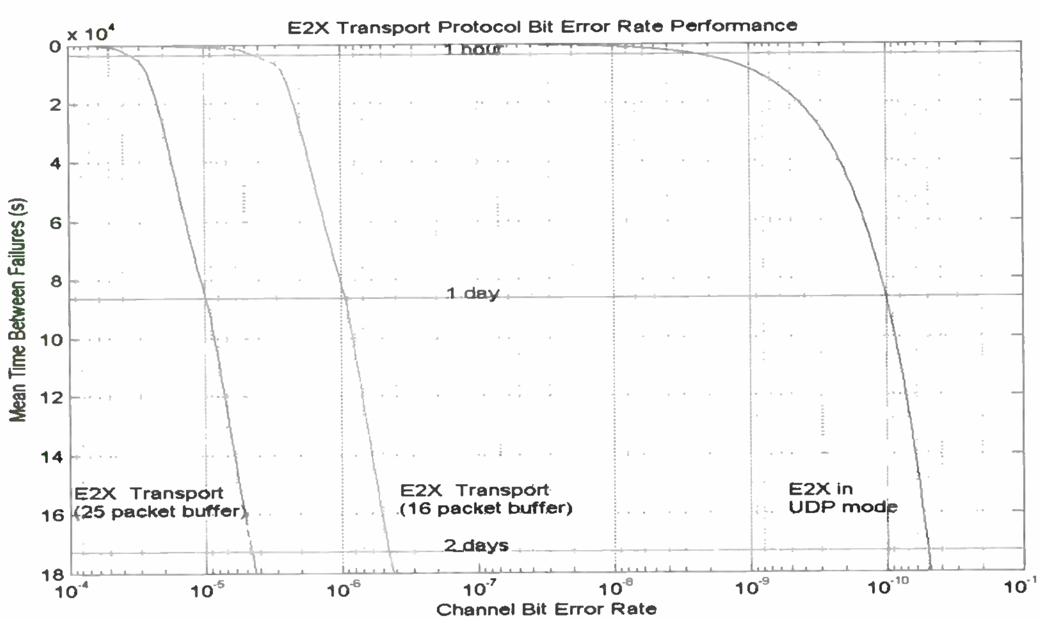


Illustration 8: Bit Error Performance

transport protocol. These symptoms will manifest themselves as packet loss which can be lumped into an effective bit error rate when averaging over a long enough time period.

Bit errors are a good way of characterizing the effectiveness of this protocol, but uniform random bit errors are almost an ideal error distribution, if they happen infrequently enough. In reality, bit errors are not uniformly distributed and usually occur due to an interference condition or fading. So it makes sense to investigate the maximum amount of time the exciter can maintain a steady data stream and replenish the data due to a complete STL interruption and associated packet loss. The following table outlines the burst tolerance of the protocol with respect to a configured receive buffer depth.

Buffer Depth (packets)	Buffer Depth	Maximum Error Burst	Max Aggressor Traffic (300 kbps)
16	1.48 s	200 ms	7.3 kB
25	2.32 s	600 ms	22.0 kB
35	3.2 s	1300 ms	47.6 kB
50	4.64 s	2100 ms	76.9 kB
75	6.96 s	3700 ms	135.5 kB

The leading reason for an STL interruption may not be the RF path at all. Instead the STL may experience a congestion condition due to other traffic, which could lead to packet loss or severe packet delay with the same consequences. The last column shows how much aggressor traffic can cause a corresponding STL interruption given an effective throughput rate of 300 kbps. The speed differential between common 10/100 Mbps LAN installations and common STL bandwidths can easily allow these amounts of traffic to be injected into the STL if no precautions are taken.

Quality of Service (QoS) Considerations

In order to effectively deal with aggressor traffic and share the available STL bandwidth with other traffic, absolute priority must be given to the E2X data stream. The IP protocol provides the Type of Service (ToS) octet to differentiate network packets and assign them different priorities. The IP ToS octet has the following definition:

0	1	2	3	4	5	6	7
precedence			delay	throughput	reliability	reserved	

The precedence field identifies whether this packet is routine traffic, immediate, or critical traffic. The delay, throughput, and reliability settings may either be set to high or low. However, these settings only have an effect if the intermediate network equipment provides the appropriate class of service queuing. Some types of STLs, such as the DMD20 satellite modem from Radyne provides a queue manager unit in its Ethernet bridge option that prioritizes traffic flow across four output queues using the IP4 ToS field to differentiate packets. Ideally more STL manufacturers would include QoS support in their STLs.

With QoS in place we now have a solid transport solution that allows the E2X stream to share the STL bandwidth with other E2X streams or even unrelated traffic, such as transmitter remote control.

APPLICATION EXAMPLES

With this information let's look at two application examples that use the E2X transport protocol.

Dual Station Topology

This topology is ideal for multiple HD stations that are co-located at the same studio and transmitter site. It demonstrates how multiple STL paths can be combined on the same network and is shown in Illustration 9. Two links with sufficient forward error correction to allow a reasonable quality of service, such as the Moseley Starlink, are chosen to carry the HD traffic to the transmitter site. These links have enough bandwidth to carry two HD streams simultaneously. Therefore, each link can serve as a backup STL to the other link. Because the E2X transport protocol is structured in such a way as to minimize impact on the other data stream, the presence or non-presence of the other data stream only affects the average clock packet throughput delay by about 4 ms; an amount that can easily be adjusted for to match the analog diversity delay. In conjunction with this protocol and a dedicated link, the StarLink provides clock packet jitter under 1 ms, this allows the exciter to synchronize to an error under 0.1 ppm. Therefore, no GPS locked synchronizer is required at the transmitter site to attain IBOC level II synchronization.

The network is broken into two distinct subnets the studio subnet and the transmitter subnet. This is necessary in order to isolate the three parallel paths between the two sites. Ethernet topology intrinsically does not like to be configured in closed loop systems as it could lead to unnecessary packet forwarding. Spanning tree protocols implemented in many Ethernet devices will intentionally sever Ethernet links if they detect multiple paths between two locations. By placing layer 3 routers between the two subnets not only are the subnets isolated, but we can also prevent any unnecessary network traffic, such as broadcast traffic, from traversing the link.

The forward links are unidirectional by nature, but the exciters for both stations can share the same bidirectional Moseley LanLink STL for retransmission requests. A host at the studio can select which path to take to the transmitter site by pointing its gateway setting to the appropriate router connected to the desired STL path. Only the Exporter points its gateway to the StarLink routers, the Importer or other

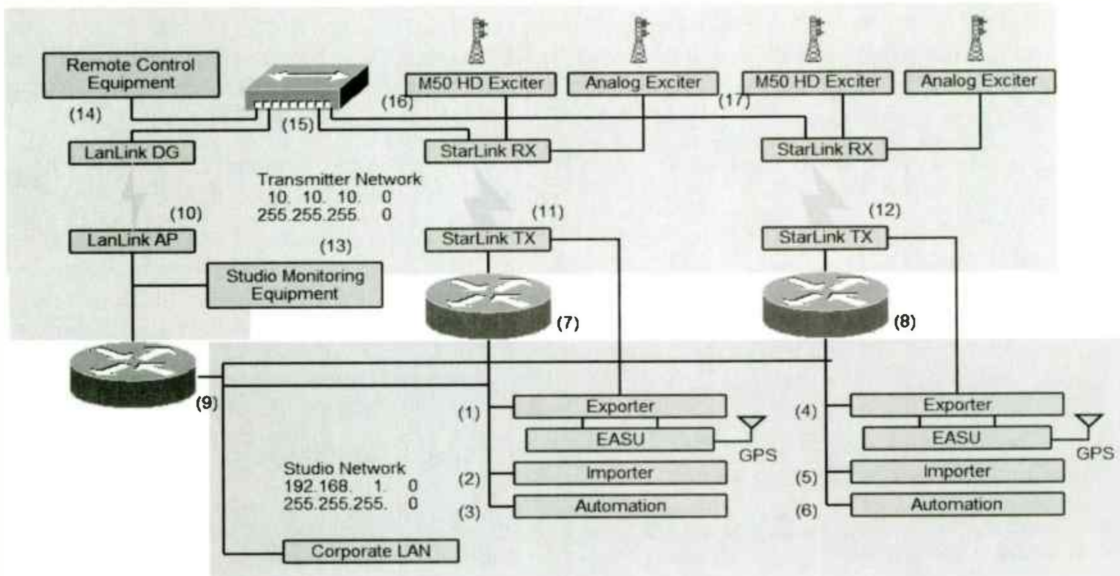


Illustration 9: Dual Station Topology

studio equipment would use the LanLink router to access the exciter or other control equipment.

The exciters, on the other hand, point to the LanLink router as their gateway in order to reach the Exporter for retransmission requests or serve control and status information.

This topology has many advantages:

- Redundant HD forward path allows the Exporter to select which link to use by simply reconfiguring its gateway. It can be demonstrated that this switch can be made quickly enough that with retransmissions enabled no on-air data loss occurs allowing STLs to be serviced and maintained.
- The reverse path can be shared by multiple stations reducing infrastructure costs.
- The LanLink STL is not on the critical HD path allowing it to be shared with other network applications without the use of quality of service queuing.
- Unlike TCP/IP, the E2X transport protocol is not dependent on the reverse path. For TCP/IP a failure (including congestion) in the reverse path would terminate the forward connection, as well. The E2X transport protocol loses the ability to retransmit data while continuing operation.
- Using directed broadcast, each station can have a standby HD exciter, provided that the studio side switch can prevent the broadcast stream from traversing the bidirectional LanLink STL back to the studio.
- Link monitoring can be included into the E2X transport protocol with support for automatic link switch over on link failure.

Satellite Distribution Network

Satellite modems are a convenient way of extending IP networks to remote locations, such as transmitter sites. Many satellite modems are bidirectional, such as the DMD20 satellite modem by Radyne. Even with the increased throughput delay, the proposed E2X transport protocol is well suited for satellite distribution.

Illustration 10 shows a proposed satellite distribution topology based on DMD20 Ethernet Bridge Option White Paper by Radyne[2]. In order to protect the satellite bandwidth, each satellite modem is connected to a router, effectively placing the studio and each transmitter site on its own subnet. This prevents broadcast or directed broadcast from being used, since broadcast traffic is only active within a given subnet. Broadcast E2X distribution could still be used by removing the routers at the transmitter sites, but using IP multicast is the preferred solution. The Ethernet standard provides for point-to-multipoint communications by assigning a special class of MAC addresses for multicast communications.

The IP multicast standard makes use of this fact and dedicates a class of IP addresses for multicast use that directly map to multicast MAC addresses. In this example, the IP address 232.0.0.1 is used, as it is a generic address that does not have to be registered with the Internet Assigned Numbers Authority (IANA). Multicast communications and multicast routing is fundamentally different from standard unicast communications, the multicast IP address does not identify the host, but rather identifies a data stream. So the assigned multicast IP address does not have to match the subnet IP address.

Different multicast routing protocols exist to pass multicast traffic from router to router. However, for our application it is sufficient to configure static multicast routes, for example "ip igmp static-group 232.0.0.1

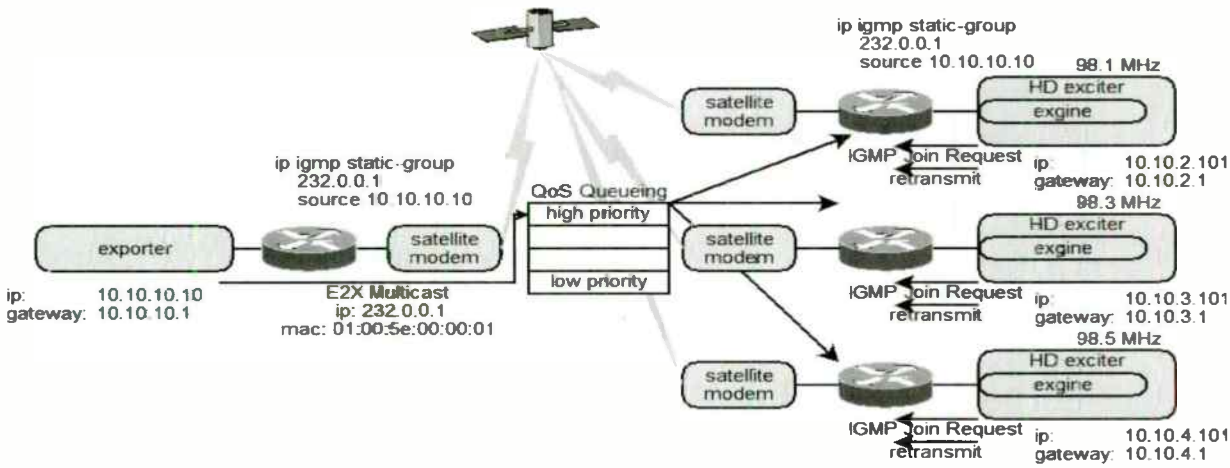


Illustration 10: Multi-frequency Satellite Distribution Network

source 10.10.10.10" configures a static route for Cisco routers.

If an exciter wants to join the E2X group, it issues a join request to its gateway using the Internet Group Management Protocol (IGMP). If the request is granted, the router will forward the multicast traffic to the exciter subnet. Only a single multicast data stream is transferred via satellite. Since the satellite modems employ quality of service queuing, we configure the E2X transport protocol to send clock packets at the highest priority and data segments at the next highest priority. All other traffic should be routed using a lower priority.

Each exciter will request data retransmissions using unicast communications back to the Exporter. So the intermediate routers must configure their appropriate routes back to the Exporter. All of these routes may also be defined statically.

In order to compensate for the increased throughput delays of satellite communications, the exciter buffer depth may need to be adjusted accordingly. To guarantee N=2 retransmissions per segment, which should be sufficient considering the high quality of satellite communications, we need to take the worst case round trip to and from the satellite into account. One way delays to a geostationary earth orbit (GEO) satellite is normally budgeted around 125 ms[3]. So for a successful retransmission, we require two trips to the satellite and two trips back to earth. Note that the initial throughput delay of the original packet is irrelevant. A single retransmission may take 500ms, adding 250ms of processing delays, leads to 750ms per retransmission. In order to ensure retransmissions across an entire L1 IBOC frame, we should add 1.48s. A buffer depth of 2.98s or 32 packets should be sufficient to ensure reliable retransmissions.

This proposed E2X transport protocol provides the reliable data transfer infrastructure to allow the deployment of multi frequency radio networks with each transmitter broadcasting the same content at different frequencies. Single frequency networks may provide an effective way of extending digital coverage in the future, but before they can become a reality, the IBOC standard must define L1 absolute frame alignment, ensure exciter synchronization, and define signal delay characteristics. It is clear, however, that single frequency networks will require a robust studio to transmitter data path, since the loss of a single data packet on a particular path could cause the digital modulators to assume different state potentially requiring the entire network to be restarted.

Single frequency networks will likely require a GPS reference at each transmitter site, but depending on the variance of the satellite throughput delays, multi frequency networks may not require a GPS reference.

CONCLUSION

The proposed E2X transport protocol provides drastic improvements over current E2X deployments, reducing bandwidth and bit error rate requirements on the studio transmitter link while allowing for Ethernet based synchronization. It provides additional flexibility in the layout of STLs by allowing multiple streams to be co-located on the same STL. By implementing a feedback path back to the studio automatic repeat requests can be used to greatly improve quality of service.

By incorporating this protocol into the radio system broadcast architecture, various multiple exciter configurations can be implemented, such as main standby exciters, multi frequency networks and single frequency networks.

REFERENCES

1 iBiquity Digital Corporation, October 2006
HD Radio™ Data Network Requirements
http://www.ibiquity.com/hd_radio/iboc_white_papers

2 Radyne ComStream, Inc, March 2004
DMD20 Ethernet Bridge Option

3 Freeman, R. L., 1997 Radio System Design for Telecommunications 2nd ed. New York: John Wiley & Sons

HD Radio™ is a trademark of iBiquity Digital Corporation, Columbia, Maryland.

NETWORKING AND STL ISSUES WHEN IMPLEMENTING MULTICASTING FOR HD RADIO

Richard Hinkle
Broadcast Electronics, Inc.
Quincy, Illinois

ABSTRACT

HD Radio provides broadcasters with a way to deliver programming digitally, but that's just the beginning. In addition to main program audio, the underlying design of HD Radio allows for a variety of additional advanced application services, most notably the ability to deliver additional channels of programming using the same signal—an application called "multicasting". Listeners with HD Radio-enabled receivers can listen to a station's main program service (which is delivered concurrently with the conventional analog FM signal) or adjust their receivers to listen to the secondary program services, commonly referred to as HD2 and HD3.

By the time an HD Radio signal reaches a listener's receiver, the complete payload is made up of multiple parts:

Main Program Service (MPS)
This is your main HD Radio signal. This will carry the same audio as analog FM.

Secondary Program Services (SPS)
These are your additional channels of audio programming, commonly called multicast channels.

Program Associated Data (PAD)
Data associated with program playing on air, including artist/title information. PAD can be specific to either the Main Program Service (MPS PAD) or to a Secondary Program Service (SPS PAD).

Additionally, the HD Radio specification allows for the delivery of other types of data including information like traffic information. In the following, I discuss the integration and delivery of MPS, SPS and PAD and the common questions and issues facing broadcasters as they prepare to "go HD".

ARCHITECTURE CONSIDERATIONS

While MPS audio can be delivered from the studio to the transmitter site by conventional means as AES serial digital audio, SPS and PAD are delivered as data-only over an Ethernet connection. This transition to the world of data as opposed to audio forces engineers to at least venture into the world normally consigned to network and IT specialists. Discourses in Information Technology are beyond the scope of this document, but you will need to know some basic terms to carry on a conversation:

Duplex
Bidirectional communication, used to describe a link between devices that allows for two-way, back-and-forth communication.

Simplex
Unidirectional communication, used to describe a link between devices that limits data flow to one-way, send-receive communication.

User Datagram Protocol (UDP)
A protocol that is part of the Internet Protocol Suite, UDP is a broadcast protocol that lacks provisions for guaranteed packet delivery or the ability to request packets to be resent. UDP communication can be bidirectional (duplex) or unidirectional (simplex). Broadcast protocol as described here means that the packets can be sent to all devices on a network, not just the one intended.

Transmission Control Protocol (TCP)
TCP is a connection-based protocol that allows for more reliable communication between devices. TCP sends data to a destination in numbered packets, and requires the destination device to acknowledge delivery of each packet. The trade-off is in overhead and speed as TCP is slightly costlier in bandwidth and slower than UDP.

Bandwidth
In this context, bandwidth refers to data transmission rates or capacity when communicating over certain media between devices. In this paper, we will differentiate between **average bandwidth**, which is the actual average bandwidth the HD Radio payload requires as measured over a 24-hour period, and **provisioned bandwidth**, which is the maximum bandwidth peak including data, data bursts and overhead.

HD RADIO FUNDAMENTALS

When HD Radio was first introduced, broadcasters were limited to Main Program Service audio at 96 Kbps, broadcasting the same programming as the FM analog signal. The HD Radio specification also included provisions for very low 400 bps worth of MPS PAD, which was primarily limited to Artist/Title information.

MP2

Extended Hybrid HD Mode of Operation: 108 Kbps available to partition between main and multicast channels.

MP3

Extended Hybrid HD Mode of Operation: 120 Kbps available to partition between main and multicast channels.

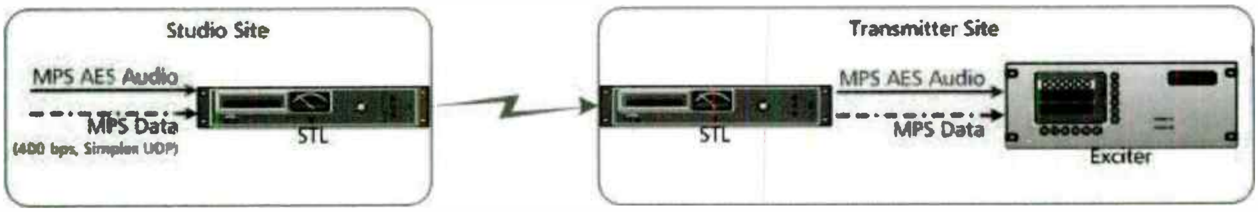


Figure 1: Basic HD STL Configuration, No Multicast

Main AES audio and this small data payload were delivered to the transmitter site where the exciter handled all the HD Radio generation. Today, almost all standard Studio to Transmitter Links (STLs) can handle this implementation.

STL requirements for basic HD Radio systems are quite simple, and practically unchanged from pre-HD Radio requirements:

AES Audio Transport Capability

- 44.1 kHz AES Audio = 1.411 Mbps (uncompressed)
- 32 kHz AES Audio = 1.024 Mbps (uncompressed)

Ethernet Data Transport Capability

- MPS PAD – 400 bps Simplex (unidirectional) UDP

Network requirements for these original systems are practically nonexistent. In most cases the studio automation data interface connects directly to the STL, but even in cases where the data interface connects to the STL using existing/shared network connections, impact on other network traffic is negligible because of the low data rate.

As HD Radio evolved and multicasting was introduced, the structure of the HD Radio signal became more flexible and more complex. Currently, there are three modes of HD Radio operation:

MP1

Standard Hybrid HD Mode of Operation: 96 Kbps available to partition between main and multicast channels.

There are two different ways to implement multicasting: **I2E** (Importer to Exporter/Exciter) and **E2X** (Exporter to Exgine), the primary difference being the location of the individual hardware components. Link and network requirements are different for each approach, with benefits and tradeoffs for each.

Multicasting I2E: HD Generation at TX Site

When multicasting was first introduced, engineers were working within a framework that placed all HD Radio generation equipment at the transmitter site. The introduction of multicasting increased the flexibility of HD Radio, but also increased the complexity.

To allow for multicasting, a piece of equipment called an Importer was added to the system which required a duplex (bidirectional) link.

Studio Site Equipment: Importer

An Importer is a Windows-based computer that:

- converts secondary program audio to a data stream
- multiplexes all secondary audio data and program associated data into a single data stream
- sends that data stream to your Exporter/HD Radio signal generator

The Importer allows the further introduction of MP2 and MP3 modes. It adds the capability to adjust the bandwidth allocation of MPS and multicast audio, a process called *bandwidth provisioning*. MPS audio can be scaled to between 48 Kbps and 96 Kbps, and up to two multicast channels can be added, between 12 Kbps and 48 Kbps each. The Importer is the point of control for the HD Radio bandwidth provisioning process.

Transmitter Site Equipment: HD Generator

After the Importer integrates multicast audio and data into a single Ethernet stream, the coded data stream is delivered to the HD Generator, such as BE's FSi 10 or an XPi 10 depending on system architecture (other manufacturers may have different products that serve a similar purpose). The data stream is sent over a duplex link using either UDP or TCP (coming in version 2.0). The HD Generator encodes the main program audio (which is delivered to the transmitter site as AES audio), and integrates it with MPS PAD and the combined multicast data stream from the Importer (which includes multicast audio and PAD). This process creates a single, properly-provisioned HD Radio signal payload.

STL Considerations for I2E

STL transport capability requirements for I2E systems are more demanding than pre-HD Radio requirements.

In systems where the main HD coding is taking place at the transmitter site, you *must* have an STL capable of a duplex Ethernet link between the Importer at studio site and the HD Generator/Exporter at the transmitter site. Today, that bidirectional link must support UDP, but with the 2.0 release of the Importer software from iBiquity this will change to TCP.

Only the MPS PAD, SPS Audio, and SPS PAD are sent via this bidirectional link, meaning that if the Ethernet connection is lost, only PAD and multicast services are lost.

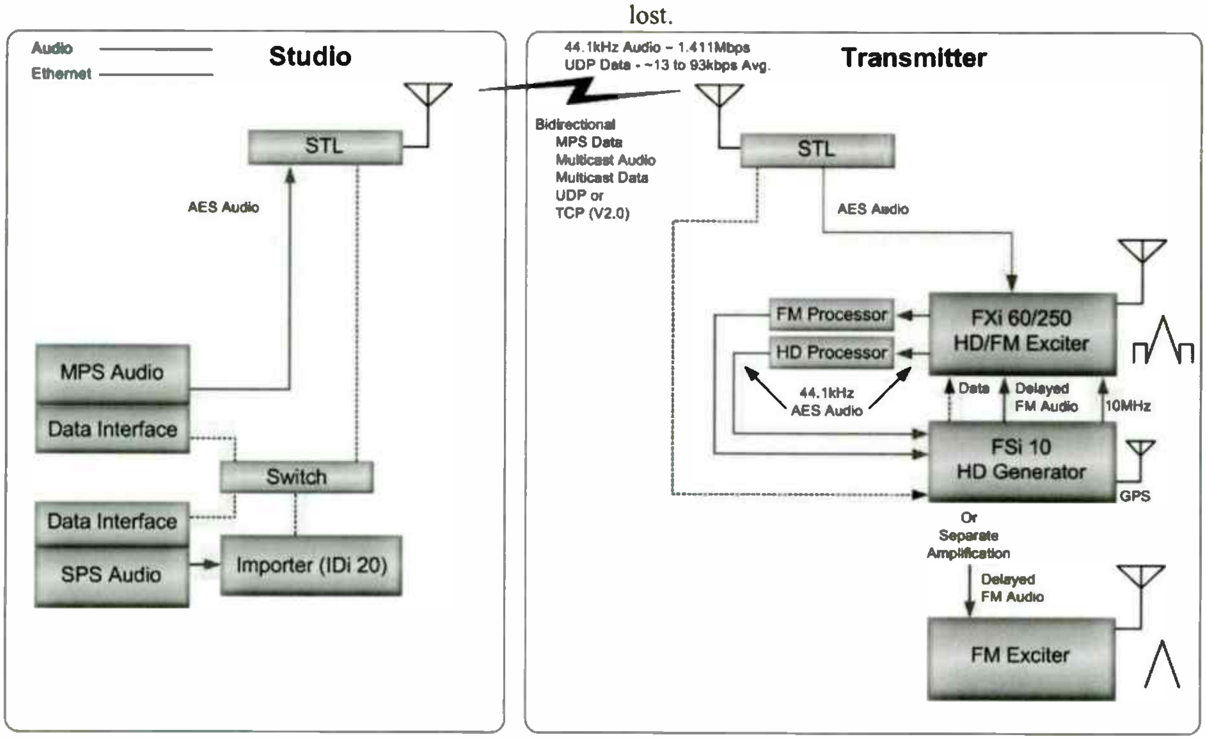


Figure 2: Multicast Architecture, HD Generation at TX Site

The HD Radio Generator also delays your conventional FM audio so that as listeners switch between analog and HD Radio the audio is synchronized.

Transmitter Site Equipment: Exciter

Also at the transmitter site is the HD Radio/FM Exciter. The exciter digitally up-converts the complete HD Radio payload to the FM band and corrects for non-linearities in the transmitter. The FXi series of exciters from Broadcast Electronics is capable of HD Radio-only, FM-only and simultaneous FM+HD Radio operation.

The Ethernet data capabilities of the STL should be capable of at least the provisioned data rate which ranges from ~17 kbps to ~156 kbps depending on the specific configuration. Keep in mind that much higher data bursts are inherent to the process.

Audio should also be considered when calculating the necessary overall bandwidth capacity of your STL. Bandwidth for AES audio can be calculated with the following simple formula:

(C x R) S = B

C = Number of channels - for stereo, this is 2

R = Resolution in bits - standard AES is 16 bits

S = Sample rate in bits per second

B = Bandwidth

Using this formula, stereo 44.1 kHz audio would require 1.4112 Mbps:

(2 x 16) 44100 = 1411200 bps = 1.4112 Mbps

Table 1 shows the data requirements of just the multicast data sent by the Importer to the transmitter site. The average bandwidth column indicates (in kbps) the average requirement to deliver the multicast data. The provisioned bandwidth column provides some guidelines that indicate the required capability to pass the multicast data without losing packets. If other traffic is being delivered over this link, that bandwidth should be taken into account as you plan your system.

Provisioned bandwidth for multicast data is calculated using the following formula:

X * (HD + Other) = P

X = Protocol overhead constant - 1.33 for UDP; 1.67 for TCP

HD = Combined average measured multicast bandwidth in kbps

Other = All other WAN traffic in kbps

P = Minimum bandwidth recommendation in kbps

Summary of STL Bandwidth Requirements I2E:
AES Audio Transport Capability:

- 44.1 kHz AES = 1.411 Mbps (uncompressed)
- 32 kHz AES = 1.024 Mbps (uncompressed)

Ethernet Data Transport Capability:

- Multiplexed PAD and SPS Audio: ~17 to ~155 kbps duplex (bidirectional) UDP
- "Other" Ethernet capability as demanded by system configuration

Data Rates and Provisioning Required for Modes and Services (Importer to Exporter or HD Generator)

Interface	Direction	IP Protocol	Service Mode	Avg. Bandwidth Kb/s	Provision Bandwidth Kb/s
Importer to Exporter (Xpi 10) or HD Generator (FSi 10)	Bi-Directional (Duplex)	UDP	MP1, SPS1 = 12Kb	13.0	17.3
			MP1, SPS1 = 32Kb	34.9	46.6
			MP1, SPS1 = 48Kb	43.5	58.0
			MP2, SPS1 = 12Kb	21.5	28.7
			MP2, SPS1 = 32Kb, SPS2 = 12Kb	57.0	76.0
			MP2, SPS1 = 48Kb, SPS2 = 12Kb	65.2	87.0
			MP3, SPS1 = 24Kb	36.5	48.6
			MP3, SPS1 = 32Kb, SPS2 = 24Kb	69.4	92.5
			MP3, SPS1 = 48Kb, SPS2 = 24Kb	77.7	103.6
		TCP	MP1, SPS1 = 12Kb	16.3	27.2
			MP1, SPS1 = 32Kb	37.6	62.7
			MP1, SPS1 = 48Kb	53.8	89.6
			MP2, SPS1 = 12Kb	29.8	49.7
			MP2, SPS1 = 32Kb, SPS2 = 12Kb	65.2	108.7
			MP2, SPS1 = 48Kb, SPS2 = 12Kb	80.6	134.2
			MP3, SPS1 = 24Kb	42.3	70.4
			MP3, SPS1 = 32Kb, SPS2 = 24Kb	78.2	130.3
			MP3, SPS1 = 48Kb, SPS2 = 24Kb	93.3	155.5

Table 1: Data Rate Requirements of Multicast Data, I2E

Data Transport Considerations for I2E

Data transport considerations are far more important with the addition of multicast services because of the increased data that is transported via the network. Encoded data is sent from the Importer in 1.48 second intervals, with each data burst referred to as a "frame".

Fig. 3 captures graphically the nature of these data bursts as captured by the Ethernet sniffer software EtherReal. Bursts are delivered every 1.48 seconds, so your network must be capable of properly spooling and transmitting the data while waiting for the next burst.

Factor in at least an additional 67% capability to allow for overhead related to TCP communication and be sure to include all anticipated traffic when considering the data requirements of the system, not just the HD Radio component. *(Provisioned bandwidth entries in Table 1 include this overhead calculation.)*

When using TCP, network latency should still be less than 100 ms, but use of TCP allows for a higher STL BER, up to 10^{-3} or 1.0%. Higher packet loss will result in multicast service drop-outs.

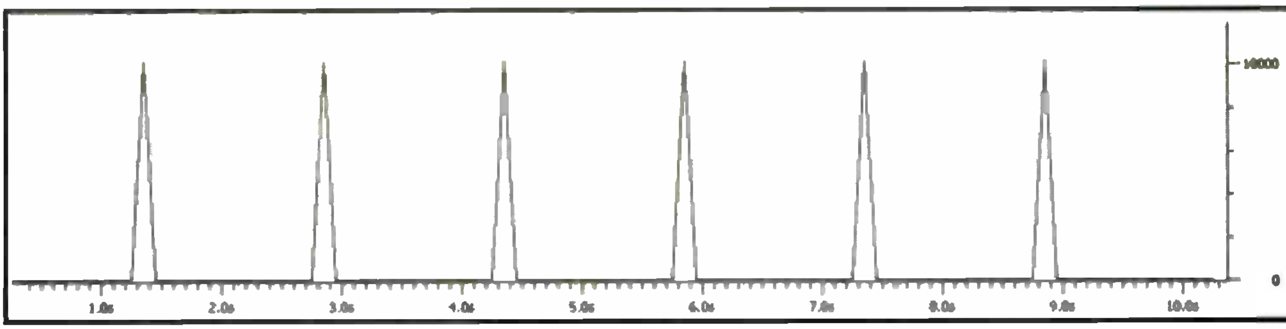


Figure 3: Importer Output Data Bursts

Currently, all traffic is sent using duplex (bidirectional) UDP. Care should be taken when designing your network. When considering the data requirements of the system, include all anticipated traffic, not just the HD Radio component. Factor in at least an additional 33% capability to allow for overhead related to UDP communication. *(Provisioned bandwidth entries in Table 1 include this overhead calculation.)*

Additionally, the Bit Error Rate of your STL should be 10^{-5} (0.01%) or better. Higher packet loss will result in multicast service drop-outs. Network link latency should be less than 100 ms. Latency can be measured using the Windows PING command.

Ping with 64 bytes of data (while running traffic) 20 times using this command line:

ping –l 64 –n 20 [address of HD generator]

With the anticipated release of version 2.0 of the Importer software, TCP will be supported instead of UDP. A duplex link will still be required, but TCP has some distinct advantages over UDP. Unlike UDP, TCP is not a broadcast protocol, which means it includes the provisions for verified packet delivery and buffering. TCP allows for higher packet loss, but requires higher bandwidth allocation.

Network Considerations for I2E

Networking is as important as the data capacity of your design. UDP is a broadcast protocol. It can easily interfere with other components on a network, and other devices on the network can interfere with your transport capabilities.

One solution is to put the transport devices on a different subnet. Installing a layer 2 switch, you can create separate virtual networks, and isolate their traffic. It is also possible to put all HD Radio components on a separate network which connects to the automation network through the use of a managed router.

As you design your network, here are some basic guidelines to follow:

- Use Static IP Addresses for Importer, Exporter, & Exciter
- HD Equipment should be on own subnet
- Subnet *must* be separate from rest of facility through use of VLAN's or physically separated networks
- Use a Switch or a Router for communication between VLAN's or network segments
- Utilize IT resources in System Planning
- Utilize shielded CAT6 cable at the transmitter site

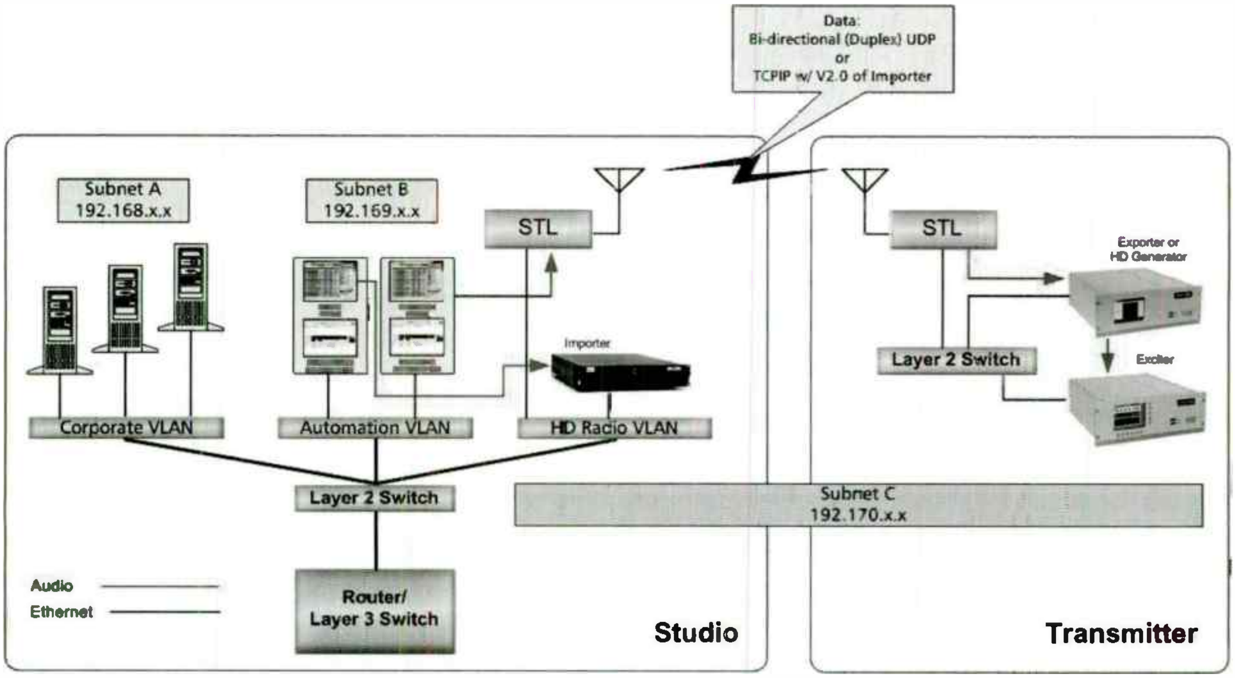


Figure 5: Networking Configuration, HD Generation at TX Site

Summary I2E

As you plan your I2E HD Radio system, evaluate your STL for data just as critically you did for audio. Remember to allow an extra 33% overhead when using UDP, and if possible prepare for TCP implementation by allowing for an extra 67% overhead. Don't forget to include "other" network traffic demands when putting together your numbers.

With the I2E configuration, audio processing, HDC audio coding and HD carrier generation remain at transmitter site. In the event of an Ethernet link failure, only PAD and multicast services will be affected. Link failure will *not* affect Main FM audio or Main HD audio in the I2E configuration.

HD equipment should be networked on a subnet or separate network, with each device assigned a static IP address. High-quality CAT6 cable and managed switches or routers should be used.

Multicasting E2X: Exporter to Exgine

E2X (Exporter to Exgine) allows multicasting in cases where bi-directional communication with the transmitter site is not an option and works well if the simplex connection to the transmitter site is exceptionally stable. Under this architecture the HD coding for multicast *and* primary audio is done at the studio site. This means that failures or deficiencies of the network connection can affect your main audio programming, not just the multicast programming as with I2E. All of the same basic capabilities of I2E (MPS, multicast, etc) are available with the E2X architecture.

In addition to all audio processing moving to the studio site, key changes in this architecture are the addition of the *Exporter* and the *Exgine card* installed in the HD exciter. The Exporter serves double-duty, not only multiplexing multicast programming into a single data stream, but also HD encoding all audio for delivery over an Ethernet connection to the transmitter site. This is the only architecture allowing you to send HD Radio data to the transmitter site over a simplex (unidirectional) UDP connection.

Data is still sent out at 1.48 second intervals, but the data rate is higher. The average simplex UDP data stream increases with E2X to approximately 120-150 kbps due to the addition of the Main HD Radio programming to the data stream. Analog audio goes to the transmitter site as before. In fact, analog audio can be transmitted by a second STL, as long as the audio reaches the transmitter site at approximately the same time.

Studio Site Equipment: Importer

The function of the Importer is identical in both I2E and E2E implementations. The Importer:

- converts secondary program audio to a data stream
- multiplexes all secondary audio data and program associated data into a single data stream
- controls the HD Radio bandwidth provisioning process

Instead of sending the multiplexed data stream directly to an HD Generator, in this configuration the data stream is delivered to an Exporter.

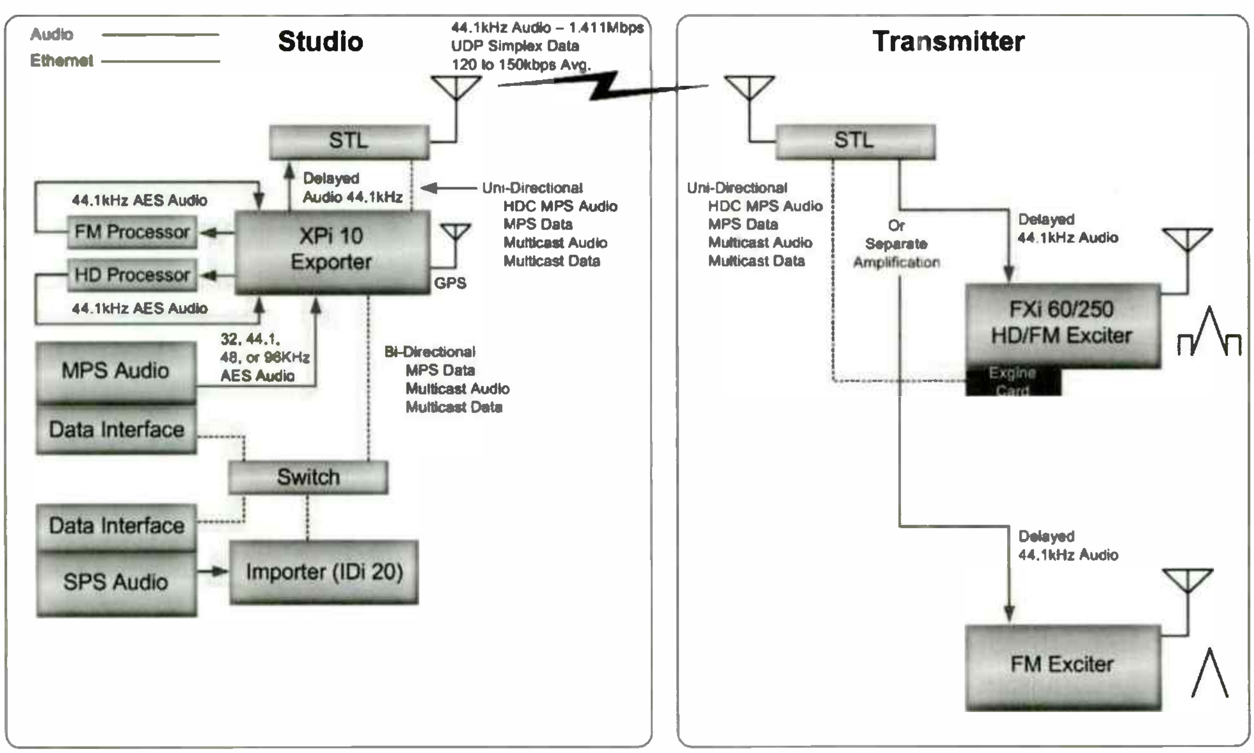


Figure 6: Multicast Architecture, HD Generation at Studio Site

Studio Site Equipment: Exporter

The Exporter codes Main program audio, delays FM analog audio, and integrates all HD Radio audio and data into a single, simplex UDP Ethernet stream. This complete HD Radio payload is then sent to the HD Radio Exciter at the transmitter site.

Transmitter Site Equipment: Exciter

Enabling simplex communication with the Exporter based at the studio site requires the simple installation of an Exgine card in the exciter. Once properly upgraded, the exciter digitally up-converts HD carriers to the FM band, corrects for non-linearities in the transmitter, and generates the Orthogonal Frequency Division Multiplex HD carriers.

No GPS signal is required at the transmitter site in this configuration—GPS timing is integrated at the studio site into the multiplexed signal that is transmitted to the Exciter.

$(C \times R)\, S = B$

C = Number of channels - for stereo, this is 2

R = Resolution in bits - standard AES is 16 bits

S = Sample rate in bits per second

B = Bandwidth

Using this formula, stereo 44.1 kHz audio would require 1.4112 Mbps:

$(2 \times 16)\, 44100 = 1411200$ bps = 1.4112 Mbps

Table 2 shows the data requirements of the HD Radio data sent by the Exporter to the transmitter site. The average bandwidth column indicates (in kbps) the average requirement to deliver the combined payload. The provisioned bandwidth column provides some guidelines that indicate the required capability to pass the data without losing packets. If other traffic is being delivered over this link, that bandwidth should be taken into account as you plan your system.

Data Rates and Provisioning Required for Modes and Services (Exporter to Exgine)

Interface	Direction	IP Protocol	Service Mode	Avg. Bandwidth Kb/s	Provision Bandwidth Kb/s
Exporter to Exgine	Bi-Directional (Duplex)	UDP	MP1	119.7	159.5
			MP2	132.1	176.1
			MP3	149.3	199.0
	Uni-Directional (Simplex)	TCP	MP1	139.3	232.0
			MP2	155.6	259.2
			MP3	167.8	279.5

Table 2: Data Rate Requirements, Exporter to TX Site

STL Considerations for E2X

When the main HD coding is taking place at the studio site, your STL must be capable of a stable simplex Ethernet link between the Exporter at studio site and the Exgine-enabled exciter at the transmitter site.

MPS audio, MPS PAD, SPS Audio, and SPS PAD are sent as a combined data stream via this unidirectional link, meaning that if the Ethernet connection is lost all HD Radio services are lost. The Ethernet data capabilities of the STL should be capable of at least the average provisioned data rate which ranges from ~160 kbps to ~280 kbps depending on the specific configuration. Keep in mind that high data bursts are inherent to the process. Audio should also be considered when calculating the necessary overall bandwidth capacity of your STL.

As with the I2E architecture bandwidth for AES audio can be calculated with the following simple formula:

Provisioned bandwidth for multicast data is calculated using the following formula as was the case in the I2E architecture:

$X * (HD + Other) = P$

X = Protocol overhead constant - 1.33 for UDP; 1.67 for TCP

HD = Combined average measured multicast bandwidth in kbps

Other = All other WAN traffic in kbps

P = Minimum bandwidth recommendation in kbps

Overall, STL transport capability requirements for E2X systems focus on Ethernet transport since in this configuration, all HD Radio services are delivered as a single data payload.

Ethernet Data Transport Capability:

- Multiplexed PAD and SPS Audio: ~160 to ~280 kbps simplex (uni-directional) UDP
- "Other" Ethernet capability as demanded by system configuration

Data Transport Considerations for E2X

The reliability and data transport capacity of the STL is far more important with the addition of multicast services because of the increased data that is transported via the network. Encoded data is sent from the Exporter in 1.48 second intervals, with each data burst referred to as a "frame".

Fig. 7 captures graphically the nature of these data frames as captured by the Ethernet sniffer software EtherReal. Data bursts still occur at the same rate, but the impact is greater due to the increased amount of information being transmitted.

Network Considerations for E2X

As with the I2E configuration, physical or virtual separation of network segments is recommended. Since data bursts in this configuration are even more extreme, it is more important to properly plan and manage your network to minimize the risk of overloading your network and risking packet loss (see Fig. 8). The same network guidelines as explained for I2E systems should be followed. HD equipment should be networked on a subnet or separate network, with each device assigned a static IP address. High-quality CAT6 cable and managed switches or routers should be used.

Summary E2X

As you plan your E2X HD Radio system, evaluate your STL for data just as critically you did for audio. Remember to allow an extra 33% overhead when using UDP, and if possible prepare for TCP implementation by allowing for an extra 67% overhead. Don't forget to include "other" network traffic demands when putting together your numbers.

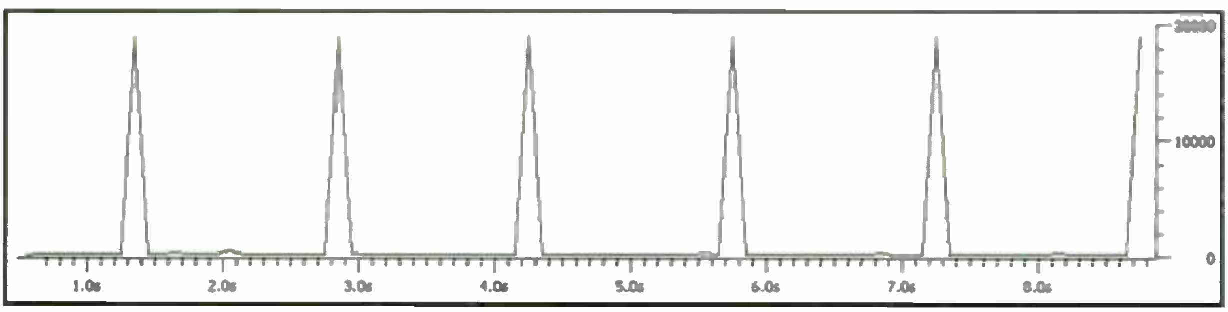


Figure 7: Exporter Output Data Bursts

Even operating at peak efficiency, the 1.544 Mbps capability of a T1 is not sufficient to transport all HD radio data *and* digital audio for the analog channel. If you are transporting 44.1 kHz audio (which requires 1.411Mbps), that leaves only approximately 137 Kbps—lower than the lowest recommended provisioned bandwidth recommendation.

Currently in the E2X configuration, all traffic is sent using simplex (unidirectional) UDP. Care should be taken when designing your network. When considering the data requirements of the system, include all anticipated traffic, not just the HD Radio component. Factor in an at least an additional 33% capability to allow for overhead related to UDP communication.

Additionally, the BER of your STL should be 10^{-6} (0.001%) or better. Higher packet loss will result in complete service drop-outs. This is a slightly higher specification than with I2E; if you have a simplex link, the reliability must be better.

With the E2X configuration, audio processing and HDC audio coding takes place at the *studio* site while HD carrier generation remains at the transmitter site. In the event of an Ethernet link failure, all HD Radio services will be affected. If only the data link fails it will *not* affect Main FM audio in the E2X configuration as it may be transported separately. If both the audio and HD services are sent over a common link and that entire link fails, it will, of course, affect both services

A word or two about point-to-multipoint requirements for large translator networks or multi-feed satellite systems: you can use one Exporter to send to multiple exciters due to the broadcast nature of UDP. At the transmitter site, the incoming UDP stream can be sent to two exciters allowing for implementation of a backup solution. It is also possible to split the outbound payload at the studio site between two different STL's to send your data to two separate transmitter sites. STL and network requirements are similar as with the more common I2E and E2X configurations, but it becomes even more necessary to isolate sub-networks as the

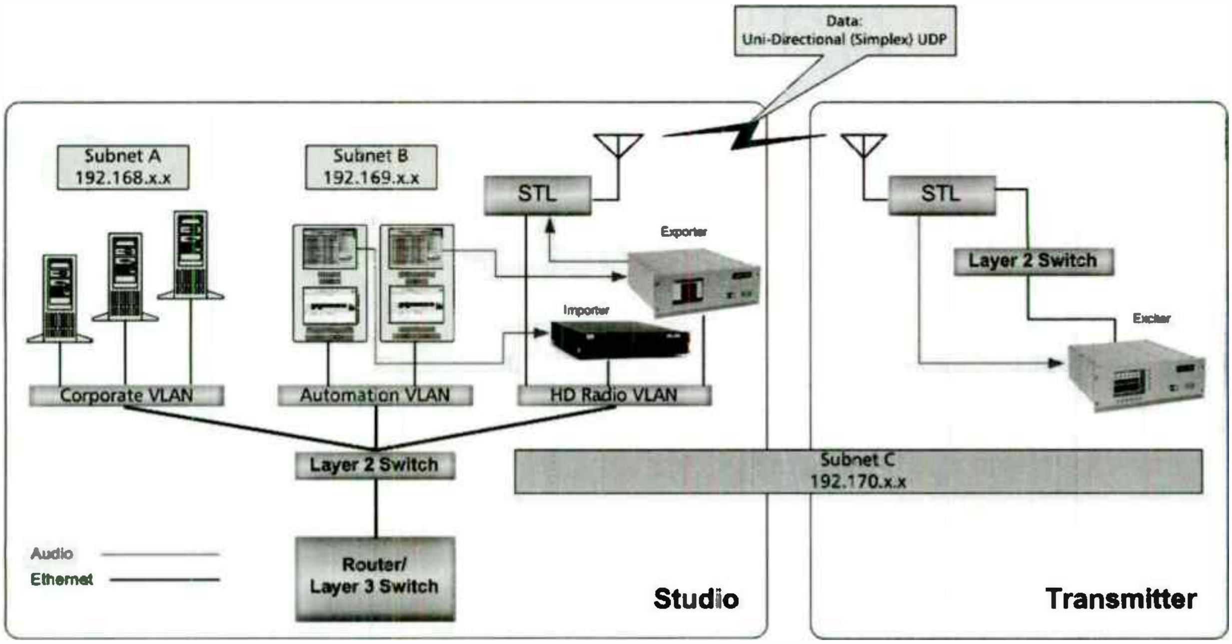


Figure 8: Multicast Network Architecture, HD Generation at Studio Side

Exporter broadcasts the UDP payload to all devices on the subnet. Due to the broadcast nature of the stream, the HD Radio payload would be identical on all receiving exciters.

When planning for these larger systems as well as the point-to-point network, it's always a good idea to consult an IT expert. HD, particularly the multicasting piece of it, adds a layer of complexity to the radio operation and transitions your stations into the world of data. This is a world markedly different than that of the audio world historically occupied by radio engineers. Knowing all the options before you will help you determine the ideal architecture (I2E or E2X) and ensure that you have built in enough bandwidth and reliability to handle the extra payload required of HD and multicasting for the foreseeable future.

Broadcast Content – Ingest to Playout

Monday, April 16, 2007
1:00 PM – 5:00 PM

Chairperson: Lewis Zager
PBS DTV Strategic Services Group, Arlington, VA

An Introduction to Networked Storage
Drew Meyer, StoreVault, a NetApp Division, Sunnyvale, CA

***Building a Broadcast Datacenter**
Ian Fletcher, OmniBus Systems, Loughborough, United Kingdom

***Media Asset Management: Simple and Cost Effective**
Michael Conway, Focus Enhancements, Campbell, CA

The Next Evolution of Statistical Multiplexing
John Bach, Scientific Atlanta, A Cisco Company, Lawrenceville, GA

Distributed Enterprise Workflow Systems from Sales to Playout
Andrew McCulloch, Harris Corporation, Colorado Springs, CO

Distributed Encoding Architectures
Kenelm Deen, TANDBERG Television, Southampton, United Kingdom

New Applications for Monitoring Your TV Broadcast – It's More Than Just Replacing Tape
Gary Learner, Volicon, Burlington, MA

---

*Paper not available at the time of publication

AN INTRODUCTION TO NETWORK STORAGE

DREW MEYER
StoreVault, a NetApp Division
Sunnyvale, CA

Introduction

This paper provides an introduction to the basics of networked storage, and its relevance to smaller business operations where storage specialists are typically not found. It discusses the origins and development of the latest advances in digital data storage technology, while illustrating the real-world advantages they bring.

Further reading and more advanced topics may be found at www.netapp.com

Historical Perspective

Mainframes

The development of networked storage has its roots in early computing models and has been enabled by technology developments of the past ten years. Early in the mainframe era, computers controlled their own disks and allocated space to applications based on priority. Disk storage was very expensive, so it was a prized computing resource and very carefully managed by the operating system and system administrators. Desktop computers were nothing more than dumb terminals, competing for the same resources on the central computer. This model became more difficult to manage as bottlenecks formed on the central computer.

Client-server computing

As computing and storage costs fell, we moved into the networked client-server model: smaller, but powerful computers called servers managed specific applications like email, database, print and file sharing; while the emergence of desktop PCs enabled local processing and local storage. These computing resources were interconnected to form a network so that the servers and the PCs (or clients) were able to communicate with each other and exchange data. This network was known as a Local Area Network (LAN) and was based on Ethernet technology. As computing and storage costs fell further; servers proliferated, PCs got more powerful and gained more storage. This decentralization of processing power and storage led to a realization that management of the storage had become significantly more difficult. As more servers were added, backups became more challenging; adding storage often meant adding another server, which just made management worse.

While there had been great advances in the availability of processing power, storage resources were poorly used, management of the various machines was time-consuming and assessing what data was valuable became nearly impossible.

Today's environment

Today's computing networks typically have very powerful client computers with significant amounts of captive storage attached to each one. Processing power is very easily available and the Internet interconnects a huge number of private networks and individual computers. Captive storage for clients, servers and dedicated applications has exploded. New regulations are affecting some industries, requiring more IT spending and planning around storage than ever before. In addition, increased dependence on computer applications for daily business operations has heightened the focus on data availability across all industries. Successfully executed backups and (more importantly) successful restores, may mean the difference between gracefully recovering from data loss and starting the countdown to filing for bankruptcy.

The Challenges of Decentralized Storage

As processing power proliferates, and dedicated captive storage follows along with each new machine, storage "islands" are gradually formed. Islands are dedicated storage arrays or volumes that are only used by a single host or a single application. One example of a common island is an Exchange server with an external SCSI array; another could be a SQL server with some dedicated RAID-protected drives. Captive locally attached storage for a database cannot be shared out to other computers nor can it be easily grown, moved or duplicated without placing a significant burden on or even causing downtime for the host.

These islands represent wasted hardware investment dollars, since many times some are underutilized while others are overflowing. Also, as these islands propagate over time they increasingly burden the systems administrators with management of the new growth. Storage islands often cause backup problems, requiring dedicated tape drives and software to manage multiple individual backup tasks – not to mention the time and burden on the host when a restore is necessary.

Other examples of storage islands can be found with even simple applications. General purpose file and print servers that are dedicated to sharing files out to individual computers require backup management, anti-virus and client software licensing and are subject to similar capacity upgrade and downtime concerns as more specific application servers.

The Benefits of Centralized storage

Application data can be combined into a single storage device or pool of devices called a SAN. SANs provide very high performance data storage that can be scaled as needed. File sharing services, usually proved by partially or wholly dedicated servers, can be consolidated into a NAS server. In the simplest cases, a single application server can share files as well as host dedicated application data, as well as run software for backup and restore tasks. But as servers grow in capacity while narrowing their specialization, the advantages of consolidating become clearer.

With increasing pressures of security, regulatory compliance and corporate governance, management of business data is becoming ever more complex and important. Technology is allowing businesses to create and store exponentially more data; the key to being successful is managing that data explosion.

What would happen if storage were combined? Consolidation means reducing the amount of time spent managing tasks, jobs, applications and data growth. Moving all types of data to a single, central, redundant and high-performing storage server could mean less time spent handling repeated storage-related tasks and worrying about backup failures.

Let us now consider some of the key technologies involved in networked storage.

Networked Storage Technologies

Direct Attached Storage

The most common form of server storage today is still direct attached storage. The disks may be internal to the server or they may be in an array that is connected directly to the server. Either way, the storage may only be accessed through that server. An application server will have its own storage; the next application server will have its own storage; and the file and print servers will each have their own storage. Backups must either be performed on each individual server with a dedicated tape drive or across the LAN to a shared tape device consuming a significant amount of bandwidth. Storage can only be added by taking down the application server, adding physical disks and rebuilding the storage array. When a server is upgraded then its data needs to be migrated to the new server.

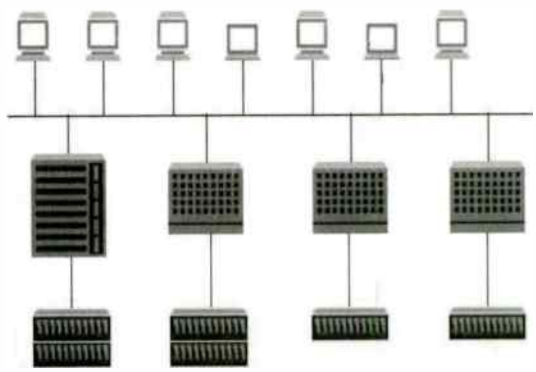


Figure 1: DAS Architecture

In small installations this setup can work well but it gets very much more difficult to manage as the number of servers increases. Backups become more challenging and because storage is not shared anywhere, storage utilization is typically very low in some servers and overflowing in others. Disk storage is physically added to a server based on the predicted needs of the application. If that application is under utilized, then capital cost has been unnecessarily tied up. If the application runs out of storage then it must be taken down and rebuilt after more disk storage has been added. In the case of file services, the typical response to a full server is to add another file server. This adds more storage, but all the clients must now be reset to point to the new network location, adding complexity to the client side. Additional cost in the form of Client Access Licenses (CAL) must also be taken into account.

Storage Area Networks

A SAN allows more than one application server to share storage. Data is stored at a block level and can therefore be accessed by an application, not directly by clients. The physical elements of the SAN (servers, switches, storage arrays etc.) are typically connected with Fibre-Channel – an interconnect technology that permits high performance resource sharing. Backups can be performed centrally and can more easily be managed to avoid interrupting the applications. The primary advantage of a SAN is its scalability and flexibility. Storage may be added without disrupting the applications and different types of storage may be added to the pool.

With the advent of storage area networks, adding storage capacity has become simplified for systems administrators, it is no longer necessary to bring down the application server. Additional storage is simply be added to the SAN and the new storage can then be configured and made immediately available to those applications that need it. Upgrading the application server is also simplified; the data can remain on the disk arrays, the new server just needs to point to the appropriate data set. Backups can be centralized, reducing workload and providing greater assurance that the backups are complete. The time taken for backups is dramatically reduced because the backup is performed over the high-speed SAN and no backup traffic ever impacts users on the LAN.

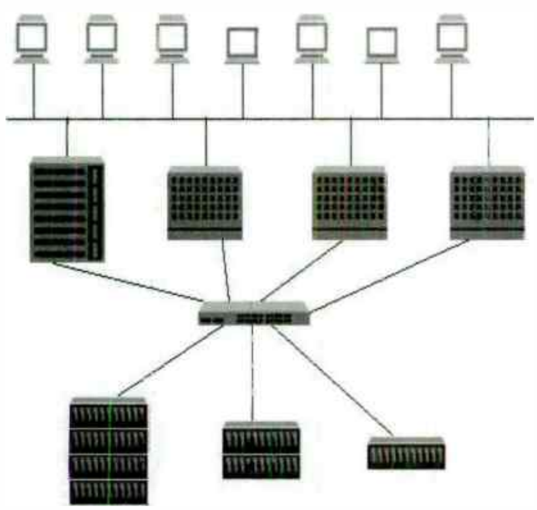


Figure 2: SAN Architecture

However, the actual implementation of a SAN can be quite daunting given the cost and complexity of Fibre-Channel infrastructure components. For this reason, SAN installations have primarily been confined to large organizations with dedicated storage management resources.
The last few years have seen the emergence of iSCSI (which means SCSI over IP or Internet Protocol) as a new interconnect for a SAN. iSCSI is a lower cost alternative to Fibre-Channel SAN infrastructure and is an ideal solution for many small and medium sized businesses. Essentially all of the same capability of FC-SAN is provided, but the interconnect is Ethernet cable and the switches are Gigabit Ethernet – the same low-cost technology you use today for your LAN. The trade-off is slightly lower performance but most businesses simply will not notice. iSCSI also provides a simple migration path for businesses to use more comprehensive data storage management technologies without the need for a "fork-lift" upgrade.

Network Attached Storage

A NAS appliance is a simplified form of file server; it is optimized for file sharing in an organization. Authorized clients can see folders and files on the NAS device just as they can on their local hard drive. NAS appliances are so called because they have all of the required software preloaded and they are easy to install and simple to use. Installation consists of rack mounting, connecting power and Ethernet, and configuring via a simple browser-based tool. Installation is typically achieved in less than half an hour.

NAS devices are frequently used to consolidate file services. To prevent the proliferation of file servers, a single NAS appliance can replace many regular file servers, simplifying management and reducing cost and workload for the systems administrator. NAS appliances are also multi-protocol, which means that they can share files among clients using Windows and UNIX-based operating systems.

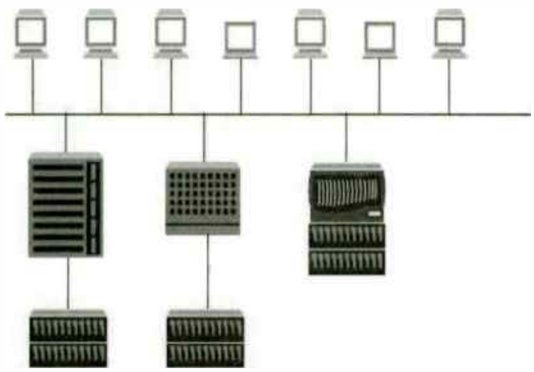


Figure 3: NAS Architecture

Administrators manage the NAS device via a browser window from anywhere they have network access and can assign shares, security settings etc, just as they would for a regular file server except there are no Client Access Licenses to manage or pay for. NAS fits right in with existing security and network management tools. As the business grows and needs more capacity, more storage can be added to the NAS device without disrupting users.

One of the most common requests to a systems administrator is to restore a single file or group of files. With traditional backup and restore processes, it is necessary to sort through the backup tapes, restore the correct directory and look for the required file. Current storage products generally offer some type of snapshot function that provides an almost instantaneous way for the systems administrator (or even an authorized user) to recover lost, deleted or corrupted files. Snapshots are point-in-time copies of the live file system, shifting tape to a more archival role and dramatically improving both backup and recovery times.

Figure 4 below illustrates some of the key differences between these technologies.

	DAS	SAN	NAS
Applications	Storage is dedicated to a single host	Provides shared storage for multiple hosts and their applications	Provides shared storage for multiple clients and their files
Adding storage	Shut down host, add drives, rebuild, reconfigure and restart applications	Add storage, configure and allocate centrally with no user disruption	Add storage, configure and allocate centrally with no user disruption
Backup	Through host or over LAN	Centralized	Centralized
Snapshots	Requires 3rd party software	Possible	Yes
Typical storage interconnect	SCSI	Fibre Channel or iSCSI	Ethernet
Multi-protocol file sharing?	Not likely	Application specific	Yes

Figure 4: Architecture Comparison

Summary

We have seen how storage networks have developed to help businesses manage their data. The evolution of SAN and NAS have led to a realization that the best features of both can be combined into a single powerful device that is capable of storing application data (as a SAN) and file data (as a NAS) at the same time; while also providing immediate backup and restore options using snapshots.

Seek an all-in-one scalable network storage solution that allows first-time NAS and SAN users to simplify not only the installation process, but also the day-to-day protection and management of their data. A storage product should provide access to the powerful data storage management tools and capabilities that have previously been only available to large IT shops. For the best in fault-tolerance and highest degree of data protection and recoverability, look for RAID 6 implementations that do not sacrifice performance and capacity. High numbers of snapshots should built-in, making backup and recovery virtually instantaneous, and giving your network users access to data with unprecedented simplicity and reliability.

For more information, visit http://www.storevault.com

Glossary of Networked Storage Terms

Backup window
The time it takes to complete a backup of the data. Often done at nights and weekends when the systems are not in use.

Ethernet
The most common protocol for communicating over Local Area Networks.

Fibre Channel
A high performance interconnect used in Storage Area Networks.

iSCSI
SCSI over IP or Internet Protocol: a lower cost alternative interconnect to Fibre Channel for Storage Area Networks.

LAN
Local Area Network: the connection between servers and clients or desktop PCs.

NAS
Network Attached Storage: a simple means of providing shared file storage.

RAID
Redundant Array of Independent Disks (originally Redundant Array of Inexpensive Disks) is a series of methodologies for providing data protection

RAID levels 0, 1, 4, 5 and 6
RAID 0: data is striped across at least 2 drives to improve performance - this is not strictly RAID as there is no redundancy and the loss of just one drive causes all data to be lost. All of the disk capacity is available for data.

RAID 1: often called mirroring, the same data is written to 2 drives at the same time – if one drive fails then the data is still available on the other drive. Only half of the total disk capacity is available for data.

RAID 4: blocks of data are written to a group of drives and parity is written to an extra drive. If any drive fails, the data can be reconstructed from the surviving drives. In a group of **n** drives, **n-1** drives of capacity are available for data as one drive is dedicated to parity.

RAID 5: Similar to RAID 4 except that the parity is distributed among all the drives. The same capacity compromise also applies.

RAID 6: Any form of RAID that can continue to execute read and write requests to all of a RAID array's virtual disks in the presence of any two concurrent disk failures. Several methods, including dual check data computations (parity and Reed Solomon), orthogonal dual parity check data and diagonal parity have been used to implement RAID Level 6.

Compound RAID
RAID groups may be combined to improve reliability or performance.
Raid 10 combines striping for performance with mirroring for redundancy, RAID 51 combines RAID 5 with mirroring to provide the capability to survive and recover from multiple drive failures. The capacity compromise reflects the combined redundancy.

SAN
Storage Area Network: a high performance network that is dedicated to sharing storage between application servers.

Snapshot
An instantaneous, point-in-time image of the file system that allows individual files, selected directories or an entire file system to be immediately captured or recovered.

The Next Evolution of Statistical Multiplexing

JOHN BACH
Scientific Atlanta, A Cisco Company

INTRODUCTION

Video distribution systems have taken advantage of statistical multiplexing for many years. In fact, statistical multiplexing, combined with efficient encoders, has become the standard for efficient video distribution via satellite, cable or terrestrial links.

Contemporary video distribution systems typically use one of two primary architectures to accomplish statistical multiplexing. The first of these, closed-loop architecture, relies on tight and direct control of video encoders to optimize the performance of the overall system across all channels. The second, open-loop architecture, allows non-collocated encoders and depends on transrating, sometimes called grooming, or rate shaping, to change the bit rate of the video streams. For the last decade statistical multiplexing technology has been deployed with MPEG-2 encoders and video streams, and during the last two years AVC based encoding technology has been used together with a closed-loop system. AVC transrating technology is not ready for field deployment, and AVC-based open-loop systems are not yet feasible.

Now, high speed IP-based networks make it possible to combine the benefits of closed-loop statistical multiplexing with the open-loop advantage of being able to distribute encoders. This paper discusses measurement results from open- and closed-loop systems and considerations for an architecture that combines non-collocated encoders in a statistical multiplexing architecture while maintaining the efficiency of a tightly coupled closed-loop statistical multiplexing architecture.

THE BASICS OF STATISTICAL MULTIPLEXING

A statistical multiplexer (stat-mux) uses statistical knowledge about the users and the system to allocate bandwidth. Unlike fixed multiplexing systems, in which each channels is assigned a constant bit rate, a statistical multiplexing system uses data gathered from all encoders within the system to allocate bandwidth.

Statistical multiplexing dynamically distributes available channel bandwidth among video programs to maximize the overall picture quality of the system. A joint rate-control algorithm guides individual encoders based on continuous monitoring of the scene content of each video source.

Statistical multiplexing gives the broadcaster the ability to allocate bandwidth on an as-needed basis. For example, if there is a news program on one channel and a basketball game on another, a statistical multiplexer can analyze the needs of the two channels and allocate more bandwidth to the game, where there are crowd shots and fast movement, and less to the relatively static news program. Statistical multiplexing also allows each individual channel to peak the bit rate when needed without compromising quality of service. The net result is more efficient bandwidth use and a good quality picture for each channel.

CLOSED-LOOP STATISTICAL MULTIPLEXING

A typical closed-loop system consists of some number of encoders in a single location interconnected to a multiplexer in the same location aggregating the different streams to a multiplex with a constant payload. Additionally, a statistical multiplex controller is needed to manage the bit rate of each encoder. The controller may be integrated into the multiplexer, or in one of the encoders, or in a separate controller. The controller looks at information about video bit rates and complexity of the information across all encoded content and compares the bits against each other. When the controller sees content that can be broadcast with fewer bits without compromising video quality, it reallocates the extra bits to content that needs more bandwidth. The controller then provides the information to the encoders, where it is encoded based on this new information.

The closed loop system requires instantaneous direct feedback, which creates two significant issues for the broadcaster. First, there must be a communications path from the controller to the encoder. In practice, this means that the encoding and controller/multiplexing equipment need to be at the same location. The other issue is that the encoders' output must be multiplexed (aggregated) without introducing any different delays in the path from the encoder to the multiplexer. Otherwise, the multiplexed stream will not sum up to a constant bit rate that can be transported within the available channel capacity.

OPEN LOOP STATISTICAL MULTIPLEXING

More recent versions of statistical multiplexers can shift the demand peaks so that they do not align, exploiting the typically short duration of demand peaks. These advanced multiplexers maintain a buffer model of the decoder receiving the stream for each program being statistically multiplexed. Since the decoder has one or more pictures (frames) in its buffer ready to present to the viewer, the multiplexer can allocate the output bandwidth on a priority basis to those programs for which the decoders have the least amount of video in their buffers. They also offer re-quantization. In an open-loop architecture, transrating typically performs re-quantization of individual MPEG-2 encoded programs. Re-quantization works by reducing complexity within each frame, removing some of the high-frequency coefficients. Quantizing is one of the last encoding steps from uncompressed video to an MPEG-2 stream. Open loop transrating reverses the quantization process, decides on a compression level and performs a new quantization.

A side benefit of this capability is that this process eliminates the need for a feedback loop, paving the way for an open-loop architecture in which remotely encoded material can be statistically multiplexed. This means that, in open-loop systems, material may originate from another studio in the same town, across the country, or even across the ocean via an international satellite link.

Re-encoding and Transrating in Systems with Non-collocated Encoders

Non-collocated encoders can be supported via either re-encoding or transrating. When re-encoding is used, the open-loop encoder is usually a single task, analog on one side, or baseband, and MPEG-2 on the other side. Open-loop encoding is a one-pass type device that converts content into MPEG-2. Often, the type of content dictates open- or closed-loop encoding, with multiplexing split across multiple channels and multiple channels in a single multiplex.

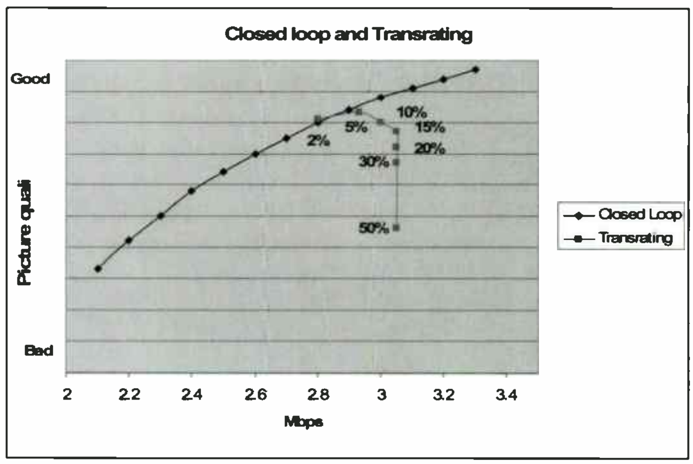
Transrating, actually just one member of a family of transcoding processes, is the process of changing the bit rate on an MPEG-2 stream of one or more programs to a different, and generally lower, bit rate. In the broader transcoding arena, content is converted from one format to another. This may be conversion between PAL/NTSC and SDI or between PAL and NTSC, for example. When referring to compressed MPEG-2 streams, it could be performed between 4:2:2 and 4:2:0 formats, between I-frame-only and Long-GOP streams or between audio format such as Dolby Digital and MPEG LayerII.

When an operator wants to re-multiplex programs from different statistical multiplexes, there is no way to know the total bit rates of the different programs in the new combinations. Sometimes the total transponder bit rate can be lower than the channel capacity and sometimes it can be much higher. Transrating allows an operator to reduce the bit rate of the different streams, sometimes by a significant amount. Transrating can also be used in the headend to lower the bit rate of a full transponder or of certain programs with the decoding or re-encoding and therefore, without the need for the multiplexer and the encoder to be collocated.

While this process reduces the video bit rate, it can lead to progressively decreasing quality. The I-and P-frames within the Group of Pictures (GOP) degrade during re-quantization, resulting in progressive video quality degradation through each sequence since the P- and B-frames have the "wrong" reference frames. The result, called "breathing," shows in the video as complex objects pulsing with MPEG-induced noise. For acceptable video quality, the transrating percentage generally needs to be 20 percent or less.

COMPARISON BETWEEN OPEN AND CLOSED LOOP SYSTEMS

The following experiment compares closed-loop and open-loop statistical multiplexing: A traditional closed-loop system with six channels fed with uncompressed video sources is compared to an open-loop system fed with the same six services from identical encoders in Variable Bit Rate (VBR) mode. The transrater is configured for a constant output bit rate of 3Mpbs pr channel (average bit rate). By changing the quality target on the VBR encoders, bit rate reductions between 2% and 50% are performed by the transrater. Output picture quality is measured across all channels with a Digital Video Quality (DVQ) analyzer.



The blue curve represents the picture quality of the closed-loop system, where the quality steadily improves as the average bit rate is increased from 2.1 to 3.3 Mbps. The pink curve is the quality measured on the transrated content where the average bit rate reductions required changes from 2% to 50%.

At the 2% point, the average bit rate of the VBR encoders are below the 3Mbps target and there is only an occasional need for transrating. The quality is close to that of the closed-loop system at a reduced bit rate, and there are ~7% null packets in the stream which could have been used for video quality improvements. As the VBR encoders' bit rate are increased the quality improves as null packets are replaced with video, and the needed transrating increases. The transrater provides comparable quality to the closed loop system at correspong reduced bit rate. When the VBR encoders' bit rates are further increased and the need for transrating increases, the quality falls off despite the fact that the number of null packets is minimized. At an average reduction rate of 50%, the transrater has severe problems and its performance is far below that of the closed-loop system.

Key take-aways from this experiment:

If the VBR encoders can be trimmed to the content and available bit rate, then tansrating represents a minor loss (5% to 8%) as compared to closed-loop encoding. However, in practice it would be very difficult to find the right spot, as the quality target in the encoders depends on the other video source and hence must change over time. In practice, an open-loop system will be 15% to 20% less efficient than a closed-loop system, and only works well at moderate bit rate reductions.

Open- and closed-loop systems each have advantages and disadvantages. For closed loop systems, the biggest advantage is consistently acceptable video quality with maximum efficiency of bandwidth utilization. Their principle disadvantage is inflexibility due to the need for collocated encoders and statistical multiplexers and the consequent equipment expense.

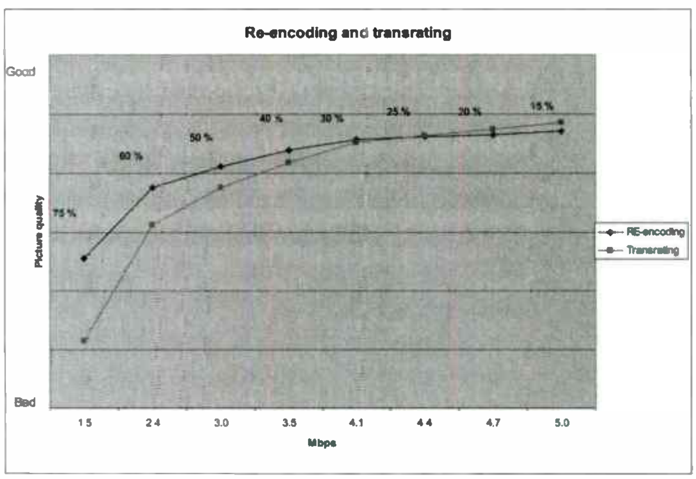
COMPARISON BETWEEN RE-ENCODING AND TRANSRATING

In open-loop systems, the principle advantage is the flexibility resulting from the elimination of the need for re-encoding. This means that encoders can be located remotely from the statistical multiplexer. The downside is that there is more opportunity for video degradation that can produce unacceptable video quality.

Traditionally, non-collocated encoders can be supported via either re-encoding or transrating. When re-encoding is used, the content needs to be decoded back to baseband before a closed loop encoding system is used to re-encode the content optimized for the final distribution system. Alternatively, multiplexing content from non-collocated encoders can be accomplished with transrating, avoiding the costly decode and re-encode stages.

The comparison in the experiment focused on transrating as compared to closed-loop encoding. In a non-collocated environment, a single closed-loop encoding stage has traditionally not been an option. Hence we really need to compare encoding followed with a transrating stage with two encoding stages.

In the figure below, a concatenated encoding system is compared to a transrating solution for an MPEG-2 based system.



The pink line represents the transrated results of a 6Mbps Constant Bit Rate (CBR) encoded stream across four channels. The blue line is the results obtained from re-encoding. The re-encoding is done while maintaining picture resolution and the video content is passed digitally from decoder to encoder without any filtering applied. GOP is maintained the same, and the encoder used has a special auto-concatenation feature that ensures that each video frame is encoded using the same MPEG-2 picture type as the first encoder used.

The numbers along the graph represent the rate reduction in the final compressing stage.

Key take-aways from this experiment:

For small to medium rate reductions, transrating and re-encoding are very comparable. In these situations, transrating is a good choice, as it is a simpler and cheaper solution than re-encoding. As the bit-rate reduction increases, the transrating solution suffers, and re-encoding is clearly better. For the typical bit rates used for statistical multiplexing, 2.5-3.0 Mbps, re-encoding is clearly the better choice.

IP-BASED TRANSPORT OF VIDEO AND MULTIPLEX CONTROL DATA

Before discussing IP networking issues, first let me repeat a couple of conclusion from the previous sections. As expected closed-loop encoding performs better than open-loop. Even if the open-loop system is optimized for a given set of sources, it still performs less efficiently than a closed-loop system. For non-collocated encoders, transrating is a realistic alternative to re-encoding, provided that the reduction in bit rate is on the smaller side. The measurement results show earlier in this paper are based on MPEG-2 video encoding technology, but the basic characteristics are fairly similar to AVC encoding, hence the conclusion would apply equally well for AVC based systems. Additionally, as AVC transrating technology is not generally available at the moment, closed-loop encoding would seem to be the best solution.

Based on that, it is fairly obvious that if we face an application where content is generated in multiple different locations, supporting closed-loop encoding across the multiple locations would be the best architecture. And, as IP networks have evolved to provide high bandwidth in a cost-efficient manner and at the same time provide both video and data transport, it is an obvious choice as a transport medium for both control and video transport for a non-collocated closed-loop encoding system.

Using IP interfaces for a collocated closed-loop encoding system is not an issue, as both control information and video can be transported over IP networks in a reliable and timely manner. IP transport of video adds overhead and jitter, but by using a dedicated IP networ, this can be maintained at low levels and will not represent a problem for the MPEG multiplexer to deal with in creating the outgoing transport stream.

The real challenge starts when the encoder is moved away from the multiplexer/controller and then interconnected via an IP network. First of all, the IP network introduces delay and jitter that are not present in the small dedicated network of a collocated implementation. Typical closed-loop systems rely on timely and accurate interchange of data, and in a properly optimized system, data is exchanged for every frame or even more frequently. Reducing that update time limits the system's capabilities to react rapidly to changes in encoded video. Significant delays in the video links from encoders to the multiplexer and the control path between encoder and controller, need to be accounted for in the algorithm. Network jitter, especially, makes it very difficult to predict the timing and optimize the algorithm.

Introducing jitter and delay in a closed-loop system, will lead to quality reductions. Eventually it may be better to transrate than used closed-loop encoding if delays become too great.

IP network categories

The term "IP networks" covers a very wide range of networks. In the scope of this document, I will divide them into three different kinds and look at them from a closed-loop perspective.

Wide general IP networks (WAN's)

These are general-purpose IP networks used for multiple purposes including such applications as corporate LAN interconnect and voice. A network like that will typically include very significant delay and jitter, depending on factors such as load and time of day, and will likely grow worse in the future as more and more data are loaded onto the network. The ultimate example of this is the Internet. Trying to use a network like this, in which there is no control of traffic, for a closed-loop encoding system, is not recommended. Even through it may work initially, delay and load will change over time. It will be very difficult or even impossible to ensure the predictable delay to allow for statistical multiplexing.

Closed dedicated network

If a dedicated network is allocated for the closed-loop system, where control data and video is passed on dedicated bandwidth using dedicated routers, delays and jitter can be minimized. In this situation, non-collocated closed-loop systems can be built using only a few routers, and can maintain picture quality comparable to a collocated system.

The above scenario may seem somewhat idealistic, but it is often achievable using existing SONET or PDH infrastructure to build a dedicated IP network that will meet the requirements for low jitter and delay. Such a network should accommodate a closed-loop encoding system.

Closed IP network with mixed data traffic

In the presences of a closed IP network for video and control data, it is often very attractive to use the same network for other data transport as well. This data might include management traffic for the equipment involved in the closed-loop system itself plus other equipment monitored at each location and other data services that are part of daily workflow.

Additional data traffic creates increased delay and jitter through the routers, as the time- critical traffic needs to wait for IP packets that have arrived previously. The problem escalates as the network is loaded with more traffic. The increased delay and jitter will cause problems for the closed-loop system if not controlled. With Differentiated Services Code Point (DSCP), video traffic and control data can be given priority over other lower- priority traffic, and delays and jitter can be minimized. In fact, in modern routers, the latency through the router for the highest-priority packet is determined by the delay it takes to transmit the worst-case previous packet. So, provided the network is not loaded with extremely large IP packets and the number of routers is limited, delays can be limited, facilitating closed loop encoding with minimum quality losses.

DISTRIBUTED ENTERPRISE WORKFLOW SYSTEMS FROM SALES TO PLAYOUT

ANDREW MCCULLOCH, CHUCK KOCSIS
Harris Corporation
Denver, Colorado

ABSTRACT

The world of broadcast is evolving as never before, from high level organizational consolidations right down to changes in delivery from linear to non-linear. The model of disconnected systems interfaced through custom interfaces no longer satisfies the demand for near real-time interactions between business functions and the need to control and monitor workflow throughout the delivery chain. Workflows are no longer constrained within single systems, increasingly they must bring systems together in a way that breaks down the traditional boundaries and adapt to changing business models in a cost effective and timely manner. Mandatory regulations such as Sarbanes Oxley are requiring broadcasters to monitor changes across all systems, particularly content and its associated rights. The ability to provide this type of information quickly and easily is critical to successfully managing a broadcast operation.

AN EVOLVING ENVIRONMENT

With advances in technology and improvements in connectivity between systems, the increased demand for content has changed the way broadcasters think , manage and distributed content as well as how to control the associated rights.

Broadcasters are now selling outside the traditional: 30 spot, finding new advertising revenue streams and delivering content through traditional and non-traditional means such as video on demand, podcast, cell phones etc. There are numerous systems on the market that manage and consume content and any organization that delivers and consumes content is using one of those systems today for their sales, scheduling, electronic program guides, play-out or rights management.

The way we are today

Getting these systems to integrate, share information and remain synchronized is not always simple at the best of times. Having these systems rapidly adapt to the changing environment is all together a different dilemma.

Disconnected Systems

If you look within broadcast organizations today you will see many differing systems, both new and old, all performing one or more targeted functions. One common function they will all have is they are required to integrate with each other. These integrations are almost always done through interfaces, be these binary or text files, database, and most recently XML and web services.

Looking closer you will see these system integrations are frequently proprietary, custom and point-to-point. As new integrations are required the vendors are gathered together usually with a 3rd party integrator/negotiator, to get everyone to agree to how this integration will take place.

Vendors inherently do not really want to play together with common interfaces unless industry standards or dominant players force a standard. There are internal needs and drivers that will usually result in a complex integration with one or more parties down the line conceding and implementing some other proprietary format. This invariably focuses everyone on just the problem at hand and does little to address the overarching integration and workflows across the whole broadcast operation, perpetuating the lack of visibility of the big picture.

The introduction of new workflows and functionality to any individual system in itself is not difficult, but getting all systems to share common approaches and manage cross system workflows is a more daunting task, given the desire to utilize existing proprietary interfaces.

Today's broadcast integrations are largely batch in nature, with a limited concept of real time interactions or common exchange formats with multiple point-to-point connections as shown in figure 1.

The triggering of an integration in many cases requires a user to select a function, press a few buttons and a flat file is produced in a given location. This file is then copied to another location and manually imported by the receiving system. It's then the receiving systems responsibility to manage problems encountered as well as handle situations where long gaps occur between delivery and final acknowledgement and processing.

Recent mandatory rules introduced with Sarbanes Oxley require much greater accountability of who made changes where and when. This information, including interactions between systems must be accounted for and reported. Systems today log this information locally which makes following the trail between systems difficult. Where systems do not share common user accounts it is especially difficult to monitor at the user level. It is commonly left to the broadcaster to sift through the myriad of audits to provide meaningful information.

Adding further systems into the integration process and providing them the same data usually requires the user to select another function and/or produce a file in yet another format for that system to consume.

Traditional System Boundaries

Many systems today have some sort of workflow available for the processing of a series of tasks in a given or variable order. These workflows are in nearly all cases constrained within the boundaries of that system, be that by existing data types and structures or physically within the applications themselves.

Broadcast operations are changing, content is all-pervasive and workflows must span multiple systems incorporating both human and system tasks. Workflows can no longer be constrained by the boundaries of any single system, it is no longer inconceivable that in some instances traffic could control automation, or that any single system is responsible for all content and the associated metadata.

Cross system workflows today differ considerably, involve many manual tasks, passing around of spreadsheets and often require data re-entry into systems. Much of this can and should be automated to free valuable resources to be more productive, but requires us to get away from the custom point-to-point interfacing we have today.

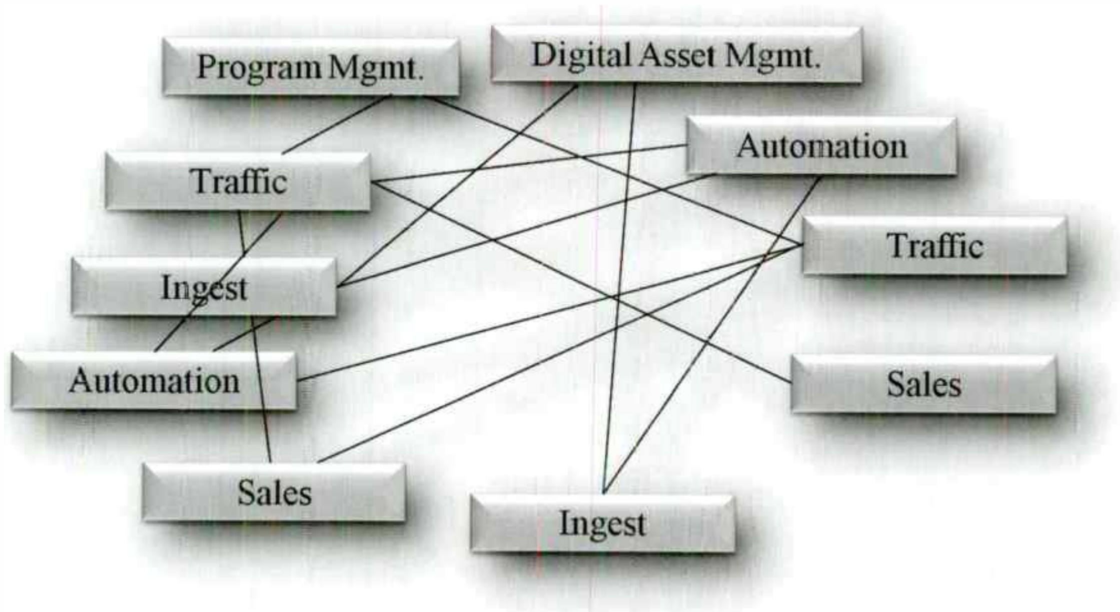


Figure 1 A maze of connections

Where is my Content?

New content and metadata can be created in multiple locations meeting the needs of those systems. Automated Ingest processes and digital asset management systems can now process speech to text, apply sophisticated shape recognition techniques and incorporate geo-spatial information at the frame level.

All of this metadata is not stored in every system, nor does it need to be as data needs differ for each system. However, it is useful to know that it exists and where. Any system should have the capability to know everything about the metadata that exists within an organization and decide if it is required to be stored locally within that system.

We continue to see consolidations within the industry and this inevitably leads to more integrations. In a multi-national organization it is not uncommon to have different Sales and scheduling systems for domestic and international markets, multiple automation systems and multiple Digital Asset Management systems. As shown in figure 1 this rather simplified view shows the maze of different connections that exist between these systems making effective and efficient control difficult and content synchronization far more problematic. This can become much more difficult in more complex scenarios.

Common questions to ask are:

- What content do I have?
- Where is this content and which system is the master owner for each element?
- Do I have unnecessary duplicates?
- Is this content needed or used and can I archive or delete it?

There are few vendors today who can offer a complete end-to-end solution for broadcasters and for historic, merger and other reasons the resulting solutions are typically from multiple vendors.

Systems typically perform specific functions and when combined they form an end-to-end solution. Within this there is increasingly more overlap between them in the area of Content which is central to everything within broadcast, from Sales to play-out. Content is used globally; Rights are established domestically and globally for linear and non-linear. Digital content can be distributed throughout the organization and utilized by multiple systems. With all of this, systems today still think and operate locally. Sales, traffic, digital asset, library management and automations systems have evolved independently and have content overlap which presents us with significant hurdles for building efficient workflows at an enterprise level.

THE WAY FORWARD

The control of content and workflows throughout an organization can no longer occur at the local system level. Systems must act more as services to be plugged together into workflows combining both human and system tasks. Every vendor will claim they have embraced service orientated architecture (SOA) and make available key aspects of their systems as services to facilitate effective integrations.

SOA is an enabling technology that facilitates better connectivity, but its effectiveness is diminished if these services still require proprietary data formats specific to the way each system sees the data. This understanding and interpretation of data is the difficult aspect of any integration and just having services does not change this.

To efficiently manage content at the enterprise level requires we address what causes many of the problems we see today with integrations. We must (a) be able to have better accountability of the flow of information and track changes for all data exchanges at the user level; (b) we must know of the existence and location of all the logical metadata and physical essence; (c) we must be able to share this information readily with any other system, (d) we must know that all the systems are synchronized and share a common identification scheme and (e) we must know which system is the master or owner for specific metadata elements for conflict resolution.

Common content repository

Content metadata can originate from and be changed by many systems, each having different usage requirements at specific points in time. The logical metadata and physical essence are distributed throughout the enterprise stored in Sales, Traffic, DAM, Automation, video servers, libraries etc. This information is utilized locally, be it for linear and non-linear scheduling, editing, dubbing or copying to make available for play-out. Today, as in figure 1, systems are mostly connected point-to-point with custom information exchange formats and means of uniquely identifying content. This makes it extremely difficult to process the information between the systems in a consistent manner, including managing and automating workflows.

In an ideal world all content metadata would be stored in a single common repository and all other systems would reference it for all of their data needs, storing nothing locally. However, because of the distributed nature of many of these systems, the need to store data locally for performance and the practical constraints of exchanging information between business and operational networks it is unlikely that this would ever become a reality. Therefore a more practical solution is required.

This lies in part with a common synchronized repository for metadata that accommodates local system storage, and in part by exchanging commonly understood information through a central messaging/platform. The platform must be able to audit all the information flows between systems to provide cross system accountability at the user level.

Rather than exchanging messages point-to-point, systems need to be able to exchange information through a common messaging platform and have this information distributed to the appropriate destinations via publish and subscribe rather than each system managing connections to every other system (figure 2). This approach allows for additional systems to be added over time with little or no impact to existing systems, unlike most environments today.

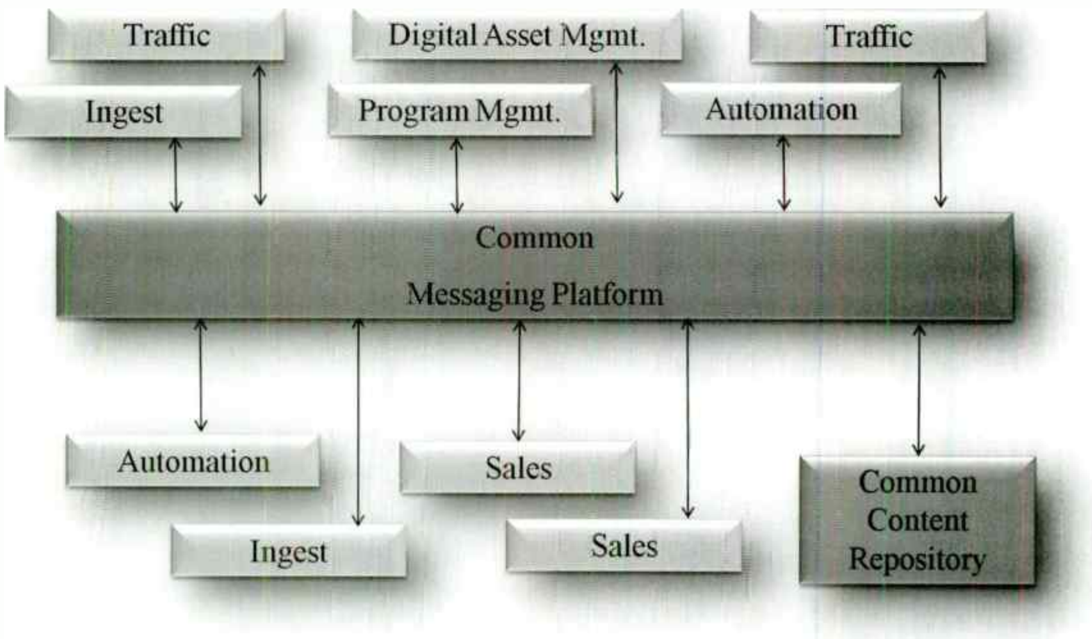


Figure 2 Common Messaging and content repository

Common metadata standards

For a common repository to be successful as well as facilitate enterprise wide workflows, systems must move away from the proprietary formats for data exchange and embrace more open standards. There have been several attempts in the past to introduce standards; some have been relatively successful while others have not. One recent standard that has gained momentum is the Broadcast eXchange Format (BXF) from the SMPTE S22-10 data exchange group. This format is squarely aimed at the current and future systems needing to exchange Schedule and content information utilizing common XML based messages. Several leading vendors including Harris have already embraced this standard and others are in the process of adopting it. The benefits of open standards such as these are the ability to rapidly introduce unknown systems into the environment and have a common understanding of data, utilize and share the information with minimal impact.

Putting the various elements together such as a common repository and messaging platform and less custom proprietary interfaces we are better positioned to start managing and building enterprise workflows based around the content and share this information with any connected system. One element that is still missing from the equation is common identification of content across each system.

Common Content identification

Common content identification has been a challenge for broadcasters with content distributed throughout the enterprise, each having their own means for identifying content, with differing identification schemes and combinations of elements being used to uniquely identify and match content. Several global content identification schemes exist like ISAN (www.isan.org) and Ad-ID (www.ad-id.org) in an attempt to address this for commercial and audiovisual work.

Today there are hundreds of thousands of pieces of content that do not follow common identification schemes and must be matched up between systems. The role of a common content repository in combination with a common messaging platform is to ensure that new schemes along with old are supported. This includes historic metadata being matched through algorithms where no standard scheme exists. Information can then be matched and exchanged with minimal impact to existing systems.

ENTERPRISE WORKFLOWS

While it's always been possible to implement human workflow management and individual systems have had workflow tools for years, establishing a smooth end-to-end process has remained challenging. Disparate views on business data and batch-oriented interfaces have hindered attempts to pass a task from one application to another efficiently.

Tying it all together

Using common metadata standards, content repository and messaging platform allows a workflow orchestration layer between systems to be implemented. Since all the systems are speaking the same language while sharing data through a publish/subscribe messaging platform, it becomes possible to automate decisions using the data created by any of the systems. Furthermore, because the relevant information is gathered in one place (the common content repository), the orchestration layer does not always need to communicate directly with each system to make content related decisions. The existing systems can be used to carry out and automate the functions they do best, while the orchestration layer fills in automation between functions and provides the overall view of the entire business process.

For discussion purposes, consider a basic example of an operation with several functional areas, each with their own applications and tasks to perform. Each task plays an integral role in the overall objective of the business, and the boundaries between functions are where the process breaks down today.

1) Programming: Researches new program content, negotiates a contract with the content provider, requests materials, creates a programming schedule, establishes distribution policy, and orders promotions.

2) Traffic: Sells commercial inventory, processes electronic orders, processes paper orders, creates invoices, receives payment, and manages advertiser credit.

3) Edit: Creates promotions, edits content for time, and edits content for standards and practices.

4) Digital Asset Management: Catalogs content and creates sub-content records to represent copyrighted material.

5) Ingest: Receives digital material from delivery services, processes materials for broadcast, and creates copies for VOD and Broadband.

6) Automation: Applies presentation rules to the schedule, executes play-out, manages discrepancies, and manages media server storage.

A closer look at the tasks each function performs reveals dependencies on the other functions. This is where batch processes and spreadsheets have prevailed in the past. Now, each task becomes part of a service that the orchestration layer will use to create the overall flow of tasks.

For example, the programming group needs to request materials once a program contract is in place. In the past, this was accomplished by a spreadsheet or flat-file sent to ingest. If you were lucky the ingest system would have the ability to process the file and retrieve some of the data. If not, someone would need to re-enter the information into the ingest system.

Now, this can be accomplished by the ingest system exposing the "ingest service" to the orchestration layer, which in turn is available to receive through the messaging platform ingest requests from the programming system. Utilizing a standard such as BXF, these "requests" come in the form of a schedule message that identifies that the content has been scheduled. From this, actual requests can be made of the ingest system to perform the required ingest, without each system communicating directly.

If however, we can identify through the common repository that the material already exists we can avoid unnecessary ingests by not forwarding the request to the ingest system. If ingest is required the ingest service would perform a well-defined task, in this case "process materials for broadcast". When it is completed, the media is ready to play on air, timing is accurate, and quality control has been done, the ingest system need only publish the existence of the metadata via a new content message and any connected system can be notified of the ingest, including the originating programming system which will now know everything has been performed. The details of how it is done and the workflows are transparent from the programming system.

What are the advantages? For one, requests can be submitted as they occur, instead of collecting them and submitting a list once per day as has been typical. Any delays in the process are due to the response of humans, not systems. Secondly, the ability to change the way parts of the process are implemented without impacting other systems, including introducing new systems or workflows.

Suppose you are ready to transition to a completely file-based ingest process. Instead of receiving tapes and performing the traditional ingest process; files appear on a "catch server" in your facility, complete with additional metadata. The files still need to be made ready for broadcast, have the content timing updated, and be checked for quality control. The result of the "ingest request" has not changed, therefore the capabilities of the ingest service do not change.

What has changed is the implementation behind the ingest service. Instead of presenting a task to a user to find a tape and ingest it, the ingest system can acquire the material directly from the catch server and processes it automatically. The metadata can be captured and other systems brought up-to-date with the new information through publishing a content message (provided the data is from a trusted source, no human intervention is needed). Systems exist today that can analyze the media for technical flaws and perform quality control automatically. This can easily be introduced into the process by subscribing to the content message, performing the automated Quality Assurance and publishing another content message with updated status.

While the digital ingest process is significantly different than the tape ingest process (it could be completely automated), the end result is the same from the programming systems perspective. In the end the media is ready to play to air, the timing is accurate, and the quality control has been done.

Customizing workflows

Consider another scenario where an unforeseen event, such as a live news cut-in, preempts that critical commercial from a brand new account that a salesperson has worked so hard to sell.

In the past, the as-run log would reflect that the commercial was missed, as much as 24 hours after the fact. It is received by the traffic department just in time to stop that invoice from being mailed to the client. The revenue is lost, the salesperson, who had no idea this happened, has already received an irate phone call from the client. If you're lucky, the master control operator would have realized that the commercial was important and started making calls... but what if this happened on Saturday?

Now suppose the same scenario occurs where the systems are connected through a common messaging platform. At the moment the preemption occurs, automation reports it to the traffic system (through messaging and common metadata standards). The traffic system receives the message and performs an analysis of the contract – there is a half-hour run window. The traffic system finds a run-of-schedule spot that can be removed within the next half hour and replaces it with the important preempted commercial. The traffic system then publishes the late schedule change which is picked up by automation.

The salesperson receives a message on their Blackberry from the orchestration layer to inform them that the commercial was preempted, and that the system has taken action to preserve the contractual obligations to air it within the half-hour run window.

Time is of the essence in these scenarios, and the addition of automation between functional areas in the organization using an orchestration layer serves to eliminate costly delays. Additionally, exceptions to the normal process are brought directly to the attention of the right people.

The power of comprehensive data

Did that tape come in yet? Did someone already make a QuickTime copy of it? Is it ready to air? Has it been checked for quality control and standards & practices? Do we have the rights to broadcast it? Stream it on the web? Does it contain music that we can't make available as a web download? How many times has it been aired? How many downloads? We need space on the server – can we archive it now?

Historically, the business data needed to answer all these questions has been spread across systems with no one place to find the answers you needed related to the state of the available content. Added to that the fact that most interactions are point-to-point and with custom formats, broadcasters have and continue to look to multiple systems to implement workflows requiring customizations in each of those systems for every new change in business process.

By utilizing the data made available through a common messaging platform, common standards and the common content repository and by moving applications away from limited modes of operation to one where they make information available as events take place within the systems allows broadcasters to develop true enterprise wide distributed workflows across all systems from Sales to play-out placing control back into the hands of the broadcaster to change and manage workflows.

ACKNOWLEDGEMENTS

Chris Lennon
Harris Corporation
8415 Explorer Drive
Colorado Springs
Colorado 80920

DISTRIBUTED ENCODING ARCHITECTURES

Kenelm Deen
TANDBERG Television
Southampton, United Kingdom

ABSTRACT

The introduction of Internet Protocol (IP) technology into broadcast headends has enabled a move from co-located encoding and multiplexing solutions to new solutions where the encoding stage is located at the content provider. This removes the requirements for contribution encoding and re-encoding for final distribution.

In a distributed encode solutions content is encoded with the philosophy 'encoded once - distribute many'. This saves an additional encode stage, transmission bandwidth between the content provider and headend/s, reduced network demand lowering distribution cost while increasing capacity. Direct-to-Home (DTH) headends typically operate statistical multiplexing to maximise quality and content delivery to the subscriber over the distribution pipe

CENTRALIZED ARCHITECTURES

Typically central headends consist of four distinct stages.

- content acquisition
- encoding
- multiplexing
- modulation

Content acquisition and modulation are often located in diverse locations to the encoding and multiplexing and are typical connected by Telco networks. The encoding and multiplexing stages are traditional co-located. In DTH applications encoding efficiency is important, the lower the bit rate per a service, the more services can be delivered for the same cost over the distribution pipe.

This drives DTH headends to operate Statistical Multiplexing (SM). Video content can be encoded using either Constant Bit Rate (CBR) or Variable Bit Rate (VBR). SM is the process of controlling VBR encoding to squeeze video components into a group bit rate. Bit rate is dynamically allocated across all components in the group, controlled by a SM controller.

Figure 1 show the 4 stages typically found in a central headend.

Content Acquisition

Content acquisition is the process of collecting content for encoding and delivery to the subscriber. Content acquisition means different things to different people. It could mean collecting content off-air using banks of receivers for remultiplexing/re-encoding or could take the form of a contribution feed from a remote encode site feeding high quality high bit rate content directly to central headend for re-encoding. On main interest is contribution acquisition were the headend operates the remote encoders and has control.

Contribution feeds are mainly Single Program Transport Streams (SPTS) encoded using CBR. The feeds are encoded at a higher bit rate than the final distribution bit rate to minimise loss of quality in the second encode pass. Alternatively Multi Program Transport Streams (MPTS) can be used combined with SM to reduce distribution costs to the Central Head End (CHE). MPTS provides greater efficiency than SPTS but adds complexity at the remote site. Figure 2 shows a mixed headend architecture.

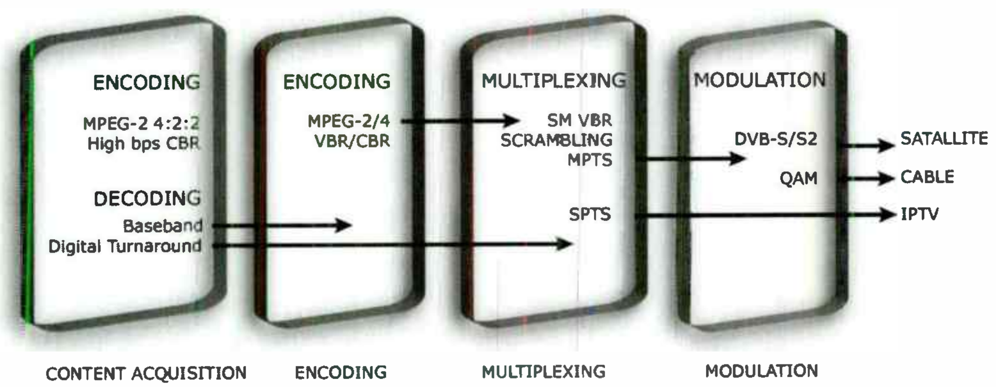


Figure 1 Stage within a Central Headend

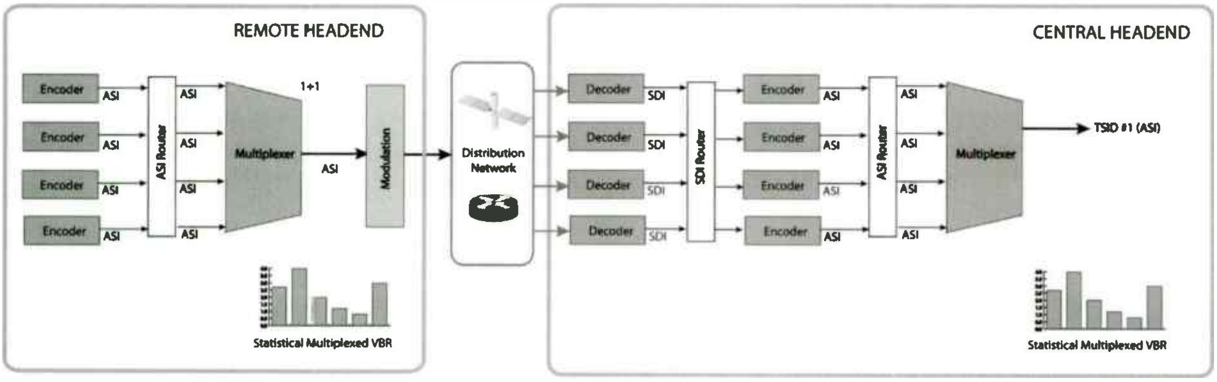


Figure 2 Mixed Headend Architecture

Encoding Stage

The encoding stage encodes the content for delivery to the subscriber population. If the content was previously encoded the bit rate or format may not be suitable for delivery to the home. Contribution feeds are normally encoded at higher bit rates than final distribution rates to maintain quality for editing. Statistical multiplexing is common in central headends, maintaining picture quality while increasing delivery efficiency, enabling additional content delivery without increasing the delivery bandwidth.

The use of ASI between the encoder and multiplexer is ideally suited to MPEG transport and SM systems were the SM controller (often inside the multiplexer) controls the VBR encoding rate of a group of encoder based on collective demand. This assumes closed loop statistical multiplexing systems which deliver increased efficiency over open loop solutions were the multiplexer has no control over the encoder. Open loop SM has limited ability to alter the bit rate while maintaining quality.

Statistical multiplexing places stringent requirements on the interaction between the encoder and multiplexer and SM controller. We shall examine the impact of using IP networks as an interconnection technology on statistical multiplexing and highlight how these issues are resolved to provide high performance SM over IP networks.

Multiplexing Stage

The multiplexing stage packages the content into a service offering for delivery to the consumer. Multiplexing is normally associated with Satellite and Cable platforms were MPTS are used to deliver content to the home via shared distribution pips on a broadcast model. In IPTV headends content is distributed using CBR or capped VBR SPTS due to restricted last mile bandwidth pipes to the home.

The multiplexer stage will often incorporate security, in the application of DVB-CSA (Common Scrambling Algorithm) to the content for secure delivery to the home. The multiplexing stage generates a MPEG-2 transport stream for modulation and delivery.

Modulation Stage

The modulation stage applies the relevant modulation (DVB-S/S2 or QAM) for distribution to the consumer. The modulated signal is then provide to the distribution system for delivery. The modulation stage is normally located with the distribution system and not co-located with the multiplexers.

We shall now examine the interconnection technology focusing on ASI as this is used between the encoder, multiplexer and modulator.

ASI TRANSPORT

ASI is the default standard for interconnecting broadcast compression equipment and is widely adopted within the industry. It provides ideal characteristics for the carriage of MPEG compressed content between devices.

- low (almost zero) latency
- deterministic bandwidth
- constant error rates

These are all ideal characteristics for compressed content were bandwidth capacity is deterministic and known, constant propagation delay and minimal error rates that are correctable using Reed Solomon error correction techniques which add minimal latency to the transmission and are transparent in terms of overall latency.

This is not surprising as ASI (Asynchronous Serial Interface) was developed for carriage of MPEG-2 transport stream packets for the purpose of interconnecting broadcast equipment. ASI was the evolution of the original LVDS/SPI (Low Voltage

Differential Signalling/Serial Parallel Interface) interface.

Transport over ASI provides sufficient bandwidth for compressed content distribution with a maximum bit rate of 216 Mbps. Combined with a low error rate correctable with Reed Solomon error correction with almost zero prorogation latency it is ideally suited to transporting MPEG compressed content.

The main restriction in using ASI to transport MPEG content is:

- 200 meter interconnect limitation
- expensive routing equipment

These restrictions impact today's business models which have evolved requiring dynamic behaviour over static solution that were the norm only a few years ago. ASI imposes restrictions on the design due to high equipment cost and the fundamental requirement of co-locating devices within 200 meters of each other.

IP ENTERS CENTRAL HEADENDS

The industry is currently shifting from ASI to IP enabled headends. Within the last 12 months a shift from standard Telco networks such as ATM and DS3 to IP has occurred within the contribution market. This has fuelled the acceptance of IP within the broadcast industry and now a move to IP enabled headends as started.

The switch to IP enabled headends is being driven by the increased flexibility in using IP over ASI and the reduction in equipment cost. The use of ASI routers and distribution amplifiers add significant cost to headend designs while imposing interconnection and switching restrictions. IP inherently provides flexible switching and routing of content. The cost of IP network equipment is constantly falling compared to ASI equipment which is in a niche market. IP provides the following benefits.

- Increased capacity GigE and 10GigE over 216Mbps limited ASI capacity.
- No distance limitations
- Full duplex (bi-directional) transport
- Flexible routing
- Improved redundancy solutions

Compression vendors are now including IP interfaces on their equipment in addition to ASI with the emphasis on IP connections moving forward. IP headends remove single points of failure, typically introduced by the placement of an ASI router between encoders and multiplexers for redundancy switching, by enabling the multiplexer to select the content streams dynamically from any encoder on the IP network using standard multicast IP protocols.

This flexible content selection, allows a shift in paradigm from isolated encoding pools dedicated to a single transport stream/multiplexer to one of a single encoder stage generating content. Figure 3 depicts an IP headend containing equipment pools.

HD THE DRIVING FORCE

The requirement to broadcast High Definition (HD)

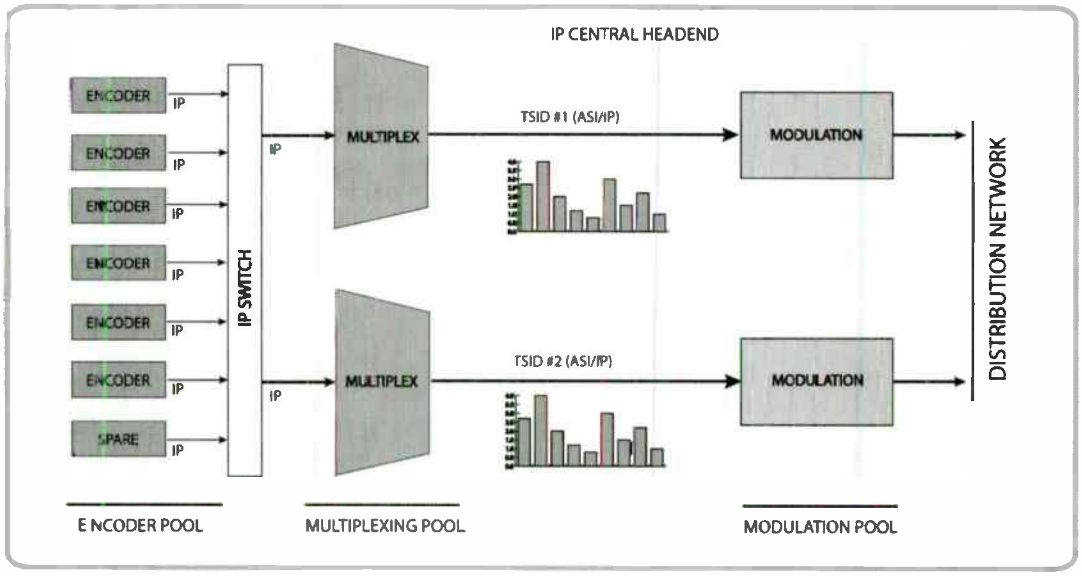


Figure 3 IP Central Headend

content requires substantially greater transmission bandwidth for contribution feeds than Standard Definition (SD). To transmit HD uncompressed content requires 1.44Gbps compared to 270Mbps for SD, while the use of light compression standards such as JPEG2000 reduce the bit rate to around 100Mbps offering a substantial reduction.

- 1.44Gbps uncompressed 1080i HD
- 100Mbps JPEG2000 compressed HD
- 80-85Mbps MPEG-2 4:2:2 HD

Contribution rate MPEG-2 SD requires approximately 18Mbps using 4:2:2 encoding to prevent chroma loss on re-encode. Although MPEG-4 AVC is the chosen format for delivery to the home due to the low bit rate, there are no implementations of the MPEG-4 contribution profiles currently. This leaves MPEG-2 HD for contribution encoding requiring 80-85 Mbps for high quality 4:2:2 transmissions. This high bit rate increases the network transmission costs making the solution unacceptable for contribution/distribution of HD content between remote sites and the CHE.

HD broadcasts are nearly always VBR encoded to squeeze the maximum number of service into the available bandwidth. Typically 3-4 HD services are broadcast in 38Mbps averaging 6Mbps per service in Europe and Asia with 4-5 HD service in the same bandwidth in the US.

Taking this example a contribution link would require a bandwidth of 340Mbps (4x85Mbps). This would require a STM-4 (622Mbps) pipe which would be extremely costly and not to mention access restrictions. This has forced broadcasters to place remote compression headends closer to the content to minimise the contribution link bandwidth. This adds complexity, adds additional equipment costs and compromises the video encoding efficiency and quality due to encoding the content twice at similar bit rates. This is not ideal but is currently the only solution available.

Distributed Encoding To the Rescue

The use of distributed encoding over IP networks enables broadcasters to encode HD content once in the correct format (MPEG-4 AVC) for transmission to the CHE for multiplexing and distribution to the subscriber.

Distributed encoding benefits:

- Single encode stage
- Lower contribution bandwidth between sites
- Simplified remote sites

This achieves the goal of maintaining the encoding efficiency, simplifying the equipment at the remote site and reducing network costs.

DISTRIBUTED ARCHITECTURES

IP enabled headends remove the requirement to co - locate the encoding and multiplexing stage, enabling the encoders to be distributed. This allows the broadcaster to place the final emission encoders at the content provider, removing the need to encode multiple times and reducing the transmission bandwidth required between sites.

In pushing the encoding stage out to the content providers the need for traditional contribution encoding disappears. Now only a single encoding stage is required removing the need to encode twice while improving the quality and reducing the contribution bandwidth. A distributed architecture is shown in Figure 4.

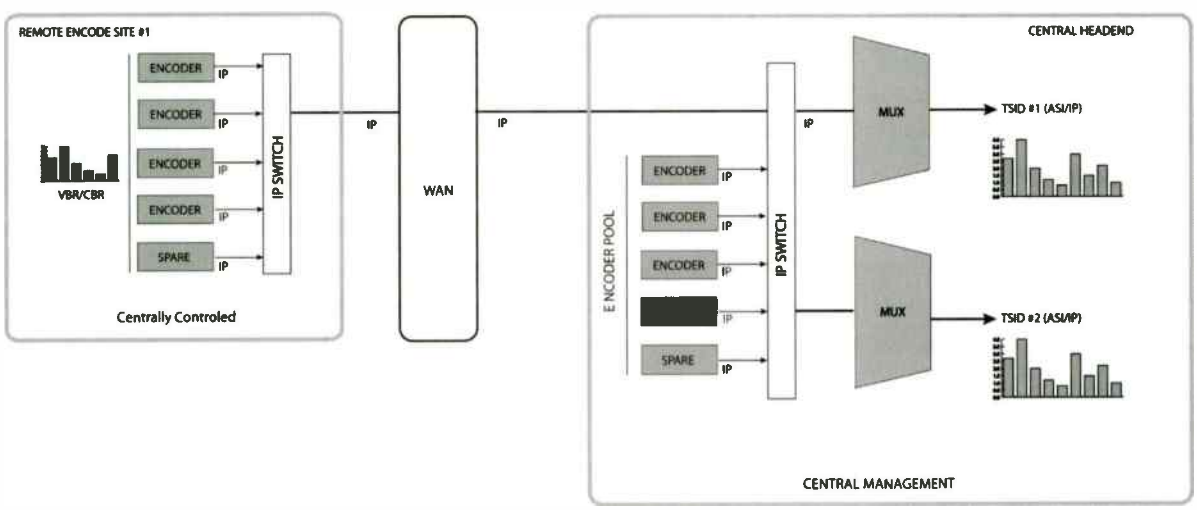


Figure 4 Distributed Encoding Architecture

Now rather than duplicating the functionality of the CHE at the remote site the CHE becomes distributed leaving only the multiplexing and possibly modulation.

An IP network is used to link the encoding and multiplexing stage. In distributing the encoders the IP network characteristics become more important. The move from a Local Area Network (LAN) with specific performance characteristics to a Wide Area Network (WAN) with variable characteristics emphasizes the affects of carriage over IP network. Network introduced impairments are:

- high network latency (>50ms)
- packet loss and reordering
- variable inter-arrival-packet-time (jitter)

These network characteristics add additional complexity to IP multiplexers as they must compensate for impairments, otherwise services disruption will occur. IP networks can provide the required characteristics for transporting MPEG compressed content but require careful design.

IP NETWORK CHARACTERISTICS

The carriage of real-time content places stringent requirements on the underlying network. A real-time broadcast requires constant bandwidth, consistent latency, and a low known bit error rate (BER). Traditional transport mediums such as ASI and ATM provide deterministic characteristics which are ideal for broadcast transmissions, whereas Ethernet network are inherently un-deterministic. We draw the comparison with ATM technology as this was the preferred Telco network for contribution/distribution before the move to IP.

IP is a packet switched technology unlike ASI which contained no framing and purely carries content between two points (point-to-point). In contrast ATM is a circuit switched technology which requires setting up of dedicated routes (connections between two nodes) for the duration of the transmission. IP packets are routed to their destination determined by network routing algorithms and do not require pre defined routes to be established.

Transmission characteristics required for real-time content.

- 100% network availability
- guaranteed bandwidth
- consistent network latency
- defined error rate allowing real-time recovery

If any of these characteristics are not meet, service disruption will occur. For real-time transmissions fixed bandwidth is essential for the duration of the transmission, if the network is unable to sustain bandwidth demand then service disruption will occur as packets are discarded. With network introduced packet loss or variable latency the receive device will not be able to replace missing or delayed packets without affecting the service.

ASI was specifically designed for the carriage of MPEG-2 transport packets, unlike IP networks which were designed for general purpose traffic utilising a 'best effort' approach. IP was defined for flexible routing and this is a key strength in its adoption within the broadcast industry combined with its flexible bandwidth allocations.

Broadcast IP networks are designed for the distribution of broadcast content with specific characteristics and sustained bandwidth availability. With the demand on IP networks to provide a consistent quality of service comparable to traditional networks, a number of protocols can be applied to provide the required Quality of Service (QoS).

Let us now examine the possible impairments introduced by carriage over an IP network.

Transmission Impairment

The encoder generates an IP packet flow which is routed over the IP network until it reaches the multiplexer. The process of transmitting the IP packet across the network may introduce impairment when the MPEG-2 stream in reconstructed. The following impairments are typical.

- packet loss due to random or burst errors
- jitter from queuing within the network
- sequence errors (out of order packets)
- variable latency between TX and RX sites

The MPEG-2 Transport Stream (TS) packet is the preferred container for carriage of MPEG compressed content between broadcast devices. An alternative is the encapsulation of the Elementary Streams (ES) containing video, audio etc directly into IP packets for delivery to consumer devices such as Set-Top-Box (STB) and handled devices.

The MPEG-2 TS packet must be encapsulated in an IP frame for transmission over the network. A standard IPv4 frame is 1500bytes and an MPEG-2 TS packet is 188 bytes. This allows 7 MPEG-2 TS packets to be encapsulated into a single IPv4 frame. IPv6 frames are generally not used unless you can guarantee the whole network is IPv6 enabled otherwise the larger IPv6 frames will be fragmented resulting in jitter on the MPEG stream. RFC2250 defines the use of RTP (Real Time Protocol) for the transmission of MPEG content as this allows ordering errors to be detected and possibly corrected.

Impairments result in service disruption unless corrected. Depending on the level of impairment, correctional techniques can be applied to recover and provide a stable transmission. Once such technique is the application of Forward Error Correction (FEC) which can allow the receive device to recover missing packets. In addition network protocols such as traffic classification and prioritization can be used to apply edge-to-edge Quality of Service (QoS) based on Traffic Engineering (TE) to ensure bandwidth and minimise jitter and sequence errors.

We shall now examine the correction techniques available to protect against these impairments.

MPEG-2 TRANSPORT STREAM RECOVERY

The receiver device must recover the original transport stream transmitted by the encoder without the need for retransmission. We shall examine the impairments introduced during transmission over the IP network.

Packet Jitter

The received MPEG-2 TS will have had jitter introduced during the transmission over the IP network. MPEG-2 TS packets are transmitted in IP frames which contain a maximum of 7 MPEG-2 TS packets.

This encapsulation process introduces a small amount of jitter but carriage over the network introduces the largest jitter component. This IP jitter (inter-packet arrival-time variance) results in MPEG PCR jitter which if not removed may affect the service depending on the level of jitter introduced. MPEG specifies the input signal should be within ±500ns of PCR jitter and multiplexers normally are designed to remove a few ms of jitter from the received MPEG-2 TS.

IP packets are not transmitted at a constant rate. Packets are sent in bursts between network devices as they transverse the network. If the network is experiencing high utilization the network will hold packets in buffers adding variable latency resulting in PCR jitter. The multiplexer must smooth out the introduced jitter. This is typically achieved using a de-jitter buffer. The de-jitter buffer absorbs input jitter by buffering the received IP packet flow and emptying the buffer at a constant rate equal to the original transmission rate. Figure 5 depicts an IP de-jitter buffer and the smoothing process.

This requires the multiplexer to lock to the incoming input flow and determine the incoming rate in order to set the correct empty rate to removing the jitter. A reliable method of determining the TX rate of the IP flow is to use the 27MHz PCR clock contained within the MPEG-2 TS packets. Taking successive samples from the PCR value the output rate can be determined quickly and accurately. Alternative approaches derive the TX rate using the RTP 90Khz clock which has a lower resolution than the 27MHz PCR clock significantly increasing the locking time.

The size of buffer is dependent on the amount of jitter that must be removed. The addition of de-jitter buffers add additional latency to the transmission path which maybe undesirable. It is essential to minimise network introduced jitter in order to reduce end-to-end latency through the transmission system.

Latency

Latency is added between the encoders and multiplexer due to the propagation time of the IP flow over the network. This additional latency component forms part of the overall end-to-end system latency. This latency component cannot be removed once introduced. It is important to minimise network latency to keep end-to-end delays latency down and minimise redundancy switching time between remotely located encoders.

In CBR encoding the addition of latency between the encoder and multiplexer does not add complexity to the solution. However in a VBR statistical multiplexing solutions the design must compensate for the presence of this latency otherwise the performance of the statistical multiplexing will be impaired dependent on the amount of latency introduced.

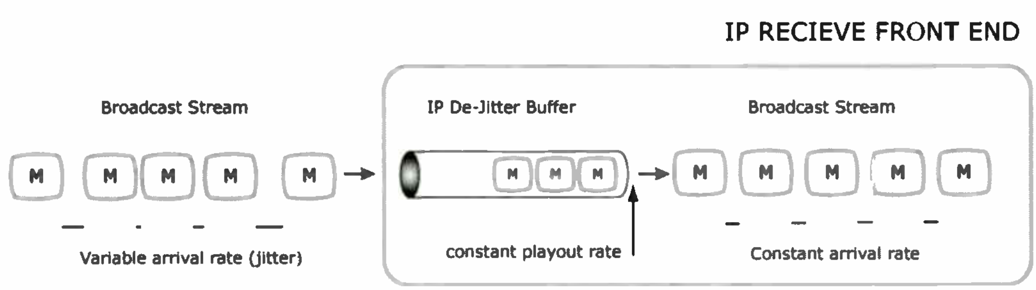


Figure 5 De-Jitter Buffer Model

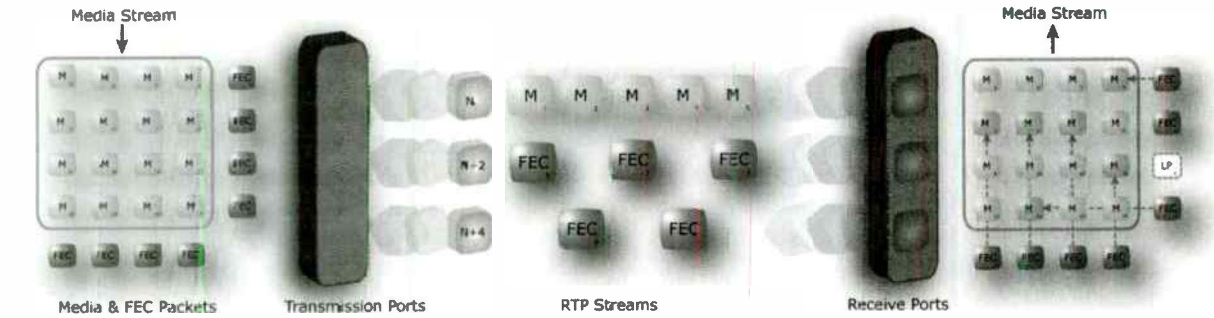


Figure 6 FEC Transmission Model

Packet Loss

During transmission over the IP network packet loss may occur. Missing IP packets will result in lost MPEG-2 TS packets and service disruption. Remember that a single IP packet can contain 7 MPEG-2 TS packets. A single lost IP packet causes a service outage. For this reason the application of error correction data is essential for the recovery of missing packets.

The Pro-MPEG Code of Practice (CoP) Forward Error Correction (FEC)/SMPTE 2022-1 scheme defines an application were additional FEC data streams are transmitted in separate IP flows to the content enabling real-time recovery at receive devices. Figure 6 shows the process of FEC transmission and recovery.

The Pro-MPEG CoP FEC provides a flexible implementation whereby the level of error protection can be tailored to meet the networks performance. Figure 7 shows the parameter trades possible.

FEC provides:

- Burst and random packet loss recovery
- Variable protection levels

The use of FEC introduces additional considerations.

- Increased TX bandwidth.
- Additional transmission latency due to the decode process.

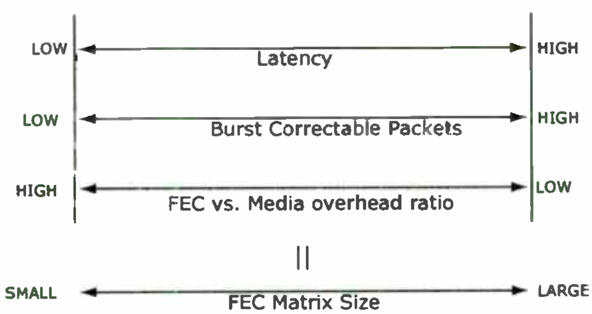


Figure 7 FEC Parameter Trade Offs

In CBR headend implementations the addition of FEC between the encoder and multiplexer adds no additional complexity other than FEC processing. In VBR statistical multiplexing solutions the addition of latency between the encoder and multiplexer has implications on the design of the statistical multiplexing system to handle the additional delay introduced by the FEC scheme. We shall examine the affects of latency on statistical multiplexer later on.

Sequence Errors

The other main error common in IP transmission is the possibility of receiving packets in a different sequence than they were transmitted. This is termed sequence errors and is a result of the network sending different IP packets over different routes through the network.

RFC2250 defines the use of Real Time Protocol (RTP) for transmission which includes a sequence number in the IP header allowing a receive device to determine out of sequence packets. Figure 8 shows received sequence errors.

In combination with the de-jitter buffer this allows the receiver to reorder out of sequence packets. Excessive out of order packets can not be recovered without the use of FEC since the receiver can only wait a limited time before signalling a packet as missing.

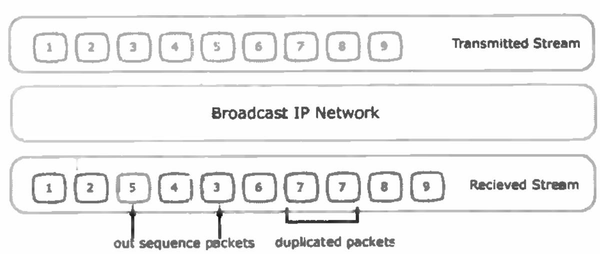


Figure 8 Receive Packet Sequence Errors

STATISTICAL MULTIPLEXING

Statistical multiplexing is the process of VBR encoding multiple video components within a multiplex, allocating dynamically bit rate per a video component from a shared bit rate pool. There are two fundamental concepts, open and closed loop systems. In open loop the downstream device has no interaction with the encoder and can not adjust the encoded bit rate. In closed loop systems the multiplexer interacts with the encoders and controls the encoders bit rate based on encoder bit rate requests across all the encoders.

We shall examine closed loop statistical multiplexing as this is the most common format for encoder multiplexer solutions. Statistical multiplexing is nearly always deployed as it enables the broadcaster to deliver additional services within the same transport stream bandwidth than if using CBR encoding. This reduces the cost per a service by typically around 30%.

ASI Based Statistical Multiplexing

Statistical multiplexing has been around for many years with the first generation solutions operating on proprietary interfaces between the encoder and multiplexer such as TAXI (Transparent Asynchronous Transceiver Interface) due to the high speed bi-directional properties were data and control could share the same interface.

Statistical multiplexing traditionally requires the following characteristics between the encoder and multiplexer/statistical multiplexing controller.

- Constant low latency MPEG interconnect
- Reliable high speed low latency communications channel

These characteristics are essential in statistical multiplexing in order to change the encoder bit rate within short accurate time intervals to allow precise and co-timed changes in bit rate between various encoders.

With the introduction of ASI interconnects between the encoders and multiplexer vendors moved to split solutions were MPEG-2 TS packets were sent over ASI and statistical multiplexing messages over another interface such as IP/Optical/High Speed RS232.

SM Message Cycle

The following cycle takes place between the encoders and the SM controller. In a SM system the encoders will encode using VBR at the bit rate allocated by the SM controller.

- The encoders will typically request an encode bit rate dependent on the difficulty of the source material based on an initial first pass encode.

- The SM controller will allocate a bit rate to each encoder taking the overall bandwidth available and the demands from the individual encoders into account. Some priority rules such as target encode quality and bit rate limits will apply in setting the bit rate. The priority rules are often set dependent on service importance to the broadcaster.

- The encoder will then adjust the compression level to meet the assigned target bit rate set by the SM controller.

Figure 9 shows this continuous cycle. The interval between encoder bit rate request is driven by the need to alter the bit rate to ensure consistent encode quality while efficiently allocating bit rate across all encoders within the group. This cycle of messages must happen quickly and reliably in order to maintain rapid bit rate changes.

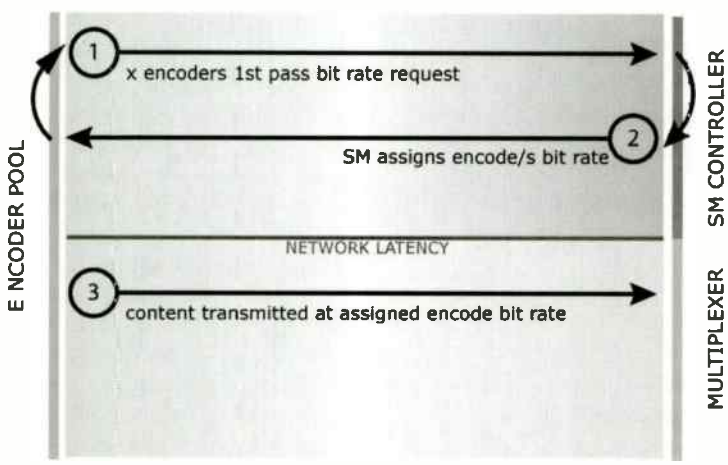


Figure 9 SM Message Cycle

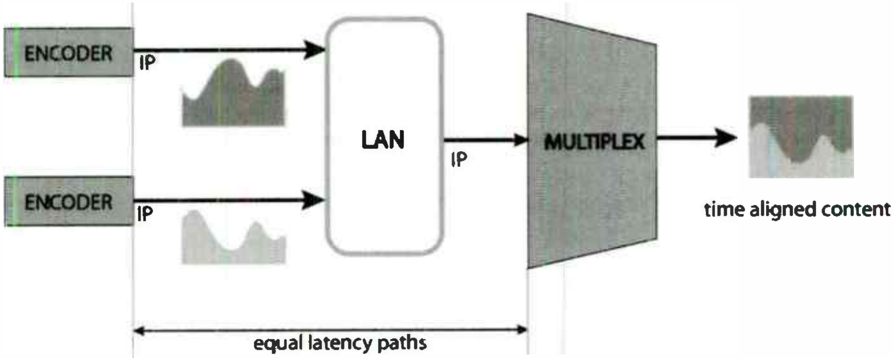


Figure 10 Equal Latency Paths

If we assume that the bit rate should change at the field/frame rate of the source content in order to maintain encode quality. Taking an NTSC frame rate of 30fps we are required to change the encoder bit rate every 33.33ms based on frame level refresh or every 16.66ms for field level refresh. This means that the round trip time between the encoders and SM controller must be less than 16.66ms for message exchange and internal processing time and the MPEG-2 TS interconnects must have a latency less than 16.66ms for the encoded content to arrive at the multiplex in time.

It is essential that the interconnection between the individual encoders and multiplexer have the same latency to ensure that the multiplexer receives the correct bit rate from each encoder at the correct time. Figure 10 shows SM with equal latency paths.

With ASI the latency is almost zero and constant between all encoders. This assumes that all encoders are located in the same physical location ensuring propagation delay is not relevant. These characteristics are easily achievable in headends with co-located encoders and multiplexers. This becomes an issue in distributed solution were the encoders are located in spilt locations and separated from the multiplexer.

LAN IP Based Statistical Multiplexing

The introduction of IP networks connecting encoders and multiplexer within a central headend places stringent requirements on the networks performance for statistical multiplexing.

For SM the IP network must have.

- Low latency unidirectional transmission from the encoders to multiplexer.
- A bi-directional communication channel for SM messaging with a response time of less than 16.66ms in order to avoid imposing a performance restriction on the bit rate refresh rate.

Figure 11 shows different latency components in a CHE.

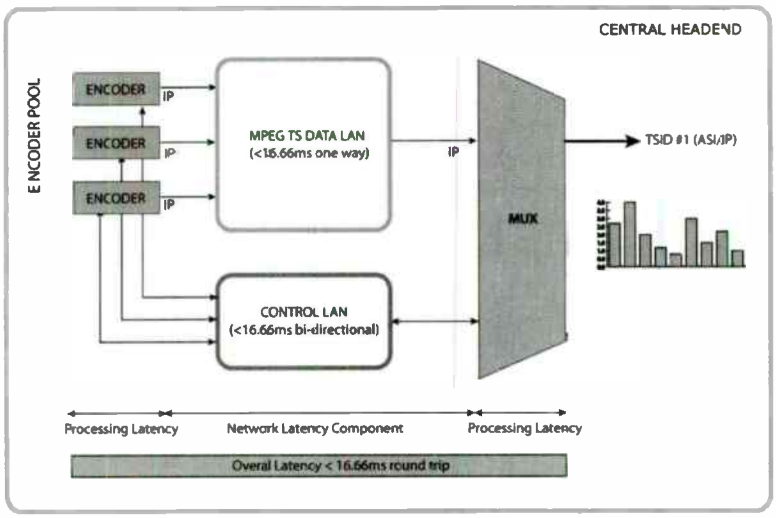


Figure 11 SM CHE Timing

In addition to the characteristics required for real-time content transmission.

- 100% network availability
- guaranteed bandwidth
- consistent latency
- defined error rate allowing real-time recovery

A well engineered Local Area Network (LAN) can easily meet these requirements ensuring the same performance level associated with ASI implementations. GigE networks are normally used as they provide sufficient bandwidth without becoming congested which would compromise performance. The LAN will be an isolated network solely for the interconnection of the compression equipment. A dedicated network should not introduce packet loss if correctly designed and does not require the use of FEC protection schemes. We have now seen that IP enabled headends can meet the performance demands required for statistical multiplexing.

We will now examine the affects of distributing the encoders at different locations to the multiplexer connected by an IP Wide Area Network (WAN).

WAN IP Based Statistical Multiplexing

Operating statistical multiplexing over a WAN introduces additional complexity. The introduction of a WAN between the encoders and multiplexer can introduce:

- latency in excess of 16ms
- variable latency paths between sites
- increased susceptibility to packet loss

The introduction of increased latency exceeding 16ms on the communication path will affect the rate at which the bit rate can be altered on the encoders if the same model used in a LAN is implemented I used. In order to maintain performance and rate of change of encoding bit rate a different solution is required.

For a high performance SM solution the following issues need to be resolved.

- Latency equalisation between encoders sharing a common bit rate pool, ensuring all content arrives at the multiplexer is time aligned.

- The SM algorithm adjusts for variable latency between the encoders and multiplexer and encoder and SM controller.

We shall now examine the affects these have on statistical multiplexing.

Latency Equalisation

In SM headends the latency between all encoders sharing a common bit rate pool must be the same. If the encoders do not share a common latency the summation of all encoded bit rates arriving at the multiplexer will either overflow or underflow the group allocated bit rate. An underflow wastes bit rate and reduces the efficiency, while an overflow breaks the transmission pipe potentially resulting in service disruption.

The effect of different latency paths between two encoders and a multiplexer is shown in Figure 12.

We can see the effect of different latencies between encoders in the same bit rate pool and the need to equalise the latency to a common value. In co-located encoding and multiplexing headends latency equalisation is not required as the latency is constant across all paths.

With the introduction of IP, headends containing local encoding and multiplexing do not require additional

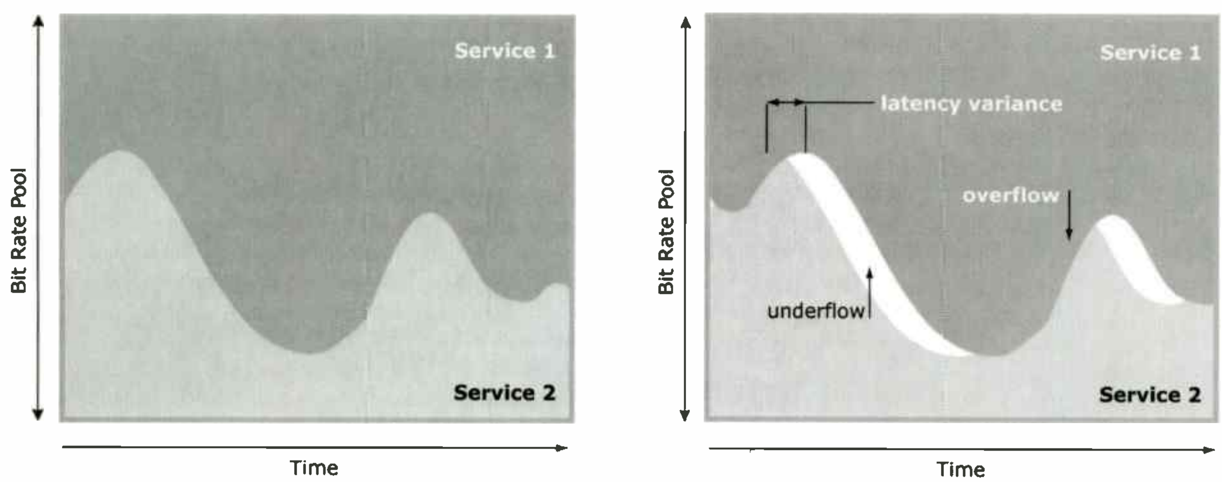


Figure 12 Latency Path Effects

design changes due to constant latency across all encoders (assume same location). Most multiplexers are capable of removing a small amount of latency variance using the de-jitter buffer.

With a WAN implementation if all encoders are located at the same location then it is likely that the delay will be reasonably consistent across all the encoders. If the encoders are located in more than one location and are contained within the same SM pool then latency equalisation is essential.

The latency variance between all encoders must be calculated and then removed to ensure content arrives time aligned at the final multiplex engine within the multiplexer. Once the longest latency has been calculated all encoders can be equalised to this value through additional buffering. It should be noted that this adds additional latency to the end-to-end system latency.

SM Algorithm

With the introduction of variable latency paths between the encoders and the multiplexer for transporting the compressed content and the increase in latency between the encoders and multiplexer for communication the SM algorithm must be changed in order to maintain the SM performance associated with current ASI solutions.

Variable Latency Paths

The introduction of variable latency paths between the encoders and the multiplexer is the main complication with distributed statistical multiplexing solutions. The removal of latency is simple enough with the use of buffers to equalise the individual IP flows from the encoders. The calculation of the amount of latency between the encoders and the multiplexer is the complicated part.

A distributed SM solution must calculate the latency paths and allow adjustment to equalise the content delivery from the encoders at the multiplexing engine within the multiplexer. A common reference point is required between the encoders and the multiplexer in order to determine the latency of each RTP flow. The more accurate the calculation of the latency paths the more accurate co-timed bit rate changes between the encoders can be managed.

The use of a common time source across the encoders and multiplexer provides a reference point for the calculation of latency paths. Once the latency paths have been calculated, all the paths must be equalised to the longest latency path to ensure that all content arrives at the multiplexing engine co-timed. The calculation of latency paths can be static or dynamic, since these paths should not vary much over time. Only slight variance is expected due to jitter on the network.

Communication Latency Path

The latency between the encoders and SM controller is also important. The SM controller should always be co-located with the multiplexer or be aware of the worst latency path between the encoders and multiplexer.

The increase in latency between the encoders and SM controller must be accounted for in designing a solution. The encoders should constantly send bit rate requests to the SM controller and must have sufficient buffering to wait for a response on the encoding bit rate to encode at. The amount of buffering will depend on the latency between the encoders and the SM controller.

CONCLUSION

The introduction of IP networks into broadcast headends has not been without its challenges, both commercially and technically. Broadcasters have had to gain confidence in MPEG transport over IP networks. Broadcasters are now embracing IP enabled headends over traditional ASI infrastructures.

The main drive in adopting IP has been its ability to offer new headend architectures which are more flexible than ASI solutions, removing single points of failure and simplify the process of contribution encoding with the placement of distributed encoding. The introduction of distributed encoding has simplified remote sites by removing the need for two stage encoding and multiplexing, reducing headend cost and lowering network transport bandwidth between facilities. This enables an "encode once – distribute many" model to be adopted.

IP networks are able to provide the quality of service and availability required in the broadcast environment. The use of engineered IP networks combined with techniques such as SMPTE 2022-1 FEC allows reliable transmission over wide area networks.

The introduction of IP into headends has posed many technical challenges in providing the flexibility associated with IP while maintaining traditional performance characteristics set by ASI. These technical challenges have been meet enabling broadcasters to take advantage of the benefits IP has to offer. The biggest challenge has been the introduction of latency and packet loss and the affect they have on statistical multiplexing and service delivery.

IP Headends have arrived and are being deployed worldwide replacing ASI solutions. The adoption of IP will eventually result in ASI disappearing from broadcast equipment as standard and becoming an optional legacy interface.

NEW APPLICATIONS FOR MONITORING YOUR TV BROADCAST – IT'S MORE THAN JUST REPLACING TAPE

Gary Learner
Volicon, Inc.
Burlington, MA, USA

ABSTRACT

Monitoring of television broadcasts traditionally has been used to ensure compliance with FCC regulations on indecency monitoring, verify closed-captioning, confirm music licensing rights and use, and to verify ad playout. In the VHS tape-based world of monitoring, broadcasters and cable operators were limited to this valuable but very basic functionality.

MONITORING IN THE ANALOG ENVIRONMENT

In that analog environment, broadcasters used multiple VTR machines to record their own broadcasts and those of their competitors, a process that required the constant changing of tapes and acceptance of a relatively poor-quality recording. Beyond these issues were other problems including the limited shelf life of tape and the gradual degradation of video each time the tape is used.

The duplication of tape-based media for monitoring also is a cumbersome process. It's a slow process, and the end result is never as good as the original. Furthermore, the content being copied is unavailable to other users until the duplication process is complete and the tape restored to the library.

The greatest limitation of the tape-based model, however, is in broadcasters' ability to access and distribute recorded content. To find a specific piece of content on tape means going into the tape library, looking up a tape in the library catalog, tracking down that tape, and using a tape playback machine to scan for that content or move to the desired point in the broadcast. In some cases, this is like looking through the pages of a book, trying to find a passage without knowing where it is.

When recording to tape, the broadcaster doesn't have access to that content until the tape is finished. If the broadcast is being recorded to an eight-hour tape, then it can take nearly that long to be able to provide internal quality control or external clients with verification of the content aired.

MONITORING IN THE DIGITAL ENVIRONMENT

The shift of the broadcast industry toward digital operations has allowed for the introduction of new technologies, which in turn offer much greater functionality than earlier analog systems. The award-winning Volicon Observer is one such system.

The Observer broadcast monitoring, streaming, and archive solution delivers the power and capabilities of professional AV control rooms to ordinary desktop computers. Multiple channels of content are available on-demand, live or archived, from any desktop, where users can search, retrieve, view, analyze, annotate, share, and export video. A Web-based GUI turns mission-critical video broadcasts and analysis into work material to increase efficiency and streamline workflows.

Technology like this has been made possible through advances in digital video compression and in the development of Internet-based tools. At one time the argument for tape-based monitoring was the low cost of tapes and ease of integrating VTRs into existing analog environments. Now, however, the cost of digital storage and infrastructure has dropped. It's easier than ever to store video, and continuous improvements in storage technology have brought down storage costs, making it feasible to store digital data inexpensively over long periods of time.

Web-based work tools used over high-speed Internet connections now enable easy sharing and distribution of digital media to just about anywhere. Here, too, costs have dropped, and many companies already have access to the necessary network infrastructure needed to make digital media available to remote users.

Among the many benefits of monitoring broadcasts using digital media is that everything is instantaneous. Any queries or complaints about a particular broadcast or ad can be addressed instantly because data becomes available immediately, as soon as it's recorded, not four or eight hours later. Users across a facility, or working remotely over a high-speed network, can access content simultaneously. Thus, media is available to any authorized user, anywhere, at any time following the broadcast.

Duplication of recorded media takes place much more rapidly in the digital environment, without the loss associated with VHS tape. Nonlinear or random access to content, as opposed to sequential access on tape, offers tremendous time savings in finding and accessing the desired portion of a media asset.

EXPANDING THE APPLICATIONS OF DIGITAL MONITORING

In addition to giving greater functionality and flexibility to the individual operator, advanced monitoring systems allow for centralized monitoring, competitive news analysis, quality of service (QoS) monitoring, automated import of as-run logs, and advanced ad and content verification though closed-captioning text and keyword searches. This new generation of monitoring technology also adds value to content repurposing operations, as recorded content can be easily clipped and published on a Web site or pushed to mobile devices.

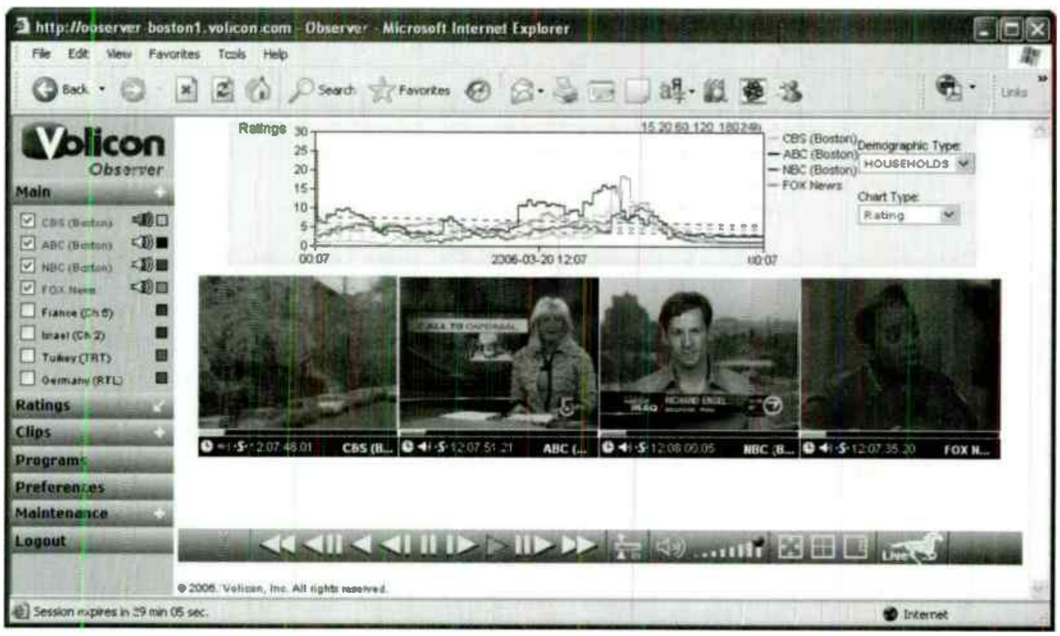
Competitive News Analysis

One of the key drivers behind any investment for a broadcast facility is the need to remain competitive. New monitoring technology goes beyond compliance, closed-caption verification, ad verification, and licensing to offer competitive news analysis, with the ability to compare and analyze program content across local or national news networks. For major networks this capability provides critical, up-to-the-minute information about breaking newscasts on any monitored station and viewer ratings figures indicating how well their own programming is faring with viewers.

During multichannel monitoring, broadcasters can use a mouse and Web browser-based interface to add or subtract channels for monitoring, toggling among video-window layouts – windowed, split, and full-screen – as needed. The system records and displays selected stations simultaneously and in sync. Thus, a broadcaster can clearly see which station or network broke a news story first and how the story unfolds on each network.

Broadcasters can take this evaluation one step farther and correlate all of the recorded broadcasts with ratings numbers to learn more about how viewers respond to the content and style of different broadcasts. Once broadcasters begin analyzing themselves in relation to competitors, they can discover which components or features work well to attract viewers, and they can make use of this information to restructure their programming, make it more attractive to viewers, and therefore increase not only ratings, but also the amount they can charge for ads with higher viewership.

This comprehensive analytical tool offers other valuable information about the network's competitiveness within the market. If a particular network is able to establish that it broke a story first or provided details unavailable on any other network, it can take that information to viewers and also use it as positive press and a sales tool in attracting more advertisers or negotiating higher rates.



Viewer's ratings synchronized with broadcasted content

Sales Communication and Ad Generation

Advanced video monitoring systems give broadcasters the ability to provide proof of performance of sponsored elements during a broadcast and, in turn, to expedite the revenue-collection cycle.

In addition to conventional commercial breaks, many networks – typically sports networks – air sponsored elements during the actual event. Unlike a commercial, which is routed through the traffic system and carries a special ID, sponsored elements such as "Tonight's game is sponsored by …" are not so easy to identify. Closed-caption search functionality incorporated into systems such as the Volicon Observer enable users to find key words, such as a sponsor's name, within a specified window of time and then locate exactly when and where the sponsored element ran, create a clip of it, and email or ftp it to the advertiser or sponsor. These clips serve as a literal record of what went to air.

The exporting of data is a highly automated process, making it easy for operators to send or post clips to an ftp site in the size and format required by each client. The broadcaster's ability to communicate this information rapidly increases the advertiser's confidence that their promotion, whether 30-second clip or quick reference during the event, was in fact aired according to contract. Because clear proof of this is provided very shortly after the broadcast, there is little room for argument and advertisers tend to pay up more quickly. Anecdotal data suggest that this style of sales communication shortens the collection cycle dramatically, cutting it from an average of 90 days to just 30.

Quality of Service (QoS) Monitoring

Sophisticated monitoring systems not only allow broadcasters to show their clients that material went to air as intended, but also provides a valuable tool for evaluating the integrity and quality of the broadcast audio and video. Staff members can review aired material for on-air discrepancies, signal problems, and other issues that may interfere with the quality of the programming. Immediate access to any content that aired and the ability to access that content quickly significantly enhance the quality of broadcast productions.

Current monitoring technology can detect faulty of false signals, reporting on errors and alerting the broadcaster via email, SMS or SNMP trap that something is wrong with the broadcast. Immediate detection of low audio, loss of audio, a blank or frozen screen (no motion detected), and loss of closed captioning helps minimize down time, thus preserving the integrity of the on-air product, meeting the expectations of viewers and advertisers, and protecting station or network revenues.

Other functions possible with the move from analog to digital monitoring systems include ensuring the presence of a station ID or V-chip rating (G, PG-13, MA, etc.).

Fault List

Search Clips

Details: Time range: None Search

ID	Encoder	Channel	Details	Defect	Audio	QAM	Detected
410	E04	60	HGTV-Home and Garden Television	All back	14300	QAM	09-21-06 06 45:20 am
409	E04	60	HGTV-Home and Garden Television	Black scr.	35	QAM	09-21-06 06:36:12 am
408	E04	14	WGBX-PBS (44) (14) Encoder E04	All back	6047	QAM	09-21-06 05 28:58 am
407	E04		WGBX-PBS (44) (14). Encoder: E04	Black scr.	37	QAM	09-21-06 05:02:20 am
406	E04	14	WGBX-PBS (44) (14). Encoder E04	Video back	-	QAM	09-21-06 04:08:36 am
405	E04	14	WGBX-PBS (44) (14). Encoder: E04	Black scr.	39	QAM	09-21-06 04:02:38 am
404	E04	33	CNN-Cable News Network (33)	Video back	-	QAM	09-21-06 01:59:38 am
403	E04	33	CNN-Cable News Network (33).	No video	-	QAM	09-21-06 01:59:06 am
391	E04	14	WGBX-PBS (44) (14). Encoder: E04	Video back	-	QAM	09-19-06 03:09 51 am
390	E04	14	WGBX-PBS (44) (14). Encoder: E04	Black scr.	40	QAM	09-19-06 03:06:22 am

1-20 of 349 found Pages 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

List of on-air discrepancies, descriptions and links to content

Automated As-Run-Log Import

Recording all of the content is by its nature chronologically structured, but without information such as an as-run log or electronic program guide (EPG), there is no way to know where programs begin and end. This knowledge is quite valuable in allowing the operator to segment content and/or sort it according to various criteria including not only type of program or program name, but also length of the program. Segments of similar duration often fall into similar categories. For example, segments of just 15 or 30 seconds most likely are commercials. The ability to identify and sort segments in this manner offers users a quick way of finding or isolating specific content.

Monitoring systems such as the Observer are able to take the as-run logs created by automation systems and match that log data to the video being recorded based on time-code data. Until recently, this has been a labor-intensive manual process requiring a significant amount of time, especially for multichannel broadcasters such as station groups and networks. The automatic import and segmenting of video based on as-run information saves time, eliminates the potential for operator error, and makes the segmented content available for review much faster.

Content Repurposing

Distribution of media assets to viewers over new platforms such as the Web and mobile devices offers broadcasters the opportunity to increase revenues leveraging assets they already own. Video monitoring, recording, archiving, and streaming solutions – including the Observer system – feature tools through which recorded content can be clipped and easily published to a Web site or pushed to mobile customers via PDAs, handheld video players, and cell phones.

One of the most common examples of how stations repurpose content is the publishing of news clips on the station Web site. By putting current events, sports highlights, and breaking news online, the station promotes its televised news programming while providing a valued local resource for both existing and potential viewers. With analog systems the operator had to wait until the end of a newscast to begin generating and publishing clips and images to the Web site.

Digital monitoring systems, however, make content available as the news develops, so the user can begin work on Web content immediately. Furthermore, because the system is timecode-based and nonlinear, it's much easier for staff to find and cut out segments of the broadcast, even while the news is still being broadcast. Time-to-Web, if you will, can be reduced dramatically, allowing the station to make the news available to its Web audience much faster and, as a result, deliver the most up-to-date and accurate information. The very same process also speeds delivery to the growing mobile market, streamlining the repurposing workflow and giving broadcasters more flexibility in leveraging their unique news content.

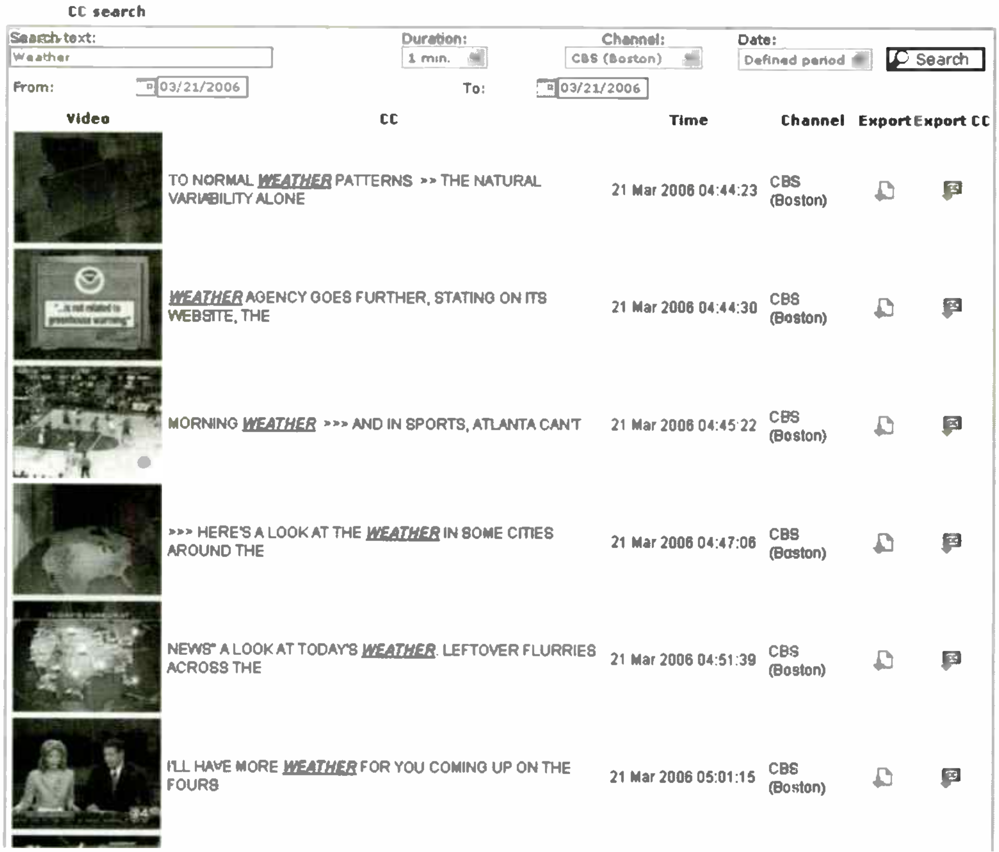
Keyword Recognition

Keyword recognition allows the monitoring system and its users to recognize and report keywords in closed-captioning for breaking news and indecency. From the broadcaster's point of view, this capability can provide critical cues when competing networks or stations break a news story first. Alerted to any "breaking news" on other networks, the broadcaster can adapt news programming quickly to get up to speed.

While broadcasters can use the keyword recognition feature to make sure they're on top of the latest news or to make sure they're aware of any possible indecency violations, this tool is also valuable to media monitoring companies or large corporations that want to monitor themselves in the media.

With the monitoring system set to any number of channels – typically the major news and business networks – the user can configure the system to send an alert every time the company name is mentioned. With a text message or email delivered immediately to the designated contact, authorized users at the company can log in to the monitoring system and view content to find out exactly what was said. This instant insight into media buzz gives the company's officers and public relations department the tools they needs to better prepare for announcements, press conferences, or other publicity.

Real-time monitoring is also valuable to political and government agencies, which can use it to monitor topics related to their field of interest. For example, the mayor's office could use the system to ensure it's alerted any time a certain issue or topic is mentioned on a local channel. Armed with the latest knowledge about the issues and what's being said about his or her actions, the mayor won't be blindsided when, 15 minutes later, reporters ask for a comment or response. Keyword monitoring also allows staffers to gather a library of clips related to different issues.



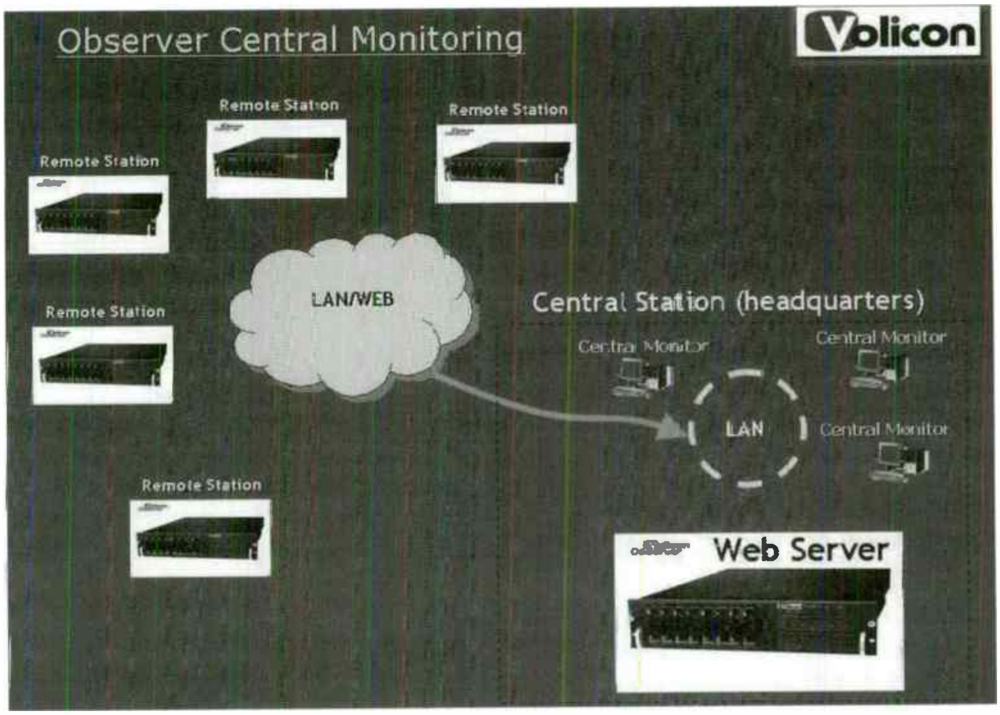
Close-captioning search results with thumbnails, descriptions and content links

Centralized Monitoring

For those networks or station groups dealing with channels all over the country, centralized monitoring allows for detailed analysis of multiple channels, all from one location. Highly scalable monitoring systems can be used both at the individual station level and as part of a centralized monitoring scheme that leverages digital compression technologies and the Internet to provide simultaneous and instant access to video and audio being recorded remotely. Thus, quality of service can be monitored all across a group of stations, with the same access to alerts and fault detection, to gain a full picture of the network broadcast.

Global monitoring capability from a single location makes it easier to identify whether a problem is unique to a particular station, or if it's a recurring issue across stations that must be solved at a higher level. The sharing of media in real time also simplifies analysis because it allows all parties involved to see what's actually going on when there is a problem. A picture is worth a thousand words, and a visual demonstration of the problem often gives technical staff the details they need to help resolve the situation.

Centralized monitoring systems can also be used for evaluation of content across all station groups and inform decision-making at the managerial and executive level, providing a valuable tool for competitive analysis both internally an in comparison to competing networks or stations.



Typical architecture for Central Monitoring

LOOKING AHEAD

Because digital video technologies are evolving every day, systems such as the Observer grow even more powerful and robust in providing tools for improving the on-air product and, ultimately, adding to the bottom line. For example, future refinements of the system will enable broadcasters to locate not only sponsored elements and traditional ad spots, but also the presence and duration of corporate logos on-screen. Broadcasters will be able to verify content automatically for purposes of license validation and for detection of information such as emergency alert system (EAS) updates. Use of digital content opens the door to many other enhancements to ad verification and broadcast analysis.

The shift toward monitoring systems based on digital video and data offers more than convenience; it provides a new way of looking at and evaluating aired content. A few of the broadcasters still relying on tape-based monitoring may be hesitant to shake things up by introducing a new system into their operations and disturbing existing workflows. More and more broadcasters are aware, however, that once they get a digital system in place, they can take advantage of it in so many different ways just not possible with a tape-based monitoring model.

As the industry continues to adopt advanced monitoring solutions, the trend resembles the earlier adoption of email or the cell phone. The first few adopters understood the potential of the new technology, and subsequent adopters in greater numbers found that they couldn't survive without it. The competitive edge delivered by the Observer allows broadcasters to make significant improvements to their broadcasts, ratings, and ad revenues. Once one broadcaster in a market implements this powerful tool, competitors likely will find it to be a critical addition to their operations, as well.

Digital monitoring systems such as the Volicon Observer continue to evolve in functionality and, as they are developed further, in the value-added features they provide. With these developments, broadcasters with advanced digital monitoring systems in place will gain the opportunity to extend the use of the system for even more streamlined, cost-efficient, and competitive broadcast operations.

Audio Solutions for Radio

Monday, April 16, 2007
1:00 PM – 5:00 PM

Chairperson: Talmage Ball
Bonneville International, Salt Lake City, UT

High Quality Speech Synthesis System Using Speech Rate Conversion for Stock-price Bulletins

Hiroyuki Segi, NHK (Japan Broadcasting Corporation), Tokyo, JAPAN

Reliable Audio Transport over IP Networks

Bob Band, Harris Corporation, Mason, OH
Josh Sparks, VT Communications, London, United Kingdom

Sonic Tonic for Audio Coding

Frank Foti, Omnia Audio, Cleveland, OH

*Digital Radio - Audio & Beyond

Rich Redmond, Harris Corporation, Mason, OH

Audio Compression – The Trade Offs Between Quality, Data Reduction and Latency

Guy Gampell, APT, Belfast, Northern Ireland

Radio Remotes Using New Satellite Technology

Paul Shulins, Greater Media, Boston, MA

IBOC Analog-Digital Blend Time Alignment Quality

Michael Bergman, Kenwood, Princeton Junction, NJ

---

*Paper not available at the time of publication

High Quality Speech Synthesis System Using Speech Rate Conversion for Stock-price Bulletins

HIROYUKI SEGI, NOBUMASA SEIYAMA, REIKO TAKO, TOHRU TAKAGI, HIDEO TODA and RYUJI KOYAMA
Japan Broadcasting Corporation (NHK)
Tokyo, Japan

ABSTRACT

We have been developing a high quality speech synthesis system that realizes natural-sounding speech through the use of a massive speech database, consisting of recordings of an announcer reading from past news programs. The system selects the combination with the smoothest acoustic flow of words and phrases to compose an entire sentence based on the database.
As a practical application we have developed a prototype voice synthesizer for stock-price bulletins, using numerical speech synthesis and automatic speech rate conversion. In the stock market report program broadcast on NHK Radio 2, the closing prices and change on the day of about 830 stocks listed on the Tokyo Stock Exchange are read out over 45 minutes. It is hard for announcers to read numerical values without making mistakes and to pace themselves to fit the allotted broadcast time exactly.
In the case of automation, synthesizing natural-sounding numerical speech as a compilation of recorded elements is difficult, and recording all the combinations of possible patterns of utterance in advance is impossible because there are prohibitively many. However, by modeling the effects of co-articulation in the processing of the voice waveform, synthesized speech can be created that approaches the natural sound of the real thing. Additionally applying automatic speech rate conversion enables the synthesized program to be fitted precisely into the allotted broadcast time frame.
Our prototype system is currently used in the experimental broadcasting of digital terrestrial radio.

1. INTRODUCTION

NHK Science and Technical Research Laboratories (STRL) is researching a high quality speech synthesis system to achieve text-to-speech conversion (TTS) that sounds more natural. If a broadcast station were to transmit supplementary information such as pronunciation, accent, and intonation in addition to text in data broadcasts, receivers would be capable of performing high quality speech synthesis superior to that of conventional systems. This would enable the visually impaired, automobile drivers, and others to listen to data broadcasts with natural sounding synthesized speech [1].
Additionally, the creation of natural synthesized speech from original news copy, announcements, or stock-price bulletins on the broadcast-station side would enable automatic radio broadcasts, and combining this with a technology like TVML [2] that automatically produces computer-generated images from a script could produce automatic TV broadcasts.

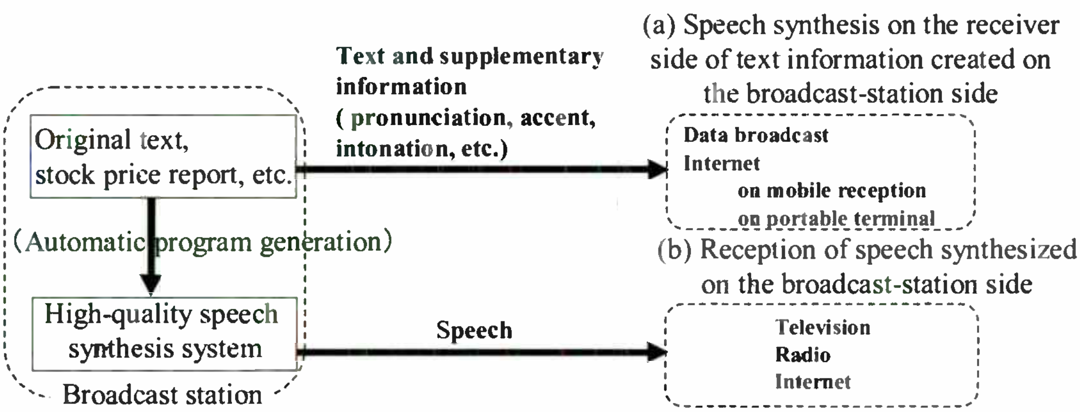


Figure 1. A block diagram of speech synthesis services

Figure 1 shows a block diagram of such a speech synthesis service.

This paper describes a high quality speech synthesis system developed at STRL for the above purpose and a speech synthesis system for stock-price bulletins that aims to achieve an even higher level of quality by limiting the task to be performed.

2. HIGH QUALITY SPEECH SYNTHESIS SYSTEM

2.1 Outline of high quality speech synthesis system

A number of concatenative speech synthesis systems have been proposed that select and concatenate voice waveforms having smooth connections from a large speech database [3-6]. The sound quality obtained with this approach, while relatively good, is not sufficient for broadcast applications. One way to improve quality here is to make the speech database even larger, but as the search units used for voice waveforms are usually short units such as syllables and phonemes (vowels and consonants), the number of voice waveform combinations for composing a sentence can be huge requiring substantial approximations to be made if a solution is to be found in a realistic period of time.

In response to this problem, the proposed high quality speech synthesis system concatenates voice waveform data using units longer than syllables, namely, "phoneme sequences" (strings of phonemes) that appear with high frequency in the speech database. This approach significantly decreases combinatory calculations enabling a large speech database to be used and sound quality to be improved.

Figure 2 shows the operation procedure in this high quality speech synthesis system. The following describes its process flow referring to the numbers indicated in the figure.

(1) On inputting text to be synthesized for audio output, the speech synthesis system automatically generates a string of phonemes by analyzing the text and determining pronunciation with respect to the mixture of Chinese characters and Japanese phonetic

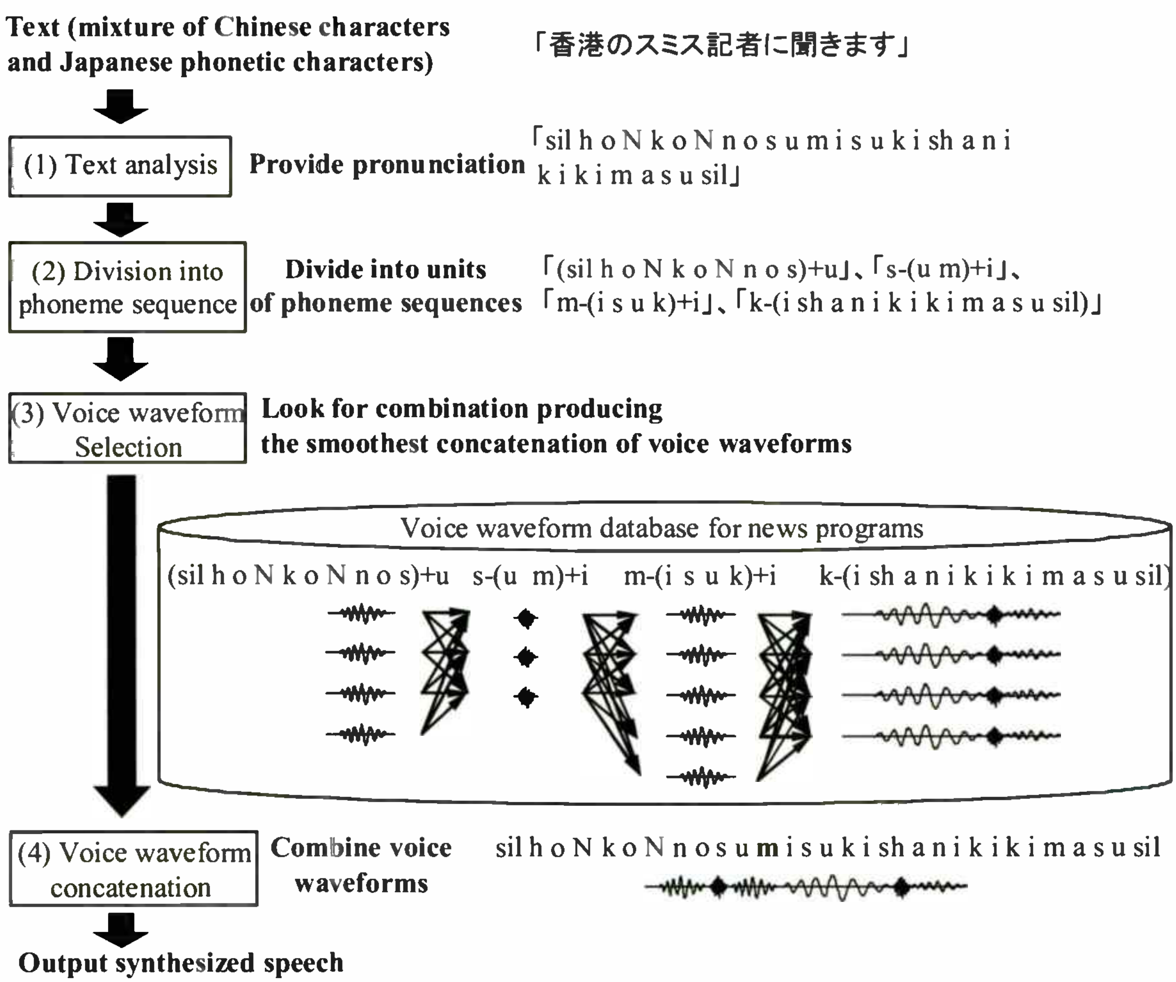


Figure 2. Operation procedure of high quality speech synthesis system

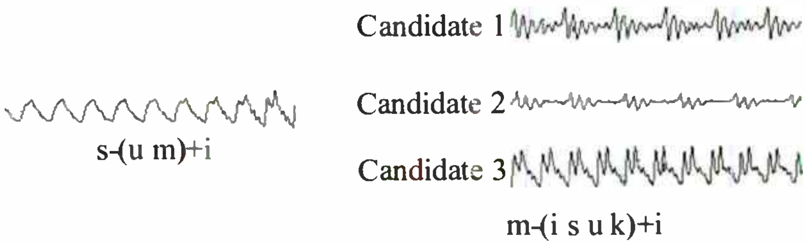


Figure 3. Selection of similar waveforms

characters in the sentence. For example, given the Japanese text "香港のスミス記者に聞きます", which means "Mr. Smith reporting from Hong Kong", the string of phonemes obtained from text analysis is "sil h o N k o N n o s u m i s u k i sh a n i k i k i m a s u sil".

(2) The input text so denoted by phonemes is now divided into phoneme sequences that appear frequently in the speech database. The reason for this is as follows. If there are only a few voice waveform candidates that configure the same phoneme sequence at the time of speech synthesis, smooth speech connections are less probable and speech deterioration in that portion of the text is more likely. The division of text into phoneme sequences aims to minimize the number of divisions and to make the length of each division as uniform as possible. The method used here for determining phoneme sequences that appear with high frequency in the speech database is described in section 2.2. The output from the text analysis section in this case may be "(sil h o N k o N n o s)+u", "s-(u m)+i", "m-(i s u k)+i", and "k-(i sh a n i k i k i m a s u sil)". Here, the notation "s-(u m)+i" means the phoneme sequence "(u m)" preceded by phoneme "s" and followed by phoneme "i".

(3) Next, the system searches for all combinations of voice waveforms for the phoneme sequences obtained from text analysis and selects the combination for which the acoustical features of consecutive voice waveforms are as similar as possible. If concatenation is not performed with similar voice waveforms, synthesized speech can be unnatural in pitch and sound quality. Here, the systems determines the degree of similarity between consecutive voice waveforms using the similarity of fundamental frequency expressing pitch (repeated period of a waveform) and the similarity of cepstrum, a feature quantity of the spectral envelope expressing the content (phonemes) of an utterance. Figure 3 shows a voice waveform for the phoneme sequence "s-(u m)+i" and three voice-waveform candidates for the subsequent phoneme sequence "m-(i s u k)+i". Candidates 1 and 2 have different fundamental frequencies and cepstrum than that of "s-(u m)+i" and consequently exhibit low similarity. Candidate 3 is most similar.

(4) Finally, the system concatenates the voice waveforms for which the degree of similarity from the beginning to the end of the original text is largest and outputs the results as synthesized speech.

2.2 Creation of phoneme sequence list

Before performing speech synthesis, a list of phoneme sequences that appear with high frequency in the speech database must be prepared. To this end, we decide beforehand on a minimally required number of appearances and use only phoneme sequences having a number of appearances equal to or greater than that value. Here, we treat all phoneme sequences having different previous and subsequent phonemes as different phoneme sequences. On applying this technique to the speech database described below having about 30,000 sentences, we generated about 40,000 phoneme sequences. Of these, the longest phoneme sequence at 55 phonemes turned out to be "(sil i Q p o- t o- ky o- k a b u sh i k i sh i j o- sh u y o- m e- g a r a n o h e- k i N k a b u k a n o o w a r i n e w a)+sp", which means "The closing stock price average of the major companies on the Tokyo stock exchange". This shows how a set expression stored in the speech database can be searched for at one time. In the above phoneme sequence, "sil" refers to silence at the beginning and end of a sentence and "sp" refers to short pause within a sentence.

2.3 Evaluation of synthesized speech

We performed a subjective quality-evaluation test on a 5-point scale to assess the naturalness of synthesized speech created by the proposed method.

To build a speech database for the test, we used 27,888 sentences running a total of 86 hours with little background noise from a news program broadcast from June 3, 1996 to June 22, 2001 [7]. These sentences, which were spoken by the same female announcer, consisted of 3,840,000 phonemes in total. Because speech to be submitted to the broadcast station is digitally recorded at 16-kHz/16-bits, there was no transmission-related noise to deal with, but speech from other people conversing with the announcer as well as sound effects happened to be mixed in. The portion of speech spoken only by the announcer, which has relatively little noise, was manually extracted. After this processing, we divided the recorded speech into units of sentences and described the utterances made with a mixture of Chinese characters and Japanese phonetic characters in speech files. The temporal correspondence between the recorded speech and utterance descriptions was established automatically and fundamental-frequency values were also extracted automatically [8].

The text used for evaluation consisted of 40 sentences (1,444 words and 5,927 phonemes) not included in the speech database. Four types of speech databases, 11 hours, 22 hours, 43 hours, and 86 hours of speech extracted from the total 86-hour database, were prepared to synthesize speech for evaluation. A total of 200 test samples consisting of 160 samples of synthesized speech by the proposed method and 40 samples of natural speech were prepared against the above 40 sentences used as evaluation text. These test samples were tested for quality by subjective evaluation on a 5-point scale.

To conduct the test, we set up a loudspeaker in a soundproof room and asked subjects consisting of five males and four females to listen to the speech samples. For each trial, we presented test samples in random order and had the subject evaluate the presented sound in terms of perceived naturalness on a scale from 1 to 5, with 5 being "natural," 4 "not natural but negligible," 3 "slightly noticeable," 2 "noticeable," and 1 "very noticeable." Before beginning an evaluation, we had the subject listen to three sentences of speech from the speech database and instructed the subject to consider the degree of naturalness of that speech to be level 5 (natural).

[Test Results]

Figure 4 shows the distribution of scores for each of the speech database sizes described above. In these results, 98.6% of all scores for natural speech gave it a 5 (natural), which could not even be plotted in the graph shown. Total score for natural speech was 4.99. The average score for synthesized speech by the 86-hour speech database was 3.94 indicating that synthesized speech with a naturalness perceived as "not natural but negligible" could be obtained.

In the score distribution for the 86-hour speech database, synthesized speech received a score of 5 (natural) 36.7% of the time indicating that synthesized speech with a level of quality the same as natural speech could be prepared quite frequently. But for broadcasting purposes, the quality of synthesized speech should be uniform regardless of synthesized content. It is important that a means be found to improve synthesized speech that receives a score of 2 or 1 using the 86-hour speech database (which, in this case, is 11.1% of total score).

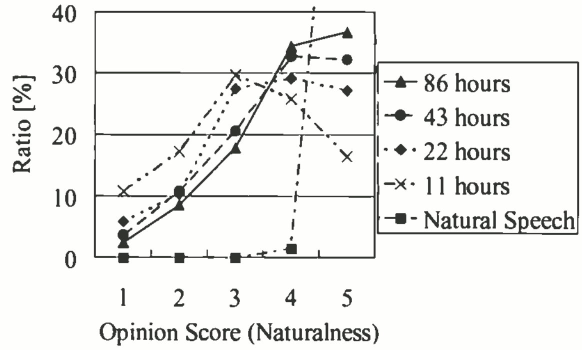


Figure 4. Histogram for various speech databases

3. SPEECH SYNTHESIS SYSTEM FOR STOCK-PRICE BULLETINS

3.1 "Stock Market" program and speech synthesis

As shown in section 2.3, the naturalness of synthesized speech for arbitrary text is quite good but insufficient for broadcasting. Considering that limiting the content for speech synthesis to numbers might enhance the quality of synthesized speech, we attempted to apply our speech synthesis method to NHK's "Stock Market" program.

This program, which is broadcast on NHK Radio 2, reads out, over a period of 45 minutes, the closing prices and change from the previous day's price for about 830 stocks from among the 1500 or so stock issues listed on the First Section of the Tokyo Stock Exchange. Only announcers with extensive experience can handle this program because they must speak as quickly as possible to fit all of the information in the allotted time. Because of the difficulty for a single announcer to read out this information continuously for 45 minutes, this is performed by two announcers that take turns during the program. Nevertheless, this is a difficult task for announcers as they are required to 1) read out numbers without mistakes and 2) adjust their reading speed to ensure that all information fits into the allotted broadcast time frame.

As for requirement 1) above, the possibility of making a mistake when reading out numbers cannot be made zero for humans but can be for machines. For the latter, speech synthesis would be required since the recording of each number up to the current maximum of several million yen would be prohibitively time consuming and expensive. The system described in section 2 could be used for speech synthesis, but even though it achieves more natural than conventional systems, it is still sufficient for broadcasting quality. Accordingly, with the aim of achieving a speech synthesis system that can satisfy broadcast quality, we concatenate voice waveforms assuming concatenation

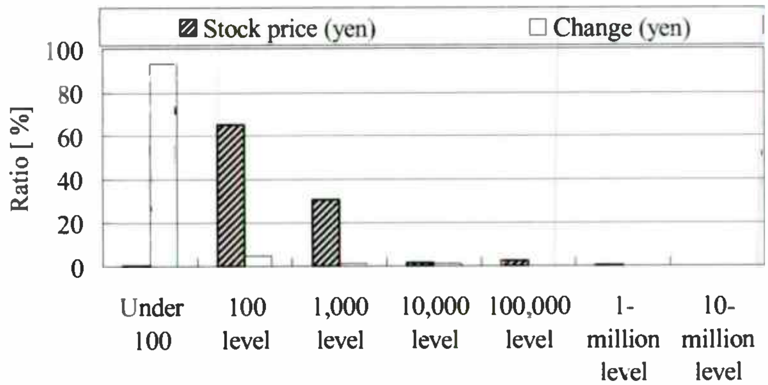


Figure 5. Histogram for the closing prices and movements on May 6th in 2004

in units of numerical digits (place values) taking coarticulation, which express the temporal movement of phonemes, into account.

For requirement 2), the time required for playback can be calculated at the time of receiving stock price data in the case of a machine. This means that time-related adjustments can be made to some extent by inserting equivalent intervals of silence between neighboring instances of speech. But this technique can adjust time only so much. Furthermore, recording original speech at the speech rate required by the "Stock Market" program is difficult taking reading mistakes and reading stability into account. In this study, we propose expanding the range for which time can be adjusted by varying the length of vocalized sections using a speech-rate conversion technique.

Taking the above two points into account, we constructed a prototype speech synthesis system for stock-price bulletins. This prototype system has been operating on a test basis every Friday since October 2006 as part of experimental broadcasting of digital terrestrial radio.

3.2 Basic units of speech synthesis

One way of creating a voice waveform database here would be to use actual broadcast speech as described in section 2. However, utterances in the "Stock Market" program tend to be unstable because of rapid speaking and a bias in the distribution of numbers indicating stock prices in broadcast speech. Figure 5 shows this bias. For these reasons, we created the voice waveform database for our prototype speech synthesis system for stock-price bulletins by asking an announcer to read out sentences having no bias in numbers.

Given that the purpose here is to synthesize numbers, the basic database units are taken to be the digits (place values) of a number (under 100, 100, 1000, 10,000, 100,000, 1,000,000, and 10,000,000) as opposed to phoneme sequences as described in section 2. At the same time, using simply digits would degrade the naturalness of synthesized speech since silence cannot be inserted between digits, which means that coarticulation must be taken into account. Here, it would be desirable to use "digits taking the previous and subsequent digits into account," but this would result in about 40,000 basic units presenting another problem. We therefore treat different combinations of a digit with previous/subsequent digits that satisfy the following conditions as the same digit in a process called "clustering." In short, we use "digits taking previous and subsequent clusterized digits into account."

(1) Combinations of a digit with a previous digit that end in "ju", the Japanese word for "10" (namely, ju (10), niju (20), sanju (30)...kyuju (90)) are clustered together as the same digit. The same applies to combinations of a digit with a previous digit that end in "yaku", the Japanese word for "100" (hyaku (100), nihyaku (200), sanbyaku (300)...kyuhyaku (900)), that ends in "en", the Japanese word for "1000" (i'sen (1000), nisen (2000), sanzen (3000)...kyusen (9000)), that ends in "man", the Japanese word for "10,000" (ichiman (10,000), niman (20,000, sanman (30,000)...kyuman (90,000)), and so on. In the case of English, this rule is almost the same as Japanese. Combinations of a digit with a previous digit that end in "ty", the English suffix for "10" (namely, twenty (20), thirty (30), forty (40)...ninety (90)) are clustered together as the same digit. The same applies to combinations of a digit with a previous digit that end in "hundred", the English word for "100" (one hundred (100), two hundred (200), three hundred (300)...nine hundred (900)), that ends in "thousand", the English word for "1000" (one thousand (1000), two thousand (2000), three thousand (3000)...nine thousand (9000)), and so on.

(2) Combinations of a digit with a subsequent digit that begins with "ju" or 10 (juichi (11), juni (12), jusan (13)...jukyu (19)) are clustered together as the

same digit. The same applies to combinations of a digit with a subsequent digit that begins with "niju" (niju (20), nijuichi (21), nijuni (22), nijusan (23)...nijukyu (29)), and so on for sanju (30), yonju (40), goju (50), rokuju (60), nanaju (70), hachiju (80), and kyuju (90). In the case of English, this rule is again almost the same as Japanese. Combinations of a digit with a subsequent digit that begins with "twenty" or 20 (twenty (20), twenty-one (21), twenty-two (22), twenty-three (23)...twenty-nine (29)) are clustered together as the same digit, and so on for thirty (30), forty (40), fifty (50), sixty (60), seventy (70), eighty (80), and ninety (90).

With the above type of clustering, only 3033 basic units are needed to include the integers from 1 to 99,999,999.

3.3 Evaluation of synthesized speech

We conducted a subjective quality-evaluation test on a 5-point scale to assess the naturalness of synthesized speech by the proposed method.

The speech database was prepared as follows. First, for stock prices, we had text read out that included all of the basic units for stock-price speech synthesis as described in section 3.2 so as to enable the synthesis of integers from 1 to 99,999,999. This text consisted of 2,106 integers suffixed with the word "yen". Next, for change in stock price with respect to the previous day, we had text read out that included all of the basic units for speech synthesis of integers from 1 to 9,999. Stock prices and changes were not read out and recorded independently of each other. Rather, price and change for each issue were recorded in one utterance starting with the issue name followed by stock price and change in the format "X company, 15,632 yen, up 217 yen". The person reading out the text was a male announcer, who was also in charge of the "Stock Market" program. Recordings were made in a soundproof room, and on completion, recorded speech was manually divided into issue name, stock-price, and change portions.

For the evaluation, considering that a random selection of numbers from 1 to 99,999,999 could result in only huge numbers in the range of 10,000,000 and up, 40 numbers were selected randomly for each place value from numbers not in the speech database. Given the six place values of 100, 1000, 10,000, 100,000, 1,000,000, and 10,000,000, this comes out to 240 numbers. A set of 240 numbers for price changes was selected in the same way, and adding in 40 stock prices and 40 price changes in natural speech, a total of 560 test samples were prepared for evaluation.

To conduct the test, we asked subjects consisting of three females to listen to the samples. The evaluation conditions were otherwise the same as described in section 2.3.

[Test Results]

Figure 6 shows test results. It can be seen that synthesized speech achieved a score of either "natural" or "not natural but negligible" regardless of the place value. Comparing this result with the score of 4.97 achieved by natural speech, the naturalness of synthesized speech is slightly inferior to that of natural speech, but it can still be seen that the naturalness of synthesized speech created by the proposed method is relatively high. These results also reveal that naturalness degrades for higher place values. This is because more speech data is required for numbers with more digits, which means more locations where speech data must be combined.

3.4 Outline of a speech synthesis system for experimental broadcasting of digital terrestrial radio

Figure 7 shows a prototype speech synthesis system for stock-price bulletins that we constructed for experimental broadcasting of digital terrestrial radio using the proposed speech synthesis method. This system consists of an FTP server that receives stock

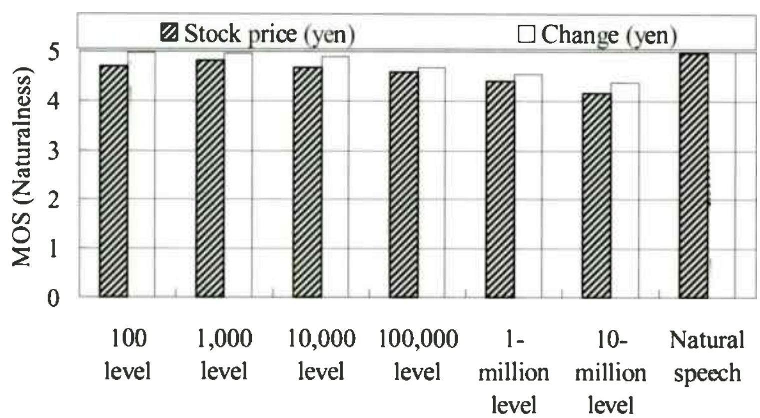


Figure 6. MOS for speech samples synthesized by proposed method

price data from the outside and a speech-synthesis server that runs speech-data production and broadcast applications. The copy to be read out on the "Stock Market" program broadcast on NHK's Radio 2 is currently printed out by a printer connected via a leased line (LAN) to a news agency that provides the stock price data. In the proposed system, we connect the FTP server to the above LAN to enable stock price data to be obtained electronically.

The FTP server and speech-synthesis server are capable of file sharing, which enables the speech-data production application running on the speech-synthesis server to automatically detect the receiving of stock price data from the news agency at the FTP server. On detecting this data, the speech-synthesis server can proceed to create a speech file based on that stock price data. This speech file includes speech for the issue name, stock price, and price change as given in the received data. The speech-synthesis server may also create a speech file at the time of the speech-data broadcast application, but creating the file beforehand avoids the chance occurrence of any file I/O errors during broadcasting. Here, the speech for the issue name, which is able to be synthesized by the method described in Chapter 2, is simply recorded speech since speech of broadcast-quality level cannot be synthesized for it. The speech for stock prices and price changes are, of course, synthesized speech.

The speech-data broadcast application running on the speech-synthesis server can automatically re-input a speech file in the event that the speech file output from the speech-data production application is updated. Referring to Figure 7, it can be seen that the entire system has a redundant configuration. This prevents a breakdown of a LAN circuit or server on one side from impacting a live broadcast.

In current operation, stock price data is forwarded from the news agency at 3:30 in the afternoon. The speech-data production application begins to create a speech file at this time. Since it takes about two minutes to create a file, the server enters a broadcast-ready state from 3:32 on. The program director then performs a 3-point check, and after checking for any subsequent updates, begins to broadcast the data at 5:05.

When broadcast time arrives, the speech-data broadcast application automatically begins to transmit the speech file that it has read in. At this time, the application plays back the speech data while continuously adjusting speech rate so that the reports for 830 stock issues can be presented within the allotted broadcast time frame. In this way, the system can complete the reading out of all data within the specified time even if the playback for certain issues were to be canceled or temporarily stopped or if completion time were to be changed along the way.

3.5 Speech rate conversion

In the prototype speech synthesis system for stock-price bulletins introduced above, recorded utterances are slower than broadcast ones. As a result, total time of synthesized speech for broadcasting is greater than program frame time. The system therefore performs speech rate conversion to speed up synthesized speech before output. When speeding up speech, the system culls speech in units of fundamental periods corresponding to periodic repetition as in spoken vowels and semivowels and in units of pseudo fundamental periods corresponding to consonants and silence as depicted in Figure 8. This mechanism enables speech rate to be converted without a loss in pitch or quality. When slowing down speech, the system inserts voice waveforms in units of fundamental periods instead of performing culling.

This speech synthesis system for stock-price bulletins with speech-rate control makes listening easier by making the beginning of each issue name, stock-price, and price-change utterance slower than its second half [9]. Though the rate of speech rate conversion is dependent on the program time frame, stock price data,

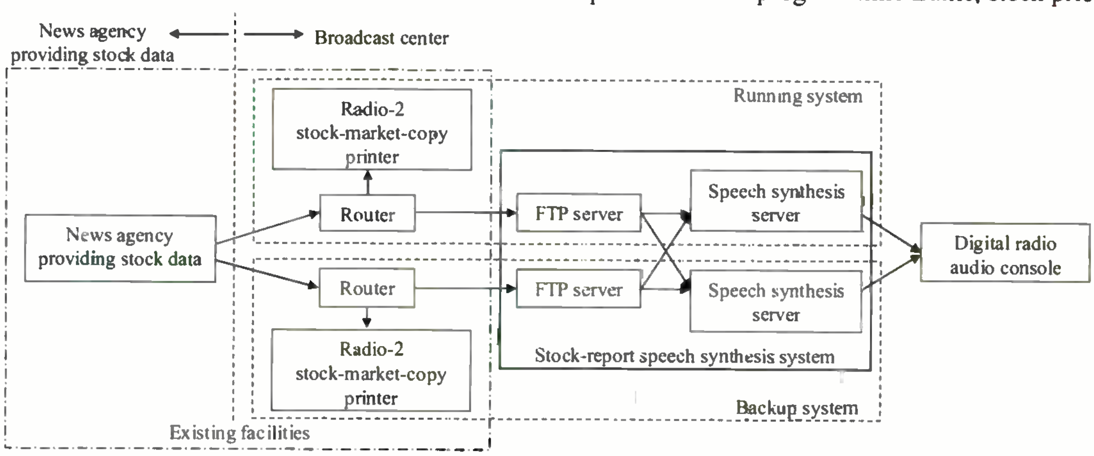


Figure 7. Stock speech synthesis system for digital terrestrial radio

and number of stock issues to report, this speech synthesis system for stock-price bulletins used for digital terrestrial radio shortens the original voice waveform by about 16%.

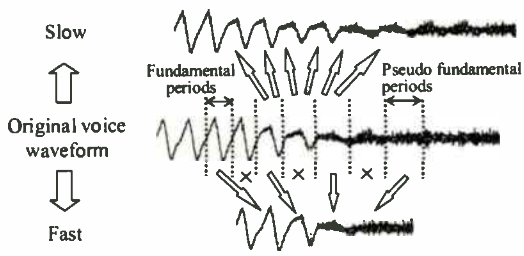


Figure 8. The conceptual figure of speech rate conversion

4. CONCLUSION

We described a high quality speech synthesis system and a speech synthesis system for stock-price bulletins and described how we obtained synthesized speech with a high degree of naturalness by enhancing the former system to achieve the latter. Though the effect of speech rate conversion has not yet been evaluated, user's found the resulting speech to be relaxed and unhurried compared to an announcer's reading speed as there is no need in a speech synthesis system with speech rate conversion to take breaths or to clear one's throat.

In future research, we will address the issue of synthesizing speech for the names of stock issue. Stock-issue names can undergo a change in the event of a company's name change, a merger, or a new listing, which occurs at a rate of about once every two months. In the present system, we ask the same announcer who reads out the text for the speech database to record required speech when such a change occurs. To prevent this inconvenience, we look forward to the development of speech synthesis technology for issue names.

5. REFERENCES

[1] K. Matsumura, K. Kai, H. Hamada and N. Yagi, "Transforming Data Broadcast Contents to Fit Different User Interfaces Generating a Readout Service for Mobile DTV Receiver," *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services [MobileHCI 05]*, ACM, p323-324, Sept. 2005.

[2] M. Hayashi, M. Douke, and N. Hamaguchi, "TV-like Presentation for Agent System," *13th IEEE International Workshop on Robot and Human Interactive Communication [RO-MAN 2004]*, 2B1, Sept. 2004.

[3] A. Black and N. Campbell, "Optimizing selection of units from speech database for concatenative synthesis," *Proc. of EUROSPEECH'95*, vol. 1, pp. 581-584, Sept. 1995.

[4] A. J. Hunt and A. W. Black, "Unit selection in a concatenative speech synthesis system using a large speech database," *Proc. of ICASSP*, vol. 1, pp. 373-376, 1996.

[5] A. Conkie, "A robust unit selection system for speech synthesis," *Proc. 137th meet. ASA/Forum Acusticum*, 1999.

[6] T. Toda, H. Kawai, M. Tsuzaki, and K. Shikano, "Unit selection algorithm for Japanese speech synthesis based on both phoneme unit and diphone unit," *Proc. of ICASSP*, vol. 1, pp. 465-468, May 2002.

[7] A. Ando, T. Imai, A. Kobayashi, S. Homma, J. Goto, N. Seiyama, T. Mishima, T. Kobayakawa, S. Sato, K. Onoe, H. Segi, A. Imai, A. Matsui, A. Nakamura, H. Tanaka, T. Takagi, E. Miyasaka and H. Isono, "Simultaneous Subtitling System for Broadcast News Programs with a Speech Recognizer", *IEICE Transactions on Information and Systems*, vol.E86-D, no.1, p.15-25, Jan. 2003.

[8] S. J. Young, G. Evermann, T. Hain, D. Kershaw, G. L. Moore, J. J. Odell, D. Ollason, D. Povey, V. Valtchev, and P. C. Woodland, "The HTK BOOK," 2002.

[9] N. Seiyama, A. Nakamura, A. Imai, T. Takagi, and E. Miyasaka, "Portabel Speech Rate Conversion System," *Proc. of EUROSPEECH'95*, vol. 1, pp. 1717-1720, Sept. 1995.

RELIABLE AUDIO TRANSPORT OVER IP NETWORKS

Bob Band
Harris Corporation
Mason, Ohio

Josh Sparks
VT Communications
London, UK

INTRODUCTION

The broadcast world is moving inexorably toward the use of IP networks for real-time media transport. In this paper we address the use of IP networks for reliable, professional audio transport for STLs, studio-to-studio links, and other 24/7 real-time applications. We cover such questions such as: What types of networks are suitable? What is required of the equipment to operate successfully in the packet environment? What can users expect in terms of bandwidth usage and delay?

We look briefly at the various types of hardware and software tools available to ensure optimum performance on IP audio networks, such as QoS tagging, FEC, packet size, and jitter buffering.

Beyond the theoretical, we also examine in detail the real–world example of VT Communications' recent upgrade of its worldwide audio contribution and distribution network from a combination of services to IP multicast.

The Challenges of the IP Environment

Internet Protocol (IP) networks are packet-based networks, originally designed to work as shared resources for non-real-time data transport in such applications as file transfers, e-mail and Web browsing. The shared aspect of these networks provides highly efficient use of the available bandwidth, and offers considerable cost savings to users who are able to place all their communication services over a single link.

However, it also presents a challenge in the world of real-time audio delivery, as ways must be found of ensuring that the audio packets arrive with minimal delay, in the correct sequence, and with no packets missing.

Before placing audio onto an IP network it must be packetized, or broken up into small segments and placed into packets. Along with user data, each packet contains several pieces of housekeeping data such as the destination address, a sequential packet ID, and its Quality of Service (QoS) information.

Proper QoS tagging is critical to ensuring reliable delivery of real-time audio. Every IP packet has a byte in its header reserved for information about its priority in the network, referred to as the Type of Service (ToS) byte. All IP networks support some form of prioritization, with the most advanced using a system called DiffServ, or Differential Services. At the highest level are packets containing information critical to the operation of the network itself. Just below that are packets containing time-critical information, such as real-time audio or video, or VoIP telephony. Further down the priority chain are packets containing data from non-real-time applications like file transfers and Web pages. Depending on the network, as many as seven different levels of prioritization may be used.

Once the data is packetized, the packets are given to the network for delivery. In the network, the packets are sent through to their destination via one or more switches and routers. Each device looks at each packet, makes a decision on the best route available at that moment in time, and sends the packet on its way. Gateways, which allow packets to move between networks using different protocols, can easily become sources of short-term congestion, as they must handle all packets moving between the networks involved.

For the audio to be successfully decoded at the far end, it is essential that all the packets reach their destination, and that they are played out in the same sequence as they originated. For this to happen, we have to address two inherent challenges presented by IP networks: jitter and packet loss.

Jitter is a condition where packets arrive at irregular intervals, or even out of sequence altogether. It occurs because there is no way to pre-determine the exact route any given packet will take through the network. As a packet reaches each switch, that device sends it on to the next one based on constantly-updated information about the congestion conditions on the various possible paths. Thus, sequential packets may be sent along different routes.

While this is one of the great strengths of IP networks, ensuring that packets can get through to their destination by multiple possible routes and thus eliminating the possibility of link failure due to a single node failure, it also mean that routes taken by adjacent packets may vary considerably. As the routes vary, so too will arrival times.

Differences in arrival time translate into packet jitter, and as packet jitter increases, packets may begin to arrive out of order. To cope with this, a well-designed IP audio transport system should have a receive jitter buffer deep enough to hold the incoming audio packets until any delayed ones have arrived, so they can be played out in proper sequence. While in many cases a jitter buffer of less than 100 ms depth is adequate, at times the buffer must be set to as much as one second in order to cope with actual network conditions.

Packet loss can occur in several ways. One relates to the foregoing discussion of jitter. If the jitter is so high that a packet arrives too late to be inserted back in its proper place in the jitter buffer before playout, then it is essentially lost. Another is that IP networks are designed such that if a switch along the way becomes congested, packets may be dropped as needed to reduce the congestion.

QoS tagging is the primary tool for ensuring that real-time audio packets are allowed through even when lower-priority packets are dropped, but it alone cannot prevent excessive congestion from causing problems. Accordingly, it is important that adequate bandwidth to provision all real-time streams be considered when designing and operating IP networks.

Despite everyone's best efforts, there may still be times when for one reason or another some packets do not arrive at their destination. One way to deal with this is through the use of Forward Error Correction (FEC). FEC rearranges the transmitted bits and adds redundant data such that a missing packet can be reconstructed based on nearby packets. The amount of FEC needed depends on the conditions in the network (especially the last mile), and an analysis of traffic over time, but a usable FEC scheme typically adds 50% to 100% to the audio bandwidth.

The right amount of FEC can take a reasonable network and make it quite robust. On the other hand, too much FEC is a waste of bandwidth, as increased FEC levels mean increased overhead. Also, when packets are being dropped due to limited bandwidth on a given link, the additional bandwidth required by FEC can exacerbate rather than alleviate the problem.

An ideal system allows the user to adjust the amount of FEC on each stream to best suit the actual conditions, and to turn it off altogether when it is inappropriate.

Bandwidth Requirements

When planning an IP audio transport network, a critical question is, how much bandwidth do you need? The answer depends on five factors:

1. The frequency range of the audio to be transported.
2. The sample size and coding rate used in digitizing the audio.
3. The type of digital compression (if any) being applied
4. The amount of packetization overhead.
5. The amount of FEC (if any) being added.

For example, let's look at a typical FM broadcast program. The audio starts out in the analog realm as two 15 kHz channels. Sampling this at 32 kHz (following the Nyquist theorem that states we must sample at a frequency at least twice that of the highest audio frequency), using 16-bit samples, we get 16 x 32,000 = 512,000 bits per channel per second, or a total of just over 1 Mbps for stereo. Audio with a wider frequency range, such as 20 kHz contribution feeds, will require proportionally more bandwidth; lower quality audio, such as 7 kHz or 10 kHz mono signals suitable for AM radio, proportionally less.

Audio compression can take this 1 Mbps and reduce it considerably. Many options exist, from the relatively mild companding of the J.41 algorithm which reduces that 1 Mbps to 768 kbps, to the various forms of MPEG compression (MPEG II, MPEG III, AAC, etc), to the proprietary options such as apt-X-100, Enhanced apt, and Dolby AC-3. Which of these to choose is itself dependent on several factors, which shall remain outside the scope of this paper. But among them they are capable of reducing the 1 Mbps of stereo audio to 384 kbps, 256 kbps, 128 kbps, or even lower, while retaining acceptable audio quality for most practical purposes.

At the conclusion of this compression stage we have the "raw" digital audio bitstream, but to put it onto the IP network, we have to break up this bitstream into segments and package each segment in an IP packet, which adds about 45 bytes of metadata as described above. If you put a large amount of audio data into each packet - 1000 bytes, say - then the overhead is a relatively small proportion of the total: 45 ÷ 1045 = < 5%. On the other hand, if you use smaller packets, then the proportion of overhead to audio grows larger, and can become a significant factor in the total bandwidth requirement.

The question of packet size is worth a careful look, because it is in effect a tradeoff between latency (delay) and bandwidth usage. The fewer audio bits in each packet, the faster the packet fills up and is ready to send, minimizing packetization delay but maximizing bandwidth; the more audio bits per packet, the greater the delay, but bandwidth usage is minimized.

Finally, FEC, if you decide to use it, can as much as double the final packetized bandwidth.

So, what happened to our original 1 Mbps of stereo audio? At one extreme (no compression, small packets, high FEC) it can become as much as 4 Mbps; at the other (high compression, larger packets, no FEC) it shrinks to less than 100 kbps.

The key is to choose the options that best suit your audio quality requirements and network bandwidth availability.

Delays in an IP Audio Network

Another question frequently asked is, how much delay should I expect when I send audio across an IP network? Here again there are several factors involved:

1. The choice of audio compression.
2. Packetization delay.
3. Actual network throughput delay.
4. Jitter buffering.

Uncompressed or J.41 audio incurs very little delay in the digitization process, normally less than 5 ms. Some compression methods, like the apt algorithms and other ADPCM-based processes, can be nearly as fast, with delays in the 5 ms to 10 ms range. Dolby and MPEG algorithms, on the other hand, can add hundreds of milliseconds of delay.

Packetization delay can range from below 3 ms to over 100 ms. As noted above, the delay incurred here is inversely proportional to the increase in bandwidth caused by additional packetization overhead.

Network throughput delay can vary widely, but typical figures on well-run networks with adequate bandwidth are in the range of 10 ms to 50 ms.

Finally the depth of the jitter buffer should be set to accommodate the actual performance of the network.

Taken all together we see that these factors combine to create overall trans-network delays from as little as 20 ms to well over 1 second in a worst-case scenario.

Unicast vs Multicast

In a unicast IP stream, as its name would imply, packets are sent from one address to another, intended for receipt by that location only. A multicast stream, on the other hand, is made available to the network on a special IP address that allows an unlimited number of users to gain access to it. This is ideal, for example, for distribution of a program from a studio to multiple transmitter sites. It is also highly bandwidth-efficient; the multicast protocol allows the same program to be made available in multiple locations without the need to reduplicate the bandwidth more than once on any given link.

Unfortunately, not all networks support IP multicast, and even in those than can, it must be properly activated on all the network devices to operate effectively. Experience has taught us that it is always best to thoroughly test this function on the actual network where you plan to use it before going live.

What Type of Network?

While the public internet can be adequate for certain types of low-bandwidth occasional requirements such as highly-compressed remote feeds where no more reliable alternative exists, it is too unreliable for mission-critical full-time links. It also does not support IP multicast.

For these purposes it is necessary to have some form of controlled IP environment, such as a corporate WAN or a contract service provided by a communications network provider.

When preparing to transport audio in a corporate IT environment, it is essential to ensure at the outset that the network administrators understand your bandwidth requirements and can support the amount of high-priority traffic that you are planning to hand them.

With a contract network service, it is important to have an adequate Service Level Agreement (SLA) with your provider. An SLA is a service contract in which the provider agrees to transport specified amounts of traffic at specified Class of Service (CoS) levels across their network.

CoS is achieved using the QoS tagging described earlier. A low-level CoS is fine for traffic such as file transfers, where if a packet is missing, the receiver simply asks for it to be sent again. However, in real-time media transport, there is no tolerance for the system to wait while requesting a packet to be re-sent, so a higher-level CoS is required.

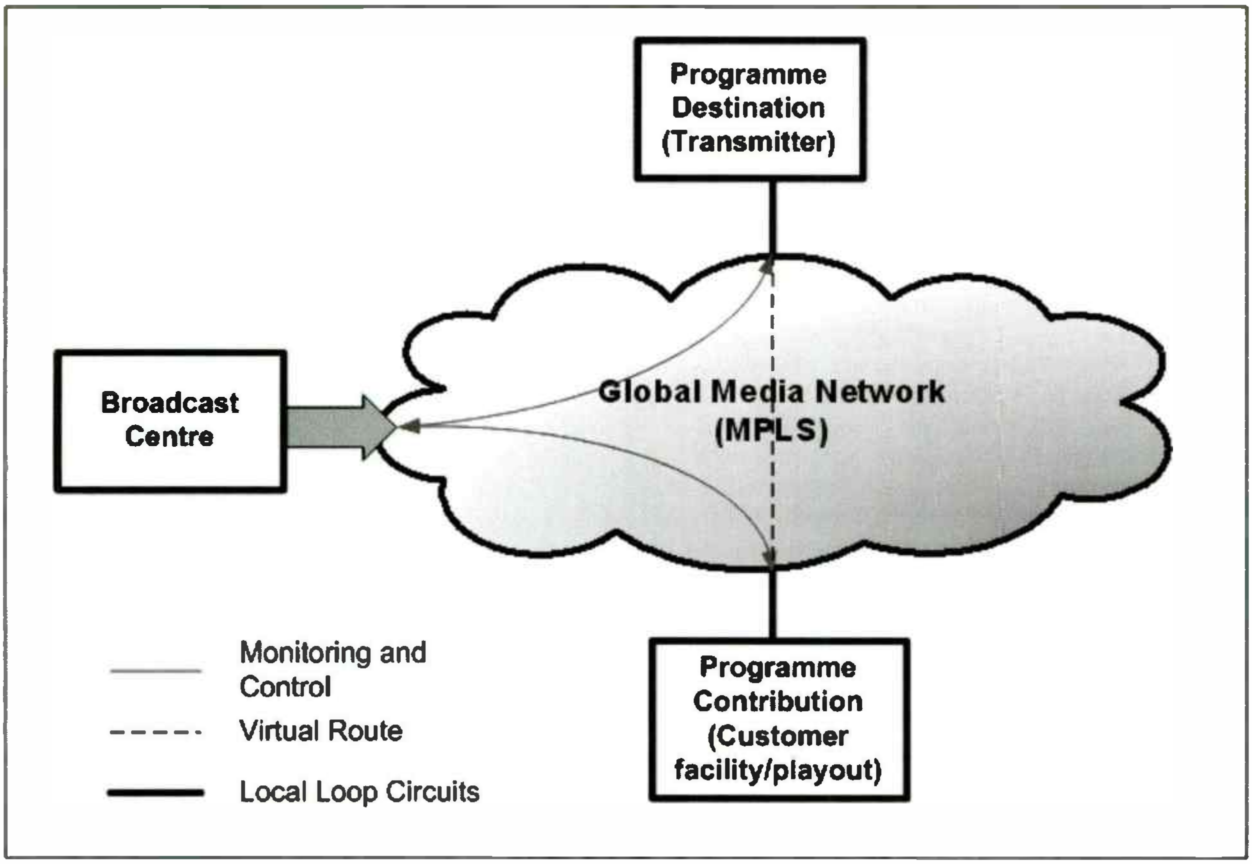
A data transport mechanism called Multi-Protocol Label Switching (MPLS) adds to the overall reliability of IP audio transport networks by emulating some of the properties of circuit-switched (i.e., "nailed-up") networks. However, not all networks support MPLS.

IP Audio Transport in the Real World – the VT Communications Experience

VT Communications (VTC) is a UK-based company that designs, builds, maintains and supports mission critical communications for many of the world's most demanding organizations, including governments, defence organisations, NGOs and broadcasters.

As one of its capabilities, the Broadcast division of VT Communications provides content handling and transmission services to the international broadcasting community. Until recently, this has been largely focused on radio programming utilising AM, namely shortwave and medium wave. We currently transmit around 1,200 hours per day for customers such as the BBC, VOA, NHK and other state and specialist broadcasters.

In 2005, we made a strategic decision to replace and enhance our infrastructure so that we could extend our content handling services onto the emerging New-Media platforms. At the same time, we had identified the market requirement for value-for-money, fully managed services and we needed to implement a solution that would support the onerous Service Level Agreements that underpin such a requirement.



The original distribution network connecting our UK transmission sites was based upon a hybrid of satellite and ATM platforms. Both had been reliable, but the ATM infrastructure was expensive to operate and was reaching the end of its design life. Our international distribution and contribution was largely satellite based, with some ISDN and private leased circuits. Programming for non-live playout and transmission was generally supplied over the Internet via FTP.

This centralised contribution/aggregation/distribution workflow naturally leads to a "star" network topology, with a central Control room acting as a focus for both the workflow and the physical routing and handling of content. Whilst operationally efficient, this exposes services to a critical single point of failure which is resource-hungry and inefficient to mitigate. As a final factor, there was an operational need for reliable, high-availability, duplex connectivity at our own and some of our international partner sites.

The Design

With the large telecommunication corporations driving the rapid development of IP networks, IP seemed to provide the answer to many of VT Communications requirements for its new infrastructure. Specifically, the one-to-many benefits of multicasting and the QoS/CoS structures implemented over MPLS offered us a way to develop scalable capabilities to higher service levels than were previously available to us.

With the addition of multicast enabled audio codecs at the edge of the network, comes the ability to create and manage audio streams which no longer need to pass though a unique point on the network. Instead, the Control Room manages the destination codecs, pointing them at the correct multicast source address as required. In a disaster recovery scenario, with the Control Room incapacitated, any node on the network with the correct management tools, would be able to control the destination codecs to maintain normal switching operations.

The codec which most fitted our requirement was the Harris NetXpress. This leading-edge IP codec supported multicast and our QoS requirements, but also incorporated the tried and tested Intraplex backplane and audio capability, reducing the technology risk of implementing a new solution. We worked with Harris and our pilot network provider, Cable & Wireless, on a proof-of-concept solution linking Kranji in Singapore, with Portsmouth and London in the UK.

The Testing

There were, of course, issues which needed to be resolved. We fully expected this; every new implementation of technology will encounter teething troubles, which is why we ran the proof-of-concept pilot before finalising on a real-world solution. The nature of the issues fell into two camps, although it was sometimes difficult to distinguish exactly in which camp the issue lay. The following paragraphs illustrate some issues which we encountered.

Although the packets will be subject to jitter over the network, they must still be clocked out from the source and clocked in at the destination at the same rate, or buffer over/under-runs will occur and audio loss will result. The simplest solution on paper would have been to lock all of the units to local Stratum 1 clock sources, such as GPS receivers. However, the nature of high-power transmitter sites can make the siting and routing of the GPS receiver extremely difficult.

Our preference was to have the codecs synchronise with each other across the network, but NTP mechanisms were not accurate to the necessary degree. The Harris solution was to develop a timing streams mechanism with primary and secondary stream sources. The trade-off has been bandwidth and hence cost, but this is a small overhead on the overall network operational expenditure and provides the benefit of three tiers of timing failover.

There were core network configuration issues which required identifying and refining. Cable & Wireless took on the task of providing a multicast network to the high levels of reliability and availability required for professional broadcasting. Before the roll-out of the pilot network between Kranji and the UK, the configuration was bench-tested by Cable & Wireless at their offices in the UK, which enabled some teething problems with the timing streams approach to be identified and solved with the co-operation of Harris.

The audio streams run over Expedited Forwarding (EF) CoS, and are segregated from the other traffic on our network by using Virtual Routing and Forwarding (VRF) instances. The control and system monitoring requirements of the NetXpress edge devices are catered for by allocating the highest level of Assured Fowarding (AF) CoS to the SNMP and HTTP protocols which support these functions. This control and monitoring network is also run within its own VRF instance for security purposes. All of the PIM multicast rendezvous points are set up to allow us to switch routes with minimum latency. The whole core network is MPLS, with the service levels and capabilities maintained right into our facility LANs, the NetXpress itself having the ability to drive QoS, should we require it.

The Implementation

With the success of the proof-of-concept MPLS network and NetXpress devices, VT Communications has started the roll-out of the Global Media Network (GMN).

The first phase, which went live in September 2006, replaced all of our UK terrestrial site connectivity with the MPLS/NetXpress solution. This encompasses four remote transmitter sites at Skelton in Cumbria, Woofferton in Shropshire, Rampisham in Dorset and Orfordness in Suffolk. We are currently in the process of moving the control of this network to the new Broadcast Centre in London.

The current phase expands the network to Moosbrunn in Austria, Moscow in Russia, Taiwan, Cyprus Creek in Florida and our UK satellite uplink provider, with these coming on-net by May 2007.

In Summary

IP networks have come a long way in the past few years. New audio coding equipment with specialized tools to enhance the reliability of audio over IP, combined with service providers offering robust MPLS networks, mean IP can be as solid a transport medium as T1 and E1 were before it.

With the right technology and the right network, reliable audio transport over IP networks is not only possible, it's happening today.

Sonic Tonic For Audio Coding

Frank Foti
Omnia Audio
Cleveland, Ohio

ABSTRACT

Achieving great sounding coded audio is easier said than done. What are the critical elements that set apart great sounding digital channels or streams, especially at lower bitrates? Improving codec performance is accomplished through innovative signal conditioning and processing means. Coded audio is now a way of life in the professional and consumer sound industry. It is commonplace in all forms of media that utilizes sound in one form or another. This paper is not written to sell the reader on the benefits of coding, or how it works. The writer assumes this to be a *given,* that the reader understands what audio coding is. (If not, you might consider moving onto the next paper in the proceedings!) The purpose here is to investigate what transpires within the coded transmission system, where the stumbling blocks are, and seek methods that provide smooth sonic sailing. This presentation provides detailed and comprehensive information regarding the causes of perceptable problems in audio coding, how to avoid them, along with a method that improves the sound quality of coded audio.

CODECS: THEY'RE HERE TO STAY

Codecs are common in practically every audio transmission system throughout the world: FM, AM, HD-Radio, DAB, DRM, television, multicasting, podcasting, netcasting, satcasting, and just about every form of *casting* you can think of. Quality sound, especially at low bitrates, requires comprehensive understanding of the coded system, as well as knowing what must be applied, prior, to audio content that insures consistent sound performance. It is more than just plugging sound gear together, configuring the applications and sha-zzam, great sound appears. For the codec naysayers out there, if you feel that Life gave us lemons, well, we're about to make lemonade!

PACKING THE PIPELINE

Data reduced audio systems have changed our world! With the rapid growth of ISO/MPEG Layer-III (MP3) and subsequent additional methods, the capability of transmitting multiple channels of audio is commonplace in this day and age. Within the data payload that once contained a single stereo pair, many stereo channels now exist. Almost seems like yesterday when high quality stereo feeds at 128kbps. via ISDN, were thought to be the best our world could ever expect. Typical of technology, the bar just kept rising and it continues to do so. Today, in the year 2007, a listener can experience high quality stereo, as well as surround, digital broadcasting at bitrates much lower than what we felt were the maximum a few short years ago.

This paper/presentation grew out of efforts to seek improved performance of coded audio at lower bitrates (24kbps – 48kbps). Critical listening to the performance of a new conditioning algorithm, designed to improve vocal intelligibility, revealed two significant results: voice reproduction was noticeably improved due to the new algorithm…but…the enhanced midrange uncovered and/or disclosed negative aural discolorations in the presence and high frequency range.

MOVING TARGETS

Since the early 1990's, audio coding has been around the professional sound industry. Codec developers have been on a fast track, and they continue to be. Audio quality, judged by MPEG (Motion Picture Experts Group) to be excellent at 256kbps and 128 kbps, are now offering the same judgment at bitrates much lower. As codecs improve payload efficiency, it becomes possible to add more transmission channels to the existing infrastructure. It's much easier to improve the data payload, as compared to expanding the pipe. This is how program services are able to expand their range of content offerings with additional channels.

To accomplish this requires using lower bitrate codecs. Lowering the birate increases potential degradation of audio performance. Advancement of codec design has allowed lower bitrates to be employed, and most codecs sound *decent* at these rates, but they are much more fragile with regards to distortion, and susceptible to artifacts. Due to the various types of codecs, and lower bitrates, makes getting a handle on the issues that annoy these functions a moving target, so to speak. The goal of this dissertation is to seek out the gremlins, then offer ways and means to avoid them. Getting under the hood and removing the veil of the codec itself is beyond the scope of this paper.

HISTORY

All transmission systems suffer from some form of problem, one way or another. Doesn't matter if the system is linear or not, suffice it to say that all of them have something to overcome. The simple phrase *no free lunch* applies.

The key to improving audio quality through a coded system is in understanding where the challenges are located, and what can be done to avoid causing them. Performance advancements in prior transmission methods came about due to investigating what caused the *ills* in that particular method. By example the FM-Stereo system: High frequency distortion and peak level overshoots were very common in early FM-Stereo generators. Both the preemphasis boost, and sharp cutoff of the required low pass filters, caused severe problems within the system. In-depth analysis of the system lead to the discovery of embedded preemphasis management and non-overshooting low pass filters, which dramatically improved FM-Stereo performance. By researching the difficulties within the system, and then utilizing the gathered information, it enabled new means by which the challenges were overcome. The same applies to coded audio systems too.

While the concern for FM-Stereo was distortion and overshoot, coded audio suffers from what are referred to as *sonic artifacts*. These are the perceptible annoyances that bother the listener. Most sound anomalies are categorized as one form of distortion or another. Most common are harmonic distortion (THD) and intermodulation distortion (IMD). Coding artifacts are neither. When they are perceived, they occur due to inadequacies of the coding algorithm. Basically, this is the point where the coder runs out of capability to reduce the audio data without the process of data reduction being heard. While there have not been specific technical terms assigned to describe these artifacts, they can be referred to as *swishy-swirly, underwater-like, gurgle-like,* and sometimes *synthetic-metallic*. All of these characteristics degrade sound quality, and reduce intelligibility.

PRIOR ATTEMPTS

Dynamic signal processing does provide benefits to coded audio. Dedicated audio processors that utilize look-ahead limiting and bandwidth control do improve sound performance, but they still do not reduce artifacts enough, at low bitrates…especially below 48kbps. HD-Radio, satcasters, podcasters, and netcasters employ bitrates at 24kbps, and lower in some instances. Reducing artifacts at these low rates usually requires severe bandwidth reduction, which in turn dulls the sound quality.

BUT WAIT, THERE'S MORE!

Additional, careful listening to lower bitrate coded audio revealed another underlying discoloration of the signal. Not necessarily artifact-like, and not really distortion, but the audio quality within the presence range, sounds like there is some type of degrading *ghost-like product* being carried along with the signal. Attempts to remove it via signal processing seemed to increase this characteristic. Careful listening to the output of the audio processor, prior to the encode/decode section sounded very clean. Upon adding the codec to the scenario, this annoyance returned. **Note:** This problem was observed with use of a common known codec for HD Radio…and…various audio processors of different designers/companies were used. All of them produced the same results.

A clue to the problem was revealed when the timing in one of the audio processors was modified to reduce the amount of fast-limiting applied to presence and high frequencies. (This did not remove the limiting in this spectra, but changed the manner in which the limiter's timing responded to transient signals.) The audio immediately opened up, along with clarity in the presence and high frequency range. The ghost-like products were gone. So…what's going on?

UNDER THE SCOPE: TRANSIENT IMD

Considering that the modification to the timing of the audio processor lead to the change in sound, consideration was given to the effect of processor induced IMD within the codec. The following simple test was crafted to observe the effects of IMD through a codec.

Figure- 1 illustrates the test setup. A multi-tone sinewave generator creates the source signals to stress the audio processor and codec. Frequencies were set to 400 Hz and 11.5kHz. The output from the audio processor was routed in two directions: to the input of a multi-channel spectrum analyzer, and to the input of an HD Radio encoder. The encoder was routed directly to a corresponding decoder, and its output was connected to the other input of the spectrum analyzer.

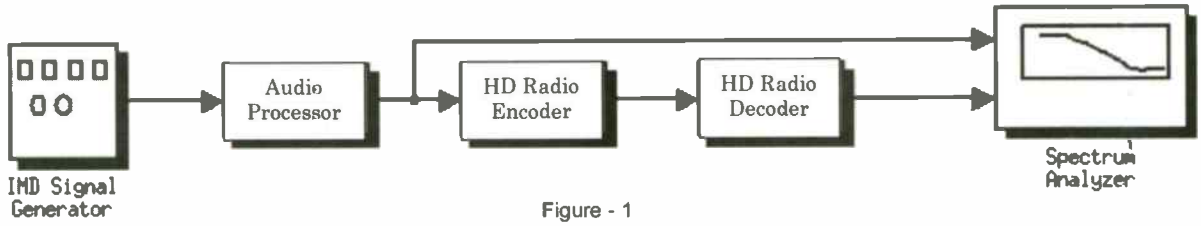


Figure - 1

The objective of this test is to observe whether or not any part of the dynamics function will generate distortion via the codec. The audio processor employed for the test is designed to condition audio in a coded environment. The back end processing utilizes look-ahead limiting, in place of hard limiting/clipping. This reduces THD components in the codec and eliminates aliasing in the system. Tone bursts of the twin tones were used, as this would simulate the effects of transient activity in the source signal, as well as activate the fast-limiting functions in the audio processor.

Figure Ð2 is the spectral illustration of the tone bursts at the output of the audio processor. The twin-tones appear as would be expected. This is also the result when observed at the output of the codec when steady state tones are passed through the processor and codec together.

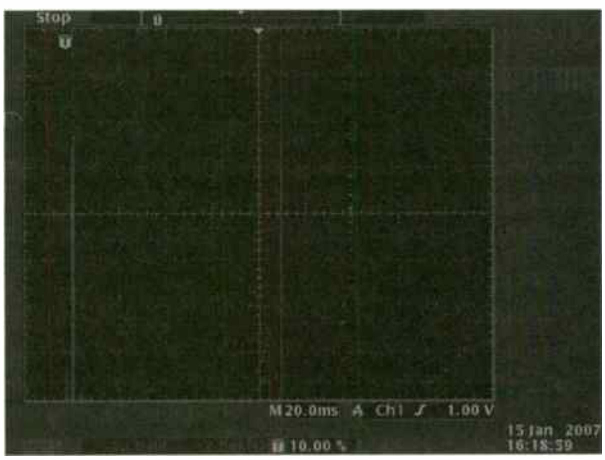


Figure - 2

Figure Ð3 illustrates the output of the codecÕs decoder. Ah, Houston, weÕve got a problem! Notice the significant spectra around the upper frequency of 11.5kHz. Further investigation of the situation revealed that the transient activity upset the encoder and caused added modulation in the upper frequency domain. This is what was causing the added ghost-like product heard prior. Possibly the effect of the SBR function becoming upset at transient information? This is only a hunch, and again, trying to diagnose the issue is beyond the scope of this paper.

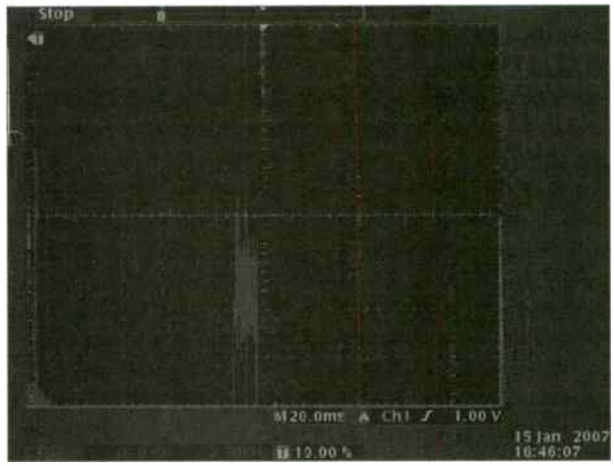


Figure - 3

The rigor of this test exhibits what appears to be severe IMD in the signal. While broadcast source material does not contain transient twin-tones, it does contain plenty of dynamically transient signals within this frequency range. The extent of this added IMD is dependent upon the transients embedded in the source material. Additionally, fast-limiting time constants in the audio processor are capable of exaggerating, and/or creating this problem.

LoIMD

As with most discoveries, thereÕs an answer. In this case, further study of the presence and high-frequency limiting algorithms yielded a method to reduce processor induced IMD. Utilizing a proprietary new function known as *LoIMD*, the algorithm is capable of providing fast-limiting to control transients, yet without agitating the encoder and causing the afore mentioned IMD. When normal source content material is applied, the audio through the entire coded system is devoid of the ghost-like annoyances that were mentioned earlier.

The LoIMD function modifies the control function within a dynamics algorithm. Through internal analysis of the incoming dynamics, and IMD characteristics, the architecture of the control method is rearranged to provide a control signal that reduces, and sometimes eliminates IMD in the processed signal. The sonic result is cleaner sound for a given amount of gain control.

While the LoIMD function answers the issue regarding coding artifacts tied to induced IMD, another critical issue is achieving improved voice intelligibility. For low bitrate applications, this is vital.

CODEC PROVISIONING

Traditional dynamics processors are designed to fulfill the requirements of a medium where the functions are *static*, such as precision peak control and bandwidth limiting for conventional broadcasting, or the normalization needed for recording and mastering. Each of these functions is a known static entity. They are singular, one-dimensional functions where the target is known and the audio processor is designed to accommodate this.

The audio codec, on the other hand, is a moving target. No two codecs are alike, or sound the same. They vary in sonic quality based upon bitrate…AND…more importantly they vary within the same architecture based upon audio content! Here is where conventional audio processors fall short when used in a coding environment.

Until now, dynamics processing has been able to address *some* of the hurdles and artifacts generated by audio coding. The codec has the ability to adapt and modify its algorithm internally, in order to provide maximum throughput, and this alters the sonic artifacts created by the coding process. Unless an audio processor can do the same, it will *hit and miss* regarding how well it provisions the audio to avoid artifacts. Sometimes coded audio sounds acceptable, and sometimes it doesn't. Conventional processors play games with HF limiters and static low pass filtering to minimize coding anomalies. In order to condition audio in hopes of artifact avoidance, the processing will over-compensate audio bandwidth and dynamics. The result is dull, lifeless sounding audio that still contains audible gremlins.

SENSUS EXPLAINED

Sensus technology takes dynamics processing into a new realm. Instead of two-dimensional static architecture and functionality, Sensus adds a third domain where it modifies processing algorithms, architecture, and functions based upon conditions that are understood by the system. Simply stated, Sensus has the ability to sense what must be done to a signal, and then "rearrange the furniture" to accomplish its goal. There are numerous derivatives to this innovative tech, and it can be scaled to many different applications. Following is a discussion of how this method is applied to a processor used in a coded audio environment.

The Sensus algorithm detects troublesome content for a codec, modifies the processor's architecture, and then makes the appropriate changes. These could be dynamics, bandwidth adjustment, a combination of both, or the elimination of a not needed function. The result is consistent quality through the coded transmission system, even at low bitrates; i.e. 18kbps – 21kbps. Voice by example, especially without any other accompaniment, is very difficult to code at low bitrates without the quality and intelligibility suffering. This new process generates clean, smooth, intelligible, and clear audio that is consistent sounding no matter what the content is.

HEADROOM CONSIDERATIONS

Another important factor regarding the coded system is headroom. Digital systems have an absolute maximum ceiling of 0dBfs. Theoretically, audio levels for transmission should be able to be set right up to this level. But, depending upon the encode/decode implementation, overshoots may occur. This is not consistent from codec to codec, but more so due to the implementation of the codec by various manufacturers. Additional input low pass filters in the encoder may cause headroom difficulties. A well designed encoder will insure that any added input filter possess the same headroom as the system, along without generating overshot that reduces headroom. **Note:** Most filter overshot is of the 2dB – 3dB magnitude, but can exceed this amount depending upon filter characteristics.

It would be wise to test any codecs within a specified infrastructure to make sure that 0dBfs, is attainable without system overload or clipping. For this reason, setting the absolute peak level 2dB – 3dB below 0dBfs, offers insurance to avoid clipping.

PROCESSING FOR MULTICAST, STREAMS, DAB, DRM, ETC: SONIC TONIC

The advent of HD Radio[R] has introduced the capability to broadcast multiple content streams within the 96kbps digital channel. To facilitate multicast requires the use of lower bitrate audio coding. The broadcaster can choose the bitrate for each content channel, as well as the number of desired channels, with a maximum limit of seven. Therefore it is possible that extremely low bitrate audio channels will exist, and those will require dynamics processing capable of consistent sound quality that yields low, or no sonic artifacts.

Research, of which this paper is based upon, has yielded a new audio processor for multicast An innovative codec provisioning algorithm - using Sensus Technology, and LoIMD limiters, yields consistent audio quality that contains little, if any, coding artifacts. Yet, audio quality does not suffer the dull or muffled quality due to extreme bandwidth reduction that would normally be employed to mask codec "nasties."

Now it is possible for lower bitrate channerls to offer high quality and clear intelligibility through the use of a dedicated processor that employs the means to understand and handle the challenges of the coded audio path.

For those who wish to tweak on their own, with existing processing equipment, the following should be observed:

1. Avoid dense processing that contains fast limiting time constants. Try to reduce the attack time on functions when 5dB, or more, depth-of-compression is desired. This will reduce upper frequency processor induced IMD.

2. Make sure that the coding system provides full headroom. If the system clips, on its own before 0dBfs, then reset the maximum input level to avoid system headroom problems.

3. Low bitrates will benefit from bandwidth control. A static low pass filter will reduce artifacts. The tradeoff will be perceived high frequencies vs. quality. A specialized processor for coded audio will offer some dynamic method to accomplish this.

4. Do not use *any* final limiter that contains a clipper. The THD generated by the clipping function will cause more trouble than itÕs worth. Precision peak control is needed in the coded system. As mentioned prior, specialized processing for this medium will provide a look-ahead limiter to accomplish this task.

If the above four items are followed, improved coded audio will result. The following section offers further insight as to why a hard-limiter/clipper is a bad application for coded audio.

CODECS AND CLIPPING

Sound mediums require peak control to avoid loss of headroom and eventual system distortion. Precsion peak limiting is employed to accomplish this. Hard limiting, or peak clipping is used in conventional broadcasting, and it works quite well. The method does not technically degrade the system. (Overuse of final limiting is a subjective adjustment, and too much can degrade performance.) Suffice it to say that hard limiting does work as a precision peak controller within FM-Stereo and AM transmission.

The coded path offers a different set of challenges. It is not possible to overmodulate the system, as there is a precise peak ceiling of 0dBfs. Sorry, but +6dBfs is not possible! (Last statement provided for the Programming brethren.) Precision peak control is required, but the conventional method of clipping creates systemic problems, and those occur as aliasing products within the encoder. *Figure Ð4* is an example of what happens to a 2kHz tone, when clipped and 15kHz low pass filtered in a conventional audio processor, used for FM-Stereo, and passed through the HD Radio codec. This problem is consistent with other codecs too.

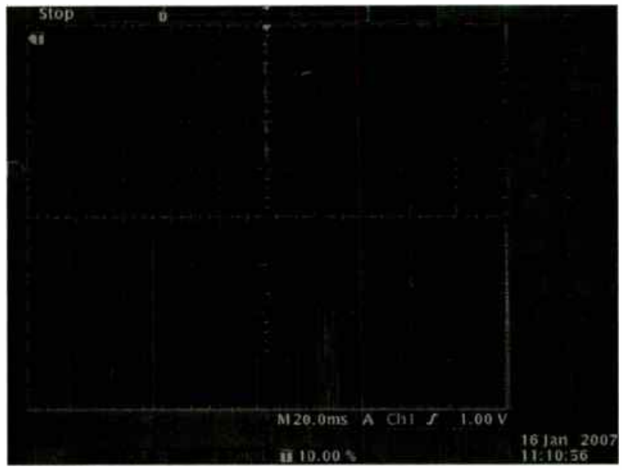


Figure - 4

The cluster of energy that appears around 15kHz are aliasing components. These were caused by the 2kHz clipped signal from a conventional audio processor, as the hard limited signal was routed to the codec. This is proof that all peak limiting for coded audio must employ a limiting means that is void of THD content. Clipped waveforms are exceedingly high in THD. This is why the use of look-ahead limiting is the preferred limiting mechanism for coders. This style of limiter yields very low THD, and will not alias the system.

For reference purposes, *Figure Ð5* is the same signal, prior to the codec. Notice how the odd harmonics line up as would be expected from a clipped waveform. The added strange content that appears around 15kHz in *Figure Ð4* is what exaggerates coding artifacts when conventional style processing is applied to coded audio.

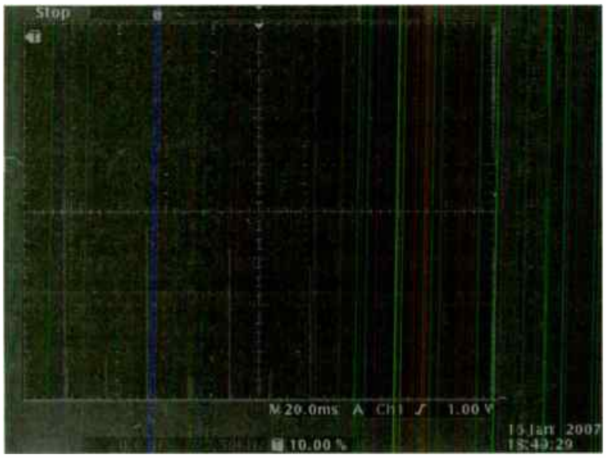


Figure - 5

CONCLUSIONS

Research, testing, development, and hopefully sound reasoning offered here, now explain why coded audio performs as it does. Various signal processing and conditioning means can be used to bring to life coded sound. The test results illustrated here reveal that conventional compressors and limiters exaggerate artifacts. While signal processing, conditioning and peak limiting is required for coded audio, the processing must employ methods that do not contribute additional distortion aspects, as this is what degrades clarity and quality at low bitrates, and sometimes even at moderate to higher rates.

AUDIO COMPRESSION – THE TRADE OFFS BETWEEN QUALITY, DATA REDUCTION AND LATENCY

Guy Gampell
APT, Belfast, Northern Ireland

1. INTRODUCTION

Digital Audio Compression technologies are utilized in a wide variety of audio distribution and storage applications. While these technologies bring many benefits, compression also contributes latency and can have a negative impact on the perceived audio quality.

In its most directly encoded linear form, digital audio requires 1.411kbps - almost the entire capacity of a 1.5Mb per sec T1 circuit to achieve CD resolution (16 bit stereo, 44.1Khz sample rate) audio transmission. This carries significant penalties in terms of the operating cost of circuits and distribution and is particularly inefficient when dealing with multiple channels. Therefore, for many applications, data compression is essential.

With a single encode-decode cycle of high quality audio compression, an average listener can rarely detect coding. However, problems develop when audio is passed through multiple encode-decode cycles - this effect is referred to as concatenation.

In research conducted by the EBU, it was identified that typically, a broadcast chain consists of the following elements:

- Source
- Contribution circuit
- Broadcast studio installation
- Secondary distribution
- Emission

Due to this growing number of 'jumps' and storage points used in the broadcast chain, and the varied compression techniques employed, it is important to choose your audio compression technology wisely.

This paper looks at the options for compression, their relative benefits and disadvantages and suitable applications.

2. COMPRESSION TYPES

The basic criteria for assessing the suitability of an audio compression algorithm or codec for a broadcast application lie in qualifying its latency, robustness and its audio transparency.

There are two fundamental audio compression processes currently employed:

Perceptual: These are based on psychoacoustic models of hearing.

Predictive: These employ a system of predictive coding and are known as ADPCM codecs.

Perceptual codecs – An Overview

The theory of perceptual coders is that of aural and temporal masking, taking advantage of two properties of the human auditory system: Spectral masking (whereby frequencies of high amplitude "drown out" nearby higher or lower frequencies of low amplitude) and temporal masking (whereby loud aural events "drown out" quiet events immediately before or afterwards). For the purpose of Spectral Masking, most psychoacoustic algorithms split an audio signal into a small number of critical bands. Within each band the human ear has difficulty differentiating between co-occurring sounds, so that sounds of higher volume or lower frequency will mask others within the same band. Temporal masking is dealt with by dynamically adjusting the input between long and short signal blocks to identify and mask any desirable effects of sharp temporal sound.

Development of perceptual codecs has been largely driven by the desire to lower bit rates further and further, in part to cater for the consumer market which has seen an explosion in the adoption of portable devices (not just MP3 players and iPods but phones too) for audio storage.

Predictive codecs – An Overview

The theory of ADPCM or predictive coding operates in a fundamentally different way, not least because it does not rely upon the removal of certain parts of the audio signal in order to reduce the bit requirement for transport. Hence, predictive coding is non-destructive.

ADPCM coding takes advantage of the fact that it takes less information to code the difference between two separate and successive audio samples compared to using their actual values. Since the early days of ADPCM algorithms such as G.722, the technology has developed significantly and, in the form of Enhanced apt-X, it is now capable of operating at multiple bit widths of 16, 20 or 24bits.

Development of predictive codecs has been largely driven either by the desire to maintain low delay (latency) or the need to keep audio quality at its highest possible.

3. APPLICATIONS

Low delay being one of the key attributes of ADPCM coding makes it particularly suitable for use on remote live links. A pronounced delay in this application can cause confusion in the contributor and affect the consistency of the live broadcast. The encode/decode cycle is not the only contributor of delay in broadcast applications: the additional interfaces that need to be bridged in order to get the signal onto an ISDN or leased line introduce their own latency into the process - usually about 15ms. Given this fixed delay, it is important to minimize the additional variable delay from audio coding. Predictive codecs will introduce far less latency into the signal than a perceptual codec i.e. less than 1.98ms.

Perceptual codecs take a significantly longer time to complete the encode/decode cycle, for example, modules utilizing MPEG compression types will typically create a latency of up to 120ms in each cycle and should be avoided in circumstances where a low delay is crucial to the quality of the broadcast.

One of the most popular forms of compression applied in such situations is the ADPCM codec Enhanced apt-X. Enhanced apt-X can transport 4 x 20K stereo signals of 384kbps across a 1.536Mbps T1 circuit and maintains 98% of the original signal.

Where latency isn't an issue, and the pursuit of lowest possible data rate is paramount, perceptual coders are often chosen. An obvious example of this would be for final delivery to the consumer (e.g. DAB, IBOC, and streaming over the internet etc). Where such delivery is distributed via terrestrial or satellite using a psychoacoustic algorithm, special care needs to be taken over contribution and distribution circuits so as to avoid concatenation (the detrimental effects of multiple encode/decode cycles).

As shown earlier, the modern broadcast infrastructure has at least 5 "hops" or storage points where several different audio codecs can be used. The EBU identified that for example, in the contribution element as many as 13 different codecs were in use in a sample of broadcasters. Taking into account all possible scenarios, one would arrive at a total of over 50,000 possible codec combinations! Clearly, when choosing a codec, the broadcaster needs to think, not just about the stand-alone performance but the entire broadcast chain.

This factor will be examined further in the paper.

4. THE EFFECTS OF AUDIO COMPRESSION

The purpose of this section of the paper is to investigate different manifestations of audio artifacts that occur as a result of compression based on the dynamic characteristics of audio content. I will also examine how different types of program material can be affected by audio compression and how these processes can affect the logistics of audio transport and application demands.

Content and Artifacts

Loosely speaking, in audio compression, simpler is not easier. To encode and decode simple sounds without detection is more difficult than to encode and decode a complex audio signal such as an entire orchestra or rock group. When testing audio compression algorithms, the most unforgiving source material to use and that used by many researchers is that of isolated pitch pipes, glockenspiels and other sound sources of similar dynamic characteristics.

Classical music is hard to compress due to the high transience between the bass and treble, this could also be said for an acapella female voice and a male voice with low tones. Even local language dialects and speech attributes can cause anomalies in audio compression.

Musical instruments such as the tin whistle generate audio signals that are mostly tonal in appearance. Such signals contain a high level of redundant information in that most of the energy of the signal is concentrated in a small range of frequencies, principally at the fundamental note and higher frequency harmonics. On the other hand, the signal from a cymbal is more like a noise signal in appearance and contains very little redundancy in that the energy of the signal is evenly distributed across a wide range of frequencies.

It is important to note that the waveform of an analog audio signal is time and amplitude continuous whereas a digital signal is time and amplitude discrete according to the sample rate and quantization resolution (number of bits).

Before proceeding, some notes of explanation are required:

Sample rate is a measure of how often in a given period (1 second), a quantization measurement takes place. A minimum of two samples is required to encode a given pure frequency waveform therefore an audio bandwidth "f" requires a sample rate "2xf" to encode it. In practice it is common to use higher sample rates, since the process of sampling also creates noise (called aliasing) which is an image frequency of actual sample rate +/- f. A filter is required to remove this out of band noise and, if too aggressive, such filters can affect the performance "in band". A higher sample rate requires a less aggressive anti-aliasing filter for the same audio bandwidth or can be used to provide a wider audio bandwidth

Quantization is the process of encoding the amplitude of an analogue audio signal: the higher the number of possible quantization values offered, the more accurate the systems encoding is. A 16-bit system such as CD offers 2^{16} possible discrete quantization values. i.e. 65,536 possible values whereas a 24 bit system offers 16,772,216 possible values. Of course the actual signal level will usually fall somewhere between two possible quantization levels. The difference between the two levels (actual and quantized) in itself is an error which is a form of distortion known as quantization noise and in unavoidably induced into the audio signal during the conversion to digital. As this noise determines the maximum dynamic range achievable, the higher the quantization resolution the better the dynamic range.

Figures are as follows:

16 bit
Theoretical 96 dB, in practice usually around 85– 90dB.

24 bit
Theoretical 144 db, in practice usually around 110-120 dB

Data compression can be seen as another stage of quantization added to an already quantized signal and, depending on the accuracy of the quantizer, will simply induce more quantization noise. This is an accumulative process and it is easy to see that, if the process is continually repeated, then there will come a time when the quantization noise is audible.

When encoding a PCM signal, a perceptual coder allocates the reduced bit rate in a way that attempts to maintain a separation band or threshold between the wanted audio signal and the quantization noise - a masking effect.

However, repeated passes of perceptual encoding may very quickly erode this threshold and soon, the quantization noise breaks through this threshold and is audible. In predictive encoders the quantization noise is restricted within each of a number of subbands. The subband quantizers, processing only very small differential signals, are effectively insulated from each other and the quantization noise induced in one subband does not influence any of the other subbands. The overall effect is that an audio signal can withstand many more passes of predictive encoding than an equivalent perceptual encoder. Awareness of these effects is of ever-increasing importance as more instances of digital compression are added to the broadcast delivery chain.

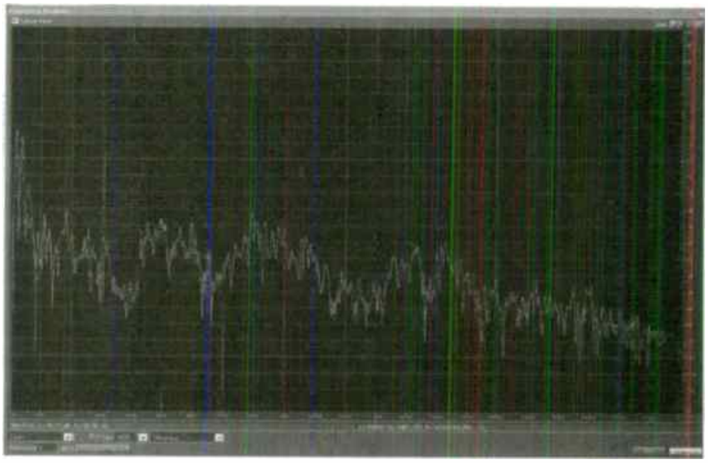
Perceptual encoders operate in the frequency domain using the theory that loud sounds will mask out the quieter sounds. The analyzed sounds are subjected to a masking process that removes all those irrelevant parts of the audio that would never be heard. It will depend on the nature of the audio signal as to how much irrelevancy is removed but it could be as much as 80% of the original signal. If the audio signal is subjected to other passes of compression that involve other types of perceptual encoding then there can be a conflict between the competing masking models. When different types and levels of perceptual encoding and decoding are applied over a number of 'jumps' in the broadcast chain, the problem of concatenation may occur and very quickly degrade the audio signal.

Predictive encoding operates solely in the time domain. Filtered subbands of the PCM signal are analyzed to identify those parts of the signal that are repetitive, a common feature of all PCM signals. Once identified, the sample levels are measurable and readily predictable and this prediction can then be compared with the actual sample. The resulting low-level differential signal is now all that is required to be re-quantized to a lower bit resolution. This is then multiplexed with the reduced data from the other subbands to represent the original linear PCM audio signal. Such an approach is known as redundancy removal and this theory is extended to the decoder which uses the same prediction process to re-introduce a similar amount of data to the decoded differential signal, thus reconstructing the original linear PCM signal. In an end-to-end process, the accuracy of both prediction processes delivers around 98% of the original signal. The minimum signal loss at each pass of predictive compression means that many more passes can be tolerated without audible degradation.

Illustration and Demonstration

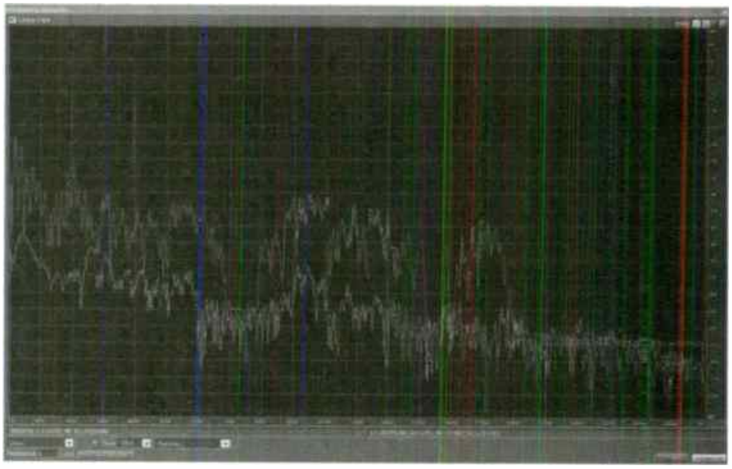
By way of illustration lets look at a recording of the singer Suzanne Vega singing a well know song of hers, solo, unaccompanied.

[A] This is the original after PCM encoding (linear uncompressed audio at CD resolution)



Note the surprising complexity of harmonic content from a single source (solo singer), the complex harmonic structure continues to 20Khz with a gentle roll off from 20Khz to 21Khz.

[B] Next is a comparison between the original PCM of above, and a single Encode/Decode cycle using MPEG.



High the level of harmonic distortion between 20Hz and 2Khz apparent by the wide variation in frequency content shown above between the original and the MPEG encoded signal.

A very large amount of spectral content has been completely removed between 6Khz and 9Khz, with almost total content loss between 7Khz and 8Khz. This is repeated with a very large amount of spectral content removed between 12Khz and 15Khz and almost total content loss between 13Khz and 14Khz. There is a high level of harmonic distortion between 14Khz and 15.5Khz, a steep roll-off above 15Khz and no content over 16Khz.

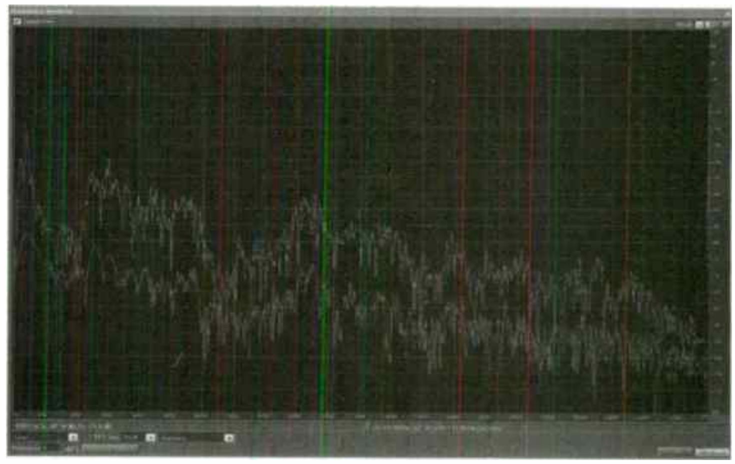
Listen to MPEG 1 time encoded/decoded.

Why does it work? Well it works for the reasons explained above, in that content removed has been done so according to well researched psychoacoustic prediction that this is the content that will not be audible. However this does not bode well for a second, third or fourth encode/decode cycle. **Since it is accepted that a typical broadcast chain involves up to 5 hops or storage points, this is critical.**

Let us have a listen to what happens.

Listen to MPEG, 5 times encoded/decoded.

[C] Finally, a comparison between the initial PCM encoding and a popular ADPCM type of algorithm (Enhanced apt-X).



The preceding is an FFT plot of the same audio after one encode and decode cycle of Enhanced apt-X data compression running at 16bit, 44.1Khz, data rate of 384Kbps.

Note that Spectral content is preserved very accurately including complex harmonic structure and high frequency content.

Listen to apt-X 1 time encode/decode

Listen to Enhanced apt-X after 5 times encode/decode.

5. CONCLUSION AND SUGGESTIONS

Broadcasters seeking the best performance from their infrastructure investment will inevitably have to seek the best possible trade-off between the quality, data reduction and latency offered by their audio delivery compression.

Digital Transmission systems such as DAB (Eureka 147) or HD Radio (formerly IBOC) specify the use of perceptual codecs, such as MPEG or AAC respectively, for final transmission. They also specify the desired compression ratio, which in the case of HD Radio is 12:1. Given these constraints, care needs to be taken with the choice of compression used in any prior stage of the program chain. Any loss incurred earlier in the chain cannot be recovered and incremental losses can be disastrous.

The use of perceptual algorithms elsewhere in the production chain (location "Flash Memory" recorders, storage and play out systems, presenter's personal MP3 players, Mini Disc etc) also adds to the requirement for an ADPCM algorithm to be used for distribution.

Therefore, the safest choice is to use the highest bit rate possible or, put another way, the lowest compression ratio. In parallel with this, the processing delay should also be kept at a minimum for broadcasting consistency.

Radio Remotes Using New Satellite Technology

PAUL SHULINS
Greater Media
Boston, Massachusetts

ABSTRACT

Broadcasting radio remotes has always been a challenge. From getting all the people together, making sure there is power and phone service or a clean remote pick up (RPU) shot available at the location, and putting it all together far enough in advance to allow a good chance of actually pulling off a good sounding and reliable broadcast on time.

Over the past few years that task has become just a bit more difficult with the advent of HD Radio, and the resultant delays required to be worked around. As many radio broadcasters have found, it is no longer practical to use a tuner at the remote site to bring back air product. Instead a full mix minus was required from the studio. This was just another path that needed to be created and depended upon for air talent to get their cues and a sense of what the people back at the studio were doing. Back in 2005 when IP Codecs specifically designed for transporting high fidelity (albeit compressed) audio over the internet started to appear, more options opened up for radio broadcasters. Suddenly the internet could be a part of a transport conduit for getting audio back home (and out to the remote also).

Recently EVDO Cards offered by various cellular companies have provided a way to wirelessly connect to the internet, using data rates approaching a few hundred kilobits/second. The problems are that you have to be in an area where the cellular service was available, and that the data rates varied widely depending on the provider, time of day, and location. This often lead to drop outs and very long delay times, sometimes exceeding 10 seconds or more. Using a hard wired internet connection from a location that had wide bandwidth, and an open firewall turned out to be a good alternative, but even this required careful planning and a thorough check out days before the broadcast was scheduled. Just this year, a new satellite service has become available to broadcasters in North America from Inmarsat. This new service opened up some interesting opportunities for both radio and television broadcasters.

THE CHALLANGE

Greater Media owns five Commercial Class "B" FM Radio Stations in the Boston Massachusetts Market. As Director of Technical Operations for the Boston Group, I am responsible for making sure that remote broadcasts get done successfully. With five FM Stations, the remote schedule can get quite busy, especially during certain times of the year. Programmers (especially those in the news business) are demanding high quality remotes, often on short notice in some cases in areas where telephone facilities or line of site RPU Signals are non existent. Looking for alternative solutions, I purchased an IP Codec along with a portable satellite terminal to allow me to get the data back to the studio via satellite and the internet. The satellite terminal was portable (about the size of a laptop computer) and was easy and quick to set up and align. Data rates available on the hardware I used ranged up to 256 kilobits per second. Latency times including the coding, decoding, uplink, downlink, and internet transit time together totaled about one second. While this was certainly not optimum, it was definitely workable for two way communications.

UP IN THE SKY

Inmarsat owns and operates a network of three communications satellites that form the bgan (Broadband Global Area Network Service) Network. bgan offers affordable connectivity to the internet, pots telephone lines, and the global ISDN network. One of the advantages of deploying this technology for broadcasters is the fact that the service is available just about anywhere on the planet (figure one), and connection to it requires no planning or advance notice once the user is registered as a subscriber to the service. Depending on the location, satellite look angles can vary quite a bit, but from most locations in North America, look angles are high giving more flexibility to choosing a location without fear of terrain, buildings, or trees being in the way.

THE BIRDS

The Inmarsat satellites are positioned in geostationary orbit. They follow a circular orbit in

the plane of the Equator at a height of 35,600km, so they appear stationary relative to a point on the Earth's surface.
They are controlled from the Satellite Control Center at Inmarsat HQ in London, which is responsible for keeping the satellites in position and for ensuring the onboard systems are fully functional at all times. Data on the status of the Inmarsat satellites is supplied to the SCC by four tracking, telemetry and control stations located at Fucino, Italy; Beijing in China; Lake Cowichan, Western Canada; and Pennant Point, Eastern Canada. A call from an Inmarsat mobile terminal goes directly to the satellite overhead, and is routed it back to a gateway on the ground called a land earth station. From there the call is passed into the public phone network or the public internet. A total of eighty five percent of the world's land mass is illuminated by the Inmarsat footprints. This translates into a signal that can be received by ninety eight per cent of the world's population. As you can see, going just about anywhere to do a remote is possible with this technology.

Although Inmarsat owns and operates the satellites, it is the distribution partners and Service Providers that provide the value added connections of the service and the circuit once the transmission hits the ground.

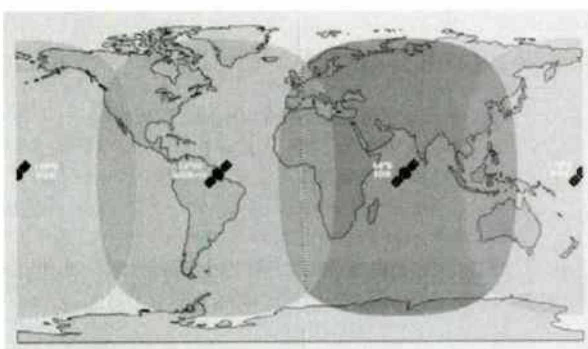


Figure 1. Inmarsat Coverage footprint

HARDWARE OPTIONS

In order to participate in this technology, the hardware on the user end required is usually purchased and the airtime plan is selected. Once registered with the service provider, the system can be used at any time without prior coordination.

There are several hardware manufacturers that offer options for ground based terminals. Your selection will depend on several factors including, size and portability, durability, ease of setup, maximum data rates desired, and of course price.

There are some terminals available that are literally not much larger than a paperback book, and others that are as large as a larger laptop computer. Figure four shows a smaller form factor terminal that is quite easy to transport and set up. It lacks ISDN capabilities but is capable of streaming data at rates up to 64 kilobits per second.

One unit that I decided to gain some experience with is the Thrane Explorer Model 700.

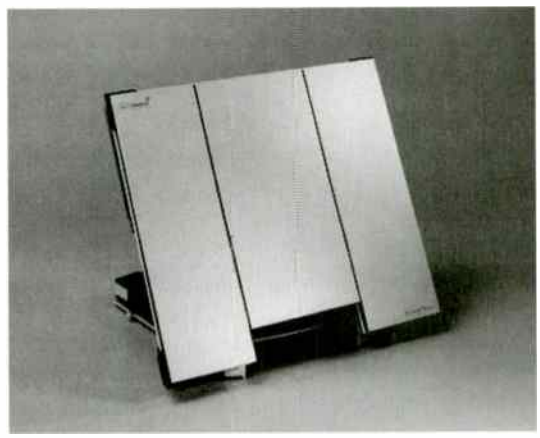


Figure 2. Thrane Explorer Model 700 Satellite Terminal

This particular terminal fit my needs for radio remote broadcasts due to its light weight, ease of setup, and high speed data streaming capabilities. This model has some attractive features including a rechargeable built in battery, detachable uplink/downlink panel antenna, Pots Phone line capability, ISDN capability, and guaranteed data streaming capabilities up to 256 kilobits per second. In addition the unit has the ability to access the internet at speeds of up to 492 kilobits per second.

DATA SERVICE OPTIONS

The core of the bgan service is IP data connections provided as on-demand streaming or standard IP connections. Streaming services are provided with guaranteed data rates over the satellite of 32,64,128, or 256 Kbps. Streaming transmissions over the satellite are prioritized requests and are billed by the minute. Most broadcast or UDP IP transmissions would be suited for this type of service. Standard IP transmissions can be as high as 492 Kbps, but utilize a shared channel that can be contended down based on the number of users in the particular spot beam of the satellite. Standard IP is billed by the Kb for all data transmitted to or from the terminal. Other service options are 64 Kbps ISDN, 4 Kbps and POTS voice that are also billed by the minute.

Establishing a link to the network is generally a simple matter that should normally take no more than a minute or two. The earth terminal is powered up either by battery or by the included AC power adapter. After a short boot time, the panel antenna can be oriented to look to the sky (generally to the equator), and the azimuth and elevation can be adjusted by hand to maximize the signal strength as indicated on the LCD display on the terminal. The signal strength is also referenced by an audible tone that changed in pitch in proportion to the quality of the signal. After a short delay, the unit will register with the service provider, and a ready indication will appear on the terminal's display. At this point in time you may issue a command to initiate a streaming data connection to the internet, place a voice or an ISDN telephone call at 64 kilobits per second. An IP address will be assigned for background IP internet browsing.

It important to realize that registering and aligning the terminal will not incur any charges to your account, however the moment you initiate a streaming IP Connection, the clock starts running and per minute data usage charges will accumulate even if you are not "using" the connection for meaningful audio transfer.

Figure three illustrates a typical configuration used for my remote broadcasts. The diagram shows both an IP Codec and a typical ISDN Codec. The ISDN codec operates using a single "B" channel at 64 kilobits per second, and the IP codec works well at data rates as low as 32 kilobits per second.

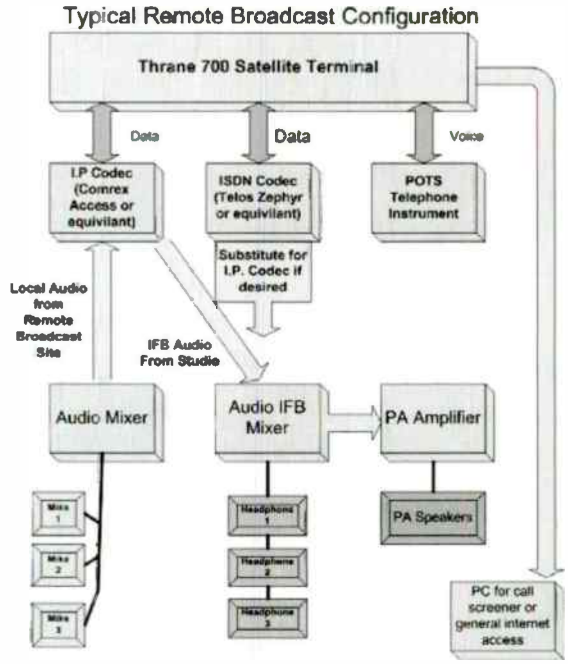


Figure 3. Typical Broadcast Configuration

Figure 4. The Wide Eye Sabre I from Addvalue Communications is a small form factor satellite terminal capable of streaming at rates as high as 64 kilobits per second.

COSTS TO OPERATE

At this writing, the cost of acquiring the hardware (satellite terminal) can range from about $2,000 to $4,000 (US Dollars) depending on the type of equipment desired. The streaming data costs at 32 kilobits per second are running about $2.40 /minute and 64 kilobit ISDN rates are near $5.00/minute. As you can see there is fairly significant cost at this time associated with using this equipment for broadcasts. For example a four hour remote using 32 kilobits streaming will incur about $576.00 in satellite time. When this cost is compared to cost paid to the phone company for the installation and usage charge for a land line ISDN line the pricing seems to be more competitive. The flexibility and reliability of the service can easily outweigh the cost if the broadcasts are critical and other options are not available.

REMOTE CONTROL

Since costs can add up fairly quickly for the streaming option of broadcasting, it may be desirable to have an easy way to start and stop the streaming process during certain types of remotes. For example if a station is jut doing "drop ins" twice an hour, there may be no need to keep the stream running during the entire remote. By supplying a switch closure to the auxiliary port of the satellite terminal you may be able to more conveniently start and stop the streaming process just before and after the talent needs it, thus saving a significant amount of money.

NO LAPTOP

One of the concerns I had before deploying this technology was the simplicity of setting up and operating the equipment. If this was going to be a system that required a substantial amount of wiring, and IT support on site, it was not practical for me to entertain that type of scenario based on how my company operates, and the amount of technical support I would be able to dedicate to a remote broadcast (particularly on short notice). This is one reason why I selected a terminal that would be able to be pre-configured with a standard set of parameters to allow the unit to be used in the field immediately after boot up, without the need to bring along a laptop computer, and a person who is capable of configuring the terminal. Not all terminals may offer this feature, so it is a good idea to research this before investing in a particular set of hardware.

DEALING WITH THE DELAY

This subject comes up more and more often these days. We know from dealing with new digital audio processing, compression, and transport, that listening off the air, or even to a mix minus from the studio can be a big problem. Anytime the delay exceeds 15 to 20 milliseconds, a feed of the talent's voice back into his or her headphones can be distracting or confusing. Therefore a mix minus feed containing not only the programming from the studio without the remote audio, but also a mix of the local remote audio must be fed to the talent's earphones at the remote site. Even still with a delay of one second or more fluent, communications between the talent and the studio will not be possible. This makes doing a radio talk show difficult, because interacting with phone callers or studio talent cannot be accomplished in a natural way due to the resultant delays. Dealing with a one second delay for cuing and passing basic information to the studio and remote site can however be easily worked around since this type of communication is generally not broadcast on the air, and a slight delay in communicating with the talent or the producer can be tolerated.

VSAT

VSAT (Very Small Aperture Terminal) Communications satellites have been around for decades. The first commercial terminals were 6 GHz receive only systems, but today most are Ku band (12-14 GHz) used for among other things corporate communications, and high speed internet access.

Many companies offer relatively inexpensive 1.2 meter "fly away" systems that can be used for remotes. While not as convenient, or as inexpensive at the outset to purchase the hardware for, these systems offer reliable service and upload and download speeds that exceed those of the bgan Network. The monthly rates for unlimited data transfer can be in the price range of what may be spent on only one or two remotes using the bgan Technology. Although I have not had any experience with these terminals for portable remotes, the company I work for has used them quite successfully for dedicated point to point links to carry program material continuously across the continent at very reasonable rates.

CONCLUSIONS

The idea of using portable satellite terminals for remote broadcasts has finally come of age. Being able to instantly get on the air from just about anywhere on the globe with broadcast quality audio is an attractive option that has not been available to broadcasters until now. While the operating costs are still a bit high, having the hardware on hand along with an active data account allows a good measure of insurance for radio stations in case the primary method of broadcasting does not work, or there is just no other way of doing the broadcast. This is especially true during those times when there is no advance notice of the remote broadcast, and a line of site RPU shot will not work.

Then there are those times when you think you have all the bases covered, you show up and the phone lines don't work or your RPU Shot is being interfered with. With just minutes to go before the broadcast scheduled start time, you deploy your portable satellite terminal, connect to the internet, and within minutes you are on the air. That is when this technology really shines!

Acknowledgements

The author wishes to thank his colleagues at Greater Media, GMPCS, and Comrex Corporation for their assistance.

IBOC ANALOG-DIGITAL BLEND TIME ALIGNMENT QUALITY

Michael Bergman
Kenwood Corporation

John Kean
NPR Labs - National Public Radio

ABSTRACT

Analog-digital blend is a unique feature of HD Radio™. Proper time alignment of the audio chains appears to be a key factor for good performance in signal fading conditions. Poor alignment leads to audible level shifts during blend, echo, and comb-filter artifacts. This paper looks in depth at time alignment measurement techniques and the resulting accuracy, presents IBOC receiver alignment results across several brands, and considers audio quality issues

BACKGROUND

In the United States, and increasingly in other parts of the world, terrestrial digital audio broadcast is done using NRSC-5-A IBOC Digital Radio Broadcasting Standard (the National Radio Systems Committee specification for In-Band-On-Channel transmission). Currently, the only implementation of this is HD Radio™, which comes from iBiquity Digital Corporation, the technology's creator. IBOC carries digital audio via OFDM modulation, but also allows for simultaneous transmission of legacy analog AM and FM.

NRSC-5-A IBOC has the unusual ability to "blend" between analog host audio and Primary Digital audio as reception conditions warrant. For the FM band, in multipath fading environments, the receiver smoothly switches from the OFDM-transported digital audio to legacy FM analog audio until conditions improve; it then blends back to digital. This blending is accomplished by synchronizing the analog and digital audio waveforms at the transmitter, carefully calibrating the transmitter's analog-to-digital time alignment, taking into account a receiver's standard analog-to-digital time alignment—all to result in a maximum time differential between the analog and digital of 68µS.

This explanation begs several questions:

- What can be used as a standard reference for calibration?
- How does one calibrate hardware at each end?
- What calibration accuracy is required at each end?
- How does calibration vary with time or manufacturing process?
- What happens if calibration is off?

Through a series of investigations at NPR Labs, quite a bit of detail on these questions comes to light.

CREATING A REFERENCE

Calibration for anything presupposes the existence of a calibration standard. When iBiquity established its standard for time alignment it was realized that it could not be made traceable back to NIST as it is a relative, not absolute, standard. Time alignment depends on matching the delay through the analog path to the delay through the digital path. These paths start with the audio feed into the exciter or audio processor, and end with the receiver audio output.

Since both paths have to go through the IBOC technology, there isn't a single fixed reference. The analog path includes exciter processing and a 7 second delay; the digital path includes a great deal of buffering and processing. The only way to have a standard reference is to create one and then ensure that each transmitter and receiver works within specification.

Originally, iBiquity created a test vector. This vector is a coded fixed signal that can be played back by a signal generator. The over-the-air RF signal is therefore very consistent and can be used as a fixed reference for receiver designers. They used this to calibrate the iBiquity reference receiver. Next, iBiquity and Kenwood calibrated the original Kenwood KTC-HR100.

With a calibrated receiver reference, the transmitter's analog delay could then be determined. To ease this, many receivers have a special mode which splits the stereo outputs, with the stereo left channel output typically carrying digital audio left channel signal, and the stereo right channel output carrying analog audio right channel signal. Time delay between the digital left and analog right waveforms can then be compared.

THINGS GOING WRONG

This "split-mode" test assumes that the original audio was suitable for a tight calibration. As a trivial example, if the digital content does not match the analog content at all, this method will not work. For example, if the analog audio path carries a hip-hop track, and the digital audio track carries jazz, there will be no way to match waveforms. (This example is of course academic, since the FCC's authorization for IBOC requires that broadcasters simulcast the main channel audio on the analog and digital signals.)

A more likely example occurs when the audio processing varies between the two paths. Experience with measurement of actual stations indicates this is a common issue. Waveform phase can vary severely if the analog and the digital paths go through the separate processors.

Assuming simulcast audio, and joint audio processing (analog and digital are jointly processed in a way that ensures some level of phase matching), we can consider calibration of the equipment.

TRANSMITTER CALIBRATION

There are several methods available, each with its own characteristics and with varying accuracy:

1. Calibration by Ear
2. Calibration by Test Vector
3. Calibration by Waveform

The easiest method is calibration by ear. The broadcast engineer simply listens to the audio, forcing blends back and forth. This is not recommended.

The second method, by Test Vector, requires test data from iBiquity and a signal generator. Since this is not commonly available broadcast equipment, the third method, Calibration by Waveform, is preferred.

The rest of this paper investigates the method, challenges and accuracy of calibrating using waveform analysis.

TEST METHODOLOGY

First we calibrated a test Harris Dexstar exciter with a Kenwood KTC-HR100TR IBOC receiver. The method for this was to assume the receiver was already calibrated, which the manufacturer indicates is the case.

The procedure included the following steps:

- *Put test receiver in split mode*

We put the Kenwood receiver into "split mode" (analog out of the right channel, digital out of the left). Split mode on the Kenwood product is enabled with a key sequence available from the manufacturer or in an application note at www.broadcastsignallab.com.

- *Use a L+R tone burst as audio source material*

A sinusoidal rectangular burst tone (1000 Hz tone burst, 5 ms duration, 5 sec interval) is a good alignment test signal for HD Radio and analog host FM channels. This audio file is Track 99 of the "NAB Broadcast and Audio System Test CD", available from the NAB Store online.

- *Capture audio waveforms from the receiver using a high-quality audio input sampler*

A Lynx professional PC sound card was used to digitize the analog audio from the receivers under test

Less-costly PC sound boards (internal, standard line-in hardware) may be usable but must be validated by the user for stability. For example, one laptop's line-in hardware exhibited poor audio SNR in the recordings. One laptop's line-in hardware was mono, not stereo. Some inexpensive audio input hardware samples right and left channels alternately, not simultaneously

We therefore concluded that there may be some laptop or desktop PCs with clean stereo line-in hardware, but engineers should not assume that their PC qualifies as such. A CD source with the 1kHz tone burst fed directly into the PC should be used to verify that the PC audio system is high enough quality.

- *View the captured audio in a sample-level waveform editor*

In our setup, we used Cool Edit 2000 to view waveforms, adjust audio levels and measure timing differences

This overall approach is a good indicator of time alignment, providing resolution down to one sample (1/44,100th of a second or approximately 22.7 µS).

- *Make processing as close as possible to real-life, but watch for processing-induced issues with testing.*

We used an Audemat-Aztec FMX410 stereo generator to provide the appropriate 75 µS pre-emphasis and group delay vs. frequency in the analog FM channel. We found that the generator's audio processing caused a slight amount of clipping on the tone bursts. The audio processing was turned off to eliminate the clipping.

With the transmitter at a known point relative to a (presumably) calibrated receiver, we were able to begin testing more receivers.

We choose the initial rise time point of the burst waveform as the place to measure. iBiquity uses this in their specification for such testing; however, they specify a digital oscilloscope as mentioned above.

Figure 1 shows the digital (upper) and analog (lower) channels of a calibrated transmitter-receiver pair. The vertical black line shows the initial rise time point of the waveforms. This is the basis of the "rise-time" method of blend alignment.

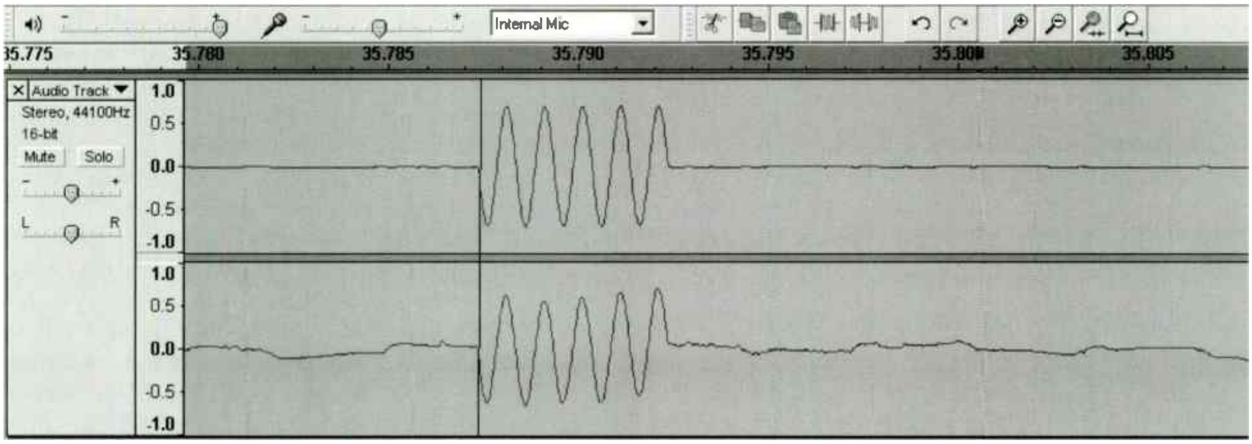


Figure 1: Digital (upper) and analog (lower) output from a calibrated transmitter-receiver pair.

RECEIVER-TO-RECEIVER VARIATIONS

We looked for receiver-to-receiver variation in analog-digital blend time alignment within a single model (manufacturing process variations in alignment). That is, does a single receiver model calibration hold for all the units coming off the production line? We tested four Kenwood KTC-HR100TR units, built over a 10-month period, to see if they had similar calibrations relative to our calibrated transmitter.

Table 1 shows the results when the calibration of a single model receiver is checked over multiple units at a given time. This tight match was repeated three times during testing. In this table, "Samples analog lags digital" refers to how many 44.1 kHz samples (about 22.7μS) that the analog audio is behind the digital. The iBiquity specification for alignment is 68 μS; this is about three audio sample times.

Model	S/N	Samples analog lags digital
Kenwood KTC-HR100MC/TR	# 30900166	0
Kenwood KTC-HR100MC/TR	# 30900867	1
Kenwood KTC-HR100MC/TR	# 40700143	0
Kenwood KTC-HR100MC/TR	# 40700157	0

Table 1: Receiver-to-Receiver variation test results

In this test, we are not comparing receivers tested hours apart due to concerns about transmitter jitter. For a particular point in time, receivers from a common design remain within one 44.1 kHz sample of each other.

Of course, this is not exhaustive testing. However, iBiquity maintains that the receivers should in theory be identical across production lots, so this was simply confirmation that receivers do not vary across production lots.

MANUFACTURER-TO-MANUFACTURER VARIATIONS

We next looked for manufacturer-to-manufacturer variations in receiver blend time alignment. That is, are all manufacturers calibrating to the same standard?

In addition to the Kenwood model, we tested models from JVC, Boston Acoustics, and a professional receiver from Day-Sequerra. Table 2 shows the summary results for five receivers and a monitor. Each of the receivers was well within the iBiquity specification of 68 μS (three samples).

Model	S/N	Samples analog lags digital
Kenwood KTC-HR100MC/TR	# 30900166	0
JVC KD-HDR1	# 119X0336	0
JVC KD-HDR1	# 139X0172	0
Boston Acoustics Recepter	# AFQ 5D001544	-1
Boston Acoustics Recepter	# AFQ 5D002623	0
Day-Sequerra M4	# D70002	0

Table 2: Receiver-to-Receiver variation test results

We concluded that for the receivers tested there is no significant manufacturer-to-manufacturer variation. As a result, any of these receivers could be used for calibration.

TRANSMITTER TIMING JITTER

Table 2 is a summary of the "typical" results. What is not shown in Table 2 is that occasionally—albeit rarely—we did see up to ±2 samples, rather than the worst-case value of -1 in the table. That is, we would sometimes see a value of +2 or -2 in our testing.

At certain times during the day, all receivers would shift one or two samples to the right or left. This added variation appears to be due to the exciter. That is, the receivers remained quite consistent to each other, but all would simultaneously shift one or two samples. As a result, the data varied slightly during the day, even for a single receiver tested in the morning compared to the same receiver in the afternoon. This was observed on two separate days.

Since all receivers of that design behaved the same way at any particular time, we concluded that most likely the transmission chain has one or two samples' worth of delay variation, or large-scale timing jitter, exhibited over a span of hours.

RECEIVER ABSOLUTE LATENCY

The tested receivers had consistent RF-in-to-audio-out latency. That is, assuming the delay through the transmitter was consistent, the delay *through a receiver* was always the same. However, different receivers were not equal in this respect. In fact Kenwood and the Boston Acoustics (BA) Recepter Radio® HD differed by 5120 samples (116 mS). This measurement was very repeatable; three successive tests were run.

The first (baseline) test involved measuring the RF-audio latency between the Kenwood (digital) audio out, left channel, and the BA (digital) audio out, right channel.

The second test was identical to the first, except that the BA was powered down and back up before running the second test.

The third test was identical to the second, except that the Kenwood was power-cycled first.

In all cases, the difference between the Kenwood and BA receivers was the same: exactly 5120 samples.

We did not find any variations in receiver latency from boot to boot, based on comparing the Kenwood to the BA. However, absolute latency through a single receiver is not guaranteed in general—only the relative latency from analog to digital (i.e., blend time alignment) is guaranteed.

This means that it is inadvisable to compare a receiver in digital mode to an identical model in analog mode for blend time alignment testing. Instead, the test should be performed with one receiver in split mode.

MODULATION MONITORS AND REAL-TIME ALIGNMENT IN THE LAB

Ideally, the broadcaster would like to be able to monitor blend time alignment in real-time, off the air. Several monitor makers have incorporated correlation-based time alignment monitors into their products. We tested the Audemat-Aztec Golden Eagle FM-HD and the Belar FM-HD1 modulation monitors.

Our first tests were to verify the monitors under lab conditions. We began with minimal processing enabled, a stereo music source, and a calibrated Dexstar. We observed the effect of changes in blend time alignment on the Belar FM-HD1.

It appears that even small adjustments in timing offset (performed by adjusting the Dexstar) such as 0.16 mS (about 7 samples) were sensed by the Belar. Moreover, it took the Belar only a few seconds to resolve these small errors. The Audemat performed almost identically in all cases.

Frankly, this was not a real-world challenge. The program material was an unprocessed cello solo performance, which provided identical audio for both HD and analog. Processing was minimal. There were no channel artifacts, since this was a lab test with no impairments. The next test was to introduce real-world conditions.

OFF-AIR ALIGNMENT METHODOLOGY

Previously, we described a test methodology for lab alignment, using a tone burst mono signal. To consider off-air alignment, we needed to modify this procedure to look at stereo, processed audio from real stations.

We chose the following approach:

1. Put the receiver in split mode;
2. Capture audio with a PC and a professional quality sound card (must sample both channels each sample time);
3. Examine the difference between waveforms, left (digital) and right (analog), at the sample level, to determine the blend time alignment offset.

MODULATION MONITORS AND OFF-AIR ALIGNMENT: THE REAL WORLD

Using an outdoor antenna, we scanned the Washington DC metro area airwaves for HD Radio stations. Although the reception conditions were not optimal, we were pleased to find six stations with good digital signals.

We next verified the blend time alignment for each station in three ways:

1. Off-air capture of waveforms, then examination with Cool Edit 2000;
2. Using the Belar FM HD1;
3. Using the Audemat-Aztec.

A first conclusion: the Belar and Audemat-Aztec consistently produced nearly identical results, but only for five of the six stations. For those five stations, we had excellent matches between these three methods. However, the last station proved to be challenging. For the last station, with a talk format, it was nearly impossible to reliably compare the waveforms on the PC. Figure 2 shows one example. The upper trace is digital; the lower is analog. There are three peaks or troughs marked with straight lines as possible points for comparison.

The first issue found is that the station has phase inversion between analog and digital. This was not

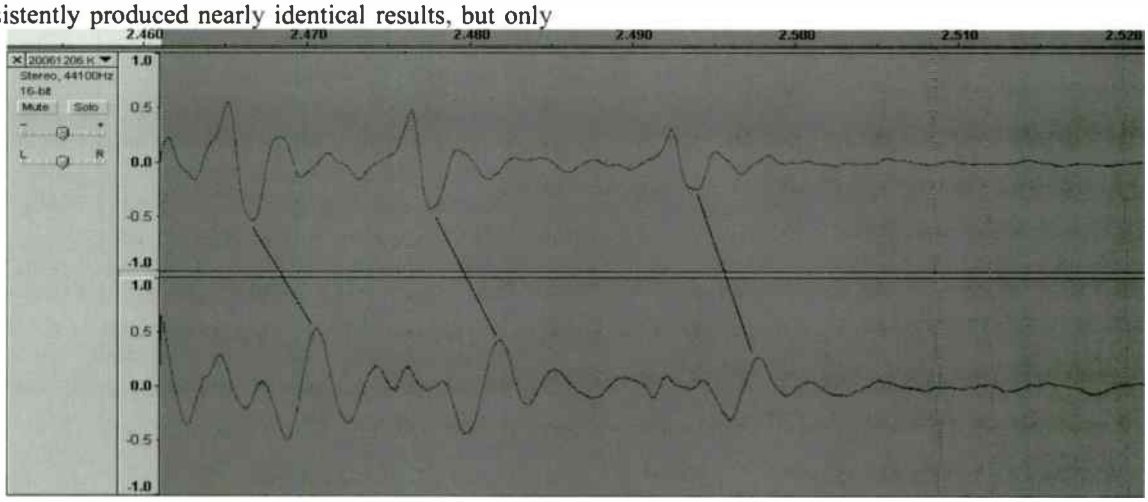


Figure 2: Off-air captured sample, in split mode (left is digital, right analog), with three possible points of comparison, all with different offsets at the sample level.

observed with any of the other stations, of course. But as a result for this station, we had to compare the troughs of the digital waveform with the peaks of the analog, and vice-versa.

With that established, we looked at the points of comparison. Again referring to Figure 2, a detailed comparison of each of these points yielded offset values that differed from the Belar, from the Audemat, and from each other.

The reason is that the compression and audio processing chains are not phase-coherent, analog to digital. We are looking at the tail of some speech component in Figure 2—phase differences between analog and digital are changing the results.

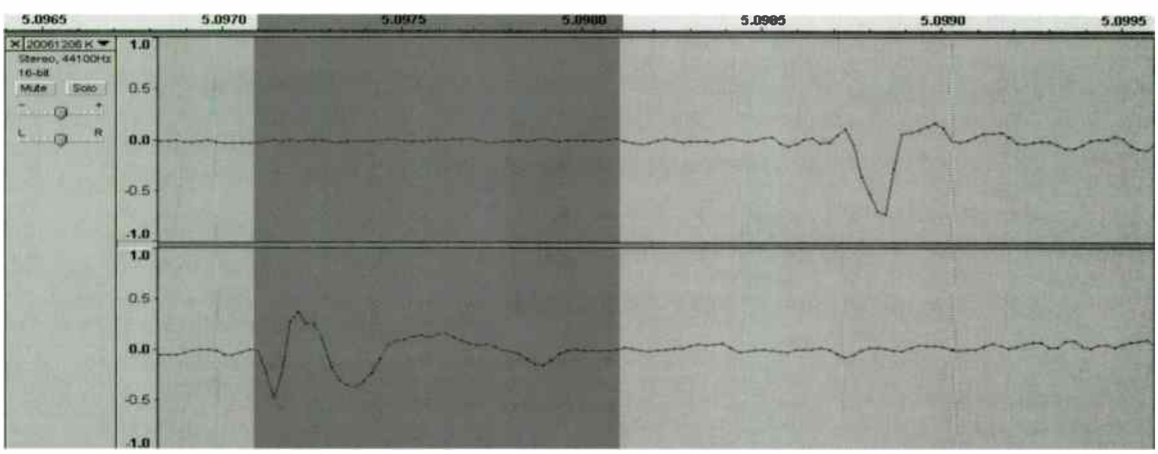


Figure 3: An isolated impulse in digital (upper trace) and analog (lower trace), with an apparent 71 sample offset (analog lags digital). The shaded area represents 46 samples offset, which was measured by the Belar and Audemat during this capture.

We looked instead for isolated impulses in the off-air waveform. Isolated impulse energy appeared to be easier to verify. However, it is by no means a certain result. Figure 3 shows just how difficult this is.

In Figure 3, we can see an impulse which appeared during a moderate pause in the audio; it is fairly isolated and easy to examine by itself. This helps solve the problem of trying to find which sinusoidal peak corresponds to which. However, it still isn't clear. The analog appears to be ringing a bit, and one cannot determine with certainty where to compare.

Worse, the Audemat and Belar both chose 46 samples offset for this waveform; we recorded the monitors' levels during this waveform capture. The shaded area in Figure 3 shows what the Belar and Audemat thought was the offset. There is no way to match the visual of this impulse with the actual value of the monitors.

We note that iBiquity also examined these waveforms in detail and came to different conclusions. We could not reach consensus on whether the off-air monitors had the correct calibration, or the rise-time method was correct.

This issue—that a single station could be so difficult to verify either with the monitors or by hand—highlights a key point. The reason this station is so difficult to calibrate off-air is that they have separate processing for analog and digital, and that processing is heavy enough to make the waveforms quite different.

Therefore, if the station uses separate processing for analog and digital, waveform differences between analog and digital can make it very difficult to reliably determine blend time alignment.

We were not able to determine the effect of the phase inversion on the blend time alignment. It may be that correcting the phase would put everything in order. Unfortunately, the logical step—phase-coherent joint audio processing for analog and digital—is not really practical for many stations.

CONCLUSIONS

The rise-time alignment procedure using a PC for audio capture and on-screen waveform review is verified as a valid receiver calibration method, provided a good sound card is used in the capture. Receivers calibrated by the manufacturer consistently show identical results, although the exciter did exhibit some jitter over a period of hours.

Off-air monitoring is an excellent approach as well, unless the station uses separate and heavy audio processing for analog and digital; in this case it is not yet possible to verify the off-air calibration to the rise-time calibration. Presumably, the human-perceptible differences in audio between the analog and digital would help mask the timing offset differences in such a case.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the help of Jeff Detweiler of iBiquity Digital, Tony Peterle of Audemat-Aztec, and Mark Grant of Belar Labs.

HD Radio is a registered trademark of iBiquity Digital Corporation.

Radio Facilities

Tuesday, April 17, 2007
9:00 AM – 12:00 PM

Chairperson: Paul Shulins
Greater Media, Boston, MA

Rebuilding a Legend: Rebuilding WOR Radio From the Tip of the Mic to the Top of the Tower
Thomas Ray III, Buckley Broadcasting/WOR Radio, New York, NY

Working with the Free Software Community to Replace a Radio Broadcast Automation System
Federico Grau, Radio Free Asia, Washington, DC

Systems Integration Basics for Radio Engineers
Mario Hieb, P.E., Consultant, Salt Lake City, UT

An Embedded Systems Approach to IP-Based Broadcast Facility Remote Control
Stephen Dinkel, Burk Technology, Lees Summit, MO

***100 Things You Forgot**
Stephen Lampen, Belden, San Francisco, CA

Packets Everywhere: How IP-Audio and Ethernet Are Transforming Modern Radio Facilities
Martin Sacks, Axia Audio, Cleveland, OH

*Paper not available at the time of publication

REBUILDING A LEGEND: REBUILDING WOR RADIO FROM THE TIP OF THE MIC TO THE TOP OF THE TOWER

Thomas R. Ray, III CPBE
Vice President/Corporate Director of Engineering
Buckley Broadcasting/WOR Radio

ABSTRACT

WOR Radio is one of the oldest radio stations in the United States, signing on for the first time February 21, 1922, from the radio department of Bamberger's Department Store, Newark, New Jersey. The ensuing 80+ years saw the WOR studios move to Chickering Hall uptown, to 1440 Broadway in 1926 where WOR remained until 2005.

On the transmitter side, WOR went from 500 Watts to 5,000 Watts with a transmitter site move from the roof of Bamberger's to a site up the road from Bamberger's. In 1936, the Carteret, NJ facility was built, along with a power boost to 50,000 Watts and a directional antenna. President Franklin Delano Roosevelt threw the ceremonial switch to launch a high power Golden Age of radio for WOR.

In the early 1960's, it had become obvious that WOR's three tower, quarter wave directional antenna, about eighteen miles out of Manhattan, was just not able to cut through the noise of the rapidly growing metropolis called New York City, and the site was moved to a recently closed dump in Lyndhurst, New Jersey: the Jersey Meadowlands. This site featured towers almost a half-wave in height, arranged in a "dog-leg" arrangement to put the major lobe directly over Mid-town Manhattan and the secondary lobe straight down through New Jersey towards Philadelphia. This site served WOR well until September 8, 2006, when the newest facility in Rutherford, New Jersey, was placed into operation.

This paper will take us through the process of relocating the studios and offices of WOR, as well as the transmitter facility in Jersey.

THE NEED FOR OFFICE SPACE

In the spring of 2004, WOR was starting the renewal process of the lease of our space on the 23rd and part of the 22nd floor of 1440 Broadway. The building had been part of the Helmsley/Spears conglomerate, and had been sold. The new owners had made many upgrades to the building, including a completely renovated lobby and updated elevators. But the new owners thought they had built the Taz-Mahal. Their offer to WOR would have increased our rent immediately by a factor of 2.8 times, almost tripling our rent. The new building owners would not bend from their demands, unless WOR agreed to move to a lower floor in the building. We were told that we could have the basement for the amount we were spending for rent at that time. Psychologically, it did not sit well with Buckley Broadcasting President Richard Buckley: taking a legendary radio station and relegating it to the basement of the building. Secondly, the N/R/W subway lines run under Broadway as close as 50 feet under the building foundation. There was simply no way from a noise standpoint that radio studios could be located in the basement of the building. They would be prohibitively expensive to noise-proof.

THE SEARCH BEGINS

In a City like New York, one simply does not go out on their own and rent office space. You need a real estate broker, and Rick Buckley knew a very good one, Miles Mermel from Tenant-Wise. First, we needed to establish our criteria:

- The rent range we were willing to accept.
- We needed 23,000 square feet of space on one floor, rather than having space on two floors.
- We needed adequate power for the studio area.
- We would need roof rights to install antennas for reception and STL.
- We would need to either tie in to the building's emergency generator, if it had one, or would need the right to install one on our own.
- We would need our own air conditioning system for the studio area.
- The building would need to have the proper telephone facilities, in particular, fiber, to make it work.
- The building would need to be easily accessible.
- We would need 24/7 access.

Miles located several buildings. Rick Buckley, VP/General Manager Bob Bruno, VP/CFO Joe Bilotta and I went on hours of tours. Numerous buildings were immediately vetoed for various reasons. Reasons like, they would not allow the installation of a generator, or having to go through metal detectors and x-ray machines to get in. Can you imagine someone like Lewis Black having to go through that for an interview? He'd have new material for his show.

We finally settled on three buildings. The overall deciding factors on 111 Broadway were:

- The building was willing to work with us to make our move happen – whatever it takes.
- The building is located two blocks north of Wall Street in a very good and accessible area of New York.
- The rent agreement was very good.
- The building is a restored turn-of-the-century building that is absolutely beautiful, with gold trimmed ceiling ornamentation in the lobby and stained glass windows in most of our offices. It truly has the class that is WOR.

The entry hallway into WOR's new home, 111 Broadway, New York City.

An agreement was reached, and planning began.

PLANNING THE BUILDOUT

Things began to move rapidly in the fall of 2004. We decided to keep a good portion of the offices and the lobby as they were on the south side of the floor. The west end, where producers and talent were to be located; the east end where sales was to be located; and the north side, where the studios and Master Control Room would be located, would pretty much need to be gutted.

Architectural drawings were made, studios were laid out, with the following caveat: the studios at 1440 Broadway were huge, having been built in another era. The typical studio at 1440 was 20 feet by 30 feet. The typical control room was 20 feet by 15 feet. The new control rooms are 15 feet by 8 feet, and studios are 15 feet by 17 feet.

We decided to put the technical area, studios and Master, on a 6" raised computer floor. This would facilitate wire runs and keep them out of the ceiling, which would be packed with air conditioning ducts, as the building systems are shut down from 6PM to 6AM and on weekends. We needed to provide our own air conditioning for the studios and Master.

That presented another problem. The building at 111 Broadway is a Landmark building, being one of the original high rise buildings in Manhattan. There are strict conditions in regards to what you can and cannot do to Landmark buildings. The floor we would be on, floor number three, has louvers in place for the building air handlers and, luckily, louver space which could be used for the studio air conditioning. This space is limited, and the air conditioning units barely fit in the space allotted.

WOR's new home: 111 Broadway, New York City.

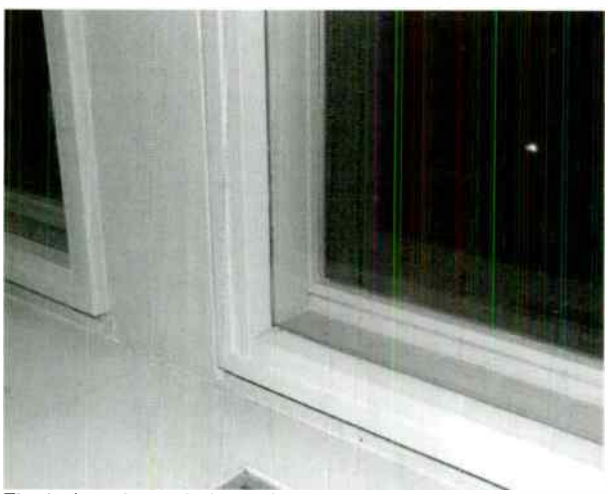
SPECIFYING SOUNDPROOFING

For construction of the studios, we literally took a book out of the Lucasfilms THX manual, and designed the studios to their specifications. They are constructed of two layers of 5/8 inch sheetrock attached to the studs, fiberglass insulation blankets, another set of studs (not connected to the first set of studs), two more layers of 5/8 inch sheetrock, and Armstrong SoundSoak. The sheetrock goes deck to deck and is sealed to the decks with sound rated caulk. The ceilings are two inch thick fiberglass tile. The space above the ceiling is lined with "duct liner", a thick fabric that is normally used for lining air conditioning ducts to reduce noise transmission. The entire space is lined. The air conditioning returns cascade through the studio ceilings. There is a duct in the shape of an "S" between each room, the "S" points down in one room, and points up toward the lining material in the next. All air conditioning ducts are lined. This has effectively kept noise transmission between rooms to almost zero.

WINDOWS

One advantage the studios at 111 Broadway have over the studios at 1440 Broadway is that they have windows. Granted, they look out onto Thames Street, a very narrow street that runs between 111 Broadway and its sister building at 115 Broadway, but there are windows none the less.

The building replaced every window on the floor when we started construction. And I distinctly remember the conversation about having to sound proof the windows I had with Rick Buckley. I was not pleased with the street noise I was hearing come through the windows. He responded by saying that "this is New York. People are used to hearing noise in New York. It's not that bad". As luck would have it, the high school on Church Street let out as we were having this conversation. It quickly became obvious that I was right, as we could clearly hear every "F" word that was shouted. That would have played really well with the FCC's crackdown on obscene words.



The inside window in front of the outside window. This piece of plate glass reduced noise from outside by greater than 60dB.

We decided to leave about an inch or so of air space to the building's windows, and placed a ½ inch thick piece of plate glass over the window in a frame. This effectively took the outside noise down a good 70dB. And it was fortuitous we did this: we were building in winter. When summer comes, the restaurant downstairs in our building hires local rock bands to play on Thames Street. That would have made the studios completely unuseable.

This area would become the new WOR Master Control Room.

CATCHING AN "OOPS"

We decided to go with Studio Technologies furniture. First, the quality is very good. Second, being in the Philly area, they are local and could do the installation.

Vince Fiola came in one fine afternoon to do a final measurement on the rooms to make absolutely sure our furniture would fit. This was before the computer floor was in, and the mock up of the studio windows was in place. There was only one problem. It looked too high in the wall between the control room and studio. Vince measured it. It was a foot too high when the computer floor was taken into account.

Just then, our architect came around the corner. So we told him the window needed to come down a foot. He said he put it at that height so "if anyone puts stuff on the counter, it won't be seen in the other room". So I asked him what I was going to do about our five foot tall producer that wouldn't be able to see above the bottom window frame and therefore couldn't communicate with the host visually. Even though he had sat in on our studio operations at 1440, it didn't strike him that the personnel in the control room and the talent in the studio needed to and relied on seeing each other. I'm glad we caught this while it was only roughed in.

PLANNING TO MOVE A MONSTER

WOR wanted out of 1440 Broadway as soon as possible. But because of construction schedules and the fact that Engineering is responsible for not only the studio area but all phone systems and computer systems, and because moving 100 people in and of itself would be a major undertaking, we didn't want to attempt to get the studios online at the same time as the office move. It was decided to move the offices first.

We met with Verizon, the local phone company, and our reps from AT&T. The PBX is fed from 4-T1 circuits through AT&T. We needed two T1 circuits to ABC Radio Networks at 125 West End Avenue uptown for audio transmission for uplink of The WOR Radio Networks. We needed two T1 circuits, taking completely different paths out of the City, to the transmitter site(s), one in Lyndhurst, New Jersey, and the one being constructed in Rutherford, New Jersey. We needed 3 other T1 circuits for the various POTS lines from AT&T. We needed about 18 POTS lines from Verizon. And we needed 30 ISDN circuits, as we use Telos 2X12 phone systems and feed them with ISDN. Verizon got busy installing fiber and a mux on the third floor.

Decisions were also being made as to what furniture would travel from 1440 to the new location, and what would be purchased new. Offices were being laid out. Mine was to be in the same suite with the General Manager, as opposed to being in the studio area, which I thought was odd. Later, after the move of the studios and offices was complete, Rick Buckley named me a Vice President, so there was a method to his placement of my office all along.

WHAT ABOUT THE TECHNICAL SIDE?

While all the above was going on, it was time to figure out how to outfit the studios. I personally went to visit Wheatstone, Harris/Pacific, talked with SAS, and had in depth conversations with Kirk Harnack, Frank Foti, Mike Dosch and Steve Church of Telos/Axia/Omnia.

I decided to take a chance on being the first large scale integration of the Axia System. There were several things I liked about the Axia system, among them having no single point of failure, being able to basically design the amount of redundancy I wanted into the layout, and having to only run one cable to each studio. This would be a major plus, as we were short on time and didn't need cumbersome multipair to run and punch down. Plus, the Axia system interfaced directly with the Radio Systems Studio Hub system, greatly simplifying and speeding up installation as readily available off-the-shelf Cat-6 patch cords could be used to connect studio equipment along with Radio Systems XLR-to-RJ45 dongles.

We also would be in need of a systems integrator, because of all the work that needed to be done with the office side, plus laying out of the studio side, plus having to constantly check on what was happening at the transmitter site, and on top of our busy remote schedule. I chose Andrew Rosenberg from Creative Studio Solutions of Wheat Ridge, Colorado to perform the integration. I had worked with Andrew previously and knew that we could work well together, as I had specific ways I wanted certain things done.

Right after Christmas, 2004, while the studios were still being put together construction wise, we started integration.

BACK TO THE OFFICE

Because the studios would be remaining at 1440 Broadway to start, certain personnel would be left up there, among them the Program Director, all the Producers and Talent, the Operations Director, the News Director and Newsroom, and the Technical Shop. Since the PBX would be coming over to 111 Broadway with the offices, we needed a plan so people would still have telephones. The studio call in lines would not be a problem, as they were separate from the PBX.

We ordered in 35 Centrex lines from Verizon with voice mail attached. When the PBX moved, we would cut over these lines to the remaining phones, then "sling-shot" the normal phone numbers for the persons left at 1440 through the PBX. In this way, there would be no number change for anyone dialing those left behind.

The Mitel PBX in its new home, the utility room near the Producer/Talent end of the floor.

All phones and computer network connections in the office areas were pre-wired before anyone moved, including several Wi-Fi access points. We mapped out the PBX and made sure that the wiring to the punchblocks matched up with the correct extension numbers.

Meanwhile, construction on the office area was about to be complete, and we were ready to move the offices. On Friday, February 25, 2005, everyone except those to be left behind were told to fill boxes provided, General Services wrapped and packed computers and telephones, and at 5PM, the PBX was disconnected.

The elevator lobby on the 3rd floor of 111 Broadway, the entrance to WOR.

The trucks were being filled all day. It still took two hours from the time the PBX was disconnected until it rolled out of the freight elevator at 111 Broadway (keep in mind the new location is only 3.6 miles from the old location). The PBX was connected, the T1's lit up, and a call to AT&T showed that all was well. I went to my office and made a phone call. Then went to a hotel for the night.

The following day was spent plugging in phones, correcting any errant wiring, and getting computers up and running, while the staff, who had been instructed that it would be in their best interests to come in on Saturday and unpack, were setting themselves up. First day of business at 111 Broadway was February 28, 2005.

WOR's VP/Corporate Director of Engineering Tom Ray in a soon-to-be-completed studio 4.

AS THE STUDIOS PROGRESS

Meanwhile, Creative Studio Solutions was making progress on cable runs and putting things together in the studios. I tackled Master Control, and started populating racks. Axia was on site briefly to get us going with the system, but we were far from ready to program things and make them work.

It was also determined early on that we would not be able to see the transmitter site from our new location. We rented space on 4 Times Square for an STL repeater, and would be shooting through a narrow window up to Mid-town, then over to Jersey. The Gods of RF were smiling on me. I found what could be the last available STL frequency in the City, and that would be our link to 4 Times Square.

The view to the 4 Times Square building from the roof of 111 Broadway for WOR's STL shot. The building is in the center of the photo.

The T1's to the transmitter sites were installed. We brought the backup Intraplex unit over to 111 Broadway and were soon "talking" to Jersey. All of our satellite equipment is at the transmitter, so "Y" adaptors were added to feed both Intraplex units. We now had external audio coming into the studios! Now we could play and start programming the Axia system.

I should mention at this point that the "Y" adaptors were added around 3AM one fine morning near the end of March. As luck would have it, the power company in Jersey whacked us good. At 11PM one night, the Harris DX-50 main transmitter failed. We found 3 shorted rectifiers, then needed to run the transmitter through its paces before buttoning it up. So, since the opportunity to be at the transmitter had arisen (along with myself and another Engineer), we cut the satellite channels over.

STILL MUCH TO DO

They say the devil is in the details, and that is exactly what needed attention. Arrangements were being made to put the STL antennas on the roof of 111 Broadway and 4 Times Square, and get the link from 111 Broadway aimed. We wouldn't be able to aim the link from 4 Times Square to Jersey until last minute. All this was done by the first week of April, and the link to Jersey was aligned the Friday morning of the studio move at the end of April.

The STL antenna, AM receive loop, and DirecTV dish on the roof of 111 Broadway. Note that the Mark grid dish is painted the same color as the penthouse. The AM loop antenna should be pointing at the camera, but offers no signal in that direction.

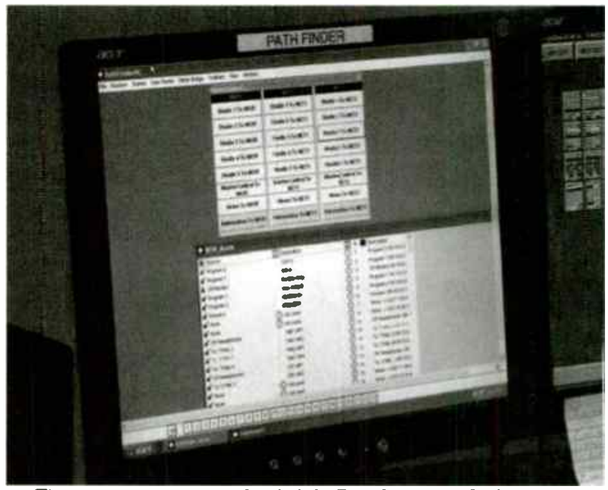
One problem we had were the T-1 circuits to ABC Satellite Services. The first one was installed and we then tested and set levels with ABC the day before we were to move. The second T-1 came up two weeks after the move. We unfortunately had to run without hot backup for two weeks.

We were using Axia Pathfinder as a studio switcher, and right after NAB, John Makely, Mike Dosch and crew arrived for the final touches. We spent countless late hours designing paths, testing studios, testing switching, deciding what sources each studio needed access to and creating those sources in the system. At one point, someone moved my laptop out of their way and Rick Buckley quipped, "Careful! He's got the whole damn radio station in that thing!"

We found that there was an audible "pop" when changing studios. Axia needed to and found the source of the software glitch causing the problem and off we went. Audio chain levels were being aligned. Studio telephone numbers were being tested. There was too much happening at once.

WOR also has unique uses for Pathfinder. WOR is a talk station, and we don't believe in dumping delay unless someone says something really bad. We'll bleep out the word or the offending call letters of our competition. At 1440 Broadway, we had an analog switcher in the delay line in each studio. If something needed to be bleeped, the operator would switch to listen to the outgoing audio, press the bleep button which would put an oscillator on the air, and release the bleep button when it was safe to do so, putting program audio back on line.

Having everything in the digital domain would not allow for analog switching. The Axia crew designed a method for Pathfinder to intercept the button press for a bleep, determine if this studio were on the air or feeding one of the Radio Networks, and route a bleep tone to the appropriate output buss. When the operator releases the button, Pathfinder remembers which studio was on what buss and puts the studio audio back online.



The computer screen for Axia's Pathfinder, WOR's studio switcher and all around system information shuttling device.

We also use Pathfinder to create various intercom paths that are normally not standard in any console system, such as giving the producer direct access to a talent's headphones. All of this is possible through the GPIO nodes in each room.

We have Pathfinder route the cues for The WOR Radio Networks from the studio or Network refeed ENCO to the cueing encoders which feed to ABC and then to the local station's Starguide receiver. This was a big change from 1440 Broadway, where we generated subaudible tones in each studio which were decoded in Master, then pulsed the cueing encoders to ABC. Because we operate a live show in profanity delay, the cue buttons in each room are routed through the Air Tools delay units, which will delay contact closures by the delay amount.

This solved another problem we had at 1440 – we couldn't go to break on either Network if we had dumped delay and were ramping back in. The delay units would pitch shift the subaudible tones just enough that they would not decode in Master. By being able to send straight contact closures, and have them delayed correctly, life has been made much easier.

The WOR Master Control Room at 111 Broadway.

WOR uses an ENCO Digital Audio Delivery System, which was not replaced for the move. The majority of the studio equipment at 111 Broadway was brand new. The ENCO system would not be. On the Thursday before the move, we brought the backup server to 111 Broadway. Operators were busily making backup CD's of the entire log for Friday night, Saturday, Sunday and Monday, just in case. We also planned on operating with entirely recorded programming after the Morning Show on Saturday, and again on Sunday, so we wouldn't have Talent tripping over us and vice versa. Only the newscasts would be live. On Friday morning of the move, after the Morning Show signed off the air, we disabled several studios at 1440 and brought the ENCO workstations which were freed up to 111 Broadway.

Control Room 4 during Henican and White, an intimate setting. Producer Naomi Gabay, foreground, Engineer in Training Jerryl Bell, and Engineer John Feret at the digital editor.

After the ENCO workstations were connected and levels verified, we had the operators come in and touch the consoles – the first time they had been able to do that when at least three studios were fully equipped. Talent came in and got used to the sight lines and the sound of the new facility. What could be packed up at 1440 was in the studio area: the Producer's desks and other areas were not a problem to pack.

Lynne White and Ellis Henican of Henican and White in WOR's Studio 4.

On Saturday, April 30, 2005, I called ABC Satellite Services at 10AM, and they cut WOR Radio Network one over to 111 Broadway. The first program aired on the WOR Radio Network one at 10:06:40. At 11AM, I called ABC again, and we cut over Network two. At 11:06:40, both WOR Radio Networks were operating out of 111 Broadway. I called Master at 1440 for the last time at noon. Near the end of the live newscast, I told the operator to position the mouse over the correct button on the transmitter remote control screen, as with the HD delay, I wouldn't know when to hit the button to switch locations.

Bob Gibson had the final words from 1440, announcing that he was doing the final newscast from those studios, ending with, "and you're listening to WOR, New York". The operator hit the button on the remote control, the audio switcher at the transmitter site switched to the Intraplex from 111 Broadway, and the new facility was completely on the air at 12:06:40. We missed zero airtime.

BUT WAIT – THERE'S MORE!

Eloise Maroney, our Operations Director, got lunch for those of us slaving away. Meantime, the Engineers at 1440 were gingerly removing the main server and remaining workstations of the ENCO system, then gutting the audio racks. Some equipment would be reused. Some would be sent to other Buckley stations. Some would be sold. Some would go into dumpsters.

The ENCO and other equipment finally arrived at 111Broadway by 7PM. We had ENCO up and running fully by 10PM, everyone retired to the hotel before we all passed out, only to be back at 9AM Sunday to tie up loose ends, correct problems, reposition things.

I needed to go to the transmitter Sunday morning and cut over the microwave units and remaining Intraplex to 111 Broadway, and our STL redundancy was back. There were still some minor issues when The WOR Morning Show hit the air at 5AM on Monday, May 2, but overall, the entire move was rather smooth, and more importantly, was transparent to the WOR listener.

AND THEN THERE WAS SADNESS

I visited the 1440 location near the end of the first week of May to make sure everything that needed to be saved was out of the racks. I also had to make sure everything that needed to be saved was out of my old office, which I hadn't seen since January. It was sad looking at the old, original, 25 year old Pacific Recorders consoles that would be sent to the dumpster.

It was amazing to see the bundles of wire in the racks – God knows how any equipment ever fit in there. There had been a tremendous amount of history for WOR at 1440 Broadway, but the new studios have worked out very well.

INSTALLING THE GENERATOR

I wanted to install a 50 kiloWatt generator on the roof of the building. But there was a problem. There is a subway line that runs behind the building under Church Street. There is a subway line that runs under Broadway at the front of the building. The Trinity Church graveyard is on the south side, and Thames Street is too narrow for a crane. Because of the subway lines, a crane is not allowed on Church Street or Broadway.

Any generator would need to be disassembled and brought up the freight elevator. And a 50kiloWatt unit wouldn't fit. A 35 kiloWatt unit would.

So we installed a 35 kiloWatt unit with a 125 gallon diesel tank underneath. There are no natural gas pipes going to the roof of the building, and the City will only allow 125 gallons of diesel. Propane is a no-no.

It was quite the sight seeing this genny dismantled, being crammed into the elevator, being removed on the 21^{st} floor, then being hoisted with a block and tackle one flight to the roof. And then being moved over and around air conditioning units to its final resting place.

Of course, the generator has a story all its own. I mentioned previously that this is a Landmark building. You cannot modify the building outside in any way. Period. As it is, our STL antenna is painted the same color as the equipment penthouse. We had to perform physical measurements, put up a plywood mock-up of the generator, and photograph the building twenty-five ways to Sunday to get approval from the Landmarks Commission to put the genny on the roof.

The generator sits in the south-west corner of the roof, hidden from street view by the facade. Ironically, the generator is almost the same color as the penthouse.

Incidentally, we have used it twice, as the building had to replace two 5,000 Amp switches for the electric utility in the basement – switches that had been in service since 1919! We discovered that 125 gallons will last in excess of 45 hours under full load. To refuel, we need to bring the diesel up the freight elevator in 5 gallon cans. We can bring it as far as the 21^{st} floor, then need to carry it up one flight to the roof.

SOME QUIRKS

I mentioned the windows in the studios previously. We found a few quirks along the way, the funniest of which was the newsroom.

Originally, we were considering doing newscasts in the center of the newsroom. Unfortunately, the newsroom windows are right over where the rock bands play, and fairly close to Broadway. So rather than build an edit booth in News, we chose to use it as the air studio.

What no one caught was that on the other side of the wall in this booth is a bathroom, located in sales. You can imagine the look on everyone's face in Master Control the first time someone in the bathroom flushed the 'ol commode and we heard it on the air! Needless to say, there is now a modified on-the-air light in the bathroom that says "DON'T FLUSH" when the newscaster's mic is on.

Another quirk is that, with the AM loop antenna on the roof pointed in the direction of the transmitter site, we have absolutely no signal. The antenna is 45 degrees off axis where the signal peaks. All I can figure is that it is some phenomenon with the way the signal propagates around the buildings.

Yet another quirk was discovered when the building air conditioning ramped up in May. Our off-air signal became unlistenable. Yet on the roof, the signal coming off the loop was clean. I built a small AM amplifier, phantom powered through the coax from the Master Control racks, and mounted it in the penthouse before the cable left the building to the loop antenna. This raised the signal available in Master from 10mV to almost 100mV and eliminated the interference.

THE SUMMER OF NO AIR CONDITIONING

Summer of 2005 arrived with a very hot spell – and it got very hot in the studios and Master Control room. Something was obviously wrong. And what was wrong was that the Air Conditioning Engineering firm screwed up.

When we were planning, I repeatedly told them I wanted 20 Tons of air conditioning for the studios and Master. They repeatedly asked for heat load on all pieces of equipment – impossible for something like a CD player or computer. The manufacturers just don't provide that information. So, I added up the total power consumption of everything in the room and presented them with a number in kiloWatts for just the equipment, as a piece of equipment can't generate any more heat than the number of Watts it is consuming. This did not include a calculation for people or the sizes of the rooms.

The Air Conditioning Engineers had only put 10 Tons of air conditioning in. And told Rick Buckley, when I was out of town, that I told them that each room would be occupied by one person and a laptop computer. Luckily, Rick knows me better than that, and it then became the problem of the Engineering firm to correct the problem. They, however, had no clue.

So I did the calculations, and determined that the 10 tons that were installed were perfect – for the studios only. Add Master Control and the newsroom to the equation, and we were 15 tons short. Since we had no louver space on the building to play with, it was decided to tap into the building's cooling water supply and use a water cooled air conditioner to feed the Master Control Room and News. The existing units were rerouted to feed only the studios. The new unit feeds Master and News. All is well.

We also found that we needed to install filters between the outside air and the air conditioning units. The air in Manhattan is so dirty that the condenser coils were becoming clogged and severely reducing efficiency once per month.

EQUIPMENT COMPLIMENT

Each "studio suite" at WOR includes a talent studio with an Axia SmartSurface, and a control room with an Axia SmartSurface. Some talent at WOR are capable of running their own consoles to a point. For most shows, however, the operator runs the board in the control room.

Each talent studio contains:

- 1-Axia SmartSurface console
- 1-Axia microphone node
- 1-Axia audio node
- 1-Axia GPIO node
- 1-Axia Engine
- 1-HHB Burn-It CD Recorder
- 2-Denon CD players
- 4-Electrovoice RE-27 microphones with shock mounts on OC White mic arms
- Radio Systems headphone amplifiers
- JBL 4408 speakers
- 1-Crown D-75 amplifier
- 2-Titus Technological Labs on-air lights
- 2-ESE Analog clocks (one is offset 8 seconds)
- 2-ESE digital clocks
- 1-Fostex speaker on the cue circuit
- Numerous computers
- Furniture by Studio Technologies
- 2-Samsung LCD televisions
- 1-DirecTV receiver

Each Control Room contains:

- 1-Axia SmartSurface console
- 2-Axia audio nodes
- 1-Axia GPIO node
- 1-Axia Engine
- 2-Teledyne 100MB Network Switches
- 3-HHB Burn-It CD burners
- 1-Air Tools digital audio delay
- 1-Telos Zephyr
- 1-Denon cassette recorder
- 2-Electrovoice RE-20 microphones with shock mounts on OC White mic arms
- JBL 4408 speakers
- 1-Crown D-75 amplifier
- 1-Fostex speaker for cueing
- 1-custom studio switcher panel
- 1-custom EAS control panel
- 1-computer running Adobe Audition, utilizing a Delta audio card
- Various computers and monitors for call screening, ENCO, and Internet browsers.
- 1-Samsung LCD television
- 1-DirecTV receiver
- 1-Telos 2X12 telephone system fed with ISDN lines and utilizing two Call Directors and two in-console control units.

Engineer Bob Iorio at the Axia SmartSurface in Control Room three.

The Axia core in the Master Control Room contains numerous Axia Audio nodes, a couple of GPIO nodes, two Teledyne 100MB switches, and two Cisco 1 GB Network switches, linked.

We are using Moseley Starlink 9003 microwave transmitters through a Moseley transfer panel as on-air STL, Moseley T1 units for transport of Network audio to ABC Satellite Services, and are using Intraplex units to transport audio to and satellite audio from the WOR Transmitter site plus data and telco.

Pre-processing for WOR, the networks, and the WOR Internet stream are Aphex Compellors. We use an Air Tools Orion to process headphone audio for the Morning Show and to process pre-delay audio for the newsroom.

Master also has a Telos Zephyr Extreme, two Comrex G.722 codecs, and a CCS Micro 56 codec.

Master has their own console, an AudioArts R-60/12, along with numerous monitors and computers, and is home to the entire ENCO system.

IT BASED SYSTEM OBSERVATIONS

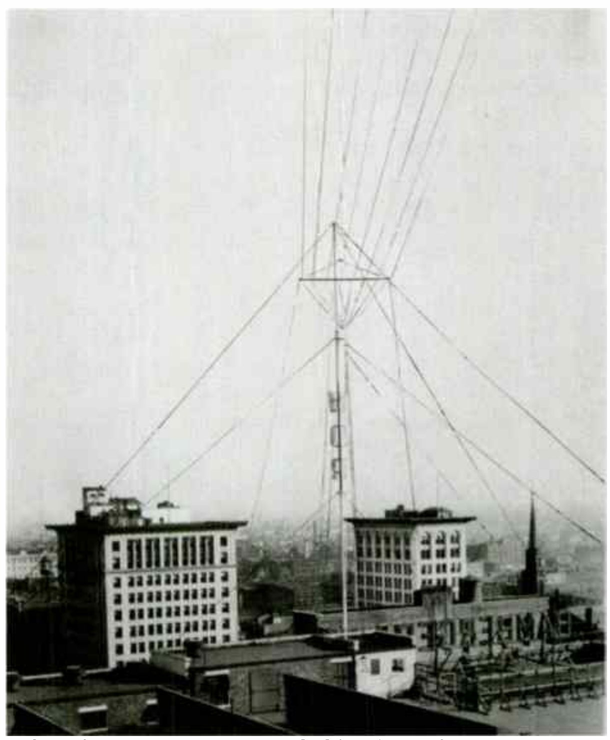
The overall experience using an IP based system has been very positive. The flexibility is wonderful. This became obvious when The Dr. Joy Browne Show started simulcasting with Discovery Health television. The TV crew wanted all the mics, the telephone mix, and certain other items in the studio broken out. Seven items in all. I simply handed them an Axia audio node, programmed seven outputs and made them very happy. Total time? About ten minutes. And no multipair was involved, nor did we have to physically tap into the devices they wanted audio breakouts from.

It is also very easy to reconfigure the system in any way with a standard web browser. We can check system parameters on the audio engines this way, add or subtract what the operators can see for sources, change monitoring assignments. And each operator can set up his console the way he likes it for his particular show, then simply load his or her own profile when they set up a studio. It has proven to be a good choice.

From concept to on air, we built an entire facility in six months. The facility is working out very well and gives WOR the flexibility for whatever may come down the road.

REBUILDING THE WOR TRANSMITTER FACILITY

WOR's first transmitter facility was on the rooftop of Bamberger's Department Store in Newark, New Jersey in 1922. This site utilized a Western Electric five hundred Watt transmitter feeding an open wire horizontal antenna. Around 1924, WOR's power was increased to five thousand Watts, and the site moved up the street from Bamberger's, still utilizing an open wire horizontal antenna.



WOR's first antenna on the roof of Bamberger's Department Store in Newark, New Jersey.

In 1936, WOR moved to a state of the art facility in Carteret, New Jersey, increasing power to fifty thousand Watts and utilizing a three tower in-line directional antenna, one of the first in the United States. This produced a figure eight pattern with one lobe towards Manhattan and the other towards Philadelphia. The site also featured a Western Electric transmitter.

In the late 1950's, New York City was growing, and the additional tall buildings and electric gadgets started causing trouble for the WOR signal in Manhattan. The Carteret site was 18 or so miles from Manhattan, but the site utilized quarter wave towers. Around the beginning of the 1960's, a rather large dump was closed in an area of New Jersey known as the Meadowlands. This area was only 6 miles from Manhattan, and a site in Lyndhurst, New Jersey was chosen.

The Western Electric 50KW transmitter at the Carteret, New Jersey transmitter site. The Engineer in the photo is unidentified.

The Lyndhurst site utilized a "dog-leg" three tower directional antenna and half wave towers. The dog-leg design allowed a large lobe of power directed over Manhattan and Long Island, and a smaller but respectable lobe directly towards Philadelphia. This site gave WOR the most concentration of RF over Manhattan of any New York radio station. It also featured detuning networks and traps in the tower tuning houses because of other AM transmitters nearby. This site served WOR from 1967 until September of 2006.

REDEVELOPMENT

In 2001, the State of New Jersey decided to redevelop the part of the Meadowlands where the WOR Lyndhurst transmitter was located. An initial public hearing was held where concerned parties could voice their opinion. I represented the station, and our message was basically that we supported the efforts to redevelop the Meadowlands, but please leave WOR alone. It was obvious during this session that things had already been decided and the State was simply going through the legal motions – I had to interrupt myself several times to remind members of the board, who were clearly not paying attention and talking amongst themselves, that I was there to present our opinion and that it was their job to listen.

To make a long story short, the State decided to redevelop the area which included WOR's property, and WOR entered into an agreement with the developer, EnCap Golf Holdings, to rebuild the WOR facility. Let the games begin!

CHOOSING A LOCATION

WOR needed about forty acres of land to accommodate our three tower dog-leg array. The good news is that the Jersey Meadowlands area is very wet – great for an AM ground system. The bad news is that this area is congested and there is potential for re-radiation everywhere, so finding property was going to be a challenge.

I traveled around with the representatives from EnCap for two weeks. I was presented with land that strictly wasn't big enough (I was actually told, "well, just run your ground radials around that building next door – I'm sure they won't mind!"). I was shown land that simply wouldn't work because of all the potential re-radiation sources nearby (I counted 10 steel electrical towers on one of the pieces of property!). EnCap sighed and then pointed me to a section of land just north of our Lyndhurst site that was virgin swamp that they had intended to remain completely natural as a buffer. I was able to carve out a 40 acre site to make the system work, and we were off and running.

There was only one problem. EnCap wanted us to start work RIGHT NOW, in 2002. We had not yet filed with the FCC. We were building in a protected wetland and did not yet have the Army Corps permits. There were many questions that needed to be answered.

Carl T. Jones Corporation was chosen to design and commission the new antenna system for WOR. The system was designed, and paperwork submitted to the FCC. To the FCC's credit, we received a conditional approval on the application within 9 months, pending FAA action. That is when the trouble started.

FIGHTING THE FAA

The WOR Lyndhurst antenna system was designed to be overly efficient due to the traps and detuning networks that were required. There are three other AM facilities within 1-1/2 miles of the WOR antenna, 620 kHz, 1190 kHz, and 1010 kHz. In the 1960's, when the antenna was designed, there were no methods for computer modeling and it was very difficult to tell how the detunes and traps would affect the efficiency of the system. So the system was designed with 177 degree towers, 681 feet tall. It was found that the detunes and traps really did not affect the efficiency, and the system was grandfathered as overly efficient.

The new facility would need to meet the RMS requirements of the pattern, and was designed with slightly shorter towers, 171 degrees, 658 feet tall. The new location was ½ mile north of the existing WOR Lyndhurst facility, and therefore was ½ mile closer to Teterboro Regional Airport in Teterboro, New Jersey. Interestingly, even though most traffic into Newark Liberty International Airport in Newark, New Jersey goes straight over the WOR facility (and most planes begin dropping their landing gear at that point), Newark was not an issue.

The FAA initially came back and told us we needed to shorten our tower height by 30 feet. This would have put the efficiency of the antenna system below that required for a Class A AM facility, meaning that WOR would lose its Class A status. This clearly was not acceptable. We decided to fight the FAA. Carl T. Jones Corp. hired an airspace consultant.

What the consultant found was interesting. The Rate of Climb for a missed approach at Teterboro was set at 300 feet per nautical mile, meaning that a plane which needs to abort its landing for any reason must be able to climb out of the airport clear of any obstruction at 300 vertical feet per nautical mile. Just about every other airport in the country has a Rate of Climb set at 400 feet per nautical mile. Increasing this Rate of Climb would make the proposed WOR tower heights fine. There was no justification found as to why the Teterboro Rate of Climb was at 300 feet.

During this process, our consultant suddenly was not able to reach the agent we were assigned at the FAA regional office. After roughly a month, he was finally able to reach this agent's supervisor, and found that the agent had undergone emergency kidney surgery and would be out for another two months. Our consultant explained the urgency of our request, and asked that the case be assigned to another agent. Being bureaucracy at its best, the FAA supervisor stated that this would not be a problem – but the case would need to be restarted from the beginning. We decided to wait for the agent to return from sick leave.

Upon the agent's return from sick leave, it was agreed that there was no reason the Teterboro Rate of Climb could not be set to 400 feet per nautical mile. A circular was issued to the pilots who frequent Teterboro for comment. There were no negative comments, and the Rate of Climb was changed 30 days later. The FCC issued WOR's Construction Permit within two weeks of the FAA's approval, in March, 2003.

STARTING CONSTRUCTION

While waiting for the FCC and FAA, we were able to perform certain tasks before all the paperwork was completed. Since the area we were building in was a swamp, we were allowed to drill test borings to find out where bedrock was, as we would want the tower bases, guy anchor points, and transmitter building to be sitting on something solid. Our tests showed that the ground was literal muck until bedrock was found at 150 feet deep.

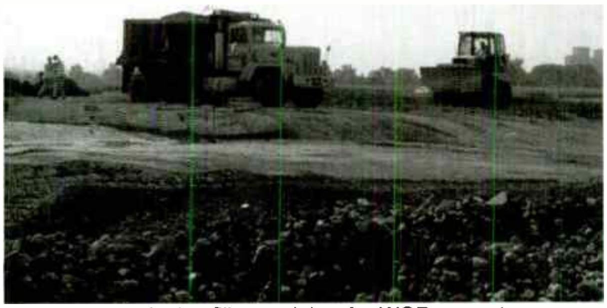
Cutting the Phragmities was the first order of business before conducting a survey of the location for the new WOR transmitter site.

We also were able to get all of our drawings made and approved, and obtain building permits. In another bureaucratic mess, the area is controlled by two entities: the New Jersey Meadowlands Commission and the Borough of Rutherford. NJMC issues certain permits, Rutherford others. Of course, their game is that they tell you what is required, you make drawings and produce seventeen (!) copies, submit them, and find out they "forgot" to tell you something that was required. They "forgot" 15 times. Oh. And I also needed to get a "soil conservation" permit. I guess the fact that we needed to fill, rather than remove soil, was overlooked.

Having paperwork in hand, we were ready to start construction. The contractor mobilized and was on site. And was then told to stop. Because the funding mechanism from EnCap was not in place, and no funding was available at this time. This was May, 2003. We would not be able to start until April, 2004. In the mean time, however, we were using a small contracting firm who had not taken on other work because of the extent of the WOR build. They now needed to be paid, as it was too late to bid on contracts for summer work. Needless to say, the time between May, 2003 and April, 2004 was a mess.

REALLY STARTING CONSTRUCTION

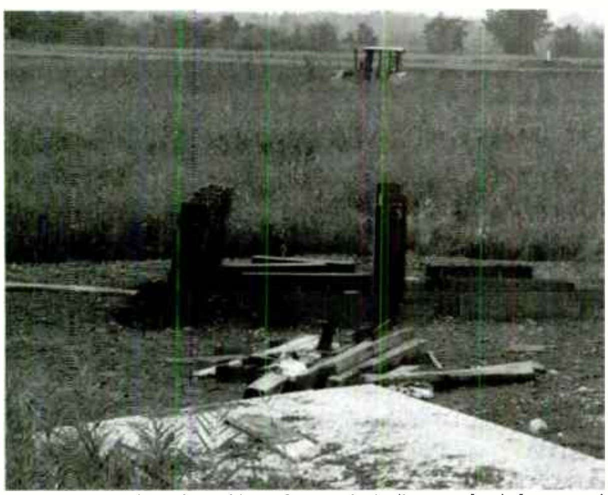
When May of 2004 came around, we started in earnest. A total of 1.45 acres was to be filled to a height of 12 feet above swamp level (100 year flood stage is 9 feet) for the area for the transmitter building, tower bases/tuning houses, guy anchors and roadways to get to the towers. Because of the height and the fact that the fill actually sunk two feet, we used over 80,000 cubic yards of fill material.



Dump trucks bringing fill material to the WOR property.

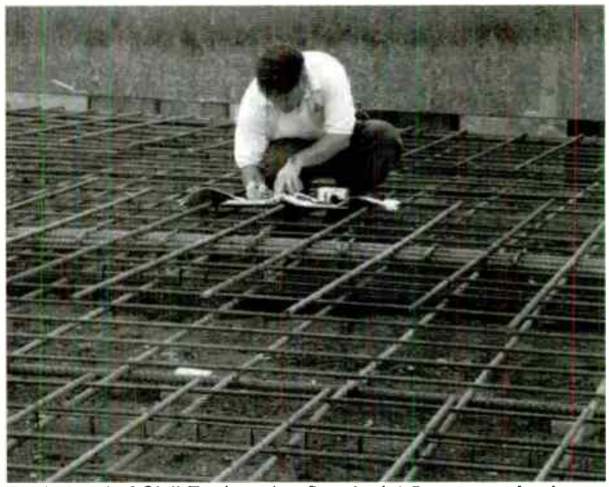
After the filling was done, we needed to put in piles to the bedrock at the location of the transmitter building, the guy anchors, the tower bases, and the tuning houses. All of those mentioned sit on steel I-beam piles. The bedrock depth varies across the 40 acre site. In the area of the transmitter building, the depth is 165 feet. As you move east, the depth decreases to 110 feet. At the transmitter building, I had the contractor braise four inch copper strap to the piles in all four corners of the building, a pig tail for inside the building and one for outside the building. WOR has 165 foot deep ground rods going through wet, sticky muck. Perfect.

At the guy anchor locations, a set of piles was put in straight down, and a second set was put in angled. The angled piles were set so that they sloped downward in the direction of guy wire pull, so that the pulling force would be pulling towards the bedrock rather than pulling against a pile driven straight down.



A guy anchor location. Note the angled piles on the left, toward the direction of guy wire pull.

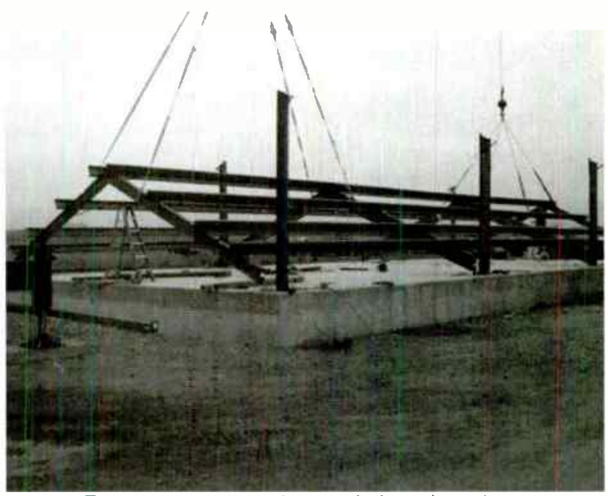
After the piles were in place, the foundations were formed and poured. The slab for the transmitter building alone took eleven trucks of concrete. Bolts to hold the guy anchors and the tower insulators were placed in the forms and welded to the rebar before concrete was poured.



Paul Mucci of Civil Engineering firm Aerial Spectrum checks foundation rebar placement against the drawings before the transmitter building foundation is poured.

UP GOES THE BUILDING

To save time, WOR decided to put up steel buildings in all locations. This would significantly cut down on the time needed to erect a building, and would provide some shielding in the harsh RF environment that is the New Jersey Meadowlands. It's bad enough that WOR's 50,000 Watt transmitter is on site. We have three Volts per Meter of signal from WINS to contend with, as well as signal from other stations in the area. While the building does not provide complete shielding, it cuts down the RF in the building, and I have not experienced an RF problem with any of the equipment.



The roof trusses ready to be hoisted into place.

At the time the superstructure of the transmitter building was put together, we had the Onan 300 kiloWatt generator delivered and placed on the slab with a crane. I guess you could say that the building was built around the generator.

A worker enlarges anchor holes on the transmitter building support beams.

It was now December, 2004. The building shells were up. Insulation had been applied to the transmitter building interior. Framing had been put up for the bathroom , studio area, garage, generator and utility rooms. The electrician had started roughing in the electrical. But there was a massive leak in the building and the contractor was taking his sweet time. And he had requested that we put the transmitters in the building. After looking at the extremely large puddle of water on the slab, I flatly said "no".

Insulation has been blown in and the framing is in for the bathroom and studio areas.

I was then extremely involved with the WOR studio project, and the contractor was moving slower. Right before I left for NAB, 2005, I showed up at the building one morning to look at progress to find the contractor nowhere to be found. All his materials and equipment were gone. He had moved out in the middle of the night. As it turns out, he had underbid the job, realized his mistake, and didn't like my response when I told him we intended to hold him to his contract. After NAB, I started working on finding a new contractor.

ENTER THE NEW

We ended up hiring a contractor who had performed work for a couple of other stations in the area. I liked his work. He seemed like a good guy. The reviews I got of him were glowing. In the end, there were things that people "forgot" to tell me about him. Like the fact that he had a habit of getting himself in financial trouble and disappearing.

But, we got the building to the point where we could accept transmitters, got the electrical up to snuff, and started construction on what is known as tower number two in the array.

The transmitters and phasor have just been delivered to the WOR site.

TOWER CONSTRUCTION

Tower construction in the swamp is very difficult. The crew can't simply walk out into the swamp to do things like pull guy wires because they sink in up to their waists. What should have taken six weeks took five months.

A boat had to be used to take the 500 pound guy anchors out to the anchor points. The crew had to be creative in how they ran their load lines, as they couldn't achieve the optimal angles they wanted (and didn't want their winch to sink into the swamp). A crane was used to set the first 100 feet of tower on the base insulator. This 100 foot stub was guy wired in place, and the rest of the tower was built with a gin pole.

What really amazes me about things like towers is that, here we had a 100 foot stub sitting on top of a ceramic insulator with guy wires attached. The weight of the tower sections was approximately 15,000 pounds, and you could probably tack close to 5,000 pounds onto the weight for the dead weight of the guy wires. Yet, if I grabbed onto the tower and yanked to one side or the other, I was able to twist the tower on its pivot point.

Workers from Northeast Towers remove sections of tower 2 from the delivery truck.

Another problem the tower crew had was the RF in the area. The old WOR site was one-half mile away operating at fifty kiloWatts and putting two Volts per Meter in that direction, WINS was putting three Volts per Meter in that direction, WLIB was putting one point five Volts per Meter in that direction, and WJWR was putting about one Volt per Meter. The site was very hot. And the tower guys were getting belted at various times. It was interesting that the tower wasn't very hot at all until it hit around 300 feet. It was very hot from 300 to 500 feet, then was OK again. And then there were the skirts that had to be put on the towers.

Each of the WOR towers has two wire skirts. The purpose of these skirts is to minimize current flow in the tower at 1010 kHz and 1190 kHz, effectively shortening the electrical length of the tower at these frequencies. Of course, no matter where you jumped the skirt to the tower, which was grounded, they were still warm. The tower crew was glad when tower 2 was complete. The tower was complete around the first of November, 2005.

Meanwhile, after much arguing with the general contractor, who couldn't seem to get the attention of the electrical inspector or the power company, I stepped in and got power turned on to the building so that we could power the tower lights. I also had dialtone installed in the building so we could monitor those lights.

The first one hundred feet of tower 2 stands as the crane is lowered.

MOTHER NATURE HAS HER SAY

The morning of November 7, 2005, found me with a flu bug. I was supposed to be in Washington, DC for a committee meeting at NAB. Our other two towers were to be delivered to Rutherford this day. But I felt really awful when I got up, immediately sent an email to John Marino at NAB requesting a phone number I could call to attend the meeting via phone, and wandered into the kitchen to get something to eat. My wife had the CBS Morning News on TV, and they were showing images of a tornado that went through Evansville and Newburg, Indiana. Evansville. That name struck a bell with me in my stupor. Then it came to me. That's where our towers were coming from! But, they were supposed to be delivered at 8AM, and it was 8:10AM. No call from the tower crew. I guess things were OK.

Then the phone rang. The tower crew wanted to know if I heard from the tower company because the trucks weren't there yet. I called Steve Savino at Northeast Towers, told him what I was seeing on TV, and said, "they DID get the towers out of there Friday night, didn't they?" Steve tried to call Central Tower (now Tower Innovations) with no luck. A chill went through both of us.

Grabbing a bowl of cereal and a glass of juice, I proceeded to put together a "nest" for myself in the recliner in my home office, then booted up my laptop. It was going to be a long day, even with the flu. By 11AM, I was attending my committee meeting on the phone and frantically emailing people. I finally managed to reach, via email, the Northeast Sales Manager of Central Tower who stated that "things are a mess there right now and there isn't much in the yard that survived". By 1PM, it was confirmed that there was an issue with the trucking company and that the towers were to leave the yard at 7AM Monday morning, November 7. Our remaining two towers had been destroyed in a tornado.

I then emailed our consultant, Tom Jones, along with our FCC attorney, David Oxenford. At this point, Central Tower was picking up the pieces and assessing damage. I had no date, nor could I expect one, for refabrication and shipment of the two towers. It was November, 2005. Our Construction Permit expired in March, 2006. And they are not renewable.

As it turns out, the FCC has a mechanism in place for things like natural disasters which, through no fault of the station and being out of the station's control, the station cannot fulfill the Construction Permit by the end date. A tornado destroying the towers the day before shipment from the manufacturer fit their definition perfectly. By the end of the day on November 7, 2005, we had filed with the FCC for Tolling of our Construction Permit because of an Act of God, out of the control of the radio station.

CONSTRUCTION CONTINUES – MAYBE

Things had now slowed down on the completion of the building because, with our last two towers destroyed, the pressure was off to get everything done. Our contractor asked if it were OK if he took on a 2 month job building a high-class dog boarding facility, and I said OK. There were things I could do, like getting equipment racks in place, putting the phasor together, getting transmitters put together. But when those two months were done, in January 2006, it was getting hard to get the contractor's attention. By this time, I had a delivery date for one of the towers for the end of March, so I needed to ramp up and certain things needed to be done, like power needed to be run to the transmitters so I could start testing into the dummy load and align the HD signal.

After much arguing and cajoling, I was finally able to get the electrician in to wire the transmitters. Then the tower came in and erection started. I was wiring the audio racks and transmitters, Verizon put in the fiber and the two T-1 circuits to 111 Broadway, the tower crew freed up a guy to hang the STL dishes on a 50 foot section of Rohn 40 they provided and ran the cables inside (there is a 2 degree aiming difference between Lyndhurst and Rutherford which provided great STL signal). Then, it started becoming harder to get the contractor's attention.

The WOR twin Harris 3DX50 transmitters, with HD rack in between.

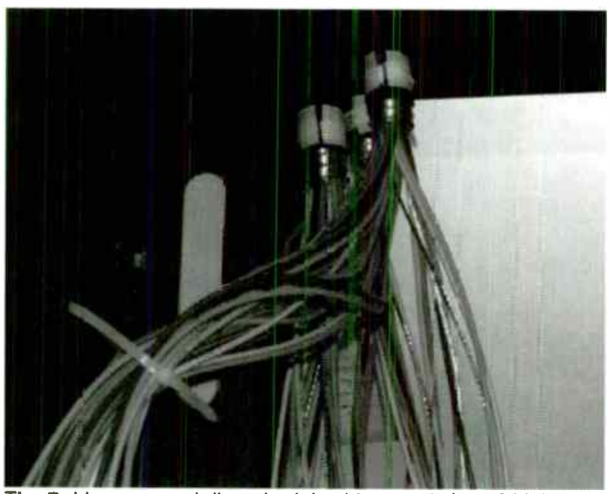
As it turned out, he had gotten himself into financial trouble. Our Construction Permit Tolling stated that the clock on the CP would start running again when the second tower was on site. With the FCC having a New York office, I did not want to have the tower delivered and not inform them it was on site, only to have someone come out and see what was going on, so everything was done on the "up and up". I was given a date for delivery of the second tower which coincided nicely with the completion of the tower known as number three. The tower showed up and was brought around to the peninsula, the home of tower number one. To get there, we need to travel down the New Jersey Transit railway right of way, and it took almost two days to bring the tower sections over to the location, as we needed to be conscious of trains. I managed to reach our contractor and told him I needed a crane on a particular Friday morning at 8AM. He told me he had no money. I told him again I expected to see a crane at 8AM Friday and hung up.

On Thursday, I called the crane company. No crane had been reserved, which was no surprise to me. So I made arrangements with the crane company to supply a crane the next morning with an operator.

On Friday, the crane showed up right on time at 8AM. By 9AM, 100 feet of tower number one was standing. At 9:10AM the contractor decided to show up and had a conniption. I apparently had "no right" to get the job done. He was given his walking papers right then and there. As it turns out, he had not paid the electrician or for the crane when tower three went up, so we took care of the electrician and crane company. When you do that, it's amazing how cooperative people become. I have no problem getting the electrician in, and the crane company will show up whenever I need them, like when I had to set the 4,000 gallon diesel tank. As of this moment, I was the general contractor.

The 4,000 gallon diesel tank in place.

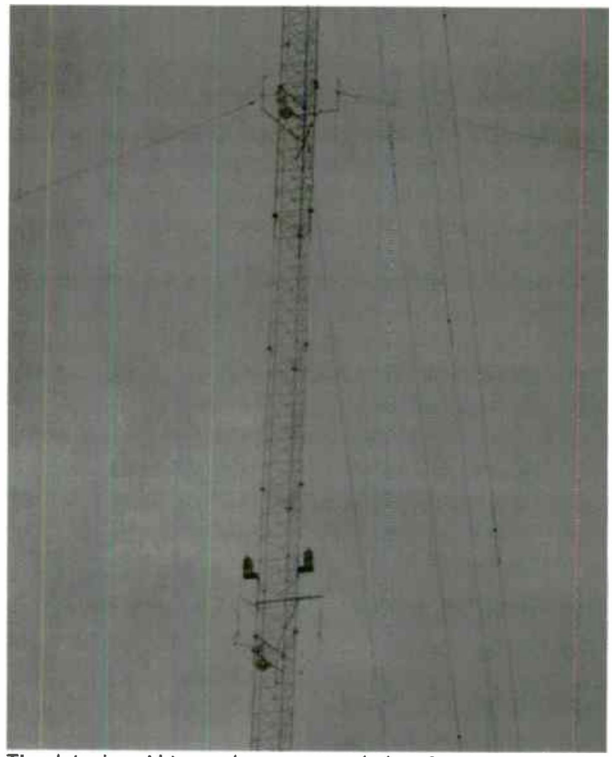
I spent the next several weeks getting things ready to start the proof. Getting control cables terminated at the towers and inside the building, the cable being a Belden steel armored, PVC jacketed cable rated for direct burial, with twenty 14 gauge wires. Making certain I had the correct transmission line connected to the correct place in the phasor. Verifying I had the sample lines labeled correctly. And taking care of various projects like getting the generator started for the first time, arranging for outside work to be done, getting our station handyman company to take care of other things. By the end of June, I was ready to bring in John Hidle of Carl T. Jones Corporation so we could "fire the mother up".



The Belden armored direct burial cable, consisting of 20 #14 gauge wires, which was used as control cabling to the tuning houses.

TUNING AND MATCHING

John Hidle and crew arrived and started putting the finishing touches on the tuning component cabinets in each tuning house. We needed to be a tad creative in getting at least one transmission line into the tuning house and it's first cabinet, but overall, this went smoothly. The next item on the list was adjustment of the detuning skirts on each tower.



The detuning skirts can be seen, one below the center beacons, one just above where the orange section starts above the beacons. You can also see the capacitor boxes.

Each tower in the array has two tuning skirts made of copperweld cable held out from the tower legs eighteen inches on insulators. These skirts are joined together and bonded to the tower on one end. They are attached to a capacitor enclosure which is bonded to the tower on the other end.

To adjust these skirts, Carl T. Jones Corporation constructed a giant toroidal transformer out of PVC hose. The hose was wrapped with fourteen gauge wire, and the diameter is roughly 9 feet, enough to easily fit around the tower and detuning skirts. The ends of this wire were attached to a BNC adaptor which was connected to a Potomac FIM-41 that the tower crew had hoisted up the tower. NEC modeling showed that the maximum current point in the tower was roughly 250 feet up, right under the sample loop.

The tower crew then set the FIM-41 for the correct frequency, either 1010 kHz or 1190 kHz depending on which skirt they were tuning, then cranked on the capacitor in the capacitor enclosure until they reached minimum indication on the FIM-41. It took two days to adjust all the skirts.

The next step was to detune the towers for 620 kHz, 1010 kHz, 1190 kHz, and 710 kHz, our frequency. While it sounds strange to detune for our frequency, it makes sense not only for the non-directional Proof of Performance, but for the fact that we can operate non-directional off of any of the three towers in the array for maintenance. While the detuning skirts minimized the current flow in the tower at 1010 kHz and 1190 kHz, the job wasn't finished. What this accomplished was that the electrical height of the towers was reduced at these two frequencies. We still needed to detune at the tower base.

To detune, the tower was grounded, and the FIM-41 connected to the sample line running down from the sample loop on any given tower. The FIM-41 was set to the correct frequency, and the detuning adjustments made to minimize signal on the FIM-41. We were able to achieve close to a complete null on all of the frequencies. Ungrounding the towers had very little effect on the signals being received on the FIM-41, and we took this as the detuning was good. Incidentally, the final settings were not that far off from the theoretical settings.

FIRING UP THE SYSTEM

Carl T. Jones had preset all of the phasor networks and tower tuning networks. We decided to bring the system up in directional mode slowly, and brought the transmitter up to one kiloWatt, obviously after we had signed the Lyndhurst facility off the air and made sure the Lyndhurst towers were detuned at 710 kHz. We saw that we had ratio and phase on tower number two, but nothing for tower number three. After verifying that the antenna monitor was working correctly by moving the tower two cable to the tower three position, we called it a night.

The phasor, where many a night was spent making adjustments.

The following day showed a short on the tower three sample line, which is what you would expect with the line connected to the sample loop. So we called Northeast Towers who immediately sent a crew down. They disconnected the sample loop, and we were still seeing a short at the base of the tower. The tower crew removed the connector from the line at the sample loop, and I removed the connector from the line at the tower base. We still had a short. I had a spare coil of sample line, and the tower crew began preparing to remove the sample line from the tower to install a new one.

When John Hidle and I attempted to remove the sample line from the feed-thru bowl on the output cabinet in tuning house number three, we found we couldn't budge the cable. We disassembled the bowl, and it took two strong tower guys pulling on the cable and bowl assembly to yank the sample line out. It appeared we took a lightning strike the day before, it arced over the sample line and welded it to the bowl feed thru, shorting the cable. The tower crew put the connector back together at the sample loop, but the connector in the output cabinet gave us grief. We finally made a nasty splice just to get us through adjustment, and it worked very well.

We came back that night and fired the system up again. Considering we didn't have the final cable run lengths when the system was designed (and the installed version varied slightly), the system came up with tower two being about ten degrees off in phase and about fifteen percent low in ratio. Tower three was roughly the same. Common point impedance was 54 ohms and +j20. Not bad. A little cranking brought the parameters right in and the common point to where it should be.

We then brought the system up to twelve point five kiloWatts and let it cook for about 30 minutes. When all was well, we brought it up to twenty five kiloWatts and let that cook. Finally. we went for the whole magilla and brought it up to 52,650 Watts. All looked good. Then we tried modulating it. The transmitter was most unhappy, so we backed power down to twelve and one half kiloWatts, left it on for about 45 minutes, then called it a night.

Equipment racks at the transmitter site.

TRANSMITTER ISSUES

WOR purchased two new Harris 3DX50 transmitters for installation in our Rutherford location. These transmitters unfortunately had some minor issues which I needed to work out. First, they were far too sensitive to VSWR, and did not care for the HD signal (or, from the paragraph above, they did not care for regular audio into a slightly funky load, either). I put at least 30 hours into adjusting the output monitor boards on both transmitters before they would operate reliably with HD.

The HD rack sits between the transmitters. Each transmitter has its own HD exciter.

Additionally, what we call transmitter A was set up for minimum settings at the factory, a far cry from how its sister transmitter was set up. Turns out we had a bad backplane for the exciter modules in this transmitter. Once the problems were ironed out, they have performed flawlessly.

MORE ANTENNA SET UP

The next numerous nights were spent balancing the system and setting the impedances looking into the tuning units to 50J0 on both directional and non-directional operation. After this was done, we proceeded to measure every conceivable parameter under operational conditions and made a record of them. The actual trick was that the impedance would shift when the Antenna Tuning Unit cabinet doors would be closed. We ended up measuring for 50J0 where the transmission lines come into the phasor cabinet.

The Harris 3DX50 is interesting. Rather than disconnect the transmission line between the transmitter and phasor and connect a very low power oscillator to perform an impedance sweep, we simply changed the operating frequency on the 3DX and ran the sweep at one kiloWatt. We found that the VSWR +/- 30 kHz was less than 1.4:1, which meets the IBOC specifications.

The sample system is interesting. We tuned the system up with sample loops, then performed the non-directional Proof. We then went out and did a rough proof on the directional. We found that the pattern appeared it would satisfy the CP, so we went to work finalizing the sample system.

Normally, the sample line comes off the tower and goes through an isolation coil such that the outer conductor at the top of the coil is at tower potential and the bottom is grounded. The bottom then feeds the sample line to the transmitter building. In the case of the WOR system, however, the isolation coils prove problematic. Leaving the coils in the circuit causes two things to happen. First, the detuning on 1190 kHz is not as good as it can be. Second, it narrows the bandwidth of the WOR signal, not good on any count.

After getting the system set up with the sample loops, we then needed to transition from the sample loops to toroidal transformers to feed the antenna monitor. This would take the isolation coils out of the circuits. We did this in three stages.

The base of Tower 3, the tower closest to the transmitter building.

We went one tower at a time, and disconnected the sample loop going to the building from the isolation coil, and connected it to a toridal transformer placed temporarily over the output pipe. We turned the system on and measured the parameters on the antenna monitor, then shut the system down. We did this on each of the three towers so that we now had the system running on the temporary toroids and also had a complete set of readings.

We then went to all the towers and, one by one, removed the isolation coil from the circuit, then made slight adjustments to the phasor to return to the readings we started with.

Once that was done and the readings were the same, we transitioned the sample lines to the permanent toroids located in the antenna tuning unit. Once this was done and readings were logged, we removed the temporary toroids and made slight adjustments to the phasor to bring the parameters back in. Mission accomplished.

WOR is inserting Program Associated Data (PAD) into the HD stream through a computer program written by SBE 15 student member Dennis Graiani.

DOING THE PROOF

Doing Proof measurements in northern New Jersey is a challenge. The close in points for the non-directional proof were a real challenge, as most of them were in the redevelopment area around the transmitter building. It's a challenge dodging dump trucks while you're trying to make measurements. All of the close in points were done with GPS because there simply are no landmarks if we had to do them over again.

The other challenging points were down the train tracks and one that was only accessible by walking one-quarter mile out into the swamp. It was ninety degrees that day, and it was a load of fun stomping out in boots, holding an FIM-41, a GPS unit and a notebook.

The other issues with Northern New Jersey are the simple density of population which makes it hard to move around efficiently, and the suspicious nature of most people. I had a point that I measured several times on a public sidewalk near a house. After the third time I was there, the owner of the house came out and told me to get lost. I handed him my business card, told him what I was doing, and informed him that I was on a public sidewalk and had every right to be there. He said he would call the cops. I said "go right ahead".

So when the police showed up, I handed them my business card and told them what I was doing. The officer happened to be listener, and said to the homeowner, "he has a right to be here and isn't hurting you." The homeowner became indignant, so I commented to the officer that "perhaps he thinks I'm from the cable company and is trying to hide something. Maybe we should get them out here to run some tests." The homeowner went white and returned to his domicile. Maybe he *was* trying to hide something.

The proof came out fine, and we needed to do a bit of augmenting because of the nature of re-radiation in northern Jersey.

A QUANDRY AT THE FCC

We were ready to file paperwork with the FCC, but had a problem. Actually, the FCC had a problem. The FCC can't issue a license for the new site or authority to operate until we present them with the "after proofs" on the six other AM stations our antenna system could potentially affect. But we can't do the after proofs until the array in Lyndhurst is taken down.

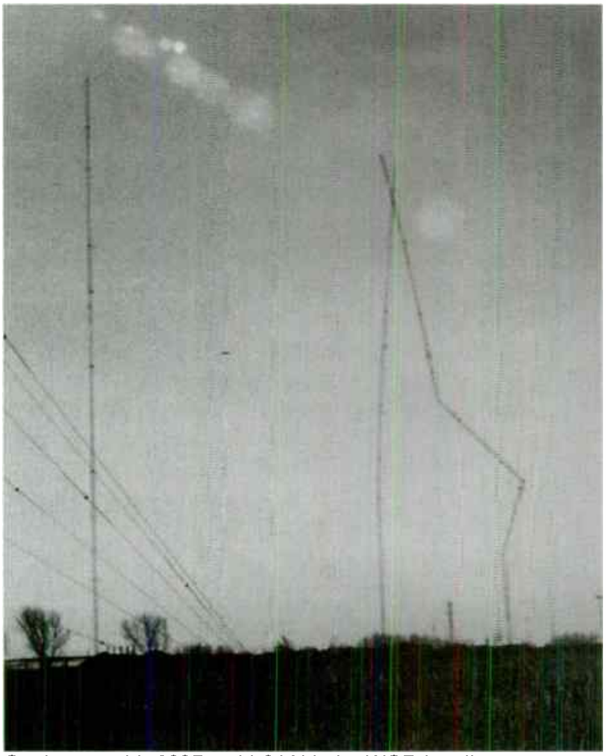
After much discussion between our attorney, Tom Jones and the FCC, it was decided that an STA could be issued to allow operation of the Rutherford facility providing that the filed proof on the new antenna satisfied the FCC and that no interference claims were filed against the new system. We would then be free to take down the towers in Lyndhurst. An STA was granted on September 8, 2006, and operation from Rutherford began full time.

TAKING DOWN THE TOWERS – NOT

A date of September 20, 2006 was scheduled for demolition of the towers in Lyndhurst. Meetings were held. WOR had not owned the Lyndhurst property for two years, so even though we were the licensee of the towers, it was not our responsibility to make notifications of the demolition. The towers are located in such an area that the New Jersey Turnpike is only ¾ of a mile from the location, and the towers were a landmark along the Turnpike. There is also a 100 year old water main for Jersey City that travels not far from the site, in addition to a gas main, and New Jersey Transit railway. I was assured that all notifications would be given, and we scheduled the date.

On September 20, we had 50 members of the press in New York City and New Jersey assembled, along with

Society of Broadcast Engineers members and invited guests. The towers were to be cut down at 10AM. Unfortunately, the owner of the property had decided that the notifications they were to give consisted of them closing Valley Brook Avenue, the location of the towers, on their own temporarily. WOR was on the air all morning talking about the event, and would be broadcasting it live.



On January 11, 2007 at 11:01AM, the WOR Lyndhurst towers fell gracefully to the ground.

The police chief in Lyndhurst walked into his office at 9AM to a ringing phone. It was the New Jersey Meadowlands Commission who heard us talking about the tower demolition on the air. "Who the hell OK'd this, and who issued an explosives permit?" The chief took off across town to the tower location and put a stop to it. Incidentally, we were not using explosives, and no one seems to know where that came from.

After several months of meetings, building officials, safety officials and the police were satisfied. All notifications went out. News stories hit every newspaper in northern New Jersey, and I must have given a dozen interviews.

On the crisp, clear morning of January 11, 2007, the guy wires were cut on Lyndhurst towers 2 and 3. All the Lyndhurst towers were 689 feet tall, were 42" face, and had dead weights of over 85,000 pounds with guy wires attached. The reason both towers 2 and 3 were taken down at the same time is that one set of guy wires from each tower interlaced. The tower crew did not want either tower to lose momentum on the way down by getting caught in the wires, or snapping the wires on the opposite tower which could have been catastrophic. So they cut the guy wires on both towers on the legs away from the center of the tower field. Both towers fell gracefully into heaps in the tower field. Total time for all three towers was approximately 35 seconds.

The remains of Lyndhurst tower 3.

The crew then moved onto tower number 1. It did a graceful foldover, hit the ground lawn-dart style, and proceeded to collapse. It is interesting to note that the tuning houses for towers 2 and 3 did not have a scratch on them. Tuning house number one was hit almost dead center by guy wires and looked like a tin can that had a cherry bomb placed under it.

Lyndhurst tuning house 1 was hit by guy wires. The other two tuning houses survived without a scratch.

With the Lyndhurst towers gone, the after proofs could be done and the FCC paperwork finally completely filed.

SATISTYING THE RUTHERFORD BUILDING OFFICIALS

At the time we turned on the Rutherford site permanently, we did not have a Certificate of Occupancy. Very simply, we were playing a game of "Beat the Clock" with time running out on the FCC Construction Permit. I could either pay attention to the building or to getting the FCC work done. If we missed the deadline, the Construction Permit would not be renewed.

Luckily, the state and town officials are the type who want to work with you rather than against you. As of this writing, we are completing several items that the Borough of Rutherford wants done on our building to get their Certificate of Occupancy. We have a temporary CO from the New Jersey Meadowlands Commission with requirements that we install the permanent culvert at the crossing of the waterway into our facility, pave the parking area, and fence the transmitter building compound.

One issue that we still haven't yet figured out is what to do about E911. The "road" our site is on is a construction road, and won't be a real town road for at least another year. So we don't have a street address. As I said to the fire marshall, "this has been in the works for five years – why is it a surprise and a problem now?" We shall see. There are always ways around these things.

CONCLUSION

Completely rebuilding one of the country's Heritage radio stations in a period of two years has been a challenge. It is definitely something that is not for the faint of heart.

The author, Tom Ray, with the remains of Lyndhurst tower 1. A twisted man with his twisted tower.

WOR is now completely new from the tips of the microphones to the tops of the towers. The studio facility is completely digital and state of the art. The audio produced in the studio facility is superb.

The WOR transmitter site was designed from the ground up with digital in mind. and in fact signed on with an HD signal (we did the proof with the HD on, as well). The transmitter site has several audio paths from the City, and a studio, complete with call in lines, just in case. Easily, WOR is prepared for whatever may come down the road in the future.

I'm glad I was able to be the one to make it happen.

WORKING WITH THE FREE SOFTWARE COMMUNITY TO REPLACE A RADIO BROADCAST AUTOMATION SYSTEM

Federico Grau
Radio Free Asia
Washington DC
2007-01-19

ABSTRACT

Like many radio stations, Radio Free Asia (RFA) has a minimal operating budget and an aging radio broadcast automation system. While their current system is all digital and has served them well, it is reaching the end of its life cycle. For their next generation replacement they searched the Internet for available Free Software components. The most mature project found is the Rivendell radio broadcast automation system authored by Fred Gleason of Salem Radio Labs, and released under the GNU General Public License (GPL).

This paper will start with a review of what Free Software is, how an organization can use it, and how and why to contribute back. In the case of Radio Free Asia, one example is their investment into modifying the Rivendell system to work in a multi-user environment. As part of their deployment, RFA has integrated several Free Software components such as the Linux High Availability (HA) project and the Linux Terminal Server Project (LTSP) to produce a robust working solution.

The paper will also elaborate on the differences in working with a Free Software project versus a commercial solution. For example, when it comes to support, with a commercial solution a problem will be approached by dialing a dedicated support phone number (and possibly waiting for a solution); with a Free Software project more effort and interaction is possible and expected. Avenues of support with Free Software include participating in mailing lists, reading documentation on community wikis (web sites that allow visitors to easily edit the content of the web pages), reviewing the source code, and working with consultants who know the field. In addition to these steps, RFA has helped establish a dedicated Internet Relay Chat (IRC) channel for the project, where users and developers can help support each other in real time.

HISTORY

Radio Free Asia

RFA is a large organization with over 50 studios, a staff of 300, transmitting 36 hours per day in 9 different languages. They are a non-profit organization and therefore have a fixed budget to work with. As a result there is a strong "do it yourself" attitude and an openness to trying new ideas and technologies. The facility has been all digital since its inception in the 1996. RFA has a history of using and working with Free Software since the late 1990s, documenting and sharing their efforts with the community in the Radio Broadcast Open Source System (RBOSS) hosted on the techweb.rfa.org server.

Their current broadcast automation system is a commercial all digital solution that has worked well, but it is almost 10 years old. There are increasing hardware failures and software support is reaching its end of life cycle. Because of the scale of the organization, the project of upgrading the facility is a large and challenging endeavor.

Free Software

The ideas behind Free Software were published by Richard Stallman (RMS) during the early 1980s as he observed software distribution practices changing. Where previously computer source code was commonly distributed along with software, allowing one to study and improve a computer program, this was becoming less and less the case with the onset of the microcomputer. RMS saw the importance in allowing one to see computer code as society and the economy become more dependent on computers.

The key ideas behind Free Software are to ensure freedoms as in liberty, "libre", or free speech. They do not necessarily guarantee free as in cost, "gratis", or free beer. Those freedoms are listed in Table 1 below. They encourage sharing and cooperation and result in continuously improving software as long as a community thrives around it. Software has a fixed amount of effort that is required to produce a quality product; one example is the office suite which was functionally complete around the release of Microsoft® Office version 95. With the mass collaboration possible via the Internet, very high quality results can come from Free Software, some projects more further along than others.

Table 1: Guarantees of Free Software

• to run the program, for any purpose
• to modify the program to suit one's needs (this means access to the source code)
• to redistribute copies, either gratis or for a fee
• to distribute modified versions of the program

Recognizing the importance of Free Software, in 1984 RMS started the GNU's Not Unix (GNU) project to create a complete UNIX-like operating system comprised of Free Software. To this software project he applied the GNU GPL which ensures the guarantees listed in Table 1 above. The GNU project has evolved into one of the key components in many GNU/Linux distributions available today such as Debian, Ubuntu, SUSE, Redhat.

Free Software has grown in popularity with the development of the Internet – both acting as catalysts to help the other grow. The Internet has allowed groups from all over the world to cooperate and produce more software projects. Similarly, the Internet is based on open standards and many portions run on Free Software, which has a large collection of programs that work with those standards. The most popular example in the media today is the Linux kernel, most commonly shipped in GNU/Linux distributions. The number of Free Software projects is in the thousands and growing (there are 42,000 projects registered at freshmeat.net and 139,000 at sourceforge.net).

So what drives one to use and develop Free Software, especially when there is less of a guarantee to profit? Often the motivation is to "scratch an itch": solving one's own computing problem by writing some software. A good programmer knows how to write code, a better programmer knows how to reuse code. Using Free Software, there is a growing base of work done by others on which to build upon and a culture that encourages further cooperation.

The result is a work of passion, necessity or both. Unlike commercial ventures, it is not restricted by budgets or marketing deadlines. What ties the community together is not profit, but a common interest and goal (such as broadcasting effectively).

Radio And Free Software

Radio and Free Software are well suited to one another with commonalities in cultures. Hacking, or modifying a tool to better fit ones needs, has been a long term practice in the broadcast arena. When a new piece of equipment would come in, engineers would study the technical manuals and schematics that were standard with pro-audio grade gear. This would enable them to learn how the unit worked and allow them to support it. They might make improvements such as adding buttons to a console, replacing the heads on a tape machine, or adding a remote control. In the 21st century an increasing number of the tools in the broadcast field are virtual and software based. Consoles are now Internet Protocol (IP) based; recorders and editors are digital software programs instead of physical machines. The lines between engineer, system administrator, and software developer are blurring. Using Free Software there is opportunity for modern radio broadcast engineers to see and hack the code of the tools they use to better suit them. The source code is the ultimate technical manual and schematic defining how a software system works.

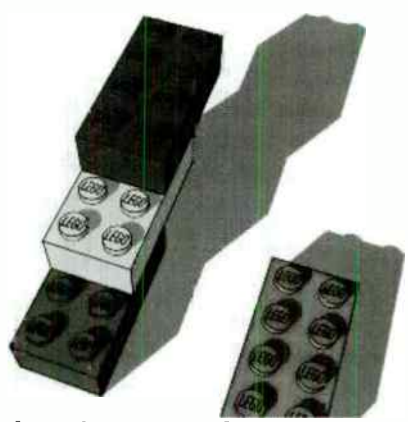
The radio industry has a thriving community of organizations using, sharing, and developing Free Software. They cluster around projects that share common goals, such as the Rivendell broadcast automation system. The #rivendell IRC channel, has participants from Eastern, Central, and Western US, the Caribbean, Spain, England, France, Germany, Hungary, Mexico, and the Philippines, to name a few.

THE SWITCH

Design

Before starting the Rivendell project, RFA had been working with a commercial vendor to replace their broadcast automation system. During the testing phase, the project came to a halt when, due to a faulty driver that would not be corrected in the foreseeable future, a cost prohibitive hardware replacement was required. RFA was in an adverse situation beyond their control. They reviewed their options, and while the use of Free Software was not as feature-rich as the commercial offerings and it required some initial effort, it positioned them on a path that guaranteed security and control over their organization. The software has proven to be robust enough for the broadcast environment, and its feature set continues to improve.

Rather than working with a vendor and adapting their work model to the the vendor's product line, RFA began a twofold study. First a review of audio related projects available upon which to build. Members of the community, like Dave Phillips and his site linux-sound.org, have filled the role of exploring, reviewing, and summarizing audio software that the community develops. With the open development process that is common of this community, one can review and track the progress on most any component of the software system, even down to the most low level driver. While it is highly technical with a very steep learning curve, it does allow for direct contact with the developers when trying to expedite the fix of a critical bug.



The second study RFA performed was a review of their current work flow. A migration plan for current and long term deployment was developed. Features were evaluated and prioritized whether they were critical requiring completion before the initial deployment or if they could be postponed as a later improvement.

The next step was to pick out the best components and assemble them into a working system, a bit like building with LEGOs. For simple configurations the blocks to assemble are straightforward; download and install the SUSE Linux DVD iso, and then download and install the Salem Radio Labs packages for Rivendell. After a little configuration, most of it GUI based, the system pretty much takes care of itself. A more complex solution was required for the RFA deployment to work well with their environment, provide redundancy, and better support a multi-user environment.

Components that needed to be improved for the delivery plan were identified by RFA staff. Some effort was done by in house engineers; examples include creating a Plugable Authentication Module (PAM) library to better control Rivendell authentication for multi-user environments, and improving the installation by adding wizard dialogs rather than requiring editing of a configuration file. Other efforts were outsourced to consultants, the main example is an improved integration between the Audacity multi track audio editor and the Rivendell system. When completed these improvements are contributed back to the community, improving the product for the group at large as a result of RFA scratching its itch.

To the good surprise of RFA, other items that had been planned as "later improvements" for their deployment system have been contributed to by other members of the Rivendell community. The ability to store audio in the compressed audio formats Ogg Vorbis and FLAC are examples of new features contributed by others in the community while RFA was in their testing phase.

With commercial software one has to use their products in the ways designed and planned for by them. With some thought and Free Software a better fitting end solution, one that is capable of evolving with an organizations needs, can be delivered.

Building Blocks

How does one obtain Free Software? Individual projects will often host information, source code, and documentation on Internet web pages available for download. Distributions are projects that collect many components and then prepare them by compiling and packaging them to be easily installable. RFA used both types of resources in their deployment.

The centerpiece of the RFA deployment plan is the Rivendell broadcast automation system. This is the portion that does the actual play-out to broadcast. Table 2 lists many of the pieces used in the deployment.

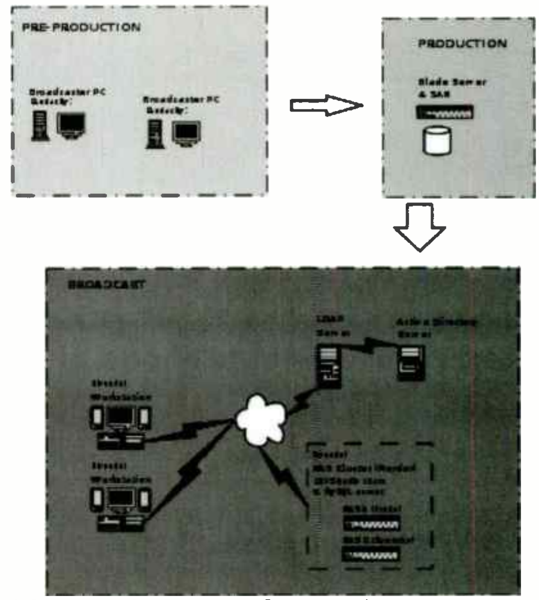
Table 2: Software Components Used

Rivendell – Broadcast automation system with a virtual library of carts, log creation, and play-out.
Audacity – Multi track audio editor, used to create and modify content.
KDE – User friendly desktop environment, with a user interface similar to the MS Windows® desktop.
Samba, NFS – Network file sharing for MS Windows® and GNU/Linux computers.
Linux High Availability project – used to provide redundancy with the back end servers.
LTSP – Linux Terminal Server project, that allows for disk-less deployment of workstations enabling easier administration in large deployments.
Linux Kernel – Core software component that talks to the hardware.
ALSA – Sound driver portion of the Linux kernel.
JACK – Low-latency audio server, routing audio amongst multiple programs and a sound card.
Portaudio - Portable audio library/API, used by Audacity.
Ardour – A full featured digital audio workstation.
OpenOffice – Free office suite.
Firefox – Web browser.
Gaim – Instant messenger chat utility.
Debian – Distribution that groups most of these packages together.

Support

A great part of the value that is paid for in a commercial product is the support. It is to receive support and continued updates that recurring yearly maintenance contracts are made.

Rivendell Workflow and LAN diagram



With Free Software there exist more avenues for support. One can purchase support contracts for Free Software products from large organizations such as SUSE, IBM, MySQL. Support can also be purchased from consultants specialized in the field on an as needed or contractual basis.

Users also have the option of helping themselves. The same tools of the Internet that allow developers to cooperate and build the software, allow users to interact and support each other in a distributed fashion. Some of those tools include email lists for discussion of problems and wikis. Wikis are excellent tools for allowing collaboration on documentation. IRC is a common group chat forum used to discuss problems in real time.

Finally, with Free Software there is always the source code available to review. While the idea may seem daunting to the uninitiated, good tools make exploring easier. Much can be learned by searching the source code for an error message a program produces, and then scanning the nearby code for comments written by the developer.

On the flip side, ways of supporting a project include micro-payments or sponsorship on web sites hosting the project. Similarly, some projects sell CDs with their software. There are also non-monetary ways of supporting projects such as offering bandwidth or server space to host a project on the Internet. Hardware manufacturers have incentive to share sample models with established developers so they can author drivers. This results in a larger group of people buying and using the hardware device.

Returning to the idea of scratching an itch, RFA recognized a need for an improved support community around the Rivendell project. They researched what was already present and found the mailing list and community wiki. They began with participation in the email list and adding documentation to the wiki. Seeking a more live support forum, they fostered the IRC channel #rivendell on irc.freenode.net . The result has been a growth in participation on the IRC channel from 2 or 3 people one year ago to over a dozen regular users currently. With more users, the greater the speed of getting a response to a problem.

CONCLUSION

With the switch to using Free Software for their broadcast automation system, RFA has transitioned from being in an adverse situation they could not control, to now having direct influence in the development of their core infrastructure. No longer will they have to worry about product availability changing, or worse disappearing, due to market fluctuations beyond their control such as a company being sold or going out of business. Artificial constraints that limit or lock in a user to a piece of software are no longer a concern. No key dongle is required for the system to function. File formats that are used are open and well documented allowing interoperability, not restricted with expensive patents.

Benefits include a tight integration path throughout their organization In time the broadcast system will have bindings to other systems such as the WWW Content Management System (CMS). Support now includes dedicating some staff to the project and use of consultants. Response time to problems is quickened and there is some say in the future direction of the product. Their investment in thinking has saved them money. To an organization this ensures security of ownership, control. Combining multiple Free Software components and working with the community creates a powerful synergy that has helped deliver a working broadcast solution for RFA.

REFERENCES

Open Sources
O'reilly 1st Edition January 1999
http://www.oreilly.com/catalog/opensources/book/stallman.html
http://www.oreilly.com/catalog/opensources/book/perens.html
ISBN 1-56592-582-3, 280 pages

Special thanks to the diverse staff at RFA who have contributed to this project.

Systems Integration Basics for Radio Engineers

Mario Hieb, P.E.
Consulting Engineer
Salt Lake City, UT

You cannot plan too much. Nothing will speed up the process, simplify the work, save you money, get you on budget and in on time, and allow you to sleep at night as good planning. Planning is the virtual installation, the installation in your brain, and the brains of those around you, before the work begins.

Stephen Lampen, Audio/Video Cable Installers Pocket Guide

ABSTRACT

Systems integration project management is common in large-scale, high-dollar broadcasting projects (typically video systems) but is not as common in radio projects. These time-tested techniques can help radio engineers develop higher performance systems, become more organized, work more efficiently, meet deadlines and budgets, and be more professional.

Systems integration techniques can help the radio engineer with difficult projects, such as moving a station while staying on the air or the transition to HD radio.

Systems integration techniques include:

- Project management.
- "Scoping" the project.
- Developing and maintaining systems documentation.
- Systematic design, construction and testing of audio, RF, communications, control, data and other systems.
- Systems integration standards.
- Design approvals.
- Estimates and budgets.
- Managing expectations.
- Managing sub-contractors.
- Managing "your client."

PROJECT MANAGEMENT

Project Management is a vast, complex subject unto itself. There are tools used by professional project managers such as Gantt Charts, Critical Path Analysis, etc. that, although useful, typically are not necessary for smaller broadcast projects.

For the purposes of this discussion, Project Management will be defined as the role the radio engineer plays in systematically executing the project. Also in this discussion, your boss, manager, client, etc. will all be referred to herein as "the client."

"SCOPING" THE PROJECT

The first step in any project is to define, or "scope" the project. The following questions need to be asked and answered:

- What is the project?
- How big is the project?
- Where is the project?
- What is the time frame of the project?
- What is the budget for the project?

All this may sound trivial, but the person requesting the project and the person managing the project need to be in complete agreement on these points, and the "scope" needs to be documented.

"Scope creep" is when a project is "scoped", yet the parameters somehow change as the project progresses. One of the prime tasks of a project manager is to manage this "scope creep." It is important that once the project is "scoped", changes will occur only if they are defined in writing or on drawings, and are approved by the client.

Other people, such as architects and other engineers, may also be involved in the project. Work with them early to define the project scope.

SYSTEMS DESIGN AND DOCUMENTATION

The single most important aspect of project management and systems integration is documentation. This documentation locks in the scope of the project and defines the deliverables.

The project needs to be "built" on paper before any other task can occur. This documentation package usually consists of:

1. Architectural drawings, including floor plans, elevation drawings, etc.

2. Engineering drawings, including electrical drawings, mechanical drawings, etc.

3. Broadcast systems drawings. Including:
a. Audio signal flow drawings.
b. RF signal flow drawings.
c. Control signal flow drawings.
d. Communications signal flow drawings.
e. Computer network signal flow drawings.
f. Power system drawings.
g. Rack/cabinet elevation drawings

4. Custom design. Including:
a. Panels.
b. Cabinets.
c. Special fixtures.

5. Equipment spreadsheets.

6. Wire-run spreadsheets.

7. Systems test documentation.

8. Project timelines.

9. Contracts.

10. Permits.

11. System integration standards.

12. Equipment data sheets.

13. Labels.

14. Change order form.

15. Punchlists.

A complete documentation package locks in the project scope and is critical in the next stages of systems integration. Drawings can be created on AutoCAD or equivalent software. The drawings can be created for any size paper; personally I find 11" X 17" the most manageable size.

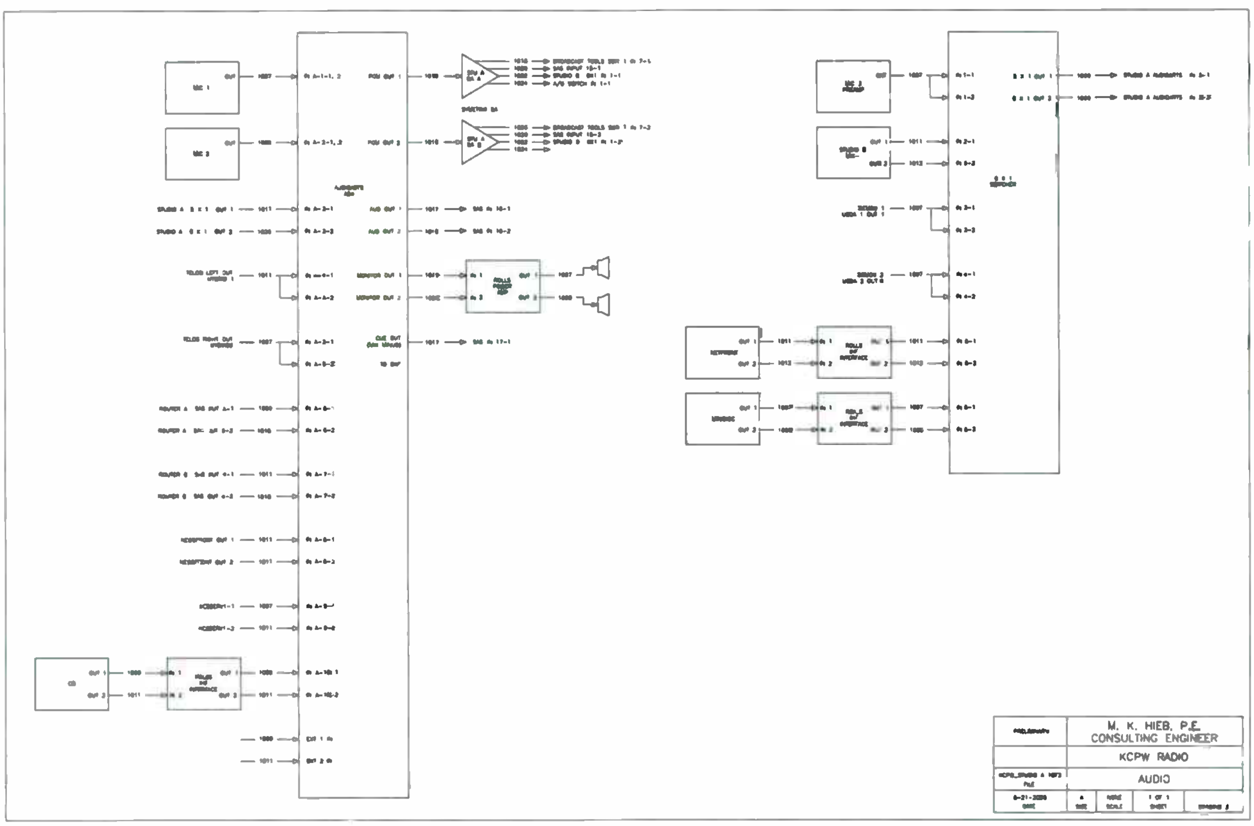


FIGURE 1: Audio Flow Diagram

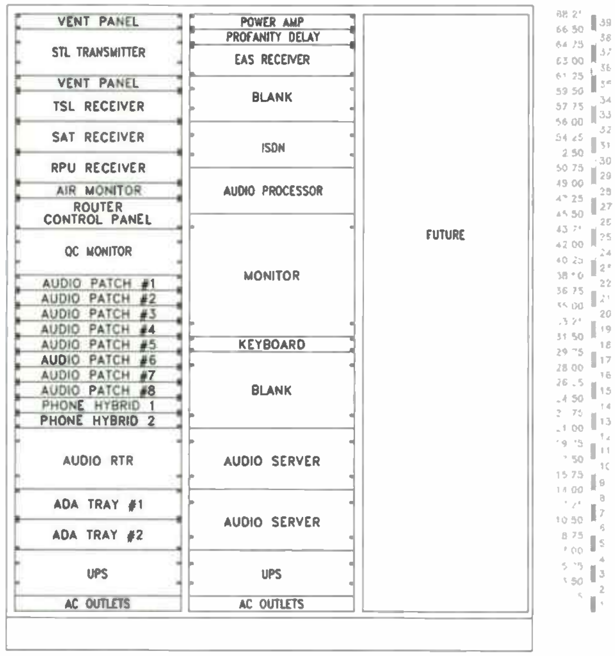


FIGURE 2: Rack Elevation Drawing

If you don't know how to use drafting software, you can often find a freelance or student CAD operator to help you out.

Figure 1 is an example of an audio flow diagram created with AutoCAD.

Signal flow drawings don't need to show every detail of the installation, but should communicate the essentials. All equipment should be shown as blocks on the flows, with interconnect wiring and wire numbers. The signal flows should read from left to right and from top to bottom. Rack elevations show where the equipment is located in a cabinet or metal rack.

Figure 2 is an example of a rack elevation drawing created with AutoCAD.

After the system is assembled, these drawings can serve as a document to track testing of the system. Once system testing is complete, the amended drawings, or "as-builts" will serve as the final documentation.

Every equipment item involved in the project should be included in an equipment list, including any existing equipment that is to be integrated. Don't forget equipment racks, plug strips, blank panels, rack screws, connectors, rack shelves and other associated hardware. A typical equipment list includes item, quantity, location, system, make, model, vendor, part number, price each and price total.

Wire run spreadsheets provide the detailed information that the flow drawings don't provide. This information can consist of cable source and destination, wire number, cable type, cable length, connector type, pin-out, etc.

SYSTEMS INTEGRATION STANDARDS

Specific details of the project should be included in the project documentation such as:

- Audio level standards: i.e. 0 VU = + 4dBu
- Connector wiring standards: i.e. XLR pin 2 = +
- Signal grounding standards: i.e. telescopic ground, etc.
- Power grounding standards: i.e. star ground, isolated ground outlets, etc.
- Patch panel wiring: i.e. half-normalled, etc.
- Cable types: 9451, RG-8, etc.
- Acoustical ratings.

Defining standards is useful when more than one person is doing the installation.

DESIGN APPROVALS

I like to involve station personnel, including program directors, air staff, etc., in the design process. Your client should approve your design drawings before they are let for bid.

BUDGETS AND ESTIMATES

Because you have already built the project "on paper," estimating the cost of the project is fairly straightforward. Contractors can prepare a bid from the drawing package you provide them.

Your equipment list can be given to several vendors for competitive bidding and package discounts.

Don't forget that some items require longer lead times than others.

MANAGING EXPECTATIONS

Often, a project manager is expected to do the impossible, such as meeting unreasonable deadlines or working with a budget that doesn't match the scope of the project. A well prepared documentation package can help communicate the cost or complexity of a project. If the project scope exceeds the budget, work with your client to eliminate items, or relegate them to future status.

Include a contingency amount in your estimates; 5 to 10 percent is typical. It is always better to bring a project in under budget rather than over budget.

Too many engineers are asked by their client, "what does a studio cost" or "give me a ballpark estimate". I usually don't answer these questions directly because, unless I've done a complete design, whatever number I come up with will most likely be wrong. Instead, I may answer "when WXYZ built their facility, it cost them $300,000."

The best thing is to design the facility to a budget number, or to give an accurate estimate based on the predetermined "scope" of the project.

"When will it be done" is another question that is usually brought up. A realistic timeline is a good document to show management; it helps communicate the complexity of the project.

Figure 3 is an example of a project timeline.

The most important thing is to communicate the project to your client. Showing them the drawings, equipment lists, time lines, etc. shows that you are organized and in control of the project.

CONSTRUCTION

Because you've planned well, construction should be smooth and uneventful.

MANAGING CONTRACTORS

Unscrupulous contractors can spot an amateur project manager a mile away. The more professional you are in your preparations, the less the likelihood they will take advantage of you. Insist that their bid be based on your drawing package and include in your bid package language stating that any changes to the project will require a change order signed by the client.

Contractors can also promise a completion date, yet deliver another. If timing is an issue, put the completion date in the contract. Another tactic is to ask for completion a week or two before you really need it. Having a timeline on paper also shows that you are serious about timely completion and it can even be made part of the contract.

Before you sign off on the construction, go through the construction drawings and make a "punch list" of any items not completed, or completed satisfactorily. Insist that these items be completed before final payment is issued.

SYSTEMS INSTALLATION

Systems installation consists of mounting the equipment, running the appropriate wiring or transmission line, terminating the wiring to punch blocks, patch panels, connectors, etc.

One often overlooked aspect of systems wiring is cable labeling. Each cable in a system should have a unique wire number label (on both cable ends) that correlates with the flow drawings and wire run list.

Other information can be included on the wire label, such as source, destination, input, output, etc. Labels can be printed on most inkjet or laser printers. Some installers prefer printed heat shrink labels.

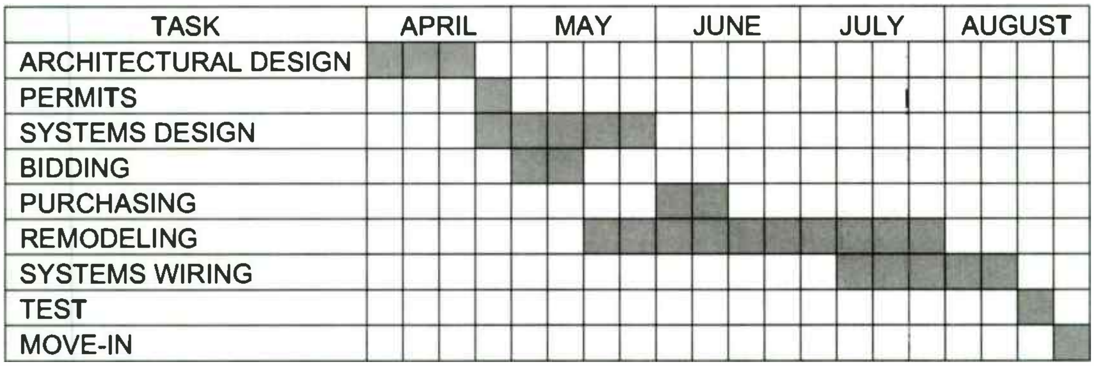


Figure 3: Basic Project Timeline.

Patch panel labels can be created with AutoCAD and can be printed on paper or Mylar.

Once all wiring is in place, it is tested. Then at a selected time, often in the middle of the night, the old equipment is moved to the new facility, mounted in the racks or cabinetry, and connected to the previously installed cabling.

Occasionally, a system will be constructed where the majority of equipment is currently in use. In this case, a "phantom" installation is done, where the wiring is run to a rack or cabinet location before the equipment is installed.

Doing a professional job of installation communicates your competence to your client. They may not understand what the equipment does, but if it looks nice, they'll feel better about the money they are spending and their decision to hire you

MANAGING SYSTEMS INSTALLERS

Sometimes the engineer installs the systems, but in other cases it is quicker and more cost effective to hire an install team.

When installers are used, it is important to have all of the documentation in order and that everyone have the same set of plans. The installer should not need to make decisions about cable type, connector wiring, etc.; this should all be pre-determined by the systems designer.

Good free-lance broadcast installers can be found in most major cities, and will often travel to a jobsite. Individuals with no radio background, but good soldering and cabling skills can be employed with proper supervision. An unskilled individual I once hired for a radio station build-out is now the installation supervisor for a major video systems integration firm.

SYSTEMS TESTING

Every cable in a new facility should be tested. Wiring errors, like a lifted leg in a balanced line, can create phase and level issues even though the audio is present.

First, configure any pertinent equipment and set audio levels. Next, analog audio cabling should be tested for level and phase. This is easy to do, just take your "as-built" system flow drawings and highlight each wire as it's checked, working from one end of the facility to other in the direction of signal flow.

If half-normalled patch panels are used, the top row can be used to bridge signals with a test set. An "end-to-end" test is useful to determine signal noise floor, aggregate distortion level, etc.

RF cables can be tested with an RF source, dummy load and VSWR meter. Network cables can be checked with a test set.

FINAL DOCUMENTATION

Any changes made during the system installation should be documented, with changes written in red ink on the flow drawing. Later, these changes should be made to the master software copy and marked as "as-built."

As-built copies should be given to your client as one of the deliverables.

CONCLUSIONS

Systems integration techniques may seem like a great deal of work, but they save time and money. The resulting systems are better planned, better built, more reliable and easier to troubleshoot. More importantly, the engineer has better control of the project and is viewed by the client as more competent, effective and professional.

BIBLIOGRAPHY

Stephen Lampen, *Audio/Video Cable Installers Pocket Guide, 2002*

AN EMBEDDED SYSTEMS APPROACH TO IP-BASED BROADCAST FACILITY REMOTE CONTROL

Stephen Dinkel
Burk Technology
Lee's Summit, MO

Nathan Burk
Burk Technology
Littleton, MA

Jonathan Burk
Burk Technology
Littleton, MA

ABSTRACT

HD Radio, DTV and multicasting have combined with the realities of downsized engineering departments to raise the expectations for broadcast facility remote control. At the same time, LAN/WAN connectivity has invited vast improvements in the way broadcasters collect, distribute and manage data. An embedded systems approach to IP-based broadcast facility remote control is the outcome of these dynamics. Combining IP architecture with a self-contained, embedded design offers broadcasters major advancements in alarm propagation, alarm aggregation and multi-site operability. Exploring this approach explains how to adapt broadcast facility management to today's technologies while satisfying reliability requirements in ways where PC and server-based solutions have fallen short.

DESIGNING FOR REMOTE DATA NETWORKS

Consolidation presents a major challenge to traditional remote control. Bringing digital buildouts into the mix further underscores the need for scalability and a broader approach to managing remote sites. With engineering attention divided among more stations and increased need for high level corporate overview, the point-to-point approach to site management had to be expanded while still maintaining the reliability of a closed system.

The ability of new solutions to meet the need for scalability is directly linked to the availability of faster, more efficient communication paths to remote sites. Solutions are available that leverage existing broadcast infrastructure to lower upfront costs. Satellite IP and cellular GSM/GPRS links have also gained acceptance in the broadcast environment, and new technologies such as WiMAX are gaining ground. Moreover, many groups have already invested in infrastructure for email, mobile messaging, corporate LANs, VPNs, Blackberries, etc. Bringing remote control onto the same canvas adds to the return on such an investment. The challenge lies in delivering these solutions to the realm of RF engineering, demonstrating broadcast-grade reliability and proving future readiness.

Navigating the IP Landscape

For all the advancements in IP deployment, migrating to an IP-based system for remote broadcast facility management can be more complicated than the service providers, equipment manufacturers and end users would like it to be. Even if a high speed link exists between studio and transmitter, program needs must be met first, often leaving little bandwidth for other applications. Dropped packets, network downtime and latency all challenge a broadcaster's confidence in using IP for broadcast remote control.

However, these obstacles to using IP for remote facility management can be overcome. Broadcasters understand this as one more extension of the marriage of RF engineering and IT engineering. For manufacturers, the challenge is to understand that IP-based equipment needs to live in the reality of broadcast networking, not the "universal broadband" model of the software industry.

MAKING LOW BANDWIDTH WORK

This distinction between developing IP-based products for the broadcast industry and designing for the corporate/IT environment is critical, and is what led us to the embedded system model for IP-based broadcast facility remote control.

In a networked environment, it is easy to think of bandwidth as unlimited. However, there are two problems with this assumption. First, networks are shared by many devices that compete for bandwidth. Second, bandwidth at remote sites is far more limited than it is in the office. Any real-time control device running over an IP network cannot assume that 100Mbps are available for its use.

In designing an IP-based remote facility management system, the realities of bandwidth limitations required us to design a protocol that allowed maximum data content in each packet. Comparing this approach to SNMP showed that OID strings generated larger packet sizes, consumed more processing power, and required substantially more bandwidth to achieve scalability. Fig. 1 compares the embedded system to an SNMP-based system to show that even if the pipe is narrowed to modem or ISDN speeds, the embedded system remains exceptionally scalable. Calculations at DSL speeds and faster revealed similar results: over a T1, the embedded system showed more than six times the site capacity of an SNMP system.

TCP/IP, although reliable, requires too much processing overhead to allow for ample scalability. We opted to build a design that utilizes UDP/IP, taking advantage of its leaner connectionless model. The downside is that UDP packets are not acknowledged by the receiving network node. To overcome this obstacle, we added application layer acknowledgements to the protocol to ensure commands and data updates are received by their intended recipients. This results in lower overhead, while still safeguarding against dropped packets.

Using a peer-to-peer architecture further economizes communications. Unlike an approach where passive interfaces report telemetry back to a central server, a peer-to-peer network takes the remote server out of the equation, removes a point of failure and eliminates the extra "hop" in the data path that a server creates. This model places additional burden on protocol efficiency because without a central server, there needs to be a way for multiple clients to access the same data simultaneously. This includes software applications, web browsers, or other units in the system. To address this requirement, a subscription model is used, where the embedded system adds to the data stream only those parameters requested by the clients. For example, if a user at the front panel selects a new site for display, the system terminates the subscription to the previous site's data and provides the data for the newly selected site instead. Data is not added to the stream if it is not needed for display at that moment. This is how an embedded system, without the aid of PCs or servers, can operate at modem speeds and still encompass more than 3,000 discrete monitoring and control parameters.

The efficiency of such a design supports site-to-site control over low speeds, which is important when considering system operability over existing data links.

IP Over Serial Links

Where IP connectivity is not available, a serial link may be used with the addition of a learning serial-to-Ethernet bridge. The serial-to-Ethernet bridge extends a LAN to a remote location over a serial link, at speeds as low as 4800 baud. Unlike a serial-to-Ethernet converter, which locally transforms serial data into TCP/IP and requires an existing IP path, a serial-to-Ethernet bridge connects two LANs and operates on a serial link such as one provided by a digital STL or dedicated telco

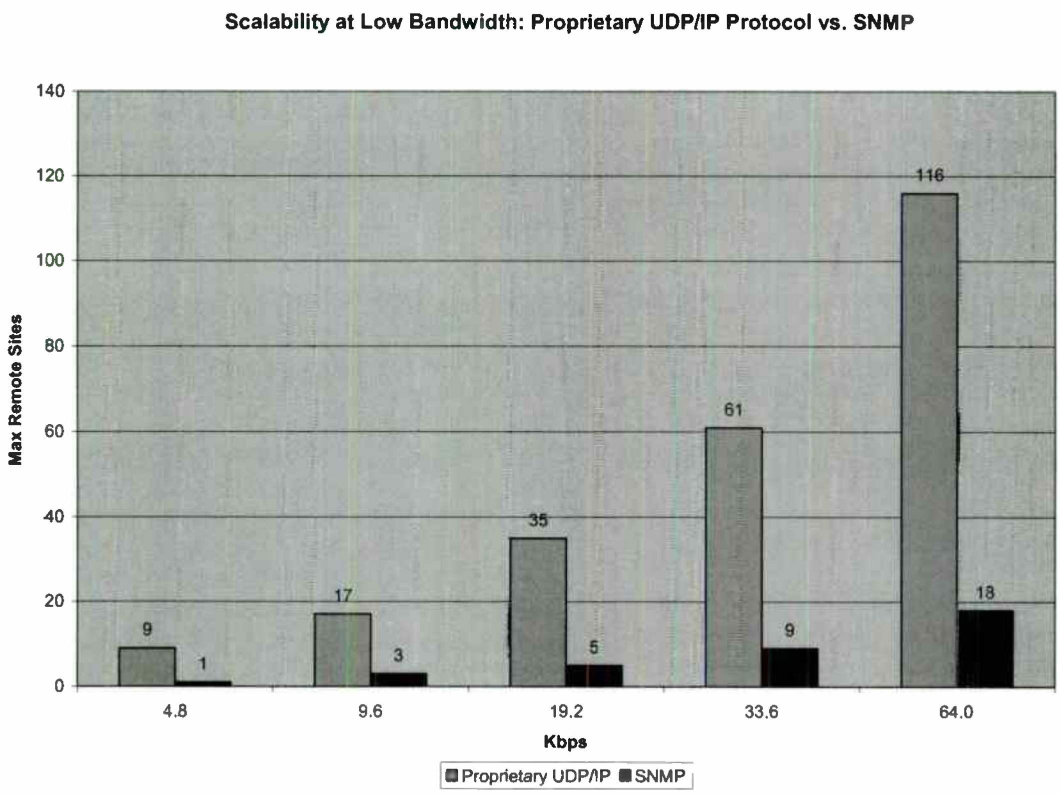


Fig. 1: Embedded System Scalability. Scalability at various modem/ISDN speeds of an embedded system with proprietary UDP/IP protocol compared to SNMP monitoring. Calculations assume 32 channels of analog metering updating every second.

link. To optimize throughput, the bridge "learns" the location of each MAC address, and transports only that network traffic which is destined to the other side of the bridge. With such a solution in place, the convenience of IP addressability at the remote site is achieved without the expense of broadband connectivity.

Integrating SNMP

With so much emphasis on using custom protocol for bandwidth economy, it was important for us to keep the door open for integration with standard protocols like SNMP. Even though the larger packet size and slower data polling make SNMP less suited for real-time monitoring and control of the broadcast network, SNMP bridges the IT and broadcast engineering worlds and provides access to a broader range of equipment than before. The main difference is that a real-time embedded system can sample data very quickly, which allows it to handle data that naturally fluctuates, such as power output, line voltage, temperature, etc. With typical polling rates in minutes instead of milliseconds, SNMP is often used for monitoring network infrastructure and reporting performance metrics where data granularity is less critical.

Because of the differences in SNMP monitoring and real-time reporting, combining the two models onboard a single platform was a unique challenge. It ultimately led to the development of an SNMP manager embedded in the client software, allowing SNMP monitoring alongside real-time broadcast data without inflating the native protocol or requiring greater bandwidth.

EMBEDDED SYSTEM FOR STABILITY

An embedded system offers advantages beyond protocol and bandwidth efficiency. Because the system is self-contained and independent from a PC, disruptions induced by changes in operating system

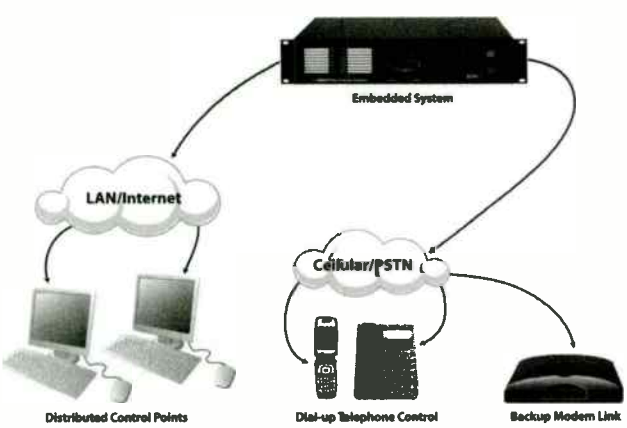


Fig. 2: Server-Client Topology. An embedded system remains connected to the outside world without requiring a central server or computer for data management. Client-side access to the system is flexible.

versions, service packs, security updates, etc. are not an issue. Because a PC is used for multiple purposes, there are often conflicts between applications and competition for CPU time. Desktop operating systems like Windows and Linux are non-deterministic, which further challenges implementation of real-time, mission critical monitoring and control functions. PC solutions also present the problem of inconsistent performance from one installation to the next because each PC is configured with different computing resources and a different complement of software. Finally, because the PC relies on moving parts, its MTBF is lower than an embedded device.

To avoid these pitfalls, one of our primary design requirements was to manage alarms, macros and system configuration onboard the system hardware instead of depending on PCs. We looked to an industrial real-time operating system that offers several advantages over an embedded version of a consumer OS. The result is that the system only requires sufficient memory and processing power to handle its specific tasks. There is also very little OS overhead, which is not the case with an embedded Windows OS. It also means that it was possible to precisely control process priorities to prevent, for example, an urgent alarm notification from being queued behind a configuration update.

Running a remote control system in an embedded device does not, however, leave it disconnected from the PC-driven world, as shown in Fig. 2. In the IP environment, software and web browsers can connect to the system using a client-server model. The hardware at the remote site acts as a server, leaving the desktop PCs to the purpose for which they are best suited: handling the user interface features as a remote client. The upshot is if a service pack is released for the PCs operating system, if a hard drive fails, or if the computer is reallocated for a new use, the integrity of the broadcast facility remote control system remains uninterrupted. This is in contrast to a system that depends on remote commands from a PC or server for routine events like tower light monitoring, AC power management, backup transmitter switching, etc.

PUTTING SCALABILITY TO WORK

A system built on IP architecture breaks the barrier of point-to-point communications, enabling site-to-site connectivity on a far greater scale than before. Being able to monitor and control one remote site from any other fits perfectly into the model of consolidation, where there are more sites to manage and fewer staff to share the responsibility. The ability to automatically (or manually) issue commands from one site directly to another brings the entire broadcast network within easy reach of any location and resolves problems of "orphaned" sites due to traditional communication limitations.

A logical follow-on to the site-to-site capability is being able to integrate unattended studios in the broadcast facility management system. A trouble condition at either location can result in an off-the-air event. Especially with trends toward centralcasting and using the Network Operations Center model for site management, the ability of IP to handle point-to-multipoint communication provides an effective means to expand the scope of broadcast facility management to include studios.

Simplifying Data Presentation

With a system designed to transport larger volumes of data, the need to present that data to the broadcaster in a manageable format becomes more important. All of the advantages of scalability only compound the problem of workload management unless there is a system in place to reduce the volume of data to a reasonable size, deliver it using an effective medium, and apply user-assigned rules to ensure information is directed to the appropriate personnel.

Alarm notification is the area where this concept is most critical. A real-time system can report events quickly, but the alarm report is of little value if it does not prompt an effective resolution. Being able to direct alarm notifications to only the appropriate personnel is a major improvement toward facilitating an effective response to the alarm. By routing transmission system alarms to the RF engineer, audio alarms to the PD, and so on, each system user has a more clearly defined role and there is less potential for alarm conditions to remain unresolved due to communication errors. All operators can still receive notifications via email while, for example, the personnel with primary responsibility receive a pager alert or telephone call from the remote management system. This approach, which maximizes data sharing within the workgroup while directing data distribution based on area of responsibility, is shown in Fig. 3. Client software, as always, can log and track all data for future review.

An extension of selective alarm notification is the capability to analyze multiple out-of-tolerance events so that system reports a small number of primary alarm conditions instead of every alarm that results as a consequence of the primary problem. As shown in Fig. 4, a traditional reporting scheme results in numerous alarm calls to the engineer, even though the alarms are the result of a single event. If the system instead collects the out-of-tolerance data over the span of a few

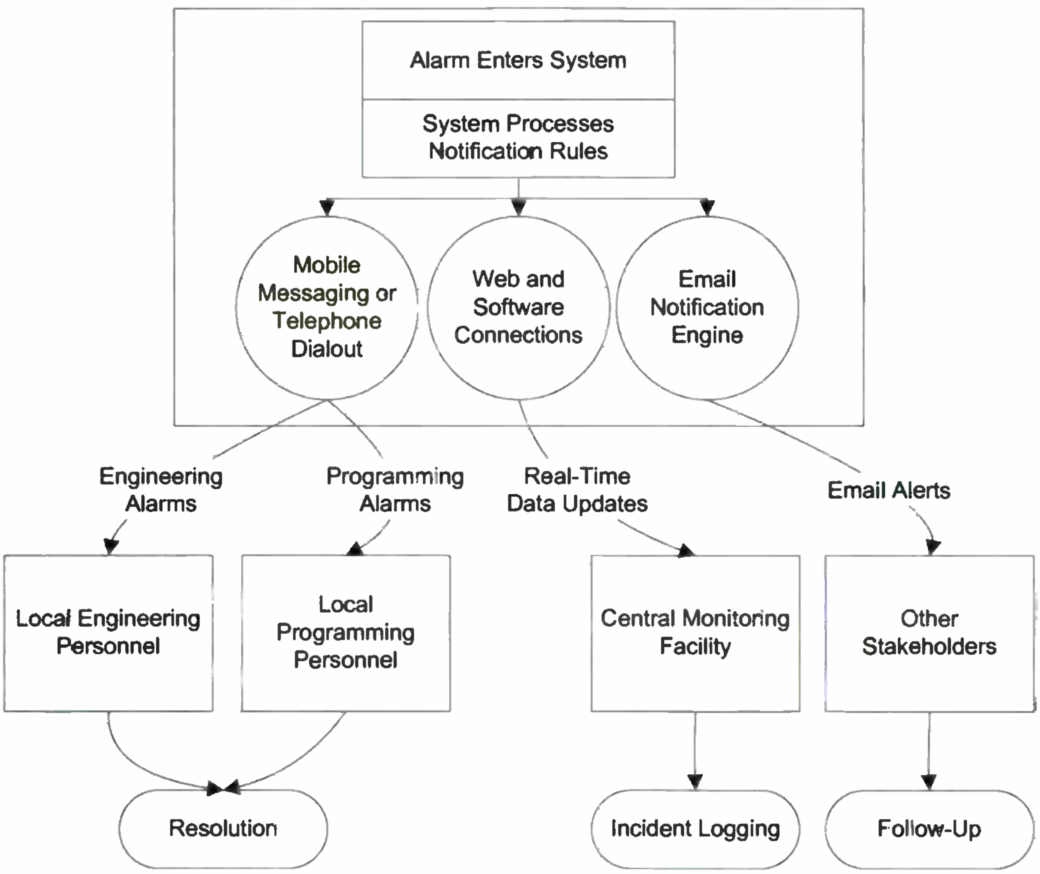


Fig. 3: Alarm Propagation. Selective alarm notification, combined with flexible vehicles for incident reporting, results in an effective means of sharing data among stakeholders and resolving alarm conditions efficiently.

seconds and applies aggregation logic to the data, a more informed alarm notification can be made.

An embedded system is well positioned to handle alarm aggregation because the extra processes involved can occur in virtually the same moment as the raw data enters the system. A remote server is never tasked with receiving the data, processing rules, and determining how to report the alarm.

FROM CONCEPT TO REALITY

In bringing all of these concepts together, a guiding requirement was to achieve relevance to real-world broadcast applications. A solution that leverages a station or group's investment in networking communications is of little use if it simply leaves the broadcaster with a new set of problems to solve.

In an IP-based system for broadcast facility management, the major concern is over redundancy in the communication path. The embedded approach is designed to keep processing power at the remote site to protect the system's integrity if there is LAN/WAN failure. However, a means to access the site during that failure remains a critical requirement. Maintaining a dial-up modem link and dial-up telephone control to the site is good practice even with the most robust IP link. A system with a protocol that operates solely in an IP environment precludes this type of redundancy. This was a consideration that led us to support dial-up in addition to IP.

Outlining a Reasonable Migration Path

The transition to IP has tempted an "out-with-the-old" attitude on the design side, but on the implementation side, some level of backwards compatibility with older equipment and older link types is needed to ensure an easy transition. This consideration played a major role in development and resulted in a smooth integration with existing remote control equipment, as shown in Fig. 5.

Alarm Aggregation: Main AC Failure	
Out of tolerance conditions	**Alarms reported**
MAIN AC OFF	MAIN AC OFF
XMTR OFF AIR	XMTR OFF AIR
SUPPLY VOLTAGE LOW	
LINE VOLTAGE LOW	
LINE VOLTAGE OFF	
XMTR POWER LOW	
HVAC OFF	
STL SIGNAL LOW	
SILENCE SENSOR ALARM	
EXCITER POWER LOW	
EXCITER OFF	

Fig. 4: Alarm aggregation. As the scalability of a broadcast facility management system increases, so does the need for more intelligent alarm reporting. Alarm aggregation generates alarms on actionable conditions, not the conditions that occur as the result of a primary failure.

Fig. 5: Backwards Compatibility. Two generations of broadcast facility remote control equipment at a central studio facility in Hooksett, NH. Photo courtesy of Dirk Nadon, Director of Engineering NH, Nassau Broadcasting.

The advancements in network technologies have moved broadcast forward in impressive ways. When we began work on a broadcast facility remote control system to take advantage of these advancements, it was evident that the technology would only provide value if the system would work in real-world applications using tools that broadcasters can access now. Wrapping the functionality into an embedded system that thrives on the LAN/WAN removed the burden of OS maintenance and software conflict resolution from the broadcaster and provided a system that would deliver stability in an unattended environment over a lifespan that meets the expectations of broadcast.

Packets Everywhere: How IP-Audio and Ethernet Are Transforming Modern Radio Facilities

Marty Sacks
Vice President, Axia Audio
Cleveland, Ohio
marty@AxiaAudio.com

ABSTRACT

For years, Ethernet has been on the periphery of radio station infrastructure. We've used it for traffic management, file sharing, allowing the Internet into our studios, and some audio playback — and there's where the story ended.

Now, however, broadcasters are finding interesting new uses for Ethernet, thanks to the emergent audio standard called IP-Audio. IP-Audio opens up new capabilities and possibilities for broadcasters because it enables real-time transmission of uncompressed, linear audio alongside direct machine control of audio devices and playout computers, transmission of PAD / metadata streams, direct playout of digital PC content sans audio cards, wide-area connectivity (separate floors or separate buildings), high-spectrum STL data links and more — all using the dramatically reduced wiring infrastructure that the Ethernet environment provides.

This paper will examine some of the innovative ways radio broadcasters are employing IP-Audio and Ethernet for applications such as alternative STL, large-scale studio connectivity, networked audio monitoring, and remote system administration.

HOW IP-AUDIO WORKS

The mechanics of IP-Audio are fairly well known, so I'll give just a brief overview.

To begin, audio sources are connected to "audio nodes" located in studios, server areas or anywhere audio sources exist. The nodes convert analog or AES/EBU signals to uncompressed 48 kHz, 24-bit digital audio, which is then packetized for transport via Ethernet. Audio streams from the nodes are assigned unique IP addresses for identification and routing.

After sources are connected, logic ports on audio devices are connected to "GPIO nodes" which convert start/stop/status commands to packet data also.

Each node connects to a QoS-compliant Ethernet switch, which connect to other switches around the facility, and each node's audio and control data are "advertised" to the network, for consumption by anyone who needs them. Gigabit Ethernet or fiber-optic links between switches provide a fat pipe that can handle tens of thousands of stereo signals per system.

IP-Audio is extremely automation-friendly. Once audio and machine-control commands are converted to data, automated software control of routing paths and switching configurations is not only possible but desirable, allowing routes or entire networks to be remapped easily using a simple software interface.

Of course, all this business concerning nodes is simply the way in which IP-Audio must be implemented *today*. In the near future, broadcast equipment will have IP-Audio interfaces built in, eliminating the audio/data conversion altogether.

Picture how a built-in IP-Audio interface would simplify setup of new audio devices. Traditional broadcast phone systems, for instance, typically need wiring for two audio inputs and outputs, program-on-hold input, logic connections to the console, a serial control connection for PC screening software — somewhere near 20 separately soldered connections.

Now picture that same phone system with an IP-Audio interface. First, you plug an Ethernet cable into the jack on the hybrid; you plug the other end into a network switch. That's it! All that I/O and logic travel the same cable. Just think of how much installation time would be saved during new studio builds!

A side note: IP-Audio technology not only delivers a higher performance / cost ratio ("bang for the buck") than traditional methods of studio building, it saves money outright. Short- and long-term cost benefits can be realized in quite a few different areas, like materials (cabling and mainframe router/switcher hardware), installation (reduced labor), maintenance and troubleshooting (simpler infrastructure) and even in time spent assembling system documentation. Users of IP-

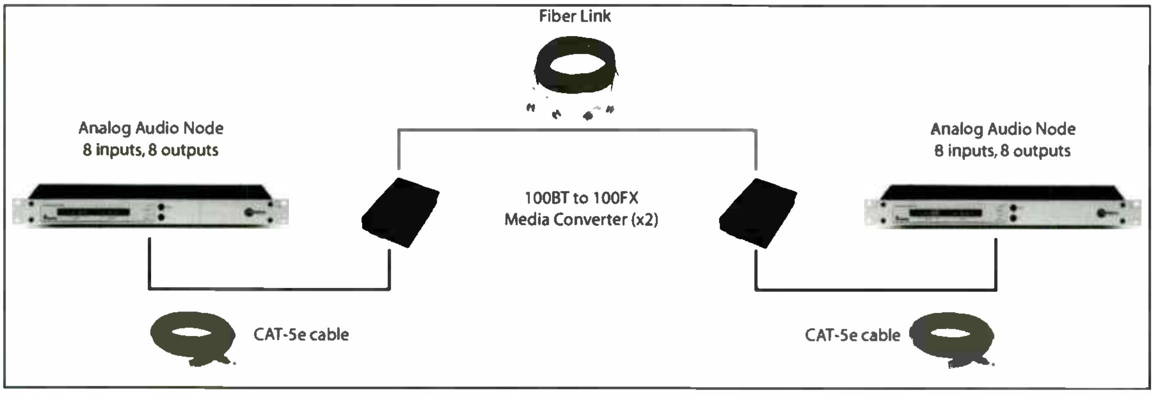


Figure 1

Audio networks are finding their installed costs on new studios or remodels to be between 20% and 35% less traditional hardwired studios.

NOT JUST CONSOLES

When most broadcasters think of IP-Audio, they tend to relate it to studio replacement. And why not? With over 500 console / routing systems on the air at this writing, IP-Audio has certainly proven itself in this area.

But broadcast engineers are nothing if not innovative. They take new technology and invent ingenious new ways of using it that even its designers might not have conceived of. And sharp broadcasters have figured out that once you get audio into the Ethernet domain, you can do many different things with it, thanks to IP-Audio's amazingly easy, flexible, scalable transport system that can use CAT-5 or CAT-6 cable, fiber or Ethernet radios – or whatever comes down the pike next – to economically transport audio and associated data wherever it's needed. From the most basic broadcast plant to the most sophisticated, IP-Audio can be a great problem solver.

SLIMMER SNAKES

The most basic use for IP-Audio is for constructing audio snakes. Snakes are useful at remotes and at venues of all kinds (including in-house performance areas) because, like their reptilian namesakes, their relatively slim profiles enable them to go almost anywhere and squeeze audio into some pretty tight areas.

In today's multiple-studio installation, however, the snake is more likely to resemble an Anaconda than the garden variety. Multi-pair cables comprised of 50, 75 or 100 discrete copper audio pairs are not only bulky, they're expensive. And yet, we've got to move audio around.

IP-Audio offers a handy alternative to the traditional audio snake. Copper bundles are not so easily scalable! And when you run out of capacity, you've got to pull another. But Ethernet capacity is easily scalable; a Gigabit or Fiber Ethernet link can carry *multiple hundreds* of audio channels at a time.

In fact, one of the very first installations of IP-Audio was for exactly this application, at Clear Channel's WREO. They needed a way to move signals between adjacent buildings, and found that using IP-Audio nodes in each building, connected by a fiber link, worked perfectly — and avoided the cost of having to trench in a conduit for multi-pair cable.

Figure 1 shows a typical IP-Audio snake, similar to what WREO deployed. IP-Audio nodes placed in each building are connected to media converters, and then fiber. Each audio node has 8 inputs and 8 outputs, well under the capacity of a 100 Mbit link, which can transport up to 32 stereo signals. If more than 8 streams are needed in each direction, another audio node and an Ethernet switch are added to each end. More audio nodes can be added until link capacity is reached. Gigabit Ethernet can carry up to 250 stereo audio channels.

Of course, building connections aren't the only reason for audio snakes. Live performance venues, worship sound and other staged events are natural applications. Nor do all IP-Audio snakes have to be permanent: remote broadcast setups at sporting venues or other outdoor events can certainly benefit from IP-Audio's single-cable setup.

FLOOR AND BUILDING INTERCONNECT

An IP-Audio application that's a natural extension of the snake is the interconnection of studios and technical operation centers in multi-floors or multi-building facilities. Minnesota Public Radio's

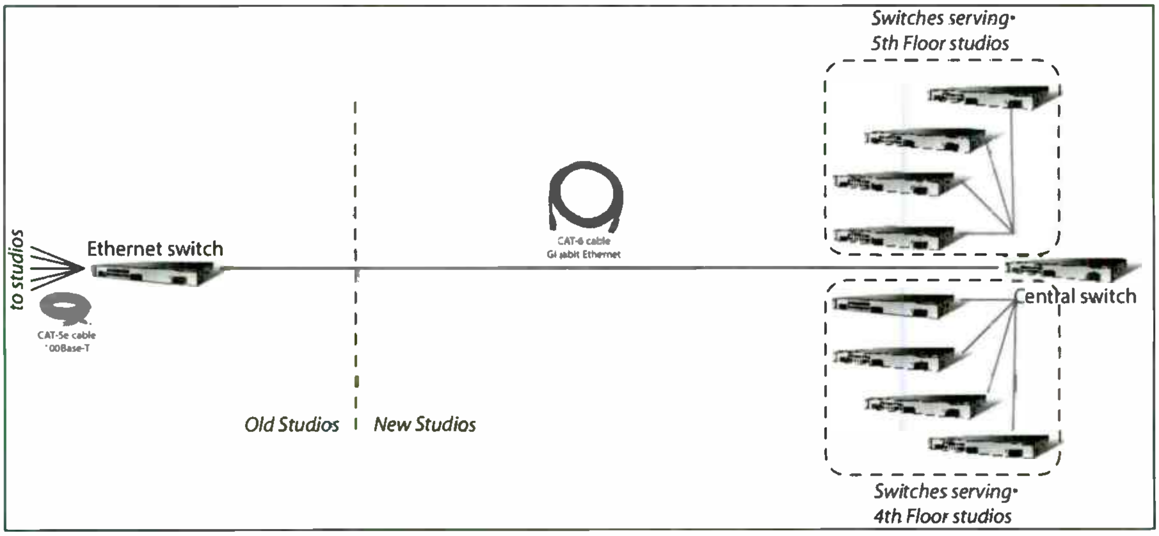


Figure 2

recent studio expansion is a perfect example: their studios span two floors of a new building, plus two floors of an adjacent building.

Users have discovered that IP-Audio is a natural for getting audio to and from transmitter shacks or where satellite or RPU receivers are located. Some have told me that fiber is an excellent connection method for this application, because it greatly reduces the potential for transmitted lightning damage — fiber doesn't conduct!

Let's look at the Minnesota Public Radio example first. MPR put a central switch in their new building, a load-balancing bladeserver from Cisco. The old studios that needed connection with the new ones are located in an adjacent building with a shared wall, so a Gigabit Ethernet link was run from the central switch through the wall to an edge switch that served the audio nodes in the old studios. These nodes use CAT-5e to connect to their edge switch, which connects to the central switch in turn with CAT-6 (for Gigabit).

The central switch also connects via Gigabit to the edge switches on both floors of the new studio complex, as shown in the (much simplified) diagram above.

Now let's look at the transmitter-to-doghouse mentioned previously as a solution for solving lightning isolation problems. Cumulus Media's Youngstown, Ohio facility is made up of on-air and production studios for 6 stations scattered across two buildings. Also on site are two transmitter buildings containing transmitters for two of those stations along with the STL and RPU gear for the rest, all attached to a 700-foot tower. The distance between the buildings is between 100 and 200 feet, and all are interconnected by copper in a triangle arrangement. Each building has a separate power feed as well. Lightning has been a recurring issue, not to mention simple surge and ground differentials.

The solution was to build a processing / data rack room and run fiber between that facility and the other studio and transmitter building, eliminating ground loop problems and helping attenuate lightning attraction, as well as breaking the conductive link between buildings.

A combination of bidirectional AES, analog, and Ethernet has to be carried over this fiber, and Cumulus engineers found, after extensive research of fiber-only systems, that Ethernet combined with

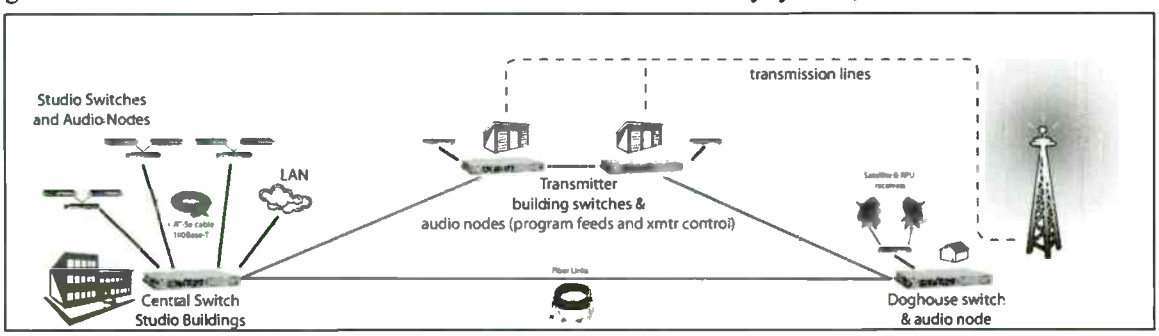


Figure 3

fiber provided more flexibility.

As shown in Figure 3, IP-Audio nodes and Ethernet switches equipped with GBIC (Gigabit Interface Converters) are placed in each location. Connecting all of the switches with fiber creates an IP-Audio network across all of the buildings, allowing sharing and routing of audio, device control and LAN traffic, eliminating ground loop problems and effectively limiting potential lightning damage to very localized areas.

SOUND CARD ALTERNATIVES

The phrase "paradigm shift" is overused, so I apologize in advance for using it here! Overuse aside, however, it's an accurate description for an IP-Audio application that has the potential to completely revolutionize one segment of broadcast studio architecture.

Since IP-Audio converts audio and control into routable Ethernet data, and since Ethernet was designed to facilitate computer users' exchange of PC data, it stands to reason that PC audio workstations should be able to exchange audio with the rest of the network directly, right? And they can. PC audio workstations or audio delivery systems can be equipped with an IP-Audio Driver that emulates a soundcard. The driver can then route multiple channels of audio playback, record and data/logic command streams via the PC's Network Interface Card (NIC), which connects with the local Ethernet switch.

Many IP-Audio users have discovered that they can save substantial amounts of money by using this approach. Not only is there money saved on sound card purchases, but also on associated wiring. Not only that, the cost of the router switcher port (or console input module) that a traditional hardwired router system would require for these sound card inputs is eliminated as well. Once PC audio sources are part of the Ethernet data flow, they are universally available as IP streams and can be switched or mixed as needed.

Users have been quick to realize the benefits of this approach, and many well-known delivery system providers have announced that they have IP-Audio drivers available for popular playout systems. These providers include the likes of Broadcast Electronics, BSI, D.A.V.I.D. Systems, ENCO Systems, Google/dMarc, iMediaTouch, Netia, Pristine Systems and Prophet Systems, to name a few. Media giants such as Univision, Clear Channel and France's Lagardere Group have already built facilities using this tight PC/router model.

Of course, this approach isn't mandated. IP-Audio systems coexist quite nicely with traditional sound cards (by simply plugging their inputs and outputs into Audio Nodes).

Even sound card providers themselves realize the benefits of IP-Audio compatibility. Respected audio-card maker AudioScience recently announced their intention to manufacture a broadcast-quality card that works directly with IP-Audio networks. Instead of the usual "pigtail" and tangle of I/O and logic connectors, multiple channels of audio and control data will travel from the card directly to the network over a single thin Ethernet cable.

ROUTING SWITCHERS, REMIXED

While the routing switcher is a pretty familiar sight, especially in larger markets, IP-Audio users have found some new ways to build them.

Traditional routers are usually fairly large and complicated, with topologies of 64x64 crosspoints or larger. IP-Audio users have found that, thanks to Ethernet's inherent scalability, it's not only possible but practical to assemble an audio routing switcher as small as 8x8 stereo inputs and outputs (analog and/or digital, with or without program associated data) that could scale to a very large system without huge jumps of costs or complexity.

Conventional routing systems, while serving their purpose well, nonetheless force a choice during initial planning and building: the "small frame" and the "big frame." And this choice must be made before even one cable is run. Studio projects are constantly in flux during construction, and requirements often change mid-stream. The result: if you purchase more capacity than you thought you'd need, money is wasted. If you purchase too little capacity, costs escalate rapidly when it comes time to correct that mistake: most routing switcher frames are custom-built, so if you wind up with more inputs than your router frame is capable of taking, the only solution is a second, expensive frame.

Another consideration is the cost of the multiple plug-in I/O cards that hardwired routers require. And it's not like you can re-use them: cards from one system – even from the same manufacturer – often cannot be used in a different card cage.

IP-Audio could completely change the way routing applications are constructed. Thanks to the scalability of their Ethernet backbone, small systems can be built with relatively little initial cash outlay. Then, when more capacity is required, the network can be expanded, again and again if need

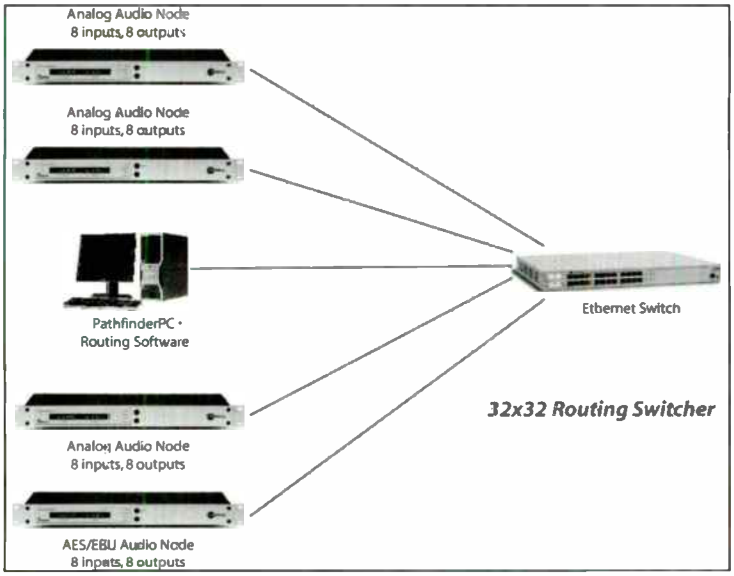


Figure 4

be, without ever discarding the original parts of the system. An IP-Audio router can thus grow from very small to very large at a very predictable and linear cost.

Figure 4 underscores this point, demonstrating how, with very little equipment, you can easily construct a 32x32 routing switcher using IP-Audio components.

The diagram shows four Audio Nodes, each of which have 8 stereo inputs and 8 stereo outputs. The nodes are connected to a 100/1000Base-T Ethernet switch using inexpensive CAT-5e cable. Only one Ethernet cable is needed to connect each node to the switch, because each 100Base-T link has enough capacity to carry 32 stereo signals — all running at once! Automated routing control and scene changes can be handled by software running on a connected PC.

32x32 is a pretty respectable routing system, easily capable of serving two or three studios, but times change and needs grow. IP-Audio users have found that it's especially easy to add capacity to their routing systems because all they have to do is connect more nodes to the existing setup, again using standard CAT-5e.

As you can see in Figure 5, the original 32x32 system is still intact — but with more Audio nodes added to accommodate increased routing demand, doubling the size of the original routing switcher system to 66x66.

Should the need for more capacity arise, the same procedure is repeated, adding more nodes where they're needed. Figure 6 shows how the original system has expanded yet again to a 132x132 router.

As it's somewhat related, this is a good place to mention another neat benefit of IP-Audio systems: portability.

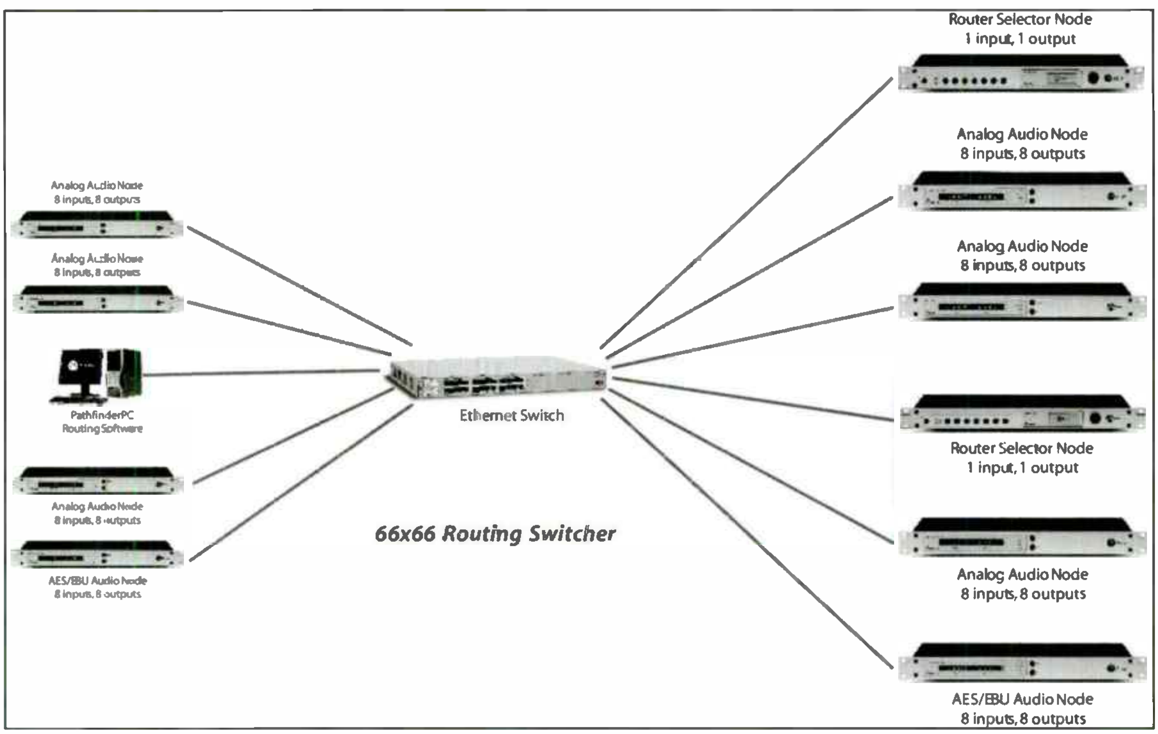


Figure 5

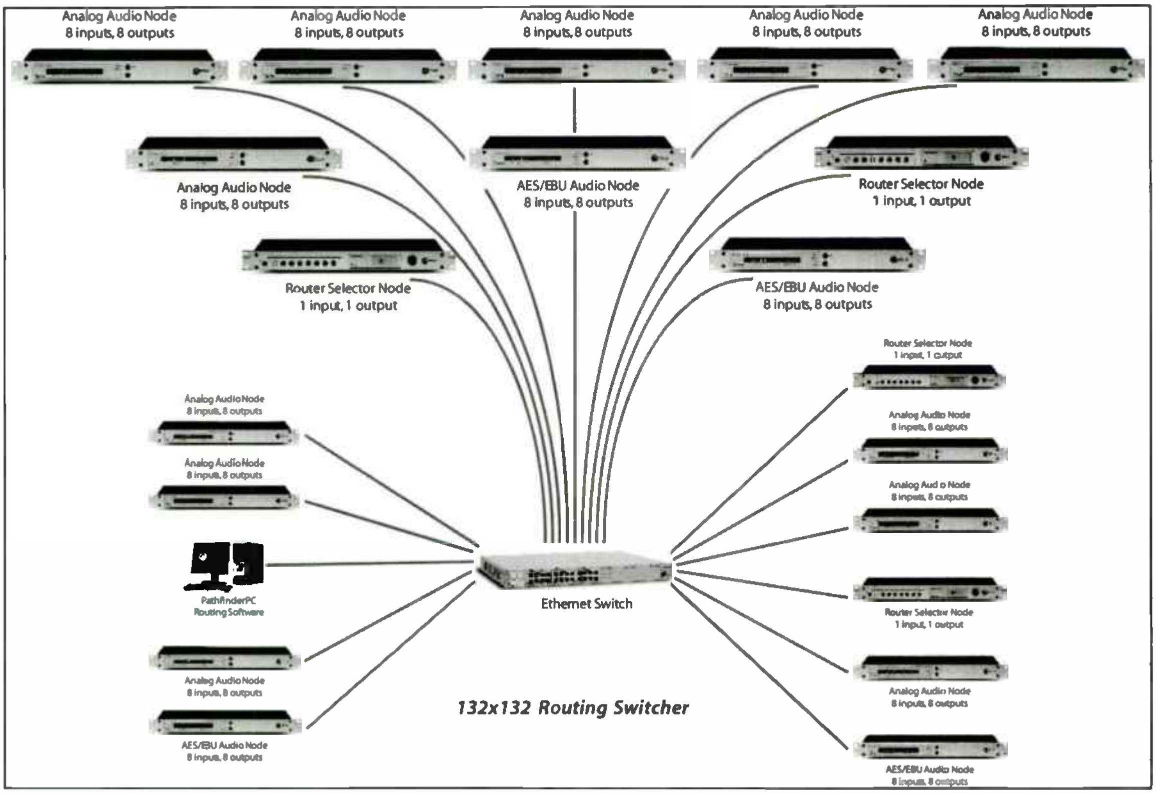


Figure 6

With traditional router systems, it's nearly impossible to re-use hardware in a new location. A different building means different card cages, different locations for inputs and outputs — not to mention all the expensive multi-pair cable that is thrown away.

In sharp contrast, IP-Audio routing systems can literally be picked up and moved to a new location. Since audio nodes are rackmount devices, they can simply be racked in the new facilities and connected with Ethernet. A little bit of software reconfiguration, and the entire system is online again.

UNCOMPRESSED STL

Everybody knows that the 950 MHz STL band is terribly crowded in all but the smallest markets. Hardly a month goes by where there isn't a story about a market where somebody was knocked off the air by another station turning on an STL, or where STL bandwidth is reduced by the frequency coordinator.

An innovative use of IP-Audio is helping to solve these problems in several markets. Broadcasters are using IP-audio nodes, combined with Ethernet radios from providers such as Broadcast Electronics, Dragonwave and Orthogon to replace 950 MHz transmissions with license-free data channels in higher-frequency bands.

These combinations of IP-Audio gear and Ethernet radios can provide multiple channels of bi-directional analog or AES audio (as well as GPIO commands for remote control of transmitter rack equipment such as audio processors satellite receivers). And, unlike traditional STL, these audio channels are uncompressed, so transmission-related coding artifacts are eliminated.

(A side benefit to this transmission method is that the link can be easily reconfigured to add and subtract audio channels — an excellent fit for HD Radio applications with multiple concurrent program channels.)

An example of such an installation using Ethernet radios is found in use at the cluster of radio stations operated by Clear Channel in Birmingham, Alabama.

Chief Engineer Bob Newberry found that the portion of the local spectrum he'd used for years for STL was unsuitable for future HD Radio implementation. Also, his stations had actually been taken off the air due to interference when another station accidentally fired up transmissions on the same frequency.

Eyes to the future, Bob considered IP-Audio as a way to solve not only frequency crowding, but also to consolidate multiple audio and data channels into a single transmission path.

The Birmingham cluster was already feeding four transmitter sites with two 4-channel STL sys-

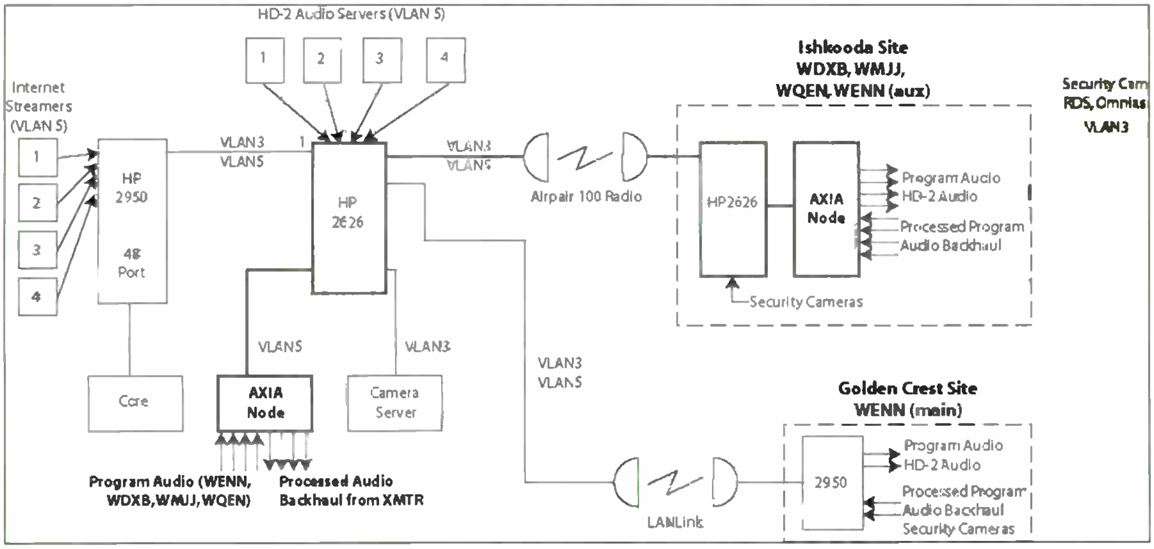


Figure 7

tems, and employed a LAN extender to get RDS data and transmitter remote control to the transmitter site. With three HD-2 channels planned for broadcast, the question arose: how would they get the additional audio (plus associated RDS and remote control) to the transmitter site without overloading the existing gear?

Investigation showed that an Ethernet radio with a 100 MB/s bidirectional link would have enough capacity for audio, PAD and control data, plus room for more audio/data channels should future plans so necessitate.

To implement the system, audio nodes were placed on either end of an 18 GHz link; the simplified diagram shown in Figure 7 illustrates how program content, audio backhaul and even streaming images from security cameras coexist on the same "STL" link.

At the time of this writing, the system will be celebrating its first anniversary in service; performance has been flawless.

A second example comes from Radio Skonto in Riga, Latvia. They'd received a license for transmitters in two other cities outside their main service area, and management was quite pleased. But for engineer Ivo Bankovs, the problem was one of distance: how would he get the signal from the studio in Riga to the cities of Rezekne and Liepaja, each about 250km away?

Traditional Telco audio service was very expensive. Installing his own STL radio system was impractical due to the need for intermediate repeaters and towers — if he could get a license. So he asked a local Internet Service Provider, Latnet, if it would be possible to use IP links to the two sites.

Fortunately, the ISP was able to offer a guaranteed bandwidth service to the two sites at a cost much lower than the alternatives. To avoid paying Telco costs, Latnet installed a 26GHz IP radio link with equipment made by Netro to Radio Skonto's studios. Radio Skonto contracted for 384kbps from the Riga studio and 256kbps at each of the remote sites, providing plenty of margin for packet overhead.

Radio Skonto decided to use Telos Zephyrs to provide MPEG compression and IP conversion. The station is a fully IP-Audio networked facility, so they wanted to input audio to the Zephyr at the main site via IP. This was accomplished by simply plugging it in to a spare port on the Ethernet switch and configuring the main IP-Audio program channel as its input source. A low-cost ($50) consumer IP router was used to connect the Latnet IP radio to another port on the Ethernet switch. This router was "locked-down" to pass only the required audio signals. and thus to isolate the IP-Audio network from any other traffic on the external network. The Zephyr provides two streams, one for each site.

At each remote site is a small studio set-up for local programming. The Zephyrs deliver their audio outputs via analog connections to a console fader input; DSL lines provide the IP connection. As before, these are firewalled with a low-cost router inserted between the DSL line and the Zephyr; see Figure 8 for a simplified diagram. The routers have additional LAN ports that are connected to some other PCs used for non-real-time audio transfer;

The station was naturally concerned about delay, wanting the lowest possible value so that people listening to live telephone calls would not be confused (the station runs without a profanity delay). To accomplish this, they set their Zephyrs to

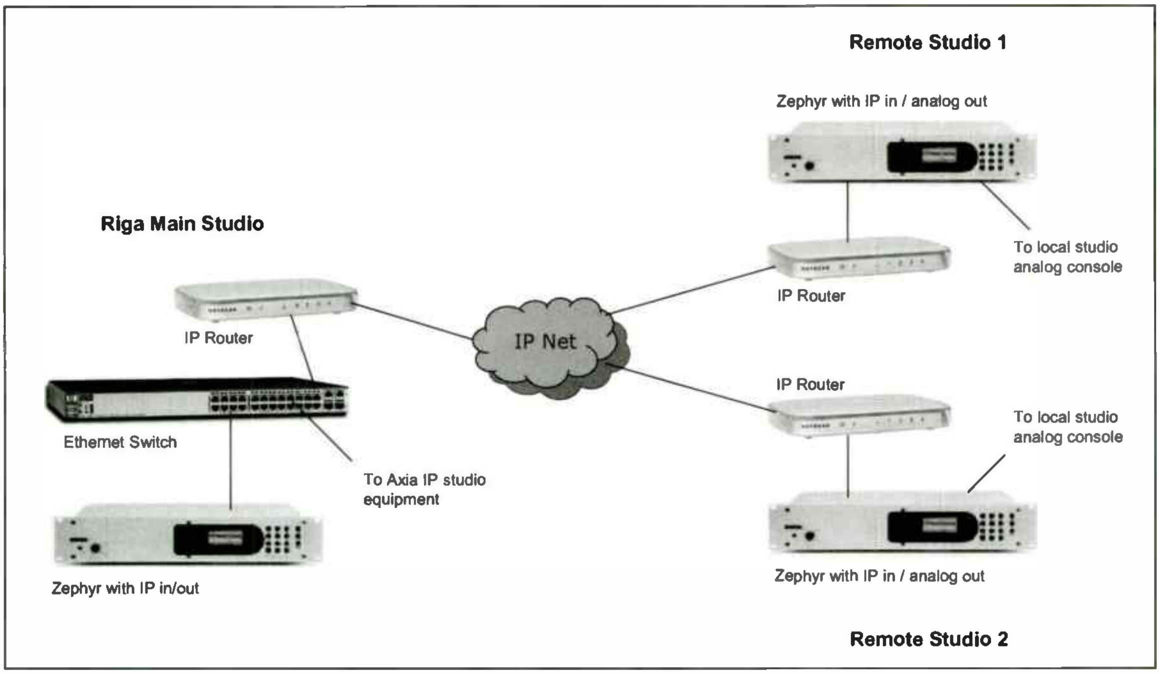


Figure 8

the lowest buffer setting: 250ms. They were not sure if a setting that low would work, but decided that the best strategy would be to start with a minimum value and increase it in small steps until any audio drop-outs stopped.

As it turned out, the minimum setting worked without problem — fortunately, the IP network had very low jitter.

Ivo and his assistant engineer, Karlis Malkavs, wondered if 128kbps would provide sufficient quality and were ready to increase to a higher rate if compression artifacts were evident. But after a month of on-air operation, they decided to stay with this lower rate. MPEG AAC has been officially designated as "indistinguishable from the source" at 128kbps by the European Broadcasting Union; this real-world application certainly proved it to be so!

The system has been in operation for a number of months and is working well.

There was an outage at one of the sites caused by a power failure affecting the ISP's equipment center. A UPS was supposed to provide back-up, but a switch in the path was mistakenly not connected to the UPS. Other than that, Ivo reports no problems. The ISP's promise of guaranteed bandwidth has been kept, and the IP option has proven to be a satisfactory studio-to-transmitter link.

AUDIO MONITORING FOR ALL

Listening stations are another area being transformed by IP-Audio. Until very recently, making listening stations available for GMs, PDs, Sales Directors – even the DOE! – meant filling a TOC rack with distribution amplifiers, tuners, line selectors and other gear… and then running more cable to individual offices throughout the facility, outfitting them with speakers, volume controls, et cetera.

Needless to say, this can be an expensive and time-consuming endeavor. As a result, these projects end up being very costly or providing less monitoring points than desired.

IP-Audio users have found that they can use PCs connected to their audio network to provide monitoring for practically everyone in a networked facility by using software applications that translate networked audio streams to PC audio playable on any local computer.

In this way, anyone connected to that network can monitor air – or any other available audio stream – using the speakers already attached to their computer. GMs, Program Directors, salespeople and programming staff can all listen to their choice of air monitors instantly, from any location, with no additional equipment to purchase and install or extra cable to run.

Quite a change from the old Altec corner speaker hung in the GM's office.

FACILITY AUTOMATION MATURES

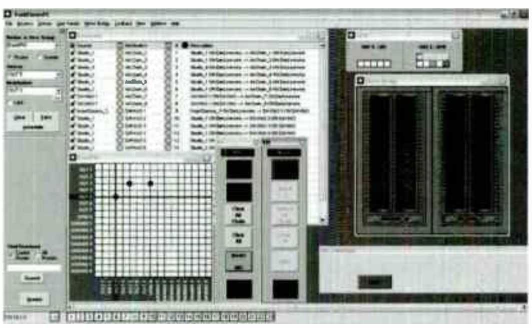
The more activities (routine and emergency) we can simplify or take out of the hands of operators, the more reliable our plants are likely to be.

IP-Audio, thanks to its computer-tech base and subsequent natural capability for integrating and interacting with PCs, is working changes in this area as well.

Certainly, we've had some ability to automate aspects of the broadcast environment before: audio processors with dayparting capabilities are a good example; PC-based audio delivery systems are now so ubiquitous that the service they provide (with smoothness and accuracy unthought-of in the Schaeffer days) is routinely taken for granted.

But now, with computers tied even more closely to hardware, software applications can take control of nearly every audio device throughout the plant. They can administer pre-scheduled routing scene changes, react in predetermined ways to trigger events, even take emergency action in the absence of station personnel.

PathfinderPC is such an application. It can work with several different kinds of routing switchers, and is optimized for control of IP-Audio networks. Pathfinder contains powerful scripting capabilities that use simple and familiar Boolean logic to enable users to construct logic trees, building applications that react either to direct user interaction such as a button press or contact closure activation, or to pre-defined fault states such as dead air, loss of program feed or diminished audio levels. Scheduled routing changes can also be programmed.



Of all the ways in the world to be awakened, a phone call at 3AM is one of the least enjoyable. Pathfinder can help take the urgency out of unexpected events by sensing and reacting to loss of program audio.

Let's say that (at 3AM, of course) a satellite channel supplying programming suddenly goes dark. Pathfinder's silence-sense function can be programmed ahead of time to deal with such a problem; when audio drops, it takes an alternate air feed. There's e-mail notification capability too – conditional routing scenes can be defined so that if, for instance, a primary ISDN feed gets knocked out, the system switches to backup and sends an e-mail to your Blackberry® telling you what happened.

Another aspect of PathfinderPC is its support for touch screen monitors. On-screen "virtual button panels" can be constructed to run on local PCs so that board ops can execute simple (changing satellite receivers) or complex (transferring program feed origination to another studio) switching functions with just one tap.

An informal poll reveals that roughly 90% of broadcasters who adopt IP-Audio studio architecture opt to include this sort of advanced automated control; such systems are in use at stations belonging to Univision, Clear Channel, Cumulus Media, MPR and many others.

INTERESTING CONSOLE APPLICATIONS

There are some unique benefits associated with IP-Audio mixing consoles that can help simplify life in the studio. First, a console in an IP-Audio system makes the distribution of I/O, logic and backfeeds easier than ever.

Sharing a standard audio connection in a traditional studio means two pairs for the stereo audio, another two pairs (at least) for machine control logic, and another two pairs for a backfeed or mix-minus. That's at least 6 pairs of wire for each audio connection, multiplied by the number of places you need to send it.

By contrast, an IP-Audio studio setup reduces the amount of wiring needed by an order of magnitude, because audio I/O, machine logic and even mix-minuses are all converted to digital streams, which are routed together in a single "bundle" of packets using Ethernet cable which can be shared with hundreds of other like signals. But elimination of wiring isn't the best part — there's also the fact that since audio, logic and backfeed are now part of the same digital package, you will *always receive the correct mix-minus* when routing remote or telephone audio. In other words, you can bring up a codec or hybrid on any console and it always works right.

Speaking of phones, IP-Audio also provides much tighter phone/console interoperability than has ever been possible before. With traditional consoles, phone operations have required outboard controllers – switch panels or phone-like devices – to control the hybrid. This interrupts the workflow in the control room, since the jock has to take his hands and eyes off the console to work another device. Drop-in control panels for consoles mitigate this somewhat, but these controls are often not adjacent to the audio faders themselves, and in any case require discrete wiring for audio, hybrid control and mix-minuses.

As noted previously, IP-Audio networks route I/O, control and backfeeds together, so the phone controller can live in the console, right next to the audio faders, and the console can receive audio, generate hybrid control logic and send mix-minus, all over the same Ethernet cable. Another benefit is that the hybrid itself doesn't need to be located in the studio anymore — it can be placed in the TOC or some other central rack room, safe from curious hands.

Finally, something I consider to be a very unique console application: the ability for two consoles to work simultaneously with shared sources.

Imagine a scenario in which talent in the talk studio wants their own level control of some audio devices. But the control room operator needs to keep control of the levels in order to "assist" the talent should they run their audio at the wrong level.

The solution: link motorized faders on the control room and talent consoles, designating the CR as the "master" console, and the studio as the "slave." This allows talent to set levels as desired, but also allow the board op to override those levels as needed – something that can't be done with a traditional console setup.

CONCLUSION

Historically, IP-based systems have worked a sea change in every industry into which they have been introduced.

The World Wide Web would not exist if not for IP. Defense computers connected with serial data transmissions formed rudimentary networks in the 1960s, and those networks expanded with ARPANET in the 1970s. But the development and implementation of TCP/IP on ARPANET in 1983 led directly to the growth of what is now the Internet.

Banking networks using ATM (Asynchronous Transfer Mode) revolutionized that industry in the 1970s by allowing computers to take on the tedious accountancy handwork that was the foundation of banking. But the implementation of TCP/IP over ATM made possible instant money transfers between far-flung locations. Remember the days of hurrying to the bank on Fridays (between 10 and 4, of course) to get cash for the weekend? Now, instant banking and automated tellers are available 24/7 thanks to IP.

Telephone systems have undergone tremendous change thanks to IP. Traditional PBX systems and Centrex lines considered advanced only a few short years ago are now being abandoned in favor of VoIP (Voice over Internet Protocol) systems that are more flexible and more cost effective than the old services could ever be. The migration to VoIP has been so universal that in early 2006, Cisco announced that it had sold its 7.5 millionth IP telephone. And the adoption by consumers of mass-market VoIP services over broadband Internet access which began in 2004 continues unabated. According to research firm Cahners In-Stat, more than 9 million US households now have at least one active VoIP user; it was estimated at the end of Q4 2006 that nearly 8% of U.S. households now use a VoIP telephone service —up from 6.5% just six months prior.[1]

Finally, the television industry has embraced Video-over-IP with a fervor. Gone are the days of sending dubbed U-Matic cartridges to other stations by courier, or even satellite uplinks: TV production facilities now share content via video gateways that link production facilities and television stations with each other directly via IP connection. Writing in Broadcast Engineering, Brad Gilmer asserts that "Video over IP is destined to become the predominant technology for the transport of professional video over WANs."[2] And, as with IP telephony, this service is spilling over into the consumer arena with IPTV reaching millions of homes via digital cable services.

With this data in mind, it's certainly no stretch to predict that IP-Audio is on a course to revolutionize the broadcasting industry. More and more applications for this technology are being discovered as the numbers of IP-Audio installations rise. The question for radio broadcasters is not whether to transition to IP-Audio, but when.

---

[1] "Report Sees Rise in Residential VoIP Service", Cabling Installation & Maintenance Online, December 22, 2006 (http://tinyurl.com/3cmvcw)

[2] "Video Over IP", Broadcast Engineering, August 1, 2006 (http://tinyurl.com/23mwky)

Media Asset Management for TV

Tuesday, April 17, 2007
9:00 AM 12:00 PM

Chairperson: Alan Popkin
KLCS - TV, Los Angeles, CA

Transitioning to a Digital Media Distribution Strategy
Tom Ohanian, Signiant, Cranston, RI

Service Oriented Architectures for Broadcast Media Asset Management
Peter Thomas, Blue Order Solutions AG, Kaiserslautern, Germany

Parts That Are More Than the Whole: Asset Management in the IT Age
Michael Atkin, BroadView Software, Toronto, Canada

Content Storage Management in the Newsroom
Brian Campanotti, Front Porch Digital, Boulder, CO

Systems Integration – The Federal Enterprise Architecture Approach to Media Asset Management System Design
Sam Stevens, Radio Free Asia, Washington, DC

Transitioning to a Digital Media Distribution Strategy

Tom Ohanian
Signiant
Burlington, Massachusetts

ABSTRACT

The Media and Entertainment (M&E) industry continues to adopt digital media in all stages of acquisition, manipulation, and distribution. Whether the focus is Broadcast, Film, Animation, and so forth, it is clear that across all M&E segments, digital technologies for creating, managing, and distributing content are in various phases of adoption. While digital nonlinear editing (DNLE) systems are furthest along in their adoption cycle, digital distribution mechanisms and strategies are just now being examined and implemented.

Why investigate and institute a digital media distribution strategy? The M&E industry is facing tremendous requirements that are often summed up in the often used phrase, "Create Once, Consume Anywhere". This, of course, refers to the ability to create a program master and then use that master to provide all required versions for a variety of view types, ranging from satellite, terrestrial, broadband, mobile, and portable devices. Implementing a digital media distribution strategy is critical to meeting these business-driven initiatives.

MOVING CONTENT

When was the last time you consciously thought about how content moves inside and outside your organization? Do you think about it every minute? Every day? Not at all? Have you clicked the "send" button and crossed your fingers that the content will actually get there? Are you still sneaker-netting drives? Are you copying files onto 300 and 500 Gigabyte drives and handing them off to bicycle-bound couriers? Worse yet, have you ever had to recopy 300 GB of data on a drive because the media wouldn't mount at the receiving location because of the dreaded "format unrecognized" error message?

Do you use logistics delivery services for guaranteed overnight delivery? Do you use FTP? SFTP? Do you have to send, receive, and coordinate the movement of content among many locations? Are you worried about security and piracy? Have you missed a transmission deadline? Come close? Do you constantly wonder how you will keep up with moving around the growing amount of digital media that your company generates?

In the end, can you make all the new deliverables that are required and get them to where they must be and do it on time?

ADOPTION OF FILE-BASED WORKFLOWS

While there is still an enormous amount of both film and tape-based acquisition of content, we are also seeing a concerted and inevitable move to file-based acquisition systems. While these camera system recording media range from optical disc to solid-state, the creation of disk-based files has spurred the growth of file-based workflows where acquisition, manipulation, mastering, and delivery are completely file based.

Today, we are squarely in the midst of this conversion and while a significant amount of content is still shot on both film and videotape and while there are still thousands of hours of tapes still being copied to satisfy all the constituents that require the content, moving to file-based workflows is inevitable. In the film world, the adoption of DI (digital intermediate) workflows has resulted in a drastic reduction in print conforming and negative cutting. Instead, film negative is scanned and forever remains in a digital file form until final recording back to film negative.

In broadcast news, each and every day we see continued conversion from tape-based workflows to digital file-based workflows. At an average yearly conversion of approximately 100-125 stations (and increasing) we are arguably in the fifth year of what will probably be a ten year conversion from tape-based to file-based workflows for broadcast.

The simple notion of "a frame is now a file" requires a digital infrastructure to support file-based workflows. It is often helpful to classify workflow methodologies. Consider for example, feature film creation. We could classify this as:

A-A-A-A: Analog film acquisition, analog manipulation (editing), and analog film mastering and distribution.

However, consider that workflow when a DI process is utilized. The workflow becomes:

A-D-D-A/D: Where film is analog acquired, digitally scanned, digitally manipulated, and, finally distributed both analog and digital.

With no analog film acquisition, but with solid state acquisition, the workflow finally achieves a complete digital workflow which becomes:

D-D-D-D: Where film is digitally acquired, digitally scanned, digitally manipulated, and, finally digitally distributed.

In broadcast news situations, we have long functioned under this workflow:

A-A-A-A: Where video is analog acquired, analog manipulated, analog mastered, and analog transmitted.

Clearly, however, broadcast news is rapidly entering the complete digital workflow of:

D-D-D-D: Where video is acquired digitally and file based, digitally manipulated, digitally mastered, and digitally transmitted.

TRENDS

Trending, then, takes the form of the following combined areas:

1. File-based acquisition systems
2. Adoption of Storage Area Network (SAN) and Network Attached Storage (NAS)
3. SAN- and NAS-connected workstations for content creation and manipulation applications
4. Media Processing applications (e.g. encoding. Transcoding, color correction, etc.)
5. Digital Delivery systems
6. Adoption of standards-based Information Technologies (IT) such as storage, routing, and protocols.

NEW WORKFLOWS AND NEW REQUIREMENTS

Time to market demands and multiple delivery venues bring forth new challenges. They force us to adopt new technology and to adapt our workflows. Why is it so important to carry programming in so many different forms and to so many different screen types? Let's consider some developments:

The average amount of time from start to finish of a major motion picture in 1994 was 12 months. Today it is seven months.

- In 1973, according to John Kramer of EMAK Worldwide Marketing, a company wanting to reach 95% of women aged 25-49 would need only three network commercials. Today, 92 commercials would be needed to reach that same group.
- In 2001, the Theatrical-to-Video window was 165 days. In 2005, the window was 129 days.
- Year-over-Year growth of User Generated Content (UGC) on YouTube was 2,662% and placed YouTube in the top 15 of Internet sites.
- In 2006, it was estimated that an average of 62% of Internet traffic consisted of un-monetized video.

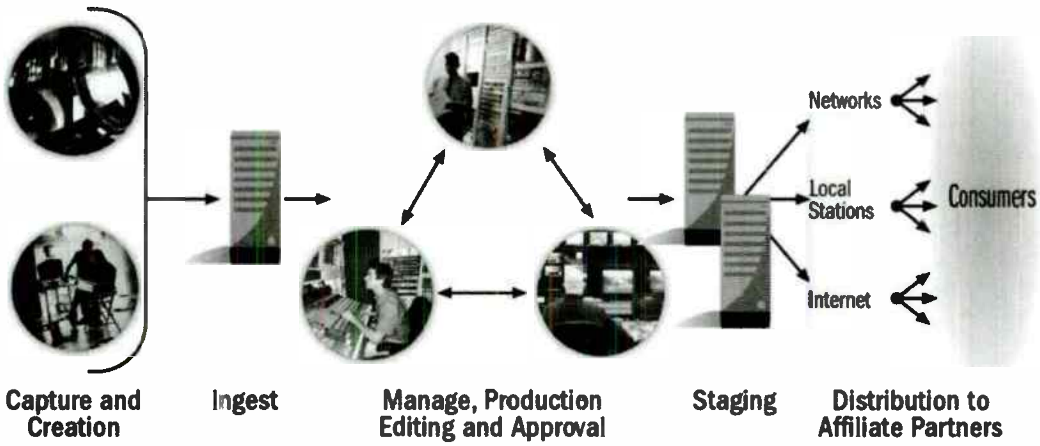


Figure 1: A typical content creation lifecycle from capture to mass market distribution.

However, while the above examples are only further indications of the necessity of file-based workflows to meet these various demands, implementing such workflows are not without complications. Figure 1 outlines the process of content acquisition to delivery and it is important to note that each of the distinct processes is, in itself, cyclical.

REQUIREMENTS OF A DIGITAL MEDIA DISTRIBUTION SYSTEM

What are some of the functional areas and technologies that form the basis in implementing a digital media distribution system? There are six main areas that typically enter into the research and discovery phase. They are as follows:

Centralized Management

This refers to capabilities that allow users to determine rules and policies by which content is transmitted on the network. An example of this could is the use of tiered pricing where content is assigned a High, Medium, and Low priority designation for scheduled delivery. Centralized control of media transfers also provides the ability to look into the past, present, and future of media transfers. When the questions are: "what has run, what is running, and what is scheduled to run?" a dashboard view of the media delivery system can easily provide that feedback. Figure 2 is an example of a simple view of media transfers that can be centrally managed.

Centralized management of the movement of digital media is indispensable in order to gain access to extremely valuable information such as: asset tracking, reporting, network and storage utilization, and billing generation.

Network Resource Management

This refers to capabilities that allow users to manage the allocation of network resources and tie those resources to media transfers that utilize all forms of networking present in the client's ecosystem. For example, this could take the form of real-time bandwidth allocation and

	Job Name	State	Completed	Frequency		
	Central	Running	85%	Every 1 week(s)		
	EastCoast	Next Run Scheduled	85%	Every 1 week(s)		
	Central	Running	85%	Every 1 week(s)		
	EastCoast	Next Run Scheduled	85%	Every 1 week(s)		
	WestCoast	Running	85%	Every 1 week(s)		

Figure 2: A dashboard view of a centrally-managed digital media distribution system.

readjustment to media transfers. Or, perhaps it is important that a series of media transfers being sent over a media path should not exceed certain bandwidth limits imposed for that path and which must be respected.

In addition to standard TCP/IP networking capabilities, there should be the capability of compensating for highly latent networks and provide for WAN (wide area network) acceleration techniques via employing UDP (user datagram protocol).
Other standard networking techniques such as checkpointing, snapshots, and auto-resumption in case of line disruption are all desirable.

Figure 3 is an example of how bandwidth could be adjusted in real-time in order to facilitate transfers of greater or lesser importance.

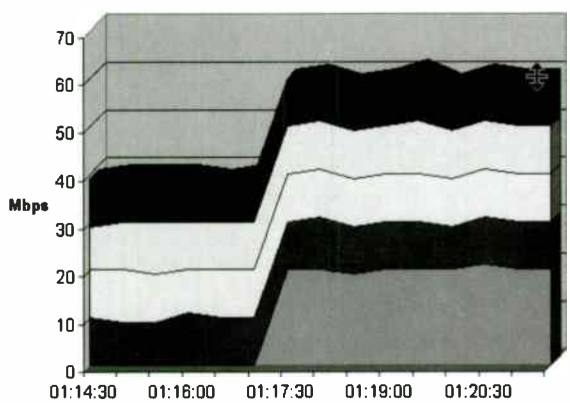


Figure 3 Real-time bandwidth allocation for transfers of greater or lesser importance.

Security

Data integrity and data confidentially are implicit requirements of any digital media distribution system. Through a combination of SSL (secure sockets layer), additional media content encryption, and the use of a PKI (public key infrastructure), digital certificates can be issued, signed, mutual port-to-port authentication can occur. Transmission of content can result in the issuance of a digital certificate for each piece of content sent and reports can be generated for non-repudiation purposes, proving that a specific piece of content was, in fact sent and was, in fact received.

Automation

As file-based workflows replace traditional film and tape-based workflows, it will become increasingly important to link the movement of content to and from the various devices that will be used to process the content. The automated movement and processing of content must be accounted for in the implementation of a digital media distribution system. For example, a common workflow can be characterized by the following steps:

1. Content is ingested and deposited into a folder on a storage volume
2. A process is run which interrogates the storage volume folder and upon detection of a file moves the file to
3. An encoding system which implements existing profiles to encode the file into multiple formats. These newly encoded files are deposited in a specific set of folders
4. Finally, another process is automatically run which interrogates that set of folders, detects that new content is present and then transmits the files to their pre-designated recipients.

Performance and Scalability

An important consideration in the deployment of a digital media distribution system is how many points of interest need to be served and whether the deployment is accomplished in a client-server or managed peer-to-peer architecture. Fundamentally, how does the system need to be designed such that it can process not only the required job submissions to handle the business aspects but to also handle the actual movement of content which satisfies the operational requirements.

Integration and Interchange

The utilization of SOAP (Simple Object Access Protocol) and XML (Extendable Markup Language) - based APIs (Application Programming Interfaces) further facilitates the interchange with and integration to systems found in the M&E ecosystem for metadata, essence, and device communications. SOAP interfaces enable powerful web services implementations for interoperation between systems.

Deployment

Actual deployments of a digital media distribution system typically begin with content that is generated from within the facility and that must

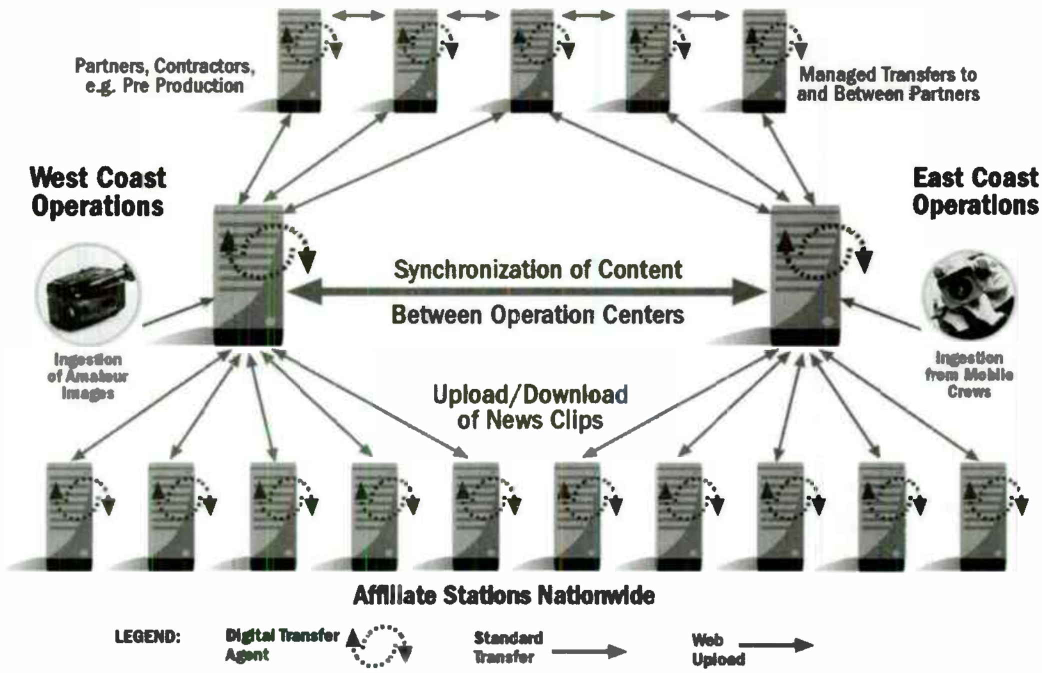


Figure 4 A deployment of a Digital Media Distribution System for a major television Network

move to areas outside of the facility. The growth of user-generated content and the timeliness with which field-based personnel must transmit content to reviewers who then decide on whether the content should be aired are both areas that need to be addressed.

Figure 4 is an example of an actual deployment of a digital media distribution system. In this example, central management is a mirrored configuration between east and west coast operations. Data mover software is placed at specific locations representing O&O (Owned and Operated) stations and at key affiliates. For transfers to and from vendor partners, data mover software is also located at key areas. Finally, for ingest and review of both user-generated content and in-field staff videographers, a web-based upload/download capability which provides for WAN-accelerated file transfers is provided.

Representative Schema of a Managed Digital Media Distribution System

Figure 5 is intended as a representative schematic layout of the components of a managed digital media distribution system. In this model, the central media manager is not in the data path of the actual movement of digital media but, instead, directs a distributed set of media agents to perform the movement.

SUMMARY

In summary, increasingly competitive and stringent business demands placed on the Media and Entertainment industry to create more content faster—to deliver it in more formats—to more places—and to more devices has raised the necessity to extend the management and control oversight of the process to a centrally managed digital media distribution system.

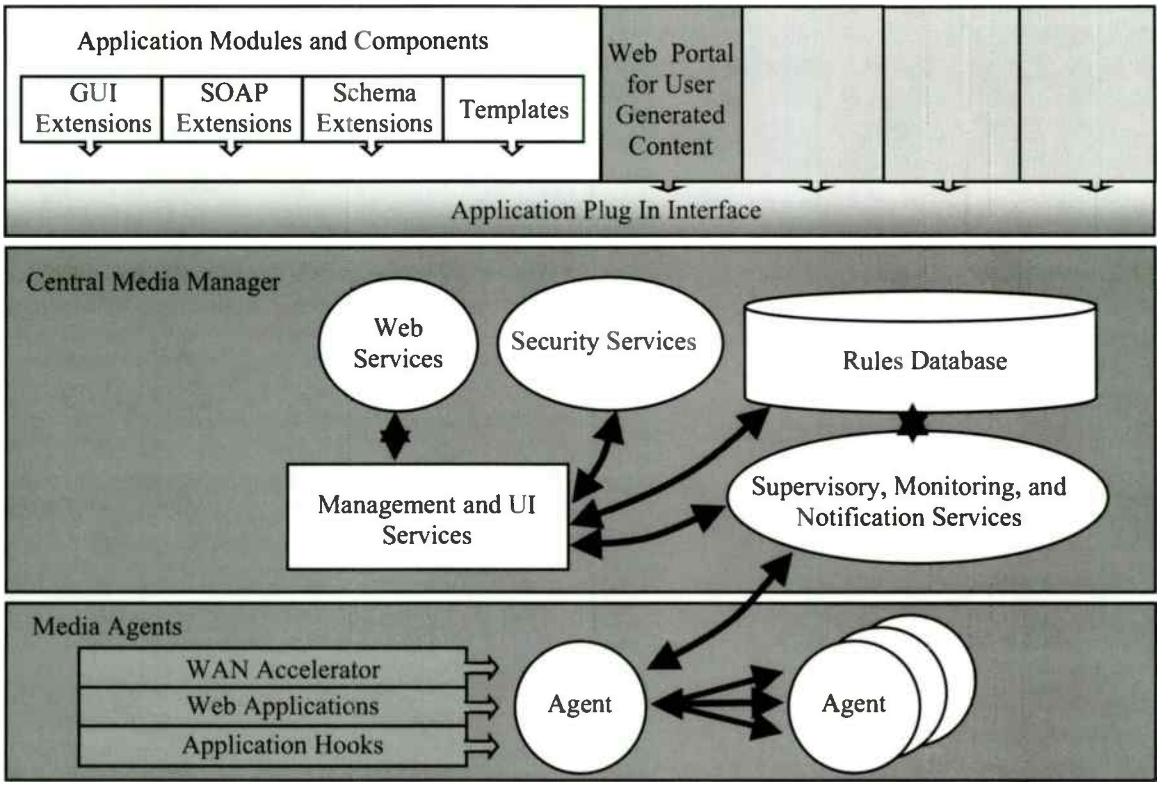


Figure 5 Components of a managed digital media distribution system

ACKNOWLEDGMENTS

A variety of research sources were utilized in the presentation of industry trends, including Morgan Stanley research and Gartner research.

Service Oriented Architectures For Broadcast Media Asset Management

PETER THOMAS
Blue Order Solutions AG
Kaiserslautern, Germany

ANDREAS MAUTHE
Lancaster University
Lancaster, United Kingdom

ABSTRACT

As broadcasters move from traditional to file-based production, distribution, and archiving, Enterprise Media Asset Management Systems (EMAM) become vital for managing business processes and content. MAM solutions that initially were dedicated to support departments like News, Programming or Sports, evolve into generic EMAM platforms managing entire business processes across those departments, from planning & drafting to commissioning, ingest, logging, cataloguing, search & retrieval, browsing, rough cut, collecting, ordering, editing, conforming, transmission, archiving, content exchange and content sales. They do this by offering workflow management, collaboration tools, content transfer, transcoding, migration and monitoring capabilities, and integrations with third party systems.

This ongoing evolution makes EMAM increasingly valuable for broadcasters, but puts considerable demands on the software architecture with respect to extensibility and integratability. EMAM must provide very high availability, scalability and resilience. Failures of system components must not result in a major loss in quality of service. Downtimes are difficult to tolerate, hence even software and hardware maintenance must be possible in operation, without perceivable loss in quality of service.

This paper explores how Service Oriented Architectures, an architectural style becoming increasingly popular in business process integration projects, can be applied to broadcast EMAM. Strategies for scalability and n+x redundancy, self-organisation, dynamic resource detection and allocation, and zero downtime strategies for maintenance and emergency situations are discussed. Finally, the proposal is validated against real-world workflows, and benefits compared to traditional system architectures are given.

INTRODUCTION

As broadcasters migrate to file-based production, distribution and archiving, Enterprise Media Asset Management (EMAM) systems evolve into indispensable systems for managing user workflows, system workflows, and content. EMAM systems initially designed to support formats like News, Programming, or Sports, continue to expand their reach and evolve into generic platforms providing integrated services across genres, formats and delivery channels. They also expand their scope, striving to embrace the entire broadcast workflow, from planning and drafting to commissioning, ingest, logging, cataloguing, search and retrieval, browsing, rough cut, collecting, ordering, editing, conforming, transmission, archiving, content exchange and content sales. In addition EMAM systems offer workflow management, collaboration tools, content transfer, transcoding, migration and monitoring capabilities, cross-system communication and integration with 3rd party systems.

This ongoing expansion of reach and scope makes EMAM increasingly valuable for broadcasters, but puts considerable demands on the architecture of EMAM systems with respect to extensibility and integration capabilities. EMAM systems also must provide a very high degree of availability, scalability and resilience. Failures of system components must not result in a severe loss in Quality of Service (QoS). Downtimes are difficult to tolerate, hence even software and hardware maintenance and updates must be possible in operation, without perceivable loss in quality of service.

In the past various EMAM system architectures have been used, all of which have their advantages and disadvantages. Examples are application server based systems and distributed systems using traditional communications layers such as CORBA [1].

In the following sections, we describe a system architecture style called "Service Oriented Architecture" (SOA) which has been applied very successfully in many business integration projects outside of EMAM. Adopting service-orientation ensures that the systems are open and can be integrated within upcoming infrastructures. This is also being done in other areas (e.g. eFinance [2] and the Grid [3]). We will show that this architectural style can be applied to broadcast EMAM, providing scalability, *n+x* redundancy, self-organisation, dynamic resource detection and allocation, and zero downtime strategies for maintenance and emergency situations. Finally, we will validate such architectures against real-world broadcast workflows, and identify benefits in comparison to traditional system architectures.

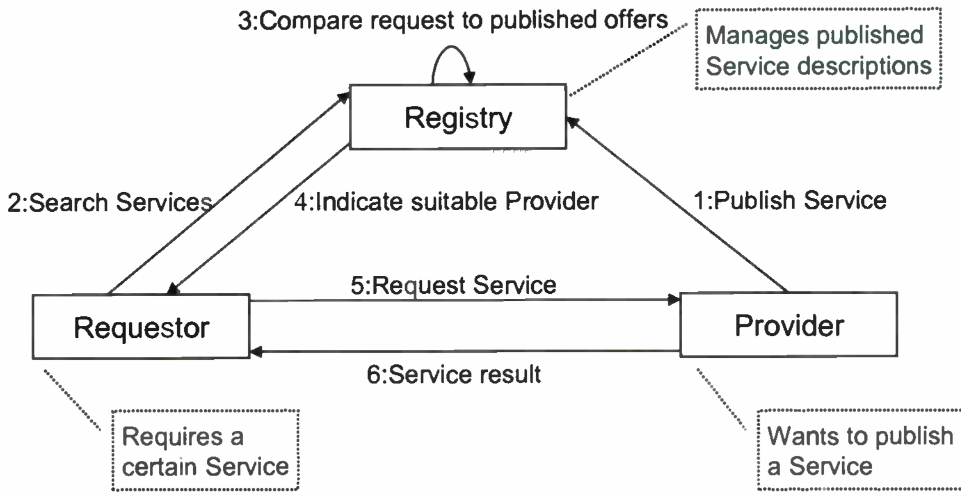


Figure – 1

SERVICE ORIENTED ARCHITECTURE

Basic Concepts

SOA is an architectural style specifically proposed for enterprise-class, distributed, IT systems. It is a promising concept for EMAM that already has been successfully introduced in other domains to achieve optimal support for business processes [2, 4]. In order to better understand how SOA can improve scalability and resilience it is important to understand some basic concepts.

SOA takes a business process oriented systems engineering view and translate it into an architectural style to build distributed software systems. In [12] SOA is defined as follows:

"An SOA is an architectural style that is reflected in an enterprise-scale IT architecture. It enables access to distributed resources, on demand. An SOA consists of a set of business-aligned IT services that collectively fulfil an organization's business processes and goals. These services can be choreographed into composite applications and can be invoked through open protocols."

Different aspects of systems engineering and software engineering are coming together in SOA. Crucial concepts in this context are the notions of *system* and *service*. According to [10], a system consists of elements:

- Elements are described by their properties, i.e. attributes and functions;
- Elements can interact via relations;
- An element can again be a system.

Interestingly, this matches a general definition for software architectures, showing that with SOA, systems engineering and software engineering finally merge [11]:

"The software architecture of a program or computing system is the structure or structures of the system, which comprise software elements, the externally visible properties of those elements, and the relationships among them".

Services in the context of a SOA are self-contained, re-usable, application functions that are offered by a provider and can be applied to fulfil tasks in business processes. They are loosely coupled, stateless and autonomous [4]. Further, a service provides business data and/or transform business data from one consistent state into another. Service interfaces usually are based on open technology, and are well-defined, standardised, and platform independent. Over such interfaces services publish data and application functions but hide the implementation details. Interdependencies between services are managed in overlaying business processes and underlying data models — the individual service has no knowledge of such interdependencies.

At present the SOA paradigm is mostly realised using Web Services [5]. A Web Service is a software component that uses well-defined interfaces and bindings through which it can interact with other Web Services, using XML messages. The Web Service Definition Language (WSDL) [6], Simple Object Access Protocol (SOAP) [7] and UDDI (Universal Description, Discovery and Integration) [8] form the core of the Web Services standard suite. BPEL4WS (Business Process Execution Language for Web Services) has been devised to model (complex) business processes that are executed by Web Services [9].

Dynamic Resource Detection and Allocation

SOA can provide dynamic resource detection and allocation using an approach comparable to the yellow pages (Figure – 1). A provider publishes services to a registry. This publication contains a full specification of the service interface and service contract. If a requestor wants to use a service offered by a provider, it first locates the provider by stating a request to the registry. The registry compares requests and delivers locations of suitable providers. The requestor selects a provider, requests the service, and receives the service result. A provider can also be a requestor of other services. The implementation of a service is hidden from the requestor. This allows choreographing higher-level services from lower-level services.

Service choreography typically is done using a business process definition language, e.g., BPEL4WS [9]. Thus, large-scale business processes are hierarchically constructed from lower-level business processes, while the basic actors stay unchanged.

Being centered on business processes instead of objects and application functions, SOA allows organizations to easily accommodate changes in their business, and it facilitates consistent communication between business and technology experts. In addition, the strong component orientation and the detailed interface specification ease the transparent integration of legacy and third party systems (a key requirement for EMAM). To do so, wrappers are provided that integrate the "alien" systems into the SOA as a well-formed service. Then, business process descriptions are adapted to include the new service. Hence, only n instead of $n*(n-1)$ interfaces have to be developed and maintained.

Revoking a Service Publication

In SOA a provider offers autonomous and stateless services by publishing them in a registry. This implies that the provider can also revoke this publication. If the service completes its current tasks before going offline, and if sufficient compatible resources are offered by other services, this has no negative impact on the QoS of the entire system, as new tasks are automatically assigned to other available services. In short:

- A provider can, at any time, add service offerings to any registry without impairing the overall system;
- A provider can, at any time, revoke services without impairing the overall system, provided the service completes open tasks before ceasing operation.

Essentially, revoking a service publication results in the service receiving no more service requests. After completing all active or queued tasks, the service is idle and can be shut down, e.g., for maintenance, and can be re-launched at any time.

Resilience and Zero Downtime Maintenance

SOA provides a very reliable and efficient model for resilience and maintenance. Business processes that invoke services typically are managed by workflow engines, which themselves again are services (offering workflow management). Multiple workflow engines may access the same workflow related data as long as it is shared via underlying persistent and safe information storage (such as a database), and all tasks are processed as transactions. Hence, a workflow engine is no single point of failure, as in case of failure another workflow engine can take up unmanaged workflows from the database and can continue processing them.

The workflow description processed by such a workflow engine specifies which resources are required to complete the various tasks in the workflow. According to this description, the workflow engine locates resources in the registry and assigns the tasks. This is augmented by requesting that a workflow engine also:

- Monitors the progress of each task as it is being completed by a service;
- Can, at will, revoke an assigned task from a certain service and re-assign the task to another service.

This introduces a high level of resilience. In case of failure or underperformance of a service, a workflow engine in a SOA can re-assign work on a task level. Depending on the task and the underlying data model, tasks may have to be restarted or can be continued. In short, as long as there are spare resources (i.e., services) available, a SOA can automatically and gracefully recover from unexpected events.

This allows to employ a very cost effective $n+x$ redundancy model (instead of $2n$ for classic mirroring or clustering), where n is the number or resources (i.e., services) required to sustain a given QoS, and x is the number of services that may fail before the desired QoS can no longer be maintained. In practice, $x = 2$ often is a good value, allowing one service to be offline for maintenance while still allowing that another service may fail without impairing QoS. Thus, zero downtime maintenance can be achieved simply by maintaining one service after the other and bringing it online again before moving on to the next.

Load Balancing

The key benefit of an SOA is the fact that it is inherently able to provide a very powerful model for scalability. This can be achieved by augmenting the service lookup mechanism as described above by a dynamic resource allocation strategy. When multiple identical services are available, load-balancing techniques can be applied to optimally distribute requests over the available resources. To do so, two basic questions have to be answered: How are the decisions criteria derived that are used to distribute requests over services, and where is the load balancing decision taken?

Decision criteria can be derived both from parameters that are associated with the respective task to be

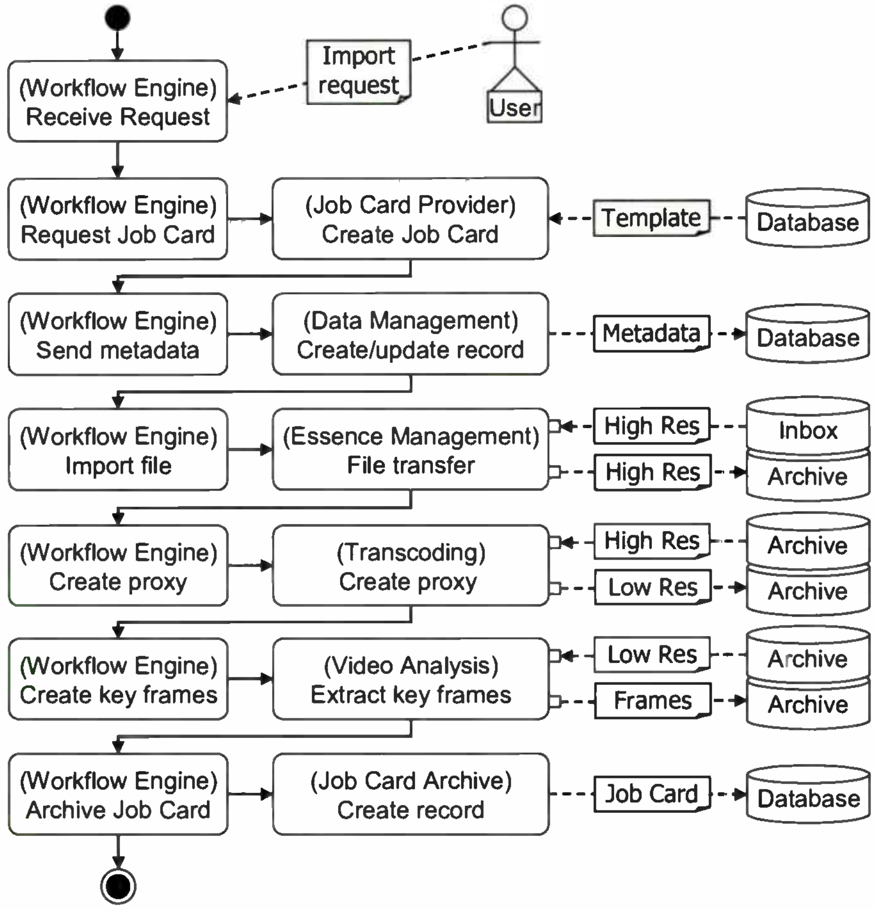


Figure – 2

performed, and from parameters related to the service itself. Parameters associated with the task are, for instance, user or task priorities (e.g., a user working for News may have higher priority in restoring a video file from the archive to the newsroom production server than a user working for a current affairs magazine), or locations (local vs. remote). Parameters associated to the service are, for instance, the current load of the service (e.g., 15 of 20 available resources for streaming video are in use at one service, while another service is mostly idle), or the available network bandwidth between provider and requestor. Apart from this there may be other system specific parameters guiding the decision to which service a new task is assigned. The decision itself is based on the costs for delivering the service. This cost is derived via a cost function that considers all decision criteria. The task is assigned to the service that bids at the lowest cost.

There are basically four locations where the cost function can be applied and the decision can be taken: The registry, the requestor, a dedicated cost service, and the service itself. A number of drawbacks have been identified in the case where the registry takes the decision. The registry would have to know all cost models for all services, which implies configuration changes in case of system reconfigurations (changes in business models and adding or removing services), or the obligation to communicate for each incoming request with each candidate service in order to request a quote. Further, the registry may see considerable load that could impair its core function of ultra-fast service location.

A good approach is to use a dedicated cost service and combine it with the desired service as decision-making instance itself. In this case each service is its own dispatcher. As it knows its internal implementation it can easily derive costs that apply to itself, and it can use a dedicated cost service to derive task related cost. The cost service could for example derive the network cost from the provider IP and the service IP, or could request priorities from User Management in order to derive priority cost.

As a service also knows its own interface specification, it can query peer services published in the registry and request quotes from these services. Based on the result of the cost analysis, the service can then decide to either perform the task itself, or to forward the execution to a "cheaper" peer.

Scalability

Provided that the individual services comply with the requirements for a dynamic resource allocation, load balancing is native to SOA. Scalability is achieved by adding more hosts and redistributing the services accordingly when more computing power is required, and more services when more resources are required. This is possible at runtime without impairing QoS.

SOA and Rich Clients

For EMAM the SOA concept is applied at the architectural level such that the EMAM provides an encapsulation of all its different architectural modules into individual services. The summary of all services and their respective functionality then represents the overall functionality that the backend of the EMAM can offer.

Additional functionality may be embedded in rich clients. However, as client applications do not offer their functionality as services, such functionality is unavailable without human participation and can only be used in workflows that include related user tasks and activities. Rich clients hence are undesirable in SOA.

SAMPLE WORKFLOW: VIDEO FILE IMPORT

Consider an EMAM that offers the following services: Workflow Engine, Job Card Provider, Data Management, Essence Management, Transcoding, Video Analysis and Job Card Archive. Figure – 2 shows a high resolution video import workflow processed by that EMAM.

A user launches an import job by providing the file name, source location, priority etc. to the Workflow Engine. As the services in an SOA hide their implementation from the requestor, the requestor (in this case the user or his/her application) has no knowledge of how the job can be completed. Hence, the first step in fulfilling the request is to create a job card that prescribes in detail how the job shall be processed. This is done by passing the parameters to the Job Card Provider, which creates the job card from a matching job card template and returns it to the requestor (i.e., the Workflow Engine). The language in which workflows are prescribed in job cards is BPEL4WS (see above).

Following this the Workflow Engine steps through the various tasks as prescribed on the job card. In our example the first task is creating or updating metadata in the Data Management. After successfully completing this task, the job card states that the Essence Management shall import the high-resolution video file from a specific video server. Hence, the Workflow Engine sends this request to the Essence Management, which performs a file transfer from the inbox server to the archive. After completion of the transfer, the Essence Management acknowledges success to the requestor.

The next step on the job card is creating a browse proxy from the high-resolution video file via transcoding. The workflow engine sends a request for proxy creation to a Transcoding Service that offers this function. The Transcoding Service, being both a service provider and a service requestor, requests a copy of the high-resolution video file from the Essence Management, transcodes the file into a proxy format, sends the proxy back to the Essence Management, and subsequently deletes the high resolution copy. The Workflow Engine is not aware of this internal sub-workflow as, according to the philosophy of SOA, the Transcoding Service hides implementation details from requestors. After completing the transcoding the Service acknowledges success to the requestor.

Subsequently, the job card requires that a key frame selection and extraction shall be performed. As the Video Analysis Service offers this feature, the workflow engine sends the request to this service. Again, the Video Analysis may employ an internal sub-workflow (request streaming access to the proxy from the Essence Management, perform analysis, store timecodes in Data Management, store key frames in Essence Management). After completion, the Video Analysis Service confirms success. Finally, the Workflow Engine archives the job card, including all timing information.

Clearly, in an SOA, system workflows are closely aligned with business processes. If changing requirements force organizations to change their business processes, the implementation can easily be adapted by modifying the according job card templates, and if necessary adding services to accomplish tasks that the system can otherwise not perform.

OTHER ARCHITECTURAL STYLES

Application server-based architectures tend to be simple to install and maintain. However, with respect to scalability, resilience and maintenance, the system depends on the capabilities of the application server and hence is very much application server specific. An EMAM built on top of a classic application server can not be expected to provide levels of scalability and resilience as a system that uses SOA.

Classic distributed systems architectures can provide comparable levels of scalability and resilience. However, interfaces to services in distributed systems tend to be proprietary and rarely are fully disclosed. Vendors tend to withhold interfaces and information, both to protect their product and market position and to avoid long term commitments to maintain them in future releases. This makes it difficult for third party companies or customer IT departments to customize and enhance such a solution. SOA requires that the information required to invoke a service is disclosed in the registry, typically in WSDL [6].

Thus, the information required to invoke a service is well-defined and accessible. It is also straightforward to add additional services and include them in workflow descriptions. This allows customers to enhance the system by adding third party tools or to implement integration services to include their in-house systems.

Finally, being very much focused on mapping business process and bridging the gap between business process engineering, systems engineering and software engineering, SOA is very well suited to evolve and adapt to the changes of an enterprise while the enterprise itself evolves and adapts to changes in the market.

CONCLUSION

Enterprise Media Asset Management Systems (EMAM) are rapidly evolving into the central information and content hub for a broadcaster, becoming increasingly vital to the operation. Thus, it is imperative that these systems not only make provisions to prevent failures but also can recover gracefully from malfunctioning components or even breakdowns of parts of the system. System resilience is the key to maintain a 24x7 operation. Further, EMAM need to be able to adapt to increasing demands and changes in business processes and organisational structures. Therefore, scalability is another crucial feature of EMAM.

In this paper, using Service Oriented Architectures (SOA) for building EMAM has been discussed. SOA is inherently self-organizing and very flexible with respect to adding, locating and using additional resources. Thus, failures leading to unavailability of certain resources are covered by automatically involving spare resources, while scalability is addressed by adding specific resources to handle the increased or modified work.

Complementing SOA with a dynamic resource allocation strategy provides an even more elegant solution to scalability, resilience and zero downtime maintenance in a large-scale, distributed, 24x7 systems. Provided that the individual services comply with the requirements for a dynamic SOA, it is sufficient to provide more services than actually needed to guarantee the desired QoS. Distributing additional services over multiple hosts increases resilience. Scalability is achieved by adding more hosts when more computing power is required, and by then redistributing the services accordingly. Zero downtime for maintenance also can be facilitated by installing more services than needed, and by making sure that these services are distributed such that an entire host or provider can be taken offline without loss in QoS.

The core concept of building EMAM from individual components and choreographing them into a collaborating system that distributes the load dynamically over multiple component instances is natural to a SOA. Being an architectural style for distributed, heterogeneous enterprise IT solutions, SOA is a very promising approach for designing and building broadcast EMAM, and has considerable benefits compared to other architectural styles.

REFERENCES

1. Mauthe, A. and Thomas, P.: Professional Content Management Systems – Handling Digital Media Assets, John Wiley & Sons Ltd, Chichester, UK, 2004.
2. Berbner, R., Heckmann, O., Steinmetz, R.: "An Architecture for a QoS driven Composition of Web Service based Workflows", in Proceedings of 2005 Networking and Electronic Commerce Research Conference (NAEC2005) Riva del Garda, Italy, 2005.
3. Foster, I.: "Service Oriented Science", Science, Vol. 308 (5723), 2005.
4. Papazoglou, M.P.: "Service-Oriented Computing: Concepts, Characteristics and Directions", in Proceedings of 4th International Conference on Web Information Systems Engineering (WISE 2003), Rome, Italy, 2003.
5. Austin, D., Barbi, A. , Ferris, C., Garge, S.: "Web Services Architecture Requirements", W3C Working Group Note, http://www.w3.or/TR/wsa-reqs/, W3C, 2004.
6. Chinnici, R. et al.: Web Services Description Language, version 2.0 (candidate recommendation), http://www.w3.org/2002/ws/desc/, Januay 2006.
7. W3C: "SOAP Version 1.2", W3C Rec., http://www.w3.org/TR/wsa-reqs/, 2004.
8. UDDI Spec Technical Committee: "UDDI Version 3.0.1" Technical Committee Specification, http://uddi.org/pubs/uddi_v3.htm, 2003.
9. Andrews, T. et al.: Business Process Execution Language for Web Services, version 1.1, http://www-106.ibm.com/developerworks/library/ws-bpel/, May 2003
10. Igenbergs, E.: Grundlagen der Systemtechnik, Lecture, Fachgebiet Raumfahrttechnik, Munich Technical University, 1993.
11. Bass, L., Clements, P., and Kazman, R.: Software Architectures in Practice, 2nd edition, Addison Wesley, Reading (MA), 2003.
12. Asarjani, A.: The SOA Best Practices Blog, http://www-106.ibm.com/developerworks/blogs/dw_blog.jspa?blog=490, March 2005.

Parts that are more than the Whole: Asset Management in the IT Age

MICHAEL ATKIN
BroadView Software Inc.
Toronto, Canada

ABSTRACT

Not long ago, "asset management" meant properly cataloging and shelving tapes containing finished productions. In this legacy tape library environment, metadata consisted of having the right labels on the tape boxes.

Now, with digital media located in server and file-based environments, the very notion of shelving a "finished production" evaporates. To maximize the value of media assets, elements that make for finished production must be available for re-use, which means they must be ingested properly at high bitrates, then identified, catalogued, and indexed by weaving the appropriate metadata into the media.

Since content owners must make a major investment to bring legacy content into the digital domain, it is essential that these steps be done properly to avoid having to repeat the process in the years ahead. Once properly prepared, content elements facilitate metadata-driven "flattening" - automated packaging and repackaging in an array of formats for distribution on a growing variety of platforms such as iPods, Video-on-Demand, and over-the-air playout. The metadata also provides the technical infrastructure needed for the emerging micro-media industry now coming to the fore with outlets like YouTube.

Micro-media is the natural complement to existing mass media and will become as ubiquitous as written messaging is today. Empowered by metadata, micro-media will possess exceptional utility and flexibility enabling rights management for content owners and value to consumers. Like content, promotion and advertising materials must also exist in this ready-to-mix form.

The key to realizing this asset management paradigm is through a new concept in broadcast management software that goes beyond traditional programming, traffic and billing software to integrate systems across the entire enterprise. Such Media Fabric Management (MFM) tools facilitate micro-media creation. They optimize workflow to weave together media assets however needed, track where they are allocated and monetize use while delivering searchable, customizable content that consumers will pay a premium to access.

An effective Media Fabric Management system allows content creators, broadcasters and other media distributors to seek out and serve today's mobile and increasingly fragmented audiences. Media companies that do not adopt asset management systems that are MFM capable risk having to start over in the mid-term and play catch-up with those already moving forward with them today.

WHERE'S THE SHOW?

Until recently, media consumers have been a remarkably compliant group. They have accepted how, what and where media companies have wanted to deliver content to them. Now, changes in technology and society have created a need to allow for greater mobility and diversity in how media is consumed. But the legacy technical infrastructure cannot hope to keep up with this emerging marketplace.

New tools and concepts for asset management are necessary to accomplish this. But it will take far more than just a simple upgrade of tape libraries into server-based digital media storage. These systems can no longer be thought of as standalone operations. Instead, they must be tightly integrated elements of an enterprise-wide Media Fabric Management system. In this way, asset management allows for the marriage of mass and micro media with shared content resources serving a multitude of distribution platforms.

FROM SIMPLICITY TO COMPLEXITY

For many years, audiences received content only two ways. The film industry released its content through movie theaters while television offered "broadcast quality" programs delivered over-the-air. In this setting, audiences would dutifully gather when and where they were told and would gladly accept what was provided to them. Here, the only repurposing of material came with broadcast of the theatrical releases. Managing assets with tape libraries that mimicked traditional libraries for books worked well in such a media market. Relatively simple tools managed the tape catalog and manual intervention was the rule.

Then, the arrival of cable TV and home VCRs set in motion the fundamental changes we see today. Initially, these simply expanded the existing model by opening new delivery platforms for the same "broadcast quality" programs. But under the surface, a major shift had occurred. The traditional theatrical and broadcast media marketplaces were defined by a scarcity in the number of available time slots to present content. Suddenly, cable and video opened up a broad horizon limited only by the audience's attention and interest. VCRs, too, freed viewers to watch at their own convenience according to their own schedule.

Initial efforts to take advantage of this abundance began with "extras" offered on video to add value. Production elements, trailers and other bits and pieces associated with a program were tossed in providing a "bonus" above and beyond what was available through traditional broadcast outlets. Likewise, the need for massive amounts of content on cable TV offered opportunities for outtakes and other ancillary material to find use.

Distribution, however, remained locked to the same broadcast quality content for viewing on standard home televisions. The only additional complexity was repackaging content for home videocassette for playback on the same standard definition TV platform. For all the growth of content available for VCRs and cable TV, it could still be managed with traditional tools amplified by advances in computing technology. A variety of standalone applications focusing on particular functions including programming and traffic software addressed specific issues to alleviate local workflow bottlenecks.

PUSHING BEYOND THE BOUNDARIES

Throughout this evolution, these standalone systems assisted by manual intervention have grown up to carry out operations. Today, such side-by-side operations integrated by manual efforts are typical in even major cable TV networks. But the new demands for delivering on multiple platforms stretches such systems past the breaking point.

Legacy operations simply cannot manage the new complexity to take advantage of emerging opportunities. With advances in technology and changes in society, audiences are increasingly unwilling to follow the dictates of content distributors. To continue to be successful and to maximize the value of content, media companies must understand and appreciate how and where audiences wish to access content. Then, the challenge is to deliver it there.

This is a practical impossibility with the tangle of unintegrated systems creating an unbreakable bottleneck and strangling efficiency. Beyond accomplishing these basic production issues, the existing infrastructure lacks the ability to add fundamental value essential for the emerging "micro-media" marketplace.

But metadata-driven tools allow for new efficiencies in asset management by empowering multi-platform production automation. These tools also provide content owners the ability to track distribution and manage rights. Likewise, they give content important value-adds for consumers by providing text-like search capabilities.

The conspicuous success of online outlets like YouTube just scratches the surface of the coming opportunities. This next-wave Dot-Com success is to the coming micro-media market what the first generation (circa 1995) online search engines are to today's Google. Leaps in consumer utility and acceptance drive a thriving next-generation market that is as unexpected as it is exceptional.

DEFINING MEDIA FABRIC MANAGEMENT

To take advantage of these opportunities requires far more than simply trading tapes for server-based digital archives to shuffle programs for broadcast more efficiently. It requires rethinking and reengineering the entire enterprise to place operations under an end-to-end tightly integrated software solution. Today's standalone systems must be knit together into a single fabric that runs through the entire media enterprise.

This new category of systems, dubbed Media Fabric Management (MFM) software, is designed to bring previous standalone packages including programming, traffic, scheduling, billing, digital library management, and automation into a unified, interoperable whole. This advance is akin to the move away from disparate word processing, database and spreadsheet applications in the DOS era to today's integrated office suites.

THE INGEST INVESTMENT

Under this new operational interface, although a wholesale replacement of existing operations would be ideal, realistically; piecemeal replacements of individual systems are likely to be the rule. However, individual upgrades should be carried out within the context of the emerging MFM environment.

It is especially important that existing media assets are ingested at a high bit rate to facilitate repackaging and repurposing in whatever platform that may provide opportunities in the years to come. At this stage in the transition, the goal is to carry out an ingest process that avoids having to go back to the original tapes. Ingesting the legacy content today with an eye to future multiplatform use can achieve that.

The reason for focusing on ingest is simple. Ingesting content is probably the most expense step in operating an asset management system. Ingest requires extensive manual intervention requiring direct supervision throughout the process. Monitoring for and applying an organization's technical requirements (tech evaluation) requires that content be ingested in real-time so that it can be monitored frame-by-frame. Should any technical issues emerge (e.g. audio level problems), then manual intervention is required. Occasionally, content may actually need to be passed through an edit suite for preparation before another, final ingest can be completed. All of this effort takes time and significant expense.

Resolution of the ingested components needs to be high enough to support both the end operations and formats that the organization anticipates using. If the organization expects to edit and manipulate content, then an I-Frame, frame-accurate ingest format must be used. If editing is not important, then the bit rate of the ingest must still be high enough to support all of the anticipated end formats and bit rates without creating any noticeable compression or transcoding artifacts.

KEYS TO MICRO MEDIA: "FLATTENING", RIGHTS MANAGEMENT & UTILITY

Moving beyond the traditional broadcast model to develop the nascent micro-media marketplace requires two factors. First, the cost of repurposing content must be minimal. Second, a new metadata / rights management infrastructure needs to allow for communication between rights holders and those using their content. Together, these will create new markets and new opportunities that will reshape the media industry much like the arrival of VCRs and DVDs.

Rather than replace traditional outlets as initially feared, VCRs and DVDs created an unexpected bonanza for repackaging media assets. Micro-media, too, will exist alongside mass media to reinvigorate the entire media industry.

COST-EFFECTIVE REPACKAGING

Once ingested, the cost of repackaging digital content for multiple platforms and target audiences must be negligible. Beyond prepping content for playout in the traditional broadcast setting, the newly ingested content must also be suitable for "flattening", a process where tools dynamically package content combining desired elements and then finishing them in whatever format needed. Producers simply select what they want and for what platform. The MFM system automatically pulls content elements from archives, weaves them together, and then turns them into a finished product. Ideally, the process becomes completely automated and there is no need for intervention by human editors or producers.

An emerging generation of products and systems that anticipate the larger MFM architecture by carrying out this flattening process are already deployed. The SeaChange QuickSilver workstation is an example of an off-the-shelf MFM-ready component delivering real value today. At PBS member station WHYY in the Philadelphia area, the SeaChange system is in use preparing and delivering VOD content for distribution to cable operators and Web-based material for online audiences. This is a simple, seamless way for WHYY to program and repurpose content for multiple platform distribution eliminating the need for manual intervention. WHYY is able to effectively leverage server-based programming assets for additional market opportunities without adding staff to manage it. It is an example of an emerging generation of products and solutions.

A customized, enterprise-scale MFM installation is also in place at the PBS Media Operations Center (see case study below). The PBS corporate solution to content preparation or flattening utilizes an automation driven approach designed to handle more complex content preparation than the QuickSilver or transcoder manages. PBS utilizes Harris automation control of a device chain consisting of a high resolution server playout port, an aspect ratio converter, audio and image branding tools, closed captioning system, Nielsen encoder and finally an encode port capable of generating video files in the desired resolutions and formats. This design has allowed PBS to not only prepare files for distribution but to also modify them, brand them and monetize them all under automatic control of the MFM system.

METADATA OPENS NEW MARKETS WITH REPURPOSED CONTENT

Once content has been successfully ingested, it can be repurposed on various platforms for use within a media organization as well as for external distribution. Internally, proxy versions may have many uses. For example, many organizations carry out content screening and review (often referred to as Standards and Practices or S&P). This is an ideal use of a medium / high resolution proxy generated from a successful ingest. This digital file (produced as an MPEG, Windows Media, Flash, etc. file) can be shared instantly throughout the organization and viewed concurrently.

Note that screening / S&P is also an ideal time to gather detailed metadata about the content including language, violence, nudity, embedded objects like web tags and product placements, etc. Once collected and stored within the MFM system, this screening information can then be consistently and easily shared within and outside of the organization. For networks, such information can help support affiliates by making it simpler to review content and apply whatever local Standards and Practices that may exist.

Metadata's utility extends far beyond such typical broadcast uses and opens up many opportunities. Metadata provides the technical infrastructure to unlock the full utility of micro-media for consumers by adding the same search capabilities as text. It also enables rights management to support whatever business model can properly leverage this value-add for consumers.

A key challenge for the media industry as a whole is developing standards to facilitate tracking metadata about assets and maintaining the accuracy of this information. As Media Fabric Management systems advance, they should address this issue directly. Metadata creation and modification should be straightforward and accessible throughout the organization. In many cases, it can even be carried out automatically. This must happen alongside organizational change to make sure that proper focus, procedures and workflow are in place to support effective management of metadata.

New concepts in external metadata use could prove revolutionary, opening new horizons for the media industry. Today's approach is to embed limited metadata within distributed assets / objects. Through the emergence of new standards and ubiquitous interconnectivity through the growth of the Internet, capabilities for a uniform and consistent method of querying metadata would greatly increase the power of that information. Ultimately, any external individual could use the embedded information within an asset to query the originating organization for metadata updates.

This advance would parallel the arrival of the networked computer environment. Before, software updates had to be delivered manually on physical media. Now, they can be accomplished automatically through any Internet connection. In the same way, metadata is quickly evolving into a dynamic communications tool, to the point that it could be delivered in an XML document format, ready either for display, indexing or consumption by other systems.

ASSET MANAGEMENT MORE THAN JUST ARCHIVES

The arrival of inexpensive content repurposing along with a new rights management infrastructure will realize the promise of the emerging MFM model for an IT-based multiplatform media distribution enterprise. This extends far beyond the traditional broadcast concept.

Broadcasting relies on the compliance of the audience to accept content in a set fashion. But the monolithic audience eagerly sitting in front of televisions according to the network schedule is fast evaporating. Instead, media companies must now vie for the attention of mobile individuals interested in a wide variety of entertainment on multiple platforms, from gaming consoles, to online attractions, to content downloaded on iPods. The challenge is to identify who and where customized packages of content tailored to specific interests can be sent. As promised in the heyday of the Dot Com boom, the future consists in "narrowcasting" to gain the attention of specific market segments.

Another business model that parallels MFM's likely development is the direct mail marketing industry. Desktop publishing transformed direct mail marketing by eliminating much of the traditional production costs associated with it. The technology made producing the "content" that goes into mailings simple and inexpensive. There is no longer any need for a full-fledged art department utilizing a large creative team to construct pieces. Instead, a talented, technology-enabled individual can operate a successful operation. The limit here is not the cost of production but rather printing costs, mailing expenses and the investment in developing the database of existing and potential customers.

An MFM-enabled media operation leveraging an extensive library of media assets should operate much the same way. The cost of assembling and repackaging content components for a target audience must be negligible. Then, tracking rights to monetize the delivery of assets provides the financial incentive for content creators to produce for the emerging medium.

The possibilities for micro-media marketing are enormous. Unlike direct mail, it is not hindered by material costs and the cost of delivery can also be negligible. Properly realized, the MFM model becomes a pure marketing play. Success will come from gathering detailed information about target audiences and custom crafting content to meet their lifestyle needs.

While Media Fabric Management appears to take us far beyond what we now typically think of as "asset management", it is a logical consequence of the concept placed in context of the rising diversity of media platforms. While still a visionary concept, the reality is far closer than many might imagine. Most importantly, anyone planning to ingest legacy content into the digital domain is well advised to consider the MFM context and to be prepared for it.

MEDIA FABRIC MANAGEMENT CASE STUDIES:

PBS' PREP FOR ASSET MANAGEMENT

BroadView Software has pioneered the new category of fully integrated Media Fabric Management solutions at numerous notable broadcast operations. At the PBS Media Operation Center, BroadView designed and implemented an MFM framework over a next-generation server- and file-based broadcast engineering environment. The system totally integrates scheduling and trafficking with post-production functions and forms a foundation that can easily add distribution platforms as needed, including Video-on-Demand.

The three-year project has brought all operations under BroadView's unified graphical user interface. The system is fully automated from receiving content through initial encoding and archiving, then to playout and repurposing of material for multiplatform distribution.

The new system is the result of close collaboration between BroadView, PBS technical staff, and numerous other vendors and consultants to ensure both the technical integrity of content alongside accurate, complete metadata. The end result is a step beyond traditional engineering architecture and processes to achieve workflow innovation. It achieved the design goal of radically reducing the need for re-entry of content-related data and the data-entry errors that ensue.

The cornerstone of this MFM approach is accurate and updated metadata. Metadata enables the flattening process for repurposing and repackaging of programs. The end result is a substantial savings in production time, materials and human resources while raising the technical quality of programs delivered to PBS member stations. At the same time it enables seamless production and delivery to other distribution platforms such as IP delivery, VOD and Web-based media.

AN END-TO-END SOLUTION

The MFM system for the PBS Media Operation Center begins with how producers submit content. Before any source material arrives at PBS, content creators send a Media Inventory via the Internet that provides program metadata including frame-accurate timing. Once reviewed and approved by PBS, the MFM system generates a printable bar code label that is attached to the physical media and then is scanned upon delivery. Once marked as "arrived" in the system, the MFM system generates an advanced authoring format (AAF) file along with a work order so that technicians have everything needed to prep the material so that it is properly ingested. The last step at ingest is archiving the original material in various formats so that it can be reworked as needed. This makes it simple to add additional features such as closed captioning, DVI, program-specific tagging, Nielsen monitoring information, etc. Once ingested, requests for such changes submitted through the MFM system create a work order for editors while summoning up all the required program material. This streamlines manual intervention. Editors can get to work immediately instead of having to prep by gathering necessary elements.

The ingest process requires a great investment in staff time to carry out the technical evaluation of the content. But the return for this investment is enormous. It becomes possible to explore endless opportunities for repurposing and repackaging with negligible production costs. In most cases, the MFM system allows for repackaging programs automatically by manipulating metadata, eliminating the need for costly manual intervention. For example, switching out an advertiser's embedded spot or promotional item can be accomplished completely within the system. The MFM system removes existing elements and stitches in replacements by referencing the metadata for the different pieces to ensure a seamless whole. This "flattening" saves hours of content re-ingestion, editing room manipulation and tape recreation. Through the same flattening process, the MFM system automatically creates the final program file for distribution to PBS member stations.

While Media Fabric Management marks a fundamental departure from traditional broadcast engineering, it is adapted from well-established IT-driven manufacturing methodologies. It delivers the same kind of mass-customization applied to both durable goods and consumer products allowing for carefully tailored, individual production packages created without significant losses in efficiency or cost effectiveness.

This initial implementation at PBS can easily be advanced without major changes in the new infrastructure. For example, metadata creation will be pushed to the edges of the workflow. Content creators will eventually encode metadata as part of the production process prior to assembling the finished content. Rather than deliver content on legacy tape media, sending digital files fully described with metadata tied to the production elements will streamline PBS' ingest process.

BET BRINGS TOGETHER STANDALONE APPS UNDER MEDIA FABRIC MANAGEMENT

Viacom subsidiary Black Entertainment Television (BET), the leading African American multi-media entertainment company, has committed to adopt

BroadView's MFM strategy to enhance its ability to leverage metadata. This opens new opportunities across multiple distribution platforms to add value for its audience by effectively managing its growing library of both created and acquired content.

BET's situation was typical for many organizations. The operation has been run by a patchwork of standalone department-based applications cobbled together with various office software packages. These informal, idiosyncratic systems had become an embedded feature of the organization over its 25-year history. Finally, the organization was pushed to a point of decision because these homegrown methods had become cumbersome and unwieldy. Using spreadsheets and other general office systems adapted to the specialized work only supported individual departments without considering the value to the overall organization. Having to reconcile these standalone operations left them constantly playing catch-up to meet daily deadlines, leaving little time to take advantage of emerging opportunities.

THE GENESIS FOR MEDIA FABRIC MANAGEMENT IN PROGRAMMING

The genesis of the MFM system at BET started with an interest in establishing a new system for programming in early 2005. A key driver was the way different groups within programming carried out different tasks – finance, payments, and amortization. Difficulties in sharing information between these groups multiplied as each group added detailed information in different ways. For example, notes added to a spreadsheet about amortization required some consultation before the impact on finance could be made explicit. As challenging as this was for existing staff experienced with these methods, introducing newcomers to this idiosyncratic style was monumental. With the programming library set to expand significantly, the added overhead of this inherently inefficient system was no longer acceptable.

To address this, BroadView began implementing its core MFM-ready solution to knit together the different aspects of the programming department in early 2006. With the success of the installation, BET now seeks to extend the new Media Fabric Management paradigm across the enterprise to unify production, master control, and the digital library under the same common interface.

BET's goals are to reduce the number of active applications utilized to accomplish the core business functions of the organization. In addition, the aim is to enable interactions between applications to allow for data sharing, reduce duplicate data entry and provide for significant efficiency and workflow gains. Then, with the total asset management that MFM architecture provides, the whole focus of the organization can shift.

Instead of scrambling to maintain the existing operation, BET becomes a forward-looking enterprise able to effectively implement an aggressive program of marketing across multiple platforms. Ultimately, the relationships it will be able to develop and leverage with its audience will serve as a model for the new micro-media industry.

ACKNOWLEDGEMENTS

Carl Lindemann
Cyberscene.com

Bill Weber
WHYY Philadelphia

Deirdre Jackson
Black Entertainment Television

Michael Hunt
PBS

Seachange International

iPod is a registered trademark of Apple Inc.

G. Davenport, B. Barry, A. Kelliher & P. Nemirovsky, "Media fabric – a process-oriented approach to media creation and exchange". MIT Media Labe, BT Technology Journal, Vol. 22 No 4, October 2004.

CONTENT STORAGE MANAGEMENT IN THE NEWSROOM

BRIAN CAMPANOTTI, P.Eng.
CTO, Front Porch Digital Inc.

ABSTRACT

In the fast-paced and demanding newsroom environment, videotape-based workflows are no longer an option for progressive content creators. Web delivery, IPTV, podcasting and other non-traditional delivery mechanisms are voracious consumers of content that simply cannot be satisfied using traditional management methods. Effective Content Storage Management (CSM) systems and digital storage infrastructures have evolved to provide near immediate and reliable access to limitless content repositories and allow content creators to draw on an almost infinite amount of material. Married with advanced media content management systems, producers and journalists no longer need to concern themselves with the formats or location of content and can focus on their competitive challenges of creating compelling news stories.

NEW CHALLENGES IN BROADCAST

In this ever-changing broadcast landscape, the adage "Content is King" could not be more true. Content owners now have the ability to directly reach consumers, essentially bypassing traditional carriage paths. Services like Apple iTunes and YouTube will only get stronger as bandwidth to the home allows more immediate and high quality content to be consumed.

"In the near future an increasing number of viewers will seek the flexibility offered by online video and abandon conventional broadcast television, with its fixed program slots and commercials that interrupt shows", said Microsoft chairman Bill Gates at the recent World Economic Forum in Davos, Switzerland

Consumers are less tolerant of the fixed-time broadcast schedules of television and are using personal video recorders and media PCs to make television fit their lifestyles rather than the other way around. Younger viewers are watching what they want, when they want on services such as YouTube rather than sitting in front of their television sets. Times are changing.

Viewers are asking new questions: Why do I have to wait until 6:00pm to watch my local newscast? Why can't I customize the types of stories I see? Why can't I watch my custom newscast on my mobile phone on my way to work in the morning?

With the advent of advanced CSM systems in the newsroom, immediate access to infinite content as well as alternative delivery mechanisms and revenue-generating models can be fully realized.

TRANSITION FROM VIDEOTAPE TO FILES

In traditional newsrooms videotape is still the most prevalent acquisition and storage medium. Reporters and journalists rely on this linear medium for camera acquisition in the field because of its pervasive nature and familiarity. Typically, videotape also becomes the long-term storage medium for finished stories as well as for the field tapes containing all of the raw material. This means that videotape feeds the news generation process, and also generates more videotape at the output of the process with the finished stories and raw footage.

Sitting in the midst of this videotape input and output process typically sits a nonlinear editing or digital newsroom system because of the obvious benefits these "file-based" tools offer to production including creative flexibility and quick story turn around. Still, despite this file-based stage in the news creation process, the finished pieces are often recorded back to videotape for longer term storage and potential repurposing (Figure 1).

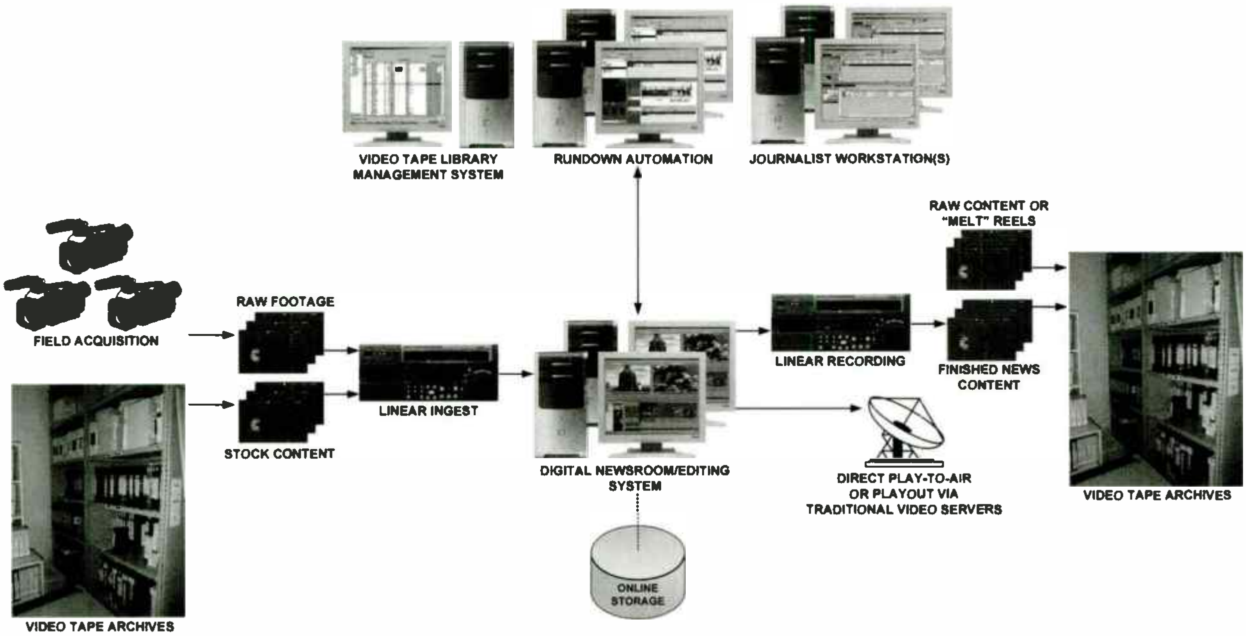


Figure 1 - The Endless Cycle of Videotape

Raw footage arriving from the field must be digitized or ingested into these file-based editing systems, but most allow editing to commence almost immediately once encoding begins. It is important to note that this ingest process cannot occur any faster than realtime, meaning that 10 minutes of raw footage will take at least 10 minutes to digitize fully into the system. If shot lists were generated in the field, efficiencies could be gained during this ingest process as only the best acquired shots need be digitized. During the ingest process, a technical operator must pay specific attention to the quality of the encoded content to ensure the editors have the best possible quality content with which to work. This is a very important step in the overall file-based workflow.

Because of increasing disk-storage density and lower associated costs, file-based editing systems typically include quite a large amount of storage and are often networked in a shared storage topology allowing many edit stations access to the same content simultaneously. As content ages in a newsroom, it typically diminishes in importance and can be removed from the storage arrays to free capacity for newly arriving and consequently more relevant material. This removal process can be facilitated by the generation of "melt reels" that consolidate the best extractions from the digitized raw footage and finished pieces back to videotape for later access. Ideally, some metadata is also entered for this videotape-based content into a searchable database for later accessibility.

Although the cost of disk-based storage has dropped significantly year after year, capacities required for high quality SD production, and now HD, are simply out of reach for most newsroom operations.

Once a news story has been produced, some systems allow play-to-air directly from the nonlinear editing/newsroom system, but others require rendering and possibly even output to videotape beforehand. These generated videotapes can then be used to play directly to air under control of a newsroom automation system, or ingested into another nonlinear device such as a broadcast video server for this purpose.

These resultant stories, perhaps in many forms and durations can then be placed in the videotape archives along with the raw footage used to comprise the items for later access and potential reuse. Of course, the value of this content is fully dependent on the "metadata" or descriptive information associated with the videotapes and the ability for the journalists to access this content again in the future. Because these finished stories have to be fully rendered as they are output to videotape, they cannot be later disassembled for easy repurposing once committed to the archives.

In addition to the obvious pitfalls in this hybrid linear and nonlinear newsroom workflow, it is worth reiterating the key "no-faster-than-realtime" limitation of videotape. Each of the individual steps described that involve videotape are, in fact, typically several times slower than realtime. The best way to quantify this statement is by example.

Suppose a journalist is looking to include some archive footage of significant historical regional weather events in a news story on the topic of global warming. The first step would be for her to search the videotape catalog for relevant keywords and the associated videotapes. This search might generate pointers to 50 archived videotapes and hopefully include some measure of relevancy to allow her to filter the selection down to say five of the tapes. She must then visit the tape archive shelves and hopefully find each of these desired videotapes. Assuming the newsroom policy allows for original videotapes to be borrowed from their archives without requiring a dub to be made, she can then take these to a screening room for review. By shuttling through this linear media, the journalist generates a shot list with timecode marks identifying the two-to-three minutes of raw content, which will be cut down and included in her piece. Once the relevant footage is located using this manually intensive linear screening process, it must then be re-ingested into the edit platform and quality checked as it gets rolled into the current story production. The news story can then be produced and then likely output to videotape for play-to-air, again adding another time-consuming step to this creation process.

At this point it is painfully obvious that the mere collection of a couple of minutes of videotape footage to generate a new piece of content can take several manually intensive hours outside of the actual editing process. Of course, there is some risk that the "best" shots are overlooked during this process as the time crunch limits the depth of the manual screening process.

In addition to all of these creative and time-based pitfalls, there is another more severe ramification of these newsroom implementations. The more recent issues surrounding disaster recovery and content protection introduce significant challenges when looking at a hard-drive-based solution and legacy videotape-based archives. If we support the concept that content is king, why does almost every newsroom in the world maintain only a single copy of all of their valuable content in basement videotape archive rooms?

Despite its effectiveness in the past, this model simply cannot support the dynamic expectations of the content consumer of today, multiple delivery formats, and hunger for content in faster-than-realtime fashion.

QUICK TURN AND FILE-BASED ACQUISITION

There are several issues that rise to the surface when analyzing the efficiency of these evolving file-based systems. The first is that field acquisition is still predominantly videotape-based although new non linear acquisition devices from companies such as Thomson and Sony are beginning to become more and more prevalent.

These nonlinear acquisition devices rely on optical, hard drive, or flash media to avoid the realtime ingest process typical in traditional videotape-based newsrooms. Shot lists automatically produced during the actual recording allow editors to begin work on the content almost immediately. In most cases, editing can be performed directly on the acquisition media and does not even require a transfer or ingest into the editing platform storage arrays. Also, because these are file-based assets and no ingest is required, the time intensive quality check can also be skipped or at least minimized. By alleviating the initial linear step in allowing immediate access to newly acquired content, the benefits to quick turn-around content is obvious.

By introducing file-based acquisition to the front end of the newsroom process, amazing efficiencies can be gained. However, it still does not address the issues relating to long term-storage, access, and protection of this valuable station-owned content. This is where CSM can help.

CONTENT STORAGE MANAGEMENT IN THE NEWSROOM

CSM systems introduce other "tiers" of physical storage into the active newsroom workflow. The key to these systems is that they do not interfere with the content creation workflow, but simply act as a near infinite background content repository. That is, editors and journalists continue to create new content but now have limitless access to all archived content. No longer are vague queries passed to the videotape archivist for fulfillment, feeding a manually intensive screen process to find the right footage. Instead, commands are sent directly from the newsroom workstation to the CSM system, which will automatically pass the audio and video content back to the newsroom system for repurposing.

A true file-based newsroom model comprises four necessary components (Figure 2): the nonlinear *newsroom editing system*; the *content storage management system*; the *physical storage infrastructure* including high-speed networks and any mix of disk-storage arrays, data-tape or optical libraries; and the higher level *newsroom production and play-to-air system.*

These components work in tandem to facilitate the end-to-end, file-based digital content workflow in the newsroom. At the heart of this concept is the "essence" storage, handling, and management infrastructure, or simply, the CSM system.

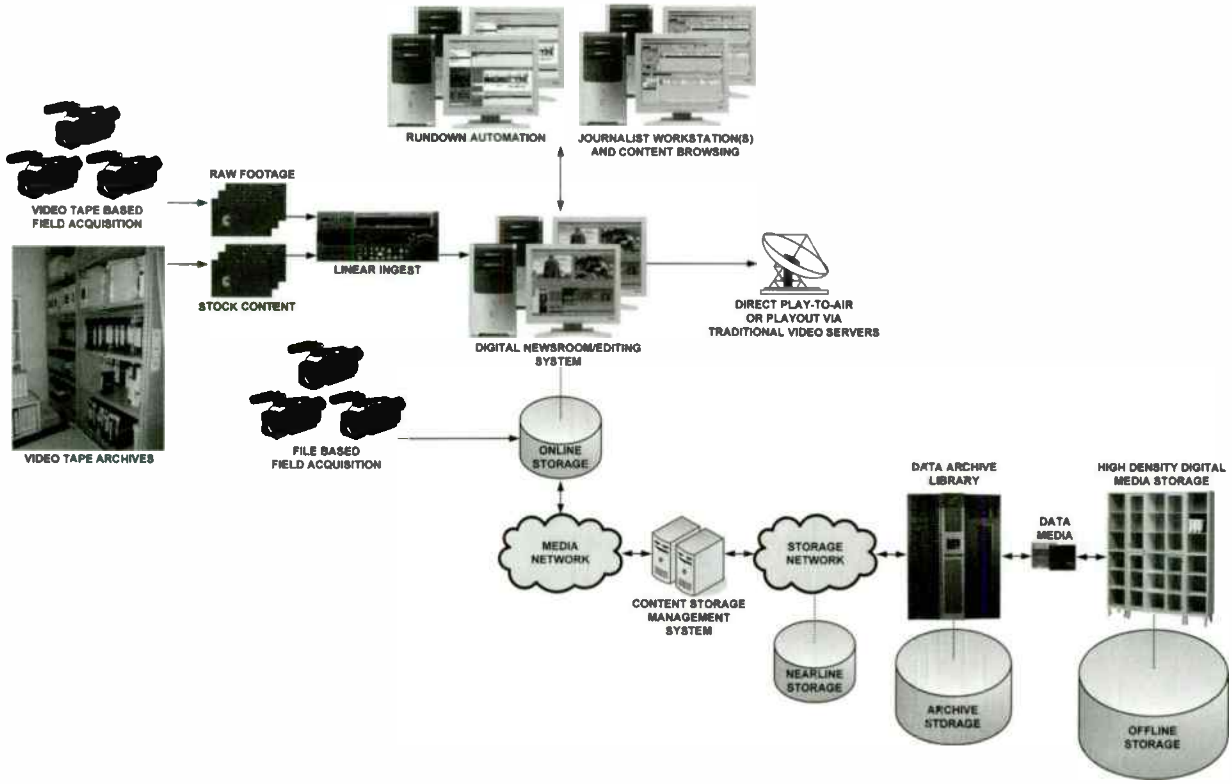


Figure 2 – File-Based Newsroom Workflow

Most digital storage infrastructures (DSI) feature four distinct storage tiers: *online, nearline, archive,* and *offline*. The major benefit of this tiered storage model is the cost of storage decreases significantly as content is migrated from online to nearline to archive and finally to offline storage (Table 1). Each tier of storage also provides certain workflow advantages and disadvantages, which must also be taken into careful consideration during the planning and architecture stage.

	Storage Tier	Description	Cost and Access
Newsroom System	ONLINE	High-speed and vendor-specific spinning disk connected to the digital newsroom system as "active" storage. Allows for immediate content access, editing, and playout.	**Cost:** $$$$$ **Access:** Immediate
Content Storage Management System	NEARLINE	Less expensive generic "IT" spinning disk technology controlled by the CSM System. Requires a disk-to-disk-based fast "copy" operation to online storage for content accessibility.	**Cost:** $$$ **Access:** Very Fast
	ARCHIVE	Large and expandable robotic data-tape or optical storage library controlled by the CSM system providing near infinite storage and bandwidth capabilities.	**Cost:** $$ **Access:** Fast
	OFFLINE	File-based media from the robotic storage library can be ejected and stored on shelves (or duplicated and moved offsite for content protection) while still actively tracked by the CSM system.	**Cost:** Very Low **Access:** Slow

Table 1 - Storage Tier Summary

The CSM system is responsible for the effective and transparent management of these various storage tiers and seamless interface into the newsroom environment.

As the accelerating adoption of digital, file-based infrastructures continues in the broadcast world, the need for storage capacity is also growing at a phenomenal rate. Digital storage "silos" independently serving post production, on-air playout, graphics, and newsroom systems are becoming common to fit these unique workflows, despite the obvious disadvantages. Effective unification of these distinct digital silos are necessary if a facility's storage and content management capabilities are to facilitate next generation, file-based collaborative workflows.

CSM AT THE HEART OF THE NEWSROOM

CSM systems, traditionally referred to as archive management systems, are becoming less focused on media "archiving" and more on active, file-based digital workflows, distribution, content exchange, interoperability, and collaboration. The term *archive management* no longer does these advanced systems sufficient justice.

CSM systems include direct connectivity to various broadcast devices and provide intelligent digital storage infrastructure management and abstraction. Advanced solutions such as those from Front Porch Digital can also facilitate digital content repurposing via timecode-based partial restore, in-path high- and low-bitrate content transcoding for tapeless interoperability between systems, as well as automatic site-to-site content distribution and replication for disaster recovery and business continuance.

At a high level, CSM solutions interface directly into broadcast devices – the digital newsroom system in this case – manage all additional tiers of storage and provide a unified and intelligent view of the storage infrastructure to various control system such as the newsroom production and play-to-air systems. CSM systems are less focused on expensive storage tier capacity maximization, and more focused on the complex collaborative and accessibility requirements of a newsroom or overall broadcast facility.

In general, CSM systems are server-based software solutions that reside between the so-called "media network," which connects various broadcast devices and the "storage network," which connects the nearline and archive storage tiers (Figure 3). CSM systems take responsibility for the broadcast asset directly from online storage through nearline, archive, offline, and back. This eliminates the need to have an additional proprietary control layer act as an intermediary between online and nearline storage to facilitate storage management.

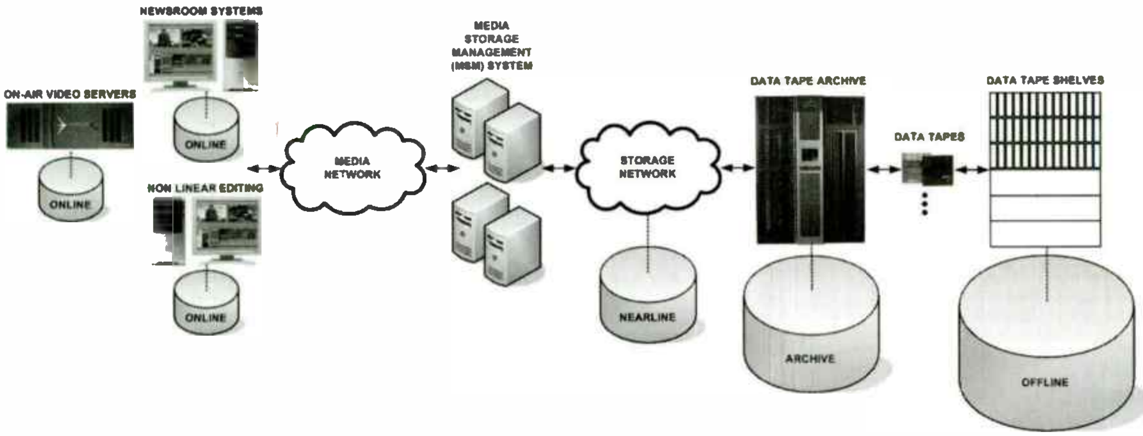


Figure 3 - Content Storage Management Overview

In an advanced file-based newsroom, a balance between the storage tiers managed by the CSM system can gain efficiencies and cost benefits unrecognized in legacy newsroom systems. For example:

- The expensive nonlinear, **online** shared storage may have sufficient capacity for three to seven days worth of immediately available raw footage and the various news stories/serials in production

- The less expensive **nearline** spinning disk arrays may hold 30 days worth of quickly accessible content

- The high capacity **archive** library may store years of content, which can be accessed in minutes

- The **offline** data media shelves can hold an infinite amount of content, which can be accessed as quickly as the librarian can insert the required piece of data media into the archive library.

The important factors to note about CSM-based newsroom infrastructures are:

Exact Copies
File-based content is always exactly the same as when it was first acquired or ingested and does not require any additional human quality checks. This effectively eliminates any human intervention once the content has been checked during the initial ingest process.

Faster Than Realtime
File-based content can be moved many times faster than realtime, eliminating long lead times on accessing content stored in nearline or archive storage. Relevant raw footage can be restored for use in new productions many times faster than realtime facilitating very fast production turnaround.

Integrated In-Path Transcoding
File-based content can be transcoded or reformatted by the CSM system as it moves between devices allowing reuse in systems other than those in which it was originally encoded. This allows finished news stories to be digitally moved to play-to-air servers skipping over any videotape-based steps completely.

Desktop Proxy Browsing
File-based content can be transcoded automatically by the CSM system to generate frame accurate, low bitrate proxies for desktop browsing using the newsroom production system or other simple Web-based Media Asset Management (MAM) systems. Journalists can now find relevant content, review it, and produce shotlists directly from their desktop using these proxies and have the CSM system deliver only these best shots directly to the nonlinear editors. Again, this operation can occur many times faster than realtime, which is key in a progressive newsroom.

Inexpensive Content Replication
Content can be replicated very quickly producing extremely inexpensive content protection with no human intervention required. This replicated media can be easily "externalized" or taken out of the data library and moved to another protected offsite facility for disaster recovery. The station-owned content is no longer at risk of destruction as exact copies can be immediately and inexpensively generated.

Timecode-Based Partial Restore

Content can be fully restored or partially restored by the CSM referencing only timecode mark-in and mark-out points allowing only the most relevant footage to be restored to online storage for repurposing. This allows the journalist to skip the videotape screening process and can drive this directly from their desktop computer leveraging the proxies to identify only the content they need.

High Density Storage Media
Data media is extremely high density and can significantly reduce the real estate required by traditional videotape shelves. A single piece of high density digital storage media, such as LTO3, can store more content than 20 analog videotapes!

Content Repurposing
Transcoding can also facilitate automatic delivery to alternate distribution systems such as VOD, IPTV, 3G mobile, etc., all drawn from the original material. Revenue generation can be driven by the content rather than simply relying on commercial revenue during the newscast. Alternate delivery mechanisms such as mobile phone and IPTV delivery will put distribution control into the hands of the content owners.

Content Lifecycle Control
CSM systems can automatically handle file-based content using advanced, context-sensitive policies to define replication factors, migration between storage tiers, as well as automatic transcoding for delivery to other distribution channels (mobile phone, IPTV, VOD, etc). Finished news stories can be automatically replicated within the CSM system, transcoded to 3G format, and passed to a mobile distribution service for consumer access. Raw footage can be automatically maintained on nearline spinning disk arrays for 30 days and then migrated to the data library.

Resource Prioritization
Advanced CSM systems can automatically manage storage resources and bandwidth to ensure the creative process is not slowed by technology. Rather, facilitating access to a limitless pool of content the CSM system will allow the creation of more inspired content.

Unflattened Content Storage
Content created on nonlinear editing platforms can be stored in the CSM system in an unflattened fashion allowing for components such as keys and voiceovers to be changed in the future without having to recreate the entire piece.

As an additional benefit, CSM systems can provide a truly unified asset storage infrastructure, spanning any mix of newsroom systems, broadcast devices including on-air video servers, post-production editing platforms, etc. In addition to the obvious cost and workflow benefits of this single unified storage infrastructure, advanced CSM systems can also offer in-path content transcoding allowing tapeless interoperability between these distinct silos. Add timecode-based partial file restoration to this mix and the true benefits of a collaborative, file-based broadcast workflow become obvious (Figure 4).

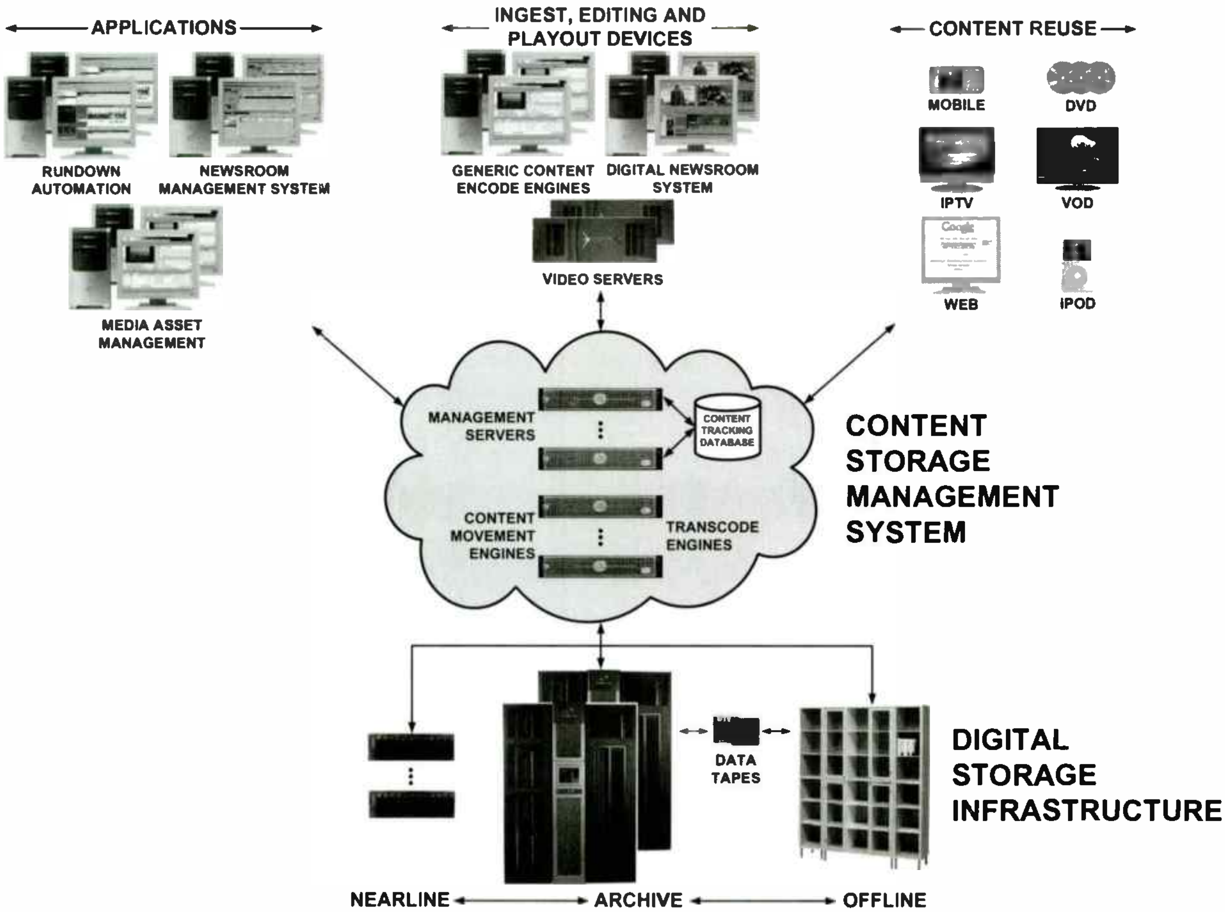


Figure 4 - CSM at the Heart of the File-Based Newsroom

WHAT THIS ALL MEANS...

In this very brief overview, we identified some key advantages to CSM solutions in newsroom environments. This is by no means an exhaustive list of the factors that should play a part in this complex decision making process. It is important to partner with a solutions provider that not only fully understands the business, but who focuses exclusively on the complex, file-based storage management needs of global broadcasters.

Effective unification of the distinct digital "silos" are necessary if a facility is to benefit from next-generation, file-based collaborative workflows. Implementation of an advanced CSM system, built specifically to meet broadcasters' needs, gives users the power and flexibility required to get the most out of their assets and to achieve the goal of a unified media storage infrastructure.

ABOUT THE AUTHOR

Brian Campanotti is Chief Technology Officer at Front Porch Digital Inc. responsible for leading industry advancement in the area of global media Content Storage Management (CSM), file-based workflows, and intelligent content management. DIVArchive is Front Porch Digital's flagship advanced CSM product offering. For more information please visit our web site at www.fpdigital.com

SYSTEMS INTEGRATION – THE FEDERAL ENTERPRISE ARCHITECTURE APPROACH TO MEDIA ASSET MANAGEMENT SYSTEM DESIGN

SAM STEVENS
Radio Free Asia
Washington, DC
2007-01-19

ABSTRACT

While rapid technological advancement brings many benefits, it also presents many challenges. Today the broadcaster is experiencing the same information technology (IT) issues that face other large "non-broadcast" organizations, such as the growing number of segmented systems and distinct networks, databases, applications that defy unification, and timely and appropriate access to the tens of hundreds of terabytes of information and content generated by both the organization and the audience participants with various security access policies in place to protect network and data.

Content Management System (CMS), Media Asset Management (MAM) and Digital Asset Management (DAM) systems all share a ferocious appetite for storage, bandwidth and the user's expectation of having instant access to information from anywhere anytime. We will take a vendor-neutral look at options available to both cutting edge and not-so-cutting edge operations to aid in design and management of content utilization. We will examine one tool to be considered for use in system design, Federal Enterprise Architecture (FEA). Whether procuring or designing, FEA provides a blueprint to "engineer" Digital Asset Management, Media Asset Management or Content Management Systems and the subsystems that support them - storage, networking and archival workflows. In conclusion, we will investigate the "use and reuse" of systems, hardware and workflows.

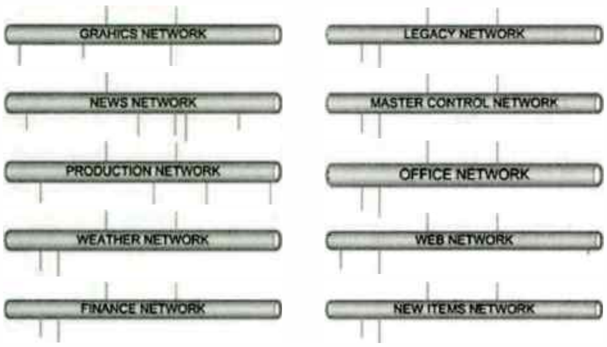
THE FACTS

Broadcasters are experiencing the same information technology (IT) issues that face other large "non-broadcast" organizations:

- Growing numbers of segmented systems and distinct networks
- Adjusting to new business models that require cutting edge technologies (such as streaming media to do more)
- Network demands from desktop multimedia production and editing
- Extending network capabilities to enable expansion of services within the global community
- Securing and protecting data and databases
- Unifying applications that defy unification
- Multimedia-based applications that must be integrated into the current PC workstations
- Mitigating legacy systems and software applications that are not compliant with current security policies and updates
- Consolidating and managing storage systems, managing backup and archival data, and providing a comprehensive plan for disaster recovery
- Meeting the technological demands generated by the business and audience participants of business revenue streams from WEB hosting, POD cast, and increases in remote connectivity, such as the mobile workforce, virtual workplaces and expanded bureaus
- Providing improved detection methods that protect against intrusion and malicious attacks without compromising the ability to grant appropriate access to the tens or hundreds of terabytes of information and content. If the commonalities between Broadcast and IT sound familiar, they should. As we take on more of an IP-centric platform for systems and designs, this once blurred "line" is will become common to all. Non-broadcasters are becoming distributors of content and share a common audience with us; we need not look any further than the commercials we broadcast

THE APPROACHES

The need for systems that can manage our volume of content and media is greater than before.



Segmented Network

Each segmented network supports a specific system and its own distinct storage. Sharing may be "routed across the network but they remain a "stovepipe" system.

Business and Performance-Driven Approach

The Federal Enterprise Architecture Project Management Office (FEAPMO) reference model is a set of guides to locate reusable services, not an explicit model for direct implementation. Each organization needs to establish an explicit reference-model based on the FEAPMO model to implement its enterprise architecture. In discovery a variation of the FEA may be used to support the business model for broadcasters. The similarities are there. If we dissect the Federal Government models we will see direct correlation to what broadcasters do. Before undertaking a project of this magnitude the first step is justifying the need, then planning and defining workflows. This screening FEAPMO process brings to focus key items such as:

Part I: Screening questions and summary of spending for project stages;

- Project Description
- Justification Performance Goals and Measures
- Program Management
- Alternative Analysis
- Risk Management
- Acquisition Strategy
- Project and Funding Plan

Part II: Additional Criteria for IT

- Enterprise Architecture
- Security and Privacy

Part I validates the need and value-added, and should be considered the "front" end of the project.

Part II focuses on the IT/Broadcast enterprise infrastructure and security.

Before moving on to the "blueprint" of what our system should include based on the outcome from Part I above, let's explore in detail two areas that hit close to "home" for us as broadcasters, the Service Component Reference Model (SRM) and the Technical Reference Model (TRM).

SERVICE COMPONENT REFERENCE MODEL (SRM)

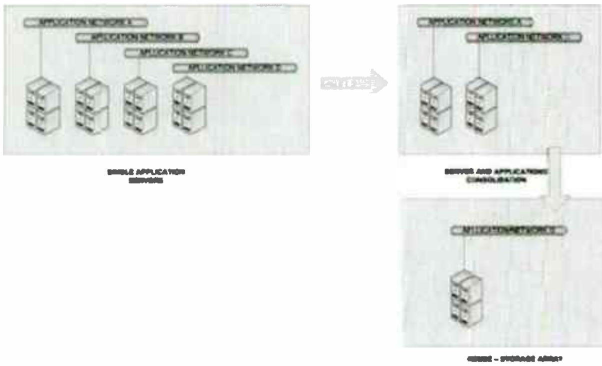
"The Service Component Reference Model (SRM) is a business and performance-driven functional framework that classifies Service Components with respect to how they support business and/or performance objectives."

The SRM is intended for use to support the discovery of government-wide business and application Service Components in IT investments and assets. The SRM is structured across horizontal and vertical service domains that, independent of the business functions, can provide a leverage-able foundation to support the reuse of applications, application capabilities, components, and business services. The following are examples of reuse:

- Reuse-Redeploy the IT network infrastructure to include broadcast network traffic such as low latency traffic, low resolution multimedia
- Consolidate segmented storage into a storage array solution or a build your own by repurposing servers or fixed storage into storage systems
- Consolidate similar broadcast applications onto a redundant server pair (such as news systems and graphics)
- Explore the use of server virtualization that afford reuse existing hardware

Customer Services

The Services For Citizens Business Area describes the mission and purpose of the United States government in terms of the services it provides both to and on behalf of the American citizen. It includes the delivery of citizen-focused, public, and collective goods and/or benefits as a service and/or obligation of the Federal Government to the benefit and protection of the nation's general population.



Server Hardware Reuse for Storage Solution

Process Automation Services

The Process Automation Services Domain defines the set of capabilities that support the automation of process and management activities that assist in effectively managing the business. The Process Automation Services domain represents those services and capabilities that serve to automate and facilitate the processes associated with tracking, monitoring and maintaining liaison throughout the business cycle of an organization.

Business Management Services

The Business Management Services Domain defines the set of capabilities that support the management of business functions and organizational activities that maintain continuity across the business and value-chain participants. The Domain represents those capabilities and services necessary for projects, programs and planning within a business operation to be successfully managed.

Digital Asset Services

The Digital Asset Services Domain defines the set of capabilities that support the generation, management and distribution of intellectual capital and electronic media across the business and extended enterprise.

Business Analytical Services

The Business Analytical Services Domain defines the set of capabilities supporting the extraction, aggregation and presentation of information to facilitate decision analysis and business evaluation.

Back Office Services

Define the set of capabilities that support the management of enterprise planning and transactional-based functions.

Support Services

The Support Services Domain defines the set of cross-functional capabilities that can be leveraged independent of Service Domain objective and/or mission.

TECHNICAL REFERENCE MODEL (TRM)

The TRM is a component-driven technical framework used to categorize the standards, specifications, and technologies that support and enable the delivery of service components and capabilities.

The Technical Reference Model (TRM) provides a foundation to categorize the standards, specifications, and technologies to support the construction, delivery, and exchange of business and application components (Service Components) that may be used and leveraged in a Component-Based or Service-Oriented Architecture. The TRM unifies existing Agency TRMs and E-Gov guidance by providing a foundation to advance the re-use of technology and component services from a government-wide perspective.

Service Access & Delivery

Service Access and Delivery refers to the collection standard and specifications to support external access, exchange, and delivery of Service Components or capabilities. This area also includes the Legislative and Regulator requirements governing the access and usage of the specific Service Component.

Service Platforms & Infrastructure

The Service Platform and Infrastructure Area defines the collection of platforms, hardware and infrastructure specifications that enable Component-Based Architectures and Service Component reuse.

Component Framework

The Component Framework Area defines the underlying foundation and technical elements by which Service Components are built, integrated and deployed across Component-Based and Distributed Architectures.

The Component Framework consist of the design of application or system software that incorporates interfaces for interacting with other programs and for future flexibility and expandability. This includes, but is not limited to, modules that are designed to interoperate with each other at runtime. Components can be large or small, written by different programmers using different development environments, and may be platform-independent. Components can be executed on stand-alone machines, a LAN, Intranet or on the Internet.

Service Interface & Integration

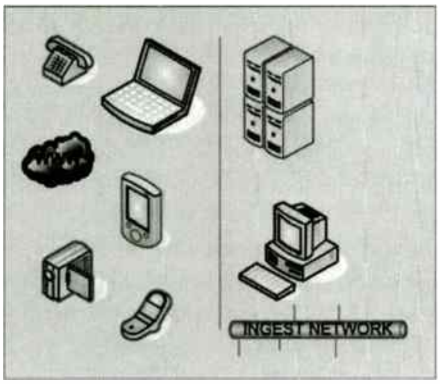
The Service Interface and Integration Area defines the discovery, interaction and communication technologies joining distinct systems and information providers. Component-based architectures leverage and incorporate Service Interface and Integration specifications to provide interoperability and scalability.

Service Access & Delivery

Access Channels
Delivery Channels
Service Requirements
Service Transport

Note the TRM service access and delivery component is describing an Asset Management System.

Access and Delivery of Managed Content and Media, sharing and reuse of workflows.



ACCESS & DELIVERY

Access Channels

Access Channels define the interface between an application and its users, whether it is a browser, personal digital assistant or other medium.

Collaboration Communications
Define the forms of electronic exchange of messages, documents, or other information. Electronic communication provides efficiency through expedited time-of-delivery.

Other Electronic Channels
Define the other various mediums of information exchange and interface between a user and an application.

Web Browser
Define the program that serves as your front end to the World Wide Web on the Internet. In order to view a site, you type its address (URL) into the browser's location field.

Wireless / PDA
Technology that uses transmission via the airwaves. A Personal Digital Assistant (PDA) is a handheld computer that serves as an organizer for personal information. It generally includes at least a name and address database, to-do list and note taker.

Delivery Channels

Delivery channels define the level of access to applications and systems based upon the type of network used to deliver them.
Extranet
An extranet is a private network that uses the Internet protocol and the public telecommunication system to securely share part of a business's information or operations with suppliers, vendors, partners, customers, or other businesses. An extranet can be viewed as part of a company's intranet that is extended to users outside the company.

Internet
The internet is a worldwide system of computer networks in which users at any one computer can, if they have permission, get information from any other computer.

Intranet
An intranet is a private network that is contained within an enterprise. It may consist of many inter-linked local area networks and is used to share company information and resources among employees.

Peer to Peer (P2P)
Peer to Peer is a class of applications that operate outside the DNS system and have significant or total autonomy from central servers that take advantage of resources available on the Internet.

Virtual Private Network (VPN)
A Private Data Network that makes use of the public telecommunication infrastructure, maintaining privacy through the use of a tunneling protocol and security procedures.

Service Requirements

Service Requirements define the necessary aspects of an application, system or service to include legislative, performance and hosting requirements.

Authentication / Single Sign-on (SSO)
Refers a method that provides users with the ability to log-in one time, getting authenticated access to all their applications and resources.

Hosting
Hosting refers to the service provider who manages and provides availability to a web site or application, often bound to a Service Level Agreement (SLA). The Hosting entity generally maintains a server farm with network support, power backup, fault tolerance, load-balancing, and storage backup.

Legislative/Compliance
Legislative/Compliance defines the pre-requisites that an application, system or service must have mandated by congress or governing bodies.

Service Transport

Service Transport defines the end-to-end management of the communications session to include the access and delivery protocols.

Service Transport
Protocols that define the format and structure of data and information that is either accessed from a directory or exchanged through communications.

Service Platform and Infrastructure

The Service Platform and Infrastructure Area defines the collection of platforms, hardware and infrastructure specifications that enable Component-Based Architectures and Service Component re-use.

Database/Storage
Delivery Servers
Hardware/Infrastructure
Software Engineering
Supporting Platforms

Database / Storage

Database/Storage refers to a collection of programs that enables storage, modification, and extraction of information from a database, and various techniques and devices for storing large amounts of data.

Database
Database refers to a collection of information organized in such a way that a computer program can quickly select desired pieces of data. A database management system (DBMS) is a software application providing management, administration, performance, and analysis tools for databases.

Storage
Storage devices are designed to provide shared storage access across a network. These devices provide extended storage capabilities to the network with reduced costs compared to traditional file servers.

Delivery Servers

Delivery Servers are front-end platforms that provide information to a requesting application. It includes the hardware, operating system, server software, and networking protocols.

Application Servers
In a three-tier environment, a separate computer (application server) performs the business logic, although some part may still be handled by the user's machine. After the Web exploded in the mid 1990s, application servers became Web based.

Media Servers
Provide optimized management of media-based files such as audio and video streams and digital images.

Portal Servers
Portals represent focus points for interaction, providing integration and single-source corporate information.

Web Servers
A computer that provides World Wide Web services on the Internet. It includes the hardware, operating system, Web server software, TCP/IP protocols and the Web site content (Web pages). If the Web server is used internally and not by the public, it may be known as an "intranet server."

Hardware/Infrastructure

Hardware Infrastructure defines the physical devices, facilities and standards that provide the computing and networking within and between enterprises.

Embedded Technology Devices
Embedded Technology Devices refers to the various devices and parts that make up a Server or Computer as well as devices that perform specific functionality outside of a Server or Computer.

Local Area Network (LAN)
A local area network is a network that interconnects devices over a geographically small area, typically in one building or a part of a building. The most popular LAN type is Ethernet. LANs allow the sharing of resources and the exchange of both video and data.

Network Devices/Standards
A group of stations (computers, telephones, or other devices) connected by communications facilities for exchanging information. Connection can be permanent, via cable, or temporary, through telephone or other communications links. The transmission medium can be physical (i.e., fiber optic cable) or wireless (i.e., satellite).

Peripherals
Computer devices that are not part of the essential computer (i.e. the memory and microprocessor). Peripheral devices can be external and internal.

Servers / Computers
This refers to the various types of programmable machines which are capable of responding to sets of instructions and executing programs.

Video Conferencing
Communication across long distances with video and audio contact that may also include graphics and data exchange. Digital video transmission systems typically consist of camera, codec (coder-decoder), network access equipment, network, and audio system.

Wide Area Network (WAN)
A data network typically extending a LAN outside a building or beyond a campus. Typically created by using bridges or routers to connect geographically separated LANs. WANs include commercial or educational dial-up networks.

Software Engineering

Software engineering covers not only the technical aspects of building software systems, but also management issues, such as testing, modeling and versioning.

Integrated Development Environment (IDE)
This consists of the hardware, software and supporting services that facilitate the development of software applications and systems.

Modeling
The process of representing entities, data, business logic, and capabilities for aiding in software engineering.

Software Configuration Management
Applicable to all aspects of software development from design to delivery specifically focused on the control of all work products and artifacts generated during the development process. Several solutions on the market provide the integration of the software configuration management functions.
Test Management
Test Management is the consolidation of all testing activities and results. Test Management activities include test planning, designing (test cases), execution, reporting, code coverage, and heuristic and harness development.

Supporting Platforms

Supporting platforms are hardware or software architectures. The term originally dealt with only hardware, and it is still used to refer to a CPU model or computer family.

Platform Dependent
Consists of the programming languages and methods for developing software on a specific operating system or platform.

Platform Independent
Defines the programming languages that are able to execute and run on any platform or operating system.

Wireless / Mobile
Radio transmission via the airwaves. Various communications techniques are used to provide wireless transmission including infrared line of sight, cellular, microwave, satellite, packet radio and spread spectrum.

Component Framework

The Component Framework Area defines the underlying foundation and technical elements by which Service Components are built, integrated and deployed across Component-Based and Distributed Architectures. The Component Framework consists of the design of application or system software that incorporates interfaces for interacting with other programs and for future flexibility and expandability.

This includes, but is not limited to, modules that are designed to interoperate with each other at runtime. Components can be large or small, written by different programmers using different development environments and may be platform independent. Components can be executed on stand-alone machines, a LAN, Intranet or on the Internet.

Business Logic
Data Interchange
Data Management
Presentation/Interface
Security

Business Logic

Defines the software, protocol or method in which business rules are enforced within applications.

Platform Dependent
Consists of the programming languages and methods for developing software on a specific operating system or platform.

Platform Independent
Consists of all software languages that are able to execute and run on any type of operating system or platform.

Data Interchange

Data Interchange defines the methods in which data is transferred and represented in and between software applications.

Data Exchange
Data Exchange is concerned with the sending of data over a communications network and the definition of data communicated from one application to another. Data Exchange provides the communications common denominator between disparate systems.

Data Management

The management of all data/information in an organization. It includes data administration, the standards for defining data, and the way in which people perceive and use it.

Database Connectivity
Defines the protocol or method in which an application connects to a data store or database.

Reporting and Analysis
Consists of the tools, languages and protocols used to extract data from a data store and process it into useful information.

Presentation/Interface

This defines the connection between the user and the software, consisting of the presentation that is physically represented on the screen.

Content Rendering
This defines the software and protocols used for transforming data for presentation in a graphical user interface (GUI).

Dynamic/Server-Side Display
This consists of the software that is used to create GUIs with the ability to change while the program is running.

Static Display
Static Display consists of the software protocols that are used to create a pre-defined, unchanging graphical interface between the user and the software.

Wireless/Mobile/Voice
Consists of the software and protocols used for wireless and voice-enabled presentation devices.

Security

Security defines the methods of protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide integrity, confidentiality and availability. Biometrics, two-factor identification, encryption, and other evolving areas of technology.

Certificates / Digital Signature
Software used by a certification authority (CA) to issue digital certificates and secure access to information.

Supporting Security Services
These consist of the different protocols and components to be used in addition to certificates and digital signatures.

Service Interface and Integration

The Service Interface and Integration Area define the discovery, interaction and communication technologies joining disparate systems and information providers. Component-based architectures leverage and incorporate Service Interface and Integration specifications to provide interoperability and scalability.

Integration
Interface
Interoperability

Integration

Integration defines the software services enabling elements of distributed business applications to interoperate. These elements can share function, content, and communications across heterogeneous computing environments. In particular, service integration offers a set of architecture services such as platform and service location transparency, transaction management, basic messaging between two points, and guaranteed message delivery.

Enterprise Application Integration (EAI)
Refers to the processes and tools specializing in updating and consolidating applications and data within an enterprise. EAI focuses on leveraging existing legacy applications and data sources so that enterprises can add and migrate to current technologies.

Middleware
Middleware increases the flexibility, interoperability, and portability of existing infrastructure by linking or "gluing" two otherwise separate applications.

Interface

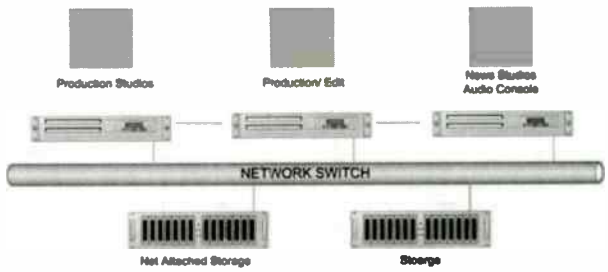
Interface defines the capabilities of communicating, transporting and exchanging information through a common dialog or method. Delivery Channels provide the information to reach the intended destination, whereas Interfaces allow the interaction to occur based on a predetermined framework.

Service Description/Interface
Service Description/Interface defines the method for publishing the way in which web services or applications can be used.

Service Discovery
Defines the method in which applications, systems or web services are registered and discovered.

Interoperability

Interoperability defines the capabilities of discovering and sharing data and services across distinct systems and vendors.



Shared resources through discovery of systems affords opportunity to use and reuse such as hardware, data and applications.

Data Format/Classification
Data Format/Classification defines the structure of a file. There are hundreds of formats, and every application has many different variations (database, word processing, graphics, executable programs, etc.). Each format defines its own layout of the data. The file format for text is the simplest.

Data Transformation
Data Transformation consists of the protocols and languages that change the presentation of data within a GUI or application.

Data Types/Validation
Data Types/Validation refers to specifications used in identifying and affirming common structures and processing rules. This technique is referenced and abstracted from the content document or source data.

CONCLUSION

Enterprise Architecture (EA) describes a comprehensive framework to identify opportunities to reuse and re-deploy IT assets through discovery.

For broadcasters it offers an improved effectiveness of technology spending by re-alignment of the information technology and broadcast technology in support of a single business initiative through use and reuse.

References

Federal Enterprise Architecture Program Management Office http://www.whitehouse.gov/omb/

Special thanks to Radio Free Asia's CTO David Baden, and Management Systems Designers, Inc., Vienna, Virginia for establishing in 2002 the Digital Media Group, a diverse group of broadcast technologists and engineers (Phillip Adams, Marty Martin, Lou Brown and me).

MXF for Television

Tuesday, April 17, 2007
1:00 PM – 5:00 PM

Chairperson: Bill Napier
Bahakel Communications, Charlotte, NC

Operational Patterns...the MXF Flavors?
Ernesto Santos, MOG Solutions, Maia, Portugal

***Managing Program Inventories Using MXF**
Oliver Morgan, Metaglue Corporation, Lexington, MA

MXF Based Workflows – A Practical Implementation
Stephen Fish, Turner Broadcasting Systems Europe Ltd, London, United Kingdom

***Business Efficiency Using Managed MXF, IT and Web Services**
Bruce Devlin, Snell & Wilcox, Liss, Hants, UK

***Enabling MXF Workflows and Flexible MXF File Exchange by Enhancing Data Tape Archives with MXF Processing Capabilities**
Mark Ostlund, Quantum Corp, Pleasanton, CA, US

Ensuring Media Exchange Format (MXF) Assets' Availability
Bill Gage, EMC Corporation, Hopkinton, MA,

Achieving MXF Interoperability
Todd Brunhoff, Omneon, Beaverton, OR

New Tools for MXF-Analysis
Matthias Schnöll and Prof. Dr. Rolf Hedtke, University Wiesbaden Television Technology and Electronic Media, Wiesbaden, Germany

*Paper not available at the time of publication

Operational Patterns… the MXF Flavors?

ERNESTO SANTOS

MOG Solutions, Portugal

ABSTRACT

MXF, the Material eXchange Format from SMPTE, is today a huge success. The market was eager for a standard file format that would serve as a basis for the new file based production paradigm, and MXF delivered precisely that.

As a result, the market is being flooded with MXF products, each using MXF in the configuration that suits best that product's application. For this reason, questions such as "which MXF flavor is this?" or "whose MXF is this?" are becoming quite common. And whenever these questions come up, so does the expression "MXF Operational Pattern".

This document discusses MXF Operational Patterns and other MXF technicalities, that relate to the way the tools in the MXF standard can be configured. These configurations usually end up interpreted by newcomers as MXF flavors. The document points out strengths and weaknesses of these flavors, and provides guidance on the building of MXF systems, as more complex flavors (with higher level MXF Operational Patterns) start being used.

The paper further explores the integration of different MXF flavors into the workflow, therefore defying the sometimes established notion that you must choose one and only one MXF Operational Pattern to work with.

SETTING THE SCENE

Every time a new technology arrives in the market, it takes its time to settle down. The deeper the impact on the everyday operations of the potential users, or their strategies for the future, or both, the more work needs to go into the way that technology is presented, and the way products and services are placed in the market.

This means that the detailed technicalities need to be refined and knowledge needs to travel from the people developing the systems, to the people marketing solutions and, of course, to the potential users of those solutions.

So what is the impact of MXF on everyday operations and strategies for the future? It could hardly be any deeper. The MXF file format enables the carriage of both the audiovisual material and related information, usually referred to as Metadata.

Metadata ranges from structural information of the material such as compression settings, to geo-localization information, to general descriptive information, including transcripts of the Content. This is why MXF represents more than a file format. It is a technology that enables the gathering of crucial information, as the audiovisual material travels through the workflow. It is therefore the foundation technology driving the adoption of Information Technology in the Professional Media market.

This also means that we are witnessing the convergence of these two different technology fields (IT and Professional Media market technologies, such as Broadcast technology), that evolved mostly separately over several decades. MXF is providing the vehicle for experts of both fields to meet half way, and its success proves the surprising speed at which this convergence is taking place.

Having the technical people meeting half way and understanding each other is a fantastic first step. However, their messages need to be refined towards a higher level of abstraction so they can be understood by potential users of the technology. These users don't know, don't need to know and don't want to know the technical details. They need just the right amount of information that will allow them to make strategic decisions for their businesses.

NEW KID ON THE BLOCK

Now, if you are a placing a new product in the market, you need to take into account:

- the current world view of your customers;
- where they want to go;
- the route they will take to get there.

The current world view of a lot of users is still based on conventional technology, such as tape/film based workflows. Most of them, if not all, have already demonstrated they want to move towards a new paradigm of file based production, management and distribution. This made it quite easy for manufacturers and service providers to figure out that MXF file based technology was the way to go. This is why you see MXF support popping up all over the place in the professional media market today.

Therefore, the key question is: which route will users take to get there?

So far in this document, the professional media market is being pointed out as the market for MXF... That is quite broad. That's because MXF was thought up to apply to virtually any scenario where professional audiovisual Content is being produced, managed and distributed.

This implies that there is a huge variety of potential MXF users, with different applications and different business models. And even those users that compete within the same application area, will most likely have different strategies and operational models. After all, they have to differentiate from one another in a world that becomes more demanding each day on the multimedia products quality, cost and time-to-market.

Strategies and operational models, in the end, map down to the workflows inside of an organization. Therefore, MXF had to be thought up as a configurable technology able to adapt to different workflows. And since strategies and operational models evolve over time as new ideas come along, MXF also had to be thought up as an extensible technology.

EXTENSIBILITY

A tutorial on MXF is outside the scope of this document. The reader should refer to [1] for detailed information on MXF, [2] for a discussion on how MXF integrates with related technologies such as AAF and XML, [3] and [4] for the entry points into the MXF Standard's document suite itself.

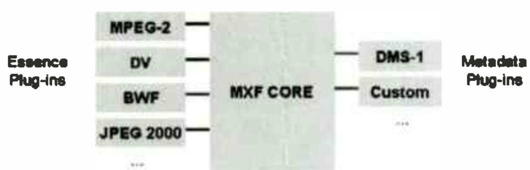


Figure 1

For the scope of this document, it suffices to say that the MXF standard defines a core set of tools which include the way the bits are encoded within the file, the way Metadata structures are logically organized, the way audiovisual material can be indexed for efficient random access… among other.

In addition, recognizing the need for extensibility, MXF is designed around a plug-in mechanism. The basic idea is that MXF is a wrapper of the audiovisual material. Therefore, if the material is compressed in MPEG-2 for example, that is how it will be represented inside the file, and that is the format that you get again once you unwrap the file. Then, for each format (MPEG, DV, JPEG2000, BWF …) the MXF document suite includes one additional document that tells you how to wrap that format efficiently in MXF. This additional document will also tell you how to represent compression settings (for compressed formats) in the MXF Metadata, and how to index that specific format.

Therefore, for each new format that comes up in the future, you just need to write down a document using this extensibility mechanism, and plug it into the MXF document suite.

The same holds true for Metadata extensibility. At the MXF core, a structural Metadata model is defined that lets you know which clips[1] you have in the file, how many tracks, their bitrates, timecode information… Then, a plug-in mechanism allows you to extend this Metadata with additional descriptions, following Metadata models that you can design yourself. It also enables you to frame accurately synchronize your Metadata with the audiovisual material you are describing. These Metadata extensions are usually called Descriptive Metadata.

SMPTE DMS-1 [4] is one example of a Metadata model plugged into MXF, using the above mechanism. In this case, it is a standard Metadata model, however you can use the same mechanisms to plug in proprietary Metadata models as well.

ADAPTABILITY

The described plug-in mechanisms provide proven support of MXF extensibility. However, all these mechanisms imply a complex MXF core. This means that decoders would have to be very complex, to support all possible variations of MXF configurations thrown at them.

[1] For the purpose of this text, clips should be considered a sequence of contiguously stored audiovisual material.

So how do you adapt MXF to application needs? The main aspects to be taking into account in MXF are the audiovisual material encoding, the Metadata model being followed, and the configuration of the whole.

Audiovisual material encoding is not something defined by MXF. And its adaptation to application needs is also not a new subject. There are several possible encodings for uncompressed material and several compression systems as well. The concept itself of compression was once new in the market as well but, although there is always evolution, the concept settled down already. There is currently extensive know-how on the criteria you need to follow to decide on which compression system, or which uncompressed representation to use. There are also several well established parameters you may tweak to get to the expected results.

Therefore, even when a new compression system arrives in this field, you can always plug it into this knowledge framework and compare with existing systems. An example is perceptual quality measurements of new codecs against the ones in the installed base.

Then we move to Metadata models... and that is a new world on itself. More than a few descriptions on a file, this is the representation of knowledge and processes, or in the end, the capturing of information flow at its highest level. A detailed discussion on Metadata models and their adaptation, although very much related to the subject of this document, deserves a document on its own and will therefore be considered out of the scope of this discussion.

So finally we need to take into account that, having all components in a system understand all possible variations of MXF configurations, would make such a system extremely complex and therefore extremely expensive.

Therefore, MXF was though up with configuration control knobs. These are parameters that enable you to "speak MXF" with more or less complexity, and enable you to signal which level of complexity you are using.

You will often come across several of those parameters in the MXF lingo: Clip Wraping, Frame Wrappping, sparse or full Index Tables, interleaving and multiplexing... these are just a few of the control knobs that you can tweak in the MXF toolset to configure it.

The most popular control knob these days is the Operational Pattern. And there is a reason for this. The Operational Pattern control knob is the highest level one. It tells you if the file has one clip or more. And if it has more, it tells you upfront what your decoder should expect in terms of the complexity of the relationship between those clips.

OPERATIONAL PATTERN

For the purpose of Operational Pattern classification, you need to regard an MXF file as a file that holds clips of audiovisual material that you want to play out in some specific way.

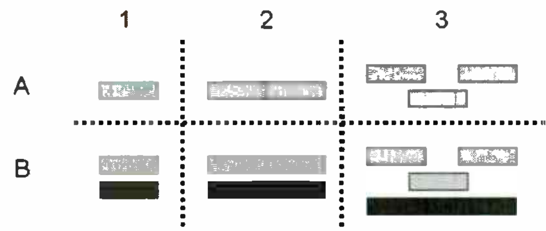


Figure 2

You may want to play-out (notice the relation to the numbering in Figure 2):

1. the full duration of one clip;
2. a sequence of full clips;
3. a sequence of selected portions of clips.

In addition, at each point in time you may want to be playing from:

a. one clip;
b. more than one clip.

In any of the cases above, a clip might mean just picture information or just sound information. However, a clip may also mean picture and sound information if both are stored together inside the file[2].

An additional level of complexity may be signaled in MXF Operational Patterns. You could see this as row 'C' that could be added in Figure 2. The row was not added because the semantics of this 'C' row follows a concept that is quite different from the one in 'A' and 'B'.

For rows 'A' and 'B' the decoder is being told the (single) way to play the clips. In Operational Patterns of the 'C' kind, the decoder can choose from 2 or more ways to play-out the file.

[2] If a clip has both picture and sound information, this will mean that the picture and the sound samples are interleaved within the file.

So for OP1C, for example, you can choose from different versions on how to play the file, each individually following the rules in OP1A or OP1B. And a similar rule applies to OP2C and OP3C. A simple OP1C example would be a file holding both a full quality version of the material and a proxy[3] version.

All the above Operational Patterns are called Generalized Operational Patterns, since they represent a general model for application scenarios.

The industry identified the need for additional Operational Patterns, targeting specific scenarios. In MXF terminology these are Specialized Operational Patterns, and OPAtom is a well know example.

OPAtom was designed to be the simplest way of placing single track material within MXF. Therefore, an OPAtom MXF file can have a picture track or a sound track; but not both. You can then use MXF Metadata to link together files that are part of the same clip. To have both in the same file, you should use OP1A.

Summarizing, classes of applications were thought up and patterns were identified on the way that these applications would need to organize their MXF files. This way, through the Operational Pattern, an MXF file can signal the decoder its "class" and therefore the level of complexity within. This way decoders can be built to support different levels of complexity, depending on the application field.

MXF CONTROL KNOBS... A THREAT?

All the other MXF control knobs quickly mentioned above, and quite a few others not mentioned here, are also important. However, the knowledge on their meaning and implications didn't travel from the people coding the bits to the general public knowledge, yet. So you can regard the Operational Pattern control knob as the first one to surface...

Is there a threat in there? How long will these other control knobs take to surface? And when they do, how complex will this get?

Actually if you are already familiar with MXF, you know that the fact that MXF is being widely deployed, is leading its usage patterns to be identified by a number of users and manufacturers working together at SMPTE and other forums. This is leading MXF experts towards an understanding of the best way to surface those other control knobs. Nothing new, just business as usual, that is: learn from your customer feedback, and improve technology.

If you are reading this document as a newcomer to MXF, then you should see this as an opportunity. You need to start using this technology today. And you have the opportunity to feedback to manufacturers on your expectations for the functionalities of the next products' versions, so the control knobs can surface taking your input into account as well.

FLAVORS

The discussion in this document finally got us to the current state in the market. The concepts above should provide some insight on the why several manufacturers in the market, even within the same application areas, are deploying slightly different MXF configurations.

Some of these turned the Operational Pattern control knob to OPAtom. Others turned the control knob to OP1A. Others still, identified they needed to give their users the freedom to handle the control knob, so they let the user choose the Operational Pattern.

Now, here comes a problem: you look at a product's datasheet and see MXF; then you look at another product's datasheet, and read MXF as well. So, you can just connect them and it will all work, right? Not necessarily. As discussed above, internal control knobs may be set differently. And as discussed above, there are very good reasons for that from a manufacturer point of view, which make all sense.

Nevertheless, there are also very good reasons for users to become upset. After all, if there are differences, that is, if there are flavors, people should find a simple way to explain those differences. This is where we come back to the previous discussion, on the need to refine the technicalities of new technology into a higher abstraction that is understandable by the general public.

Manufacturers understood this and steps have been taken to make the users aware of the fact that MXF is not monolithic. Users are now being made aware that we are talking about a toolset, and products using this toolset expose control knobs. And that is

[3] Proxy is one of terms commonly used to mean a lower quality version of the content that can be handled with fewer resources. E.g: highly compressed and reduced resolution version for playback over the Internet.

why today, even newcomers to the MXF world, immediately learn a few technical terms like OPAtom and OP1A. Datasheets are now staring to include these terms as well…

A side note: which other control knobs will come up next? Probably Index Table related control knobs will surface in the near future… but the real show will be the coming of the Metadata control knobs.

Metadata handling was the goal of MXF from the beginning. That is where the real power of MXF will be unleashed. Several pioneer projects are being deployed today worldwide setting the scene for this. However, as stated above, that is a world on its own that deserves a number of documents dedicated to it so we will keep it outside the scope of this document.

OP1A AND OPATOM… IS THAT IT?

Most products in the market today use either OP1A or OPAtom, and it is not surprising that these "OPs" are so widely used. These are the simplest MXF configurations, and it was perfectly expectable that these would be the first ones to reach the market.

A number of projects are popping around the globe starting to use the higher level Operational Patterns described above. Of course that represents a dramatic increase in functionality, but also an increase in complexity. Therefore, manufacturers need to be very careful in designing products from solid user feedback.

This is also why most products out there supporting all Operational Patterns are development tools, such as the MOG Solutions MXF::SDK[4]. And projects are using those tools to implement pioneer deployments that gather feedback on users' needs, and feed those into the design of the next generation MXF products.

WHICH "OP" TO CHOOSE

So given all this, which Operational Pattern should you choose for your system?

The MXF purely technical answer to that is that you shouldn't choose an Operational Pattern. The Operational Pattern should be a consequence of your design choice for a system.

[4] Developed in collaboration with IRT – www.irt.de.

Each organization has its own specific internal processes. Some always treat picture information and audio information separately, and have deployed systems to keep track of the location of the different assets, and the relationship between them. In this kind of system, maybe you want to make sure you keep roughly the same processes when switching to MXF, so that your staff can handle the change easier. So if you conclude from your analysis that you want to keep the picture and sound information separately, you will probably conclude that OPAtom is the way to go. And you will then gradually improve the linkage and synchronization between the different assets, by using mechanisms available in MXF Structural Metadata that enable you to link the OPAtom files together.

A notable example here is the Digital Cinema distribution. There is great tradition in the Motion Picture industry of separating the treatment of picture and sound information. Therefore, it is not surprising that the DCI specification [5] adopted OPAtom with separate MXF files for picture and sound information.

On the other hand, suppose the staff in your facility is used to handle the Content contained within a package (such as a tape containing all the video and audio to be used). In this case, the least disruptive way to introduce MXF to your facility, would probably be to keep the files with both picture and sound information… a kind of virtual tape.
Therefore, in this case, you would probably want to have MXF files which interleave picture and sound information, which would probably lead you to the OP1A side[5].

A simple example here would be a department that keeps a short term archive of the rushes coming in from the camera crews, and that doesn't want to invest into file management databases to keep links between assets. They just want to have a file repository where each file contains the full package. OP1A would work as the full package in this case.

Now suppose you need to archive your material using MXF, and you would like each file to have both a full resolution version and a proxy version. Well, in this case you would conclude you need to

[5] MXF experts will probably note at this point that you could also perfectly have separate sound and picture with OP1A files… please note though that this document is all about guidance, general trends and keeping it on the mainstream. You always have the choice to go for specialized solutions, of course.

have multiple versions of Content in the same file, and it would probably lead you to OP1C.

One final example: suppose you want to send to a dubbing facility an MXF file, with some English-spoken Content for which you need an additional sound track with the dubbing to the Portuguese language. If you want to have the new track interleaved with the picture information, and therefore immediately optimized for streaming, this would mean a re-wrapping scenario for the dubbing facility, and you would probably get back OP1A.

On the other hand, if you are going to keep the remote facility technical requirements down, and just use MXF as the container that holds all the assets so these don't get lost, then you probably want to get back an OP1B file.

The way it could work is:

- Dubbing facility receives the OP1A;
- Dubbing facility appends the soundtrack to the MXF file (appropriately encoded using the MXF encoding mechanisms, of course);
- Dubbing facility performs some MXF OP1B "magic" in the Metadata of the MXF file, so that when you play it out, you listen to the new Portuguese soundtrack rather than the original English one... However the original English version will still be intact in the file, in case you need to reuse it.

ADAPT MXF TO YOUR NEEDS

The previous section focused on the purely technical answer to the question on which "OP" to use. Of course in reality we need to take into account more than that.

In the end, when you design a system you will be using products in the market. So, you may want OP1C, or OP2B or some other...but if the products you need support only OP1A and/or OPAtom, than you have to take that in consideration.

Does that mean you need to live with what you get from those products? Absolutely not. When a new technology enters a market, initially users have to live with what they get in the product box. As the manufacturers get more feedback and learn more about users' expectations, the ones that win the race are the ones that adapt the technology to the users needs... not the other way around.

So the key here is: you should adapt MXF to your workflow, not adapt your requirements to what specific MXF products give you. How do you do that? MXF adaptation processes. Different MXF configurations will work better on different sections of your workflow. So you need to take into account that you will need processes in your workflows that automatically turn control knobs for you to adapt MXF from one operation to the next.

Just look at how it was done with the compression systems. Of course initial products were deployed supporting very specific compression operating points. Nevertheless, soon a full industry was born providing converters that let you choose which format to use at each stage.

To capture video you probably use something like MPEG-2 I-Frame 50Mbps or DV 25Mbps... You distribute it to people's homes using MPEG Long GOP... You store proxies for browsing using MPEG-1 in your media asset management ... and probably even post a few clips on your company's website using MPEG-4, for example...

Conversions are all over the place in your workflow already, adapting your Content to each specific section of it. Doesn't it make sense that similar adaptations apply also to the MXF wrapper[6]?

SO WHY DO WE NEED MXF?

I guess the obvious question after that would be: why do we need a standard called MXF if you need to do adaptations anyway? Why not just different formats at each stage then? Because no matter which conversions, for example from an OP to a different OP, at both sides of the converter you still have the same technology: MXF. And there are a number of MXF tools[7] that you will always be able to use, irrespective of the Operational Pattern to perform defined sets of operations with the files.

Additionally, on the manufacturer side, it means engineering teams can concentrate on developing and researching using one base set of tools (the MXF standard document suite) rather then having to reinvent the wheel for each specific step in the workflow. This means lower costs for the manufacturer, improved product quality and

[6] If you think about it, transport/wrapper format conversions already occur all the time (think of current workflows mixing tapes, SDI connections, satellite links...).

[7] An example is a MXF player and Metadata editor like MOG Solutions theScribe. Some of its functionalities apply only to specific "OPs". However, something like playing and Metadata editing applies to any Operational Patterns you throw at it.

reduced time-to-market, with the obvious advantages this implies for users as well.

Finally, keep in mind that a major goal of MXF is to provide the framework for Metadata handling. The model of constant adaptation throughout the workflow, always using MXF, enables you to deploy the foundation framework you will need to tunnel your Metadata through. You can hardly do this if you have different formats across the workflow.
This model also prepares you for the next step, where your system becomes more comprehensive. You will then get to a stage where you identify the need to plug-in Metadata adaptation as well, at specific stages of the workflow.

SAMPLE SCENARIO

Before we conclude, maybe it will be helpful to the reader to consider a sample scenario that uses some of the examples in the text above, and where it becomes clear how MXF interoperability doesn't mean that every box in a system needs to understand every MXF flavor. It just means that the same MXF toolset is adapted to fit the needs of different points in the workflow.
In this simple case, and since we have been focusing on the Operational Pattern control knobs, adaptation tools being deployed are Operational Pattern converters integrated, in some cases, with compression codecs. Don't forget though that you can have adaptation tools automatically turn for you a lot more MXF control knobs than just the OP control knob.

In the figure, Content can get in from different providers. Some deliver OP1A some deliver OPAtom.

All input is stored in short term archive and, to simplify file management, all Content is kept as OP1A files.

The editors in this particular facility always treat video and audio separately, so all OP1A Content is split into OPAtom files before entering that department. The output is a OP3B file, which is basically an MXF file which includes the editing list and all referenced clips.

The OP3B Content is delivered to a proxy production system, which transcodes all Content to a proxy version packaged as an OP1A file.

The proxy is sent to a dubbing facility over the Internet, where an alternative language audio track is created, and then added to the file making it now an OP1B proxy file.

Back at the post-production facility, the OP3B file is combined with the dubbed soundtrack of the proxy OP1B file to produce a final OP1A file, which holds the final product.

Notice that throughout this process you have Essence transcoding, just as you already have today. Notice that you can also have Metadata transcoding since different Metadata fields will be relevant at each point. As an example, maybe you want to have the source camera serial number and cameraman contact information on all rushes... however it probably doesn't make sense to keep that information on the proxy being sent to the dubbing house.

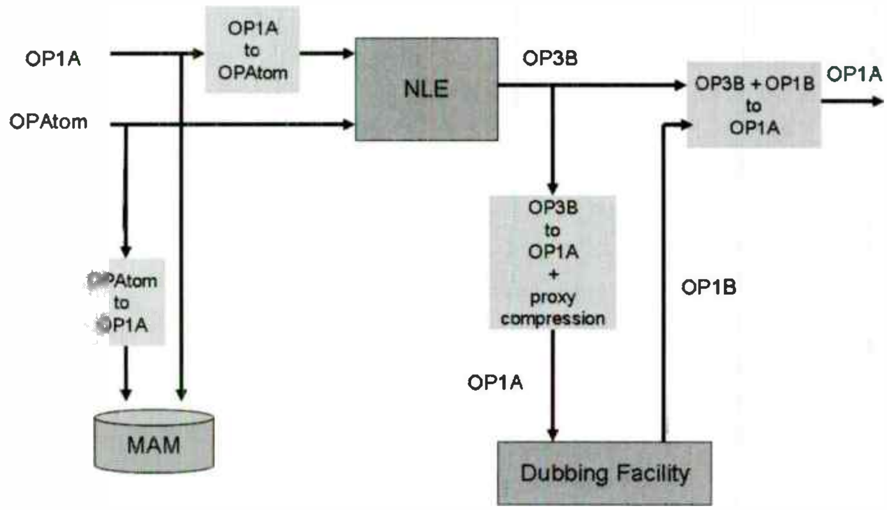


Figure 3

On the other hand, Metadata such as episodic information, rights information, transcripts of the soundtracks, among other, may be very useful to keep across the workflow in the different versions of the Content. And you can tunnel this Metadata throughout the all system, because it is using the same MXF wrapping technology everywhere.
In the figure, you could easily make that Metadata travel through all the MXF paths and make it editable at any point you wish... While still taking advantage of the processes automatically tweaking MXF control knobs for you, to adapt that wrapping technology to most suitable configuration at each point.

CONCLUSIONS

MXF was developed to be extensible and to be adaptable. This enables this technology to be used in a variety of different scenarios.

Adaptability is achieved through control knobs that enable configuration of MXF for each scenario. Some of these control knobs are already common knowledge, with Operational Patterns getting the popularity prize there. Other control knobs will emerge, as the market gains more experience on this technology, and more projects experiment with higher level OPs and rich Metadata workflows.

This document seeks to project a image of a technology that is replacing a number of other technologies used across complex workflows. This document puts forward that it is not feasible to reach such an ambitious goal with a monolithic approach. Therefore, MXF should be regarded as a toolset that is expected to be constantly adapted throughout the workflow.

This means that not all blocks in the system need to support all possible MXF configurations, therefore simplifying their design and reducing their costs.

However, since a common toolset is being used, this enables new products to be developed and easily integrated in the systems. It also enables rich Metadata applications, the principal goal of MXF, to tunnel their Metadata throughout the system.

Manufacturers need to do a better job at outlining their products functionality and interfaces. They also need to provide more Metadata functionality on their systems, so that the tunneling of that Metadata through complex workflows, with constant adaptations as in the scenario above, clearly demonstrates the functionality MXF aims to provide.

Users need to be more proactive. Yes... it is true that big-budget projects can change complete workflows and deploy spectacular systems. But small-budget projects can't just wait to see what happens next. It is already a fact that MXF is the way to go. The only question is which route each user will take.

Results from other projects will for sure provide guidance on the route to be taken. Nevertheless, in the end, the technology needs to be adapted to each users needs, not the other way around.
Therefore, step by step, each user needs to proceed in the adoption of this technology, bringing in the adaptation tools being made available, and deploying these with a strategy of adaptation to their own way of doing business.

REFERENCES

[1] Devlin, B., Wilkinson, J., Beard, M., Tudor, P. and Wells, N., 2006. The MXF Book. Focal Press.

[2] Santos, E., 2006, AAF, MXF and XML... Putting it all together, IBC 2006 Conference Publication

[3] SMPTE, 2004. S377 - Material Exchange Format (MXF) - File Format Specification. http://www.smpte.org.

[4] SMPTE, 2004. S380 - Material Exchange Format (MXF) Descriptive Metadata Scheme-1. http://www.smpte.org.

[5] Digital Cinema Initives, LLC, 2005. Digital Cinema System SpecificationV1.0. http://www.dcimovies.com

MXF based workflows – a practical implementation.

Steve Fish
TBSEL
London, England

ABSTRACT

This paper will present recent and ongoing work by Turner Broadcasting in Europe to select and apply higher MXF operational patterns to deliver positive business benefits. The paper will focus on constructing use case models of our working systems and using these to review the many choices that an end user must consider in implementing an integrated end to end MXF based work flow, including which operational patterns and metadata standards to use for ingest, playout, post production and archive based storage systems and how these will deliver positive business benefits and improved functionality inside existing systems. In showing how the patterns were chosen consideration will be given to both essence and metadata handling throughout the entire cycle.

WHY?

Once upon a time, tape management consisted of a physical tape library, a library asset number and knowledge of what was stored on each tape. Checking this information was slow and cumbersome in that to confirm what was on the tape you needed to compare the record report with the slate or clock by reviewing the tape. The system worked, but was conditional upon good librarians blessed with a photographic memory and few requirements to generate complex reports. There was significant cost in editing or generating copies in time, technology and finance terms. Careful thought and regulation therefore needed to be integral to all working patterns and processes. The system was self limiting due to the fact that no matter how much care was taken with bandwidth, noise and resolution every copy suffered a generation loss. As generation loss could be described accurately in technical terms, it was clear at which point the quality dropped below the broadcast standard and for most tape formats this was only a few generations. In all cases the known deteriorations ensured that the work delivered was of the lowest possible generation.

This generation loss is a significant problem in the multi-channel, multi-use broadcast world of today. The need to have multiple different copies across multiple different regions means that videotape does not provide enough flexibility. The advent of digital tape reduced the issue of generation loss, but it did not address the problems of expensive essentially "*lossey*" tape systems that were designed on the basis of a defined amount of essence storage and covering up missing 'bits' rather than methods of reliably storing all data.

File based storage has removed some of these problems, for instance a copy with no generation loss became a clone and practically free. Despite the benefits, new technology brings new challenges and problems have arisen. Easily generated multiple clones do not have a physical representation on which to stick an asset number. Therefore more stringent requirements exist to manage these virtual assets in order to keep track of both their content and location. No longer can a scrappy piece of paper with the asset details on, be the asset management. This data must exist with the file and remain with it throughout its life, thus there is a whole new class of data handling requirements….and so metadata was born.

Unfortunately the new world still mirrors the old. Different formats still exist; each manufacturer has used different compression systems and wrappers. This means that while everything is constructed to recognised standards there is no true interoperability. Manufacturer A's system does create files immediately useable on Manufactures B's system. In order for the systems to work together transcoding, and / or re-wrapping is required. This process may introduce generation loss and thus limitations still exist in how many times an asset can be moved around a system and given the myriad of systems currently in use, maintaining quality throughout an asset lifecycle is not realistically possible, given that concatenation of different coding systems is not deterministic in understanding the deterioration to the quality of the signals. In order to manage this issue complexity is added to the management and control system.

Combining the facts that file based world has limitations and that Videotape doesn't meet our needs and (in Standard Definition form at least) doesn't have a long shelf life led to the conclusion that a more consistent system of technologies was needed in order to remove these technology based bottlenecks. It had to be one that incorporated the best of the old and new designs. It needed to allow for interchange of media without limitations, but within standards. Ideally, it also had to be more deterministic in regard to the control of quality. In short it required the industry to look at re-working its systems so they became consistent, logical and well defined. This would then allow for more deterministic and easily changed processes where and when business requirements dictated.

WHAT?

All technology projects are ultimately driven by business needs. TBSEL realised that a full understanding of current and future requirements were essential if they were going to drive towards full interoperability of all systems. These requirements needed to be analysed and built into a formal specification that in turn could be implemented by manufacturers, thus enabling the manufacturers to deliver new products with minimum changes to equipment, software, development timescales, cost and support expectations. TBSEL undertook extensive research into the needs, workflows and working practices at local, regional and international levels. The conclusions reached showed that evolutionary technology that could be phased into systems as part of the natural upgrade and replacement cycles was necessary, and that all changes needed to encompass positive business benefits that would lead to better efficiency and flexibility. These studies included a careful examination of the options available in the broadcast file handling world. The only option deemed complete enough was the set of MXF standards based on SMPTE 377M et al.

Whilst SMPTE377m has flexibility built in, it has too much flexibility and to date most implementations use simpler operating patterns within well bounded systems. Manufacturers were understandably nervous of undertaking development work that involved handling in full every single operating pattern above the chosen one, and as it became clear that practical operations would need to use higher order patterns it was realised that some solid guidance would be necessary to steer developments. With this in mind and the fact that there were no other large implementations to learn from TBSEL realised that we needed to write this guidance in the form of a recommended practice. The aim was to gather together the relevant parts of the existing MXF standards in a way that could easily guide manufacturers as to our needs, and also put a solid stake in the ground as a possible way forward for others with similar processes and workflows.

WISHES

As part of the requirements gathering TBSEL devised a hit list of key items that they either already enjoyed or wanted to extend the system to do. The following is part of that list and helps to show why there is a need to work with the higher operational patterns;

- Data entry. Data should be entered once at the appropriate point in the asset life cycle. This data should be available to any system looking at that asset.
- Data updates. There should be the ability to revise or edit data. Any revisions or edits must be updated across the 'Inventory'
- Files usable before being closed (i.e. partitioned.) This allows preview work to start before the write cycle is completed and the file closed. The time between first byte arriving and first use should be reasonable short.
- The concept of an asset 'Inventory'. This is where everything related to a single asset is stored. Asset in this context means a single episode of a show.
- The ability to extract a copy of this 'Inventory' and be able to send it to an independent site. Once received, it must be importable to the asset management system (which may be different to the source system), based solely on the supplied file. This means it must be self describing with all essence and metadata bound in. One benefit of this is that best of breed local solutions can be employed and reduce reliance on single monolithic systems across organisations.
- Storage system independent. As assets are expected to be used in many different places at different times there must be no reliance or limit on particular physical storage devices or systems.
- Flexible restores. A need was identified for 'Partial Restore'. This is where a selection of essence tracks from the inventory would be restored to form a new version of the asset. This functionality would be further extended to give 'Partial Partial Restore' where a time sliced selection of essence tracks would be restored. Finally, a 'Partial Partial Partial Restore' requirement was identified where a more than one time sliced sections of a selection of essence tracks would be restored, this last requirement was particularly aimed at clip selection for edit purposes.
- Flexible stores. The need to add tracks (essence or metadata) to an 'Inventory' without the overhead of having to re-write the entire 'Inventory'.
- Metadata must reside with, and be updated throughout an assets lifecycle.
- Essence files to be external. As flexibility of adding (and removing!) of essence was required it was realised that for this to be most efficient all essence tracks should be external.
- Where requirements have already been addressed in compatible standards, such as work done in DC28 efforts should be made to harmonise with these standards.

THE CHALLENGE

As part of the requirements capture, the challenge was to be able to present the results and resultant requests in a clearly defined way. Traditionally, broadcast systems have started life on the back of an envelope. A good System Integrator could take this and produce a set of design documents from which a healthy discussion would take place to finalise the drawings and hence the system requirements. From this the system and its features and limitations would be understood. With software of course it is all different there are no meaningful 'system diagrams' in the conventional broadcast sense just a collection of computers, a couple of servers, some network switches, a few screen shots and maybe a data flow diagram of some sort. There is no way to describe adequately the functionality expected of multiple software programs, running on multiple computers with multiple different interfaces that does not involve multiple degrees in computer science! The quest was to generate documentation that made sense to the broadcast industry. In order to do this we adopted the following approach.

WORKING DESCRIPTION

The approach taken was broken down into the following three steps:

1) Put all functional blocks on a diagram, label them and link them at a physical / networking level.
2) For each major workflow create a table that shows the individual actions required in that workflow. Each action should be documented to describe what happens at the physical, essence and data levels.
3) Convert each action into a UML activity diagram that shows in sufficient detail what is required from the action, and the messages that are likely to flow.

The benefits of this process meant that it was easy to start at relatively high level and quickly focus on the functionality. It also became very easy to see where re-use of actions would take place and avoid unnecessary duplication of work.

A model of our system was drawn up, showing the major system components and labelling each with a unique reference. An example of a simplified version is shown in figure 1. Actual interfaces are shown as solid interconnecting lines, network connectivity shown as dotted lines, we ignored all other connections as they only added complexity without adding clarity.

Having generated the diagram the next step was to break down each major workflow into a sequence of steps that dealt with one action at a time. The major workflows identified were Ingest, Playout, Add Additional Essence and Edit. The action information was simply captured in a spreadsheet. Table 1 shows the fields we captured.

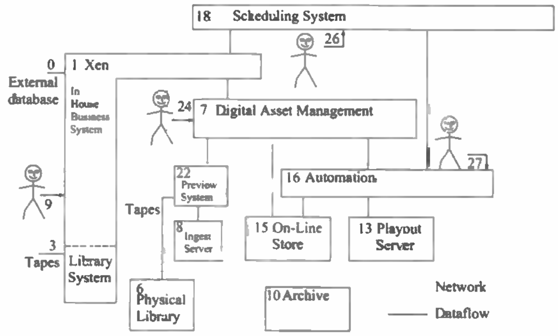


Figure 1

Field	Description
Source	Source data (Meta or Essence) upon which action is to take place
Destination	Destination for data once action completed
Level	In the event that there are multiple different actions acting upon the same source and destination this differentiates between them
Brief Description	
Type	What type of data is being acted upon (Essence, Metadata or both)
Request Mechanism	Which system or process initiated the action
Metadata Source	Which process or external agent did this data come from
Metadata Destination	Which process or external agent is this data going to
Schema detail	The schema under which the data will lie
Essence Source	Which process or external agent did the data come from
Essence Destination	Which process or external agent is the data going to.
Notes	Any final notes needed

Table 1

These fields became column headers and each action was entered as a row. To help describe this better an example is shown below of a copy function. Here the task is to copy a file from the On_Line store (15) to the Playout server (13). It is assumed that a request has been made by the automation system (16). The high level requirements are indicated by the red arrows in Figure 2, and actual copy action can be seen in Table 2.

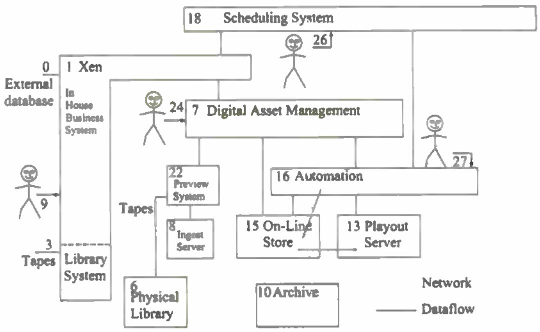


Figure 2

Field	Value	Notes
Source	15	MXF file in On_Line store
Destination	13	Playout server
Level	A	
Brief Description	Copy file	
Type	Essence & data	
Request Mechanism	Automation (16)	Playlist request
Metadata Source	N/A	
Metadata Destination	N/A	
Schema detail	Internal to Automation system	
Essence Source	On_Line server	
Essence Destination	Playout server	
Notes		

(It should be noted that not all fields are always relevant and can be omitted as appropriate.)

Table 2

The analysis highlighted specific processes that involved MXF files and any interoperability between manufacturers. In the example discussed here, the process is a simple copy between two similar systems (a clone!). However, it could as easily be between two dissimilar systems and therefore the copying may involve a transcode and/or re-wrap as well as the actual copy. As this is a two action requirement the actual action table would have two entries showing a transcode / wrap as one action and copy as a second action. It is necessary to be very specific about the individual steps taken in pure process terms in order to ensure that all steps are captured. It would be left to individual manufacturers to implement an action or group of actions in the most efficient way.

With all the actions captured, the next step is to model this in UML. Continuing with the copy example, the activity diagram is shown as figure 3. This shows the request mechanism as Automation (16). The request is sent to the 'Processor' with parameters P0-P2. In this case the processor is a module within the Automation system; however it could equally be another device independent of the automation. The parameter data will include source and destination device locations. In this example the processor is shown broken down into an interface, a controller and the core worker code, this allows us to describe the messages available and ensure that these are adequate and properly handled.

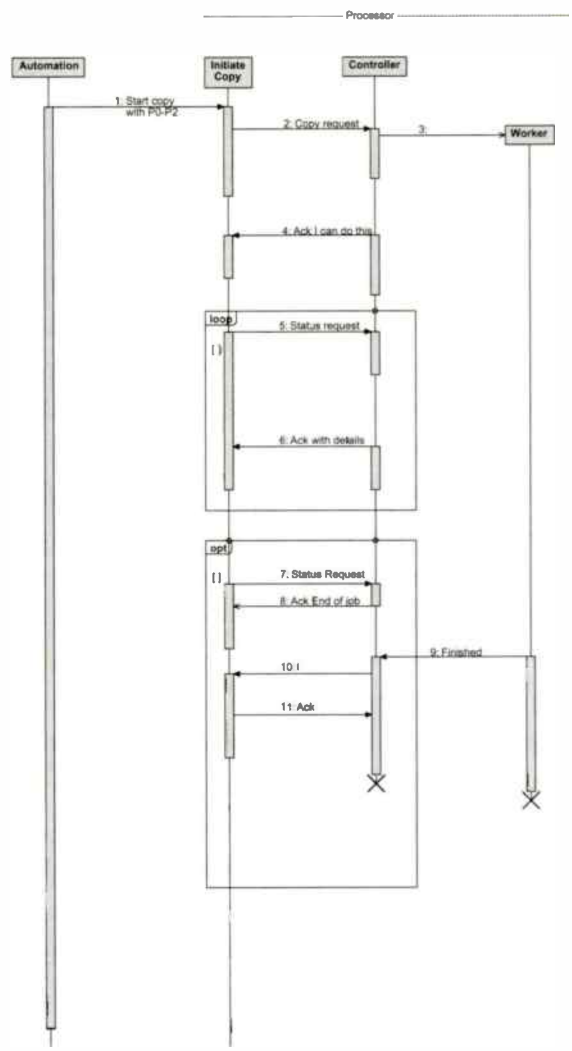


Figure 3

Upon building these diagrams from the action tables, it quickly became clear that, at the interface level there was a common set of parameters being used for almost every process. This led to the conclusion that a standard interface could be specified thus simplifying the specifications and providing a strong foundation on which to base the control level interoperability. It also meant the control and worker layers of the processor could be abstracted, allowing any MXF toolkit manufacturer to provide these tools, if they built an interface to the common control set. This had the advantage that as there are a few good toolkits already in existence, these could be built on and help speed up the development process.

A further benefit is that as new requirements came along they could be modelled in the same way. It is then easy to recognise re-useable actions from the original analysis. In many cases these workflows created no new actions as they were simply a re-use of existing actions. If new equipment was introduced the main diagram would need to be updated and the new actions related to this equipment would be added in the same way.

INVENTORIES, VERSIONS and COMPONENTS

MXF recognised that different businesses have different file handling requirements and reflected this in allowing many different levels of complexity in the types of files and their use. These are all brought together in a table of Operating Patterns or 'OP's' which describes the functionality available at each operating point. When trying to apply this to programming it is easy to get confused, especially as the specification expects that the lowest operating point be used for any file, this has the property that a particular asset starts at one 'OP' and following some process has changed to a different 'OP'. While it is important to understand this from a technical position it does not lend itself to accurate descriptions of programming in the broadcasting world. For this reason it was decided to use 'Inventories', 'Versions' and 'Components' when talking about assets.

- 'Components' are single essence or data tracks e.g. a Video track, Audio track or Metadata track.
- 'Versions' are pointers to Components and describe how a collection of Components may be played. There can be many different Versions pointing to the same Components.
- 'Inventories' are a repository for all the Versions and Components for a single asset. There should be no duplication of Components within the Inventory. There maybe many similar Components of the same type, for instance there could be three different Videos – High Resolution edit master, Transmission master and a Browse resolution copy. At least three versions would exist to allow each to be played.

These definitions can be loosely related to Operating Patterns in the following way:

- Components OP1a.
- Versions OP2B or OP3B depending on how complex the program structure is.
- Inventories could OP3B or OP3C depending on how many versions are in the Inventory. This shows that talking in 'OPs' is variable and also that all Operating Patterns are likely to be met in a complex system.

METADATA

From the hit list of key design requirements it will be remembered that Metadata must reside with, and be updated throughout an assets lifecycle, thus definitions were required.

The key questions were;

- What metadata needs to be stored at Inventory, Version and Component level?
- How is this incorporate in a standard?
- How are legacy business systems handled?

Early analysis had shown that metadata definition is personal, with every group, division and region having different definitions for common terms. As the MXF standard allowed multiple metadata tracks in the same way as multiple essence tracks, this diversity could be accommodated by the storage of site specific metadata from existing business systems on different metadata tracks without having to force a difficult conversion of the existing data into a standard container. This it was felt provided a suitable evolutionary approach to implementation that realistically dealt with existing, often complex legacy business systems. In the short term, selected existing data could be mapped into DMS1 and updated whenever an update occurred or, Version extracted from the Inventory, thus allowing interoperability testing to start and provide a realistic view of the problems faced in the real world.

In terms of practically implementing this there were three design goals.

1) Allow each division within the organisation a metadata track at Inventory, Version and Component level that effectively could be used as a private or dark track. This would enable all divisions to keep their current asset management systems and store all their data in an easy and consistent way with evolution in mind.
2) Where possible it should provide a mapping to the DMS1 schema. This would be automatically generated upon any extract from the inventory, thus keeping the DMS1 data relevant and up to date.
3) Migration of all data to a standard taxonomy and lexicon is a long term goal.

A major discussion point was how the metadata should be described in a standard way. MXF is intimately linked to DMS 1 and it was felt this would provide a good platform to work off. There was little confidence that we understood all the issues and problems that would be encountered as the systems grew to encompass all workflows within an organisation and with the ability to hold private data tracks it was felt that taking this route acknowledged that standardisation was needed and could be built in over time.

AUDIO

Within the multiregional broadcast world is the need to handle many different audio languages. A major challenge within this area is how to accurately describe an audio track or Component. This is complex as there are many options; for example there is a program 'Gone with the Wind' for which the rights are purchased throughout the Europe Africa and Middle East region. Initially the asset is supplied with Video, an English Stereo and a M&E Stereo. At this stage it is relatively simple to describe these audios as shown in Table 3

Language	Type	Tag
English	Stereo	ENG
M&E	Stereo	ME

Table 3

The Tag column is used to describe how the track is routed in the playout system. i.e. where the signal exists in the SDI stream. This needs to be fixed to ensure it correctly appears in the Satellite multiplex and when the customer selects 'English'. However is the language description of 'English' suitable? Is it American English or English English? At this point it is sensible to add an additional field which describes the dialect. Table 4 shows this.

Language	Type	Spoken	Tag
English	Stereo	American	ENG
M&E	Stereo	M&E	ME

Table 4

While this is relatively simple for English, it becomes far more complex in the Spanish market where there is the option of Catalan, Latin (South American) and Castilian. To further add to the complexity sometimes an acquired program may come in with the main audio in the language of the original acquisition, for example Japanese. In this case the Language details can be seen in Table 5

Language	Type	Spoken	Tag
Japanese	Stereo	Japanese	SPA
Spanish	Stereo	Castilian	CAS

Table 5

Regionalised Audio Description for Access Services need to be described in the same way as the main language, as it broadcast as a separate track in its own right.

Once the main track requirements have been described further information may still be needed for business reasons. An example is does the track contain any profanity, or mention anything that would give a regional censor cause for concern? What happens if there is a scene in a language different to the main narrative? In transmission terms the system will follow the Tag and ensure the track is routed correctly, but there maybe good reasons to know that the track is not a single language. A more systematic approach to the labelling of the audio tracks improves the ability to search and repurpose libraries and more effectively use purchased rights; it may also benefit content distribution if the descriptions can be adopted universally.

DMS 1 has multiple fields for describing audio content, these are many and varied and yet none fully addressed the complexity of describing multilingual operations RFC 4646 was considered a much better alternative as it had the extensibility to bring together the potentially complex description requirements discussed previously in a reasonably human readable format

In order to demonstrate how this works take the example from Table 6.

Language	Type	Spoken	Tag
English	Stereo	US	ENG
French	Dolby E	French	FRA

Table 6

Converts to:
en-US-x-Stereo-a-noprofanity
fr-FRA-x-DolbyE-a-noprofanity

The actual example used shows safe audio for any audience. If there is also available, an 'after watershed' version, then the table could be extended as shown in Table 7.

Language	Type	Spoken	Tag
English	Stereo	US	ENG
French	Dolby E	French	FRA
English	Stereo	US-AW	ENG1
French	Dolby E	French-AW	FRA2

Table 7

Converts to:
en-US-x-Stereo-a-noprofanity
fr-FRA-x-DolbyE-a-noprofanity
en-US-x-Stereo-a-profanity
fr-FRA-x-DolbyE-a-profanity

It is seen that this is more flexible and robust in coping with the demands of the business.

WHERE ARE WE

TBSEL has recently built a playout facility in London which is the base on which the workflows for this project have been modelled. Workflows showing Ingest, Archive and Playout have been documented and turned into code; demonstrations of this software have taken place. There have been many challenges along the route as when the project started we had no methodology on which to document the workflows and actions, as this has been built up in parallel with the descriptions time has been wasted going blind alleys. We now feel confident that there is a solid system on which to base future work. Beyond the challenges of describing workflows and actions there have been very real challenges in ensuring technical decisions enable flexible business decisions with the right level of control, for instance how tightly does the Inventory keep track of how the versions reference essence, is this an MXF issue or Asset management system issue. Not all the answers are known at this stage and some refinement is expected as the demonstrations continue.

The project has now reached a point where demonstration software has been developed and is available for viewing at NAB.

CONCLUSION

The broadcasting industry needs to evolve its strategy and products if it is to survive and thrive in the face of unprecedented technical evolution and commoditisation. This paper has presented the ongoing work by TBSEL in bringing together the increasingly complex requirements of a multi-channel broadcaster with analysis techniques that allows accurate and rapid description of those requirements in a work flow and file dominant world. MXF has been shown offering a consistent and adaptable solution to media handling that is not available in other file based systems. Bringing the two together enables solutions at local, regional and enterprise level for complete and consistent workflows and improved efficiencies, with the ability to ensure that these visions can be correctly described to manufacturers.

Acknowledgements:

Granby Patrick – Marquis
Oliver Morgan – Metaglue
Paul Turner - Omneon
Benoit Fevrier – Open Cube
Dave Collins – Pro-Bel
Richard Harlen – Pro-Bel
Howard Twine – Pro-Bel
Mark Ostlund – Quantum
Bruce Devlin – Snell & Wilcox
Russ Wood – Softel
Clyde Smith - TBS Inc
John Morgan – TBS Inc
Ron Tarasoff – TBS Inc
Quincy Johnson – TBS Inc
John Bukowski – TBS Inc
Oliver Gallienne – TBSEL
Tony Taylor – TMD

ENSURING MEDIA EXCHANGE FORMAT (MXF) ASSETS' AVAILABILITY

Bill Gage
EMC Corporation
Louisville, Colorado

ABSTRACT

The future will bring use of complex, interoperable formats such as MXF to the video broadcaster's world. This paper discusses key factors which must be considered by archive systems to ensure the accessibility and availability of MXF assets. Highlights include techniques to ensure that such assets are immutable through time and across tiers of storage. Implications for archive design and interface to meet the performance needs of the broadcast system will be explored. Further, the paper addresses the fit of an archive system into a broadcast station's end-to-end workflow. The ability of a central or distributed archive system to accomplish the capture and storage of meta-data associated with video assets and provide data management policies will become critical to success of broadcast operations.

BORING IS BEAUTIFUL

In broadcast engineering, as opposed to content creation, boring or reliability or consistency is beautiful. What makes a good broadcast system and good components of a broadcast system is that it works: simply and reliably. When material is played-to-air, creative decisions have been made. The material is (or should be) immutable. It needs to be played out as intended, without modification, repeatedly. No broadcast engineer ever wants to have the "excitement" of explaining why the material scheduled to play is different than the play list or, heaven forbid, is "black".

In content creation, the creativity should be reserved for the people who create and produce the content, not the systems they use to create content. Material, once edited, should be what a human intended, not the system.

The process of moving material from creation to playout had better be boring or reliable, as well. There should be no exciting surprises involving damaged tracks, incompatible formats, or transfer delays impacting getting material to air.

The archive components should be the most boring part of the systems. Any material stored to the archive repository should be brought back to online playout or editing systems' storage just as it was. No video frame, no corresponding audio, no associated meta-data should differ.

COMPLEXITY OF ASSETS

Not so many years ago in a broadcaster's world, video assets were simpler to manage: analog video tapes with attached hand-written notes describing the tape contents. The assets were moved around the broadcast facility via "sneaker net" from news / production to broadcast / playout. Catalogs of information about assets were maintained in a variety of "interesting" fashions, more often than not with manual re-entry of only some of the descriptive information (meta-data) about the assets.

Things have "matured" considerably. Assets are now stored in digitally encoded formats. They make their way around broadcast facilities as digital data streams. Meta-data about the assets is captured at many steps in the business processes: creation, editing, post-production, and playout.

Fortunately, there are standards of digital encoding for video. But unfortunately, there are many standards and many interpretations of these same standards. A video asset with a standard format (i.e. MPEG), bit rate (i.e. 15Mbps), a standard audio format (i.e. AES), number of audio channels, etc. can be played out on the equipment on which it was created or ingested. However, it is unlikely to be editable or playable on different equipment. To make this happen, more passes through the material are required to re-format and/or re-wrap the material to a compatible format. These are performed by additional software and/or hardware transcoder components.

Beyond the "simple" issues of transformations to provide interoperability among software and hardware systems, modern video assets often require additional attributes to make them compatible for different delivery. For example, an asset in NTSC (or PAL) format suitable for broadcast needs media transformation to prepare it for IPTV/VOD servers, cell phone content distribution platforms, or for low resolution preview editing by producers. Further, some broadcasting stations require additional language tracks, captions, or altered video content to comply with local customs and/or communications regulations.

To keep track of all the meta-data about the video asset, including the multiple versions, the discerning broadcaster will want to keep descriptive metadata, including portions of the change history with the video asset itself.

As we have heard for a few years, the Media eXchange Format MXF, promises to help with this complexity. The standards documenting this format are, in part, a reflection of this complexity. There are several SMPTE standards, recommended practices, and engineering guidelines needed to document MXF.

If an archive storage system is to support the needed complexity, it must support a major portion of MXF. The MXF-aware archive should be able to store and restore video assets using a wide variety of MXF operational patterns. It should demonstrate interoperability with a wide variety of video server, transcoder, asset management, and broadcast automation software and hardware. In the rich world of MXF, not all implementations by video equipment and software manufacturers will be the same.

INTERFACING WITH THE ARCHIVE

To move material into or out of an archive should be simple and once again reliable or boring. Asset Management software systems and Broadcast Automation software systems should basically "work" with the MXF-aware archive, without extensive modifications. The interface (programmatic API) presented to these archive "users" should be a level of abstraction above the details of the video asset. For most operations, one should be able to refer to an asset by name and have the archive perform operations on the entire collection of container wrapper(s), essence tracks, and meta-data comprising the asset.

When more granular access is needed, the MXF-aware archive should present an additional interface should present functions to access individual essence or meta-data components without burdening the user with any more details about the data than are necessary. (See "MXF API" for more on this aspect.")

The data interface to the archive must be independent from the command and monitoring interface. The very large byte count of video material, particularly HD, will consume network infrastructure well beyond the expectations of typical IT systems. Within a Broadcaster's archive, there are usually a small number of files: between 100,000 and 10,000,000. However, the files themselves are extremely large: from a few hundred megabytes to tens of gigabytes. If data flows involving many such files are running concurrently, they can soon overwhelm the file handling software normal to most common IT applications.

Transferring such large amounts of data, with sustained data rates of several hundred megabytes per second is only possible if the archive system's data interface allows as many independent data paths as needed. Further, the archive system must provide governance on the use of data paths to avoid overloading source and target devices and network infrastructure.

Some things needed within an MXF-aware archive should not be presented via the API. The archive should not burden archive users (asset managers and broadcast automation) with explicit data management operations. Without explicit instructions, the archive must be capable of moving material among storage tiers (both to less expensive long-term storage and back to near-line storage). Copies of video assets on different storage tiers should be automatically removed when established data management policies dictate. The archive user should not have to instruct the archive on these data management polices.

Similarly, archive users should not be burdened with storage management details. The archive should conduct media management, device management (drive and library), and infrastructure governance without attention via the API. When a new asset is introduced into the archive repository, the appropriate storage should be provided without any specific instruction needed. When an asset is retrieved, the user should not have to ask that a tape be mounted or a set of files be moved among storage tiers.

AUTOMATIC DATA TRANSFORMATIONS

In a typical workflow, a video asset might go through a number of transformations from ingest to playout:

- Initial encode from analog to digital format
- A pass over the video to add an essence track at a different bit rate
- A pass over the video to add captions in an additional language (either replacing those in the VBI or adding a separate track.)
- The addition of a wrapper defining a new reference to the changed bit rate track and the new captions
- A pass over the data to make the result compatible with a play-out server.

Each of these passes through the video data may require data movement thorough a separate process and even to a separate server.

Surprisingly, the computing power needed to build and certify humble TCP/IP network checksums can be come quite large. Information Technology engineers noted this phenomenon many years ago and have been working on techniques to improve the situation. So far, there are no easy answers. In recent tests, we have seen that these checksum computations on network transfers can require more CPU than the transcoding processes for video material.

In an ideal world, an archive storage system can be tightly integrated with the transcoding software needed to make many of data transformations shown above. This will enable the transformation of video as the material is "in flight" from the archive to a target video server. If in the end-to-end workflow, the archive system can combine multiple operations in each pass over the asset, the performance of the complete system can be increased significantly. Less passes over the data means better throughput and less equipment on the floor.

An MXF-aware archive system that is closely integrated with other needed processes in the workflow can be used to bring equipment costs in line with budgets and provide a more efficient operation.

"REMOTE" REPOSITORY INGEST AND EXPORT FROM THE REPOSITORY

An MXF-aware archive repository can enable the efficient movement of material within a broadcast facility. What about the exchange of video material with other facilities and organizations?

If other organizations, e.g. small video production companies, can prepare video material as MXF files, these files can be stored directly into an MXF-aware archive. This can save processing steps involved in decoding video to analog and re-encoding back to digital format at the broadcast station: an error-prone and time consuming operation. Video can be written from production systems as MXF to some recently developed tape devices. Then, the material can be read directly from the MXF tape into the archive.

From time to time, material must be "exported" from the archive repository for shipment to other facilities. Again, using MXF-aware devices, material can be written to tape for transport to any facility or system that is also MXF-aware.

FUTURE PROOFING EXISTING ASSETS

Let's suppose a Broadcaster has many thousands of hours of video digitally encoded in their existing system. To take advantage of the rich functionality that MXF provides, they certainly will not want to play out (decode) and re-ingest (encode) that material to MXF.

The ideal, MXF-aware archive repository should be capable of maintaining the existing material in its native format while adding new material in the format of choice. When older, non-MXF material is retrieved, the archive should be able to simply and efficiently re-wrap (and potentially transcode) the material to MXF as side effect of the data movement from the archive to the intended target. If desired, the "new" MXF version of material can then be introduced to the repository either in addition to the existing version of the material or replacing the previous, non-MXF version.

These two techniques, MXF conversion when needed and re-introducing MXF versions can be used to smoothly change a legacy video asset base to the MXF world. The archive capable of maintaining both MXF and legacy material enables a transition over time rather than as a sudden "surge" event.

ASSET MODIFICATION: UNDER CONTROL

The fundamental nature of an archive repository is to preserve data integrity. If the material going into and out of the archive is simple, this is relatively simple as well. An asset is put into the archive as a single object, travels among the archive's storage tiers as a single object, and is brought back out as a single object. With more complex objects, many aspects of the archive must become more complex.

The initial consideration for archiving complex assets is that among components comprising a video asset, there must be "locality of reference". When the asset is introduced, all components must be successfully brought in together. As the asset migrates among storage tiers, the archive system must ensure that all relevant components are kept together. When the asset is retrieved, the components must all be retrieved.

In the MXF-aware archive, there must be certain transformations allowed to add tracks, wrappers, and meta-data to assets. To preserve the integrity of the archive, these transformations must be under strict control so nothing is lost during the processing and assets are accessible continuously. It should not matter how the archive achieves this end. The archive might maintain the asset as single object within archive or as a set of structures within the archive that reference component parts. With either approach, the archive must never lose any component or allow an asset's state to become inconsistent.

When dealing with complex MXF assets, the archive system must apply data management policies to the components of assets in a consistent and rational manner. Here is a seemingly simple example that illustrates the kinds of questions that must be answered. Suppose the policy for a particular asset is to initially store the asset on disk, then move the asset to tape 30 days after initial archive. This same policy should apply to new components added to the asset. What if the addition of the component comes when the asset is 60 days old? In this case, it would seem reasonable to bring the asset to disk with the new component and then move it back to tape when another 30 days have passed. But wait; why not add the new component to what is already on tape? We could attempt to reserve room on the tape or we write the entire asset (with the new component) to a different tape. There scenarios the archive must correctly anticipate is sometimes frightening.

AN MXF API

For many of the modifications to and transformations of video assets, the simple interfaces to archive repositories mentioned above are not sufficient. Interfaces are needed beyond those used to automatically deal with MXF video assets in their entirety by a simple reference to an asset name.

Over the past couple years, a group of vendors brought together by Turner Broadcasting has been working to develop an application program interface, or API, to provide a language to manipulate MXF video assets. Details of this API are beyond the scope or purview of this paper. This API encompasses all the functions needed for the MXF modification and data movement described above. An MXF-aware archive should provide support for this MXF API along with other interfaces that the archive presents.

CHALLENGES ACCEPTED, BACK TO BOREDOM

There are certainly challenges in architecting, designing, and implementing broadcast systems for the complex, interoperable world enabled by MXF. We can, however, look to good archive systems to help remove some of this complexity and return reliability or boredom to these tasks. With that in mind, broadcast engineers can concentrate on the excitement of bringing their programs to new audiences.

Achieving MXF Interoperability

Todd Brunhoff
Omneon Video Networks

Overture

MXF is the first widely accepted standard that has support for a variety of essence types, multi-clip assembly, descriptive metadata and ability to maintain a traceable record over multiple generations of edit sessions. The standard is largely based on SMPTE 377M, but also dozens of other interlocking SMPTE standards.

There are a number of ways to construct a movie based on MXF: the number of files in the finished movie, the presence or absence of values, and the size and order of high- and low-level elements. Many documents comprise the standard. Simple statements in one may conflict with or mutate clauses in another. Sometimes the writer is ambiguous, other times the reader may misinterpret. This was never the intent.

Two themes of interoperability are presented here: the evolution that leads to a problem, and the path to a solution through education or process.

What is MXF?

MXF is a wrapper file format. It describes ways for a file to contain multiple tracks of audio, video and data. It specifies how to compose metadata that describes the contents of each track and the timing relationships among them. An important distinction for a pure container format is that it does not specify a compression scheme or limit the types of media that it can contain. DV is a counter example to this in that it is both a compression specification for video, and a container format to hold the compressed video and uncompressed audio.

Other pure container formats have preceded MXF. For example, OMF (Open Media Framework) was developed by Avid in the late 1980's but lacked broad acceptance.

Quicktime is also a container format, but for this discussion that name is a bit muddy because it refers to three different things: a file format, a rich API, and a player executable distributed by Apple Computer. Confusion aside, the file format is generally open and can contain a variety of media types. It even includes the notion of effects, expressed as *tweens* that capture the parameters of the effect.

MXF is like these, and adds file streaming, rich descriptive metadata support, traceable material ids and others. Above all, it is a public standard.

A Brief Technical Introduction

Although you will find the complete list of acronyms used in this paper defined at the end of the paper, some terms need a few more words. Here are a couple of key terms the reader needs to understand in the description of a system, a problem or some interaction.

Edit unit is a period of time equal to 1/(edit rate), which usally turns out to be one frame. It also names the collection essence elements that comprise one frame.

KLV, or Key-Length-Value is the most basic storage unit in an MXF file. The key is a 16-byte value taken from one of the dictionaries in the SMPTE standards. The L is an encoding of the length of the value field. The encoding can be one or several bytes, whatever is necessary to express the length with enough bits. Each of the metadata sets in the structural metadata is a KLV, where the V is the value. In the essence container, the V is the essence payload.

Essence – MXF uses the term *essence* where everyone else in the world uses *media*; you will find both terms here, but they mean the same thing. Essence is used when the point of interest is MXF-specific.

Frame-wrapped – When each edit unit or frame occupies a single KLV, the V holds that single frame of essence. The term for this is frame-wrapping.

Clip-wrapped – When all edit units from a single track of essence is wrapped in a single KLV, it is said to be clip-wrapped.

CBE – Constant bytes per element. This is in contrast to VBE (variable bytes per element) and describes the character of the essence in the files. That is, every edit unit in a CBE stream is a constant size. A DV stream is CBE because every frame of 25 hz DV25 is 144,000 bytes long. VBE elements can be made CBE by adding a filler pack to pad to a constant boundary. Determining whether essence is CBE is a factor for file layout. Such files can be simpler because the number of partitions can be minimized to two or three and the index table becomes a tiny fixed length set.

GC – generic container. This is also called "essence container" or just "container." Logically, the GC holds all the essence in a movie. If it frame-wrapped it may be multiplexed over one or more partitions, and if it is clip-wrapped It is one of three possible parts of a partition, as illustrated in figure 3. A key decision when creating an MXF file is the location of the GC (internal or external) and how it is split among partitions.

KAG – KLV alignment grid. This was introduced to address concerns on alignment of important file elements. A KAG value of 0 or 1 means the grid is undefined. Any other positive value requires that the MXF encoder align the start and end of a partition on a KAG multiple. The encoder is allowed it to align the index table segment and edit units.

RIP – The *R*andom *I*ndex *P*ack is an optional table that is placed in the final bytes of the file. The table provides offsets to the first byte of each partition plus an indication about which GC is in that partition. When a GC is multiplexed into two or more body partition, this set usually becomes an important component for random access. Without it, the location of body and footer partitions can only be located by parsing all KLV elements prior that that.

SMD – The SMD acronym is not defined in MXF, but I use this to replace the term *S*tructural *M*eta*D*ata. It has come to be common in the MXF discussion groups. SMD defines the capabilities and construction of the file. It is also a term that contrasts with descriptive metadata.

VBE – Variable bytes per element. This kind of element varies in size from one edit unit to the next. Even if the size of only a single element in a GC is different, the stream of elements is considered VBE. The most common VBE stream is MPEG video essence.

W25 – There are several committees in SMPTE with different tasks. W25 is the label applied to the SMPTE committee on Metadata and Wrapper Technology. That committee is the one largely responsible for guiding the development of MXF.

On Brevity

Web addresses appear without the preceding http://. This paper is about MXF as defined by a set of SMPTE standards. Those documents are usually cited as SMPTE 377M or SMPTE RP224, but for this paper, SMPTE is omitted. In this context 377M and RP224 is sufficient. Acronyms are heavily used and care has been taken to define each in a section at the end.

Solutions

There are three solutions to interoperability: education, analysis tools, and cooperative refinement. Education is in papers like this, good books like The MXF Book, the SMPTE standards themselves, and especially the refinements to those standards that are making their way through the SMPTE process.

Analysis tools have been produced by every organization that implements an MXF codec. Some of the best tools are available free or for purchase, and are described at the end of the paper.

The paper also describes work sponsored by Turner Broadcast Systems. The product of the work is commonly called an AS (application specification). The Turner-AS selects a proper subset of the MXF standard, and provides additional guidance on file constructs. This allows Turner to assemble its media assets in a standard file format, and simultaneously allows vendors enough breadth that the implementation does not become user-specific.

Types of interoperability

In all cases, interoperability is achieved through education and information exchange. Early on in MXF development, the exchange occurred among a few early adopters and their organizations. More recently, the group of serious participants has grown to include several companies, organizations, consultants and a large number of end users.

The greater exposure has brought in lots of clever people, each finding some salient point in the standards that either is incorrect, conflicts with other details, or is vague enough for conflicting interpretations.

An influential contributor to this exposure is the IRT, a European R&D institute. It acts as a clearinghouse for sample files, and it provides file analyzers to test for compliance. These play a crucial role for vendors and users by being a central point of convergence.

Based on this activity over the past few years, there are a number of ways to classify how and where interoperability problems arise.

Standards compliance – the majority of the MXF file format details is in 377M, with other details, such as essence containers and descriptive metadata in a few dozen other SMPTE documents. Given a file of any particular structure, many of these documents are needed to create a compliant file. Pulling these details together is a difficult task.

Essence mix – One MXF decoder may work fine with one DV track and a two channel audio tracks, but the same decoder may stumble on two tracks of single-channel audio. Similarly, handling long-GOP MPEG is very different from handling DV. Validating all possible input is important.

Container elements, wrapping and indexing – With a frame-wrapped essence container, each essence element is wrapped in its own KLV. There are constraints on the value of K and the order of each KLV that have never been clearly specified. This will be clarified in future 377M editions. Index tables are an optional part of MXF, but are typically important when used for high-speed random access and identifying MPEG picture coding types. In exactly this usage, the possible variations on the indexing theme are daunting. With both essence-wrapping and indexing, there are many ways to say the same thing. Something is bound to go wrong.

Operational pattern – There is some misplaced belief that citing op1a or op3c captures all the relevant features of a file (see section 7.3 in 377M on generalized operational patterns). In reality, the OP only expresses one dimension of a file. The vast majority of files that exist today are variations of op1a. Higher OPs allow cut edits and optional tracks. When they begin to appear in large numbers, effort will be needed to interoperate.

File layout policies – Even when a file is compliant with MXF, there are a number of choices that an MXF encoder makes that can disorient a foreign decoder. The policies that have the most impact include partition size, placement or presence of the index table, selection of the indexing method, repetition of the structural metadata, and presence of the RIP.

Descriptive metadata – there is significant work among a number of groups to define usable metadata schemes, ways for multiple schemes to be present in a single file, and possible semantics for descriptive metadata that affect playout; e.g., tagging an audio track with the language in it.

Parse-while-write – This can be characterized by an existing MXF file being copied to a new location, and the destination file parsed before it has fully arrived. The file arrives in strictly left-to-right order. For example, many customers have MXF files stored on an archive that are copied to another server for editing or playout. A parsing problem is introduced when editing or playout begins before the transfer is complete. Normally, an MXF parser starts by reading the RIP at the end of the file. If unavailable, the structure of the file may require an alternate parsing strategy, one that could be expensive.

Parse-while-record – This issue is different than *parse-while-write* in subtle, important ways. The primary difference is that the source file is actively being recorded, and not written strictly left-to right. If such a file were to be copied blindly, the read pointer may pass a portion of the source file that is not in its final state. In some cases, it is possible for the recorder to write a file left-to-right, but not all. If not, the parser of that file must know enough about MXF to handle a changing file. In addition, the MXF recorder can make changes to the way a file is written to ease the burden on downstream parsers.

The sections that follow treat these areas in detail. This may help the reader to identify, predict or fix problems that condense to an issue of interoperability.

Standards Compliance

The list of MXF standards at the end of this paper makes it clear that compliance is not easy. There is a great volume of material that must be digested by the implementer of an MXF codec. Accurate specs are misinterpreted, and poorly written ones spawn multiple interpretations. There is not much of a solution for the great many documents unless you start from scratch. But to add clarity and remove ambiguity, the W25 committee continues to work through all the known issues to produce better guidance. Most SMPTE participants think the planned revisions of 377M and others will greatly improve the standard.

To help refine the interpretation of the specs, the IRT has contributed significantly. It has collected sample MXF files from vendors so that others in the community can easily obtain the various flavors of files. They have also developed an MXF file analyzer that looks for strict compliance for proper values, structure, presence of required values, proper universal labels listed in RP224, and many others.

The collection of sample files at the IRT is divided into four groups: golden, pathological, manufacturer's and custom. I have often kidded the folks at the IRT because they have only four files in the "pathological" bucket, and several dozen files labeled "golden" that are more pathological than golden. In spite of the name, the golden files present many of the possible combinations of values, or lack thereof, in the header. The combination of files presents a worthwhile obstacle course for any decoder. The collection underscores the many ways to write the same file and can reveal design assumptions in any new decoder.

The IRT also provides a free analyzer that, while not pretty to look at, provides an honest and highly detailed breakdown of the structural metadata. Recently, they added a professional version of the analyzer that is expensive but does far more detailed review of metadata, partition multiplexing, index tables and essence containers. The pro version is distributed as a DLL. This form is valuable because it can be tightly integrated into an operational environment.

For every vendor whose decoder can tolerate the IRT's collection of files, and whose encoder passes muster with the IRT's analyzer, it is likely that the codec will interoperate well other vendors that have made the same effort.

By far the most frequent interoperability failure is caused by MXF encoders creating files that are mostly correct, but not quite. Sometimes the SMPTE documents were not quite clear enough, and early implementations adopted different interpretations. Some examples:

- For containers with CBE essence, the index table segment may have a single edit unit size. This should be the fixed distance in bytes from one edit unit to the next. A common error just gets the number wrong. Although it is a number that can be discovered and corrected by scanning the essence container, it can trip a decoder.
- The generic picture essence descriptor contains a *stored height* that was interpreted two different ways: some MXF encoders stored the field height if the material was interlaced, and others always stored the full picture height, interlaced or not. Both types of encoders exist today in popular products. There is no general agreement as to how to solve this yet, because it is difficult to both correct the spec and avoid suddenly causing many files to become non-compliant. Fortunately, if there is an underlying standard for the video signal (e.g. 625-line 25hz), a decoder can decide which interpretation has been applied. That is, if $2*H_{picture} <= H_{standard}$, then $2*H$ is the correct height.
- 377M classifies structural metadata values four ways: required, optional, decoder-required, and best-effort. One of many examples is in the CDCI picture essence descriptor. The optional *vertical subsampling* is used with another value to determine whether the video essence is 4:2:0, 4:2:2, etc. Omitting this value when the essence is MPEG video might work if the decoder is willing to get the correct value from the MPEG essence sequence header, but when the essence is uncompressed YCrCb, there is no header to provide that information, and in fact, the essence cannot be decoded without guessing at the vertical subsampling value.

The same sort of near miss can occur in decoders that have built-in assumptions about the structure, content, or specific values in an MXF file.

- One of the first devices that injected many MXF files into the world is Sony's EVTR. An observer that sampled these files might conclude that the size of the header metadata in such a file is always the same. That is true in most cases, but not all. There is no guarantee for this in the SMPTE documents. Yet I have seen decoders that make this assumption and are later surprised by the presence of descriptive metadata, which changes the size of the header.
- A SourceClip duration (see 377M, annex B.10) can be the correct duration or -1. This is a *distinguished value*, which signals a decoder to discover the duration some other way. Sometimes decoders are written in the presence of MXF files with only correct durations, and fall apart on first contact with duration of -1.

All of these compliance issues can be addressed. Ambiguity will be greatly reduced by future SMPTE standards. Incorrect encodings can be fixed by vendors and customers scrutinizing the files they create with the available analysis tools. Decoders can be improved by exposing them to the files in the IRT repository.

The remaining sections look beyond standards compliance and syntax, into an area beyond merely correct MXF parsing.

Essence Mix

Only two SMPTE standards dictate a certain mix of essence types: 386M calls for a constrained MPEG video and AES-3 audio; 383M deals with DV, which in itself defines compressed video, and PCM audio.

MXF allows any combination of essence, but it is certain that no MXF encoder or decoder handles all possible combinations. There are two ways to address interoperability here. One is to test beyond the system limits to see what happens. If the MXF codec works well with four tracks of audio, try sixteen; if 3 MBit/second MPEG works, try 80 MBit/second. Tests of this sort expose systemic assumptions.

A variation of this is to introduce previously unknown essence. 436M is only about a year old and describes a way to carry VBI. The carriage of JPEG-2000 is now described in 422M. I always find it entertaining to point an MXF decoder at a file with essence that is unknown to the parser.

MXF codecs can minimize problems if the operational environment adopts a specific set of essence types. Many customers of ours already do this for process reasons. Later in this paper, I present broader goals for adopting proper subsets of MXF.

Container Elements, Wrapping and Indexing

If something is *wrapped* then there exists a KLV where the V is the thing wrapped. Wrapping is described in section 5 of 379M, but not very well in the current edition. Even so, I suggest reading it. The essence container can be frame-wrapped, clip-wrapped, raw or custom. Frame- and clip-wrapping are described above. Custom wrapping remains vague, but might be characterized as any wrapping the does not fit the frame- or clip-wrap rules. Fortunately, custom wrapping does not appear to be in common use. Finally, if the essence is not wrapped, then it is raw.

While structural metadata does not span partitions, essence containers frequently do. Index tables segments, too, can appear in one partition or among several. Whether and how the container is distributed among partitions depends on the wrapping type. The possibilities with wrapping and indexing present a large multi-dimensional grid, which might be sized along the following edges:

1. The GC can be internal or external. If it is internal, it can exist in a single partition or multiple partitions.
2. The essence stream can be frame-wrapped, clip-wrapped, raw or custom. If frame-wrapped, the individual elements can be padded or not.
3. Indexing is optional, but if essence elements are frame-wrapped, any, all or none of the elements may be indexed.
4. The indexing an individual element can be done with one of several methods: with a delta offset, a stream offset or a slice offset. And if the essence is CBE, indexing can be done with a single value called *edit unit size*, in the index table.

There are good reasons for exercising all of the dimensions listed above. For example, if your goal is to always have a single file for each movie, frame-wrapping into an Op1a file is a common thing. If you have one essence stream per file, CBE essence like audio is efficiently stored in clip-wrapped containers; if your MXF file is only intended as a transmission format that is never directly played, omitting the index table may be useful.

There is no end to the reasons for different strategies, and consequently, no way for decoders to avoid testing as many combinations as are possible. There are, however, more dimensions to the grid than listed above. Over the last year, the W25 committee has identified several details in the specs whose presence was innocent but whose effect was complex. For example, current specs allow a clip-wrapped GC with multiple streams. Each clip-wrapped track appears in a separate KLV, but all belong to the same stream. Combine this with the indexing mechanism that uses a 32-bit stream offset for all but the first element. This implies indexing such a container is difficult or impossible, if not unexpected. The new spec warns of these shortcomings and recommend against multi-essence clip-wrapping.

Another provision allows frame-wrapped elements in a generic container to change in their order and presence. This too can make indexing impossible and add complexity to every decoder that tries to handle the case. A future edition of 379M clarifies the rules to guarantee presence and order for essence elements for the full GC duration.

Another set of rules makes it clear that index table segments should appear in sequence, but is unclear about whether gaps are allowed, the intended usage for so-called sparse index tables, and the rules governing repetitions of index segments.

The solution to interoperability in GCs and indexing is two fold. The first is for encoders to avoid the untrodden path. The well-known methods include frame-wrapping and mono-essence clip-wrapping. With indexing, there are examples of just about all combinations of constructing an index table. Nevertheless, the common use of indexing is a single set of index segments appearing in order, without gaps. I have yet to see an Op1b file with two internal essence streams and two index streams.

The second part of interoperability here is for users to select a subset that meets their needs and work with vendors to make it formal. Users that want to store assets as a single file should claim that Op1a is one measure of their file requirements. Others that want to allow editing on material should specify Op1b and external essence. For index tables, there appears to be a tacit agreement among vendors that a single index table is included, no copies, in sequence and no gaps.

There is more on the subset theme below.

Operational Pattern

The Operational Pattern is a concept described in chapter 7 of 377M, and makes a well-defined grid of constraints in two dimensions: item complexity and package complexity. Studying that chapter is difficult but worthwhile. A simplified version of this grid appears in figure 1. Note that MP is Material Package and SP is Source Package.

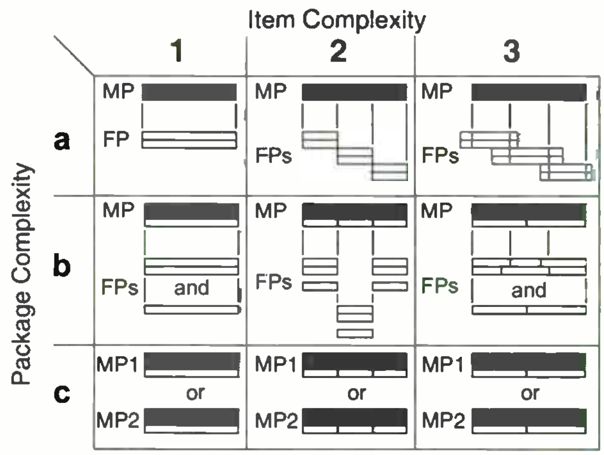


Figure 1: The Operational Patterns

The vast majority of MXF files are Op1a which means that there is a single package (the thing that holds the essence), and a single item (the number of edits in the clip). A highly constrained version of an Op1a file is specified in 390M, known as OpAtom, and is in common use as well.

An Op1b file is a distant second in popularity, and moves one notch down the package complexity to allow multiple packages. This means that different tracks can appear in different packages.

That much is all you really need to know. The rest of the operational patterns merely add edits and more packages. Note that the illustration of Op1a and Op1b shows no edit, making it a simple movie with all tracks starting and ending together.

Although MXF does not specify it, I believe that the selection of OP strongly suggests whether essence is internal or external, and to some extent, the method of wrapping. Consider this: Op1a contains a single package and a single clip. It is certainly allowed to put the essence outside the file, but because there is only one package, the external file would end up either being raw or as another Op1a file. What is the value of an Op1a file that points to another Op1a file? In contrast, the convenience of putting it inside the Op1a file is very compelling. Witness the plethora of Op1a files. With internal essence you can either clip-wrap or frame-wrap a single track. With multiple essence tracks, you are forced to frame-wrap.

Now consider Op1b. This allows multiple packages. You could put the essence inside the file, but the governing rules for partitions and essence containers and index tables make this an awkward choice. Each package must go in a separate essence container; each essence container may be indexed by an index stream; each partition may have one essence container and/or segments from one index stream. The structure of such a file with just two packages is shown in figure 2. Potentially, the indexes and GCs could occupy four large regions. To play it would require reading all four regions of the file in parallel. This has no locality of reference for all the data for each frame. Few file systems today can do a good job at preventing the file system cache from thrashing with this usage. To make matters worse, recording a movie of unknown length would continue this pattern of alternating partitions.

Figure 2: Op1b with internal essence

If you put these packages in external files, things get much simpler. The external files can be raw and the index tables can be multiplexed among the partitions. An alternative is to put the external essence in Op1a files, which can have their own index table, making the referencing Op1b file very simple.

These facts make a strong argument that Op1a files have internal essence and Op1b files have external essence. Stated from a package view, a single package ought to go inside an MXF file, and multiple packages should all go outside the file.

The other operational patterns are in use but not widely. This suggests that the best interoperability is achieved with Op1a, OpAtom and Op1b.

File Layout

The MXF specifications are pretty good at spelling out the syntax for every KLV, and good at the next level up, for the contents of a partition header. At a higher level where these components are combined into a file, there is a great deal of latitude. Do not be surprised that engineers have exercised that freedom.

There are rules for the composition of index table segments, for header, body and footer partitions, and for RIP tables. Other than rudimentary rules that the order in a partition is header followed by index, followed by GC, there is no guidance for sizes of each or relative placement.

For CBE material such as DV, partition layout is straightforward. The index table needs no index entry array, so files can be written left to right. When creating an Op1a file, this has led every encoder I know to choose a structure with a header, footer, and all the essence in the header. Sometimes the essence is placed in a single body partition, but there ends the variation. This is a good thing.

With VBE material, such as MPEG video, several applications require an index table with an index entry array to describe the location and frame type for every frame. From an MXF encoder view, recording an Op1a file of this sort requires simultaneously updating two different places in the same file: the essence container and the index table. The file layout for this depends on choices along three different vectors:

Selecting a partition size. If a GC is composed entirely of CBE essence, its possible to create an non-growing, simple index table, and then allow the GC to grow unbounded. But with VBE essence, the encoder must select a partition size. A partition is measured in bytes, but the choice is more often selected by the number of seconds or frames of essence in each partition. Implementations vary between a few seconds to several minutes.

Selecting a policy for the distribution of index table segments. Through a quirk in the MXF specs, an index segment can be arbitrarily large, but the index array inside it is limited to 64K bytes. In practical terms, a single index entry array can address at most about four minutes of VBE essence; less as the frame rate and the number of tracks increase. A long movie will require several index segments and the encoder must decide where to place the segments. There is no guidance in 377M for index distribution, or placement relative to essence indexed.

Fortunately, most implementations distill to two: either all index segments appear in the footer, or they are distributed among the partitions. When distributed, each partition will hold a collection of index segments, and that collection will index the essence of a single partition. That is, if a partition's GC contains ten minutes of essence, then the corresponding set of index table segments will address the frames in that ten minute span. It is good that this has been the natural choice in so many implementations.

Selecting a placement of index tables relative to the essence. When index segments for VBE essence are distributed, the group of segments addressing a single partition is placed (a) in the same partition as the indexed essence, (b) in the partition following the partition with the indexed essence, or (c) in separate partitions not shared with any essence.

Illustrating some of the possibilities is enlightening. The simplest form of indexing for VBE essence is show in figure 3 where the footer contains the entire index table and the essence container is in the body partition (the header partition is not shown). This layout is required in OpAtom, 390M. Assuming the movie is long, it requires that the MXF encoder either know the duration of the movie before beginning to encode it, or must buffer the index table segments until the GC is complete. If, on the other hand the essence is CBE, the index in the footer collapses to a single segment without index entry array.

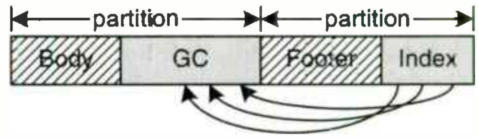


Figure 3: a self-contained partition with backward indexing for VBE essence

Single-use partitions are shown in figure 4. This is a forward-looking layout, in anticipation of a proposed improvement in 377M. The RIP contains an array of file offsets and body stream ids. By making the partitions contain only either index or GC, a body stream id of zero could imply that the partition contains an index. If an initial goal of random access were to find all index tables, this would help the decoder avoid partitions without index tables.

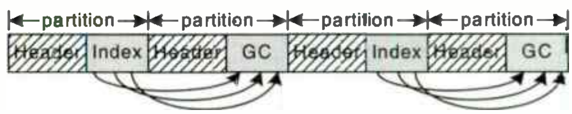


Figure 4: inter-partition, forward indexing in single-use partitions

A variation of this layout with distributed index tables appears in figure 5. Here, the partitions contain both index and GC, except for the first, and index segments point to essence in the previous partition.

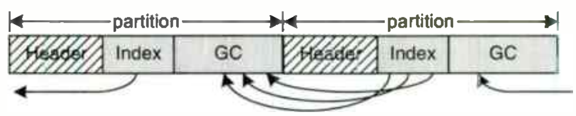


Figure 5: backward indexing

This presentation of file layout, driven by essence type, index and GC placement, should underscore that MXF allows many possibilities. Some of these are good because they play well into certain applications. However, there many are other possible combinations, some of which could be considered pathological.

For example, index tables could be distributed among partitions without regard to GC placement, as shown in figure 4. It is also possible to have multiple index tables (with different stream IDs) that index the same essence. MXF also describes the notion of sparse index tables (each segment containing a single index entry), but the meaning of what each index entry designates is undefined. They could indicate a partition start, an edit point or a scene change. If a meaning is adopted, it must be universal.

Systems do not interoperate well in the presence of the unexpected. Only Op1a has been discussed here, but it should be clear that while the possible file layouts are many, fewer are better. Exercising all layouts creates more validation tests for MXF decoders and less reliability for customers.

Again, the solution is to adopt well-defined subsets of the MXF standard; not with the goal of corseting an MXF encoder, but rather to prevent the unexpected.

Descriptive Metadata

There are three areas where DM (descriptive metadata) may cause, or suffer from interoperability problems.

Lack of space in the header. A monolithic Op1a file frequently has a header partition followed by index tables and essence in the same or following partitions. What do you do when a client requests adding DM to a header that is followed by 20 Gbytes of essence?

Two solutions comply with MXF. The simplest is preventative. During encoding, the encoder can add a large filler pack in the header partition after the SMD, and before the index table and GC. Within constraints set by the KAG, space for adding metadata can be absorbed from the filler pack without rewriting the rest of the file.

Where practical, this is a simple solution, but it is not sufficient, because it is still possible to run out of space. A second solution is a complex multi-step process. (a) Mark the header as open, and (b) copy or update the SMD in the footer and mark it closed. For many decoders today, this is too subtle a move.

Multiple Descriptive Metadata schemes (DMS) in one file. There is a well-defined mechanism for identifying which DMS are in use in the header, but the metadata values themselves are not marked in a way that associates them with one DMS or another. This means for example, there is no way to identify or remove all DM associated with a single DMS. There is some dialog in committee meetings about this issue, but standardization will take time.

Structural semantics for descriptive metadata – MXF does not include any specification for marking audio channels in a way that associates a language with a given track. The OpZero spec, described in *Selecting a Proper Subset* uses descriptive metadata to do this with a DMS-1 label that indicates the primary spoken language.

The first issue is solvable with the details outlined above. The latter two are too new for any consensus to be clear, so a simple solution is more difficult. It may be sufficient to write an AS to either define the semantics or prevent the usage.

Parse-While-Write

In the professional video environment, MXF files may contain minutes to hours of material that ranges from 12 MBit/sec up to 100 MBit/sec or more. For broadcasters, these assets move among off-line, near-line and on-line storage. The favorite transport for moving files of this sort appears to be the age-old FTP protocol, while distributed file systems like NFS, AFP and SMB are a second choice. The reason is simply bandwidth: the FTP protocol, once initial handshakes are complete, can usually attain speeds that are near the limit of the underlying hardware. Distributed file systems, while more generally useful, use a protocol with higher overhead and thus lower bandwidth. In any case, most transfers are done by clients and servers that know nothing about the type of data being transferred.

This introduces the concept of parse-while-write, which is subtly and substantively different from parse-while-record, treated in a section below. I limit the scope of this subject to MXF files where:

- The source file in the transfer already exists in its entirety (or the transfer agent makes it seem so).
- The movie structure is Op1a with internal, VBE essence, and a complete index table.
- The transfer is strictly left-to-right, in the same way that FTP would deliver a file
- The file is large enough that it cannot be fully transferred for a long time.
- The destination copy is being parsed by an MXF decoder before the transfer is complete.

An Op1b movie or any higher-level operation pattern that points to external essence files is not an issue here. Having no essence, the file is likely to be so small it can be transmitted in milliseconds. Now, it is possible that the Op1b's external essence files are Op1a, and that is the same problem as a single Op1a movie.

The problem set is further trimmed by eliminating CBE material because an index table for CBE can be constructed with a fixed edit unit size and no index entries. Although not specified in MXF, most encoders place the essence for this type of file in a single partition. Assuming this, a decoder can play every frame and perform random access, all the way up to the last available byte, with simple arithmetic.

When an MXF decoder encounters an Op1a file with VBE essence and that has not fully arrived on disk, the implementation must address several issues, yet those details are well outside MXF standards.

- The MXF file format contains no forward pointers. While there are pointers in each body and footer to the previous partition offset, there are no such pointers to the next partition. The RIP is at the end of the file, and it is unavailable in this situation. Without forward pointers or RIP, an index table or partition beyond the header partition can only be discovered by parsing the file from left-to-right. And if the index table is unavailable for the essence in any partition, the essence KLV must be parsed left-to-right to find each edit unit. A decoder must be prepared to read a great many bytes to accomplish what could be done easily with a RIP.
- If the essence is clip-wrapped, the MXF rules state that the entire essence resides in a single KLV. Although possible, a rational MXF encoder will not create more than two or three partitions. You can conclude that at the KLV level, the file can be parsed quickly, because the GC consists of a single, huge KLV. However, with VBE essence the decoder may depend on the index entry array to find the individual frames. If the index table segment is in the footer and not yet available, the decoder will need an alternate plan.
- If the VBE essence container is frame-wrapped, getting to the second and following partitions requires that every essence element in every edit unit must be parsed. In addition, if the index entry array appears in a partition following the essence, it may not be available. This is a particular problem for MPEG streams, because the index entry array may contain frame types needed by the decoder. If so the decoder will require an alternate strategy for the non-indexed part of the essence.
- Given VBE essence and no RIP, random access is possible, but frequently too

expensive to be feasible. To do it, you could jump to a random location and then begin a search for one of several expected ULs. When found, the parser must continue searching forward for a partition header with essence, which will provide the stream offset of the GC. Finally, an index table must then be found that can correlate the stream offset with an edit unit.

Circumventing interoperability in this case requires that an MXF decoder implementation observe the issues above. This is also a prelude to the next issue, a refinement of this one.

Parse-while-record

This subject is distinguished from the last by guidance for the agent recording the file and for the parser reading the file.

It is possible that the recorder can write the file left to right, thereby reducing the problem to parse-while-write. This is most possible with CBE, because a simple index table segment without an index entry array can be used. Sometimes, VBE can reasonably be crafted into CBE. For example, PCM and DV alone are both CBE, but when the frame rate is 29.97 Hz and DV and PCM appear in the same GC, the number of audio samples per frame oscillates in a 5-frame cycle, making it VBE. Padding each PCM element to a fixed boundary restores its CBE-ness.

Although MPEG is usually considered VBE, intra-coded MPEG like that described in 356M, can be made into CBE. Such essence usually varies little in size, and depending on the encoder, may have a well-known maximum size for a given bit rate. Padding these frames out to a fixed maximum size makes these CBE. 386M does exactly that.

The difference between parse-while-write and parse-while-record is introduced when more than one region of a file must be updated by the encoder. The reasons for this appear below. These regions are illustrated in Figure 6 as part of a single partition, although each region could reside in different partitions. There are several things to note about how the file may be updated by a recorder over time, as seen by the downstream parser.

- The sequence and source clip are in the header and may have changing duration values.
- The index segment contains an index entry array that grows with the addition of every frame.
- The GC grows left-to-right with the addition of every frame.
- The footer and RIP must be relocated and partition offsets updated whenever the GC overruns the footer.

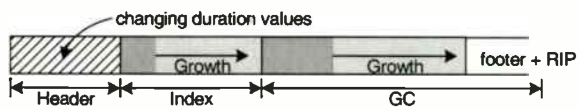


Figure 6: possible regions of change

If essence is CBE, then strict left-to-right is within reach. First, the index table collapses to a fixed-size set. Next, the header duration values can be set to -1. Then the index table duration may be set to zero, a value allowed in the case of CBE to indicate "the entire container." Finally, writing the footer and the RIP can be delayed until the recording is complete. Once these guidelines are followed, a reader would not be able to discern the written file from the parse-while-write scenario.

There are several reasons that an encoder might not be able to render a file strictly left to right. These assume that simplifying encoder design, minimizing resource usage and system performance is paramount. These concerns usually push an implementation to minimize buffering, reduce temporary storage and avoid multiple copies of data. These goals may lead to the reasons that require multiple regions outlined above:

- First, assume the essence is VBE and requires an index entry array.
- Choosing a small GC partition might allow a recorder to buffer the index table

for that part of the container and then write it into the next partition. Small partitions might also allow random access when a RIP is unavailable as described under parse-while-write. However, small partitions in MXF make random access more expensive because a RIP does not identify which edit units reside in a partition. Random access, then, calls for some amount of trial and error. A large partition improves the random access performance but suggests that the corresponding index table cannot be efficiently held for later writing. So a choice for random access performance drives a large partition size, which may make index buffering out of the question.

- If the goal of the written movie is to maximize random access for a parser during the recording, a RIP must be available. This introduces a number of problems for left-to-right writing. If the duration of the recording is not known, the correct size for the GC or index table cannot be chosen when it begins. Since the index table must be available for random access, the GC, the index, the footer and the RIP must be written in parallel.

Where an encoder cannot write a file left-to-right, it should follow some guidelines that minimize issues for downstream parsers. I call these smart-writers.

- The encoder should make some effort to maintain coherency in all regions of the file. This includes adding filler packs after incomplete index tables and GCs.

- Once a file is no longer left-to-right, maintaining the footer and RIP is an aid to parsers during the recording. Note that writing a footer and RIP immediately after the last recorded frame is not effective if it is will immediately be overwritten by another frame. It is a good idea to pick a location where the footer and RIP can remain for at least a few seconds.

- If the duration of the recording is unknown when it begins, the header cannot be written with final duration values (sequence and source clip). Write these with the *distinguished* value -1, meaning unknown. This relies on the decoder to either reconstruct the correct value or look for a closed footer with complete SMD.

- If the SMD is located in the header, a partial left-to-right transfer, described below, can be performed if the header is marked open, durations are written as -1, and the final SMD is written to the closed footer.

- The smart-writer should consider placing the index table for VBE essence after the GC. Since the duration for an index segment does not have a distinguished value and must be accurate, readers could resample this value from time to time to discover the valid extent of the GC preceding it.

It is easy to see that the reader of such a file must be more intelligent than an FTP server. Such readers could be divided into two groups; call them MXF-aware and smart readers.

A reader that is MXF-aware need only discover the extent of the file that is no longer changing. If the guidelines for smart-writers are followed, an MXF-aware reader could easily identify the last partitions containing index and GC and avoid copying until more partitions are present or the recording is complete. If the index table follows the GC, this reader could improve this by resampling the index duration and index entry array and copy the GC elements that are referenced by index entries.

A smart-reader goes well beyond this and assumes nothing more than a coherent file. This reader re-reads portions of the file as needed, much the same way a demand-paged memory system retrieves pages only when needed. The value of this reader is in performance and the ability to follow closely behind the recorder. The typical use for an agent of this source is a player that shares only a file system with the recorder.

This section goes well beyond MXF in its definition and application. Fortunately, some of my colleagues in the SMPTE community agree with these guidelines for MXF-aware readers, smart-readers and smart-writers.

Selecting a Proper Subset

The preceding topics should be convincing arguments that interoperability can be partly solved with improved specs, education and good analysis tools. The remaining problems lie in details that are above the MXF specification: the selection of operational pattern, file layout, restraint in the selection of essence and options, and cooperation during file transfer.

Some of the best interoperability work I have seen is an effort sponsored by Turner Broadcasting to produce an end-to-end workflow that specifies everything from the way material makes its way through production, all the way down to the details of the MXF file. For the latter, significant contributions came from a number for companies, including Snell & Wilcox, Metaglue and Omneon. Two of the documents in this AS call out specific features of MXF and describe how and when they are used. For details that require specific numbers, such as partition size and number of audio channels, these are parameterized and the value of the parameter is placed in a Turner-specific annex in the document.

The documents, informally called OpZero, emphasize a particular workflow that demands several things:

- Essence is stored in individual files to aid in editing and reassembly.
- Every file is MXF wrapped to leverage the detail that can be stored in the metadata and the traceability of the UMIDs.
- General file rules above those specified by MXF are defined to provide decoder predictability, defined semantics for language tagging, partition sizes, location of the index tables and the selection of essence types.

Efforts like this are a part of what will make MXF successful as a professional file format. An accompanying part must also come from end users that subscribe to the standards and to a few well-defined subsets like this.

The Future...

The SMPTE specifications are not controlled by a single company and are being refined among some excellent contributors from a variety of sources. The collection of tools, organizations, consultants and sample files continues to grow. This alone provides momentum toward a well defined MXF.

I continue to find decoders that cannot handle certain MXF files created by Omneon, as well as files that the Omneon decoder cannot understand, in whole or in part. In general, I always find a willingness among vendors that have a different MXF codec to compromise on the basis of the MXF standard itself. That is, nearly every problem can be reduced to an error in the encoder or the decoder.

In rare cases, I have encountered foreign files that have unexpected consequences in our decoder in the context of a real-time playout system. This requires revisiting first principles and refining goals for our MXF codec.

Analysis Tools

www.freemxf.org This is maintained by Matt Beard, a worthy contributor to the SMPTE standards, and developer for the open-source MXF implementation.

www.irt.de/mxf The IRT is the research and development institute of the public broadcasters of Germany, Austria and Switzerland. They have long made contributions to a variety of standards efforts. This is the home of the IRT Analyzer Light, the Analyzer Pro, many sample MXF files, testing services, and consultancy.

www.metaglue.com This company was founded in part by Oliver Morgan, an important contributor to the MXF standards. They provide consultancy and a number of MXF tools.

www.mog-solutions.com A small in Portugal, they still have time to make significant contributions to the MXF standards. Besides providing the underlying toolkit for the IRT Analyzers, they sell a number of useful MXF tools, training and consultancy.

www.opencube.fr This company has long been involved in the MXF community. They sell an MXF SDK, MXF tools and provide consultancy.

www.themxfbook.com This is a pointer to a book written by four very important contributors to MXF and SMPTE: Bruce Devlin, Jim Wilkinson, Phil Tudor, Matt Beard. It is a must-read for anyone looking to understand MXF.

Acronyms

AFP – Appletalk filing protocol.

AS – Application specification; a document that describes a selection of a subset of MXF and provides guidance on file construction.

CBE – Constant bytes per element; see 377M.

DLL – Dynamic-link library; this term is specific to the Windows OS.

DM – Descriptive metadata. To be symmetric with SMD, you'd think that this should be DMD, but DM is the usage typically coined in committee.

DMS-1 – Descriptive metadata data scheme 1; see 380M.

DV – Digital Video, defined in ISO/IEC 61834.

GC – generic container; also: essence container, although I've rarely seen an EC acronym.

IEC – The International Electrotechnical Commission at www.iec.ch)

FTP – File transfer protocol; defined by RFC 959 at www.ietf.org. Although it is really a protocol, it is commonly thought of a the application that performs FTP.

IRT – Institut für Rundfunktechnik at www.irt.de.

ISO – The International Organization for Standardization at www.iso.org.

KAG – KLV alignment grid; see 377M.

KLV – Key-length-value used extensively in 377M and defined in detail in 336M.

MPEG – Motion Picture Experts Group as defined in ISO/IEC 13818; www.mpeg.org is a good starting place.

MXF – Material Exchange Format; largely defined by 377M by SMPTE.

NFS – Network file system; Originally developed by Sun Microsystems.

OP – Operational Pattern; e.g., op1a.; see section 7.3 in 377M.

PCM – Pulse coded modulation; for the common man, this is just uncompressed audio.

PDF – Portable document format defined by adobe.com.

RIP – Random index pack. An important component for random access; see 377M.

SMB – Commonly called Samba, SMB stands for Server Message Block, and is a protocol for sharing files, printers, serial ports, and communications abstractions such as named pipes and mail slots between computers.

SMD – Structural metadata; generally all the detail in the MXF header that describes the tracks and their properties. See notes on DM.

SMPTE – Society of Motion Picture and Television Engineers at www.smpte.org.

UL – a 16-byte universal label; used as the K in every KLV and to mark nearly all properties of an MXF file.

UMID – universal material id; commonly 32 bytes long, although there are shorter and longer UMIDs possible; see 377M.

VBE – Variable bytes per element; see 377M.

W25 – The SMPTE committee on Metadata and Wrapper Technology.

The MXF Document Set

To make a point of the mass of MXF, if not the significance, it is worthwhile to list the SMPTE documents that comprise it in its entirety. The basic format is described in 377M. There is a long, informative description written about the same time in EG41. Many other documents provide details about operational patterns, essence, containers, metadata, and labels. They are:

EG42	386M	394M
378M	387M	405M
380M	388M	407M
381M	389M	408M
382M	390M	422M
383M	391M	436M
384M	392M	RP224
385M	393M	RP210

Other relevant documents provide underlying mechanisms that are used by MXF.

298M	ISO/IEC 11578-1:1998
336M	ISO/IEC 8824-1:1995
395M	ISO/IEC 8825-1:1995
400M	ISO/IEC 8879:1986
	ISO/IEC 9070:1991

New Tools for MXF-Analysis

Matthias Schnöll, Rolf Hedtke,
Fachhochschule Wiesbaden, University of Applied Sciences
Department Ingenieurwissenschaften,
Television Technology and Electronic Media
65195 Wiesbaden, Unter den Eichen 5,Germany

Abstract

MXF is widely accepted as the main exchange format for programme material in a networked studio environment. For monitoring the files and the essence parameters for video and audio new tools for monitoring are needed. In this paper an analyser tool is presented, which is based on the software development kit of the freemxf organisation. This analyser has the capability of offline and online analysis so it can be integrated in the networked management system. The MXF files are analysed according to the individual settings of the different operators and the results are displayed in an easily readable form, so even operators who are not familiar with MXF technology can use this tool.

Introduction

In a conventional SDI digital studio environment the signals are monitored via television screens or in particular cases also by error rate measurements. Monitoring and measurements are not as that easy in a networked environment. That is because on the one hand the number of different formats and compression methods has increased, and on the other hand a lot of IT technology is used: Besides the standard television formats with 4:3 and 16:9, HDTV formats with different resolutions and scanning formats have to be considered here. While in SDTV mainly DV and MPEG-2 compression formats were used, there are a lot more formats available in HDTV: MPEG-4/AVC (H264), JPEG2000, VC-1, VC-2, and VC-3.

The introduction of IT-technology has changed the infrastructure of a television studio at a large extend. With file transfers the transmission is guaranteed to be error-free by the network protocols itself. But the main aspect with networked studio environment is the change in the workflow. So in the whole production chain new workflows have to be introduced - within the studio as well as for programme exchange. Main goal is to increase the efficiency and creativity, which leads to lower operation cost in nearly all fields. To fulfil the demands of the professional networked TV-studio technology not only a new data exchange format but also an intensive use of metadata is necessary. The creation and handling of metadata, from ingest to distribution, is the key for the improvement of the workflow.

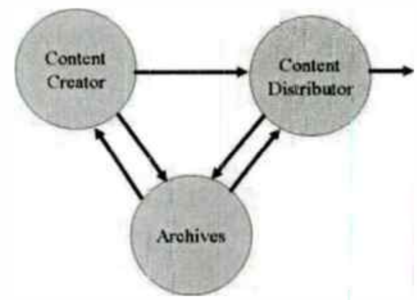


Figure - 1: Data model for programme exchange

The data model (*Figure-1*) shows three logical units where the program exchange takes place:

- Content creation
- Archives
- Content distribution

Within the archives not only the new programmes are stored, but also already existing material is re-used for creating new programme material. To ease the worldwide programme exchange, metadata are very important. They enable

- a clear identification of the material together with information about the history,
- a description of the processing steps,
- clarification of rights
- Description of the technical parameters to ease the exchange of the material.

MXF - Format

MXF („Material eXchange Format") is an open international standard published by the SMPTE. This is the prerequisite for the worldwide programme exchange of audio visual material.

In general there are two classes of metadata:

1. structural metadata, which are necessary for the *technical* description of the content,
2. Descriptive metadata which represent the user information to describe the content.

MXF is a fileformat, which is independent of the particular content, i.e. the coding format; however the standard must define how the data are to be mapped into the MXF container to achieve interoperability. For storage or handling of the data it is not necessary to decode the video or audio data because the metadata contain all information about the material.

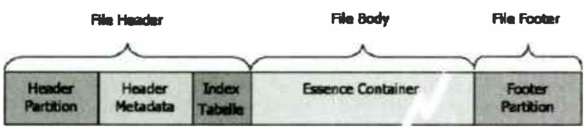


Figure - 2: MXF-Fileformat

The structure of a MXF-file is shown in *Figure-2*:

- *Header und Footer Partition* define the begin and end of the file
- *Header Partition Pack* contains the essential data about the file structure (coded in KLV)
- In the *Header Metadata* the structural and descriptive metadata are inserted
- *Index Tabeles* are optional but offer the advantage for the user to get access to parts of the file without the need to decode the whole MXF file.
- Inside the *Essence Container* are the video and audio data. MXF is independent of the coding format. This concept is very important because these data files can be handled inside the networked studio environment even if the essence is not decoded. The user can evaluate the metadata in the MXF Header to identify the content. So in general the material is completely identified by the metadata.

Besides in the header, metadata there can also be inserted into the body part of the MXF-file:

- Header Metadata are used for identification of the content and the scene description and are valid for the whole file.
- Dynamic metadata, which change during the time, may be coded into the body partition as a separate metadata track.

MXF-Analysis

In *Figure-3* the basic structure of a networked studio is shown. Each of the function blocks in *Figure-3* must offer the possibility to monitor and analyse the MXF datastreams.

Figure - 3: Structure of a networked studio

As mentioned above it is not always necessary to decode the audio and video sequences. In many cases it is sufficient to analyse the header metadata for getting the information about the content of the file. It is even not useful to decode all the metadata contained in the header.

At the different areas in the studio also different metadata are of interest. For playback video and audio, per example, the structural metadata which specifies the parameters for the decoder are if interests, while for the journalists or the rights management the descriptive metadata are of major importance. Therefore the analyser must be configured to the respective application within the studio environment.

Data-analyser

The MXF data-analyser presented in this paper is based on the open-source software development kit of the freemxf organisation. This SDK is a C^{++} lib which can be used with different operation systems. It is object oriented and offers classes for KLV encoding, for encoding physical elements of the MXF datastreams and header metadata. For the analyser the following tools were used:

MXFdump includes software for decoding and analysing MXF metadata. Here the MXF file is searched for metadata codes and these were listed in the order of their appearance. Only those keys which are listed in the dictionary are also decoded.

The *MXFSplit* tool de-packs the MXF file. and allows access to the original data.

An important part of the software is the *MXFdictionary.* This provides the keys for the KLV coding and the universal labels (UL). To each of these there is information available about the type of data in the appropriate value field. The XML parser includes the dictionary and the software can be easily adapted to the further development of the format or the application. The XML syntax shows the hierarchy of the different metadata sets. *Figure-4* shows an overview over the software elements used for the analyser.

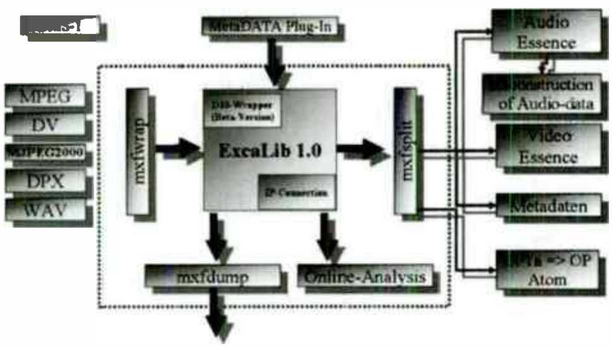


Figure - 4: Structure of the MXF software tools

Offline Analysis

Offline analysis is used to monitor MXF files which are already archived. These data are stored on a server, and the analysing tool tests these for their structure and consistency. The software modul *mxfdump* is used for testing the keys in the header as well as in the footer. The KLV packets are translated with the help of the MXF metadata dictionary of SMPTE RP210 and then stored in a so called "dump". The analyser software displays the content of the dump in a structural window so the user can easily look at the logical items he is interested in *Figure-5*. In the explorer window a MXF file can be selected for testing. Then the software analyses the actual header metadata and displays the decoded values in a window. For better visualization the information about the coding of video and audio is shown in different windows. In the video window the resolution of the picture and the aspect ratio is shown as well. For audio essence the sampling frequency and the channel assignment are the most important technical parameters which are displayed in the audio window.

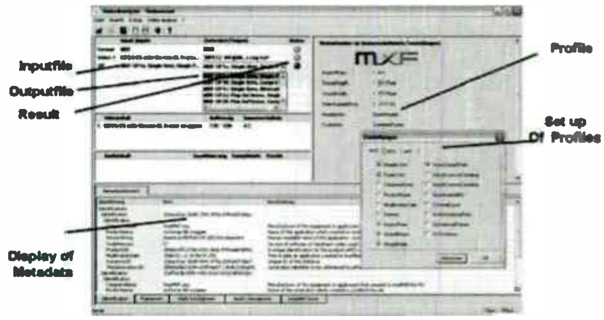


Figure - 5: Analyser main window

More information about the structure of the MXF file is given in the metadata window. Here all the structural metadata for video and audio can be seen on demand, and in addition the MXF file structure is shown here. An important feature for the user is the profile window. Here he can define which of the metadata shall be displayed. So for a very quick analysis the users at the different parts of the networked studio can choose a limited set of metadata for their personalized display. Another important function for a quick and easy file analysis is the comparison of the MXF file against a user defined profile. In this profile the operator can freely select a set of format parameters he expects the file should have. The file under test is checked against this profile and the result is shown in a form of a traffic light: green is correct and red is false. So without knowing the MXF format in detail even an unskilled person can see if the file fulfils the need of the studio requirements.

Online Analysis

In a networked studio environment the functionality of the software application must be enhanced by control and management functions. Within the analyser software new modules were developed to address this demand in LAN and WAN networks. These modules enable to control the analyser via remote workstations.

A client-server structure is used for this application. Before sending a MXF file from a remote location, the data is analysed and compared with the defined format requirements. A number of clients can be controlled and managed by this server function. There is a fixed IP and Port assignment of the client function, so a stable connection can be set up even in a large network. By this online analysis first the header and footer of the MXF file is extracted and send over the network - without the file body- to the analyser. The client structure starts the analysis of the header and sends back the results to the server MXF analyser. Only if the MXF header matches the defined parameters, the whole MXF file is transferred to the play-out server. This interaction is shown if *Figure-6.*

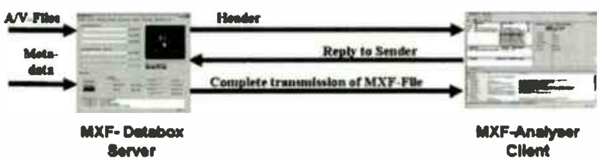


Figure - 6: Server-Client Structure

All analyser actions on the network are recorded into a log-file with details about date, time and IP address of the workstation which initiated the task. For increasing the transparency an additional tool was implemented which allows to search the log-file for specific actions or other items, like IP addresses.

In contrast to the offline analysis, where the user controls the analyser software, the online analysis is a measurement tool which needs no user interaction. The online analyser can automatically check a whole folder of a data archive and start the analysis of the MXF files automatically whenever a new file is stored. The result then is reported back to the remote site.

The *Remote Control* function is split into three parts where the different settings are displayed on the screen *(Figure - 7)* and the user can change them.

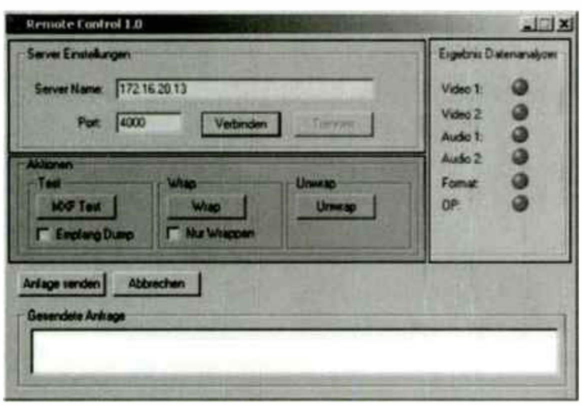


Figure -7: Remote control function

The first window contains the settings about the connection to the server software, the second lets the user choose the analyser action to be performed, and the third shows the result of the analysing tool. So the user at the client side has a very fast indication whether the selected MXF file matches the desired parameters or not.

Conclusions

After MXF is adopted as the main exchange format for programme material in a networked studio environment, new tools for monitoring the content of the data and the parameters have to be developed. Based on the software development kit of the freemxf organisation an analyser for offline and online analysis was developed, which can effectively test the MXF files and show the results in an easily readable form even for operators who are not familiar with MXF technology.

References

[1] John Ive: The Material eXchange Format and the Workflow Revolution
Pro-MPEG Forum Publication

[2] Katsumi Tahara, Hugo Gaggioni: MXF: Technology enabler for IT-based Broadcast Operations
Sony White Paper 2002

[3] SMPTE 336M-2000: Television-Data Encoding Protocol Using Key-Length-Value

[4] SMPTE RP210-2001: Metadata Dictionary Contents

[5] Pro-MPEG-Forum: www.pro-mpeg.org

[6] Bruce Devlin: MXF -the Material eXchange Format
EBU Technical Review No.291 Juli 2002

[7] Jim Wilkinson: An Overview of the Material Exchange Format (MXF) Specifications
Proceedings of the 143. SMPTE Technical Conference Nov.2001, S. 118 – 126

[8] Jim Wilkinson, Bruce Devlin: The Material Exchange Format (MXF) and its Applications
SMPTE Motion Imaging Sept. 2002, S. 378 –384

[9] Oliver Morgan: Metadata Systems Architecture
SMPTE Motion Imaging April 2003, S. 129 – 135

[10] Dennis Hartmann : Entwickeln eines MXF-Essencecontainers für Motion JPEG 2000 Daten
Diploma thesis at FH Wiesbaden University of applied sciences, 2004

[11] Jan Christoph Simon: Entwicklung eines MXF-Systems zur Einbindung von Metadaten
Diploma thesis at FH Wiesbaden University of applied sciences, 2005

[12] Patrick Bongart
Entwicklung von DPX-Tools für das MXF-Format

Diploma thesis at FH Wiesbaden University of applied sciences, 2005

[13] Matthias Schnöll
Einmal produzieren – mehrfach nutzen
Vortrag auf dem 3.Thüringer Medienseminar, Juni 2005

[14] Matthias Elser
Analyse von MXF-Dateien über eine Client Server Architektur
Diploma thesis at FH Wiesbaden University of applied sciences, 2006

[15] Jan Giese
Entwicklung von TCP/IP Control- und Management Modulen für die MXF Software
Diploma thesis at FH Wiesbaden University of applied sciences, 2006

[16] Matthias Schnöll, Rolf Hedtke
Die vernetzte Studioumgebung mit dem Austauschformat MXF
11. Dortmunder Fernsehseminar, Universität Dortmund, 26. September bis 28. September 2005, ITG Fachbericht 188

Broadcast Facility Safety and Compliance

Tuesday, April 17, 2007
1:00 PM – 2:00 PM

Chairperson: John Lyons
The Durst Organization, New York, NY

What Broadcast Tower Owners and Tenants Need to Know About 222-G

Thomas Hoenninger, Stainless LLC, North Wales, PA

ENG and All-Hazards Safety

Matt Bruskotter, The University of Findlay, Findlay, OH

What Broadcast Tower Owners and Tenants Need To Know About 222-G

Thomas Hoenninger
VP Operations and Chief Engineer, Stainless LLC
North Wales, PA, USA

What are the top 10 items that Broadcast Tower Owners and Tenants need to know about 222-G?

1. INTERNATIONAL BUILDING CODE (IBC) TOWER REQUIREMENTS

When tower construction requires a building permit most jurisdictions have adopted the International Building Code (IBC) for their design/construction requirements. The International Code Council (ICC), which maintains the IBC, publishes the code every three years with supplements being published in between the code releases. Currently 2000 IBC, 2003 IBC and 2006 IBC have been published. Once a code is published each jurisdiction has to approve the code before it can be adopted in that jurisdiction. There is normally a time lag between when a code is published and when it is adopted by jurisdictions. This time lag varies from jurisdiction to jurisdiction so some jurisdictions currently build to 2000 IBC, some to 2003 IBC and some to 2006 IBC. 2000 IBC, 2003 IBC and 2006 IBC all accept tower designs using a TIA-222 standard. However, 2000 IBC accepts tower designs to 222-E and 2003 IBC and 2006 IBC accept tower designs to 222-F. None of the IBC codes accept 222-G.

The TIA subcommittee TR14.7 has submitted a proposal to ICC recommending changing the TIA reference standard from 222-F to 222-G. This proposal was approved as submitted at the ICC Public Hearings in September 2006. If there are no objections at the ICC Final Action Hearings, which will be held in May 2007, this proposal will be incorporated into the 2007 Supplement of the 2006 IBC. Jurisdictions that already approved 2006 IBC do not automatically adopt the 2007 Supplement of 2006 IBC. In addition, supplements are normally bypassed by jurisdictions and only main code changes are adopted. However, the 2007 Supplement could be adopted by jurisdictions on a project-by-project basis by presenting the 2007 supplement to jurisdictions that have already adopted the 2006 IBC for acceptance. The next main code change is scheduled for 2009 and if this proposal is adopted in the 2007 Supplement of 2006 IBC it will then automatically be included in 2009 IBC.

2. ANNEX A: PROCUREMENT AND USER GUIDELINES

Annex A provides the broadcast tower owner and tenant a list and explanation of the procurement specifications required for purchasing a new tower and for purchasing an analysis/modification to existing towers. All of the procurement specification explanations in Annex A also refer to sections in the body of the standard for additional information.

The major items that need to be included in the procurement specifications for new towers and for analysis/modification of existing towers are:

a. Structure classification
b. Three-second-gust basic wind speed and design ice thickness
c. Exposure category
d. Topographic category
e. Earthquake site class
f. Serviceability requirements
g. Climbing and working facilities classification
h. Supplemental requirements

Structure classification

Three structure classifications have been established to differentiate reliability criteria. Structure classifications determine the nominal wind, ice and earthquake loads. This is accomplished through an importance factor. An importance factor is applied to the wind load, ice thickness and earthquake load.

Table 2-1 of 222-G provides a detailed description of each classification and is also shown in Figure 1. The default structure classification is Class II and most broadcast facilities would be Class II.

Basically Class III increases the resistance (reliability) of the tower with respect to Class II and Class I decreases the resistance (reliability) of the tower with respect to Class II. The importance factors are listed in Table 2-3 of 222-G and are also shown in Figure 2.

If the structure classification is not specified in the procurement specifications the default classification will be used.

Three-second-gust basic wind speed and design ice thickness

Wind load induced on an object is a function of wind pressure multiplied by the effective projected area (EPA) of the object. In addition, wind pressure is proportional to the square of wind speed.

The 222-G standard follows the recommendations of the American Society of Civil Engineers Standard "Minimum Design Loads for Buildings and Other Structures", ASCE 7, for basic wind speed and design ice thickness. ASCE 7 has changed from fastest mile to three second-gust basic wind speeds. Basic wind speed is measured 10 meters above grade. Most regions in the United States are in the 90 MPH three second gust basic wind speed zone. Figure A1-1a through A1-1e of 222-G show the basic wind speed without ice contour map. 90 MPH three second gust basic wind speed converts to 75 MPH fastest mile basic wind speed. The 90 MPH three second gust basic wind speed is defined as the 50 year return period basic wind speed averaged over three seconds and 75 MPH fastest mile basic wind speed is defined as the 50 year return period basic wind speed averaged over 48 seconds. Annex L of 222-G and Figure 3 show the conversion from fastest mile to three second-gust wind speeds for various wind speeds and also equivalent wind speeds for other time intervals.

222-G is the first 222 standard that includes a mandatory ice loading. However, ice may be ignored for Class I structures and for structures located in regions where the design ice thickness is less than or equal to 0.25 inches. Ice thickness escalates with the height of the structure above ground and this escalation is shown in Figure 4.

Annex B of 222-G includes a county listing of minimum basic wind speed without ice, minimum basic wind speed with ice and minimum design ice thickness. Jurisdictions may require a higher basic wind speed and/or design ice thickness than 222-G. If higher values are to be used they must be specified in the procurement specifications. Otherwise the values in Annex B will be used.

Exposure category

Wind speed varies with respect to the height of the structure above ground. The escalation of wind speed with height is dependent on the ground surface irregularities at the site. 222-G includes three exposure categories. A description of each category is presented in Section 2.6.5.1 of 222-G and also in Figure 5. Figure 6 shows the escalation of wind speed with height for the three exposure categories. Wind speed and, therefore wind load, is increased for exposure D with regards to exposure C below 900 ft and decreased for exposure B with regards to exposure C below 1200 ft. The default exposure category is category C.

If the exposure category is not specified in the procurement specifications the default category will be used.

Topographic category

When a structure is located on a hill, ridge or escarpment (Figure 7) the structure could be exposed to an increase in wind load due to an increase in wind speed caused by the topography at the site. 222-G describes this wind speed-up effect with five topographic categories. The five categories are described in section 2.6.6.2 of 222-G and also are shown in Figure 8. A topographic factor (Kzt) is calculated based on the topographic parameters at the site and is used to determine wind load and ice thickness. Topographic category 1 is the default category and is used when there is no wind speed-up effect. Kzt for category 1 is 1.00. If wind speed up effect is to be considered the topographic category and height of the topographic feature is to be included in the procurement specifications. If category 5 is being considered detailed site information is required with the procurement specifications.

Earthquake site class

The earthquake site class is determined by the soil parameters at the site. The soil parameters and, therefore the earthquake site class, affect the earthquake loading on the structure. The earthquake site classes are shown in Table 2-11 of 222-G and are also shown in Figure 9. The default earthquake site class is Class D.

If sufficient geotechnical information is not included with the procurement specifications to determine the earthquake site class the default site class will be used.

Another earthquake parameter that affects earthquake loading is the earthquake spectral response acceleration at short periods (Ss). Values for Ss are presented, by county, in Annex B of 222-G. Earthquake loading may be ignored for Class I structures, structures in regions where Ss is less than or equal to 1.00 or where the total calculated horizontal earthquake load is less than 50% of the total calculated horizontal wind load without ice. Based on the above requirements earthquake loading may be ignored for most broadcast tower sites in the United States.

If Ss is not specified in the procurement specifications the values in Annex B will be used.

Serviceability requirements

Serviceability requirements are included in Section 2.8 of 222-G. These requirements are minimum requirements. If more stringent specifications are required they must be stated in the procurement specifications. Refer to Section A2.8 of 222-G.

Climbing and working facilities classification

Climbing and working facilities are classified for the purpose of determining dimensional requirements. Two classifications have been established. Class A facilities can be accessed by both Authorized (Basic) and Competent (Skilled) climbers. Class B facilities can be accessed by only Competent (Skilled) climbers. Class A facilities provide greater climbing clearances and additional climbing safety items. Section 12 of 222-G describes the two classes and minimum requirements. The default classification is Class B. If additional requirements are necessary they must be included in the procurement specifications.

Supplemental requirements

The following chart lists supplemental requirements that includes:

a. A reference to the section in 222-G that describes the parameter
b. An Annex A reference noted in ()
c. The default condition when supplemental requirements are not specified.

Supplemental Requirement	Section Reference	Default
Corrosion Control	5.6 (A5.6.6)	Non-corrosive soil
High Frequency Dampers	7.5 (A7.5)	>1200 FT in HT
Low Frequency Dampers	7.5 (A7.5)	N/A
Foundation and Anchorage	9.0 (A9.0) Annex F	Clay presumptive soil parameters
Protective Grounding	10.0 (A10.0)	Section A10.0
Obstruction Marking	11.0 (A11.0)	None

The following chart summarizes the major 222-G procurement specifications and the associated 222-F requirement:

	222-G	222-F
Structure Classification	I, II or III Default is II	None
Basic Wind Speed	3-second gust	Fastest mile
Ice	Mandatory Follows ASCE7-05	Optional – 75% no ice wind pressure w/ ½" ice
Exposure Category	B, C or D Default is C	C
Topographic Category	1 to 5 Default is 1 Follows ASCE7-05	None
Earthquake loading	Section 2.7 Ignore if Class I, Ss<1.00 or total seismic shear is less than 50% total wind shear for structure w/o irregularities	None
Climbing and Working Facility Classes	Updated/Comprehensive A or B Default is B	Minimal

3. BASIS FOR PROCEEDING WITH EXISTING TOWER MODIFICATIONS

222-G includes two types of structural analyses that can be performed. They are:

a. Feasibility Structural Analysis
b. Rigorous Structural Analysis

A feasibility structural analysis is described in Section 15.5.1 of 222-G and is to be used as a preliminary review to evaluate a changed condition. Only the overall tower stability and main structural members are evaluated. Adequacy of connections and foundations are ignored. It may not be used as final acceptance of a changed condition.

A rigorous structural analysis is described in Section 15.5.2 of 222-G and is to be used as the final acceptance of a changed condition. All structural members, connections, foundations and associated details must be evaluated.

The structural analysis report must specify that it is either a feasibility or rigorous analysis. A feasibility analysis report must state that final acceptance of the changed condition shall be based on a rigorous structural analysis. All assumptions must be resolved in order for the structural analysis to be specified as a rigorous structural analysis. Not following this requirement could expose both the engineering firm, performing the structural analysis, and the tower owner/tenant to risk.

Broadcast towers originally designed to a previous standard are exempt from some requirements of 222-G. The exemptions are listed in Section 15.6 of 222-G.

The modification design must be based on a rigorous structural analysis. Design documents must be prepared to show the modification members and connections. Any data specified in the design documents to be verified must be validated prior to implementing. This is described in detail in Section 15.7 of 222-G.

4. RIGGING AND TEMPORARY SUPPORT REQUIREMENTS DURING INSTALLATION

There are two rigging requirements that must be performed prior to installation. They are:

a. A rigging plan must be documented that adequately determines the installation sequence and confirms the rigging equipment will have the proper safety factors when under installation loads. The rigging plan is the responsibility of the erector.
b. The structure must be evaluated for structural adequacy from the rigging loads that are applied to it from the installation operation.

This is also the responsibility of the erector and in most cases will be outsourced to the Engineer of Record (EOR) of the rigorous analysis report or a competent structural engineer with structural engineering experience in broadcast towers.

5. MAINTENANCE AND CONDITION ASSESSMENT

The maximum intervals between maintenance and condition assessments are listed in Section 14.2 of 222-G and also shown in Figure 10.

Annex J of 222-G provides guidelines for assessment, structure mapping and equipment mapping.

6. DESIGN REQUIREMENTS

a. ASD vs. LRFD

One of the major changes in 222-G was to convert from an Allowable Stress Design (ASD) to Load and Resistance Factor Design (LRFD). Basically a LRFD design means the strength provided in the design must be at least equal to the factored load acting on it. The load factors were derived based on the probability of an overload. An ASD design means the strength provided in the design must be at least greater, by a certain margin, to the load acting on it. This "certain margin" is the safety factor. In ASD the load is a service load and therefore is not factored. The load factors in LRFD were determined by probabilistic methods and the safety factors in ASD were determined by experience and judgment. A LRFD design should result in a more reliable design. The major reason for this is that the overload condition is calculated in a LRFD design and only assumed in an ASD design.

b. Dynamic effects from wind gusts

Even though wind involves kinetic energy, their effects are simulated by equivalent static loads determined in accordance with the design standards. In recent years there have been more sophisticated efforts to investigate the actual response of tower structures to the dynamic aspects of wind gusts. A conclusion drawn from these studies is that the tower member internal forces in the upper portions of tall, guyed towers are considerably higher than those determined by the usual static analysis. Consequently, 222-G has addressed this issue by including pattern loading in the static analysis of towers, which emulates the dynamic load and the dynamic response of the tower and provides for a more reliable design. Pattern loading is described in detail in Section 3.0 of 222-G.

c. Transmission line wind load

Another major change in 222-G is how transmission line (T/L) wind load is treated. Section 2.3.6.3 of 222-F allowed T/L's that were located in the vicinity of tower faces to be treated as structural components when calculating wind load. The use of Section 2.3.6.3 decreased the overall wind load on towers and resulted in lighter tower designs. Towers originally designed to 222-F utilizing section 2.3.6.3 for most of the T/L's will see a higher stress level and possible strengthening when analyzing the tower to 222-G. Towers originally designed to 222-F without utilizing section 2.3.6.3 will see a lower stress level when analyzing the tower to 222-G. Case studies will be presented later to confirm this statement.

7. STRUCTURAL STEEL MATERIAL

Structural steel material is covered in section 5.4 of 222-G. The major changes are:

a. Pre-qualified steel list, which is included in Table 5-1 of 222-G and also shown in Figure 11.
b. Non pre-qualified steel must meet additional requirements, which are:
   - Maximum carbon equivalency
   - Minimum elongation
   - Location of mechanical tests for round bars
   - Minimum Charpy V-notch values

   These requirements affect the weldability, ductility and strength of the steel.

8. FOUNDATION COST BUDGETING

Many budgets for new tower projects were based on preliminary foundation designs using "normal soil" as defined in section 7.1.3 of 222-F. "Normal soil" is a very strong cohesive soil and in many cases results in liberal preliminary foundation designs. When final designs were prepared, after the geotechnical investigation was completed, the cost of the final designs was higher than the preliminary designs based on "normal soil". 222-G has attempted to solve this issue by creating two presumptive soil types to be used in the absence of geotechnical investigation. One is a clay type soil that is weaker than the "normal soil" specified in 222-F. The other is a sand type soil to be used in certain regions where sand is probably the actual soil type. The default presumptive soil type is clay. If a sand type soil is to be used it must be specified in the procurement specifications. If the actual geotechnical investigation information is available it should be included in the procurement specifications.

9. GROUNDING

Section 10.0 includes the minimum grounding requirements. Default requirements are included in section A10.0.

There is a maximum electrical resistance requirement that the primary grounding system must meet. It is the owner's responsibility to specify a grounding system that will meet this requirement. It is intended that the owner or owner's representative will verify that the requirement is met. If it does not modifications to the grounding system will be required.

10. HOW TO PURCHASE 222-G

The 222-G standard can be purchased through the TIA website,
www.tiaonline.org/standards/search_n_order.cfm
or
can be purchased through Global Engineering Documents,
(1-800-854-7179) International (303-397-7956)

CASE STUDIES

Four case studies are presented to show the effects from evaluating a tower to 222-G when the original tower design was from a previous 222 standard. For these studies the guy assembly and leg member stress levels are evaluated and compared. Stress level is defined as the ratio (%) of the actual force in the member to the strength of the member. The design parameters for each case study are shown below:

Original Design	Feasibility Analysis
222-C	222-G
222-F with no ice and utilizing section 2.3.6.3	222-G
222-F with ice and utilizing section 2.3.6.3	222-G
222-F with no ice and not utilizing section 2.3.6.3	222-G

The procurement specifications for the feasibility structural analysis are:

Structure Classification II
90 MPH basic wind speed without ice
30 MPH basic wind speed with ice
0.5 inches design ice thickness
Topographic category 1
Earthquake loading site class D
Earthquake spectral response acceleration at short periods (Ss) of 1.00
Climbing and working facilities class (N/A - existing tower)
Equipment

1. ATW22H3-HTS-26S top mounted antenna fed with a 6-1/8" T/L.
2. SHPX-12AC antenna, with radomes, side mounted between the 887' and 1013' levels fed with a 3" flex line.
3. TFU-22JTH antenna side mounted between the 811' and 848' levels fed with 3-1/8" line.
4. Three 6' diameter dishes, with radome, at the 325' level each fed with one (1) EW63.
5. Nine panel cell system at the 200' level each fed with one (1) 1-5/8" line.
6. One (1) inside climbing ladder with cable type safety device for the full height of the tower.
7. One (1) 1-1/4" lighting conduit for the full height of the tower.

8. One (1) 1" support conduit to the 950' level.
9. One (1) 1-1/2" support conduit to the 325' level.
10. Two (2) 1-1/2" support conduit to the 200' level.
11. Dual lighting system

The tower characteristics are:

Overall height = 1116.0 ft
Steel Height = 1067.0 ft
Face width = 8.0 ft
Guy level elevations:
(1) 154.0 ft
(2) 316.0 ft
(3) 490.0 ft
(4) 664.0 ft
(5) 850.0 ft
(6) 1042.0 ft
Inner anchor distance = 375.0 ft
Outer Anchor distance = 785.0 ft

Case Study No. 1

Tower originally designed to 50 psf in accordance with 222-C. The results of the leg member stress levels show that the overall stress levels for 222-C are higher than for 222-G except for an isolated area just above the 5th guy level. This overstress is due to the no ice pattern loading requirement of 222-G. The 222-G ice loading did not govern. The leg stress level comparison is shown in Figure 12. The guy assembly stress level comparison is shown in Figure 13.

To upgrade the tower to 222-G requires adjusting the initial tensions of the guy assemblies at guy levels 4 and 5.

Case Study No. 2

Tower originally designed to 70 mph with no ice in accordance with 222-F utilizing section 2.3.6.3 of 222-F. The results show leg overstresses in the vicinity of the 2nd and 4th guy levels. Also, guy assembly levels 1 through 4 are overstressed. These overstresses are due to the original design utilizing section 2.3.6.3 of 222-F and also the original design not including an ice loading. Using these two design parameters reduced both the wind load and weight on the tower and resulted in a very economical 222-F tower design but also reduced the reliability of the tower if exposed to an overload condition. In this case study the 222-G T/L wind load requirements created more load on the tower than the 222-F T/L wind load requirements. The 222-G ice loading did not govern. The leg stress level comparison is shown in Figure 14. The guy stress level comparison is shown in Figure 15.

To upgrade the tower to 222-G requires increasing the guy assembly size at guy levels 1 through 4, adjusting the initial tensions at guy levels 1 through 4 and installing redundant horizontal, to brace the tower legs, at 17 levels. 15 in the vicinity of the 4th guy level and 2 in the vicinity of the 2nd guy level.

Case Study No. 3

Tower originally designed to 70 MPH with no ice and 61 mph with ½" radial ice in accordance with 222-F utilizing section 2.3.6.3. The results show leg overstresses in two tower sections just above the 5th guy level and guy overstresses at level #2. The overstress is not as significant as Study 2 due to the original design including an ice load case, which governed the 222-F design in some areas of the tower and therefore increased the size of some of the leg members and guy assemblies. The 222-G ice loading did not govern. The leg stress level comparison is shown in Figure 16. The guy stress level comparison is shown in Figure 17.

To upgrade the tower to 222-G requires increasing the guy size at guy level #3 and adjusting the initial tensions at guy levels 1, 2, 3 and 4.

Case Study No. 4

Tower originally designed to 70 mph with no ice in accordance with 222-F not utilizing section 2.3.6.3. The results show leg overstresses in five tower sections in the vicinity of guy levels 4 and 5. The guy assemblies are not overstressed. The overstresses are minor in this case study due to the 222-F T/L wind load being higher which was caused by not utilizing section 2.3.6.3 in the original design. This resulted in the original design having larger guy assemblies and leg members. The 222-G ice loading did not govern. The leg stress level comparison is shown in Figure 18. The guy stress level comparison is shown in Figure 19.

To upgrade the tower to 222-G requires adjusting the initial tensions at guy levels 4,5 and 6.

A summary of the modifications are shown below:

	Leg Reinforcement	Guy Change out	Initial Tension Changes
Study 1	None	None	3 Levels
Study 2	17 levels	1st thru 4th Level	4 levels
Study 3	None	3rd Level	4 Levels
Study 4	None	None	3 Levels

The following conclusions may be drawn:

a. Even if the 222-G feasibility analysis shows remaining overall capacity the pattern loading may cause overstresses in the top spans of the tower.
b. Utilizing section 2.3.6.3 in 222-F designs may cause overstresses when analyzing the tower for 222-G.
c. 222-G 30 MPH and ½" radial ice load does not govern. However, if a higher ice thickness would have been considered leg overstresses in the bottom tower spans may have occurred.
d. 222-F ice loading governs some guy assemblies and leg members in the original design so as to result in less stress levels when analyzed to 222-G.

INITIAL STANDARD AND EQUIPMENT REVIEW

There are methods available that can provide broadcast owners and tenants a quick estimate on the feasibility of upgrading the tower from the original tower design parameters to 222-G. In addition, there are also methods available to quickly determine the feasibility on equipment inventory changes to the tower.

The following is one such method that has been developed for Stainless towers but can also be used for non-Stainless towers. The information required to perform a standard and equipment review of a Stainless Tower is:

a. Stainless tower ID number
b. Tower location (county)
c. Tower cross section showing existing T/L, ladder and conduit location
d. Proposed equipment

An initial standard and equipment review was performed for Case Study #1. The input page to the program is shown in Figure 20. Based on the information provided from the design parameters listed above, the tower effective projected area (EPA) and the linear appurtenance EPA are determined. The program takes into account wind load, load factors (222-G), allowable stresses and resistances (222-G) to determine the comparison results. The output is shown in Figure 21. A wind loading index is calculated for both the original and proposed design parameters and is shown for tower heights from 33 ft to 2000 ft. Case Study #1 is an 1116 ft overall height structure. At that height the 222-C wind loading index is higher than the 222-G wind loading index. This means that the overall 222-C tower member stress level is higher than the overall 222-G tower member stress level. The 222-C wind loading index is 16.5 and the 222-G wind-loading index is 14.9. Since the difference is 10% it can be estimated that the tower will either meet or require minor modification to meet 222-G. This is very close to the result of the Case Study 1 feasibility analysis. Please note that this wind loading index tool does not investigate an ice loading.

An initial standard and equipment review was performed for Case Study 2. The output is shown in Figure 22. The 222-G wind loading index is higher than the 222-F wind loading index. Therefore, the overall 222-G tower member stress level is higher than the overall 222-F tower member stress level. The 222-G wind loading index is 14.9 and the 222-F wind-loading index is 13.0. Since the difference is 15% it can be estimated that the tower will require significant modifications to meet 222-G. This is very close to the result of the Case Study 2 feasibility analysis. Again, please note that this wind loading index tool does not investigate an ice loading.

An initial standard and equipment review was performed for Case Study 3. Since the initial standard and equipment review does not investigate ice loading and the only difference between Case Study 2 and Case Study 3 is that ice loading was included in the original design the results of the initial standard and equipment review will be the same as Case Study #2. However, since the original design included an ice loading and that ice loading governed parts of the original design the initial standard and equipment review results in a slightly conservative estimate. Modifications are still required to upgrade the tower to 222-G. The upgrades are just not as significant as in Case Study 2.

An initial standard and equipment review was performed for Case Study 4. The output is shown in Figure 23. The 222-F wind loading index is higher than the 222-G wind loading index. Therefore, the overall 222-F tower member stress level is higher than the overall 222-G tower member stress level. The 222-F wind loading index is 17.6 and the 222-G wind-loading index is 14.9. Since the difference is 15% it can be estimated that the tower will either meet or require minor modification to meet 222-G. This is very close to the results of the Case Study 4 feasibility analysis. Again, please note that this wind loading index tool does not investigate an ice loading.

Classification of Structures (ANSI/TIA-222-G **)

Description of Structure	Class
Structures that due to height, use or location represent a low hazard to human life and damage to property in the event of failure and/or used for services that are optional and/or where a delay in returning the services would be acceptable.	I
Structures that due to height, use or location represent a substantial hazard to human life and/or damage to property in the event of failure and/or used for services that may be provided by other means.	II
Structures that due to height, use or location represent a high hazard to human life and/or damage to property in the event of failure and/or used primarily for essential communications.	III

Figure 1

Importance Factors (ANSI/TIA-222-G **)

Structure Class	Wind Load Without Ice	Wind Load With Ice	Ice Thickness	Earthquake
I	0.87	N/A	N/A	N/A
II	1.00	1.00	1.00	1.00
III	1.15	1.00	1.25	1.50
Note: Ice and earthquake loads do not apply to Class I structures				

Figure 2

Annex L: Wind SPEED CONVERSIONS (Normative) (ANSI/TIA-222-G **)

This annex provides conversion of wind speeds based on various averaging periods to the 3-sec gust wind speed. Wind data based on other averaging periods are to be converted to a 3-sec gust wind speed for use with the Standard.

3-sec gust (mph)	Fastest-mile		10-min avg. (mph)	Hourly Mean (mph)
	Wind Speed (mph)	Averaging period (sec)		
60	50	72	42	40
70	58	62	49	46
80	66	55	56	53
85	70	51	59	56
90	75	48	62	60
95	78	46	66	63
100	80	45	69	66
105	85	42	73	70
110	90	40	76	73
115	95	38	80	76
120	100	36	83	79
125	105	34	87	83
130	110	33	90	86
135	115	31	94	89
140	120	30	97	93
145	125	29	101	96
150	130	28	104	99
155	135	27	108	103
160	140	26	111	106
165	145	25	115	109
170	150	24	118	113

Notes: 1. For conversion to [m/s] multiply the above values by 0.447.
2. Linear interpolation may be used between the values shown.

Figure 3

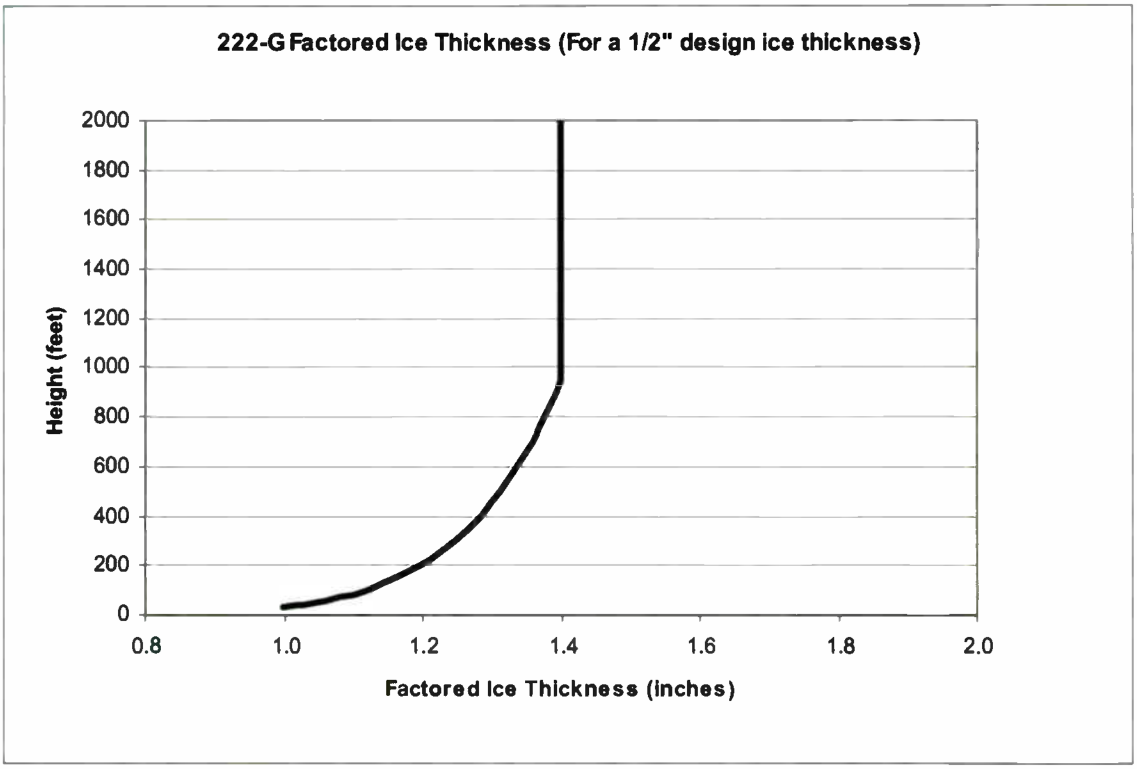


Figure 4

2.6.5 Exposure Categories (ANSI/TIA-222-G **)
2.6.5.1 General

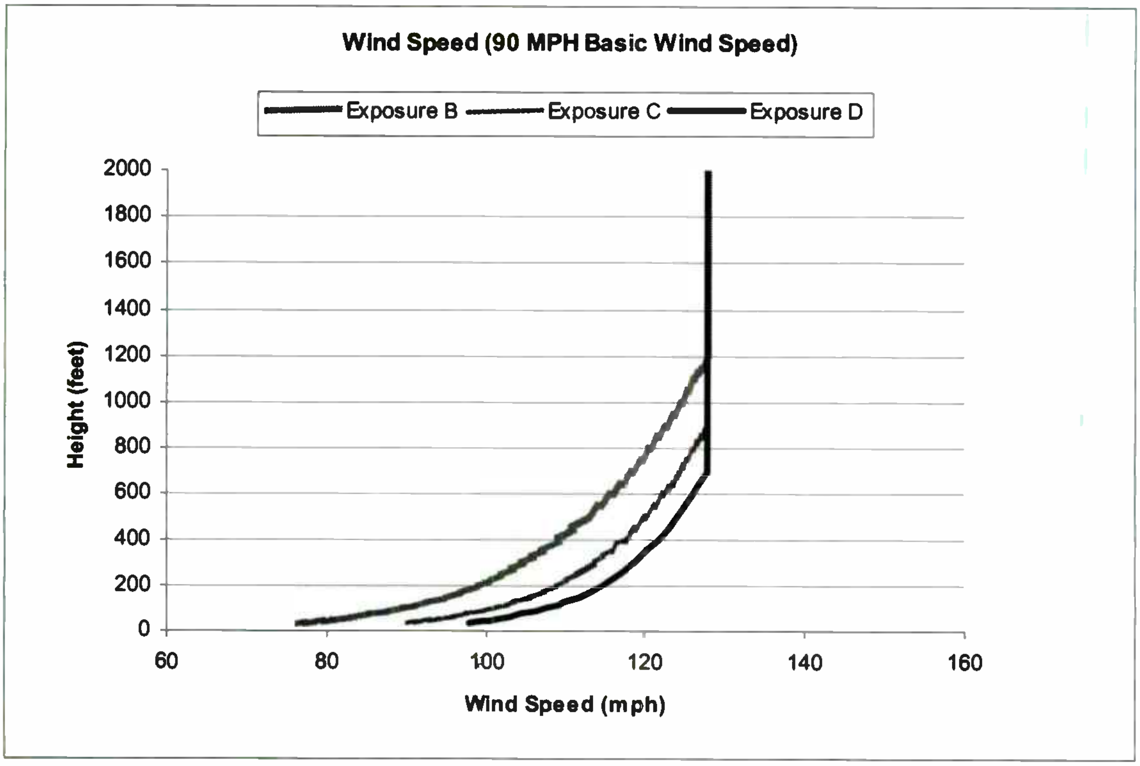
An exposure category that adequately reflects the characteristics of ground surface irregularities at the site shall be determined. Account shall be taken of variations in ground surface roughness that arise from natural topography and vegetation as well as from constructed features. The exposure category for a structure shall be assessed as being one of the following:

1. **Exposure B:**Urban and suburban areas, wooded areas, or other terrain with numerous closely spaced obstructions having the size of single-family dwellings or larger. Use of this exposure shall be limited to those areas for which terrain representative of Exposure B surrounds the structure in all directions for a distance of at least 2,630 ft [800 m] or ten times the height of the structure, whichever is greater.

2. **Exposure C:**Open terrain with scattered obstructions having heights generally less than 30 ft [9.1 m]. This category includes flat, open country, grasslands and shorelines in hurricane prone regions.

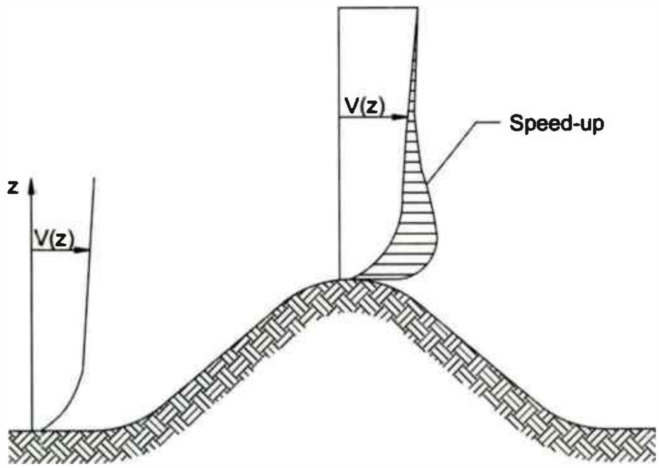
3. **Exposure D:**Flat, unobstructed shorelines exposed to wind flowing over open water (excluding shorelines in hurricane prone regions) for a distance of at least 1 mile [1.61 km]. Shorelines in Exposure D include inland waterways, lakes and non-hurricane coastal areas. Exposure D extends inland a distance of 660 ft [200 m] or ten times the height of the structure, whichever is greater. Smooth mud flats, salt flats and other similar terrain shall be considered as Exposure D.

Figure 5

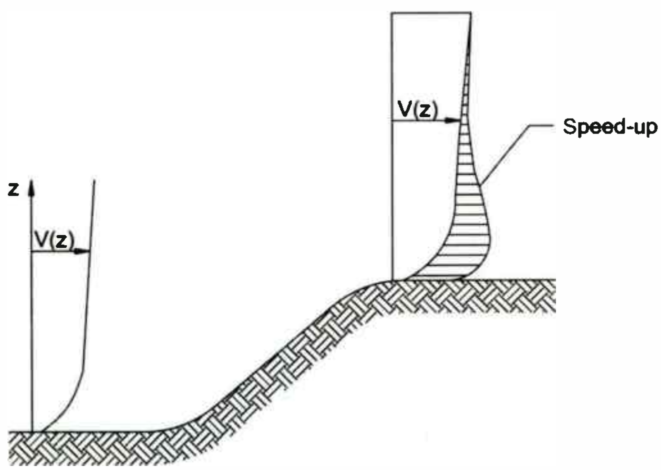


Wind Speed (90 MPH Basic Wind Speed)
Exposure B
Exposure C
Exposure D
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Height (feet)
60
80
100
120
140
160
Wind Speed (mph)

Figure 6



2-D RIDGE OR 3-D AXISYMMETRICAL HILL



ESCARPMENT

FIGURE 7

2.6.6.2 Topographic Categories (ANSI/TIA-222-G **)
The topographic category for a structure shall be assessed as being one of the following:

1. **Category 1:** No abrupt changes in general topography, e.g. flat or rolling terrain, no wind speed-up consideration shall be required.

2. **Category 2:**Structures located at or near the crest of an escarpment. Wind speed-up shall be considered to occur in all directions. Structures located vertically on the lower half of an escarpment or horizontally beyond 8 times the height of the escarpment from its crest, shall be permitted to be considered as Topographic Category 1.

3. **Category 3:**Structures located in the upper half of a hill. Wind speed-up shall be considered to occur in all directions. Structures located vertically on the lower half of a hill shall be permitted to be considered as Topographic Category 1.

4. **Category 4:**Structures located in the upper half of a ridge. Wind speed-up shall be considered to occur in all directions. Structures located vertically on the lower half of a ridge shall be permitted to be considered as Topographic Category 1.

5. **Category 5:**Wind speed-up criteria based on a site-specific investigation.

Figure 8

Table 2-11 Site Class Definitions (ANSI/TIA-222-G **)

<table>
<tr><th>Site Class</th><th>Description of Upper 100 ft [30.5 m] of Soil for the Site Location</th><th>Standard Penetration Resistance, N Cohesionless Soils PI ≤20</th><th>Undrained Shear Strength, Su Cohesive Soils PI > 20</th></tr>
<tr><td>A</td><td>Hard rock with 10 ft [3 m] or less of soil overburden.</td><td>N/A</td><td>N/A</td></tr>
<tr><td>B</td><td>Competent rock with moderate fracturing and weathering with 10 ft [3 m] or less of soil overburden.</td><td>N/A</td><td>N/A</td></tr>
<tr><td>C</td><td>Very dense soil, soft rock or highly fractured and weathered rock.</td><td>> 50</td><td>> 2 ksf [100 kPa]</td></tr>
<tr><td>D</td><td>Stiff soil.</td><td>15 to 50</td><td>1.0 to 2.0 ksf [50 to 100 kPa]</td></tr>
<tr><td rowspan="2">E</td><td rowspan="2">Weak soil (excluding site class F).</td><td>< 15</td><td>< 1.0 ksf [50 kPa]</td></tr>
<tr><td colspan="2">Soil profiles over 10 ft [3 m] thick with PI ≥20, moisture content ≥ 40%, Su< 0.5 ksf [25 kPa]</td></tr>
<tr><td>F</td><td>Soils vulnerable to potential failure or collapse under seismic loading.</td><td colspan="2">Soil profiles containing any of the following: peat and/or highly organic clays over 10 ft [3 m] thick, very high plasticity clays (PI > 75) over 25 ft [7.6 m] thick, soft/medium clays over 120 ft [36.6 m] thick, liquefiable soils, quick and highly sensitive clays, collapsible weakly cemented soils.</td></tr>
</table>

Figure 9

14.0 MAINTENANCE AND CONDITION ASSESSMENT (ANSI/TIA-222-G **)

14.1 Scope

This section addresses the maintenance and condition assessment of structures.

14.2 Maximum Intervals

Maintenance and condition assessment shall be performed as follows:

a) Three-year intervals for guyed masts and five-year intervals for self-supporting structures.

b) After severe wind and /or ice storms or other extreme conditions.

c) Shorter inspection intervals may be required for Class III structures and structures in coastal regions, in corrosive environments, and in areas subject to frequent vandalism.

Maintenance and condition assessment guidelines are provided in Annex J.

Figure 10

Table 5-1 (ANSI/TIA-222-G **)
Pre-Qualified Structural Steels

API – 5L	API Specification (Spec. 5L)
ASTM A36	Structural Steel.
ASTM A53	Pipe, Steel, Black and Hot-Dipped, Zinc-Coated Welded and Seamless.
ASTM A106	Seamless Carbon Steel Pipe.
ASTM A242	High-Strength Low-Alloy Structural Steel.
ASTM A500	Cold-Formed Welded and Seamless Carbon Steel Structural Tubing in Round and Shapes.
ASTM A501	Hot-Formed Welded and Seamless Carbon Steel Structural Tubing.
ASTM A514	High-Yield Strength, Quenched and Tempered Alloy Steel Plate, Suitable for Welding.
ASTM A529	High-Strength Carbon-Manganese Steel of Structural Quality.
ASTM A572	High-Strength Low-Alloy Columbium-Vanadium Steels of Structural Quality.
ASTM A588	High-Strength Low-Alloy Structural Steel with 50 ksi Minimum Yield
ASTM A606	Steel, Sheet and Strip, High-strength, Low-Alloy, Hot-Rolled and Cold-Rolled, with Improved Atmospheric Corrosion Resistance.
ASTM A618	Hot-Formed Welded and Seamless High-strength Low-Alloy Structural Tubing.
ASTM A633	Normalized High-Strength Low-Alloy Structural Steel Plates.
ASTM A871	High-Strength Low-Alloy Structural Steel Plate with Atmospheric Corrosion Resistance.
ASTM A913	High-Strength Low-Alloy Steel Shapes of Structural Quality.
ASTM A992	Steel for Structural Shapes for Use in Building Framing.
ASTM A1008/A	Steel, Sheet, Cold-Rolled, Carbon, Structural, High-Strength Low-Alloy and High Strength Low-Alloy with Improved Formability
ASTM A1011	Steel, Sheet and Strip, Carbon, Hot-rolled Structural Quality.
CSA G40.20/.21	General Requirements for Rolled or Welded Structural Quality Steels.

Note: Shapes and sizes outside the scope of an ASTM standard shall be considered as non-pre-qualified steel.

<u>Figure 11</u>

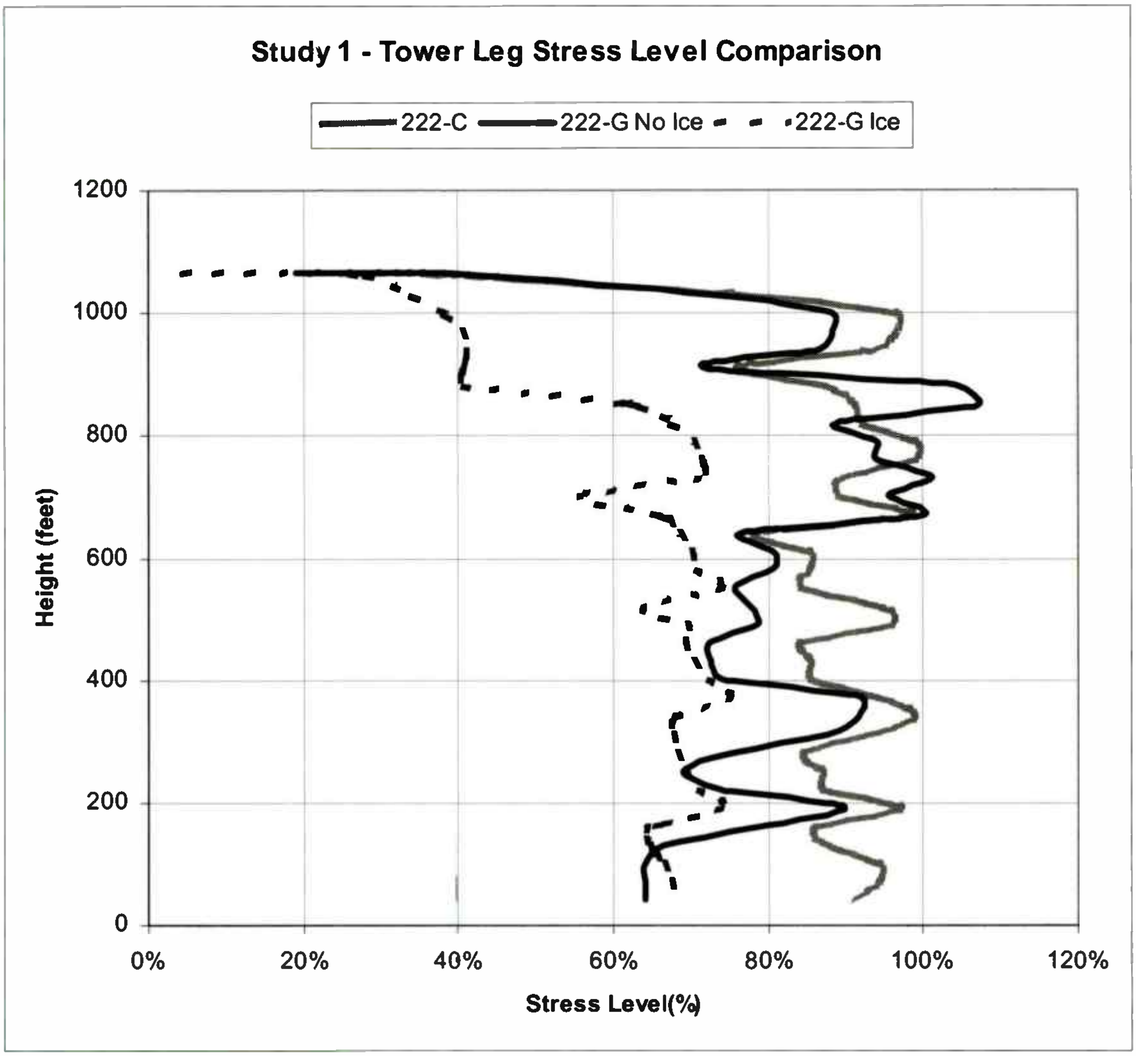

Study 1 - Tower Leg Stress Level Comparison
222-C
222-G No Ice
222-G Ice
Height (feet)
1200
1000
800
600
400
200
0
0%
20%
40%
60%
80%
100%
120%
Stress Level(%)

Figure 12

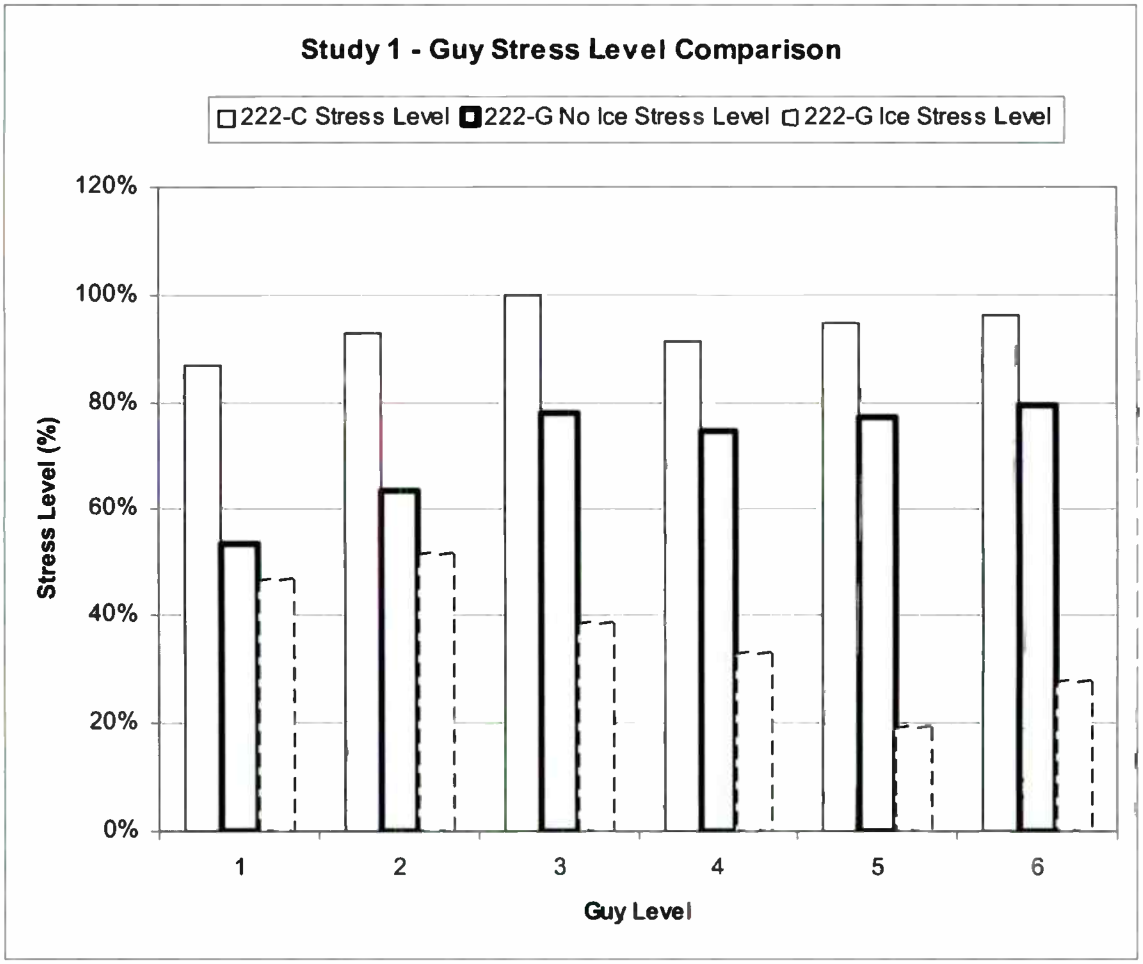

Study 1 - Guy Stress Level Comparison
222-C Stress Level
222-G No Ice Stress Level
222-G Ice Stress Level
120%
100%
80%
60%
40%
20%
0%
Stress Level (%)
1
2
3
4
5
6
Guy Level

Figure 13

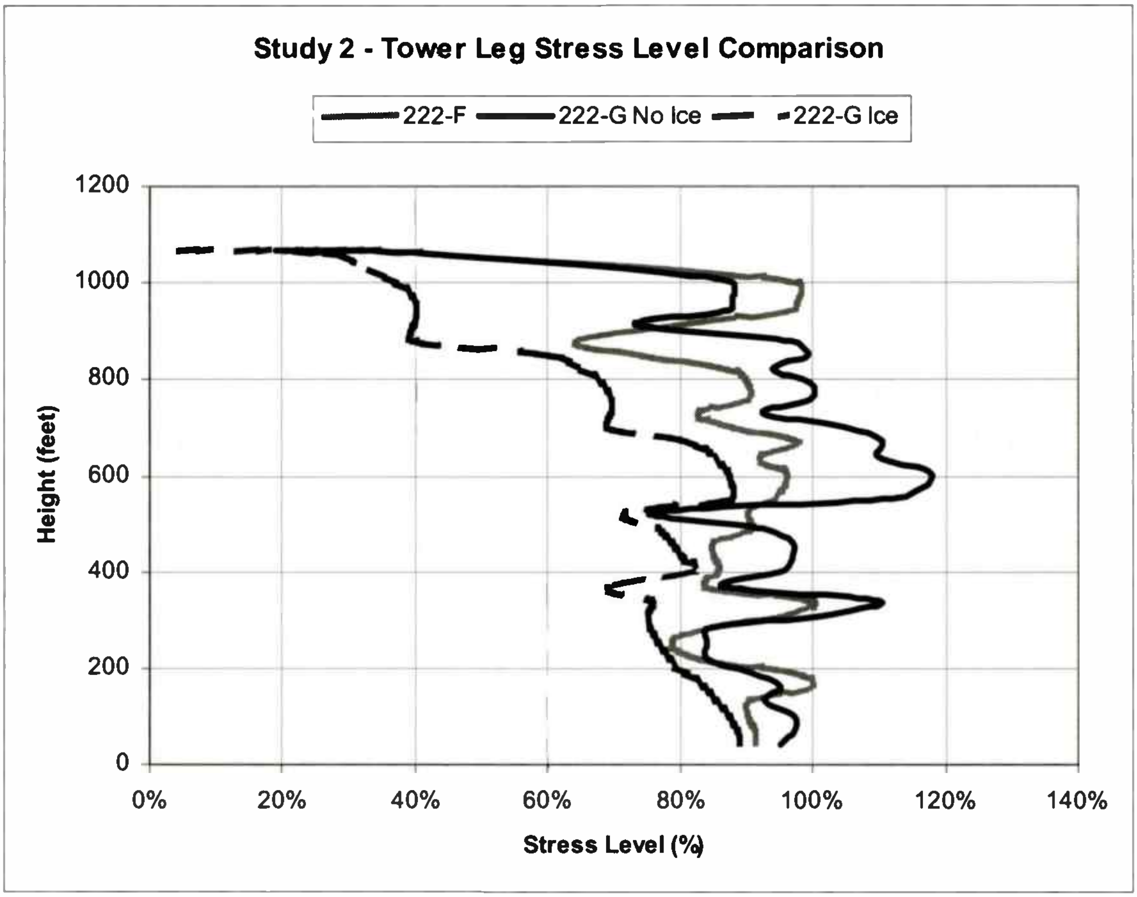

Study 2 - Tower Leg Stress Level Comparison
222-F
222-G No Ice
222-G Ice
1200
1000
800
600
400
200
0
Height (feet)
0%
20%
40%
60%
80%
100%
120%
140%
Stress Level (%)

Figure 14

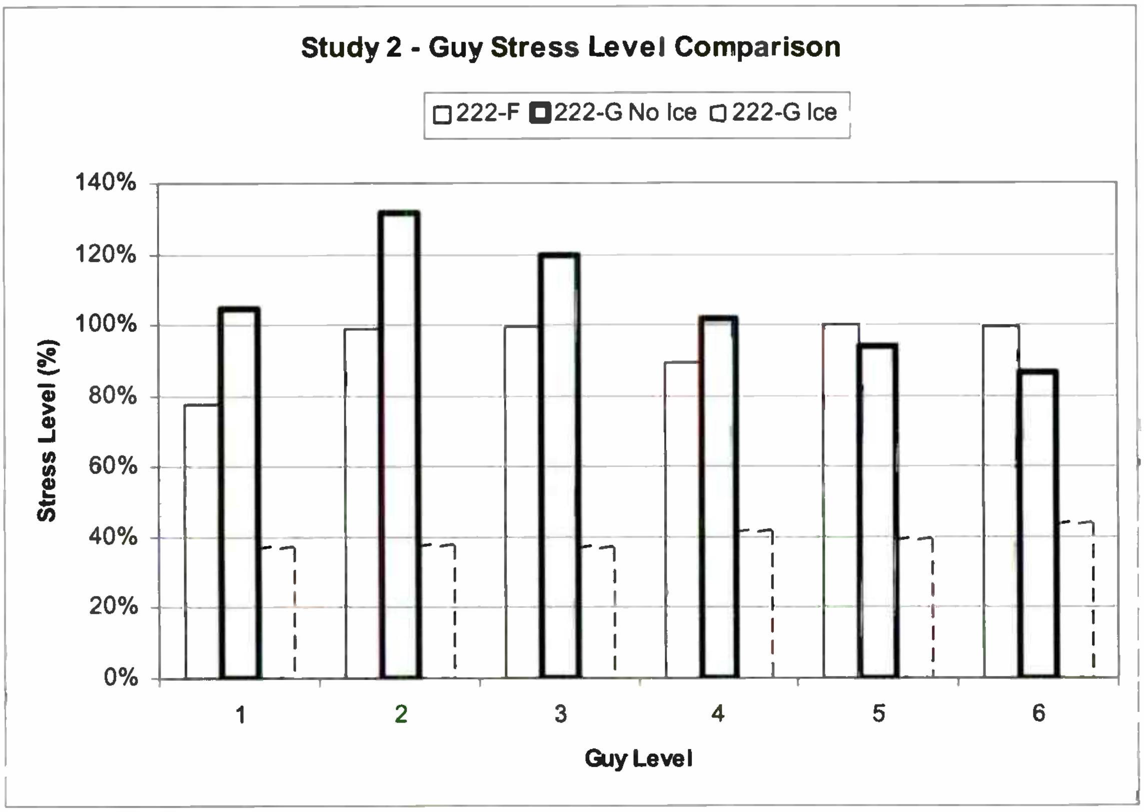

Study 2 - Guy Stress Level Comparison
222-F
222-G No Ice
222-G Ice
Stress Level (%)
140%
120%
100%
80%
60%
40%
20%
0%
1
2
3
4
5
6
Guy Level

Figure 15

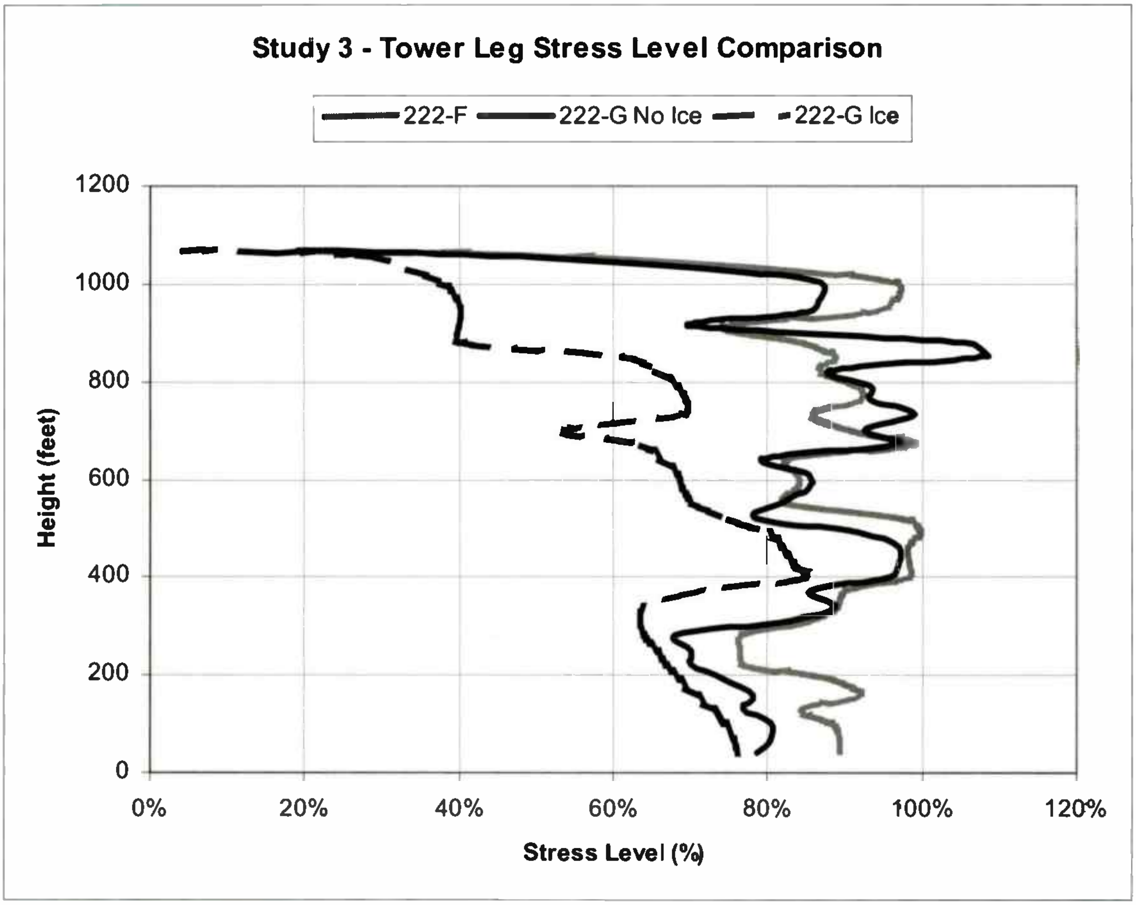

Study 3 - Tower Leg Stress Level Comparison
222-F
222-G No Ice
222-G Ice
1200
1000
800
600
400
200
0
Height (feet)
0%
20%
40%
60%
80%
100%
120%
Stress Level (%)

Figure 16

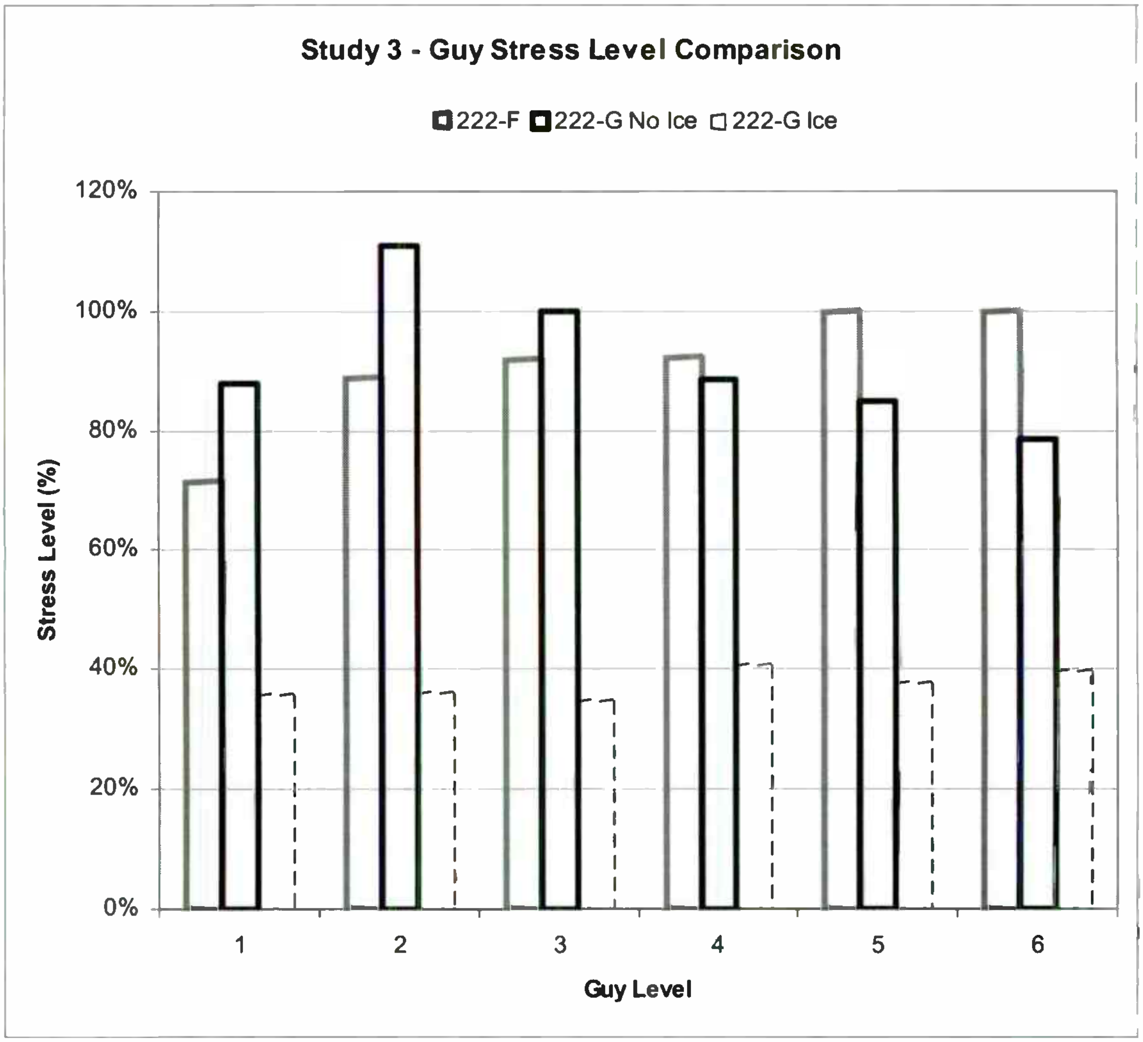

Study 3 - Guy Stress Level Comparison
222-F
222-G No Ice
222-G Ice
120%
100%
80%
60%
40%
20%
0%
Stress Level (%)
1
2
3
4
5
6
Guy Level

Figure 17

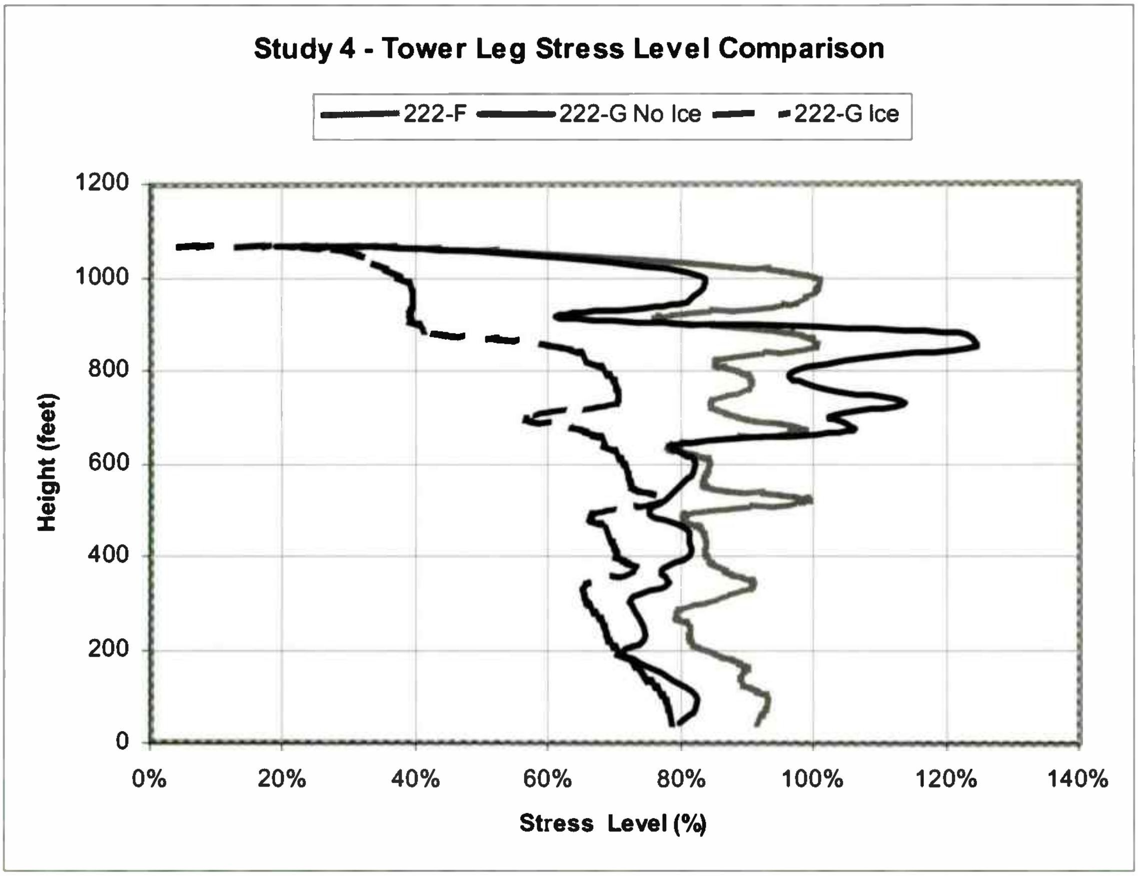

Study 4 - Tower Leg Stress Level Comparison
222-F
222-G No Ice
222-G Ice
1200
1000
800
600
400
200
0
Height (feet)
0%
20%
40%
60%
80%
100%
120%
140%
Stress Level (%)

Figure 18

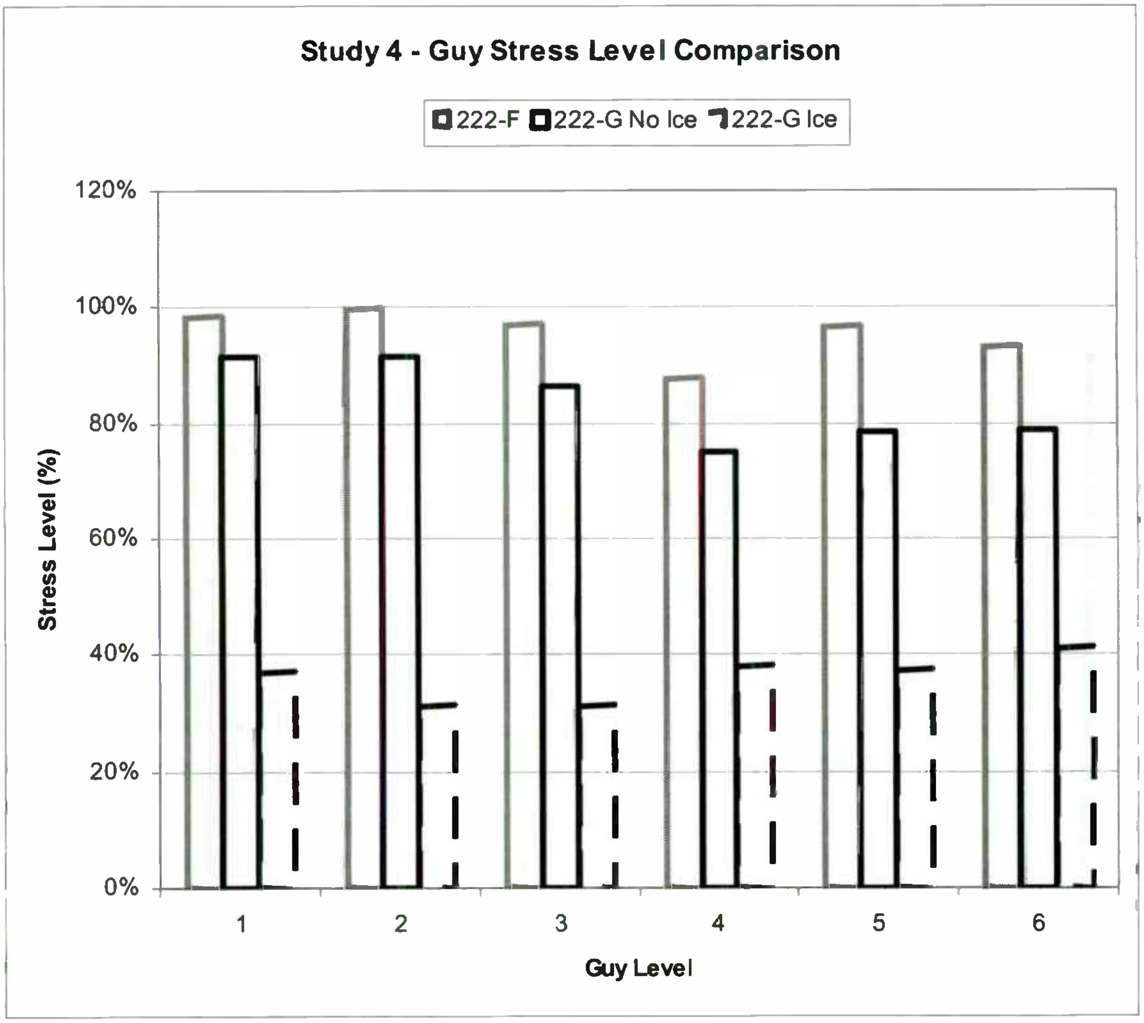

Study 4 - Guy Stress Level Comparison
222-F
222-G No Ice
222-G Ice
120%
100%
80%
60%
40%
20%
0%
Stress Level (%)
1
2
3
4
5
6
Guy Level

Figure 19

Customer Name:	NAB2007
Telephone Number:	
Manufacturer/Tower ID:	Stainless LLC
Station Call Letters:	
County:	
Structure Height:	1116

CURRENT EQUIPMENT				**ORIGINAL EQUIPMENT**		
BASIC WIND SPEED =	70	222-F		**BASIC WIND SPEED =**	70	222-F
BASIC WIND SPEED =	90	222-G		**BASIC WIND SPEED =**	90	222-G
WIND PRESSURE =	50	222-C		**WIND PRESSURE =**	50	222-C
MODIFICATION FACTOR	1			**MODIFICATION FACTOR**	1	
TOWER FACE WIDTH	8	feet		**TOWER FACE WIDTH**	8	feet
G/F Stress Factor	1.1			**G/F Stress Factor**	1.1	
TOWER and APPURTENANCES				**TOWER and APPURTENANCES**		
Ar(str)	0.79	sq. ft		**Ar(str)**	0.79	sq. ft
Ar(linear) 222-F	0.00	sq. ft		**Ar(linear) 222-F**	0.00	sq. ft
Af	0.56	sq. ft		**Af 222**	0.56	sq. ft
LINEAR APPURT				**LINEAR APPURT**		
Ac (round)	1.51	sq. ft		**Ac (round)**	1.51	sq. ft

Bay Height	6	FT
Average Leg diameter	3.5	inches
Average Horizontal size	3	inches
Average Diagonal size	0.75	inches
Average gusset area	100	sq inches
Round/FT	0.79	SQFT/FT
Flat/FT	0.56	SQFT/FT

Figure 20

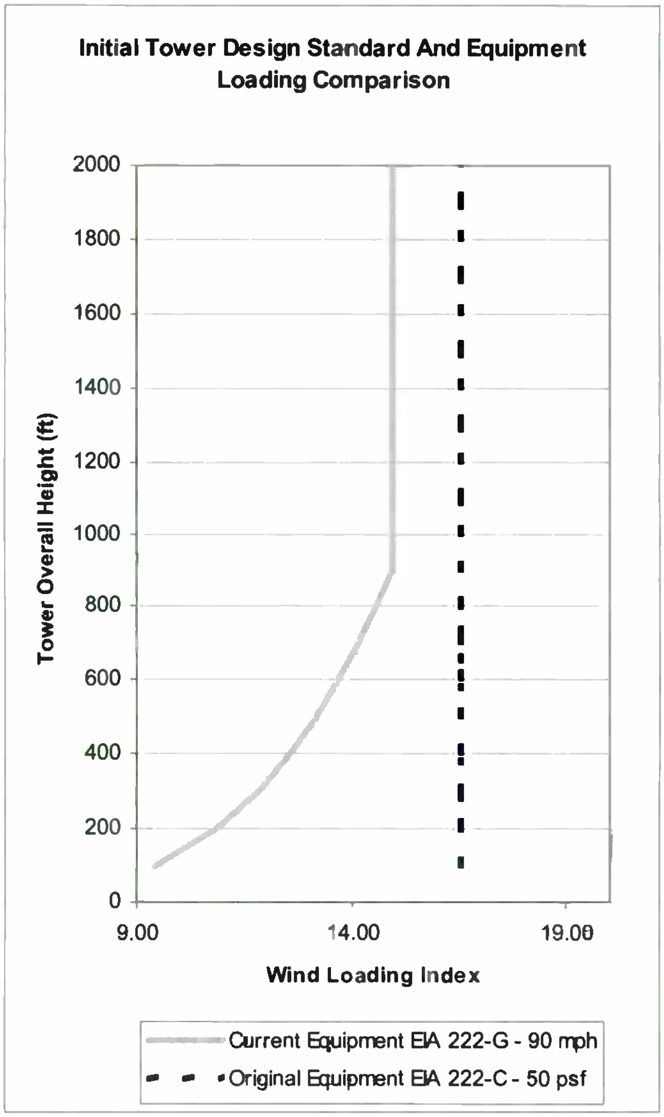

Initial Tower Design Standard And Equipment Loading Comparison
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Tower Overall Height (ft)
9.00
14.00
19.00
Wind Loading Index
Current Equipment EIA 222-G - 90 mph
Original Equipment EIA 222-C - 50 psf

Figure 21

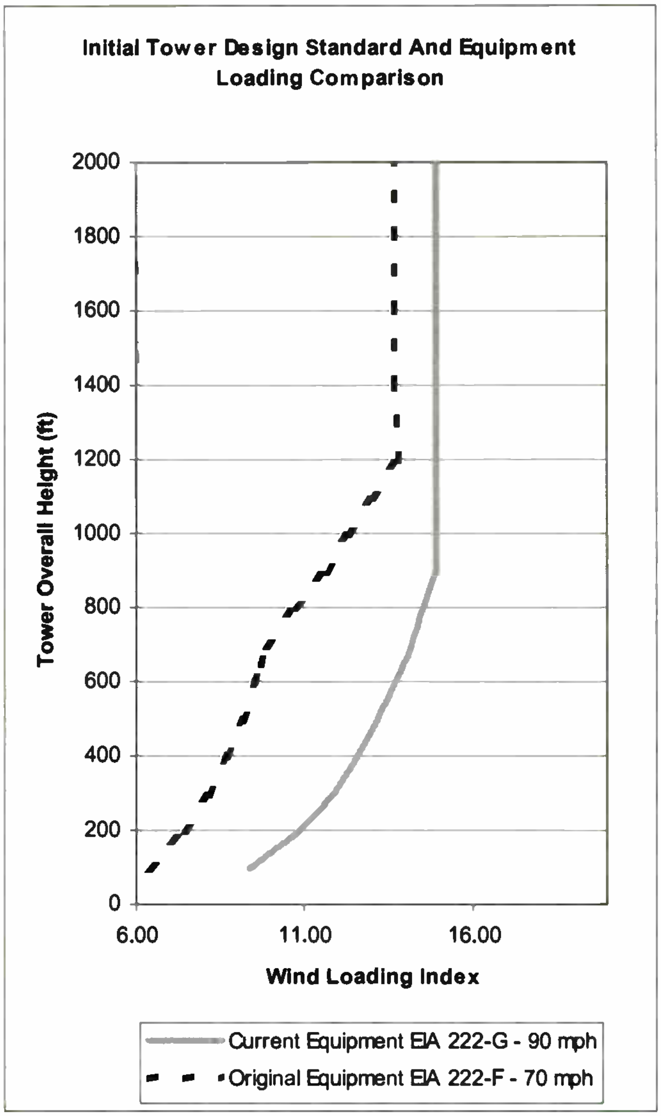

Initial Tower Design Standard And Equipment Loading Comparison
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Tower Overall Height (ft)
6.00
11.00
16.00
Wind Loading Index
Current Equipment EIA 222-G - 90 mph
Original Equipment EIA 222-F - 70 mph

Figure 22

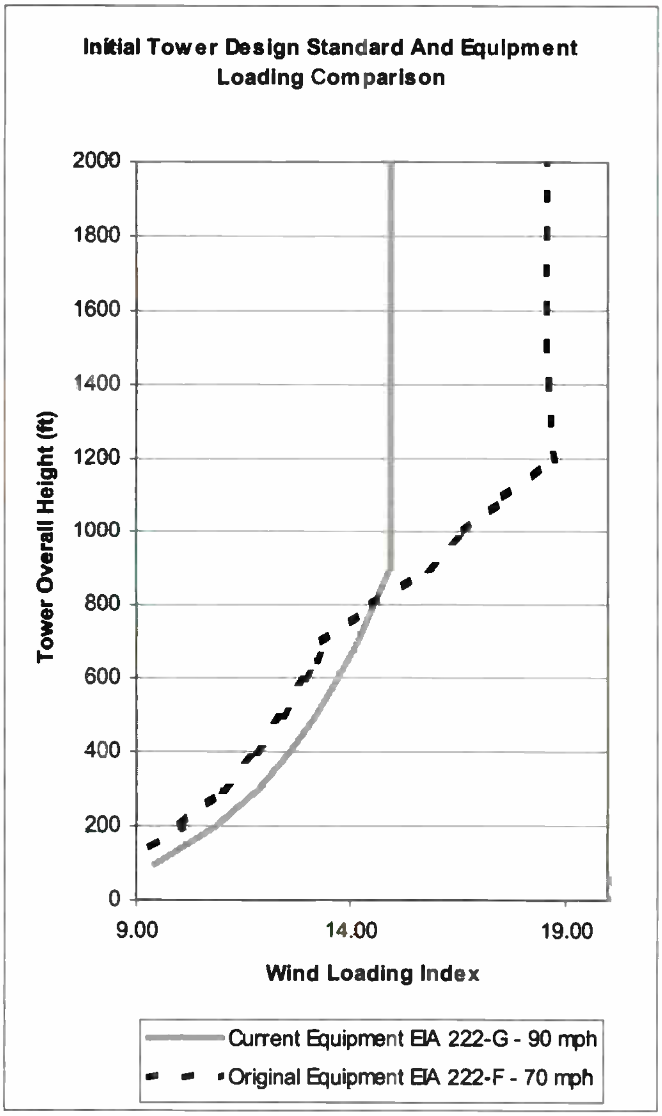

Initial Tower Design Standard And Equipment Loading Comparison
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Tower Overall Height (ft)
9.00
14.00
19.00
Wind Loading Index
Current Equipment EIA 222-G - 90 mph
Original Equipment EIA 222-F - 70 mph

Figure 23

ENG & ALL HAZARDS SAFETY

MATT BRUSKOTTER
The University of Findlay
Findlay, Ohio

ABSTRACT

Currently, in the United States, much funding and training is being offered to response agencies that could face various incidents from severe weather to hazardous material spills as well as terrorist incidents. Less consideration has been given to other groups who are first on the scene, such as the media, who respond for the purpose of informing the public. These news crews are at risk when reporting on these types of incidents and should be trained in how to remain safe and still get the story.

HAZARD? WHAT HAZARD?

As Don Hayford so eloquently put it; "Those of us who work in day to day, minute to minute news operations understand that news events have no schedules or pre-determined locations. And all too often we are going to an unknown location with not enough time to get there, let alone enough time to set up the shot. But we get it on the air anyway".

Identifying and understanding the hazards and risks in a work space can be a difficult task. A task made considerably more difficult for news crews who are working in the field. The myriad of environments in which they may find themselves each day can create a situation open to injury, exposure to hazardous materials, car accidents, raising the mast into power lines and violence from an angry crowd or person just to name a few.

News crews need to be aware that they could become a part of the story if not properly trained in how to recognize and avoid hazards in the field. Hazards and risks that may include downed power lines, fallen tree limbs, mud slides, sink holes and hazards posed by animal populations. Natural disasters such as tornados, earthquakes, hurricanes and tsunamis are real possibilities and present a variety of hazards to any personnel sent to report on either the event or the aftermath.

Chemical spills resulting from train derailments, accidents at chemical plants or other incidents can present exposure hazards as well as fire or explosion risks to personnel and their equipment. Terrorism targeted at First Responders such as fire fighters and police can also be directed at news crews arriving on scene. Secondary devices are of great concern to emergency services in the United States and will continue to be a consideration on many scenes.

The ability to identify hazards indigenous to the area in which these news crews live is typically taken for granted when dispatched to another part of the country. Many individuals learned a hard lesson about fire ants, heat stress, humidity and other hazards when dispatched to assist the areas hit by Hurricane Katrina.

A sound training program that provides the tools needed to make sound decisions in the field and increase situational awareness can pay immense dividends. It can result in a safer, happier and more productive work force and contribute to a continuous improvement of safe work practices throughout the organization.

ENG/SNG Safety

The requirements found under California OSHA Article 40 (California CCR Title 8 Section 2982 Subchapter 5 Group 2 Article 40) are familiar to the news gathering industry. These requirements provide an outline of the training required in California for "all employees operating or working in the proximity of ENG vehicles under field conditions". The training requirements also extend to supervisory personnel who "assign or supervise field personnel and have those individuals under their immediate and direct control".

The training must be conducted at least annually and includes many elements, all of which are covered in The University of Findlay's ENG and All Hazards Training program.

The training elements include topics ranging from overhead and downed power lines to wind, lightning and other severe weather conditions to safe vehicle operation to environmental hazards such as animals, insects, plant life and harmful substances.

Identifying and establishing safe work practices that are followed throughout the course of each work day can help news gathering operations maintain a safe work place for the employee. One of the other requirements of the California OSHA Article 40 regulation is the "establishment, implementation and maintenance of a written Code of Safe Practices for ENG Operations". This document is to be maintained in the ENG vehicle at all times.

Once again, these requirements apply only to California operations; however, they make excellent best practices that could be easily implemented across a national organization. These safe work practices could include actions such as pre-trip inspections of the ENG vehicle.

A simple pre-trip takes approximately five to ten minutes and can prevent a host of issues from affecting the crew's work day. Inspection of the vehicle's condition to insure that it is roadworthy, a check of emergency equipment such as flares, first aid kit and weather related equipment as well as the function of any safety or warning devices on the vehicle all work to insure a safe day in the field.

Other issues such as simply dressing for the weather conditions can play a large role in the safety and comfort of the crew in the field. Personnel are more likely to have an accident or become injured when they are not prepared for the conditions in the field. In cold or inclement weather, if an employee is not dressed appropriately, short cuts may be taken to reduce the amount of time spent outside. Further, the employee could be risking hypothermia, frost bite or heat exhaustion / stroke depending on the conditions outside.

Prior to being sent to a location in another part of the country, crews should investigate what hazards they may face. There are different hazards to be found in the various portions of the United States and being prepared for them will certainly make for a more comfortable and safer trip.

Other safe work practices such as positioning the ENG/SNG vehicle can provide more than one safety benefit. This particular safe work practice insures that the vehicle and its occupants are not a traffic hazard and will find a location that is out of the traffic lanes and flow of traffic. It also can help reduce or prevent exposure to carbon monoxide.

All Hazards Safety

The term All Hazards has been adopted by most of the country to encompass all manner of incidents including natural disasters, man-made disasters (large industrial accidents or hazardous material releases and civil disturbance or riots), incidents of terrorism, and global issues such as a pandemic whether it appears in the form of avian flu, SARS or some other virus or disease.

In an effort to prepare news gathering crews, The University of Findlay has developed several All Hazards training modules for the news gathering industry to compliment the ENG portion of the training program. These training modules cover a wide variety of issues that news crews in the field currently or could potentially face.

The primary goal of this type of training is to provide the trainees with the knowledge necessary to make sound decisions regarding safety while in the field before, during and after a disaster has occurred.

Floods, earthquakes, tornados and hurricanes are types of natural disasters where many health and safety risks may be increased. Disease, chemical contamination, unstable structures and encounters with indigenous animal populations are just a few of the hazards that may be faced during in the aftermath of an all hazards event.

News crews need to be aware that the danger in reporting on natural disasters is not encompassed entirely in the event itself. There are dangers before the disaster occurs as well as in the aftermath of the incident. For example, prior to a hurricane there may be evacuations causing traffic back-ups, accidents and with large amounts of people leaving an area there is always frustration that can boil over to violence.

Even before the hurricane makes landfall there will be increased winds and the potential for tornados to be spawned. Emergency services will be strained and immediate help may not be available if the news crew gets into trouble. Localized flooding may occur due to increased tidal activity or due to the large amounts of rain falling on the area. These flood waters can cut off escape routes, force a vehicle off the road, knock an individual from their feet and wreak havoc with electrical equipment.

During this period, animal populations will also begin to seek safety and potentially high ground. The probability of coming into contact with an undomesticated animal increases as they are typically seeking the same types of shelter the news crew is looking for.

Crews should avoid filming in open areas where they are being blown around by the wind. The hazard of high wind speeds is not solely in the fact that the wind is blowing. Many individuals are injured by the debris that is pushed around by the strong winds of a hurricane or other strong wind event. Crew members can be blown off their feet or struck by flying debris and suffer severe injuries. Filming from just inside the lee side of a building or just inside the door of the building itself is enough protection from the wind and still allows the crew to get a shot of whatever the wind may be pushing around in the background.

Developing tips for protecting the crews while in the field can provide a guideline and reference for determining where and what is safe while in the field. Each natural disaster presents its own set of hazards and the safe way to react may not be clear to the individual without proper training.

The aftermath of these disasters can be equally hazardous. These dangers could include unstable structures, ruptured gas lines, damaged industrial facilities that could be leaking hazardous materials, fires and potential civil unrest. All of these could present a danger to the crew let alone the potential to impair their ability to perform their job.

Unstable structures can present more hazards than simply the potential for collapse or ruptured gas lines and exposed electrical service. The amount of hazardous substances that can be found in the average home can easily reach several gallons. Home businesses or buildings undergoing remodeling can contain substantially more putting the crew at risk to exposure.

Older buildings and homes can also present unique hazards. Many of these structures were built using asbestos containing materials such as acoustical ceilings, pipe insulation, shingles, vinyl flooring, transite siding and plaster or wallboard. The asbestos could be "friable" meaning dry and able to be crumbled into a dust that is easily airborne leading to exposure and the potential for disease such as emphysema and asbestosis. Many of these structures also contain lead paint which can easily chip off of surfaces and become an airborne dust if the paint chips are crushed.

News gathering crews must also be aware that they could be the target of violence directed by a group of protestors, terrorists or others such as Eric Rudolph the Olympic Park bomber or the Uni-bomber Ted Kaczynski.

A major concern to first responders is the secondary device. In this scenario an initial device is detonated at a location designed to draw a large number of first responders. However, a second device is located in an area where these emergency service agencies are likely to stage. The second device is detonated or released upon their arrival creating a situation where those trained and designated to help cannot provide assistance to those who need it.

Likewise, the secondary device could be located in an area where the news gathering crews may set up. These crews could also be attempting to speak with first responders when the device is detonated.

Some training on typical agents or devices used by terrorists and others to enact violence on a population can provide crews with insight that can help protect them. Some basic information on topics such as biological and chemical weapons, radiological agents and explosives can help in identifying a device in the area. Knowledge on the basic routes of entry and signs and symptoms of exposure can also help crews to know when they may have become exposed, how that exposure can occur and methods that can be used to avoid the exposure.

Training on what to look for, identifying suspicious, unusual or out of place items or materials can help crews identify a secondary device before it is too late. This situational awareness can also assist them in identifying other potential hazards.

Other violence such as that exhibited by some protestors, rioters or other large groups of people can be equally hazardous and very unpredictable. Crews should work to avoid getting between the police line and the protestors. This is a no man's land where any violence that occurs will certainly either start or spill into and surround anyone in this area.

People beginning to show up with their faces covered are good indicators that violence could be about to start. Anonymity breeds violence in these situations because people feel less inhibited when the think they can not be identified. Crews should attempt to work in pairs so that each can watch any blind spots that may be created by carrying a camera, working around structures or vehicles or those caused by tunnel vision. At times in exciting situations individuals can become focused on the action in front of them without paying attention to what is happening on either side of them or behind. A partner can help keep watch on those areas to avoid being injured or targeted by the crowd or police.

The changing work environment and increased potential for harm means that any training program needs to be dynamic and updated frequently as new information becomes available. New technologies or variations of existing technologies and materials can provide new levels of safety for the crews in the field. Advances in monitoring technology for carbon monoxide to evaluate exposures and warn crews of increased levels of this harmful gas.

Materials such as personal protective equipment (PPE) can provide an emergency means of escape should conditions at the scene change faster than the crew can react. Appropriate, *escape only*, PPE such as a protective garment and escape hood can be used to provide the crew the opportunity to escape from an area that is suddenly contaminated or an accident or intentional release of a hazardous substance occurs. The PPE should not be used to enter any areas identified as or considered to be potentially contaminated as only personnel with appropriate levels of training are allowed in those areas due to the dangers posed by any hazards present.

Conclusion

Media personnel are typically among the first to arrive on the scene preceded only by emergency services such as police and fire departments. While those emergency services receive extensive training on a regular basis, the media personnel are very often provided with little to no training regarding the hazards present at the events on which they are reporting.

A training program should provide news crews with the knowledge necessary to avoid those hazards and recognize when to retreat to a safer location. The potential for harm coming to these crews continues to increase as new and emerging issues are identified, severe weather enters a period of increased activity and individuals and groups take more extreme or violent measures to make their voices heard.

With appropriate training and increased awareness, crews in the field should be able to increase their ability to remain safe while still getting the story and reporting the news.

Technical Planning for Emergencies

Tuesday, April 17, 2007
2:00 PM – 5:30 PM

Chairperson: Martin Hadfield
Entercom Communications Corporation, Seattle, WA

Access to Emergency and Non-Emergency Broadcast Information for People with Disabilities
Geoff Freed, WGBH National Center for Accessible Media, Boston, MA

Media Monitoring by the Military in the Electronic Age
Duncan Ayre, Department of National Defence - Canada, Victoria, BC, Canada

Wireless Data Services Save the Day at a Mount Hood Transmitter Facility
James Dalke, Dalke Broadcast Services, Inc., Bellevue, WA

Improve EAS by Adding Details to the Message
Bruce W. Robertson, Digital Alert Systems, LLC, Oracle. AZ

Police Communications and DTV, Pushing Data to the Field
Mark O'Brien, SpectraRep, Chantilly, VA

MTV Networks' Disaster Recovery Facility – Case Study
Lionel Hightower, Viacom/MTV Networks, Hauppauge, NY

Flywheel Technology as a Superior Solution for Reliable, Continuous Backup Power
Gary Rackow, P.E., Active Power, Austin, TX

Access to Emergency and Non-Emergency Broadcast Information for People with Disabilities

Geoff Freed
Carl and Ruth Shapiro National Center for Accessible Media
at WGBH (NCAM), Boston, MA

INTRODUCTION

People with disabilities need equal and timely access to emergency and non-emergency information and warnings about events in the community. This includes information relating to tornadoes, hurricanes, floods, tidal waves, earthquakes, icing conditions, heavy snows, widespread fires, warnings and watches of impending changes in weather, news of discharge of toxic gases, widespread power failures, industrial explosions, civil disorders, traffic problems and school closings. A primary source for this information-- television broadcasts-- does not consistently serve the needs of the 28 million people who are deaf or hard-of-hearing, or the needs of the 11 million people who are blind or have low vision.

The need for an accessible, consistent, reliable and redundant multi-platform emergency notification system that effectively serves people with disabilities is recognized by both the Federal Communications Commission (FCC) and the Department of Homeland Security (DHS). FCC rules (47 C.F.R. Section 79.2 (b) (1), established in 2000, require all broadcasters, cable operators and satellite television services to make local emergency information accessible to persons who are deaf or hard of hearing, as well as to persons who are blind or have visual disabilities. These rules apply to information given during regular programming, an unscheduled break, as part of continuing coverage or any other means of televising an emergency. [i]

This requirement is rarely met. Failures to comply occur so frequently that the FCC issued the document "Reminder to Video Programming Distributors of Obligation to Make Emergency Information Accessible to Persons with Hearing and Vision Disabilities" (Public Notice, DA 03-2361, July 18, 2003) that explicitly instructs broadcasters in the approaches that should be followed to make information accessible.[ii] Still, compliance is rare.

In addition to the obvious need to make emergency information accessible to people with disabilities, it is equally important to ensure that non-emergency information is accessible. A typical local newscast is made up of dozens of graphics conveying information to the viewer, including daily or weekly weather previews, maps, time and temperature statistics, lottery numbers, stock updates and sports scores. This type of information is integral to the viewer's experience and is frequently not repeated orally by the newscaster. In many cases, the newscaster may make a reference to the data, such as, "On your screen you can see the five-day forecast, and it doesn't look good!" or "Today's stock finals are on the screen, as you can see," leaving blind or visually impaired viewers at a total loss.

Two projects are currently in progress at the Carl and Ruth Shapiro Family National Center for Accessible Media at WGBH (NCAM) that are addressing problems related to access to both on-screen information (emergency and non-emergency) as well as emergency alerts:

1. **Access to Locally Televised On-Screen Information (http://ncam.wgbh.org/onscreen)**
   This project is exploring solutions to enable local television stations to convey both emergency and non-emergency information, conventionally displayed on the screen, in a manner that meets the communication needs of people with sensory disabilities. NCAM is, for example, investigating procedures for enabling real-time conversion of on-screen text into speech output and integrating this new audio seamlessly into the broadcast stream via the secondary audio program (SAP) channel or auxiliary DTV audio channels. NCAM is also addressing display conflicts between captions and on-screen graphics by developing methods of tagging and prioritizing text and graphics messages within automated display systems.

2. **Access to Emergency Alerts for People with Disabilities (http://ncam.wgbh.org/alerts)**
   The Access to Emergency Alerts project is addressing the need to develop and encourage adoption of standardized methods, systems and services to identify, filter and present content in ways that are meaningful to people with disabilities. The project is developing an information model that provides recommended

accessibility extensions to emergency system protocols, technologies and services for wired, wireless, DTV- and IP-based delivery; and is identifying key usability factors that must be addressed to serve people with disabilities, including cross platform and cross-environment issues. A public reference repository will be established for summary documents of user needs, design requirements for accessible products and services, and usability research.

BROADCASTERS' REQUIREMENTS AND CHALLENGES

In order to serve people who are blind or visually impaired, broadcasters are required to describe within the main audio all emergency information that is presented visually on screen during the newscast. When broadcasters present emergency information as a text crawl superimposed over regular programming, for example, they must make sure it is accompanied by an audio tone. This tone is intended to alert consumers who are blind or have low vision to seek information about a local emergency from other sources such as radio. However, this tone is rarely provided. And on the occasion when it is sounded, this solution clearly does not provide equal access. In addition to providing no information about the type or severity of the emergency (is it a thunderstorm or a biological attack?) a blind or visually impaired person may not have ready access to accurate information via alternate sources such as radio or telephone, and may not even be able to contact a neighbor, friend or relative to find out what emergency is in progress.

Similarly, in order to reach people who are deaf or hard of hearing, broadcasters are instructed to provide critical details about an emergency in a visual format such as open captions, closed captions, or a text scroll crawl. However, emergency text displays often end up blocking closed captions and/or closed captions block the emergency information, rendering the emergency information partially obscured or completely useless. While deaf viewers might know how to turn off their caption displays so that they may read an emergency crawl hidden by the captions, in the time it takes to turn off the captions (which may involve navigating through several choices from an on-screen menu) the emergency display may already be gone.

The importance of access to on-screen information is not limited to times of dire need, however. A blind or visually impaired person getting ready for work in the morning is just as likely to tune in to a local morning news program as a sighted person, and will be equally interested to know about the day's weather forecast, traffic tie-ups and school closings. Consider the following example:

There are at least three streams of important information available here: the weather, a list of school closings, and the time and temperature. The sighted viewer can take in the entire screen at once and instantly filter the desired information. A blind viewer, however, will have no idea that any of this information is available unless a) the newscaster reads it aloud, b) a sighted companion reads it aloud, or c) it is delivered in an alternative manner. Assume for the moment that the newscaster does in fact repeat the on-screen weather summary aloud, rendering that stream accessible. However, it is probably safe to assume that the long list of school closings shown at the bottom of the screen will not be read aloud, and will probably be referred to only once or twice during the broadcast when the newscaster says something such as, "The weather is playing havoc with local school schedules. See the list of school closings at the bottom of the screen."

OPPORTUNITIES PRESENTED BY DTV TECHNOLOGIES TO ADDRESS BARRIERS

The core technologies of traditional broadcasting have converged with those of computer networking, digital processing and the Internet to create dramatically new communication services and business models. National networks and local stations alike are transforming their operations into completely digital environments, often blurring the previously clear boundaries between television, radio, print and Web-based communication. As a result, broadcasters and programmers are rethinking every aspect of content creation, from design and display to distribution, control, archiving and asset management. Myriad new workflow concepts are being created, supported by electronic client/server and browser-based technologies and systems. Program content and information have become software objects, easily exchanged in faster-than-real-time sequences from network to station, among stations themselves, and even on- demand directly by the individual viewer.

With the advent of DTV and widespread use of software-based graphics and automation systems, there are opportunities to develop and integrate solutions into broadcast industry products and procedures that will greatly improve the delivery of televised public warnings and alerts to deaf, hard-of-hearing, blind and visually impaired people. Software that prioritizes caption data so that it may be relocated on the screen, or an application that transforms the text source of an on-screen crawl into an audio file, could be integrated within broadcast graphic and automation systems. An audio conversion of a text crawl could, for example, be transmitted via the SAP channel in today's analog broadcast environment or through one of the multiple audio channels available in digital television broadcasts. However, these capabilities are not currently integrated into broadcast equipment and local broadcasters have not developed procedures to identify and prioritize accessibility to critical information that is subject to FCC requirements.

Broadcasters and equipment manufacturers are increasingly concerned with interoperability-- the need to manage, customize and integrate different and sometimes conflicting software and hardware systems. New entrants into the marketplace confuse matters even further. The central engineering challenge is to guarantee that new broadcast systems can process and exchange complex information successfully, while mixing proprietary commercial products with custom software. Across these industries, there is a movement to develop harmonized standards to support complementary methods for data exchange. In recent years, common file formats for video, audio, text and data have been proposed, discussed and at times even adopted, but much work needs to be done to reach common ground.

Typically, the graphics and text appearing on a television screen are the result of a complex system of text, backgrounds, photos, logos, symbols and animated elements originating in a variety of graphics, video and word processors. These elements can be gathered together and assembled in a series of layers on the screen, including scrolling or crawling text, either synchronized with or in addition to associated audio and video. For the viewer at home, they provide everything from a simple identification of a speaker to additional program details, time, weather, stock quotes, breaking news, school closings and emergency information.

Often, local stations use multiple systems to generate a variety of on-screen text displays and services. Some may use a system to generate a single-row "headline news" crawl at the bottom of the screen during morning network news programs. This system automatically generates news headlines from the stories being followed and developed by the newsroom. A second text system might be fed by National Weather Service information, generating weather-alert notices and updates. A third system could be used to manage and display school closings and event-cancellation notices, based on information provided directly to the station by local schools and community organizations. Each of these systems uses a computer server that gathers raw data from the information source, and uses software "templates" chosen by the local station to select and format the data in text form for a related graphics generation system.

Once formatted, this text information is passed from the server to the graphics system, which creates a video element that can be selected by the station's master control for display on the screen. One common graphics system used by broadcasters is the DekoCast from Pinnacle Systems. The DekoCast uses a text file to generate all information that is in any way changeable in each "page" of design. This text file can originate from virtually any source that is available to the DekoCast.

With all this in mind, consider the same example discussed earlier-- a constantly changing list of school closings displayed at the bottom of the screen:

In order to make the list accessible to blind or visually impaired viewers, one approach would be to convert the content of the source file (a simple text file) to speech on the fly using text-to-speech software running on a stand-alone computer. Once the audio conversion has occurred-- a process that may take half a second or several seconds, depending on the length of the text file and the number or school closings involved-- this new audio file can be merged by the DekoCast into the broadcast stream at the same time the school closings appear on the screen. To avoid competition between this new audio stream and the regular program audio, the system would also automatically lower (duck) the program audio to a pre-determined level. Once the school-closings audio was finished, the system would raise the program audio back to normal levels.

NCAM is currently testing this theory; once tests are complete, descriptions and examples will be available at the project Web site: http://ncam.wgbh.org/onscreen.

NOTIFICATION OF EMERGENCY ALERTS

There is also vital work to be done to ensure that people with disabilities are provided with appropriate and flexible methods of receiving emergency information from television as well as other sources, such as radio, and there is a need for research and suggested structures for tailoring messages to be effective. For example, processing voice messages can be time consuming for blind users, which means the volume of information provided must ensure that consumers are well-informed without overwhelming them so much that the information becomes obscured by frustration. Further, users with disabilities may not be well-served by terse alerts that direct them to inaccessible sources, such as poorly constructed Web sites, for further information. And people with disabilities may require evacuation and recovery information that is substantially different than what is delivered to other consumers.

It is not uncommon for television stations to offer, via their Web sites, subscription services to notify viewers of weather bulletins via e-mail, text messages or voice mail. However, the Web sites where viewers may sign up for these services are often poorly accessible or completely inaccessible to users of access technology, such as screen readers or screen magnifiers. In addition, new digital television and consumer-electronics-based public alert systems are designed to produce multimodal output, as needed by a range of users with sensory disabilities. However, many of these systems are themselves embedded within consumer equipment that cannot be operated by blind or visually impaired users: they often require the ability to use touch screens or on-screen menus, for example, something blind and low vision users cannot currently do unless the devices provide speech output not only to convey content but also to provide navigational information.

There is a critical need for a collaborative effort to research, develop and disseminate practical solutions for these problems in concert with newscasters and equipment manufacturers. Since September 11th, 2001, the need for equal and effective access to televised communications about public safety has been widely acknowledged on the policy level, but there has been little progress made to develop the capability to provide it within broadcast systems and procedures. Most local newscasters are truly committed to and interested in serving the needs of all the constituents in their community, and are themselves frustrated by lack of attention to this urgent need within current technologies. Broadcast technologies and procedures must be modified to better meet the needs of people with sensory disabilities for accessible warning and alerts related to local emergencies or to other important community information.

About NCAM

The WGBH National Center for Accessible Media (NCAM) was founded in 1993. NCAM acts as the research and development arm of WGBH's Media Access Group and is involved in technology, policy and program development to assure that the nation's media and technologies are fully accessible to people with disabilities. In 2006, the Boston- and Palm Beach-based Carl and Ruth Shapiro Family Foundation donated funding to support NCAM's efforts, and in recognition of their contribution, the center has been renamed the Carl and Ruth Shapiro Family National Center for Accessible Media at WGBH.

NCAM is an extension of public broadcasting's ground-breaking work in media access that began in 1972 with the establishment of The Caption Center at WGBH and its groundbreaking development of captioning for deaf and hard-of-hearing television viewers. More recently, in 1990, public broadcasting's access mission resulted in the development of video description for blind and visually impaired audiences. NCAM and its sister organizations, The Caption Center and Descriptive Video Service® (DVS®), make up the Media Access department of the WGBH Educational Foundation.

NCAM strives to make media more accessible in schools, the workplace, the home, and the community. In addition to a focus on the retrofitting of existing media, such as television, radio, newspapers, and theatrical movies, NCAM is designing access into emerging telecommunications such as digital television, convergent media and Web-based multimedia. NCAM's mission is to ensure that the 45 million Americans with little or no access to media's sights and sounds will not be left out of the Information Age.

[i] *Accessibility of Emergency Programming*; http://ftp.fcc.gov/cgb/dro/caption.html

[ii] *Reminder to Video Programming Distributors of Obligation to Make Emergency Information Accessible to Persons with Hearing or Vision Disabilities*, http://www.fcc.gov/cgb/dro/emergency_access.html

MEDIA MONITORING BY THE MILITARY IN THE ELECTRONIC AGE

Duncan Ayre
Department of National Defence
Canadian Forces Base Esquimalt, Maritime Forces Pacific Headquarters

INTRODUCTION

We all know what happened on September 11, 2001.

Shortly afterwards the Admiral commanding Maritime Forces Pacific met with the families of the crew of Her Majesty's Canadian Ship Vancouver, a Canadian Patrol Frigate then deployed to the Persian Gulf as part of the Canadian contribution to the campaign against terrorism. During this briefing, the Admiral said to the families "I am sending your loved ones into harm's way. I don't know what will happen while they are out there but we will tell you as much as we can as soon as we can. But in these uncertain times, you will see it on TV first before we get the chance to tell you."

From that moment our office knew that we would have to monitor the news like we never had before.

Our office is the Public Affairs section of the Maritime Forces Pacific, which is the West Coast Canadian Navy. We traditionally got information out to the media through news releases, media advisories and a website. Due to our relationship with the media and the need to verify that which was reported, we had already positioned ourselves to perform a more intensive media monitoring function.

We started with two VCRs, one radio with a tape recorder and a CRT TV. Other then passing the audio and video tapes to the people who needed to see them, we had no way to efficiently distribute the information to those who needed to see it.

We have now grown to 7 PVR computers, Sling boxes, a radio scanner, LCD and plasma TVs, CD and DVD burners, thumb drives, the ability to capture local radio broadcasts digitally, and to capture audio and video off the internet. We receive AP, Canadian Press, and Reuters news wires, and RSS feeds from over 100 sites.

To distribute information to the media, we have stand-alone fax machines, a portable fax machine, computer fax cards, e-mail services from 4 different providers, Blackberries, a website with mirror servers, FTP services, access to emergency radio channels and in extreme emergencies a group of runners to hand deliver reports to the local media. Hopefully with all of these resources we can get the information to the media in a timely manner.

The Canadian military is divided up into a number of different reporting lines. At the top is National Defence Headquarters, the Canadian version of the Pentagon. There are of course the elements; Navy, Army and Air Force and their respective bases.

We also have divided up the country into different joint task forces; Atlantic, East, Central, West, North and where I'm from: Joint Task Force (Pacific). This division allows the commanders in their individual areas to quickly and efficiently respond to civil emergencies including floods, major forest fires, devastating earthquakes or any other emergency that the civilian authorities require military assistance for. The JTF commander can order any military asset within their area to assist the civilian authorities under this arrangement.

Within Joint Task Force (Pacific) the major base is Canadian Forces Base Esquimalt that houses Maritime Forces Pacific, the command for the Navy in the Pacific. Throughout all these commands there are various departments that receive general and specific media reports from our office providing them with better situational awareness. These range from my boss (the senior public affairs officer) to the Admiral, the Commander of Canadian Fleet Pacific, Construction Engineers, the Maritime Operations Centre, the Joint Rescue Coordination Centre and the Joint Operations Centre to name but a few of the section that receive information.

At National Defence Headquarters in Ottawa, Ontario, there is a section for media monitoring. Unfortunately they have limitations due to equipment and staffing that allow them to currently provide only transcripts in a timely manner. On many occasions our office has been asked to record and distribute broadcasts from television and radio stations as far away as Halifax, Nova Scotia. Many bases across Canada have media monitoring capabilities, but to a much lesser extent than what our office currently has. We are currently working on improving everyone's ability to capture broadcasts within their own area.

THE PROCESS

We have two separate but similar systems to handle media monitoring. We have a bank of PVR computers preprogrammed to record the various news broadcasts that happen outside of normal work hours. These recordings are reviewed every morning with the items of interest edited out, and copied to our internal network. These clips are then catalogued and a short description of each clip and link is sent to those people on our base who by position would be interested in the subject of the report. The description of the clip is also sent by email to those individuals within the military outside of our location with a link to an internal FTP location to download the selected clips from.

The second system is an almost real time distribution system. Due to security restrictions, a PVR system cannot be attached to our internal network. To counter this we have a second bank of PVR computers maintaining watch over the main 24 hour news channels available in Canada; CNN, CBC Newsworld and CTV NewsNet. These systems all have their own LCD/Plasma TVs constantly on during the workday. When we note a story of interest, we start recording, doing line editing as the story progresses. On completion we transfer the clip to the internal network using a thumbdrive. The distribution system is the same as the overnight clips, just sent from a different email account.

For real time events that occur outside of normal office hours, the operations centre at the base, which is manned 24/7, have TV's constantly on the news channels that allow them to maintain a watch for events. They do not have the same capability to record and distribute a clip, but can recognize an event and start informing the senior commanders if the event warrants.

We monitor one local and one regional talk radio broadcast using streaming audio and an AM/FM radio antenna attached to our computer. This allows us to digitally capture audio broadcasts and distribute them quickly to those people who are interested for their situational awareness.

Most senior commanders have Blackberries allowing communication when they are away from the office. They may not be able to see the video clip but will know by the short description what the item is about. If they need further clarification they often call our office asking for further information.

After a major event, where there are multiple broadcasts about the event, we produce a CD of all the audio/visual clips, allowing a full review of the media coverage now and in the future. We also maintain a database of all clips so that past reports can be reviewed.

RSS feeds have become integral to our ability to quickly receive breaking news. We receive feeds from over 100 different sites, covering all types of information.

The National Oceanic and Atmospheric Administration's West Coast & Alaska Tsunami Warning Center and the Alaska Volcano Observatory both send out alerts using RSS. In the past we would surf every few hours to both of these sites, plus we were on their respective e-mail/fax list. Now that we have subscribed to their RSS feeds we receive alerts wherever they provide an update.

RSS feeds cut down on our workload and stress from having to constantly search websites for updates and has improved our speed-reading ability from having to scan the over 1600 messages that we receive daily.

We have subscribed to Google News Feeds using a number of keywords for our searches. We have then subscribed to the news RSS feed allowing for specific news articles to be sent to our RSS reader. This system has assisted us in finding obscure stories of interest that we would have not normally have found so easily.

We are interested in local stories about the military.

Military Public Affairs Officers generally work with reporters on local items in the media. For this reason we record local stories to check for clarity in the report, if the information has been relayed correctly or if there is any slant to the story. On occasion a Public Affairs Officer does contact the reporter to clarify or correct any misinformation that has been reported.

Many times bloggers and non-mainstream media publish articles and videos about the local military. These stories are sometimes controversial and are often more of a concern then the mainstream local media. In most cases the person doing the report has not contacted the Public Affairs Office, and thus we have no advanced knowledge of the report, nor are we able to provide correct information to the reporter. These reports are captured for review and for our situational awareness. With many of these stories, there is no response from the military, as it could serve only to inflame the issue.

We are also interested in regional stories.

These are generally the realm of JTFP: stories dealing with emergency preparedness, potential or actual disasters, major changes to transportation routes and the local government and community's relationship with the military. This information is brought in for review, comment, and situational awareness and/or for planning purposes. During an actual emergency a Public Affairs Officer will be part of the deployed task group. We have a pack-up box containing all the material needed to set up a remote office and maintain communications between the deployed Public Affairs Officer, the media and our office.

The role of media monitoring during a natural disaster would be to provide the decision makers in the military with media reports, giving them an added source of information for their situational awareness. As we have seen during Hurricane Katrina and the confusion around the Superdome, the media can sometimes have important information that the local powers don't or reports that substantiate information that we currently have. These added pieces of information are an important part of the situational awareness that the commanders must have to make an informed decision.

We are also interested in national and international stories.

With national events we coordinate with National Defence Headquarters media monitoring group to see what they need assistance with. These stories generally are of national interest and will have a wide distribution. We have recorded a number of reports coming out of Afghanistan last year, as our current boss was the Public Affairs officer for the coalition headquarters in Kandahar.

The Army bases from which the troops are currently deployed are recording most of the stories from Afghanistan. We only take an interest in stories out of Afghanistan if a local person is in the report. In a number of cases this has been the death of a member from our area. In these cases we record and catalogue all the media clips until after the funeral and make a presentation CD for the family.

The Canadian Navy coordinates with many countries around the Pacific in military exercises, United Nation missions and humanitarian assistance. When a ship from our base is deployed we are constantly searching for news items about their activities. This includes mainstream news items, blogs and other services such as YouTube and Google Video services. Not only are these report captured for our records they are also forwarded to the organization in charge of that activity.

With media events in North America we are interested on how the participants handle news conferences and how the media report the event. We also look at how the event affects the ability to get the information to the media.

During the start of Operation Iraqi Freedom, a Major at Fort Bliss, Texas had to hold a news conference to announce the death of seven members from that base. It now unfortunately has become commonplace, but at that time it was new to everyone. We recorded the event and passed it on to senior command and the Public Affairs training cell so that those officers who would have to go through the same event would have some knowledge as to the questions put to the speaker.

MEDIA MONITORING AND CRISIS COMMUNICATIONS

In the aftermath of Hurricane Katrina, General Honore gave a media brief and uttered the now famous line, "You are stuck on stupid." This event is now part of the training for Public Affairs officers to recognize when a senior commander is having trouble dealing with the media and to assist the senior officer to get the information needed to the media before such incidents happen.

The need for recording and communicating what is being reported came to the fore during the Chicoutimi tragedy. HMCS Chicoutimi is a Canadian submarine purchased from the Royal Navy. After a refit in England she was being sailed over to Halifax in the fall of 2004. During the transit a fire broke out on board. The fire caused a serious disruption in the communications systems leaving the crew with only a satellite phone to communicate to the outside world. The call for help was received by Search and Rescue in Ireland and relayed to the Canadian Navy and Search and Rescue base in Halifax, Nova Scotia.

During the ensuing hours media from Ireland, the UK and Canada as well as from around the world picked up the story. The media was able on a number of occasions get access to information that we in the military did not have or which had not been released. The only way the senior commanders and the public affairs officers knew what was being reported was through the efforts of our office on the west coast of Canada as our counterparts on the east coast were too busy handling the media, family of the crew and trying to do media conferences with timely information. Our office was able to communicate with senior Canadian commanders at National Defence Headquarters in Ottawa, Halifax, and in the UK to advise them of what information the media was reporting. This added information allowed them to adjust and respond to media inquires quickly and efficiently.

We have an exercise every quarter, to prepare for an event that in all likelihood will never happen, a release of nuclear particulates from a visiting submarine or aircraft carrier. When we first started running the exercises we had slide after slide of what the senior leadership determined to be important information. Once we started reviewing what others have been presenting during their respective emergencies, we have now done away with all PowerPoint slides. All that we will be presenting is what the public needs to do to be safe. The detailed information on how the accident happened and all the related information on radiation doses etc. can be presented after people are safe.

During these exercises, we practice dealing with the media, sometimes using real media. Other times we have others play media who grill us with questions both in person and over the phone. We practice getting the right information out and answering their questions without causing the general population to panic but keeping the media on the topic of what the public should do now. Once the initial media conference, (informing the media that there is an incident) is complete, the main communications will come from the Provincial Emergency Program, which has the responsibility to determine the evacuation area and routes and determine the shelters to be used.

In this type of major emergency we assume that the normal communication networks will become useless as unnecessary traffic jams the system. To communicate with the media, we will use emergency radios, which the media monitor and we will conscript various members of the military to run messages between our Joint Information Bureau and the major media offices in our area.

Numerous events across North America have demonstrated that the public communication systems can easily fail. The shootings at Douglas College in Montreal, Quebec this past September caused the cell phone networks in that city to be disrupted. Although we are all told to stay off the phones in emergency situations, many people ignore this in order to contact loved ones.

A couple of years ago there was a mild earthquake in Victoria BC. During one of the morning radio call in shows just after the shaking stopped, some yahoo called in asking why his house was shaking. We cannot depend on people listening to the authorities and keeping off the phone so that emergency calls could get priority.

Planning must take into account that after the first few minutes in an emergency we may not be able to depend on our normal communication systems. Other alternative methods must be planned for and exercised.

The most likely major event on the West Coast of North America that will knock out normal communications will be an earthquake. Depending on the damage that is done, normal communications could be disrupted for days if not weeks. Many radio stations have generators, but as we have seen from a series of major storms that hit Vancouver Island in November and December 2006, power to whole regions can be cut from a small tree limb falling onto a power line in one area.

There are two sets of submarine cables supplying 65% of the power requirements for Vancouver Island. These lines could be easily damaged during a major earthquake. Local hydroelectric dams supply the rest of the power to the Island. How these dams and infrastructure will hold up in a major earthquake is unknown.

We work in an earthquake resistant building with shock absorbers, reinforced concrete, and earthquake kits for all personnel. Generators supply emergency power to the building. The biggest concern, after initial survival, is getting to the building if the earthquake occurring outside of work hours. There are only a few bridges connecting major parts of the city to the area where the base is located. With a failed infrastructure, the ability to respond to emergences outside of work hours will be challenging.

If the worst happens and the majority of the communication infrastructure is rendered unusable, we will depend on the Navy ships to communicate to the outside world. All the major warships are equipped with communication systems that allow them to communicate around the world. The military would play a major role in providing support to the Provincial Emergency Program. Communicating to the media would again fall to the Provincial Emergency Program officials with our office providing whatever support we can.

THE WAY AHEAD

In February 2010 the Winter Olympics will be held in and around Vancouver British Columbia, a short ferry ride from our offices. There are expected to be over 5,000 Olympic Games athletes and officials; 1,700 Paralympic Games athletes and officials; 10,000 media representatives, 3 billion worldwide television viewers and an unknown number of visitors.

The Canadian Forces will be working closely with the Royal Canadian Mounted Police, the lead security agency for the games. Part of our offer is to provide media monitoring services to the security group. This would be exactly what we do now, except expanded to capture more channels that may have news of interest. As described before this service provides the decision makes with added situational awareness that they would not normally have.

The reports that we would be interest in would not be the actual sports as these would all be from one pool but would be the human interest and filler stories about the people and area as well as the preparation and organization of the Games. These stories and live broadcasts would provide insight into any event of concern that is occurring.

We also are part of the Amber Alert system. When an Amber Alert is raised we send out a message to all the networked computer users on the base with the relevant information. This assists the police with an extra 6000 set of eyes as we commute to and from work.

With over 6000 people who work on base, we are very connected to the surrounding community. We raise 13% of the United Way's annual campaign goal, providing $65,000 worth of toys to the Salvation Army, countless boxes of food and funds to the local food bank and supporting many other charities in our community.

There have been many cases of members of the Canadian Forces outside of work helping in situations that have be covered in the media. From assisting with Habitat for Humanity, saving a child from a fiery traffic accident, to helping deliver a baby on a roadside. All these actions and more have been captured by the media and by doing so have allowed us to bring these Canadian Forces members to our Admirals attention. This allows the Canadian Forces members to be recognized by the leadership for their actions.

We are constantly reviewing new products and services to improve our productivity and the services that we can offer. Currently we are reviewing portable video media players, portable ebook readers, voice recognition capturing systems and a variety of screen capture software. Some of our options are limited due to security restrictions. We have a modified version of Microsoft Windows 2000 Professional as our operating system and must live within the limitations of this environment.

Media Monitoring has an important role to play in providing an overall situational awareness to the commanding officers who make the decisions on how best to help the civilian population in an emergency situation. Being able to capture all the reports and turn them over to the officers in a timely manner does and will greatly assist making the right decision during the fog of battle, be that with a foreign adversary or a civil emergency.

WIRELESS DATA SERVICES SAVE THE DAY AT A MOUNT HOOD TRANSMITTER FACILITY

JIM DALKE
Dalke Broadcast Services, Inc.
Bellevue, Washington

ABSTRACT

This is an account of the installation of an FM transmitter facility for KPFR on Mount Hood in Oregon with 15 feet of snow on the ground and no landline telephone facility. Remote control for the transmitter was originally a simple interface with a cellular telephone connected to the remote control system, using touch-tone commands for control and voice response for monitoring. Eventually, a control system was developed for the site with a wireless wideband data service provider with Internet access that not only provided flexible transmitter control and monitoring, but real-time program delivery as well.

Timberline at Mount Hood

THE CHALLENGE

Late in 2004, Family Radio in Oakland asked for help installing a new FM station in Oregon. Family had a Construction Permit for the facility that was set to expire in six months. They also said the site location listed in the CP was unsuitable, and asked if a better site could be found. Pine Grove, Oregon was listed as the city of license. When looking at the FCC Tiger Map locator referenced in the application, it appeared the transmitter site was located between two glaciers on Mount Hood. Referencing USGS topographic maps clarified the location. The site was in fact on Mount Hood, but located between ski runs, not glaciers, about a half mile west of the historic and picturesque Timberline Lodge, at an elevation of 6000 feet on the South side of the mountain.

The city of license, Pine Grove, is located in arid Eastern Oregon, 30 miles east the transmitter facility. The town is not much more than a wide spot on State Highway 216 with a population of 162.

The transmitter site is literally at timberline on the mountain. On a reasonably clear day, you can see Pine Grove to the East, and metropolitan Portland 60 miles to the West. It became obvious why this was an ideal transmitter site with some 2 million people within view in the Willamette Valley to the west and southwest.

A suitable building was located for the transmitter facility, a twenty by twenty foot wood "A" frame structure with two forty foot poles at each end for antennas. The building housed vhf and uhf repeaters for the Oregon State Police, Clackamas County Public Safety, and several other two-way radio and paging companies.

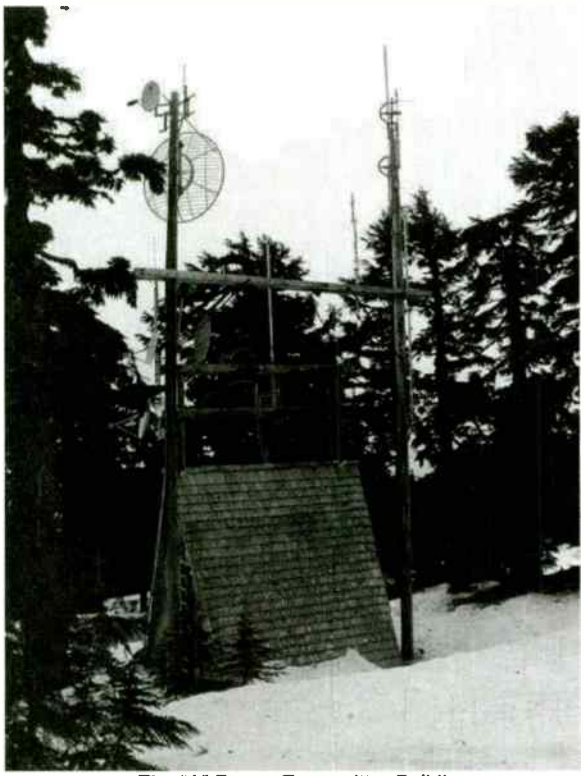
The building is located on US Forest Service land in the Mount Hood National Forest. The building had adequate 200 amp, 240 volt underground power service, more than adequate for the 5 kW transmitter. The good power service was in part a result of the location between ski lifts on the mountain. The Mount Hood ski facilities are among the best in the Northwest. Timberline has the longest ski season in North America with one of the chair lifts taking skiers up to the 8,500 foot level into July and even August for skiing and snowboarding on the mountain glaciers.

No Landline Telephone

One problem persisted: no landline telephone service. Timberline Lodge, a half mile away, had T1 facilities, providing telephone and data service for the staff and guests of the 13,000 square foot rustic lodge. The entire thousand acre Timberline facility is on Federal Forest Service property, and as might be expected, is considered a fragile alpine environmental area. All of the power and communications lines are underground. To get new underground telephone service would have been prohibitively expensive and likely would not be permitted, to say nothing of the time required or digging ditches in the snow.

After the transmitter site had been located and secured began the task of getting equipment delivered and installed. An electrical contractor in Hood River, 50 miles northeast of the site on the Columbia River provided a good staging location for equipment deliveries and the electrical contractor was able to provide transportation for the equipment to the Timberline Lodge maintenance facility which was located about a mile downhill from the transmitter site.

Once delivered to the maintenance building, the equipment was transferred to sleds that were towed by snowmobiles to the transmitter building. The largest and heaviest delivery was the 900 pound, 5 kW Nautel transmitter. After delivery by truck to the maintenance facility, the transmitter was loaded on the back of one of Timberline's utility snow cats for the last mile over the snow up the hill. The snow cat is normally used for transporting the ski facility maintenance crews and equipment to support the nine ski lifts spread over a thousand acres of ski area.



The "A" Frame Transmitter Building

There was over 10 feet of snow on the ground when the installation began. When the snowmobiles were not available for taxi service, the half mile hike from the Timberline Lodge Parking lot to the transmitter building was good aerobic exercise, especially at 6,000 feet. The air is noticeably thin.

Program Delivery

KPFR is licensed as an educational station and operates under a waiver of the FCC rules requiring a local studio in the city of license (Pine Grove, Oregon.) The station operates as a satellite of its parent station, KUFR-FM in Salt Lake City. Most of the programming for the station is delivered by satellite from Family Radio Headquarters in Oakland, California. The station also is required to provide local programming which is produced at the KUFR facility in Salt Lake City.

Similar to most networked broadcast stations, most of the programs for the Family stations originate in Oakland, California and are distributed by satellite to their 50 plus stations scattered around the country. The original plan called for a 3.8 meter dish to be installed at the site for receiving Family's network programming. The 20 plus feet of snow in mid-winter meant a ground mounted dish was out of the question. Mounting a large dish high enough on the aging wooden pole structure over the A frame building would have been risky at best with 100 mph winds from winter storms. Heavy snow and icing presented further problems.

The problem was neatly solved when it was discovered that a 1.8 meter dish used for program reception from an earlier tenant turned out to be pointed at the correct satellite and the received quality was more than adequate for a primary program source. The rain fade that can plague reception in other parts of the country with a small dish, occurs rarely in the Northwest. Because of the heavy snow fall and low temperatures, anti-icing equipment was suggested for the minimal antenna.

The station was licensed for a single bay directional antenna, so the site was surveyed for proper orientation and the transmitting antenna installed. When the new antenna was connected to the transmitter, there was literally no reflected power. The transmitter SWR circuit was carefully checked for proper operation, and it was confirmed: it was a perfect match from the start.

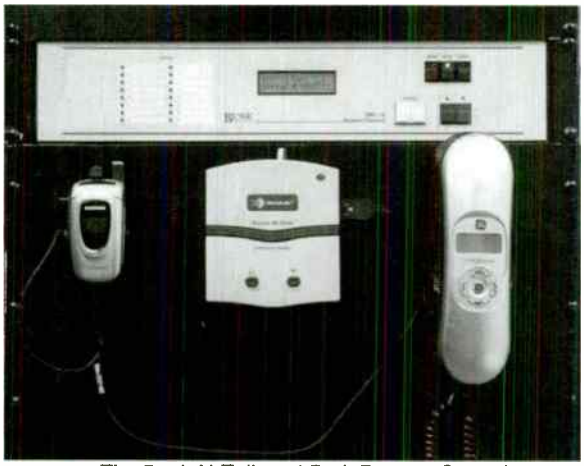
THE WIRELESS CONTROL SOLUTION

With the transmitter now on the air, the only real FCC requirement for transmitter remote control for the site was the ability to simply turn off the transmitter in the event of a malfunction or at the request of the FCC. A Burk ARC-16 with the ESI (Enhanced Speech Interface) option installed provided the remote control interface. The ESI normally allows a regular land-line telephone line to be used for transmitter control and monitoring. With the ESI, touch-tone commands can be used to interrogate and control the transmitter power. A simple, limited vocabulary speech response system provides appropriate prompts and return readings. But there was no land-line available for the Burk.

In the early part of the site search, it was obvious the area had good cell phone coverage. With the tourist traffic in the summer and ski enthusiasts in the winter, this was cell phone territory. It seemed with the proper interface, the cell phone system could provide the necessary connectivity for the site.

After extensive research into the problem, a device called "Dock-N-Talk™" was installed. This device, little larger than a paper back book, had an interface that plugged into the accessory jack on the cell phone on one side, and a RJ11 jack on the other. It also had a mini jack that would interconnect with the headsetmicrophone jack on the cell phone. With a standard corded phone plugged in, it operates exactly as if it was plugged into a wall jack of a land-line.

The cable connected to the cell phone accessory jack is used to send an RI (Ring Indicate) signal to the Dock-N-Talk when an incoming call was received and an OH (Off Hook) signal when the corded phone was picked up. These signals are provided in many of the cell phones compatible with the Dock-N-Talk.



The Dock-N-Talk and Burk Remote Control

With an incoming call on the cell phone, the RI signal would trigger a ringer signal that is sent to the corded phone. When the corded phone handset was picked up, a signal went back to the cell phone to answer the call.

On an outgoing call, the handset is picked up on the corded phone and the DOCK-N-TALK would provide battery and generate a dial tone. As the touch-tones are dialed on the corded phone, the Dock-N-Talk would store the dialed digits, and when the dialer quit sending tones or ten digits are received, the Dock-N-Talk forwards the dialed digits digitally to the cell phone through the interface cable.

Except for a short delay when the DOCK-N-TALK redialed the touch tones, the corded phone user is not aware of the cell phone service. The Dock-N-Talk people suggested that one of their devices could be plugged into a regular telephone system in a house for example, and when the resident arrived home, he could just connect his cell phone to the Dock-N-Talk, and any corded phones in the house could be used as if connected to a land-line.

When contacting the Dock-N-Talk sales people about dealers in the Seattle area, they said there were a number of dealers. All of them were boat dealers or ship builders. It turned out the Dock-N-Talk has been very successfully marketed to the yachting industry. Yacht owners can install corded phones throughout the boat, and when they come aboard, they connect their cell phone to the Dock-N-Talk, and all of the corded phones on board operate like their landline connected phones at home.

The Dock-N-Talk worked almost flawlessly with the ARC-16. The only problem was occasional errors with the touch-tone commands from the remote telephone. The digital processing of the analog touch tones in the cell phone causes distortion that in turn, can cause errors in the relatively unsophisticated touch tone detector of the ARC-16. A minor modification in the detector circuit improved the detector performance.

This approach to interfacing the cell phone with the ARC-16 would probably have been much better with the original, first generation (1G), analog 900 mHz AMPS (Advanced Mobile Phone System) phones which have been replaced with the all digital cell phone network. In fact, some of the original analog cell phones were equipped with an RJ11 jack for the purpose of plugging in a corded phone or slow speed modem.

With cell phone access to control and monitor the transmitter installed, the station was fully FCC compliant and licensed. A simple FM radio was installed at a location 30 some miles from the transmitter near Portland. An auto-coupler was installed so a remote telephone served as an off-the-air monitor.

SECOND GENERATION

The Dock-N-Talk system provided adequate control and monitoring of the transmitter operation at the beginning, but another problem required a little more sophistication. KPFR is licensed as an educational station, and is authorized to operate without a studio in the city of license, Pine Grove, OR. The station is obligated however to provide local programming for the community.

Local programming for the station is produced at Family Radio's Salt Lake City facility, KUFR and then broadcast in KPFR. Delivering this local programming to the KPFR site without land line access posed a significant problem. Manually delivering a CD with the local programming is a real problem with twenty feet of snow covering the transmitter building. The solution was to move on to second generation (2G) cellular data technology.

As the cellular industry has transitioned from analog to digital technology over the last several years, the cellular network has also had the bandwidth and capacity to handle non-voice data. Several modem manufactures have developed product to use this new wireless data network to provide data services with voice phone service.

These 2G modems originally provided service ranging from 9.6 to 64 Kbit/s, significantly more than the 2400 baud of the analog system. This opened up a whole range of applications for wireless data service, from telemetry and remote metering applications in the energy production industry to digital transportation dispatching for emergency services and truckers. The 2G modem provided Family with the ability to download their local programs produced in Salt Lake City directly to the KPFR transmitter site.

When the installation contract was completed with Family for the KPFR project, the third generation wideband data services were just emerging.

THIRD GENERATION

These new third generation cellular data systems can be truly classified as wideband wireless data services. The latest wireless speeds are approaching T1 delivery rates, allowing real-time audio program delivery to backup the present satellite delivery system as well as downloading remotely produced local programming.

Some of the latest 3G offerings from the major cellular companies include Cingular (AT&T) with their launch of "BroadbandConnect," and Verizon Wireless and Sprint's EV-DO networks. These high speed networks can not only delivery high quality stereo audio, but video content as well.

Generation	Mode	BPS	Connection
1G	Analog	2400	dial-up
2G	Digital	64k	dial-up/internet
3G	Digital	1 mb	internet

Wireless Data Services Bandwidth

One of the latest entries to the wideband wireless market is a Seattle based company, Clearwire. Clearwire is using the 2.5 gHz IPTV band and recently began marketing their product in several major markets around the country. The Clearwire approach to the wireless market is antithetical to that of the cellular companies. They are providing wideband data services first, and will soon be offering VOIP telephone service as an optional add-on to their wireless data services.

This 3G wideband data service has also become known as WiMax. WiMax is cellular delivery of WiFi over a much greater geographical area and, of course, subscription based.

EMERGENCY BACKUP APPLICATIONS

In December 2006, the worst windstorm in a decade hit the Northwest. At the transmitter facility in Kirkland, Washington that is used by Family Radio KARR and nighttime facility for Radio Disney KKDZ, all power and landline communication was destroyed as falling trees brought down a quarter mile of power and phone cables.

Windstorm Damage Blocking Transmitter Site

It took the utility companies two weeks to clear the debris and restore power and communications. During that time, a generator provided power for the KARR transmitter, but the only communication and data into the site was the cell phone and a wide band data connection.

Had the windstorm damaged the network satellite dish, the site would not have been able to broadcast. The wideband wireless services continued to operate and in those areas where there was damage to the wireless facilities, they were the first to be restored, and were available for real time program delivery.

The areas cellular telephone service continued to operate through the emergency. There were certainly some delays with the high call demand, but it was possible to communicate.

NON-BROADCAST APPLICATIONS

The development of wideband wireless has many applications in the broadcast industry and beyond. It is interesting to speculate on how the industry has developed from providing basic wireless telephone service to a broad base of wireless data applications, and now even video on demand.

In the research process for this paper, some of the more unique applications for wireless wideband data services include the following two examples:

Wimax Espresso

Concordia Coffee Systems is a Bellevue, Washington company which manufactures automatic espresso machines. (It is surely only a coincidence that the company is not far away from the ubiquitous Starbucks coffee company headquarters in Seattle.) Concordia's espresso equipment can operate in a fully automatic vending machine mode or it can be operated by an attendant. Concordia claims that at the push of a button, the machine automatically grinds fresh whole espresso beans to make the perfect espresso shot while the milk is steamed to make a perfect cappuccino, latte, or just black coffee in less than a minute.

A wireless modem is installed in the machine to provide real-time data such as drink counts, equipment diagnostics and programming that allows the operator to precisely control the perfect latte. In addition, a remote paging system notifies the operator with system warning alerts when the machine needs milk, beans or cash to continue operation virtually eliminating machine downtime.

Wireless Parking Meters

Recently, the city of Seattle began a program to remove all of the city's parking meters, and replace them with 1,600 wireless pay stations. Each pay station has an integrated solar power panel on top and a wireless modem. The wireless connection allows the city to have real time communication for maintenance, payment information, credit card verification and usage statistics.

CONCLUSION

The wireless communication technology has progressed from the analog cell phone to today's wideband digital technology with incredibly small phones and wireless modems with speeds well into the megabit range.

Virtually every metropolitan area and interstate highway corridor now has access to the new technology, and coverage extends to many remote areas as well.

This paper has been a glimpse at just one segment of this new technology in broadcast applications, including transmitter control and monitoring as well as audio program and content delivery.

The technology can serve in those areas where the conventional landline telephone is unavailable and as a backup in event landlines and satellite connections fail.

Biography

James "Jim" Dalke is a veteran Broadcast Engineer. Jim is a Senior Member of the Society of Broadcast Engineers, a Certified Professional Broadcast Engineer and a Certified AM Directional Specialist. Jim has served the past several years on the executive board of SBE Chapter 16 in Seattle. Jim holds Amateur Extra Class Radio License W7PB and is a Coast Guard Licensed Merchant Marine Radio Officer. Jim is also an FCC Licensed Maritime Radio Inspector. Jim's recent projects include the only licensed broadcast station in the United States on a ship with temporary transmitter facilities for KKOL, Seattle, installed on a 175 foot cargo ship. Jim holds several patents on television, radio, and communications inventions and he has presented several papers at the annual NAB Engineering Conference. Jim also serves as a volunteer for emergency communications services with the Bellevue Washington Fire Department and provides technical support for multimedia services at Westminster Chapel in Bellevue.

Improve EAS by Adding Details to the Message

Bruce W. Robertson
Digital Alert Systems, LLC (DAS)
Oracle, Arizona

ABSTRACT

The need to improve EAS is obvious. It's most apparent shortcomings are the lack of detail in EAS messages, the lack of involvement by state and local authorities, and the fact that it does not utilize the power of the Internet. Most improvements to emergency communications are aimed in the direction of newer technologies. However, the recommendation to be presented here by Digital Alert Systems focuses on a new method to improve the protocol of EAS with established technology, which we believe can economically and efficiently bring EAS in line with the Internet, provide the message details EAS has always lacked, and extend the life and value of what is our most widely implemented emergency communications system. This presentation defines an open "method" called TDX, or Textual Data eXchange, and comments on different applications.

Background

The federal government maintained and perpetuated the EAS mandate after the latest Notice of Proposed Rule Making (NPRM), where it stated, *"We agree with those commenters who argue that EAS should remain an important component of any future alert and warning system."* Additionally, in his June 2006, Executive Order, President Bush's directions to the Secretary of Homeland Security stated, *"administer the Emergency Alert System (EAS) as a critical component of the public alert and warning system."*

The timetable charts in the 1994 EAS Report and Order mandated that broadcast stations and cable systems implement EAS, and stated that EAS "may be used to provide the heads of state and local government with a means of emergency communications." A very small percentage of states had submitted any formal plan to the FCC by the deadline for state plan submission. You would think that a government-mandated plan would include designated emergency management personnel. To make EAS "a critical component," state and local authorities must invest in its operation.

If we are going to keep EAS, as it appears we are, then we need to consider how it can be improved. Our recommended method functionally adds details to the EAS message without changing the protocol, without violating the FCC rules, without making presently installed equipment obsolete, and without objectionably adding to the EAS tones. It can be incrementally incorporated into any EAS configuration without requiring changes to the remaining EAS components, and without the need for additional federal criteria.

If indeed this enhanced method can improve EAS as stated, broadcasters may be able to use this "improved EAS" to solicit a more active participation by state and local government authorities. This participation is crucial to any real improvement to EAS and is a critical component in any future emergency communications scenario. One hundred percent of emergencies that could benefit from the use of an "improved EAS" are local or state events. EAN (the National Alert), if it is ever used, has been defined by, and belongs to the President. This method does not interfere with the EAN, as it does not change the protocol, does not violate the ruling, and does not render the installed base obsolete. Implementation is left completely to the state or local authorities. I am sure most broadcasters would consider helping to expedite and assist in this implementation.

TDX defined

Operationally, our proposed method simply uses presently defined EAS header tones to create an audio data packet, and places that packet at the front end of the audio envelope in the EAS message. The data packet can contain additional event-specific text or it can contain one or more URLs. If the packet contains text it makes that text available for use or forwarding, completely independent of the Web. The duration of packets containing text will of course vary depending on the amount of text.

The duration of the packet containing one or more URLs is approximately 1.5 seconds. It sounds like an additional header burst so it does not objectionably add to the EAS interruption. This method can be fully tested during weekly (RWT) or monthly (RMT) tests. With the RWT it sounds like four header bursts. With RMT it sounds like an additional "quack" after the two tone. This information could contain pictures (as in AMBER or evacuation alerts), additional text, and any additional audio with source-related accuracy.

The method leverages unused bandwidth in the EAS protocol. Very simply, with this method, EAS becomes IP compatible by placing IP-compatible data inside the EAS message.

Nonproprietary technology

TDX was defined and tested by Digital Alert Systems, LLC, and is offered as an open and published method to invite constructive comment or criticism, make changes as may be required, develop a competitive economic base, and encourage a broader implementation to improve EAS.

NPRM response

We could spend considerable time quoting and referencing various individuals and groups regarding how to improve EAS. As of this writing, there have been 431 replies, comments, letters, and other responses to FCC NPRM Proceeding:04-296. The important essentials are most satisfactorily covered in the three responses from the Society of Broadcast Engineers (SBE). The SBE comments are recommended reading, and provide an intelligent, logical, and functional look at EAS.

We have summarized the NPRM responses as follows:

- EAS is in place so don't do away with it.
- To improve EAS add text to the EAS message.
- Real EAS improvements require a very active commitment by state and local authority.

Ken Putkovich's reply to the NPRM stated, *"There was nearly unanimous agreement that the current EAS doesn't work very well, the EAS serves a useful purpose and should be fixed, a single Government department should be responsible for national emergency warning matters, and the Department of Homeland Security/Federal Emergency Management Agency is a logical choice."*

Ken's comments, included to support our conclusions, identified an agency he feels should take the lead. Hopefully that agency will accept. To realize a timely benefit from an improved EAS, we encourage immediate action by state and local authorities, which means these authorities must in the end use EAS.

EAS for people with disabilities

Groups like CPB/WGBH National Center for Accessible Media, the Rehabilitation Engineering Research Center for Wireless technologies (RERC) at Georgia Tech, and Gallaudet University are working to define more effective emergency communications for people with a variety of disabilities. While these discussions are by necessity very broad, most look to EAS or "improvements" to EAS as the authority on alerting and informing. Information provided in the current EAS message must be questioned and supplemented by even the most able, so the availability of more information in the EAS message would naturally help these groups to better define worthwhile methods. Improvements like DEAS (Digital EAS being developed for FEMA in partnership with the Public Broadcasting Community by SpectraRep) or HazCollect (a current effort of the National Weather Service) are focused on making the additional information they can carry available and responsive to this community.

The data packet used by the proposed method could reference a URL containing information aimed at a particular group in this community. The EAS authority working with one of the groups above may be able to effectively configure such a URL.

Testing in Kansas

On Friday August 18, 2006, Sedgwick County Emergency Management (SCEM) in Wichita, KS, met with area broadcasters, National Weather Service personnel, and Tom Wood and Bruce Robertson from Digital Alert Systems to plan area-wide testing of a preliminary implementation of the proposed method. The testing, which started in October 2006, consists of SCEM encoding the enhanced EAS messages that are output on a VHF transmitter for reception and decoding by KAKE-TV, KFDI-FM/KFTI-AM, KNSS, the National Weather Service, and the SECC at Broadcast Technical Associates.

Dick Elder, meteorologist-in-charge at the National Weather Service, Wichita, KS, commented, *"TDX gives emergency managers an effective tool to get information out to media outlets and other agencies on potentially life-threatening situations. It is reliable and multi-faceted, which allows it to interact with many different platforms. In this day and age where technology, especially in the area of communications, is expanding and growing almost daily, having a system that can support and change to meet customer needs is vital."*

We appreciate all participants in this testing and plan a full report for distribution at NAB. John Crosby, Emergency Management Operations officer, has taken the lead in this testing and was instrumental in helping to write and edit the following on Emergency Operation Centers.

Emergency Operations Center (EOC)

Emergency managers take advantage of any and all available means to get their message out to the public. Broadcast faxes to newsrooms, radio alerts aimed at the media over Public Safety Radio, e-mail lists, and even direct phone calls with automated notification systems are all utilized. However, none of these systems take advantage of EAS. You may wonder why this is the case, but the answer is really quite simple – most jurisdictions don't have a local EAS plan promoting a single- to multi-point warning system. Many look at the EAS system from the perspective of its original design – to allow the President to address the nation in times of emergencies. Reality, however, demonstrates that all emergencies start locally and good emergency managers can take advantage of an improved EAS method to make sure the public receives adequate information about the warning.

Systems can fail – telephones, cellular phones, radio systems and other means of communicating information to the public are all vulnerable. That's why it is so important to take advantage of every communications tool possible to get vitally important information to the public – including multiple and redundant systems. It is also important that these multiple and redundant systems be operational and in place before the disaster strikes.

Let's consider the example of a chemical release from the XYZ Chemical Company – just down the street – which happened a few moments ago as the result of a power outage in the area. Normal control of the facility is not possible and crews are responding to take action; however, a hazardous substance is being released. Many of the systems that would be called upon to get the warning out to the affected area have been compromised by the power outage. Residents of the community know that they should monitor their battery operated radios and televisions when there is such an outage.

The Incident Commander at the chemical release was just told that manual control is not working and the chemical release is going off-site to the surrounding neighborhood. He requests that a Shelter in Place Warning (SPW) be sent out by all possible means. The EOC utilizing the enhanced EAS method provides details to define "Shelter in Place" by making those details available to the local broadcasters. This creates an appropriate warning message, including audio and graphical elements (like a plume dispersion depiction) to help warn the public.

As conditions change, the Local Area Emergency (LAE) event code could be placed in another EAS release. Additional EAS messages could be released using the Administrative Message (ADR) event code providing additional information through the data packet used by the enhanced EAS method as conditions change. Additional ADR EAS messages could cancel the warning and advise when it is safe to return to normal activity.

The flexibility of this enhanced EAS system allows for a number of new uses. For example, use of the SPW and LAE EAS codes ensure that the information goes directly to the public and to the media (if there's a functional local EAS plan in place). Utilizing the EAS codes in this fashion allows distribution of the most important information directly to the public – and the follow-up information under the ADR releases is available to the media to pass along in their format. EAS decoders without the enhancement are in compliance by responding to the SPW and LAE event codes and ignoring the ADR releases if so programmed. The key issue to making this flexible system work is the local EAS plan. Without that, you might as well yell the warning information out the window.

DEAS (Digital EAS)

The AlertManager system from SpectraRep is a next-generation emergency notification system that enables emergency managers to generate EAS with Common Alerting Protocol messaging, along with audio, video, and multiple file attachments. This system is being deployed by FEMA in a pilot project in nine states.

The proposed enhanced EAS method can be used in such a system to enable EAS to include more of the critical information that a CAP message may contain. In other words, without the enhanced method, an EAS encoder/decoder will not be able to forward information that an emergency manager is attempting to relay using the Common Alerting Protocol. For this reason alone, the new method can be a critical bridge between the legacy FCC Part 11 EAS system and next-generation systems already being deployed by FEMA in IPAWS, the National Weather Service in HazCollect, and by systems integrators like SpectraRep with its AlertManager.

National Weather Service (NWS)

One of the National Weather Service's latest efforts is HazCollect, which is defined as follows:

"The HazCollect system is a comprehensive solution for the centralized collection and efficient distribution of Non-Weather Emergency Messages (NWEMs) to the NWS dissemination infrastructure, the Emergency Alert System (EAS), and other national systems. Emergency managers will use the Disaster Management Interoperability Services (DMIS) desktop client to write NWEMs in Common Alerting Protocol (CAP) format and send them through DMIS for further processing.

DMIS will send authorized and authenticated CAP-formatted NWEM messages to the new HazCollect server for conversion to the NWS World Meteorological Organization (WMO) format. The HazCollect server will send the NWEMs to existing NWS dissemination systems, including the Console Replacement System (CRS) wherein these messages are broadcasted by the NOAA transmitters and subsequently heard by the general public via NOAA weather radios."

This is certainly a step in the right direction as it uses the NWS dissemination infrastructure to communicate Non-Weather Emergency Messages in Common Alerting Protocol. The way I understand it, this system communicates detailed information as an EAS "statement." At one time the FCC told me a "statement" (i.e., Severe Weather Statement – SVS) is talk between NWS personnel, so no one else should be concerned about enabling a decoder to respond to a "statement." Here is all the information in a defined format, and except for the Console Replacement System, which does provide audio directly to the public over NOAA weather radio, most of this information would be lost in an EAS message. The data packet used by this enhanced method could identify these data locations, making that information available not only to the NWS personnel, but to any decoder site enabled to decode the packet.

SBE considerations

I again acknowledge the SBE's experience, knowledge, and application of logic to provide EAS guidelines by referencing their January 2006, comments that stated, *"It is SBE's view that the EAS has reached a point where simple modifications or band-aid approaches are no longer applicable. The most recent R&O augmenting the EAS is a clear call for the application of additional technology to not only correct existing issues, but make changes in the way the systems works in order to enable the EAS to move forward and serve more U.S. citizens via an ever-growing number of electronic communication systems."*

In addition to other recommendations SBE made a proposal that the Commission consider the adoption of CAP (Common Alerting Protocol), creation of EAS performance standards, elimination of broadcasters as an EAS origination source, and federal funding and training for a national system.

Conclusion

Our recommended method bridges the most widely implemented emergency communications system, EAS, with today's most widely used communications channel, the Internet. Placing event-specific details in the EAS message directly addresses the numerous requests for "the addition of textual transmission capability." Those details can contain URLs, making all event-related data, gathered and used in the EAS release decision, immediately available at the decode site. Making all event-related information available means, or at least certainly implies, that the proposed method makes EAS IP-compatible. If this improvement moves EAS this far into the digital domain, then with a little handling it can be CAP compatible. Can you imagine an EAS message communicating pictures, unlimited additional details as text, and accessible event specific audio?

EOCs should implement this method as part of an EAS plan. Economically this puts a cost burden on the EOC and reduces the cost burden for broadcasters who may upgrade their system, or decide not to enable the packet decoder.

There is no doubt EAS is broken, but it can be fixed. We can't ignore the potential and value of the installed EAS base, and if the numerous recommendations to improve EAS are to be more than lip service, it is up to us to consider how EAS can be improved, and it is up to us to instigate the implementation. Emergency messages are often labeled as critical, crucial, or life saving, and if these adjectives really apply to the message then they should apply equally to the methods we incorporate to communicate the message.

What have we said?

1. EAS is in place so don't do away with it.
2. Adding text to the message will improve EAS.
3. Real improvement requires active participation by state and local EOCs.
4. EAS performance standards need better definition for state and local participation.
5. CAP adoption is critical.
6. Broadcasters and EOCs need to start an active dialogue aimed at passing the EAS baton.
7. DHS and FEMA should consider funding to improve EAS and the future EAS environment.

We have at least three ways to go:

1. Operate as we have and continue to improve critical communications independent of EAS.
2. Improve EAS and use that improvement to actively encourage state and local participation.
3. Wait until DHS or some other entity tells us what to do.

EAS is in place, EAS is mandated, and EAS can be made better. To improve emergency communications in your area, please consider this method to bring EAS in line with today's communication abilities.

Thank you to NAB for making this presentation possible, and to broadcasters, state or local EOCs, EAS vendors, and other interested parties for their time and consideration.

Police Communications and DTV, Pushing Data to the Field

MARK O'BRIEN
SPECTRAREP
CHANTILLY, VIRGINIA

ABSTRACT

While first responders already have good communications infrastructure, they are typically bandwidth constrained and, many times, cannot communicate with other agencies at a response scene. Digital Television is increasingly being used to bridge this gap by providing access to critical data in the field.

The Clark County School District police are using DTV datacasting to deliver content, including live security camera video, directly to their vehicles. This paper will provide a detailed look at that system and explore what is possible in other markets.

JUST TALKING AMONG FRIENDS

Interoperability is a term that comes up recurrently among first responders. Police cannot talk to the paramedics, fire cannot communicate with the hospitals, and the list goes on.

Most existing first responder communication uses cloistered radio frequencies and has limited bandwidth. While they are increasingly converting to digital transport, data rates are typically very low (9.6kbps), which limits the ability to send files and especially video.

WiFi and WiMax are on the horizon to help address this problem. The cellular carriers are building high speed data networks and proprietary systems are emerging. However, every transmission system has its own advantages and disadvantages. The challenge is to find a way to take advantage of the positive elements of each system, while minimizing the impact of the limitations. Sometimes this results in a system of systems, rather than a single solution.

Columbine

On April 20, 1999 fifteen people were killed and twenty three wounded at Columbine High School in Colorado. It was the worst school massacre in American history. While school violence did not start on that day, it was a wakeup call for the country.

Communities began evaluating what they were doing to prevent and combat this type of threat and most found many areas that could be improved. In Las Vegas, the Clark County School District partnered with Vegas PBS (KLVX-DT), the local public television station, to provide information directly to the school district police.

Using datacasting technology, the station is able to deliver security camera video from inside school buildings to squad cars outside. They can also send data directly from the school computers so that responders have the latest information on attendance, special needs, response plans and other critical information.

CLARK COUNTY SCHOOL DISTRICT PILOT

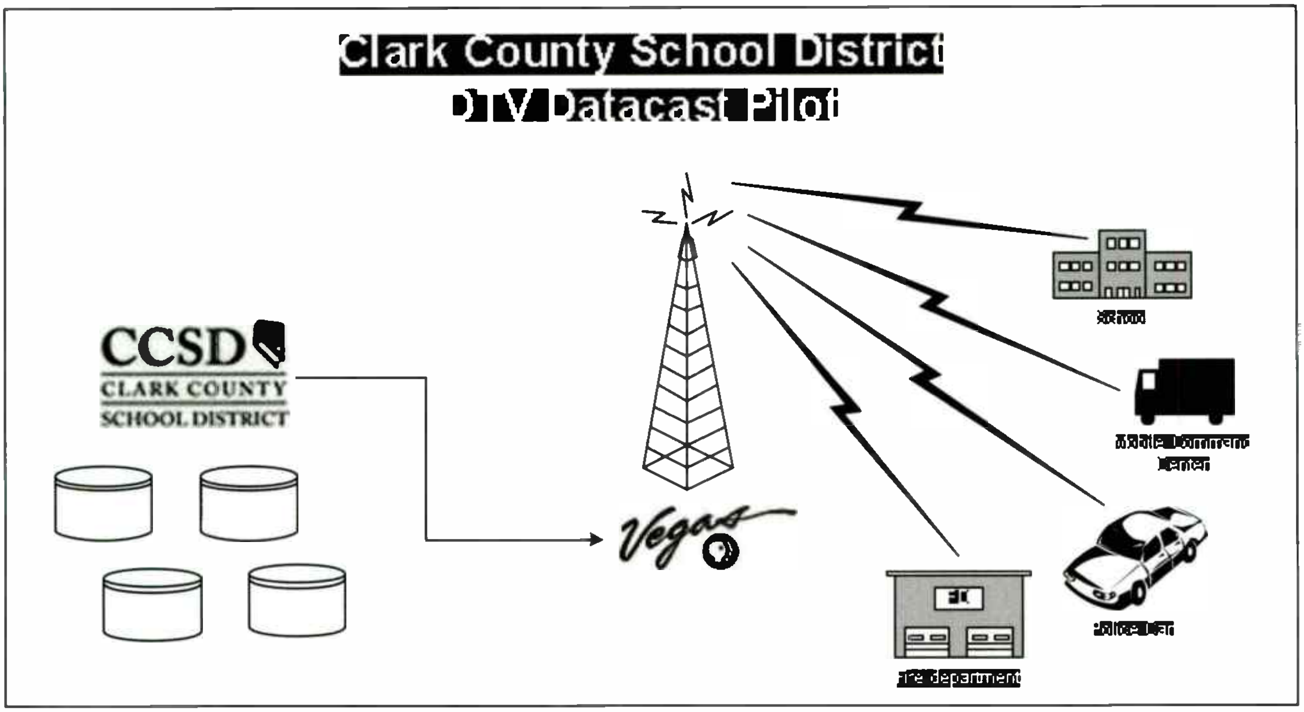
KLVX-DT is the local public broadcasting member station for Las Vegas. Many years ago, KLVX installed datacasting equipment and experimented with ways to use this technology to help the community.

Recently, they were awarded a grant from the Corporation for Public Broadcasting to initiate a service that will help with school safety. This pilot project is well underway and is already showing benefits to the community.

In the pilot, approximately 75 county school police vehicles are or will be equipped with DTV data receivers. Our challenge was to receive IP content delivered over the KLVX-DT signal in unknown and unpredictable receive locations.

Initially, we assumed that a smart antenna that could electronically orient itself would be required. These antennas exist, but are rare. The EIA/CIA 909 standard is one smart antenna specification. Much of this intellectual property was acquired by ATI not long ago and they have not pushed for mass production.

We experimented with other "smart" antennas and found that they all worked exceptionally well. The challenge was going to be mounting one of them on a car. They tend to be bulky and have high wind resistance. To our surprise, we discovered that a simple omni directional whip antenna provided enough signal strength to receive the station.



The problem with omni antennas is that they are just as sensitive to multipath coming in from a bounced signal as they are to the main signal.

5th Generation to the Rescue

We also discovered that 5th generation receivers had improved to the point that they could deal with the multipath being received by the omni-directional antenna. We expected this would not hold up as we tested further away and in more challenging locations, but it did.

This is partly due to the fact that, unlike most digital stations, KLVX broadcasts its digital signal on VHF. VHF waves are larger and tend to go around obstacles better than UHF. Regardless, the results were impressive.

Content Management

Part of the challenge was to identify, aggregate and organize useful content so that it could be distributed. This turned out to be more difficult than we expected. When an incident occurs, responders need information fast. They do not have time to wait for someone to find what they need and then find a way to get it to them.

Through interviews with responders and school officials, we attempted to identify the data that would most likely be needed in an emergency. Identifying the data was easy, finding who controlled it was harder and actually getting access to it was the most difficult. Some of the most useful data is also the most dangerous if it gets into the wrong hands. We had to go to great lengths to maintain security and protect the private nature of this sensitive information.

A system is being put in place now to allow individuals with proper access to identify, locate and distribute critical data during a time of need.

DTV DELIVERS

Digital television offers many advantages when delivering content to the field but, like every other distribution mechanism, it has its limitations too. Let's take a look at why DTV datacasting is of interest to public safety officials.

Why Datacast?

The Internet is obviously well suited to IP data distribution, so what value does datacasting add? The answer typically comes down to two issues, access and reliability. The internet may eventually become ubiquitous. WiMax in particular holds the promise of eventually providing metropolitan wide access to the Internet. However, it is not available today and, even if it were, the Internet is a useful research and information tool, but it is not well suited to delivering critical response information to mobile users.

The Internet is, for the most part, a unicast medium. This means that each user on the Internet gets his own (unicast) stream. This works perfectly for things like email, web surfing and other applications where each user is working with different content.

However, the only way to deliver video over the Internet is to send an individual stream to each end user. As the number of users grows, so does the bandwidth and server requirements. For example, if 1,000 people try to watch a live event at 1 Mbps, you need 1 Gbps of bandwidth and servers capable of serving that many streams. This becomes even more problematic as audience size increases. Eventually, you run out of capacity and cannot serve any more users.

One needs look no further than September 11, 2001. In both New York and Washington, DC it became almost impossible to initiate a cell phone call or get to many web sites. This is not because either of those networks failed on that day. It is because they became so congested that, for all practical purposes, they became unreliable.

DTV datacasting takes advantage of the same multicasting nature of broadcast TV. It does not take any more bandwidth to reach 1,000,000 people than it does to reach 1 person. As a result, it is impossible to overload the system capacity.

Digital television is also wireless, which is a benefit for many applications. With DTV, if you can receive the television signal, you get the benefit of high bandwidth no matter where you are. This is especially useful where no wired infrastructure exists, like police vehicles in the middle of an outdoor incident scene.

One of the Internet's strengths is that it is 2-way. Television is a 1-way push medium. It is possible to create a two-way network by using similar technology to the old cable modems. These systems used a telephone line for upstream communication and the cable feed for downstream. This same model works with datacasting. We can use low data rate packet radios to communicate to a server and then broadcast the requested data over the digital television signal.

Because security is always a concern with public safety applications, we can take this one step further and eliminate the digital back channel altogether.

One advantage of a one-way push delivery system is that it cannot be hacked. After all, you cannot hack a server if you cannot get to it. It is also possible to blend these models and provide no direct link to the datacast content server, but through voice requests over existing radios, have someone at the server manually initiate pushing content out to the field.

In Las Vegas, officers inside the network will have the ability to initiate data broadcasts as soon as they know that an incident has started and also respond to voice requests from officers in the field.

What is datacasting?

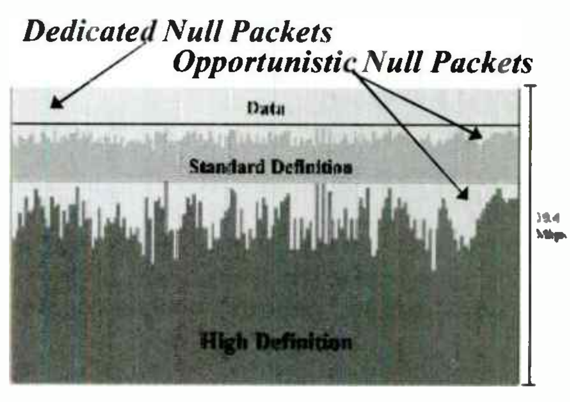
DTV datacasting, is the process of inserting IP computer data into an ATSC digital television transport stream. This data can be real-time IP video or file transfers to a client PC hard drive. Simply put, anything you can send over the Internet, you can send over a digital television signal. Just like the Internet, you can layer additional technology into the mix to encrypt content, conditionally grant access, manage digital rights, etc.

Data broadcasting uses digital television's MPEG transport packets, but is not limited to MPEG video like standard broadcasts targeting a DTV television set. In fact the receive device is typically not a TV at all. It can be a computer, special set top box or dedicated appliance. As long as it has an ATSC digital tuner that supports data, the IP packets can be identified and saved.

Just like you can have multiple MPEG video services at the same time, you can have multiple data services running concurrently. This means that multiple files going to the same customer or different content designated to multiple customers is possible.

Datacasting uses the null packets that are left over after adding MPEG digital TV content to the transport stream. The most efficient bandwidth utilization comes from using just unclaimed null packets. However, this opportunistic data insertion has limitations and may not be supported by all IP encapsulators. With opportunistic data insertion, file transfers will speed up when more bandwidth is available and slow down when less is available.

A more common approach is to allocate bandwidth specifically for data broadcasting. Because video is encoded at a specific data rate, this is a requirement if you are streaming video that is meant to be watched in real time. For example, if your IP video was recorded at 1 Mbps, then you must make 1 Mbps of null packets available in the stream or you will drop frames.



Digital television receivers are designed to display MPEG-2 video. When a packet comes in that has IP data in it, the television receiver ignores it.

Unlike digital television receivers that can only be used to watch TV programming, some data receivers can also decode DTV signals so you can watch digital television on the computer screen.

This can be quite valuable to public safety officials. Imagine not only being able to receive relevant computer files on a laptop in the police vehicle, but also being able to watch local news coverage of an event. Many times helicopter video being broadcast live conveys more information to first responders than they get through their own sources.

IP Video vs. MPEG Video

Since digital television is already sending video and audio, why would you want to send IP video over a digital television signal? There are several cases where it makes sense to send video programming as a DTV datacast, rather than as TV programming.

First, DTV television programming is designed to be viewed on a digital television set. IP video is designed to be viewed on a computer. This simple distinction allows us to "hide" IP data inside an ATSC transport stream. Televisions at home have no idea it is there. In fact, the only way to even see that it exists is with a properly configured data receiver or a stream analyzer.

When the video is sent as IP, you gain all of the advantages of any other IP content. It becomes encryptable, storable and routable. IP video is also more compressible than MPEG-2 digital television video. The IP video landscape is very competitive with new and improved codecs being released regularly. DTV MPEG-2 video quality is also improving, but at a much slower pace. In short, IP video at 1 Mbps is roughly comparable to DTV MPEG-2 video at 3 Mbps.

Security

Another advantage of IP data distribution is security. Television broadcasters want the largest possible audience watching their programming. However, when you are sending corporate information, police, fire and other sensitive data, you may want to limit who can see it. This is as easy to do with DTV datacasting as it is with a wired distribution system.

The IP packets are encrypted before they are transmitted, distribution can be tightly controlled. This can be done using anything from free software only systems, all the way up to high-end hardware solutions that meet strict government standards. You can also additionally encrypt the MPEG transport layer, making it even more secure.

Security options also include receiver targeting and conditional access, which uses passwords, smartcards or biometrics to allow access to the content by some people, but not others. Emergency content might be targeted differently to police, fire, rescue, battalion chiefs, decision makers, etc. You can also target content to specific individuals, rather than groups of users.

Reliability

A standard digital link assumes two-way communication and uses that connection to acknowledge each packet as it is received. This ack (acknowledgement) or nac (negative acknowledgement) tells the server if the client received the packet or if it needs to be sent again.

Since there is no return path for ACK/NAK, forward error correction (FEC) must be used to make sure that all packets arrive at each of the clients. This can be combined with carouselling, sending the file more than once, to assure delivery.

By increasing the redundant data in each transmission (FEC), we can assure that the file will be received reliably even if there are signal dropouts during transmission. The amount of FEC required to assure that all packets are received depends on reception conditions. 10% FEC is usually plenty, but it can be increased if some clients are susceptible to occasional signal dropouts.

File carouseling works by scheduling multiple transmissions of the same file. If a packet is dropped on the first pass, the client can pick it up the second time around. Only the missing packet needs to be received on the second transmission, not the whole file. FEC and carouseling can be used at the same time to increase reliability. Using these systems, files can be sent reliably even in the worst reception conditions.

Common Alerting Protocol

Common Alerting Protocol (CAP) is a new file structure for emergency alert massages. It is an XML schema that dictates how compatible messages must be formatted. The advantage is that properly formatted messages can then be read and understood by any CAP compliant device.

Prior to CAP, proprietary interfaces were used to communicate with devices. For example, one siren system might require pressing the red button to sound the siren, while another system might require flipping the blue switch. This would be analogous to not knowing were the gas pedal was going to be when you got into a car that you had not driven before.

CAP fixes this by standardizing the structure so that any device that reads CAP messages will know what to do automatically. CAP is also designed so that one message can initiate a response in multiple compliant devices. This speeds delivery of alerts.

CAP is an international standard that has been ratified by OASIS. See http://www.oasis-open.org/committees/download.php/14759/emergency-CAPv1.1.pdf for the full specification.

PORTABLE RECEPTION

Everyone knows that ATSC broadcasts cannot be received in moving vehicles, but several recent advances are gaining on that objective. In Las Vegas, we are installing DTV data receives in police vehicles. However, we are using the system as a portable service, not a mobile one.

The difference is that our initial testing has shown that a vehicle can travel to just about anywhere in Metropolitan Las Vegas and lock on the KLVX signal as soon as they stop. This portability has provides great benefit over previous options for content distribution.

Advances like A-VSB being demonstrated at this years NAB, hold even more promise for true mobile reception. The tradeoff is a reduction in total data rate, which means that fewer television programs may be possible. This is a small price to pay for some stations if it means that they can deliver critical data to first responders in the field. In a true crisis, the focus may be on data delivery rather than multiple channels of television content.

Sixth generation receivers are also on the horizon. With each generation of receiver technology, we see tremendous improvements in the ability to receive ATSC signals in challenging conditions.

Another future possibility for mobile reception is digital radio. Many public television broadcasters also operate radio facilities. HD Radio, or iBiquity as it used to be known, is the standard for digital radio broadcasts. While there are no receivers currently on the market that can receive IP data over a digital radio broadcast and pass it on to a connected computer, these devices are coming.

Microsoft, for example, at this year's Consumer Electronics Association show announced a plan to use digital radio to deliver IP data to specialized devices. This same technology could be used to deliver low data rates (about 50kbps) to vehicles on the move. Those same vehicles could then begin receiving higher data rates and video sent over digital television signals once they stop.

SUMMARY

Digital television datacasting provides many capabilities that are difficult to replicate with other distribution systems. Multicast delivery, high data rates, digital wireless ubiquitous metropolitan coverage, and the ability to layer on existing elements like conditional access, routing and highly compressed video provides real benefits to end users. In the public safety space, these attributes are difficult to replicate.

While the technology works well today, future enhancements like mobile reception and higher compression just add to the advantages over other systems. However, the challenge is not in just having the technical capability to broadcast data. Customer benefit comes from integrating multiple systems into a functioning whole.

Providing a service that solves customer problems involves sourcing and securing the data that will be delivered, designing an ingest architecture for that data, installing and configuring the encapsulation equipment, installing receive antennas and software to make the information accessible. Training and technical support are also critical to a complete system.

When we start thinking in these terms and not just about sending the data out, broadcasters can begin to realize the benefits, both public service and financial, that digital television datacasting makes possible.

Mark O'Brien can be reached at 703-802-2975 or mobrien@spectrarep.com

MTV Networks' Disaster Recovery Facility – Case Study

LIONEL HIGHTOWER
MIKE SCHLESIER
SCOTT SILBERLUST
MTV Networks
Hauppauge, New York

INTRODUCTION

The paper will describe the decision making process and implementation that resulted in the recently activated MTV Networks Disaster Recovery Facility. We will discuss the many decision points involved in conceiving and building a DR facility and what made sense for us in the context of MTV Networks business requirements.

OVERVIEW

Viacom/MTV Networks

Viacom, the parent company of MTV Networks, is a leading global entertainment content company, with prominent and respected brands in focused demographics. The company employs approximately 10,000, with revenues from its cable properties of $9.61 Billion in 2005. MTV Networks reaches more than 480 million households in more than 175 territories and more than 25 languages via more than 130 locally programmed and operated TV channels, more than 155 Web sites and broadband services, and more than 25 mobile services

MTV Networks Network Operations Center

The MTV Networks Network Operations Center is located approximately 50 miles East of New York City in Hauppauge, Long Island, literally across of the street from our neighbors at HBO. The NOC employs approximately 300, in a facility of around 10,000 square feet. The NOC presently uplinks approximately 100 feeds, including both xx for MTV Networks and xx for Showtime Networks. Networks we originate include MTV, VH1, Nickelodeon, Spike, CMT, TV Land, and Comedy Central, for the US domestic market, as well as feeds to Latin America and Canada. The revenue generated from the facility is roughly $5Billion a year. The NOC is an all digital, automation driven, server playout based facility and has approximately 300TB of archived material that is stored on a combination of disk and tape storage devices.

The Business Justification for a DR Facility

With the number of networks originating from the NOC continued to grow at a rapid rate, it became increasingly obvious that the business impact to MTV Networks resulting from a damaged, destroyed or uninhabitable NOC was reaching an unacceptable level. The NOC, because of its proximity to New York City and exposure to hurricanes, was clearly at risk.

Although a rudimentary DR facility was put in place at a remote uplink site in the aftermath of 9/11, it was clearly not capable of providing much more than an on air presence with stale, incorrect programming and incorrect commercial inventory. The ability to provide long term resumption of the MTV Networks' feeds was simply not achievable, only a 48 hour loop that could not be modified or updated.

It was also obvious that the every growing archive of digital assets was at risk as well. Although material did exist on tape at other sites, the labor and time necessary to reingest 50,000 hours of content demanded that multiple instances of the content had to be stored and that it made sense to make this a requirement for an enhanced DR strategy.

Finally, the NOC was bursting at the seams with the high rate of network launches; in 2006 alone there were 14 network facilities added. Any DR strategy needed to incorporate a capability of providing "value-add" type services to the NOC, such as off air monitoring, media QC and Media ingest. Finally, launches of new networks at the DR facility could help to alleviate space limitations at the NOC.

OVERVIEW – DESIGN CRITERIA

Location

One of the main criteria of a DR facility is that it is located a sufficient distance from the main facility such that it is highly unlikely that weather or other impacts will not affect both sites. When MTV Networks started the search for the DR facility location, a minimum distance from the NOC of 500 miles was set as the minimum. After that distance was set, we hoped to find a site that was also accessible in a few hours by car or plane, allowing staff to be quickly mobilized if that was needed and possible. Also, a minimum of 200 miles from coastal areas to minimize hurricane impact was added to the site selection process.

Facility Model

In considering a DR site, a major fundamental decision that has wide ranging impact is how the DR site will operate, both under emergency and non emergency periods. In other words, will the site be "Hot", ready to go to air at a moment's notice, fully staffed 24/7 with lists running in parallel with the main NOC? Will it be "warm", with equipment on and tested, but with minimal staffing and no systems actually attempting to mirror the main NOC? Or will the site be "cold", totally shut down and not staffed until needed in an emergency?

In addition to the operational mode decision, the next decision point is whether the DR facility will be owned and staffed by internal company staff or will it be set up and operated by a third party service provider?

Finally, how many of the networks would be protected by the DR facility? What were the criteria for inclusion? Where did the cost/benefit equation make sense?

The decisions made by MTV Networks were as follows:

1. The site would be "hot", staffed as close to 24/7 as possible, given budget constraints and would run playlists in parallel with the NOC to allow us to resume origination on all protected networks within 2 hours.

2. The site would be in MTV Networks' space with owned origination equipment, with MTV Networks' staff. The driver was the strong conviction that we wanted to be sure that we were not subordinated in restore operations in a shared DR service provider environment. In additional, it set the stage for possible expansion and/or support activities to occur at the DR site to be a value add to the NOC.

3. The DR site would back up the top 7 networks originated by the NOC, as determined by revenue. These 7 networks and their West Coast version all reside on a single Digicipher Multiplex, making inclusion beyond these 7 a large incremental spend on encoding and transmission equipment. The West Coast feeds were desirable, but only if a cost effective solution could be identified to generate and distribute the feeds at the DR Site.

Technical Capabilities

The DR site technical design criteria started to take shape after the facility model questions were answered. The facility would feature master control facilities to originate 7 networks under automation control, with full graphics and time of air effects. If possible a cost effective method to original the 7 West coast feeds of these networks was desirable.

In addition, an archive system sized to hold a copy of all 300TB of NOC long form digital media would be included. Provisions would be made to include digital file based delivery of commercial inventory, so the site would be always current with spots, and could run independently from the NOC when necessary.

Media Ingest and Media QC functionality were necessary to allow independent operation and ongoing media preparation.

A Digicipher encoder for transmission and RF system to satellite return feeds were necessary to create the transmission multiplex and to monitor the satellite transponder activity.

It was desired to include some level of remote control and monitoring of the facility from the NOC in case of staffing issues at DR or for remote troubleshooting of DR systems by NOC personnel.

The DR facility, in order to accommodate possible future growth, should be designed to support up 32 channels of origination, but equipped and integrated up to the channel count the budget constraints would allow.

Essentially, the decision was made to build a adjunct Network Operations Center that was technically capable of assuming origination of the top 7 MTVN services for an indefinite period of time.

THE LOCATION SEARCH

The most difficult aspect of implementing the MTV Networks' DR facility was finding and securing the real estate necessary. Multiple sites were pursued and numerous real estate deals fell through at the last minute.

Things started looking more promising when we began to focus on data center space that was already technical in nature and had HVAC and power systems that were appropriate to broadcast facility requirements. In addition, an early requirement that the space be adjacent to uplink centers was replaced with a new approach that used high speed connectivity to link the DR site with a teleport anywhere in the country.

Finally, in mid 2006, a site was identified and the process of lease negotiation was begun. As the process began to wind down, the NOC Engineering group began to ramp up for the integration of the new facility.

Technical Architecture

The DR facility was designed to support up to 32 channels of origination The integration plan was to integrate 12 of these networks, but equipping and commissioning only 7 networks to support the 7 key NOC network services.

The Master Control Room, the heart of the DR facility, was designed to handle 32 channels when completely built out and equipped, with 7 channels MCR implemented for facility launch. The room would be built around a "Pod "concept, each pod would be staffed by an operator who would have responsibility for 4 channels of automation driven origination. The Master Control Room would also feature a supervisory console that would allow a complete overview of all channels and take care of any live events. The Maser Control Room would feature a KVM over IP system that would allow any network control surface to be moved to any pod in the MCR.

The Equipment Room would house all the electronic systems that support the MCR, as well basic support systems such as clock, time code and sync distribution. The facility routing switcher, a 512x512 matrix would also be installed in the equipment room to handle all SDI routing. A stream reorder system, which would record in a destructive loop the last 48 hours of on air feeds of the seven top networks would provided. In addition, a West Coast delay system would be installed. All RF systems, such as the satellite return IRDs and Motorola DigicipherII encoder would also be located in this room. Finally, the Disk Cache and Robotic tape archive that supports all media storage would be located in the portion of the facility.

An Ingest area, with room for 4 multiple ingest stations would be located in a separate area with VTR and Flexicarts to support tape to disk transfers

Three QC Rooms would be provided, individual small rooms that lend themselves to careful audio review, isolated from equipment noise and other QC activity.

For satellite return feeds, two 3.8 meter dishes would be installed on the roof of the building and fed to the rack room to the IRDs.

In order to meet to design criteria that the DR facility back up all NOC media a WAN was designed by the MTV Networks IS&T Department that would support the movement of media files across a 1000 mile hop. A calculation of the daily amount of media traffic was used to size the WAN and formulate a strategy to accomplish the initial synchronization of the NOC and DR Archives.

The Engineering Project

The NOC Engineering Team began formulating the plan necessary to bring the DR facility to reality in Mid Summer of 2006. The desired "in service" date for the new facility was set for a very aggressive end of 2006 completion.

This aggressive date moved the team in a direction that involved reusing technical system design that heavily relied on designs that had already been deployed in the NOC. The Engineering team had recently engineered and integrated the "SuperSuite" facility on a tight time line. The SuperSuite is a MultiChannel facility that was very versatile, but also very cost effective and very expandable. This approach had a number of strong advantages:

- The majority of the equipment list was already complete
- The majority of the cable run listing was done
- The majority of documentation was done and reflected as built conditions
- Configuration files for most systems were reusable with minimal editing

Challenges

However, some challenges existed, even with this head start strategy. The DR facility chosen had under floor HAVC air supply, so plenum rated cable was required that represented a large unexpected budget impact. This required some value engineering decisions to be made to keep the project within financial constraints.

In addition, the network launch schedule at the NOC was also very full during the scheduled DR integration schedule, resulting in major conflicts with Engineering staff schedules. Luckily, the SuperSuite, where most network launches were to be placed, had been "preintegrated" so that the strain on Staff was minimized.

Finally, the challenge of managing a large project at a far flung site required significant planning in scheduling travel and accommodations arrangements to make sure the project was properly supervised and managed.

The cloned SuperSuite Design was incorporated into an RFP that was released to system integration firms in June 2006

Project Integration

Integration began at the System Integrator's site almost immediately following the award of the RFP in June 2006. The technical system began to arrive on site at the DR facility in September 2006, immediately following the completion of the necessary HVAC, Power, and space modifications.

After the Integrator achieved substantial completion in November of 2006, the NOC Engineering Team and numerous key vendor field support staff began to arrive November 7 to begin testing, configuration, and commissioning. Because many of the configuration files had been saved a reused from the SuperSuite project, the team was able to move quickly, working hand in had with the integrator to verify the systems.

The NOC Operations staff arrived on site on December 7 to begin User Acceptance Testing. Armed with on air playlists, the master control system was put through its paces and discrepancies reported to the Engineering team. Graphics were loaded and verified, and the ability to mimic the on air look of the 7 networks was checked. By December 18, the majority of the punch list items were resolved, and the DR Operations Staff began dry runs following the Christmas Holidays.

The faculty went on line on January 2, 2007, as planned.

Future Plans

With the DR facility now on line, our planning has begun for the next phases of the facility.

There has been some discussion regarding extending the DR site to encompass not only the Domestic US, but the European and Asian operations as well. The issues to accomplish such a feat are significant, but the value to the company's worldwide operation is very high.

The DR site, because of its designed–in expandability and access to media archives, can certainly be the home to new linear networks or to the new emerging non linear and non traditional delivery media initiatives.

Lessons Learned

WAN Data Transfer across a large geographic distance is not as easy one might believe. Latency issues, negotiations with carriers in unfamiliar areas of the country, and the multitude of vendor hardware and software involved in a high speed WAN are all issues that make moving 90-100Mb/sec over 1000 miles a challenge.

The emotional and physical toll on the entire technical and operations staff to accomplish a task of this magnitude while attempting to support on going operations and network growth at the main facility was high. This project spanned a number of normal vacation periods and holidays, making the demand even higher on staff.

The strategy of reusing both the design and integrator form the SuperSuite project was pivotal in the facility staying on schedule and going on line on the promised date. We effectively cut 6-10 weeks of time off the timeline by leveraged all the work that was cloned.

Summary

After 3 years of effort, MTV Networks now can continue generating revenue on our top networks, regardless of the viability of our main Network Operations Center. In addition, we have backed up all our digital media, set the stage for additional network growth, and provided for additional ongoing value add support of our main NOC.

In a little over 6 months, the NOC Engineering team assembled an entire secondary NOC that can grow substantially over time.

Although we all hope that our DR facility will never be pressed in to long term service, we are confident that it can fulfill its mission to keep our audience entertained and informed, and fulfill our business obligations to our advertisers in a disaster scenario.

Acknowledgements

This project was a total team effort, involving all groups at the NOC.

Scott Silberlust, VP of Finance and Facilities, and Mike Miglino, VP of Business Development persevered and found the real estate.

The Engineering Team: Tom Fumante, Mike McMackin, Frank Burgert, Brendan Peterson put huge amounts of time and effort into making the DR facility a reality.

The Operations Team, led by Mike Schlesier, put together the staffing and facility testing plans.

Finally, thanks to Paul Sartain, NOC SVP/GM for the vision and setting this great opportunity in front of the NOC team.

Flywheel Technology as Superior Solution for Reliable, Continuous Backup Power

Gary Rackow
Active Power
Austin, Texas

ABSTRACT

Continuous power is vital to the broadcast industry. People rely on radio and television in emergency situations, and being off the air as the result of a powerful weather or electrical event can result in not only lost viewers and listeners, but also lost advertising revenue. A sudden loss of power can also prove costly in terms of the damage that sensitive broadcast equipment can suffer due to an unexpected hard shutdown.

As broadcast engineers continually seek better and more efficient solutions to address this situation, it turns out that a new twist on an established technology is fast becoming a serious option when it comes to choosing the best Uninterruptible Power Supply (UPS) solution.

This presentation will illustrate that while the technology behind a new wave of flywheel-based UPS products involves relatively simple mechanics, the final products themselves have proven to be highly reliable, extremely efficient, and a more environmentally sound alternative to traditional lead-acid battery-based systems. It will also detail how broadcasters are employing products based on this technology.

THE POWER GAP

On October 17, 1989 the Oakland A's were squaring off against the San Francisco Giants in game three of the World Series. Oakland won the first two games and the Giants were anxious to wrest some momentum back from their bay area rival.

As game time approached the weather was beautiful. It was a typical Northern California afternoon, the diamond was meticulously groomed for the most important game of the baseball season, and game time excitement was building. As the players sat in their locker rooms contemplating the game or lazily pitched the ball on the Candlestick field, the broadcast announcers ran through statistics and the various story angles that infuse baseball with its sense of drama.

Suddenly, people began to sense that something was wrong. The earth began to vibrate and the fans assembling in the stadium became uneasy at what was happening around them. It became clear within moments that nature was intervening in America's past time and that this was a significant event.

Al Michaels, the ABC Sports play-by-play commentator, addressed his audience with the following: "I'll tell you what; we're having an earth…" Al's commentary was cut short due to a power failure. ABC quickly had the power up and running and began broadcasting news coverage of the unfolding seismic events.

Al's words were lost in a "power gap". A "power gap" is the space between full power and backup power caused by the temporary interruption of power flow. The interruption can last for milliseconds, seconds, or even a matter of minutes, but it can be disproportionably costly to broadcasters. With the sensitive electronic equipment now in use in almost every facet of broadcasting, even the slightest interruption of power, or the provision of poor quality power, can cause capital loss and mounting difficulties for engineers. Historically, batteries have been the primary energy storage medium for UPS systems but increasingly the broadcast industry is turning to UPS systems using flywheel technology. This change is occurring because flywheel technology has proven to be more reliable, more cost-effective, and a more environmentally sound alternative to traditional lead-acid battery-based systems.

UPS SYSTEMS MEETING POWER GRID CHALLENGES

According to a report created by the North American Electric Reliability Council, over the next 10 years the demand for electricity is expected to increase 19 percent while power generation will only grow by 6 percent. This alone points to the potential forces working against broadcasters who face considerable challenges and cost in maintaining continuous power for their operations. Power grid infrastructure shortcomings coupled with erratic weather conditions, power surges, and brown outs are all serious sources of concern to those charged with beaming the messages and images that make up the fabric of American culture.

Fortunately, UPS systems are easily justified in studio broadcast applications that cannot tolerate any interruptions, including maintenance interruptions that are scheduled in advance. In these situations the cost of the UPS system is usually a small part of the total infrastructure cost and the potential loss in revenue due to the interruption in program transmission.

Advantages of Flywheel-based UPS Systems

Flywheel-based UPS systems are financially attractive to a wide range of broadcasting studios and transmitters sites. They can: 1) reduce maintenance costs, particularly those related to batteries; 2) reduce additional utility costs that result from UPS system efficiency losses; 3) handle crowbar events and keep the transmitter protected; and 4) improve overall UPS system reliability.

Batteries

Today, valve-regulated lead-acid (VRLA) batteries are commonly used at all UPS power levels. They are the predominant UPS battery used in studio operations because they cost much less and require much less space than conventional vented lead-acid batteries (wet cells). Batteries, particularly VRLA batteries, are recognized as being the highest failure component of UPS systems.

For a typical 240-cell battery system, one would expect a few batteries to fail in the first couple of years, and then about 50 or more to fail in each of the next two years. Since VRLA batteries predominantly fail in open-circuits, any single cell failure results in failure of the entire battery system. Therefore, the mean time between failures of the entire battery system will be measured in months or weeks rather than years. Additionally, batteries leak, are toxic, and must be disposed of in a costly and environmentally sound way. Thus, eliminating batteries greatly improves overall reliability, and at the same time, eliminates the continuing costs associated with battery maintenance and replacement.

Some flywheel UPS systems use a line-interactive design that makes them much more efficient than conventional double-conversion, battery-based UPS systems. As a result, utility costs resulting from UPS system efficiency losses are lower. UPS systems designed with integrated flywheel energy storage can achieve efficiencies of 98 percent compared to 93 percent or 94 percent for conventional battery-based UPS system. In the 1000 kVA size range, such as at large transmitter sites, the difference can easily amount to an annual energy savings of $18,000 to $20,000. In addition to the electrical characteristics, the smaller space requirements and wider operating temperature range of flywheel UPS systems make them easier and less costly to retrofit and install.

Operating Costs

UPS system operating costs can vary significantly, depending on the energy storage chosen. Annual operating costs for a 900-kVA flywheel UPS system, in current dollars, will amount to about $29,000, assuming electricity cost of seven cents per kilowatt-hour. Battery-based UPS system operating costs vary between

$51,000, in years when battery replacements are minimal, to about $105,000 in years when full battery system replacements are required. In most cases, conventional UPS systems have not been able to meet the technical or financial criteria required to justify capital expenditures in broadcast transmitting environments. There are several reasons for this. First the total cost of the UPS system installation, including facility modifications to create a suitable environment for the batteries, has been too high, even when the cost of UPS equipment alone is well within required capital cost limits. Second, UPS system maintenance costs, including those resulting from the ongoing costs of scheduled and unscheduled battery replacements, significantly increase annual operating expenditures. By comparison, most flywheels have a design life of 20 years or more, so replacement costs are non-relevant. Utility costs increase due to electrical loss of the UPS equipment. Finally, the possibility of transmission interruptions due to crowbar events offsets any upfront financial advantages held by traditional UPS systems.

What is a crowbar event?

A crowbar event is the automatic shutdown method used in high-power transmitters as a safety circuit. The shutdown protects the transmitter amplifier tube or the inductive output tube (IOT) in the event of an arc-over inside the IOT. The function of the crowbar is to remove the high voltage from the amplifier as quickly as possible, typically within a few microseconds of the detected problem. The crowbar circuit shorts out the high-voltage DC power supply to the IOT for a brief period of time, often in the range of several milliseconds. This function is typically performed with a device called a thyratron, which is a gas-filled tube that is similar in construction to a vacuum tube. The thyratron is connected directly across the high-voltage DC supply. It is a very fast, high-voltage switch, comparable to a silicon-controlled rectifier but with much higher voltage ratings and speed. When a problem is detected in the amplifier the crowbar acts quickly, otherwise the IOT would be destroyed, resulting in significant capital expense. The crowbar action produces a current of several thousand amps on the AC input of the high-voltage supply and, therefore, on the output of the UPS system supplying the transmitter. The event is equivalent to a short-circuit applied directly to the output of the UPS system, which can draw up to 20 times rated current. Assuming the input power supply to the transmitter can supply the large current that is demanded, the action of the crowbar does no harm to the transmitting equipment and after a few seconds the high-voltage supply returns to normal. In the event of an overload from a crowbar event, an integrated flywheel-based UPS system switches to bypass in order to help supply the desired current from the lowest impedance source—and it does so without disturbing the operation of other transmitters on the same circuit.

APPLICATIONS: TRANSMITTER PROTECTION

The architecture of an integrated, line-interactive flywheel-based UPS system is designed to handle overloads and large step loads that could disrupt studio transmissions or crowbar events that are inherent to transmitter sites. Given the severe nature of the overload from a crowbar event, a flywheel-based UPS system will seamlessly switch to bypass and supply the desired current from the lowest impedance source. Other typical power events are easily taken care of with the flywheel UPS design. Two examples of these types of installations are given below.

Telemundo

Telemundo' KVEA, a Los Angeles, California-based TV station, knew that the slightest power disturbance could cause interruptions in studio transmission of programming and possible loss of views in a very competitive Spanish-speaking market.

As the fastest growing network in the country, Telemundo Communications Group, Inc. wanted an expandable power quality system that could provide maximum protection for the station's sensitive digital equipment while withstanding the existing environmental conditions in the switchgear control room. Legacy battery-based systems could not tolerate the environmental requirements due to the need for constant battery cooling. The station satisfied its requirements with an integrated flywheel UPS system along with a generator set inside an outdoor enclosure that included a base tank. KVEA is now assured critical load protection from all IEEE 587 power anomalies—from transients to long-term outages—without disruption to the studio's production.

After their first flywheel UPS system successfully protected the critical load from numerous utility disturbances, Telemundo Communication Group chose to place additional flywheel UPS systems across the country. Active Power's compact system fits easily into small studios, computer and switchgear rooms or outside containers. As a result, when Telemundo needed uninterruptible power to protect multiple transmitters at the top of the as Empire State Building, two Active Power flywheel UPS systems were the obvious choice. The efficiency, space savings, and environmental aspects of the flywheel UPS system made it the perfect solution.

Christian Television Network Makes a Choice

When evaluating a UPS system for broadcast applications, programming loss is only one consideration. The costs of replacing transmitter tubes, operational efficiency, as well as maintenance are important factors in operational efficiency and budgets. When Christian Television Network's (CTN) Chief Engineer needed uninterruptible power supply systems for CTN's transmitters, he recalled working with another station that had a large, battery-based UPS system for its transmitters. His recollection was that batteries caused problems so he steered his acquisition efforts toward flywheels. All of CTN's transmitters are high-current devices with voltages in excess of 20,000 volts. This was key in driving CTNs capital expenditures for its UPS system.

CTN's transmitters put out 60,000 watts of power, 24 hours a day. A power interruption shuts off the transmitter and stops cooling water flow. This is a non-standard shut down that dramatically shortens tube life. And since the various tubes that CTN uses to power its equipment cost anywhere from under $30,000 to more than $40,000, the initial cost of a mid-power flywheel UPS system can potentially pay for the costs incurred in just one outage in addition to the constant savings due to low maintenance and efficient operations.

Christian Television Network, which has a potential audience of more than 15 million viewers, has several Active Power flywheel UPS systems installed at different transmitter sites throughout the Southeastern United States. At a site near Tampa, crowbar testing was performed on the output of one of its flywheel UPS systems. This transmission site has a primary 60 kW analog transmitter with a backup 30 kW transmitter that shares the output of the UPS system with a new 65 kW digital transmitter equipped with a crowbar circuit. The UPS system was tested several times in order to view the response of the system during a digital transmitter crowbar event. The current delivered by the UPS system during the crowbar test peaked at 3000 to 4000 amps, varying due to the position of the voltage sine wave when the crowbar fired. In one of the crowbar tests, the UPS system only discharged for a short period and did not have to go to bypass. In five of the six crowbar tests, the system went to bypass via the static bypass switch and displayed a warning message due to the severe step load. The UPS system then returned to normal on-line operation within a few seconds. The UPS system is configured so that if this same warning message occurs more than once per hour, the system will remain in bypass mode and will require an operator to reset it. Additionally, an external contact signal from a delay relay may be used to automatically return the UPS system to on-line operation if multiple crowbar events are expected within one hour. The other transmitter at

CTN's site, which was connected to the output of the UPS system, stayed on the air without a glitch during all of the crowbar tests. This is partly because the UPS static bypass switch allows transfer to the bypass source without affecting downstream equipment; otherwise the critical loads could be subject to power disturbance during the event. The crowbar test at CTN's transmitter site indicates that flywheel-based UPS systems perform exceptionally well under extreme conditions and deliver stable, uninterrupted power to sensitive loads even during severe transient conditions.

The Future of Flywheel Technology

These two examples illustrate some of the common power supply experiences of broadcast stations everywhere and demonstrate the compelling benefits of flywheel-based UPS systems. There is little question that the provision of high-quality, uninterrupted power for broadcast facilities will be a continued source of concern for broadcast engineers for the foreseeable future. In the absence of alternatives, batteries have been the prevailing energy storage for decades, but UPS systems based on flywheel technologies have been proven, are more widely adopted today, and offer many advantages. UPS systems based on flywheel technology are inherently more reliable than battery-based systems, particularly in broadcast applications, where they handle crowbar events effectively. They also cost less over the life of the system, due principally to the lower installation and maintenance costs and reduced utility costs that result from UPS system efficiency losses.

And they are an environmentally friendly solution, which makes them easier to manage and ideal for transmission sites. In short, flywheel UPS systems help broadcasters stay "On Air", and be more competitive in a world of increasing power challenges and proliferating sources for news and entertainment.

Surround Sound For Radio

Wednesday, April 18, 2007
9:00 AM – 11:30 AM

Chairperson: Steve Fluker
Cox Radio/Orlando, Orlando, FL

Implementing Surround Sound in a Broadcast Facility
Steve Fluker, Cox Radio / Orlando, Orlando, FL

5.1 Radio – Too Much, Too Soon?
Jon McClintock, APT, Belfast, Northern Ireland

***Surround Sound Radio Broadcasting is Reality**
Dan Danko, Cincinnati Public Radio, Cincinnati, OH

***Radio Surround Gets Ready for Primetime: New Techniques for Surround Mixing**
Alex Kosiorek, Cleveland Institute of Music, Cleveland, OH

Internet Radio in Surround Sound
Nikolaus Färber, Fraunhofer IIS, Erlangen, Germany

*Paper not available at the time of publication

IMPLEMENTING SURROUND SOUND IN A BROADCAST FACILITY

Steve Fluker
Director of Engineering
Cox Radio, Orlando

ABSTRACT

As radio operators continue to try to attract new listeners through new technologies such as HD Radio and multicasting, the idea of broadcasting in Surround Sound has made its way into the spotlight. We have been enjoying the surround experience for many years in movie theaters and more recently through our home theaters and HDTV sets, but it has yet to really make its presence felt on the radio. This could be about to change though as the public awareness for surround sound has grown, and as broadcasters look for ways to keep listeners from seeking out alternative entertainment sources such as mp3 players and satellite radio.

Implementing surround in today's studio facilities can seem confusing up front, but really doesn't need to be. This paper will focus on getting surround sound on the air from the simplest and most economical way, to the most complex systems available and everything in between. Discussions of the various levels of complexity will help guide the broadcaster to choose what is right for their specific situation.

BASICS OF SURROUND SOUND

Surround Sound is the term that has been given to today's multi-channel audio systems. The most common configurations of surround are 5.1 and 7.1. Figure 1 shows the layout of a typical 5.1 surround system, which for now will be the focus for FM and HD Radio broadcasts. This configuration consists of three speakers in front of the listener, the Left (L), Right (R), and Center (C), and two surround speakers located slightly behind the listener. These are the Left Surround (Ls), and Right Surround (Rs). There is also a sixth channel, which is a very narrowband channel providing low frequency audio to a subwoofer. The bandwidth of this sixth channel is about 1/10th of that of the other five channels, thus is represented by the ".1" in the 5.1 name. A 7.1 surround sound system simply adds two additional speakers to even further the surround sound experience.

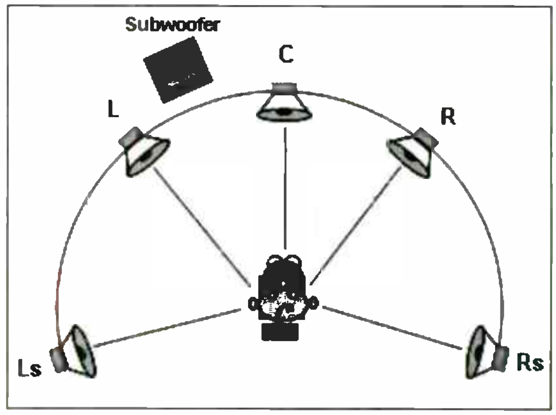


Figure 1 – Typical 5.1 Surround Sound Speaker Configuration

BROADCASTING SURROUND ON RADIO

Pure 5.1 surround sound source material consists of five independently recorded audio channels, each with their own separate amplifiers for reproduction, just as a stereo signal is recorded and played back on two separate channels. This is referred to as discrete surround, and while this would provide the most accurate reproduction of the original material, it is not practical to broadcast over an analog FM radio station due to a limited bandwidth and only two audio channels to work with. To resolve this issue, several manufacturers have developed methods to encode the five channel surround audio into a two channel downmixed version which can easily be transmitted over a standard analog FM signal. When this two channel broadcast is received, a surround decoder will be able to split the audio back out to recreate the original five channel surround audio experience with relatively good accuracy. This method of encoding the five individual audio channels into a two channel downmixed version is known as a composite type of surround sound system. It's important to note that this downmixed signal is compatible with existing stereo and mono receivers.

HD Radio has now opened the door for a new type of surround sound system called spatial or component surround. Component surround takes advantage of the available data streams on the digital carrier. With this method, the original stereo audio is transmitted along with a digital data spatial bitstream which carries the multichannel information separately to the receiver. A benefit of this method is that it can very accurately reproduce the original five channel discrete source material. The down side to this approach is that it will only work on the digital carrier side and will not provide surround to the analog signal, unless a second composite signal is used.

IMPLEMENTING SURROUND IN A STUDIO

With all of the above information we now come to the key question. How do we get surround sound on the air? As with most new technologies, there are quick and relatively inexpensive ways as well as more complex and elaborate ways to incorporate surround sound into an existing broadcast facility. The first step is to decide how big of an investment the station is willing to commit to the project. If your radio station just wants to be able to promote the fact that they are in surround yet keep the cost to a minimum, a "music only" method might be the answer. This won't be the case however for the station serious about creating a more impressive surround experience.

MUSIC ONLY STUDIO

Let's start with the simplest solution, the music only scenario. Figure 2 shows the basic needs for your studio. This is the easiest solution because it requires the addition of only a few pieces of equipment and little or no changes to the overall infrastructure of your studio facilities. For this method you first need to add a CD, DVD player, or some other source capable of playing the original discrete 5 channel surround material. Next, the output of this source is connected to the input of a Composite Surround Sound Encoder from any of the manufacturers on the market. Currently the primary surround systems available for broadcasting are the Dolby ProLogic II, Neural Surround, and SRS Circle Surround. The surround encoder will take the five channels of audio and create a two channel downmix which will then be compatible with your existing production console, audio storage system, control room, and the entire air chain of the radio station.

While you can get away with just this setup, it's advisable to also upgrade your studio monitoring It's important to have the ability to hear the final surround product after passing through the encoder and decoder so that you know exactly what your audience will hear. Without it you are producing blindly. We will further discuss monitoring in a later section of this paper.

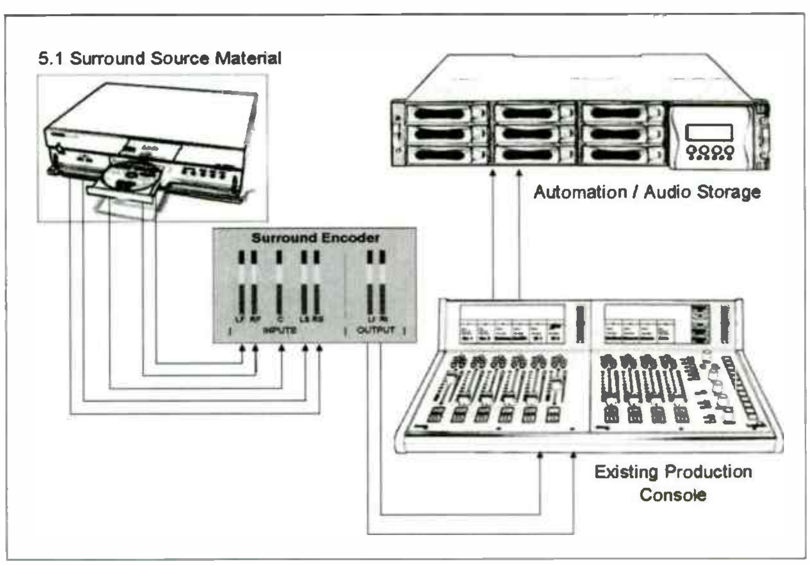


Figure 2 – Music Only Surround Studio Configuration

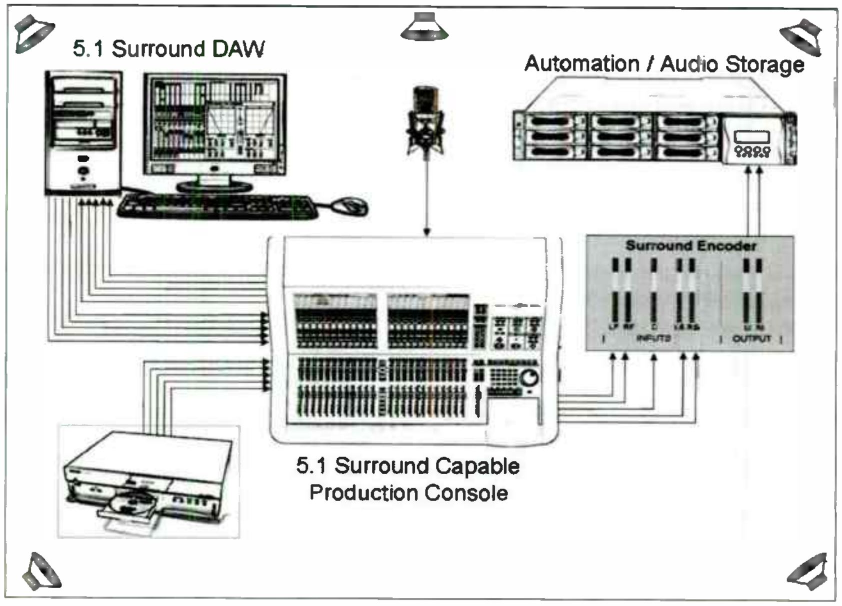


Figure 2 – Full Surround Sound Production Room

DON'T MISS THE "WOW" FACTOR

The music only configuration will allow you to get on the air in surround sound, but with it you just might be missing out on what could be one of the biggest hidden benefits of broadcasting in surround. That would be wowing your audience with your local production elements. Consider this; you're driving in your car, listening to a station in surround. The song sounds great, but at the end of it you hear a station sweeper play. Granted, it's very well produced, but not in surround. What happens? The audio collapses and the station's energy just fell flat. What about commercials? Creative spot production in surround may be the ticket to keep your audience listening through your breaks, rather than pushing the button to your competition. Remember when you upgraded to your new HDTV? You probably were so excited about your new picture quality that you would sit and watch a channel in HD over another channel, even if the program wasn't quite as good. Surround on radio will be new to your listeners, and they will be excited to hear just how good it can sound. Give them what they want!
Wow them with your production elements and they'll look forward to hearing your commercials. This is the next level of implementing surround in your facility; creating a studio where you can do full production recording and editing in surround sound.

FULL LOCAL SURROUND PRODUCTION

To accomplish the task of producing local elements in surround sound you will need to upgrade and dedicate at least one production room for the job. Figure 2 above illustrates the basic infrastructure of the surround production room. As with the music only setup, you will need a primary source audio player with the five discrete audio channel outputs. You will also need to have a digital audio workstation / editor capable of recording and editing in full surround. In some cases you may be able to simply upgrade the audio cards in your existing computer, but you may also discover the need to upgrade your editing software as well. This is an area you shouldn't compromise in, as most of your editing and production will occur in the editor.

Next you will need to have a production board capable of mixing five independent audio buses. Many of the newer audio consoles and router surfaces already have this feature available which may just need to be activated or programmed in. Even some older consoles may have Program, Audition, and Utility audio busses which can be reconfigured for your five channel output. This is a good temporary fix, but you are losing some flexibility in the control board. As you can see from the figure, the digital audio workstation must have five input and output channels connected to the production console. With this setup, the surround encoder will be placed at the final output of the mixing board so that everything you feed through it will be downmixed into the encoded two channel stream. From this point forward your facility will remain unchanged.

Please don't overlook training of the station production staff. Many production managers in radio have been around for a long time so producing content in surround sound will be something new for them. Don't assume that they will automatically know just what to do. While you can wow your listeners with well produced surround material, you can also chase them away just as easily with production work that sounds bad.

MONITORING CONSIDERATIONS

As briefly mentioned earlier, monitoring is a very important aspect of your surround sound facilities. You need to know what your programming will sound like before you actually put it on the air.

Figure 3 Surround Sound Production Room with Surround Monitor Spearkers

The photo above in figure 3 shows a typical layout and monitoring system for a surround sound production room. It would also be advisable to outfit the on air control room with surround monitoring as well. First of all, it's great to let you announcers enjoy the experience, but it also gives you another set of ears listening to your final product for possible problems.

The monitoring system does not have to be elaborate. It can simply be an off the shelf home theater receiver with built in surround capabilities. Your production room monitor should also be able to decode all of the primary surround systems on the market today. This is important for cross compatibility checks. Currently there are three major manufacturers of composite surround sound encoders. While they are all similar in nature, they do have their differences. Figure 4 illustrates what can happen.

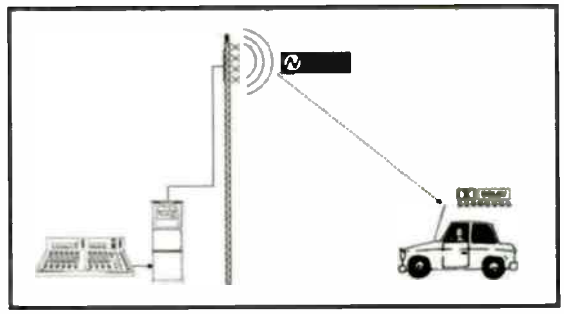


Figure 4 – Cross Compatibility Example

In this example, a radio station is broadcasting in surround using the Neural Surround system. Your listener, driving around in his car is listening, but has his radio set to Dolby ProLogic II. What is he going to hear? For this reason it's a good idea to have the ability in a production room to listen to all of the major surround systems. When you produce a surround element, listen to the final playback over each of the surround decoders to be certain it

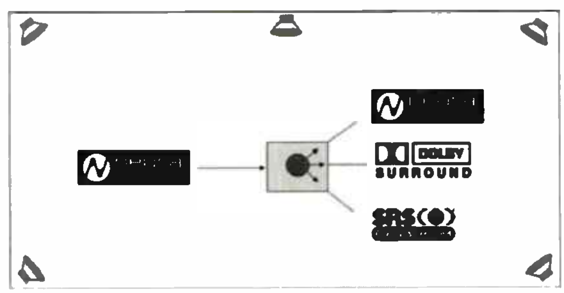


Figure 5 – Cross Compatibility Monitor Switch

will sound good. It's better for you to catch and correct problems in the studio before they make it out over the air and cause a listener complaint. Tests have shown that in most cases the listener will have a pleasant listening experience, even when listening through the wrong decoder. In the HD Radio platform, the radio station will be able to transmit a data flag to the radio directing it to decode in the proper format.

COMPONENT SURROUND SOUND

Up until now the discussion has been focusing on the composite surround sound systems. A more recent development in surround sound technology

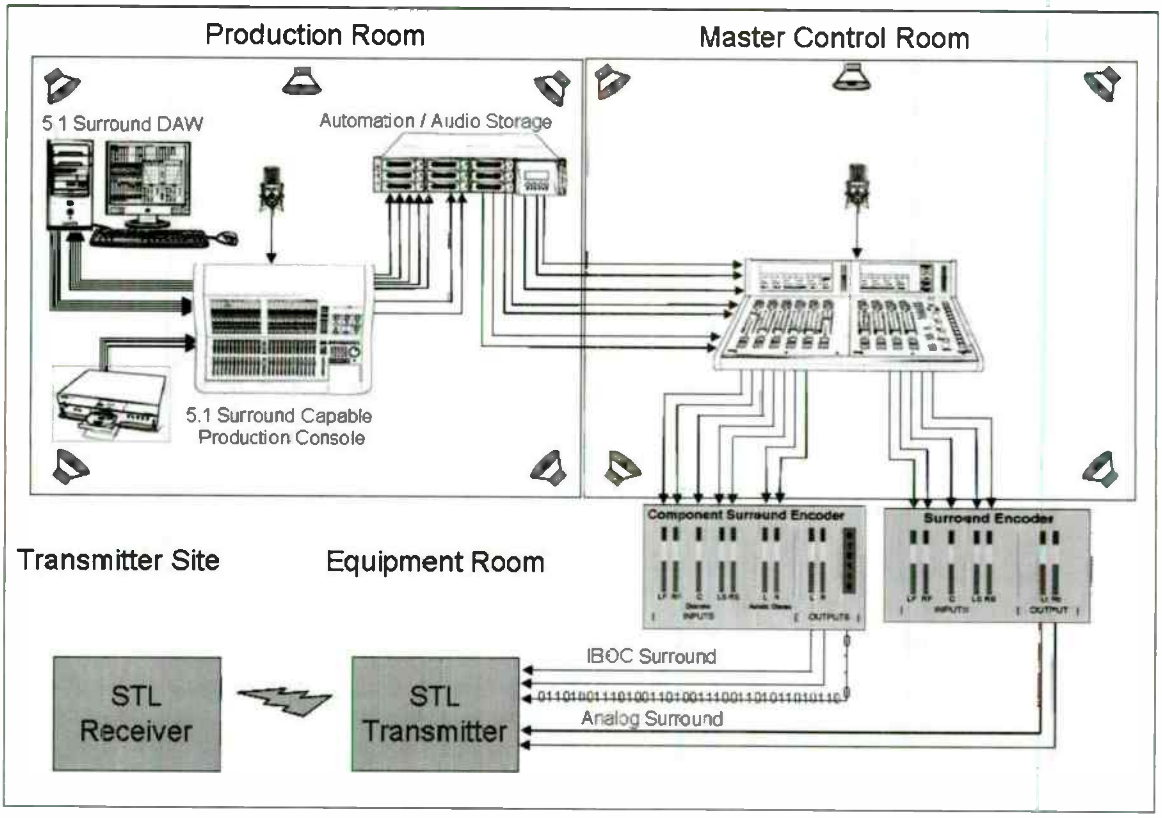


Figure 6 – Studio Infrastructure for a Component Surround Sound Air Chain

has taken advantage of the data stream in the HD Radio platform to take surround to the next level. This method is referred to as Spatial Surround and will also be called component surround in this paper. Figure 6 above is a block diagram for a broadcast facility using a component surround system. With a spatial or component surround system, the surround encoder is placed at the end of the air chain where it takes the discrete five channel surround audio, and a stereo mix and outputs the normal artistic stereo programming, plus a spatial data channel which provides the multichannel data to the receiver. This method of surround encoding may produce a more accurate reproduction of the original discrete recording than the composite counterparts, but requires considerably more changes to the infrastructure of the studios. The station automation / audio storage method will need to be able to record and store all five channels, plus possibly the stereo mix as well. It will most likely be necessary to increase the hard drive capacity to accommodate the additional channels. Routing switchers and interconnection wiring must also have the extra capacity for the increased channels. Now, not only does the production room console need to have surround capability, but so does the console in the master control room. The final output of the master control room will then feed the surround sound encoder and output the artistic stereo feed, plus the data channel. Be sure you have the ability to carry the data channel from the surround encoder to the HD Radio Exciter either through your STL, or directly if you have the HD exporter at your studio.

Since the component method requires the data stream to get the spatial information, it becomes limited to HD Radio only. This creates another issue when a listener is driving to the fringe areas of a radio station. As the digital signal fades, the radio will blend back and forth between the analog and digital carriers. Currently the blend is acceptable with only a loss in audio quality. If component surround is incorporated, the audio would suffer a separation collapse during this blending and may become more objectionable. To compensate for this a radio station may elect to add a composite surround encoder at the end of the air chain to ultimately feed the analog carrier reducing the shock of the blend.

SUMMARY

Surround Sound can become one of the "killer applications" for HD Radio. Most people who have heard demonstrations of surround in a car have walked away impressed and wanting it today. It's now in the hands of the radio station owners to make the commitment to incorporate surround on the air. Unfortunately many of the major broadcasting companies have spent millions of dollars already to consolidate multiple radio stations under one roof and may not be willing to spend even more money this soon to upgrade the facilities again. For those stations, the less complicated installations may make surround feasible. Stations in the preparation stages of a studio rebuild should definitely design for a total 5.1 surround infrastructure, even if they plan on using a composite system where the full five channel backbone would not be necessary. Remember that your facility will be in place for a long time and things change. Be ready with the latest in technology as it stands today. Don't get caught outdated from the starting gate.

Don't overlook your HD Radio Multicasting either. As the surround technology is developing, manufacturers are finding ways to make their systems work well at reduced data bit rates. This could allow radio stations to broadcast not only their main channels in surround, but their HD2 channels as well.

5.1 RADIO-TOO MUCH, TOO SOON?

Jon McClintock
APT
Belfast, Northern Ireland

ABSTRACT

Broadcast technology is advancing at great speed. Are we developing so fast that the listening public can't keep up? Or are we working towards the killer application that will enable radio to compete more effectively with the vast array of mediums vying for the consumer's attention?

Surround Sound for Radio has been touted as this next killer application, a key differentiator for those whose radio experience is about music quality and audio integrity. Both in the US and in Europe various projects have been undertaken -some on a pilot basis and others carrying regular programming bringing 5.1 for radio live into the home. This paper will look at the technical challenges faced by those who have established networks to support audio contribution and distribution for surround sound broadcasts, the approaches they adopted and the success of their ventures.

The paper will incorporate case studies where APT has worked with the broadcasters on new and innovative approaches to Surround Sound Radio Broadcasting.

COMPETITION

In recent years, 5.1 Surround Sound has overtaken simple two-channel stereo as the accepted mode for audio in the Film, TV and Music industries. This increase in audio channels has been driven largely by consumer demand with most households now owning a DVD player complete with a 5.1 speaker configuration. Add to this HDTV audiences, cinema-goers and young "gamers" who grew up with the Surround Sound experience and it quickly becomes clear that that there is now a level of consumer expectation that has to be met by content providers.

WHY RADIO NEEDS SURROUND

Now that the film industry has refined the concept, it is a simple fact that the Radio Broadcast industry needs to provide 5.1 Multi Channel Audio content. Radio broadcasters are in a highly pressurised market competing against information and entertainment mediums such as TV, Internet, DVD, MP3 players and iPods. An old two-channel programme is simply not going to meet the demands of their target audience. It may be argued that a stereo signal is abnormal. Although we have two ears, the brain has the ability to interpolate many signals simultaneously, which then creates a multidimensional image. In comparison, a two-channel signal will be both flat and unnatural.

Digital Radio has made the incremental step of improving quality by increasing audio bandwidth ("crystal clear CD quality") and adding ancillary services (scrolling text), but this is primarily still focused on delivering two-channel content. To make a significant step, the next generation of digital radio services needs to offer the Surround Sound experience.

Several 5.1-for-Radio pilot projects have been completed and the war of standards is raging for a suitable transmission protocol. Although several parties are involved, the primary participants are Dolby, DTS and SRS who between them have many millions of decoders in the market place. However, the actual methodology of the transmission is an interesting topic in that it offers up the options of Discrete, Matrix and now a third alternative known as Watermarking.

BUT HOW MANY CAN LISTEN?

Some simple pieces of information can be examined to validate the potential number of listeners. At this time we are evaluating homes that have Set Top Boxes and Home Theatres.

1. There are 300M people in the US.
2. In 2005 there were 108Million households (according to the US Census Bureau). Therefore, 100 Million homes make up 100% Market Share.
3. Up to 40% of all homes in the US have some form of Home Theatre.
4. Therefore up to 40Million Homes in the US could have Set-Top Box's with S/PDIF (Sony/Philips Digital Interface)
5. A report on CNN Money in 2003, Stated US consumer spending accounted for 20% of world GDP. We could multiply US Market by 5 for world markets.
6. Therefore for a World Market: 200Million Homes with the theoretical potential of receiving 5.1 for Radio

However, figures generated by ORF in Austria and BR in Germany would indicate that the listening audience is less than 1%. In Austria, the listening figures for the Cultural Program is 10% of the total potential. The number that can actual decode Surround Sound is 10% of that figure. Assuming a total potential of 10 million, then the actual figure is 100,000. This is in a country where the national broadcaster has been providing content since 2004, which at this time equates to 10 hours per week. With BR this figure is less with 3 hours per week content being generated.

The above shows a huge disparity between the potential and the actual.

EXAMINING THE TRANSMISSION OPTIONS

Discrete is self-explanatory in that a broadcaster transmits five separate channels plus the LFE – an approach that is well received amongst audio purists.

The Matrix solution has two principal parties - Fraunhofer and Coding Technologies -pushing this agenda. This approach advocates the use of a parametric steering channel as a side channel to an existing stereo platform and uses a perceptual coding technique to achieve the bit rate reduction. Multichannel decoders can reconstruct a surround mix from the Left/Right and Steering channel signal. This process assumes that content should be kept in discrete format up to the point of transmission.

The third alternative is watermarking, which generates a parametric steering channel upstream during the broadcast production phase. This signal encoded as a watermark and is perceptually hidden in an uncompressed digital stereo audio signal. Neural Audio and Harris Broadcast are offering this solution to broadcasters, which should ensure that the watermark remains imperceptible after, broadcast processing, and can be correctly decoded by a Surround Sound system.

The manufacturers of digital receivers will probably take the well-trodden route of spinning-in a variety of solutions to decode Surround Sound. As such, in principle, listeners will be able to get 5.1 content into their homes (through Set Top Boxes). Hopefully, the various alternatives for transmission will be largely irrelevant to the listening audience as the Set Top Boxes will be sophisticated enough to decode content regardless of the transmission algorithms/protocols. However the home listener is only a small minority of the target audience. The segment that really gets the juices flowing for broadcasters and technology providers is the automotive industry -those people in their cars listening to the radio on their way to and from work and generally going about their daily business.

Now comes the bigger challenge: how does the broadcaster move live content from remote locations (concert halls and sporting venues), through to their studios and then out to their transmitter sites? This is one of the main issues for Surround Sound, especially in light of the fact that failing to provide surround sound content was the primary reason why Quadraphonic failed in the 1970's.

COMPRESSION, CODECS AND ALGORITHMS

One solution is to process all five audio channels and the LFE in linear PCM. For the record, a 24 bit, 48kHz sampled program requires a data rate of almost 6 Mbit/s for 5.1 channels. For most broadcasters, such an option is completely cost prohibitive and will kill any contribution and distribution projects at birth.

This brings us to our old friend digital audio data compression and the balancing act between low bit rate algorithms using perceptual coders and slightly higher bit rates that use ADPCM principles. Perceptual coders will certainly reduce the network costs but add substantial latency and run the risk of destroying the phase relationship between the individual channels. Given the loss of stereo separation caused by perceptual coders in stereo signals, this destruction is almost a certainty. It is also worth noting that the final transmission algorithm, whether DTS, Dolby or SRS, will be highly bit rate reduced and all efforts toward conserving content should be made prior to the final transmission. Many lessons should have been learned following DAB and HD Radio, not least that the artefacts in the content was directly due to the number of perceptual coding passes in the broadcast chain.

An ADPCM-based algorithm such as Enhanced apt-X® will offer a much lower coding delay and will retain the phasing between the channels. Enhanced apt-X® is generally considered by the Broadcast and Post Industry (and cinema audiences listening to films in the DTS format) to be relatively non-destructive in nature. In addition ADPCM derivatives tend to offer an end-to end system latency of fewer than five milliseconds making it a suitable solution for 5.1 contribution and distribution. Another option is Dolby E, which has been touted as a solution given its success in DTV. Dolby E was specifically designed to ease the transition for DTV broadcasters from two channels to Multi Channel Audio for distribution applications using existing AES3 infrastructures. However, radio broadcast does not suffer from the same constraints as DTV i.e. video frame rates. As such Dolby E has a few fundamental issues that would have to be overcome by radio broadcasters. These include a latency of over 60 milliseconds, bit polling between channels, an inability to process individual channels, a word resolution limited to approximately 22 bits and a set sampling frequency of 48 kHz. Also, and perhaps most fundamentally, Dolby E is based on perceptual coding techniques.

REAL WORLD SOLUTIONS:

There have been a few pioneering broadcasters who have put to air 5.1 pilot projects using live source material. In most cases the live material has been either been Classical or Jazz content. Interestingly, these projects are happening in parallel on both sides of the Atlantic – at ORF in Austria and NPR in the USA being two examples. Unsurprisingly, the Europeans and the Americans have explored two fundamentally different techniques.

TCP/IP

ORF, under the guidance of Karl Petermichl (and closely observed by his fellow EBU members) chose Enhanced apt-X® wrapped up in the WorldNet SkyLink. This unit is a codec with eight discrete channels (5.1 and a Stereo pair) and it uses an Ethernet port to present the compressed data to the outside world. The IP codecs were used for projects that included "Night of the Long Radio" and ORF's New Years Eve broadcast. The discrete channel approach enabled ORF to process individual channels and keep to a minimum the amount of hardware used in the broadcast chain. Data capacity was provided by Austria Telecom, which supplied a 2Mbit/s ADSL circuit.

E1 / T1

It is worth outlining that both WDR and BR in Germany are also using Enhanced apt-X® for 5.1 but are using an E1 interface (for these projects the broadcasters used the APT WorldNet Oslo). The actual transport medium, Synchronous or IP, would appear to be a decision based on what service the local Telco provides.

Required bit rates over a synchronous T1 or E1 circuit will depend on the word depth and Sampling Frequency. Assuming that the Sampling Frequency is 48kHz, the figures below related to word depths of 16 bit and 24 bit:

- Bit Rate for 5.1 @ 16 bit will require:
  5 x 192kBit/s + 64kBit/s = 1024kBit/s

- Bit Rate for 5.1 @ 24 bit will require:
  5 x 256kBit/s + 64kBit/s = 1344kBit/s

Note: the 64kBit/s bit rate is for the Low Frequency Effect channel. This will provide an audio bandwidth from 0Hz to 7Khz. The algorithm uses 4 sub bands and allocates bits to each sub band. The lowest sub band i.e. 0Hz to 1750Hz, will have either 7 or 9 bits allocated (depending on the word depth i.e. 16 bit or 24 bit) as this is where the majority of energy will be focused for a .1 channel.

Over an IP network, the required bit rates are approximately the same as synchronous networks but it is prudent to add an additional 10% for overhead.

ISDN

In the USA, NPR, directed by Mike Papas, has put to air a number of broadcasts including a Diane Reeves concert and the "Toast of the Nation" New Years Eve event in December 2004 and the broadcast from New Orleans in 2005.

Papas used the fundamentally different approach of Watermarking for moving Surround Sound content. In this case, at the concert site Mike and his Production team down mixed the 5.1 channels to 2 channels using the Neural Audio 5225 system. He then transported the two channels using ISDN codecs. The ISDN codec in question was the APT WorldNet Tokyo, which bonds together 4 x ISDN lines to create a 512kBit/s data pipe and uses Enhanced apt-X® at 24 bit word resolution, 48kHz sampling frequency. The algorithm is relatively non-destructive and with a gentle compression ratio. Benefiting from these features removed the incidence of any artefacts affecting the down mix (something that an MPEG algorithm could not achieve in an ISDN codec). At the receiving end, NPR reconstituted the two channels back up to a 5.1 signal.

The results of pilot projects undertaken by these pioneering broadcasters are likely to shape and influence the decisions taken when large-scale deployment of 5.1 multi-channel broadcasting begins in earnest. While ORF and NPR adopted two fundamentally alternative approaches using different equipment and transport mediums (and this presenter does not claim that one approach is superior to another), the constant in both scenarios was the choice of coding algorithm. All parties agreed that only the codecs and compression algorithm mentioned, provided the right balance of network efficiency, latency and flawless audio quality. Regardless of the coding or production techniques used, the radio industry can now put to air live 5.1 material, which should offer a compelling reason for listeners to stay with radio as an entertainment and information medium.

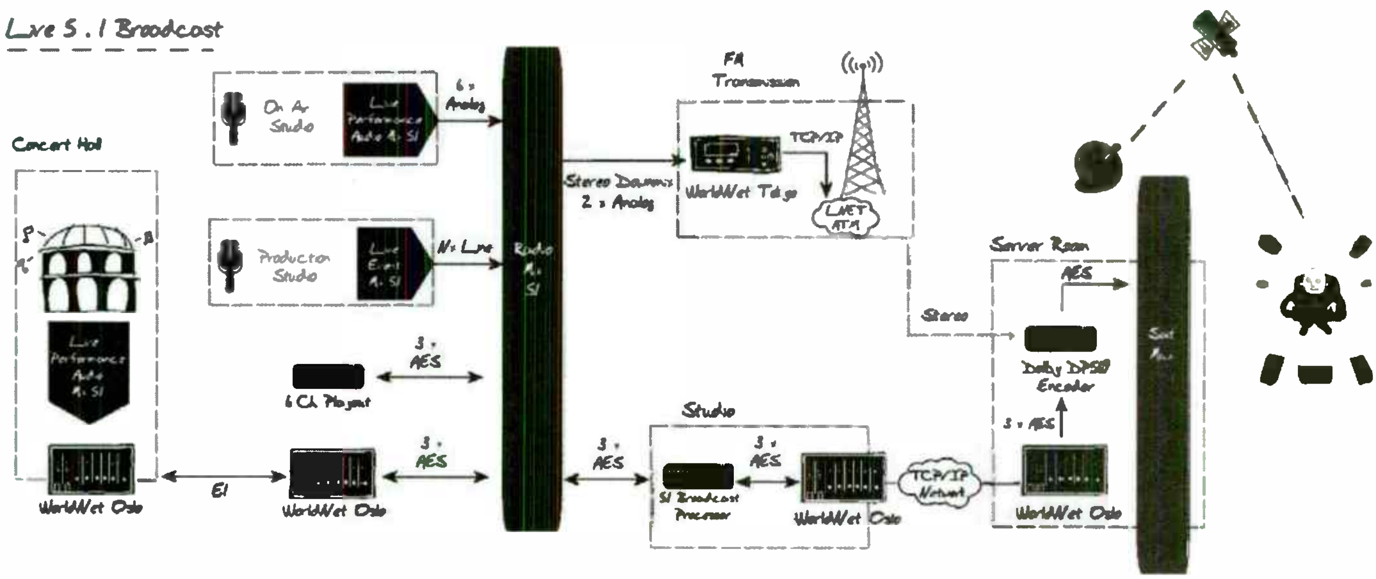
SUMMARY

Now broadcasters can design and build networks capable of capturing live 5.1 content and bringing this to the family home. The next important step will be to take this service from the hands of the Engineers and putting it into the hands of the marketers to ensure the message of 5.1 Surround Sound for Radio is being highlighted to the vast populace.

The future for radio has to be 5.1, it is proven, and we now need to capture the audience.

Acknowledgements :
Mike Pappas, KUVO
Karl Petermichl, ORF

Figure 1



INTERNET RADIO IN SURROUND SOUND

Nikolaus Färber
Fraunhofer IIS, Erlangen, Germany

ABSTRACT

In this paper we describe how Internet Radio in surround sound can be realized today with minimum effort. The system is based on two technology components. On the one hand, Spatial Audio Coding (SAC) allows a fully backward compatible representation of multi-channel audio at bit rates that are only slightly higher than bit rates currently used for stereo. Though SAC can be combined with any audio codec, we focus on MP3 Surround, building on the most widely used audio format on the Internet today. On the other hand, Shoutcast is one of the most commonly used streaming platforms for Internet Radio. The Shoutcast-Server and Winamp-Player are available free and enjoy a very broad user community. Furthermore, Streaming-Hosts and Content Distribution Networks (CDNs) are well established and can be used without any changes. We describe how the platform can be extended to support surround sound and discuss the functionality and availability of the involved components. Based on MP3 Surround and Shoutcast, broadcasting surround sound over the Internet is becoming very easy to implement.

1. INTRODUCTION

There are two current industry trends which make surround sound for Internet Radio an attractive topic. On the one hand, the number of 5.1 speaker installations is increasing constantly in homes – mainly driven through the success of the DVD and home theatre systems (HTS). Based on a report by the China Audio Industry Association (CAIA), the HTS annual output from the mainland is forecast to reach 58 million sets in 2010. For 2006, CAIA projects mainland China sales and export volume, to reach $1.97 billion and 4 million units, respectively. China is of particular interest, since it is the world's major supply center for home theater systems with over 700 manufacturers. Based on industry statistics, more than half of HTS products under famous worldwide brands are made in mainland China factories and enter the overseas market through OEM contracts. These projections are also in line with new figures released by the Consumer Electronics Association (CEA), suggesting that factory-to-dealer sales of consumer electronics are projected to exceed $155 billion in 2007, up 7% from last year.

On the other hand, an increasing amount of music is consumed online, i.e. purchased or streamed from the Internet. This is also true for Internet Radio. In 2006, the American media research company Arbitron, estimated that approximately 30 million Americans have listened to Internet Radio per week, i.e. 12% of the total population [1]. This corresponds to a 50% growth compared to 2005, see **Fig. 1**. Similar numbers are also available for Europe and Germany, where the growth rate for Internet Radio in 2006 is 35% and 45% respectively [2, 3].

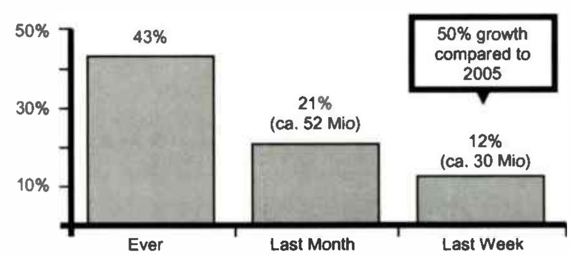


Figure 1: Percentage of US citizens listening to Internet Radio in 2006 [1].

Hence, there are literally millions of 5.1 speaker installations in the homes and their owners are also listening more and more to Internet Radio. Unfortunately, they cannot take full advantage of their new investment in this case because Internet Radio can only deliver mono or stereo music – up to now.

The missing technology link is an effective compression scheme for 5.1 surround audio, such that streaming over the Internet becomes feasible and affordable. Considering that surround sound, as e.g. encoded with AC-3 on a DVD, requires up to 384 kbps it is clear that the distribution costs and Internet access rates are a major obstacle. In particular, since the radio station would e.g. still need to simulcast the usual 128-192 kbps MP3 stereo stream in addition to keep the current audience.

An elegant solution to this problem is Spatial Audio Coding (SAC). SAC is an advanced compression technology which provides a fully backward compatible representation of multi-channel audio at bit rates that are only slightly higher than bit rates currently used for stereo. Backward compatibility allows extending an existing Internet Radio system seamlessly with surround sound. For example, consider the popular MP3/Shoutcast architecture. With SAC MP3 can be extended to MP3 Surround and streamed over existing Content Distribution Networks (CDNs) without any required changes. Legacy players, such as e.g. Winamp can decode the stereo stream. However, after installing a free plugin, surround sound can be decoded from the identical stream. There is no need for simulcasting.

In this paper we first describe the technical background for MP3 Surround and the Shoutcast streaming platform in Section 2 and 3 respectively. Then we explain in Section 4 how the existing components and infrastructure can be extended to support surround sound.

2. MP3 SURROUND CODING

This section briefly summarizes the MP3 Surround format – a backwards compatible extension of the popular MP3 audio coding format for multi-channel audio signals [4, 5]. After explaining the basic idea and underlying technologies, the key features which are relevant for Internet Radio are highlighted.

2.1 Spatial Audio Coding

MP3 Surround is based on the concept of Spatial Audio Coding (SAC) [6] which was developed on the grounds of Binaural Cue Coding (BCC) [7, 8]. The basic idea of this approach is to capture the spatial image of a multi-channel audio signal into a compact set of parameters that can be used to synthesize a high quality multi-channel representation from a transmitted downmix signal as illustrated in **Fig 2**. In the encoding process, the *spatial cues* are extracted from the multi-channel input signal. These parameters typically include Inter-Channel Time Differences (ICTD), Inter-Channel Level Differences (ICLD) and Inter-Channel Coherence (ICC) which can be represented in an extremely compact way. At the same time, a monophonic or two-channel downmix signal of the sound material is created and transmitted to the decoder together with the spatial cue information. Instead of these encoder generated downmixes, externally created downmixes ('artistic downmix') may be used as well which allows the sound engineer complete control over separate stereo and surround productions. On the decoding side, the cue parameters are used to expand the downmix signal into a high quality multi-channel output.

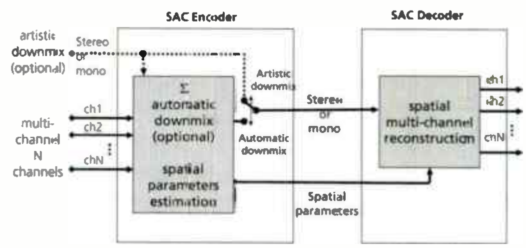


Figure 2: Principle of Spatial Audio Coding

As a result, this approach provides on the one hand an extremely efficient representation of multi-channel audio signals due to the reduced number of audio channels to be transmitted (e.g. just the stereo downmix signal instead of 5 discrete channels). On the other hand, a receiver device without a spatial audio decoder will still be able to produce a meaningful output signal by simply presenting the downmix signal.

2.2 MP3 Surround

The general principle of SAC can be used to extend any audio coding format by applying the codec to the downmix signal. For the specific case of MP3 the resulting format is called *MP3 Surround* [4] which is explained in more detail below.

Fig. 3 illustrates the general structure of an MP3 Surround encoder for the case of encoding a 5.1 multi-channel signal (L, R, C, Ls, Rs, LFE). As a first step, a two-channel compatible stereo downmix (Lc, Rc) is generated from the multichannel material by a downmixing processor or other suitable means. The resulting stereo signal is encoded by a conventional MP3 encoder in a fully standards compliant way. At the same time, a set of spatial parameters (ICLD, ICTD, ICC) are extracted from the multi-channel signal, possibly considering the stereo downmix signals. These spatial parameters are encoded and embedded as surround enhancement data into the ancillary data field of the MP3 bitstream within a suitable data container that unambiguously identifies the presence of such data. Hence, any legacy decoder which is not aware of the extension will gracefully skip the data and simply decode the stereo stream. The bit rate required for the side information is only about 16 kbps. Hence, MP3 Surround allows a representation of multi-channel audio at bit rates that are identical to the bit rates currently used for the coding of mono or stereo, i.e. typically 128-192 kbps. In this case the required bit rate for the side information is typically subtracted from the core codec which then has 112-176 kbps available for the stereo stream.

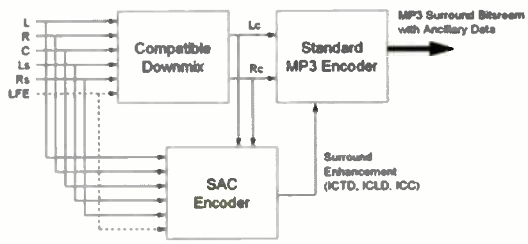


Figure 3: General structure of an MP3 Surround encoder.

Fig. 4 shows the decoder side of the transmission chain. The MP3 Surround bitstream is decoded into a compatible stereo downmix signal that is ready for presentation over a conventional 2-channel reproduction setup (speakers or headphones). Since this step is based on a fully compliant MP3 bitstream, any existing MP3 decoding device can perform this step and thus produce stereo output. MP3 Surround enabled decoders will furthermore detect the presence of the embedded surround enhancement information and, if available, expand the compatible stereo signal into a full multi-channel audio signal using a SAC-decoder.

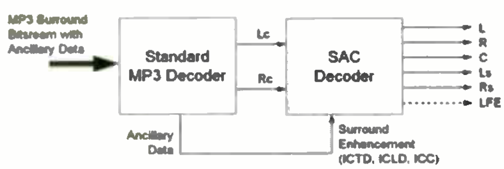


Figure 4: General structure of an MP3 Surround decoder.

Fig. 5 compares the subjective quality of MP3 Surround with Dolby Prologic II [9] which is a matrix-based surround coding scheme commonly used in multi-channel audio broadcasting systems today. While MP3 Surrounds uses a total bit rate of 192 kbps (including the side information), Dolby Prologic II operates on the uncoded stereo PCM data. Despite this advantage, listeners frequently report a change in both the general perception of the sound stage as well as the positioning of certain sound components for the matrix-based approach. On a MUSHRA quality scale from 0 to 100 [9] MP3 Surround was found to have a score above 80 and is therefore in the "exellent" range, while Dolby Prologic II was scored below 70 and therefore only "good". Hence, despite compression to 192 kbps a significant improvement is made resulting in excellent surround quality. For a more detailed description of the test setup we refer to [5].

The use of MP3 Surround for Internet Radio is particular attractive because of the following two properties.

1) Low bit rates. The required bit rates are identical to that of stereo codecs. Typically only 16 kbps are required and the total bit rate for surround sound can be identical to stereo streams, i.e. 128-192 kbps. Hence, Internet Radio station streaming MP3 today can go to surround without extra cost for distribution. Furthermore, these bit rates are totally feasible for listeners who have broadband Internet access, which is becoming very common these days.

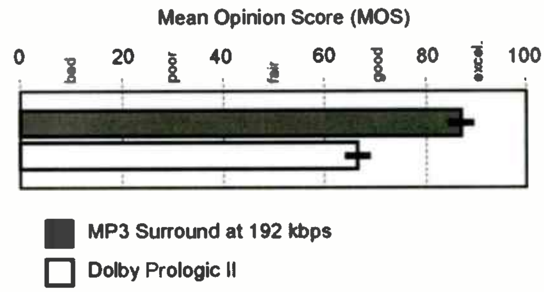


Figure 5: MP3 Surround Subjective Quality [5].

2) Backward compatibility. MP3 Surround is fully backwards compatible with MP3. Already deployed players can decode the stereo stream in uncompromised quality whereas MP3 Surround enabled players are able to generate surround sound from the same stream. Hence, enhanced quality can be introduced without losing existing audience. Furthermore there is no need for simulcasting a new surround stream in parallel to the established stereo stream. This offers a seamless and cost neutral transition from stereo to surround.

For more details on MP3 Surround the reader is referred to our technology overview paper [5]. Finally, it should be noted that MP3 Surround is also getting adopted in other markets. For example, it is in use in audio editing software (MusikMaker by Magix, Cubase4 by Steinberg) and in the new DivX Player. Other companies who have recently licensed MP3 Surround include Samsung and Sony. Hence, the format is being well adopted in the industry which will result in a wide availability of content and the required support for content creation.

3. SHOUTCAST

In this section we describe how audio is distributed over the Internet based on Shoutcast. Though there are many other common streaming platforms – the most important ones being Windows Media (Microsoft), QuickTime/iTunes (Apple), Real (RealNetworks), and Flash (Adobe) – Shoutcast is one of the earliest platforms that were available for Internet Radio. Because of its simplicity, free availability, flexibility, and open interfaces, it is adopted by a broad user community.

It should also be noted that the open source project Icecast [12] is based on the same technology and is mostly compatible with Shoutcast. Interestingly, but not too surprising for open source projects, Icecast has been reported to perform better than the closed source Shoutcast server. The availability of source code and the use of open interfaces makes this technology future prove and extensible. Furthermore, all major player applications (Windows Media Player, QuickTime Player, iTunes, Real Player) are capable of receiving and playing Shoutcast streams.

Before we explain the individual Shoutcast components, we first discuss why broadcasting on the internet is a general problem and how it is solved in practice.

3.1 Broadcasting on a Unicast-Network

The distribution of radio on the Internet is fundamentally different from the distribution of classic radio broadcasting. While the Internet Protocol (IP) is intended for one-to-one communication, i.e. *unicasting*, classic radio relies on *broadcasting*, i.e. one-to-many communication. For classic radio broadcasting it is basically the same technical effort to support one listener than it is to support several thousands. The signal is *on air* anyways and can be received by as many listeners as interested. Since, IP is unicast by nature, it is much more difficult to distribute a signal to a broad audience. Broadcasting actually pushes the Internet to its technical limits and is a severe challenge. This is because unicasting requires that an extra copy is sent over the network for each extra listener – even when the IP packets contain identical audio data and are transported over the same link. This multiplies the load on the sender and bit rate on the network and becomes quickly infeasible when reaching the 1000 listener range.

One solution to overcome this problem is *IP-multicasting,* where special routers duplicate the packets as late as possible along the transmission path. Even though this approach scales very well and is an elegant solution it requires the availability of multicast-enabled routers. Since these are not generally available on the Internet, IP-multicasting is not feasible solution in practice.

Instead, *Content Caching* and *Edge-Servers* have proven efficient to distribute content on the Internet to many users. Here, a single copy of the content – in our case the audio stream – is transmitted from the actual source to an edge-server. The server is responsible to distribute the streams to many receivers by providing a copy from local storage – the cache. In the case of streaming, the cache is dynamically updated with the incoming stream and may hold several seconds of audio data in a ring buffer. In order to save bit rate, the server has to be placed close to the receivers, i.e. at the edge of the network. Therefore, these servers are also known as *edge-servers*. This is illustrated in **Fig. 6** for a sender with four receivers. Even though all data is transmitted using unicasting, the traffic can be reduced significantly in most parts of the network. There are companies hosting edge-servers all over the world, aggregated into *Content Distribution Networks* (CDNs). In order to distribute a stream on the Internet, a radio station can rent a certain capacity in a certain geographic region.

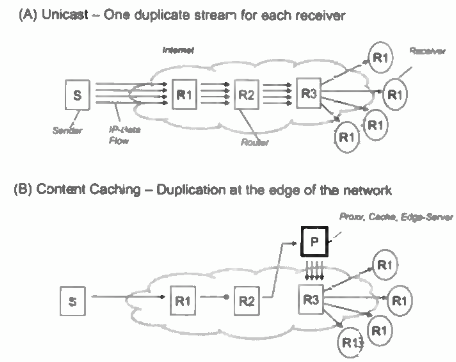


Figure 6: Principle of content caching based on an edge-server compared to unicasting.

3.2 Shoutcast Components

The same principle is also exploited in the Shoutcast streaming platform, which consists of three components: the *source*, the *server*, and the *client*. These components are illustrated in **Fig. 7** and are discussed in more detail below.

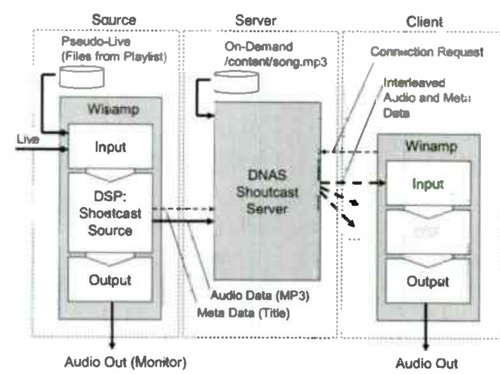


Figure 7: The three components of an Internet Radio system based on Shoutcast: Source, Server, and Client.

Shoutcast Source

The source is where the audio stream actually originates from. For *Live-Streaming*, the source is responsible for audio input, encoding, and transmission to the server. Since music is often transmitted from files during a show, the source also allows to stream titles from a playlist. This is called *Pseudo-Live* since the following components cannot distinguish these operating modes from a protocol perspective. Furthermore, also the user cannot e.g. seek or start a title on-demand if the sources streams from a playlist. It is important to note that even in this case the files are decoded to PCM and then re-encoded or transcoded again to the format used in the transmitted stream. This has the advantage that files of different formats can be used in the playlist while the stream is always of the same format and bit rate. Furthermore, it allows a DJ to do transitions and voice-overs as desired.

For professional studio application there are dedicated hardware units available for this functionality. However, a basic PC with a reasonable sound card is totally sufficient. The common PC-software to run a Shoutcast source is Winamp [10]. Though Winamp is usually known as a client (or player) application, it can be extended to become a source by installing the *Shoutcast Source DSP-Plugin* from Nullsoft. As illustrated in **Fig. 7**, the plugin takes the decoded PCM samples from the input plugin, re-encodes the content with the specified format and bit rate and then transmits the stream to the server. In addition, the audio is also played out locally for monitoring. All required software to operate a Shoutcast source is available free at [10] and [11].

Shoutcast Server

The central component of the Shoutcast platform is the shoutcast server, also known as the Distributed Network Audio Server (DNAS). The server is responsible for an effective duplication of the incomming audio stream and can be used as an edge-server for content distribution (see above). It has a single input stream and can typically handle several 1000 output streams on a standard PC platform.

PC-software implementing a DNAS is available free from Nullsoft [11] and provides a smart user interface on a multitude of operating systems. Even though this makes it very easy to run a Shoutcast server on your own, this is not always the best solution because the accumulated bit rate for all outgoing streams is usually the limiting factor. For example, consider that for the support of only 16 streams at 128 kbps one would need a dedicated 2 Mbit/s uplink at a typical cost of more than 280 USD/month (price based on SDSL offering from T-Systems, dated Jan 2007). However, as noted above, it is also possible to contact a streaming server provider and rent an edge-server instead. Examples for CDN-providers offering Shoutcast servers include StreamGuys (http://www.streamguys.com), ViaStreaming (http://www.viastreaming.com) or StreamMusic (http://stream-music.net). The rent per month is e.g. 4 USD for a 128 kbps stream, which scales proportional to the bit rate (price based on offering from StreamGuys, dated Jan 2007).

When the source connects to the server, it has to provide a password before sending the audio stream. Hence, access to the server can be controlled and it can be assured that nobody is taking a free ride on the server that you are running or renting. In addition, the source can update the metadata, typically the title of the song which is played next.

A client who wants to connect to the audio stream, makes a request to the server. The server will then establish a connection to the client and stream the audio data. For each new client a new connection is established and the stream is sent in multiple copies to the clients using unicast. The stream which is transmitted from the server to the source includes the metadata in an interleaved manner, i.e. after a fixed number of bytes with audio data several bytes of metadata are inserted.

In addition, the Shoutcast server offers the possibility to stream stored files on-demand. For this purpose the audio files have to be placed in the content directory of the server.

Shoutcast Client

As mentioned above, all major client applications can receive and play Shoutcast streams. However, the classic Shoutcast-Client is Winamp [10] which is available for free in its basic version. A screenshot is given in **Fig. 8**. As other client applications, Winamp has many useful features and can be extended through plugins. For example, new codecs can be installed and the visual appearance can be controlled through skins. Furthermore, a huge collection of Internet Radio stations are accessible through the media library of Winamp under the category *Shoutcast Radio*.

Figure 8: Winamp is the classic Shoutcast Client.

Shoutcast Protocol

Unlike many other streaming platforms which rely on RTP/UDP for transmitting audio over IP, Shoutcast is based on TCP with a HTML-like syntax. The fact that TCP is used instead of UDP may seem irritating at first but has many advantages in practice. First, TCP provides a reliable, error-free service. Hence, there are no transmission errors or lost packets which need to be concealed. The related drawbacks of 1) increased delay through retransmissions and 2) variable transfer rates through TCP rate control can be absorbed by a bigger client buffer since Internet Radio can accept several seconds of delay. Second, TCP travels more easily through firewalls, in particular since the client requests and server response used by Shoutcast look like HTTP to the firewall.

A Shoutcast stream starts with a textual header followed by the actual audio data. The header is comprised of several lines of text including key-value pairs such as

```
icy-metaint:1024
```

which e.g. indicates that after every 1024 byte of audio data several bytes of metadata are interleaved (meta data interval). All Shoutcast related keys have the prefix "icy-". Without going into more detail we show a typical communication setup between server and client in **Fig. 9**. The similarity of the client request with an HTTP request is clearly visible. However, there is no strict compliance with HTTP and the term *HTTP-streaming* has to be used with care.

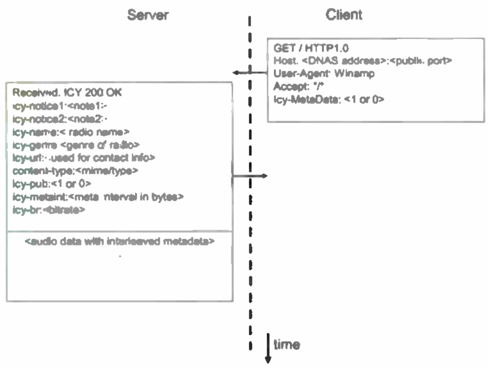


Figure 9: Shoutcast protocol between server and client.

4. SHOUTCASTING MP3 SURROUND

Having provided the background on MP3 Surround and Shoutcast in Section 3 and 4 respectively, we now describe now how these two technologies can be combined for Shoutcasting MP3 Surround.

Since MP3 Surround is basically a new audio format (though fully backwards compatible to MP3) we need to extend functionality at the source and at the client. Interestingly, no changes are required at the server component. This is because the side information is embedded in the MP3 stream, i.e. the Shoutcast server does not even realize that it is carrying MP3 Surround. From his perspective, he is just carrying and duplicating MP3 streams. He does not care about the embedded extension inside. This fact is very important to point out, because it means that the existing distribution infrastructure can be used as is. For example, if a radio station is renting several Shoutcast servers from a CDN-provider to distribute MP3 today, it does not even have to contact him when switching to MP3 Surround. The same servers, the same contracts, and the same passwords can be used. The CDN-provider is completely out of the loop for this format transition. Furthermore, if you want to run your own Shoutcast server you can use the same software as before – available free from Nullsoft. You do not need any adaptation support form Nullsoft. The standard Shoutcast server will be totally sufficient. Hence, we can check off the central and most important component.

Next, let's look at the client. Since the Winamp player can be extended with plugins, all we need is an MP3 Surround plugin capable of receiving a Shoutcast stream. As it turns out, Fraunhofer IIS is providing this plugin for free download from it's web site [4]. **Fig. 10** shows a screenshot of its configuration menu. This plugin is also capable of playing MP3 Surround files and even allows to rendered surround sound via headphones using a virtual surround technology called *Ensonido* [13].

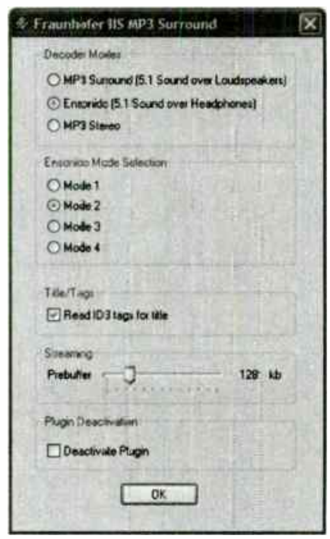


Figure 10: Configuration page of the MP3 Surround Winamp plugin.

Ensonido is a technology for enjoying surround sound on any set of stereo headphones, simulating the natural reception of multi-channel sound by the human ear. Ensonido calculates the sounds that reach the ears from the 5.1 loudspeakers and the reflections and echoes of a listening room. The "out-of-head" localization and the feeling of spatiality achieved that way increases the listening comfort noticeably compared to traditional stereo headphone listening. By combining this technology into the Winamp plugin, it is possible to experience surround sound even when no 5.1 speaker setup is available. The plugin also allows fine tuning to personal preferences by selecting one of 4 modes (see **Fig. 10**).

In summary, the Winamp plugin for MP3 Surround from Fraunhofer IIS provides all required functionality for extending the client. With file playback and Ensonido it even adds significant extra value. Hence, we can also check off the client component.

Finally, we need to consider the source. Also this component is available from Fraunhofer IIS. Either as a DSP-Plugin for Winamp or as a stand-alone application program with a simple GUI. Because the technology is still new and running a Shoutcast source with surround sound may need extra support, this component is not yet available for free download. However, it is provided for evaluation and trials to Internet Radio stations and interested parties are encouraged to contact the Audio and Multimedia (AMM) team at Fraunhofer IIS [14]. Hence, interested radio stations can also get access to the source component and we can also check off this last item on our requirements list.

Initial trials are already ongoing and show that the transition from MP3 to MP3 Surround can be done with little effort. The worldwide first Internet Radio station streaming MP3 Surround is the German campus radio *bit eXpress* which is operated by the University of Erlangen-Nuremberg [15]. Hence, if you want to experience Internet Radio in Surround Sound today, you simply need to download and install the Winamp plugin and enter

http://streaming.bitexpress.de:7002

under *File: Play URL*. If you do not have a 5.1 speaker setup attached to your PC you can still experience surround sound with your headphones by enabling Ensonido. However, in any case, you can listen to the backwards compatible stereo downmix – all from the same URL. Further trails are expected to launch in Q1/Q2 2007 and include e.g. WZLZ - Boston's Classic Rock [16].

5. CONCLUSIONS

In this paper we have described how the popular Shoutcast platform for Internet Radio can be extended to support surround sound.

A key element in this transition is the efficient compression of surround sound which is provided by MP3 Surround. This format allows a representation of multichannel audio at bit rates that are identical to bit rates currently used for MP3 stereo, i.e. 128-192 kbps. In addition, MP3 Surround has the valuable feature of being fully backwards compatible, which allows for a seamless transition from stereo to surround. On the one hand, a radio station can continue to stream using the MP3 format and therefore build on the most widely used audio format on the Internet today. On the other hand, backwards compatibility avoids the need for simulcasting a stereo stream in parallel to a surround stream. Only one stream is needed, no re-design of the web page, no extra cost for distribution. Any player who does not know about the surround extension is still able to decode a perfect stereo stream and therefore the station does not have to be afraid to loose audience in the transition phase.

For the streaming of MP3 Surround over the Internet we have focused on Shoutcast as one of the most popular Internet Radio platforms today. The Shoutcast server, which plays the central role in the distribution of the streams, does not have to be modified at all. Hence, when a radio station relies on a CDN-provider to distribute its MP3 stream to edge-servers all over the world, it can use this arrangement and even existing contracts unchanged. The required infrastructure, which may otherwise be difficult to establish, is already in place and CDN-providers are well experienced with the technology.

The source and client components can also be extended easily. For Winamp, which is the most common client application for Shoutcast radio, a free plugin is available. Hence, any interested user can download this software and become a surround listener. Besides receiving Shoutcast streams and decoding MP3 Surround, the plugin is also capable to playback MP3 Surround files and render surround over standard stereo headphones using Ensonido technology. Finally, also the source component is available upon request for interested radio stations. Therfore, the complete transmission chain can be extended to support MP3 Surround and the generation, distribution, and consumption of surround sound is now as easy a never before.

Initial trials are already ongoing and we expect that in 2007 many other radio stations will start to stream surround sound based on the described technology. An important aspect for radio stations is that streaming over the Internet can be launched very quickly whereas surround sound over terrestrial or satellite broadcast depends on the existence of enhanced receivers. Since those are not as easily introduced as the download and installation of a Winamp plugin, the willingness to establish surround technology in the studio can be low. With Internet Radio, however, the enhanced experience can be delivered immediately to listeners and studios can make valuable experiences for content production and studio equipment. Hence, the operation of Internet Radio in surround sound can also help to prepare the successful launch of surround sound over other radio systems.

LITERATURE AND LINKS

[1] „Internet and Multimedia 2006", Arbitron/Edison Research, available at http://arbitronresearch.com/downloads/im2006pres.pdf

[2] BITCOM press release based on market research by Goldmedia, available at http://bitkom.de/41381_41368.aspx
[3] „ARD/ZDF-Online Studie 2006“, ARD/ZDF, available at http://www.daserste.de/service/studie.asp
[4] www.mp3surround.com
[5] J. Herre et al., “An Introduction To MP3 Surround”, Fraunhofer IIS White Paper available at www.mp3surround.com.
[6] J. Herre, C. Faller, S. Disch, C. Ertel, J. Hilpert, A. Hoelzer, K. Linzmeier, C. Spenger, P. Kroon: "Spatial Audio Coding: Next-Generation Efficient and Compatible Coding of Multi-Channel Audio", 117th AES Convention, San Francisco 2004, Preprint 6186.
[7] C. Faller, F. Baumgarte: “Efficient Representation of Spatial Audio Using Perceptual Parametrization”, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York 2001.
[8] C. Faller: “Parametric Coding of Spatial Audio”, 7th International Conference on Audio Effects (DAFX-04), Naples, Italy, October 2004.
[9] Dolby Publication, Roger Dressler: “Dolby Surround Prologic Decoder - Principles of Operation”, http://www.dolby.com/tech/whtppr.html
[10] www.winamp.com
[11] www.shoutcast.com
[12] www.icecast.org
[13] www.ensonido.com
[14] amm-info@iis.fraunhofer.de
[15] www.bitexpress.de
[16] www.wzlx.com

Emerging Broadcast Technologies
Wednesday, April 18, 2007
9:00 AM – 12:00 PM

Chairperson: Graham Jones
NAB, Washington, DC

DRM Progress Toward Standardization in the Broadcast Frequencies Between 30 and 108 MHz
H. Donald Messer, Digital Radio Mondiale, Washington, DC

***Interactive Radio Has Arrived**
T.A. McCann, Jump2Go, Bellevue, WA

***Improved Data-Follow-Audio in Multiplatform Radio**
Neil Glassman, Broadcast Electronics, Quincy, IL

Exploring the Full Potential of H.264
Faouzi Kossentini, Scientific Atlanta, A Cisco Company, Lawrenceville, GA

Joining Content Production to the DVR
David White, NDS Limited, Hampshire, United Kingdom

A2/M2 Digital Audience Measurement
Scott Brown, Nielsen Media Research, Oldsmar, FL

---

*Paper not available at the time of publication

DRM Progress toward Standardization in the Broadcast Frequencies between 30 and 108 MHz

FRANK HOFMANN
Robert Bosch, GmbH, CR/AEM
Hildersheim, Germany
Frank.Hofmann2@de.bosch.com

H. DONALD MESSER
DRM Steering Board Member
messerhdonald@aol.com

ABSTRACT

Digital Radio Mondiale will offer new opportunities for broadcasters to migrate from analog to digital transmissions. The existing DRM standard developed for the long-, medium- and shortwave bands is being extended to frequencies up to 108 MHz, which is called DRM+. In this paper, we describe the preliminary system architecture, the transmission channel and a solution to avoid flat fading in narrowband systems.

INTRODUCTION

DRM is drawing more and more the attention of the broadcast world when regarding the frequency bands below 30 MHz. To respond to new challenges the consortium decided to extend the system to frequencies up to 108 MHz. Thereafter it can be seen as an alternative to the proprietary system HD-Radio which is not adequate to many countries because of spectrum mask requirements.

The article describes the first attempt of the system architecture and the performance in different mobile channels. The goal is the development of a narrowband system with the flexibility for an introduction also in the crowded frequency band II (88 – 108 MHz). In addition, frequency band I (segments of 47 – 72 MHz, depending on the ITU-R Region) induces the opportunity for new broadcasts. Further only a small number of services should be within one multiplex to prevent complex service multiplex organization at the transmitter side. To cope with the resulting requirements a bandwidth of 100 kHz is planned. This leads to the flat fading effect in urban mobile channels. New diversity techniques help to solve this challenge.

DRM+ development is addressing any country where there is a marketing opportunity either to add new program material or to transmit, on a simulcast basis, a digital version of the existing FM signal. It does this in existing (traditional) broadcasting bands up to 108 MHz, and therefore can be considered to be complementary to DAB (Digital Audio Broadcasting) in countries where using higher VHF and L-band frequencies are permitted.

SYSTEM REQUIREMENTS

The introduction of a new technology into a world market requires a careful analysis of all the needs of consumers, broadcasters, network operators and manufacturer. DRM collected the requirements which are the basis for the technical development. The system will be an open international extension to the already existing DRM standard [1].

The most important requirements considered for the development are:

- Better spectral efficiency than analog FM. It is expected that one audio program needs less than a tenth of the spectrum than analog FM.
- Higher service reliability than analog FM especially in mobile environments.
- Improved audio quality including the support of surround sound.
- Possibility for extended service information features including frequency switching to other systems.
- Capabilities of conveying audio and data to allow new multimedia features.
- System approach should facilitate the development of low cost receivers.
- Support of single frequency networks.

TRANSMISSION CHANNEL

The basis for each communication system is the transmission channel. With the knowledge of the channel characteristics, a channel simulator can be applied to allow performance comparison between different system proposals under predefined conditions. The frequencies in VHF bands I and II are characterized by diffraction, scattering and reflection of the electromagnetic waves between transmitter and receiver. In addition, the movement of the receiver has to be taken into account. This results in multipath propagation and Doppler spread.

Since no deterministic description of the time-variance of the channel is feasible, a stochastic approach is needed. Therefore the WSSUS-model (Wide Sense Stationary Uncorrelated Scattering) was developed and well proven in the last decade [2]. The tap-delay line structure is a suitable representation whereas the relation between input signal *e(t)* and output signal *s(t)* in complex base band is given by:

$$s(t) = \sum_{k=1}^{N} \rho_k \cdot c_k(t) \cdot e(t - \tau_k)$$

N is the number of paths with the respective average power ρ_k and the relative delay τ_k. The time-variant tap-weights $c_k(t)$ are zero mean complex-valued stationary Gaussian random processes with Rayleigh- or Ricean-distributed magnitudes and uniformly distributed phases.

The definition of different channel profiles allows the evaluation of the system in good, typical and worst case propagation scenarios.

Table 1 shows an overview of the relevant channel profiles for DRM+. Each profile is described by the number of paths with their individual delays and Doppler spreads.

SYSTEM ARCHITECTURE

The basis for the extension of DRM to higher frequency bands is to reach as much communality as possible with the existing standard. This approach facilitates the receiver and modulator development and supports a fast market introduction. Since the transmission channel has different characteristics in the VHF bands than in the frequencies below 30 MHz new evaluations become necessary.

System overview

In the following the most important aspects of the system will be explained. All information has to be seen as preliminary since the system is still in development phase. Figure 1 gives an overview of the system architecture.

Several audio and data services are source encoded or pre-encoded. The same MPEG-4 source encoders as used in DRM are available. These are the audio encoder AAC (Advanced Audio Coding) and the two speech encoders CELP (Code Excited Linear Prediction) and HVXC (Harmonic Vector eXcitation Coding). All encoders can be extended by SBR (Spectral Bandwidth Replication), an enhancement tool, which uses information of the spectral envelope of the audio signal to enlarge the audio bandwidth.

The two information channels FAC (Fast Access Channel) and SDC (Service Description Channel) are also maintained from DRM. By this the consistency of multiplex information can be ensured.

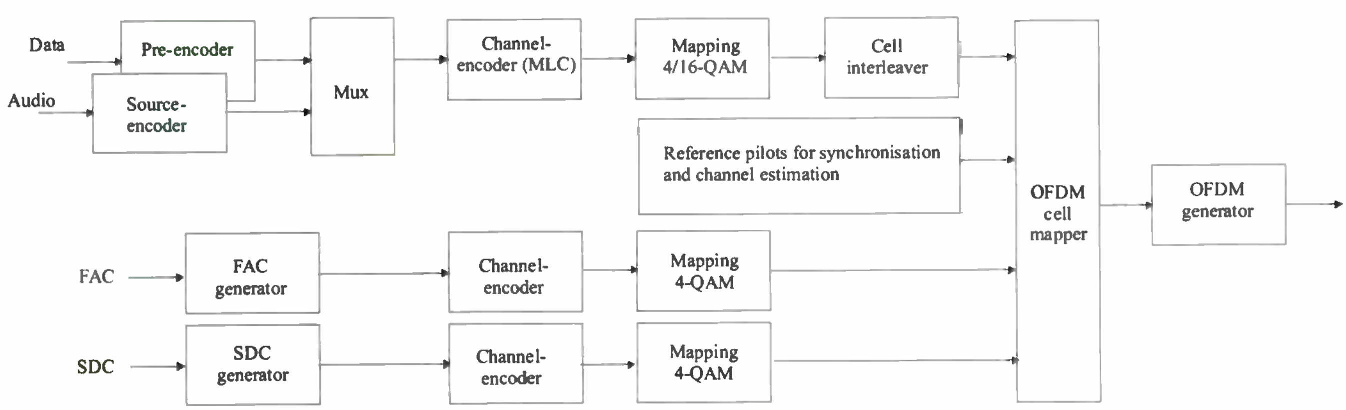


Figure 1. System architecture

Table 1. Channel profiles

No.	Profile Name	Speed v [km/h]	Max. Doppler frequency * $f_{D,max}$ [Hz]	No. paths N	Maximum delay τ_{max} [µs]
1	AWGN (additive white Gaussian noise)	0	0	1	0
2	Urban	2 / 60	0.2 / 5.6	9	3
3	Rural	150	13.9	9	3
4	Terrain obstructed	60	5.6	9	16
5	Hilly terrain	100	9.3	12	82.7
6	SFN (single frequency network)	150	13.9	7	600

* at 100 MHz transmission frequency

Channel coding is done with the aid of convolutional codes and MLC (Multilevel Coding) for the higher constellations. Respectively, for the MSC (Main Service Channel), which contains the useful data, two different constellations namely 4-QAM and 16-QAM, are planned. Several code rates are specified between R=0.25 for 4-QAM and R=0.625 for 16-QAM to reach flexibility concerning robustness and data rate. The implemented interleaver improves the performance of the channel coding system. The MSC and the two information channels FAC and SDC are accomplished with reference pilots which are inevitable for channel estimation and synchronization in the receiver. This whole multiplex runs through the OFDM (Orthogonal Frequency Division Multiplexing) generator to build the time signal which can then be upconverted to the transmission frequency.

OFDM transmission parameters

The idea of the OFDM transmission techniques is the division of the data rate on a plurality of modulated subcarriers. As a consequence the symbol time increases, which in turn decreases the inter-symbol interference in multipath environment. With the insertion of an additional guard interval before each symbol in time domain, the inter-symbol interference can be avoided completely. The individual subcarriers are spaced by the frequency distance $\Delta F = 1/T_u$ to avoid inter-carrier interference where T_u is the time interval of the useful part of one OFDM symbol. Together with the guard time T_g the complete symbol duration $T_s = T_u + T_g$ is defined.

To define the OFDM parameters several design requirements have to be taken into account [3]:

- $T_g > \tau_{max}$: corresponds to the fact that the longest delay of the channel profile should be shorter than the guard interval. With the choice of $T_g = 250\mu s$ the requirement was kept for all channel profiles except No.6. But the very late echoes in the SFN profile have a negligible power.

- $T_g < T_u$: the guard interval should be much shorter than the useful symbol duration to obtain an efficient system. With the choice of $T_u = 2.25$ms the power loss caused by the guard interval is less than 0.5 dB.
- $\Delta F >> f_D$: a rule of thumb says the subcarrier spacing has to be 10 times larger than the maximum occurring Doppler frequencies to avoid inter-carrier interference. With the choice of $\Delta F = 444$ Hz very high speeds could be reached without significant loss in signal-to-noise ratio.

In addition several practical requirements have to be taken into account. One example is the choice of the OFDM transform length (IFFT) which should be a convenient number for implementation.

Pilot pattern

In the following we will have a closer look to the pilot pattern for channel estimation. Figure 2 shows the distribution of the pilots in the time-frequency grid.

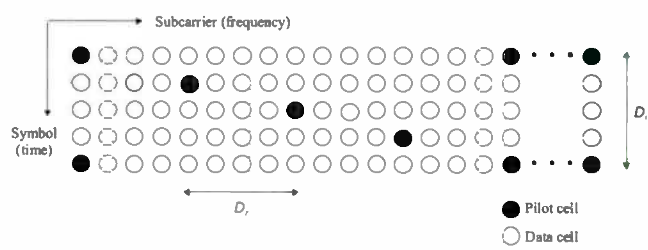


Figure 2. Pilot pattern in the time-frequency grid

The pilot pattern has to be adapted to the transmission channel to ensure good channel estimation in the receiver. Therefore the sampling theorem has to be kept according to the formulas:

$$D_t = \frac{1}{2 f_{D,\max} \cdot T_s}$$

and

$$D_f = \frac{1}{\tau_{\max} \cdot \Delta F}.$$

In addition some room for non-ideal channel interpolation filters has to be taken into account. With $D_t = 4$ and $D_f = 4$

a good compromise between pilot overhead and performance has been defined. This choice allows the receiver to track the channel also at high speed for all channel profiles.

Data rates

The needs in capacity of the broadcasters differ from one high quality audio program to several programs with surround sound and additional data services. The system should provide data rates in the range of 40 kbit/s and 190 kbit/s to satisfy all interests. The two constellations 4-QAM and 16-QAM with code rates between 0.25 and 0.625 deliver adequate choices.

Figure 3 shows the spectral efficiency of example several code rates and the two chosen constellations. (For the 64 QAM plots, the code rates vary from 0.625, the uppermost point to 0.35, the lowermost point; for the 4 QAM plots, the code rates vary from 0.5 to 0.25.) That means the plot present the number of bits which could be transmitted per symbol and subcarrier in dependency of the signal-to-noise ratio (SNR). The dashed line and the dashed dotted line correspond to the union bounds for the Rayleigh channel. They show the theoretically maximum achievable data rate for a bit error rate of 10^{-4}. The other lines show simulation results for the AWGN and Rural area (RA) channel. One can see that the performance for the RA channel is close to the theoretical performance of the Rayleigh channel and the simulation of the AWGN channel even outperforms these limits as expected. In addition the huge range of achievable data rates in dependency of the SNR will become evident.

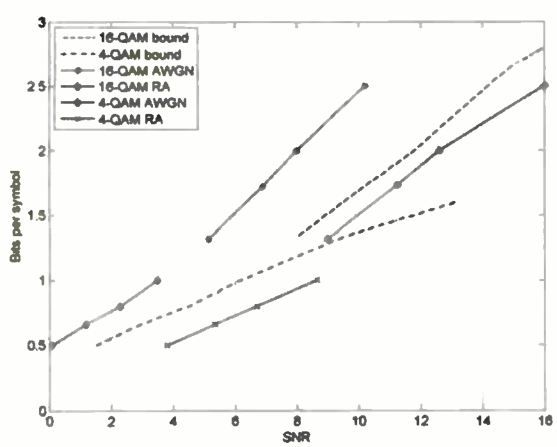


Figure 3. Spectral effeciency plot

TRANSMIT DIVERSITY

The big challenge in system design is the achievement of a good performance in the urban channel profile at low speed. This is caused by the problem of flat fading which occurs when the coherence bandwidth $1/\tau_{max}$ of the channel is smaller than the signal bandwidth. That means a 100 kHz coherence bandwidth corresponds to a delay spread of 10 µs. The channel profiles No.2 and 3 have smaller delay spreads resulting in flat fading. At high receiver speeds as defined for channel profile No.3 the time interleaver is able to solve this issue. But for low speeds other methods should be found to increase the statistical variety of the channel, which we call channel diversity.

Different diversity techniques are worthy of being evaluated, including use of two or more antennas at the transmitter or receiver side. Receive-antenna diversity works well but increases cost and size of all receivers, this is why we concentrate on transmit-antenna diversity. Different techniques can be found in literature [4].

A very interesting attempt is space-time block coding (STBC) with two antennas [5]. Both antennas transmit the same content with different orthogonal waveforms. A good performance could be proofed in simulations but also two drawbacks were found. First, the OFDM symbol duration has to be short in comparison to the time variance of the channel because the fading has to be constant across two consecutive symbols. Second, the amount of pilots doubles because both channel transfer functions according to the two transmit antennas have to be estimated in the receiver.

A better alternative for our application is delay diversity (DD) [6]. The OFDM modulated signal is transmitted over several antennas whereas the particular signals differ only in an antenna specific delay. A more sophisticated method is the cyclic delay diversity (CDD). Here the delay d_l for each antenna is added between the IFFT and the insertion of the guard interval (GI). This relation is depicted in Figure 4.

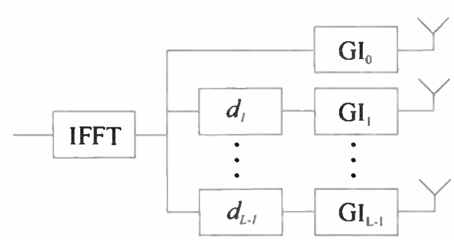


Figure 4. OFDM transmitter with CDD

Both methods DD and CDD have the same performance as far as no inter-symbol interference occurs. The advantage of CDD is that it does not consume a part of the guard interval. The increase in frequency selectivity of the channel reduces the effect of flat fading. The performance in some channel profiles can be increased dramatically.

To judge the system performance the bit error rate (BER) over the ratio of energy per useful bit E_b to the noise density N_0 is the adequate approach. Figure 5 compares the BER of one transmitter with two transmitters CDD.

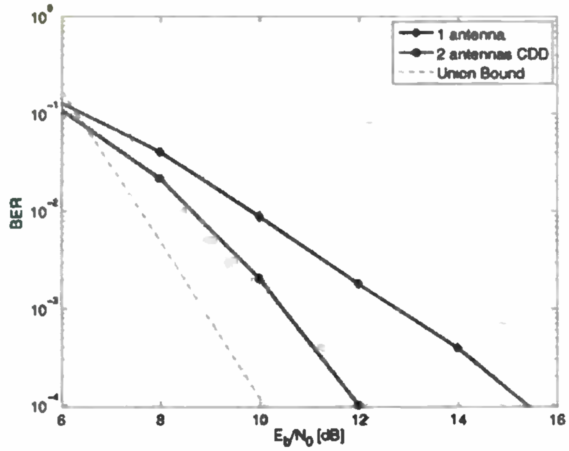


Figure 5. Performance Curves

For the simulation a 16-QAM with code rate $r = 0.5$ and 400 ms interleaving time was used. Channel profile No.2 with 60 km/h was chosen as a difficult but important transmission case. The gain at the target BER of 10^{-4} is more than 3 dB for CDD which is remarkable. This gain decreases with higher speed and for channel profiles with longer echoes.

REFERENCES

[1] ETSI ES 201 980: *Digital Radio Mondiale (DRM); System Specification* V2.2.1 (2005).

[2] Höher, P., "A statistical discrete-time model for the WSSUS multipath channel", IEEE Trans. Vehicular Technol., No. 30, pp 557-558, (1992).

[3] Hofmann, F.,"Multilevel-Codierung und Kanalschätzung für OFDM in der Lang-, Mittel- und Kurzwelle", PhD Thesis, University of Hannover (2003).

[4] Schulze, H.; Lüders, C., *Theory and Applications of OFDM and CDMA*, Wiley (2005).

[5] Alamouti, S., "A simple transmit diversity techniques for wireless communications", IEEE J. Select. Areas Comm., vol. 16, pp 1452-1458 (1998)

[6] Dammann, A.; Kaiser, S., "Transmit/receive antenna diversity techniques for OFDM systems", Europ. Trans. Telecomm., vol. 13, pp 531-538 (2002)

EXPLORING THE FULL POTENTIAL OF H.264

Dr. Faouzi Kossentini
Dr. Ali Jerbi
Scientific Atlanta,
A Cisco Company
Lawrenceville, GA

OVERVIEW

The H.264 video coding standard, also known as MPEG-4 Advanced Video Coding or MPEG-4 Part 10, is poised to revolutionize the video industry in many ways. The most significant way is that even though the H.264 standard currently allows an approximately 2:1 advantage in terms of bandwidth savings over the MPEG-2 standard, it has the potential to allow further bandwidth savings of 3:1 and beyond. It also includes a network adaptation layer as well as several application-targeted profiles which enable its successful use in different video applications and markets. Despite the fact that it employs essentially the same block-based motion-compensation and transform coding structure as previous coding standards, H.264 also supports many important features such as the use of multiple block sizes for motion prediction (from 16x16 to 4x4), intra prediction (16x16, 8x8 and 4x4) and transform coding (4x4 and 8x8), the use of hierarchical Groups-of-Pictures (GOPs) with complex picture dependencies, picture- and macroblock-adaptive field/frame coding, and weighted prediction.

This paper will provide a technical description of the H.264 video coding standard, discuss its most popular profiles and present in detail two of the advanced H.264 coding features: hierarchical GOPs and weighted prediction. Since the proper implementation and optimization of all of the relevant H.264 features will take many more years of research, Scientific Atlanta, a Cisco company, has developed a unique single-slice H.264 encoding architecture with the horsepower and flexibility needed not only to implement all of the relevant H.264 features but also to accommodate all of the future video coding improvements. The Scientific Atlanta architecture, which is discussed at the end of this paper, will help ensure that broadcasters will provide their customers with not only the best possible video quality at the lowest bandwidth at all times, but also the ultimate video quality promised by the H.264 standard.

1. INTRODUCTION

Broadcast video applications are currently undergoing a transition where more and more content is produced and distributed in digital format as opposed to the traditional analog format. The transition is fuelled by the flexibility digital video offers in handling video content at all stages from production to transport/storage to decoding/display. The infrastructure needed to enable this "digital transformation" is being put in place at a rapid pace, allowing both content and service providers to significantly expand their market reach over the next few years.

Service providers in the broadcast industry face unprecedented competition for viewers. In a landscape that was previously dominated by the cable and satellite industries, IP Television (IPTV) companies are now competing for the same customers using their Digital Subscriber Line (DSL) infrastructure. The competitive landscape is driving service providers to develop ways to differentiate their services and to adopt new technological solutions for the encoding and delivery of digital video. Customers would like to receive an ever-expanding number of channels from a given service, implying the need for a more efficient use of bandwidth. At the same time, they are becoming more demanding in terms of the video quality they expect to see on their televisions, requiring sophisticated video pre-processing on the broadcasting side. Despite the increasingly stringent compression and video quality requirements placed on service providers, customers still demand that the cost of service be as low as possible. Therefore, companies such as Scientific Atlanta, a Cisco company, have been developing advanced pre-processing and compression solutions for the most efficient delivery of high-quality digital video.

Early solutions were based on a standard developed by the Moving Pictures Experts Group (MPEG), known as MPEG-2 video, which has been the standard of choice for the broadcast video industry. MPEG-2 video coding operates at bit rates from around 3 to 6 Mbps for Standard Definition (SD) resolution and from around 12 to 20 Mbps for High Definition (HD) resolution. The H.264 standard (also know as MPEG-4 Advanced Video Coding or MPEG-4 Part 10), however, is revolutionizing the broadcast video industry by presenting a viable alternative to MPEG-2 video. The H.264 standard allows encoders to achieve more than 50 percent additional savings in bandwidth as compared to the MPEG-2 standard, hence its strong value proposition. In fact, H.264 has now enabled a substantial market shift, allowing new service providers (e.g., IPTV) to become more competitive in providing SD/HD video services over their networks.

However, H.264 is also a substantially more complex standard than MPEG-2 video, and both the H.264 encoders and decoders are much more demanding in terms of computations and memory than their MPEG-2 counterparts. This, coupled with the substantial amount of research needed to properly implement and optimize all of the relevant H.264 features, makes the development of high-quality H.264 SD/HD encoders a daunting task.

Despite the above-mentioned challenges, Scientific Atlanta was able to develop the world's first single-slice (i.e., one slice per picture) encoder architecture for the development of H.264 High-Profile (HP) encoders. This encoder architecture (henceforth called "the Scientific Atlanta architecture") has set an unprecedented benchmark in video quality at relatively low bit rates, as low as 1.5 Mbps for SD and 6.0 Mbps for HD. The Scientific Atlanta architecture provides huge horsepower and much needed flexibility to implement the full set of the H.264 HP features as well as all of Scientific Atlanta's optimized video pre-processing and compression algorithms. The Scientific Atlanta architecture also allows future video quality enhancements through only software upgrades (i.e., by downloading new DSP/FPGA code).

In what follows, we first provide a technical description of the H.264 standard and its most popular profiles. We then describe in more detail the following two features: hierarchical GOPs and weighted prediction. We conclude our paper with a brief discussion of the Scientific Atlanta architecture.

2. THE H.264 STANDARD

The H.264 standard (reference [1]) consists of an intricate web of the latest video compression advancements that together allow H.264-compliant encoders to achieve huge improvements in both coding efficiency and video quality compared to previous standards-based encoders. Current H.264 encoders can already achieve the same quality as MPEG-2 at essentially half the bit rate; however, more important is the fact that H.264 promises substantial additional improvements in bandwidth efficiency as compared to competing standards (e.g., more that 3:1 advantage over MPEG-2). Moreover, unlike encoders that are compliant with previous standards, well-designed H.264 encoders do not break down when the bandwidth constraints become too severe; rather, they degrade gracefully, gradually softening the reproduced video pictures as the bit rate is reduced.

The H.264 standard was developed as a result of a partnership project that was completed by the Joint Video Team (JVT) which was formed in 2001 by members of ITU-T VCEG and ISO/IEC MPEG. The H.264 original specification, with three defined profiles (Baseline, Main and Extended), became an International Standard in 2003. The so-called Fidelity Range Extensions (FRExt) were standardized later to provide a number of enhanced capabilities relative to the original specification, as well as to define new profiles (a.k.a. the High Profiles).

The main objective behind H.264 was to allow video researchers and engineers to develop high-performance video encoders/decoders that would address the needs of the many different video applications (from broadcast to mobile phone and essentially everything in-between) and delivery networks (e.g., satellite, Internet, 3G) that would be used to carry the coded video data. In order to facilitate this, the H.264 standard is conceptually divided into a Video Coding Layer (VCL) and a Network Abstraction Layer (NAL). The VCL, which is discussed in this paper and is presented in greater detail in reference [2], defines a decoding process that would provide an efficient representation of the video. The NAL, which is also presented in detail in reference [2], enables network friendliness by mapping VCL data to a variety of transport layers such as RTP/IP, MPEG-2 systems, and H.323.

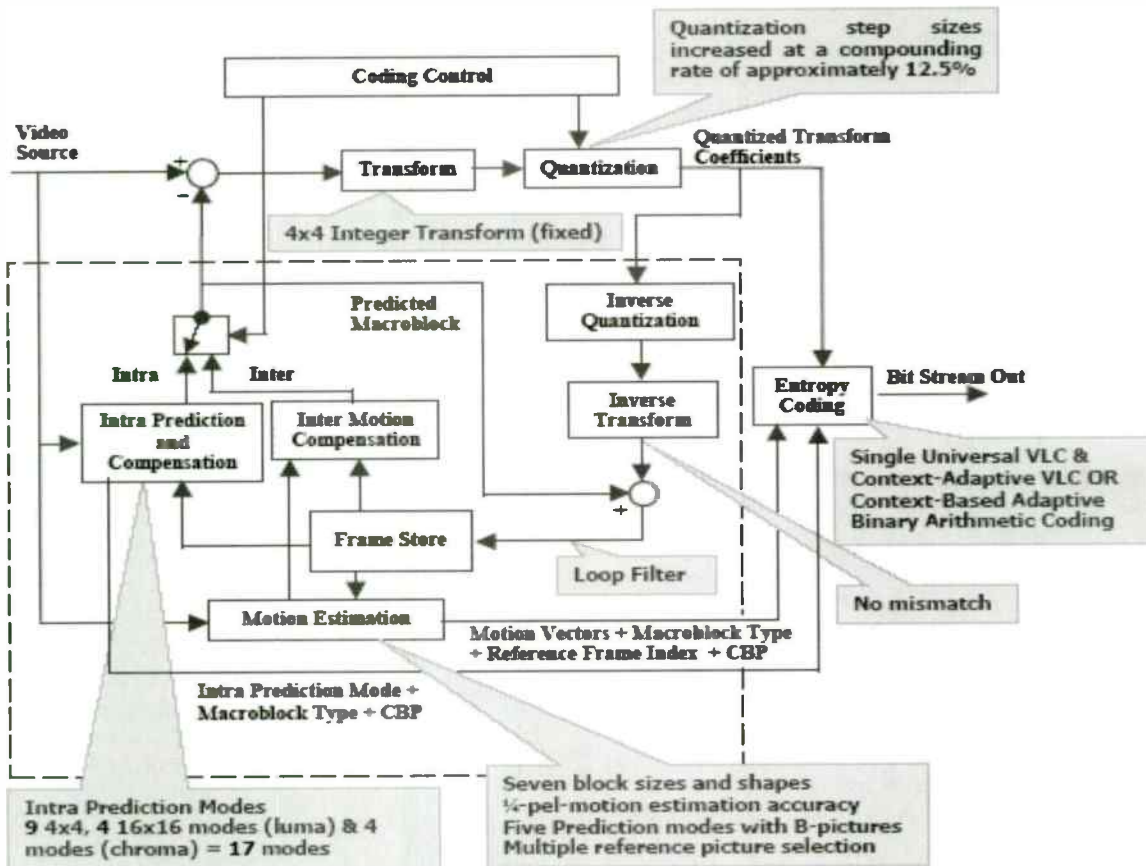


Figure 1. Block Diagram of the H.264 encoder (Original Specification, 2003)

2.1 TECHNICAL DESCRIPTION

Figure 1 shows a block diagram of the H.264 encoding engine described in the original specification. The underlying approach of H.264 is similar to that adopted in previous standards, such as MPEG-2, and consists of the following four main stages:

A. Dividing each video picture into slices, which are then divided into macroblocks (containing 16x16 luminance pixels and, at most, 2 x 16x16 chrominance pixels) so that processing of the video picture can be conducted at the macroblock level.

B. Exploiting the spatial redundancies that exist within the video picture by performing intra prediction and compensation. For any given macroblock, a search is performed to determine the intra prediction modes that are then similarly used by the encoder and decoder to construct a spatial prediction for the subject macroblock.

C. Exploiting the temporal dependencies that exist between macroblocks in temporally neighboring video pictures so that only changes between such pictures need to be encoded. This is accomplished by using motion estimation and compensation. For any given macroblock, a search is performed in the previously coded one or more pictures to determine the motion vectors that are then similarly used by the encoder and decoder to construct a temporal prediction for the subject macroblock.

D. Exploiting any remaining spatial redundancies that exist within the video picture by encoding the residual macroblocks (i.e., the difference between the original macroblocks and the corresponding predicted macroblocks) through transform, quantization and entropy coding.

However, H.264 also supports new features such as the use of different block partitions for intra prediction, inter prediction and integer-based transform coding of a macroblock, video-adaptive de-blocking filtering, hierarchical GOPs with complex inter-picture dependencies, picture- and macroblock-adaptive field/frame coding, and weighted prediction. Moreover, entropy coding in H.264 can be performed using either a combination of a single Universal Variable Length Codes (UVLC) table with Context Adaptive Variable Length Codes (CAVLC) for the transform coefficients *or* using Context-based Adaptive Binary Arithmetic Coding (CABAC).

To encode color images, H.264 uses the YCbCr color space like its predecessors, separating the image into luminance (or "luma," brightness) and chrominance (or "chroma," color) planes. While the original specification supports only the 4:2:0 format (where each of the chroma channels has half the horizontal and vertical resolutions of the luma channel), the recent extensions allow for the support of the 4:2:2 and 4:4:4 formats.

A. Organization of the Bit Stream

An H.264 video stream is placed in discrete packets called NAL units. Each of these packets can contain a part of a slice, that is, there may be one or more NAL units per slice. But not all NAL units contain slice data. There are also NAL unit types for other purposes, such as signaling, headers and additional data. The slices, in turn, contain a part of a video picture. In normal bit streams, each picture consists of a single slice whose data is stored in a single NAL unit. Nevertheless, the possibility to spread pictures over an almost arbitrary number of NAL units can be useful if the stream is transmitted over an error prone network. The decoder may resynchronize after each NAL unit instead of skipping a whole picture if a single error occurs.

The Sequence Parameter Set (SPS) and Picture Parameter Set (PPS) contain the basic stream headers. Each of these parameter sets is stored in its own NAL unit, usually occupying only a few bytes. Both parameter sets have their own ID values so that multiple video streams can be transferred in only one H.264 elementary stream.

The most important fields of an SPS contain:

• A profile and level indicator, signaling conformance to a profile/level combination that is specified in H.264 Annex A of reference [1]
• Information about the decoding method of the Picture Order Count (POC)
• The number of reference pictures
• The picture size in macroblocks, as well as the interlaced encoding flag
• Frame cropping information for enabling non-multiple-of-16 picture sizes
• Video Usability Information (VUI) parameters, such as aspect ratio or color space details

The most important fields of a PPS contain:

• A flag indicating which entropy coding mode is used
• Information about slice data partitioning and macroblock reordering
• The maximum reference picture list index
• Flags indicating the usage of weighted bi-prediction
• The initial quantization parameters as well as the luma/chroma quantization offset
• A flag indicating whether inter-predicted macroblocks may be used for intra prediction (constrained intra prediction)

H.264 defines three slice types (I, P and B), as well as two un-common slice types (SI and SP), which may be used for transitions between two different H.264 video streams. I slices or "intra" slices describe a full still image, containing only references to it. A video stream may consist only of I slices, but this implementation is typically not used. However, the first picture of a sequence always needs to be built out of I slices. P slices or "Predicted" slices use one or more recently decoded (past and/or future) slices to generate a single reference for prediction. B slices or "Bi-predicted" slices work like P slices with a major exception that two references are used for prediction. That is, macroblocks in P slices are predicted from one list (List 0), while macroblocks in B slices use two lists (List 0 and List 1) for prediction. Note that, since List 0 and List 1 can contain exactly the same past/future reference pictures, it is possible to obtain the same predicted macroblock regardless of whether the subject macroblock belongs to a P or B slice.

B. Intra Prediction

Intra prediction refers to the case where only spatial redundancies within a video slice (which is usually a video picture) are exploited. The resulting slice is referred to as an intra slice (or I slice). In order to increase the accuracy of the intra prediction process in H.264, spatial correlation between adjacent macroblocks in a given picture is exploited. The idea is based on the observation that adjacent macroblocks tend to have similar properties. Therefore, as a first step in the encoding process for a given macroblock, one may predict the macroblock of interest from the surrounding macroblocks (typically the ones located on top and to the left of the macroblock of interest, since those macroblocks would have already been encoded). The difference between the actual macroblock and its prediction is then encoded, which results in fewer bits to represent the macroblock of interest compared to when encoding directly to the macroblock itself.

In order to perform the intra prediction mentioned above, H.264 offers nine modes for prediction of the 16 4x4 blocks within the luminance component of a macroblock, including DC prediction (i.e., average of neighboring samples in Figure 2) or Mode 2, and eight directional modes, labeled 0, 1, 3, 4, 5, 6, 7, and 8 in Figure 2.

Figure 2. Intra prediction modes for the 4x4 luminance blocks

Pixels A to M from neighboring blocks have already been encoded and may be used for prediction. For example, if Mode 0 (vertical prediction) is selected, then the values of the pixels a to p are assigned as follows:

The values of a, e, i and m are set to the value of A
The values of b, f, j and n are set to the value of B
The values of c, g, k and o are set to the value of C
The values of d, h, l and p are set to the value of D

For regions with less spatial detail (i.e., flat regions), H.264 supports both 8x8 and 16x16 intra prediction (with the former supported only in FRExt). As in 4x4 prediction, H.264 offers nine modes for prediction of the 4 8x8 blocks within the luminance component of a macroblock, except that low-pass filtering is first applied. However, in 16x16 intra prediction, one of four prediction modes (DC, Vertical, Horizontal and Planar) is chosen for the prediction of the entire luminance component of the macroblock. In addition to the above, H.264 supports intra prediction for the 8x8 chrominance blocks also using four prediction modes (DC, Vertical, Horizontal and Planar). Finally, the prediction mode for each block is efficiently encoded by assigning shorter symbols to more likely modes, where the probability of each mode is determined based on the modes used for encoding the surrounding blocks.

C. Inter Prediction

Inter prediction is based on using motion estimation and compensation to take advantage of the temporal redundancies that exist between successive pictures. Motion estimation and compensation for each 16x16 macroblock can be performed using a number of different block sizes and shapes (Figure 3). Individual motion vectors can be transmitted for blocks as small as 4x4, so up to 16 motion vectors may be transmitted for a single macroblock.

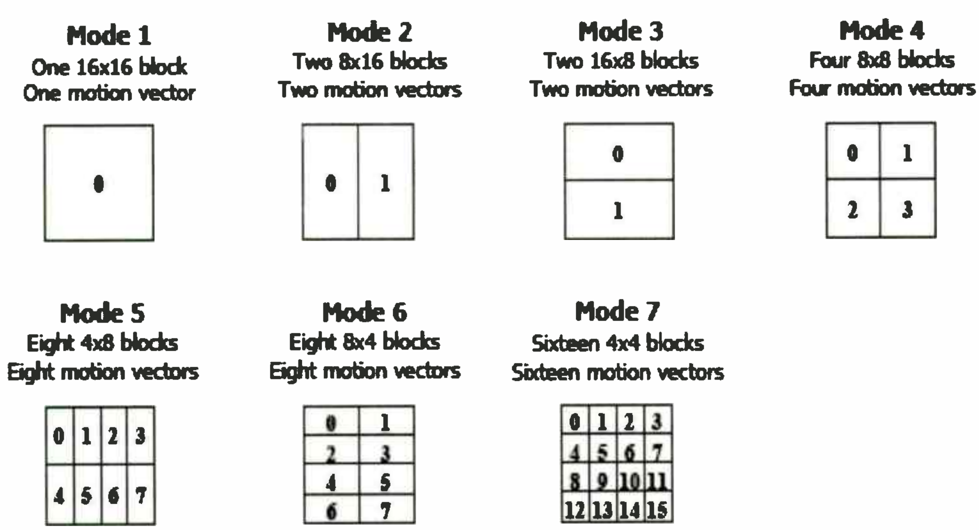


Figure 3. Dividing a macroblock for motion estimation and compensation

Block sizes of 16x8, 8x16, 8x8, 8x4, and 4x8 are also supported as shown. The availability of smaller motion compensation blocks improves prediction in general, and, in particular, the small blocks improve the ability of the model to handle fine motion detail and result in better subjective viewing quality because they do not produce large blocking artifacts. H.264 allows the use of multiple partitions of 16x8, 8x16 and 8x8 blocks, where each can use a different reference picture for prediction. Moreover, through the tree structure segmentation method, it is also possible to have a combination of 4x8, 8x4, or 4x4 sub-blocks within an 8x8 sub-block (see Figure 4). However, for P slices, all sub-partitions within an 8x8 partition must use the same reference picture for prediction. In the case of B slices, partitioning of 8x8 partitions is not allowed.

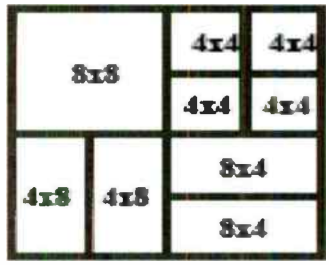


Figure 4. A partitioned 16x16 macroblock

The inter prediction capability of H.264 is further improved by allowing motion vectors to be determined with higher spatial accuracy than in existing standards. Quarter-pixel accurate motion compensation is the lowest-accuracy form of motion compensation in H.264, in contrast with prior standards based primarily on half-pel accuracy, with quarter-pel accuracy only available in the newest version of MPEG-4 Part 2.

H.264 also offers the option of having multiple reference pictures in inter picture coding, resulting in higher coding efficiency, better subjective video quality and higher error resiliency. The H.264 reference picture selection process, which is an enhanced version of the process described in the latest extension of the H.263 standard, involves a mechanism whereby the encoder and decoder maintain one or two lists of reference pictures that can contain multiple previously past/future decoded pictures. The new H.264 multiple reference picture design is considered a significant shift compared to those of previous standards in the way reference buffer lists are constructed and managed. The behavior of the Decoded Picture Buffer (DPB) is capable of adding new reference pictures within a short-term reference list or a long-term reference list, or deleting existing reference pictures through issuing Memory Management Control Operations (MMCO) commands.

Finally, H.264 supports two unique "Direct Mode" (spatial and temporal) prediction modes in B slices in which the motion vectors for the macroblock are interpolated based on the motion vectors used for the encoding of some spatio-temporally neighboring macroblocks. For more details on such modes, please see references [1] and [2].

D. Transform Coding

The information contained in a prediction error macroblock resulting from either intra prediction or inter prediction is then re-expressed in the form of transform coefficients. H.264 is unique in that it employs purely integer spatial transforms (rough approximations of the DCT) which are either 4x4 or 8x8 (see Figure 5) in shape, as opposed to the usual floating-point 8x8 DCT specified with rounding-error tolerances as used in earlier standards. The precise integer specification eliminates any mismatch issues between the encoder and decoder in the inverse transform.

$$\begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 12 & 10 & 6 & 3 & -3 & -6 & -10 & -12 \\ 8 & 4 & -4 & -8 & -8 & -4 & 4 & 8 \\ 10 & -3 & -12 & -6 & 6 & 12 & 3 & -10 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -12 & 3 & 10 & -10 & -3 & 12 & -6 \\ 4 & -8 & 8 & -4 & -4 & 8 & -8 & 4 \\ 3 & -6 & 10 & -12 & 12 & -10 & 6 & -3 \end{bmatrix} \cdot /8$$

Figure 5 - 8x8 transform.

A significant portion of data compression takes place at the quantization. In H.264, the transform coefficients are quantized using scalar quantization with no widened dead-zone. Fifty-two different quantization step sizes can be chosen on a macroblock basis – this being different from prior standards. Moreover, the step sizes in H.264 are increased at a compounding rate of approximately 12.5 percent, as opposed to their being increased by a constant increment. The fidelity of chrominance components is improved by using finer quantization step sizes as compared to those used for the luminance coefficients, particularly when the luminance coefficients are coarsely quantized.

As part of the FRExt extensions for the High profiles and similar in concept to MPEG-2, scaling matrices are supported to allow the tuning of the quantization fidelity to achieve a better subjective optimization. There are two different matrices for 4x4 and 8x8 transforms, and there are two different matrices for inter and intra coding. While Figure 6 illustrates the default scaling matrices, the encoder has the option to use any matrix whose elements range from 1 to 255.

$$\begin{bmatrix} 6 & 13 & 20 & 28 \\ 13 & 20 & 28 & 32 \\ 20 & 28 & 32 & 37 \\ 28 & 32 & 37 & 42 \end{bmatrix}$$

Default_4x4_Intra

$$\begin{bmatrix} 10 & 14 & 20 & 24 \\ 14 & 20 & 24 & 27 \\ 20 & 24 & 27 & 30 \\ 24 & 27 & 30 & 34 \end{bmatrix}$$

Default_4x4_Inter

$$\begin{bmatrix} 6 & 10 & 13 & 16 & 18 & 23 & 25 & 27 \\ 10 & 11 & 16 & 18 & 23 & 25 & 27 & 29 \\ 13 & 16 & 18 & 23 & 25 & 27 & 29 & 31 \\ 16 & 18 & 23 & 25 & 27 & 29 & 31 & 33 \\ 18 & 23 & 25 & 27 & 29 & 31 & 33 & 36 \\ 23 & 25 & 27 & 29 & 31 & 33 & 36 & 38 \\ 25 & 27 & 29 & 31 & 33 & 36 & 38 & 40 \\ 27 & 29 & 31 & 33 & 36 & 38 & 40 & 42 \end{bmatrix}$$

Default_8x8_Intra

$$\begin{bmatrix} 9 & 13 & 15 & 17 & 19 & 21 & 22 & 24 \\ 13 & 13 & 17 & 19 & 21 & 22 & 24 & 25 \\ 15 & 17 & 19 & 21 & 22 & 24 & 25 & 27 \\ 17 & 19 & 21 & 22 & 24 & 25 & 27 & 28 \\ 19 & 21 & 22 & 24 & 25 & 27 & 28 & 30 \\ 21 & 22 & 24 & 25 & 27 & 28 & 30 & 32 \\ 22 & 24 & 25 & 27 & 28 & 30 & 32 & 33 \\ 24 & 25 & 27 & 28 & 30 & 32 & 33 & 35 \end{bmatrix}$$

Default_8x8_Inter

Figure 6. Default scaling matrices defined in High profiles

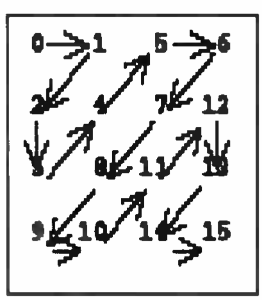


Figure 7. Scan pattern for a transformed 4x4 block for frame coding in H.264

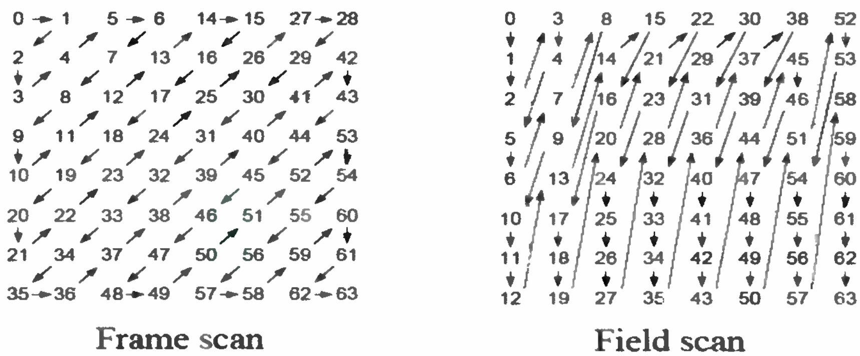


Figure 8. Frame and field Scans for 8x8 blocks in H.264

The quantized transform coefficients correspond to different frequencies, with the coefficient at the top left corner in Figure 7 or Figure 8 representing the DC value, and the rest of the coefficients corresponding to different (usually) nonzero frequency values. The next step in the encoding process is to arrange the quantized coefficients in an array, starting with the DC coefficient. The zigzag coefficient-scanning patterns presented in Figures 7 and 8 arrange the coefficients in an ascending order of the corresponding frequencies, so to maximize the number of consecutive runs of 0s, +1s and -1s..

E. Entropy Coding

The last step in the H.264 video coding process is entropy coding. Entropy coding is based on assigning shorter codewords to symbols with higher probabilities of occurrence, and longer codewords to symbols with less frequent occurrences. Some of the parameters to be entropy coded include transform coefficients for the residual data, motion vectors and other encoder information. Two types of entropy coding have been adopted in the original specification. The first method represents a combination of Universal Variable Length Coding (UVLC) and Context Adaptive Variable-Length Coding (CAVLC). The second method consists of Context Adaptive Binary Arithmetic Coding (CABAC). In FRExt; however, only CABAC is supported.

In the UVLC/CAVLC entropy coding mode, H.264 offers a single Universal VLC (UVLC) table that is to be used in entropy coding of all symbols in the encoder except for the transform coefficients. The transform coefficients are coded using CAVLC, which is designed to take advantage of several characteristics of the quantized 4x4 blocks. First, non-zero coefficients at the end of the zigzag scan are often equal to +/- 1. CAVLC encodes the number of these coefficients (trailing 1s) in a compact way. Second, CAVLC employs run-level coding efficiently to represent the string of zeros in a quantized 4x4 block. Moreover, the numbers of non-zero coefficients in neighboring blocks are usually correlated. Thus, the number of non-zero coefficients is encoded using a look-up table that depends on the numbers of non-zero coefficients in neighboring blocks. Finally, the magnitude (level) of non-zero coefficients increases near the DC coefficient and decreases around the high-frequency coefficients. CAVLC takes advantage of this by making the choice of the VLC look-up table for the level dependent on the recently coded levels.

CABAC makes use of a probability model at both the encoder and decoder for all the syntax elements (transform coefficients, motion vectors, etc.). To increase the coding efficiency of arithmetic coding, the underlying probability model is adapted to the changing statistics within a video picture through a process called context modeling which provides estimates of conditional probabilities of the symbols being encoded. Using suitable context models, given inter-symbol redundancy can be exploited by switching between different probability models according to already encoded symbols in the neighborhood of the current symbol being encoded. The context modeling is responsible for most of CABAC's advantage in coding efficiency over UVLC/CAVLC.

F. De-blocking (loop) filtering

H.264 specifies the use of an adaptive de-blocking (loop) filter that operates on the horizontal and vertical block edges within the prediction loop in order to remove artifacts caused by block prediction errors. The filtering is generally based on 4x4 block boundaries, in which two pixels on either side of the boundary may be updated using a different filter. The rules for applying the de-blocking filter are intricate and quite complex; however, the use of the loop filter is optional for each slice. Nonetheless, the improvement in subjective quality often more than justifies the increase in complexity.

2.2. H.264 PROFILES

Figure 9 provides a compact representation of all of the defined H.264 profiles. The original H.264 specification describes three profiles: (1) Baseline, mainly for video conferencing and telephony/mobile applications; (2) Main, primarily for broadcast video applications; and (3) Extended, mainly for streaming applications. The Baseline profile allows the use of Arbitrary Slice Ordering (ASO) to reduce the latency in real-time communication applications, as well as the use of Flexible Macroblock Ordering (FMO) and redundant slices to improve error resilience in the coded bit stream. The Main profile enables an additional reduction in bandwidth over the Baseline profile primarily through the use of hierarchical GOPs (that include B slices) with complex inter-dependencies, picture- and macroblock-adaptive field/frame coding, weighted prediction and CABAC. Some of these advanced features are discussed next.

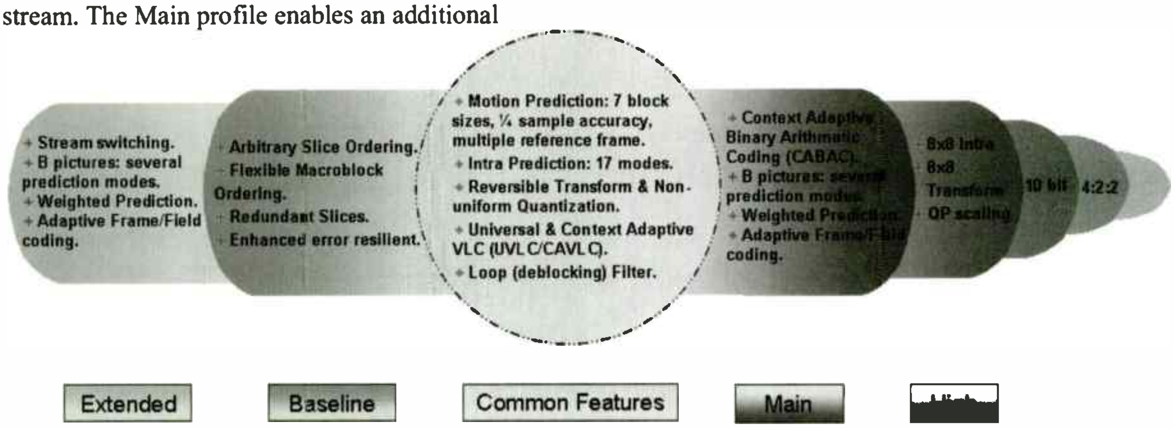


Figure 9. Features for all profiles

The FRExt extensions allow an H.264 encoder to (among several additions) support non-4:2:0 chrominance formats, include a new 8x8 integer transform, switch adaptively (on a macroblock basis) between 4x4 and 8x8 transforms, support 8x8 intra prediction, use one of multiple quant scaling matrices, and determine independently the value of the chrominance quantization parameter. These extensions have led to the development of four additional profiles:

1. The High Profile (HP) with up to 8 bit video and 4:2:0 chroma sampling, addressing applications using high-resolution video without extended chroma formats or extended sample accuracy.
2. The High 10 Profile (Hi10P) with up to 10 bit video and 4:2:0 chroma sampling.
3. The High 4:2:2 Profile (H422P) with up to 10 bit video and 4:2:2 chroma sampling.
4. The High 4:4:4 Profile (H444P) with up to 10 bit video and 4:4:4 chroma sampling, and additional support of efficient lossless region coding and an integer residual color transform for coding RGB video while avoiding color-space transformation error.

The High Profile (HP), the most popular in the broadcast market, supports all of the features of the Main profile and additionally supports other features such as 8x8 intra prediction, an adaptive transform (8x8 or 4x4) and perceptual quantization scaling matrices. The High profile reduces the bit rate further (as compared to the Main profile) by ~10 percent, with higher gains having been observed for HD progressive video content. The subjective quality improvement is even larger, mainly because of the new encoder's support of quantization weighting matrices. Since the High profile includes the Main profile as a *subset*, while also not adding a significant amount of additional implementation complexity, market leaders such as Scientific Atlanta are adopting HP.

2.3 H.264 ADVANCED FEATURES

As stated above, H.264 provides many advanced features that provide additional flexibility that can be exploited by compliant encoders to achieve substantially better video quality and higher coding efficiency. Next, we will focus on two such features: hierarchical GOPs and weighted prediction.

A. Hierarchical GOPs

The concept of Group of Pictures (GOPs) was introduced in previous standards to allow random access points when switching between different channels. In MPEG-2, a typical GOP would consist of one I picture and several P and B pictures, where each B picture uses the nearest past I or P picture and the future P picture as references for prediction. B pictures are not allowed to be used as references. H.264 removes this major restriction, providing the encoder the flexibility to choose whether to use B pictures as references for prediction.

The ability to use B pictures as references for prediction (denoted in the standard as stored B pictures) leads to new ways of defining GOP structures. It is now possible to use only one I picture, one P picture and multiple B pictures in a GOP, where some B pictures would act as reference pictures, thereby creating what is known as hierarchical B pictures. A typical hierarchical

GOP with 4 dyadic hierarchical stages is depicted in Figures 10 and 11. The first picture of the GOP is intra coded, while the last picture in the GOP is predictively coded (P picture). The remaining pictures of the GOP are hierarchically predicted as illustrated in Figure 10 and Figure 11.

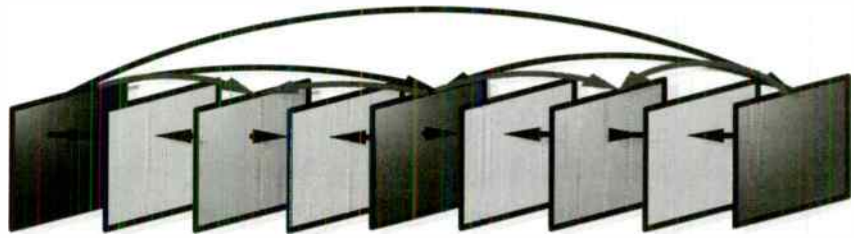


Figure 10. A typical hierarchical prediction structure with 4 dyadic hierarchy stages

As illustrated in Figure 11, the B picture of the Level 1 (the picture labeled as B1) uses only the surrounding I/P pictures as references for prediction. The B pictures Bi of Level $i > 1$ can use the surrounding I/P references as well as the B pictures Bj with Level $j < i$ that are located at the same group of pictures for prediction. The coding order of the pictures in a GOP is "I0 P0 B1 B2 B2 B3 B3 B3 B3…"

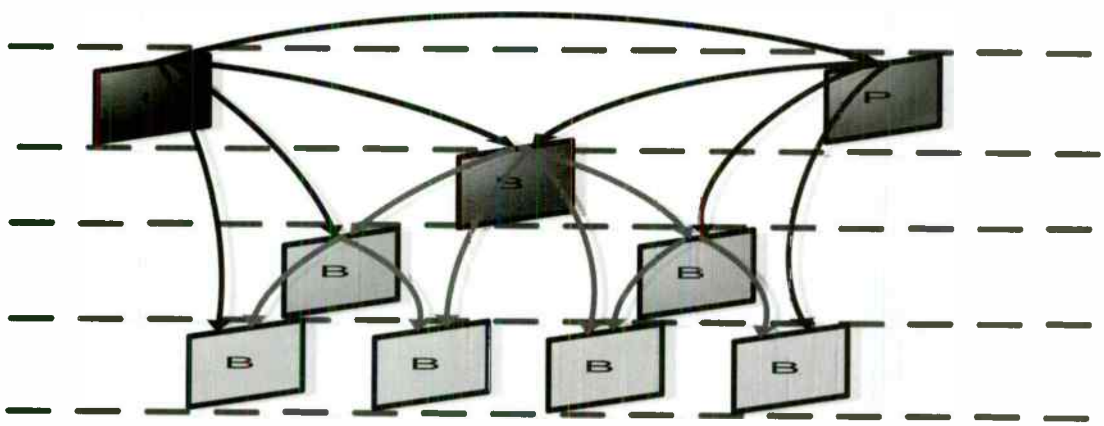


Figure 11. Hierarchical levels

B. Weighted Prediction

Weighted prediction, a unique feature of H.264, allows the modification of motion compensated predicted sample intensities using a global multiplier and a global offset before being added to produce the final predicted sample intensities. The multiplier and offset may be explicitly sent or implicitly inferred (in B slices only). The use of the multiplier and the offset aims at reducing the prediction residuals due, for example, to global brightness changes (see Figure 12 as an example), fades and other special effects. This would in turn enhance coding efficiency and improve video quality in such special cases, besides making an H.264 encoder more stable.

Figure 12: Example of global brightness changes

Explicit weighted prediction, which is allowed in either P or B slices, uses weighting factors and additive offsets that will be applied prior to motion compensation. For each allowable reference picture index in List 0, and for B slices also in List 1, flags are coded which indicate whether or not weighting parameters are present in the slice header for that reference picture index, separately for the luma and chroma components. If the weighting parameters are not present in the slice header for a given reference picture index and color component, a default weighting factor (1 for scaling, 0 for offset) is used.

For fades that are uniformly applied across the entire picture, a single weighting factor and offset are sufficient to efficiently encode all macroblocks in a picture that are predicted from the same reference picture. When multiple reference pictures are used, the best weighting factor and offsets generally differ during a fade for the different reference pictures, as brightness levels are more different for more temporally distant pictures. Use of the reference picture index in the selection of the weighting parameters allows the coding efficiency gain that results from multiple reference picture prediction to be added to the coding efficiency gain that results from weighted prediction. Figure 13 illustrates the quality gain that can be achieved by using the above weighted prediction method.

(a) (b) (c)

Figure 13: Quality gain due to weighted prediction: (a) original picture, (b) with weighted prediction, (c) without weighted prediction

On the other hand, for fades that are non-uniformly applied spatially across an image sequence (e.g. for lighting changes or camera flashes), different slices in the same picture could use different weighting factors even when predicted from the same reference picture.

Use of the implicit mode of weighted prediction is allowed only for bi-predictively coded macroblock partitions in B slices. In such mode, the weighted prediction parameters are not transmitted in the slice header but instead are derived as illustrated in Figure 14. In Figure 14, POC is the Picture Order Count, which indicates the display order of a picture in an H.264 bit stream. Moreover, *Td* is the difference in the POC values between the List 1 reference pictures and the List 0 reference pictures, clipped to the range [-128, 127] and *Tb* is the difference in the POC values between the current picture and the List 0 reference picture, clipped to the same range. In implicit weighted prediction, the weighting factors are proportional to ratio of how far the current picture is to the List 0 reference picture relative to the distance between the List 0 and List 1 references.

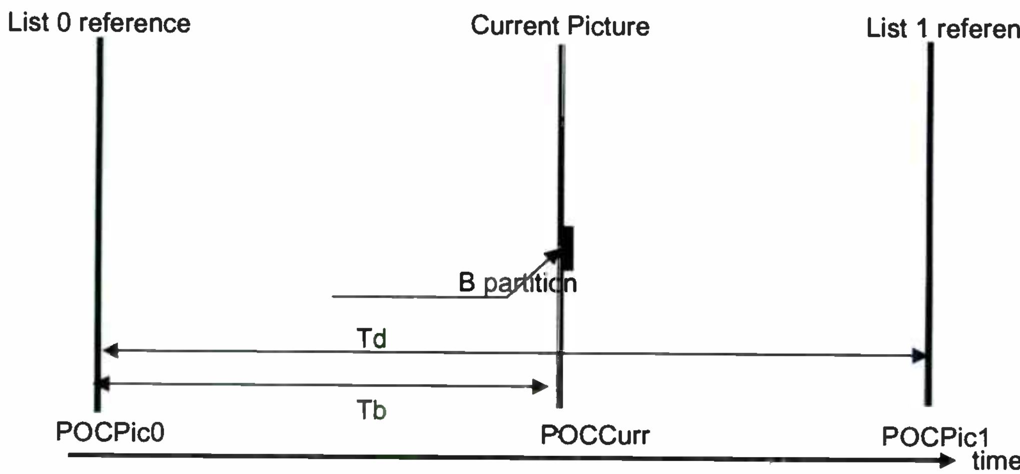


Figure 14. Temporal distance determination

3. THE SCIENTIFIC ATLANTA ENCODER ARCHITECTURE

The Scientific Atlanta H.264 SD/HD encoder is designed to support resolutions up to 1080i30 and bit rates up to 25Mbit/s. This, coupled with the need to implement/optimize all of the above H.264 HP features as well as our pre-processing, intra/inter prediction, mode decision and rate control algorithms, has led to the development of the Scientific Atlanta architecture that is based on dedicated Field Programmable Gate Arrays (FPGAs) and Digital Signal Processors (DSPs). The FPGAs are used to perform all of the "heavy lifting" functions, while the DSPs perform the key decisions in functions such as pre-analysis and rate control.

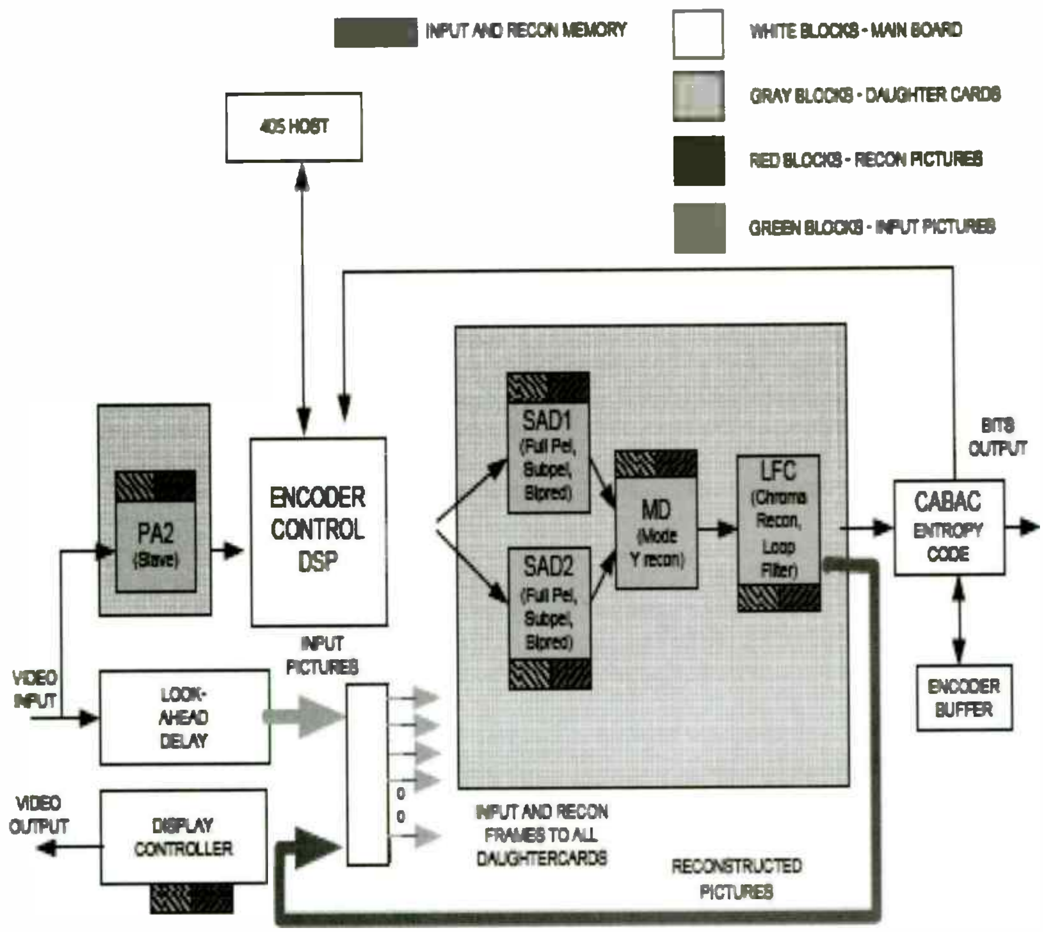


Figure 15. Overview of the SA H.264 SD/HD encoder architecture

As shown in Figure 15, the Scientific Atlanta architecture consists of several sub-systems where each sub-system performs a certain function of the core video encoder. The Pre-Analysis (PA) sub-system implements all of the pre-analysis algorithms, which produce scene complexity measures, as well as macroblock (MB) bit production and motion vector estimates, needed for GOP selection, motion estimation and rate control. The SAD sub-system performs extensive full-pel and sub-pel motion search, and the refined full-resolution motion vector estimates are communicated to a Mode Decision (MD) sub-system that determines primarily whether the subject macroblock should be skipped, intra-coded or inter-coded , as well as selecting all of the many underlying sub-modes. The MD sub-system also performs the reconstruction of the luminance pixels once the mode decision process is completed. Finally, the Loop-Filter-Chrominance (LFC) sub-system implements the reconstruction of the chrominance pixels and the H.264 de-blocking (loop) filtering of the luminance and chrominance reconstructed pixels.

The SA architecture performs all of the above complex operations without splitting the picture into slices, maintaining consistently good video quality for two reasons. First, rate control is inherently more effective because it will have easy access to the statistics of the whole picture. Second, not having to re-set statistics at the beginning of each slice within a picture will maintain high CABAC efficiency.

3.1 PERFORMANCE OF THE SCIENTIFIC ATLANTA ARCHITECTURE

The Scientific Atlanta architecture enables the resulting SD/HD H.264 encoders to achieve not only the best possible video quality at the lowest bandwidth, but also eventually the full potential of H.264. This is due to its huge horsepower and flexibility, both of which are discussed next.

A. Horsepower

In order for video quality never to be compromised, the Scientific Atlanta architecture provides unprecedented horse-power to implement (1) all of the above-discussed H.264 HP features, (2) our current pre-processing, intra/inter prediction, mode decision and rate control algorithms, and (3) our future video coding improvements, all without splitting the picture into multiple slices during the video encoding process.

Our pre-processing algorithms include powerful noise reduction techniques, as well as multi-pass look-ahead spatio-temporal analysis methods. The spatial analysis methods, which involve the computation of spatial activity measures, are used for frame/field coding decision and prediction of bit rate. The temporal analysis methods, which involve the computation of motion activity measures, are used for prediction of bit rate, detection of scene changes and detection of fades.

The Scientific Atlanta SD/HD encoder selects one intra-coding mode for the luminance part of an MB from four 16x16, nine 8x8 and nine 4x4 intra coding modes. It also selects one intra coding mode for the chrominance component of an MB from four 8x8 intra coding modes. Moreover, the Scientific Atlanta encoder performs a three-stage hierarchical motion search, with an adaptive search window size and a large motion search range.

The Scientific Atlanta encoder also performs both high- and low-level rate control to guarantee a stable encoder operation while maintaining consistently high video quality. The high-level rate control takes into account the target bit rate, the scene complexity results of the pre-analysis sub-system and some feedback information to distribute the bits among the various pictures of the subject GOP and determine an average quantization value for each picture, among other operations. The low-level rate control module is a complex MB-adaptive rate control system that maximizes subjective quality using locally computed spatio-temporal complexity measures.

The Scientific Atlanta encoder performs all of the above complex operations without splitting the picture into slices, maintaining consistently good video quality for the following reasons. First, rate control is inherently more effective because it will have easy access to the statistics of the whole picture. Second, not having to re-set statistics at the beginning of each slice within a picture will maintain high CABAC efficiency. In addition, loop filtering, which cannot be applied across multiple slice boundaries, will be applied to the full picture, reducing significantly blocking artifacts. Finally, the Scientific Atlanta encoder will not have to perform motion compensation across slice boundaries, which would have been needed to avoid unnecessary blocking artifacts in multiple-slice encoding.

The Scientific Atlanta architecture provides incredible horsepower, through a design that guarantees not only the maximum number of operations per second but also the fastest data flow. In fact, a recent analysis of the current loading percentages for some of the encoder sub-systems has shown that most DSP/FPGS devices are less than 50 percent full, demonstrating that there is abundant processing room for additional features (including all of the 8x8 HP features) and for future video coding improvements over the next 3-5 years.

B. Flexibility

Being based on fully programmable devices, the Scientific Atlanta architecture provides maximum flexibility for future video quality enhancements through simply software up-grades via downloading new DSP/FPGA code.

The DSPs, which perform all key decisions and implement the most intelligent features, can be easily re-programmed when we have developed (for example) better coding decisions and more advanced high- and low-level rate control algorithms. The FPGAs, which perform all of the remaining tasks, can also be re-programmed to implement new features and algorithms towards more advanced spatio-temporal pre-analysis, more effective intra/inter prediction, and more accurate rate/quality mode selection.

4. CONCLUSION

In this paper, we presented a technical overview of the H.264 standard, with particular emphasis on advanced features such as hierarchical GOPs and weighted prediction. Unlike previous visual coding standards, H.264 is a complex standard offering unprecedented flexibility through numerous features and allowing much greater innovation in the development of H.264 compliant encoders. In fact, current H.264 encoders already vary considerably in their achievable levels of video quality, and it will likely take many years to achieve the full potential of H.264, which is expected to allow compliant encoders to attain the same video quality level of the current MPEG2 encoders at a small fraction of the bit rate.

Since the development of well-optimized H.264 HP SD/HD encoders will take many more years of research, Scientific Atlanta, a Cisco company, has developed a unique single-slice encoding architecture with the horsepower and flexibility needed to properly implement and optimize all of the relevant H.264

features and to accommodate all of the future video coding improvements. The Scientific Atlanta architecture, which is the subject of the second part of this paper, will help ensure that broadcasters will provide their customers with not only the best possible video quality at the lowest bandwidth at all times, but also the ultimate video quality promised by the H.264 standard.

5. REFERENCES

[1] ISO/IEC 14496-10, "Information technology - coding of audio-visual objects – Part 10: Advanced video coding"; also ITU-T Recommendation H.264: "Advanced video coding for generic video services", 2003.

[2] Special issue on the H.264/AVC video coding standard, IEEE Transactions on Circuits and Video Technology (CSVT), volume 13, number 7, July 2003.

[3] The H.264 Advanced Video Coding standard: Overview and introduction to the Fidelity Range Extensions (FRExt), presented at the SPIE conference on Applications of Digital Image Processing XXVII, special session on Advances in the new emerging standard: H.264/AVC, August 2004.

Authors' Biographies

Dr. Faouzi Kossentini | Director of Compression Technology, Digital Media Networks
Scientific Atlanta, A Cisco Company

Faouzi Kossentini is presently the Director of Compression Technology for Digital Media Networks, Scientific Atlanta, A Cisco Company. Prior to SA, Dr. Kossentini served as the Founder, President and CEO of UB Video, and has made UB Video a worldwide leader in H.264 video encoding technology. Dr. Kossentini also served as an assistant professor and associate professor at the Department of Electrical and Computer Engineering at the University of British Columbia, doing research in the areas of signal processing, communications and multimedia. Dr. Kossentini is a senior member of the IEEE and has co-authored more than 200 journal papers, conference papers, book chapters and popular technical white papers. He's also led numerous international ISO and ITU-T activities involving the standardization of JBIG-2, JPEG-2000, H.263 and H.264. He holds B.S., M.S., and Ph.D. degrees from the Georgia Institute of Technology, Atlanta.

Dr. Ali Jerbi | Engineering Manager, Digital Media Networks
Scientific Atlanta,
A Cisco Company

Ali Jerbi is presently an Engineering Manager for Digital Media Networks, Scientific Atlanta, A Cisco Company. Prior to SA, Dr. Jerbi served as the R&D Manager and General Manager of UB Video, where he was responsible for the development of standards-based video encoding products, particularly H.264-based products for the video conferencing and broadcast markets. Dr. Jerbi also served as a Research Officer in the Vision and sensors group of the National Research Council of Canada where he led many industrial projects involving machine vision. Dr. Jerbi is an IEEE member, and he co-authored many journal, conference and technical papers in the areas of video analysis and coding. He holds the B.S., M.S., and Ph.D. degrees from the Georgia Institute of Technology, Atlanta.

Joining content production to the DVR

DAVID WHITE
NDS, UK

ABSTRACT

Digital video recorders (DVRs), also known as personal video recorders (PVRs), are increasingly popular products which record video content onto a mass storage device, such as a hard disk.

Since the concept of the DVR was first introduced it has had the potential to revolutionise our television viewing, in particular making recording easier and offering new ways of accessing content. Now that DVRs are here they have certainly had an impact with viewers, freeing them from the schedule to watch the shows they want, when they want them. However, the ability of DVR random access technology to jump to points of interest, or splice content together, has not been fully exploited. What is missing is not the capability in the DVR, it's the availability of the appropriate metadata.

Joined Up PVR (JuP) is a new industry group aiming to change this by defining how the right metadata, produced in current content production systems, can be passed through the broadcast chain all the way to the PVR/ DVR. By integrating this technology with content production it will be possible to enable viewers to skip back to review a point in a live football game, browse between sections of a recorded magazine programme, or navigate to a particular point in a film, as we can already do with DVDs.

DVRS AND MEDIA PLAYER DEVICES

The first consumer DVRs were unveiled at NAB in 1999 and were touted as revolutionising television viewing. DVRs are now a truly mass-market device and consumers are waking up to life with hard disks rather than tapes. DVRs don't exist in isolation, they are an important part of a wave of new media player devices for viewing content.

DVRs featuring NDS technology. DVR playback can skip to any point in a recording.

Media player devices differ enormously in form factor and capability. Some devices are portable, others are installed in the home. Some are multi-function devices, others are not. There are also different display capabilities, from screens a couple of inches across to enormous high definition displays.

Although the capabilities of these devices differ, they do have some important functionality in common.

1. DVRs and media players can store content for viewing when the viewer wishes, not limiting the viewer to watching when online or when content is scheduled for broadcast.
2. DVRs and media players support metadata for describing the content. This metadata always includes the title but may also carry additional information such as synopsis, cast list, etc. This metadata allows viewers to select items from a "library" of content held on the device. This navigation to an item of content is termed "***macro-***navigation".
3. DVRs and media players have random access storage media. Therefore content playback can jump to particular points in the content. Metadata describing points of interest in the content allows viewers to navigate within content. To distinguish this type of navigation from macro-navigation, navigation within content can be termed "***micro-***navigation".

Media players come in many form factors but random-access memory means micro-navigation.

Many broadcasters are designing their content production and playout systems to cater for this plethora of new devices, distributing content via mobile phone networks, the internet, DVDs and broadcast. However, to take advantage of the full range of functionality of these devices additional information to the audio and video is required. The problems of describing content to enable macro-navigation have been largely solved due to the roll-out of digital television. However, these new devices also support micro-navigation, requiring an approach to metadata more tightly integrated with the production and broadcast of content than previously necessary.

Even devices with limited interfaces can enable micro-navigation using skip forward / back keys.

APPLICATIONS OF MICRO-NAVIGATION

So, what can be done with micro-navigation metadata? Micro-navigation metadata can be used to enable a number of applications, including navigating films, live broadcasts and generating a précis view of content.

There are two main categories of use of micro-navigation: live content, and off-line content. These categories have differing characteristics, effecting what applications are possible and the ways in which the metadata is produced and disseminated.

The most basic form of metadata required for micro-navigation is the time offsets of points of interest within the content. Once this basic segmentation information is available to a DVR it is possible to enable a viewer to navigate between these points of interest simply by pressing a "skip forward" or "skip backward" key. Other, more advanced, functionality builds on this basic metadata by adding more information. Adding text to each segment allows a menu to be displayed: adding end points to segment metadata allows parts of the content to be viewed discretely, etc.

Micro-navigation example: NFL game

Live generation of micro-navigation data alongside a live sports event, such as an NFL game, is a powerful application. With this metadata a DVR can enable viewers to navigate to significant events in the game once they have occurred.

So, for example, if there is a particularly bad tackle the content producer could immediately send micro-navigation metadata referencing the time the incident occurred. When a viewer then pressed "skip back" the playback would jump back and the viewer would watch the incident again, no doubt protesting down the phone to his friend who supports the other team!

Another application of micro-navigation possible in live production is catching up with real-time. Once the viewer has finished watching an event in the past he will need to catch up with what is happening now. This can be achieved today by pressing fast forward and skipping through content – and through sponsors' graphics. However, if the content producer is marking sections of interest as the game unfolds, the DVR can catch up with real time by only watching those interesting parts, leaving out the dull bits. A simple extension of this can also leave in sponsors' graphics.

Micro-navigation example: magazine programme

Magazine programmes are a staple of television viewing. They also have a structure that makes them ideal for adding micro-navigation application to.

The simplest application is just to provide references to the start of each "article" of the programme. This would allow someone to quickly navigate to their favourite part or maybe skip something that doesn't interest them and would otherwise cause them to stop watching.

By enhancing the presentation a little a magazine programme can become a truly magazine-like experience. Enhancing the bare micro-navigation data with graphics and interactivity can make the content engage with viewers, enabling them to browse around the content as they would flick through a magazine

If a viewer records the entire series then micro-navigation metadata can be used to provide alternative views on the content, providing the opportunity to watch all the instances of a regular feature. For a series on motoring, for example, a viewer could choose to watch all of the super-car reviews, or all the features on celebrities and their cars.

THE SOURCE OF MICRO-NAVIGATION METADATA

There is a classic "chicken and egg" problem to making micro-navigation work. DVRs are unlikely to support micro-navigation in their software (their hardware already supports it) unless the micro-navigation information is provided to the broadcast platform. However, the micro-navigation information is unlikely to be provided if DVRs are unable to make use of it. What is needed is to bring together people involved in both ends of this problem to find a solution.

It is not enough to standardise the format for delivering micro-navigation metadata from the headend to the DVR, or even the format for providing micro-navigation metadata from the broadcaster to the broadcast platform headend . This job has been done by organisations such as TV-Anytime. There is a whole chain of devices before the broadcast stage which have to understand, perhaps manipulate, then pass on this metadata. This chain of devices goes right back to content production,. It is content production where this information is generated most efficiently and where this metadata in many cases actually exists today. Sourcing this metadata would require very little change to current workflows, if any.

Sourcing metadata in live content production

Modern news production systems are based around a disk-based video server which is driven using a high degree of automation, generally based around the MOS (Media Object Server) protocol. Before a story is broadcast, the news reader's script, video clips, images and on-screen graphics are prepared by journalists. Stories are then placed in the running order for a news broadcast. During the broadcast human intervention is typically used to trigger actions to start, such as displaying a graphic during a story, but it is the automation system which lined up the correct resources and then responded to the human intervention to cause them to be displayed.

An implication of this level of automation is that these systems often know the precise frame at which each story starts. This means that many news automation systems have the metadata needed to enable micro-navigation. The same is true for other types of live production including many live sports production systems. For systems such as this the only remaining problem, once the metadata is extracted, is getting the metadata to the DVR .

For other automated production systems some sort of additional input would be required to generate the metadata for micro-navigation. This could be as simple as a teleprompter operator having an additional button to press as he monitors the news reader moving from one script to the next.

Sourcing metadata in off-line production

DVDs are the most prominent content format today supporting micro-navigation. Although content for broadcast is regularly released on DVD, this is usually a post-production process done after the content has been delivered to the broadcaster.

However, in order to support DVD formats, many editing systems support the production of the type of information required for micro-navigation. This leaves two things which are required that is required to produce this metadata, these are a format in which to export micro-navigation metadata, and a contractual agreement between the content production company and the broadcaster to generate micro-navigation. In a competitive environment where micro-navigation is supported in target devices (DVRs, media players, etc), this relatively simple additional editing task can add value to content, therefore providing the incentive to produce the micro-navigation metadata. As an added benefit, once this metadata has been produced for the broadcaster it can be re-used in DVD production, thereby reducing costs.

JOINING UP DVRS WITH CONTENT PRODUCTION

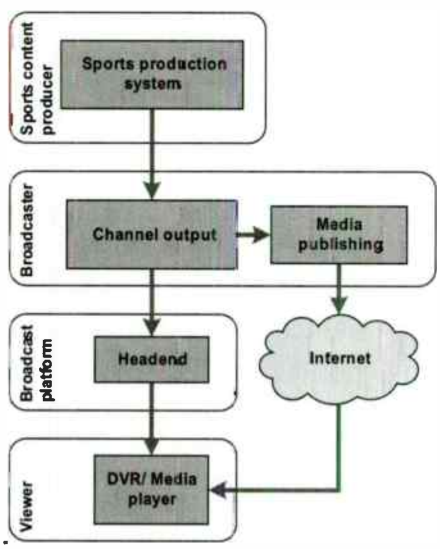
So far we have seen that the technology exists now to enable micro-navigation, it exists in mass-market devices such as DVRs, portable media players and PC media players. In fact this functionality is already available in some cases. We have also seen that it is efficient to produce micro-navigation metadata in content production and that in many cases the information is already there, in the production system. What is required is to *join-up* content producers with DVRs by building the metadata chain.

Because micro-navigation metadata is so closely associated with content, it follows that this metadata follows the path of content through broadcasters' playout systems.

Joining up DVRs with live content production

Of the two categories of use of micro-navigation, the one that is easier to achieve is live content. That is, joining up DVRs with live content production.

Live content is broadcast as it is produced and so it bypasses playout and scheduling systems. Instead, it is fed directly through to the channel output, from here the content is fed to the various platforms carrying this channel. The micro-navigation metadata must also be produced live and must follow the same path as the content. The reason joining up DVRs with live content production is easiest is because live metadata will also bypass playout and scheduling, therefore there is no need to update those systems for this use case

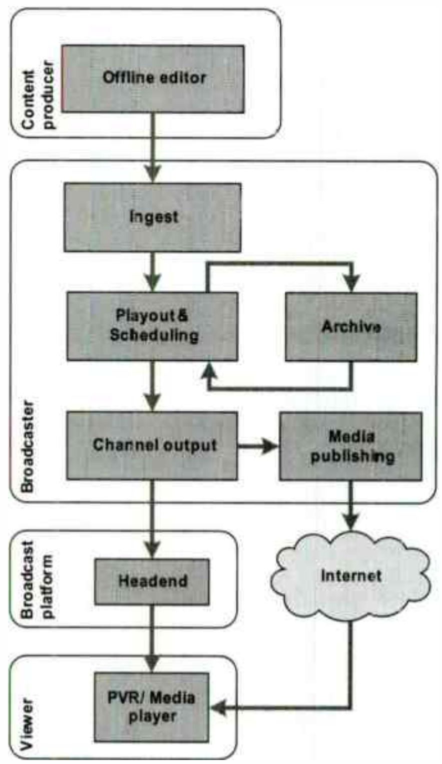


Content and micro-navigation metadata flows for live content

Different platforms have different capabilities for the audio and video: some will support HD, others SD; some will support 5.1 surround sound and others will not. Similarly, it is probable that different platforms support differing levels of micro-navigation functionality. This can be handled in each platform by filtering out unsupported information before adapting and formatting it for transmission. In spite of inevitable differences between platforms, it is essential to make the transfer of metadata from broadcaster to platform work; to achieve this efficiently a standard business-to-business protocol is needed.

Joining up DVRs with off-line content production

Broadcasting content which has been produced off-line is usually broadcast using automated systems running from tape robots or, increasingly, from video servers. Behind the actual scheduling and playout systems is often a set of complex systems including content procurement, ingest, post-production editing and archiving. If the micro-navigation data sourced at production is to be broadcast then it needs to be tightly bound to the content as it is ingested, archived, edited and scheduled in the broadcaster's content management system.



Content and micro-navigation flows for offline production.

Developing the metadata chain for off-line content requires equipment vendors to support the transfer of micro-navigation metadata with content. Whenever the content is manipulated, for example to modify the length of a film to fit in a slot in a broadcast schedule, the micro-navigation metadata will also need to be manipulated to maintain the correct time data for points of interest. The product used to do these edits will need to understand the metadata and be able to export a modified version.

JUP: ENGINEERING THE SOLUTION

As has been shown, many players must be involved in the process of ensuring movement and manipulation of the right metadata from content production to eventual consumption by the end viewer. Joined-up PVR (JuP) takes the role of bringing together these actors to develop, design and define the required interfaces and protocols to enable micro-navigation metadata to benefit end users. JuP is an industry group including equipment vendors, broadcasters and broadcast platform integrators.

The JuP group is committed to work collaboratively to develop understanding and to build the metadata chain. JuP's approach is to not "re-invent the wheel", there are many widely accepted standards relevant to JuP and part of the task of the group is to work with the standards bodies to create profiles applicable to JuP. There may also be a place for defining new standards if no standard is suitable for a particular link in the chain.

Standards which are relevant to JuP include:

TV-Anytime (www.tv-anytime.org)

The TV-Anytime forum have defined a number of specifications for DVR metadata which are published by ETSI (ETSI TS 102 822). Included in these specifications are data models for micro-navigation (or "segmentation") metadata. TV-Anytime is being widely adopted as a business-to-business format for schedule data and could also be used as a means to exchange micro-navigation metadata.

AAF (www.aafassociation.org)

Advanced Authoring Forma is defined by the AAF Association and is designed as a post-production interchange format, supporting external content links and effects such as fades, etc. Edit suites using AAF for interchange could export and import micro-navigation metadata.

MXF (www.mxf.info)

Media eXchange Format (MXF) is a SMPTE standard and is widely used for electronic exchange of content for storage, broadcast and playout. MXF defines a subset of the metadata defined by AAF. Content exchanged by MXF could contain micro-navigation metadata.

MOS (www.mosprotocol.com)

MOS stands for Media Object Server and is a protocol developed by the MOS group. MOS is widely used for news room automation. It is increasingly also being used for sports production. MOS systems can act as a source of micro-navigation metadata.

BENEFITS OF JOINING UP THE METADATA CHAIN

Adding value using existing content production systems

Micro-navigation adds value to content and can be implemented within today's workflows in content production and in broadcast content management.

Existing content production and content management systems represent a considerable investment by broadcasters and content production companies. By extending systems to extract and exchange information which is already present, content can be produced with the micro-navigation metadata.

Adding JuP support to existing systems can in most places be completely transparent to the operator. For example, content ingest and archiving systems will need to support metadata but operators should not need to take any actions to make this happen. If content is adapted in any way prior to broadcast then an edit suite which supports JuP would handle modifying the micro-navigation seamlessly, although some human intervention may be required to check that the points of interest are still valid after editing.

Supporting new content delivery networks

Building the metadata chain for micro-navigation benefits the new media player devices as well as DVRs.

In this age of a growing number of competing distribution channels, media players will increasingly support micro-navigation as standard. There are important media player platforms which include this functionality today.

Generating exciting new possibilities

Micro-navigation is the application that will join-up content production with DVRs and media player devices by building the metadata chain. Once this chain has been established it becomes easy to add other functionality.

Micro-navigation data can become richer, offering a wider range of uses. Content can be broadcast with segmentation data which gives alternative versions of the same content: adapting it for age, giving highlights of sports games, etc. Advertisement links associated with particular scenes within the content become possible, giving the viewer the opportunity to buy a product or service described during a programme. All this is made possible by joining up content production with DVRs.

CONCLUSION

Micro-navigation is a key way of enriching both the functionality of DVRs and the content offered by providers. For micro-navigation, the right metadata must be provided, manipulated and delivered all the way to the DVR, and for this to be achieved many organisations must be brought together. This is the role Joined-up PVR fulfils.

A2/M2 Digital Audience Measurement

Scott L. Brown
Nielsen Media Research, Inc.
Oldsmar, Florida

ABSTRACT

It has been said that there have been more changes in television in the past five years than there were evident in the previous fifty. With the changing face of television comes significant change in audience measurement, methodologies and research techniques. Nielsen Media Research is embracing these challenges by launching its most comprehensive and inclusive new data collection systems for collection of viewing, crediting, and reporting services in its history. From time shifting and place shifting to mobile and Internet delivery of television content, Nielsen Media Research will continue to provide broadcasters, networks, advertisers and agencies this "must have" audience information to support advertising, and the exciting business of television.

THE CHANGING FACE OF TV

Within the last 5 years, the face of television distribution and measurement has changed in very profound ways ... and as a direct consequence of the digital television revolution. Nielsen has found itself directly in the crosshairs of change and has faced the daunting task of measuring emerging HDTV, digital multi-cast, DVR-based timeshifting and most recently Video on Demand. Such new challenges were not unexpected as Nielsen began preparing for the digital television age back in the mid-nineties, through development of a new measurement system called the Active/Passive metering system (A/P). The system engages television content providers in applying discreet audio and video codes to the content via real time encoding devices.. The audio codes are especially key in that they are psychoaccoustically and perceptually masked, and therefore inaudible. The audio codes are particularly well suited for digital implementations and survive the rigors of digital compression systems of today. Once an A/P meter is installed within a sample home at a television site, a TV that is turned on will be detected and the source of the programming identified by virtue of either the audio or video codes or through audio signatures. The audio signature capability is primarily a backup methodology in that codes are primary, however, in cases where codes for one reason or another do not survive, Nielsen has invested heavily in an advanced signature or "fingerprinting" technology. The A/P meter will capture audio signatures based upon the audio available at the television set. At the same time, Nielsen has advanced market level monitoring equipment in every DMA to monitor all off air stations, cable networks via satellite, syndication etc., and to create a reference library of all known audio signatures. The subsequent matching process provides the identity of signatures captured within the home and provide a marvelous safety net. But is that all?

AUDIENCE RESEARCH TODAY

No one could have anticipated the relatively slow role out of the DTV broadcasting in the late 90's, and the slow consumer uptake through 2004. Beginning in 2005, the acceleration of digital television really advanced coincidentally with DTV price reductions and with the growth of Digital Video Recorders. Much of this growth was driven by DBS offering HTDV signals, the union of DVR with DBS receivers and with cable making the move to HD and DVR integration. Today DVR penetration hovers at about 15% of US TV homes and is likely to see more growth spurts as CE manufacturers pursue the dream of the connected home. VOD has followed the path of digital cable and is available to 25 million homes or more. As these aforementioned technologies became increasingly meaningful, Nielsen initiated DVR measurement in 2006 and then followed with VOD measurement in late 2006. This coincides with aggressive moves by content providers to license content to new platforms such as the iPod, and to rapidly make certain content available on Web sites. Perhaps the most significant move was to license content to both commercial and distinctly non commercial venues; a major break with tradition. Mid 2006 however was an industry tipping point in that proliferation of wireless, handheld, and portable video devices finally gained acceptance as platforms befitting additional content deals. At that precise moment, Nielsen polled clients directly to assess and survey the needs for measurement going forward. The questions were direct; what platforms make good sense to measure? What are the important measurement issues as the digital television age unfolds? What do clients most need to know about unfolding viewing patterns?

WHERE WE'RE GOING

The answers were clear:

(1) be certain in -home measurement is all that it can be,

(2) measure the broadband consumption of TV content as the next great outlet,
(3) begin measuring out of home ... bars, work place, other places beyond the home where viewing occurs,
(4) measure the multimedia cell phones ... the 3G and 4G phones of today, and
(5) measure the iPod and its brother and sisters.

All television consumption is key and vital to media players dealing with the effects of viewing fragmentation and diversification of platform delivery. Every rating point, every 10th or 100th is vital in a world of thousands of programming choices .... and perhaps most of all, gaining directional insight into emerging consumer behaviors is key. The who, what, why, and where of television consumption.

And so launched the most ambitious program, from Nielsen, ever fielded. It is known as the Anytime Anywhere Media Measurement plan, or to most, the A2/M2 effort. It commenced recently with fresh client feedback in hand and with the creation of special Nielsen project teams to embrace some newly developed measurement architecture that in many ways are "derivatives" from the original A/P metering system. Put differently, Nielsen has developed several new meters designed for new environments that leverage existing audio and video encoding, or makes use of the audio signaturing capability which has been refined further still, or lastly, that use either software engines or off the shelf technology for capturing other identifying characteristics. Sound interesting? It is, but the new meters were actually proceeded by the metering expansion of Local People Meter technology from markets 1-10 to now include a deployment plan for markets 11-25 … thereby embodying measurement of 75% of commercial inventory.

LOCAL PEOPLEMETER EXPANSION

Nielsen's first component of the A2/M2 plan was to dramatically expand electronic demographic collection capability to the top 25 markets. This is a means of improving existing measurement but also helps address some of the negative effects of viewer fragmentation. Viewer fragmentation brought by simple channel expansion has resulted in turbulent impact to an industry accustomed to large share of viewing results. With more viewer choices, audiences to individual shows have dropped, even though overall total usage continues to grow. Through placement of peoplemeter devices in the top 25 markets, Nielsen improves measurement by providing larger samples, overnight demographics and daily audience estimates. This enables our media clients to better understand viewing patterns, emerging behaviors, amid a continuous flow key data metrics. It should be noted that LPM expansion actually replaces the previous system comprised of household meters, diary sweeps for demographics, and this is quite an industry advancement. Lastly, that buildup of local people meter homes actually enlarges the national reporting sample which grows effectively from 5,000 daily reporting homes with peoplemeter to a total of 21,000 reporting homes.

Nielsen Peoplemeter

Hence the first major commitment by Nielsen via A2M2 is the expansion of our finest demographic metering technology to key markets 11-25. The expansion of peoplemeter technology benefits clients by enabling much finer targeting of demographics and tracking of smaller audiences to lesser viewed shows; a key issue in the age of viewer and technology driven fragmentation.

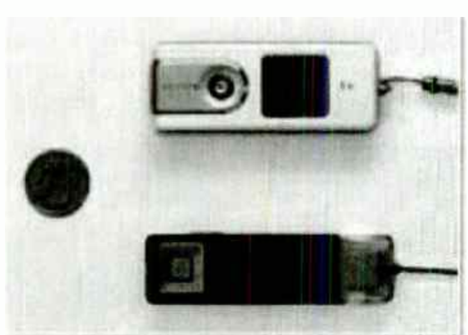
ANYTIME, ANYWHERE / MEDIA MEASUREMENT (A2/M2)

Of the A2/M2 efforts, perhaps none is more important at the moment than the effort focused on measuring Internet television. With all the major broadcasters and many cable networks using the web as an outlet, clients are highly interested in viewing levels and what can be learned about non-linear consumption. Of course a web log can give some idea, but the all important demographics of who is watching and the pre/post history of tuning usage is really key to understanding this new phenomena. Nielsen NetRatings focuses on Internet content measurement via a large opt-in panel.

Today, we have a dedicated team working diligently to bring Internet TV measurement to the Nielsen audience estimates. To be certain, there are unresolved questions of how best to integrate the Internet–based television experience with peoplemeter based audience traditional television estimates. The target is for market trials later this year to confront and seek client input on how best to integrate such usage with present in home measurement. This brings a unique concept to fore. How to measure many new technologies where the exact same metering system is not possible across all platforms. Take a cell phone, how would a meter and possible wiring look to the user? Not serious of course and so Nielsen has engaged in efforts to initialize where possible, software engines which might reside on small handheld appliances such as cell phones and provide measurement data. There are many nested business challenges to go with that. Then there are new techniques whereby log files might be gleaned from such devices to reproduce the viewing events.

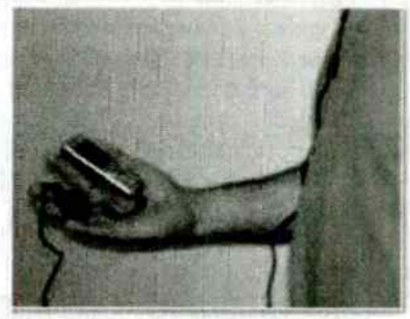
New Metering Techniques

Nielsen has initiated development of several new meters which build upon the existing strategy of audio codes and audio signatures. After all, this encoding is in place across the commercial television industry today and the signature reference sites number 600+ market sites spanning analog and digital transmission. With the codes and signatures actively collected 24x7, it was highly appropriate to develop new meters specifically targeted to new personal media consumption platforms. The first is referred to as the "Solo Meter", which is designed for measurement of personal media devices and players.



The Go Meters identify viewing by collecting audio signatures (MP3-based and cell phone-based).

This device is "off the shelf" and utilizes a microphone which is quite small and adept at capturing audio signatures of possible television viewing events. Recall from earlier that Nielsen has an enormous signature reference library with which to match off programming and viewing events. As the solo meter is active and carried by a cooperating sample member, the device can capture audio signatures of televised events which can be subsequently compared to the audio signature reference library for identification. Solo Meter is also augmented by another meter called the "Go Meter" for out of home viewing events in bars, restaurants, and other locations normally considered very difficult to measure. Much like the Solo Meter, this meter requires extensive testing and validation along with field trials to prove the efficacy of measurement. But the point is that audio signatures can again be very useful for specific applications.



Developing "Solo Meters" that will be able to measure video on portable media devices such as MP3 players and cell phones.

Clearly A2/M2 is enormous in scope and investment, but it doesn't stop there. Nielsen is simultaneously evaluating digital set top box data in concert with the MSO's to determine if that data represents a resource with potential. At the moment, there is actually very little of such data available as the operators have not largely deployed collection software with two way data extraction capabilities. That will change. Deployment of set top box collection is inevitable and with it comes the opportunity for Nielsen to engage in processing non-Nielsen originated data. Of course the data must be cleansed, stored and put to some practical use. That could involve possible integration with Nielsen sample-based data because the potential of set top data is found in sheer volume potential as well as in click-stream data. Nielsen finds itself right in the middle of the set top box discussion and had formed a new business service to provide for the opportunity and the possible merging of that data with other Nielsen data metrics. Nielsen is already at work processing trial data for many VOD applications with the MSO's. That form of processing is considered transactional VOD server-based data which presents its own data value equation. Indeed, the sources of data in the digital age are diverse.

Nielsen is also investing in the interactive space as an area that is likely to finally emerge. Remember that purchase that required some limited configuration, but in the absence of those, we'll go out and pay someone a **lot** of money to assemble technologies for us in ways that might be new, untested and extreme. In the computing world, SGI, Cray, and more recently a host of cluster companies have been trying to solve these problems and then drag the technology (kicking and screaming) back down into the COTS world. The Holy Grail being to drag them all the way back into the retail realm (If you worked at a national lab or a university, that would be an out of the box cluster that magically runs LINPAK and inserts itself in the top five hundred list at the touch of a button!)

Retail, COTS and Custom Technology

Items that truly reach the status of retail explode.

Customers do not want to buy custom technology. They do not want to buy COTS technology. They want to make retail purchases of items that they understand, intuitively know how to use, and are comfortable comparing. This is really the heart of "retail."

What constitutes a retail purchase varies for each of us. Some people are quite comfortable purchasing a motherboard, power-supply and various other components. They know how to compare, they know what parameters are important, and they end their buying session feeling good about the purchase. Others buy from DELL or Apple because they would not feel comfortable making these component purchases. It would take too much time, involve too much risk, or be too much hassle.

Storage is no different. There are numerous technologies out there, all of which have advantages and disadvantages. Customers want to share more. They want redundancy. They want security. They want scalability. And they want all this to be intuitive. It should be automatic.

iSCSI, AoE (ATA over Ethernet) and Fibre Channel

To make a storage purchase, a customer has to consider too many elements right now.

1. How can I share this storage? (Clustered file system, volume sharing, sneaker-net?)
2. How will my users interact with this storage? (Will it be seamless, or will they have to use some special GUI or interface?)
3. What sort of adapters will I need to purchase to connect this storage? Do they exist for all of my platforms? Will they work together?
4. What sort of cabling will I need to run to lash all this together?
5. What infrastructure will I need to buy to meet my requirements? (Security, remote access, flexibility, scalability)
6. What types of storage options exist for this choice? What features do they offer? (Redundancy, performance, administration)

The majority of the customers Small Tree talks to approach us because they fear Fibre Channel. They have never used it and they don't see Fibre Channel switches at Best Buy.

However, Ethernet is something they are quite comfortable with. They buy their own switches (even if they aren't good ones). They can browse, test, try, upgrade and eventually get to where they want to be.

This gives iSCSI and AoE a huge advantage over Fibre Channel. In places where customers have not deployed either one, they would much rather deploy something they understand vs. something they have never seen before. They would much rather upgrade their entire Ethernet infrastructure with a "better" switch than adding a second type of switch to the equation.

Let's look quickly at these elements and consider what customers think about them and more importantly, how vendors are marketing these elements

How can I share this stuff?

Customers are clearly unhappy with their clustered file system options today. They have proven to be complicated, expensive, and unreliable. They do not scale particularly well.

As the CTO of a company that writes Mac drivers for networking and storage, I have spent considerable time talking to customers and iSCSI vendors to understand where things are converging.

Dynamic, real-time sharing (whether through NAS or clustered file system SAN) isn't cutting it. Once share groups get larger than one server can deal with, choices become few and far between. There's nothing in the retail space at all. Customers that have yet to deploy are still in "wait and see" mode.

iSCSI and AoE storage vendors avoid the problem entirely. Their products are marketed to attach to application servers. Performance and configuration presentations deal purely with a single system connection.

Interaction

I once heard a saying that the best interface to a product is not to have one. The product just "blends in."

The product should seamlessly drop in and any changes or configuration should be intuitive and automatic and built right into how customers interact with the product.

Direct attach storage options like Firewire and USB have all reached a plateau here. Customers intuitively understand how to connect them and use them.

Interfaces

Very few systems have built in Fibre Channel interfaces. Some have built in fire wire, some don't.

All systems have built in Ethernet interfaces. If these systems have free IO slots, they all have available Gigabit Ethernet adapters. All of these adapters have been in use for long periods across multiple switch brands from Cisco, to Foundry, to Linksys and Dell.
Of all the potential interfaces, Ethernet is definitely "retail."

The main concern here is performance and we should talk about that.

Fibre Channel implements SCSI over a proprietary network interface. They do this because Ethernet and TCP are not efficient enough for the large IO sizes preferred by block storage solutions. This is not the first time this has been tried.

Years ago, FDDI looked much better than Ethernet. It had a larger block size. It degraded more uniformly (no collision storms to worry about) and it was faster. It ran long distances over fiber and offered redundant ring topologies. I can recall vividly my boss showing me the network hardware and telling me that Ethernet was dead. FDDI and other token ring topologies would eventually take over.

TCP is still a problem and vendors have been trying to deal with it in several ways. In the Windows/Linux environment, there are TCP Offload engines like those from Alacritech, Chelsio and Netxen. However their primary feature (processing TCP so the CPU doesn't have to) becomes less valuable as CPU speeds continue to climb and the number of available cores continues to go up. As things stand, the OS vendors aren't making adequate use of those extra cores now. So why not simply call one of those extra cores a "TCP offload engine" that's built right into the system and be done with it?

Another important element that FC has over Ethernet is 0 copy. Ethernet still has the problem of the CPUS needing to memcopy each word of data from the users space into kernel space so it can be kept in case of retransmit requests.

There's another irony in this in that sockets were originally set up such that I could "fire and forget" my data. I write it into the socket and it takes my data and returns control to me immediately (presumably so I can do other things). So what do customer apps do? They immediately write over the buffer of data! This forces the stack into the memcopy mode, even in cases where the OS tries to avoid it (SGI had page flipping implemented for many years and it never did much good for normal apps).

TCP offload cards avoid this problem by putting hooks into the stack. They can try to pull the pages in directly. However an implementation like this requires intimate knowledge of the stack as well as serious modifications. This puts them in the unenviable position of spinning a release each time the vendor changes something in the stack. It makes it difficult, if not impossible for developers to keep up.

Intel is taking a serious whack at this problem with their latest chipset. They have a feature (built into every new Mac Pro and Intel Xserve) called Quick-data. This is basically a built in DMA engine that can offload the user-kernel space DMA. It's not particularly useful in our opinion for "normal" tcp networking. 1500 byte frames just aren't large enough to make the DMA setup worth the effort.

However, when one looks at iSCSI, where jumbo frames would be the norm and internal buffers can be 4k, this feature starts to make a lot more sense.

Another option here is to avoid the TCP stack entirely. AoE (ATA over Ethernet) uses this option. They simply use a layer 2 protocol. The benefit is that packets can be written directly to the network without using the tcp/sockets interface. This avoids a great deal of the traditional overhead, makes for a much simpler driver model, and provides excellent efficiency over the wire.

The only serious drawback to this model is that the packets aren't routable. So any storage must be sitting on the same subnet as the clients. I argue this is hardly a liability given that very few iSCSI products offer the security or performance required to allow for routed disk traffic.

Cabling choices

Setting aside your own building for a second, I suggest you pay close attention next time you are in any other building. What cabling is in the walls?

Take a look at any commercial building and you'll find they are probably plumbed with cat5. We've visited numerous video/audio/printing customers. In each case, the buildings were plumbed with cat5.

This doesn't on its face mean that iSCSI/Ethernet gets a free ride. In most cases, there wasn't enough Ethernet to easily deploy a storage solution. We would have to at least double the number of wires in the wall. However finding someone to do this is a snap. Selecting the wire is a snap. Putting the wall panels and concentrators in is a snap. Not so with optical.

Cat5 is the retail interconnect bar none and as interesting as multimode fiber is, customers don't clearly understand it and unless they are in a serious COTS/technology buying mode, they are not interested in buying it.

Infrastructure

Assuming a customer does want to connect storage to some sort of shared network, such that potentially sensitive block data is moving around on the wire for all to snoop, how are they going to do it? How are they going to put the storage in the data center and share it to the office and even workers' home offices without creating a security nightmare?

There's only one interconnect that has been solving these problems since the early 90's: Ethernet.

Ethernet has the encryption, bridging hardware, routing hardware and large scale switching hardware to do just about anything. Not so Fibre Channel.

RAID Features

Companies like Left Hand Networks and Equallogic have created extremely scalable, redundant and performant iSCSI RAIDS that are quite capable of displacing the likes of EMC as a storage base for an enterprise customer. Additionally, since these RAIDS connect via Ethernet, they offer the ability to work directly with laptops and other Fibre Channel challenged platforms.

Ready for Prime Time?

So is the Storage over Ethernet Market ready for Prime Time? I argue that it's closer, but still a ways from displacing Fibre Channel. While iSCSI vendors offer feature rich products that have numerous storage management functions, they still do not offer the performance and security necessary to move into the "retail" space.

TOE cards necessary to provide good performance from client to server are still difficult for developers to integrate with rapidly changing operating systems. This leaves lots of opportunity for customers to hit "gaps" where a particular adapter, application and operating system don't overlap such that they can be deployed!

AOE vendors are much closer from a performance standpoint, however they are still adding the more powerful storage management features that customers demand (like live snapshots and dynamic growth of volumes).

Now more than ever, there is a lot of investment capital focused on the iSCSI performance problem. I expect that in the next year, many of these problems will be solved.

My expectation is that towards the end of 2007, all of these technologies (AoE, iSCSI and TOE/offload products) will start to show up at your local big box store sitting on a retail store shelf. They may be "lower end" products, but they will be the harbinger that the top end products are finally reaching maturity. If you don't see these technologies dribbling down into common retail shelves by then, that will be a sign a large amount of that investment capital is in trouble.

MEDIA STORAGE AND ASSOCIATED FILE-BASED WORKFLOWS

Stavros Hilaris, Bob Timpone, Bill Van Bloom, Josef Marc
Ascent Media Systems & Technology Services

ABSTRACT

Media storage has become a key element in designing and operating a broadcast facility, accelerated by the popularity of file-based workflows now that videotape is no longer the only practical way to store broadcasters' assets. There are a few basic categories of storage available today, including a trend toward media-aware storage. In order to make the best choices in media storage and gain the most advantages, workflow is the place to start. Workflow engines also come in a few basic categories, including a trend toward applet-style toolsets as opposed to traditionally-hardened enterprise automation. After media storage is installed and a workflow engine is built, future improvements in media storage and workflow are managed within the workflow rules and storage interfaces. In other words, rewiring is replaced by data entry. These general principles apply to many kinds of media operations. To address a wide NAB audience, this paper illustrates a typical Ascent Media customer building a multi-channel, multi-destination, broadcast origination facility with integrated post-production.

MEDIA STORAGE

The good news is that broadcast engineering teams have options in media storage products, with clearly distinguishable technical attributes. Unfortunately, that's also the bad news because those attributes may not be as obvious to the eye as a videotape's size and color, shelf location, and machine/human-readable label. Operations teams may become disappointed when media files aren't as easy to find, share and use as their spreadsheets and email, or worse – not as easy as the videotapes they were already using before the media storage appeared.

Nonetheless, with appropriate storage choices and workflow policies, experience shows that file-based workflows are faster, cheaper and more powerful than tape-based workflows for most broadcast applications. Proper training for the operations and creative staff is essential, of course, to achieve efficiencies. First let's review the essential media storage options a broadcaster can choose from.

Common Components

Media storage systems are built from a variety of underlying technologies originally invented for general IT purposes, each with their own pros and cons in any given media application. Below are some key examples used in media storage equipment.

FCAL (Fibre Channel) and **10GbE** (10 Gigabits per second Ethernet) are the main choices for networking media storage hardware. **SAN** (Storage Area Network) and **NAS** (Network Attached Storage), **Cluster** and **Grid** are the traditional architectural options for grouping large numbers of disks. **TCP/IP** is the dominant protocol for network messaging between devices, although both TCP and IP are implemented somewhat differently in a media environment than on the Internet. **NFS/SMB** (Sun), **CIFS** (Microsoft), and **Samba** (free SMB) are popular application-level protocols for finding, reading and writing files on shared storage. **JBOD** (Just a Bunch Of Disks), **SBOD** (Switched Bunch Of Disks), and **RAID** (Redundant Array of Independent/Inexpensive Disks) represent different layers of complexity for spreading a file's bits between different disks – or not. Parallel and Serial **SCSI** (Small Computer Serial Interface), **SATA** (Serial Advanced Technology Attachment), older **PATA** (Parallel ATA), **FireWire / IEEE 1394**, and **USB** are common technologies for connecting disks to their host computers.

Discrete Storage

By default or by history, media devices in a broadcast facility can each have their own media storage. A videocassette recorder/player (VCR) has its cassettes, and a playout video server has its disks. A non-linear editor (NLE) has its own disks, and so does a second NLE. A graphics station and a closed-caption inserter have their own disks, and maybe they share with the NLEs some companion VCRs with their own cassettes. Any of these devices can employ a file format usable by the others – or not.

We can call this a discrete storage environment because all storage is separate. Media files are moved by streaming or by duplication in some form that can be shared. This can be called a House Format, a term that can apply equally to streaming through a router as to file-replication on tape or disk.

A discrete storage environment tends to employ a lot of format conversion (transcoding) which consumes time, transcoders, and quality assurance people and equipment. When a discrete storage environment grows linearly (adds a new device), transcoding tends to expand geometrically. By the way, transcoders have their own disks too.

Single-Manufacturer Shared Storage

Many broadcasters' early experiences with file-based workflow are associated with shared media storage from two popular NLE manufacturers, AVID's Unity and Apple's xSAN (Final Cut Studio). We can call this single-manufacturer shared storage. Producers and editors in this kind of environment can "edit-in-place," meaning they can share media files without having to wait for House Format transcoding and/or duplication.

Single-manufacturer shared storage saves time and money (tapes, VCRs, transcoders). It also enables some workflows which aren't feasible in a discrete storage environment, such as starting closed captioning before a final cut is published.

But wherever a broadcaster needs to apply equipment or a workflow that is not supported by that single manufacturer, they revert to discrete storage solutions. Most broadcast facilities fit this description.

Figure 1 illustrates a common hybrid scenario where tapes and VCRs feed a playout server with discrete storage, other VCRs feed NLEs with single-manufacturer shared storage, and more VCRs feed ingest and transcoding workstations for mobile TV, broadband TV and proxy servers, each with their own discrete storage.

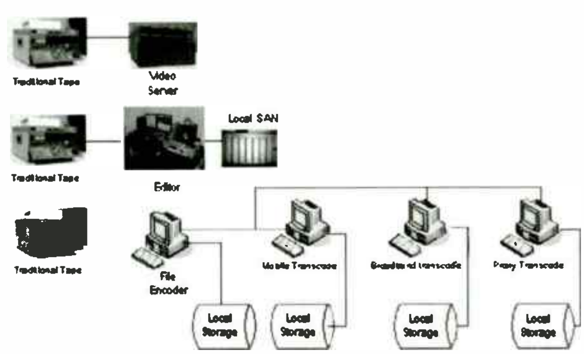


Figure 1. A common hybrid of discrete and shared storage

Video Archive Systems

Videotape libraries are often supplemented by robotic data-tape solutions adapted from the general IT universe. These can be considered part of a Hierarchical Storage Management (HSM) solution because of their place in the implied hierarchy of on-line, near-line and off-line storage. They can also be thought of as single-manufacturer shared storage optimized for archiving, although multiple manufacturers tend to be required for a complete solution.

Video archive systems introduce a paradigm shift in how broadcasters view their media assets. Like discrete storage environments, transcoding and House Formats are the keys to archive storage and retrieval. Like single-manufacturer shared storage, archive systems save time and money. Unlike single-manufacturer shared storage, however, video archives must preserve and recall media assets produced on equipment from nearly any manufacturer, including single-manufacturer shared storage.

This requirement may introduce a broadcaster to new layers of software integration between network file systems, computer operating systems, metadata cross-walking between databases, tape library management applications, media asset management (MAM) applications, air-compliance recording applications, browsers, automated video testing applications, Keyboard-Video-Monitor (KVM) switchers etc. Products like ADIC's StorNext and Front Porch Digital's DIVArchive may be popular because of how well they serve multiple computer operating systems and applications during the integration process.

Media-Aware Centralized Storage

Many drawbacks mentioned above are addressed in the new "media-aware" shared storage solutions from manufacturers like Omneon and AVID. In essence, they combine storage, networking, transcoding, high availability and expansion in an integrated package optimized for media assets. This greatly simplifies all five functions.

Omneon's MediaGrid exploits the benefits of parallel computing. Each box of disks includes processors that can collaborate on tasks besides reading and writing files.

But like video archive systems, media-aware centralized storage can also introduce a broadcaster to new layers of software integration.

In large facilities we typically see a mix of discrete storage, single-manufacturer shared storage (usually a

SAN for a single workgroup), video archiving, and media-aware centralized storage. We call this a hybrid or tiered media storage environment.

Figure 2 illustrates how media-aware storage changes the landscape from the discrete-and-shared storage model in Figure 1.

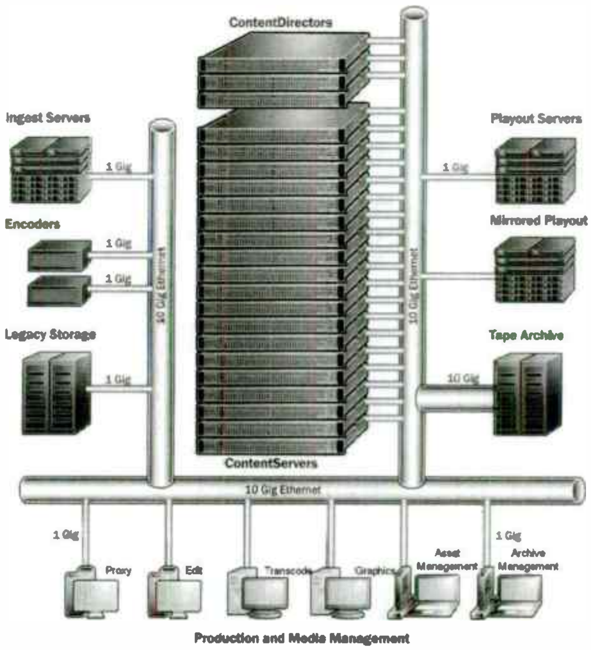


Figure 2. A typical application of media-aware storage (Omneon's MediaGrid)

WORKFLOWS

To help a broadcaster make their best choices from such a variety of media storage options, we recommend deriving their media storage solution from an analysis of their workflows.

A workflow is a solution of all the relevant answers to the question: "What does X need from Y to do its job?" X and Y can be people, applications (software), or equipment. The logical possibilities are:

1) What an application needs from an operator, to do its job.

2) What a machine needs from another machine, to do the first machine's job.

3) What an operator needs from an application, to do the operator's job.

4) What an operator needs from another operator, to do the first operator's job.

5) Exception-handling for all of the above.

Some workflows may only share metadata, such as associating promo IDs and ISCI numbers. Other workflows may only change the status of media assets, such as whether or not they have passed a quality check. Some workflows change the nature of media assets, such as embedding closed captioning data. Most workflows need to move media assets from one workplace to another: from pitcher to catcher, acquisition/ingest to quality assurance, satellite reception to production editing, graphics and production to master control, playout to compliance recording, editing to archive, archive to marketing, archive to editing etc.

Our workflow specialists examine all metadata and media movements in search of ways to make them work more efficiently, meaning more production for less cost. Efficiencies are typically found in the elimination of manual processes like tape duplication, they're found in increased speed because work doesn't have to wait for tapes, and they're found in increased throughput because workplaces handle more assets at one time when tapes don't have to be moved and loaded.

Workflows and Interfaces

The first stage in the process is called discovery: learning how things work now, with an emphasis on identifying what, when, where, why and how information is exchanged between business units, workgroups, and systems (equipment). We record this discovery in graphical and textual documents, often based on Business Process Modeling Notation techniques developed and maintained by the Object Management Group (a standards body analogous to SMPTE, see bpmn.org).

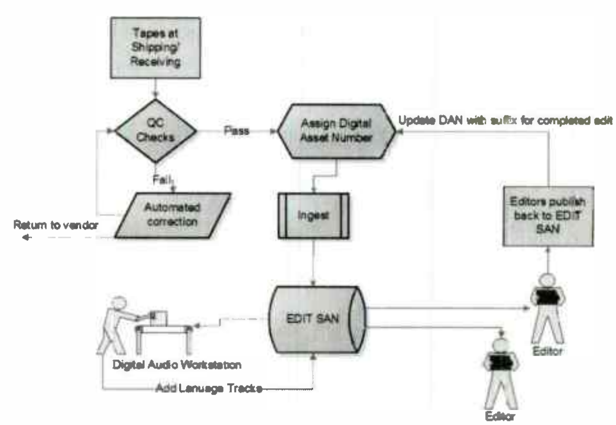


Figure 3. Workflow diagram for a fairly simple editing operation

The next step is to develop new workflows. Proposed workflows are drafted in diagrams and text, describing the rules, data schemas, and even XML documents, that would support an improvement in speed, cost and/or throughput. Many broadcasters are also looking at new media distribution opportunities such as mobile or Internet TV, requiring the invention of new and incremental workflows.

Proposed workflows are then evaluated in workshop meetings with user groups to gain their feedback and acceptance of new practices. We call these sessions Workflow Walkthroughs.

When proposed workflows are accepted and approved by management, we derive from those workflows a set of Interface Control Documents (ICDs) giving technical definition to the information that passes between business units, workgroups and equipment.

The third stage is to apply the approved workflows and ICDs to vendors' scope of work documents, operational training plans, interface and system acceptance test plans, and deployment (project) plans. We have found these workflows and ICDs to be extremely powerful tools in practice.

Workflow Engines

Although broadcasters can buy a single-manufacturer workflow engine, we use the term more generically to mean any toolset that automates or augments workflows. Broadcasters are familiar with station automation as one example of a single-manufacturer workflow engine. But file-based workflows extend beyond the reach of station automation.

Workflow engines aren't meant to work out-of-the-box, so to speak. They are software environments requiring setup and commissioning, which means that following the workflow analyses and ICDs, data and business rules are entered into the software for the broadcaster's particular operations and workflows. Most Media Asset Management (MAM) products include a workflow module.

A file-based workflow's primary objective is efficiency but it can also affect how the broadcaster approaches new technologies and business opportunities. This is because changes to the business tend to be achieved by changes in the workflow engine's rules and/or ICDs, allowing choices about media storage equipment to be made separately.

In other words, media storage can be expanded or swapped without changing the workflow engine's rules, just change the ICD. Or conversely, change a workflow engine's rules to redirect media assets to different media storage, leaving the ICDs and equipment unchanged.

Workflow engines are typically optimized for a certain type of media operation: post-production, local station playout, newsgathering, channel origination, wide area content distribution such as VOD, event recording etc. Even though the underlying media storage architectures may be similar, they differ widely in their toolsets.

Workflow engines optimized for discrete storage environments will have more tools for transcoding and duplication while others, targeting environments with more shared storage, will have more tools for logging, proxy viewing, desktop rough-cut editing, approvals, playout scheduling etc.

MEDIA STORAGE AND WORKFLOWS IN PRACTICE

The above generalities apply to many kinds of media operations, including many kinds of broadcasters. The example below illustrates a typical Ascent Media customer building a multi-channel, multi-destination, broadcast origination facility with integrated post-production.

We want to stress the importance of two principles at the outset:

1) Media storage components and architecture should be derived from workflows. The choices illustrated here are well-suited to this operation; applying them elsewhere depends on "elsewhere's" workflows.

2) In a file-based environment, you can modify storage, workflows and software independently. The storage, workflows and software tools illustrated here are meant to be improved, whenever and wherever practical and beneficial – some maybe even before the publication of this paper.

From Tape To Air

One basic and key function of this facility is playing out long-form programs to air. The workflow diagram in Figure 4 illustrates the steps involved.

In this case, the program elements are in the tape library. They may be old or new, previously aired or not, captioned or not. At the end of this workflow, the desired program has been assembled, checked for readiness and quality, and saved to a playout server. A copy is also placed in the archive.

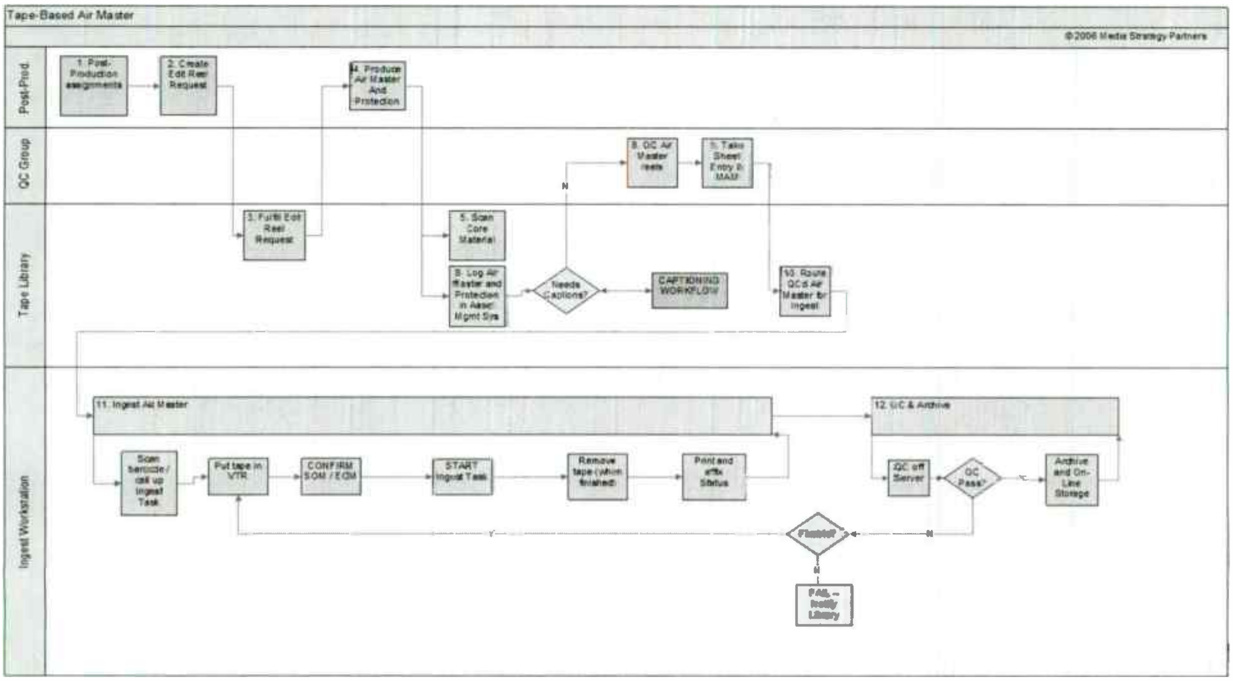


Figure 4. Workflow diagram to produce a tape-based air master

Note that in the center of this diagram there is a diamond-shaped text box: "Needs Captions?" This leads to "Captioning Workflow" which refers to a separate workflow diagram. This is not only a convenience for making the diagrams manageable, it is also an example of how workflows can be developed independently, and can change independently later on.

This workflow does not define nor depend on any particular media storage, although media storage is implied throughout. See, for example, the text box at the lower right corner: "Archive and On Line Storage." Whether this broadcaster uses a discrete storage environment, single-manufacturer shared storage, or media-aware storage, this highly repetitive workflow remains the same and can therefore be largely automated in a workflow engine.

Each of the boxes and arrows in Figure 4 are also described in a companion text, called a Workflow (capitalized to indicate a document, as opposed to the generic term). Each Workflow is numbered, catalogued and controlled much like technical system drawings and wire run lists. The Workflow contains many notations such as the names of the workstations and software applications to be used, locations and addresses for media movements, names and style sheets for filling in forms and database fields, reporting and escalation protocols, mnemonic naming conventions (or references to their online repository), a revision history, references to precedent and subsequent Workflows, and other tips and tricks to get the job done.

Our media storage design engineer uses this diagram and Workflow to identify the necessary interfaces (ICDs). After some triangulation arithmetic to factor in the facility's number of program hours going to air, the engineer can begin to define the sizes of those ICDs (bandwidths) and the media storage (bytes). Figures 5 and 6 illustrate some typical results.

SD ingest	100 hours x 8 Mbps =	500 GB
HD ingest	40 hours x 50 Mbps =	1 TB
Proxy ingest	580 hours x 1 Mbps =	1 TB
Editing (IMX)	20 hours x 50 Mbps =	500 GB
Compositing	20 Hours x 50 Mbps=	500 GB
Graphics	=	500 GB
	Total:	**4 TB**

Figure 5. Sample storage requirements

SD ingest	8 Mbps x 4 channels =	32 Mbps
HD ingest	50 Mbps x 4 streams =	200 Mbps
Proxy ingest	1 Mbps x 20 clients =	20 Mbps
Editing (IMX)	50 Mbps x 4 stations =	200 Mbps
Compositing	50 Mbps x 1 station =	50 Mbps
Graphics	10 Mbps x 4 stations =	40 Mbps
		Total: 542 Mbps

Figure 6. Sample bandwidth requirements

INTERFACES, WORKFLOWS & MEDIA STORAGE

Taking information from the Workflows and bandwidth requirements, our engineers can then design a media storage solution. The speeds and protocols for each interface are defined in ICDs, and the storage requirements are defined in equipment specifications. Figure 7 illustrates the project interfaces in this typical Ascent customer facility, where each arrow represents an ICD.

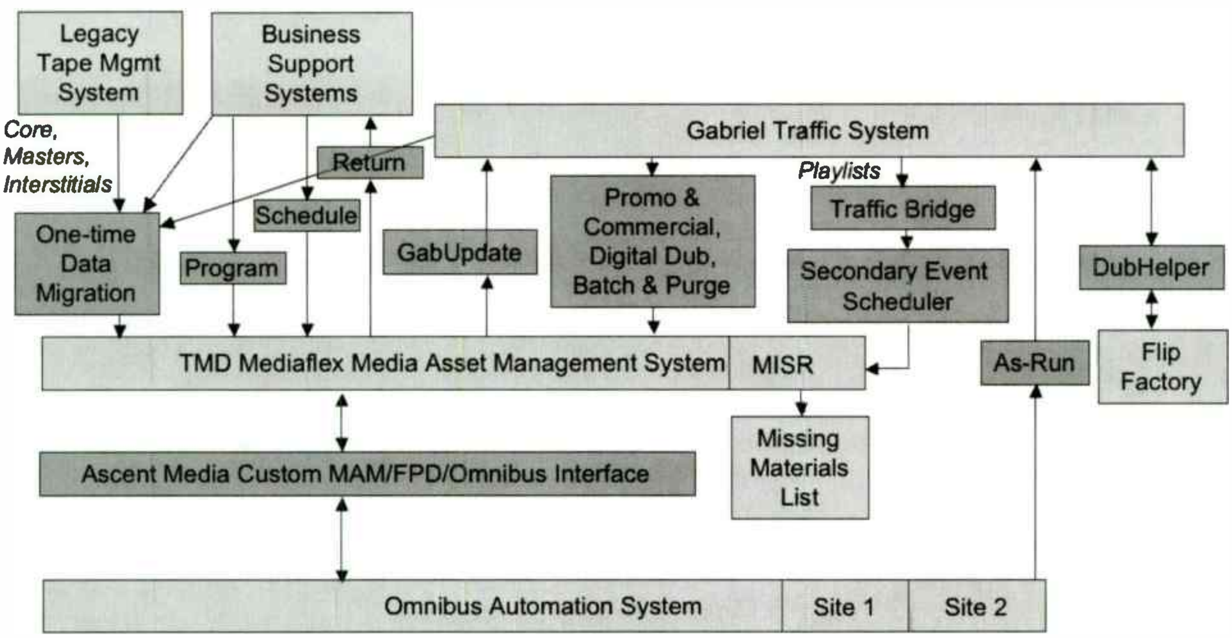


Figure 7. Interfaces

Figure 7 doesn't show all of the interfaces in this facility, it shows only those requiring an ICD and business rules in a workflow engine in order to achieve this transition to a file-based workflow.

Gabriel is the name of the traffic software, and Omnibus is the automation system. MediaFlex (MFX) is the name of the media asset management suite, and MISR is the module which reports missing media; both are by TransMedia Dynamics (TMD). Front Porch Digital (FPD) supplied mezzanine applications for managing the video archives. Marquis' Medway is a mezzanine application for managing the exchange of media and metadata with AVID systems.

In Figures 8, 9 and 10 we see how these Workflows are mapped to media storage and to applications (software).

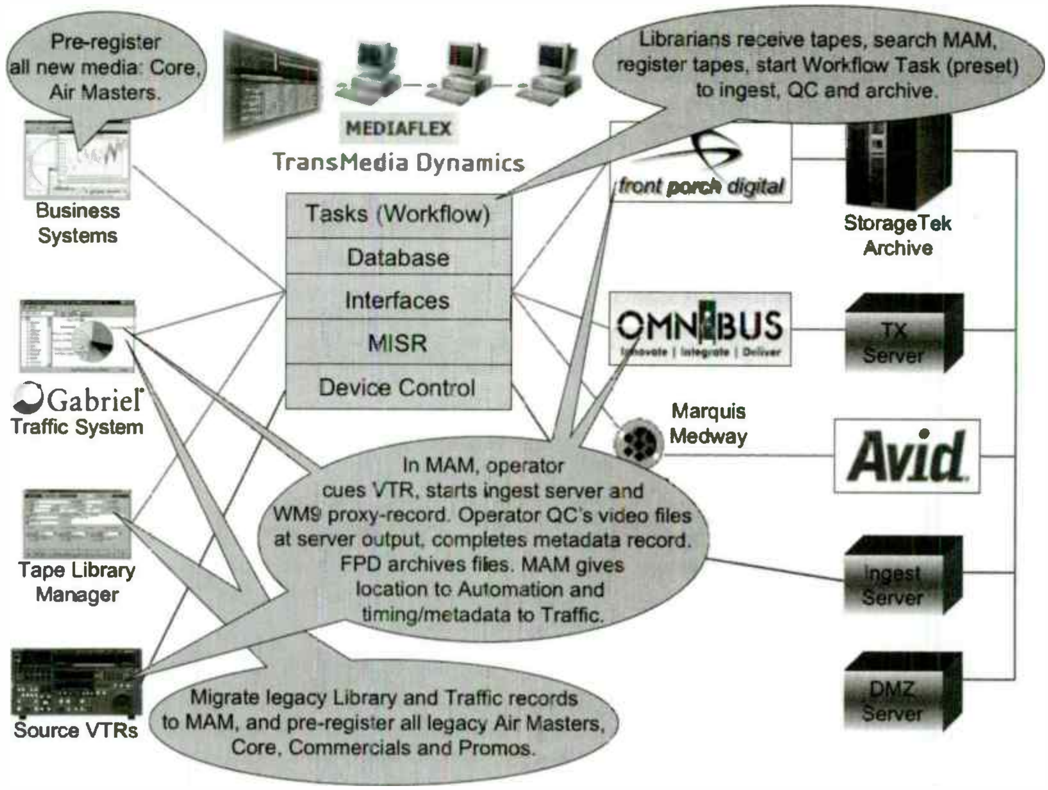


Figure 8. Long-form Program Workflows

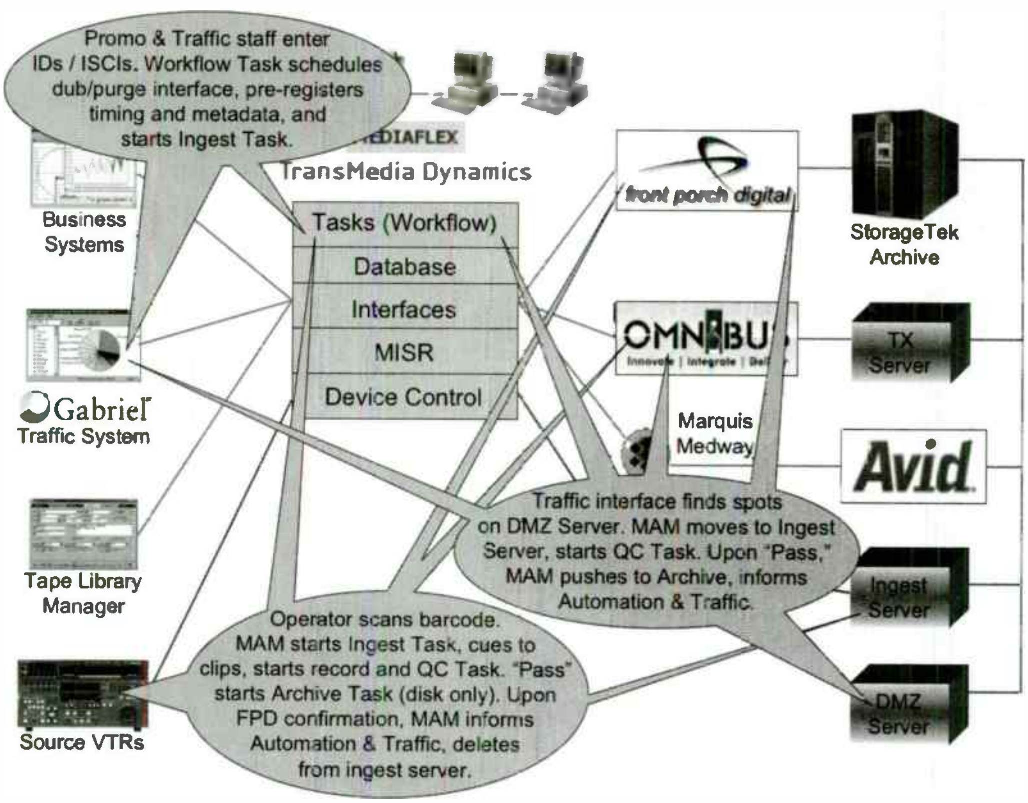


Figure 9. Commercial and Promo Workflows

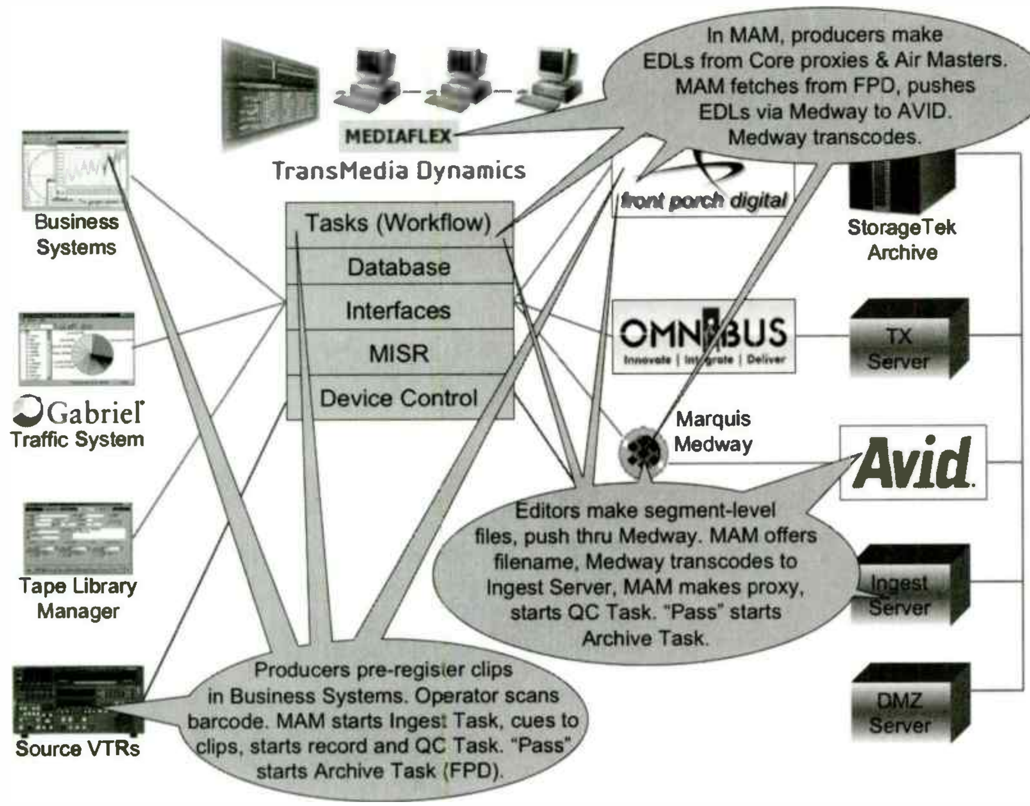


Figure 10. Post-Production Workflows (Editing)

Figure 8 is closely related to the diagram in Figure 4 and its associated Workflows. In Figures 9 and 10 we can see how two more Workflows are mapped over the same set of equipment, applications and ICDs.

Comparing the Workflows in Figures 8 and 9 we can see that the Commercial and Promo Workflows don't involve a librarian; the Long-Form Program Workflow does.

In Figure 9 is a reference to Archive "disk only." Such references are used when entering business rules to the workflow engine. We can call this one a "business rule" because it is a business choice, not a technical one; these commercials and promos could be saved to robotic data tape just by changing the business rule in the workflow engine.

Results

We use this combination of diagrams, Workflows and traditional broadcast engineering techniques to build and train file-based workflows. With a workflow engine holding the business rules, supported by ICDs holding the technical specifications, the system is faster and more powerful thanks to automation, yet flexible thanks to the separation of workflows from devices. And the overall enterprise can respond better to new business opportunities as well as improvements in media storage.

The Money Shot: Postproduction Workflows for Delivering HDTV Sports Content for Mobile TV

Michael Biremboim

CloseVU

Sdei Boker, ISRAEL

ABSTRACT

Viewers expect to see their sports in high-definition and more and more broadcasters are delivering the big file format size with stunning results. The challenge remains in re-purposing this sought after content for engaging and interactive viewing via mobile devices. Small mobile device screens are here to stay and broadcasters need to re-evaluate postproduction strategies and workflows for delivering compelling broadcast quality content to this new breed of sports viewers. This white paper discusses new postproduction technologies and workflows for re-purposing HDTV Sports content for Mobile devices.

1. Making the Money Shot - the new protocol for editors assembling sporting event highlights and fast action shots for small screen viewing. The white paper will take you through how it can be done and what tools can be used to achieve it.

2. Brand opportunities interactive mobile generation creates new avenues of revenue. Creative teams can turn to special real-time 3D motion graphics tools designed to easily create and transmit graphic overlays for titles and promotions for the small screen.

IMPROVING THE TELEVISION SPORTS EXPERIENCE

Mobile TV channels can capture audiences by providing unique - differentiated content. To do so, they will need to offer editorial enhancements that both inform and entertain sports fans. The demand is strong for feature-rich multimedia applications. By delivering targeted content and data, the new generation of portable devices enables viewers to personalize schedules, content, and presentation to maximize viewing enjoyment. These devices will allow viewers to receive sports news as an animation on top of a video stream or video-overlay ticker.

Mobile TV phones will come with the cutting edge graphic engines such as Nokia SVG Tiny or Adobe Flash Lite, allowing the synchronization of rich media content with the video streams. Mobile TV broadcast platforms such as DVB-H or MediaFLO allow broadcasting associated data along with the video channels. Graphics content will be rendered locally on the cellular phone or mobile device and adapted to each device resolution capabilities. Graphics will no longer be part of the compressed video signals as they are today on 3G video clips, for low quality graphic

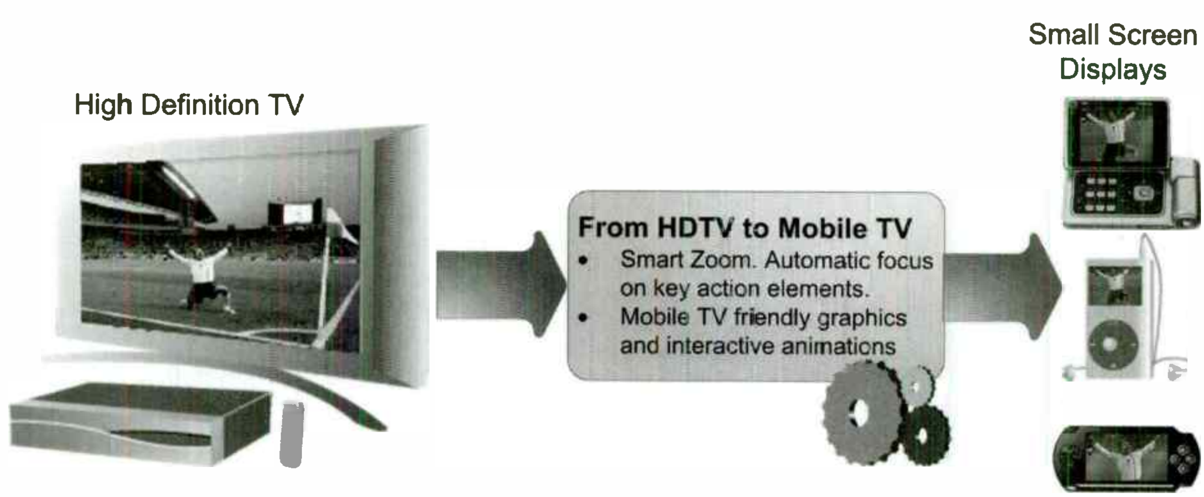


Figure 1: From HDTV to Mobile TV Diagram

capabilities makes any text overlay unreadable. Graphics therefore will now have a much higher quality.

MOBILE TV OPPORTUNITIES

Predictions of rapid growth for Mobile TV

Most consulting firms predict fast and dramatic growth for Mobile TV. McKinsey research expects that mobile television in Europe could be a €24 billion market and reach more than 190 Million users by 2025.
Datamonitor estimates the market for Mobile TV and Video for Mobile Devices growing exponentially in just the next few years – going from relatively non-existent to 69 million people worldwide by 2009, generating total revenue of around $5.5 billion.
A strong consortium of industry forces are stakeholders in the push toward Mobile TV over the next few years. Mobile operators are facing market saturation; their average revenues per user (ARPU) are under pressure. They see new revenue sources in mobile TV. Vodafone expects that by 2010, Mobile TV should represent 10% of their infotainment revenue. Manufacturers such as Nokia, Motorola, Samsung, and LG see Mobile TV as a way to sell more of their new and improved handsets. Television stations whose advertising revenues are challenged by new media see Mobile TV as a way to repurpose their content, extend their brands, and target on-the-go users at times when they are not watching their home televisions.

I want my Mobile TV ... my way!

Research shows that user expectations for Mobile TV are high in terms of content quality and presentation. Content is still king, but the look, feel, touch and usability are key considerations. Viewers want quick and easy access to unlimited content. Consumers expect a full, digital experience on their phones similar to those they get on the Web and other devices like iPods, and IPTV set-top boxes. Mobile TV will need services such as personal video recorders, "one click" downloads, podcasts and interactivity.
The interactivity capabilities of mobile TV are unique, thanks to rich media functionalities and the availability of a "one click" return path over the cellular network. This opens up hundreds of opportunities to create a call for action to the viewers (more info, voting, polling, gaming, e-commerce, etc.). Interactive services are essential to provide additional revenue to the Mobile Operator by generating data usage. In order to attract viewers, these interactive services should be contextual to the program and provide added value.
With a mobile phone, the viewer now has many more choices of when, what and where he wants to see his favorite programs.

CHALLENGES FOR TELEVISION STATIONS

Branding

Television Stations and Mobile Operators have both invested heavily to establish their brands. While entering a new business, channel or medium, branding and editorial control becomes essential in order to remain competitive. Television Stations and Mobile Operators will continue to provide high quality-content for their brands in Mobile TV in order to protect their established brand excellence and maintain as much visibility as possible. Television Stations and Mobile Operators will also need to ensure that the viewers' experience will be consistent across their media offerings, be it on the Internet, television, or now Mobile TV.
Brand recognition represents a strong competitive differentiator. Mobile TV channels will need to match high quality standards in terms of image quality and service in order to provide a multi-channel consistent customer experience. Otherwise, as happened for some on the web, they will mislead customer expectations which can be costly and difficult to recover from. Television stations will also have to insist that their Mobile TV channels maintain the highest quality standard in order to not jeopardize their brand image in other outlets.
Therefore, in order to produce the highest quality possible, television stations and mobile operators will have to adopt production tools dedicated specifically to mobile TV content production.

Appropriate effort and investment in the mobile TV business

For broadcasters and Mobile Operators, Mobile TV is a challenge. As mobile TV will definitely be a "hot topic" during the coming years, generating a lot of buzz and garnering attention; content broadcasters are eager to demonstrate their capabilities to deliver the best mobile TV offering. Investments will be made even for expensive capital such as dedicated production units or the creation of specific content. However, mobile TV will not be their core business and main revenue source. Moreover, business models remain unclear and market perspectives (mobile TV capable handsets diffusion, offer adoption rates) are still uncertain. Hence, expected

revenue is quite difficult to forecast now, making it sound to scale investments carefully.
Finding the right balance will be a tough job for mobile TV content players. However, since they must take some action in order to launch a mobile TV experience, they will certainly look for "simple" solutions with:

1. **Strong focus on mobile TV needs**, rather than other solutions derived from gaming or rich-media environments to support video and TV;
2. **Lightweight, low-cost working environments**, in order to lower capital expenditures;
3. **Software based** solutions, operated on a

Capitalization on existing content capabilities, creating mobile TV programs at minimal cost, and leveraging on existing material (e.g. through a shared access to the Media Asset Manager) will be the true challenges for the success of Mobile TV content production.

From HDTV to Mobile TV

Today, more and more television shows are being shot in High Definition for viewing in large size screens. The small size screens of cell phones makes it hard for viewers to read text overlays on the screen such as names or sports scores, as well as seeing long distant shots which are typical in HD Television.

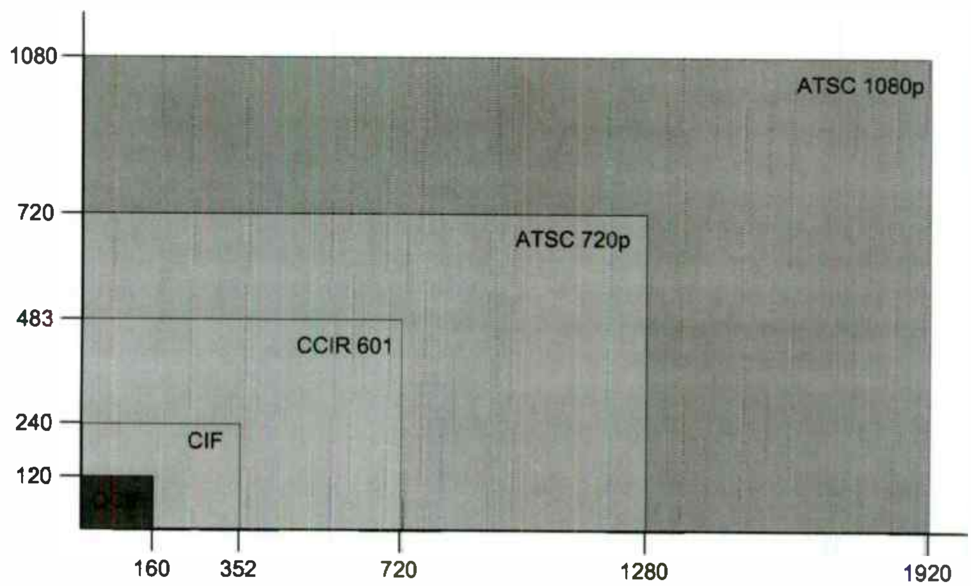


Figure 2: Screen display sizes

classic hardware (PC or Mac) and operating system configuration, which must be both easy to roll-out and easy to manage and support;
4. **Ease of use** by non-dedicated/non-specialist teams. The solution must require neither strong training investments nor expensive operations people (mobile TV tools experts) to run the system, ensuring minimized OPEX (operational expenditures) for the mobile TV activity; and
5. **Strong integration with existing infrastructure, architecture and systems.** The solution should not reinvent or duplicate basic pieces of the TV architecture, like the Media Asset Manager, but rather capitalize on it through smart integration.

Display Device	Display Sizes
Cell phone GPRS	QCIF
Cell phone Mobile TV	CIF 240 x 352
PDA/Smartphone	256 / 768 kbps CIF – VGA
Automotive / Portable / Home	VGA-D1 768 – 1 Mb/s
HD	1080x1920

Table 1: Screen display sizes

Great attention must be given to the viewing experience and programming needs of the small screen to make it user-friendly and attractive for the Mobile TV watcher. With a television in one's hands, studies show that customers tend to zap from one channel to the next more often than with their regular television. Due to

this short period of watching, television groups wish to create fast paced content to offer sports highlights.
It is very unlikely that television stations will hire expensive separate crews for shooting for small screens. They will look for tools to easily resize and fine-tune HD content to make it suitable for small screen viewing.
There is a strong need for a set of tools, some automatic for live sports broadcasts and others semi-automatic where content creators guided by some very easy-to-use controls manage the optimization of the image quality. These tools will be based on advanced object-tracking technologies that will identify the desired objects for image processing.

INTEGRATING MOBILE TV PRODUCTION INTO THE TELEVISION WORKFLOW

Content Repurposing

Material available at TV stations can be effectively repurposed into a Mobile TV channel. A small team of producers could have access to content files at the TV stations and integrate current television production into content suitable for mobile TV. File-based systems are being used across all aspects of the digital content workflow including acquisition (file-based cameras), production (video editing platforms), and distribution (file-based delivery services).
Content can pass more efficiently through the value chain and there is greater flexibility in how to deliver programs to the consumers. File-based production and content management capabilities allow for easy search, retrieval, movement and repurposing of digital media files.
In addition, any authorized contributor can push content to the mobile producers from any computer. Only a minimal software installation on the PC is needed. Viewing content in new forms and on new devices will require content to be reformatted specifically for each platform, a process that takes time and consumes significant processing resources. Ultimately, there will be an increase in value of the digital assets and an increase in revenues for the content producers.

Specificities of Mobile TV content

The small screens of cell phones makes it hard for viewers to read writing on the screen such as player names and sports scores, as well as seeing long distant shots. To solve this problem, Mobile television producers are being trained to re-edit current content to focus on 'close up' shots and increase the size of the text on the screen.

Mobile TV playlists generally consist of a TV program sequences between 5 to 20 minutes long which will be played on a loop and be updated continuously by a small team of editors.

Mobile TV sequences consist of video clips, audio and graphic animations that are synchronized with transitions, audio insertions and specific graphics to be rendered locally on the handset for better look and interactivity applications. Digital media files lend themselves to this process since they can be easily broken up and reassembled as needed.

Specific Tools for Mobile TV Production

Scenario 1: A TV producer is usually not a professional video editor, but more likely a journalist or a content producer, with no extensive technical editing skills. He needs an easy to use, simple and fast tool to repurpose content for new platforms, including mobile phones. The production tools required to generate workflow contain several steps as shown in figure 3.

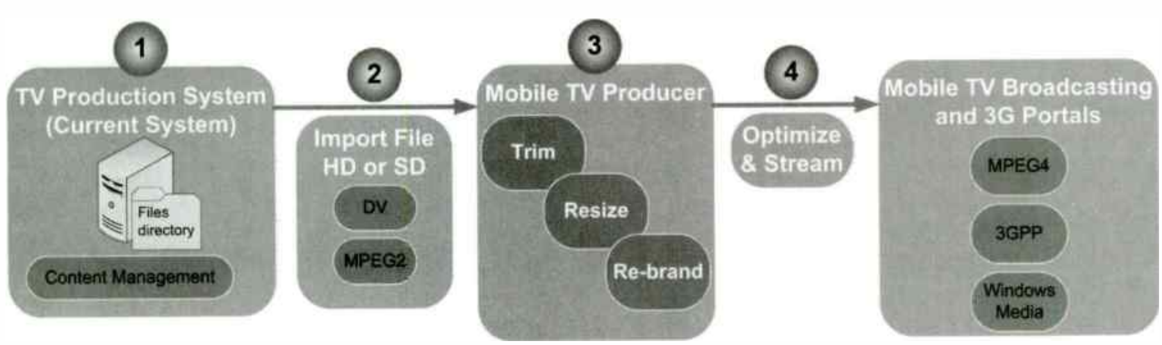


Figure 3 Mobile TV production tool as a stand alone application

- Step 1: The current production system used by the television station serves as a files directory, having the access of files through the network or local repository. The Mobile TV producer pulls the files from the system while considering the compatibility of files formats.
- Step 2: The production tool enables the importation of files from third party systems, using tools for conversion of HD and SD formats like DV or MPEG2, and various file wrappers.
- Step 3: The set production tool for Mobile TV should enable the Mobile TV producer to resize frames, meaning focus on the action through close-ups rather than large-angle shots. In addition, he can insert specific branding which is more suitable for the Mobile TV small screen.
- Step 4: To complete the Mobile TV product, the file needs to be exported and optimally compressed into Mobile TV formats so this file can be later on be streamed directly into mobile devices.

This approach offers several benefits:

- Simple and fast repurposing of content to a mobile TV product due to a dedicated tool with all the functions required.
- No need of previous video editing technical skills.
- Production of high quality image for graphics and video material.

Scenario 2: The Mobile TV producer uses traditional video editing tools (NLEs) within the current production editor. The Mobile TV producer is a professional video editor. He deals with heavy production when repurposing content. The production tool is a plug in that opens within the video editor (NLE), and enables the resizing and re-branding of video content. When the production tool is applied, the changes are saved and the work continues in the NLE.

With this approach the dedicated Mobile TV production tool is mainly used for smart resizing the video content. It does not deal with file formats conversion and integrate with the video editor workflow. However, when graphics are rendered on high resolution and then compressed, lower quality of graphics are produced.

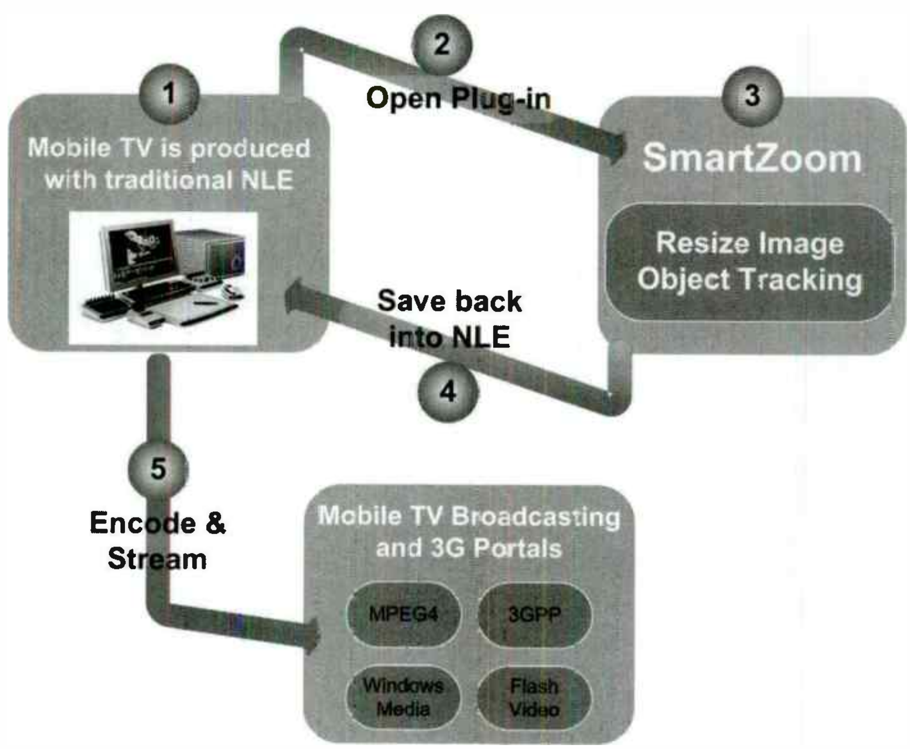


Figure 4: Mobile TV production tool as a NLE plug-in

NEED FOR NEW TECHNOLOGIES

New production tools are emerging with image processing features which rely on state-of-art technology algorithms to aid the operator to adapt the video content to formats that are optimally viewed on the small screen.

For example, a tool for automatic or semi-automatic resizing would enable the operator to mark the desired frame size on key frames while performing smooth transformation between them, and displaying the end product to the operator for quality and viewing experience examination. These systems will also perform moving objects identification and tracking, and help the operator to mark such objects to for the resizing design. Image processing algorithms are capable of automatically keeping the moving or other "significant" objects marked by the operator to remain within the interpolated frame borders throughout the generated clip.

Michael Birenboim

Michael is Chief Technology Officer at CloseVU, an innovative development company focusing on Mobile TV content production tools. An expert in image processing, 3D visualization, and object-tracking technologies, and a seasoned conference speaker, Michael has developed and implemented various algorithms and patents for leading 3D and Image technology firms such as N-Trig, Solidimension and the Military.

Michael holds a B.Sc in Physics and Computer Science.

About CloseVU

CloseVU is an innovative start-up that develops post-production tools to aid video editors to adapt the video content to formats and forms that are optimally viewed on the small screen for Mobile TV. CloseVU products are aimed mainly to content creators, television stations, and mobile telecom operators adapting and repurposing original sports video content to Mobile TV. CloseVU has developed unique image processing algorithms for object tracking. For more info, check www.closevu.com

GLOSSARY

3G Services – The services associated with 3G (third generation of wireless communication technology), provide the ability to transfer simultaneously both voice data (a telephone call) and non-voice data (downloading information, exchanging email). 3G services include broad bandwidth and high speed (upwards of 2 Mbps) for TV streaming.

Mobile TV – Watching TV on a cell phone, Service to subscribers via mobile telecommunications networks. There are several mobile TV air interfaces: Digital Multimedia Broadcasting (DMB) is based on the Digital Audio Broadcasting radio standard (DAB); Digital Video Broadcast-Handheld (DVB-H) is the mobile version of the international digital TV standard (DVB), and Forward Link Only (FLO) is based on QUALCOMM's popular CDMA technology (MediaFLO).

Signal Processing – The processing of signals (such as detection, shaping, converting, coding, and time positioning), that results in their transformation into other forms.

Smart Zoom – Based on smart object tracking algorithms, crops a portion of a video clip taken at the maximum image size to obtain a zoomed image.

** All brands and trademarks are the property of their respective owners.*

MEDIA-CENTRIC DATA-DRIVEN WORKFLOW IN BROADCAST OPERATIONS

John Price
VCI Solutions
Austin, Texas

ABSTRACT

A data-driven workflow relies on the system to apply policies, rules and conditions to automate the transfer, transformation and delivery of content – guided by the data attributes and pre-packaged metadata imbedded in the content.. Current technology, development tools, various data warehousing techniques and standardization efforts like the work of the SMPTE S22-10 Data Exchange Group – enable an architectural framework and model to accomplish a rules-based, data-driven workflow for broadcast operations. In this data-driven workflow modeling framework, traffic, programming, automation and content delivery systems work together to transfer, store, manage and deliver content. This reduces the amount of human intervention and automates the processes required to provision and execute play-out schedules. A media-centric system manages content classes, categories, and attributes with a sophisticated database. When tightly integrated with engines for command, control, rules, workflow and business intelligence, now a data-driven framework facilitates automated and seamless execution of schedule-based, content event transactions.

THE PREMISE

Mainstream media – broadcasters – are still the 800-pound gorilla in terms of advertising revenue, mass audience, and impact. That's the good news. The problem with 800-pound gorillas is they tend not to move quickly – their size and rich diet makes them tend to be a bit lethargic. Drat, those lean, mean little monkeys are jumping on their backs and always stealing away little pieces of the pie. That is the television industry.today. The intent here is not to resolve the well-documented and endlessly discussed issue of audience fragmentation, convergence and the Internet opportunity or threat. However, we will provide some ideas and concepts to work smarter. The whole point of the paper is to explore a framework to use metadata more intelligently.

First, let's set a premise. A long established maxim is that CONTENT IS KING. In today's digital world, content may indeed still be king, but "METADATA RULES!" While compelling content is certainly important, it's now a million channel universe, with ubiquitous content in myriad formats, distribution channels, delivery devices, free and ad-supported content, mainstream and social networks – content and media are everywhere. The most compelling content in the world has little value if no one knows about it. Metadata provides that meaningful information and enables the tools to filter the universe of content to get what you want when you want it.

Content and Metadata

For broadcasters, according to SMPTE, content has two elements: content essence and content metadata. The essence is *the encoded information that directly represents the actual message… Essence is the physical representation of content in different forms and formats used for different purposes*[1].

Content metadata, can be classified into:

- *Content-related metadata,* giving a description of the actual content or subject matter;
- *Material-related metadata,* describing available formats, encoding parameters, and recording specific information;
- *Location-related metadata,* describing location and number of copies, condition of carriers, etc.[2]

---

[1] *Professional Content Management Systems: Handling Digital Media Assets* A.Mauthe, P. Thomas, (2004), John Wiley & Sons, Ltd, Chichester. England.
[2] *Ibid.*

Metadata is generated at every stage of the content lifecycle. It describes characteristics and workflow related to each stage – creation, production, contribution, storage, planning, scheduling, approval, transform, transfer, distribution, delivery archive, etc. A content object and all its related metadata are stored in a database. At each stage of the content lifecycle metadata is needed to describe the content object and act on it to fulfill workflow tasks and requests.

Content Lifecycle and the Workflow

A modern television operation functions much like a factory. It takes raw material (video footage, images, commercials, programs, etc.) and processes it for eventual distribution and delivery. The workflows are the steps and tasks taken to prepare the media content for eventual delivery. In a broadcast operation, every piece of content has its own lifecycle. Syndicated programs are ingested, played to air and deleted. Promos, bumpers and stings are created, stored, used many times and archived to re-use again. Commercials are ingested, stored and played for a certain time period. Content can be re-used, re-purposed and re-created, and the lifecycle begins again.

As content moves through its lifecycle, schedules are created and different devices are used to store, move and play content. Multiple broadcast devices execute sophisticated schedules with multi-layered events and elements, all integrated and controlled by play-out automation.

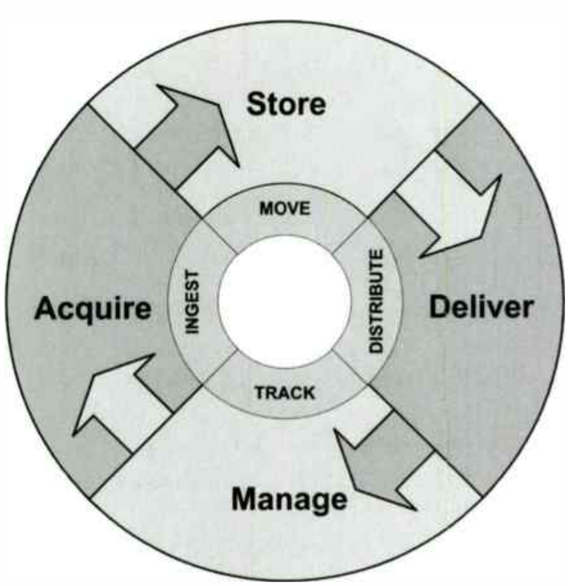


Figure 1. Content Lifecycle

Figure 1 depicts the content lifecycle. Each of the identified stages – acquisition-ingest, storage, delivery, and management – can have many processes. Each process is accomplished with activities and tasks applied to the content – ingest, move, distribute and track.

Different processes are applied as content moves through the broadcast operation, each adding more metadata. These processes are usually repeatable, creating a series of related interactions between people and the system. This is the workflow

Workflow is defined as a series of tasks to accomplish or move through a work process or path. More specifically, according to Wiki (www.wikipedia.org), *workflow is the operational aspect of a work procedure: how tasks are structured, who performs them, what their relative order is, how they are synchronized, how information flows to support the tasks and how tasks are being tracked.*

These processes can be initiated in many ways. In a data-driven workflow, the goal is to invoke a process, based on the content attributes, the task sequence, the state and any conditions related to the content object.

WHAT IS A DATA-DRIVEN WORKFLOW?

Ultimately, a data-driven workflow minimizes the amount of human intervention needed to move content through the various processes in the content lifecycle. Operators and managers would be conducting and monitoring the flow, intervening only when necessary

A workflow consists of states, events, and transitions that define the path a content object (metadata and/or essence elements) takes from initial acquisition to delivery and completion. The workflow guides users through ingest, conformance, tracking and auditing processes. The complexity of the workflow depends on the processes you are following.

Workflow States, Events and Transitions

States indicate a step in the workflow process and the current status of the content object. Events specify the action or task that can be performed at each step (or state) in the workflow. Transitions are points at which content moves from one state to another and are governed by rules for that transition. States can also include additional rules for processing, assignment, access and availability.

Table 1 shows an example of workflow states and events for ingest of content.

STATE	EVENTS
Ingest pending	Content Source Identification Schedule for Recording Assign Encoding Path Assign Storage Location
Transition	
Ingest complete	Notification of Delivery Recording Status, Percent Complete Confirmation of Recording
Transition	
Conform	Assign Workflow Tasks: Transcode Timing Versioning Segmenting Create Browse Copy Create Archive Copy Create Safety Copy Quality Check
Transition	
Conform Complete	Content Available

Table 1. Example of State – Event – Transition Workflow.

In a data-driven workflow, the system is aware of the state and can apply policies, rules and conditions to the content object, directing specific events, paths or transitions to take place.

Challenges and Obstacles

Many media companies, station groups and networks are saddled with entrenched automation systems and applications. Isolated proprietary data silos and legacy technology limit the ability to integrate effectively with new applications, business models and technology. Modern digital television station infrastructure is quickly evolving to network-based systems.

In today's world, metadata information resides in multiple database repositories and silos, creating islands of information. In many cases these assets are unmanaged and unstructured, ignoring the inherent value of the content.

There are key challenges for content management and automated workflows in evolving digital broadcast operations.

- *Multiple formats.* The same content can exist in many different formats, each instance compatible with some, but not all, devices and routing paths, each serving its own purpose and processes in the content lifecycle.
- *Multiple databases.* Each database maintains metadata but there is limited interoperability between systems. Problems quickly arise between different schemas, nomenclature and normalization.
- *Multiple devices and streams.* Managing content for distribution and delivery requires the orchestration and provisioning of schedules, streams, devices and secondary event contributions, each with their own metadata and workflow processes.

- *Multiple workflows.* Different users have different requirements for content and different stages of the lifecycle create many different workflows and activities. Each are controlled and directed by different processes and interact with multiple systems.

Proprietary, monolithic broadcast products create integration roadblocks causing costly workarounds, repetitive tasks and lost productivity. Legacy and obsolete systems lacking scalability usually have complex installations, proprietary hardware and software components, and often require custom development for simple changes.

EXAMPLE – AUTOMATING CONTENT DELIVERY

Before getting lost in the weeds, let us examine a data-driven, automated workflow process that replaced a manual process, minimizing the need for human intervention. This example of an ingest process shows the operational impact from the practical application of these concepts. It requires communication, data exchange and coordination among systems in the broadcast operation: Pathfire, a content delivery service, VCI Automation autoXe™, an automation system, and the station's traffic system.

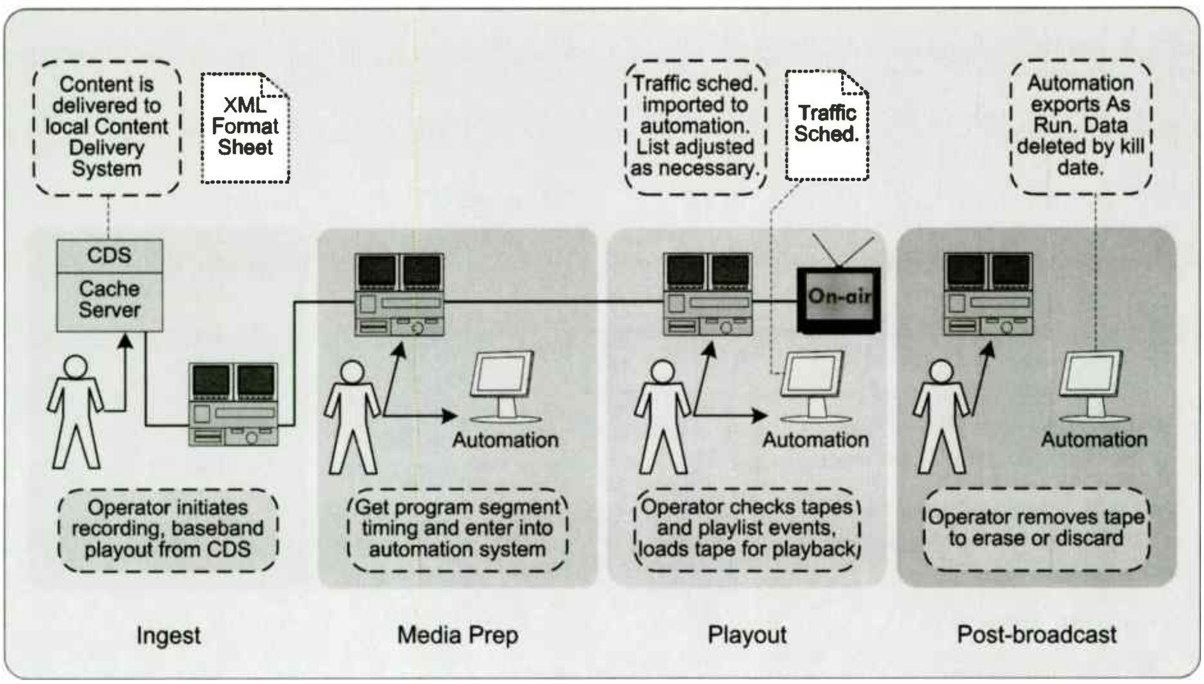


Figure 2. Manual Content Delivery Process and Workflow Tasks

A Manual Process

A broadcast station with six channels receives up to 75 syndicated programs a week from the Pathfire content delivery service, delivered directly to an on-site cache server. The 30 and 60 minute programs consist of several components – program segments, barter advertisements, slates, etc. Programs are scheduled for broadcast based on their segments, which may or may not have barter advertising spots associated with them. Pathfire provided a format sheet for the programs with content ID's and timing information. The manual process requires operators initiate the digital file for baseband play-out. The signal is then recorded on videotape. When the recording is done the operator preps the program, entering the content ID, program name, program episode, and title in the master control automation system. The operator then finds the start of each segment recorded on the tape and enters that metadata plus the duration from the format sheet.

When it is time for play-out, the operator loads the tape into the VTR assigned to that channel and makes sure the playlist is ready. Traffic has already scheduled local commercial avails in the allotted slots for the program. The local commercials are played from a video server. Operators use timers and alarms to remind them to start the next recording and the entire process usually takes two operators from 1:00 PM to 10:00 AM the next day.

This process is not uncommon. The cache server cannot play the content to air so the programs must be made to conform to the station's on-air play-out format, in this case play-out on videotape. The process is time and labor intensive, prone to human error at multiple points, the videotape is consumable stock and mechanical VTRs need regular maintenance and repair.

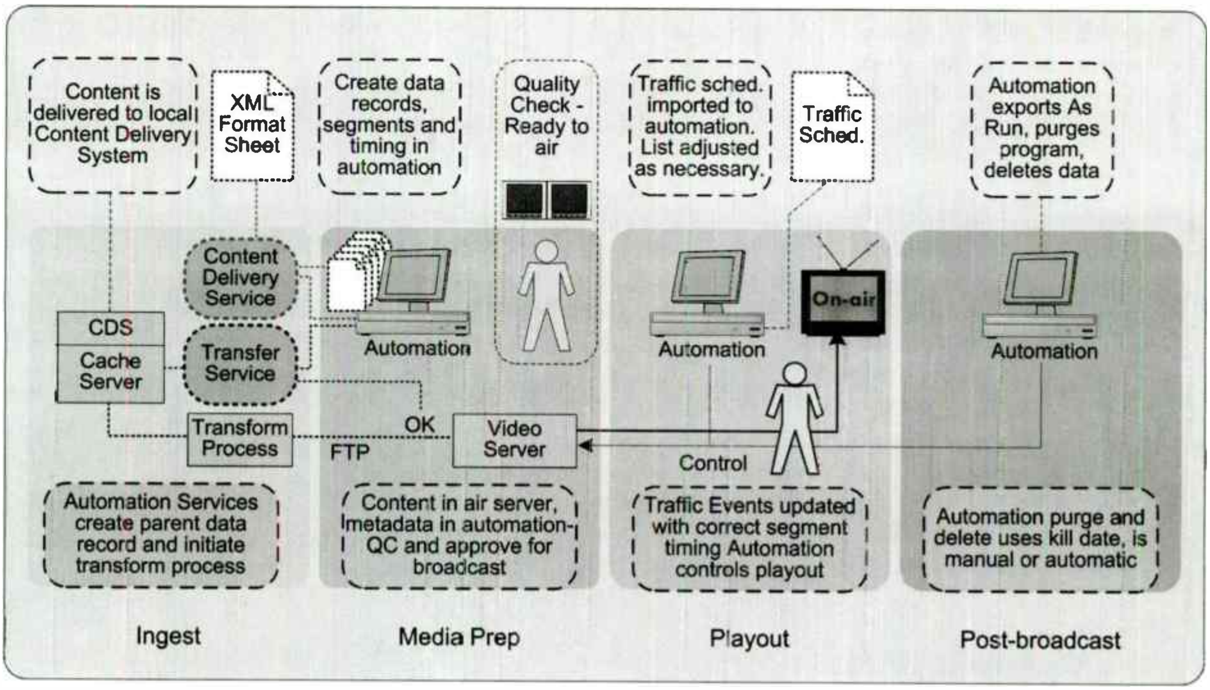


Figure 3. Automated Data-Driven Content Delivery Process and Workflow Tasks

An Automated Process

In the data-driven, automated process, the content arrives at the Pathfire System (CDS) cache server and the autoXe automation system gets notification. autoXe *Content Delivery Service* creates a new database record using the Pathfire Content ID. This is a unique ID assigned by Pathfire for the entire program, which includes segments, ads, barter spots, and fee spots. autoXe then uses *Transfer Service* to request the content transfer to the play-to-air server.

Using the Pathfire Server Connect, the file on the cache server is first localized, making a "container" with all the program segments, promos, ads, barter spots, and fee spots, created as a contiguous digital file with continuous timecode. This file is sent to a transcode engine to transform the file format to be compatible with the station's play-to-air server. When transcoding is completed, the file is moved to the air-server via FTP.

When the Transfer Service completes its job, autoXe *Content Delivery Service* requests the XML format sheet from Pathfire's Automation Connect. The format sheet includes metadata for all the items transferred in the "container" Pathfire content file. The *Content Delivery Service* gives the data record a kill date (+*n* days) and also creates an alternate ID for the program that will match the content ID in the Traffic. The format sheet provides all the timing information for the sequential "child" segments that are added to the database. When the traffic schedule is downloaded, the event ID for the sequential program segments will have the same ID and be ready for playback.

To deal with barter and fee commercial spots that are part of the syndicated program (as opposed to local Traffic Scheduling commercial break slots for Local Avails) but may have different scheduling or trafficking requirements, autoXe can process the XML format sheet in three different ways based on what the station wants:

1) Single Segments. Every item in the syndicated program is identified in the autoXe database with its own ID, SOM, DUR, Description and Item Type. When traffic schedules the program, each segment and Ad Item is an event in the schedule.

2) Mix Segments and Ads. Program segments are identified as individual Events. Ads are combined into a single block as an Ad Event. This allows logos and other secondary events to be on or off based on the type of content.

3) Joined Segments and Ads. Program segments and ads are combined into a single event up to any gap in the contiguous timecode. A gap would be an item not recognized as a program or ad element needed for on-air playback, such as <Slate> <Black> or <Local Ad>. When this syndicated program is scheduled, the program segment and contiguous ads appear as a single block in the schedule.

After the program airs, the *as run log* is exported to traffic and reconciled. The program can be purged from the on-air server manually or automatically based on the kill date. Pathfire also notifies autoXe *Content Delivery Service* to delete the data record.

Results and Observations

Increased productivity was the immediate impact for the station in our example. Operators were no longer spending two shifts a day to dub and prep programs to videotape. 75 + syndicated programs being efficiently processed each week led to the system being expanded to include commercial delivery systems such as Fast Channel, VYVX, DG Systems and others. The station was pleased with the measurable and significant return on investment.

Broadcast workflow processes are not simple. At first glance, our example is a fairly straightforward process, digital content – syndicated programming – arrives at the facility ready to be processed so it can be broadcast. Multiple systems, devices, and applications must interact and exchange metadata, information, status and notifications to deliver an automated workflow. As operations strive for greater efficiencies, interoperability, flexibility and adaptability will be the key requirements for business and operational management systems.

ARCHITECTURE

A new approach is needed for effective and appropriate interoperability to enable point-of-sale to point-of-air™ efficiencies in the broadcast arena.

A service-oriented architecture (SOA) enables different systems to interoperate and exchange data without the significant development needed for tightly integrated systems. Getting multiple specialized software systems, applications and technology platforms working together provides greater efficiencies, and delivers more capabilities. A service-oriented architecture with applications built on a core technology platform enables multiple systems to exchange data, deploy web services and orchestrate workflow processes.

Below is a Service-Oriented Architecture and system model to enable a data-driven workflow for broadcast operations. It separates functionality into a collection of processing layers or tiers operating together to contribute to the solution. The system manages the device resources for storage, staging and play-out. The proper path, media, broadcast devices and play-out channel can be aligned for delivery of content with the desired metadata attributes. The architecture allows an event-based, transactional system interacting with the data for the provisioning and fulfillment of workflow requests.

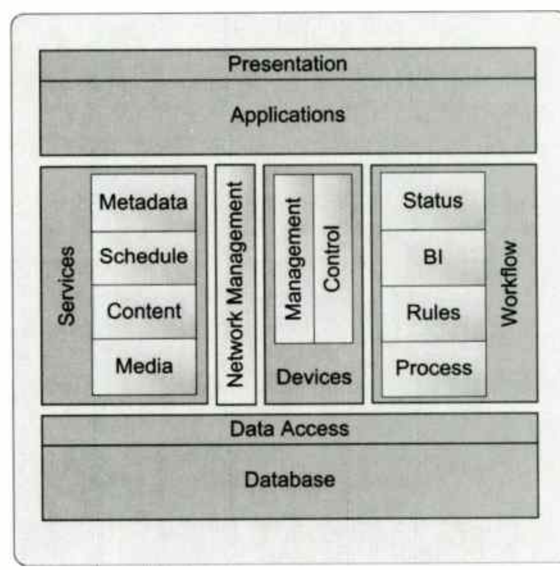


Figure 4. Service-Oriented Architecture for data-driven platform.

This architecture has many advantages over traditional, monolithic automation and content management systems. It promotes a highly adaptable and maintainable system. Each layer is dedicated to specific tasks and is capable of expansion and modification without impacting other layers.

- *Presentation.* The part of the system seen by the users, the user interface (UI). Implemented as a thin client, it merely processes the user input and displays results. Major processing is delegated to the application and workflow engines residing on centralized servers. This allows many of the UI components to be web-browser based interfaces.

- *Services, Workflow and Applications.* These layers are where the operator workflow and business logic decisions are resolved. It includes a collection of services and components that perform tasks, such as managing content creation, storage, and schedule delivery as well as security and user management.
- *Device Management and Control.* This layer provides the high level control of physical broadcast devices to the workflow and application layers. For example, a schedule-delivery service in the workflow layer can issue a command: "play content" without being concerned with the actual command byte stream that directs the playback device to start playing. The separation of the traditional automation device control and the physical device layer from presentation and database layers allows the majority of the system to use non-real-time techniques without compromising frame accuracy of content delivery.
- *Network Management.* This layer monitors the status and availability of the system, ensuring that all parts are functioning.
- *Data Access and Database.* The object relational database is the repository of the system. Designed to manage data access, synchronization, connectivity and interaction across all the layers. Access to the database is through discrete, controllable interfaces. It maintains all system state, configuration and transactional history. It also synchronizes and maintains local cached data throughout the system to increase robustness and performance of the system.

The modular architecture supports a software platform approach to vertical application development. This provides the common elements and components for new applications built on top of it. For customers and users this allows faster integration with new products, devices and services, It also provides greater flexibility, customization and configurability. This is achieved by using software development tools, technologies and standards, including object-oriented programming, software patterns, C#, SOAP, XML, ASP.NET, and COM+.

Platform Foundation: The Database

The database is the foundation for the platform architecture. In the past, broadcasters would implement separate databases for each of their various departments. For instance, the traffic department would have one independent database, while operations and programming would have another. However, there often existed many common and shared attributes (columns of data) among these databases, which would invariably result in not only the duplication of mass amounts of data, but the added expense of additional storage and processing equipment for their continued use. These common attributes, across databases, could contain different data even though they represent the same meaning. This opens the door to mistakes and confusion since it is not clear which database stores the true representation.

These database trends of the past and their many pitfalls still persist today. Many databases continue to be "flat" (single file, non-relational) or might be actual relational databases used in a "flat" manner. Within this model, there could be only one entity (table), and all the attributes would be contained in this single table.

Our approach uses Microsoft® SQL Server™ 2005 in an object-relational data model. The database is designed for platform independence and can be ported to other database systems such as Oracle or IBM DB2. The data model is compatible with others, such as Dublin Core, EBU P-META, PB Core, SMEF and SMPTE Metadata Dictionary. The database design we use encapsulates defined objects that relate to each other in an organized manner. There are many entities (tables) of each object and each of these entities would have attributes (columns) that help define the object. Defining objects instead of simply having a relational database encapsulates the data in a meaningful way so objects can relate better.

The objects relate to each other using object indexes and associations. The database is armed with triggers and rules to define the relationships between objects and to prevent these objects from becoming unrelated. For example, if a request to delete a row in one table which is related to a column in a different table, the triggers and rules would intervene and alert the requesting process of the error. This type of defense ensures the integrity of the database.

Although the database can act as a central repository, the database provides access to distributed content assets and metadata. New users, business processes and operational workflows can access and exchange metadata. The technology connects multiple repositories of data and content. To automate a data-driven workflow, based on the content's metadata, different systems need to communicate and exchange needed data and information.

STANDARDS AND INTEROPERABILITY

As next generation systems continue to evolve, a key requirement is for transparent access to other data and content repositories. The systems must be open and adaptable, able to integrate and interact with each other to accomplish workflow tasks.. In a perfect world, industry standards would be established to facilitate communication and data exchange.

The work of the SMPTE S22-10 Group to standardize metadata exchange between broadcast systems has great potential. The Group's Broadcast eXchange Format Protocol (BXF) provides a well-defined, standardized way to exchange metadata and messages between systems. The group that developed BXF is composed of broadcasters and companies directly impacted by the difficulties in integrating with different vendors and systems. The players represented develop broadcast systems for Automation Content Delivery, Digital Asset Management, Traffic, Program Management, and other vendors in the digital supply chain..

All these vendors' systems and applications have metadata critical to their own functions in the broadcast operation. Additionally, they depend on other systems for metadata and workflow instructions to accomplish processes in the broadcast environment. For customers, the lack of interoperability and limited integration adds many manual steps and repetitive data entry. For vendors, each integration project is custom development and must be supported through subsequent product releases.

The BXF Protocol provides a better way. The industry standard Protocol would replace many of the proprietary interfaces currently being used. It includes data elements that are common among all tasks, and addresses communication, transport, and security issues.

According to the BXF documentation:

The principal scope of BXF is to provide a standard way to exchange the data and metadata among professional systems, as follows:

1. *Broadcast schedules, including playout and record schedules*
2. *As-run information*
3. *Content metadata, such as Content ID, Title, Duration, etc.*
4. *Content management requests such as dub and purge requests*
5. *Requests for transfer of content*

To understand the scope of BXF, a little background is helpful. The genesis of BXF can be traced to the need for a consistent yet flexible means for exchanging schedule, as-run, and content metadata between Traffic and Automation systems. Literally hundreds of proprietary, fixed interfaces and protocols have been created over the past 20 years or so between these two types of systems. Vendors who create and develop these systems were invited to share ideas in search of a better way to facilitate the exchange of data between systems. The group quickly embraced Content Distribution and Program Management vendors as well, as they too were seeking a form of standardization and improvement in their interfacing efforts. BXF supports the exchange of single or multiple records at one time, over a variety of transport mechanisms. While endeavoring to be a comprehensive protocol, it is acknowledged that it is possible that additional data elements, or data elements that apply to a few specific systems, may need to be exchanged. For this reason, Private Information structures have been placed at various points in the schema to allow vendors the flexibility to add data elements.

The SMPTE S22-10 Group has published an XML Schema and data dictionary supporting BXF, to enable development for the interchange of messages and metadata. The efforts of a committed group of individuals representing many companies have resulted in an industry-standard protocol that offers great benefits to the broadcast industry.

The BXF Protocol will help to accelerate and facilitate the data-driven workflow processes. From our earlier example, using BXF protocol for the workflow services would provide a replicatable and repeatable service. The content delivery system can use the same BXF-based service to message automation, traffic and program management systems. An event in one system – arrival of content – is now communicated to other systems, which can respond with a new message requesting metadata. All the systems will be working on the same metadata, communicating dynamically. Broadcast organizations deploying technologies capable of service-oriented, standards-based systems will have greater agility and ability to leverage their content assets.

CONCLUSION

Content is king – Metadata rules! It is critical to recognize that broadcast systems interface with both humans and with other systems. Metadata plays a key role to facilitate the human and system interfaces. The physical asset, the actual content essence is now more often digital – a file with countless bits in a binary format. Metadata is important because it dictates and defines how that asset can be used or consumed.

In a broadcast operation, or any advertiser-supported business, point-of-sale to point-of-air encompasses scores of processes, workflows and tasks to create, manage and distribute content. In a multi-site, multi-channel, multi-format, multi-platform world, reducing manual tasks and human intervention is imperative.

The structure and architecture to enable an automated, media-centric data-driven workflow exists. Operational and business systems are focusing on technology platforms that embrace workflow and business processes. Forward thinking companies focus on and support unifying technologies that provide users and customers with more open and adaptable systems.

In a million-channel world, all of us need as much help as possible to create, manage, distribute and consume content.

Mobile Reception of ATSC DTV Signals

Wednesday, April 18, 2007
2:00 PM – 3:30 PM

Chairperson: Louis Libin
BroadComm, Inc., Woodmere, NY

TV Broadcaster Connecting with the Mobile User
Richard Chernock, Triveni Digital, Princeton Junction, NJ

Advanced VSB: A Proposed Enhancement ATSC DTV
Jungpil Yu, Samsung, Suwon, Korea

***Mobile and Handheld Reception for ATSC Broadcasters**
Wayne Bretl, Zenith Electronics Corp., Lincolnshire, IL

*Paper not available at the time of publication

TV BROADCASTER CONNECTING WITH THE MOBILE USER

RICH CHERNOCK
Triveni Digital
Princeton Junction, NJ

RICH GONG
Triveni Digital
Princeton Junction, NJ

INTRODUCTION

Billions are being spent by Qualcom, Crown Castle and others to develop and deploy a broadband broadcast network. Mobile TV revenues are expected to grow and surpass half a billion dollars by 2010. At the same time, manufacturers, content providers, and consumers alike are increasingly intrigued by the prospect of wireless, mobile delivery and reception of entertainment and information content. While much of the effort around wireless mobile content has been concentrated on hand-held devices, there has also been a significant growth in the effort to deliver high-bandwidth content to the occupants of motor vehicles. Devices which deliver limited wireless content to motor vehicles, such as Global Positioning data in order to run navigation devices, or traffic data delivered via satellite radio providers, are increasingly popular and are well on their way to becoming the norm. While it is common to find OEM or after-market products in vehicles which play high-bandwidth content (such as DVD players, interfaces with iPods and other mp3 players, etc.), it is rare to find this content delivered to the vehicle by any means other than the user physically carrying the content in and out of the vehicle.

There is an emerging spectrum of approaches to the wireless delivery of content to the motor vehicle emerging in the marketplace. On one end of the spectrum, traffic data delivered via a satellite radio provider has already been mentioned; this data is then processed and rendered in a graphical fashion to display traffic conditions to the driver, presumably aiding the driver's decision about choosing routes depending on traffic conditions. This represents a highly specific, highly targeted application for which only a small amount of data is transmitted with high frequency, as required to update a highly dynamic situation. On the other end of the spectrum, at least one manufacturer has begun to promote a device which promises to deliver continuous wireless broadband internet connectivity to the vehicle, thus supporting the ability of laptops, PDA's, etc. to connect to and download content from the internet. This represents an unbounded, and thus unstructured, universe of content that can be delivered to the occupants of the vehicle.

While the extremes in the wireless delivery spectrum described above have been commanding most of the attention as of late, there is a promising delivery mechanism which occupies the solid middle of this spectrum – in terms of content <u>and</u> cost – and up until now, has been largely ignored: Digital broadcast of data and content to the motor vehicle, using the existing ATSC 8-VSB modulation standard[1]. Many view ATSC 8-VSB as incompatible with mobile reception and have thus never seriously considered the technology as a feasible alternative for the delivery of content to mobile devices or motor vehicle occupants. However, a newly developed system leverages ASTC 8-VSB technology to solve the need for delivery of high-bandwidth information and entertainment content to motor vehicles. This system uses a mixed delivery model that pushes delivery of large data objects over ASTC 8-VSB during periods of favorable reception conditions and utilizes other techniques for smaller, real time information. This paper will describe this distribution system and provide an overview of applications.

POSSIBLE CONSIDERATIONS FOR A MOTOR VEHICLE APPLICATION

Before the use of ATSC 8-VSB technology for mobile reception can be discussed in detail, a brief outline of the types of features proposed for a motor vehicle application is required, as the features of the specific application can dictate the type of technology that is appropriate in scale and cost.

While some manufacturers are promoting devices that provide vehicle occupants with the possibility of full broadband internet access, one could argue that vehicle occupants are not necessarily well served by the unbounded universe of content available from the internet. Rather, vehicle occupants – particularly drivers – are better served by well-defined and well-structured content designed specifically for the unique environment of the motor vehicle and the critical task of driving. A possible set of considerations that affect the definition and structure of this content for motor vehicle occupants are as follows:

1. Since the vast majority of motor vehicles on the road at any given time are occupied only by the driver (79% of all workers commute to work alone, according to a 2003 Bureau of Transportation Statistics study[2]), features for a motor vehicle application should primarily focus on delivering content to aid the driving task and/or enhance the driving experience, without compromising safety. Content for other vehicle occupants (passengers) should also be considered, but is of secondary interest.

2. Since an increasing number of vehicles are equipped with navigation devices (either as original equipment from the vehicle manufacturer or as after-market devices) which contain pre-loaded static content (such as maps and routing information), delivered content and data should provide drivers with on-demand access to information that is of a dynamic, rapidly changing nature, which directly aids the driving task. Examples of such content include traffic conditions, current weather conditions and forecasts, and alert information.

3. Delivered content and data should provide drivers with on-demand access to information which also needs to be refreshed regularly but less frequently, and is of significant value to the driving experience, while less critical to the immediate driving task. Examples of such content include current gas prices, parking information, movie times, updates to maps, updates to "point-of-interest" data, etc.

4. Delivered content should be local in its nature, not only providing drivers local weather and traffic conditions, but also information and updates about local "points-of-interest". Examples of localized information include previously mentioned examples of gas prices and movie times as well as updates to local businesses (e.g. a change in phone number), and promotions for a local business (coupon or ad).

5. Delivered content could provide drivers access to ancillary content which by its nature needs to be refreshed regularly but is not directly related to the driving task or travel experience (while not hindering the driving task). An example of such content is audio podcasts, which currently must be downloaded to an external device such as an iPod or portable media player. A digital broadcast delivery mechanism can easily deliver this kind of content directly to the vehicle on a recurring basis as dictated by the podcast feed, bypassing the need for the user to transfer the podcast from another device.

6. Delivered content could provide vehicle passengers with access to other high-bandwidth content not available to the driver, such as games, videos, etc. Content in this category can be accessed through a variety of models, such as purely free content (podcasts and short video clips) or different subscription levels (for longer video clips, feature-length programs, feature-length movies, or games).

The purpose of listing this set of considerations is to demonstrate that if the specific kind of mobile application is taken into account (which in this case is a dedicated receiver in the motor vehicle), the content most beneficial for this application is clearly narrowed down to some well-defined data sets or "walled garden" of features. This feature set, in turn, can be more than adequately addressed by the current digital broadcasting technology available to us today. The specific feature set of the application being developed will be discussed in a later section of this paper.

USE OF ATSC 8-VSB FOR MOBILE RECEPTION

Common Objections to the Use of 8-VSB Modulation for Mobile Receivers

Perhaps one manner in which to examine the use of ATSC 8-VSB modulation to deliver content to motor vehicle occupants is to address the major objections up front to its use for mobile reception in general. How is it that the terms "mobile reception" and "ATSC 8-VSB" can even be used in the same sentence? The most common (and legitimate) objections to the use of 8-VSB for the reception of content at mobile receivers are as follows:

1. *The use of 8-VSB for the delivery of content to mobile receivers is unreliable at best, and nearly impossible at worst[3].* ATSC 8-VSB is designed for stationary receivers, and the reception difficulties in moving vehicles due to the Doppler effect are well known.

2. *With any broadcast scheme, including 8-VSB, there is no back channel, and thus no possibility for interactivity.* Any delivery of mobile content using digital broadcasting will necessarily be *pushed* content, and would preclude the possibility of real-time user interactivity with the content provider(s). Thus, the universe of possible content is significantly restricted, relative to the content that is possible with other delivery mechanisms, such as broadband internet access.

Let's consider each of these objections in turn.

First, "mobile reception" often does not really need to be truly mobile in its purest sense to provide a more than satisfactory user experience to the mobile receiver. In the initial investigations into the use of a digital broadcast solution, we have found that a receiver characterized by its mobility (such as a motor vehicle) spends a significant amount of time not being mobile. These non-mobile times or events, even during periods of high overall mobility, turn out to be non-trivial and lend themselves well to an opportunistic *mixed delivery model* of information and content via 8-VSB digital broadcast. This model will be described in detail in the following section.

Second, while it is true that the use of any kind of broadcast technology to deliver content to the receiver represents a push model of delivery, this model is thought to be more than sufficient to address the primary features and data sets that are optimal for a motor vehicle application, as discussed in the previous section. In the majority of cases, local interactivity is sufficient – user interactions with pre-delivered data. These considerations will be briefly reviewed below.

Finally, not only is the use of a digital broadcast overlay for the delivery of content to motor vehicles fully sufficient, there are a number of advantages to using this technology as well, for end users as well as content providers, such as broadcasters. These advantages will also be more fully discussed below.

The Mixed Delivery Model

In the mixed delivery model, content is simply and arbitrarily divided into 3 categories, roughly according to the size of the data files being delivered from the source to the receiver: small files (5-100KB), intermediate files (100-500KB), and large files (greater than 500KB, to as large as several GB). The size of the data files roughly corresponds to the frequency by which they are transmitted through digital broadcasting. Small data files can be transmitted frequently (many times an hour) and received intact by the receiver during short periods when the vehicle is stationary or nearly stationary (such as a momentary stop in traffic or at a stop sign, often not more than a few seconds); intermediate data files are transmitted regularly but less frequently (several times per hour) and may require a slightly longer stationary period, such as that which might be encountered during a full stop at a traffic light, in order to be received; and large data files may be transmitted on the order of once per day (or less frequently) and require a much lengthier vehicle stationary period to be fully received, such as when a vehicle is parked during an errand or during a much longer duration such as when parked at the workplace during the day and at home during the evening and overnight. Additionally, files may be received in pieces and reassembled. There is no requirement that the entire file be received in one session.

The considerations for the application feature set were briefly described in a previous section. In general, the features that are deemed to be of high value to motor vehicle occupants lend themselves well to the size of files and to the frequency of updates that are likely to occur. For example, real-time traffic data is contained in small files (often as small as 5KB) and can thus be updated frequently in the typical commuting scenario since only a second or two of stationary or nearly stationary vehicle status is required for reception. As one might expect, traffic conditions are highly dynamic and fluid during the course of the day and even during the course of the typical commute. Drivers would expect and need such information to be updated frequently. Because of the small data sets that are required to render this information in the vehicle, a real-time traffic feature is possible through the delivery of data by digital broadcasting technology. Preliminary studies, in which data files are broadcast via ATSC 8-VSB and are received by a vehicle driving during commuting hours on surface streets, indicate that there are sufficient opportunities to receive data sets of the size needed for traffic and weather updates, and to update these features with adequate frequency.

While traffic conditions represent a category of information that is highly dynamic, another category of information that is also dynamic but less rapidly changing is that of weather conditions and weather forecasts. Current weather conditions can be rendered in a number of ways, and people are becoming increasingly accustomed to the use and interpretation of radar and satellite images provided by any number of weather providers found on broadcast TV and the Internet. Typically these images are updated a few times per hour (typically every 10-15 minutes, depending upon the source). In addition to text weather forecasts and conditions, delivery of static radar and satellite images represents high-value, desirable content to vehicle occupants, especially drivers. These images are contained in data files that are in the "intermediate" size category, and thus are delivered less frequently than files in the "small" category, such as traffic data. In the preliminary investigations using the commuting scenario described above, we have found there are sufficient opportunities – often several times per hour – to fully receive data files of this category in the vehicle, as they are typically in the 300KB to 400KB size range when transmitted through digital broadcasting.

Finally, there are a number of other data sets that constitute the category of "large" files. These include data sets containing the current gas prices, information about local "points of interests" (such as businesses, attractions, etc.), and possibly targeted, location-based advertising content. And, certainly in this category, are the files that constitute podcasts, videos, games, etc. Depending upon the content, these files can be as small as 500KB (for daily gas price information) and be as large as several gigabytes (for full-length movies).

Content in this category may be updated as often as hourly (in the case of some news podcasts), to daily (in the case of daily gas prices and certain podcasts) to weekly (in the case of certain podcasts and subscription movies). Content of this type must be delivered to the vehicle during longer periods of time when the vehicle is stationary, such as during the day when the vehicle is parked at the work place or overnight when the vehicle is parked at home.

In summary, the mixed delivery model transmits small data files with high frequency, intermediate files with less frequency, and large files with low frequency. The file sizes and frequency correspond with the character of the content being transmitted; local traffic conditions change frequently and can be transmitted and received using small data files during commuting, while podcasts and video clips update relatively infrequently and can thus be transmitted and received during the longer periods of time when the vehicle is stationary.

The table below summarizes the mixed delivery model.

Category	**File Size**	**Examples**	**Update Frequency**
Small	Less than 100KB	Traffic conditions	Every 5 – 10 minutes
Intermediate	100KB – 500KB	Weather forecasts and conditions; localized "point of interest" information	As often as every 10-15 minutes for weather; can also be several times per day or once per day for other localized content
Large	Larger than 500KB	Daily gas price information; podcasts; video clips; games; feature-length programs and movies	Several times per day for smaller podcasts; once per day to once per week for gas prices information, other podcasts, games, videos, and movies

The Sufficiency of Content Delivered by the Mixed Delivery Model

As has been previously mentioned, even if delivering content to a mobile receiver using digital broadcasting proved to be sufficient, some might raise the objection that such a delivery mechanism still represents a sub-optimal user experience as it pushes content to the user without the possibility of user interactivity through a back channel. Some vendors are now aggressively promoting devices that promise full internet access and interactivity to the vehicle, in much the same fashion that one might access and interact with the internet in a stationary setting. Nonetheless, for a significant population of mobile users, particularly motor vehicle occupants, an unbounded, unlimited universe of content to the receiver is simply not needed, or even desirable. Considering that the majority of motor vehicles are driver-only occupied, the primary consideration is to provide well-defined and appropriately formatted fresh content to the driver. This content should aid the driving task and enhance the driving experiencing without compromising safety. Initial marketing studies have suggested that perhaps a dozen or so well-defined categories of content would be more than sufficient for a significant portion of the driving population. Properly formatted content targeted to the needs and wants of motor vehicle occupants can be delivered efficiently, reliably, and at a much lower cost – to both the content provider <u>and</u> the customer – than an unbounded universe of internet content. The mixed delivery model assures that the proper types of content are delivered during the kinds of "mobile" conditions in which it is most useful and can be reliable ingested by the receiver.

Advantages of Delivering Content to Vehicles Using ATSC 8-VSB Modulation

Previously in this paper we examined some of the common objections to using ATSC 8-VSB to transmit content and data to a mobile receiver, specifically a dedicated receiver in a motor vehicle. Not only is digital broadcast technology, specifically ATSC 8-VSB modulation, a feasible alternative for delivering content and data for this specific kind of mobile reception, but there are some clear advantages to utilizing such technology. The advantages will be described in terms of the advantages conferred to the end-user, as well as advantages to the broadcaster.

Advantages to Users

Users can benefit from a system that uses ATSC 8-VSB for content delivery to vehicles in the following ways:

- Localization: Because the broadcaster has local coverage, content can be delivered to the vehicle that is relevant to the vehicle's current location. Localization is clearly important for traffic and weather information. Other information, such as gas pricing and restaurant specials benefit from localization as well. Most users will be primarily interested in information relating to the area they are currently driving in.
- Content Filtering: Because of safety issues, "walled garden" content (as described here) is

the most appropriate for delivery to a driver. Pre-selection of types of content for delivery and optimized formatting for display will avoid driver distraction.

- Updates: Content displayed to the user will be constantly refreshed, without a need for user intervention.
- Savings: The lower cost of content delivery to the vehicle using a broadcast mechanism can be passed on to the user as a reduction in service fees.
- Efficiency: The user will benefit from the efficiencies afforded by the one-to-many broadcast distribution system by having a richer set of information available.
- Software Updates: Software updates can be delivered through the same broadcast mechanism as the content (using the ATSC A/97 Software Data Download Service standard[4]).
- Back-seat Applications: This distribution model enables numerous back-seat applications (such as subscription based movie or game download services). Content can be automatically refreshed to allow a more compelling service.

Advantages to Broadcasters

Digital bandwidth is a revenue producing resource for broadcasters. Delivering television content in a linear fashion is currently the most common use of a broadcasters bandwidth. Other forms of content delivery, such as the system described in this paper, provide additional revenue-generating opportunities to the broadcaster.

The following points summarize how broadcasters are advantaged in delivering content to automobiles over other delivery mechanisms:

- Efficiency: The broadcast channel provides a point-multipoint delivery system, which is the most efficient mechanism to deliver common content to multiple receivers.
- Localization: Broadcasters are recognized for providing local content (for example "Local News at 5"). By the nature of the signal propagation from the broadcast towers, the broadcast signal is localized. It is a natural fit to have localized content delivered by a broadcaster.
- Lower Cost: The efficiency of the broadcast mechanism mentioned above translates into a lower cost for providing a content distribution system, relative to other content delivery methods currently being promoted.

These advantages allow additional possible business models for broadcasters that have yet to be fully explored. As an example, a high value feature for a vehicle information system is one that quickly locates a business that is of interest to the vehicle occupants, such as the nearest Starbucks. This feature is certainly available in most current navigation systems. However, depending upon the current location of one's vehicle, there may be up to 4 Starbucks stores within 2 miles of one's location. An enterprising manager of one of the stores might contact the local data-casting TV station, and pay for an ad with a promotion or coupon available only at his or her store, to run for a month. Thus, the driver may notice on his or her vehicle information system that while the nearest Starbucks is 0.5 miles away, there is a Starbucks 1.2 miles away running a half-price special on Grande Lattes that is good for the entire month! This could be the deciding factor in this driver becoming a dedicated customer to that particular store, at least for the duration of the promotion.

This is an example of how this solution not only enhances the basic information to the user that is generally available from current navigation systems, but provides a previously-unrealized source of revenue for the local broadcaster as well.

THE SOLUTION

The specific ideas that have been described in this paper have been instantiated in a working prototype. Many of the assertions in this paper are being validated through real-world experience with a mobile device mounted in a test vehicle. The device receives pre-defined data through 8-VSB modulation, and renders content within a highly compelling and usable vehicle telematic application.

Broadcast Station

In the current solution, a dedicated software application gathers pre-defined sets of data and content from various sources and providers, and "bundles" this data into different packages, depending upon the size of the data files and the frequency and criticality with which they are to be delivered, according to the mixed delivery model. For example, one data bundle may solely consist of traffic data and might be scheduled to be refreshed every 5 minutes or so, due to its small size and its critical nature. Other kinds of data bundles, such as weather, gas prices, podcasts, etc. will also be refreshed and rebroadcast on a regular basis but less frequently, according to their particular criticality.

While there is no requirement as to where this "data bundling" application must reside, it will most likely reside at the local broadcast station where a data broadcast server is installed and can readily broadcast the data bundles to the receivers in motor vehicles tuned to receive the data broadcast. The data broadcast application at the broadcast station schedules the various data bundles for broadcast according to the mixed delivery model described in an earlier section of this paper.

Mobile Receiver in the Motor Vehicle

The dedicated receiver that resides in the motor vehicle contains the data receiver and a specialized application to present the content to the driver in manner which is highly usable while at the same time optimally formatted to meet existing guidelines for vehicle telematics and to minimize distraction from the driving task. The content is presented to the driver in the form of a touch screen LCD display. In the working prototype current under evaluation, the touch screen display is mounted on the dashboard as a separate display, but in a production environment, it is expected to be integrated into the same display that displays the GPS-based navigation system commonly available as original equipment from the vehicle manufacturer. Figure 1, below, shows the touch screen displaying the prototype application mounted in a test vehicle.

Figure 1 – Touch screen for the prototype application mounted in a test vehicle

The prototype application contains features that render the kinds data and content that is of high value for the driving task and driving environment. These features were referenced in an earlier section of this paper discussing the various considerations for such a task and environment. These features include:

- Current local traffic conditions
- Local and regional weather forecasts
- Current weather conditions as shown on radar
- Critical alerts, if activated:
  - Amber alerts
  - Severe weather alerts
  - Vehicle service alerts
  - Other alerts (e.g., Homeland Security)
- Location-based "Point-of-interest" information
  - Nearest ATMs
  - Nearest Attractions
  - Nearest Lodging
  - Nearest Gas Stations including current gas prices
  - Nearest Restaurants
  - Other POIs
- Podcasts
- Location-based ads and promotions

The information encompassed by these features is organized hierarchically in a fashion that is easily accessible through relatively large targets on the touch screen. An example of the top-level or "home" screen for a vehicle based in the Atlanta (GA) metro area is shown in Figure 2, below:

Figure 2 – The "Home" or top level screen showing traffic, weather, and alert overviews

For the most part, information that is deemed of high-value to the driving *task* is displayed in an overview fashion on the home screen. This includes traffic, weather, and alert information. The status of each of these is immediately apparent through the relatively large thumbnails shown on the home screen. More detailed information for each category can be accessed with a single "press" of each category button. For example, Figure 3 shows detailed local and regional weather forecast information after the forecast thumbnail on the home screen has been "pressed":

Figure 3 – Detailed weather forecast information

Information that is deemed of high-value to the driving or travel experience is accessed through the column of "soft buttons" on the right of the display.

<u>Home</u>

The "Home" soft button brings the user back to the top level or Home screen.

Services

The "Services" soft button allows users to access information about various categories of "points of interest" (POIs). A sampling of the possible categories of POIs are mentioned above. The following figures show examples for the categories of Restaurants (Figure 4) and Gas Stations (Figure 5). For each type of POI, the nearest 8 merchants are shown on a grid, based on the driver's current location (as determined by GPS). For each merchant, the merchant name, brand logo, and distance are shown. In the case of Gas Stations, the price for regular unleaded gas is shown instead of the merchant name.

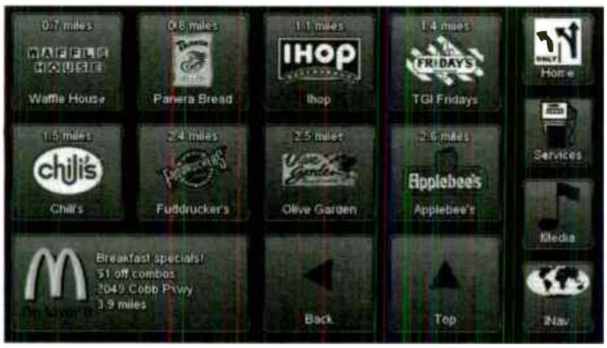


Figure 4 – Nearest restaurants based on current vehicle location

Figure 5 – Nearest gas stations (including current prices for unleaded regular gas) based on current vehicle location

If any specific cell on the grid is selected, detailed information about that merchant is displayed, as shown in Figure 6, which is the detailed information screen for the IHOP restaurant. Navigation, address, and phone information are currently displayed; there is the possibility for additional static information as well. There are "placeholder" links shown for either directly contacting the merchant via phone (some auto makers include on-board equipment which makes this possible) or for accessing the "turn by turn" directions available in most on-board navigation systems.

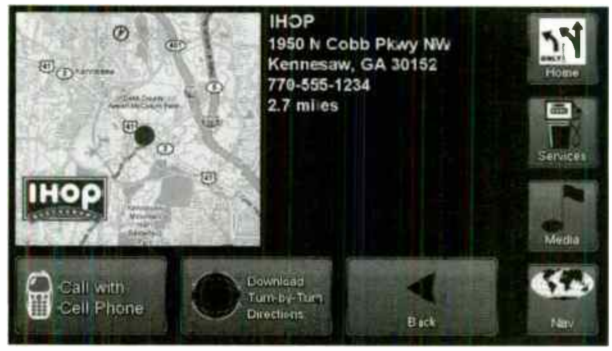


Figure 6 – Detailed information screen for selected item

Media

The "Media" soft button allows users to access the various kinds of specialized media available to the driver, and optionally passengers, of the motor vehicle. This not only includes the traditional "media" channels of AM radio, FM radio, and CD player, but also high-bandwidth media such as podcasts and videos, which can be delivered directly to the vehicle during conditions of favorable reception because of the relatively high data through-put rates of digital broadcasting. Figure 7, below, shows a screen from the "Podcasts" subcategory which provides the driver a half-dozen podcasts that he or she has designated as "favorites" and thus most readily accessible from the several dozen which could be routinely delivered to the vehicle.

Figure 7 – Examples of podcasts available in the vehicle, designated as "favorites"

Location-based Advertising and Promotions

The prototype vehicle application includes some provision for location-based ads to be displayed to the user. While it is well known that users are generally averse to intrusive, loosely targeted or untargeted advertising, some space on the Home screen and the POI screens is provided for relatively discreet, location-based advertising and promotions. An example of how the right kind of location-based ad in a vehicle environment can provide value to both the driver and the broadcaster was discussed in an earlier section outlining the advantages of using digital broadcast technology to deliver content to the vehicle.

FUTURE WORK

Prototype testing has successfully shown the feasibility of the system described in this paper. Additional work will include:

- More complete analysis of data collected
- Analysis of marketing data collected
- More rigorous testing for local reception characteristics and re-tuning when traveling between locales (roaming)
- Additional chip and driver development for the current receiver
- Emerging ATSC standards development for mobile reception

REFERENCES

1. ATSC Digital Television Standard (A/53) Revision E, with Amendments No. 1 and 2

2. http://www.bts.gov/publications/transportation_statistics_annual_report/2005/html/chapter_02/commuting_to_work.html

3. http://www.reedtelevision.com/index.asp?layout=articlePrint&articleID=CA63344

4. ATSC Standard A/97: Software Download Data Service

Advanced VSB
A Proposed Enhancement for ATSC DTV

Jungpil YU
Samsung Electronics, Korea
Jungpil.yu@samsung.com

Eui-Jun PARK
Samsung Electronics, Korea
Euijun.park@samsung.com

Yong-Sik KWON
Samsung Electronics, Korea
yskwon@samsung.com

Mike SIMON
Rohde & Schwarz, USA
Mike.Simon@rsa.rohde-schwarz.com

JoonSoo KIM
Samsung Electronics, Korea
kjoons@samsung.com

Jin Hee JEONG
Samsung Electronics, Korea
Jini.jeong@samsung.com

Jong-Hun KIM
Samsung Electronics, Korea
Jhdm45.kim@samsung.com

Kumran JI
Samsung Electronics, Korea
Kumran.ji@samsung.com

C.B. PATEL
Samsung Electronics, USA
Cbnhdtv@aol.com

Haejoo JEONG
Samsung Electronics, Korea
hjjeong@samsung.com

Abstract

In this paper, a system level enhancement to the ATSC-VSB digital TV standard is described. This technology, Advanced VSB (A-VSB) is backward compatible. A legacy receiver would ignore the added information in the signal while continuing to receive the 8-VSB TV stream. The A-VSB improves the reception of the main stream in dynamic environment. It also provides additional streams, called Turbo streams for a low SNR high dynamic environment. The A-VSB proposal also gives broadcasters the option to operate two or more synchronized transmitters known as a Single Frequency Network to provide more uniform signal strength in their service area.

Keywords
ATSC, DTV, Advanced VSB, A-VSB, VSB, 8VSB, SFN, Extensibility

INTRODUCTION

In A-VSB system, by using Adaptation Field feature of the MPEG2, a backwardly compatible Supplemental Reference Sequence (SRS) is added to the transmitted signal. The SRS (known long contiguous symbols) is available several times during a field. The A-VSB receiver utilizes SRS to enable its equalizer to track the dynamic aspects of the received signal. This helps an A-VSB receiver to maintain reception in rapidly changing multi-path interference in a fixed environment or conversely when the receiver itself is moving. Coupled with latest algorithmic advances in receiver design, the SRS provides new levels of performance in dynamic multi-path environments. The SRS is a system-level solution that makes available frequent equalizer training signals at the 8-VSB physical layer to help new receivers mitigate dynamic multipath.

The A-VSB Turbo stream is very tolerant of severe signal distortion enough to support broadcasting applications such as Indoor and Pedestrian services. The Turbo stream achieves low SNR TOV (Threshold of Visibility) by additional forward error corrections – a serial concatenated convolutional code [1][2][3]. Apart from the SNR improvement, Turbo stream also has bit-level interleaving which offers additional time-diversity. The applied coding scheme in Turbo stream is ATSC-specific in the sense that the inner encoder used is already a part of the VSB system.

In this paper, we also develop the so-called core technology. It provides the foundation on which the three applications tools (SRS, Turbo stream, and SFN) of the A-VSB proposal are based. The core technology consists of compatible changes in VSB system known as the Deterministic Frame (DF) and the Deterministic Trellis Reset (DTR). These changes make the ATSC-VSB system deterministic in its operation. This in turn enables future extensibility of the A-VSB system, which is the goal of A-VSB

ADVANCED VSB APPROACH TO ENHANCEMENT

The first objective of A-VSB is to improve reception issues of 8-VSB services in fixed and portable modes of operation. This system is backwards compatible in that existing (normal 8-VSB) receiver performance is not adversely affected by the Advanced-VSB signal.

This A-VSB system defines the following core techniques:

- Deterministic Frame (DF)
- Deterministic Trellis Reset (DTR)

And, the following application tools:

- Supplementary Reference Sequence (SRS)
- Turbo Stream
- Single Frequency Network (SFN)

These core techniques and application tools can be combined as shown in Figure 1. It shows the core techniques (DF, DTR) as the basis for all of the application tools defined here and potentially in the future. The solid lines show this dependency. Certain tools are used to mitigate propagation channel environments expected for certain broadcast services. Again the solid lines show this relationship. Tools are combined together synergistically for certain terrestrial environments and constitute a service profile for a given service type. The solid lines between "Application Tools" and "Service Environment" indicate this profile dependency. The dash boxes and lines are potential future tools and dependency is not yet defined.

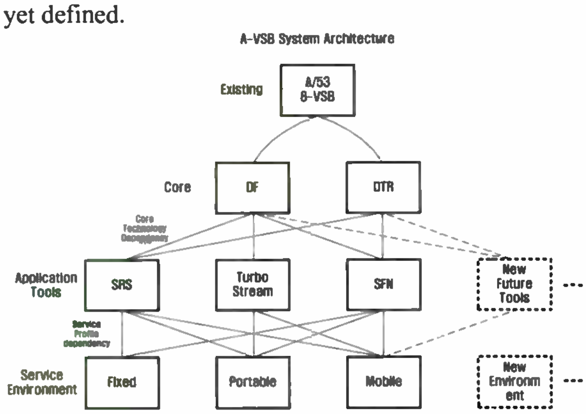


Figure 1 A-VSB Approach to System Enhancement

DETERMINISTIC FRAME & TRELLIS RESET

Deterministic Frame (DF)

The first core technique of A-VSB is to make the mapping of ATSC Transport Stream packets a deterministic or synchronous process. This is currently an asynchronous process in the A/53 standard. The current ATSC multiplexer produces a fixed rate Transport Stream with no knowledge of the 8-VSB physical layer frame structure and the mapping of packets into segments of a VSB Frame which will occur later at the transmitter site in an ATSC Modulator.

When powered on, the normal (8-VSB) ATSC modulator independently and arbitrarily determines which packet begins the frame of segments. Currently, no knowledge of this decision and hence the temporal position of any transport stream packet in the VSB frame is known to the current ATSC multiplexing system.

In the A-VSB system, the multiplexer known as an emission multiplexer has this knowledge and makes the selection for the first packet to begin the first segment of a VSB frame. This framing decision is then implicitly signaled to the A-VSB modulator via a periodic insertion of an Operation and Maintenance Packet (OMP) which has an ATSC reserved PID of 0x1FFA. The OMP syntax and semantics cause the A-VSB modulator to slave to the emission multiplexer for this framing decision, and start a VSB Frame with the selected packet in the first segment. The OMP is usually sent once every 20 frames (about 1 second).

In summary by changing this paradigm and allowing the Emission Multiplexer to make the selection of the starting packet and signal this decision to an A-VSB Modulator the emission multiplexer gains advanced knowledge or map of the temporal position that every byte of every packet will assume later when converted into a VSB Frame. This situation is shown in the Figure 2. Further, the A-VSB-enabled emission multiplexer works synchronously (Master/Slave) with the A-VSB modulator to perform some intelligent multiplexing. This then allows a calculated pre-processing of some selected bytes of a packet to occur in an A-VSB-enabled emission multiplexer and this can be designed to trigger a synchronous post-processing event in an A-VSB-enabled modulator. This core technique DF as shown in Figure 1 and is part of the basis for the extensibility and the application tools shown.

There is no negative impact to ATSC receivers from operating a station with DF and the authors believe this is a core asset of ATSC that can be leveraged to help bring value to future ATSC terrestrial services.

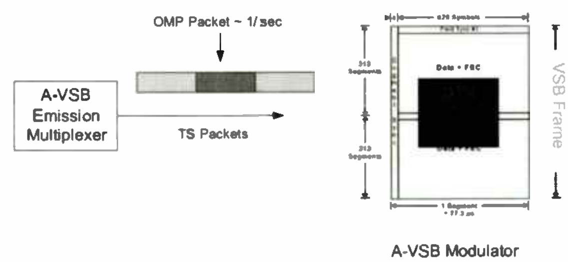


Figure 2 Deterministic Framing between A-VSB Emission Multiplexer and Modulator

Deterministic Trellis Reset (DTR)

The second core technique is the Deterministic Trellis Resetting (DTR) which resets the Trellis Coded Modulation (TCM) encoder states (the Pre-Coder and Trellis Encoder States) in the ATSC modulator. The reset signaling is triggered at selected temporal locations in the VSB Frame using the bytes deterministically inserted by Emission Multiplexer. Figure 3 shows (1 of 12) TCM Encoders used in Trellis Coded 8-VSB (8T-VSB). There are two new Mux circuits wired with reset added to existing logic gates in the circuit shown. When the reset = 1 is held for two symbol clock cycles the TCM encoder is driven to a deterministic zero state.

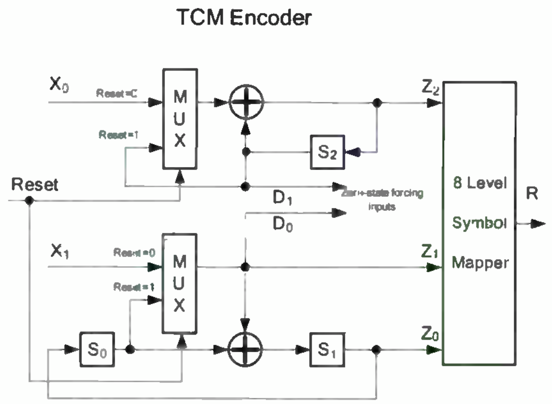


Figure 3 TCM Encoder with Deterministic Trellis Reset

The truth table of an XOR gates states, "when both inputs are at like logic levels (either 1 or 0), the output of the XOR is always 0 (Zero)." Note that there are three D-Latches (S0, S1, S2), which form the memory. The latches can be in one of two possible states (0 or 1). Therefore, independent of the starting state of the TCM, it is forced to a known Zero state (S0=S1=S2=0). Hence a Deterministic Trellis Reset (DTR) can be forced in a A-VSB modulator by post processing or triggering the reset= 1 over two symbol clock periods on select deterministic bytes that were inserted for this purpose by A-VSB emission multiplexer. When the Reset is not active (Reset = 0) the circuit performs as a normal 8-VSB TCM encoder. Figure 1 shows DTR as the second A-VSB core element that will be leveraged.

In summary the value the A-VSB core elements (DF, DTR) bring is operation of the ATSC system in a deterministic or synchronous manner at the transport and physical layers. Leveraging these core elements several new application tools are created such as SRS, Turbo, SFN and these tools bring new functionality for ATSC terrestrial services and will be discussed next.

SUPPLEMENTARY REFERENCE SEQUENCE(SRS)

An SRS-enabled ATSC DTV Transmitter is shown in Figure 4. The blocks modified for SRS processing are shown in gray (Mux and TCM encoders block) while the newly introduced block (SRS stuffer) is shown in dark gray. The other blocks are the current ATSC 8-VSB blocks. The ATSC Emission Multiplexer takes into consideration a pre-defined deterministic frame template for SRS which contains the mapping of the bytes in packets reserved for SRS. To maintain backwards compatibility these reserved bytes are placed in the adaptation field of ATSC transport packets. The generated packets are prepared for the SRS post-processing in an A-VSB modulator.

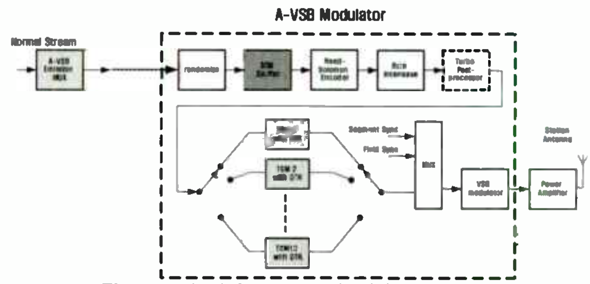


Figure 4 SRS featured ATSC Transmitter

The (Normal A/53) randomizer drops all sync bytes of incoming TS packets. The packets are then randomized. Then the SRS stuffer is triggered by knowledge of frame template used by emission multiplexer and fills the stuffing area in the adaptation fields of packets with a pre-defined byte-sequence, (the SRS-bytes). The SRS-containing packets are then processed for forward error corrections with the (207, 187) Reed-Solomon code. In the byte Interleaver, the bytes from the RS-encoder output gets distributed such that a 207 contiguous SRS-bytes are created. The unit of 207 bytes is the payload for a segment. These units are encoded in (12) TCM encoders. At the beginning of each SRS-byte sequence, the Deterministic Trellis Reset (DTR) occurs to initialize the TCM encoders this is possible because of the A-VSB deterministic nature and the A-VSB Modulator knowing the frame template that is used.

The TCM encoder states are forced to a known deterministic state by DTR at the beginning of the SRS-byte sequence. A pre-determined known byte-sequence (SRS-bytes) which is calculated based on this zero starting state of TCM encoder is inserted by the SRS Stuffer is then TCM encoded immediately. The resulting 8-level symbols at the TCM encoder output will appear as known contiguous 8-level symbol patterns in known locations in the VSB frame. This 8 level symbol-sequence is called SRS-symbols and is available to the receiver as additional equalizer training sequence. Since the SRS known pre-calculated sequence is inserted in the A-VSB modulator after the ATSC randomizer these 8 level symbols must have specific properties of a noise-like spectrum and zero dc-value which are SRS-byte design criteria.

Figure 5 shows the normal VSB frame on the left and an A-VSB frame on the right with SRS turned on. Each A-VSB frame has 12 groups of 8-level SRS-symbols. Each group can be 10, 15, 20 or 26 sequential data-segments and each data-segment will have 184 usable SRS-symbols. On MPEG-2 TS decoding, the SRS symbols appear in the adaptation field and will be ignored by a legacy receiver. Hence the backward compatibility is maintained.

Figure 5 shows 12 parallelograms in grey which have different composition depending on the number of SRS bytes. The actual SRS-bytes that are stuffed and the resulting group of SRS symbols are pre-determined and fixed. The 8-VSB standard has two

DFS each with Training sequences at a rate of 2 / frame. In addition to that A-VSB provides 184-symbol SRS tracking sequence per data-segment. There are optionally 120, 180, 240 or 312 such data-segments per frame depending on choice of SRS-10, 15, 20, or 26 respectively. More training sequences appearing frequently throughout a VSB frame help a new SRS receiver equalizer track dynamically changing channel conditions caused by motion of objects or people at a fixed location or conversely if the receiver itself is in motion.

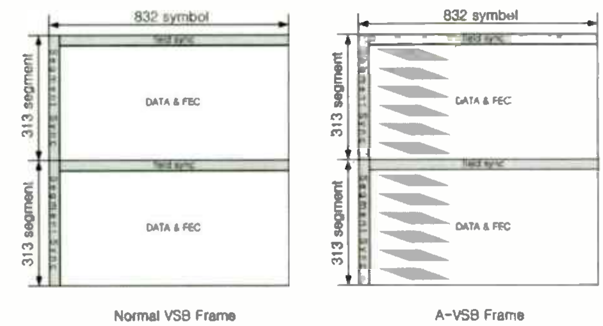


Figure 5 VSB Frame

Since these changes (DTR and that altering SRS-bytes) occur after Reed-Solomon encoding, previously calculated RS parity bytes are no longer valid. In order to correct these erroneous parity bytes, they are re-calculated. The old parity-bytes are replaced with re-calculated parity-bytes in the "TCM with DTR" block in Figure 4. The Turbo Stream post-processor in Figure 4 is active only when the Turbo Stream is used. Otherwise, it does nothing to change this process as the input just passes through to the output. The remaining blocks are the same as the standard ATSC VSB modulator.

Table 1 shows the payload loss associated with the number of SRS bytes per packet. Rough payload loss can be calculated as follows. Since 1 slice (52 segments) takes 4.03ms, the payload loss due to SRS-10 bytes is

$$\frac{(10+2)bytes \cdot 52packets}{4.03ms} \cdot 8 = 1.24Mbps$$

Table 1 Normal Stream Rate Loss due to SRS

SRS byte per packet	10 bytes	15 bytes	20 bytes	26 bytes
Normal TS Rate Loss	1.24 Mbps	1.75 Mbps	2.27 Mbps	2.89 Mbps

TURBO STREAM

Turbo Stream-supporting the ATSC DTV transmitter is functionally shown in Figure 6. The ATSC Emission MUX receives a normal stream and another stream, the Turbo Stream. After pre-processing of the Turbo Stream, the ATSC Emission Multiplexer generates a transport stream and encapsulates the Turbo Stream in the adaptation fields of TS packets compliant with a known Frame template for Turbo Stream The pre-processor expands Turbo Stream TS packets, for the selected 1/2 or 1/4-rate, by inserting the required placeholder bytes for later Turbo Stream post-processing.

In the modulator, the randomizer drops sync bytes of TS packets from a MUX and then randomizes the packets. The SRS stuffer in Figure 6 is active only when SRS is used. After being encoded in the (207, 187) Reed-Solomon coder the data stream is byte-interleaved. Then the Turbo Stream post-processor collects placeholder bytes (Turbo Stream data) reserved in the pre-processor and fills them with redundancy bits. The remaining blocks are the same as the standard ATSC VSB modulator.

An Emission Mux has to notify a modulator of the current Turbo Stream mode. The OMP (Operations & Maintenance Packet), extended to include Turbo Stream signaling, includes this information. The information about the current Turbo Stream mode is also conveyed to a receiver through signaling in the reserved area of the data field sync.

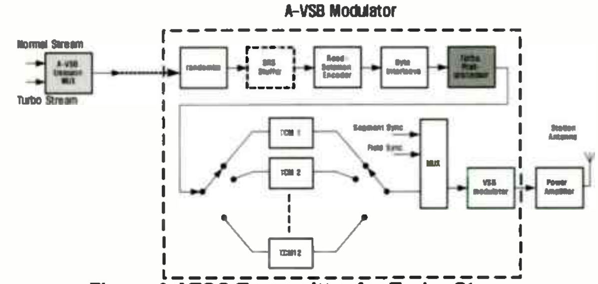


Figure 6 ATSC Transmitter for Turbo Stream

ATSC Emission MUX for Turbo Stream

ATSC Emission Mux for Turbo Stream is shown in Figure 7. There are three new blocks, Transmission Adaptor (TA), Turbo Pre-processor, and Turbo-packet Stuffer. An A-VSB Transmission Adaptor recovers all elementary streams from the normal TS and re-packetizes all elementary streams with adaptation fields in every 4th packet as one example of a Turbo frame template. This adaptation field space will serve as Turbo stream TS packet placeholders.

In the Turbo pre-processor of Figure 7, all sync bytes of multiplexed Turbo Stream TS packets are dropped. Then the Turbo packets are, RS-encoded and expanded by 2 or 4 according to the selected coding rate. The spaces created by byte expansions serve as placeholders to be used in the Turbo Stream post-processor in modulator. The Turbo Pre-processed Packet Stuffer simply cuts and pastes Turbo Stream pre-processed packets into the adaptation field of normal TA output packets, which results in the ATSC Emission Multiplexer's output.

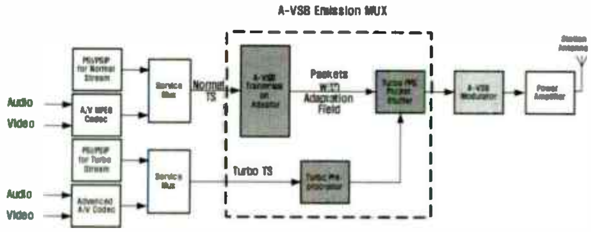


Figure 7 ATSC Emission Multiplexer

Modulator for Turbo Stream

As is shown in Figure 6, the first block in the modulator is a (normal A/53) randomizer. At first it drops sync bytes of an incoming stream from ATSC Emission Multiplexer and randomizes the stream. The SRS stuffer just passes an incoming stream to the next block when SRS is not used. The use of SRS with Turbo stream is synergistic and strongly recommended. The only added block in Figure 6 is "Turbo Post-processor" which is explained in the following.

.1.1 Turbo Post-processor

Turbo Post-processor works on only Turbo Stream data bytes. The other bytes just pass through Turbo post-processor. The block diagram of Turbo post-processor is as shown in Figure 8. Turbo Stream data bytes are extracted in the Turbo data extractor until one block of Turbo post-processing is collected. The block size for Turbo post-processing depends on the Turbo stream configuration. This block size is defined as the number of Turbo stream data bytes included in a slice (52 data-segments). Placeholders in the Turbo Stream data bytes are filled with the outer encoder's redundancy bits and Turbo Stream data is then interleaved in the Outer Interleaver bit by bit for one block of Turbo post-processing. The Turbo data stuffer puts the processed Turbo Stream data bytes into the space where Turbo Stream data are extracted.

Since the Turbo Stream data placeholders are filled here, the RS parity bytes attached in the previous block are no longer valid. These incorrect parity bytes are corrected in the Parity Correction block after the Turbo data stuffer in Figure 8.

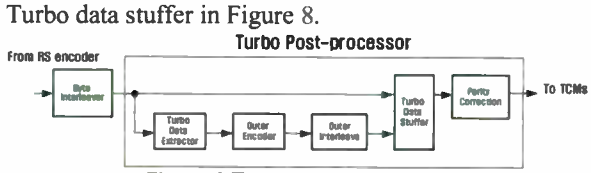


Figure 8 Turbo Post-processor

SINGLE FREQUENCY NETWORK (SFN)

When identical ATSC transport streams are distributed from a studio to multiple transmitters and when the channel coding and modulation processes in all modulators (transmitters) are synchronized the same bits in will produce the same RF symbols out from all modulators. If the emission times are then controlled these multiple coherent RF symbols will appear like natural environmental echoes to a receivers equalizer, and hence be mitigated and received.

The A-VSB application tool, Single Frequency Network (SFN) offers the option of using transmitter spatial diversity to obtain higher and more uniform signal strength throughout, and in targeted portions of a service area. An SFN can be used to improve quality of service to terrain shielded areas, "urban canyons", indoor reception environments or for future terrestrial, pedestrian or mobile services. Figure 9 shows a block diagram of a SFN.

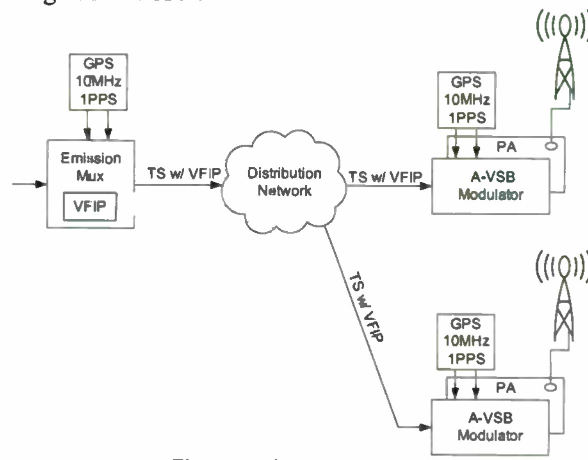


Figure 9 Block A-VSB SFN

The A-VSB application tool, SFN, requires several elements in each modulator to be synchronized. This will produce the emission of coherent symbols from all transmitters in the SFN. These synchronized elements are:

- Frequency
- Data Frame
- Pre-Coders/Trellis Coders

The frequency synchronization of all modulators' pilot frequencies and symbol clocks can be achieved by locking these to a universally available frequency reference such as the 10 MHz frequency reference from a GPS receiver.

Data frame synchronization requires that all modulators choose the same packet from the incoming transport stream to start or initialize a VSB Frame. This requirement is synergistic with the A-VSB core element Deterministic Frame (DF). A special Operations and Maintenance Packet (OMP) known as a VSB Frame Initialization Packet (VFIP) is inserted every 20 frames (~ 1/sec) as the last or 624th packet as determined by a frame cadence counter in an emission multiplexer. All modulators slave their VSB data framing when this special OMP appears in transport stream.

Synchronization of all pre-coders and Trellis Coders, known collectively as just Trellis Coders, in all modulators is achieved by leveraging synergy in the A-VSB core element Deterministic Trellis Reset (DTR) in a serial fashion over the first 4 segments in a frame after the DF has occurred.

The VFIP is shown in Figure 10 it has twelve (12) deterministic byte positions reserved for DTR to occur in all modulators.

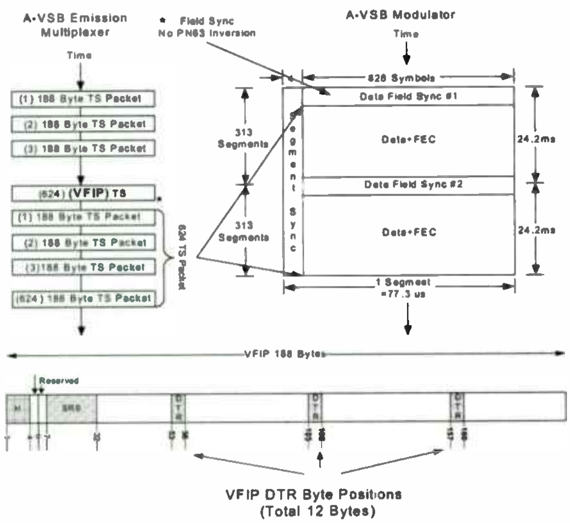


Figure 10 VFIP

The VFIP is inserted as the last (624th) packet once every 20 VSB Frames (~ 1/sec) as determined by the VSB frame cadence counter in an emission multiplexer. All the modulators initialize or start a VSB Frame by inserting a DFS with No PN 63 inversion after the last bit of the VFIP is processed. This action will synchronize all VSB frames in all modulators.

Synchronization of all Trellis coders in all modulators shall use the VFIP and the twelve DTR special reserved bytes. These shall be used to deterministically reset each one of the 12 Trellis Coders in each modulator to a zero state. This occurs in a serial fashion over the first 4 segments of the next VSB frame following insertion of the VFIP.

In summary the A-VSB SFN application tool leverages both core elements (DF, DTR) and the availability of GPS to meet the basis requirements for ATSC SFN functionality.

The VFIP also has syntax to control the emissions timing, etc from all transmitters in a SFN but this will not be discussed in this paper.

PERFORMANCE EVALUATIONS

A-VSB System Bed Setup

A FPGA prototype receiver was implemented to evaluate the proposed A-VSB system. The test system configuration is shown in Figure 11.

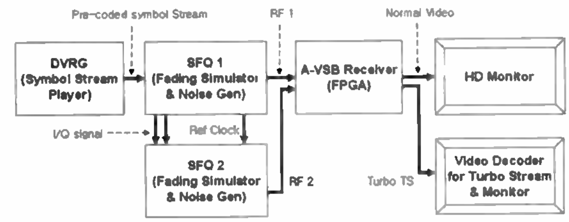


Figure 11 Test Bed for A-VSB System Evaluation

Since the Advanced VSB-supporting multiplexer and modulator are not available yet, a pre-coded symbol stream was generated by a program written in C. DVRG (DTV Recorder Generator) plays a pre-coded symbol stream and inputs it to SFQ1 [5]. In this way, DVRG simulates the signal from broadcast studios.

SFQ1 has modulator, fading simulator, and noise generator. It generates all the interference signals which are added up to the modulated signal. SFQ1 produces a final RF signal with all the interferences for the test [6].

SFQ2 is optional for the antenna diversity test. The baseband I/Q signals are connected to SFQ2 (slave) and it has an independent fading simulator and a noise generator. Finally, SFQ2 provides the modulated signal uncorrelated with SFQ1's output. Both SFQ1 and SFQ2 are required to set up the diversity test bed simulating uncorrelated multi-path phase and noise environment. A 10 MHz Reference clock is shared for the RF converters to maintain the frequency synchronization between both SFQs.

The FPGA receiver consists of an A-VSB demodulator and MPEG decoder for normal TS stream. The normal video output goes to a conventional monitor while Turbo TS stream goes out to external video decoder and a monitor.

System Evaluation Result

Several channel models are applied to an A-VSB receiver. In all simulated channels, the level of signal strength is adjusted to -53dBm which is a moderate signal level. The three channel models are considered:

1) Additive White Gaussian Noise channel
2) Modified CRC #3 (M-CRC3)
3) Typical Urban channel(TU6)

The second channel is mainly for a Normal stream test in the dynamic environment. The last channel is the dynamic 6-path Rayleigh-fading channel which is widely accepted as a typical channel model for the urban environment. The detail channel profiles are shown in the Appendix.

The two quantities are measured. One is TOV (Threshold of Visibility) and the other is the Maximum Doppler spread in Hz at SFP. SFP stands for the subjective failure point (SNR in dB) at which at least 100 packet errors are observed in 3 minutes.

The Normal stream improvement is achieved owing to SRS-helped equalizer tracking. No changes are made to the Normal stream packets except adding Adaptation Field.

The two symbol streams were generated in advance to test the 1/2 rate and 1/4 rate Turbo stream performance. In each symbol stream, the payload data rates of Normal stream and Turbo stream with the bandwidth efficiency are shown in Table 2. In the table, $Rate_{SRS}$ means the Normal TS Loss due to SRS-N bytes shown in Table 1.

Table 2 Normal Stream and Turbo Stream Details for 1/2 and 1/4 rate Turbo streams

	1/2 rate Turbo	1/4 rate Turbo
Turbo TS Rate	750Kbps	750Kbps
Normal TS Loss*	1.65Mbps	3.30Mbps
BW efficiency**	1.47 bit/Hz/s	0.74 bit/Hz/s
Normal TS Rate	17.65 – $Rate_{SRS}$	16.00 – $Rate_{SRS}$

* Normal Transport Stream Loss due to Turbo Stream

** Turbo stream bandwidth efficiency

The test results are listed in Table 3 ~ Table 7. In all tables, the Stream Type (rate) column indicates whether the TOV(or Doppler spread) measurement is made on the Normal stream or the Turbo stream (1/2 rate or 1/4 rate mode) in the A-VSB systems. In Table 5 and Table 6, 'Fail' means the failure even at a high SNR and 0 Hz Doppler spread.

As seen in Table 3, the Normal stream TOV is 15.1 dB. The SRS does not affect AWGN performance. The Turbo stream TOVs are 9.6 and 4.4 dB which are fairly low SNRs. The diversity reception technique brings roughly 3 dB TOV improvement in Table 4.

The usefulness of the SRS in the dynamic environment becomes evident in Table 5. The Normal stream in the dynamic channel (modified CRC#3) fails without the SRS.

Table 6 shows the Doppler spreads in TU channel at TOV. The Normal and 1/2 rate Turbo stream fail to be properly decoded while the 1/4 rate Turbo stream succeeds to. In addition, the assistance of SRS is obvious. This is not the case of AWGN channel.

From Table 7, the two antenna diversity technique resurrects the 1/2 rate Turbo stream in TU channel with the better performance than that of the 1/4 rate Turbo stream with a single antenna. The Doppler spread of the 1/4 rate Turbo stream in TU channel can go up to 155Hz (the carrier frequency of 475 MHz) which is experienced for example by a car running at 353 km/h.

Table 3 AWGN Channel TOV with Single Antenna

Stream Type (rate)	TOV1 (SNR) in dB				
	SRS- 0	SRS-10	SRS-15	SRS-20	SRS-26
Normal	15.1	15.1	15.1	15.0	15.0
Turbo (1/2)	9.6	9.6	9.6	9.6	9.6
Turbo (1/4)	4.4	4.4	4.4	4.4	4.4

Table 4 AWGN Channel TOV with Diversity

Stream Type (rate)	TOV (SNR) in dB				
	SRS- 0	SRS-10	SRS-15	SRS-20	SRS-26
Normal	12.0	12.0	12.0	12.0	11.9
Turbo (1/2)	7.2	7.2	7.2	7.1	7.1
Turbo (1/4)	1.7	1.7	1.6	1.6	1.6

Table 5 Modified CRC#3 Channel with Single Antenna at 5.2 dB Echo Power and 60 Hz Doppler Rate

Stream Type (rate)	TOV (SNR) in dB				
	SRS- 0	SRS-10	SRS-15	SRS-20	SRS-26
Normal	FAIL	FAIL	25.5	23.5	23.5
Turbo (1/4)	8.4	7.6	7.3	7.1	6.8

Table 6 Typical Urban Channel with Single Antenna at 25dBSNR

Stream Type (rate)	Doppler Rate at SFP (Hz)				
	SRS- 0	SRS-10	SRS-15	SRS-20	SRS-26
Turbo (1/2)	FAIL	FAIL	FAIL	FAIL	FAIL
Turbo (1/4)	45	50	50	55	70

Table 7 Typical Urban Channel with Diversity at 25dBSNR

Stream Type (rate)	Doppler Rate at SFP(Hz)				
	SRS- 0	SRS-10	SRS-15	SRS-20	SRS-26
Turbo (1/2)	55	60	70	90	110
Turbo (1/4)	125	135	135	145	155

CONCLUSION

In this paper, an enhancement to ATSC DTV standard was presented. The presented enhancement is a tool-kit based approach. The three application tools were proposed. One is an additional equalizer training sequence (SRS) and another is a robust Turbo stream. The SRS improves the ATSC-VSB performance in dynamic reception environment in backward compatible fashion. The last one is SFN functionality. To validate the proposed system, a laboratory test bed was constructed. The performance of Turbo stream and SRS under various channel models was obtained. Considering the difficulties of ATSC 8VSB in mobile-like channels, the test results are very impressive. The diversity reception, though optional receiver implementation, is presented here to illustrate the improvement it can deliver. There is still an improvement area in the system such as multiple Turbo stream support and power-saving techniques to name a few. These issues are being actively addressed by Samsung and Rohde-Schwarz.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Jaemoon JO and Mr. Dongil SONG for their guidance. They are specially grateful to Mr. Eui-Jun Park for an initial attempt to apply a Turbo code to ATSC VSB system. They are also very grateful to Yongdeok Chang, Donghoon Lee, Sungwoo Park, Chan-Sub Park, Jung-Jin Kim, Youhyun Choi, Teahoan Jung, Jun-Seok Kang, and Bo-Young Shin for their technical support.

APPENDIX

Table 8 Modified CRC#3 (M-CRC3) Profile

Simulator Parameter	Channel Simulator Parameter Value					
	Path1	Path2	Path3	Path4	Path5	Path6
Delay (us)	0	-1.8	0.15	1.8	5.7	35
Attenuation (dB)	0	14	14	4	Varie d to TOV	12
Phase or Doppler	0°	125°	80°	45°	60Hz	90°

Table 9 Typical Urban Channel (TU6) Profile

Simulator Parameter	Channel Simulator Parameter Value					
	Path1	Path2	Path3	Path4	Path5	Path6
Delay (us)	0	0.2	0.6	1.6	2.4	5.0
Attenuation (dB)	3	0	2	6	8	10

time varying profile	Rayle igh	Rayle igh	Rrayl eigh	Rayle igh	Rayle igh	Rayle igh
Coefficient PSD	Jakes	Jakes	Jakes	Jakes	Jakes	Jakes

REFERENCES

[1] S. Benedetto, D. Divsalar, G. Montorsi and F. Pollara, "Serial Concatenation of Interleaved Codes: Performance Analysis, Design, and Iterative Decoding," IEEE trans. on Information Theory, vol. 44, no. 3, pp 909-926, May. 1998.

[2] S. Benedetto and G. Montorsi, "Iterative Decoding of Serially Concatenated Convolutional Codes," Electronics Letters, vol. 32, no. 13, pp 1186-1188, June 1996.

[3] Jungpil Yu, Marie-Laure Boucheret, Robert Vallet, Alban Duverdier and Gilles Mesnager, " Interleaver Design for Serial Concatenated Convolutional Codes," IEEE Communications Letters, vol. 8, no. 8, pp 523-525, August 2004.

[4] ATSC A/53D, "ATSC Standard: Digital Television Standard (A/53), Revision D", Advanced Television Systems Committee, Washington, D.C.

[5] Product Data Sheet, *DTV Recorder Generator (DVRG)*, Rohde & Schwarz, 2004, Available in `www.rohde-schwarz.com`.

[6] Product Data Sheet, *TV Test Transmitter (SFQ)*, Rohde & Schwarz, 2004, Available in `www.rohde-schwarz.com`.

Engineering Management for the 21st Century

Wednesday, April 18, 2007
3:30 PM – 5:00 PM

Chairperson: Steve Davis
Clear Channel Radio, Tulsa, OK

"Media Facilities in Transition" Change Management Planning and the Mission Critical Media Facility

Stephen Newbold, Gensler, Boston, MA
Benjamin Coe, P.E., EYP Mission Critical Facilities, Atlanta, GA

Good Asset Management Equals Smart Scheduling

Andrew Janitschek, Radio Free Asia, Washington, DC

The Capital Budget Plan; A Tool for the Chief Engineer

Mike Seaver, Seaver Management and Consulting Services, Quincy, IL

"MEDIA FACILITIES IN TRANSITION" CHANGE MANAGEMENT PLANNING AND THE MISSION CRITICAL MEDIA FACILITY

BEN COE, PE, LEED AP
EYP Mission Critical Facilities
Atlanta, GA

STEPHEN R. NEWBOLD, AIA
Gensler
Boston, MA

The continuous and fundamental flux in the media industry forces companies to look beyond a series of "digital upgrades" to question the need for a full "business facility upgrade". We will be examining the process for realigning real estate and facilities to support your new media business models. This discussion is aimed at providing you with a variety of ways of looking at your facilities and the effect changes have on your mission and bottom line.

By "facilities transition," we include the possibilities of redesigning your work environment whether by down sizing, upsizing, renovating or relocating your facility to better suit your needs and goals. Many factors go into this decision process due to the mission critical nature of these facilities. Knowing when to make a change is difficult. But we all realize that at some point we will need to do a facility upgrade. The trick to the transition is to capitalize on what may be seen as a series of concurrent, routine "upgrades" to create a "business transformation" which will create new opportunities.

Recently, the full promise of the digital technology transition and its related markets has emerged as a driving force in changing the nature of the traditional broadcast and "content creation" facilities and businesses. At the heart of the change is migration from tape based workflow to server / networked content work flow and the move from analog to digital delivery. This change is wide reaching and has affected existing single purpose facilities as they look to move to multi-format and multi-channel capability. It has also created the need for a variety of hybrid co-location facility solutions. These include duopolies, university/ cable station alliances, sport network/ venue co-locations, facility rental income solutions etc. This digital "upgrade" with it's the related workflow change is the watershed which has allowed many firms to investigate and realize a significant transformation in their businesses, facilities and mission critical infrastructure.

Some examples of companies, which have made such a transition:

- A catholic diocese merges it's broadcast facilities (4 cable channels) with an affiliated college communication department (2 radio and 1 cable channel). Independently, the acquisition of new facilities was difficult, combining, they shared costs, acquired greater educational and work training environment and expanded the college's media reach and distance learning capabilities.
- A duopoly is agreed upon and one 50 year old news facility is retired and the other station's digital ready facility is expanded with new studios and work space to house the combined organization. The "digital" conversion of the second station is greatly simplified by moving into a building with strong infrastructure and a trained digital engineering staff. The older station buildings are sold for redevelopment, paying for a substantial portion of the relocation.
- A local station desires to upgrade their facility, and recapture under utilized building area for new production facilities. A careful multi-year plan results in the first major upgrade in 40 years. The process changes workflow, editing practices, realigns staff to proper work environments, removes all the hazardous materials in the building, upgrades the building systems and recovers 20% of the building by replacing oversized analog technical spaces with new departments and digital production environments.
- A cable network realizes that they are land-locked with no growth potential. They relocate and convert to full a digital production environment. They move from a small multi-floor windowless building to a former factory building, which allows them to locate all the staff on one high floor with an open floor plan.

The entire facility now has abundant daylight, employees love the new environment and the station was able to grow to a multi-language facility due to the planning for growth, re-bust infrastructure provided and the flexible work environment.

- A sports network located in a sports facility realized their growth within the facility had created a complicated and cramped facility, which was unable to convert to their new business mission. A planned relocation freed them from years of poor "legacy" planned decisions and tripled their production facility capacity without considerably expanding their overall space requirements by reducing corridors, updating antiquated facilities and repurposing inefficient spaces.

All of these clients also reaped a significant change in moral and human resources performance, not to mention technical capacities, and brand enhancement through the creation of new and improved facilities. The goal of this illustration is to make you aware of the potential that changes in facility realignment can have on your business. Going "digital" should not be thought of in isolation as an engineering directive, but an opportunity to achieve multiple business goals and opportunities.

"IT IS MORE THEN EQUIPMENT"- PEOPLE MAKE IT HAPPEN

Concurrent with technical change is a social and business change. The basic business models are shifting and competition is intensifying. While attendees at NAB are often told that their competitive edge is dependent on the technology they use, you have to look farther. One of the highest pressures on firms is staff recruiting and retention. As we move toward a fluid multi-generational workforce, the work place and employee satisfaction becomes "mission critical" as well. Facilities being re-examined. Are they well designed, safe, attractive, productive environments suitable for workflow change and provided with robust infrastructure? Studies have proven the appropriate work environment has significant effects on employee retention, productivity and overall job satisfaction. This is even more important in deadline driven, technical and creative environments. Spending money on new equipment and giving it to unhappy employee in poorly designed environments is not a wise investment. It is even worst in facilities lacking reliable building systems, power and infrastructure.

WHAT CLIENTS ARE LOOKING FOR

With the advent of digital technology, a flexible work flow is now possible, with it a much wider variety of work place environments which are better able to meet employee and business goals can be achieved. Critical factors we look to deliver when working with media clients include higher levels of staff communication, an increase in "teamwork" environments, a higher level of labor relations / satisfaction and productivity, greater air quality, better lighting, access to daylight, access to amenities, flexible workplace configurations, better utilization of real estate assets, increase in engineering system reliability and implementation of healthy work environments including removal of toxic materials.

On the engineering side, planning, proper redundancy provisions, N+ systems, disaster recover procedures, and well documented systems are all key to the minimum building and engineering systems facility design criteria. Digital equipment, while able to create a wide variety of new opportunities is still dependent on a high level of power and control reliability. When disruptions happen the outage and "reboot" of the system is extensive. Corporate demands for increased levels of reliability have mandated a complete rethinking of what had been fairly standard stand alone systems.

Bottom line is there are far more things to consider then just the basic idea of "going digital". You would not "digitalize" an old analog process, why would you "digitally upgrade" an antiquated work environment and plant. Look at all the factors, which contribute to the success of a "business upgrade".

BUT HOW TO CHANGE...

Hopefully you are starting to see the reasoning for concurrent physical change. But, we also understand that mission critical facilities are not easy to work within due to the "24 X 7" nature of the operations. There are very few opportunities to create significant down time for anything but the most critical changes. The result is facility stagnation. As the digital equipment demand has increased, so have the facility challenges and costs. As we migrate from single purpose TV facilities (local stations, cable production companies, news organizations etc) to more multi-format, multi-channel facilities (co-location, consolidations, alliances, new media facilities) this logistical challenge increases. Few companies initially want to consider expanding their facilities when challenged to expand the business so they add more " of the same" to the present facility-more equipment, more heat, more wire, more productions etc. The result is lower quality work environment, loss of support functions and amenities, more stress, higher staff turn over, inadequate, under rated engineering infrastructure, and increased business risk. At some point you hit a threshold either of organizational pain or corporate risk or both.

SHOULD I UPGRADE OR MOVE?

We are often asked to investigate for clients the process required to upgrade them and fix all the problems in an existing facility while they occupy it. On average, it takes approximately 3-4 times as long to remodel a working technical facility "in place" then it does to build a new one. When many clients seriously look at the complexity associated with wholesale upgrading an operating facility while remaining "on air", they often decide to look for a "new" locations to start again. Building a new facility creates opportunities while eliminating the constraints which would of limited your possible success if you remained in place- poor location, poor building configuration with antiquated space planning, legacy or inferior infrastructure, potential risk to on going operations.

We encourage clients to consider all aspects of their facility investment during the decision process. There are several factors, which should trigger a study of possible relocation/ restructuring. Each of these items should be tracked on a regular basis.

- Lease expiration/ potential rent increase
- Facility quality deterioration/ major maintenance expenditure needs
- Infrastructure upgrade required (aging equipment, loss of reliability, poor energy efficiency, poor connectivity), substantial increase in risk to ongoing operations.
- Business change- change in business mission, workflow, product delivery, business size, alternate business opportunity
- Staffing issues-staff retention, staff happiness/"wellness"
- Lack of growth potential ("land locked")
- Location- change in business environment, access, security, expansion options, amenities, utility connections, tax / incentive opportunities

Any and all of these weigh in to the decision to consider relocation. Note, I did not list new technology upgrade. If your basic financial and staffing goals are flawed, technology will not fix it but will make it worse. In sure your "solution" considers the effect on all aspects of the business.

BEST PRACTICES

Whichever way your business decides to perform their facility change, in place or by relocating, we wanted to share with you some important technical planning and execution lessons learned over the years.

Here are some pointers if decide to renovate instead of relocating.

1. Create a master plan, which addresses all of the major business decisions and desired end results. Think "blue sky" to get all the issues on the table. Confirm what motivates your business and what could defeat it.
2. Build your new work on a solid foundation of mission critical infrastructure. Try to isolate new work from older legacy systems. Remember that shutdowns to existing infrastructure are costly and a potential risk factor to be avoided
3. Make a detailed "sequence of operations" phasing plan. This should include temporary relocations, infrastructure upgrades, construction, shut downs equipment installs and re-occupancies. Remember temporary infrastructures required during the transition..
4. Select consultants and contractors who are accustomed to working in mission critical environments. This results in better planning, fewer surprises and lower project risk. Select in house staff who have the time to dedicate to the project. Nothing slows a project more then indecisive, part time client participation. We are looking for long term planning and management skills.
5. Consider your equipment planning carefully. Control rooms are cooler now but racks are hotter. Keep in mind that whatever equipment you select it will be replaced and changed. Build the flexibility in to your solutions. Spare capacity is much less expensive now then later. Rack electrical usage and heat densities have increased and cooling methods need to be matched to equipment spacing and configurations. . Take advantage of the potential opportunity to improve your infrastructure reliability and redundancy. Power quality equipment is more critical then ever in the digital arena. Improving energy efficiency of HVAC systems is becoming of more interest as prices increase.

5. Be informed: Understand everything that is to be done to your building engineering systems and why.

6. Keep everyone else informed of the renovation process. Notify your employees; keep all your managers informed of schedules, shutdowns, work disruptions etc. Keep your contractor informed of potential disruptions to his/ her work - unusual production schedules, "sweeps" related workload changes, special events etc. Put them in the project schedule upfront.

6. Commission it! Test it! Don't take some one's word that it will work in an emergency. Prepare your disaster recovery plan. Train your staff now, not after an emergency.

BEST PRACTICES FOR RELOCATION SOLUTIONS

1. Spend time to rethink. Ask a lot of questions of your business and you needs. Seek an unbiased facilitator to help you with this. Talk to your consultant about investigating your options and defiantly make careful plans before making any commitments.

2. Fully understand the effect of the digital equipment and the new content creation process. It empowers more people to do more with less facility. It also is much more infrastructure dependent.
3. Choose your site carefully; don't be lead around town by a real estate broker who does not understand all the implications to your business or legal constraints. Have your consultants evaluate all potential sites. Don't chose a site for the "unusual parts" of your need (High bay studio space, existing generator, truck parking etc) Look for site for your people and office needs first, this is probably the largest part of the space need. Insure there is land so that you can add on the special parts. This gives you more latitude and allows you to build to your unique needs.. (Remember-location, location, location. This works for price as well as employee recruiting, business relations, corporate and community image, ect.)
4. Look at possible synergies between your business and others. Consider co-locating options. Spend time to understand the effects on your brand, operations, cultures and growth prospects. Consider the benefits of being on a campus environment with other activities.
5. Remember to review surrounding microwave transmission conditions and constraints.
6. Define your mission critical design criteria to insure the facility is "hardened" to protect your operations-roof construction, relationship to the flood plain, multiple power sources, substantial building envelope construction, appropriate security provisions, hardened connectivity, generator provisions (including significant fuel storage)
7. If you are locating in a multi-tenant building, be aware of your own destiny. Craft lease terms to protect your special facility needs and insure that the activities of other tenants cannot disrupt your operations.
8. All of the above renovation issues relate to building a new facility as well.

After the engineering considerations are understood, remember that you need a superior building design and interiors to accommodate all the "soft" staff and business goals that will make or break your organization.

Great design complements great engineering to deliver superior results. Look at the Ipod.

In closing, the work we undertake in support of the media industry's change management process is both complicated and challenging, but with careful planning and an eye to look for opportunities for holistic change, the process can be much more rewarding then you may initially think.

GOOD ASSET MANAGEMENT EQUALS SMART SCHEDULING

ANDREW JANITSCHEK
RADIO FREE ASIA

ABSTRACT

Demanding work schedules are a fact of life in our society today; this is highlighted by broadcast stations the must operate 24-hours a day simply to serve their audience properly. The work schedules required for operations like this can create quite a strain on the entire staff and can affect a worker's safety and health. Mistakes by an employee that is tired can also affect the public's safety and health.

DEAL!

How does one deal with the stress of being in a difficult profession like police officer or working the graveyard shift in a 24-hour broadcasting station? The reality of these professions is shiftwork and the related stress. If this comes as a surprise to anyone entering the career field, then they should run away from this work as quickly as possible. We can assume anyone entering either of these, or similar careers, must know to expect some shiftwork and realize there are other challenges involved.

In our profession, many cannot avoid night or rotating shiftwork. Without invading anyone privacy, we will examine some ways today's supervisor can work with their people to ensure that the work schedule and work environment is minimized as a contributing factor in any health or safety issue. A night-shift worker's concerns may begin with safety when entering the building at midnight and end with falling asleep at the wheel while driving home. In between are problems with alertness, social isolation and physical distress.

According to the Bureau of Labor Statistics, in 2001 almost 15 million Americans worked evening shift, night shift, rotating shifts, or other irregular schedules. In 2003, The International Labor Office reported the amount of hours worked in the United States exceeded Japan's and most of western Europe. Both shift work and extended work hours have long been associated with health and safety risks.

In the May, 2005 Bureau of Labor statistics survey, there were almost 1.7 million employed in the arts, design, entertainment, sports and media industry. The industry groups are based upon the Standard Occupational Classification system which is used by federal agencies to classify workers into occupational categories in order to collect, calculate, or disseminate data. Everyone in the workforce can be classified into one of hundreds of occupations according to their job. To make this system easier to work with, there are 23 major groups, 96 minor groups, and 449 broad occupations which cover all Americans. Each broad occupational group includes detailed occupations that have similar duties, skills, education, and/or experience. In the same , we found that the mean annual salary of our major occupational group was $44,310; averaging a 3 percent increase per year; that would give us an annual salary of $47,008 in 2007. We can safely assume there will be a difference in annual earnings by a master control operator in Ottumwa, Iowa and someone working in Manhattan. By the way, some of the other professions also grouped with us include, painters, sculptors, window trimmers, actors, umpires and referees, and public address system announcers and fashion designers.

In many of the industries mentioned above, the work is performed generally on weekdays; yes, some weekends are included, but few are limited to round-the-clock work schedules; radio and television broadcasting are some of the professions.

TO YOUR HEALTH

Shiftworkers and night workers are often tired due to their work schedule. Being tired makes it difficult to concentrate, which increases the possibility of errors or accidents. This can be a risk both to the worker and to the public. The stress of shiftwork can also aggravate health conditions; such as heart disease or stomach disorders.

Working at night makes it difficult to get enough sleep. Sleep after night work usually is shorter and less refreshing or satisfying than sleep during the normal nighttime hours. Brain and body functions slow down during the nighttime and early morning hours. The combination of sleep loss and working at the body's low point can cause excessive fatigue and sleepiness. This not only makes it more difficult to perform well, but it increases the risk of accidents. Also, shiftwork can be stressful because of frequent switching from a day to night schedule and because of separation from family and friend. These stressors can be harmful to anyone's health.

According to the Bureau of Labor Statistics Census of Fatal Occupational Injuries, over 13,300 workers died in roadway crashes between 1992 and 2001. If we did the math, we would find that approximately four people, per day, died from these accidents. Of all causes for work related deaths, roadway crashes; this represented 22 percent of all the deaths, and was the number one cause. In 2000 alone, it cost society $61 billion in lost wages and benefits for crash victims in work and non-work related incidents. An additional $4 billion in costs was passed on to employers in the form of lost productivity. For employers and victims, a workplace crash can have additional financial, medical, and legal consequences for year to come.

Shiftwork involves most any work performed outside normal daylight hours. In a 1997 publication from the Department of Health and Human Services called "Plain Language About Shiftwork," daylight hours are generally considered to be from 7 a.m. – 6 p.m. where one has to work a 7-8 hour shift, though many in Alaska would argue that with you. Those working a shift may work in the evenings, late nights, early mornings or a combination of them. Workers might also work regular days at one time or another. Many shiftworkers simply rotate around the clock, constantly changing work times from day to evening or day to night. The rotation may occur daily, weekly or monthly and is a classic scheduling technique still used today by police, firefighters, and the military. Waiters and waitresses may be assigned to a permanent shift but can only work a night, while a night watchman may only work the graveyard shift. The need for shiftwork is understandable and if we are in the position of scheduling our staff to cover many shifts, it is important to consider the needs of the company against the needs of the job at hand and the needs of the workers. It's a paradox that since there are so many more shiftworkers these days, we also need more shiftworkers to provide services that were once only available during the day.

ANY VOLUNTEERS?

Why would an employee want to work on a shift? Some actually prefer non-day work, but most do not go looking for shiftwork. Some of the reasons for choosing shiftwork include better pay (night differential), more time during the day for child care, more daylight for personal pursuits, and more time to attend school. Some prefer shiftwork because there is less activity at work and fewer managers around; nonetheless, most of those workers would not choose shift work but do it as a requirement of the job.

Personnel and Human Resource managers are constantly asked, "What is the best schedule?" There is no perfect schedule that fits every situation. The most common schedule is five consecutive days of work followed by two consecutive days away from work for a traditional weekend break. In a rotating schedule, the employee will normally start a new shift after their weekend break. It is not unusual for workers on offshore oil rigs though to work 14 days in a row then return home for a 14 day break. With each schedule comes its good and bad points; it is always hoped the good points outweigh the bad and that there is little affect upon safety, health and productivity.

One technique many managers and supervisors use is 'creative scheduling' in order to ensure there is enough staffing for every daily requirement. Creative scheduling is simply another name for looking beyond the norms and coming up with a way to satisfy the needs of the company and the employee, especially when extraordinary circumstances force scheduling outside normal methods. For example, scheduling weekend coverage with a full time employee working 10-hours a day for 4 days with a 3-day weekend, or even three 13-hour days; there are many mixtures of schedules that can be accommodated when both employer and employee are willing to work together. These shorter work periods are also examples of compressed work weeks. A warning about creative scheduling though: About three years ago a company had an employee work a split shift for a number of months. With the employee's approval, they worked 4 hours in the morning and 4 hours in the evening. The shift was scheduled and went on for months. Soon the employee changed back to a full 8-hour shift daily. Unbeknownst to the employee and the employer, labor laws guaranteed the employee extra pay for each day worked on the split shift. Years later, when the problem was realized, the company paid the extra money.

To entice employees into working unusual, or off-hour shifts, many companies use shift differentials, or shift premiums. Most companies with non-stop operations generally rely on three shifts to maintain on-site staffing. Employees working the overnight shift, or grave shift, are the most likely recipients of extra money per hour for working the least popular work hours. Some of the normal questions that must be asked before establishing any shifts are:

- Will there be overtime allowed?
- How long should each shift last?
- How much time will be allowed between shifts?
- How much rest will be allowed during shifts?
- How many days will be worked before a weekend break?
- How many days will comprise the weekend break?

Some ways to improve the morale and health of the crew are:

- Provide time for exercise while at work.

- Have extended hours of service for HR and Finance.
- Make sure healthy snacks are available in vending machines.
- Subsidize extended hour child care.
- Subsidize secure parking when possible.
- Allow music and television during breaks.
- Eliminate shift premiums by increasing base pay.
- Be tolerant and flexible during family emergencies.
- Keep security personnel on site.
- Install emergency call boxes in secluded areas.

In their most recent report called Workers on Flexible and Shift Schedules, the US Bureau of Labor Statistics said that as of May 2004 over 27 million full-time wage and salary workers had flexible work schedules letting them vary the time they started or ended work. These workers made up 27.5 percent of the full-time work force, which was down from 28.6 percent three years earlier. The proportion who usually worked a shift other than a daytime schedule remained close to the 2001 level of 14.8 percent. With the increase of shiftwork needed for production and support services, these numbers are expected to jump. In the past, it was safe to claim that shiftworkers were generally limited to the manufacturing, mining and transportation industries. Now shift workers are found in customer service call centers, retail establishments, information technology monitoring and support centers, hospitals, hotels, casinos, emergency response services, and broadcast operations.

As the need for extended operating hours grows, more and more organizations are adopting work schedules that require longer and/or multiple shifts. If you find yourself faced with this problem, ask for assistance from your Personnel or Human Resources Department; they can help develop and implement any new work schedules. With their support, most schedules will see few problems.

Shift Length

Understandably, the majority of employees prefer to work shifts longer than eight hours in order to get a longer weekend. The benefits can be a substantial morale booster. For example, in a company that must run a 24/7 schedule, the employee that works five 8-hour shifts willing receive on average 91 days off each year. Those working four 10-hour shifts receive 146 days off per year, and those working 12-hour shifts will normally receive 182 days off per year. On average, 10-hours shifts are preferred by workers because they are easier to tolerate than 12-hours shifts while still allowing for more time off over the 8-hour work schedules.

While 10-hour shifts are great for the employee, they do not make good management sense for any 24/7 operation. When scheduling people on 10-hours shifts it takes three shifts to cover an entire day but in stead of 3 shifts of 8-hours each to cover the 24-hour day, 3 shifts of 10-hours each have workers on-site for 30-hours daily; an overage of 6-hours of double coverage per day. Assuming the workload is static throughout each workday, the overlapping hours can very well constitute wasted manpower and in the long run, cost a great deal of wasted money. As long as the workload of a 10-hour shift can fluctuate to accommodate increased staffing, then it may make very good sense, but the solution is to schedule overlaps only when work levels fluctuate to make the staffing levels match the workload. This will increase customer service levels and boost morale as staffing levels will consistently meet the demands placed upon the workers. An option to consider is possibly putting a few select people on 10-hour shifts, especially for those that must work every Saturday and Sunday as a perquisite to missing out on many of the standard weekend activities we take for granted.

If you are working in a company with consistent work loads, then consider 12-hour shifts over 8-hour shifts. There are not many that will enjoy the long work days, but that is usually forgotten when they realize how much free time they will earn. Before you launch into a new 12-hour work schedule, there are some considerations.

As long as total numbers of hours worked remains the same, most work can be accomplished as well on either shorter or longer shifts. Some jobs involving a great deal of tedium are best left to shorter shifts.

If the shiftworkers will be exposed to extreme heat, noises, or other physically draining work, it may be too much even for the best worker to endure for more than 8-hours.

Lastly, consider what your specific staff prefers. Longer shifts are generally preferred, but they may not be acceptable to everyone. Help yourself and the employees by helping them understand what will be expected of them and what shifts are available. The informed worker with the ability to have some influence on their future work schedule will be a happier worker.

THIS REQUIRES SHIFTWORK

You have considered the information above and realize that your new schedule requiring shiftwork is the best way to go. Will you stick with stable schedules or rotate people? If you rotate, will it be a forward or backwards rotation? With a stable schedule, the hours and days worked will remain predictable. An employee can remain on the shift and yet work different days of the week in order to fairly distribute the work on weekends and holidays too. This could be considered a lite version of a rotation schedule, but it does let the employee maintain a more regular schedule. Certainly, employees prefer a static, predictable shift schedule, but sometimes that is not possible, especially when trying to fairly schedule the rest of the staff. The advantages of stable shifts are:

- Every facet of the shift is learned by all.
- Employees are happier, thus morale is higher.
- Greater team building with a known group of co-workers.
- Where one shift has more work than another, the work, or the shift personnel can be redistributed to equalize all things.

The disadvantages of stable shifts are:

- Shift ownership and turf wars.
- Apathy due to familiarity.
- Personal and professional growth stymied due to a lack of challenges.
- Those on the overnight shift can feel forgotten.

The other point to make here is there are definite benefits to shift rotation, as there are deterrents. Many companies manage their productivity and operations quite well with a rotating shift schedule, while others would never consider a rotating schedule. The choice is really up to company management and what seems to work best for the company the industry and the employees. This is a good point to note that so much of the research available about rotating schedules and shiftwork leans towards providing work hours that are best for the employee. The bottom line is if you can keep the employee happy in the work place and in their personal life, then they will likely produce far and above the standard of the person who is unhappy. Schedules that are selected as part of a process where employees are involved are likely to remain in place nearly twice as long as any duty schedule that is dictated by management. The key to a successful partnership between labor and management is a collaborative labor relations program based on mutual trust and respect. These principles apply in commercial business and in government today. With rotating shifts, personnel move from on shift to another. The advantages of rotating shifts are:

- A balance of expertise on each shift.
- Everyone gets equal exposure to the most desired shifts.
- Limited training resources can be used during one shift that will eventually touch all personnel as they rotate.

Many companies go the extra mile by holding training sessions during every shift so it is accomplished during normal hours of the shiftworker and not on overtime. Some real benefits of this gesture is that workers will be more attentive and alert during their shift and will also view this effort by the company in a very positive way. When its impossible to provide training during the shift, at least schedule it when the shift begins, that way you are not trying to train workers at the end of a shift when they are their most tired.

Set The Pattern

So you have decided on whether rotation is best for your company; now you need to consider what the standard 'days on' and 'days off' pattern will be for the staff. Most prefer to work Monday through Friday with Saturday and Sunday off. With shift work, this can be extremely hard to accommodate. There are many different variations, and while the standard is 5 days-on and 2 days off, the preferred method is 4 days-on followed by 3-days off as it provides one more day of 'rest and relation time' in the employee's mind. When it comes time for you to choose the pattern of days to work for the crew, consider when they will have their weekends, give consecutive days off if at all possible, which in turn means working people for consecutive days. Working for a long stretch of time, followed by a one day break may not be acceptable to anyone; be flexible and work with people as much as possible.

STRESS AND SHIFTWORK

The stress of shiftwork can be reduced by movement from one shift to another. Given the regular chance to move off of an undesirable shift will help reduce the hours spent anxiously worrying about negative aspects of any shift. Conflicts between work and non-work related duties can be addressed through flexible scheduling or rotating schedules. Given this flexibility, people assigned to the least desirable shifts will know they can leave the shift eventually and have a fair shot at working the shift, or shifts, they prefer. As someone that must create a duty schedule for a 24/7 operation, it is important to note it is impossible to make everyone happy; nonetheless, efforts must be made towards accomplishing just that!

Common sense tells us that with any schedule, especially one requiring shiftwork, all must be treated fairly with no discrimination or favoritism. Determining what is best for the company and the employees leaves many questions to be answered. Which type of schedule is truly best for the organization? Is it truly the best for all the workers? If not, which is?

In the workplace, there are finite resources available to complete the work at hand; they are personnel, money, materials and machines, and time. The one resource difficult to quantify is that of personnel and the specific skills each employee possesses. How they are scheduled and used on the job will have dramatic effects upon the success of operations and the employee turnover rate. When faced with a 24 hour operation, someone must be working around the clock; how those personnel are scheduled can result in dramatic outcomes. For a clearer understanding, we must recognize some of the challenges and hazards of shiftwork.

Shiftworkers are often tired and sleepy because of their work schedule. Being overly tired makes it difficult to concentrate, which increases the possibility of errors or accidents. This can be a risk both to the worker and to the public. The stress of shiftwork also can aggravate health conditions, such as heart disease or digestive disorders.

Working at night makes it difficult to get enough sleep. For a shiftworker, sleep is usually shorter and less refreshing than sleep taken during normal nighttime hours. Additionally, brain and body functions slow during the nighttime and early morning hours. The combination of sleep loss and working at the body's low-point can cause fatigue and sleepiness. This makes it difficult to not only perform well but it also increases the risk of accidents. For anyone that must work a rotating schedule, shiftwork can be even more stressful because of frequent switching from a day to night schedule. When you add to that the inherent separations from family and friends, you have the makings for a greater number of stressed employees.

One point to consider with shifts is ownership, or turf wars. This is protection of one's territory within an organization, whether this territory is perceived or real. Allowed to stay on one shift, it is only natural the employees will bond and work to protect their work zones. It is not unusual for many to identify with their coworkers and to pride in the identification they give to their crew, but it can also lead to arguments with the people from other shifts. There is a line from the old Mac Davis song that goes: "Walk a mile in my shoes; before you abuse, criticize and choose; walk a mile in my shoes." This axiom can be applied to work schedules too. You cannot empathize with, or understand what another group is doing or thinking until you experience it for yourself, right? One way to solve the problems between different work crews is to rotate them so each benefits from the different hours. To many, the clear answer is to rotate individuals or entire teams. In this case, the person setting the schedule must also help set a schedule of when and where people work.

Absenteeism costs companies more and more each year. If an employee only makes $7 per hour and there are 99 other employees in the same situation, it costs the company $33000 a year to give them all 6 days of sick leave. It doesn't take into account the cost of other benefits, the cost of contractor or temporary help, overtime pay, lost productivity, impact on morale, and more. Absenteeism overall is on the rise. To keep this under control, companies must realize that helping workers stay satisfied with their jobs is of the utmost importance.

FORWARD OR BACKWARD?

How to handle rotating schedules? First of all, all humans have a built in tendency to resist change. When we feel we have reached a level of competence and confidence with our work, we want to hold onto the position; to change is to risk becoming less competent. Change not only causes us to lose the routine we have come to know, but it can also be a bit frightening to move onto something new. One of the worse things a scheduler can do is to make a significant change in someone's life and give them little or no time to adjust.

To start a rotating schedule, plan it out, know what schedule every person will move to next and do not forget to give appropriate breaks to those going from an extreme shift to another. Consider whether moving forward or backwards will have any negative effects too. When announcing changes, give the staff at least 3-months advanced notice; 90-days to adjust one life is certainly better than anything less. If you care to make adjustment to the change easier, then make the new schedule effective after 6-months. Again, people will still be shocked at the change, but you have now given them months to mentally and physically prepare for their new routine. Also provide employees sufficient outlets for voicing their concerns and lastly, remain flexible where possible. Concerns that go unexpressed lead to resistance that is difficult to overcome. For example, someone may not be able to work a shift due to the financial burden caused by the costs associated with finding a babysitter.

In general is it best to rotate backwards; the staff on the overnight shift starts earlier in the day on the mid-shift; the mid-shift moves backwards to the dayshift and the day shift crew moves backwards to the overnight shift. Studies have shown the backward rotation is easier on a body's circadian rhythms than a forward rotation. When forward rotation is used, people are then working at a time that was likely used for their sleep; it is a much more difficult adjustment on everyone's part.

ROOTS & TRENDS

The history of rotating shifts has it roots dating back to 1879 when Thomas Edison invented the light bulb. When we conquered the night, it let industries continue operations well beyond sunset. At the turn of the 20th century industrialization grew dramatically; manufacturing could literally continue around the clock with new crews coming in after so many hours to continue the process.

In the 1980's computer technologies took hold and the demand for employees with technical skills jumped. As with manufacturing, computer industries found many needs for round-the-clock operations. In the 1990's the booming economy drove increases in productivity. The key was to maximize the use of all assets while continuing to provide customer service. Just a few years ago, 24/7 operations have taken a foothold in retail too. You can shop online or by phone with Sears, JC Penny and others no matter what time of day. This trend is also highlighted by Wal-Mart opening up their 1500th 24-hour store just a few short years ago. Here in the 21st-century, the current trend is to consolidate and centralize global control and monitoring of employees and facilities. Current trends also show that the education level of today's workforce in the industries with round-the-clock operations is on the rise. In 1988, 33% of shiftworkers had some college education. In 1991 40% of the workforce had some college education and by 2001, 51% of shiftworkers had some college education. The higher education of our society reflects not only the computerization of our society but also the rapid pace technological change and automation has been accepted throughout our culture.

When considering a rotating shift schedule, it is important to know what trends currently exist in the US. As of 2002, 48% of all companies with shiftwork use 12-hour shifts; 37% used 8-hour shifts while the remaining companies use a combination of 8-hour, 10-hour and 12-hour rotating shifts. For those companies using the 8-hour shift schedule, 21% rotated their employees backwards. While there can be different names for each shift, we will identify backwards rotation as moving from the day shift to the graveyard shift, then to a midday, or late afternoon shift. 32% of the companies used a forward rotating shift schedule while the majority, 47% used fixed, or stable, shift schedules.

While rotating, an employee is bound to feel tired at one time or another. Over 35% of companies prohibiting any napping during a break also had at least one fatigue-related incident in 2001. Typically, people who work graveyard shifts and have difficulty sleeping are diagnosed with shiftwork disorder. The effects of this disorder can be a combination of headaches, problems staying focuses and of course, insomnia. While some doctors are prescribing drugs to aid those diagnosed with this disorder, it is still an issue that needs further study. At issue is determining whether it is ethical to prescribe anti-sleeping drugs to those employees who choose to stay awake when their body wants them to sleep. Companies that allowed napping during a break times had no reported incidents. If we use this data alone, we can deduce that allowing naps during breaks is good for the employee's safety and therefore good for the company. As we all know, employers have a duty to provide a safe and healthy workplace for their employees. Knowing that good health and fitness promote a safe environment in the workplace, many companies provide clinics or financial assistance for smoking cessation classes, stress management, and other wellness or lifestyle training.

Another issue to consider when opting for rotation in duty schedules is the age of the employees. When we are younger, there is usually no problem working on different shifts; our bodies seem to handle the changes better. But once a person reaches their forties or fifties, they traditionally find it more difficult to handle changes from one shift to another. This is mainly due to the constant changes in a person's sleep patterns. As we get older, we tend to wake up and go to sleep earlier than we did when we were younger. Circadian rhythms, or our body clocks, simply become less flexible and make it very difficult for older employees to make it through a graveyard shift. In a 2001 study by Business & Legal Reports, older workers did not perform as well as their younger co-workers during a simulated 12-hour graveyard shift schedule. This further confirmed what other studies have found, that the older the employee, the more likely that shiftwork will have a negative effect on the employee's quality of sleep, length of sleep, and their ability to fall asleep and stay asleep.

The Answer is Clear

Your mind is already made up. You have decided that rotation of duty schedules is the best for the company. How does one continue to promote good employee relations? Productivity, retention, morale, and the ability to implement changes and new projects are all affected by the morale of the employees. If their attitudes towards the company and their job are poor, this will likely have long term effects on all you do. Companies with round-the-clock operations have many challenges to deal with when working with their employees. Decisions that are made for the benefit of one crew may make another upset. Work-life programs can help ease the stress associated with shiftwork or even for those on stable shifts too. Working with the employee to help them meet their work and family obligations can benefit the company and employee in many ways. When a company works to provide childcare, flexible schedules, and more, the benefits are reduced absenteeism and turnover, increased productivity, and retention of customers through more satisfying contact with the company's employees.

Let us look at some of the advantages and disadvantages of rotating shifts. The advantages to rotating shifts are:

- They provide a balance of skills and work knowledge on every shift. Since all crews rotate fairly through all shifts, both the desirable and undesirable shifts, there is no motivation for the more senior workers to gather into a single crew.
- Prevent boredom and stress from repetitive work.
- Everyone is given a fair chance at being on the dayshift. As crews rotate through the day shift, they are given equal exposure and 'face time' with managers and company support personnel.
- Employees assigned to night or graveyard shifts usually qualify for shift premiums as an incentive to work during non-daytime hours.
- Training resources can be consolidated into one shift; that shift usually being the day shift. Since every employee must rotate through every shift, training specialists can concentrate their efforts on one shift without duplicating them on all shifts.
- Production standardization goes up. As a result of equitable training, exposure to management and support personnel, and equal work skills, all employees should perform their duties in a more uniform manner.

While this next example is a bit dated, it is important to know that in 1970, Exxon started a fast rotation schedule at its facility in Winnipeg, Canada that was made of 12-hour shifts. The employees worked a 3-2-2 schedule where over a two week time frame. They worked for three days, off for two, worked two more days followed by three days off, two days of work then two more days off. They ended the latter part of their schedule by working two days and then they were off two more days before the next rotation. With this schedule, the employee is guaranteed every other weekend off. Employees and management at Exxon were so happy with the scheduling that they adopted it at 20 other facilities. Monsanto also decided to test the 3-2-2 schedule and found that it increased morale and productivity while it decreased turnover and absenteeism. The disadvantages of rotating shiftwork are:

- The majority of shiftworkers, about 90% in all, prefer fixed shifts to rotating shifts.
- Uneven workloads are harder to manage with rotating shifts. As work crews change they may find the shift very easy or stressful compared to the shift to which they previously were assigned.
- It is generally more difficult for the body to adjust to rotating shifts. Very fast rotation (3 days or less on a shift) or very slow rotation (7 days are more on a shift) can help lessen the impact on a persons circadian rhythms.
- Stress on the employees family and social life. The additional physical and psychological demands of night and rotating shiftwork can be especially hard on couples with children.

In a case study held in conjunction with a plant that manufactured compact discs, the 3-2-2 monthly rotation of 12-hour shifts was adopted due to the successes of other local companies with the same type of shiftwork. An interview with 30 randomly selected employees revealed problems the company and employees did not foresee. Employees were not adjusting to the rotation as well as hoped. Conflicts arose. 70% of those interviewed stated they were having difficulty adjusting to the fatigue brought on by the rotation. 53% indicated a problem with the short cycle between rotations and over 40% of the employees had some conflicts in their social or family activities.

There are things a company can do to alleviate or reduce problems of rotating shiftwork. According to Circadian Technologies, shift schedules that accommodate a person's circadian rhythm are biocompatible. A biocompatible shift schedule does not rotate backwards, but rotates slowly forward. It lengthens the time between each rotation so there is little need for sleep adjustment. According to Circadian Technologies, this type of schedule meets the personal needs of shiftworkers and meets their preferences. The shift should not only be biocompatible, but it should also be compatible with the employee's family life. This means that shifts should be predictable, flexible and provide enough time off between each shift to allow a worker to recover from the shift change and still have some quality family or personal time.

Since the body tends to go into a shutdown mode at night, there are things we can do to help productivity. Some proven methods are:

- increase lighting.
- provide a variable noise source (radio, TV, machinery).
- lower the inside temperature.
- increase the air flow into the work areas.
- provide more physical activity.
- increase opportunities for social interaction
- vary the daily job routine.
- effectively use breaks.
- if possible, provide nutritional, low fat foods, even if no cafeteria is available.

There are many types of rotating shifts. For example, the traditional rotation of schedules is to change crews from one shift to another weekly. The slow rotation schedule can be a rotation of as little as a week (7 days) but usually a rotation every month or as much as one year between rotations. Others like an oscillating shift where two schedules rotate back and forth while a third shift stays fixed, or a partial rotation where a portion of a crew rotates while the remainder stays on their fixed shift, are used but are less popular than the traditional, or slow rotation, schedules. Fast rotating schedules, those done very quickly over a few days time warrant consideration too. The advantage to fast rotation is that a person's circadian rhythm has little time to adjust to the different shifts and therefore lets the employee stay in a day shift rhythm.

Rotating shift schedules in any 24/7 operation has always presented problems with productivity. Rotating shiftworkers normally suffer from sleep disorders due to changing sleep schedules and circadian rhythms. Attention spans suffer with the results of decreased productivity and an increased risk of fatigue-related incidents. Lastly, employees can suffer gastro-intestinal disorders from poor nutrition and higher caffeine consumption.

CONCLUSION

Communicating with employees and letting them be involved in the design of the shift schedule will help ensure the best schedule is implemented for both the company and the employees. While you may not have found all the answers in these pagers, it is sincerely hoped you now have some new information to work with and that you will develop the perfect work schedule someday soon. .

THE CAPITAL BUDGET PLAN; A TOOL FOR THE CHIEF ENGINEER

By Mike Seaver, President
Seaver Management and Consulting Services
Quincy, Illinois

THE CHIEF ENGINEER

For years, Chief Engineers were almost always the technical persons most proficient in repairing the transmitter. They knew the equipment, were good at, and really enjoyed fixing defects.

Chief Engineers were usually somewhat introverted, smart, and able to perform miracles with the equipment. They enjoyed being at the transmitter site, many were deeply involved in amateur radio, meticulous in work habits, able to repair their own cars, were great with all types of building projects, and repairing items whether at work or at home. I have always marveled at their abilities.

But working with managers and department directors, even other technical types was something most were really not overly comfortable doing. Often "people skills" were not their strengths.

In recent years, as equipment became more and more reliable, as circuit boards took the place of replaceable parts, and the FCC no longer required licensed engineers to be "on duty" at all times, the job of the Chief was either eliminated or expanded to supervision of multiple stations and more duties were assigned.

THE CHIEF ENGINEER AS A BUSINESS MAN

Chief Engineers usually don't consider themselves "businessmen", but in today's world, they must. Broadcasting, at least commercial broadcasting, is about making money! One of my former General Managers once taught me that we are, or should be, "broadcasters" first, then do whatever job in the field we like to do. Revenue is the lifeblood of the business and it takes more than just equipment to get these ad dollars to come through the door. It takes teamwork. Chief Engineers are members of the team and how valuable is up to the Chiefs, themselves.

As the business changes and you, as a Chief Engineer, spend less and less time in the shop, you will either go the way of the Edsel or adapt to new roles and practices that will again make you valuable to your employer. Becoming an "equipment manager" and "technological resource" is a good way to ensure your value to your employer.

As a businessperson, you, the Chief Engineer, must know how the station operates, at least to the point of knowing the rudiments of each department's operation. This you will need to obtain through discussions with your GM, Business Manager, or others. I recommend you establish a good working relationship with the Business Manager. For the Business Manager knows the overall financial plan and can be of great help to you in developing technology.

THE CHIEF MUST BE MULTI-LINGUAL:

In order to be a functioning part of the business team, you must become "multi-lingual". It is imperative that you develop a good working relationship with the Business Manager. To do that, you need to know what it is they do and basically, how they do it and why. These are things that most likely do not interest you, but will serve you well as you try to keep your station and department on the cutting edge. It is the Business Manager who understands the cash flow of the station. There are a select few who understand the "big picture" of the money end of the business. And it is money that you need to maintain the station to its greatest operational efficiency. The language of "money" differs from engineering. You need to learn how to talk "accounting". Not fluently, but at least to understand the terminology and be able to use it knowledgably. You must ASK if you don't already know…and don't assume you can figure it all out. Most likely, you cannot.

Also, you must be able to explain what you do in terms the business people can understand. This requires you to be able to "teach" the GM and/or Business Manager your language to the level that they can converse with you and understand what you are telling them.

THE CAPITAL PLAN

The first step to establish the Capital Plan as a tool is the inventory. You most likely have an inventory that was done for asset management purposes. But this one should be a unique inventory for engineering. Begin by making a list of every item for which you are responsible. I developed a unique numbering system that allowed me to easily identify each item that was simple to use and easy to employ. It isn't necessary for you to include every detail, but you should at least identify each item by some unique method such as serial number and model and should include a location. Seaver Management and Consulting Services uses a form I developed for that purpose. The idea is not to duplicate any effort by the accounting department, but to provide an overview of your responsibilities.

Next, develop an anecdotal file for each item. It doesn't need to be a volume but should include things like age of the equipment, what generation it is, availability of information and/or parts, it's approximate cost when new, including any installation costs that were incurred, and approximate lifetime of the device. It might be prudent to make note of your observations of where that particular device or system is in relation to its life cycle. Realism is the key to making this work. Ideally, from an engineer's point of view, a transmitter should be replaced every 20 years; this probably would not be realistic. Same goes for every device in your charge. Be realistic.

Once the inventory of equipment, systems, and assets is completed you need to establish a method of documentation allowing you to track each item through its lifecycle. Of course this means more paper work, but ultimately, it will have been worth the effort. With databases such as Excel, it is now quite easy to establish a good working tool for tracking equipment repairs, inventories, and lifecycles.

Seaver Management and Consulting Services has developed a great form for inventory and tracking but you can do so on your own if you so desire. It doesn't need to be fancy, just functional. After all, it is your "tool," not a "document."

Armed with this new system you can now begin to establish a "paper trail" on each piece of equipment as you repair or otherwise interface with it. While, as I said, it takes a bit of time, the system will pay off in the long run. I have used this method to either replace or add equipment that would have otherwise gone wanting. And the nice thing, it saves your station money!

SHARING THE PLAN

Now you have a decent idea of the status of the facility and all the equipment under your charge. What to do with it? First explain to the General Manager why you compiled this information and how valuable it is to the overall profit plan of the station. Now is the time to employ your new, multi-lingual skills.

While we would like to believe that as "captain of the ship," the General Manager would understand new technology and how it drives the business, the reality is most do not. They are too busy to spend the necessary time to acquire and incorporate that information. In the real world, most General Managers come from the Sales ranks. It is a fairly natural progression since making money is the lifeblood of the business. Regardless of how proficient they are at sales, they probably don't know the difference between a flux capacitor and an intermediate beacon. As an equipment manager, you must be able to inform the General Manager in layman's terms what you are talking about and why this is important enough for them to care. Here, your written and verbal skills, all in that new language we shall call "management-ese" will pay off as the General Manager, etal. come to rely on you as the technology expert and authority that you are.

NOTE: I have always said you are proficient at what you do when you can explain it to someone who doesn't have a clue about the technicalities of your work but they come away with enough comprehension to be able to at least talk to you and have a remedial understanding of what it is you do.

Now is the time to discuss the future technical direction of the station with the General Manager. This requires you to be up to speed with new equipment on the market and techniques that might enhance the revenue and operation of the station. Maintaining status quo is not always in the best interest of the station.

You also should share this information with the Business Manager. This helps them understand why you are asking for money to be allocated for your specific projects. Discussing the projects with individual department directors involved is also material to make the Capital Budget a working tool.

PREPARING THE BUDGET

The first step in preparing a Capital Expense Budget Plan is to confer with the Business Manager to determine approximately how much money will be available for capital projects in the up coming years. This sounds elementary, but without this knowledge, you are wasting your time. Knowing how much money is likely to be available will help you prioritize your needs/wants.

The next step is to do an evaluation of the station's technology plan. Good business practice is usually based on a five-year plan. This becomes a juggling act to determine the allocation of proposed money. The best method is to establish levels of priority. The system Seaver Management and Consulting Services employs is based on four.

Mandatory

This category must be used cautiously. Most items deal with either legalities or life safety issues. A rather blatant example of a mandatory item would be fencing around the tower. Without it, the station will most likely spend money on a fine AND fencing plus any legal settlements if someone is injured or killed on the tower. This needs to be presented by explaining the FCC rule in understandable terms why the station is in violation, what the fines could be, and perhaps even some examples of stations that did not adhere to this rule and did receive fines. EAS compliance, tower fencing, registration posting, tower lighting, building security matters, and the myriad of public laws that apply to all facilities fall into this category. If the tower is in ill-repair, the tower lights not working properly, roofs leaking, hazmat issues; all these fall into the mandatory category.

Also, having the appropriate licensing of software is mandatory! This should be addressed if not as "capital" somewhere in the budget cycle. Fines for failure to do so can be quite hefty and cut into the profits.

All these mandatory items, as all the Capital Budget, should be well documented and in writing with your name and title somewhere affixed and a notation to whom it is being addressed. Most Chief Engineers are also the designated "Chief Operator" and as your letter of designation declares, it is the General Manager, along with yourself, who is responsible for the proper operation of the facility. Be sure you have nicely, but adequately covered your "bases."

Required

This is the category in which most capital expenses will fall. Under this title will be equipment and items that are necessary to maintain essential services, to incorporate any new technology into the operation, or to upgrade systems and services that are necessary to maintain daily operations.

This category includes both new equipment and upgrades. An example of a "required" item would be a replacement for an audio console that is failing or otherwise unsuitable for current use. This would be something that needs replaced or added just to keep things where they are now or to enhance the product of the station in a manner that would either maintain or enhance the current revenue stream.

Necessary

As with *required,* these items are part of the workflow. But unlike *required,* the station can, if necessary, get by without them. Usually, these items are the high repair equipment/services that are "dollaring" you in repair costs and upkeep.

Desirable

This is the category that contains all the items and systems that will enhance the overall operation and allow for expansion of services and revenue. Pushing the station ahead in technology and increasing the efficiency of the operation is the purpose of these items and you need to demonstrate how that will accomplish this. Often this requires input and help from the various departments such as News, Operations, Creative, etc. Of course this requires you to learn more "languages" and demonstrate your skills at talking in their language. Special emphasis should be placed on documenting and demonstrating how the expense for these items will increase revenue. That will be of great value in getting them approved.

These are only guidelines and will need to be customized to your facility and staffing levels. From the smallest radio facility to the largest radio or TV facility these criteria should be part of the overall equipment management plan.

Now that the priorities have been established we must next designate their category.

CLASSIFICATION

Not to be confused with the justification, classification defines how the purchase will be incorporated into the business inventory. They are classified as:

New

Kind of obvious, but never-the-less, necessary to establish inventory control. "New" can include rebuilt, used, or other specifications but is a "new" service, piece of equipment, or system to the station. It doesn't replace or add to, it is new.

Additional

This, as an example, would be the criteria you would use to add an additional camera to the news pool, or an additional recorder. Additional equipment simply adds to the number of existing service.

Expansion

If, for instance, the router is too small and you wish to add on, this would be an "expansion."

Replacement

The old remote control quit working? Budgeting for a "replacement?" That is the type of equipment you put into this category.

Of course all classifications require you to demonstrate their cost benefit. This is the key to developing the Capital Budget Plan as your working tool. If it doesn't affect the bottom line in some positive manner the equipment is not needed.

Finally, you probably won't be able to buy it all at once, even if approved. Here, you need to work with your money person, be it the General Manager or the Business Manager, to determine in which quarter the purchase would best work into the projected budget. Breaking proposed expenses into quarters is the way business is done. Simply identify in which quarter you wish to spend the money. Remember, not everything sits on a shelf somewhere and there is lead-time between orders and billing. Work with the management and business office to coordinate that to everyone's satisfaction.

JUSTIFICATIONS:

All the above information leads to one thing, "why we need it." Worn out phrases such as "it is old" or "it doesn't work" won't suffice. You need to employ the linguistic skills you learned to explain why the request would be in the best financial interest of the station. This is where you attach and/or include information that documents why you want to spend the money. Remember, the accountants will be looking at value for value. Will the money spent on "capital" be worth more than it would make by investing or by spending it on other projects? While this is something that engineers don't like to do, it is necessary to address and will go far in assisting you in acquiring the capital investments you desire.

PRESENTING THE CAPITAL EXPENSE PLAN:

This will be your greatest challenge. Again, you must quit thinking like an engineer. You will not be presenting the budget to another engineer, even if it is the Corporate Chief. For even the Corporate Chief is first and foremost a manager, trying to make every aspect of the business fit within a revenue plan. Your enthusiasm for the new Model XL 5000 Flux Capacitor will mean little, even if it is 0.000001% accurate! You need to explain how that accuracy translates into money.

Here your ability to speak "managerial-ese" will pay off. Right up front you will need to establish the following: What are you proposing to buy (specs, are not all that important at this stage), why you think the station needs it (saves money/makes money are about the only two categories here), how much it will make, what will be the approximate payback time, and how much it will cost (profit to expense benefit).

I would suggest that you have this written as concisely as possible. Include all the financial data, both direct and indirect, at the beginning, and in very understandable terms. This is where you might need someone to assist you if you are not a wordsmith. You should have each item with all the above information of priority, reason(s), and revenue advantages, when it is needed and why.

Put your capital budgets in a three-ring binder with a title page. Use good paper for the financial portion and put it at the first of the report. It must be typewritten and neat. Use creative tools, such as different colors, font changes, or images to help make this report something that the GM/Owner/Business Manager would want to pick up and read.

If you wish to put pictures, specifications, and other things in the binder, do so at the end; they are of little interest to those making the money decisions.

Remember, the Capital Budget Plan is a tool, something that will help you achieve the goals for your department and enhance the overall profitability of your station. Working within the guides of this plan, you should be able to eliminate some of the problems and emergencies that arise throughout the technical arena, make for a better work environment, and increase profitability for your station by planning for those capital expenses you know will occur. When you become part of the revenue stream, regardless of how, you become a more valuable employee.

IN SUMMARY, as Chief Engineer you need to become THE source for technology information. You need to "manage" the equipment by having a written inventory of each system or device, a method for tracking the repair cost cycle, the downtime cycle, and the lifecycle of each. You need to become multi-lingual, able to converse with non-engineers using their language and be able to teach them to understand your needs are and why. You must be part of the "team"; in fact, you are the quarterback, so to speak, in that you coordinate the technical future of the station.

If you employ the tools outlined in this paper you will go far in establishing an engineering program that is extremely valuable to your management and assist your station in achieving their revenue goals.

Radio Technology Advancements

Thursday, April 19, 2007
9:00 AM – 11:30 AM

Chairperson: Gary Kline
Cumulus Media, Atlanta, GA

***RDS Enhancements: Artist and Song Title Through RT+**
Michael Bergman, New Digital Technologies, Kenwood, Princeton Junction, NJ

Updating Longley-Rice for FM Reception Area Prediction
Sid Shumate, BIA Financial Network/Dataworld, Chantilly, VA

Planning of the Digital Terrestrial Broadcasting Service in Parts of Regions 1 and 3, in the Frequency Bands 174 – 230 MHz and 470 – 862 MHz (RRC-06)
Pham Hai, International Telecommunication Union, Geneva, Switzerland

***FM-based Technologies Now and in the Future**
Matthew Straeb, Global Security Systems, Lafayette, LA

***Public Radio International's Program Service Data Project: A New Dimension of Local Service on HD Radio**
Riccardo Ruotolo, Public Radio International, Minneapolis, MN

***NRSC Report**
David Layer, NAB, Washington, DC
Skip Pizzi, Microsoft Corporation, Fairfax, VA

---

*Paper not available at the time of publication

UPDATING LONGLEY RICE FOR FM RECEPTION AREA PREDICTION

Sid Shumate
BIA Financial Network/Dataworld
Chantilly, VA

BIAfn, and its subsidiary, Dataworld, are conducting a long-term project to provide a more thorough, yet easily understood, FM broadcast vehicular reception analysis and mapping system. At the 2006 NAB engineering conference, we presented updated RF reception reference levels and calculations to support this project. In order to move forward, we found it necessary to consider updating the computer implementations of the Longley-Rice algorithms to take full advantage of currently available desktop computer capabilities, and recently improved terrain database information now available for civilian use.

WHY UPDATE LONGLEY-RICE?

The Longley-Rice methodology, developed by Anita Longley and Phil Rice at the National Telecommunications and Information Administration (NTIA) Institute for Telecommunications Sciences (ITS) in the late 1960's and 1970's, has long proven to be the gold standard for determining signal strength for FM and TV broadcast reception. However, the early practical computer implementations of the complex and data-intensive methodology were severely limited by the available computer hardware of the day, and the difficulty of manually extracting the terrain path data from USGS maps.

Fig. 1 A basic IBM 1130 CPU with hard drive cabinet.

A typical industrial-strength scientific computer available to academic, business, and government institutions in the late 1960's was an IBM Series 1130.

The preceding photo is the CPU console of the IBM Series 1130. It was intended for technical and scientific users. The series 1130 was programmed using the FORTRAN computer language, and could receive input from IBM punch cards. Its output was to a removable hard disk drive that could be as large as 1 Megabyte, or to a printer, paper tape, IBM punch cards, a plotter, or a graphical interface monitor. It was capable of floating-point calculations, had a processor clock speed of 170 kilohertz, and a magnetic core RAM that could hold up to 64 KB.

These limitations led to the creation of two modes of operation for the first practical version of a Longley-Rice computer program, an area prediction mode, and a point-to-point path prediction mode. The area mode did not use a terrain database as an information source; it generated an output based on average terrain factor inputs. This mode is little used today for FM broadcast prediction. The second mode generated a point-to-point signal loss calculation based upon a single path of equidistant terrain data points between a transmitting antenna and a receiving antenna. The terrain heights were often derived manually from USGS maps in a straight line, and input onto punch cards to feed to the computer.

This process could be slow and labor intensive. By the late 1980's, commercial updates of the Longley Rice methodology were based on a little-changed porting of the earlier computer algorithms of the point-to-point mode into C++ language, that could be compiled and run on personal computers. The punch cards were now obsolete. These improved versions repeated a point to point calculation along 8, 72, or up to 360 radials using terrain information from a computerized database, then interpolated between the radial data for map display. To do such a Longley-Rice calculation would take an IBM PC clone with a 486 processor running in DOS, about 2 hours. But this still did not come close to the computer version anticipated by the originators of the methodology.

In 1982, in "A Guide to the Use of the ITS Irregular Terrain Model in the Area Prediction Mode", George Hufford, Anita Longley, and W.A. Kissick wrote:

"In the case of a specific operational area or a specific coverage area, one is confronted with a different problem. Here one has a large multitude of possible propagation paths, each of which can presumably be described in detail. One might, therefore, want to consider point-to-point calculations for each of them. But the sheer magnitude of the required input data makes one hesitant. An alternative, which requires far less input data, is to use an area prediction model, particularly if the model provides by itself the required statistics. Even in the case of a specific communications link, the required detailed information for the propagation path may be unobtainable so that one is forced to use the less demanding area prediction model. Of course, in doing so, one expects to lose in precision and in the dependability of the results."

For reference, when this was written, the latest and greatest small business microcomputer was the IBM XT.

The point-to-point algorithms used by the current version of the NTIA/ITS Irregular Terrain Model (ITM) executable program, publicly released as ITMsetup.exe in 2003, was ported from FORTRAN to Visual C++, but otherwise remains largely the same. It is a Windows program, operable on current PC's, and the program utilizes the GLOBE 30 arc-second terrain database. Otherwise, it still only does the area and point-to-point mode calculations of the original program.

In 1999, the FCC was required to create an improved point-to-point signal reception model for DTV. The resulting program calculated individual point-to-point paths to reception points that were located at one-kilometer intervals; it was the first major active use of an implementation of the point-to-large-multitude-of-points mode anticipated by Hufford, Longley, and Kissick.

TERRAIN DATABASE LIMITATIONS

The arduous manual task of obtaining terrain data points from maps, contributed to the limitation of a 16 km (10 mile) radius being chosen for the original FM contour calculation methodology. But when calculating coverage for a Class C FM station, with 50 kW of ERP, the 10-mile limit allows the FCC Contour calculation to fail to consider the effects of terrain located more than 10 miles from the transmitter site. The slide rules and USGS 7.5 minute paper maps now gather dust, but the FCC has never revised the contour calculation methodology to increase the path length considered.

Similarly, once the FCC accepted the use of the Longley-Rice methodology, utilizing a 30 arc-second database, there has been a certain reluctance to experiment with, or change to improve, what is accepted as good enough for government work. The distance between points in the GLOBE 30 arc-second database, however, for most of the Continental U.S., are separated by more than 900 meters. In addition, radio engineers who work with digital signals are familiar with the Nyquist criterion, stating that the maximum resolution that can be derived from a modulated signal is half the sampling rate. If one thinks of the terrain database's samples as a signal modulated by height, then the smallest object that could be noted, defined as the resolution of the signal, would be no less than 2 times the sample rate; because of the other factors involved, the actual maximum resolution of a terrain database is 2.83 times the sample rate. So the resolution of the GLOBE database in the continental U. S. is approximately 2.83 times 900 meters, or about 2.5 kilometers or 1.6 miles, depending upon location and direction. Therefore, while an FCC contour calculation can miss a mountain more than 10 miles from the transmitter site, the Longley Rice path calculation using a 30 arc-second terrain database can easily miss a mile-wide hill or valley, anywhere.

Why is this not obvious to anyone looking at the data from a point-to-point calculation? If one downloads (its FREE and public domain) the NTIA ITM executable released in 2003, then tracks down and downloads the (also FREE) GLOBE database from its current web location, and creates a simple globe.dat file (instructions are in the readme.txt file) to make them work together, then one can run the ITM point to point calculation on any modern Intel clone PC running Windows. If we select a set of transmit and a receive co-ordinates that are 10 miles (16 km) apart, we find that the point-to-point mode derives about 800 data points along the path to use in calculating the Longley-Rice signal loss.

But lets use a special case; let us select a transmitter point and a receive point that have the same latitude, with a 16 km path length. Again the program will derive about 800 data points from the GLOBE terrain database. But did we not say that the points in the GLOBE terrain database are more than 900 meters apart? If we divide 16 km, or 16,000 meters, by 900 meters, we find that there are only about 18 actual height data points in the GLOBE database between the transmitter and the receiver. The rest are interpolated. In the NTIA executable, it is supposed to be using a weighting function that considers the four closest data points, creating variation in the derived terrain heights, and making them appear realistic.

One cannot be certain that the same algorithms are being used, as the publicly available source code for the Longley Rice implementation does not include the code to derive the pfl array, or terrain data profile, that originally was manually derived and input on punch cards. This has required commercial and amateur programmers to write their own code to determine the actual terrain path using either a direct line for short paths, or a great circle calculation for longer paths. The great circle calculation may use either spherical or one of several ellipsoid earth models to derive the path co-ordinates. Programmers also have to write code to determine the parameters of the weighted averaging function used to derive the height of the points that are in between the database points.

Unfortunately, as explained above, a lot of real terrain features can be missed between the actual data points in a low-resolution terrain height database. So it significantly improves the accuracy of a Longley-Rice prediction to use as fine-detailed a terrain height database as is currently available.

The current DTV calculation methodology, using points a kilometer apart, therefore suffers significantly in resolution compared to the resolution we could and should be obtaining using massive multiple point-to-point calculations with data points at 1 arc-second (90 meter) intervals.

Nationally, there are currently two publicly available databases that provide terrain height data at a 1-arc-second sample rate, and are essentially complete for the entire continental U.S.; these are the Shuttle Radar Terrain Mission SRTM-1 finished-version data, and the National Elevation Database (NED) terrain database.

There are differences in height values between these two databases, owing to the methods used to obtain the data for each. The SRTM data was obtained over a few days in February 2000, using a radar system built for the STS-99 shuttle mission. Its data is based on the height of the reflected radio waves, and therefore incorporates the height of objects above ground level that reflected the radar signal, be they man-made buildings or thick foliage. The finished version has been processed and edited to fill most voids, to flatten the data over water, and to better show shorelines, and is referenced to the WGS84 (EGM96) co-ordinates and ellipsoid earth shape model.

The NED data is a ground-height database, generated from a patchwork of best-available information going back as far as 75 years, referenced to the NAD83 co-ordinates. If what you want is the actual ground height for a tower base, or the actual ground height under a vehicle, then the NED data is preferable. If what you want is the actual height of a diffraction point for a Longley-Rice calculation, the SRTM data provides the actual radio wave reflection height, effectively including man-made and natural features.

Since most roadways are at ground height most of the time, with high bridges, overpasses, and the roof level of parking garages being special exceptions, the proper implementation of Longley-Rice should utilize both databases. The program should derive the vehicular reception points height, and the transmit point ground height, (when not obtained from the FCC database), from the NED data. The remainder of the path data points should be taken from the SRTM data, as the critical height there is the height at which radio waves reflect.

OTHER IMPROVEMENTS TO CONSIDER

The current algorithm requires an estimate of the path length to the horizon be made for short path runs. However, the data is available to extend the path profile on short paths to beyond the expected horizon, allowing a more accurate determination of the actual horizon information to be used.

Hammett & Edison asked the FCC in October, 2004, to modify OET-69 software to ignore EC3, to correctly calculate depression angles, and to give stations the opportunity to submit their actual elevation patterns, and, if mechanical beam tilt is used, their main beam-elevation patterns. While this primarily refers to DTV television measurements, as beam tilt is less often used in FM antennas, it should be considered in this process.

Dr. Oded Bendov, designer of the new TV antenna system to be placed on the Freedom Tower, also issued a paper, "On the Validity of the Longley-Rice (50, 90/10) Propagation Model for HDTV Coverage and Interference analysis; this paper is available from the Dielectric Communications website. Again, while the criticisms leveled at the Longley Rice computed results apply primarily to DTV in this paper, many of the issues mentioned may also apply to FM broadcast.

Last year, Doug Lung also discussed problems with DTV coverage computed by Longley-Rice at the NAB convention, in his "RF Delusions" presentation.

Work has been done to correct errors in Longley-Rice computation. Anita Longley issued, as an Appendix to Radio Propagation in Urban Areas, a "1977 Modification of the Longley-Rice Computer Model.

The Propagation Model Development annual summaries from the Institute for Telecommunications Sciences (ITS) also document ongoing development work relating to the Irregular Terrain Model (ITM), including analysis performed to support the FCC regarding the Satellite Home Viewer Act. In 2003 and 2004, the summaries reported on intercomparison and harmonization of the ITM and TIREM models.

In 2005, the ITS Propagation Model Development summary reported on improvements of the effective height algorithm in the ITS ITM, but it is unlikely that the currently available public ITM executable program, released in 2003, has incorporated these improvements.

SUMMARY

Today's desktop PC's are far faster, and more capable, and have much larger RAM and hard drive memory, than the mainframes of 1982. The "sheer magnitude of data" still perceived as a stumbling block in 1982, can now easily be handled by today's newest desktop computers.

We are revisiting the Longley-Rice algorithms and the published FORTRAN and (Visual) C++ source code for the ITM. The goal is to create, as suggested by Hufford, Longley and Kissick in 1982, an executable program that provides a detailed point-to-point Windows-based implementation of the Longley-Rice algorithm. This would be used commercially to provide a more accurate alternative to the "area prediction mode" and current radial-based point-to-point mode implementations, as a basis for FM broadcast radio reception and interference prediction, for both analog and IBOC. The NAB 2007 conference presentation will include a progress report.

Acknowledgements:

Photo credit: IBM 1130 CPU photo by permission from: www.ibm1130.org.

IBM is a registered trademark of International Business Machines, Inc., Intel is a registered trademark of the Intel Corporation, and Windows is a registered trademark of Microsoft.

The U. S. Dept. of Commerce, National Telecommunications and Information Administration Institute for Telecommunications Sciences, for the programs, source code, and Longley-Rice documentation accessible via: http: flattop.its.builderdoc/itm.html.

NASA for the Shuttle Radar Terrain Mission and the public release of 1 arc-second databases for the U.S., and the USGS for the National Elevation Dataset (NED) in 1 arc-second data.

The long-term contributions of Anita Longley, Phil Rice, G.A. Hufford, W.A. Kissick, Paul McKenna, Fred Najimy and Alaka Paul.

The previous analysis and criticisms of the current Longley-Rice computer implementations provided by Hammett & Edison, Dr. Oded Bendov, and Doug Lung.

PLANNING OF THE DIGITAL TERRESTRIAL BROADCASTING SERVICE IN PARTS OF REGIONS 1 AND 3, IN THE FREQUENCY BANDS 174-230 MHZ AND 470-862 MHZ (RRC-06)

Pham Nhu Hai,

International Telecommunication Union
Geneva, Switzerland

ABSTRACT

On 16 June 2006, one hundred and twenty (120) countries in Europe, Africa, Middle East and the Islamic Republic of Iran signed a treaty agreement at the conclusion of ITU's Regional Radiocommunication Conference (RRC-06) in Geneva Switzerland. This heralds the development of 'all-digital' terrestrial broadcast services for sound and television. The digitalization of broadcasting in Europe, Africa, Middle East and the Islamic Republic of Iran by 2015 represents a major landmark towards establishing building blocks for the Information Society.

The agreement reached at RRC-06 paves the way for utilizing the full potential of information and communication technologies to achieve the internationally recognized development goals. The date of transition to digital terrestrial broadcasting in the year 2015 is intended to coincide with the targets set by the Millennium Development Goals.

This paper discusses the results of the Conference and future implications for the development of digital services in the planning area.

INTRODUCTION

Radio and television broadcasting in Europe, Middle East and Africa, like in many other countries in the world, is going through the substantive transformation, changing from analogue to digital. Unlike the previous change of the television transmission standards from black and white to color, the change to digital is significant because digital broadcasting technology allows more programs and choices. For example it allows the transmission of four (4) or five (5) DVB-T digital television programs in the same radio-frequency channel as that of a single analogue transmission, and reception of digital services will not be limited to fixed installations in the home, but may also be received by computers, mobile phones and other hand-held receivers.

Consequently a new frequency framework needs to be established to take full advantage of what digital broadcasting can offer. The frequency planning framework that had been established for analogue television, 45 years ago for Europe (Stockholm frequency Plan, 1961) and 16 years ago for Africa (Geneva frequency Plan, 1989), are therefore no longer suitable for the digital age.

In January 2000, European countries wrote to ITU stating that efficient implementation of terrestrial television broadcasting will not be possible on the basis of the existing frequency assignment plan contained in the European Broadcasting Agreement of Stockholm 1961. These countries asked ITU to organize a conference to revise the Stockholm Plan of 1961. Their request resulted in an ITU Council decision to convene a Regional Radiocommunication Conference for the planning of the digital broadcasting service in the Bands III (174–230 MHz) and IV and V (470–862 MHz). Other countries also recognized the advantages of a structured approach to an all-digital situation. This led to further consultations with the ITU membership, which resulted in extending the scope of the conference. It was decided that RRC-04 would deal with an expanded planning area that includes 119 Member States (out of a total of 189), from Europe, Africa, Middle East and the Islamic Republic of Iran, as well as some Asian countries bordering the Russian Federation, Kyrgiz Republic, Tajikistan, Turkmenistan and Uzbekistan.

The Council also decided that the conference would be held in two sessions. The first session held in May 2004 (RRC-04), was to establish the necessary technical bases to facilitate planning activities prior to the second session as well as to prepare a report that would provide the basis for the work of the second session. The second session (RRC-06), was to establish a digital frequency plan for the planning area and to adopt an agreement that set out the procedures for making changes to the frequency plan and to bring the frequency assignments into operation.

THE FIRST SESSION – TECHNICAL BASES

Over 800 participants from 95 countries in the planning area attended the first session, held from 10 to 28 May 2004 in Geneva Switzerland. The first session adopted the planning principles and criteria that are used as the basis for the planning activities between the two sessions and for the development of a draft frequency plan to be submitted to the second session for its consideration. Some important planning decisions are discussed below.

Digital Radio and Television Systems

It was agreed that common digital radio and television systems should be used by all countries in the planning area.

- T-DAB, Terrestrial digital sound broadcasting, for radio with two reception modes of mobile and portable.
- DVB-T, Digital video broadcasting – terrestrial, for television with four reception modes of fixed, mobile, portable indoor and outdoor.

Assig nment and Allotment Planning

In assignment planning, a specific channel is assigned to an individual transmitter site location with defined transmission characteristics (for example, radiated power, antenna height, etc.). At the completion of the assignment plan, the locations and characteristics of all transmitters are known and the transmitters can be brought into service without further coordination.

In allotment planning, a specific channel is planned for an administration to provide coverage over a defined area within its service area, called the allotment area. Transmitter sites and their characteristics are unknown at the planning stage and should be defined at the time of the conversion of the allotment into one or more assignments.

Digital requirements submitted from administrations could be either assignment or allotment or a combination of the two. This results in five possible types of digital requirements:

- a single assignment;
- a set of assignments making up a single frequency network;
- a single allotment;
- an allotment with one or more linked assignments;
- a single assignment and a single allotment.

Frequency Bands

Band III (174-230 MHz) is available for DVB-T and T-DAB planning and there should not be a rigid splitting of Band III between DVB-T and T-DAB, unless this is proposed on a national basis and only depending on national requirements. However, it was considered that the efficient use of Band III may be facilitated by the separation of T-DAB and DVB-T services, as well as a harmonized 7 MHz bandwidth for all of Band III.

In Band III, different television channel spacings are used across the planning area. In Eastern Europe, France, Ireland and some African countries, channels are 8 MHz wide but they are aligned differently. In other countries the channel width is 7 MHz, also with different channel alignments. In some countries using a 7 MHz channel width (e.g. Italy and Morocco), there are also different channel spacings. This means that in the VHF bands there are many cases where channels overlap.

Bands IV and V (470-862 MHz), are used for DVB-T only and based upon an 8 MHz channel bandwidth associated with an identical 8 MHz channel spacing.

Ref erence Planning Conf ig urations (RPC)

T-DAB and DVB-T offer the freedom to implement a large variety of broadcast service options. For DVB-T in particular, several thousand planning configurations could be thought of by combining the various possible modulation schemes, code rates, Fast Fourier Transform (FFT) modes, guard intervals, reception modes, coverage quality classes, network approaches, etc. Thus, a planning configuration describes the sum of all relevant technical aspects of a broadcasting service implementation. The various aspects of a planning configuration, for the example of DVB-T, are summarized in Table 1.

Table 1 - Aspects of DVB-T planning configurations

Aspect	Element
Reception mode	Fixed roof-level Portable outdoor Portable indoor Mobile
Coverage quality (in terms of percentage of locations)	70% 95% 99%
Network structure	MFN (single transmitter) SFN Dense SFN
DVB-T system variant	from QPSK-1/2 to 64-QAM-7/8
Reference Calculation Frequency	Band III (200 MHz) Band IV (500 MHz) Band V (800 MHz)

However, a large number of these theoretically possible combinations make little or no sense, from an economic, a technical or a frequency-management point of view. Moreover, seen from the point of view of compatibility analysis, which is the major issue in producing a frequency plan, a large number of the realistic and meaningful planning configurations can be treated as equivalent, since they differ little or not at all in terms of compatibility aspects.

For frequency planning purposes, a reduction to a very small number of so-called reference planning configurations (RPCs) is possible, which then are abstract in the sense that they no longer correspond to specific real planning configurations. RPCs including reference parameters for DVB-T and T-DAB implementation are given in Table 2 and 3 respectively.

Table 2 - RPCs for DVB-T

RPC	RPC 1	RPC 2	RPC 3
Reference location probability	95%	95%	95%
Reference *C/N* (dB)	21	19	17
Reference $(E_{med})_{ref}$ (dB(μV/m)) at 200 MHz	50	67	76
Reference $(E_{med})_{ref}$ (dB(μV/m)) at 650 MHz	56	78	88

$(E_{med})_{ref}$: minimum median equivalent field strength

RPC 1: RPC for fixed roof-level reception

RPC 2: RPC for portable outdoor reception or lower coverage quality portable indoor reception or mobile reception

RPC 3: RPC for higher coverage quality for portable indoor reception.

Table 3 - RPCs for T-DAB

RPC	RPC 4	RPC 5
Reference location probability	99%	95%
Reference *C/N* (dB)	15	15
Reference $(E_{med})_{ref}$ (dB(μV/m)) at 200 MHz	60	66

$(E_{med})_{ref}$: minimum median equivalent field strength

RPC 4: RPC for mobile reception

RPC 5: RPC for portable indoor reception

Reference Networks (RN)

A basic task when establishing a frequency plan is to perform compatibility analyses between transmitters and/or networks. For such calculations, the characteristics of the transmitters have to be known. If a requirement is given in assignment form, these characteristics are available.

However, for the cases of allotments where the service area is known, but the exact transmitter characteristics of a network are not known at the time when a frequency plan is to be established, it is necessary to define generic network structures which may represent the, as yet unknown, real networks in a compatibility analysis. Such generic networks are called reference networks.

Propagation Prediction Method

The propagation prediction method used for planning is based on ITU's Recommendation (ITU-R Rec. P.1546-1) for field strength prediction, which is applicable for broadcasting; land-mobile, etc. services in the frequency range 30 to 3,000 MHz. The propagation methods consist of a set of propagation curves for each of the nine (9) different propagation zones in the planning area:

- Zone 1: temperate and subtropical regions;
- Zone 2: desert regions, displaying propagation conditions found in regions having low humidity and small annual variations in climate;
- Zone 3: equatorial regions, displaying propagation conditions found in regions with hot and humid climates;
- Zone 4: maritime regions, displaying propagation conditions found over warm seas and in one terrestrial zone (referred to as "coastal land" in Annex 2.1) at low altitude bordering warm seas, where superrefraction conditions occasionally occur (Caspian Sea* and all the seas around the African continent are Zone 4 except Zones A and B designated below);
- Zone 5: maritime regions, displaying propagation conditions found over cold seas;
- Zone A: maritime zone at low latitudes, frequently displaying superrefractivity;
- Zone B: maritime zone at low latitudes, displaying superrefractivity to a lesser extent than Zone A;
- Zone C: maritime zone from the junction of the coastline of the Islamic Republic of Iran with its border to Pakistan westward along the coastlines of the Islamic Republic of Iran and of Iraq, through point 48° E, 30° N along the coastline of Kuwait, the eastern coastline of Saudi Arabia, the coastlines of Qatar, the United Arab Emirates and Oman down to the intersection with parallel 22° N;
- Zone D: land strip of maximum depth of 100 km surrounding Zone C and the West African land region consisting of two parts. The northerly part extends no more than 50 km inland from the Atlantic Ocean but is limited to the east by a line from 30° N 10° W to 20° N 13° W and to the west by the Atlantic coast. The southerly part is the land area west of two lines, one from 20° N 15° W to 15° N 12° W and the other from 15° N 12° W to 9° N 13° W, but not extending beyond the coastline.

INTERSESSIONAL ACTIVITIES

The first session adopted a time schedule for all activities before the second session to be held in June 2006 (Figure 2). In addition, the first session also set up three (3) main groups responsible for the work required for the preparation for the second session:

- Intersessional Planning Group (IPG) consisting of all participating countries, responsible for overseeing the development of draft frequency plans for DVB-T and T-DAB broadcasting services, taking into account the digital requirements submitted by all countries and the results of bilateral and multilateral coordination and negotiations carried out between administrations.

* In Figure 1 the Black Sea shall be considered as part of Zone 4.

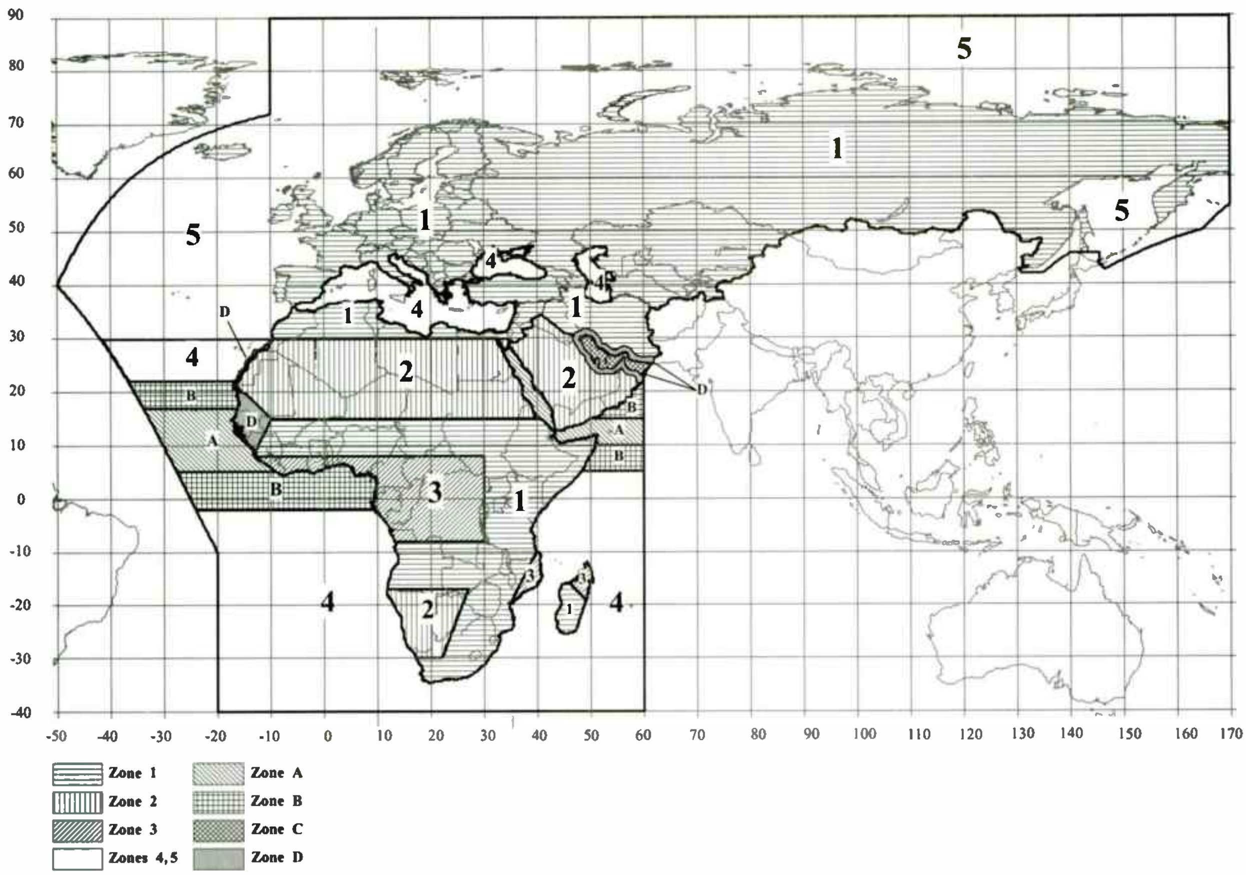


FIGURE 1 – Geographical division of the planning area into propagation zones

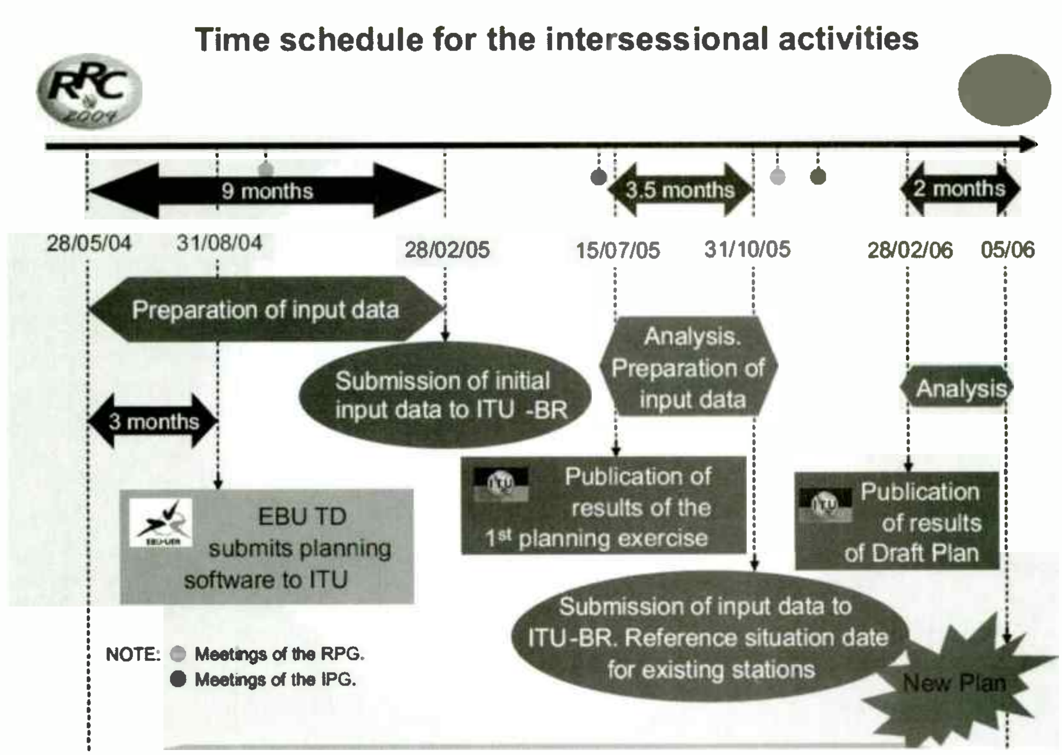


FIGURE 2 – Time schedule for main activities during the intersessional period (Source EBU)

- Regulatory and Procedural Group (RPG) also consisting of all participating countries, responsible for all regulatory and procedural questions and developing a draft text for a treaty agreement to be adopted by the second session.
- Planning Exercise Team (PXT) consisting of ITU staff and 11 planning experts from all regional groups in the planning area, responsible for implementation of the planning software system appropriate for large scale planning exercises, carrying out the planning exercises and producing a draft frequency plan.

The Planning Process and Software

The planning process (Figure 3) consists of the following steps:

- *Step 1*: submission of the input requirements for digital broadcasting;
- *Step 2*: identification of the analogue broadcasting stations and of other services that need to be taken into account;
- *Step 3*: performance of compatibility analyses;
- *Step 4*: assessment of the results from Step 3;
- *Step 5*: allowance for administrative input concerning compatibility between requirements, with a return to Step 3 if necessary;
- *Step 6*: performance of synthesis, the output of which is a plan;
- *Step 7*: review of the results, with a loop back to Step 5 and then to Step 3 if the desired result is not achieved;
- *Step 8*: adoption of the final plan.

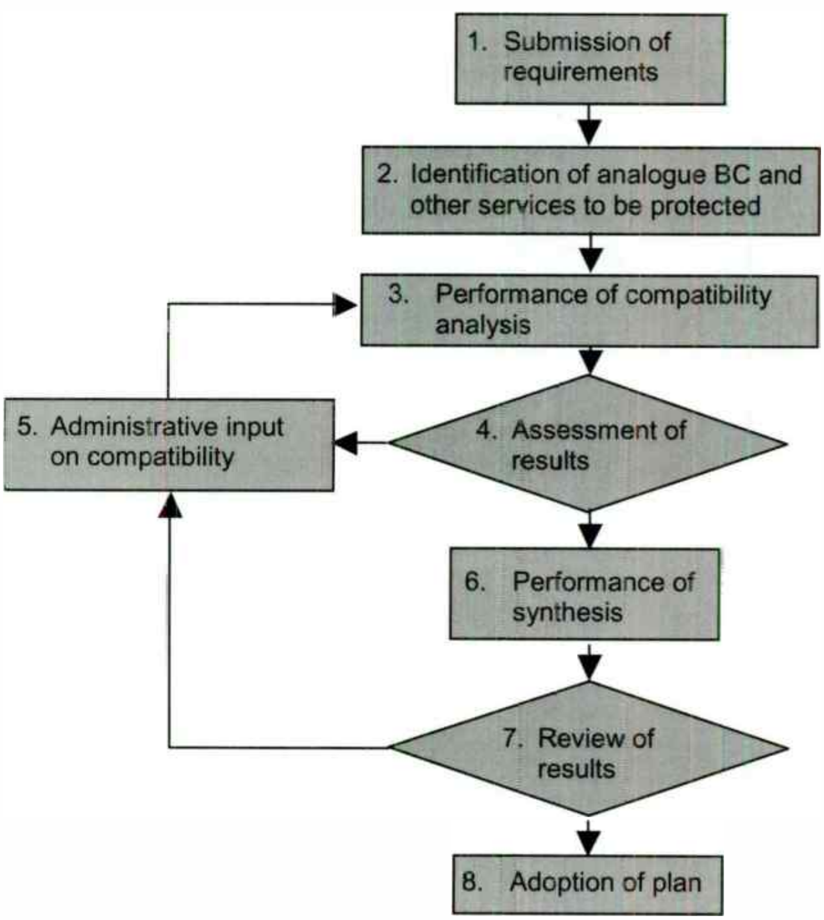


FIGURE 3 – Planning Process

The planning process can be split into "compatibility analysis" and "synthesis" stages.

Compatibility analysis calculates the required protection in order to determine which digital requirements may or may not share any given channel and which may or may not use any given channel. This involved pair-wise calculations between digital requirements, as well as between the digital requirements and other primary services other than broadcasting to be protected.

The plan synthesis determines a suitable frequency for each requirement so that no harmful interference arises, either to existing or planned stations or to the requirements themselves. Thus the results of the compatibility analysis are inputs to the synthesis process:

- the channels which are available to each requirement;
- the incompatibilities between requirements.

The planning software including compatibility and synthesis module was developed by the European Broadcasting Union. Detailed description of the methodology of compatibility and synthesis software can be found at http://www.itu.int/ITU-R/conferences/rrc/rrc-04/intersession/ipg/elec_pub_IPG-1/index.html

Coordination between Administrations

With the technical bases established in 2004, the planning activities began in earnest. For the subsequent two years there was a tremendous amount of work by all administrations in coordinating and negotiating their digital requirements with their neighbouring countries. Coordination activities were carried out in bilateral or multilateral groups of countries, for example in Europe, there were more than 16 of coordination groups established.

Regional organizations such as Arab States Broadcasting Union (ASBU), African Telecommunication Union (ATU), European Conference of Postal and Telecommunication (CEPT), European Broadcasting Union (EBU) and Regional Commonwealth in the field of Communications (RCC) played a major role in harmonizing the digital requirements in their own respective regions. Inter-regional meetings were also organised.

Results of negotiations between administrations are submitted as inputs to the planning process in the form of *administrative declarations* through which an administration effectively declares that two digital requirements or a digital requirement and an assignment relating to analogue television or other primary service are compatible even though the compatibility calculations indicate otherwise.

Implementation of the Planning System

The input data to the planning system consists of:

- Over 70,000 digital requirements of allotments and assignments.
- Nearly 100,000 assignments of existing and planned analogue television.
- Over 10,000 assignments relating stations of fixed and mobile services (other primary services other than broadcasting).
- Over 3,000,000 administrative declarations resulting from coordination and negotiation between administrations.

The processing of the input data, in particular the calculations of compatibility analysis requires a substantial amount of computer CPU time in the order of 90 days. In the first planning exercise carried out in June 2005, the whole process took 19 days using 6 dedicated high speed computers and 20 desk top PCs.

The idea of distributed computing was investigated as a possible solution for carrying out the planning calculations within a target of two days so that during the second session of the conference, at least four iterations of the frequency plans could be produced.

For the second session, two independent distributed computing systems were implemented to provide additional flexibility and reliability of the planning system

- ITU's distributed computer system consisting of 100 high speed (3.6 GHz) hyper-thread PCs, capable of running 200 parallel jobs. The system is managed by a client server system that assigns jobs, monitors the status of the running of jobs and retrieves results as soon as the calculations are completed. For each planning run, about 200,000 jobs are assigned to the distributed computer structure.
- The European Organization for Nuclear Research (CERN)'s computer grid structure (small part) using more than 300 PCs located at its member institutions in Germany, Russia, Italy, France and Spain. Its client-server system called DIANE that allows run-time load balancing, access to heterogeneous resources (Grid and local cluster at the same time) and a robust infrastructure to cope with run-time problems. DIANE allowed using in the most effective way the available resources since in this system each available worker nodes asks for the next task: while a long task will "block" a node, in the mean time the short tasks (the large majority) will flow through the other nodes.

RESULTS OF THE SECOND SESSION

The two important achievements of the conference are:

- the digital frequency plan that satisfies substantially the requirements of the participating countries and that provides a structured and equitable framework for future development of digital terrestrial broadcasting services for sound and television, and
- the treaty agreement adopted by all countries in the planning area that provides transparent, effective and flexible regulatory procedures that not only maximize the use of the radio frequency spectrum but also provides the flexibility to adapt in the changing technology and regulatory environment.

The Frequency Plans

The finally adopted digital frequency plan contains 8,379 T-DAB allotments in Band III and 62,112 DVB-T assignments and allotments in band III, IV and V. The percentages of the requirements that were submitted by the participating countries and were assigned with a frequency channel are 93% and 98% for Band III and Band IV&V respectively.

The results of the four planning iterations carried out during the second session are given in Tables 4 and 5.

Table 4 – Satisfied requirements in Band III

	Fourth Iteration			Third Iteration			Second Iteration			First Iteration		
	T-DAB	DVB-T	Total	T-DAB	DVB-T	Total	T-DAB	DVB-T	Total	T-DAB	DVB-T	Total
Total	8817	7411	16228	9061	7309	16370	10446	8402	18848	10755	8341	19096
Assigned	8379	6703	15082	8037	5745	13782	8151	5599	13750	7502	4831	12333
%	95.0	90.4	92.9	88.7	78.6	84.2	78.0	66.6	73.0	69.8	57.9	64.6

Table 5 – Satisfied requirements in Band IV&V

	Fourth Iteration	Third Iteration	Second Iteration	First Iteration
Total	56533	55876	60227	62692
Assigned	55409	52229	51222	46333
%	98.0	93.5	85.0	73.9

The improvements of the results (the percentage of number of satisfied requirements) through the successive iterations were due to a number of factors:

- reductions of the number of requirements submitted by participating countries. About 15% reduction in the number of requirements submitted for the fourth, compared to the first iteration;
- increases in the number of administrative declarations submitted as a result of continual negotiation between the participating administrations.

Generally, the majority of participating countries obtained three (3) nation-wide T-DAB coverages and one (1) nation-wide DVB-T coverage in Band III. For Bands IV and V, most of the participating countries obtained seven (7) nation-wide DVB-T coverages. As expected the results were slightly below the average level in several areas such as southern the Mediterranean countries (Algeria, Italy, Libya, Malta and Tunisia) or in the Gulf (Bahrain, Oman, United Arabs Emirates, Qatar, Iraq and Iran) because of the effect of long distance propagation due to super refraction and ducting.

Channel Usage and Reception Mode

Analysis carried out by the European Broadcasting Union (EBU) concerning the channel and reception mode usage is shown in Figures 4 to 7 and reveals some interesting general observations:

- For T-DAB the main mode of reception is mobile.
- For DVB-T television service, fixed reception represents the main mode of reception. Usage of portable indoor is negligible.

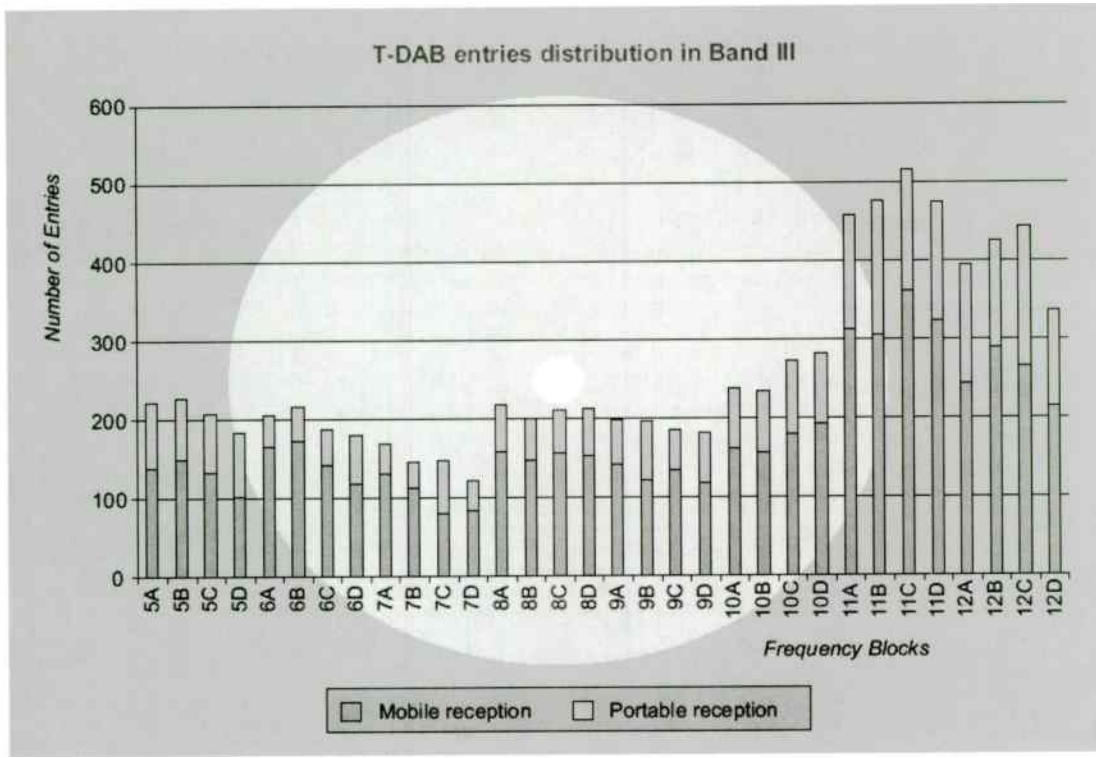


FIGURE 4 – Channel and reception mode usage for T-DAB requirements in Band III (Source EBU)

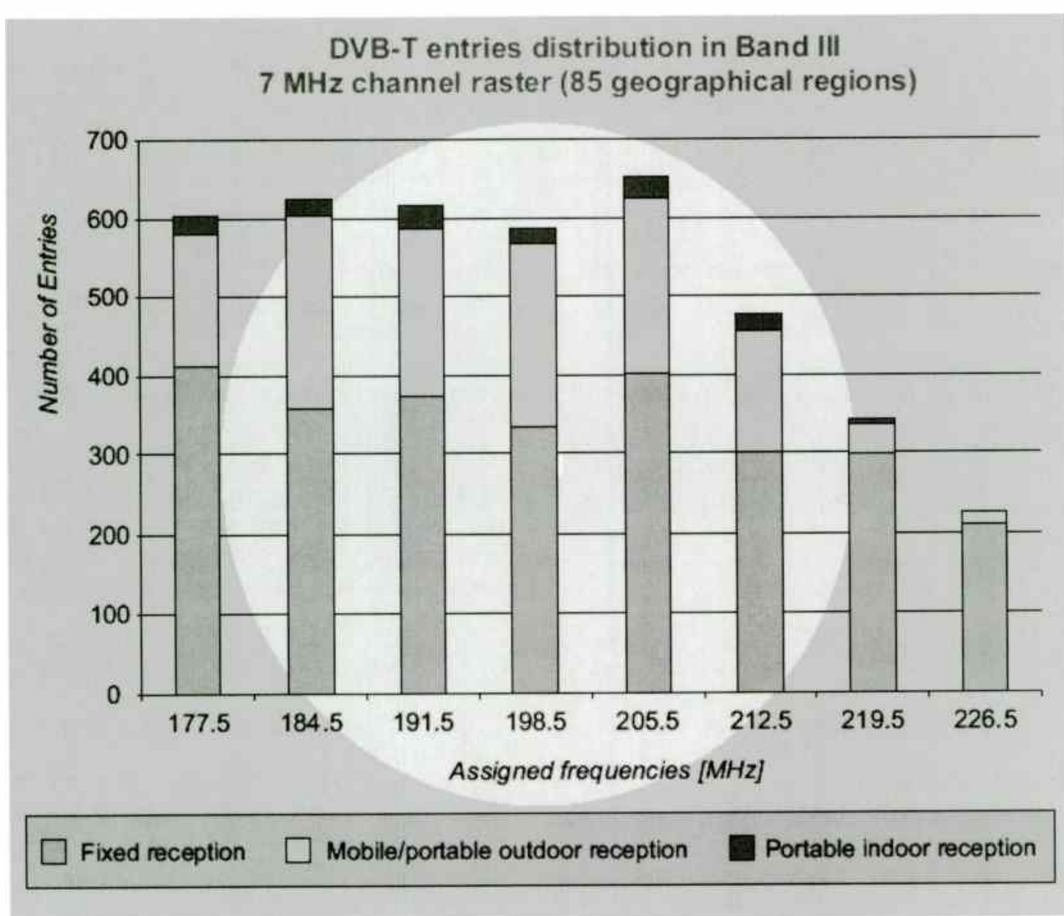


FIGURE 5 – Channel and reception mode usage for 7 MHz DVB-requirements in Band III (Source EBU)

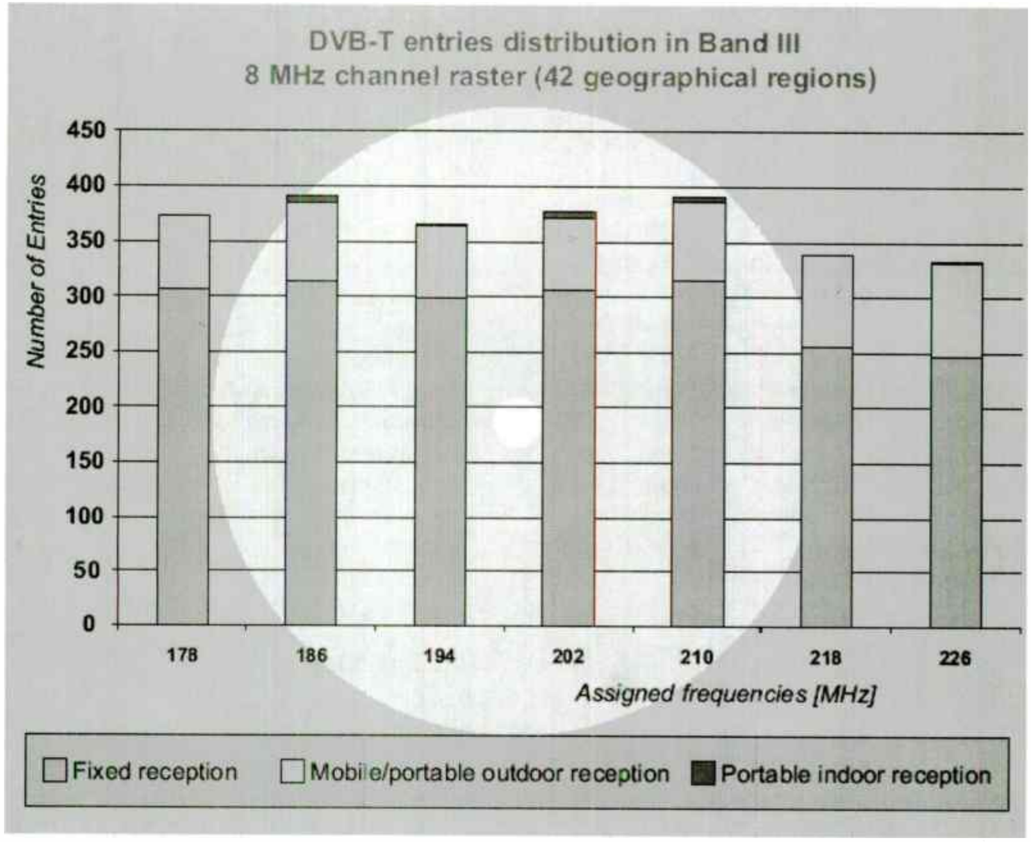


FIGURE 6 – Channel and reception mode usage for 8 MHz DVB-requirements in Band III (Source EBU)

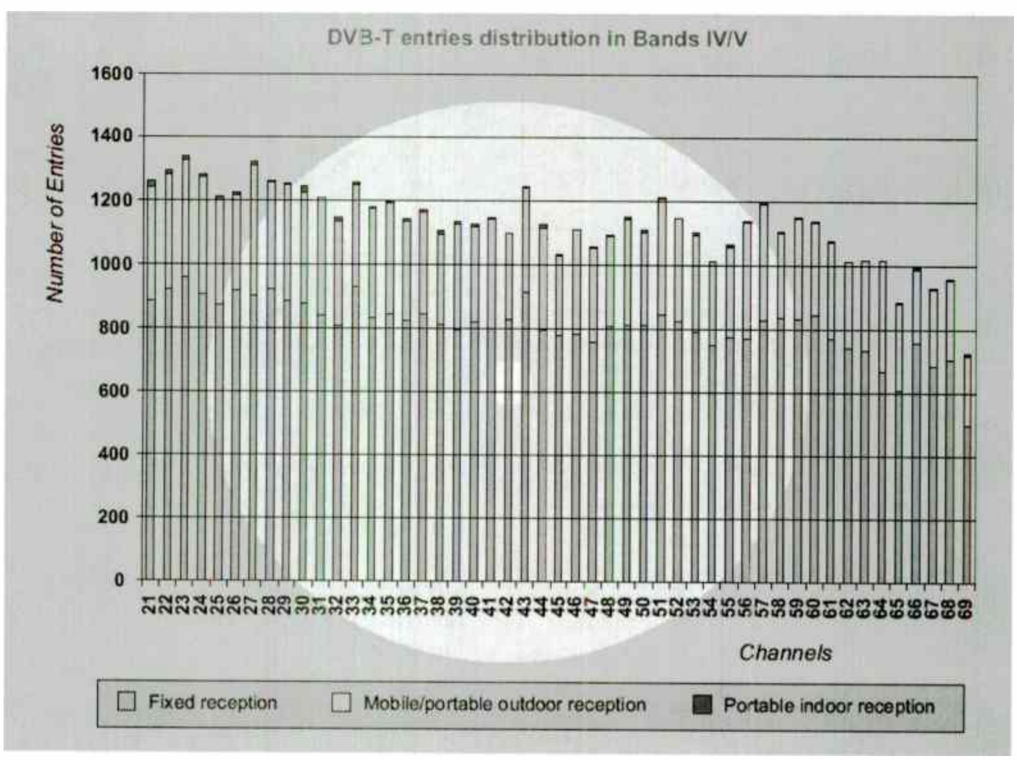


FIGURE 7 – Channel and reception mode usage for DVB-requirements in Band IV and V (Source EBU)

Flexibility - Use of a Digital Entry for Broadcasting and Other Services Applications

In establishing the Geneva 2006 frequency plan (GE06) and the Regional Agreement for digital broadcasting services in the frequency bands 174-230 MHz and 470-862 MHz, the Regional radiocommunication Conference (RRC-06) acknowledged the importance of having a regulatory framework that is responsive, flexible and long-lasting in the face of changing technological developments, future demands of spectrum from other uses and changing communication policies from member states in the planning area.

After much discussion and negotiation, the Agreement in its Article 5 under provision 5.1.3 allows for the use of a digital entry for broadcasting and other services applications with characteristics different from those appearing in the Plan:

" 5.1.3 A digital entry in the Plan may also be notified with characteristics different from those appearing in the Plan, for transmissions in the broadcasting service or in other primary terrestrial services operating in conformity with the Radio Regulations, provided that the peak power density in any 4 kHz of the above-mentioned notified assignments shall not exceed the spectral power density in the same 4 kHz of the digital entry in the Plan..... ".

The provision permits an administration, if it so wishes; to notify to the Bureau and bring into use a radiocommunication service that is neither a DVB-T nor a T-DAB broadcasting service, but it uses the spectrum envisaged for a DVB-T or T-DAB assignment/allotment recorded in the GE06 digital Plan. An example of the use of the envelope or mask concept is given in Figure 8.

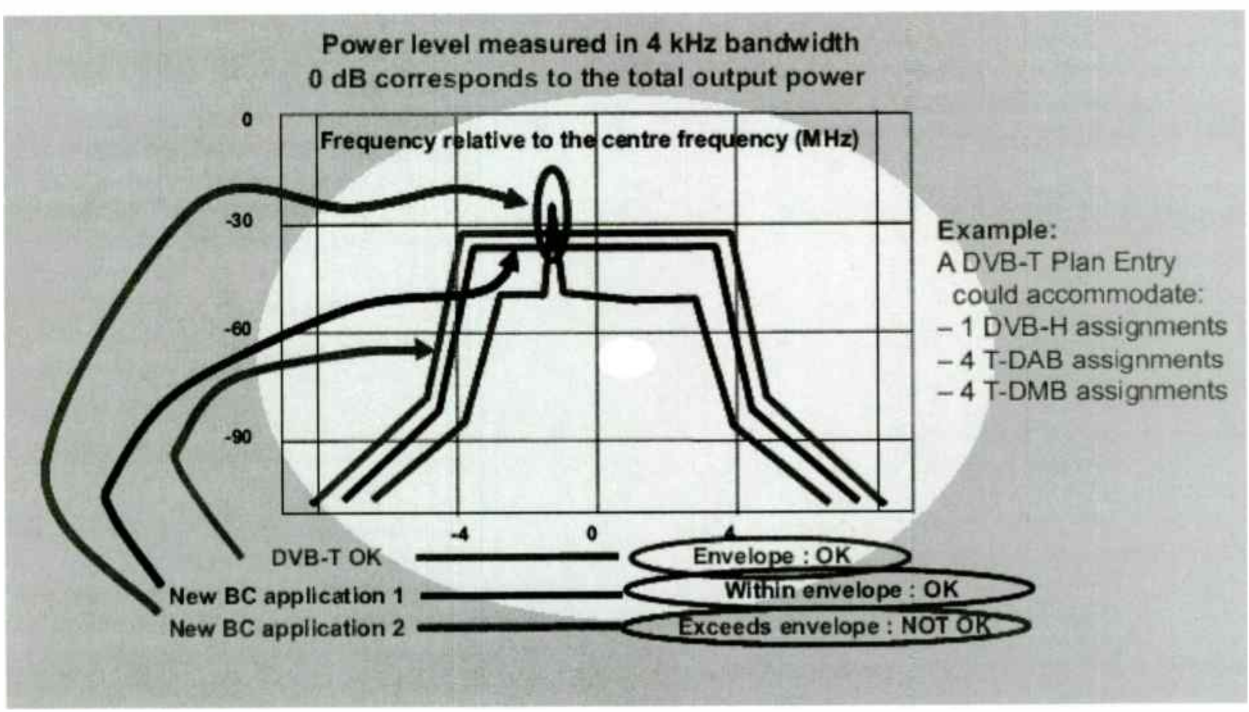


FIGURE 8 – The " envelope" or " mask" concept (Source EBU)

The Digital Dividend

The switchover from analogue to digital broadcasting will create new distribution networks and expand the potential for wireless innovation and services. The digital dividend accruing from efficiencies in spectrum usage will allow more channels to be carried across fewer airwaves and lead to greater convergence of services.

The inherent flexibility offered by digital terrestrial broadcasting will support mobile reception of video, internet and multimedia data, making applications, services and information accessible and usable anywhere and at any time. It opens the door to new innovations such as Handheld TV Broadcast (DVB-H) along with High-Definition Television (HDTV) while providing greater bandwidth to existing mobile, fixed and radionavigation services. Services ancillary to broadcasting (wireless microphones, talk back links) are also planned on a national basis and need to be extended.

The World Radiocommunication Conference (WRC-07), which will meet in the autumn of 2007, will eventually review the regulatory aspects of the usage of the spectrum for these services, based on contributions from the Member States.

Terrestrial digital broadcasting carries many advantages over the analogue system:

- Expanded services
- Higher quality video and audio
- Greater variety and faster rates of data transmission
- Consistency of data flows over long distances
- More spectrum efficiency means more channels.

The Transition Period

The Conference agreed that the transition period from analogue to digital broadcasting, which began at 0001 UTC 17 June 2006, should end on 17 June 2015, but some countries preferred an additional five-year extension for the VHF band (174-230 MHz).

At the end of the transition period, countries may continue to operate analogue broadcasting stations provided that these stations do not cause unacceptable interference to and do not claim protection from any assignments in conformity with the Agreement and its associated Plans.

CONCLUSIONS

The first session of this Conference (RRC-04) took place in May 2004 and established a solid, comprehensive and technical basis for establishing the agreement, including the framework for the intersessional studies. It has already resulted in the accelerated introduction of digital terrestrial broadcasting in many countries.

During the five weeks of deliberations which began on 15 May, RRC-06 took decisions to allow iteration of the complex software tools used by the ITU secretariat as a basis to generate the draft plan that will facilitate the coordinated and timely introduction of digital broadcasting. The Plan assures that an outstanding 70'500 digital broadcasting requirements could operate in a compatible manner. It succeeded in creating a level playing field as a new basis for competition.

A key ingredient for the success of the Conference was the unprecedented level of cooperation between ITU, the European Broadcasting Union (EBU) and the European Organization for Nuclear Research (CERN).

The complex planning activities conducted at this conference and during the intersessional period were based on the software developed by EBU, which includes hundreds of thousands of programme lines. In preparing the Plan for digital terrestrial broadcasting, ITU experts performed meticulous calculations within a limited timeframe using two independent infrastructures: the ITU distributed system with 100 PCs; and the CERN Grid infrastructure that is based on a few hundred dedicated CPUs from several European institutions.

The regional agreement for digital services has been reached in the frequency bands 174 - 230 MHz and 470 - 862 MHz in Europe, Africa, Middle East and the Islamic Republic of Iran. It marks the beginning of the end of analogue broadcasting.

Multicasting for TV

Thursday, April 19, 2007
9:00 AM – 10:00 AM

Chairperson: John Merrill
CBS 5 KPHO-TV, Phoenix, AZ

From Statistical Multiplexing to Blended Remote Statistical Multiplexing

Pascal Gravoille, Grass Valley, Rennes, France

The Advanced Service for Terrestrial DTV Multicasting

Jeong-Deok Kim, KBS DTV Service Development, Seoul, Korea
Jung-Hee Kim, KBS, Seoul, Korea

FROM STATISTICAL MULTIPLEXING TO BLENDED REMOTE STATISTICAL MULTIPLEXING

PASCAL GRAVOILLE

GRASS VALLEY
RENNES, FRANCE

OVERVIEW

Until recently, Digital TV systems have been based on a very centralized architecture (see figure 1) where all TV channels are hauled back to the Headend at high bit rate and recompressed locally using Statistical Multiplexing techniques.

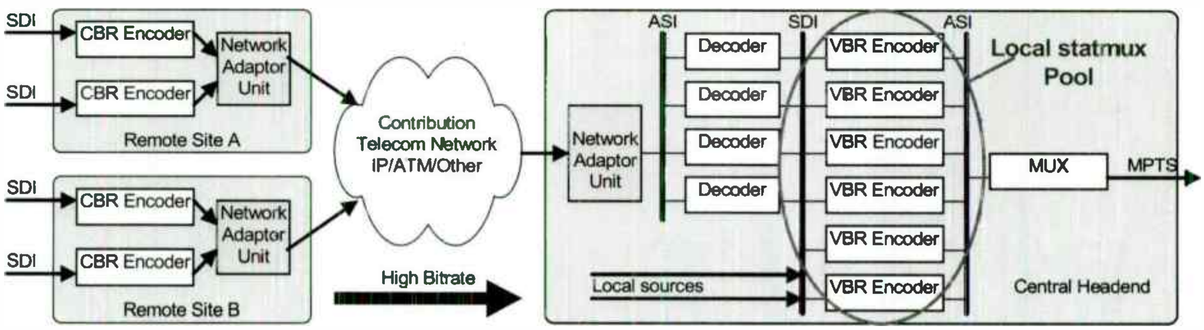


Figure - 1

Hauling back the TV channels to the Central Headend costs money both on equipment inventory, and on operation expenses. 4:2:2P or high bit rate 4:2:0 encoders are required at the remote location, and local decoders are necessary in the Central Headend in order to deliver high quality baseband video to the Statistical Multiplexing pool of encoders.

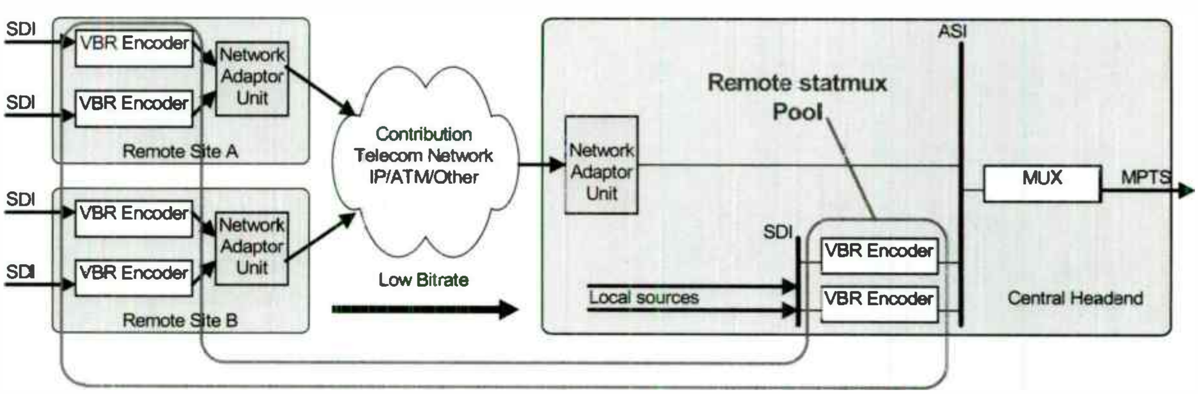


Figure - 2

Furthermore, a high bandwidth pipe must be leased from a third party company to transport the high bandwidth TV channels. On the other hand, decentralizing the encoders and relocating them on the remote sites faces the challenge of implementing Statistical Multiplexing with geographically distributed equipments.

Grass Valley has developed a new flavor of Statistical Multiplexing solution named Remote Flextream™ enabling distributed compression Headend architectures as depicted in figure 2.

This patented technology offers the following advantages:

- Encoders belonging to the same Statistical Multiplex pool can be distributed either at the central Headend collocated with the multiplexer, or at remote sites.
- High bitrate codecs are no longer necessary, leading to some CAPEX savings.
- Video being typically transported at 2Mbps average (MPEG2 SD) from remote sites to the Central Headend instead of a typical 8Mbps for back hauling, telecom pipes can be smaller, leading to OPEX savings.

- Overall compression performance is optimized since Remote Statisctical Multiplexing avoids cascading MPEG compression segments (Backhaul plus Headend).
- The video network can be IP, PDH, ATM, ASI or even any mix of them.

In addition to these short term advantages, this Remote Statistical Multiplexing concept is revisiting the way to manage the bit rate in a Digital TV system and opens the door to a new concept for Headend architecture.

MAIN ISSUE: SYNCHRONIZATION

Statistical multiplexing

The main target of the Statistical Multiplexing technique is to match component bitrate with the encoded video complexity. The outcome is a bandwidth optimization that raises the opportunity to insert more TV channels in a single Transport Stream. The result is a variable bit rate allocation for each TV program as shown in figure 3.

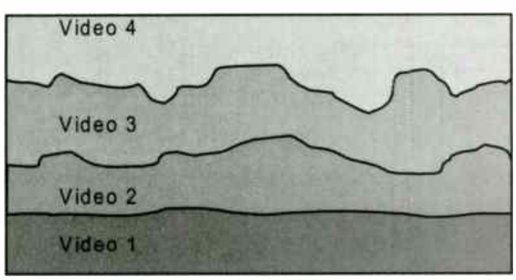


Figure - 3

Figure 3 is a macroscopic view of the result and the real transport stream actually looks more like depicted in figure 4 where all bit rates are constant during a short period. A reasonable duration period for bit rate modification is in the range of video field (16-20 ms).

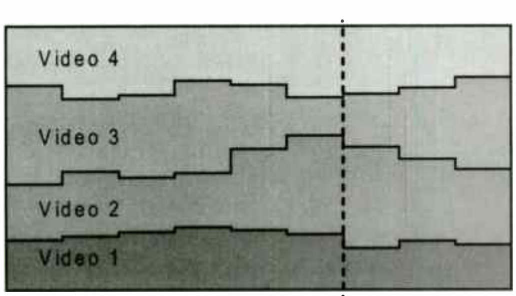


Figure - 4

In such a mechanism, synchronization is critical otherwise the total video rate will not be constant as shown in Figure 5 where Video 3 is desynchronized. Extra null packets will be required to ensure that video can be safely multiplexed and the advantage of the statistical multiplexing will be lost.

For a transport stream at 40Mbit/s, packet duration is 0,037ms. Considering a reasonable capacity to manage an extra burst of 16 packets and the Transport Stream Buffer specified by the ISO standard, the main requirement for a Statistical Multiplexing system is to synchronize video with accuracy in the range of 0.5ms.

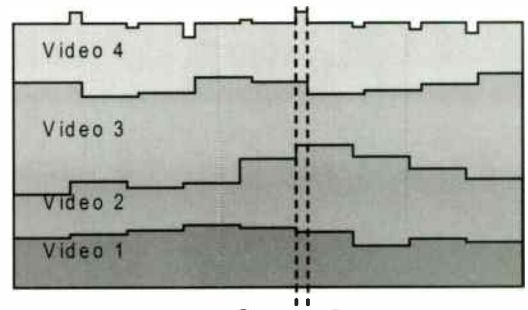


Figure - 5

Synchronization mechanism

As long as all encoders were co-located on the same site, a very simple and basic mechanism like a distributed signal was usable. Classical *Flextream*™ system is using an HDLC link to distribute statistical multiplexing related signalling and a synchronization signal. However, in a distributed architecture, the synchronization mechanism has to be reconsidered.

The very first idea that comes to mind is to synchronize all encoders. A first solution based on NTP could be elaborated but the accuracy of an NTP based synchronization is 10ms, far from our goal of 1ms. A GPS based solution gives accuracy in line with the expected 0.5ms but requires extra devices leading to a much more complicated architecture.

In addition, the real need is to have synchronized stream at the multiplexer input. The network latency will differ from site to site, and streams synchronized on departure from the remote sites will very likely be desynchronized upon arrival at the Central Headend site. Another innovative solution enables to tackle this synchronization issue.

In that mechanism, remote device synchronization is not required and each of them will have its own time base. It is just required that the multiplexer will maintain the time difference between its own time base and each encoder's time base. This will be achieved thanks to a PCR based synchronization mechanism.

PCR based synchronization mechanism

By enforcement of the MPEG2-ISO standard each encoder has to manage a 27 MHz Program Clock Reference (PCR). The MPEG2 standard provides a

mechanism enabling the decoder to synchronize to the same clock as the encoder so that it can properly display the decoded video. This mechanism relies on MPEG2 packets that contain Encoder PCR value samples which are regularly inserted in the MPTS signal (typically every 40ms).
When such a packet is flowing through a multiplexer, it is temporally disturbed and the embedded PCR shall be updated accordingly. As a consequence, by construction, a multiplexer has also to manage a Program Clock Reference. This PCR shall be based on a very accurate clock in order to minimize PCR jittering. The PCR in the multiplexer is of the same nature as the encoder one. It will be usable as a reference for the whole system and finally elects the multiplexer as the system synchronization master of a Remote Statistical Multiplexing system.
That PCR based synchronization mechanism will be used for time stamping all exchanges between the devices involved in a Remote Statistical Multiplexing system.

IMPLEMENTING REMOTE STATISTICAL MULTIPLEXING

Statistical Multiplexing principle

The Statistical Multiplexing technique relies on the following mechanism:

1) each encoder evaluates the complexity of a video sequence of images to compress,
2) each encoder sends its complexities to an allocator,
3) the allocator computes all the complexities and defines a bit rate for each encoder,
4) the allocator sends back the bit rate to each encoder,
5) the encoders use the allocated bit rate to compress the video sequence.

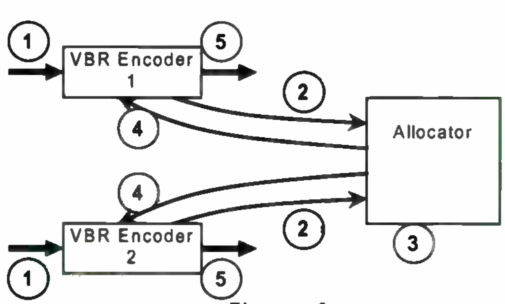


Figure - 6

This description highlights a new function "*the allocator*". In a classical statistical multiplexing architecture, the allocator function can be supported by any kind of device. Among others, one of the encoders is elected as the allocator. In a Remote Statistical Multiplexing system, every exchange (complexity and allocated bit rate) shall be timestamped by a reference based on PCR leading to the major requirement: the allocator function shall be supported by the multiplexer.

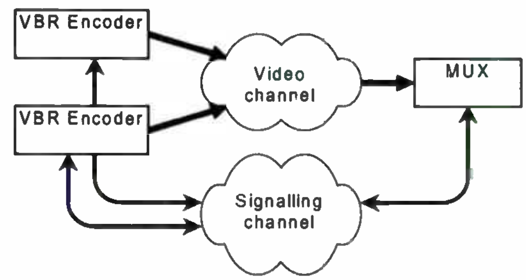


Figure - 7

It also highlights two different natures of data to be exchanged between the Remote Site and the Central Headend as depicted in figure 7:

1. The video stream that uses a contribution telecom network; let's call it the video channel
2. The Remote Statistical Multiplexing related data which will use a dedicated network; let's call it the signalling channel.

The video channel

The video channel uses a uni-directional contribution telecom network. It can be supported by any type of media as long as it enables implementation of a network with a constant delay between both sites. When the media is by essence asynchronous, non deterministic and/or not error-free, Network Adaptor devices with efficient dejittering function and advanced error correction mechanisms should be set up to restore a constant video network delay and a correct quality of service.
The video channel is not impacted by a Remote Statistical Multiplexing implementation and the existing Telecom Contribution network, if any, can be used as is today.

The signalling channel

The signalling channel is bidirectional and enables dialog between encoders and multiplexers. The requirement for this channel is less stringent than for the video one as jittering can pretty well be tolerated. The only main requirement is an acceptable PING time between sites (typically in the range of 100ms-300ms). It is typically an IP based network that offers the most comfortable protocol family support for an efficient and reliable signalling channel implementation.
In most cases such a network already exists since remote sites will be partly or fully controlled from the Central Headend. Extension to Remote Statistical

Multiplexing signalling exchange is not a big deal and mainly a matter of bandwidth.
If such a network doesn't exist, an alternative to setting up a new link is to consider an extension of the Video Telecom Contribution. Building a new bidirectional channel able to support an IP link using that way is a cost effective solution.

The Remote Flextream signalling protocol

The signalling channel is one of the major differences with regular Flextream system. In legacy Flextream system, exchange is done using an HDLC bus and exchange duration is perfectly mastered and deterministic. In a Remote Statistical Multiplexing system, exchanges are asynchronous and rely on UDP protocol.
All the process of complexity_sending/bit_rate allocation/allocated_bit_rate_sending shall be done within a maximum permitted time. If not, the encoder will not be able to encode the video sequence corresponding to the complexity with the right allocated bit rate. Because the time validity for both inputs complexity and allocated bit rate is short, the non connected UDP protocol has been selected for the Remote Flextream signalling protocol implementation.
The drawbacks of UDP are that messages can either be lost or arrive with a too long delay. Both multiplexers and encoders shall implement fall back modes based on predefined default bit rate in order to take into account that respectively, complexity and allocated bit rate are missed.
Minimizing this drawback is achieved by adding information redundancy is the exchanges. A good engineering for the signalling channel is also required to limit the transmission error and packet loss.

Managing the redundancy

The Remote Statistical Multiplexing architecture should take into account from the ground up the devices redundancy. Usually both encoder and multiplexer redundancy is achieved through a N+P redundancy mechanism.
Two solutions can be foreseen: a Unicast based solution, a Multicast based solution.

The Unicast solution consists in connecting all encoders to all multiplexers and vice versa (see figure 8). Such a solution is easy to deploy and all existing networks will support it. However some drawbacks can be highlighted:

1. The encoder shall be aware of which multiplexer is main and which is redundant.
2. The encoder shall send to more than one multiplexer its complexities, thus increasing the complexity exchange rate.
3. The multiplexer shall be aware of which encoder is main and which is redundant in order to send the relevant bit rate to the right one.
4. Redundancy in effect at one site shall be reflected in the other site.

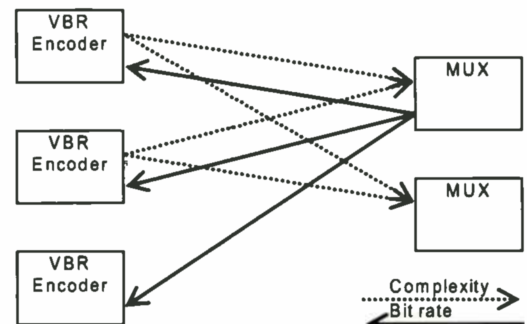


Figure - 8

The Multicast solution consists in using two multicast addresses, one for each direction (see figure 9). A multicast address can be compared to a mail box where a sender drops messages that receivers can read. Each encoder posts its complexity to one multicast address so that both multiplexers (main and redundant) will read it. On the other way round, the main multiplexer sends the allocated bit rates to a second multicast address so that encoders including redundant ones can read them.

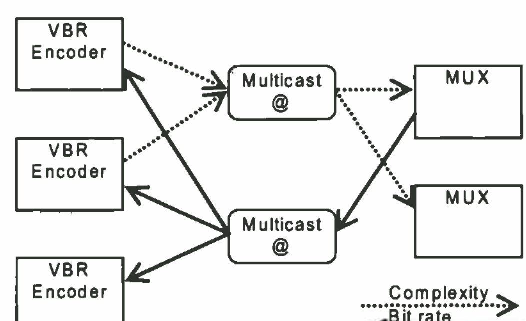


Figure - 9

The Multicast solution dramatically simplifies the redundancy mechanism since encoders and multiplexers' redundancy can be managed separately. This clearly outweighs the few drawbacks that can highlight for the sake of completeness:

1. A network supporting multicast requires a significantly more sophisticated engineering.
2. All encoders receive the bit rates for all encoders and message parsing is required to extract the appropriate encoder allocated bit rate.

USING REMOTE STATISTICAL MULTIPLEXING

Remote Flextream relies on a very powerful PCR synchronization mechanism with several advantages:

- A theoretical accuracy of 1/27 000 000 sec
- Quick resynchronization (a packet with an embedded PCR is supposed to be received every 40 ms).
- No need for additional devices other than encoders and multiplexers.
- The support of any underlying networking technology mix for the contribution network in a single Remote Flextream pool
- In addition, an MPEG2 remultiplexing stage can be considered between the encoders and the multiplexer.

In addition, Remote Statistical Multiplexing concept relies on very basic assumptions & requirements about the network. It is likely that the needed infrastructure is already in place. The investment for building a Remote Statistical Multiplexing system is thus minimized. The additional building blocks are mainly software blocks and testing the Remote Statistical Multiplexing concept before deployment can be easily and quickly done.

In fact, the only drawback is that the multiplexer becomes an essential actor in Remote Statistical Multiplexing architecture whereas in a regular Statistical Multiplexing system, it was considered as neutral. But this inconvenience can in fact become a starting point for new services based on the multiplexer as a Bit Rate manager.

EXTENSION TO TRANSRATING

Transrating as a strong requirement in new architectures

More and more Digital TV feeds are created using statistical multiplexing techniques.

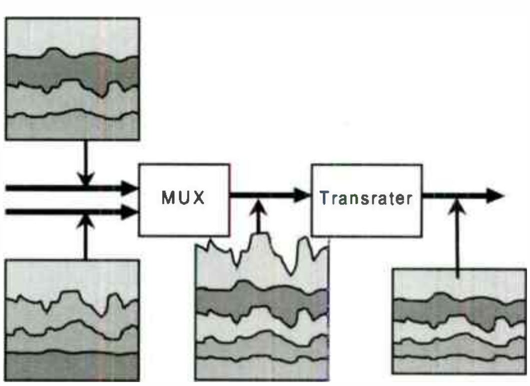


Figure - 10

Building a new Digital TV signal by picking up some variable bit rate services from different inputs (see figure 10) leads to a complex bit rate management that can only be solved through transrating mechanisms.

Transrating is the capability of a multiplexer to slightly reduce each incoming video bit rate if required so that the aggregate signal bit rate will be in line with the multiplexer maximum output bit rate.

A transrater device is composed of a collection of basic transrater engines that individually manage the video bit rate of a single component. Building a complete transrater device requires interconnecting a collection of basic transrater engines with a bit rate allocator (see figure 11).

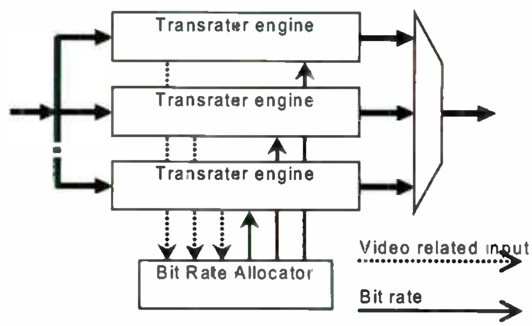


Figure - 11

The basic transrater engines analyze the incoming video sequence and determine a video related input metric such as the aggressiveness of the upstream video encoding process. This metric is sent to the allocator that finally allocates bit rates based on this basis.

Such a mechanism is very close to the genuine Remote Statistical Multiplexing Concept one, and one can consider merging them into a single one.

The major problem of merging both mechanisms is that the allocator has to process two inputs (complexity & aggressiveness) which are not of the same nature. The complexity represents the intrinsic complexity of the video sequence whereas the aggressiveness is more related to the way the video has been processed by the upstream encoder.

Let's take the following example: a single video stream is feeding two encoders. Each one is encoding the video with two different bit rates. Then both compressed videos enter a transrater device and are merged into a single transport stream. The expected result is that the allocator should allocate the same bit rate to both videos since they are identical. So from a transrater standpoint, the high bit rate compressed should be more reduced than the low one. The two aggressiveness measures from the transraters are different whereas the complexity raised by the encoder is identical.

Nevertheless one can consider merging both allocators as soon as the differentiation of input nature is taken into account.

Blended Remote Statistical Multiplexing

The natural extension to the Remote Statistical Multiplexing architecture is the Blended Remote Statistical Multiplexing concept where all video processing (encoding, transrating, but also the newcoming video transcoding building block) are under control of a unique Bit Rate manager as shown in figure 12.

PSI/SI signalling, EPG, Teletext, receiver firmware download). The bit rate control extension to that type of data can be of a great benefit for bandwidth optimization.

Multiplexers are considering data injection through the use of two mechanisms: regular data injection, opportunistic data injection. Regular data component is allocated a constant bit rate whereas opportunistic data component uses the remaining bandwidth.

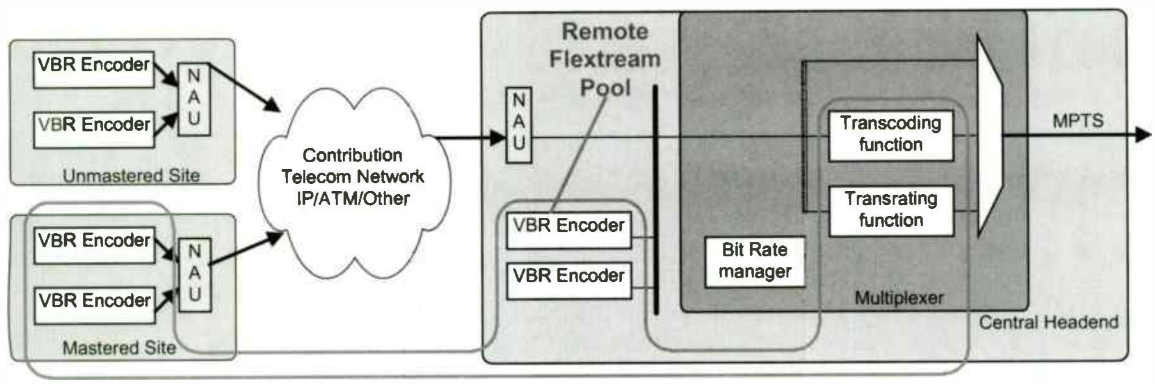


Figure - 12

The Blended Remote Statistical Multiplexing concept allows the operator to have a global view of the video rate and brings all the knobs to finely tune and balance the quality vs the bandwidth.

Thanks to this new concept, it is now possible for an operator to mix in the same bouquet some turnaround channels with its own direct encoded channels. Without this new concept, there were only two alternatives: either to decode and re-encode all the turnaround channels or to transrate all the channels. Now, the bandwidth use (and the Head-End cost!) can really be optimized without sacrificing the video quality. This solution is applicable to almost every segment of the TV distribution industry. In DTTV, the Blended Remote Statistical Multiplexing provides a way to take advantage of statistical multiplexing even when national and regional programs are broadcast in the same multiplex. For satellite operators, this new concept is particularly interesting when turnaround channels are broadcast in the same bouquet as locally encoded programs. Finally, this solution is also a great help for cable operators who want to buy decoders and encoders for the only few premium channels deserving this investment. They can simply transrate the rest of the programs and still take advantage of working in closed loop and saving about 20% of bandwidth per multiplex.

TOWARDS A UNIQUE BIT RATE MANAGER

Video is not the only component that could be controlled by a Bit Rate manager. The multiplexer is also used for ancillary data injection (private data,

Including such components into the allocation process gives additional opportunities for bit rate optimization. Moreover, playing with relative priorities between components (audio/video as well as data component) raises the opportunity for the operator to focus on some of them temporarily or permanently.

Audio components could also be subject to Bit Rate changing especially through transcoding processes. Although, today's audio standard doesn't support Variable Bit Rate, Audio bit rate will be an important input for bit bate management and allocation.

The multiplexer with its embedded Bit Rate Manager becomes the key point for a centralized and global bit rate management. It brings full control of the bit rate control to the Central Headend operator.

THE MULTIPLEXER AS A CENTRAL DEVICE

The *Remote Statistical Multiplexing Concept* technology highlights the central function of the multiplexer in up to date Headend architectures. Advanced multiplexers with embedded video processing capabilities and sophisticated bit rate manager will be in all likelihood the focal point of future Digital TV systems.

Nevertheless, this new coming role of the multiplexer is not a strong revolution but should rather be considered as a simple evolution. Architectures already in place can be used as such without any major modification. Each improvement in the central Headend can be easily made and tested before use on air and in the end with risk and cost minimization.

The Advanced Service for Terrestrial DTV Multicasting

Jeong-Deok Kim, Jung-Hee Kim, Woo-Jong Koh, Young-Sam Kim, Seong-Tae Kim, Jong-Soo Lee, Sun-Kyu Park
KBS
Seoul, Korea

ABSTRACT

The terrestrial DTV in Korea have been carried one channel of 1080i HD service in a frequency bandwidth of 6 MHz. The development of technology like MPEG-2 video compression, however, enables transmission of additional DTV channels; HD(1080i or 720p), SD(480i), audio and data channel the same bandwidth. Software Download Data Service (SDDS) can be also provided on air to meet the requirements of DTV receivers to reduce their improper operation. Therefore, a single HDTV channel expands to the service with the type of 'HDTV+SDTV+Audio+Data+SDDS'. This is an advanced multicasting service on terrestrial DTV and the service is called as the Multi-Mode Service (MMS). KBS has implemented the MMS system using advanced technologies in a DTV head-end system, the ACAP-based interactive TV system, the PSIP generating system and the SDDS system[5][8]. The technology for MMS has full compliance with ATSC standard which specifies different service types such as digital television, audio, data-only service and software download data service[1][2][3]. In addition, the MMS test streams with 21 modes are created to verify basic interfaces between the MMS transmission system and receivers for terrestrial DTV MMS broadcast[9][11].

This paper introduces the advanced service for DTV multicasting, a method of the service composition and technologies applied to the system, describing how to implement an end-to-end system to provide the MMS[10]. This paper also describes the detailed information of launching "a MMS trial broadcast" during the period of 2006 FIFA World Cup, and shows the MMS examples. In addition. this paper presents survey results from viewers after the MMS trial broadcast[4][6][7].

INTRODUCTION

The DTV MMS is fully compliant with the ATSC standard[1][2][3]. The ATSC standard specifies different service types such as digital television, audio, and data broadcasting. A 6-bit enumerated type field carried in a virtual channel table identifies the type of service as illustrated in Table 1. The service type 0x02 is defined as a digital television program that viewers usually see on a DTV. The services with types 0x03 and 0x04 are an audio program and a data service, respectively, from which the music or data can be obtained by television. Type 0x05 is a software download data service that enables a DTV receiver to upgrade its own software program from on air.

Table 1. Type of Service in DTV Virtual Channel

Service type	Meaning
0x00	[Reserved]
0x01	Analog_television
0x02	ATSC_digital_television (HDTV + SDTV)
0x03	Audio program, optional associated data
0x04	Data only service
0x05	Software download data service

There are new technical methods that are recently being applied to the DTV head-end system for bandwidth efficiency. The equipment that runs the system has many different functions such as noise-reduction, adaptive filtering, optimized rate control, and dynamic GOP. They also have more processors to run the complex operations system properly. We can improve the bandwidth by 15% from these multi-path encoding techniques.

The indispensable method for the MMS is however, to allocate bit rates to each service more efficiently. To do this, we specially use a statistical re-multiplexing technology between MPEG-2 encoders and a multiplexer that are based on Variable Bit Rate(VBR) techniques, which enable the head-end system to distribute bit rates among the encoders properly. In addition, the video format of HD is converted to 720p from 1080i for additional bandwidth efficiency. From this, we can get an additional 15% bandwidth efficiency. As a result, 30% of bandwidth efficiency has been achieved, which is 4~5Mbps in the ATSC environment.

It is possible to provide the MMS with the efficiency mentioned above which consists of not only HD channels but also additional SD, audio, and data channels. Fig. 1 shows the concept of bandwidth efficiency resulting from the technological development from a previous DTV head-end system to the present one. There exists a flexible area between HD and SD due to a VBR technical mode.

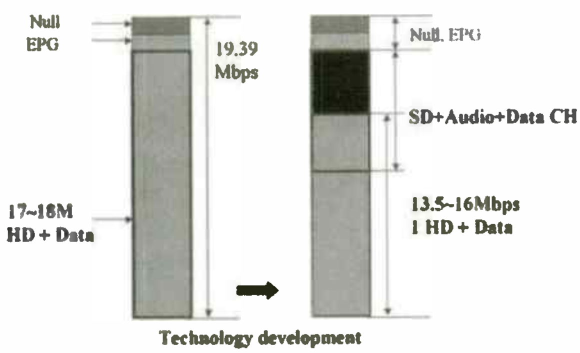


Fig. 1. Concept of Bandwidth Efficiency from Technology Development

MMS IMPLEMENTATION

"HD+SD+Audio+Data+SDDS" services

Fig. 2 illustrates the summary of the end-to-end system for the MMS that is being implemented by KBS. The system used to provide the MMS has several components such as the format converter, the remultiplexer, the SDDS server, and the PSIP/Data broadcasting server[5][7]. The system also has the HD/SD/Audio encoder.

After a HD program with a video format of 1080i is converted to 720p through a format converter, it is encoded and transferred to the remultiplexer. A SD program and an audio program are also encoded and transferred to the remultiplexer. The SDDS server encodes software download data to upgrade a DTV receiver and transmits the encoded data to the remultiplexer. The PSIP/Data broadcasting server is to provide data broadcasting service in compliance with ATSC ACAP and to extend one channel to several channels for MMS with the type, 'HD+SD+Audio+Data+SDDS'[10]. The remultiplexer not only processes the input streams as one transport stream but also plays an important role in statistical multiplexing among the encoders for bandwidth allocation properly. The generated data stream from the re-multiplxer is simply modulated and transmitted to broadcast. This procedure completes the MMS broadcasting cycle.

Table 2 describes an example of channel maps for MMS. The data broadcasting service in channels 9-1 to 9-4 is program-related and the data service in channel 9-5 is a data-only service that is not related with any A/V program. We may also transmit SDDS with an additional channel if necessary. This DTV MMS has full compliance with ATSC standard as shown in the Table 2[1][2][3].

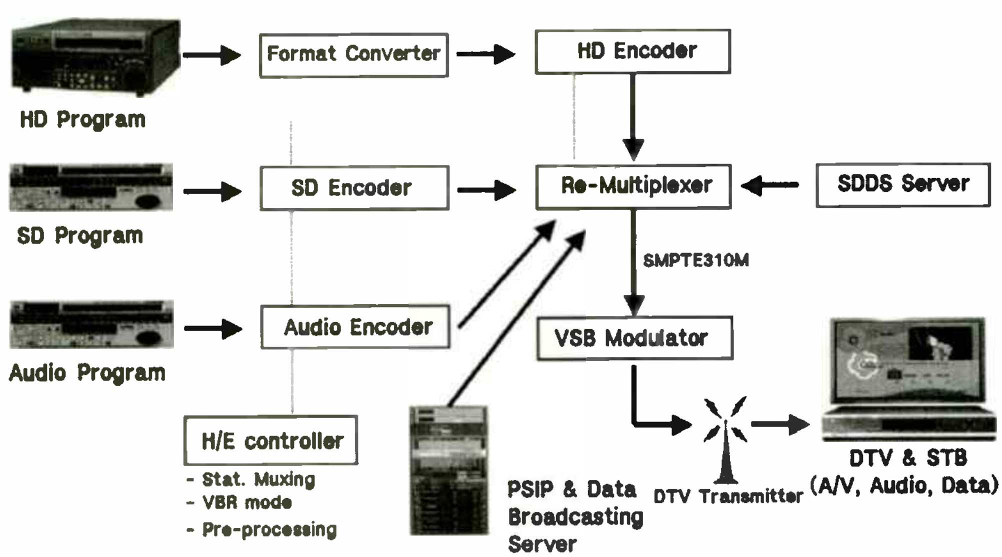


Fig. 2. MMS End-to-end System

Table 2. Channel Map of the MMS

Channel	Service	ATSC Standard	BW example (Mbps)	Total Bandwidth
9-1	HD + Data	A/53 + A/101(ACAP)	13.5~16	19.39Mbps (6MHz)
9-2	SD + Data	A/53 + A/101	2~4.5	
9-3	Audio1 + Data	A/52(AC-3) + A101	~0.2	
9-4	Audio2 + Data	A/52(AC-3) + A/101	~0.2	
9-5	Data only	A/101(ACAP)	0.2~	
9-6	Software Download Data Service (optional)	A/97	0~0.2	
PSI/PSIP(EPG) + NULL		A/65C	0.59~	

Fig. 3 depicts the resulting MMS diagram of real-time analysis for the MPEG-2 transport stream received from the remultiplexer in the MMS end-to-end system. There are five programs including HD/SD programs, two audio programs, and a data broadcasting service excluding SDDS. SDDS is an optional stand-alone service if it is necessary to be broadcast in the case of upgrading DTV receivers.

Software Download Data Service (SDDS) is to provide software data after encoding them using the method of the data carousel protocol that is compliant with ATSC A/97[2]. The system for SDDS is to transmit a transport stream including some specific data provided by manufacturers. Receivers decide whether the data is their own or not, and then they can update their own software after parsing the encoded data. Fig. 4 shows a concept of SDDS on the air, from which we can see that SDDS is a stand-alone service, so it can be controlled independently regardless of other service in MMS. After the data is encoded at the SDDS server, it is transferred to DTV remultiplexer to multiplex with TS stream from the PSI/PSIP generator, A/V encoders and the data broadcasting encoder.

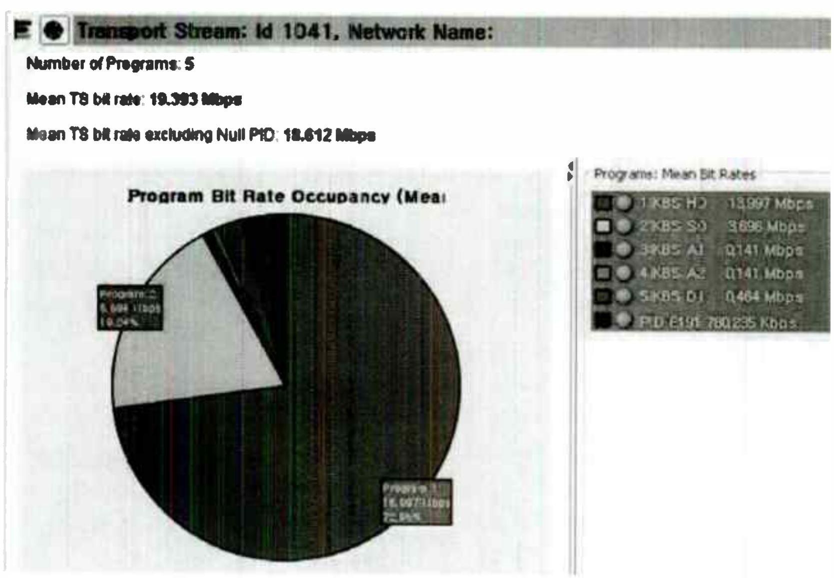


Fig. 3. MMS Bit-rate Information

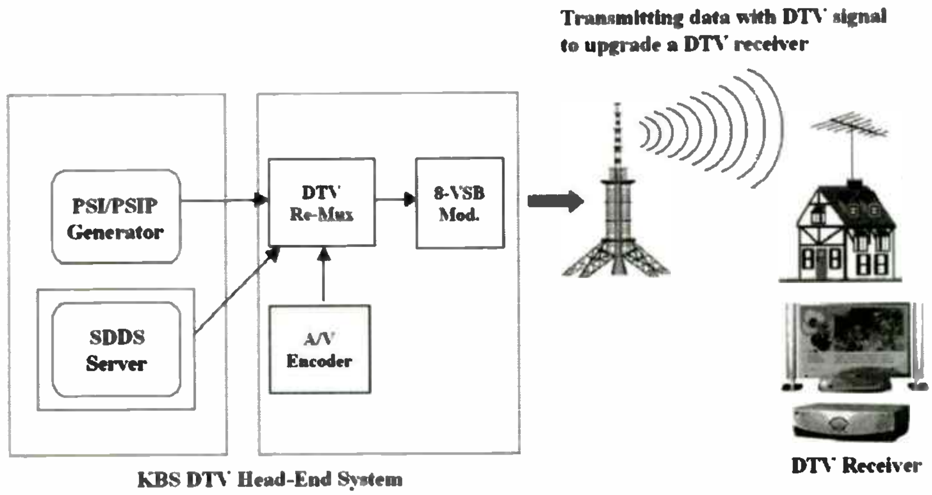


Fig. 4. Concept Diagram for SDDS

MMS TEST STREAM

The MMS implementation guideline and test streams were needed to minimize the field issues after the MMS trial broadcast during the period of 2006 FIFA World Cup[9][11]. For this requirement, Korean Broadcast Engineer Technical Association (KOBETA) led the technical action group that is composed of broadcasters and manufactures. The group has distributed the test streams to DTV manufacturers. MMS test streams consist of 21 MMS modes which are all compliant with ATSC standard. These streams are used to verify the basic interface between a MMS transmitter system and DTV receivers. Table 3 shows the composition of the first 10 streams. The number in the table indicates the number of each channel for HD, SD, data, audio, and hidden channel. The hidden channel is generally used to transmit a SDDS. Each stream has various types of multiplexing and video formats shown in the table.

Table 3. Composition of MMS Test Streams

No	HDTV	SDTV	Data	Audio	Hidden	Video Format
1	1	1	0	0	1	CBR(720P)
2		1	0	0	0	CBR(1080i)
3		2	0	0	0	CBR(720p)
4		2	2	2	0	CBR(720p)
5	0	20	2	2	0	CBR(480i)
6	1	1	0	0	0	VBR(720P)
7		1	0	0	0	VBR(1080i)
8		2	0	0	0	VBR(720P)
9		2	2	2	0	VBR(720P)
10	0	4	2	2	0	VBR(480i)

MMS TRIAL BROADCAST

During the period of 2006 FIFA World Cup, Korean Broadcasting Commission(KBC) allowed terrestrial broadcasters to launch a MMS trial service. Since the launch of the digital terrestrial broadcast in 2001, digital TV penetration rate is very low. The MMS was regarded as a best ideal option to boost the digital penetration rate in Korea. Korean government had clear strategic goals by this service, which is to increase the penetration rate by making DTV services more valuable. Therefore this service was considered as an evaluative tool to implement government policy concerning digital switch-over in future. The service area of this trial broadcast was limited in the metropolitan area, the population of which is about 20 millions. Considering the number of householders with DTV nationwide, the service could reach 2 million households. Four terrestrial TV broadcasters launched the MMS trial broadcast with 5 given digital channels at 06:00, June 5th, 2006. KBS also launched the MMS on two virtual terrestrial DTV channels. It was an historical momentum in Korean broadcasting history.

Service Configuration

At the first phase of the MMS trial broadcast, KBC didn't regulate the service configuration. Therefore, the broadcasters were willing to provide as many services as possible. Before the trial service, KBS had provided only one HD channel through 6MHz. KBS, however, transmitted one HD, one SD, one audio and two data services during the trial broadcast through the same bandwidth. Table 4 shows the MMS types for the trial service. At the middle of the trial broadcast period, KBC restricted the service configuration to one HD and one SD service due to the complaint from the pay

TV service providers such as cable and satellite operating companies. They argued the implementation of the MMS in digital terrestrial TV environments would be a threat to the fledgling pay TV industry and violation of platform neutrality.

Table 4. MMS Types of Broadcasters

Virtual Ch.	Multiplxer	MMS Service Type
6-1, 6-2	SBS	HD + SD
7-1~7-5	KBS-2	HD + SD + Audio + Data + Data
9-1~9-5	KBS-1	HD + SD + Audio + Data + Data
10-1~10-3	EBS	HD + SD + SD
11-1,11-2	MBC	HD + SD

KBS had focused on a SD service channel for actual programming since other channels were either simulcast of an analogue channel or data channels. While a World Cup game was aired on channel 9-1 with HD resolution, channel 9-2 usually showed the picture of a specific star player on the ground in SD picture quality. We call this picture a 'isolated picture' or a 'player feed' which is considered as one of the good examples of maximizing the potential of the MMS. From this service, viewers were able to watch the other big match being held at the same time. During the rest of the time, SD channel 9-2 was fed with special archive programs related to the World Cup, documentary programs, and series of soap opera.

Channel 9-3 was an audio channel. It was basically a radio channel with program-related data such as music information. Fig. 5a and 5b illustrates the examples of audio programs. The audience could listen to the best quality audio with bit rate of 200Kbps.

Channel 9-4 and 9-5 were data channels with non-interactivity. An interactive service was provided in the channel 9-1 with program-related data or additional stand-alone information. The service window of the channel 9-4 and 9-5 in Fig. 5c, 5d shows the examples of the data-only programs such as weather information and finding missing children. The basic data about missing child were continuously displayed on the monitor with police contact point. This exemplifies the extreme case of how the spectrum can be served for public interest. Table 5 shows the summary of the MMS and the contents of the service which were broadcast by KBS.

Table 5. Contents of MMS service (KBS)

Ch.	Service Type	Contents
9-1	HD + interactive data	Simulcast of Analogue Channel Program Related Interactive Data
9-2	SD	World Cup Games & Player Feed, or Archive programs related to soccer, or Reuse of Documentary and Drama
9-3	Audio + Data	Simulcast of KBS 1 Radio, and Basic Graphic Data for Program
9-4	Data only	Weather and Traffic Data
9-5	Data only	Information of Missing Children

Fig. 5a. Data Contents related with an Audio Program

Fig. 5b. Data Contents related with other Audio Programs

Fig. 5c. 9-4 Weather and traffic

Fig. 5d. 9-5 Finding missing children

Fig. 5. The examples of MMS Service (for audio and data program channel)

SURVEY OF MMS ON VIEWERS

KBS provided the MMS trial broadcast during the 2006 FIFA World Cup. After that, KBS commissioned Net Intelligence Korea to carry out a survey of MMS on viewers[6][7]. Men and women over 20 years old were surveyed nationwide from June 26th to June 30th, 2006. Valid sample numbers were 1500.

Survey results

Nearly 90% of respondents have two sets of TV or less. Viewers who have one set are at 41.6% and people who have two sets are at 47%. But people who have three sets or more are only at 11.4%. 34.9% of respondents have digital TV while the others don't. The number of people who have digital TV increased rapidly. The take-up rate for digital TV in the late 2005 was only 19.4% according to industry experts. This mainly resulted from special effect from the World Cup.

Concerning TV size, most viewers have 29-inch TV sets or the smaller for analogue TV while TV sets from 30 inch to 39 inch is the most common for digital TV. Fig. 6 depicts the percentage of TV sets by size.

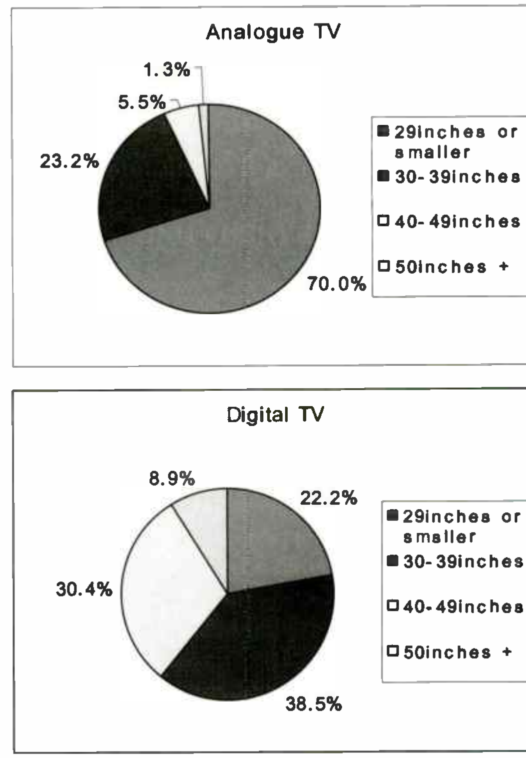


Fig. 6. Analogue TV and Digital TV by Size

Over 78% of respondents answered MMS is useful to the viewers; 21.3%, "very useful"; 56.9%, "quite useful". Over 90% of respondents answered that the MMS is desirable for the viewers; 49.1%, "very desirable"; 41.1%, "quite desirable". Over 85% of respondents answered they are likely to use MMS; 28.8%, "very likely to use"; 56.9%, "quite likely to use". Fig. 7 illustrates the percentage of evaluation of desirability of MMS.

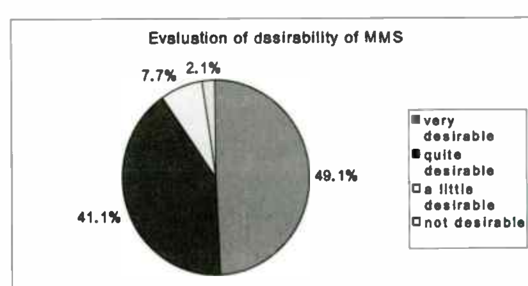


Fig. 7. Evaluation of Desirability of MMS

If MMS is to be fully adopted, preferred genres of additional channels(SD) are in order : movies, news, culture/documentary and sports. Fig. 8 shows the percentage of preferred genres of MMS additional channels.

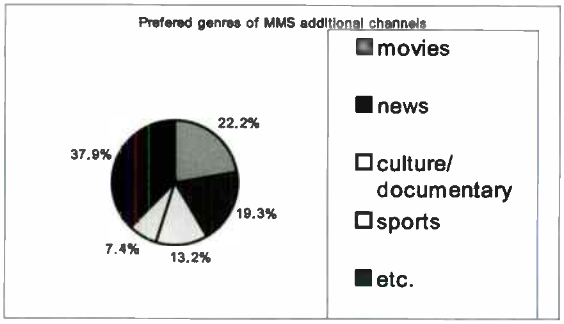


Fig. 8. Preferred Genres of MMS Additional Channels

The most common method of watching TV is via cable at 70.8%. The direct reception of terrestrial signal is at 16.3% and that of satellite broadcast is at 11.5%. If terrestrial TV provides about 10 channels free of charge, 69.1% of the respondents have intention of switching to direct reception. Also, if low-priced cable TV or satellite broadcast services are stopped, 82.5% of the respondents have intention of switching to direct reception.

The survey shows that most respondents evaluated the MMS very positively. Therefore, it can be inferred that the level of reception will be very high if the MMS is provided regularly.

CONCLUSION

KBS has developed and implemented the MMS platform. In a bandwidth of 6MHz, one HD program, one SD program, one audio program, and two data services could be provided including an optional software download service. KBS successfully verified the MMS through end-to-end experiments and its trial broadcast during the 2006 FIFA World Cup Germany in July 2006. The MMS implementation guideline and test streams were also created to minimize the field issues. 90 percent of viewers surveyed at the end of the MMS trial broadcast responded positively to the MMS. Therefore, KBS estimates that the MMS will increase the digital TV penetration rates rapidly if the MMS is fully provided.

REFERENCES

[1] A/65C; Program and System Information Protocol for Terrestrial Broadcast and Cable, Revision C, with Amendment No. 1

[2] A/97; ATSC Standard: Software Download Data Service(SDDS)

[3] A/101; ATSC Standard : Advanced Common Application Platform(ACAP)

[4] Recommendation ITU-R BT.500-11 ; Methodology for the Subjective Assessment of the Quality of Television Pictures

[5] Jeong-Deok Kim, "Developing a fully interactive TV system and application in compliance with ATSC ACAP", NAB Broadcast Engineering Conference Proceedings 2005, April, 2005

[6] Net Intelligence Korea, "The survey of the picture quality test on the MMS", June, 2006

[7] Net Intelligence Korea, "The survey of the level of reception on the viewers in the case of the MMS trial broadcast", July, 2006

[8] Jeong-Deok Kim, "The development of an interactive TV system and application in compliance with ACAP", ABU Technical Review, July-August, 2006.

[9] Korean Broadcasters Association, "MMS test streams and guideline version 1.02", November, 2006

[10] Jeong-Deok Kim, "The Development of the Multi-Mode Service on terrestrial DTV", 43th ABU(Asia-pacific Broadcasting Union) General Assembly in Beijing, November, 2006

[11] MMS technical verifying committee, "The resulting report of basic interfaces between a MMS transmission system and DTV receivers", January, 2007

Enabling New Applications
Thursday, April 19, 2007
10:00 AM – 11:30 AM

Chairperson: Joe Snelson, CPBE
VP & Director of Engineering
Meredith Broadcast Group

Hybrid TV - the Best of Both Worlds
Chris Howe, Red Bee Media, London, UK

***Cancelling the Postage Stamp, Concluded**
J. Patrick Waddell, HARMONIC INC, Sunnyvale, CA

The Data Server Based on ATSC A/76A Standard
Minsik Park, Electronics and Telecommunication Research Institute (ETRI), Daejeon, Korea

*Paper not available at the time of publication

Hybrid TV—The best of both worlds

CHRIS J. HOWE & SUE M. FARRELL
Red Bee Media
London, UK

ABSTRACT

A growing range of distribution channels for television programming is creating a new battleground between digital broadcasting and broadband television and video services. Broadcasters have to continue to support their existing viewers and want to maintain the fixed economics of broadcasting. Meanwhile, new entrants to the television market are offering sophisticated services that exploit the wider capabilities of data networks to provide television and video.

A hybrid approach, binding broadcast and broadband delivery in the set-top box, offers broadcasters the potential to preserve their position. This combination uses broadcast channels to deliver live and scheduled services together with an enhanced offering of niche or on-demand programming provided over a broadband network. Hybrid services exploit existing investments in broadcast operations and infrastructure and provide an efficient way to deliver popular, premium, special interest and enhanced programming to the home.

This paper explores the challenges and opportunities presented by combining the strengths of broadcast and broadband networks, integrating rich media services in a single device with a seamless user proposition.

INTRODUCTION

The model of free-to-air advertiser-funded broadcast television which has been eroded by multichannel, subscription and pay-per-view services is being further undermined by digital video recorders, video-on-demand services and the disruptive network effects enabled by broadband high-speed internet access.

These technologies can exploit the power of broadcast programming to provide new modes of consumption. However, without addressing the fundamental requirements of the traditional broadcast value chain, this parasitic approach risks destabilizing the historic business model that feeds them.

There is still value in broadcast distribution and it remains the most efficient means of reaching a mass audience. Nevertheless, there is considerable consumer appetite for the greater choice, convenience and control that can be offered through approaches that can address the personal preferences of the individual viewer.

HYBRID TV

Hybrid TV employs a combination of broadcast and broadband networks to deliver television and video services seamlessly to a single device.

By combining the unique capabilities of both broadcast and broadband networks, it is possible to produce a blend of services that brings the best of both worlds in a symbiotic and mutually supportive relationship.

For the user, a unified hybrid proposition offers easy access to their familiar scheduled channels with the added benefit of a far wider selection of programs available over broadband as streams or downloads. This is not in itself a benefit to the broadcaster, as it diverts the audience from scheduled programming and its supporting advertising. However, there is an obvious direct revenue opportunity in providing programming on a pay-per-view, subscription, rental or retail basis. There is also the potential to build a direct relationship with the individual viewer and household that has been unavailable to the free-to-air broadcaster in the past.

This relationship can enable targeted advertising and provide accurate audience measurement to reinforce the traditional advertising model. It can also be used to drive customer relationship management, personalized promotions and recommendations.

Individual networks, broadcasters or channels are unlikely to be able to provide such a platform in isolation. They generally rely on the presence of competing channels to deliver sufficient choice to the consumer. Equally, existing pay-television platforms and new entrants will find it difficult to entirely displace the major networks without their co-operation and support.

Some of the new entrants to the television and home entertainment business are being driven by their own commercial imperatives and are eager to own the customer relationship and dominate the value chain based on the perceived value of their distribution networks. However, those prospective service providers that are able to work with the program production and broadcast community, support their business requirements and preserve their role in the value chain are ultimately more likely to themselves be commercially successful.

PROPOSITION

The consumer proposition for hybrid TV is the presentation of services through a unified interface in an integrated set-top box, media gateway or other consumer electronics device.

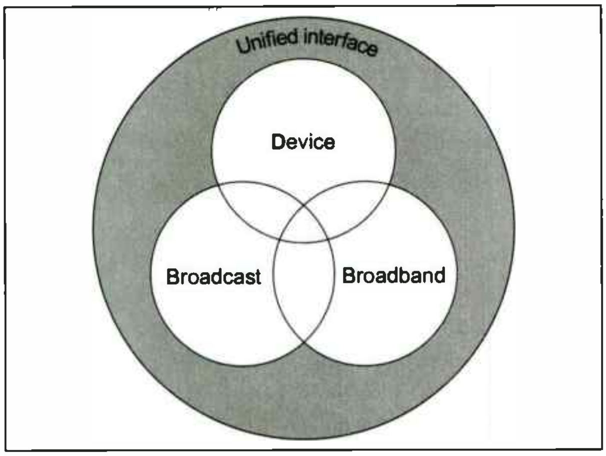


Figure 1: Key components of Hybrid TV

Unified interface

Unifying these services requires an intuitive interface allowing the range of broadcast and broadband delivered programming to be navigated seamlessly, incorporating relevant program information.

Search, collaborative filtering, promotions and editorial recommendations form a trusted guide that can be used to direct the user to relevant programs based on their stated or implicit personal preferences and previous viewing behavior.

Single device

A single device capable of receiving, displaying and optionally storing material in a variety of formats enables multiple services to be accessed through a simple remote control.

Hybrid services may emerge out of any of the existing consumer devices connected to the television set.

Set-top boxes may be provided by cable, satellite or terrestrial platform operators, or sold through retail outlets as an enhancement to free-to-air broadcast television.

Media Centers and home theatre systems combine the features of a personal computer with one or more television tuners to provide a 10ft interface with a remote control offering a more televisual experience.

These devices may contain a hard disk or use network-based media storage.

Broadcast

Cable, satellite and terrestrial broadcast channels are carried over the conventional transmission networks. The signal is delivered to a cable connection, satellite dish or conventional aerial and is presented to one or more radio frequency tuners in the device.

Broadband

On-demand and bi-directional interactive services are carried over data networks using a broadband connection to the set-top box.

INTEGRATION

The concept of an integrated television device is nothing new. The first of these was perhaps the videocassette recorder, combining access to pre-recorded and broadcast programming. Its successors were the DVD and the digital video recorder. The combined computer-based media center attempts to concentrate these capabilities in a single box. Without the need for physical media, it is not necessary or even desirable to play material from different devices if it is all available electronically.

For the consumer there is a definite attraction in reducing the number of devices connected to the television and providing seamless switching between sources destined for the same screen.

It offers viewers an incrementally improved television experience. They can continue to receive broadcast channels from known and trusted brands. In addition it provides a simple way to receive on-demand or downloaded programs on their television.

Local or network storage can selectively record broadcast programs to offer digital video recording. 'Catch-up' services, where programs remain available for some extended window after transmission, may be a key feature of the hybrid TV service to support the broadcast channels.

'Top-up', pay-per-view or pre-pay à la carte services may supplement the free-to-air offering without requiring the consumer to commit to a long-term contract or a prescribed bundle of channels.

The addition of data connectivity offers the further opportunity for integration with other communications services such as unified messaging, including caller id, voicemail, or even video calling.

Transactional services can include shopping, self-service account management, billing and payment information.

OPPORTUNITIES

These objectives are already being pursued by cable, satellite and telecommunications service providers around the world. IPTV or internet protocol television and broadband video services promise to provide the next generation of viewing experience [1].

The consumer electronics industry, from major multinationals to start-up ventures, is enthusiastically developing devices to address this latent demand and connect the television to the internet.

So far a clear winner is yet to emerge. Without a coherent service model they will be unable to maximize the opportunity and broadcasters will not benefit as a result.

In those markets with a strong tradition of free-to-air terrestrial television and a growing broadband base, new platforms combining broadcast and data networks can provide a compelling proposition. Various business models are possible, from hardware subsidized by subscription contracts, to consumer electronics products sold at retail.

Service providers

Hybrid TV services may emerge out of a number of existing platforms. Business models will vary by feature set, territory, the source of broadcast programming, and the service provider.

Pay-TV platform operators gain the opportunity to compete with the more sophisticated services emerging from telcos and broadband service providers, presenting a defensive strategy to protect their existing offering, maintain share and reduce churn.

Telcos and broadband service providers can exploit existing broadcast channels to deliver television services to consumers. This negates the need for a multicast-enabled network, maintains the broadcast economics for delivering live and popular programs and avoids the need to negotiate carriage agreements with broadcasters.

Consumer electronics companies are now launching products sold through retail outlets providing hybrid services which combine broadcast television with on-demand programming delivered via the internet, over the top of existing broadband networks.

Service management providers may partner with any of these parties to manage the service platform, provide the interactive program guide, operational media, metadata and customer relationship management as an outsourced service.

CHALLENGES

There are many challenges in delivering a hybrid television service.

Coverage

There is clearly a requirement for both broadcast and broadband networks to be available at the customer premises.

The availability of broadcast services can be limited by a range of factors. Local geography may interfere with reception and aerials may need upgrading. There may be distribution issues in multi-occupancy buildings.

While broadband availability and penetration are growing globally, not all households are necessarily able to receive broadband services and access speeds may be limited by the length of the copper cable to the home.

Coverage for hybrid TV is therefore defined by the availability of both the broadcast and broadband services at the premises, which when combined will generally be less than the availability of either service independently, as illustrated in Figure 2.

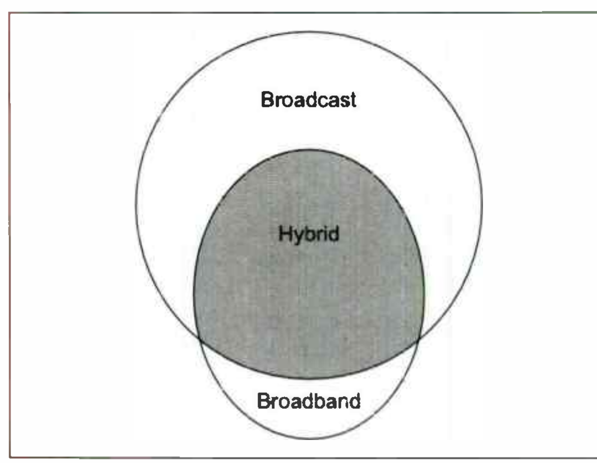


Figure 2: Hybrid TV coverage limitations

This presents considerable marketing issues. It becomes more difficult to offer services on a national basis and a geo-coded database may be needed to determine the relevant overlap of coverage for a particular consumer address. Even this may prove inadequate, because of local reception and line conditions, inevitably leading to returns by disappointed customers.

Installation

Within the customer premises, there is a requirement for both the broadcast and broadband connections to be conveniently sited near the television. This may involve the installation of additional cabling. Wireless or power line networking may be used to bring broadband to the television, but the bandwidth needs to be sufficient to maintain a consistent and acceptable user experience.

Integrating this installation with other home entertainment systems may still be a considerable challenge for the average consumer. For subscription services these issues may be addressed by an onsite installation engineer but this may not be feasible for products sold at retail. Otherwise, foolproof instructions and out-of-the-box ease of installation and configuration will be essential.

In either case, specialist telephone support will be required to troubleshoot any installation issues. These will be exacerbated by the involvement of several parties in providing different components of the end-to-end system.

Networks

While broadband may currently be adequate for web browsing and email, it is not necessarily designed to deliver home entertainment services. The network bandwidth, contention ratio and reliability need to be considered, scaled and strengthened to deliver on-demand video services.

Broadband providers may suggest that implementing quality of service is essential for reliable video streaming. However, adopting a download or peer-to-peer distribution model can satisfactorily deliver video over best efforts networks. This allows additional programming to be provided over existing broadband internet connections.

Security

The security considerations involved include ensuring appropriate and adequate content rights management, service integrity, device defense and data protection.

Conditional access, content and copy protection may control the availability and onward distribution of material. Watermarking can also be used to provide forensic identification of inappropriately distributed programming.

Normal network concerns about denial of service attacks, hacking, cracking, spoofing, spamming, sniffing or phishing—any of which could result in loss of service or abuse of personal information—need to be addressed. Existing firewall configurations may also need to be adjusted.

Standards

Convergence of functionality in a single device requires integration of technical standards across broadcast, telecommunications and information technology.

Multiple standard and high-definition video formats may need to be supported, possibly in more than one compression scheme.

Various browsers, media players and application programming interfaces will need to be implemented, all of which will inevitably be repeatedly upgraded.

The service platform will require the integration of customer relationship management, operational and billing support systems. Data may need to be translated or transformed as part of the service integration. Media may require transcoding and encryption for delivery.

Metadata

Driving the management, delivery and presentation of material, metadata—or descriptive data about media assets—is the key to efficient operation of the service platform and a good user experience. Some metadata controls the presentation of the service but is not visible to the viewer, such as aspect ratio switching and service linking data. Other metadata is exposed as part of the service, such as listings information to describe programs in the electronic program guide.

Management

There are still considerable challenges involved in providing and operating the platform, aggregating the programming proposition, running the back office systems and supporting the customer [2].

This requires core competencies in editorial, design, promotion, metadata, media logistics and service management.

The introduction of broadband connectivity and more complex software and navigation systems within the hybrid device reinforces a requirement for service management of a style more familiar to the information technology industry [3].

The increased reliance on software for differentiation will result in upgrades for functionality and security patches over the lifetime of the device. Careful management of these upgrades to middleware or user interface software is required to maintain security and supportability while avoiding interrupting the viewer in the middle of their favorite program.

Furthermore, relatively small changes can have drastic consequences if they are not carefully considered and deployed. Effective change and configuration management and quality assurance is essential for maintaining service to the consumer.

The greatest service management requirement will be providing assistance to consumers unaccustomed to reporting service defects or transactional problems with their television.

The participation of multiple parties in the delivery of service can lead to confusion, delays or blame games when there are support issues, resulting in a poor customer experience or unacceptable extended service outages.

The cost of providing this customer support, beyond the sales and billing relationship, over the lifetime of the product or service, therefore needs to be factored into the commercial model.

Programming

Needless to say, without providing compelling programming and offering sufficient choice and variety, even the greatest technology solution will not be attractive to the consumer.

The expertise in commissioning, producing, packaging, presenting and promoting programming will continue to come from the creative community that currently sustains the television industry.

Advertising

While there may be much discussion about the death of television advertising, it will inevitably evolve to exploit the capabilities of the new media environment.

The hybrid approach can offer a range of solutions to sustain advertiser-funded television. In addition to the promise of targeted advertising, there are more immediate opportunities to improve the accountability of traditional spot advertising through accurate measurement and reporting.

Commercial

The business models for these services are not well established or understood and creating the right partnerships, service offering and marketing campaign will be essential.

This may bring together companies that were formerly direct competitors and industries that were previously entirely distinct. Each participant will inevitably consider that their contribution is the most valuable. Some may believe that there is an enormous commercial opportunity to exploit through on-demand delivery direct to the consumer. The reality may be that revenues will still be required to flow back to the production industry to enable the continued creation of high-quality, popular and challenging programming. Once the costs of production and packaging are taken into account, the remaining revenue share may be marginal.

CONCLUSION

This hybrid approach offers a means of supporting the conventional broadcast television model and bringing many of the benefits of the vertically-integrated platforms being promoted by cable, satellite and telecommunications providers.

It is not without its challenges, but as countries around the world complete digital conversion it offers a pragmatic solution to many of the problems faced by traditional television networks. While much of the current activity is being led by a combination of the consumer electronics and computing industries, the active engagement of broadcasters will ultimately be essential to the long-term viability of such devices.

The day-to-day service management involved in providing an integrated platform should not be underestimated. Many of the skills involved are not the core competencies of technology and communications infrastructure providers. They are more familiar to existing pay-TV platform operators, but are largely novel to terrestrial broadcasters who may have little experience in having a direct relationship with individual viewers.

Strong partnerships, a clear consumer proposition, compelling programming, attractive pricing, unified interface, intuitive navigation and excellent customer support are all essential elements to a successful service.

Hybrid TV can exploit the strengths of both broadcast and broadband networks, offering consumers the best of both worlds.

REFERENCES

[1] Cooper, W and Lovelace, G: IPTV: Broadband Meets Broadcast—The network television revolution, informitv, 2005

[2] Howe, C and Farrell, S: Taking TV Operations into the Future, NAB Conference Proceedings, 2006

[3] Cooper, W and Farrell, S: Service Management for Digital Broadcasting, IBC Conference Proceedings, 2003

ACKNOWLEDGEMENTS

The authors would like to thank their colleagues, clients and partners at Red Bee Media.

Published by permission of Red Bee Media Limited.

The Data Server based on ATSC A/76A Standard

Minsik Park, Han-kyu Lee and Jin Woo Hong

Electronics and Telecommunications Research Institute, 161 Gajeong-dong, Yuseong-gu, Daejeon, 305-700, KOREA
{pms, hkl, jwhong}@etri.re.kr

ABSTRACT

DTV data broadcasting service offers broadcasters the new business model to acquire considerable revenue from digital broadcast. When the broadcasters integrate and maintain the station system for the data service, they encounter the unnecessary cost to be paid because there is no common interface specification among the equipments of DTV station system. This paper describes the flexible and open interface that will decrease the wasteful expenses in moving the transition to data broadcasting system.

1 INTRODUCTION

The data broadcasting service has became to be an important role in providing users with multimedia data including rich information as the broadcasting environment is evolving from analog to digital. Data service provides users with the interactive and enhanced services by sending data elements in broadcast television transport. Data applications deliver the additional and useful information to viewers with a various type of media through a broadcast channel.

To provide the data service, DTV broadcasters need several functional components to deliver data broadcast through DTV broadcasting system. The most essential component is a so-called data server to encapsulate data elements into MPEG-2 transport stream. DTV broadcasters insert the data server into their DTV broadcast systems for supporting data service.

Although all the DTV broadcast systems utilize data server with same or similar function, they have each different interface between data server and other components because there is no common interface specification for data broadcasting system. Therefore, equipment manufactures and system integrator of data broadcasting should implement the interface specific to DTV broadcasters whenever they set up the data server to data broadcasting system. This is why that the integration work of data broadcast system not only decreases the flexibility of system extension but also increases the cost of system maintenance. The inoperability interface of DTV broadcast system may be a considerable obstacle when DTV broadcasters intend to change their DTV broadcast system to data broadcast system.

ATSC(Advanced Television Systems Committee) has published A/76A standard (PMCP, Programming Metadata Communication Protocol)[1], the interface standard for exchanging PSIP(Program and System Information Protocol)[2] information among components of DTV broadcasts system. The ATSC PSIP standard, an essential element of the ATSC DTV system, provides a methodology for transporting digital television system information and electronic program guide data.

The PMCP specification will enable broadcasters and manufacturers to interconnect systems that process PSIP and other DTV metadata such as traffic, program management, listing service, automation, MPEG encoder, and PSIP generator. PMCP guarantees the interoperability among components that are required to exchange the PSIP information. PMCP is based on XML (extensible markup language) message that is flexible and usable for various system architectures.

ATSC has also released Annex D in A/76A, the extension of the existing PMCP 2.0 schema for ACAP[3] data service, which is expected to play an important role in open interface standard for data broadcasting emission system as DTV data service become more prevalent.

The PMCP 3.0 schema for data service can be developed by not only inserting data server into the PMCP reference system but also defining the new element of encapsulation, signaling and announcement for data broadcasting.

The paper describes the data broadcasting metadata to extend the existing PMCP and ACAP data server that ETRI has designed to support data service. The extension of PMCP for data service defined in A/76A can be easily applied into DTV emission system by implementing data server to support the PMCP schema elements to define encapsulation, signaling and announcement for data broadcasting. A/76A increases the flexibility and usability of DTV data broadcasting system in that it guarantees the interoperable interface among data broadcast components such as data server, PSIP server and multiplex manager.

2 DATA BROADCAST SYSTEM

In order to provide content creators with the specification necessary to ensure their data to run on all kinds of receivers in equal performance, ATSC published DASE-1(DTV Application Software Environment-Level 1)[4] and ACAP[3] which define a software layer (middleware) that allows interactive and enhanced contents to run on receiver in a platform-independent manner. In particular, ACAP has been developed in need of the harmonization between DASE and OCAP (OpenCable Application Platform)[5] published by CableLabs.
In order to provide users with data broadcast through a broadcast channel, broadcasters need the emission station system to encode the data created according to the data broadcast standards [3,4]. ATSC DIWG(Data Implementation Working Group) recommends not only functional components but also physical interfaces in the emission station for data broadcasting emission. However, ATSC DIWG doesn't define the specific interface message so as to operate each functional component.
ATSC DIWG proposed the emission station environment for data broadcast in IS/151[6]. Fig. 1 shows the essential components of the emission station environments. The emission station consists of audio/video encoder, multiplexer manager, CA(Conditional Access) generator, PSIP server and data server. The emission station generates MPEG-2 transport streams containing compressed video and audio with encapsulated data and system information tables. Data server plays an important role in encapsulating data elements into MPEG-2 transport stream. Multiplexer manager sends PSIP information into PSIP server and data information into data server for operation. Data server needs to be inputted with encoding information so as to generate data stream according to the transport protocol defined in the data broadcasting standards[3, 7-11].

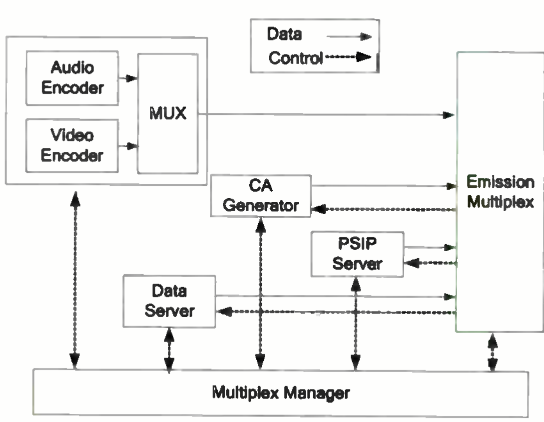


Figure -1. Data broadcast station system

The interchanging information regarding PSIP and data service among other components has been defined with various formats per broadcasters. A/76A[1] provides not only the common interface in terms of PSIP but also that with regards to data service.
For the interoperable interface between data server and other components, the existing PMCP 2.0 schema based on PSIP had been extended for data service and data server had been added into PMCP reference system.
Data server is connected into PMCP interface not only because data server needs to exchange the data information with PSIP server, but also because multiplex manager such as program management system, traffic system and automation system should control the operation of data server according to the broadcasting schedule. The following section describes A/76A Annex D, the data broadcast metadata for ACAP[3] data service.

3 DATA BROADCAST METADATA

The existing PMCP 2.0 schema had been extended to describe the information to enable data server to encode data application into transport stream. The following requirements therefore had been considered in designing the new PMCP schema for ACAP data service:

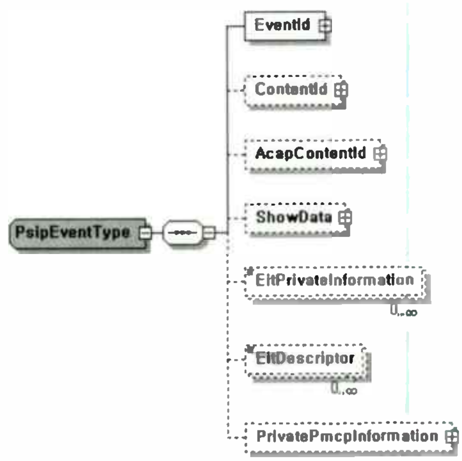
- The PMCP for data service should be backward compatible with the existing PMCP 2.0 schema so that conventional DTV station system based on PMCP 2.0 schema can easily be implemented to data broadcast station system.
- The PMCP for data service should include the announcement information to create program and system table for data broadcasting.
- The PMCP for data service should describe the encapsulation information to encapsulate data into MPEG-2 transport stream.
- The PMCP for data service should contain the signaling information to generate system table into MPEG-2 transport stream.

Data information consists of announcement, encapsulation and signaling. ATSC has defined A/101[3] for signaling and encapsulation, and A/102[12] for announcement with regard to ACAP data service.

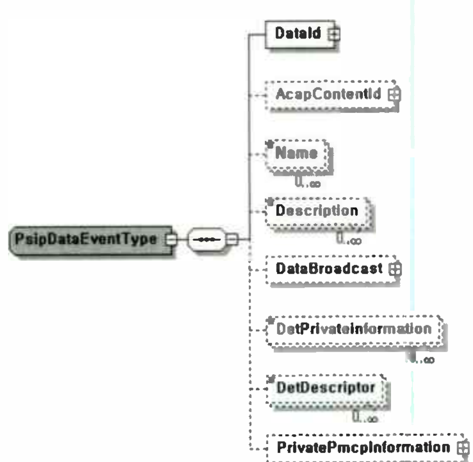
3.1 PMCP Schema for ACAP Announcement

ATSC had defined A/102[12] that describes the method to announce ACAP data service. The data broadcast descriptor, new announcement descriptor is defined in A/102[12]. The data broadcast descriptor is used to announce data service by being included in either A/65 EIT[2] or A/90 DET[7] according to data service type defined in A/102[2].

Fig. 2 shows "PsipEvent" element to describe EIT and "PsipDataEvent" to depict DET. There are three child elements required to announce data service. "AcapContentId" element included in both PsipEventType and PsipDateEventType is defined to manage data contents separately from A/V contents. The "DataId" element in PsipDataEventType is used to label or reference the event related to data. The "DataBroadcast" element in both "ShowData" element of PsipEventType and PsipDataEventType represents data broadcast descriptor used to recognize the data service type



(a) PsipEventType Diagram



(b) PsipDataEventType Diagram

Figure - 2. Announcement Element

3.2 PMCP Schema for ACAP Encapsulation

The data and attributes of one U-U object in an object carousel are transmitted in one message. The message format is specified by the BIOP(Broadcast Inter ORB Protocol) and is referred to as the BIOP Generic Object Message format. BIOP Messages are carried in Modules of data carousel[12].

The AcapObjectCarouselType therefore is divided into two children element : "DataCarousel" element and "ObjectCarousel" element as shown in Fig.3. The "DataCarousel" element represents the information of data carousel that delivers the BIOP messages in modules, and the "ObectCarousel" element contains the information of BIOP messages of the object carousel.

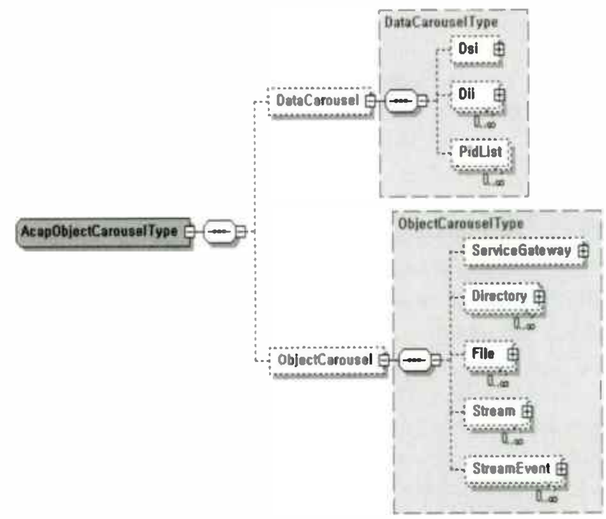


Figure - 3. AcapObjectCarouselType diagram

The DataCarouselType is composed of three elements: the "Dsi" element (DSI), the "Dii" element (DII), and the "PidList" element to describe the DSI, DII, and PIDs defined in data carousel[13].

The ObjectCarouselType represents the BIOP messages of objects such as service gateway, directory, file, stream, stream event in object carousel[3].

3.3 PMCP Schema for ACAP Signaling

The signaling information enables receivers not only to identify applications associated with a service and their location from which to recover them. ACAP standard[3] specifies PMT(Program Map Table) and AIT(Application Information Table) to signal data application into receiver. The PMCP schema for data service, therefore, includes the elements to describe the information relating to both AIT and PMT.

3.3.1 AcapApplicationType for ACAP Signaling

The AIT describes applications and their associated information. Each AIT includes one "common" descriptor loop at the top level for descriptors that are shared between application of that sub table and a loop of application. Each application in the application loop has an "application" descriptor loop containing the descriptors associated with that application.

The descriptors of AIT are categorized into three parts: generic application descriptor, application specific descriptor and application representation specific descriptor. The generic application descriptors are included in common descriptor loop. The application specific descriptors also are specific to the application instance. The application representation specific descriptors are specific not only to application instance but also to application representation.
The AIT and descriptors mentioned above can be represented with the schema as shown in Fig. 4.
The "TransportProtocol" and the "DiiLocation" describes the information of transport protocol descriptor and download info indication descriptor respectively. The "Application" element not only defines the organization_id, application_id and application_control_code with its attribute but also include the elements to describe the information of application descriptors of AIT.

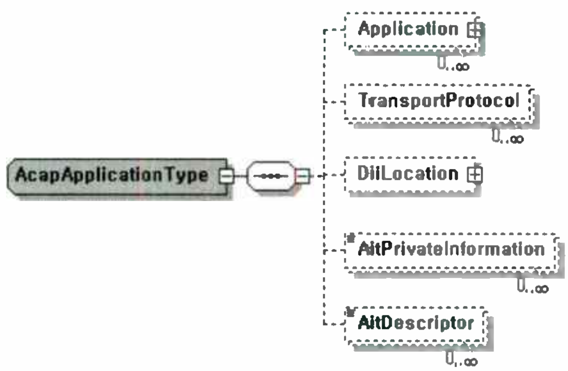


Figure – 4. AcapApplicationType diagram

3.3.2 ChannelType for PMT Signaling

In order to describe the elementary streams of the object carousel, The PMCP schema for ACAP data service includes the additional several elements to describe PMT descriptors. The first loop for PMT descriptor delivers deferred association tags descriptor; the second loop for PMT descriptor consists of carousel identifier descriptor, application signaling descriptor, data broadcast id descriptor, stream identifier descriptor and association tag descriptor.
Fig. 5 illustrates the schema of PMT descriptors for object carousel. New elements, which describes PMT descriptors for data broadcasting, can be added into "ElementaryStream" element of the ChannelType defined in PMCP[1] because "ElementaryStream" element plays a roll in describing the information of second loop descriptors for PMT.

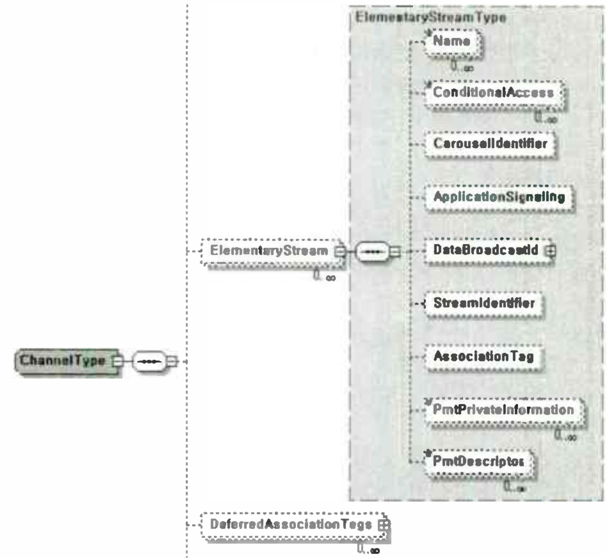


Figure – 5. ChannelType diagram for ACAP data service

3.4 PMCP Schema for ACAP Data Service

Fig. 6 illustrates the AcapDataServiceType to define the information of encapsulation and signaling for data broadcasting. The AcapDataServiceType can be composed of the "AcapContentId" element, "AcapApplication" element and "AcapObjectCarousel" element.

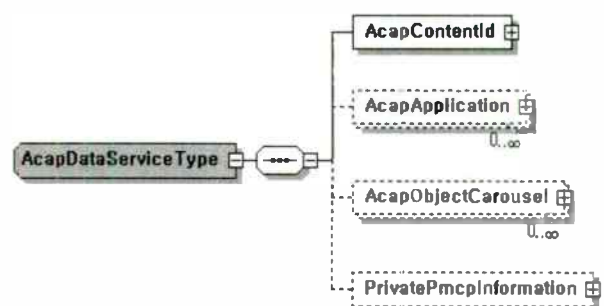


Figure - 6. AcapDataServiceType diagram for ACAP encapsulation and AIT signaling

The "AcapContentId" element provides the linkage between AcapDataServiceType and PsipEventType, or AcapDataService and PsipDataEventType. The linkage enables ACAP transport stream to obtain the schedule information from PsipEventType or PsipDataEventType that describes start time and duration with its attributes. Both the "AcapObjectCarousel" element in Section 3.2 and the "AcapApplication" element in Section 3.3.1 describe the encapsulation and the signalling information defined in ACAP[3].
As shown in Fig. 7, the PMCP schema for data service has defined "PsipDataEvent" element for ACAP announcement in Section 3.1 and "AcapDataService" element for ACAP encapsulation and AIT signaling in

Section 3.2 and 3.3.1, In addition, it has described the "Channel" element to be modified for PMT signaling specific to ACAP in Section 3.3.2. The "AcapDataService" element consists of three elements such as "AcapContentId" , "AcapObjectCarousel" and "AcapApplication". Both the "AcapObjectCarousel" and the "AcapApplication" should be combined with an "AcapDataService" element in that ACAP data service is provided with the AIT and the object carousel that delivers applications and their signaling information respectively. The "AcapDataService" element contains "AcapContentId" element defined in the "PsipEvent" or the "PsipDataEvent" in the PMCP schema to associate an ACAP transport stream with a show event. The "AcapContentId" element enables "AcapDataService" element to identify the schedule information defined in "PsipEvent" element or "PsipDataEvent" element. The schedule information is used in controlling the delivery of the transport stream of ACAP data service.

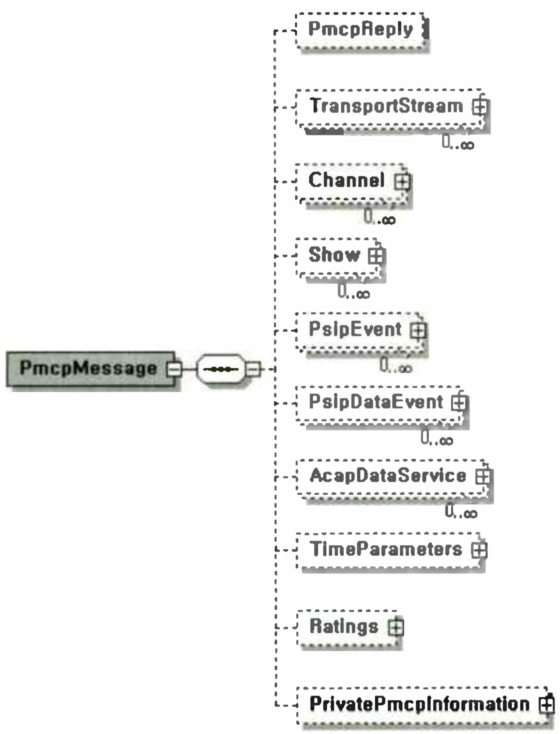


Figure – 7. PMCP Schema for ACAP Data Service

4 THE DESIGN OF DATA SERVER BASED ON A/76A

ETRI has implemented ACAP data server to be operated through PMCP schema for data service as shown in Fig. 8. PMCP schema defines all the kinds of information needed for delivering data broadcast program. Data server first receives the data broadcast metadata from multiplexer manager. The data broadcast metadata is represented with XML message according to PMCP schema defined in A/76A.

Fig. 8 illustrates the structure of data server that has been implemented according to PMCP. Data server consists of parsing block, scheduling block, storing block, encapsulation block.

- **Parsing block** receives PMCP metadata for data service and data broadcasting contents from multiplexer manager. It registers data broadcasting contents and system information, schedule information, encapsulation information extracted from PMCP metadata into storing block. It obtains the schedule information from the attribute of PSIP data event metadata and system (channel) information by using "AcapContentId". It sends schedule researching information, "AcapContentId" into scheduling block.
- **Scheduling block** obtains the schedule information stored in the storing block by "AcapContentId" delivered from the parsing block. It checks out when data contents is delivered. It sends "AcapContentId" corresponding with data contents that should be delivered at a specific time into the encapsulation block.
- **Storing block** has the database for storing the scheduling, encapsulation information and data applications. Database is designed to be associated scheduling, encapsulation, data broadcast contents through the unique identifier, "AcapContentId". Therefore, the scheduling block and encapsulation block could search data applications, scheduling, encapsulation by "AcapContentId".
- **Encapsulation block** searches and obtains data applications, encapsulation and signaling information corresponding with the "AcapContentId" to be delivered from scheduling block. It encapsulates data applications into MPEG-2 transport stream according to the system and encapsulation information.

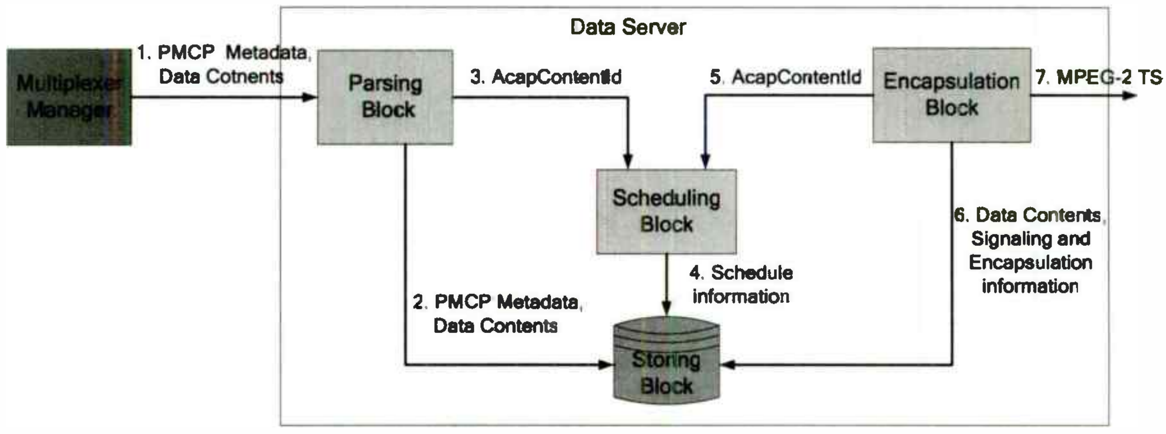


Figure – 8. Data server block structure based on PMCP schema for data service

Fig. 9 illustrates the operation flow of data server as followings;

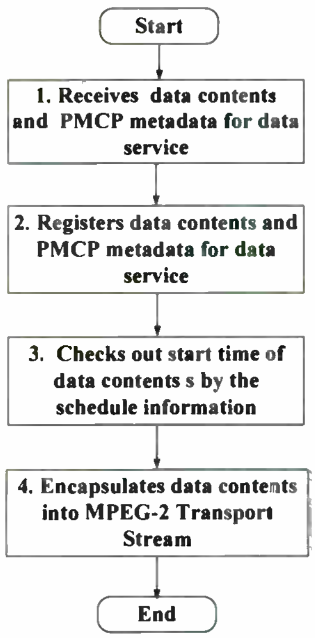


Figure – 9. The operation flow of data server

[step 1] Data server receives data contents and PMCP metadata including system and encoding information such as data type, data location, encapsulation method and signaling information from Multiplexer Manager
[Step 2] The parsing block obtains resisters data contents and system and encoding information into storing block.
[Step 3] The parsing block sends "AcapContentId" into scheduling block.
[Step 4] The scheduling block searches the schedule information stored in storing block by "AcapContentId".
[Step 5] The scheduling block checks out start time of data contents by scheduling information to be acquired from storing block.
[Step 6] The scheduling block provides the encapsulation block with "AcapContentId" corresponding with data contents to be transported.
[Step 7] The scheduling block provides the encapsulation block with "AcapContentId" corresponding with data contents to be transported.
[Step 8] The encapsulation block acquires data contents and encoding information from the storing block.
[Step 9] The encapsulation block encodes data contents into MPEG-2 transport stream according to encoding information

Data server encapsulates data contents into DSM-CC object carousel and generates AIT table according to data broadcast metadata. And then the encapsulated data application and AIT is packetized with MPEG-2 transport stream. Data broadcast system combines both the multiplexed data and AIT transport stream delivered from data server and audio/video transport stream generated from A/V encoder into a MPEG-2 transport stream, which is transferred into receiver.

5 CONCLUSIONS

ETRI has designed the data server to support data broadcasting according to PMCP for data service. The PMCP for data service can be easily developed by adding data server into DTV station system model. It increases the flexibility and usability of data broadcast system in that it guarantees the interoperable interface among data broadcast components such as data server, PSIP server and multiplex manager.
The PMCP for data service defined in ATSC A/76A[1] will be expected to be an interoperable interface of DTV station system as data service become more prevalent.

Company Index

NAB Broadcast Engineering

Conference Proceedings 2007

NAB Publications
1771 N Street, NW
Washington, DC 20036
To Order: 1-800-368-5644
www.nabstore.com

National Association of
NAB
BROADCASTERS