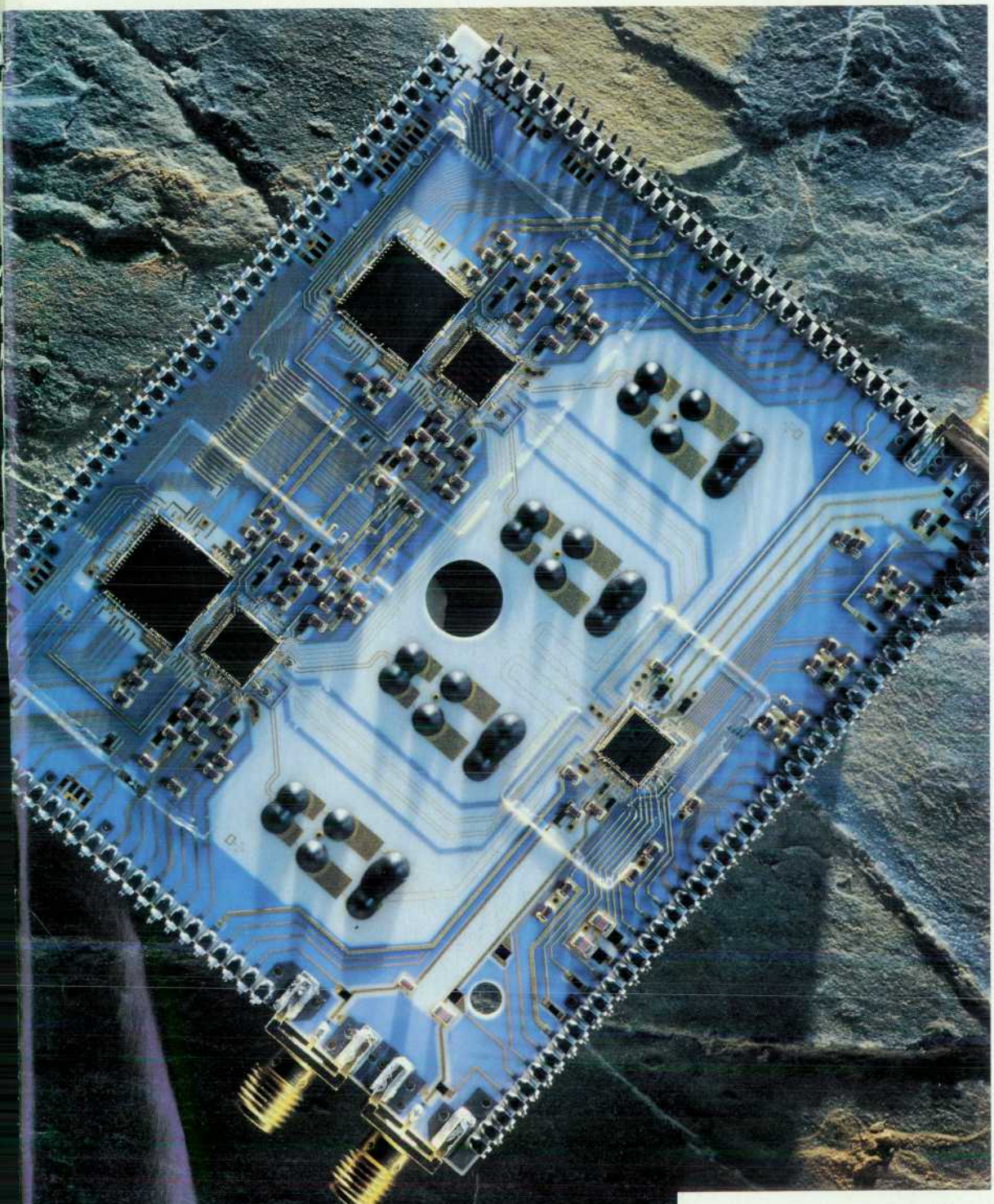
HEWLETT-PACKARD JOURNAL

October 1993



HEWLETT PACKARD

HEWLETT-PACKARD
JOURNAL

October 1993 Volume 44 • Number 5

Articles

Departments

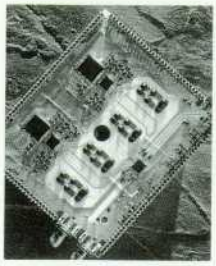
The Hewlett-Packard Journal is published bimonthly by the Hewlett-Packard Company to recognize technical contributions made by Hewlett-Packard (HP) personnel. While the information found in this publication is believed to be accurate, the Hewlett-Packard Company disclaims all warranties of merchantability and fitness for a particular purpose and all obligations and liabilities for damages, including but not limited to indirect, special, or consequential damages, attorney's and expert's fees, and court costs, arising out of or in connection with this publication.

Subscriptions: The Hewlett-Packard Journal is distributed free of charge to HP research, design and manufacturing engineering personnel, as well as to qualified non-HP individuals, libraries, and educational institutions. Please address subscription or change of address requests on printed letterhead (or include a business card) to the HP headquarters office in your country or to the HP address on the back cover. When submitting a change of address, please include your zip or postal code and a copy of your old label. Free subscriptions may not be available in all countries.

Submissions: Although articles in the Hewlett-Packard Journal are primarily authored by HP employees, articles from non-HP authors dealing with HP-related research or solutions to technical problems made possible by using HP equipment are also considered for publication. Please contact the Editor before submitting such articles. Also, the Hewlett-Packard Journal encourages technical discussions of the topics presented in recent articles and may publish letters expected to be of interest to readers. Letters should be brief, and are subject to editing by HP.

Please address inquiries, submissions, and requests to: Editor, Hewlett-Packard Journal, 3200 Hillview Avenue, Palo Alto, CA 94304 U.S.A.

In this Issue



Most laboratory oscilloscopes today are sampling digitizing oscilloscopes. Analog scopes simply can't match their precision and their storage and measurement capabilities. When you begin to look at digitizing oscilloscopes, the first thing you notice is that there are a lot of different kinds. For example, you can find bandwidth specifications from 100 megahertz all the way to 50 gigahertz and beyond. Going a little farther, you find that there are two kinds of bandwidth: one for capturing single-shot phenomena and one for capturing repetitive signals. Ignoring the input amplifier and sampler bandwidths for the moment, the most important bandwidth-determining parameter for single-shot events is the sampling rate—the faster the better. For repetitive signals, it's the equivalent sampling rate, which depends more on the scope's timing precision than on the actual real-time sampling rate. HP designers like to use a rule of thumb that says that the potential bandwidth of the oscilloscope is one fourth of the sampling rate, either real-time or equivalent, as the case may be. The reason there are so many bandwidths to choose from is that different oscilloscopes are optimized for the needs of different applications. (The scope that does everything would be so expensive that no one would buy it.) The HP 54720A/D digitizing oscilloscope, introduced in the article on page 6, is designed to be the fastest for capturing single-shot or infrequent events such as glitches, transients, bit errors, ground bounce, and timing errors in computers and communications systems, high-energy physics phenomena, electrostatic discharge, and laser pulses. Depending on how the user configures it (there are four plug-in slots), it can have a single input channel with a sampling rate of 8 gigahertz, or two, four, or eight channels with sampling rates of four, two, or two gigahertz, respectively. A smaller sibling, the HP 54710A/D, has two plug-in slots and a maximum sampling rate of 4 GHz. For repetitive signals, a precision time base and a new trigger interpolator give the HP 54720/10 an equivalent sampling rate of 1000 GHz! However, the plug-in and sampler bandwidths limit the maximum overall bandwidth to 2 GHz. The plug-ins and the data acquisition system, which features four time-interleaved analog-to-digital converter acquisition hybrid circuits and a new sample-and-filter technique, are described in the article on page 11. Accuracy issues in this interleaved system are discussed on page 38. The article on page 24 gives details of the time base and trigger system. Other aspects of the design of this oscilloscope family presented in this issue are the architectural design (page 51), the mechanical design (page 66), firmware development processes (page 59), a rugged 2.5-GHz active probe (page 31), and a probe fixture for testing the acquisition hybrid (page 73). The article on page 47 reports on a study of the pulse parameter measurement accuracy of the HP 54720A only recently made possible by the availability of a well-characterized 50-GHz oscilloscope as a standard.

A network analyzer measures the characteristics of components or networks as functions of frequency. A spectrum analyzer measures the power in a signal as a function of frequency. Because the two are frequently used together and share some functions, a combined network and spectrum analyzer makes sense. The HP 4396A network and spectrum analyzer (page 76) is designed to offer the capabilities of both instruments without compromising the performance of either, at a lower cost than two separate instruments. The HP 4396A accepts input signals up to 1.8 gigahertz. It has two measurement display channels and can display two spectrum measurements or one network and one spectrum measurement or two network measurements at once. It has the network analyzer's usual A, B, and reference input ports, and right beside them a spectrum analyzer input port. But behind the front panel, all of these inputs go to the same measurement receiver, thereby reducing the cost considerably. To make this possible, a fast-switching, high-isolation multiplexer connects the input ports to the receiver one at a time. The receiver design is described in the article on page 85 and the multiplexer design on page 95. A nice feature is the spectrum monitor mode of the network analyzer ports, which allows the user to get a rough idea of the spectrum of a signal without disconnecting it and reconnecting it to the spectrum analyzer port.

An essential component of any computer-based data acquisition or measurement system is the analog-to-digital converter, or ADC, which converts analog voltages representing real-world phenomena to digital data that can be manipulated by a computer. The basic trade-off in the design of an ADC is conversion rate—the number of samples per second that can be converted—versus the precision with which each sample is converted: a higher conversion rate costs something in precision. The HP E1430A (page 100) is an ADC module designed for instrumentation systems based on the modular VXIbus standard. It includes a ten-megasample-per-second ADC and circuitry for amplitude range switching, filtering, frequency band selection, triggering, data buffering (memory), and multichannel synchronization. While its conversion rate is considerably slower than the two-gigasample-per-second ADC hybrids with which it shares this issue, its precision is much greater. Its precision, in terms of noise, distortion, and nonlinearities (or the relative lack of these defects), is quantified in the article on page 105.

R.P. Dolan
Editor

Cover

This is the acquisition hybrid microcircuit of the HP 54720D and HP 54710D oscilloscopes. There are four of these four-channel hybrids in each instrument. This circuit differs slightly from the acquisition hybrid shown in Fig. 5 on page 14, which is from the HP 54720A and HP 54710A oscilloscopes. Both hybrids have the same function and the same 2-GHz sample rate, but the D version has bigger sample memory chips and a different filter layout.

What's Ahead

The December issue will feature the HP 89410A and 89440A vector signal analyzers, which are designed to measure the magnitude and phase of time-varying and complex modulated signals. In addition to conventional spectrum analysis, they offer a full set of measurements based on digital signal processing. Also in this issue will be the design story of the HP 71450A and 71451A optical spectrum analyzers, an article on North American cellular CDMA (code division multiple access), which is a system for packing more cellular phone users into the available frequency spectrum, an article on HP spectrum analyzer measurement capabilities for testing to the Digital European Cordless Telecommunications (DECT) standard, and an article on a standard data format used by many HP instruments for data interchange.

An 8-Gigasample-per-Second Modular Digitizing Oscilloscope System

For the first time, a digitizing sampling oscilloscope achieves single-shot bandwidths exceeding even the fastest laboratory analog oscilloscopes. The HP 54720/10 oscilloscope combines a 2-GSa/s digitizer, plug-in modularity, and software flexibility to provide the application-specific and general-purpose capabilities needed by designers of high-speed digital devices and systems.

by John A. Scharrer

The need to observe low-duty-cycle or single-shot electrical waveforms has been with us for a long time. This need has been greatly intensified with the advent of high-speed digital computer and digital communication circuits and systems. These high-speed systems are prone to glitches, ground bounce, and timing problems. These problems are usually the result of complex algorithmic processes which by their very nature result in rarely occurring problems that are hard to find and observe, but critical.

Schemes to observe fast single-shot signals have been analog in nature and generally involve storing electron beam traces on a phosphor target in a conventional cathode ray tube. If the phosphor has a fast writing rate and long enough persistence, the image can be photographed with high-speed film. Storage-tube CRT technology allows direct visual observation of the waveforms without a camera. There have been many variations on this theme, but improvements in performance have been very limited in recent years and significant disadvantages to this approach are difficult if not impossible to overcome. Among these are trace blooming, dim traces, CRT wearout, displays burned permanently into the phosphor, and small displays.

Until recently this approach was the only hope of achieving high-bandwidth single-shot capability. Now, with the introduction of the HP 54720 and 54710 digitizing oscilloscope mainframes, performance exceeding that of analog storage oscilloscopes is available. The HP 54720 (Fig. 1) and 54710 use a new high-speed digitizer methodology coupled with major improvements in computing, display technology, and product design to achieve this performance.

The HP 54720/10 system is modular. The HP 54720 mainframe provides four digitizing input slots, which accept plug-ins that offer bandwidths from 500 MHz to 2 GHz, sensitivities

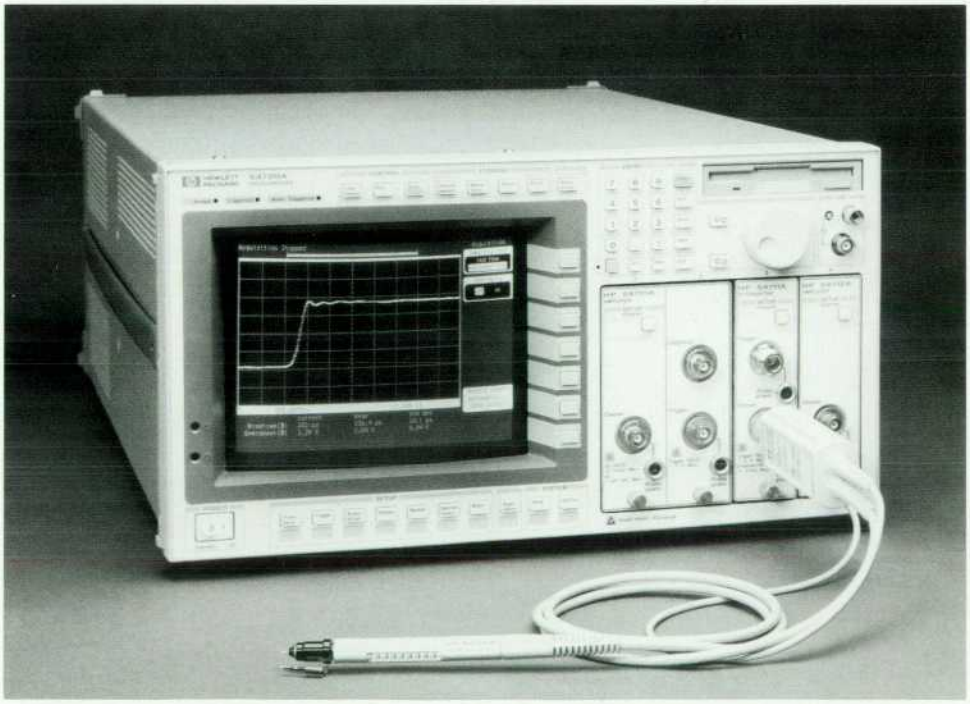


Fig. 1. The HP 54720A is a four-channel, modular digitizing oscilloscope capable of sample rates up to 8 GSa/s. It is shown here with the HP 54701A active probe.

from 2 mV/div, and input impedances of 50 ohms and 1 megohm, depending on the plug-ins selected. Each plug-in slot is matched with a 2-GSa/s (gigasample per second) analog-to-digital converter system, and by choosing a 2-slot-wide or a 4-slot-wide plug-in, sample rates of 4 GSa/s and 8 GSa/s are achieved. The HP 54710 mainframe provides two input slots. The HP 54720D and 54710D mainframes have memory depths of 64K samples per slot, and with 2-slot-wide or 4-slot-wide plug-ins, memory depth extends to 128K and 256K samples, respectively. The HP 54720A and 54710A versions of the mainframes have one quarter of the mainframe memory of the D versions.

Consistent with this performance, an active probe, the HP 54701A, provides 0.5-pF input capacitance and 100-kilohm input impedance while maintaining the bandwidth of the entire system including the probe as high as 1.3 GHz depending on the plug-in selected. The probe itself has a bandwidth of greater than 2.5 GHz and can be powered either from a plug-in or from an external supply. In the past users have been reluctant to use active probes because of their mechanical fragility and the special precautions required to avoid overvoltage at the input. The HP 54701A does not require such precautions because it is protected from damage resulting from static discharge and overvoltage. In addition, it is mechanically rugged, has replaceable tips, and is highly resistant to physical damage.

Important as it is that digital storage oscilloscopes overcome the problems of analog storage, there are far more compelling reasons to move to digital storage technology. Among these are:

- The ability to store and retrieve waveforms for further analysis or for visual observation either in the oscilloscope or other environments such as a workstation
- The ability to observe pretrigger events
- Ease of use.

In addition, the HP 54720/10 project focused on providing accuracy and precision in high-bandwidth time-domain measurements and the flexibility to configure the product in software and hardware so that specific customer application needs can be supported.

Storage and Pretrigger

The HP 54720/10 acquires waveform data in digital form and stores it in memory. The waveform can be observed, scrolled, and zoomed. Cursors can be used for automatic readout and automatic measurements can be performed on the stored data. Waveform data can be routed to internal waveform memory, to an internal flexible disk drive, or to an external computer or peripheral. Digital bus interfaces provided for this purpose include the HP-IB (IEEE 488, IEC 625), Centronix, and a parallel expansion port for very high-speed data transfer.

The HP-IB port achieves data transfer rates greater than 500 kilobytes per second. Waveforms can also be transferred by flexible disk and are formatted for use by other widely used programs including spreadsheets and graphics programs.

The analog-to-digital and memory systems run continuously. When a trigger occurs, all data in memory is stored. Therefore, data occurring before the trigger (negative time) is

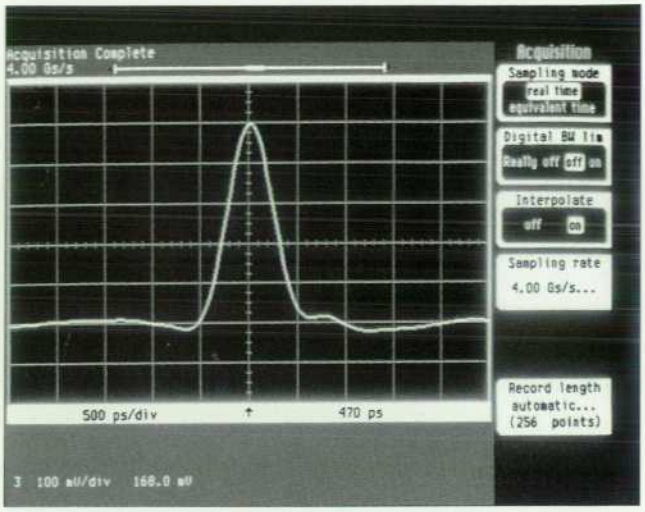


Fig. 2. The HP 54720/10 faithfully reproduces a 500-ps-wide single-shot event.

captured. This greatly facilitates troubleshooting and characterizing a system because trigger points before the point of interest are not necessary.

Of greater significance is the single-shot capability, which allows prefault observation of waveform data even if the event occurs only once. In the single-shot mode (also called the real-time mode), resolution and faithful waveform representation are determined by the analog-to-digital sample rate.

According to the Nyquist criterion, if the sample rate is twice the highest-frequency component of the signal sampled, then the signal can be faithfully reproduced. In reality, the bandwidth of an oscilloscope is down only 3 dB at the specified bandwidth and frequency components beyond the bandwidth will be sampled, causing aliasing and incorrect waveform display. To avoid this situation, Hewlett-Packard uses a rule of 4 times the specified bandwidth for the sample rate required. The HP 54720/10 achieves sample rates of 2, 4, and 8 GSa/S depending on the plug-ins selected, allowing faithful waveform reproduction at 500 MHz, 1 GHz, and 2 GHz bandwidths, respectively (Fig. 2). The maximum available sample rate is traded off against the number of channels available: 2 GSa/s allows 4 channels, 4 GSa/s allows 2 channels, and 8 GSa/s allows 1 channel.

If the signal viewed is repetitive, the equivalent time mode allows the effective sample period to be as small as one picosecond. This is achieved by accurately relating the time from the trigger event to the samples in the capture memory. After each trigger and acquisition, samples are positioned in the waveform record (and onscreen) to build up a high-resolution picture of the waveform. For repetitive waveforms, equivalent time sampling can be used and the Nyquist sample rate is not an issue, but the digitizer sample rate does directly effect the throughput to the display. Using this technique for repetitive signals with the appropriate plug-ins, four channels can be observed at 2-GHz bandwidth. Using the HP 54714A dual-channel plug-in, eight channels can be viewed at 400-MHz bandwidth for repetitive signals.

Ease of Use and Flexibility

An oscilloscope with the advanced measurement capability of the HP 54720/10 is often referred to as a laboratory oscilloscope or a high-performance oscilloscope. The use model is focused on troubleshooting and characterizing high-speed digital systems generally in the laboratory or the preproduction phase of product development. In production this class of instrument is usually used within a system in an automated way. Focus groups of users falling into these classes were used early in the project and results of this exercise indicated the need for familiar controls, high system and display throughput, and application flexibility. These requirements dictated the system architecture and user interface.

The user interface is leveraged from the HP 541xx family of high-performance oscilloscopes with significant improvements derived from user interface studies conducted on the HP 541xx family, the general-purpose HP 545xx family, and competitive products. The resulting interface uses a pop-up menu scheme with very limited depth of menus, an intuitive graphical user interface, and extensive color. Access to often-used measurements is from the front panel instead of being buried in menus or softkey levels. The setup keys such as time base, trigger, and vertical channels are fixed and always available, rather than using softkeys and hidden levels of additional keys.

Historically, ease of use in digitizing oscilloscopes has been enhanced through the use of extensive automatic pulse parameter measurements and functions such as rise time, delay, and pulse width, to name a few. The list is ever growing and the ability to add features is very powerful in extending the value of the initial investment in the product. The fast Fourier transform (FFT), mask testing, histograms, and applications such as communications and computer design were all added after the HP 54720/10 was introduced. To accommodate this flexibility the software in the HP 54720/10 is stored in flash EPROM and SRAM and can be loaded from a flexible disk. Add-on labels and shift keys allow upgrades to be made easily, yet the added features have the same direct access as existing features. The mechanical modularity provided by plug-ins allows flexibility in configuring the inputs to the digitizers and makes it easy to customize systems for specific applications.

Ease-of-use studies indicate that the responsiveness of the oscilloscope to controls and signal changes is crucial to interpreting data, adjusting the device under test, and promoting confidence in the representation of viewed signals. The HP 54720/10's high waveform display rate helps avoid visual aliasing and misinterpretation of the waveforms. To achieve a high display rate, a three-processor architecture was chosen for the HP 54720/10. A CPU controls acquisition hardware and communications, performs automatic measurements, and manages waveform data, while a graphics processor and a custom display processor present the waveform data.

The ability not only to see changes as they occur but also to observe the relative frequency and "freshness" of data was a contribution of the HP variable-persistence analog storage oscilloscopes. With this feature, old data gradually fades away while new, brighter data is written onscreen. Until the HP 54720/10 this feature eluded digitizing oscilloscopes, but

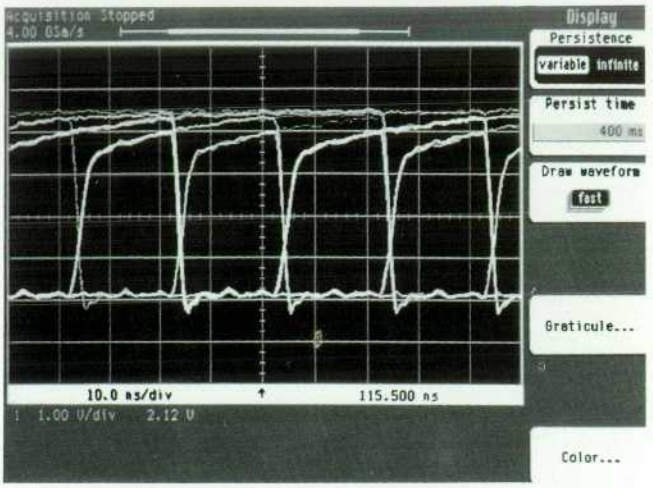


Fig. 3. Variable persistence provides a dynamic picture of waveform changes.

the custom display processor in the HP 54720/10 accommodates this feature while maintaining a high display throughput (Fig. 3).

Accuracy

Although its very high digitizing sample rates and bandwidths open the door to high precision and accuracy in voltage and time measurements over a broad band of frequencies, considerable care had to be exercised in the HP 54720/10 system design to ensure that these benefits were realized. All adjustments for gain, offset, timing, and frequency response are computer-controlled (no manual adjustments) and the necessary calibration routines are automated and use calibration resources resident in the mainframe and plug-ins. Therefore the user has a self-contained accuracy calibration system. If plug-ins and mainframes are intermixed after each has been calibrated, 3% vertical gain accuracy is achieved. If a system best-accuracy calibration is performed, 1% vertical gain accuracy is achieved.

The HP 54720/10 exhibits very low jitter on repetitive single-shot or equivalent time waveform displays because of a new trigger interpolator system. The resolution of the interpolator is 1 ps and jitter on repetitive waveforms is less than 6 ps rms. Again, the internal calibration capability ensures ±30-picosecond time interval accuracy in equivalent time mode and ±50-picosecond accuracy in single-shot mode at a sample rate of 4 GSa/s.

System Design

The HP 54720/10 is a flexible system whose characteristics are defined by the plug-in modules and software installed and the software features selected. A block diagram is shown in Fig. 4.

The boards in the main card cage are interconnected by the system interface bus, which carries address, data, control, and power. The plug-in modules are connected to the system via the module interface bus, which carries address, data, control, interrupts, and power.

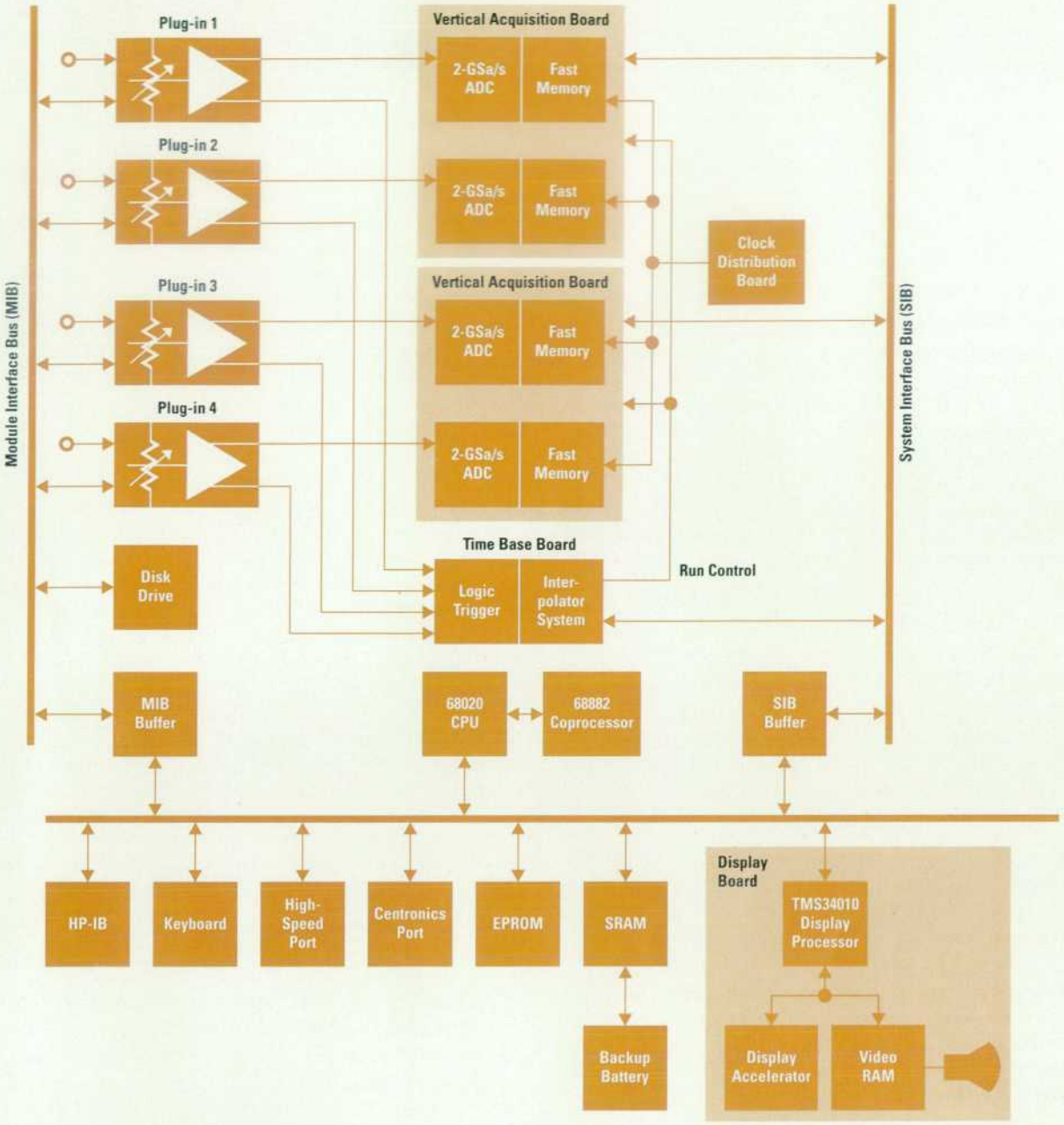


Fig. 4. Simplified block diagram of the HP 54720/10 modular oscilloscope mainframe.

Plug-ins. The plug-ins determine gain, bandwidth, maximum sample rate, and input characteristics. Their general design includes an input attenuator and overvoltage protection followed by a preamplifier which also supplies a signal split for internal trigger pickoff. In the 2-slot-wide plug-in the signal is split to drive two analog-to-digital converter inputs. This split allows interleaving two analog-to-digital converters to achieve a 4 GSa/s sample rate on two channels. A passive four-way splitter in the HP 54722A 4-slot-wide plug-in achieves 8 GSa/s on one channel. The plug-ins also provide probe power and offset voltage for the HP 54701A active probe. The plug-ins are calibrated using the mainframe calibration resources and software, and the calibration factors are stored in the plug-in.

Acquisition System. The main and trigger signals are coupled to the mainframe through very high-bandwidth connectors and are subsequently routed to the analog-to-digital hybrid circuit and the trigger system via semirigid coaxial cable. The analog-to-digital converter hybrid uses a new technique called sample-and-filter, as opposed to sample-and-hold. The sample-and-filter technique is described in the article on page 11.

There are five monolithic integrated circuit chips on the analog-to-digital converter hybrid: a sampler and two analog-to-digital converters are custom HP bipolar chips, and two fast in, slow out (FISO) memories are custom HP CMOS. The analog-to-digital subsystem including very fine-line-geometry filters are constructed on a custom Hewlett-Packard thick-film multichip module. There are two analog-to-digital hybrids on each of the vertical acquisition boards.

The trigger signals from all four slots are input to a custom logic trigger chipset, which provides numerous combinational and sequential logic trigger capabilities. The resultant

trigger is applied to an interpolator counter system, which determines the exact location of the input signal in relation to the trigger event. The waveform captured can consist of one full acquisition memory before the trigger or can be delayed up to one second after the trigger. A dual-ramp interpolator system resolves trigger location within 1 ps. The interpolation takes only six microseconds.

A clock distribution board generates a 2-GHz clock and a phase-related 100-MHz clock signal for each acquisition hybrid. The phase-related 100-MHz clock is also distributed to the time base system.

Computer and Display. The main system CPU is a 68020 with a 68882 floating-point coprocessor, but with an unusual feature. A state machine controls data flow on the CPU bus such that a complete 16-bit data word representing a captured sample can be moved from the FISO memory to CPU memory on each bus cycle by direct memory access (DMA). Similarly, data can be moved by DMA to the display processor and external ports. Coupled with the high-speed graphics subsystem this gives the HP 54720/10 a high-throughput, highly interactive display.

The display processor is a TMS34010 coupled to a custom HP display accelerator. The accelerator does the work of decrementing pixel brightness in variable-persistence mode, erasing the screen, and drawing lines.

The CPU RAM is nonvolatile battery-backed SRAM and the operating code is stored in flash EPROM. This allows completely changing or adding to the operating software using the flexible disk. The state of the entire oscilloscope, including all data, is preserved when the instrument is turned off.

Digital interfaces to the HP 54720/10 include the HP-IB, a Centronix printer port, and a high-speed parallel port which provides direct access to the CPU bus. The two other internal buses—the module interface bus and the system interface bus—interface with the plug-in modules and the internal system boards, respectively.

Software

The software system was a start-from-scratch design, and the size of the task led the design team to consider and use structured design techniques. Indeed, the first half of the software design portion of the project was design, not coding. The resulting design has been very low in defects, and adding recent features such as histograms, FFT capability, and application-specific programs went smoothly and took significantly less time than in less structured designs.

Product Design

The constraints of plug-in flexibility and high-bandwidth performance led to a new modular mechanical design that makes it possible to plug high-bandwidth amplifiers into a mating mainframe with little or no signal degradation. Plug-in modularity from the front is complemented by card modularity in the rear of the mainframe. The HP 54720/10 package is the same physical size as the previous HP 541xx high-performance digitizing oscilloscopes.

Manufacturing

Formal concurrent engineering may require elaborate discipline and tools. The simpler concept of developing manufacturing test tools concurrently with the system and circuit design not only sufficed for this project but was indispensable in achieving the project goals. There are four analog-to-digital multichip modules per HP 54720/10 acquisition system and therefore a high loaded-hybrid yield is imperative. The hybrid test system was developed in parallel with the hybrid and ensures close to 100% loaded hybrid yield. The test system was ready in time to help develop and evaluate early prototypes. Similarly, plug-in and board test tools were defined at project inception and developed concurrently with the HP 54720/10 circuits.

Acknowledgments

A project of this size had a large number of contributors and they are acknowledged in the accompanying articles. In addition, recognition must go to the original concept definition team and management staff consisting of Lynne Camp, Bill Escovitz, Mike Karin, Dave Long, Fred Rampey, Ken Rush, and John Wilson. The manufacturing engineering and process team consisted of Mike Manley, Mike Van Grouw, Van Martin, Angus Foster, Jerry Townsend, Mike Kinney, Pat Ciuba, and Don Hanlon.

An 8-Gigasample-per-Second, 8-Bit Data Acquisition System for a Sampling Digital Oscilloscope

Within the HP 54720/10 acquisition system are sixteen separate sampling and digitizing paths that can be allocated by the user to capture 16K samples at 8 GSa/s or 256K samples at 500 MSa/s or any of various other combinations of sample rate and memory depth. The sample-and-filter sampling technique is an alternative to the conventional sample-and-hold and track-and-hold techniques.

by Michael T. McTigue and Patrick J. Byrne

The data acquisition hardware for the HP 54720 and 54710 sampling digitizing oscilloscopes is designed to provide flexible and configurable signal capture hardware that can be tailored for different customers' needs. Variable-width plug-in amplifiers are used for signal conditioning, buffering, and splitting. The plug-ins deliver the conditioned input signals to the acquisition system, which has a sample-and-filter architecture for improved sample rate and noise performance and a scalable analog-to-digital converter (ADC) architecture that allows input channels to be traded for higher sample rates. The plug-ins support active probing to provide nonintrusive, high-bandwidth connections to the circuit under test.

Plug-ins

The plug-in architecture was decided upon to meet the HP 54720/10 system design goals of configurable signal conditioning and variable signal routing. The function of a plug-in in this system is to accept the input signal to the oscilloscope, provide signal conditioning (termination impedance, ac or dc coupling, and filtering), and then present a properly scaled version of the signal to the acquisition and trigger systems in the mainframe. Plug-ins also provide support and control via a connector on the front panel for active probes and other accessories that may be needed in front of the plug-in to allow interception of a signal.

The plug-in concept provides for variable signal routing by allowing multiple-width plug-ins. The HP 54720 mainframe has four ADC channels and four trigger channels that are connected to four plug-in slots. With plug-ins that are two slots wide or four slots wide, the input signal can be fed to more than one ADC so that interleaving can be done in the mainframe and higher sampling rates can be provided.

Plug-ins protect the oscilloscope's inputs from ESD (electrostatic discharge) and excessive signal levels, and store calibration factors that reflect their own accuracies so that when they are plugged into any mainframe there is a specified level of accuracy. The intent of the input protection is that ESD or overvoltage will cause no damage to any part of the oscilloscope. In practice, and especially with high-frequency circuits, it is not always possible to protect the input from all types of overstress. Specifically, it is the plug-ins' job to protect the ADCs and trigger circuits in the mainframe from any damage. If a plug-in is damaged by excessive stress the cost to the customer is moderate, but if the ADC or trigger inputs in the mainframe are damaged, the cost to the customer could be substantial.

The general block diagram of a plug-in is shown in Fig. 1. The trigger and vertical output connections are between the back of the plug-in and the bulkhead of the mainframe and use high-quality, floating, 50-ohm coaxial connectors. These connectors provide a high-bandwidth interface between the plug-in and the mainframe so that future plug-ins or mainframe acquisition systems can use the same mainframe. The power, data, and control lines of the module interface bus

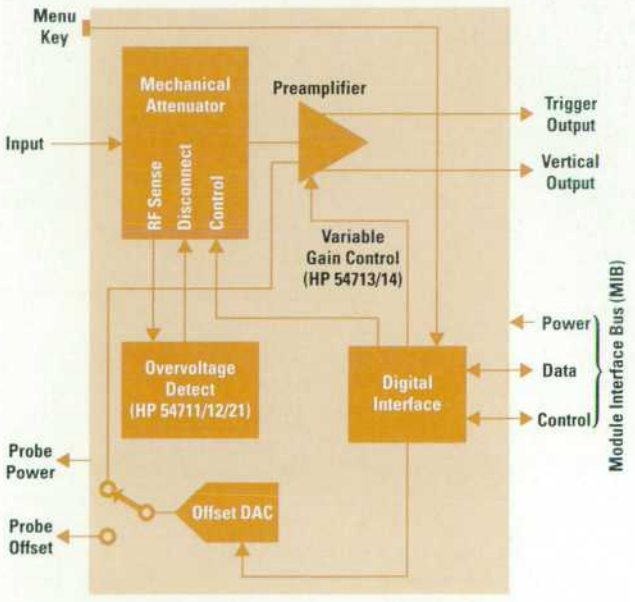


Fig. 1. General block diagram of a plug-in for the HP 54720/10 sampling digitizing oscilloscopes.

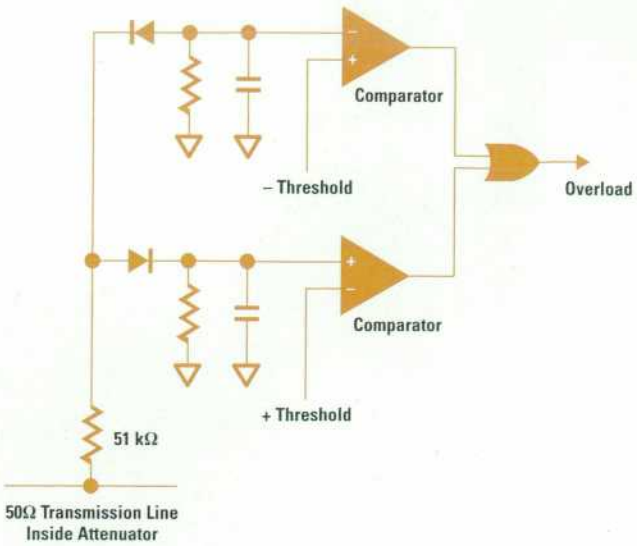


Fig. 2. Integrating peak detector. An overload indication causes the attenuator to disconnect the input signal and report the overload to the mainframe CPU.

are connected using a multipin D-shell connector. The input connector type is determined by the application for the plug-in and the bandwidth needed. Most applications at 1 GHz and below would use a BNC-type connector. Higher bandwidths use a type N connector, which is easily adapted to SMA, but much more reliable than a permanently attached SMA connector.

HP 54713A Single-Slot Amplifier Plug-in

The HP 54713A plug-in is a 500-MHz high-impedance amplifier plug-in designed to accommodate high-impedance resistive divider probes. It allows users to connect to or probe circuits that cannot drive 50-ohm loads or the low resistance of resistive divider probes (500 or 1000 ohms). The input to the HP 54713A is 1 megohm shunted by 7 pF. The 500-MHz bandwidth is adequate for many applications and allows various types of signal conditioning such as ac or dc coupling, switchable 50-ohm termination, and low-pass and high-pass filtering for removing noise when needed. Also, the 500-MHz bandwidth is only one fourth of the sample rate of 2 GHz, which helps limit aliasing for real-time measurements. The HP 54713A uses the same high-impedance attenuator/amplifier as the HP 54500 Series of benchtop oscilloscopes. This is a thick-film hybrid circuit containing a mechanical attenuator and a bipolar IC housed in a custom shield. The use of this integrated attenuator/preamplifier allows the HP 54713A to provide a general-purpose oscilloscope front end at a reasonable price per channel.

HP 54714A Single-Slot Plug-in

The HP 54714A plug-in is a dual-channel version of the HP 54713A. It allows two channels per HP 54720/10 slot. Multiplexing between the two channels is done in the plug-in, before the vertical and trigger signals are sent to the mainframe. This plug-in allows the HP 54720 system to have up to eight channels, four of which can be simultaneously sampled. This plug-in provides a higher channel count for applications such as ATE.

HP 54712A Single-Slot Amplifier Plug-in

The HP 54712A is a 1.5-GHz amplifier plug-in. It provides a system bandwidth of 1.1 GHz and has a fixed input impedance of 50 ohms. This plug-in will allow up to four 1-GHz bandwidth channels in the HP 54720 system. Since the sampling rate is only twice the bandwidth, it is normally more useful in equivalent time mode, since some aliasing may result. However, if the input signal is known to be band limited to approximately 500 MHz, then the use of this plug-in for single-shot measurements will produce a more accurate result than a lower-bandwidth plug-in.

The attenuator used in this plug-in provides the normal 1, 2, 5 attenuation sequence used in oscilloscopes and provides either ac or dc coupling and a signal pickoff for overload detection. The ac/dc selector switches a 0.047-μF capacitor into or out of the signal path. This allows ac coupling (dc rejection) of signals that are superimposed on a dc signal. The pickoff for overload protection is a 51-kilohm resistor connected to the 50-ohm transmission line. This resistor goes to an integrating peak detector. The circuit is shown in Fig. 2.

The intent of the pickoff circuitry is to disconnect the input signal if it exceeds a certain level for a period of time. Excessive dc and signals into the few-MHz range will be detected by this circuit and cause the attenuator to open one of its sections. The reason an integrating peak detector is used and not just a comparator is so that temporary spikes (such as hooking up an external dc blocking capacitor) will not cause the attenuator to open. Protection from spikes and ESD overvoltages is provided by back-to-back diodes in the signal path.

The preamplifier in the HP 54712A is a thick-film hybrid circuit. The circuit provides the buffering, gain, and signal splitting required for trigger and vertical outputs (see Fig. 1). This preamplifier achieves a typical bandwidth of 1.5 GHz or better and has a gain of six. Its schematic diagram is shown in Fig. 3. It is implemented using chip and wire technology on a thick-film substrate. The gain and bandwidth are provided by an Avantek silicon MMIC MODAMP, which is a two-transistor shunt feedback stage. Using active laser trimming, the overall gain and output impedance are adjusted using a resistor network at the output. A combination of low V_{ce} on the input transistor pair and feedback in the MODAMP minimizes thermal transients that would cause nonflat step response. Long-term flatness and dc accuracy are achieved by means of a dc loop around the amplifier with the gain of the loop matched to the gain of the amplifier.

HP 54721A Two-Slot Amplifier Plug-in

The HP 54721A is a two-slot-wide version of the HP 54712A. It uses the same preamplifier, but has two vertical outputs rather than only one (see Fig. 3). Thus the HP 54721A can split the input signal and route it to the ADC inputs of two HP 54720/10 slots so that interleaving can be done. With this plug-in a system bandwidth of of 1.1 GHz is achieved with a sampling rate of 4 GHz. This provides a sampling-rate-to-bandwidth ratio of approximately 4:1 which is adequate to limit aliasing in real-time measurements. This plug-in also provides an external trigger channel and two power and control ports for active probes.

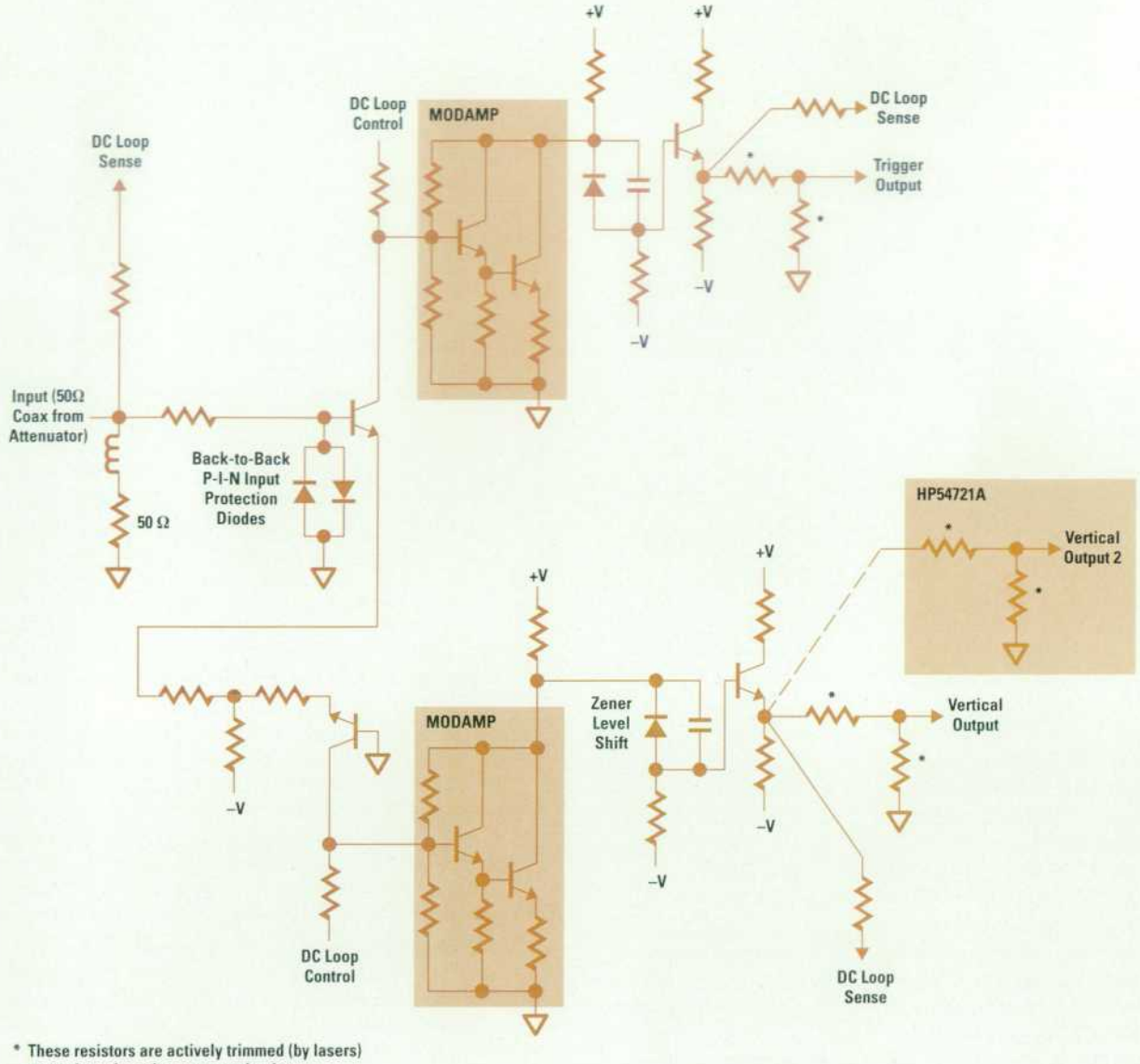


Fig. 3. Simplified schematic diagram of the HP 54712A and 54721A plug-ins.

HP 54711A Single-Slot Attenuator Plug-in

The HP 54711A plug-in is designed to allow access to the maximum bandwidth of 2 GHz in an HP 54720/10 mainframe slot. It also makes it possible to extend the trigger bandwidth of the HP 54720/10 system to 2.5 GHz and provides power and control for two active probes. The only elements in the vertical signal path through the plug-in are a 50-ohm step attenuator and an input limiter. The attenuator provides input scaling and input level pickoff like the HP 54712A described above. The limiter protects the mainframe's ADC inputs from ESD and excessive input signal. To achieve the trigger bandwidth, the HP 54711A leverages the 2.5-GHz trigger hybrid used in the HP 54120 family of sampling oscilloscopes.

HP 54722A Four-Slot Attenuator Plug-in

The HP 54722A is a four-slot-wide attenuator plug-in. It allows an input signal to be split four ways and routed to the ADCs of all four HP 54720 slots. This configuration achieves an 8-GHz sampling rate by interleaving the four 2-GHz sampling rate ADCs. The bandwidth of this configuration is 2 GHz, which is one-fourth the sampling rate. This helps limit aliasing for real-time measurements. Since passive resistive dividers are used to split the input signal, the maximum full-resolution sensitivity is 80 mV/div. A step attenuator is used to achieve a minimum sensitivity of 1V/div. Input protection is provided by the same coaxial limiter as in the HP 54711A plug-in. Triggering is external only and leverages the 2.5-GHz trigger hybrid used in the HP 54120 family of oscilloscopes. This plug-in is designed for single-channel, high-bandwidth, real-time measurements. Its bandwidth and good signal fidelity provide customers with state-of-the-art viewing of fast single-shot events.

Acquisition System

The heart of the performance contribution of the HP 54720/10 is the acquisition system that digitizes and stores the input waveform. To perform high-resolution and high-sample-rate data acquisition at the same time, an interleaved ADC system was chosen. The block diagram of the HP 54720

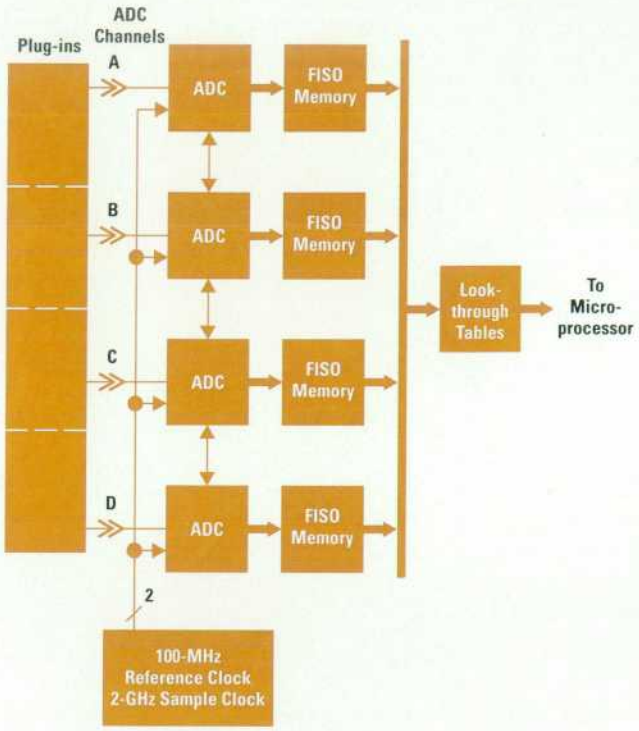


Fig. 4. Block diagram of the HP 54720 acquisition system.

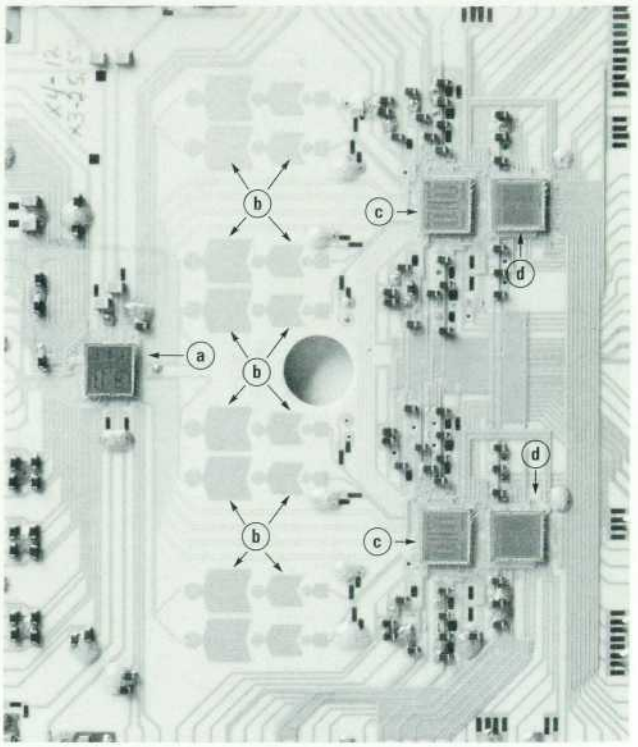


Fig. 5. Acquisition hybrid containing the sampler IC ⓐ, low-pass filters ⓑ, analog-to-digital conveter chips ⓒ, and FISO memory chips ⓓ. There are four acquisition hybrids in the HP 54720 mainframe, one per vertical channel.

acquisition system is shown in Fig. 4. There were five key goals for the ADC system:

- 2-GSa/s sample rate on each of the four channels, storing 8 μs of data
- 2-GHz bandwidth on each of the four channels
- The ability to interleave the channels to achieve sample rates up to 8 GSa/s
- The ability to trade sample rate for memory depth by interleaving memory
- The ability to trade both sample rate and memory depth for improved resolution.

These goals had to be achieved with technologies that were available and could be manufactured reliably and at competitive costs. To meet the aggressive performance goals and the manufacturing goals simultaneously, a number of techniques were employed:

- A new sampling technique called sample-and-filter, which reduces the performance requirements on the IC process technology while avoiding GaAs processes in favor of more conventional silicon bipolar processes.
- Narrow-aperture bipolar sampling circuits, which provide maximum bandwidth with minimum noise contribution.
- Thick-film ceramic hybrid technology, which provides exceptional high-frequency signal propagation, signal isolation, and passive device matching.
- A clocking and control scheme that provides for interleaving of the acquisition hardware resources, either in time (for higher sample rates) or in memory depth (for longer record length), or in some optimal combination of these alternatives.

Acquisition System Overview

The HP 54720 acquisition system is composed of four identical hybrid microcircuits and control and clocking circuits to support the interleave system. The key performance specifications of the system are determined by the hybrid microcircuits, each of which represents one ADC channel. Each hybrid microcircuit consists of five silicon integrated circuits along with passive thick-film ceramic filters combined on one ceramic substrate as shown in the photomicrograph, Fig. 5. On each hybrid are a four-channel sampler IC, four passive filters, two dual 500-MSa/s 8-bit bipolar flash ADC chips (see photomicrograph, Fig. 6), and two dual 16K-sample fast in, slow out (FISO) memory chips. The complete hybrid dissipates 17 watts and measures 3 by 4 inches. To ensure good hybrid yields, a comprehensive at-speed wafer test is employed for both bipolar chips as discussed in the article on page 73.

This hybrid ADC channel is repeated four times within the HP 54720 oscilloscope. A 2-GHz clock is distributed to each of the four ADC channels as shown in Fig. 4. The clock is phase-locked to a reference 100-MHz system clock to provide time alignment of the sample points. Discussion of the sample point phase-locking scheme will be presented later in this article.

Within an ADC channel (one hybrid), there are four separate sampling paths, as shown in Fig. 7. In each path are a bipolar narrow-aperture sampler, a low-pass filter, a 500-MSa/s 8-bit flash ADC, and a 16K-sample FISO memory. Therefore, within the HP 54720 system, there are 16 separate 500-MSa/s, 16K-deep ADC paths. These can be combined to provide a single channel with a sample rate of 8 GSa/s with 256K memory, four channels with a sample rate of 2 GSa/s with 64K memory, or some other combination of channel count, sample rate, and memory depth as preferred by the user.

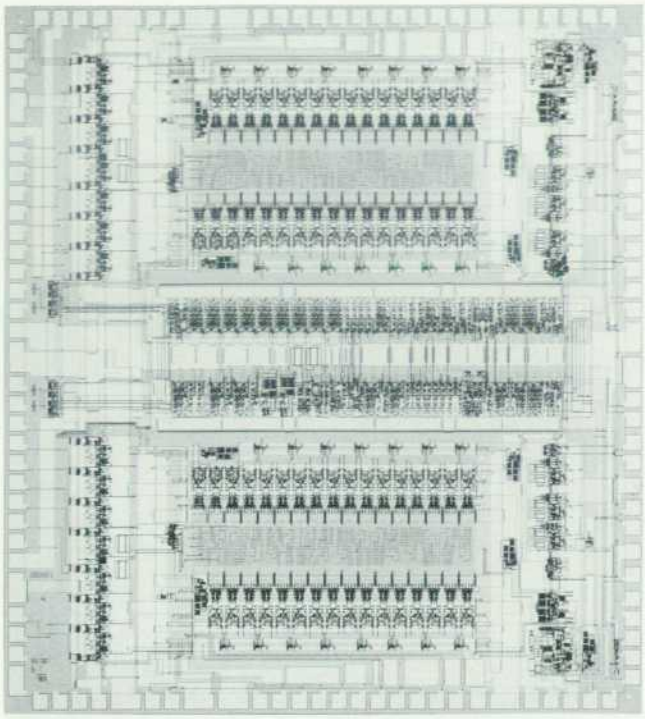


Fig. 6. Dual 500-MSa/s 8-bit bipolar flash ADC chip.

The secret to allowing this sample rate/memory depth optimization is the interleaved clocking scheme and the control of the FISO memory chips to create a chain of memory in acquisition time. The interleaved clocking scheme is provided by timing circuitry on the bipolar sampler chip, which generates four 500-MHz sample pulses that clock four sets of sampling circuitry. The phases of these sample pulses can be shifted to any one of three states, as shown in Fig. 8, depending on the desired sample rate. Fig. 8c, for example, shows full interleaving of the sample clocks within one hybrid, giving 500 ps between sample points, equivalent to 2 GSa/s. If hybrids A and B in Fig. 4 are placed in this mode and then their clocks are staggered by 250 ps, this achieves 4 GSa/s acquisition, effectively combining two channels to achieve twice the sample rate. This same technique can be extended, using the HP 54722A plug-in, to 125-ps sample points, or 8 GSa/s. Very precise time-interleaved edge control and jitter is required to realize high resolution at these sample rates. This topic will be covered in more detail later, along with the description of the FISO memory control.

Another possible use of the ADC channels is to combine the paths for more vertical resolution while sacrificing both sample rate and memory depth. In Fig. 8a, all four paths on a hybrid are sampling the same waveform at the same time. Offsetting the reference voltages within the ADC chips by one fourth of an ADC code width will yield four times the resolution. This technique for increasing vertical resolution, illustrated in Fig. 9, can be extended to multiple hybrids, limited only by the broadband noise limits of the system. The HP 54720/10 has less than 300 μV rms of noise on a 160-mV full-scale input.

Sample rate is of secondary importance in equivalent time mode, since the trigger event is repetitive and stable with respect to the incoming signal. Therefore, the high-resolution feature of the interleaved ADC system is employed in equivalent time mode when the user wants maximum voltage resolution for careful signal characterization applications.

In summary, the HP 54720 acquisition system paths can be combined as resources to achieve higher sample rates up to 8 GSa/s, higher memory depths up to 256K samples deep, or higher resolution up to the broadband noise limits of the system, or any combination of these alternatives to optimize the measurement for the user's need. Table I shows the combinations of resolution, memory depth, and sample rate that are possible with the HP 54720 acquisition system for single-slot, two-slot, and four-slot plug-ins.

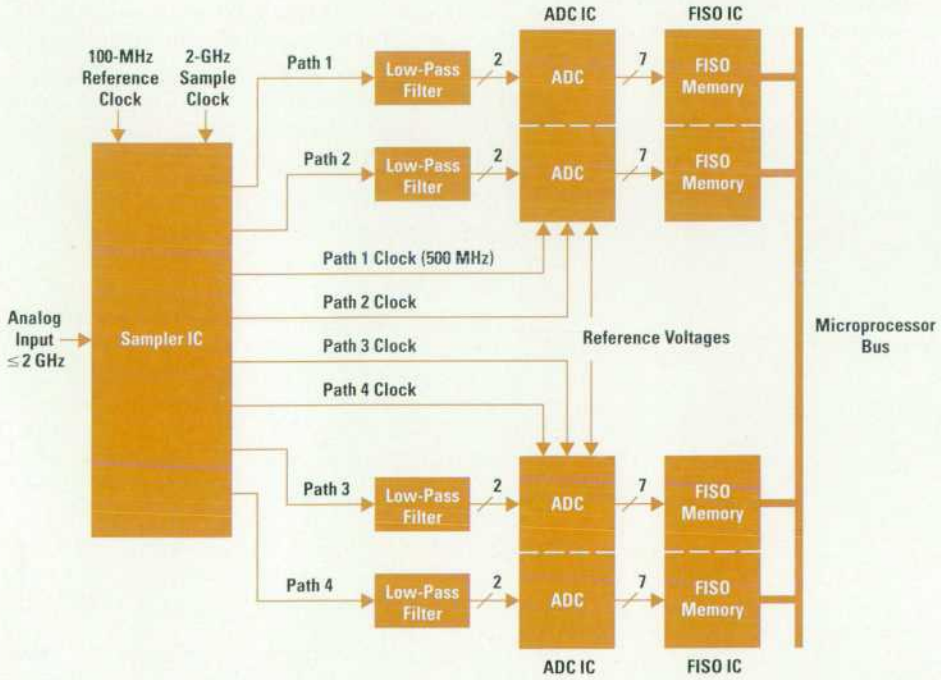


Fig. 7. Acquisition hybrid block diagram. There are four sampler channels on each hybrid.

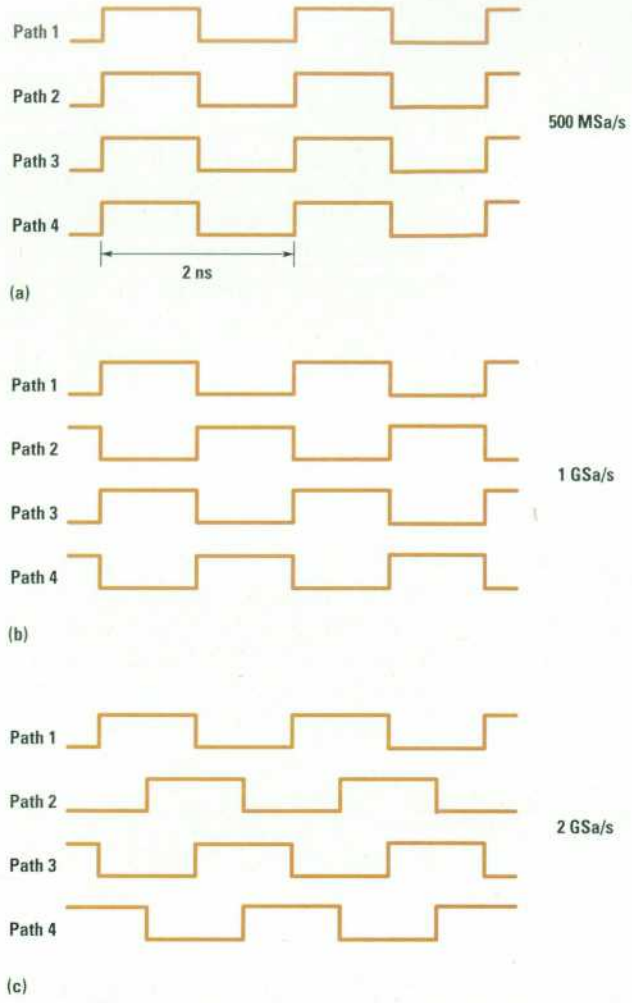


Fig. 8. Sample clock timing diagrams for 500-MSa/s, 1-GSa/s, and 2-GSa/s sample rates.

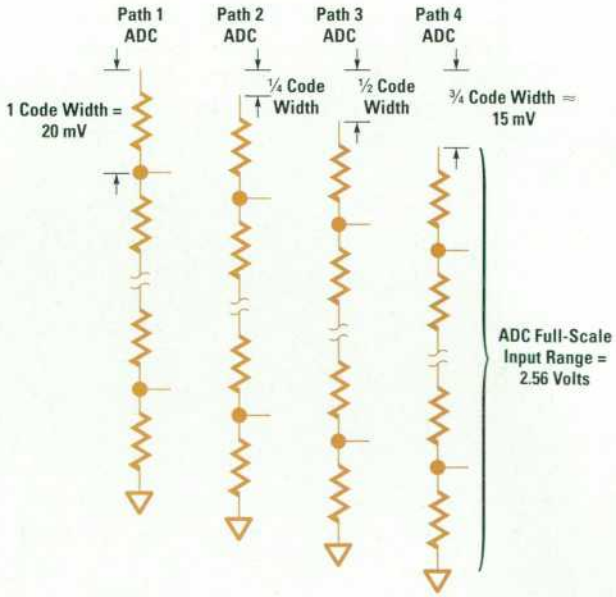


Fig. 9. Using multiple ADCs for higher resolution by offsetting their reference voltages.

Table I
HP 54720D Digitizing Oscilloscope Sample Rate, Memory Depth, and Channel Count Options

Real-Time Modes (8-Bit Resolution, BW = f_s/4)

Number of Channels	Sample Rate (f_s)	Memory Depth* (samples)
8	2 GSa/s	32K**
4	2 GSa/s	64K
2	4 GSa/s	128K
1	8 GSa/s	256K

Equivalent Time Mode (9-Bit Resolution, BW = 2 GHz)

Number of Channels	Sample Rate (f_s)	Memory Depth (samples)
8	500 MSa/s	16K**
4	500 MSa/s	32K
2	500 MSa/s	32K
1	500 MSa/s	32K

* The HP 54720A has one-fourth the memory shown here.

** There are eight input channels but only four ADC channels and four sample memories, so the system will acquire four inputs simultaneously, then acquire the other four, and so on.

Sample-and-Filter Technique

The sample-and-filter data acquisition technique overcomes two important technology barriers to high-performance interleaved ADC systems. First, it reduces the bandwidth requirements of the flash ADC chip that follows the sampler, and second, it eliminates the need for high-impedance hold circuits, which can have anomalous behavior over the operating conditions. At the same time, it imposes some new timing and gain control requirements on the ADC system.

Fig. 10 shows the signal and clock path on the ADC hybrid. At the far left, the input signal is applied to the bipolar sampler chip, where it is buffered through a transconductance amplifier that has gain G_m. A narrow pulse (175 ps in this case) of current proportional to the input voltage is applied to an RC network. During this time, which is small with respect to the RC time constant, a charge proportional to the input voltage is deposited on the capacitor, charging the capacitor to a voltage equal to:

$$V_c = V_{in}G_m(175\text{ ps})/C$$

When the sampler current is turned off, this voltage discharges through the resistor R. The choice of resistance is dictated by the sample period, in this case 2 ns between samples. The resistance must be kept low to ensure full discharge before the next sample occurs. Because the waveform just described does not reach full amplitude, there is no requirement to hold the signal on the capacitor, as there is in a sample-and-hold circuit. This allows the use of high-current buffer circuits after the sampler RC network, guaranteeing high performance by avoiding slew-rate limitations, and reducing the noise voltage contribution from the buffer input stage.

The second key advantage of the sample-and-filter technique is the relaxed intermediate frequency (IF) bandwidth requirements. This is important because flash converters have

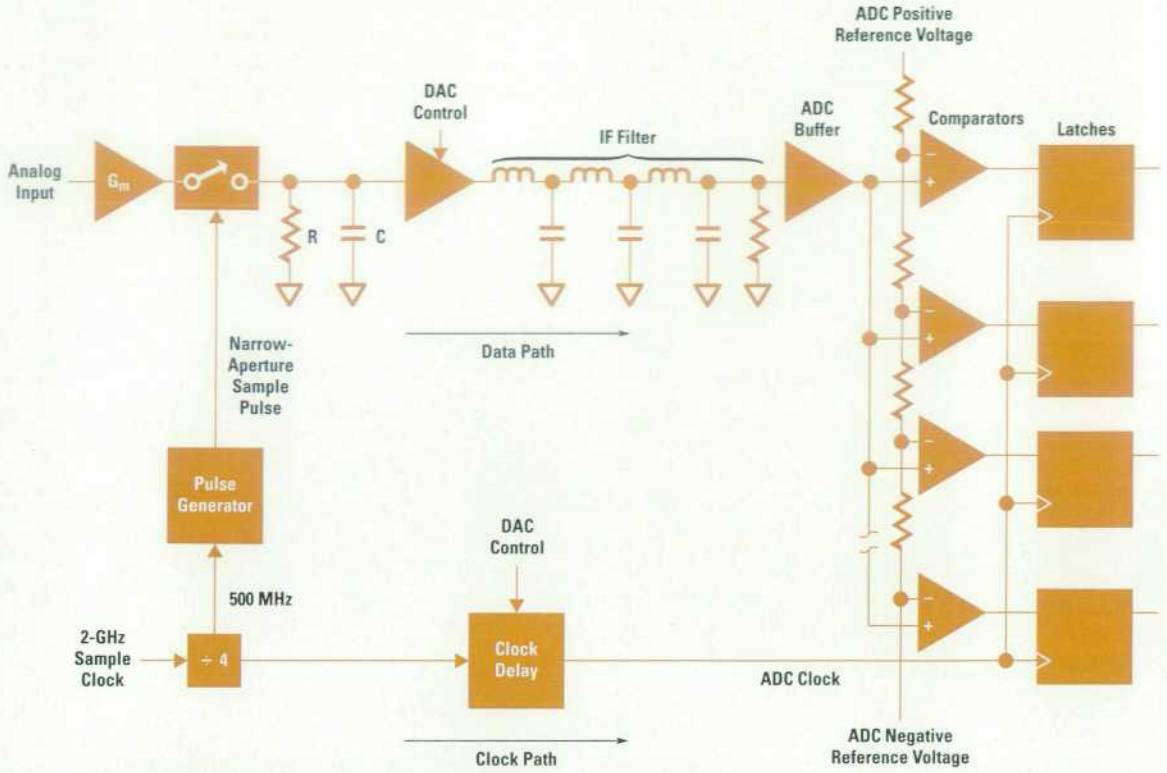


Fig. 10. Simplified schematic of the clock and data paths through the sampler, filter, and ADC of one sampler channel, illustrating the sample-and-filter technique and critical clock calibration.

a well-known characteristic of losing differential linearity at high input slew rates.[1] This is a nonlinear effect caused by the edge-rate-dependent propagation delays through the buffer and comparator circuits. Reducing the edge rates into flash ADCs by even a factor of two greatly improves the feasibility of the design within a given IC process generation. To show how the sample-and-filter technique realizes an advantage in this regard, Fig. 11 illustrates the timing of a sample-and-filter system (Fig. 11a) compared to the more traditional sample-and-hold (Fig. 11b) and track-and-hold (Fig. 11c) systems. The intermediate signal between the sampler and the flash ADC in the sample-and-filter system has one-half the bandwidth requirement of a sample-and-hold system and one-quarter that of the track-and-hold system. In the sample-and-filter system two full periods, 4 ns in this case, are allowed for the waveform to settle before the next sample is digitized by the following ADC. This long settling time is allowed because the waveform does not have to settle to a hold value but is always returned to the baseline. The product of the bandwidth and the settling time is a constant equal to one, assuming Gaussian settling to 0.1%. Fig. 12 illustrates a series of sample-and-filter waveforms that are overlapped as they are in the ADC path. Fig. 12a is a low-frequency input, Fig. 12b is a 250-MHz full-scale sine wave, and Fig. 12c is a full-scale dc input. Figs. 12b and 12c illustrate that the harmonic content of the IF composite signal is limited to 250 MHz full-scale (Fig. 12b) and 500 MHz quarter-scale (Fig. 12c). These signals have similar slew rates and illustrate the reduction in IF bandwidth requirements as a result of using the sample-and-filter technique.

Additional design requirements must be carefully analyzed to realize the advantages of the sample-and-filter technique. As shown in Fig. 10, after the sampling action, the signal is passed through a linear IF filter, where the delay, attenuation, and bandwidth reduction are controlled by precision thick-film components. The delay through this filter (approximately 1.2 ns) must be matched with a clock delay circuit on the sampler and ADC chips so that the ADC sample occurs at the peak of the IF waveform, where the slew rate is minimized. At the same time, the IF filter attenuates the peak amplitude of the IF waveform as it smears the energy captured in the narrow-aperture sampler over time, which creates poorer signal-to-noise ratios. Both of these factors are minimized by using a well-controlled ceramic thick-film printing technology to determine the filter components. Nonetheless, on-chip delay circuits and gain adjustment are incorporated to compensate for any delay and gain mismatches, some of which are inevitable because of process mismatches between the IC processes and the hybrid process.

The delay through the IF path is calibrated by firmware. A dc input equal to a full-scale positive voltage is placed on the sampling channel and the clock delay is varied by means of a control DAC to search for the peak IF voltage. Using Fig. 12c as a reference, a 500-MHz quarter-scale sine wave has approximately 160 ps during which it is within one 8-bit code of the peak value.[2] The calibration technique seeks to place the ADC sample clock within this window width.

Sampler IC

The front-end sampler integrated circuit on each hybrid microcircuit in the acquisition system contains four sampler channels. The key design objectives were:

- High bandwidth in the buffer and amplifier circuits (>3 GHz)
- Low distortion of full-scale input signals

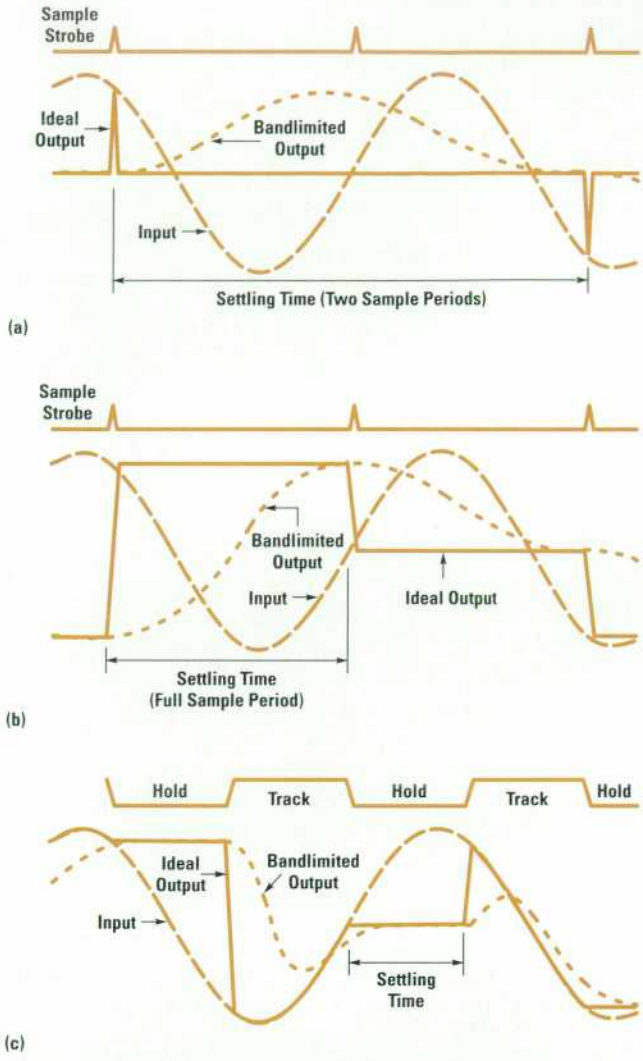


Fig. 11. The sample-and-filter technique compared with sample-and-hold and track-and-hold, showing IF bandwidth and settling time comparisons. (a) Sample-and-filter. Required bandwidth = ½ sample rate. (b) Sample-and-hold. Required BW = sample rate. (c) Track-and-hold. Required BW = 2 × sample rate.

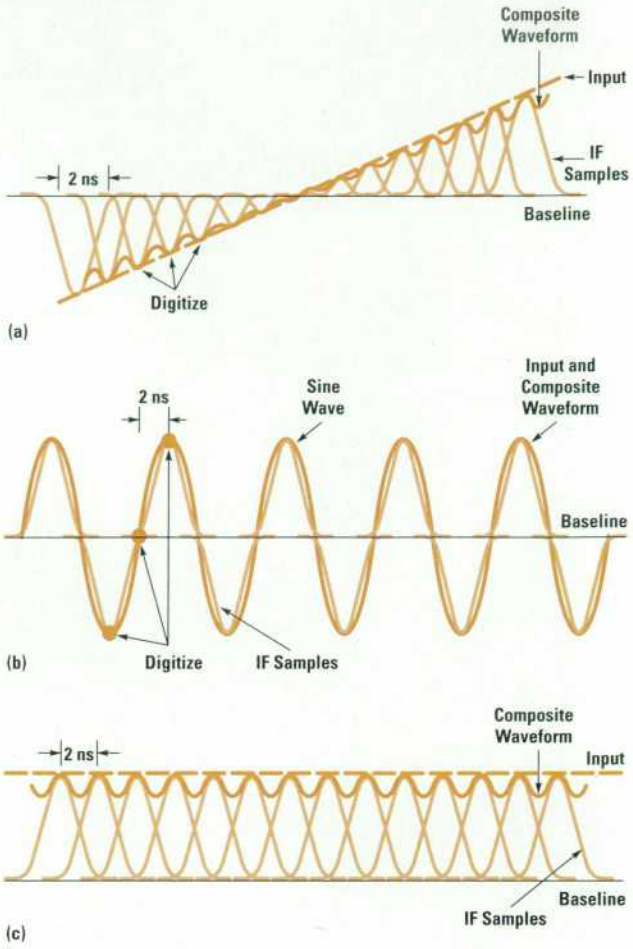


Fig. 12. Overlay of multiple pulses from a sample-and-filter output. (a) Slow-slew-rate input signal. (b) 250-MHz full-scale input. (c) Dc input.

- Low noise—200 μV rms referred to the input
- Well-controlled sampling action in narrow apertures.

The bipolar sampler circuit that produces these results is shown in Fig. 13. It is a differential current-mode sampler implemented in a 13-GHz custom bipolar process. The single-ended input voltage is applied to the V_{in} pad and converted to a differential current from the Q2/Q3 pair. The emitter degeneration of R_1 and R_2 determines the distortion of the front end and dominates the noise performance.

The differential current passes through the common-base stage formed by Q5, Q6, Q7, and Q8. Sample strobes with 175-ps pulse width are delivered to the bases of these transistors to turn on the sampling current into the sampling RC network (R_9, R_{10}, C_1, C_2). The sample pulse width is critical to the bandwidth of the system, as determined by the Fourier transform of the narrow aperture. A 175-ps pulse width corresponds to 2-GHz bandwidth.[2] The pulse width also affects the pulsed gain of the system because it determines the duty cycle of the sample pulse. It is important to deconvolve these two different effects of the sample pulse

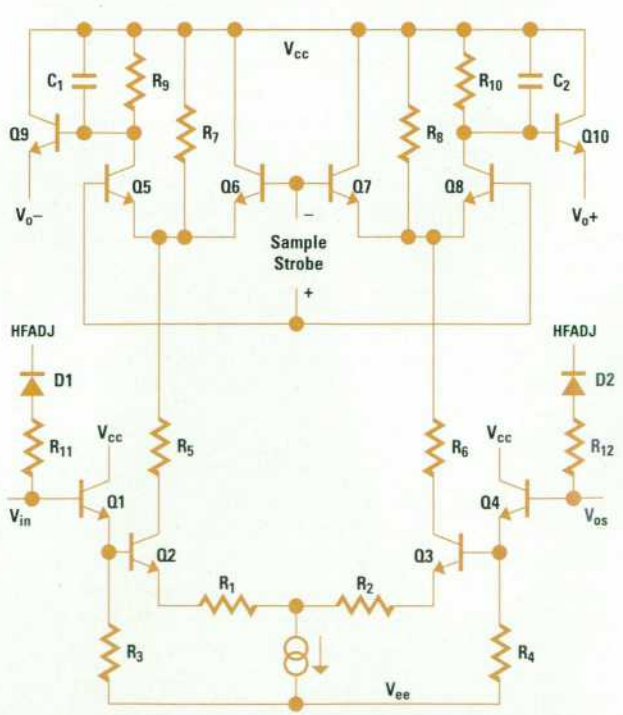


Fig. 13. Bipolar differential current-mode sampler schematic diagram.

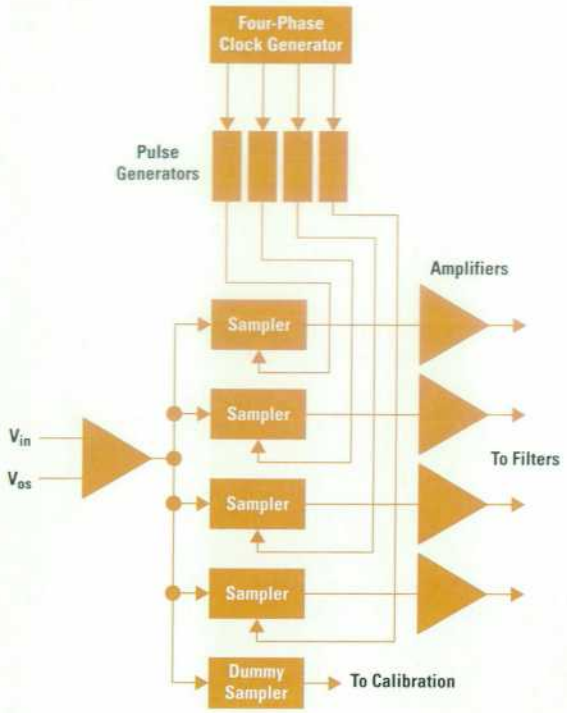


Fig. 14. Sampler channels with dummy sampler for calibration of gain and bandwidth.

width so that the gain and bandwidth of the calibrated system are both deterministic. To facilitate this deconvolution, a dummy sampler channel is added to each sampler IC as shown in Fig. 14. This dummy channel replicates the pulsed gain of the other four sampler channels but is not affected by the sample pulse duty cycle. The devices on the chip are well-matched (within a small part of one percent) and the layout of the devices is kept symmetrical to ensure that this compensation method works well.

Another compensation method is used on the input of the sampler stage. When multiple hybrids are interleaved, the bandwidth must match to ensure good performance at high input frequencies. There is a bandwidth adjustment circuit on the front end of the sampler. Under DAC control, the HFADJ line is modified to ensure that the bandwidths of the various hybrids match. The reverse-biased diode (D1 in Fig. 13) is used as a variable capacitor in this method. Fig. 15a shows an envelope of a 2-GHz input going into the HP 54722A four-slot plug-in with the sample rate set to 500 MSa/s and the memory depth set to 256K samples. The four hybrids are marked, A, B, C, and D. Hybrids A and C have their HFADJ input set to one extreme of the control range and hybrids B and D have their HFADJ input set to the other extreme of the control range. At a 2-GHz input frequency, there is a noticeable effect on the amplitude of the signal. The HFADJ control is attenuating the A and C inputs with more capacitive load than the B and D inputs. The control range exhibited here is approximately 1 dB at 2 GHz. This is enough to compensate for any process variations of the sampler IC bandwidth. Fig. 15b shows the same waveform after the bandwidth calibration factors are calculated and applied to the HFADJ inputs. The effective bandwidths of the four hybrids now match and full effective bits are delivered by the interleaved ADC system at 8 GSa/s for input signals up to 2 GHz.

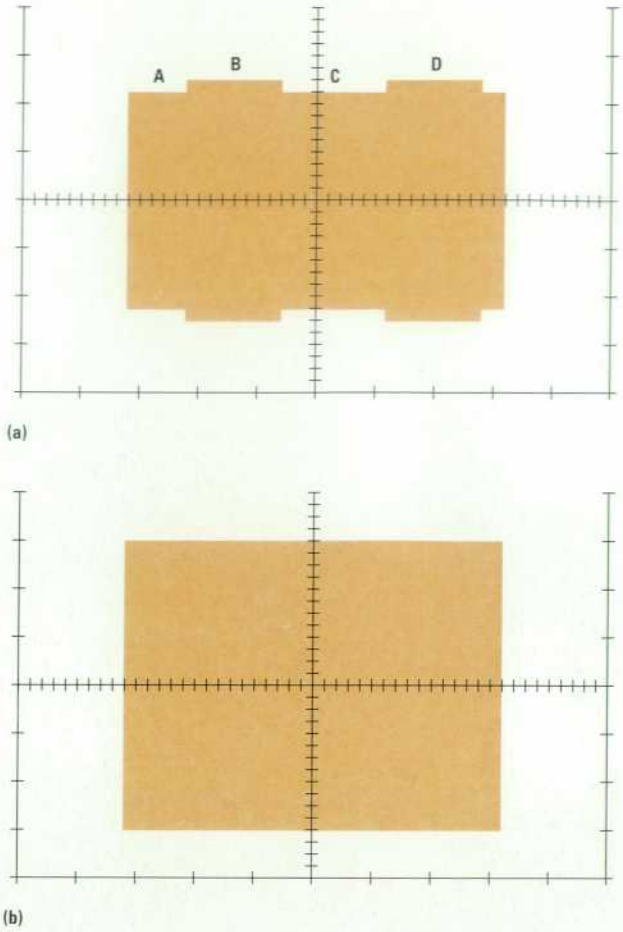


Fig. 15. Effect of HFADJ input on bandwidth matching, (a) without and (b) with calibration.

The complete sampler IC, incorporating four sampler channels with postamplifiers for gain control and adjustment and the clock and sample pulse generation circuits, is integrated on a single bipolar IC, dissipating 5.5W on a 26-mm^2 die. The chip is shown in Fig. 16.

IF Filter

The IF filter design is the single most critical design task in making the sample-and-filter technique work properly. The IF filter (labeled "Low-Pass Filter" in Fig. 7) must have well-controlled attenuation, delay, and settling time and must be as symmetrical as possible to deliver a broad peak with minimum slew rate for digitization by the ADC.

The IF filter is implemented as a passive, thick-film on ceramic, six-element Gaussian low-pass filter. It is shown photographically in Fig. 17. A differential Gaussian filter was chosen to provide exceptional settling behavior and feasible integration. The inductors range from 7 to 32 nH and are designed as spirals in ceramic while the capacitors range from 0.2 to 5 pF and are designed as plates. The key design challenges in the filter are a result of the fan-out of the signal from the sampler IC to the ADC chips. The launch path must be treated as part of the first inductor, yet accumulates resistance and capacitance en route and so must be modeled as a lossy transmission line. Another modeling issue that complicates the design of the filter components is the

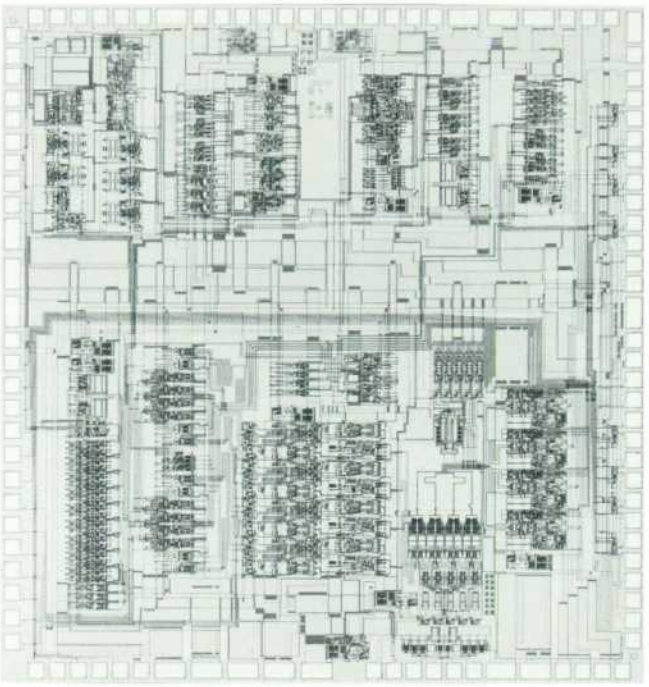


Fig. 16. Sampler IC.

coupling between differential launch paths. The result of these two physical constraints is that the launch between the sampler chip and the first inductor must be treated as a lossy coupled transmission line. The last issue is the settling behavior. If the filter waveform does not settle to within a small part, say 1%, of the peak value within 2 ns of the peak, then intersymbol interference occurs. This is a condition where the current sample being digitized in the ADC is a combination of the current input signal and some residual of the previous samples. The design goal of the IF filter design was to have an overdamped waveform settling to within 1% of the full-scale peak value within 2 ns. This required close attention to the modeling of the ceramic and IC components and additional termination within the filter. The impedance of the launch transmission line is lower at high frequencies because of the capacitive elements. This is compensated before the first spiral inductor with a transmission zero to emulate an ideal inductive launch more closely. The network used to model and implement the filter design is

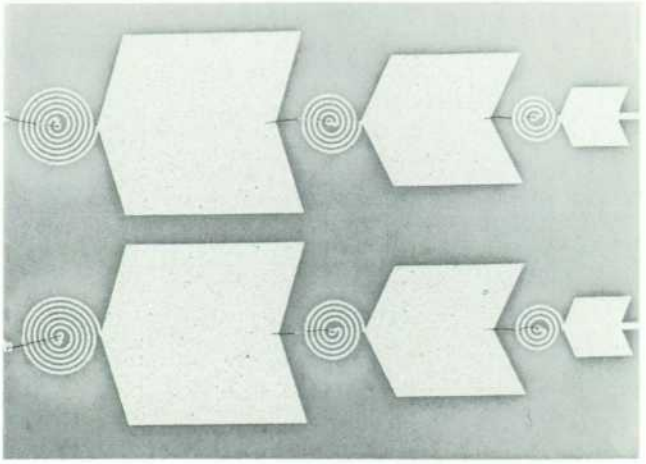


Fig. 17. IF filter.

shown in Fig. 18. As can be seen, it is more complicated than the simplified model of an ideal sixth-order differential Gaussian filter. This network is necessary to guarantee repeatable attenuation, delay, and settling behavior in the filter section of the signal path.

Variable Memory Depth

As stated earlier, a key goal for the HP 54720/10 acquisition system was the ability to scale the memory depth and sample rate based on the number of ADC channels available and the user's needs. This section and the next describe the design approaches taken to provide the variable memory depth and variable sample rate schemes. Also described is the design optimization required to realize high effective bits at very high sample rates.

Variable memory depth is provided by chaining the FISO memory chips in acquisition time. We will use one oscilloscope channel to illustrate the principle. This approach can be extended to two, three, or all four oscilloscope channels to achieve even greater memory depth at the expense of channel independence. Fig. 8a shows the timing within one ADC hybrid when all four ADC paths are sampled at the same time at 500 MSa/s. With this timing system and identical reference voltages for the four flash ADCs, all four ADC paths present the same digitized data to the four separate

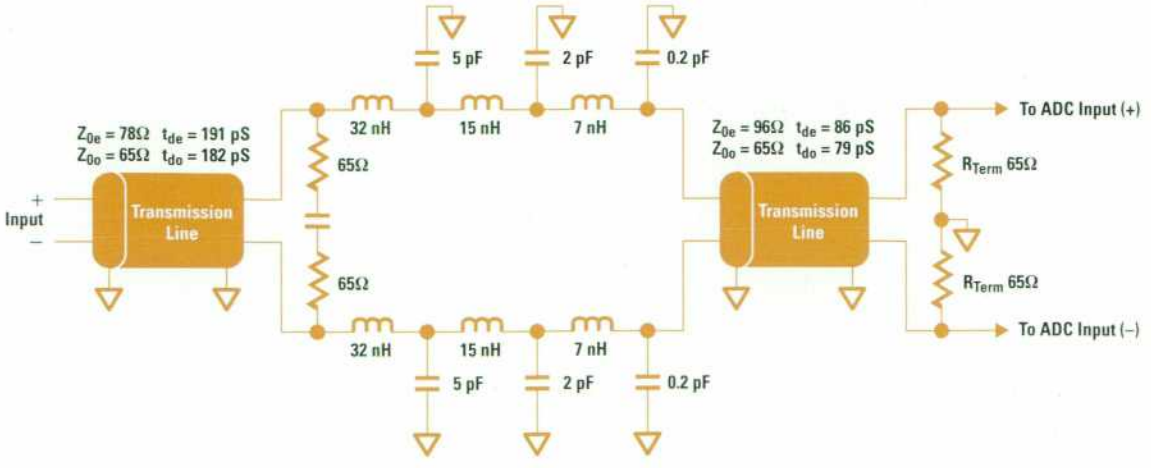


Fig. 18. IF filter model showing added termination and the complexities of hybrid integration.

memory blocks on the FISO chip at the same time. There is a total of 64K of memory on one hybrid in the HP 54720. Each of the four FISO banks is capable of storing 32 μs of acquisition time at 500 MSa/s for a total of 128 μs if all four banks are in series. The design of the FISO chips includes a chaining capability. For the first 32 μs, the first memory bank is written to. From 32 to 64 μs, the second FISO bank is written to. From 64 to 96 μs, the third bank is written to and from 96 to 128 μs, the fourth memory bank is written to. Since each of the four FISO banks receives the same data at the same time, the effect of this chain is to extend acquisition memory.

Fig. 19 shows the schematic and timing diagrams that realize this chaining of memory blocks. There are two FISO chips and dual data paths within each chip. The four 16K memory blocks are labeled A, B, C, and D. Two signals must be asserted before data can be written to a memory block, indicated by the AND gate in the WRITE control path in Fig. 18. At the start of acquisition, the RANK signal going into each FISO chip will determine the WRITE acquisition order. If this signal is high, the chaining flip-flops are PRESET to high, and if RANK is low, the flip-flops are RESET to low. The acquisition counter is PRESET to enable the first memory bank, labeled A, to be written to. The period of the acquisition counter is set to 32 μs so that only bank A is written to for the first acquisition period. After 32 μs, bank B is written to because both the counter and the chaining flip-flop outputs are high. Meanwhile, the PRESET from bank B's chaining flip-flop has propagated to the output of the chaining flop-flop for bank C. This signal propagates at a rate of 32 μs per bank, allowing large margins on memory system setup time. Thus, during the third 32-μs period, bank C receives an asserted write enable. During the fourth 32-μs period, bank D follows.

In this way, as shown by the timing diagram in Fig. 19, a circulation of write enables occurs within the memory system, making extendable memory depth possible. As stated earlier, this example is for one hybrid (oscilloscope channel) but can be extended up to four channels with the appropriate plug-in and user settings. It is also possible to increase the sample rate to 1 GSa/s within one channel and store 64 μs of acquisition time in memory. This is accomplished with a different programming of the chained write enable control path. In this case, banks A and C are written to simultaneously; then, after 32 μs, banks B and D are written to. The row-select signals point at the specific memory elements to be written to. Thus the FISO memory system is a circular memory within and between memory banks.

Variable Sample Rate

Variable sample rate capability is provided through two separate but related schemes. First, on the sampler IC within each hybrid there are divider circuits, which can produce any of the three timing diagrams shown in Fig. 8 under microprocessor control. The exact timing and jitter between these signals are critical to retaining high effective bits at high slew rates. 8 ps of time skew or 4 ps of rms jitter on one of the four sample signals shown in Fig. 8c, for example, will limit the performance of the acquisition system to 6 effective bits on a 500-MHz full-scale sine wave.[3] High-speed differential circuits are used within the sampler IC to minimize coupled jitter and time skew. Fast clock edges within the signal path reduce this effect as well, since the transistors spend less time in the critical switching regions. Systematic time skews caused by process, temperature, or voltage variations within the chip are calibrated out with variable delay circuits that provide ±30 ps of programmable delay. These techniques eliminate skew and jitter, achieving the design goals.

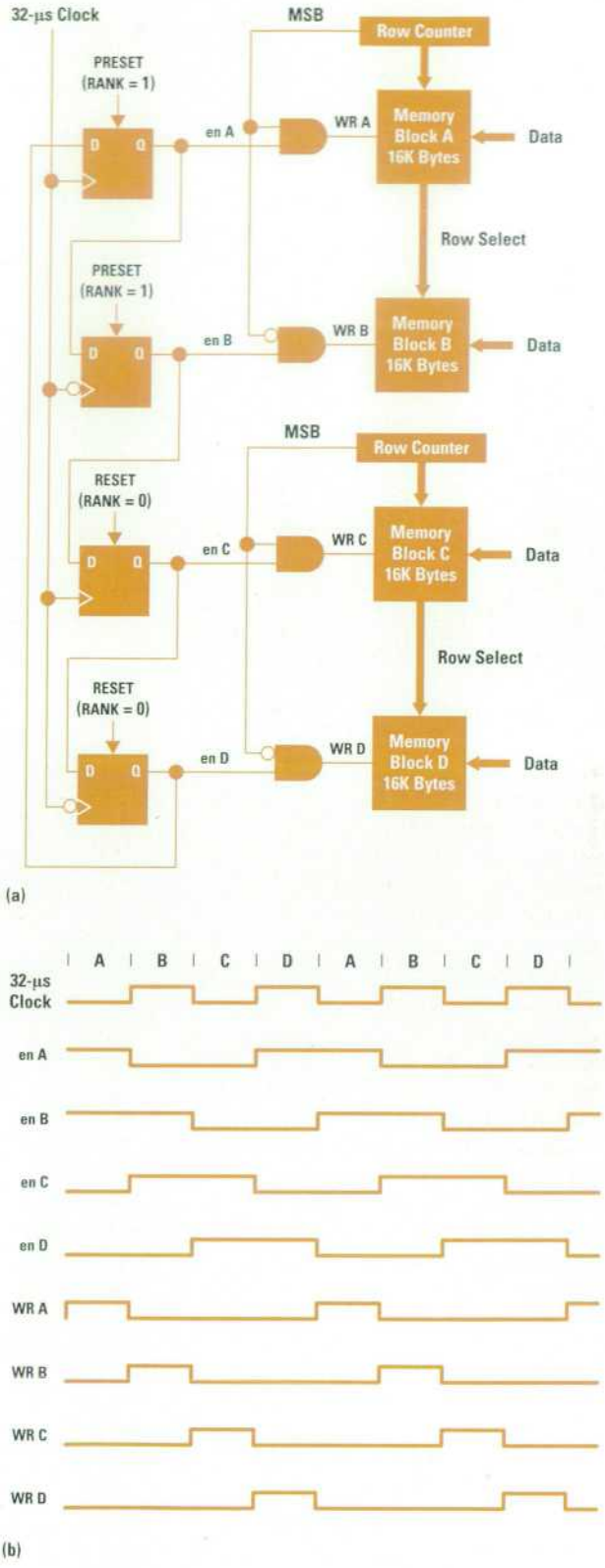


Fig. 19. (a) FISO memory write control schematic. (b) Timing diagram for chaining acquisition memory.

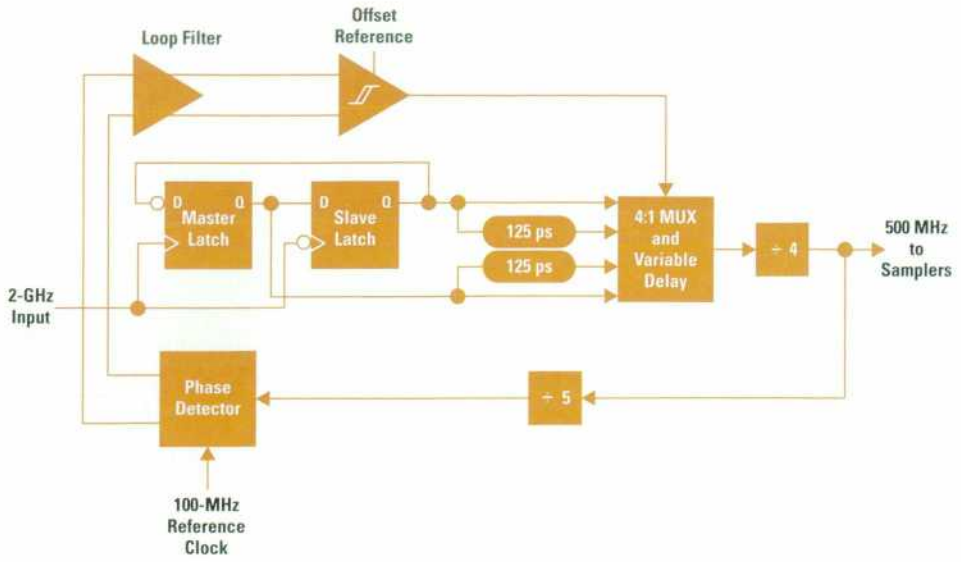


Fig. 20. Delay-locked loop schematic diagram.

The second scheme to achieve variable sample rates involves the timing of the sample clocks between oscilloscope channels. For example, in the 8 GSa/s mode supported by the HP 54722A four-slot plug-in, all four oscilloscope channels are used, with 125-ps clock timing offset between the four 2-GSa/s (500-ps) hybrids. In this case, 4 ps of time skew or 2 ps of rms jitter can be allowed between the channels to achieve 5 effective bits with a 2-GHz full-scale sine wave input.

The clock timing offset is generated by phase-locking an incoming 2-GHz oscillator to a 100-MHz reference through a delay-locked loop in such a way that different delay amounts can be added to the delay-locked loop to change the sample point offset. The schematic diagram of the circuit is shown in Fig. 20. A delay-locked loop is like a phase-locked loop in that it has a zero in the forward path, which forces the incoming signal, in the this case the 2-GHz sampler oscillator, to be phase-locked to the reference signal, the 100-MHz master clock. The difference between the two kinds of loops is that frequency lock is not required in the delay-locked loop, since the 2-GHz and 100-MHz clocks are already harmonically related and are not subject to long-term relative frequency drift. They are subject, however, to short-term phase drift, which can destroy the critical timing between sampling channels. The delay-locked loop prevents this.

Another point worth making here is the reason for using the 100-MHz clock as the reference rather than the 2-GHz clock. After all, the 2-GHz clock is the clock that all the samplers see. The problem is how to distribute the 2-GHz clock to the printed circuit boards without the possibility of phase-locking to the wrong cycle. Signals travel over semirigid coaxial cable at approximately 150 ps/inch. The physical distances required for distribution of the HP 54720/10 clock among the acquisition boards made it impossible to retain cycle coherency with a 500-ps clock. However, with a 10-ns clock, this was easily accomplished.

Referring now to Fig. 20, the incoming 2-GHz oscillator clock enters the sampler IC and passes through variable-delay circuits with 125-ps steps, then is divided by four before being sent to the samplers in four phases of 500 MHz and to the phase detector after division by five. The desired sample point offset relative to the 100-MHz reference is programmed into the offset of the error integrator. For example, if the system calls for no offset between the 100 MHz reference and the sample point, then none is entered into the offset reference. If 500 ps is desired, then it is entered into the comparator as an offset voltage and the loop forces this condition simply by adding another delay element in the path of the 2-GHz clock. Delays smaller than 500 ps are generated with a combination of coarse delay in the master-slave flip-flop and 125-ps gate delays through smaller offsets into the comparator. The gain of the feedback loop is large enough to reduce the timing uncertainty between the 100-MHz reference and the 100-MHz signal obtained by dividing the 2-GHz clock to a few picoseconds which is then calibrated out. The feedback loop is fully differential to reduce coupled sources of offset.

The remaining design goal was to reduce the drift of the sample point relative to the 100-MHz reference to a minimum so that under varying temperatures and supply voltages, the sample point would not move differently on different hybrids. This is ensured by making the gate delays from the divide-by-four circuit output through the divide-by-five circuit

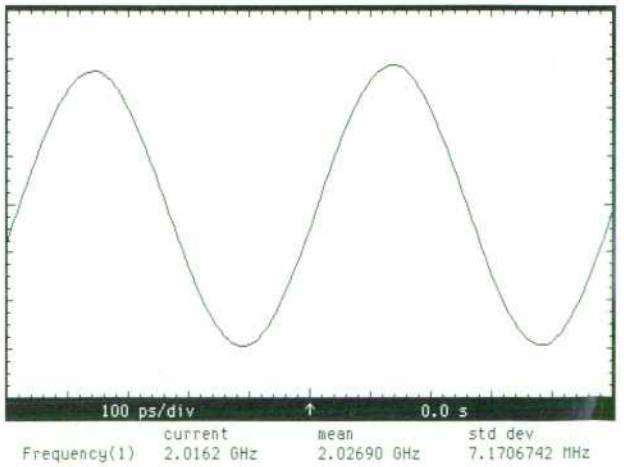


Fig. 21. 2-GHz reconstructed sine wave from an 8-GSa/s HP 54720 system.

Fig. 22. Step response of an 8-GSa/s HP 54720 system.

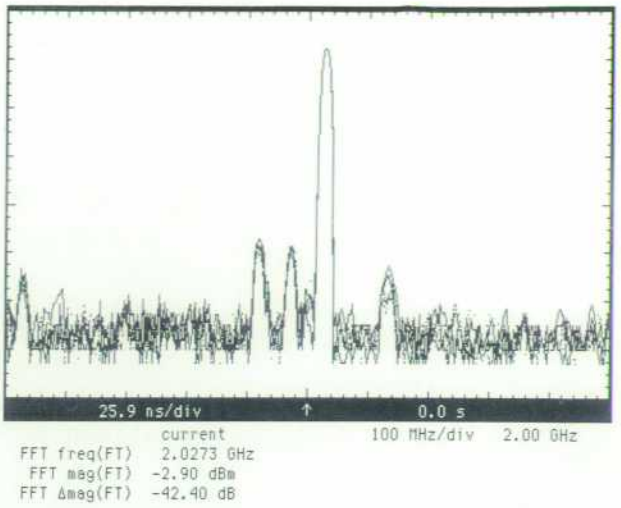


Fig. 24. FFT of the 2-GHz reconstructed sine wave of Fig. 20, showing sidebands down 42 dB from the fundamental.

equal to the gate delays from the divide-by-four circuit output to the actual sample point. Since gate delays tend to drift with the same temperature coefficient, equal delays will drift the same amount and there will be no differential drift.

Measured System Performance

Fig. 21 shows a reconstructed 2-GHz sine wave that has been digitized at 8 GSa/s. Fig. 22 shows a step response in the same configuration, showing 6% overshoot and 211-ps rise time, equivalent to 2-GHz real-time bandwidth. Fig. 23 shows effective bits for the 8-GSa/s case and for the HP 54721A plug-in at 4 GSa/s. Both cases show effective bits better than five over the entire input frequency range.

A key measure of ADC systems is the harmonic distortion caused by the nonlinearities in the amplifying and sampling circuits. These nonlinearities appear in the frequency domain as intermodulation products surrounding the fundamental. Fig. 24 is the FFT of the 2-GHz reconstructed sine wave of Fig. 21. The sidebands are down 42 dB from the fundamental. The energy in these sidebands represents the major distortion mechanisms that contribute to loss of effective bits at high frequencies.

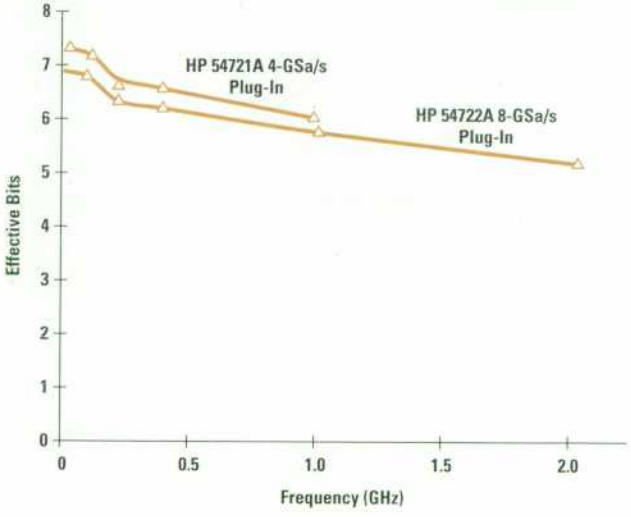


Fig. 23. Effective bits at 8 GSa/s and 4 GSa/s for an HP 54720 system with an HP 54721A plug-in.

Acknowledgments

A plug-in architecture greatly complicates the design, calibration, and performance verification of an oscilloscope. Major contributions in these areas plus specific plug-in designs were made by Marshall Boss, Kevin Loftin, and Derek Toeppen. Additional contributions in hybrid development for plug-ins were made by Jim Raney and Phil Yearsley. Major contributions to the acquisition system were made by Allen Montijo, Ken Rush, Ed Evel, Dave Dascher, Buddy Yount, Brian Gartner, Martin Guth, and Dale Walz of Hewlett-Packard Colorado Springs Division and by Chris Schiller, Bruce Domen, Bart Jansen, Bill Hillery, and Greg Smith of the Hewlett-Packard Circuit Technology Group. Joe Millard, Tracy Ireland and their teams are responsible for the bipolar process development and control.

References

1. B. Peetz, B. Hamilton, and J. Kang, "An 8-bit 250-Megasample-per-Second Analog-to-Digital Converter: Operation without a Sample-and-Hold," *IEEE Journal of Solid-State Circuits*, Vol. SC-21, no. 6, December 1986, pp. 997-1002.

2. C. Schiller and P. Byrne, "A 4-GHz 8-bit ADC System," *IEEE Journal of Solid-State Circuits*, Vol. SC-26, no. 12, December 1991, pp. 1781-1789.

3. J.J. Corcoran, K. Poulton, and K.L. Knudsen, "A One-Gigasample-per-Second Analog-to-Digital Converter," *Hewlett-Packard Journal*, Vol. 39, no. 3, June 1988, pp. 59-66.

A Digitizing Oscilloscope Time Base and Trigger System Optimized for Throughput and Low Jitter

Careful attention to low-noise coupling results in robust performance far exceeding what is normally considered possible with off-the-shelf ECL. A new interpolator design increases resolution by a factor of ten, reduces conversion time by a factor of five, and reduces jitter by a factor of more than three compared with previous designs.

by David D. Eskeldson, Reginald Kellum, and Donald A. Whiteman

The time base and trigger system of the HP 54720/10 digitizing oscilloscope has three major tasks. It controls the start and stop of the acquisition cycle, it recognizes a trigger event during an acquisition, and it measures the time from the trigger event to the next sample clock.

Fig. 1 is a block diagram of the time base and trigger system. The time base control circuit starts and stops the acquisition synchronously with the sample clock. It also keeps track of the number of pretrigger and post-trigger events that have occurred. Before the time base and trigger system can look for a trigger event, a certain number of pretrigger samples must be acquired. Similarly, a certain number of post-trigger samples must be acquired after the trigger event. Both the pretrigger and post-trigger criteria are determined by the user's setup of such things as the sweep speed and position, the sample rate, the record length, and the position of the record relative to the trigger event.

The trigger circuit detects a trigger event defined by the user (edge, pattern, glitch, etc.). When enough pretrigger samples have been acquired (determined by the time base control circuit), the trigger circuit will search for the next trigger event. When a trigger event is found, the system trigger (ATRIG) is generated and sent to the interpolator circuit.

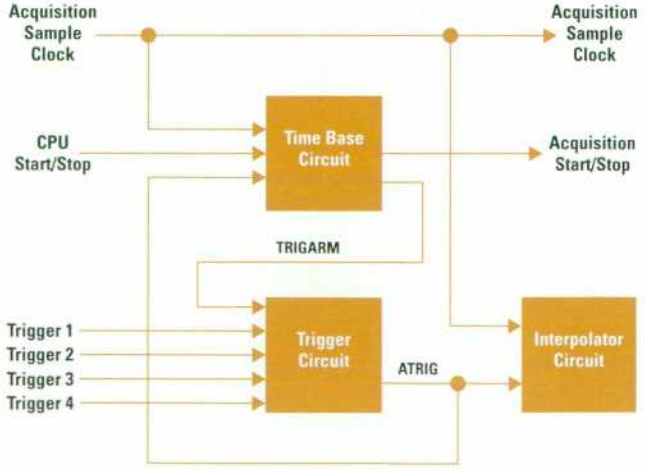


Fig. 1. Block diagram of the time base and trigger system.

The interpolator circuit is used to determine the correct placement of the samples acquired. The interpolator circuit measures the time from the system trigger to the next sample clock.

Fig. 2 shows a timing diagram for a typical acquisition cycle in the HP 54720/10 oscilloscope. It begins when the time base and trigger system receives a start signal from the CPU. The time base control circuit synchronizes the CPU start signal. The synchronized start signal enables the rest of the system and starts the pretrigger counter in the time base control circuit. When the pretrigger count is finished, the time base control circuit arms the trigger circuit to look for the next valid trigger event. When the trigger circuit recognizes a valid trigger event, it generates the system asynchronous trigger (ATRIG). ATRIG starts the post-trigger counter in the time base control circuit. It also starts the interpolator. The interpolator measures the time from ATRIG to a known edge of the sample clock. When the post-trigger counter finishes its count, the acquisition is complete.

Sources of Measurement Error

The measure of goodness for a time base and trigger system is its ability to make accurate, reliable delta-time measurements. The performance goal for the HP 54720/10 time base and trigger system was a maximum of 30 picoseconds of delta-time error. Most of this error comes from contaminated interpolator measurements, where noise coupling, in the form of either jitter or nonlinearity, has caused the trigger event to be misplaced by some amount of time.

Time base jitter is mainly caused by noise in the interpolator measurement. This noise could be introduced in the trigger circuit, in the interpolator circuit, or in the path from the trigger circuit to the interpolator circuit. In equivalent time sample mode, time base jitter can become a significant part of the error in a delta-time measurement. If the waveform is not averaged, the delta-time measurement can be distorted

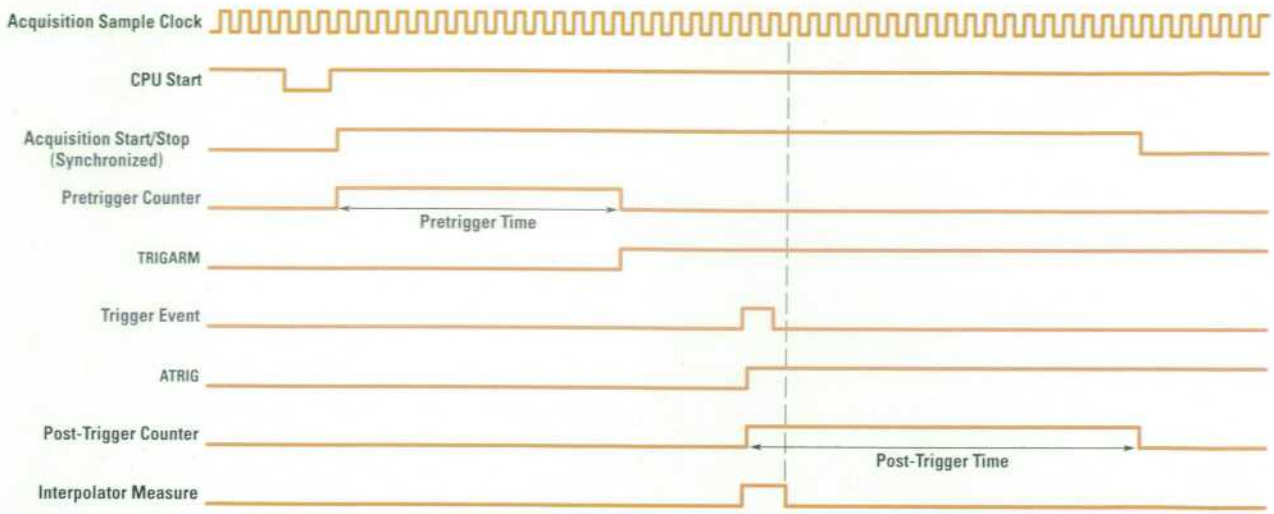


Fig. 2. Acquisition timing diagram.

by as much as the amount of peak-to-peak jitter. This degrades the repeatability of the measurements.

If averaging is used, repeatability is increased, but the averaging of a jittery waveform acts as a filter applied to edges. This filter can affect the system bandwidth. If the jitter is Gaussian, the effective filter bandwidth can be described by the equation: BW = 0.132/(rms jitter). For example, with averaging on, a system that has 20 ps of rms jitter has an effective filter bandwidth of 6.6 GHz.

Time base nonlinearities cause outright errors in time-interval measurements. A time base nonlinearity occurs when some noise source always causes the interpolator to mismeasure a particular time. Usually, nonlinearities are continuous and therefore not so easy to distinguish. Unless they are very bad, they distort the waveform by compressing and expanding it by very small amounts.

Causes of Jitter and Nonlinearities

The biggest challenge for designing the time base and trigger system became establishing a low-noise environment for the path from the trigger circuit to the interpolator ramp circuit. Both the trigger circuit and the interpolator ramp circuit are on thick-film hybrids to help isolate them from noise. However, the rest of the support circuitry from the trigger circuit to the interpolator ramp circuit is on the printed circuit board.

The path from the trigger circuit to the interpolator ramp circuit is implemented using ECL flip-flops. The ECL 10KH family was chosen because of its fast edge speeds (1 ns typical), its stable threshold voltages, and the ability to terminate outputs in a low impedance so that transmission lines can be used. The general philosophy was to propagate ATRIG from the trigger circuit to the interpolator ramp circuit with edges as fast as possible to prevent noise from coupling onto the edges.

Even with 1-ns edge speeds, the signals are still susceptible to noise. Small amounts of noise coupling onto a signal can significantly alter the timing of that signal. For example, a 10-mV spike riding on the baseline of ATRIG can shift the timing of the transition by 10 ps if the transition and the spike occur simultaneously (see Fig. 3). To achieve <30 ps error in delta-time measurements, any noise source that produced >3-mV spikes (roughly 3 ps error) was considered a problem to be dealt with.

The main thing to keep in mind when trying to reduce coupling in ECL circuits is that the ECL 10H131 flip-flops used in this design have single-ended inputs. When an input transitions from low to high (or high to low), the input transistor is abruptly turned on (or off). This causes an abrupt change in the current flowing from ground to power. The return path for the current is through the inductive ground and power leads. Since V = Ldi/dt, this change in current leads to ground and power bounce. Since the thresholds for the single-ended inputs are referenced to power and ground, they too will bounce.

Another thing to consider is the energy generated by fast edges. Even though ECL 10KH is rated for roughly 100-MHz toggle rates, the 1-ns edges generated by ECL 10KH contain plenty of 300-MHz and 500-MHz energy. While ECL 10KH doesn't have bandwidth to toggle at 300 MHz, it can pass some of the energy.

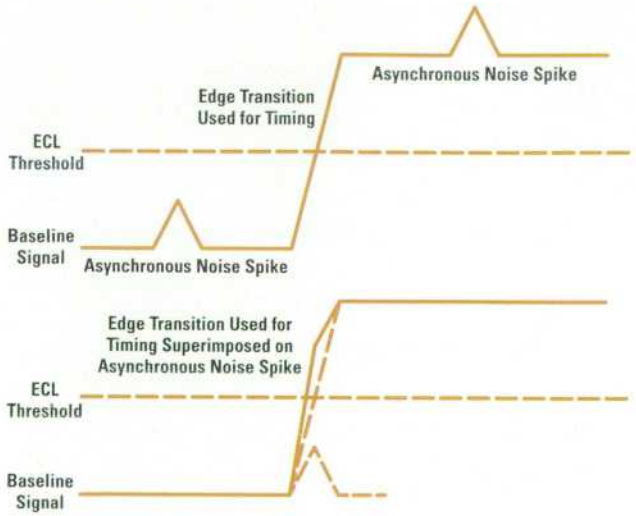


Fig. 3. Effects of noise spikes riding on a baseline signal.

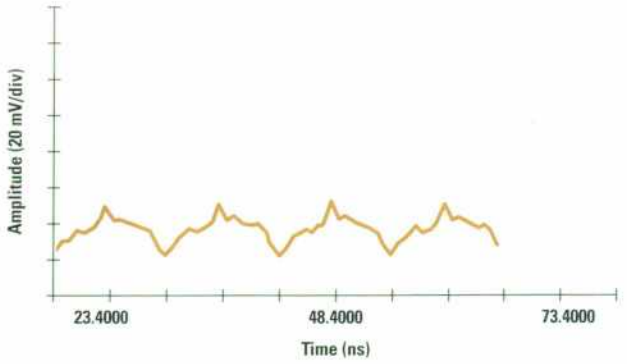


Fig. 4. Flip-flop B output, showing 100-MHz clock coupling from the input of flip-flop A.

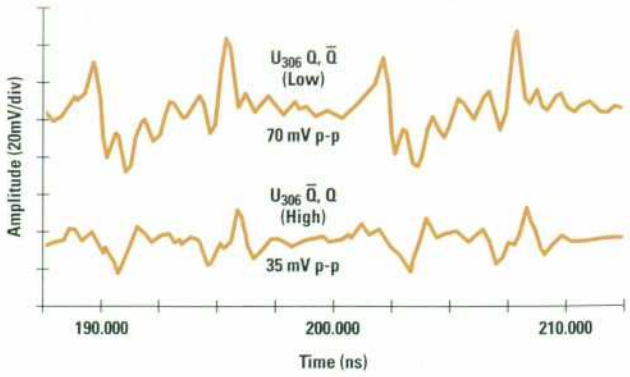


Fig. 5. Clock-to-output coupling on nonswitching outputs is higher on low outputs (top trace) than on high outputs.

Finally, the substrate capacitance of an ECL 10KH IC is enough to allow small amounts of high-frequency energy to couple in from one input, through the package, and back out onto another input.

Achieving Low-Noise Coupling

It is important to note that the performance desired from the discrete components used in this design far exceeds the performance specified by the manufacturer. Trying to achieve picosecond timing on nanosecond edges puts the design into a hypersensitive class. Nearly immeasurable amounts of noise cause problems in such a system. Through the use of several different techniques, the design achieves the desired level of performance. These techniques allow the use of cheaper off-the-shelf ECL components instead of an expensive custom IC.

The first technique is never to combine two critical functions within one IC package. The switching of one gate within a package will cause all thresholds within the package to bounce, which can influence the timing of other gates within the same package. This source of coupling was measured on a pair of ECL 10H131 flip-flops by running 100 MHz into the clock input for one flip-flop and measuring the synchronous noise on the output of the other flip-flop. Typically, this type of coupling produces 20 to 50 mV of noise (see Fig. 4).

The solution to this type of coupling is to isolate functions with more packages. For example, if function A is not related to function B, put function A in a different package from function B. Since this tends to lead to wasted gates (and power and board space), this solution is reserved only for critical signal paths where a 10-ps timing shift would degrade the desired system performance.

The second technique is to use the high output of a flip-flop and invert to get the low output. The clock-to-Q-output coupling of nonswitching outputs was measured to be five times higher on low outputs than on high outputs (see Fig. 5). This is unfortunate, since ECL flip-flops are clocked on transitions from low to high. To generate cleaner baselines for low signals, the solution is to use an inverting buffer on the high output to achieve a cleaner low signal.

The third technique is to double-buffer signals to increase reverse isolation. High-frequency energy was being generated by the clock divider circuits. Because of the substrate capacitance of the IC, small amounts of high-frequency energy would couple back through the package and cause subharmonic distortion of the main interpolator clock. Therefore, certain clock signals are double-buffered using an additional package to provide more reverse isolation.

The fourth technique is to turn circuitry off when it is not in use. Some of the circuitry spends most of its time unused. For instance, only one sample clock is used during any particular acquisition. Therefore, the signals corresponding to other clock rates don't need to be switching. By disabling these other clock outputs, another potential source of noise is eliminated.

The fifth technique is to filter inputs where edge speed isn't important. For these noncritical input signals, a small filter is added to the input using the series resistance and the input

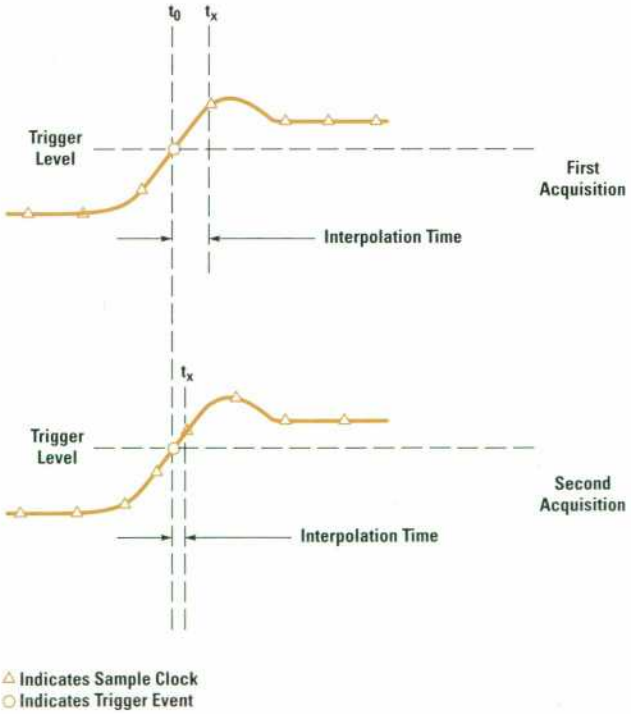


Fig. 6. The interpolator measures the time between the trigger (t_0) and the next sample clock (t_x).

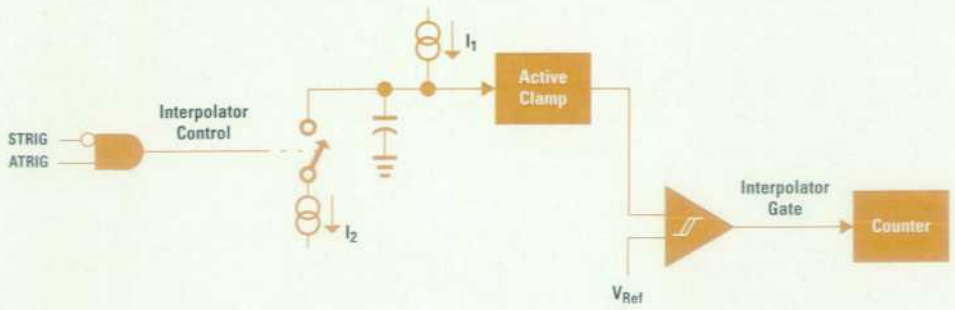


Fig. 7. Functional block diagram of the dual-slope trigger interpolator.

capacitance of the IC; this slows down the edge speed and reduces higher-frequency harmonics.

Trigger Interpolator

In real-time and equivalent time sampling systems, the sample clock is asynchronous with the input signal and the trigger event. It is the task of the trigger interpolator to measure the time interval between the trigger event at time t_0 and the next sample clock at time t_x (see Fig. 6). This measurement facilitates the accurate placement of data on the display screen. Because of the asynchronous relationship between the trigger event and the sample clock, the interpolation interval varies from acquisition to acquisition. Fig. 6 illustrates this characteristic.

In previous designs the tool for making this measurement has been the dual-slope ADC (analog-to-digital converter) or dual-slope interpolator (Fig. 7). This interpolator measures time by measuring the voltage on a capacitor. In the timing diagram of Fig. 8:

- At t_0, a trigger event occurs (ATRIG).
- From t_0 to STRIG (synchronous trigger), which is the second interpolator clock after t_0, the switch for I_2 in Fig. 7 is closed by the interpolator control signal. As a result, the capacitor is charged to a negative voltage proportional to the duration of this interval. Because of the asynchronous relationship between the trigger event and the interpolator clock, the duration of this interval will vary from acquisition to acquisition, and hence the charge on the capacitor will vary.

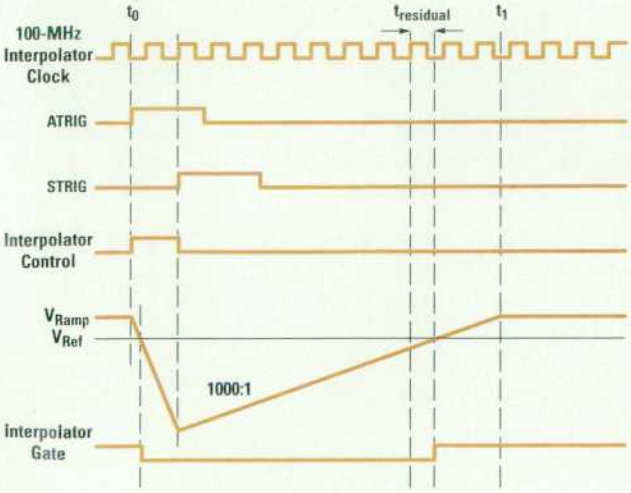


Fig. 8. Waveforms in the dual-slope interpolator.

- At STRIG, the interpolator control signal goes low, causing the switch for I_2 to open, and the capacitor is charged by I_1 to its original state. In Fig. 8 this occurs at t_1. This original state or steady-state condition is controlled by the active clamp. If the active clamp were not present I_1 would charge the capacitor up to the supply voltage of the circuit.

During this charge and discharge period the 100-MHz interpolator clock is counted. The counters are enabled when the interpolator gate signal goes low. This signal comes from the output of the comparator, which switches around the reference voltage V_{Ref}. The reference ensures that the count takes place on the linear portions of the ramp. From this count the interpolation interval is derived (the interpolator clock is synchronous with the sample clock). In our 10-ps system, I_1 is 1000 times smaller than I_2, so the time required to discharge the capacitor is 1000 times longer than the time it takes to charge it. It is this characteristic of the dual-slope interpolator that yields its resolution. Effectively, the interpolator magnifies the interpolation or uncertainty interval by the ratio of the charge and discharge currents. In the 10-ps system this magnification is 1000:1. The resolution of the system is the period of the interpolator clock divided by the ratio of the charge and discharge currents. In this system a 100-MHz interpolator clock is used, yielding an interpolation resolution of 10 ps (10 ns/1000 = 10 ps). It is important to note that there is some error in the conversion, which is denoted $t_{residual}$ in Fig. 8. This error is a result of the fact that t_1 does not occur on a rising edge of the interpolator clock, that is, t_1 is not synchronous with the interpolator clock. Therefore, the measurement is off by some fraction of one count. However, with a stretch ratio of 1000:1 and an interpolation interval of 1 to 2 interpolator clock periods, this error is small, approximately 1 count/1500 counts = 0.066%.

As customers asked for faster and more accurate measurements, the need for a faster and more accurate time base system became apparent. To achieve 1-ps resolution with the old system, two approaches are possible. One approach would be to maintain a 100-MHz interpolator clock and increase the stretch ratio to 10,000:1. This would leave the ramp so flat that noise would make accurate measurements impossible. In addition, the conversion time would be approximately 150 μs, which would make the display update rate unbearably slow. Another approach would be to increase the rate of the interpolator clock. Unfortunately, most commercially available flip-flops will only accept clocks up to 250 MHz. This still leaves the ramp ratio at 4000:1 and the conversion time at 60 μs (1.5 clock periods ×

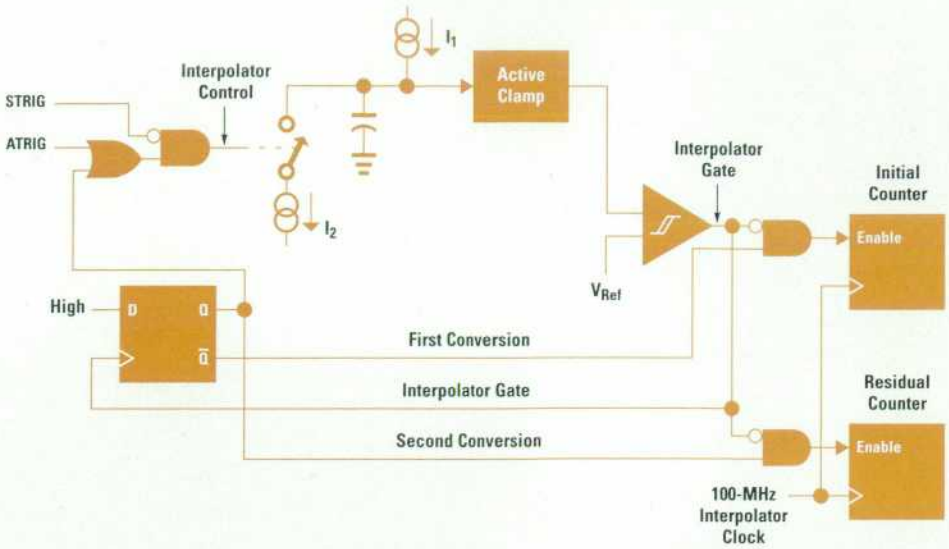


Fig. 9. Functional block diagram of the residual dual-slope trigger interpolator.

4000). Since neither of these solutions was acceptable, a new approach had to be taken.

Residual Interpolator Technique

The residual dual-slope interpolator, Fig. 9, overcomes the limitations of the old dual-slope interpolator. It functions much the same way as the dual-slope interpolator. The difference is that the residual interpolator performs a double conversion, as opposed to the single conversion performed by the previous system.

The first conversion is exactly like that of the dual-slope system, in that it interpolates the period between the trigger event and the second interpolator clock after the trigger event. The second conversion converts $t_{residual}$, which is the error left from the first conversion (see Fig. 10). Interpolation of the error maximizes the resolution. The advantage of this system is that the effective stretch ratio is now the product of the ratios from the two conversions. Therefore, to get an effective stretch ratio of 10,000:1, which would yield 1-ps resolution with a 100-MHz interpolator clock, each conversion need only have a stretch ratio of 100:1. In addition, the conversion time corresponds to a stretch ratio of 200:1, which is the sum of the two 100:1 conversions, rather than 10,000:1.

The residual dual-slope interpolator makes a threefold contribution over the 10-ps system. First, the interpolation resolution sees a 10:1 improvement, going from 10 ps to 1 ps. Second, the interpolation conversion time is improved by a factor of 5, from 15 μs to 3 μs. The significance of this contribution is that it affords faster data throughput from input to screen, and hence gives the oscilloscope's display more of an analog response. Third, the interpolator jitter is reduced by more than a factor of three, from approximately 6 ps to 1.8 ps rms. The improved jitter performance comes as a result of four primary design improvements:

- Reduced stretch ratios equate to increased ramp speeds, which equate to lower ramp jitter.
- A minimal number of gates between the trigger event and the ramp circuit yields minimal jitter, since each gate adds jitter to the signal being propagated.
- Maximized edge speeds of these same gates, obtained by using the fastest commercially available logic, equate to lower jitter because fast edges are less susceptible to noise.
- The interpolator ramp circuit is powered with an isolated supply. Thus the power supply is free from system noise and less likely to induce jitter.

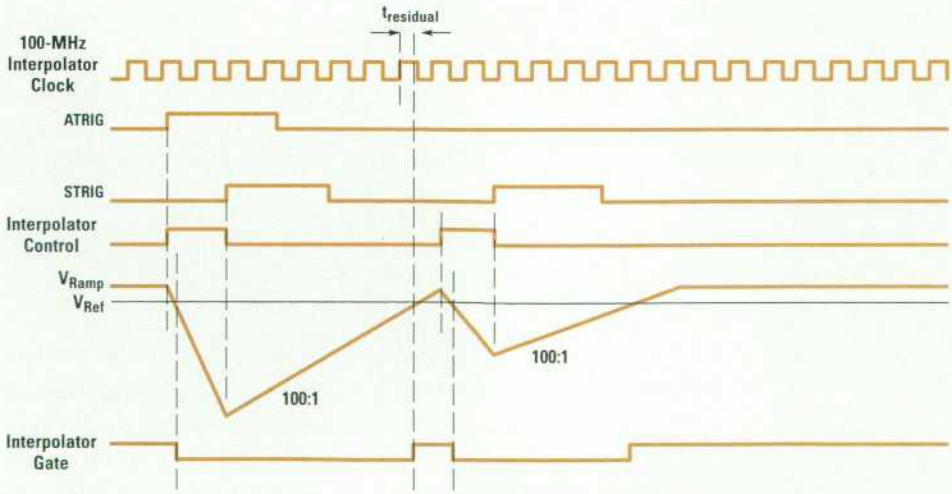


Fig. 10. Waveforms in the residual dual-slope interpolator.

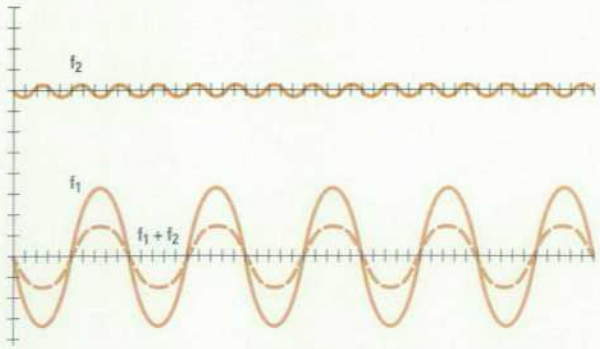


Fig. 11. Harmonics on a clock signal do not shift the position of one sample relative to the next.

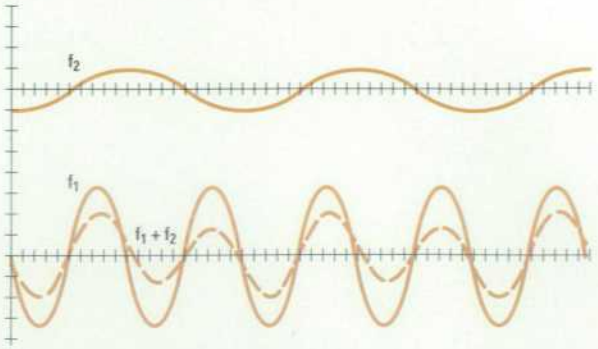


Fig. 12. Subharmonics are much more important than harmonics because they can cause periodic shifts in the sample positions.

Sample and Reference Clocks

Digital oscilloscopes, unlike their analog counterparts, require a precision clock to define the time at which the samples are taken. The HP 54720/10 oscilloscope requires two timing signals: the 2-GHz sample clock and the 100-MHz reference clock. These two clocks must be phase-locked together and distributed with a minimum of differential phase changes over the operating temperature of the instrument.

The time measurement accuracy of the HP 54720/10 is limited by the accuracy of the sample and reference clocks. There are three components of clock accuracy:

- Long-term frequency accuracy
- Periodic distortions
- Phase noise.

Long-Term Frequency Accuracy. The ability of the instrument to measure long time delays and make precision time measurements is defined by the long-term accuracy of the clock as measured by a precision counter.

Periodic Distortions. Periodic distortions consist of higher-order harmonics and lower-order subharmonics. These two types of harmonics have significantly different effects on clock accuracy.

The 2-GHz clock is a sine wave with small levels of harmonic distortion. Higher-order harmonics shift the position of the crossovers relative to the fundamental, but do not shift the position of one sample relative to the next (see Fig. 11). An extreme example of this would be a narrow pulse train, rich in harmonics, but with a constant period from one pulse to the next. The primary problems that can occur with higher-order harmonics are a reduced rate of change at the crossovers leading to a less stable sample position and increased radiated or conducted interference within or outside the instrument. The higher-order harmonics are specified to be no more than –30 dB relative to the fundamental, a readily obtainable level with the oscillator types considered.

Subharmonics, or harmonics below the sampler clock frequency, are much more significant because they cause periodic shifts in the sample positions (see Fig. 12). The subharmonics on the 2-GHz sample clock are specified at –50 dB. This corresponds to an amplitude ratio of fundamental to harmonic of 316 to 1 and a sample position error of approximately 0.5 ps. Subharmonics generated within the oscillator are not the only cause for concern. The 100-MHz reference clock is distributed from a high-speed ECL-level gate, and the edges are rich in harmonics that without careful design of the distribution network would be coupled into the sample clock.

Phase Noise. Phase noise, the random error from one sample clock crossover to the next, could cause a reduction in the effective bits of the instrument and would affect the accuracy of long delays. Phase noise is most prevalent in phase-locked oscillator designs.

Three types of oscillators were considered to generate the 2-GHz sample clock and the synchronous 100-MHz reference clock:

- A 100-MHz crystal oscillator generating the reference clock and a 20-to-1 multiplier to generate the 2-GHz clock.
- A 2-GHz STW (surface transverse wave) oscillator with a fundamental frequency of oscillation of 2 GHz. Counters would be used to generate the 100-MHz reference clock.
- A 100-MHz crystal oscillator to generate the reference clock and a phase-locked loop with a 2-GHz oscillator to generate the sample clock.

The crystal oscillator with frequency multiplier is the most stable oscillator at medium cost. A carefully designed multiplier produces a minimum of phase noise. The output stage can be designed as a filter directly driving the distribution network. Vendors can provide this kind of oscillator in a sealed enclosure with maximum shielding. However, the tuning of the multiplier stages can be difficult, leading to unpredictable output levels without careful manufacturing techniques. The production controls required suggest a vendor with this kind of manufacturing experience.

STW oscillators can be designed to have a fundamental oscillation frequency of 2 GHz. No subharmonics are produced. A counter or other technique is required to generate the 100-MHz signal. Stabilities similar to a crystal can be obtained. With integration, this would be a low-cost solution for high-volume production. However, STW oscillators at this frequency are new. Production experience has not yet been gained. Buffering must be provided to limit coupling of the 100-MHz clock into the 2-GHz clock. Phase stability is not as good as a crystal, and the circuit would probably be more complex than the crystal design unless it were integrated.

The phase-locked oscillator provides the stability of the reference oscillator without multipliers. It lends itself to integration, but its phase noise is greater than the crystal oscillator's. Buffering is required to minimize coupling of the

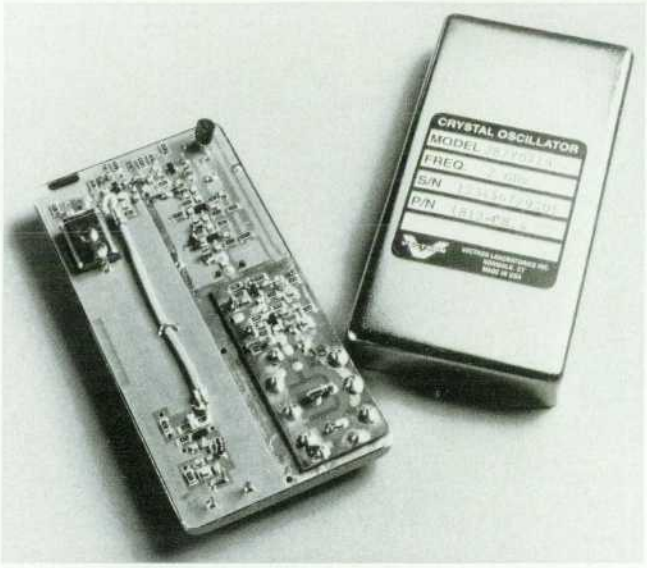


Fig. 13. Clock oscillator.

100-MHz clock into the 2-GHz output. A prescaler is required if off-the-shelf parts are used.

After careful review of the various oscillator systems it was decided to use a crystal oscillator with frequency multiplier. It is a custom oscillator designed by Vectron Laboratories (see Fig. 13). The most important specifications are:

Frequency accuracy	50 ppm
Harmonic distortion	< –30 dBc
Subharmonics	< –50 dBc
2-GHz output level into four parallel 50-ohm loads	1.0V p-p
100-MHz sine wave output into 50 ohms	0.6V p-p

Clock Distribution

For 2-GHz clock distribution, each acquisition hybrid requires a 1V p-p signal from a 50-ohm transmission line. Back termination is not required. The oscillator could have been designed to drive a power splitter or to drive the four lines in parallel. The higher voltages required to drive a power splitter complicated the oscillator output stage. The final output stage drives a transmission line of 12.5 ohms, equivalent to four 50-ohm lines in parallel. The output stage is tuned to compensate for the inductance of the output pin. This eliminates the need for an expensive connector on the oscillator. The stage is matched to drive 12.5 ohms. For the

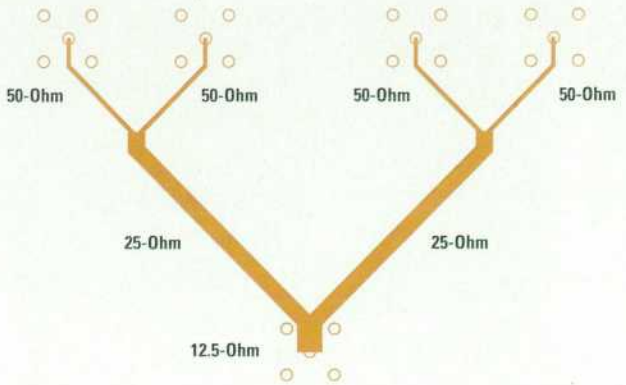


Fig. 14. The sample clock is distributed symmetrically to the four acquisition hybrids on transmission lines that transition from 50 ohms at the hybrids to 25 ohms and then to 12.5 ohms at the oscillator.

signals to each hybrid to have the same amplitude it is necessary to maintain symmetry from the oscillator output to the four 50-ohm lines coupling the sample clock to the hybrid. Fig. 14 shows how this is accomplished. To simplify the layout and reduce the requirements for layout precision the outputs are symmetrically coupled in pairs to the oscillator, making a transition from 50 ohms to 25 ohms and then to 12.5 ohms.

For 100-MHz clock distribution, the hybrids and the time base require 100-MHz ECL signals with transition times on the order of 500 ps. The oscillator output is a sine wave to minimize cross coupling from the 100-MHz signal to the 2-GHz signal. The transition from the oscillator output to the ECL output is made with a high-speed ECL clock driver with outputs paralleled to ensure that temperature drifts are common to all hybrids.

The final circuit configuration of 100-MHz oscillator and 2-GHz multiplier has proved reliable to manufacture. The 2-GHz transmission lines and the cable system for the 2-GHz clock maintain repeatable levels of signal to the hybrids and the 100-MHz clock system has had no difficulties with subharmonics.

Acknowledgment

Special credit is due Steve Draving whose experience and insight led to the excellent time base linearity and low jitter achieved in the HP 54720. Steve's skill at making millivolt and picosecond measurements is a credit to our whole lab.

A Rugged 2.5-GHz Active Oscilloscope Probe

Superior electrical performance is maintained by suspending a fragile electrical structure inside a rugged package and isolating the fragile parts from external abuse. The design required numerous trade-offs between performance, durability, aesthetics, and cost, with performance and ruggedness the primary goals.

by Thomas F. Uhling and John R. Sterner

Higher bandwidths along with more accurate responses have been the trend for new digital oscilloscopes. A necessity with this higher level of performance is a probing system that can accurately deliver signals from the device under test (DUT) to the oscilloscope. The probe must also maintain high input impedance into the gigahertz range for low loading of the DUT. These requirements set the stage for the design of the HP 54701A active probe.

High Bandwidth

It makes no sense to purchase an expensive high-performance oscilloscope like the HP 54720/10 and then lose that performance by putting a lower-performance probe in front of it. Ideally, the probe should be transparent to the oscilloscope system, thus maintaining the full performance level to the probe tip.

To obtain this kind of performance with the HP 54701A, emphasis was put on maintaining the maximum possible bandwidth in each stage of the amplifier. The impedance buffer stage or first stage was designed using n-p-n bipolar emitter followers as shown in Fig. 1. The input of the amplifier is ac coupled so the base of Q1 can be biased at a negative voltage. This allows the collectors of Q1 through Q3 to be grounded, eliminating complicated bypassing of the collectors and ensuring stability. Q1 and Q2 are self-biased with resistors R_{B1} and R_{B2}. These resistors also eliminate the need for bypassing the power supplies, thereby conserving real estate and ensuring stability. To achieve the desired response, resistors are placed in the collectors of Q1 and Q3 for damping. This approach proved to work better than series base damping and also produces a lower noise floor for the amplifier. The output stage consists of a common-base amplifier, which

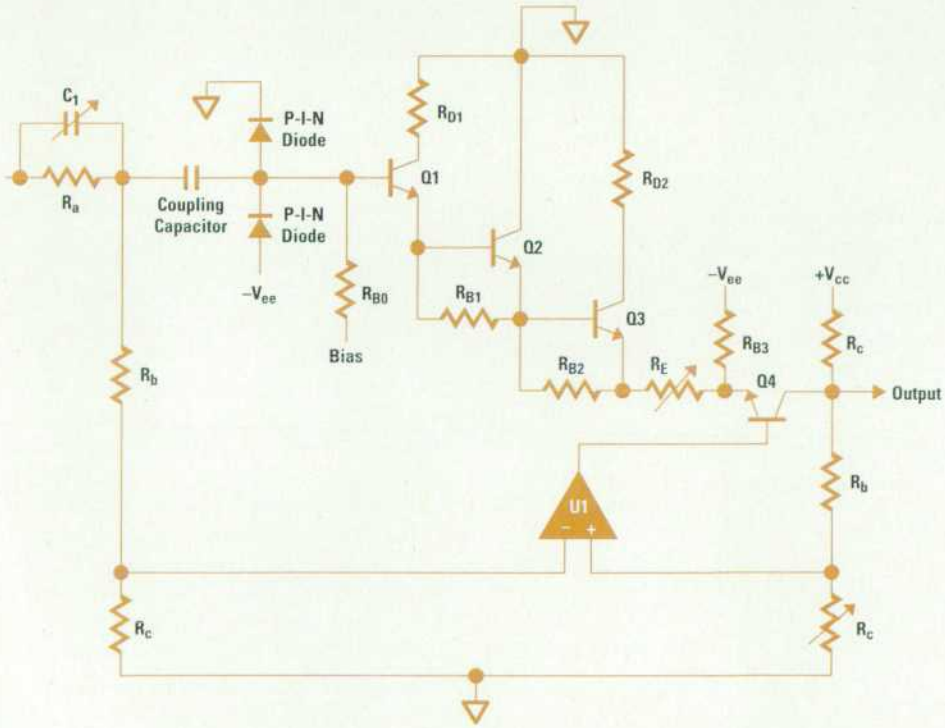


Fig. 1. Simplified schematic diagram of the HP 54701A active probe.

maximizes the bandwidth and shifts the signal level back to 0.0 volts dc. In both the input and the output stages the bandwidth is optimized, yielding a typical bandwidth of 3.4 GHz for the HP 54701A. The specified bandwidth is 2.5 GHz.

Input Protection

In the past, active probes have been notorious for being highly susceptible to electrostatic discharge (ESD), making them a specialty tool and not a general-purpose tool. Since the need for high-frequency measurement devices is increasing, a rugged active probe has become essential. This is a problem not easily solved for any microwave device. Protection schemes typically require large input devices shunted by large clamping diodes, which all have high input capacitances. The need for low probe input capacitance conflicts with these schemes, making them of no use.

The HP 54701A probe uses four protection schemes working together to protect the input up to 20 kV of ESD while maintaining a low 0.6-pF input capacitance. The four schemes consist of a dual-cut spark gap, an ac input coupling capacitor, a dual-path amplifier, and p-i-n diode clamps.

The dual-cut spark gap is the first line of protection for the probe. Fig. 2 illustrates the physical layout of the dual-cut spark gap used in the HP 54701A. It consists of a long narrow line connecting the input of the probe to ground. Two laser cuts (0.002-inch gaps) are made in the line to make the line appear as an open circuit under normal operation. When a high voltage is seen at the input these cuts will exhibit a corona breakdown, limiting the input voltage to less than 1.2 kV. The long line to ground is needed to isolate the input from ground, thus maintaining minimum probe input capacitance. Two spark gap cuts are needed to drive the LC resonance resulting from the capacitance of the gap and the inductance of the line out of the bandwidth of the probe. The original design had only one gap with a capacitance of approximately 125 fF. The inductance of the line is approximately 10 nH, putting the resonant frequency at 4.5 GHz. By cutting two gaps the capacitance is reduced by a factor of two, driving the resonance out to approximately 6 GHz. This puts the resonance far enough out of band that it does not disturb the probe's response, yet maintains a low enough breakdown voltage to protect the input.

The second line of protection is an ac coupling capacitor in front of the RF amplifier. This capacitor is sufficiently small (200 pF) that it limits the amount of energy that can be transmitted to the RF amplifier. To maintain a dc coupled amplifier a dual-path amplifier was designed with operational amplifier U1. The input to the operational amplifier is protected by a high division ratio of probe input signal to operational amplifier input signal. The overall performance of the probe is optimized by using the operational amplifier's low noise and accurate dc performance characteristics.

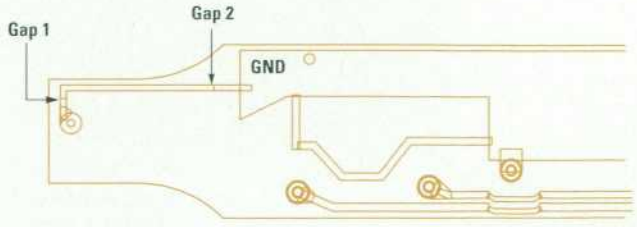


Fig. 2. Dual-cut spark gap for electrostatic discharge protection.

The last line of protection consists of two p-i-n diodes back to back, connected to the input of the RF amplifier. The two diodes also limit the amount of energy seen by the RF amplifier, making the overall system ESD-resistant up to 20 kV.

Response Accuracy

It is important for probes to have accurate responses since they become part of the overall measurement system. To achieve the specified accuracy for the HP 54701A, several techniques are used, ranging from a dual-path amplifier to a precision active trim.

A dual-path amplifier has many advantages over a single-path amplifier running open-loop. The first was mentioned above: it allows the RF amplifier to be ac coupled, thus limiting the amount of ESD energy reaching its input. Another benefit of the dual-path amplifier is dc stability. The high loop gain of the low-frequency amplifier—U1 in Fig. 1—helps eliminate the dc offset and drift of the RF amplifier. Another effect corrected by the low-frequency loop are thermal tails generated by the RF amplifier. These are caused by the instantaneous change in power of the microwave transistors when a step is applied to the probe input. The heat generated by the change in power must dissipate down through the transistor and into the substrate. As the temperature changes the V_{be} voltage of the transistors changes, causing the probe's output signal to track the thermal change. Thermal tails in the HP 54701A occur with a time constant of approximately one microsecond. This is the thermal time constant for heat to be dissipated from the junctions of the microwave transistors into the ceramic substrate. To correct this problem the low-frequency loop requires a bandwidth of 500 kHz.

Three adjustments are made to match the responses of the low-frequency loop, the RF amplifier, and the input compensation divider. The adjustments are made by an active trim method developed specifically for this probe. The dc gain is controlled by the low-frequency amplifier, which is adjusted by resistor ratios. A dc voltmeter measures the output of the probe and controls a laser trimmer, which trims the resistor. This is done in an iterative manner until the dc gain of the probe is set to 0.1 ±0.05%. The next two gains set are ac responses, which creates the problem of accurately measuring pulse parameters to an absolute level. To do this, substitution is used. The input pulse is first put into an oscilloscope for calibration through a 20-dB pad. Measurements are made of voltage levels at 3 ns, 80 ns, and 1.4 μs after the leading edge. The 1.4-μs measurement is considered to be a dc measurement, the 80-ns measurement is considered midfrequency, and the 3-ns measurement is considered high-frequency. To trim the midfrequency gain the difference between the voltage levels at 80 ns and 1.4 μs is calculated and compared to the calibration. As before, a resistor is trimmed in an iterative manner until the probe's response matches that of the calibration. To trim the high-frequency response the same method is used as for the midfrequency, except the difference is taken between the 3-ns measurement and the 80-ns measurement and the trimming is done on the input capacitor C_1 in Fig. 1. The active trim matches all three gains within ±0.25%, giving the probe a very flat, accurate response. Actively trimming the input capacitor C_1 to match the compensated divider eliminates

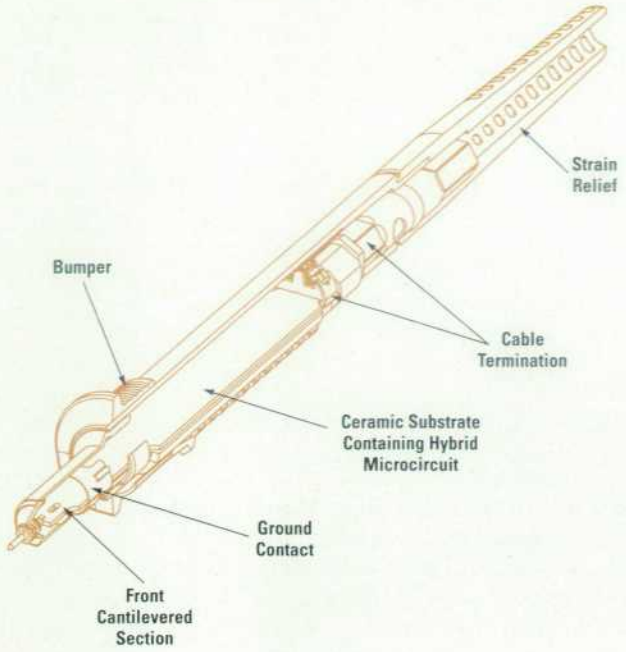


Fig. 3. Cutaway view of the active probe.

the need for a bulky mechanical adjustment. This makes it possible for C_1 to be as small as 100 fF and be trimmed to an accuracy of approximately ±0.1 fF, thus minimizing the input capacitance of the probe.

Mechanical Design

Active probes have had a reputation for being expensive and very fragile—easy to break both electrically and mechanically. The electrical performance of the HP 54701A probe requires the use of small, precise geometry and materials that are optimum for their electrical properties but have poor mechanical properties. The mechanical challenge was to create a very durable package that does not compromise the electrical performance. From both the electrical and mechanical standpoints, we wanted the user to be able to use the probe without worrying about breaking it either electrically or mechanically.

It was decided early to achieve the electrical goals by suspending the ceramic substrate that contains the probe circuitry in air in the probe (Fig. 3). The mechanical structure has to hold the probe tip in the proper position and make electrical contact with the circuitry, but not allow any external forces to impact the circuitry. These external forces can come from normal probing, dropping the probe, pulling or flexing of the cable, or environmental factors.

The final design accomplishes these objectives by isolating the circuitry from external influences by means of a protective shell and by using the backbone of the probe, the heat sink, to prestress the more fragile component, the ceramic substrate (Fig. 4). Preventing the ceramic from breaking made for a very challenging mechanical design.

Electrical Considerations

The input impedance of a probe determines how it will affect the circuit under test. The ideal probe has a very high impedance, that is, the circuit sees it as a very large resistance, and more important, as a very small capacitive load.

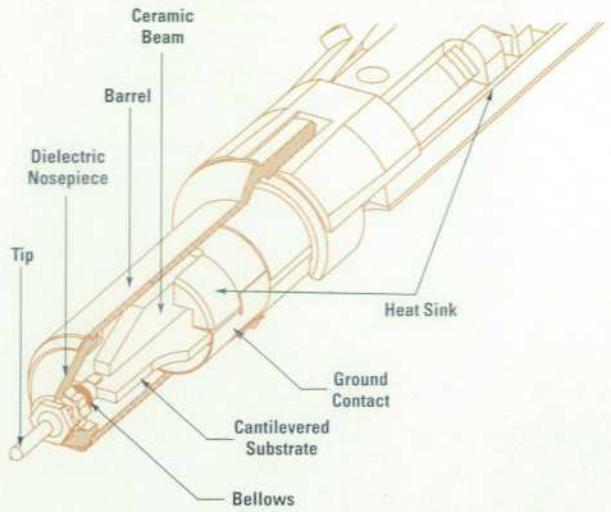


Fig. 4. Cutaway view of the probe tip.

Inductive loading comes from the ground return. Inductive and capacitive loading becomes increasingly important as the frequency increases.

The input capacitance of the probe is highly dependent on the geometry of its input circuitry. To provide low capacitive loading the input structure should have a very small geometry with a very low dielectric constant, and there should be space separating the circuit elements and the ground return. The best dielectric obtainable is a vacuum, with air being nearly as good. Unfortunately, low-dielectric-constant materials have virtually no strength or stiffness, which are needed for mechanical ruggedness. To achieve very low input capacitance and increased high-frequency performance the input circuitry is located "far" from the metal ground structure, that is, it is built on ceramic and suspended in air.

The replaceable tip screws into threads ultrasonically inserted into an acetyl dielectric nosepiece. This plastic part is then crimped into a stainless-steel barrel which provides the ground return path. The tip makes contact with the circuitry by means of a flexible bellows made of very thin nickel plated with gold. If the probe is dropped directly on its tip, the replaceable tip will be bent but the circuitry will be unaffected.

The signal return for very high frequencies requires a short path to minimize inductive loading of the circuit under test and the oscilloscope front end. Minor changes in the way the signal is returned proved to cause very large changes in the pulse response of the probe. The ground return is implemented by linking the ground structure under the first transistors to the front tube by means of a beryllium copper contact. This contact also centers the hybrid during the substrate attachment procedure and prevents movement of the front tube assembly with respect to the ceramic substrate during use.

The inductive loading is affected primarily by the shape of the conductors bringing the signal in and returning the signal to the circuit. Returning the signal typically requires some sort of external hookup because of all the possible probing configurations and circuits and the requirement that the probe can move freely around. The primary constraint on the HP 54701A was to provide a well-defined ground path in a very short distance to the barrel of the probe. We

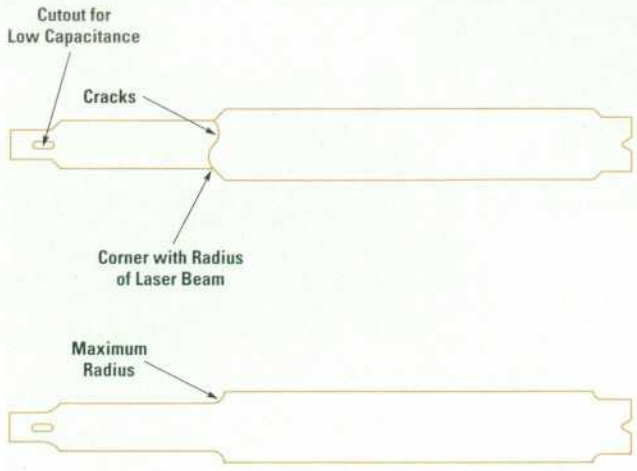


Fig. 5. In drop tests, the ceramic substrates tended to break where the laser cutting them out changed directions (top). This failure mode was eliminated by increasing the radius at that point.

started with a short barrel at the front of the probe to minimize capacitance, but had to lengthen the barrel to shield and isolate the high-impedance input circuitry. This also helps make the front end more rugged and causes only a small increase in input capacitance.

Throughout the design of the probe, signal fidelity was a primary goal. It drove the design decisions in almost all areas. A thick-film hybrid process was chosen for the ability to fire and trim very high-precision resistors and capacitors and to use bare IC chips. Unfortunately, this means that the substrate is a ceramic material, which is characteristically very brittle and prone to breakage. Ceramic has a relatively large dielectric coefficient, another reason for suspending the front circuitry in air. A cutout at the front (Fig. 5) minimizes stray capacitance in the divider network.

Ruggedness

We wanted the customer to be able to treat this probe with impunity. Accidental drops from a lab bench onto hard floors are everyday occurrences, and if a drop results in the breakage of an expensive piece of equipment, the results are an unhappy customer and a warranty problem.

Initially we used the standard drop table to test the ruggedness of the probe assembly. This is a controlled experiment in which the device under test is firmly attached to a heavy table and subjected to a controlled shock impulse—typically a half-sine pulse whose amplitude and width can be adjusted within bounds. For engineering tests we started with small drops and raised the drop height until we broke probes. This test was useful in that substrates were broken, pointing the way to design improvements. The substrate is cut out with a CO_2 laser. When dropped with the plane of the substrate vertical it tended to break where the laser changed direction (Fig. 5), leaving a radius about the size of the laser beam. Maximum radii were added as shown in Fig. 5, effectively eliminating that failure mechanism.

At this stage we periodically dropped probes on the floor and struck them against walls as a check of the more controlled tests. These probes did not appear to break, so we felt fairly good about the assembly and the validity of the standard drop tests. What we did not know is that we were not dropping them enough times and that we had not yet learned how to detect microcracks.

By the time we had functioning prototype probes, they could not be broken on the shock machine even with the table raised to the maximum possible height (approximately 60 in) which is far above the required qualification height. But as probes were distributed for use, we did have a number of prototypes that came back broken. These were probes that were dropped from bench level onto a hard floor. In other words, a six-foot controlled drop would not break the substrate but a 30-inch uncontrolled drop would.

Controlled versus Uncontrolled Drops

To understand the problem we examined the difference between the controlled drop and the uncontrolled drop. In the controlled test the probe is firmly attached to a heavy steel table and subjected to a well-defined half-sine-shaped pulse without impacting any particular point on the probe. The machine can be adjusted for various magnitudes and shapes of pulses, the limit being a full-height (60-in) drop of the table without any cushioning, resulting in a half-sine-shaped pulse of 830g with a pulse width of 1.72 ms.

To see what kind of shock a real drop presents to a probe, we attached an accelerometer to a probe and simply dropped it on the floor. Pulses as short as 0.5 ms were easily obtained. However, typically the accelerometer would overload, so we never got a good measurement of acceleration.

Analytically, we usually assume these pulses to be half sine waves. The limit as the pulse width goes to zero and the amplitude goes to infinity is an impulse function. A normalized measure of the energy in the pulse is ΔV, the integral of the pulse acceleration over time. If the pulse is assumed to be a half sine wave, then:

$$\Delta V = \int_0^\tau a(t)dt \approx \int_0^\tau A\sin\left(\frac{\pi t}{\tau}\right)dt = 2\frac{A\tau}{\pi}.$$

In a free-fall impact, we can also calculate the change in velocity at impact as the difference between the impact velocity $V_i = -\sqrt{2gh}$ and the rebound velocity $V_r = \sqrt{2gx}$, where h is the drop height and x is the rebound height. Using the concept of a coefficient of restitution e:

$$e = \left|\frac{V_r}{V_i}\right| = \sqrt{\frac{x}{h}}, \quad 0 \le e \le 1,$$

we can write:

$$\Delta V = |V_r - V_i| = e|V_i| + |V_i| = (1 + e)|V_i|$$

or

$$\Delta V = (1 + e)\sqrt{2gh}.$$

Thus we can calculate the expected ΔV for a drop on the floor given a value for e. If we assume e = 0.2 (customary for a hard surface) we can calculate a half-sine pulse of:

$$A = \frac{(1 + e)\pi}{\tau}\sqrt{2gh} = 1256g \text{ acceleration amplitude}$$

for a probe dropped 36 inches to the floor.

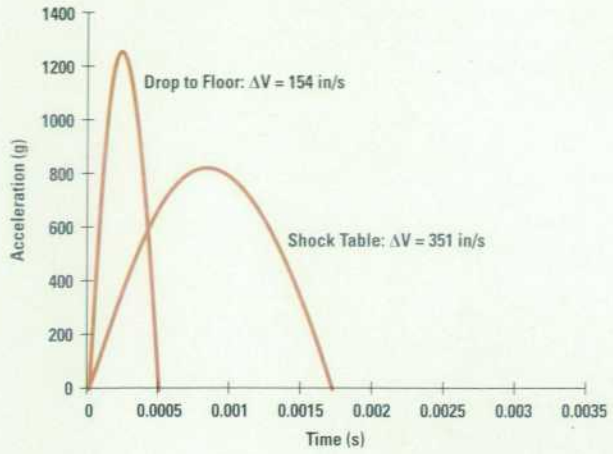


Fig. 6. Controlled and estimated uncontrolled shocks.

We now can calculate the spectrum of the pulse either analytically by the Fourier transform or simply by using a computer and a fast Fourier transform (FFT) routine.

Fig. 6 shows a representation of the ideal pulse from the shock machine at its maximum drop and the shock a probe can get simply by being dropped on the floor. Fig. 7 shows the spectra of the two pulses. Clearly there is more high-frequency energy in the drop on the floor than the shock machine is capable of delivering even though ΔV is smaller.

Substrate Dynamic Response

Experimentally we had verified that the lowest resonance of a loaded substrate is about 2.5 kHz. We thought that this was well beyond what would cause problems in shock or vibration. Now we knew why we were wrong. Our shock machine was not capable of much less than about a 1.5-ms pulse width, which produces very little energy beyond about 1 kHz.

The next step was to measure the dynamic response of the substrate by hitting it with a very small pulse and characterizing its response. The typical method is to attach an accelerometer to a structure, hit it with an impact hammer in various places, and use a spectrum analyzer to generate a transfer function that shows the frequencies at which the structure resonates. We had a real problem with this structure because even the tiniest accelerometers are very massive compared to the substrate and would not have given accurate results. The same is true of impact hammers—the available ones are just too large.

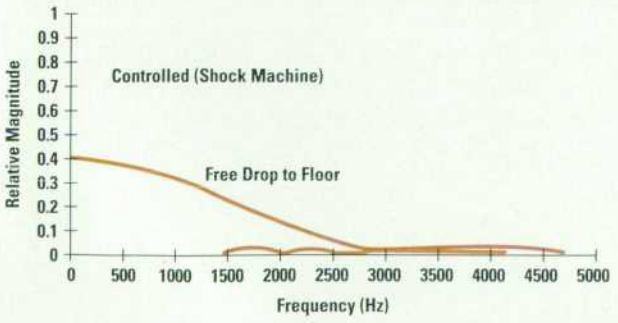


Fig. 7. Spectra of controlled and estimated uncontrolled shocks.

Instead, we used a magnetic field proximity detector (Kaman KD2400 displacement transducer), which required only a small piece of copper tape on the substrate. The detector outputs a signal proportional to position. The impact was a small piece of brass wire bounced on the ceramic. We measured the response of the ceramic but did not have any measurement of the input force so we could not measure the transfer function. However, we were able to look at the spectrum of the response and make qualitative judgments.

The substrates measured were blank rather than loaded with components (because we were testing unloaded substrates). There are a few components, notably capacitors, that add mass to the real circuit and would tend to lower the resonant frequencies.

In these tests, the primary mode of vibration was found to be at the center of the substrate—a fairly sharp peak at about 3.5 kHz. Earlier tests with loaded substrates put this resonance at 2.5 kHz. The second mode was the tip vibrating at about 8 kHz. The modes were isolated by moving the proximity sensor around to find the points of maximum response. We now had a feeling for how the internal parts move under impulse excitation as well as what frequencies were important and a way to tell if modifications changed the structure's characteristic properties.

"Feel Good" Test

We also wanted some way to test the actual breakage mechanisms, that is, to hit the probe with a hard pulse. We tried for some time to come up with some repeatable way to hit probes with sharp pulses, but eventually decided that if we could break them by simply dropping them on the floor, why not just drop them on the floor?

A "feel good" test was agreed upon by the project team. Basically, the assembled probe should survive a large number of uncontrolled drops from heights in excess of a typical lab bench. We made a number of different probes, each with different solutions, and put them through a series of drops of increasing heights. To pass the "feel good" test a probe had to survive ten three-foot and ten four-foot straight drops onto a concrete floor, and ten three-foot and ten four-foot pulls off a table onto a concrete floor. We used the technique of pulling a probe off the table by its cable because this is how some of the prototypes were broken. This proved to be a particularly effective way of consistently breaking substrates.

We prepared prototype probes and began the "feel good" test. What we saw at first was that the probes would survive a few drops, but repeated drops would cause breakage. This helped explain why we did not detect these breakages earlier. We began to experiment, brainstorming different experiments that we thought would improve performance on the "feel good" test. Occasionally we would have great successes, only to find that the results were not repeatable. We eventually amended the "feel good" test so that a sample of five probes had to pass.

The three solutions that survived the test and were implemented are:

- Change the dynamic response of the substrate to shift the primary mode of vibration up in frequency and effectively

eliminate the secondary mode. This is done with a ceramic beam.

- Absorb the energy of the drop, particularly the high-frequency energy (> 1 kHz), by means of a rubber bumper.
- Prestress the ceramic so that it is not loaded in tension in the worst breakage areas.

Ceramic Beam. The ceramic beam provides additional support to the front of the cantilevered ceramic substrate (see Fig. 4). This beam mostly eliminates any breaking of the tip of the ceramic. In dynamic measurements, the second-harmonic mode of vibration is essentially gone, and the first mode of vibration is shifted from 3.5 kHz to 4.2 kHz and reduced in amplitude. The beam is epoxied into position when the substrate is attached. Ceramic is an appropriate material because it is familiar in the assembly process area and does not introduce problems with differing coefficients of thermal expansion.

Bumper. The soft 60-durometer PVC bumper is an overmold on the probe shell (see Fig. 3). Its function is to prevent the most vulnerable part of the probe from hitting a hard surface. It also acts as a low-pass filter, effectively absorbing the high-frequency energy in the shock pulse. This makes the front of the probe larger than desired, but we decided that the gain in ruggedness offsets the gain in size, and someone who is extremely concerned with size can easily cut away the rubber bumper.

Prestressing during Gluedown. The first two solutions gave us order-of-magnitude increases in survival rates over the initial configuration, yet after several drops, the ceramic would stubbornly break at the point where all the mechanics came together. Typically the break would be in the form of a microcrack—difficult to detect visually or electrically unless the probe were somehow stressed, either by bending the probe body or by temperature induced stresses. We realized that a lot of the anomalous behavior we had seen throughout development could be explained by these microcracks—small enough to maintain electrical continuity but opening up under an applied stress, either thermal or mechanical. This also helped explain why we did not detect breakage earlier in the process—we did not know what to look for.

We amended the test for breakage from a visual inspection to an electrical check under an applied bending moment. We checked the continuity of a trace that runs the length of the substrate while bending the outside of the probe to open a crack in the substrate. This proved to be a simple and foolproof method of detecting a crack. We also applied a dye to the substrate after the tests to verify whether there were cracks.

Ceramic is an extremely brittle material. Its strength is very high in compression, and under ideal circumstances can be very high in tension, but various factors conspire to create defects and stress concentrations that cause it to fail catastrophically under what can be very low loads in tension. These tests made it clear that the best way to keep a brittle material such as ceramic from fracturing is never to allow it to be loaded in tension. The practice of prestressing has been used for many years to allow structures such as bridges and buildings to be fabricated out of concrete, another brittle material. The structure is heavily loaded in compression so that design loads that would normally bring the

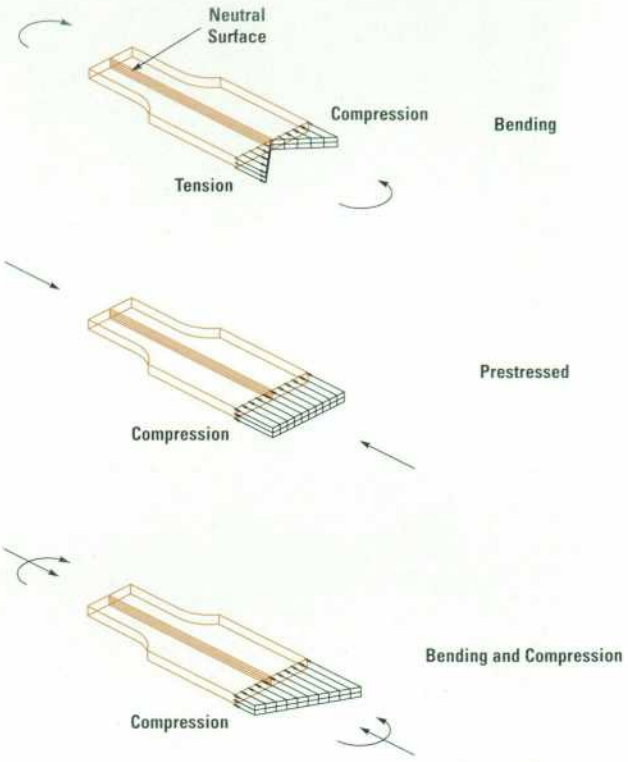


Fig. 8. Prestressing prevents tensile loads on the ceramic substrate.

structure into a tensile state of stress only cause the material to go from a highly compressive to a less compressive state of stress (see Fig. 8).

We had been using a conductive epoxy to attach the substrate to the heat sink. The epoxy minimizes the length of the thermal path from the small IC chips to the outside of the probe. It also maintains the precise mechanical tolerances from the heat sink features to the interconnect features on the hybrid for front-end interconnections. Initially we used only a small epoxy preform in the area where there was a critical thermal requirement, but we tried a large number of variations during testing. The technique that gave best results was to use a small patch at the front where the hot ICs are and another in the rear to force the prestress across the middle where the breakage typically occurred. Aluminum has about twice the thermal expansion of alumina (ceramic), so conventional wisdom would lead one away from such a situation. Joining two such dissimilar materials tends to lead to high thermally induced stresses. However, we used this to our advantage to build-in the desired prestress. The epoxy is applied as two preforms of the correct shape in an uncured condition (see Fig. 9). The substrate is precisely fixtured to locate it with respect to features on the heat sink and to compress the front and rear beryllium copper contact springs. The entire sandwich is put in a curing oven at 300°C. During this phase the epoxy is essentially a viscous fluid and the substrate and heat sink come into equilibrium at 300°C. The epoxy cures, locking the two dissimilar materials at this state. Upon cooling, the aluminum heat sink wants to shrink more than the ceramic, putting the aluminum in tension and the substrate (ceramic) in sufficient compression that cracks are never allowed to open in the brittle ceramic.

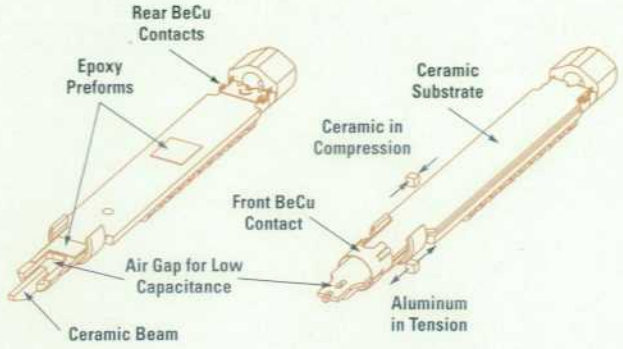


Fig. 9. Method of prestressing the substrate using epoxy.

It turns out that the force causing the prestress depends primarily on the cross sectional areas of the ceramic and aluminum and the curing temperature, and is independent of the distance separating the epoxy patches. The epoxy patches need to be large enough to maintain the shear loading, which is really the ideal way to load a glue joint. Tests showed that two patches on each end worked better than epoxy over the entire surface.

Conclusions

The electrical design of the HP 54701A active probe dictated structures less than optimum for mechanical ruggedness. Performance goals often conflicted with manufacturability and ruggedness goals. Through the process of trying to pass our "feel-good" test we tried many different configurations. In the end the three solutions of ceramic beam, rubber bumper, and prestressed gluing yielded probes that consistently passed our test. Removing any one of the solutions again gave us failures, so all three solutions were implemented. The result is a probe that greatly improves both electrical performance and mechanical ruggedness over previous designs.

Acknowledgments

We are grateful for much hard work and significant contributions from many people of the Colorado Springs Technology Center, the process development group, and the tooling group. In particular we would like to thank Mark Mathews, Dale Pittock, Mark Michaels, Phil Yearsley, Dave Kimbley, and Kim Wilson for their contributions.

Accuracy in Interleaved ADC Systems

The overall performance of the HP 54720 oscilloscope is the result of the synergistic effects of calibration, signal preconditioning, and data postprocessing.

by Allen Montijo and Kenneth Rush

Why would someone choose to implement a time-interleaved ADC system instead of simply using the latest technology to implement a single ADC with the desired performance? Simply put, a time-interleaved ADC system can achieve superior performance, given the same implementation technology. For a given technology, there is theoretically no limit to the sample rate that can be reached using interleaving methods, although there is a limit to the bandwidth and thus the usefulness of interleaving. Real-world limitations such as power and space place practical limits on the level of interleaving that can be achieved. For the HP 54720 oscilloscope, these limits allowed us to put four 500-MSa/s ADCs onto a single hybrid to achieve a maximum sample rate of 2 GSa/s. With the HP 54722A plug-in, all four hybrids are interleaved to achieve sample rates up to 8 GSa/s in one instrument. Sixteen ADCs on four hybrids work together to sample at consistent intervals of 125 ps with a signal bandwidth of nearly 2 GHz.

Interleaving is accomplished by routing the input signal to all ADCs and providing a conversion clock with a suitable phase to each ADC (see article, page 11). One ADC hybrid in this instrument contains four ADCs, each capable of sampling at 500 MSa/s. The input signal is split four ways, sampled, and fed to the four ADCs. Each of these four signal paths is furnished a clock, which is used to coordinate sampling and digitization. In this system all four clocks run at 500 MHz at all sample rates, giving an overall data rate of 2 GSa/s per hybrid. Circuitry on the sampler IC controls the phasing of the clocks while circuitry on the ADC ICs decimate the data to the correct rate (see Fig. 1).

To illustrate some of the problems that interleaving creates, look at a simple system that interleaves two 1-GSa/s ADCs to achieve 2 GSa/s. For ease in understanding, assume that the ADCs have infinite resolution. The results obtained apply to real ADCs with finite resolution on a statistical basis.

An offset difference between the two ADCs is digitized to different codes by the ADCs. Since the ADCs output data alternately, the output data values exhibit a pattern with a period equal to the sample rate of an individual ADC, or 1 GHz in this example (Fig. 2). The error signal is not a function of the input signal, but only of the difference in offset. When N ADCs are interleaved, the basic period of the error signal remains the same (the sample period of a single ADC), but the frequency content can increase because each cycle of the error signal contains N values and N arbitrary offset errors.

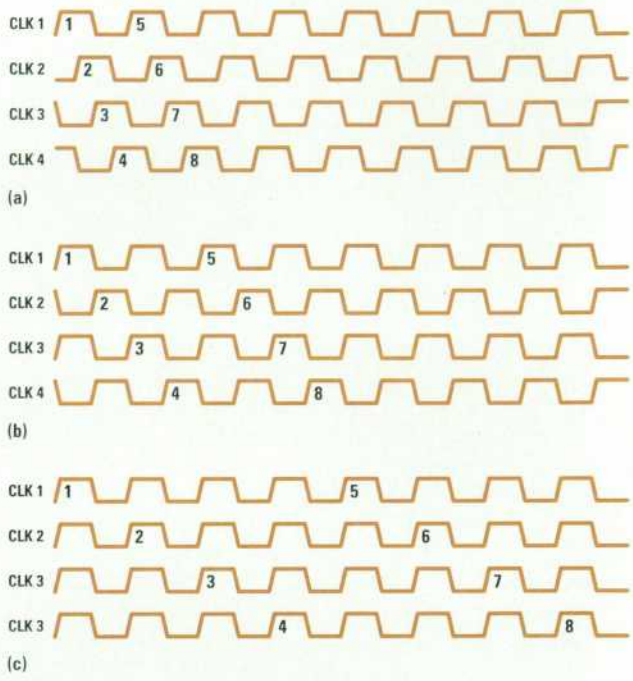


Fig. 1. Clock phasing and data decimation control the sample rate. The clocks always run at 500 MHz and the edges that control the first eight samples are indicated. (a) At 2 GSa/s the clocks are phased in increments of 500 ps (90 degrees) and data is stored on every clock. (b) At 1 GSa/s the clocks are phased at increments of 1 ns (180 degrees) and each path stores every other sample. (c) At 500 MSa/s and below, the clocks are in phase and the ADCs take turns storing data, decimating the data as required.

Fig. 3 demonstrates that if there is a difference in voltage gain between the two ADCs, mixing occurs. If a sine wave is applied to the system, the largest difference in ADC outputs occurs at the peaks of the sine wave. As with an offset error, the basic error occurs with a period equal to the sample rate of an individual ADC, but the magnitude of the error is modulated by the input frequency. In this example, if a 20-MHz signal is applied and the two ADCs have a difference in gain of 10%, then the envelope of the error signal is 5% as large as the 20-MHz input. The error signal is a 1-GHz carrier that is amplitude modulated by the 20-MHz input signal. With a full-scale sine wave input and a 0.25-dB mismatch in gain (<3%), the effective bits parameter is limited to 5.8 by this source of error alone. Increasing the number of ADCs that are interleaved increases the number of frequencies that are

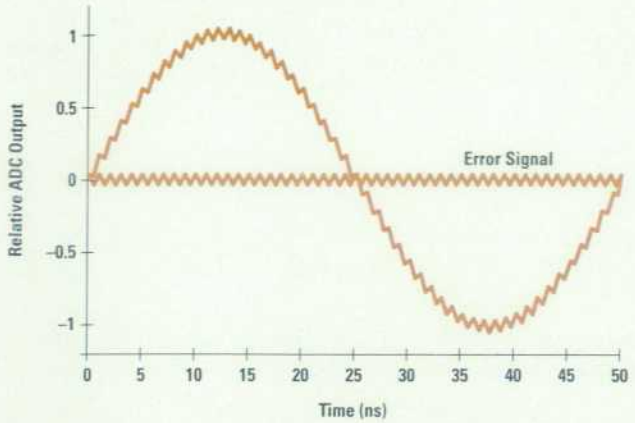


Fig. 2. The output produced by a system of two interleaved ADCs with an offset error between them is the original signal plus an error signal at the individual ADC sampling frequency, or half the interleaved sampling frequency.

Fig. 4. Output from a system of two interleaved ADCs with a timing offset between them includes an error signal at the individual ADC sampling frequency, or one half the combined sampling frequency. The carrier is amplitude modulated by the derivative of the original signal. The error is maximum on the slewing portion of the waveform.

available as carriers. All harmonics of the individual ADC sample rate can simultaneously be carriers.

An error in timing shows a similar result (Fig. 4). The samples are supposed to be precisely 500 ps apart. If they are 510 ps and 490 ps apart, then there is a 20-ps skew between the ADCs. This timing skew will show the largest error where the signal has the highest slew rate, or at the zero crossings. The envelope of the error signal will be largest at the zero crossings and will have a period equal to the period of the input signal. Once again, the basic error signal has a period equal to the sample rate of an individual ADC. The error from a timing skew is identical to the error from a gain difference, except that its envelope is shifted by 90 degrees and its magnitude is frequency dependent. In this case the 1-GHz carrier is amplitude modulated by the derivative of the input signal. With a 1-GHz input signal and a time skew of 4 ps, the number of effective bits is limited to 6.0 by this source of error alone.

The three sources of error just described are straightforward to deal with. Unfortunately, other errors are not so simple. Differences in frequency response between the signal paths create gain and phase (timing) errors that are a function of frequency. These differences can be caused by process variations, which affect the positions of poles and zeros in the response, and by nonsymmetric coupling between the signal paths.

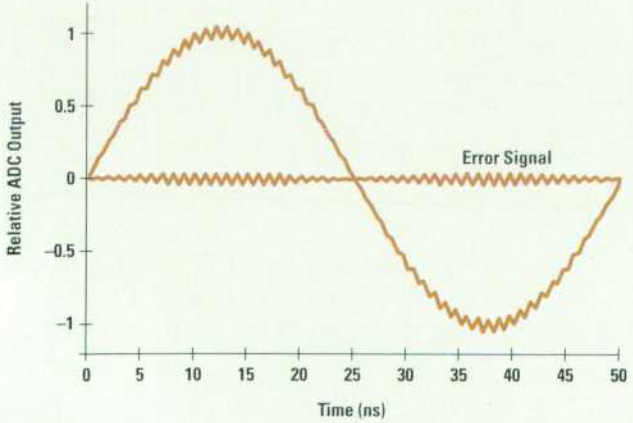


Fig. 3. A system of two interleaved ADCs with a gain error between them introduces an error signal that consists of a carrier at the individual sampling frequency that is amplitude modulated by the original signal.

Another problem with the interleaving approach is that it requires more circuitry, not just in the ADCs, but also in support circuitry such as DACs. Software complexity also grows since there is more circuitry to control, although it grows at a slower rate than the degree of interleaving.

Even with all of these pitfalls, interleaving provides the least expensive and highest-performance method for sampling at extremely high rates.

Calibration

How do you get the most out of any ADC system? The task begins with calibration; the quality and extent of the calibration is one limiting factor determining the accuracy of the data. Calibration typically includes gain and offset adjustments for the vertical system and timing alignment for systems with time-interleaved ADCs. We also calibrate the system linearity to achieve the ultimate level of performance.

When we set out to design the HP 54720 system, we had several goals in mind for calibration. First, the user must be able to calibrate the components of the system (plug-ins and mainframes) separately, and each component must have a place to store its own calibration factors, so that any calibrated plug-in can be inserted into any calibrated mainframe and the combination will typically be accurate within 3%, using only the yearly calibration data. Second, a 1% "best-accuracy" calibration must result when a plug-in is calibrated in and remains in a specific slot of a mainframe. Third, the calibration sources must be completely self-contained.

Two methods are used in the system calibration. The most common is physical adjustment of a hardware parameter such as gain, usually via DACs. The second is characterization: a parameter, such as linearity, is measured so that the instrument software can compensate after the data has been acquired.

Vertical Calibration

The signals used for most calibrations come from a calibration (**CAL**) BNC connector on the instrument front panel, which can output a variety of ac and dc signals. To calibrate the system vertically, a high-quality (±0.2%) dc source with a range of ±2.5V is used. Feedback from the **CAL** BNC allows the dc amplifier to cancel ground drops in the system and any resistance in the internal cable leading to the BNC connector.

Three statistical techniques are used to improve the accuracy and repeatability of the calibrations. The first is to take a statistically significant amount of data for each measurement. Knowing the noise sources in the instrument and the desired calibration measurement accuracy, it is a simple matter to determine how much data is required. The second technique is to dither the dc calibration source. Rather than making one measurement with a fixed input voltage, several measurements are made with different values for the dc source. For example, if a measurement is desired with the dc source at 40 mV, 21 measurements might be made with the source scanning from 39 mV to 41 mV in steps of 100 μV. Given the quantizing effects of the ADCs this is extremely important; this step eliminates uncertainty resulting from finite ADC code widths. The third technique is to randomize the order in which data is taken when determining transfer curves. This reduces the effects of low-frequency perturbations in the system by distributing them across the entire transfer curve rather than just a small section. Assume that an nth-order polynomial is being fit to a transfer curve and a disturbance causes the first seven measurements to be low. If those seven measurements are concentrated in one area of the transfer curve, then all terms of the polynomial are affected. However, if the seven measurements are spread throughout the transfer curve, then the effect is primarily limited to the y-intercept term.

The nominal input voltage range to the mainframe is 160 mV, centered about 0V. Each signal path has a gain vernier on the sampler IC with enough range to account for process variations plus an additional ±8% to compensate for gain errors in plug-ins and probes. Calibrating this vernier allows the system to use the full range of the ADCs at all major sensitivity selections.

Each signal path also has its own offset adjustment. Offset error is a function of the gain vernier position and must be calibrated as a function of the vernier. Simultaneous calibration of the gain vernier and offset is required if the system is to use the full range of the ADC. Offset calibration accomplishes two things: (1) it compensates any offset errors along the signal path and (2) it injects the dither signal (see "Dither and Bits," page 42) into the signal path.

The gain and offset are calibrated by adjusting the respective DACs until the top and bottom ADC thresholds are at the desired input voltages. For example, at nominal gain, these parameters are adjusted until the top ADC threshold is at +80 mV and the bottom ADC threshold is at –80 mV. Dither is introduced during the gain and offset calibration by modifying this calibration. The target threshold voltages are adjusted such that thresholds are in increments of 1/4 LSB (1 LSB = 1.25 mV at nominal gain). In the above example, the top and bottom thresholds would be adjusted to +80 mV + dither[n] and –80 mV + dither[n], where dither[n] = –3/8 LSB for the first path, +1/8 LSB for the second path, –1/8 LSB for the third path, and +3/8 LSB for the fourth. Thus, the ADCs are calibrated so that the thresholds are interleaved in voltage, with thresholds spaced 1/4 LSB apart.

The gain vernier is calibrated by calibrating the gain, offset, and dither at several gain settings within the ±8% range. For each signal path, polynomials are fit to the data to arrive at expressions for the gain vernier DAC code and the offset DAC code as a function of the requested gain.

Real-world ADCs are not perfectly linear. A variety of problems such as nonlinear input capacitance create system linearity errors (integral nonlinearity) while effects such as resistor and V_{be} mismatch create local linearity errors (differential nonlinearity). In the acquisition hybrid, these errors are characterized and accounted for through a linearity correction calibration. The calibration routine slowly moves the input signal across the range of the ADC to determine the position of each threshold. Each ADC code from each path is then assigned a value midway between its upper and lower thresholds. For example, if the thresholds for code 62 from the third signal path are located at 3.245 mV and 2.133 mV, then that code is assigned the value (3.245 mV + 2.133 mV)/2 = 2.689 mV. This characterization maximizes the accuracy of the system and minimizes the noise power of quantization and linearity errors.

When a plug-in has its yearly calibration performed, the mainframe in which it is calibrated acts like a voltmeter and needs to be in a state of calibration. The gains of the passive attenuator elements are determined and stored for later use. The transfer curve is determined for each preamplifier gain setting and is approximated with a third-order polynomial. After the plug-in is calibrated, the calibration data is stored in the plug-in's EEPROM so that it stays with the plug-in. When the user selects a voltage sensitivity setting, the best attenuator and preamplifier settings are determined, and a correction table is built for each ADC path which translates ADC codes to input voltage using all of the linearity characterization data for the attenuator, preamplifier, and ADC path.

The best-accuracy calibration is essentially this case; the calibrated plug-in stays in the mainframe slot in which it was calibrated. The accuracy of the calibrated combination of the plug-in and mainframe is better than the 3% accuracy specification. As the plug-in is moved to other slots or other instruments, differences in terminations and small calibration errors can add up to give a typical accuracy of 3%.

The vertical calibration portion of the best accuracy calibration is different from the plug-in calibration in the following respects: (1) an extra offset calibration is performed to cancel changes in mainframe offset as a function of sample rate, and (2) the calibration data is stored in the mainframe. Calibration errors do not add up during the best-accuracy calibration, since the important parameters are measured for the complete system. Therefore, the best-accuracy calibration achieves a 1% accuracy specification.

Horizontal Calibration

As mentioned previously, logic on the sampler IC controls the interleaving of the four paths on a hybrid to obtain the basic sample rates of 2 GSa/s, 1 GSa/s, and 500 MSa/s. Given

the tight timing required to achieve our performance goals, this logic merely provides coarse control over the sampling interval. The sample clock for each signal path on the acquisition hybrid has a variable delay control with a range of 60 ps. The sampler clock timing calibration uses this control to adjust the sampling interval precisely.

Since this is an oscilloscope, the obvious calibration method is to capture a high-speed signal (such as a fast edge) in equivalent time, building up separate waveforms for each path or hybrid. The timing difference can easily be measured and adjusted. Unfortunately, the 30-ps timing accuracy of the time base is not compatible with our subpicosecond goals. To eliminate time base effects, we are required to make our skew measurements on a single-shot basis, or effectively in the frequency domain.

One challenge in developing this calibration resulted from settling problems in the signal paths to the ADCs on the first prototype hybrids. This created a frequency dependent phase error in the system that needed to be ignored during calibration. To perform this calibration for the mainframe, the mainframe outputs an 84-MHz square wave with edge speeds of approximately 1.6 ns. The third harmonic of this signal is at 252 MHz, near the Nyquist rate for a single ADC path. This frequency was chosen because phase errors caused by incomplete settling of the signal path are nulled. An acquisition is performed and the phase of the third harmonic is determined for each path. The phase errors are calculated and used to adjust the variable delays. After a few measure-and-adjust iterations, the calibration is complete.

To achieve sample rates of 4 and 8 GSa/s, two or four acquisition hybrids are interleaved in time, each running at 2 GSa/s. Each hybrid has a timing adjustment that moves its sampling position with respect to a 100-MHz system reference clock. To interleave the hybrids in time, the timing adjustments must move the hybrids so that they are sampling 250 or 125 ps apart. Unfortunately, as the timing control for a hybrid is moved, clock coupling also changes on the hybrid. This forces us to recalibrate the timing of the sampler clocks for the individual paths so that the four paths on each hybrid are sampling at 500-ps intervals. This timing calibration works in an iterative manner to adjust the hybrids to sample 250 or 125 ps apart while adjusting the paths on each hybrid to sample in 500-ps intervals.

To perform these calibrations, a special signal was designed into the HP 54721A 4-GSa/s plug-in and the HP 54722A 8-GSa/s plug-in. This circuit outputs a narrow pulse at a 152-MHz rate. The pulse is rich in harmonics, allowing the system to optimize the timing skew over a wide band of frequencies. The calibration does not rely on any specific relationship between different harmonics (phase or magnitude), making the circuit simple to build in production quantities.

As for the sampler clock timing calibration, single-shot acquisitions are performed and the phase errors between paths or hybrids are evaluated and adjusted. Since the hybrids have slightly different frequency responses (a one-degree difference in phase response at 1 GHz amounts to an effective time skew of 2.8 ps!), the phase errors of all odd harmonics within the system bandwidth are measured and the rms error is used to adjust the system.

Signal Preconditioning

The creation of a digitizing oscilloscope that displays a true representation of the signal being measured is a system design task. The system has many elements other than the analog-to-digital converter, and describing its behavior in simple ADC terms is incomplete. We prefer to think of the signal recording system in terms of input signal preconditioning, recording, postconditioning, and display. As described in "Dither and Bits" on page 42, a perfect ADC gets the wrong answer a great deal of the time. It can be improved significantly by preconditioning the input by adding an appropriate error signal before recording and then removing that error signal systematically upon playback. Consideration must be given to the Nyquist criterion as well. We cannot allow excessive signals above half the recording sample frequency to be present because they will be aliased down into the passband by the recording process and will irreparably contaminate the record.

Dither Insertion

In an interleaved ADC system there are ways of introducing dither without having to synthesize a high-frequency signal. As can be seen from Fig. 2, an offset error between two ADCs in an interleaved system produces a systematic error in the output whose frequency, phase, and amplitude are known if we know the value of the offset. "Dither and Bits" on page 42 describes the positive benefits of dither at half the sample frequency with a value of half an LSB and known phase. One simple method to generate this dither signal is to take advantage of the offset phenomenon between two ADCs in a two-phase interleaved system by intentionally offsetting one converter by half an LSB from its partner. Because we know how we offset, we know from which ADC's record we must subtract half an LSB when we play the record back. This is simple and very effective.

In a four-phase system there are many permutations of offset between the four ADCs that can be chosen. We have chosen each individual ADC's offset relative to its neighbors to produce as much energy as possible in the reject band of our system, that is, above one-fourth the sample rate. Since the error energy in an interleaved converter system is fixed by the fixed LSB levels of the converters, maximizing the out-of-band energy will automatically minimize the in-band energy. Keeping in mind that the system includes dither extraction and low-pass filtering to one-fourth the sampling frequency, this out-of-band energy will later be removed by the digital signal processing.

Anti-Aliasing Filtering

What are the consequences of letting too much energy leak through above the Nyquist frequency? We are taught that this will introduce errors, but what kind of errors? In an oscilloscope, these errors most often appear as inconsistencies in the recorded waveform. If we record the same signal over and over we do not get the same rise time or pulse width or overshoot each time. Because repeatability of these types of measurements is so important in oscilloscopes, an effective anti-aliasing strategy is necessary for the designer. From experience, we have found that achieving bandwidth beyond one-fourth the sampling frequency with repeatability in the pulse parameter measurements is very difficult. The

Dither and Bits

The fundamental problem with analog-to-digital converters is that they have blind spots. Where? Between the codes. With a perfect ADC the analog signal can change by one part in two to the Nth (where N is the number of bits) without the output changing at all. That's why there is such interest in the number of bits. Obviously, if the number of bits is high enough the spacing between the code transitions will become so small that even thermal noise cannot slip by without tripping a comparator. Herein lies another problem. Most people have no idea what thermal noise is, much less how to determine if it is a problem. These are the people who like to talk about their fourteen-bit multigigasample-per-second data acquisition system as if the fourteen bits were actually buying them something. Practically speaking, most broadband electronic channels are limited in dynamic range to about 50 dB. The limit on the high end is typically clipping of an amplifier or the hard limit imposed by the end of an ADC's range. The limit on the low end is usually some form of thermal noise from all the resistors in the circuit and the resistive elements in the transistors.

To a designer of frequency-domain communication circuits and audio equipment 50 dB sounds rather poor. Why the limit? It has to do with bandwidth. The amount of noise power in a channel grows directly as bandwidth increases. A one-gigahertz-bandwidth system might have ten times the noise power of a one-hundred-megahertz system, and one hundred thousand times the noise of a ten-kilohertz-bandwidth audio system.

For reference, 50 dB is equivalent to the dynamic range of a perfect eight-bit ADC. Why is the dynamic range of a "perfect" converter not infinite? It's those blind spots. In an output signal record the blind spots between the code transitions generate an error term that looks like random noise about 50 dB down. Therefore, eight bits is pretty important, but more than eight rarely does any good. It becomes clear, then, why people get upset if you try to sell them a digital oscilloscope with an ADC that has less than eight bits. They can easily see that the quantization error of a six-bit or seven-bit oscilloscope is greater than the analog noise in their system, while at eight bits it becomes very difficult to tell quantization from analog noise.

But when you're trying to design the world's fastest ADC using flash converters, which require a doubling in complexity for each bit, eight sounds like an awfully big number. A designer knows that you cannot deliver eight bits at state-of-the-art speed and tends to look for "magical algorithms" that might give the performance of an eight-bit converter without requiring that one actually be designed.

Did you ever notice that most mathematical theorems only apply in the limit as something approaches an impractical bound? This is so with the Nyquist limit. Yes, you can get a bandwidth equal to half the sample rate, if you record an infinitely long record. Since I cannot design infinitely long memories that run at state-of-the-art speed any better than I can design an eight-bit ADC, this becomes a limitation on the Nyquist criterion that must be endured. But, in the case of a practical data acquisition system, these two limitations work together. If I am willing to limit the bandwidth of my signal to, say, one fourth of the sample rate (solving the Nyquist criterion problem for short records) would this be of benefit to me in solving the seven-versus-eight-bit problem?

It turns out that the quantization noise (the noise that results from the finite spacing between code transition levels) in a high-speed acquisition system when a suitably high-frequency signal is present is distributed fairly evenly between dc and half the sampling frequency, and it doesn't depend very much on the nature of the signal as long as the signal trips a lot of code transitions. Well, since we have decided to limit bandwidth to one fourth of the sampling frequency anyway, why not filter out the noise that lies above one fourth of the sampling frequency (and beyond the signal bandwidth) with a digital low-pass filter. In a perfect world we could pick up half a bit because the noise power is halved. We might reasonably expect to produce a seven-plus-bit system from a seven-bit converter. True eight-bit systems cannot deliver the theoretical 50-dB dynamic range anyway, and rarely deliver dynamic range equivalent to seven and a half bits. A seven-and-a-half-bit system could be competitive, especially if it has four times the sample rate of the nearest alternative. So, what's the catch? The catch is, this only works if the input signal is moving around tripping the code transitions and generating some noise to filter out, that is, it doesn't work at dc.

Whence dither. What if we put in a known signal on top of the original input signal to be measured and then subtract it out after we digitize? Then even a dc signal will trip the codes and generate noise. And furthermore, what if the signal that we put in is out-of-band (above one fourth of the sampling frequency)? Then the bandwidth limiting filter would help remove it. Where in the stop band is the best place to put the dither signal? What is the best way to generate it? What is the best way to remove it? What about matched filters? What about adaptive filters? We could study this subject for years, and we have. The result is the HP 54720/10 oscilloscope, an eight-bit system from seven-bit parts. The real dither signal we use and how we get rid of it is proprietary, but I will describe in some detail one primitive form of dither.

Fig. 1 illustrates the concept in as simple a form as I know. Each tick on the vertical axis represents a code transition level, and the spacing between them represents the size of a seven-bit code. An input signal that stays between two code transition levels will produce a constant ADC output code. The top trace represents a slowly changing input signal, while the second trace represents the output of a conventional seven-bit ADC. (Each trace has been offset in this figure from the previous trace just to make the figure clearer.) Note that each time the input trace crosses a code transition level the output of the normal ADC changes states to a new code.

Now suppose we add, to the input, a signal that changes between minus one quarter of a code width and plus one quarter of a code width on every other sample. This will cause the ADC output to change when the input is within a quarter of a code width of the transition level. The third trace represents this signal.

one-fourth number is a compromise that seems to produce acceptable results for most users. There is an advantage to the one-fourth number in that the roll-off characteristics of the analog anti-aliasing filter we use for preprocessing does not have to be severely steep. This helps minimize the step response overshoot of the oscilloscope. Our design method for analog anti-aliasing filters is one of setting the bandwidth of the appropriate plug-in for its intended sample rate. We use an approximation to the maximally flat amplitude response of a second-order system with several extra poles in the response well above the two primary poles. This produces a system with a few percent overshoot. Contrary to popular belief, a smooth response with no overshoot does not produce the most accurate measurements of time-domain signals. Additionally, since we are limiting bandwidth to one-fourth the sample rate, the postprocessing digital filters provide additional anti-aliasing. How effective are these techniques in achieving repeatability? The article "Pulse Parameter Accuracy" on page 47 goes into this subject in detail, outlining several measurements of performance.

Linearity Correction

After the user changes a hardware setting, such as voltage sensitivity, the microprocessor must perform a multitude of tasks before acquiring data. Many of these tasks involve using the calibration factors to adjust DACs. One task uses the linearity characterization data from calibration to build ADC code correction tables. The calibration process creates equations and tables that describe the transfer curve of every part of the system: the probe, the plug-in, and the individual ADC signal paths. A table is built for each ADC path that inverts any distortions between the probe tip and the ADC.

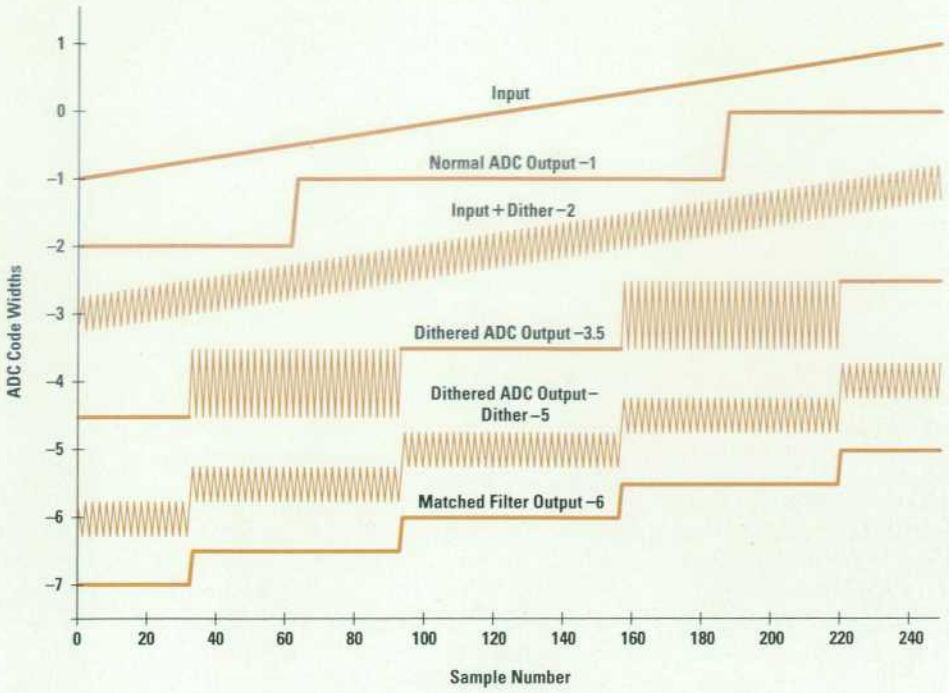


Fig. 1. Careful use of dither can double the resolution of an ADC by moving the error power out of the passband of the system where it can be removed by a matched filter.

Digitizing this new signal with a seven-bit converter produces the fourth trace. As long as the new input signal is completely bounded between two transition levels there is no change in the output, but when the signal is bouncing back and forth above and below a transition level, the output of the ADC bounces back and forth between two codes. Alternating neighborhoods of quiet and storm are produced.

Since we have prior knowledge of the dither component of the input signal, we can subtract a digital number representing this signal from the recorded ADC output. The result is the fifth trace. The amplitude of the bouncing codes has been reduced in the stormy neighborhoods because the subtraction process is out of phase in these areas, but a small storm has been added to the quiet neighborhoods. This signal looks, in the frequency domain, like suppressed carrier modulation about a carrier frequency equal to one half of the sample frequency, that is, the energy is above-band.

A simple filter matched to this dither is a two-point finite impulse response low-pass digital filter that just averages each sample with its neighbor. This filter rolls off gradually and has a three-dB bandwidth of one fourth of the sampling frequency. The output of such a filter is the bottom trace.

This now looks like an eight-bit converter. We speculated that if we simply filtered out the random noise above one fourth of the sample frequency we could pick up half a bit. This system seems to pick up a whole bit and all we did was add a simple signal and then subtract and low-pass filter. In practice this method will not recover a whole bit because the real ADC does not have uniform code widths, the dither signal cannot be so well-controlled, and there is analog noise in the system. Real systems employ more complex dither signals whose spectra are spread across the reject band and have adaptive filters matched to these spectra. The result when applied in the HP 54720/10 is preservation of the full desired signal bandwidth of one fourth of the sample frequency while delivering seven and a half effective bits at low frequencies and well above six bits at the specified bandwidth.

Careful use of dither can double the resolution of an ADC. By allowing a slowly changing signal to break up the pattern of the ADC's code regularity at a high frequency, the error power is moved out of the passband of the system where it can be removed by a matched filter.

Kenneth Rush
Engineer/Scientist
Colorado Springs Division

The correction tables are physically located on the acquisition board. As the microprocessor reads data out of the acquisition hybrids, the data passes through the appropriate "lookthrough" table in a pipelined fashion.

Intersymbol Interference Correction

After we had a prototype system working and fully calibrated, a disturbing problem remained. Differences in frequency response among the four ADC paths created a nonlinear behavior: the response of the system depended upon the phase of the four signal paths with respect to the input signal. Even though the problem was on the order of 1%, it had to be fixed. The problem was traced to two sources: intersymbol interference and path-to-path coupling. Intersymbol interference is the result of one sample (symbol) interfering with another sample by failing to settle. If a system has 1% intersymbol interference, it means that the input to the ADC is the sum of the current signal sample plus 1% of the previous sample, plus 0.01% of the second sample before, and so on. Path-to-path coupling results when a sample from one signal path couples into the sample of another signal path. Differences in layout and timing (such as 500-ps skews at 2 GSa/s) create unique errors in each path and thus a different frequency response for each signal path.

A tedious characterization process was developed to study the problem further. The process brought two important facts to light: (1) the errors are consistent from hybrid to hybrid and (2) a digital FIR (finite impulse response) filter can correct the errors. Since the response of each path is different, a unique filter is required for each path. Also, since the path-to-path coupling depends on the timing skew and thus the sample rate, separate filters were developed for each sample rate (500 MSa/s through 2 GSa/s). These filters

look at data for the path that they are correcting (intersymbol interference correction) as well as data from adjacent paths (path-to-path coupling correction).

Dither Extraction

There are two parts to dither extraction. The first part is to subtract the dither signal from the data. This step is done during calibration of the ADC linearity and is merely a matter of interpreting the ADC codes correctly. An ADC does not output volts (e.g., 23 mV), but codes. If code 14 is interpreted as 23 mV without dither, then in the presence of dither it might be interpreted as 22.75 mV for one ADC and 23.25 mV for another. Adding dither to the input of the ADC has the same effect as lowering the thresholds of the ADC and thus lowering the value, in volts, of the ADC's codes. Note that at this point in the processing, no resolution improvement has been made to the data.

The second part of the dither extraction is to remove the residual frequency components in the data. An adaptive filter acts on the data and removes high-frequency noise components. As a result of the dithering process, the lower the input frequency, the more high-frequency noise there is to remove. As more noise is removed, the signal-to-noise ratio improves and the system resolution improves! The resolution improvement of this technique depends upon the system noise levels, the dither technique, and the adaptive filter design. The dither-and-extract techniques in the HP 54720/10 improve the noise level by approximately 0.9 effective bit at low frequencies. At 2 GSa/s this improvement begins to drop off at 30 MHz and loses its effect significantly by 80 MHz. These results are for full-scale inputs; smaller inputs extend the improvement to higher bandwidths.

Waveform Reconstruction

Interpolation (waveform reconstruction) is the process of filling in missing data between actual samples of data. In an oscilloscope this is a necessary feature for two main reasons. At high sweep speeds, only a few samples of the signal can be on screen. The waveform must be reconstructed to give the user a reasonable view of the signal. Also, for many measurements it is necessary to know at what time the signal crossed a threshold, such as the 50% point on an edge. The system will almost never sample on the threshold that the measurement requires, so interpolation is necessary. The accuracy and repeatability of measurements are directly related to the quality of the waveform reconstruction. The interpolation filter is designed to optimize the frequency-domain and time-domain responses for a wide variety of signals. The bandwidth of $f_s/4$ (f_s is the sampling frequency) is required for consistent waveform reconstruction; it rejects interpolation "images" and provides some protection from aliasing (see "Filter Design for Interpolation," page 45).

Another benefit of this filter is that it removes some of the high-frequency noise in the system. At 2 GSa/s, the noise bandwidth of the filter is 740 MHz. It removes one fourth of the noise power from the system, which is an improvement of 0.2 effective bit. In practice, the improvement is not this large, since the adaptive filter that removes dither also removes much of this high-frequency noise.

The reconstruction algorithm in the instrument must work well for a wide variety of signals, but the user always has the ability to push the bandwidth closer to the Nyquist rate or to perform special-purpose filtering internally via IBASIC routines or externally by transferring data over the HP-IB or on a flexible disk.

Bandwidth Limit Filter

When the acquisition hybrid is sampling at 500 MSa/s, all four ADCs are sampling and digitizing the input signal in phase. Since the thresholds of the four converters are offset in voltage in 1/4-LSB increments, the system is essentially a 9-bit converter operating at 500 MSa/s if all of the data is saved. The logic on the acquisition hybrid is capable of saving all of this data, and this mode is in fact always used for equivalent time sampling where digital signal processing techniques are difficult to implement. We deliberately chose not to support this mode for real-time sampling, but instead provide a bandwidth limiting filter as a better solution for real-time acquisitions.

Increasing performance by one effective bit requires reduction of the system noise power by a factor of four. System noise levels prevent high-bandwidth flash ADCs from providing ideal performance without additional signal processing; accuracy and speed are at odds in such a system. High-bandwidth IC processes have a very limited range of operating voltages. Furthermore, large components are required to achieve accurate and repeatable values. As components are made larger, parasitic capacitance increases and the bandwidth is degraded. As a result, static errors such as resistor and V_{be} mismatches cut significantly into the accuracy of an ADC, even at the 7-bit level. As the target number of bits increases, resolution improving techniques such as dithering lose their effectiveness because the dither is washed out by the system noise.

With this in mind, it is not surprising that the "9-bit" mode just described improves the system performance by only one effective bit, even though it is technically a 9-bit ADC system. With an $f_s/4$ bandwidth after interpolation, the system bandwidth in 9-bit mode is limited to 125 MHz, with a Nyquist rate of 250 MHz. Furthermore, the memory depth is limited to 16K samples, or 32 μs, since four memory positions are used to store the four data values obtained for each sample.

If the system continues to sample at 2 GSa/s and a low-pass filter is applied, the effective bits are improved as the high-frequency noise is eliminated. An improvement of one effective bit requires that 3/4 of the noise power be removed, or that the filter have a noise bandwidth of $f_s/8$ or less. A filter with a bandwidth of $f_s/10$ (200 MHz) can easily achieve this goal with good signal fidelity. The important specifications for this system are: an improvement of one effective bit, a system bandwidth of 200 MHz, a Nyquist rate of 1 GHz, and a memory depth of 64K samples, or 32 μs. For a given improvement in effective bits, this approach is superior to the 9-bit mode for three reasons. First, the system bandwidth is higher by 60%. Second, aliasing problems are virtually eliminated since the Nyquist rate is five times higher than the bandwidth for the low-pass filter case, but only two times higher for the 9-bit mode. This significantly eases the design requirements of any anti-aliasing filter that might be required. Finally, the bandwidth limiting filter removes high-frequency noise contained in the user's signals, improving the signal-to-noise ratio further.

Filter Design for Interpolation

The interpolation process has some effects that are closely related to aliasing. Both types of effects are a result of mapping the s plane for a continuous time signal to the z plane for a discrete time signal. The basic problem is that mathematically there are an infinite number of mappings between the s domain and the z domain. Aliasing occurs when an incorrect mapping is used to interpret the discrete time data. For an oscilloscope, the mapping that is used assumes that the frequency content of the signal is between dc and the Nyquist rate. Any signal with significant frequency content outside of this range will have distortion when it is interpreted.

Interpolation is a process that attempts to transform discrete time data into continuous time data. Realistically, the process maps from one z plane to another, with a temporary stop in the s plane. When this transformation occurs, all possible mappings of the data are used. The results are low-pass filtered to preserve only the frequency content between dc and the original Nyquist rate.

Given these assumptions, the classic "brick wall" or rectangular-response filter will do a perfect job of rebuilding the waveform. Normalizing to a sample rate of 1 Sa/s, this ideal interpolation filter has an impulse response given by equation 1.

$$h(t) = \frac{\sin(\pi t)}{\pi t} = \mathrm{sinc}(t). \qquad (1)$$

It has a perfectly flat passband from dc to the Nyquist frequency, then crashes to zero across an infinitesimally small transition band, and has zero transmission throughout the stop band. This filter is infinite in extent and requires that an infinite amount of data be available for processing. Any filter that is not infinite in extent cannot have a perfectly flat passband, a zero-width transition band, or a zero-transmission stop band. Trade-offs must be made to create a usable filter.

Fourier analysis indicates that if all of the frequency components above any given frequency are removed from a perfect step, the response will ring on both sides of the step, with a maximum perturbation of 9% (Gibbs phenomenon). As more frequency components are included, the frequency of the ringing increases, but the amount of overshoot remains at 9%. Similarly, the frequency response of a brick wall filter shows the Gibbs phenomenon as the impulse response is truncated in time to make a usable filter. Since passband ringing of this nature is unacceptable for an oscilloscope, it is clear that a filter based on a truncated sinc function is also unacceptable.

Given the problems with implementing a sinc-based filter, what advantage is there that makes it worthwhile? The only benefit is that it does not modify the original data points under any circumstances. The characteristic of a sinc-based filter that provides this benefit is seen in the impulse response, where sinc(0) = 1 and sinc(n) = 0 for all integer values of n not equal to 0. When filtering an actual data value, this has the effect of weighting the current sample by 1 and all other samples by 0, that is, the output is equal to the input. All that is necessary to preserve the original samples is to mimic the behavior of sinc(x) whenever x is an integer.

Assume that an adequate filter has been created that has a reasonable response in the passband and does not modify the original samples. If a sine wave at the Nyquist rate is fed into such a filter, the interpolated output must be a clean sine wave at the Nyquist rate with no change in gain. Any other output would have significant distortion or would require the original samples to be modified. The point is that a signal at the Nyquist rate is not attenuated and the passband extends to the Nyquist rate.

To eliminate the ringing in the frequency response, the transition region of this filter must be gradual relative to a brick wall filter. Since the passband extends to the Nyquist rate, the soonest that the transition band of the interpolation filter can begin is at the original Nyquist rate. This delays the beginning of the stop band, creating large interpolation errors for signals around the Nyquist rate. For example, the Nyquist rate for a 2-GSa/s system is 1 GHz. If a 900-MHz sine wave is digitized, the sampling process creates an identical image at 1100 MHz that must be filtered out. For filtering purposes, 1100 MHz is only 0.14 octave from the 1-GHz Nyquist rate. If the interpolation filter has a transition region that drops off at 42 dB per octave and the filter response changes sharply from 0 dB per octave to −42 dB per octave at 1 GHz, then the 1100-MHz image will be attenuated by less than 6 dB! The reconstructed waveform shows an 1100-MHz sine wave added to the original 900-MHz sine wave, and the 1100-MHz signal is 51% as large as the 900-MHz signal. The 1100-MHz signal will cause distortion of the 900-MHz sine wave that repeats every 11 cycles, demonstrating that the system does not have a consistent response to the input; it is a time-variant system. Obviously this is undesired behavior for an oscilloscope and demonstrates two points. The first is that the reconstruction filter must be allowed to move the original samples because the passband cannot extend to the Nyquist rate. The second is that the filter must strongly reject all sampling images, implying that the filter attenuation must be adequate to do so at the Nyquist rate. This, plus the desire to have a gradual transition region, pushes the top edge of the passband well below the original Nyquist rate. The only alternative is to guarantee that the input signal does not have significant frequency content near the Nyquist rate, either by the nature of the signal (not acceptable for an oscilloscope) or by providing an anti-aliasing filter that has significant attenuation at and beyond the Nyquist rate.

There is a continuum of possible filters for waveform reconstruction. On one end is the Gaussian filter, which has the fastest rise time and settling time without overshoot. Although this filter is ideally infinite in extent, it can be realistically implemented with only one major drawback; a low bandwidth is required to achieve the desired image rejection. The brick wall, or sinc filter just discussed is at the other end of the spectrum. The reconstruction filter used in the HP 54720A oscilloscope is somewhere between these two extremes. It is designed to optimize the system response in both the time and frequency domains. Our experience shows that excellent results are obtained for a wide variety of signals by limiting the system bandwidth to $f_s/4$ (f_s = sample rate).

An additional benefit of an $f_s/4$ bandwidth is that the interpolation filter eases the requirements on the anti-aliasing filter (preamplifier). The frequency response of the interpolation filter relative to the signal input gets reflected about the Nyquist rate. At 2 GSa/s, assume that the filter has a 3-dB bandwidth of 500 MHz and is down 20 dB at 800 MHz. If the input signal is a 400-MHz square wave, then the sampling process will alias the third harmonic at 1200 MHz to 800 MHz. The digital filter can't distinguish the alias from a real signal at 800 MHz, so it will attenuate the image by 20 dB. Thus, the requirements on the anti-aliasing filter at 1200 MHz have been relaxed by 20 dB.

The interpolation algorithm should not only fill in a waveform, but provide a consistent (time-invariant) and linear response to the input. This requires the system bandwidth to be significantly less than the Nyquist rate. If a signal contains significant frequency content above $f_s/4$, then the $f_s/4$ system bandwidth provides the user with a display that is stable and consistent with the system bandwidth. Allowing the interpolation filter to modify the original samples is no different from allowing an anti-aliasing filter (which might be the preamplifier in a plug-in) to modify the analog signal before it is digitized. It is no different from allowing the amplifier chain of an analog oscilloscope to modify the signal before displaying it. In all of these systems, the signal is significantly modified only at frequencies near and above the system bandwidth.

Allen Montijo
Design Engineer
Colorado Springs Division

The bandwidth limiting filter in the HP 54720 oscilloscope has a Gaussian response and a 3-dB bandwidth of $f_s/20$, or 100 MHz at 2 GSa/s. It is capable of improving the effective bits of the acquisition hybrid by 1.6 at dc and 1.1 at full bandwidth. At 4 GSa/s and 8 GSa/s the 3-dB bandwidth increases to 200 MHz and 400 MHz, respectively.

Conclusions

A data acquisition system does not consist merely of a few pieces of hardware chained together. The overall performance of the HP 54720 system is the result of the synergistic behavior of calibration, signal preconditioning, and data postprocessing. Calibration maximizes system accuracy and minimizes system noise by correcting gain, offset, linearity, and timing errors. Signal preconditioning minimizes aliasing effects by limiting the system bandwidth to one fourth of the sample rate. Finally, data postprocessing accurately reconstructs the continuous signal from the discrete samples, and optionally provides improved resolution for low-bandwidth signals ($\leq$ sample rate/20). The dithering process combines these three techniques to further improve resolution for low-slew-rate signals.

IC processing technology places a very real limit on the useful number of quantization levels for a high-speed ADC; restricted operating voltages and component mismatches are significant, even for a 7-bit ADC. A carefully constructed dither signal, combined with digital signal processing, can improve system accuracy beyond the capability of the IC technology.

Acknowledgments

The authors would like to recognize Dana Johnson for his plug-in calibration code as well as the efforts of Marshall Boss, Kevin Loftin, and Derek Toeppen, who helped improve and verify the routines. Mike Van Grouw and his hybrid test system were critical in improving and verifying the ADC hybrid calibrations. Dave Dascher spent many tedious hours characterizing intersymbol interference and developing the compensating filters. Chris Duff's implementations of all of the digital filters were crucial to keeping the instrument throughput at an acceptable rate.

A Study of Pulse Parameter Accuracy in Real-Time Digitizing Oscilloscope Measurements

Using the well-characterized 50-GHz HP 54124T oscilloscope as a standard, HP 54720A oscilloscope errors were measured for single-shot step rise time, pulse width, and pulse height measurements. The results suggest that the errors have systematic or bias components that may be characterizable and correctable.

by Kenneth Rush

With the availability of the HP 54124T 50-GHz bandwidth oscilloscope and fundamental knowledge of its impulse response[1] comes a new era in time-domain metrology. Finally it is possible to know what a pulse signal looks like with enough precision to let that pulse signal become a transfer standard. In fact, many pulse signals can be characterized, making it possible to evaluate new real-time oscilloscopes based on the quality of their measurements. This paper describes a study that illustrates the power of knowing the "truth" about a time-domain instrument's response.

The Study

Eight steplike signal sources with rise times ranging from less than one nanosecond to greater than ten nanoseconds and eight pulse-like signal sources ranging in width from less than one nanosecond to greater than ten nanoseconds were constructed and measured by an HP 54124T oscilloscope and by an HP 54720A oscilloscope. These sources were Gaussian in basic shape, meaning that they smoothly moved from one state to the other with little overshoot and no abrupt changes in slope. The HP 54720A sampled the signals at 4 GSa/s and delivered a real-time bandwidth in excess of 1 GHz. Step rise times, impulse widths, and impulse heights were recorded. Prior characterization and calibration of the HP 54124T led us to believe that it would contribute less than 0.2% error in any of the measurements. All measurements were made using waveform averaging on the HP 54124T with the averages set to 256. All measurements made on the HP 54720A were made in the real-time mode with multiple records taken. The pulse parameter measurements were recorded for each record before the oscilloscope was armed for another single-shot acquisition. The pulse parameter measurements were evaluated for mean and standard deviation over the many different single-shot records. Over 1000 individual records were examined for each pulse parameter data point. Since the bandwidth of the HP 54720A is limited to about 1.2 GHz, it is expected to display fast signals as if they were slower than they really are. For the fast impulses, pulse height is expected to appear lower than reality.

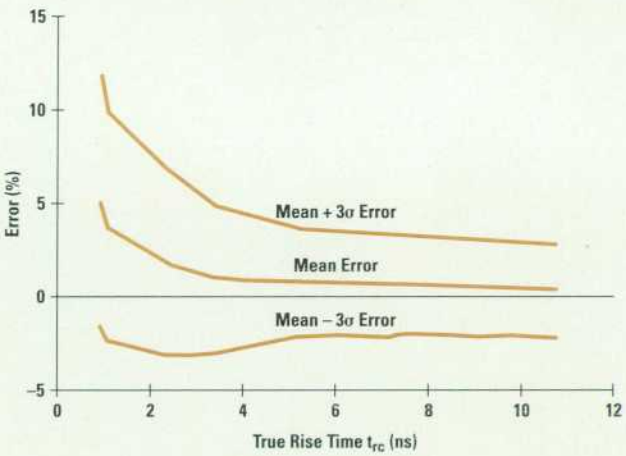


Fig. 1. HP 54720A oscilloscope single-shot step rise time error at 4 GSa/s. For the first time bias and variation in indicated 10%-to-90% rise time measurements made by faster real-time oscilloscopes like the HP 54720A can be measured. (t_{rc} is the true rise time as measured by an HP 54124T.)

Rise Time

In any high-sample-rate real-time recording device there will be significant noise and error because of imperfect anti-aliasing filters, loss of analog bandwidth because of the anti-aliasing filters, imperfect ADC quantizer alignment in time or amplitude, analog noise in the channel electronics, and other causes. The HP 54720A is no exception. Fig. 1 shows the rise time error of the HP 54720A measured in this test. Loss of bandwidth because of the anti-aliasing filters adds a bias to the rise time measurements. This error shows up as increasing mean error (average of many single measurements) as the input signal rise time becomes faster. The variation in observed rise time also increases with faster signals because of spectral leakage between different quantizers in the interleaved ADC structure. Before the HP 54124T the amount of error was unknown for faster ADC systems because there was no "known" pulse to measure. But now, we not only know the truth about our fast acquisition systems,

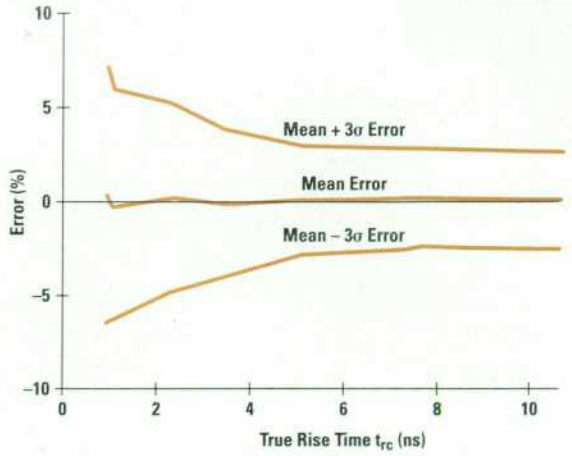


Fig. 2. HP 54720A bias-corrected rise time error at 4 GSa/s. Rise time measurement bias in the HP 54720A can be greatly reduced by correcting the readings for gain and offset. (t_{rc} is the true rise time as measured by the HP 54124T.)

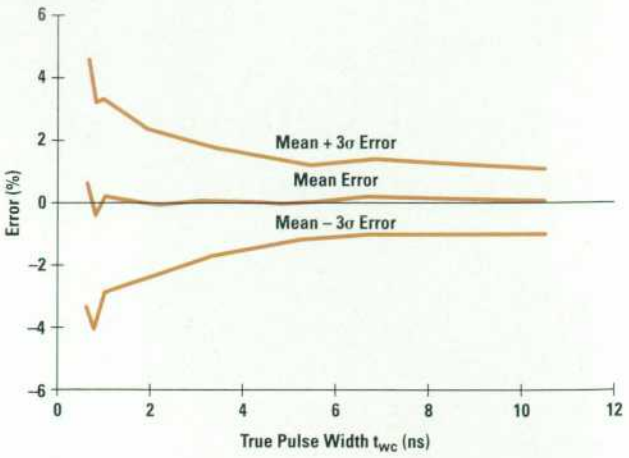


Fig. 4. HP 54720A bias-corrected impulse width error at 4 GSa/s. Bias correction equations using exponential terms can correct the HP 54720A's Gaussian pulse mean measurement accuracy to better than 0.5% and 3σ accuracy to better than 4% for most pulses.

we can also entertain analysis to correct errors. The bias in the mean rise time error of the HP 54720A can be corrected through a simple equation:

$$t_{rc} = 1.003t_{rm} - 45\text{ ps}, \quad (1)$$

where t_{rc} is the corrected rise time and t_{rm} is the measured rise time (see Fig. 2). In this case a gain correction of 0.3% and an offset correction of 45 ps are sufficient to bring the mean rise time error within one half of one percent for all steplike signals tested in this study. The bias is almost completely removed, and while the random errors are not removed, their effect is minimized by making them symmetric about zero. In this case a 3σ single-shot error of about 12% for a 0.8-ns signal has been reduced to about 7%.

Pulse Width

Measurements of several Gaussian pulses of various widths on the HP 54720A show similar bias and variation (Fig. 3), with less error for wider pulses and an abrupt increase in error as the pulses go below one nanosecond. The flatter response in the pulse width data indicates that the HP 54720A tends to maintain higher accuracy for narrow pulse widths than for quick step rise times. This tendency for pulse width errors to be less than rise time errors was observed by Draving[2] in 1991 in an idealized simulation study of time-domain measurements. He showed that maintaining a flat passband response was beneficial to the accuracy of time-domain measurements. It may be that maintaining a flat passband response is more advantageous for pulse measurements than for rise time measurements. Again, the bias in the data can be reduced by an equation:

$$t_{wc} = 1.0048t_{wm} - 48\text{ ps} + (0.18\text{ ns})e^{-t_{wm}/700\text{ ps}} - (15\text{ ns})e^{-t_{wm}/130\text{ ps}}, \quad (2)$$

where t_{wc} is the corrected pulse width and t_{wm} is the measured pulse width. Equation 2 is not as simple as equation 1. Exponential terms are required to correct the slope reversals and the abrupt rise for narrow pulses. With such extreme exponential corrections, obtaining smooth data for narrow pulses is difficult, indicating that from a confidence standpoint, corrected measurements for pulse widths below one nanosecond should be used with caution. Setting the coefficients of the exponential terms to zero leaves a 0.48% gain correction and a 48-ps offset correction. These terms alone put most of the mean errors below 0.5% for wide pulses, but there is a 1% dip at 1 ns and larger errors for even narrower pulses. With the full exponential corrections (Fig. 4), bias errors are reduced to less than 0.5% while making the random errors symmetric about zero, and 3σ errors as high as 15% are reduced to about 4.5%.

It is our hope that if these results are confirmed to be repeatable on many instruments, perhaps the correction equations can be invoked automatically by the oscilloscope's computer system before it reports the numbers. This is not done now.

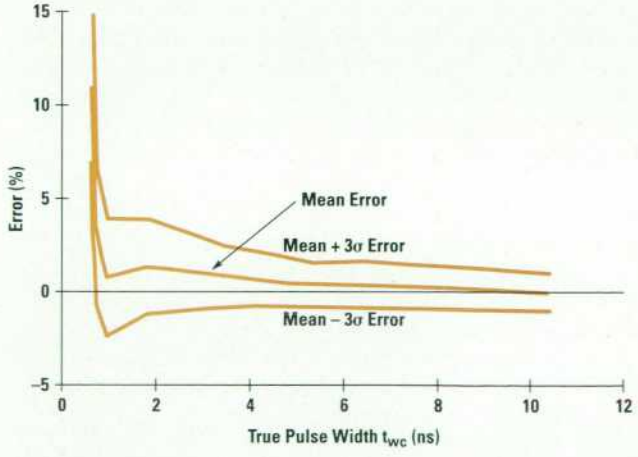


Fig. 3. HP 54720A single-shot impulse width measurement error at 4 GSa/s. Measurements of narrow Gaussian pulse widths on an HP 54720A show a dip and an abrupt increase in error below one nanosecond. (t_{wc} is the true pulse width as measured by an HP 54124T.)

Conventional Wisdom

An observation is in order here. Noting that for both Gaussian step and pulse signals the uncorrected errors increase rapidly below one nanosecond, and that correcting errors in that range is difficult and requires exponential terms, one might question the wisdom of using a one-gigahertz bandwidth

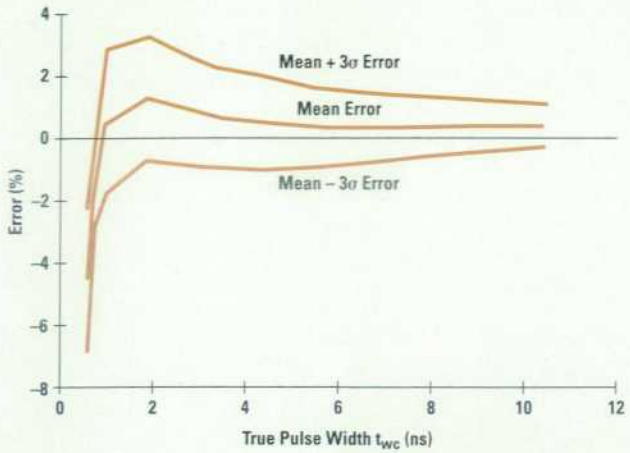


Fig. 5. HP 54720A pulse height measurement error at 4 GSa/s. Gaussian pulse height errors show a peak at 2 ns before falling abruptly below 1 ns.

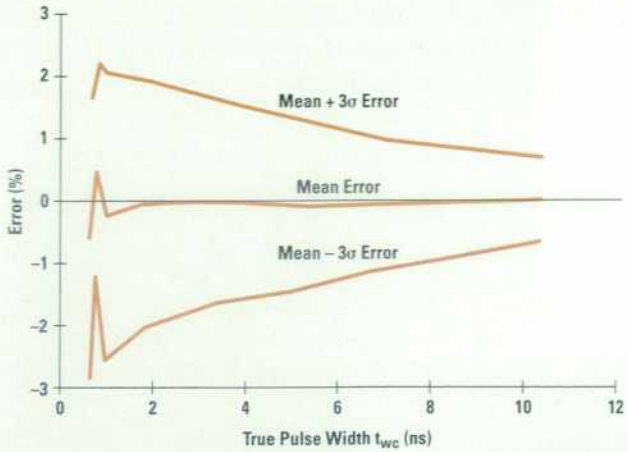


Fig. 6. HP 54720A bias-corrected pulse height error at 4 GSa/s. Bias corrections for Gaussian pulse amplitude measurements reduce the HP 54720A's bias to less than 0.5% for most of the range and to less than 0.1% for wider pulses.

oscilloscope to measure signals below one nanosecond. This reinforces the old rule of thumb that your oscilloscope should have about three times the bandwidth of the signals to be measured. A one-nanosecond rise time step signal has a bandwidth of about 350 MHz, using the relation $BW = 0.35/t_r$. Therefore, an oscilloscope with a bandwidth about three times that (1 GHz) would be expected to produce "reasonable" accuracy in a rise time measurement. Until now, "reasonable" has been defined by another rule of thumb based on a root-sum-square notion of how signals and oscilloscope responses convolve:

$$t_{r\,combined} = \sqrt{t^2_{r\,signal} + t^2_{r\,oscilloscope}}. \quad (3)$$

This is based on the assumption that both the signal and the oscilloscope have Gaussian responses, which is rarely the case. It also predicts a rapid increase in error as the signal rise time approaches the oscilloscope rise time. This rule has not proven to be a reliable assumption to use in removing bias from measurement data because it introduces its own bias through incomplete knowledge of the oscilloscope response. Draving's work[2] has brought into question the old assumption that an oscilloscope's response should be Gaussian or should have little overshoot anyway. While the Gaussian response would seem to have the fastest settling time for arbitrarily fast input signals it does not track a signal well whose speed is within the bandwidth of the oscilloscope (say 3:1 below). For signals that fall well within the bandwidth of the oscilloscope, flatness in the oscilloscope's frequency response passband is more critical. It is a little-known fact that Gaussian responses are achieved by gradually rolling off the oscilloscope's response below the 3-dB bandwidth. HP has designed its digitizing oscilloscopes to have flatter passband responses, striving for more accurate pulse parameter measurements instead of low overshoot in the step response.

Pulse Height

Single-shot pulse height measurement accuracy of the HP 54720A was also evaluated in this study. Like the pulse width measurements, the pulse height error plot shows slope reversals with a peak at 2 ns before going down below one nanosecond (Fig. 5). The peak indicates that the HP 54720A tends to overestimate the amplitude of pulses in the neighborhood of 2 ns while underestimating only when the pulses go below one nanosecond. The measurement bias and variation from one single-shot measurement to the next are less in percentage points than for the timing measurements. The bias correction equation in this case uses two exponential terms as well:

$$A_c = A_m\left(0.997 - 0.04e^{-t_{wm}/1.3\,ns} + 2.1e^{-t_{wm}/196\,ps}\right), \quad (4)$$

where A_c is the corrected amplitude, A_m is the measured amplitude, and t_{wm} is the measured pulse width. A 0.3% correction in gain is all that is needed for wider pulses, indicating the excellent low-frequency accuracy of the HP 54720A. The equation reduces the bias to less than 0.5% and the maximum 3σ error from 7% to 2.8% (Fig. 6). Again, the correction is a bit ragged at the narrow pulse end, emphasizing again that we are pushing things a bit to require precision measurements on signals faster than one nanosecond from a one-gigahertz bandwidth oscilloscope.

Unfinished Business

In any study of this kind it seems that there are as many new questions asked as there are old questions answered. While these results indicate that we may be able to correct for some errors in future oscilloscopes, do these corrections apply to other signal types, that is, does it work if the signal has overshoot? Are the oscilloscopes coming off the manufacturing line uniform enough that universal corrections can be made, or do we have to learn the response of each oscilloscope to each type of signal at each temperature? Lest we get too discouraged over all of these unknowns, let us remember that before now, we didn't even know how accurate a single measurement was. Yet now we have two different kinds of oscilloscopes agreeing on pulse parameter measurements to better than 0.5% at least for 3-to-10-ns signals. And we have cause to believe that one of those oscilloscopes is accurate within 0.2% for much faster signals.

If you find this subject interesting, then study it. Make some measurements. Report them in the journal of your choice. Time-domain measurements and standards have some

catching up to do, and it is a big job, bigger than one person or even one organization. Talk to your peers. Teach what you have learned. It is a quest worthy of your time.

Acknowledgments

Many have contributed to the growing knowledge base surrounding the HP 54120 family of digital sampling oscilloscopes, without which this study would have been in vain. Dave Long did the first signal analysis back before we understood much. Steve Draving carried the analysis to a conclusion in grand style through his unique combination of computer skill and measurement expertise. John Kerley added insight into the whole sampling process through his Spice modeling. But Jan Verspecht has carried mathematics, measurement science and data reduction to a new level with his discovery of many new error measurement and correction methods. I will always remember the all-night arguments over network reciprocity.

References

1. K. Rush, S. Draving, and J. Kerley, "Calibrating High-Speed Oscilloscopes," *IEEE Spectrum*, September 1990.

2. S. Draving, "Relationship between Waveform Recorder Frequency Response and Accuracy," *Proceedings of the Third High-Speed Digital Recorder Workshop*, Las Vegas, April 1991.

Architectural Design for a Modular Oscilloscope System

Optimum allocation of tasks to various software and hardware subsystems, a separate display processor, multiple lookthrough tables, flicker reduction techniques, and other design features support the performance of the HP 54720/10 oscilloscope and establish it as a platform for the future.

by Dana L. Johnson and Christopher J. Magnuson

Our vision for the new HP 54720/10 high-performance digitizing oscilloscope family was to design a new mainframe platform that would be both modular and easily upgradable and would provide uncompromised measurement performance. Modularity is an important concept that takes on two meanings. From the customer's standpoint, modularity means being able to select the right signal conditioning and measurement capabilities to solve a measurement problem. In the high-performance oscilloscope market especially, it is necessary to offer a configurable system to cover the wide range of measurement needs adequately. To us, as designers, modularity implies reusability and leads to economies in the development process. Therefore, modularity enables us to be more responsive to customer needs and competitive pressures. Upgradability is essential because over time a customer's needs change, and so do our capabilities. By designing the instrument to be easily upgraded, we protect the customer's investment. The instrument won't become obsolete. Instead, we can keep it current and useful and thereby reduce the customer's "cost of solution."

The measurement performance of the HP 54720/10 acquisition system is the cornerstone of the product. Every element of the mainframe design needed to be executed with an awareness of how it impacted measurement performance. This applied throughout the instrument's design from the calibration and control of the acquisition system to the analysis and display of signals. We needed to balance the broader requirements of the platform design with the specific needs of the instrument.

System Partitioning

Although Hewlett-Packard has been designing modular instruments for many years, we in the oscilloscope product development laboratory had not been building a modular product since the venerable HP 180 family. Therefore, we needed to establish a current view of how modularity was important to our customers. We reviewed our previous product experiences and studied carefully the successes and limitations of our competitors' offerings. These observations led us to develop a multitiered approach to modularity that influenced much of the product's design.

The context of the design was familiar. The oscilloscope observes input signals, and in response to the user's requests, it produces either data or images describing these inputs (Fig. 1). The oscilloscope's task can be decomposed into the four successive actions of signal conditioning, acquisition, analysis, and display. These four processes form the essential model of the oscilloscope and identify the principal partitions (Fig. 2). The flow of data through these processes is the instrument's signal path.

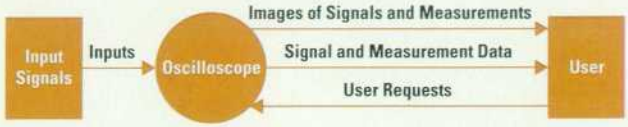


Fig. 1. Oscilloscope design context.

Signal conditioning is provided by the probe and plug-in options. These allow the user to make bandwidth, sample rate, memory depth, and feature trade-offs as needed. This modularity required that we implement run-time binding of the capability and control sets for the probes and plug-ins.

The mainframe can be configured at the factory with different acquisition systems by installing different cards in the internal cardcage. The Acquire Channel Signals process is a combination of this cardset and a corresponding software driver. Simple differences in acquisition, such as between the HP 54720 and the HP 54710, can also be handled via run-time binding. However, future systems may require installing a different acquisition driver altogether. In either case, the goal of the driver is to encapsulate the acquisition system and normalize its interface with the rest of the system.

Beyond the acquisition system the concept of modularity is important in the analysis of the acquired signals. By their nature, different acquisition systems produce fundamentally different representations of signals. For example, an acquisition based upon a real-time conversion system may yield data quantized to 8 bits that is evenly spaced in time. In contrast, through repetitive sampling techniques signals may be quantized to greater than 16 bits, and may not be evenly spaced. To reduce both signal representations to a common format would impose a data throughput, volume, or resolution penalty on the instrument. At the very least, the analysis system must be able to adapt to different classes of signal

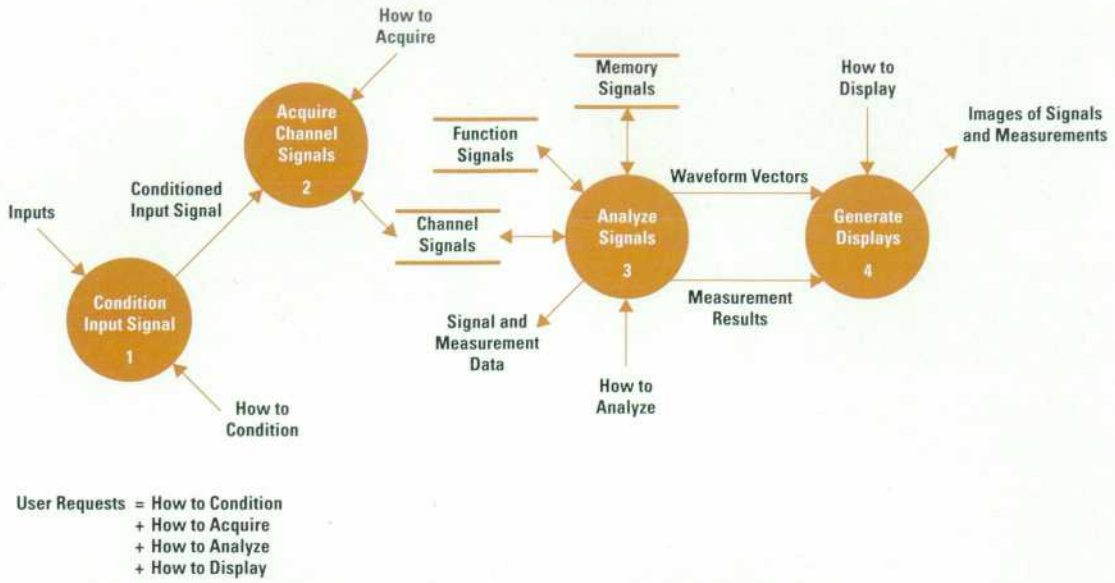


Fig. 2. The essential model of the oscilloscope has four processes that also identify the principal partitions.

representations. It must normalize the signal data by analysis method rather than by data type.

We also knew from the onset of the project that we would not be able to implement all the analysis capability that customers truly needed before the introduction of the instrument. Our customers continuously develop new measurement techniques or needs, which we would like to support in our instrument. This indicated a second dimension of the modularity of the analysis system, one that allowed for addition of new capabilities for analysis and presentation of data.

Top-Level Design

The next major issue was how to allocate functionality between hardware and software. The signal path is the essential flow of data through the system. Ideally, it would be continuous and real-time. A system completely implemented in hardware, such as an analog oscilloscope, can approach this ideal performance. In our system, of course, basic signal conditioning and acquisition had to be hardware-based, as did the ultimate generation of the display image and the data outputs. Therefore the question was, given the constraints on development time and cost, and considering the key contributions we wanted to provide, how should we allocate implementation between hardware and software through the rest of the system? We prioritized the following throughput performance goals as criteria on which to base the design:

- > 2 Mpoints/s DMA capability via a high-speed port when used as a digitizer
- > 250 kpoints/s display update rate with intensity-modulated persistence
- > 500 kbytes/s block transfer rates via the HP-IB (IEEE 488, IEC 625)
- > 100 measurements/s
- < 1 ms for command execution via the HP-IB (< 20 ms with attenuator change)
- < 100 ms total response time to a knob change
- < 5 s turn-on time.

This led us to the allocation of hardware depicted in Fig. 3. In this figure, bubbles 2, 3, and 4 from Fig. 2 are shown in greater detail. For example, bubbles 2.1, 2.2, and 2.3 contain the functionality of the Acquire Channel Signals process in Fig. 2. A Motorola MC68020 microprocessor serves as the host processor. It performs the majority of the system's signal processing and control. Certain specific tasks were offloaded to specialized hardware to allow the performance goals to be achieved.

For example, the acquisition process is subdivided into three components. Acquire Input Signal is the hardware state machine that collects the sample data record. Construct Channel Signals is the software driver that operates the acquisition system. Calibrate Data is a hardware data mapping function that efficiently translates uncalibrated data to calibrated values. This calibration entails evaluating each sample point using a third-order polynomial that represents the transfer function of the input system, after first correcting for any nonlinearity in the ADC. Traditionally, a lookup table has been precomputed as an optimization so that this complex expression does not have to be evaluated for every data point. However, to achieve the 2-Mbyte/s throughput via the high-speed port, even this was not sufficient. Instead, we implemented a lookthrough table, so named because the host processor sees the output of the acquisition system through the precomputed mapping function. As they are read, the uncalibrated acquisition data points form the address to the mapping memory. The output of the mapping memory is applied to the data bus of the host computer. The host memory system is designed to provide single-bus-cycle memory transfers between acquisition, host memory, and the I/O channels using the MC68020 as a DMA engine. The result is that the MC68020 can read a data point from acquisition memory, calibrate it, and store the corrected value in host memory in just one bus cycle. If high-speed port transfers are desired, the port can be addressed to eavesdrop on the data transfer without adding any additional clock cycles. This allows a blind transfer mode that can achieve greater than 5-Mpoint/s peak transfer rates.

Display generation was identified as a throughput bottleneck by performance testing our previous products. In these products, this process was performed by the host processor.

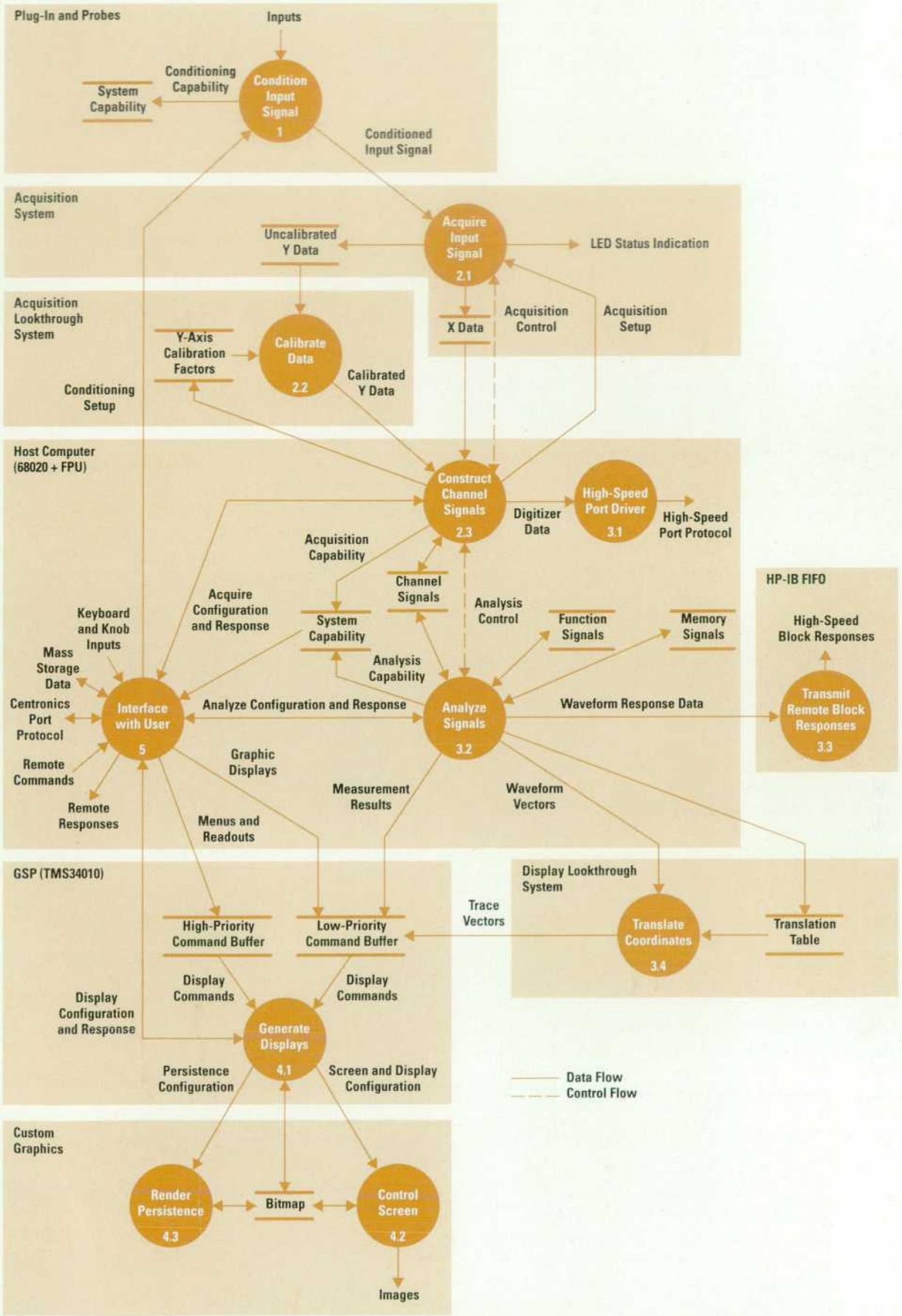


Fig. 3. Detailed flow diagram of the HP 54720/10 oscilloscope, showing processor allocation.

Our estimates indicated that to generate the trace quality we desired, including variable-persistence display, the display task would consume too much of the host processor's bandwidth. We elected to use a dedicated display processor, the Texas Instruments TMS34010, in conjunction with a custom state machine to perform fast line and area operations. This display subsystem operates under its own operating system. It executes graphics primitive operations according to commands placed by the host into the low-priority and high-priority command buffers. Partitioning the system in this way frees the host processor to acquire more data or to perform signal analysis while the display processor is rendering the trace image. The display subsystem is described in further detail later in this article.

The graphics primitives for trace displays are designed so that the host processor performs the coordinate system transformations. Voltage values are translated by the host into y-axis bitmap coordinates before being passed to the display processor. To speed this computation, a second lookthrough mapping system was devised. This allows signal data points to be translated from their internal representation to the coordinates of the display screen in a single bus cycle.

Hardware support was also needed to meet the HP-IB block transfer goals. Traditionally, the HP-IB data output process is implemented as a software loop in which the host processor transfers data to an HP-IB controller chip on a byte-by-byte basis. To an infinitely fast HP-IB listener, this method could come close to meeting the goal of >500 kbytes/s. However, to a slower receiver (on the same order of performance as the instrument) this technique would be limited to 200 to 300 kbytes/s because each device has only a portion of the total transfer cycle to accomplish its task. To alleviate this, we added a FIFO memory in front of the bus controller chip. The FIFO allows a looser coupling between the source and the receiver such that each has the full transfer cycle (or more over the short term) for its operation without impacting the other device.

Host Processor System Architecture

Because the host MC68020 processor is a key element in the signal path, it was important to architect the host processor system to maximize the time available for acquisition and analysis. The other work done by the host is primarily servicing the user. The work the host performs is event-driven in either case, but the event rates are quite different. The signal path is driven by trigger events derived from the input signals and state transitions within the acquisition system. Maximum event rates are in the range of 100 kHz to 1 MHz. In contrast, user events from the front panel are limited to rates less than 100 Hz. At the remote interface, 1-kHz event rates are a practical goal. By the criterion of temporal cohesion,* host processing was decomposed into the tasks shown in Fig. 4.

This approach to task decomposition caused the design team some confusion at first. We had been using structured analysis methods to model the system, which led us to a functional view of the system. This worked well to help identify major abstractions and objects in the system. However, the partitions of this composition did not map directly onto the partitions of the task model. An example of this is the software driver for the acquisition system. Although it deals with the control, test, and calibration of the acquisition system, portions of it run under the signal path task, others under the local or remote tasks, and some under all three tasks. We found the same considerations when partitioning the signal analyzer.

The signal path task has the lowest priority in this system. It runs continuously while enabled, driven by the status of the acquisition hardware. In this case, polling is used instead of an interrupt-driven scheme. This eliminates the overhead of interrupt service and possible additional operating system calls. All other tasks in the host system are interrupt-driven and may preempt the signal path.

The local and remote user interface processes are peers. They parse the corresponding user input streams and share control of the system via mutual exclusion. Actual execution of user inputs is performed by common execution routines. The processes are designed so that the remote task can hand off certain jobs to the local task for overlapped processing. This enables us to provide programmatic control via the remote command language over certain complex tasks such as calibration. Collectively, these user interface tasks control whether the signal path is enabled and what job it is to perform.

Plug-in communication has unique requirements. This communication is over a serial interface and uses a special protocol for read and write operations. Certain write operations to a plug-in may cause that unit to become busy, during which time no further writes are permitted. Also, when attenuator motors are energized, substantial drive current is required. This load current must not exceed the amount budgeted for the system's power supply. Because of these considerations, the module interface task is designed as a command spooler and low-level driver for the plug-ins. It contributes to the system's fast response to user inputs because it allows portions of the input processing to occur concurrently with the control of the plug-in hardware. Each plug-in has a command queue into which the user interface places hardware control instructions or other programming information. The module interface task empties these queues into the plug-ins as quickly as they will accept the data, while monitoring the amount of power in use. Control commands are not issued until sufficient power is available. The module interface also interprets plug-in system interrupts such as probe changes or key presses by validating them and translating them to standard user interface events for parsing by the local user interface task.

Interprocess communication is designed to minimize overhead in operating system calls. To reduce overhead further, we made the system clock period 100 ms. This ensures that system time management is an insignificant factor in operating system overhead. At each clock time, the operating system evaluates its ready list and may perform a context switch. At this clock rate, we determined that this operation would consume less than 0.1% of the available processor cycles. However, finer time resolution is needed for tracking hardware settling times and managing acquisition schedules. For these purposes, a pair of hardware timers are provided

* Temporal cohesion is a measure of the "time-relatedness" of functions in a system. If several functions need to execute in response to an event they exhibit a high degree of cohesion.

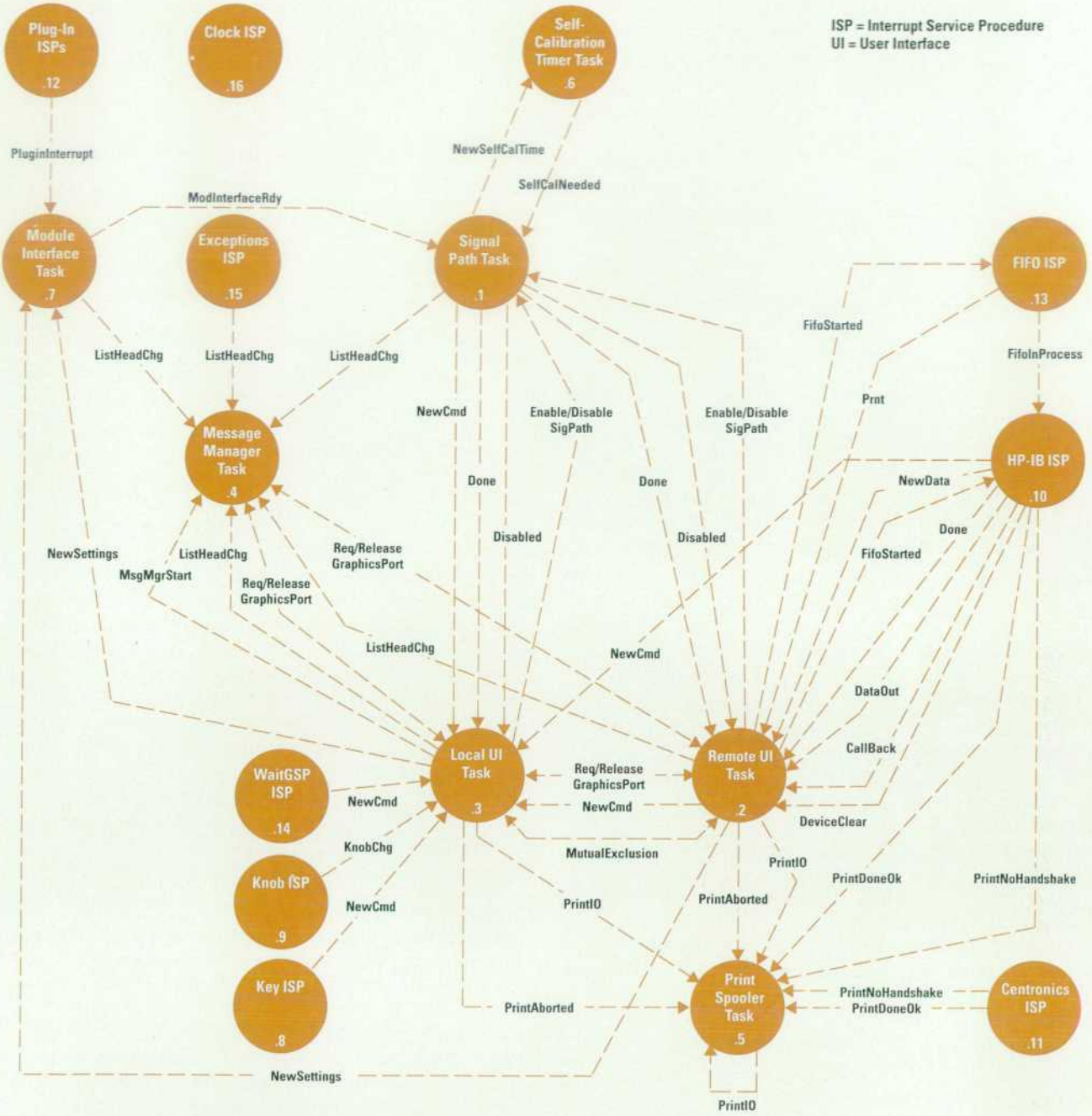


Fig. 4. Host processor tasks in the HP 54720/10 oscilloscope.

that allow timing resolutions as fine as 120 ns and durations as long as 6 s. The system clock is generated by a real-time clock which provides time and date functions as well as some power supply monitoring capability.

The settling manager is a utility function that uses one of the timers to track hardware settling times for each of the instrument's input paths and for other functions such as the system calibrator. When hardware changes are programmed by any of the tasks, they call the settling manager to record the needed settling time for the affected channels. Before an acquisition is started, the signal path task calls the settling manager to verify that the selected channels have stabilized. In previous systems, this function was implemented with microprocessor timing loops or operating system calls. Through the use of the settling manager, the instrument can ensure path settling concurrently with other processing. This design also guarantees proper operation in the future as we upgrade the CPU to use faster clocks or different processors.

The front-panel keyboard and knob are managed by a keyboard scanner and RPG (rotary pulse generator) encoder logic. This hardware is encapsulated by the key and knob interrupt service procedures, respectively. This hardware eliminates the need for the host processor to continuously monitor the keypad or measure knob position. Instead, when a change occurs, the interrupt service procedures issue standard key events for the local process to parse. Knob response is further enhanced by using a high-resolution RPG coupled to an up-down counter. The counter accumulates the net position change since it was last read. This technique eliminates backlash effects often seen in systems using

other position sampling techniques and yet provides fine sensitivity to minute changes. The knob interrupt rate is limited to a maximum of about 25 Hz to prevent overloading the system with change events during continuous slewing operations while still providing lively response.

Signal Path Design

As mentioned previously, the signal path is composed of the acquisition driver and the signal analyzer. These two components are tightly coupled within the signal path task, but they operate at significantly different abstraction levels.

The acquisition driver encapsulates the acquisition hardware and the construction of the signal data structures representing the input channels. It must follow a schedule of acquisition jobs because the acquisition system cannot necessarily perform acquisition for all channels simultaneously. For instance, if a single-slot plug-in has two channels, such as the HP 54714A, the channels must be acquired on alternate triggers. In this case, the acquisition driver first acquires channel A and then acquires channel B. A ring was chosen as the abstraction model for the acquisition schedule since it maps well to the essential model of the system (see Fig. 5).

A ring is a circular, doubly linked list. It has operations to add and delete items, and can be rotated to position a different item at the index point, known as the "top." The ring is called the scan ring, and ring items are scan stations. The scan station at the top of the ring is the one that specifies the next acquisition. Each scan station is a data structure that describes what to acquire, what acquisition method to use, exit criteria for terminating the acquisition method, and any hardware changes required. This allows a great deal of flexibility in configuring the acquisition process and is well-suited to supporting the diverse requirements of different acquisition systems as well as feature enhancements over the lifetime of the platform.

The acquisition driver rotates the scan ring, executes the acquisition for the scan station, and then calls the signal

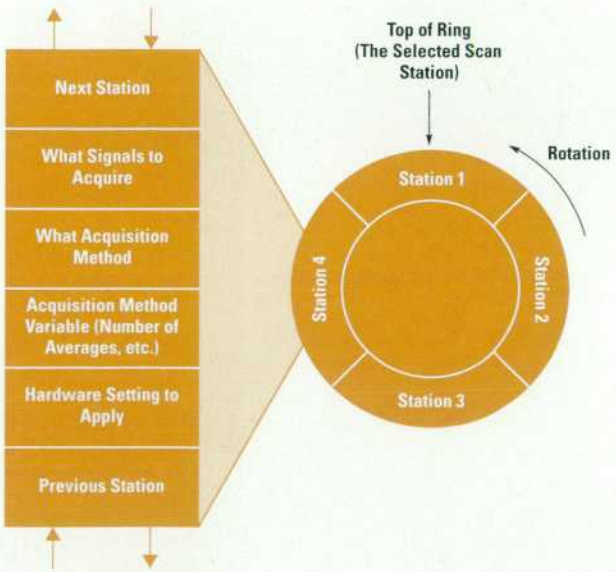


Fig. 5. The ring is a circularly linked list. It provides a good abstraction of the repetitive acquisition schedule of the oscilloscope.

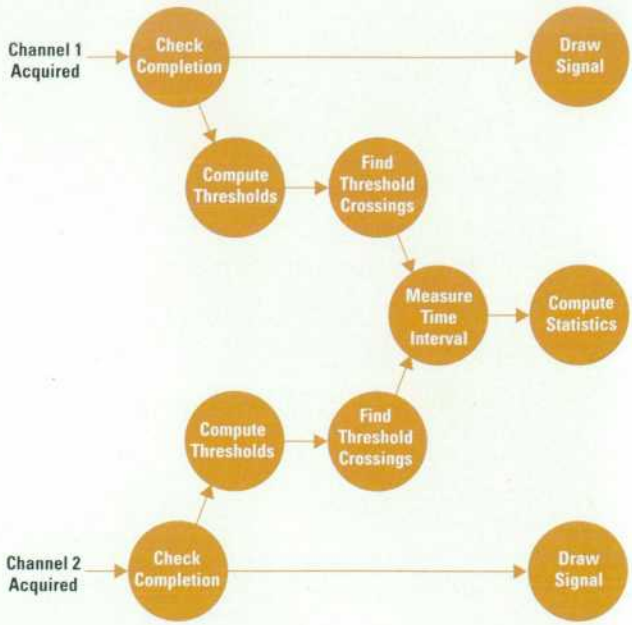


Fig. 6. A graph depicting analysis functions when two channels are acquired and a time-interval measurement is performed. The vectors show dependencies and the bubbles are processing steps.

analyzer to process each newly acquired signal. The ring is constructed based upon the channels that are selected directly or needed as operands for waveform math functions. Other considerations in construction include what signal processing is selected (such as averaging), whether real-time or equivalent time sampling is to be used, and so on. The ring construction function is part of the library of utilities the acquisition driver provides for the user interface so that it can optimize the acquisition schedule as dictated by the capabilities of the currently installed hardware.

The signal analyzer processes "normalized" signal data structures and does not need to know about the hardware that generates them. It too has a schedule of actions it must perform when it is presented with a newly acquired signal. These actions are modeled as a graph in which vertices are atomic operations and edges represent dependencies (see Fig. 6). The arrival of a new signal fires a series of actions through the graph. Some actions, such as a math operation to sum two channels, have multiple inputs that must be satisfied before they can fire. This type of action would be armed by the arrival of the first signal and fired on the arrival of the second. This model of the analyzer is attractive because it provides an effective way to optimize signal processing and is very extensible. In implementation, the graph is represented by an ordered list of the atomic operations.

Like the acquisition driver, the signal analyzer provides the construction methods for its scheduling. Each operation that can be selected has a description of its constituent operations. This description is efficiently held as a tree. When a measurement is requested, such as rise time, the leaves of the tree are visited and inserted into the list of operations according to a preferred execution order. This scheme allows new measurements to be added simply by defining the constituent operations. This design is very attractive to the implementor because only new atomic operations need to be coded.

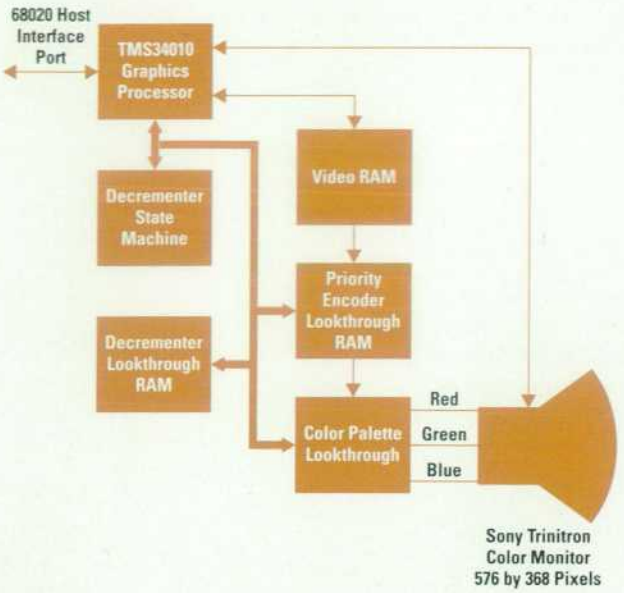


Fig. 7. Display processor system block diagram.

The signal analyzer also uses a y-axis lookthrough table when writing data to the display system via the low-priority buffer. The table is constructed beforehand to allow the rapid translation of y-axis values to pixel coordinates on the display, obviating the need for software scaling of the display data. By implementing this capability in a hardware lookthrough table the traditional software overhead associated with y-axis display data scaling is eliminated.

Display Processor System

The display processor technology was developed especially for this project. The display processor system basically consists of a series of lookthrough tables, a 576-by-368-pixel color monitor, 1M bytes of video RAM (VRAM), and a pixel-decrementing state machine all working in conjunction with a Texas Instruments TMS34010 graphics processor (Fig. 7). The TMS34010 provides control signals needed by the color monitor and for the VRAM refresh cycles. The host 68020 processor interfaces to the display system via an onboard host interface port, which acts as the portal for passing commands and data between the host and display processors. The specialized display processor system frees the host processor to concentrate on analysis and acquisition of the user's waveform data, while the low-level graphics primitives are handled by an optimized color raster display system.

The VRAM is 16 bits (one word) wide, and is addressable by the TMS34010 processor on a pixel or word basis. The VRAM serves as the storage space for the waveform pixels as well as program and data memory for the display system software. The pixel memory is also stored in VRAM. A specialized use of the VRAM is the allocation of memory buffers used for communication between the host and display systems.

The host and display processors are tightly coupled by two VRAM buffers and a series of pointers into those buffers, which are updated by both the host and display processors. One of the buffers is large and is for lower-priority commands and data, while the other buffer is small and is used for high-priority items. Both the host and the display processors must maintain the pointers to tell where to fetch and place commands. The commands placed in the lower-priority buffer are wrapped with a framing protocol, which improves display efficiency in cases where the host is sending data at a higher rate than plotting can occur (typically in excess of 500,000 points/second). In cases when this display bottleneck occurs, the framing protocol allows the display processor to find the most recent frame of data quickly, maintaining the responsiveness of the user interface.

The color palette lookthrough table is a 256-color table that outputs an RGB (red, green, blue) color combination based on a bit combination presented at its input. The output RGB combination provides a pixel of the desired color on the color monitor.

The key to the functionality of the display software is the manner in which the bits in the 16-bit VRAM word are allocated. The word is divided into several subfields consisting of from 1 to 3 bits. Some subfields are dedicated to waveform data and other subfields are multiplexed to handle multiple functions, depending on which bits are set. The waveform data subfields hold a value that eventually points to a color table entry in the color palette lookthrough table. The color palette table then outputs an RGB analog signal combination that is displayed on the CRT at the pixel location in the appropriate color. Before getting to the color palette table, portions of the pixel value are routed through the priority encoder lookthrough table, which can potentially change which color palette table entry is used.

The priority encoder lookthrough table acts as an intermediary between the pixel stored in VRAM and the color displayed. For example, if certain bits in the pixel are set, the system will display the color of a waveform memory at a given pixel instead of the color of the graticule. The priority encoder lookthrough table effectively redirects a pixel and can make it point to a different color, avoiding any software comparison operations that may be needed to accomplish the same task.

The pixel decrementing state machine shares the system bus with the TMS34010, stealing cycles as needed to accomplish pixel modifications (decrementing) during the pixel decrementing process. Decrementing consists of examining the value stored in the VRAM for a given pixel and using subfields of the pixel to point to locations in the decrementer RAM lookthrough table, effectively using the current pixel as an index into the RAM to find a new pixel value. In this manner, subfields in a pixel word can be modified by the decrementer to a different value which can signify time decay of the pixel intensity. This method is the basis upon which the pixel intensities can be varied to achieve the variable-persistence function of the system. The rate of intensity decay is determined by counting the number of interrupts generated by the horizontal blanking interval of the display and then deciding when to run the decrementer during each interrupt service call.

Some clever manipulations of the pixels during the plot and erase cycle are employed to reduce visual flicker of the display. Flicker often occurs when large changes in the display intensity are allowed to occur, and the visible intensity modulations can induce fatigue and eyestrain on the part of the

user. In variable-persistence modes, whenever the decrementer is going to be applied to the screen pixels, the software initiates a series of decrementing activities, decrementing 1-by-2-pixel areas in many different screen locations at nearly the same instant. By distributing the area of the entire screen to be decremented into many smaller areas the amount of display flicker is greatly diminished.

Another mode in which flicker reduction methods are employed is during the use of minimum persistence. In this mode the pixels are plotted invisibly and are changed into visible pixel values when the decrementer is run. The waveform pixels that were illuminated on the screen before running the decrementer will stay on the screen for one more decrementing cycle. The combination of the invisible plotting and overlapping of the previous and new waveforms combines to reduce flicker appreciably in minimum-persistence mode.

In addition to the decrementing capabilities of the decrementer state machine, the decrementer also generates x-axis plot coordinates when programmed to do so. By autoincrementing in hardware the x coordinate for each y coordinate sent by the host processor the software data stream is effectively halved, further reducing the software overhead under plot conditions involving equally spaced x-axis data points. The modes of the instrument having high throughput, such as the real-time signal capture mode, can greatly benefit from this feature.

As important as the display software is to the function of the system, the reduction of software run-time overhead through the use of the specialized hardware is the main benefit of the display system. The decrementer state machine eliminates the need to handle large numbers of pixels in software loops, and the autoincrementing nature of the decrementer and the pixel lookup and decrement operations minimize software data handling. Finally, the y lookthrough RAM used by the host software analyzer eliminates the task of y-axis scaling, enabling the display system to plot the data without scale processing. In the aggregate, the specialized hardware of the display system together with the display software produces a high-throughput system optimized for high-quality presentation of waveforms.

Acknowledgments

Special thanks should be given to Mark Heim and Dave Long for their work on the display processor system. Thanks also to Allen Montijo, who designed the display hardware, and Mark Lambuth, who helped maintain and improve the design. Bruce Schwartz and Allen Montijo worked on the main CPU development. Rick Hollon and Bdale Garbee also contributed to the CPU architecture in the areas of software and hardware feature specification and development.

A Survey of Processes Used in the Development of Firmware for a Multiprocessor Embedded System

In using structured design methods to develop a large multiprocessor embedded system, the HP 54720/10 oscilloscope design team learned that these methodologies can be very helpful if applied appropriately and supplemented with a few other processes and tools.

by David W. Long and Christopher P. Duff

The goal of the firmware design team for the HP 54720/10 oscilloscope was to develop an oscilloscope platform for the future that augments the basic acquisition hardware's capabilities and provides users with easy access to the answers they need. To have a platform for the future, the team felt the firmware needed to be easy to maintain and easy to modify so that new functionality could be added over time. While past products had limited feature sets with firmware primarily focused on acquisition control, the firmware in this product family was to add value itself by helping to make the user's job easier. Finally, for the instrument to be user-friendly, it needed to be fast, efficient, and easy to learn to use. The controls needed to be very responsive and the waveform display needed to react instantaneously to changes.

The first decision the team faced was whether to try to revitalize the existing high-performance oscilloscope mainframe or to start over. The computer was slow by current standards and not well-optimized for algorithms run in an oscilloscope. The firmware, primarily written in assembly language, had been designed with the intent that it would be used only once and was hard-coded in numerous places to support a specific set of hardware. Since it was not well-documented, it was difficult to modify or even maintain. All of the above facts made reuse of the existing system inconsistent with the goals the team had set for the platform of the future.

The team decided to start over. Computer science had advanced a great deal. Better processes and tools had been developed and the team could take advantage of lessons learned from the development of previous generations of oscilloscopes. Furthermore, to meet the throughput goals, a new computer was needed that was better optimized for running oscilloscope algorithms.

Even though significant pieces of firmware, such as an operating system, an MS-DOS® file system, and an HP-IB command parser, were available from other HP divisions, meeting the project's schedule goals was going to require process improvements. The team chose to use a combination of modern structured methods and other best practices. Table I is a summary of the tools and processes used during each phase of the project. While the use of an organized approach to system development certainly did pay off, the team learned quite a few valuable lessons along the way.

Table I
Firmware Development Tools and Processes

Project Phase	Tools and Processes
Definition	Focus Groups Usability Tests Performance Analysis of Existing Products Data Flow Diagrams Front Panel/User Interface Simulation
Design	Structure Charts
Implementation	Revision Plan Coding Standards Code Reviews Configuration Management System Compilers, Assemblers, Emulators, etc.
Performance Verification	Usability Tests Manual Firmware Abuse Tests Automated Regression Tests

Definition Phase

Early in the project definition phase, a series of focus groups were held in cities across the United States. An independent consulting firm was hired and participants were carefully selected who had varying levels of experience using a variety of brands and models of oscilloscopes. To prevent any bias in the results, participants were not told who was sponsoring the focus groups. One of the consultants, acting as a moderator, led each group through a two-hour session discussing a predefined list of topics. The design team's goals were to learn what aspects of an oscilloscope are or are not important to users, what constitutes quality, and what constitutes ease of use. We wanted insight into why people chose to buy or use a specific instrument. The focus groups were videotaped so that all team members could see the sessions in addition to reading the summary information provided by the consultant.

The focus groups provided key information for making early design decisions. For example, we learned that a touchscreen user interface, while highly desirable in other instruments such as logic analyzers, was not optimal for an oscilloscope. Further, we found that participants judged oscilloscope quality by the kind of tactile feedback they received when pushing keys and by the feel and resolution of knobs. The tactile feedback, which doesn't exist with a touchscreen, gives users confidence that they have, indeed, activated or changed the intended control. The responsiveness of the display to the controls was also important, as was the overall quality of the waveform display. Based on the information we obtained from the focus groups, we decided to use a traditional button-and-knob user interface combined with a high-resolution color display. The focus groups also led us to concentrate on high throughput and improving overall display quality to emulate more closely the displays found in analog oscilloscopes.

While the focus groups provided some general information about user interfaces, we needed more details about what makes an interface easy or hard to learn and use. We decided to run a series of usability tests on a variety of existing oscilloscope interfaces from various vendors. To ensure valid and unbiased results, independent human factors consultants with usability testing expertise were chosen to conduct our tests. Again, test participants were carefully selected to provide a representative cross section of real-world oscilloscope users. They were each given two hours to work through a set of tasks on one oscilloscope, then two hours to work through the same set of tasks on another oscilloscope. Different test subjects tested different combinations of instruments, but all attempted the same tasks. They were not told who was sponsoring the tests. The tests were videotaped for later review by the design team. Statistics were kept on how long it took each subject to complete each task and how many assists were required. After the tests, the subjects were interviewed to get their opinions about the instruments they tested.

The usability tests turned out to be very helpful to the design team. The tests showed that the usability of the existing HP oscilloscope user interface was similar to or better than competing alternatives, although there were significant opportunities for improvement. Many of the consultant's recommendations were implemented. An example of one of the more interesting findings pertained to online help. Most existing help utilities put the instrument's front panel into a help mode. When the user selects a control, the instrument provides a textual description of how to use that control. The usability tests showed, however, that users generally didn't have problems using the controls they could find. Their biggest problem was simply finding the controls they needed. As a result, the HP 54720/10 includes an alphabetized index under the Help key that tells where each control is located within the menu structure. There were many other suggestions for how features might be made easier to use, how to improve menu structures, and how to improve the front-panel layout. Several of the findings were also very similar to those from the focus groups. Examples included recommendations against using a touchscreen and in favor of using a color display and reducing menu depth.

In parallel with the focus groups and the usability tests, the firmware design team was analyzing system requirements. They focused on the digital system and firmware architectures. They had the freedom to partition functionality between hardware and firmware, as necessary, to meet their performance objectives while at the same time staying within a set of cost goals.

The first step was to analyze existing products to identify their performance bottlenecks. Emulators and logic analyzers were used to measure how firmware execution time was spent during critical tasks. Key bottlenecks included floating-point math, the display of numeric text including the process of converting the numbers to ASCII strings, and waveform display. In response to these findings, the HP 54720/10 computer design includes a floating-point processor, which not only speeds up real math but also number-to-ASCII-string conversions. The display system includes a TMS34010 graphics processor along with a custom graphics coprocessor, which speeds up waveform display.

Data Flow Analysis

The next step was to develop a model of the new system using data flow diagrams.[1] Data flow analysis, formally known as structured analysis, is intended to help engineers evaluate and specify system requirements by modeling the flow of data through the system as well as the transformations performed on that data. Emphasis is placed on exploring what the system needs to do rather than how it will do it. No assumptions are made about what is done in hardware versus firmware. It takes a little time to learn to think in this manner. It is easier and perhaps more fun for engineers to think about how to solve a problem than to fully explore the problem itself.

Two experienced engineers worked on the top-level data flow diagram for the HP 54720/10. This diagram, shown in Fig. 1, conceptually specified the system architecture. As the team grew, other engineers were assigned to analyze specific portions of the system.

Traditionally, design assignments for engineers had been feature-oriented, with each engineer designing and implementing all aspects of their assigned features. The system partitioning, done as part of the structured analysis, led to a new method for making design assignments. Each engineer was assigned responsibility for a specific technology area, like user interface or waveform analysis. This meant that engineers could become experts in specific technology areas rather than learning a little about everything required to implement a given feature. For example, the waveform marker user interface was designed by the user interface expert, the marker placement algorithms by the waveform analysis expert, and the graphics by the display expert. This organization proved to be much more efficient than the feature-focused type used in the past.

The computer design was optimized to support the system architecture as conceptualized in the data flow model. It includes custom hardware to accelerate the most time-critical data transformations and the movement of data through time-critical paths.

We learned a great deal about data flow analysis. We found data flow diagrams to be helpful in architectural development and as a communication tool within the team. However, the diagrams were not particularly meaningful to

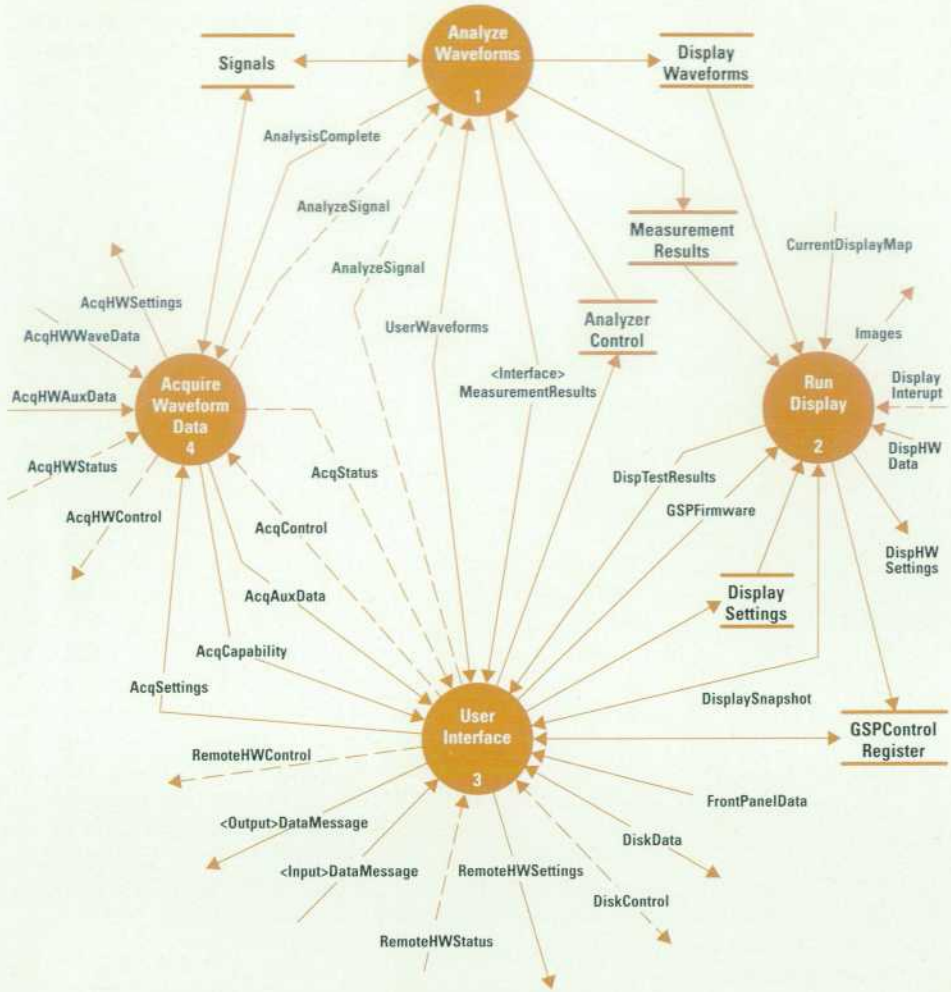


Fig. 1. Top-level data flow diagram for the HP 54720/10.

people outside of the design team or to new members of the design team without additional explanation. We found that data flow diagrams were more of a tool for exploring, learning, and modeling than for design or documentation. As the problem is decomposed to lower and lower levels, the value of further analysis declines. We found the lower-level diagrams difficult and time-consuming to maintain as we began spending more time on design and less on analysis. Since the low-level diagrams weren't used much after the analysis was completed and there was really no direct link to the design, we decided not to keep them up to date.

There were several important benefits gained from the structured analysis. The team as a whole became very familiar with the problems we needed to solve. As we explored system architectural alternatives, we discovered poor or dead-end choices early and discarded them. We could finally explain to management, in objective terms, why firmware development always seems to take so long. By presenting our model, we could show that the magnitude of the task was much greater than had been commonly believed and we were better able to justify an appropriate staffing level.

Front-Panel Simulation

In conjunction with the structured analysis and the early stages of design, we developed an interactive, graphical simulation of our instrument's front panel and display. The simulation ran under the X Window System and was written in ANSI C using X widgets. Using a mouse, people could simulate pushing front-panel keys to see how the instrument would behave. Fig. 2 shows how the front-panel simulation compares with the final front-panel design.

The simulation had multiple uses. It provided an environment for prototyping menus and various other features. It provided an excellent way of demonstrating our product concept. Most important, it provided us with a platform for doing early usability testing of our improved user interface. A large number of HP employees evaluated the new interface and refinements were made based on their inputs. The consultant who led our earlier usability tests evaluated the simulation and provided additional feedback. We were then ready to design the actual instrument front panel.

Later, the simulation turned out to be an excellent early development environment. We used it to test hardware independent code before real target system hardware was available. Because the code was written in ANSI C, porting it to the target system simply meant recompiling it.

The various processes described above helped us develop a feature set specification that balanced functionality with available design resources and that remained relatively stable throughout the remainder of the project.

Design Phase

While the analysis was still in progress several engineers began the design phase. They started by creating structured design charts.[2] Structured design involves creating a tree of

modules. The tree shows which modules can call other modules and the data that is passed between the modules. Each module contains a brief description of the tasks and transforms to be accomplished by the module. No formal approach was taken to progress from structured analysis to structured design. Rather, structure design charts were created from the top down, based on what we learned from the structured analysis.

A complete design was done for most of the firmware. Structure charts showing all modules were created including module specifications. The structured design charts helped us break the firmware into logical, functional modules. The module specifications contained a brief description or pseudocode for the module. Modules were then transformed into C functions.

Because of the technology-focused assignments, data passed between technology areas was critical to the success of the design. A great deal of time was spent refining the data that would be passed from module to module. The data dictionary, a lexical description of the data passed between modules, was used extensively to identify the type and hierarchy of key global data structures. Our structured design software used a flat name space for the data dictionary. As a result, the data dictionary quickly became a burden to maintain. New data dictionary entries were soon clashing with existing entries with different definitions. Several attempts were made to avert this problem with mixed results. A structured design package that better supported the C language would have helped.

Structured design enabled the team to identify common routines for global use. From these common routines were spawned abstract data types. An abstract data type is an association of data with the methods that can manipulate that data. Standard methods were constructed for the set, list, queue, ring, and stack data types. This abstraction enabled engineers to begin using these data types without concern for their implementation. Code size was reduced since the methods only needed to be written once. The risk of defects was reduced and the code is more consistent in the use of these types.

The structure charts were maintained for a while during the implementation phase. Once the implementation was in full force the structure charts were not kept up to date. After the implementation was completed the engineers reconstructed the structure charts, bringing them up to date. The charts have been helpful to educate new engineers on system functionality and are a good "road map" for navigation through the firmware.

Implementation Phase

After most of the structured design charts were completed and a good understanding of the data shared within and among subsystems was gained, engineers began the implementation phase. In previous products, a large portion of the code had been written in assembly language. The HP 54720/10 design team agreed to write most of the code in ANSI C. This has made the code much more maintainable. Over 85% of the HP 54720/10 system code is written in ANSI C with 15% written in assembly language. Assembly code was limited to functions that need to be optimized for speed and code that was acquired from other Hewlett-Packard divisions. The use of ANSI C enforced function prototype checking.

Since multiple engineers were working on various aspects of a feature, a firmware revision plan was constructed that coordinated our efforts to work on the same features at the same time. The firmware revision plan specified which feature would be implemented first and gave specific target dates for the completion of each phase. The features in the plan were prioritized so that the most important features would be implemented first. The plan was the first step in preparing a complete firmware project schedule. Using the plan the design team could measure their progress in each phase and estimate when future phases would be completed. In the later stages of the project, firmware revisions helped the quality assurance team track releases and ensure that engineers were testing the most recent code.

During the implementation phase, a complete project scheduling tool was used. The firmware team used the scheduling tool to measure our progress in each of the firmware revision plan phases. Each engineer created a set of tasks to be completed within each phase. Every two weeks engineers would record the time spent on these tasks and predict the time remaining. Near the end of the implementation phase, the schedule was used to assign engineers to portions of the firmware that had the greatest amount of time to completion. Using these tools, most of the engineers finished their implementation tasks within a month of each other.

After a few engineers became involved in the implementation phase, a coding standard document was agreed upon.

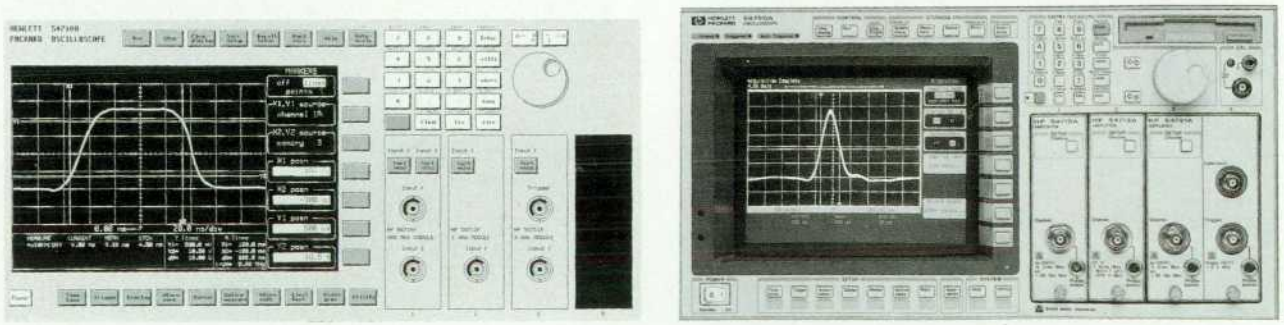


Fig. 2. Simulated and actual HP 54720/10 front panels.

The coding standard was leveraged from the Hewlett-Packard Lake Stevens Instrument Division and modified to pertain to our environment. The standard was agreed upon and supported by the team. The standard attempts to be rigorous in the definition and format of the C language but somewhat flexible in the definition of C syntax. Once a large body of code was written, the code itself became as useful as the standard since most of the code was written in a consistent format. A consistent format made the code much easier to read and understand. At the completion of the project, all of the engineers involved were enthusiastic about using the standard in developing the code.

Code reviews were held infrequently and early in the project. The reviews targeted difficult code that affected multiple engineers. Formal reviews were conducted consisting of a reader who did not write the code, a moderator, the author and several reviewers. The code reviews were sponsored by the quality assurance department and gave them early visibility to the code. The code reviews had more beneficial side affects than the accuracy of the code itself. The reviews helped to test and reinforce the coding standards. They were a forum for best practices in coding optimizations and format. They also helped some of the engineers who were unfamiliar with C learn the language.

Firmware Implementation Environment

When the implementation started, no adequate firmware implementation environment existed, so the firmware design team established several criteria for a configuration management system and developed an appropriate environment. The system supports multiple engineers working in independent directories. It supports experimental and working versions of the firmware. The environment was very beneficial in allowing multiple engineers to work collectively on the firmware. Engineers could work in their own directories on new code and post working code for other engineers to compile with. The implementation environment is very easy to use and is documented for new engineers to learn. Enhancements to the environment were made as new ideas were generated.

During the implementation phase, emulators and debuggers were used to debug code for the 68020 microprocessor. The display computer system's firmware was of low enough complexity that we decided to debug it using a logic analyzer rather than buying a more expensive emulator. We designed a logic analyzer preprocessor port directly into the display board. We were able to simplify the debugging task by downloading symbols from the link into the logic analyzer for symbolic addressing. Some emulator-like capabilities were also made available from the instrument's front panel. For example, a special debug menu was added which made it possible to look at and change the contents of memory locations in the display system, provided that the host 68020 computer was running. The logic analyzer combined with the debug menus worked well for debugging this system, but we would probably use an emulator if we were to increase the complexity of the firmware much further.

Performance Verification

As the implementation phase neared completion, we began working on performance verification. Although many aspects of performance were verified, only two are discussed here: usability and functionality.

We asked our usability consultants to conduct a final round of testing on a prototype instrument with fairly complete firmware. The test subject selection criteria were similar to those used in the initial tests and the tasks were identical. The primary difference was that subjects only tested the new instrument. The new results were compared to those for instruments tested during the first round.

The test results showed that subjects were able to complete the tasks in less time and generally with fewer assists than with the products tested during the first round. They also tended to give the instrument higher ratings on the post-test questionnaire. The new help feature received high marks and had, indeed, helped some test subjects complete their tasks.

The team had made progress towards developing an easier-to-use oscilloscope interface. There were still several suggestions for improvement, but the magnitude of the changes was smaller. For example, it was recommended (and implemented) that the menu titles be drawn in a bigger, bolder font so users would be more likely to see them and know where they were in the menu structure.

Functional verification was accomplished using two methods: automated regression testing and manual abuse testing. Fully verifying all firmware functionality was no small task. There were over 200,000 lines of source code and the team had high quality goals. To complicate matters, the firmware could only be run on the instrument's embedded computer system. Testing on an embedded computer precludes using many of the methods traditionally used for workstation or personal computer software.

The automated regression tests were written to exercise the instrument via its HP-IB (IEEE 488, IEC 625) interface. The instrument is designed such that user inputs are acted upon by the same code regardless of whether they come in from the front panel or the HP-IB. Hence, testing the instrument via HP-IB exercised much of the same code as front-panel testing. The tests were written in C and run on a UNIX* workstation. Coverage was measured using a branch analyzer in combination with an emulator.

The HP 54720/10's regression test suite had the advantages of being automated and repeatable. However, it couldn't automatically exercise some portions of the firmware. For example, a portion of the self-test and calibration firmware is only executed when a hardware problem is detected. The test development became much more time-consuming as the coverage level increased.

We considered the regression tests quite successful. The tests exercised over 65% of the system firmware. The automated tests tended to be good for finding system crashes and serious defects. Of all defects reported, 9% were found by the automated tests with an average submitter-determined severity of 5.6 on a scale of 1 to 10, where 10 is most severe.

More time was invested in manual abuse testing than in developing the automated tests. This was because the manual tests had to be repeated for each new code release. The abuse testers were given guides for testing various parts of

Developing Extensible Firmware

A primary goal of the HP 54720/10 firmware project was to develop extensible code, that is, code that can easily be modified and added to. The goal was considered important because the HP 54720/10 was the first member of a new oscilloscope family with many follow-on products expected in the future.

It is difficult to measure in quantitative terms whether the goal was achieved or not, but a good test case came up right away: adding FFT (fast Fourier transform) capability to the HP 54720/10. This proved to be an interesting case since the same feature set had just been added to the HP 54510A oscilloscope.

Three areas stood out as making it easier to add the FFT code: the vertical organization of the code, the thinking ahead that was done by the design team, and the HP-IB command parser.

Vertical versus Horizontal Organization

The organization of the code made it very easy to add FFT capability. The best way to describe the code is in terms of vertical organization versus horizontal organization. The two terms are often used to describe companies. A vertically organized company is one that performs many if not all of the steps in the manufacture and selling of a product. An example is a shoe manufacturer that does everything from processing the leather to selling the shoes. A horizontally organized company specializes in a product area and only performs one step in the manufacture and selling of the product. An example is a bottling company that produces several different brands of soft drinks.

The code in the HP 54720/10 is organized vertically. Each module performs all the tasks related to the module, such as variable definition, initialization, save and recall, menu generation, keyboard entry, and command execution. The vertical organization of the code proved to be a big time saver because most of the work of adding the FFT capability could be done in one file.

In contrast, the organization of the code in the HP 54510A oscilloscope is horizontal. There is a separate module for each of the broad tasks: variable definition, initialization, save and recall, menu generation, keyboard entry, and command execution. Because of the horizontal organization, many more files had to be understood and modified. Sixty-five modules were modified in the HP 54510A versus fifteen for the HP 54720/10 code.

The vertical organization of the code also made it easier to do concurrent development. The writing of the FFT code began several months before the HP 54720/10 code was scheduled to finish, and was completed a few months after the HP 54720/10 code. During the period when both projects were running simultaneously, it was important to keep the FFT features out of the main code and not to introduce any bugs. Since most of the code work could be done in one file, it was possible to work in a separate directory from the other developers and not have the new FFT code affect good code that was close to being released.

Thinking Ahead

Oscilloscopes can perform a wide variety of mathematical operations on the input signals such as inversion, multiplication, integration, and differentiation. The resultant waveforms are referred to as functions. The HP 54720/10 team spent quite a bit of time at the beginning of the project trying to anticipate the types of functions that might be created and trying to make functions completely general-purpose. The results paid off. Approximately two months were saved because of the thinking that went into making functions general-purpose.

Three things in particular reduced the amount of time it took to add the FFT code. The HP 54720/10 code allows waveform records to be of any length. This was especially useful for FFTs, which have weird lengths: powers of 2. The HP 54720/10 code allows waveform records to have any type of units. FFTs have units of decibels and hertz, which are different from the usual time-domain units of volts and seconds. The HP 54720/10 has a routine that automatically does the horizontal and vertical scaling of the function. This routine made it easy to implement action routines for the vertical offset, vertical scale, magnify span, and center frequency keys.

HP-IB Command Parser

Both the HP 54720/10 and the HP 54510A oscilloscopes have an HP-IB command parser, which is worth mentioning because it resulted in significant time savings. The HP-IB language is a set of commands and queries that allow a user to control the instrument with a computer over the HP-IB (IEEE 488, IEC 625). In past instruments, the HP-IB section of code has been a difficult section to extend or modify, but an HP-IB command parser that we borrowed from another HP division has made the job much easier. It took less than two weeks to add the 22 FFT commands and 22 FFT queries.

The HP-IB files consist of a language file and a set of action files. The language file defines the entire HP-IB language, and is written in almost the same format as the syntax diagrams for the commands and queries that appear in the manual. The action files contain short routines for the commands and queries that specify where the incoming data goes to and where the outgoing data comes from. Once the commands have been defined in the language file, a special compiler is executed to translate the language into code. The HP-IB compiler is an example of a software tool that greatly speeds up the firmware development process. It takes a task that has been done over and over again in HP instruments and automates it.

Documentation

One aspect of firmware development that is not fully addressed by the use of standard structured methods is complete system documentation. The HP 54720/10 team set standards for the documentation, but more documentation would have helped because a large amount of the FFT project time was spent learning about the system firmware. The typical documentation was a description at the beginning of a routine, a few comments scattered throughout the routine, and a structure chart.

Three types of documentation would have helped. First, system-level documentation that gives a big picture of the system firmware and the interaction of the main modules. Second, more comments in the source code, especially to describe what called subroutines do. Often it was necessary to go through all of the subroutines to understand what the main routine did. Third, it would have helped to have better descriptions of variables and some of the different variable parameters.

The key message is that additional time needs to be allocated specifically for system documentation beyond that generated by the structured design process itself. Good documentation is one of the areas that has the highest payback on follow-on projects, yet it is often neglected.

Rodney T. Schlater
R&D Engineer
Colorado Springs Division

the system to help ensure even coverage, although the coverage was not measured. They were also encouraged to try anything they could think of that might uncover a defect. They recorded time spent and the number of defects found.

The advantage of manual testing was that humans could easily evaluate functionality that is difficult to verify automatically. They could look at how the display responded to things like color specification changes, the responsiveness of controls, and the correctness of answers or waveform displays. We were able to include testers with varying experience levels and interests. The drawbacks were that the testing was not very repeatable, it was time-consuming, and it got monotonous very quickly.

We found that the manual testers tended to find more subtle, less severe problems than did the automated tests. Abuse testing resulted in 34% of all defect reports with an average submitter-determined severity of 4, 1.6 points lower than the average for the automated tests.

The remaining 57% of the defect reports were not the result of formal testing, but were mainly submitted by project engineers during the implementation phase.

In the past, we had focused primarily on front-panel testing. The combination of both methods produced better results. Very few defects have been reported in the firmware since the final release, even though the instrument can print a form directly to make it easy for users to report problems.

Results

The HP 54720/10 design team met its goals. The user interface was very well received by end users. The postrelease defect find rate was extremely low. The system met its performance goals. Finally, the system turned out to be well-suited to serve as a platform for the future. The first new feature added to the HP 54720/10 (see "Developing Extensible Software," page 64) required 25% less time than adding the same feature to a similar firmware system that was developed using less structured techniques. After evaluating the difficulties encountered in adding the new feature, the design team developed a more complete set of documentation to make the system even easier to maintain and enhance. Structured methods and the other processes used by the HP 54720/10 design team didn't solve all of the team's problems, but they did help the team meet its goals and develop a better product.

Acknowledgments

The authors would like to acknowledge the rest of the HP 54720 firmware team for their hard work and dedication to process improvements. Other team members included Dave Poppe, Caren Johnson, Dana Johnson, Jerry Kinsley, Rich Hollon, Mark Heim, Chris Magnuson, Allen Montijo, Greg Tarcza, Mimi Beaudoin, and Don Allison. They would also like to thank managers Mike Karin and John Scharrer for taking the risk of allowing the project team to try new development processes, Don Henry for his work on the HP 54720/10 mainframe front-panel industrial design, Denis Lambert and Lynn Lenhard for their support in project scheduling, Art Porter for his coordination of the usability tests, Cathy Habib and Barbara Rider of the Hewlett-Packard Lake Stevens Instrument Division for letting us leverage their work as the basis the the X Windows simulation of the HP 54720 front panel, and Grant Grovenburg for coordinating the firmware testing.

References

1. T. DeMarco, *Structured Analysis and System Specification*, Yourdon, Inc., 1978.
2. M. Page-Jones, *The Practical Guide to Structured Systems Design*, Yourdon, Inc., 1980.

Mechanical Design of a New Oscilloscope Mainframe for Optimum Performance

A completely new mainframe design for the HP 54720/10 oscilloscopes includes a unibody chassis and four plug-in slots that provide superior EMI performance and anticipate future enhancements.

by John W. Campbell, Kenneth W. Johnson, Wayne F. Helgoth, and William H. Escovitz

Mechanical design of the HP 54720/10 oscilloscopes began with a clean slate. The goal was to provide an improved mainframe for high-performance, state-of-the-art oscilloscopes with bandwidths greater than 1 GHz. The team found that it was time to reexamine the designs of the 1980s.

HP high-performance oscilloscopes of the 1980s used an HP System II enclosure as the mechanical mainframe. System II is a modular system consisting of standard skeletal castings and outside sheet-metal cover panels. System II made it possible for many designers in different divisions to design instruments with various outside sizes using standard parts. Each instrument used its own specific internal cardcages and sheet-metal trays and bulkheads. Cables and printed circuit board edge connectors carried the high-frequency signals from board to board. Screws attached the internal parts to the System II enclosure.

In spite of its modularity, or perhaps because of it, System II had many disadvantages for digitizing oscilloscopes. The primary disadvantage was electromagnetic interference (EMI), especially radiated interference. Digitizing oscilloscopes contain high-frequency clocks for the digitizer, computer, and display clocks and the switching power supplies. The many seams between the sheet-metal side panels and the mainframe were difficult to seal completely against radiation from these sources. It was common to spend many engineer-months at the end of the project trying to meet the EMI regulatory requirements. Furthermore, there was extensive internal coupling. Oscilloscopes contain sensitive analog input, trigger, and calibration circuits. Vertical accuracy has improved from about 3% to about 1% in the last decade, putting higher demand on calibration. Time base circuits of digitizing oscilloscopes are more sensitive (and therefore more susceptible to EMI) and much more accurate than those of analog oscilloscopes. So radiation, in addition to leaking out, could easily appear unwanted in victim circuits, often leaking through the many cables inside the instrument. It is easy to see how fixing EMI and its associated noise turns into a critical path for the project.

The previous high-performance oscilloscope mainframes had other shortcomings that increased the desire for a new mainframe. Many of the parts in the basic mainframe exoskeleton, as well as internal parts attached to the mainframe, were held in place by screws. This assembly was expensive for HP. The new oscilloscope would contain more and higher-power-density circuits. As HP bipolar processes have moved through HP5, HP10, and HP25, the new integrated circuits have had smaller transistors, more functionality, higher speed, and sometimes smaller chips. This has also led to more acquisition channels in the oscilloscope and increased requirements on the cooling, while customers were already complaining that their instruments are too loud. Thin hybrid substrates used to obtain the necessary circuit density needed special protection from shock and vibration damage.

Users of general-purpose oscilloscopes with bandwidths above 500 MHz generally need plug-in or modular systems. This allows the user to tailor the number of channels, bandwidth, sample rate, input impedance, ac/dc coupling, and other options to match the needs of the application. With present technology, one combination of attenuator and amplifier can handle these capabilities up to 500 MHz. For oscilloscopes with bandwidths above 500 MHz (the HP 54720/10's is over 2.0 GHz), it is necessary to have a plug-in system so that the attenuator/amplifier combination can be changed. However, previous plug-in interconnection schemes limited the bandwidth to about 1 GHz.

The HP 54720/10 design team decided to develop an oscilloscope mainframe with the following characteristics:

- Improved internal and external shielding so that noise and EMI problems would be minimal
- Simple assembly
- Plug-in system with higher bandwidth and more reliable, easier mating
- Higher capacity but quieter cooling.

Chassis Design

The cabinet design for the HP 54710/20 oscilloscopes represents a departure from typical packaging concepts. Historically, design teams have often expended an inordinate amount of effort attempting to design, modify, and refine electronic enclosures to comply with regulatory agency tests for radiated emissions. Many designs have required a variety of special parts as well as difficult installation procedures to achieve the required performance. The primary goal for this design was to improve shielding effectiveness against radiated energy while at the same time improving the simplicity of assembly and parts handling.

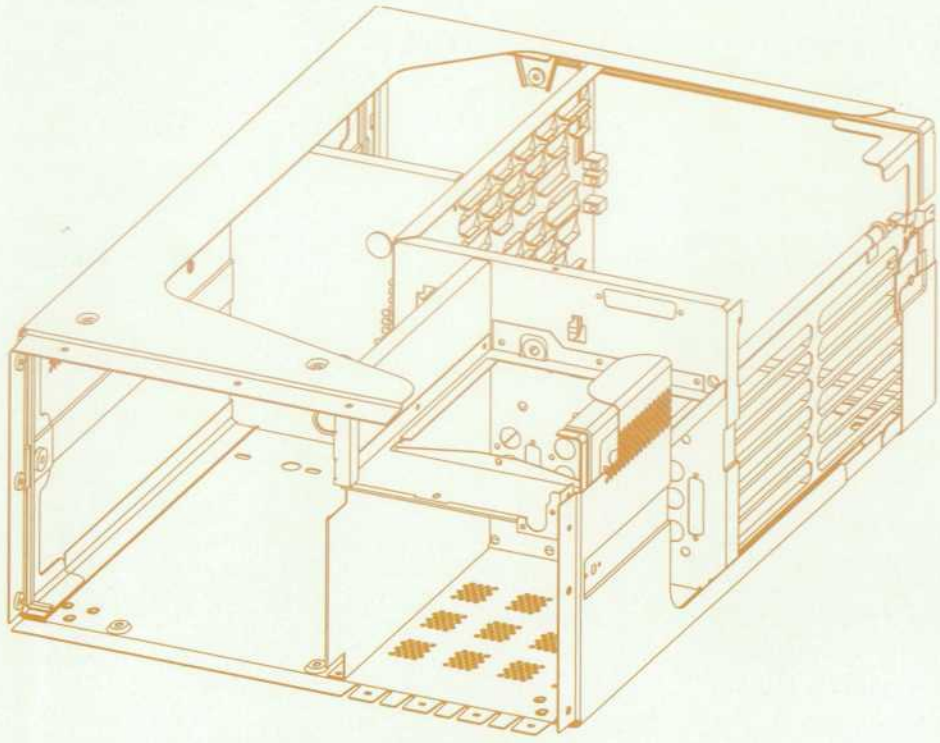


Fig. 1. View of the oscilloscope cabinet from the front. Internal compartment elements and the external cosmetic shell are welded into a single structure. Clockwise from the left front are compartments for the display, the fan, the cardcage, and the plug-ins. The compartment at the right front above the plug-ins houses the plug-in fan and disk drive.

Beginning with the assumption that the cardcage area of the product would be the primary source of radiation, our first-cut concept was to create that section as a completely sealed compartment which, if necessary, could stand alone as a shielding entity. Since internal coupling within the container can also represent problems for the circuit designers, other areas were targeted for at least some level of compartmentalizing, primarily the display. Other compartments were quickly added onto the shielded cardcage. This created the problems of how to fasten the other compartments to the cardcage and how to cover the whole thing in a cosmetic wrap. A variety of interesting modular concepts emerged. In the approach selected, the individual compartments share common internal walls as a single unit, rather than being separate, and their outer walls form the external cosmetic skin—a sort of unibody oscilloscope cabinet (Fig. 1). Why should the external skin be simply an external skin, with some sort of internal skin doing the real job?

One of the difficulties encountered in the design of electronic packaging is the configuration of joints, which must have small enough gaps to prevent radiation of electrical energy. Gasketing between chassis elements can create assembly difficulty because of the pressure acting on sliding or compression fits. The contact effectiveness can vary from unit to unit, and it is sometimes very difficult to get the gasketing to conform to all of the required locations. With the pieces welded together with appropriate weld spacing, there are no variable-seal joints to become shielding leaks.

The choice of low-carbon steel for the cabinet was driven by several issues. Spot-welding seemed appropriate to the fabrication of such an assembly. Steel spot-welds very well compared to aluminum, which has a tendency to build up on welding tips. Other important characteristics of steel compared to aluminum are its higher permeability and lower conductivity. This combination of characteristics allows steel to provide increased absorption loss at lower frequencies. This was an important consideration since internal lower-frequency magnetic coupling into sensitive circuits was a concern requiring attention in the design. Combining these factors with a lower basic material cost, spot-welded steel appeared to have the advantage.

Trying to predict a total theoretical shielding effectiveness of all of the air holes, weld seams, and other configuration features can be quite difficult. The basic rules governing slot sizes, including the weld spacings, and other configuration rules and guidelines were combined into a design that seemed reasonable. An early model of the design was then tested in the screen room.

The slogan, "One trip to the EMI site!" was used during the package development. Previous products required many trips to the outdoor test site along with lab and screen room testing, not to mention rolls of copper tape to put over troublesome cover and chassis seal leaks. The goal here was to develop a configuration that could be evaluated in the screen room with enough margin to go to the qualification step with confidence.

The chassis was tested in the screen room with a comb generator. The same test was performed on some other typical boxes. Measured shielding effectiveness was an average of 20 dB (25 to 32 dB in the 400-to-700-MHz range) higher than the previous package design (Fig. 2). No repeat trips to the EMI site were required!

To allow access to the inside of the instrument, some form of detachable seal must exist at some point. The design concept called for a controlled, repeatable seal at the rear panel to allow easy access to the cardcage compartment. The panel contains four folded sides which interface with mating U slots folded in the rear edges of the package. These U slots contain a sliding-fit spring gasket which makes contact with

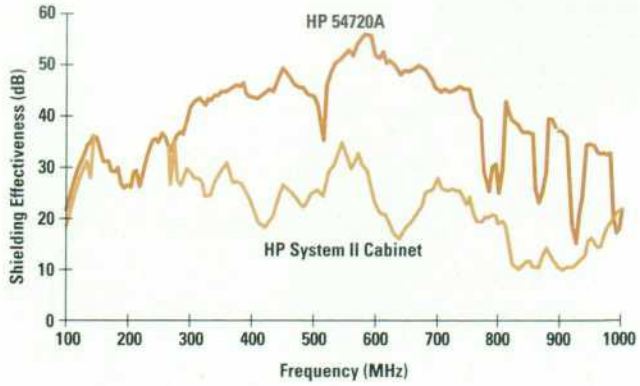


Fig. 2. Chassis shielding effectiveness was measured against other packages using a harmonic comb generator mounted inside the package.

the four sides of the panel. The panel sides are surrounded by the the gasket and the slot (Fig. 3). The quality of the contact made is not sensitive to variations of in and out positioning. At any point in its contact range, the contact is the same. Very little pressure is required to install or remove it.

Since the cardcage is sealed and the package is compartmentalized, no shielding boundary is required at the front. The unshielded plastic front panel is removable to provide access to the forward compartments.

The chassis supplier assembles eleven part numbers representing a total of 19 pieces that make up the entire raw cabinet (Fig. 4). Commercial weld nuts are used along with two custom threaded parts to provide threaded fastener attachment points. Assembly requires mostly spot and some wire welds. Two of the parts—the front deck and cardcage—are tin-electroplated to meet connection requirements while the remaining pieces are zinc-plated with a yellow chromate conversion coat. After welding, the exterior receives a sanding of the welds, primer, and a stipple-texture waterborne color paint coat.

While no specific parts count reduction was identified early in the project, the new design results in a considerable improvement on the assembly floor. To assemble a complete cabinet, EMI-sealed, with all internal supporting structures

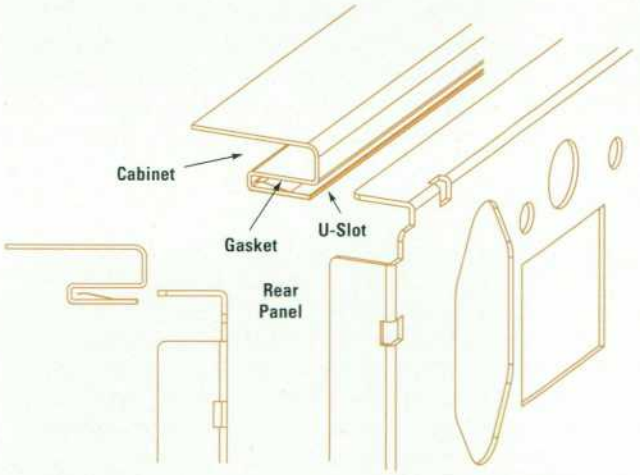


Fig. 3. A sliding-fit spring contact strip mounted in the rear U slots makes low-pressure, repeatable contact with the rear panel.

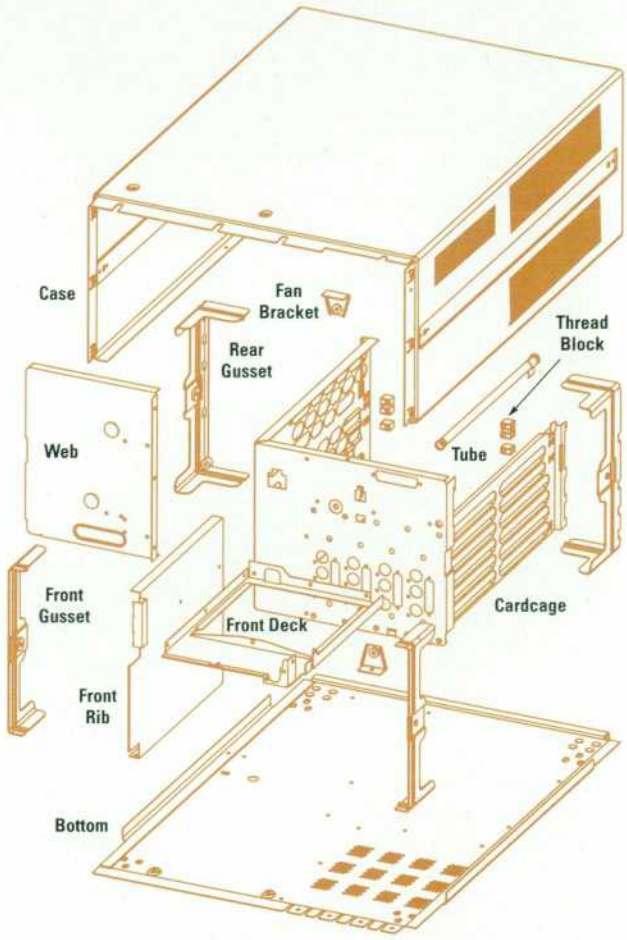


Fig. 4. Nineteen parts are welded by the supplier to complete a cabinet ready to accept circuit boards and other components. The vertical plane of sheet metal at the center of the drawing is the bulkhead, which establishes the z location of the plug-ins.

in place requires less than one-third the parts needed for the older designs. For the HP 54720/10 this amounts to five part numbers representing a total of seven individual pieces. The previous designs required 15 to 18 part numbers and 24 to 25 pieces. While many of the EMI seal features (gaskets) made cover installation difficult for previous designs, with the new design of the cabinet an outside supplier skilled at metal fabrication and welding focuses on assembly of that significant portion of the product, thereby giving HP manufacturing more time to address other technical details.

With internal access limited to the front and rear only, modularity of the subassemblies is a must. The result is a very logical modular partitioning of the internal parts (Fig. 5).

Plug-ins and Their Interconnection

One of the problems experienced in previous projects with this type of product involved the often extensive cabling between the input at the front and the acquisition section (cardcage). Coupling from circuit to circuit and interference from the display assembly presented difficult problems for the circuit designers.

The original design concept identified the use of plug-in modules as a means of allowing a variety of measurements to be configured at the front end of the instrument. The four

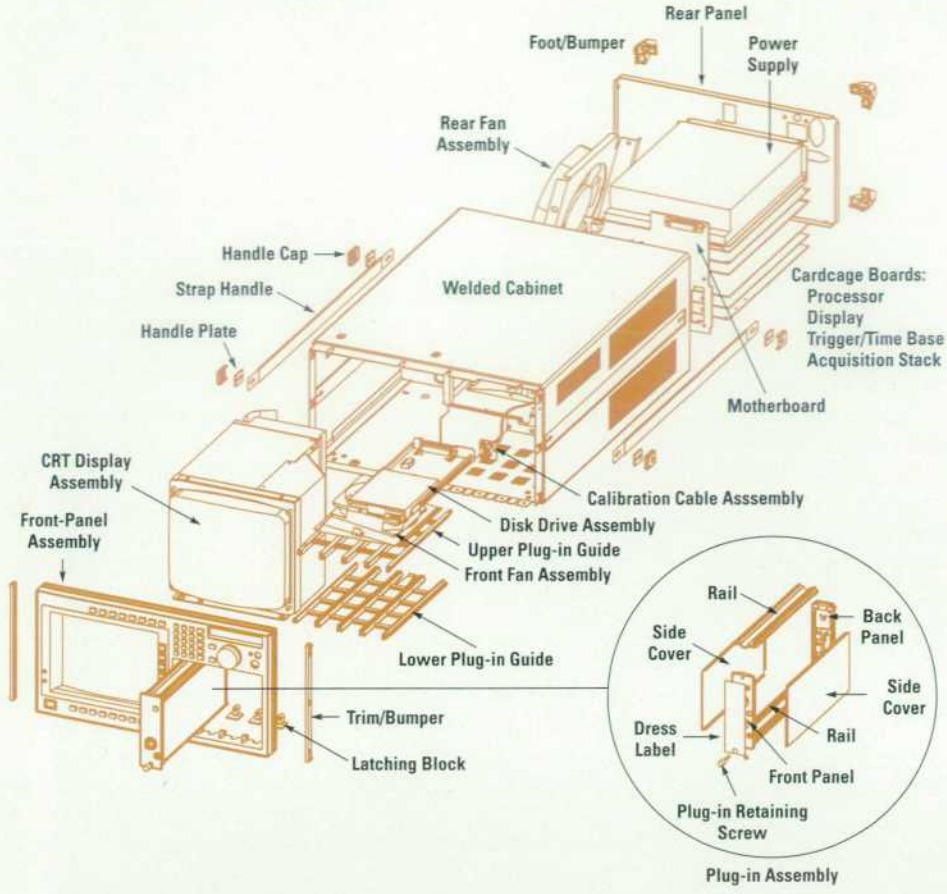


Fig. 5. Modularity of components is required for assembly into the cabinet and contributes to overall simplicity and organization of the production flow. This drawing shows the chamfered protrusions in the front panel that define the plug-in slots and the top and bottom guide rails for the plug-ins.

plug-in slots have to be able to make connection with four acquisition slots in the cardcage behind them. This is where all the variable cabling was used in previous designs. A variety of configurations dealing with a vertical or horizontal cardcage and vertical or horizontal plug-ins were considered. The result was a vertical plug-in orientation mating with a horizontal cardcage, thus allowing direct connection from each plug-in to each of four acquisition boards. The front of the cardcage forms the boundary of the sealed cardcage. The plug-ins are sealed individually and communicate with the acquisition boards through shielded connectors into the cardcage.

The plug-in compartment contains four slots for accommodating any practical combination of one-slot, two-slot, three-slot, or four-slot plug-ins. The plug-ins consist of a four-piece die-cast aluminum outer mainframe which includes a cosmetic front panel and all of the features that align and locate the plug-in. To this mainframe are mounted printed circuit boards, connectors, cable assemblies, and some sheet-metal brackets for support and damping. After assembly of the plug-in mainframe and internal hardware, two snap-on sheet-metal covers are put in place to protect the internal hardware and to create a fully EMI-shielded plug-in package (Fig. 5).

Forward Compatibility. One of the contributions a modular instrument offers the customer is the ability to upgrade or reconfigure a measurement system in the future. Therefore, one of the primary design goals was to make the plug-ins and plug-in/mainframe interface as forward-compatible and expandable as possible. This imposed extra design constraints since the design did not just need to solve existing problems but also had to anticipate future problems and incorporate features into the design so that the future problems could be solved.

Initial products were envisioned using only one-slot plug-ins, but it was an initial design decision not to rule out the use of two-slot, three-slot, or four-slot plug-ins. Future designers could use these other plug-ins to take advantage of the increased volume, to eliminate redundancies, to reroute signals from the front panel of the plug-in to the mainframe, or for some other unanticipated reason. The possibility of a plug-in that could span multiple slots meant that only the top and bottom of the plug-in compartment and the plug-ins could be used for guiding, aligning, keying, and retaining the plug-ins.

The plug-in opening in the front panel of the instrument has chamfered protrusions which delineate each plug-in slot and guide the plug-in into the instrument. Pockets in the rear of the protrusions index with the front of a set of guide rails in the top and bottom of the plug-in compartment. The guide rails are held captive between the front panel and the bulkhead at the rear of the plug-in compartment. These rails guide the plug-in from the front panel to the bulkhead where the plug-in rear panel engages the guide pins. The top and bottom rails of the plug-in have a short ramped portion near the end of the plug-in's travel which aligns the plug-in cosmetically with the instrument. The plug-in is then held in place with a thumbscrew which screws into a powdered metal receptacle in the bottom of the front-panel opening.

In anticipation of future plug-ins the die-cast front panels of the plug-ins have extra holes cast in place for additional connectors and keycaps. When these openings are not used

they are covered by the polycarbonate cosmetic front panel. The die-cast rear panels of the plug-ins have pockets and holes cast in place for additional connectors; when these are not in use they remain empty. In addition, all the die-cast parts have extra features cast in place, holes drilled and tapped, or machined surfaces in anticipation of additional hardware in future plug-ins. Although many of these features go unused, the small incremental increase in tooling costs or part costs to produce them is offset by the potential savings in designing and setting up new parts in the future.

With the potential for larger plug-ins to be designed in the future, the plug-in design needed to anticipate plug-ins with a larger mass. The design assumes that the heaviest plug-in will weigh four times as much as the heaviest existing one-slot plug-in and applies a safety factor of two. All load-bearing members and components experiencing relative motion are designed with these larger loads in mind. A mockup of a heavy plug-in was built for testing, with certain characteristics exaggerated to determine how much margin is in the design. The mass inside the test plug-in was nonuniformly distributed to add extra stress to specific load-bearing members. The test plug-in and instrument were put through shock and vibration tests and the plug-in was inserted and removed from the instrument over 1000 times. The experiment was repeated several times for statistical significance. After the tests were completed, all the components of the system were evaluated. Although some components showed significant wear, nothing had failed. The plug-in system is specified to handle 3.6-kilogram plug-ins for 500 insertion and removal cycles.

Reliable Interconnect. Previous attempts at modular oscilloscopes resulted in plug-ins that had to be persuaded to plug in and did not give a feeling of quality to the customer. A design goal for this instrument was that regardless of the instrument's orientation (excluding face down on the ground), if the opening in the front panel was accessible, the plug-in would in fact plug in and mate with all of its connectors, and it would feel good. Although the last part of the goal is only qualitative, it is obvious if it is achieved.

Each plug-in slot has provisions for one 25-pin D-shell connector for power and control and three coaxial connectors for passing signals from the plug-in to three individual mainframe boards. In the case of a four-slot plug-in, the design for the plug-in/mainframe interface guarantees that up to four D-shell connectors and 12 coaxial connectors will mate simultaneously with enough precision not to degrade the performance of any of the connectors.

The backbone of the plug-in/mainframe interconnect is the backbone of the entire instrument, the motherboard (Fig. 6). The motherboard uses press-fit connectors that are loaded onto both sides of the board. One side of the board contains all the connectors that connect to all the boards in the cardcage. The other side of the board contains the connectors that connect to the assemblies outside of the cardcage. This architecture allows all of the cardcage boards to communicate with each other and to communicate with the assemblies outside of the cardcage (plug-ins, disk drive, etc.) without extra cabling. The motherboard is attached to the bulkhead, which separates the cardcage from the plug-in compartment.

The motherboard is the x and y datum and the bulkhead is the z datum for the plug-in/mainframe interface. Eight stainless-steel guide pins, two for each plug-in slot, are press-fit into the motherboard. Because both the 25-pin D-shell connectors that interface with the plug-in and the guide pins are press-fit into the motherboard, the x-y relationship of the two is a function of the hole-to-hole capability of the motherboard vendor. Any manufacturing variation of the connectors or the guide pins is an order of magnitude less of a contributor to connector and guide pin mismatch. The guide pin has a large shoulder at the base so that it becomes captured between the motherboard and the bulkhead and cannot work itself out. The guide pin is also tied to the ground plane of the motherboard; it supplies the safety ground path for the plug-in and discharges the plug-in before the connectors mate.

Up to three of the cardcage boards can receive signals from each plug-in slot through high-quality coaxial connectors. On

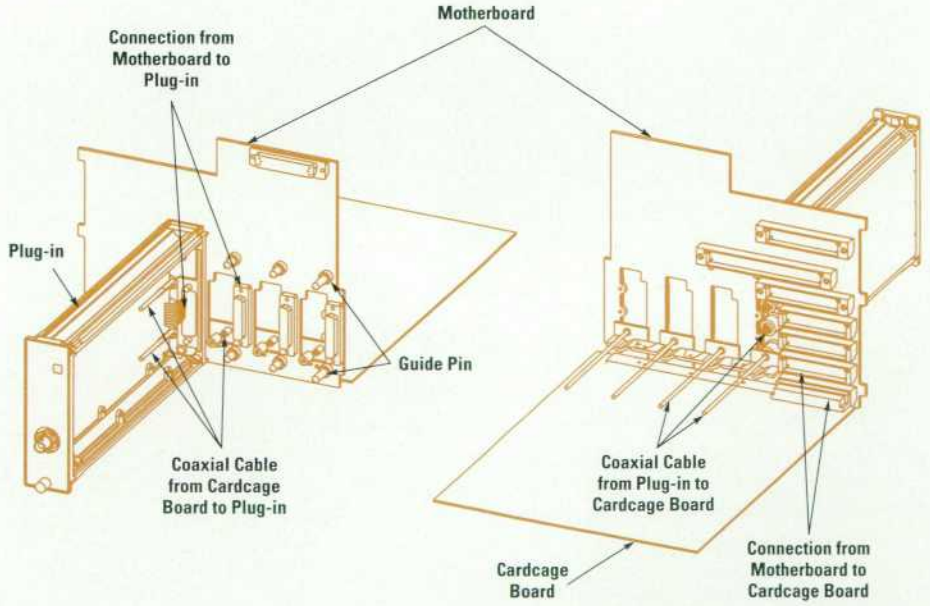


Fig. 6. Motherboard and major mechanical components of the plug-ins and cardcage boards, showing how they interconnect.

the cardcage side, the coaxial connectors are rigidly mounted on die-cast brackets mounted on the front of the boards. Two small dowel pins are press-fit into these brackets. When the board is inserted into the instrument, the dowel pins fit into holes in the motherboard and align the coaxial connectors with respect to the guide pins. The brackets also have four mounting bosses, one next to each coaxial connector. These bosses are screwed to the bulkhead from the plug-in side and pull the bosses up to the bulkhead to locate the coaxial connectors in the z direction. The screws also ground the shield of each coaxial connector where the signal passes through the bulkhead, which is an EMI shield, to minimize any radiated emissions.

The connectors in the back panel of the plug-in are all float-mounted, meaning they are compliant in x, y, and z. This is to accommodate any misalignment between the nominal positions of the plug-in and mainframe connectors because of manufacturing variations.

The back panel of each plug-in engages two guide pins, one at the top and one at the bottom. Two pins are necessary to prevent the plug-in from rotating and misaligning the connectors. Plug-ins that span more than one slot use the guide pins along the diagonal from the top to the bottom. This allows more precise alignment and minimizes any torque applied to the guide pins. The back panel contains two bronze bushings which slide over the guide pins and align the plug-in in x and y. Bronze was chosen for its bearing properties. The top bushing is round and performs the main alignment. The bottom bushing is obround to accommodate manufacturing variations and only prevents the plug-in from rotating. Around the top bushing is an annular spring which contacts the top guide pin and provides a path for safety grounding. As the plug-in slides down the guide pins the foot of the back panel contacts the bulkhead and registers the plug-in connectors in the z direction.

To ensure that all of the connectors would mate all of the time an analysis was performed on the plug-in/mainframe interface design. The analysis assumed manufacturing variables contributing to misalignment with five degrees of freedom. Most variables were assumed to have normal distributions and several had bimodal distributions. The worst-case analysis proved that the design would not work, so further analysis was undertaken, using a computer program that performs Monte Carlo analysis. The program randomly takes components from the different distributions, builds a plug-in and instrument, and measures the connector mismatch. These simulations were repeated thousands of times while varying some of the distributions. The results of the simulations showed that several custom parts could be eliminated and several manufacturing tolerances reduced. This resulted in a 66% reduction in material cost for the plug-in/mainframe interface and a predicted failure rate less than 6 parts per million.

Quality and Durability

Quality and durability are qualities that HP products are known for. To ensure that all new products live up to customer expectations they must go through many tests. One of the more severe tests is the shock test. In the shock test the amount of energy applied is inversely proportional to the mass of the product. Shock tests were performed for an instrument with plug-ins and for plug-ins alone. The plug-ins tested by themselves were subjected to a shock pulse nearly three times greater than in an instrument, which seemed reasonable because it was easy to envision a plug-in being bumped off a workbench.

Two of the plug-ins were experiencing failures in the shock test. Both contain thick-film preamplifiers. The thick-film circuits are 1-mm-thick ceramic hybrids with edge clips on two sides that are soldered into a printed circuit board. The printed circuit board is mounted with screws to the die-cast rails of the plug-in. When the plug-ins were excited with a shock pulse on either of the two faces parallel to the plane of the ceramic substrates the thick film hybrids would crack.

The transmissibility of the shock pulse was measured to be 14. Transmissibility is given as:[1]

$$\frac{F_T}{F_0} = \frac{\sqrt{1 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}}{\sqrt{\left[1-\left(\frac{\omega}{\omega_n}\right)^2\right]^2 + \left[2\zeta\frac{\omega}{\omega_n}\right]^2}}$$

where F_0 is the exciting force, F_T is the transmitted force, ω is the exciting frequency, ω_n is the natural frequency, and $\zeta = c/c_c$ is the damping ratio, where c is the damping coefficient of the system and c_c is the damping coefficient of a critically damped system. Fig. 7 is a plot of transmissibility.[1]

The above equation shows that to reduce the transmissibility, either the damping ratio must be increased, or the ratio of the exciting frequency to the natural frequency must be increased, or both. The natural frequency of a thick-film hybrid is a function of its beam properties and any change in these properties would require extensive redesign. Therefore, we added more damping to the system to increase the damping ratio. This was accomplished by applying two strips of adhesive-backed foam, one at each end of the ceramic substrate between the substrate and the printed circuit board, covering approximately 30 percent of the surface area of the hybrid. The net result is a transmissibility slightly less than 2 and no more broken thick-film hybrids. All of the plug-ins have been now subjected to a slightly less formal

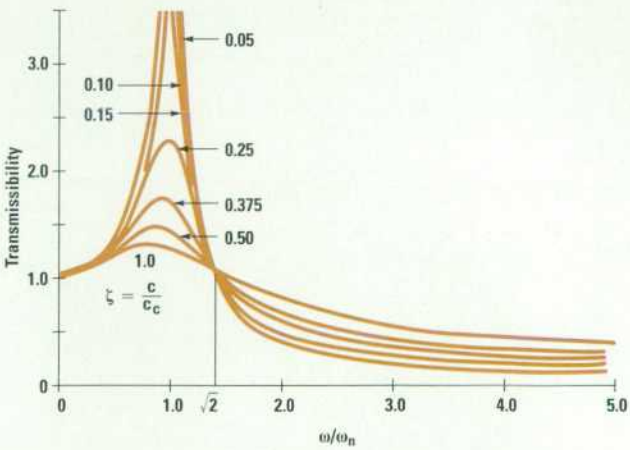


Fig. 7. Transmissibility as a function of frequency for different damping ratios.

shock test. A 6-foot-tall engineer holds the plug-in out at arm's length and drops it onto a hard floor. The result is a few alarming nicks and dings, but nothing fails.

Cooling

One of the goals for the cabinet design was to have sufficient cooling of all components using a fan quiet enough for a lab environment. For a given airflow, a large fan running slowly will be quieter than a small fan running faster, so a 6-inch-diameter fan was chosen as the main fan. We considered using a pair of 4.7-inch fans, but the single large fan was better for airflow uniformity and quietness. The speed of the fan is controlled with a thermistor to increase airflow when the oscilloscope is in a hot environment. Three compartments required cooling air: the cardcage, the display, and the plug-ins (see Figs. 1 and 5). For the cardcage, the emphasis was to have good airflow with no pockets of dead air. Placing the fan alongside the cardcage makes airflow in this area reasonably straight-through. A major part of the cardcage cooling design was for the heat sinks that cool the acquisition hybrids. There are four of these hybrids, two on each of two circuit boards, and each uses 17.5 watts. One large heat sink is used for each board. Having one heat sink for two hybrids keeps the second hybrid at about the same temperature as the first and also provides the board rigidity needed for the fragile hybrids. With this heat sink and the fan at high speed, the highest chip temperature is only 53°C above ambient.

The display compartment has air inlet holes from the fan compartment and cardcage. Initially, there was enough airflow into the compartment but it was not cooling some of the components. Much of the air was going across the back of the compartment, then along the side to the exit holes. Adding a deflector to the back of the display unit to channel air into the display puts this previously wasted air to use and keeps all components cool.

A separate fan for the plug-ins guarantees adequate airflow while not compromising cardcage airflow or EMI shielding. With this fan spanning all four plug-in slots and blowing air down into them, all plug-ins get enough air even if adjacent slots are left empty.

Shock and Vibration Resistance

Most of the design considerations for shock and vibration resistance were straightforward, but some electrical requirements combined to present a challenging problem of protecting the acquisition assembly. Circuit density and impedance considerations required that the ceramic acquisition hybrid substrates be only 0.015 inch thick. There are two acquisition circuit boards, each of which has two hybrids. A large heat sink (approximately one pound) is used on each board to keep the hybrids cool. The timing pulses are generated on a separate board and distributed to each hybrid. Timing accuracy and EMI requirements dictated that semirigid cables be used for the timing pulses. The hybrid input cables at the front of the boards had to be precisely positioned at the plug-in interface, while for ease of assembly and cooling, the hybrids are at the back of the boards. The result of these requirements was three boards and four fragile hybrids interconnected with stiff cables and no simple way to mount the heavy end of the boards rigidly to the cabinet (Fig. 8).

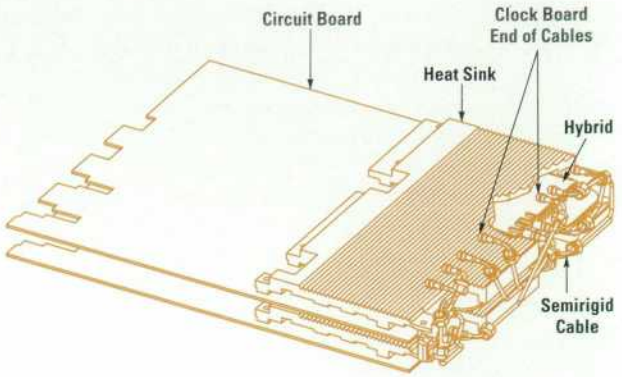


Fig. 8. Acquisition assembly with one heat sink partially cut away to show the thick-film hybrid circuit. Semirigid cables attach directly to the hybrids.

To solve the problem, the three boards are bolted together at the back with spacer blocks and stiffening brackets. After the front of the assembly is bolted into place, two stiff bars are snapped into the sides of the cardcage and bolted to the spacer bars. This constrains the assembly in the side-to-side and up-down directions while allowing enough clearance fore and aft so that the front-end alignment is maintained. One problem that was found during shock testing was that the bulkhead at the plug-in interface was being bent from slight fore and aft movement of the acquisition assembly. The assembly was bolted close to the side of the bulkhead where it was welded to the display compartment. The other side of the bulkhead (with the acquisition assembly bolt position farther from the edge) was not being damaged, so it seemed that making the problem side more flexible would solve the bending problem. Designing a clearance slot in the bulkhead and moving the welds farther away made the bulkhead resilient enough to be able to withstand shock and maintain alignment between the plug-in and the acquisition assembly.

Acknowledgments

We would like to recognize the efforts of others who made important contributions to the mechanical design effort. In our model and tool shop, Dave Burrows, Bill Wright, Ray Weddle, Herman Beeh, and Brian Hoff translated our ME 30 designs into prototypes, with a considerable amount of effort expended on the fabrication and welding of the steel package and the large, highly detailed front panel. Frank Leri and Pete Martinez provided the concurrent engineering link, creating the necessary assembly fixtures and processes, including the insertion and solder process for the 143-pin sampling hybrids. Bryan Gartner provided thermal analysis and test parts for the hybrids. Don Henry, who has left HP and started his own design business, contributed to the user interface and industrial design work. Greg Kruger provided help with the Monte Carlo analysis of the plug-in fit.

Reference

1. W. Thomson, *Theory of Vibration with Applications*, Prentice-Hall, 1981.

A Probe Fixture for Wafer Testing High-Performance Data Acquisition Integrated Circuits

This new probe fixture offers both a wide bandwidth and a high probe count, along with flexible interfacing and low maintenance. The fixture is used to perform at-speed wafer testing of the data acquisition circuits for the HP 54720/10 oscilloscope.

by Daniel T. Hamling

Advances in digitizing oscilloscope performance are necessarily accompanied by an increase in the speed and complexity of the oscilloscope's data acquisition hardware. Today's data acquisition front end, which performs the sampling and analog-to-digital conversion of the oscilloscope input, typically is a costly multichip module that supplies interfacing for several high-performance data acquisition integrated circuits (ICs). To reduce the cost of the multichip module, the IC reject rate after assembly must be reduced beyond what can be achieved with conventional low-speed IC wafer testing. However, performing the required at-speed (i.e., at operating speed) wafer tests of these complex, high-pad-count, high-bandwidth ICs is extremely difficult in an automated manufacturing test environment. Thus, testing an IC at its operating speed is now more necessary than ever, but more difficult to achieve.

While automatic test equipment (ATE) is keeping pace with most IC test needs, interface fixturing, particularly for wafer testing, is the source of most high-performance test limitations. Conventional probe technologies, while separately achieving either wide bandwidths or high probe densities, have not produced very useful combinations of these two important features. To test high-performance data acquisition ICs, a wide bandwidth and a high probe density are both needed. Therefore, an advanced probe fixture has been developed to support ongoing advancements in Hewlett-Packard's data acquisition technology.

Probe Fixture

The new probe fixture, described more extensively in reference 1,* is basically a new combination of existing technologies that provides interfacing for 28 high-frequency and 144 low-frequency signals between the ATE and the device under test (DUT). A simplified view of the probe fixture is shown in Fig. 1. The probe fixture consists of an aluminum oxide (Al_2O_3) substrate with a hole in its center from which a maximum of about 150 miniature tungsten (W) probes extend. The substrate not only serves as a probe carrier but also provides a state-of-the-art thick-film interface environment close to the DUT. Thus, with the probes being approximately the same length as a typical bond wire (i.e., equivalent to 3.5 nH), this portion of the probe fixture is electrically analogous to the multichip modules on which the data acquisition ICs are mounted. The substrate is mounted to the underside of a standard six-layer printed circuit board which provides electrical interfacing of the low-speed signals and mechanical support for the entire fixture. The 144 low-speed signals are brought from the substrate to the printed circuit board via a conductive elastomer pressed between the substrate and the board. This interface can supply up to 1A of current to each of the 144 signal paths.

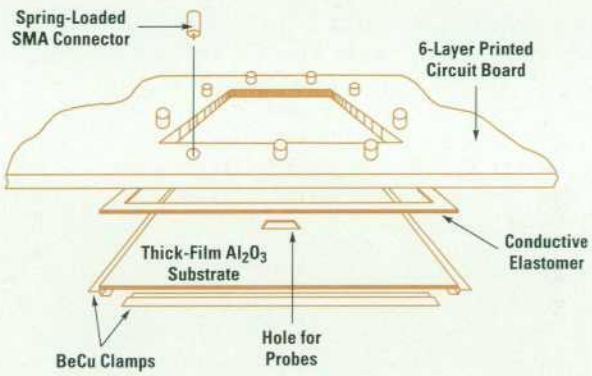


Fig. 1. Simplified perspective view of the new probe fixture.

The 28 high-speed signals are brought from the substrate to subminiature series A (SMA) connectors mounted in the printed circuit board via the spring-loaded signal pins of these connectors. This low-distortion interface, along with the controlled-impedance capability of the thick-film substrate, provides a –3-dB bandwidth of 3.2 GHz for the high-frequency ports. The threaded female barrel portions of the SMA connectors are conveniently placed about the printed circuit board so that high-frequency coaxial cables from the ATE can be attached easily to the fixture. A photograph of a partially assembled probe fixture is shown in Fig. 2. The custom substrate shown (without probes) is specifically for testing the sampler IC of the HP 54720/10 oscilloscope, as described later.

* Portions of this article were originally published in reference 1.

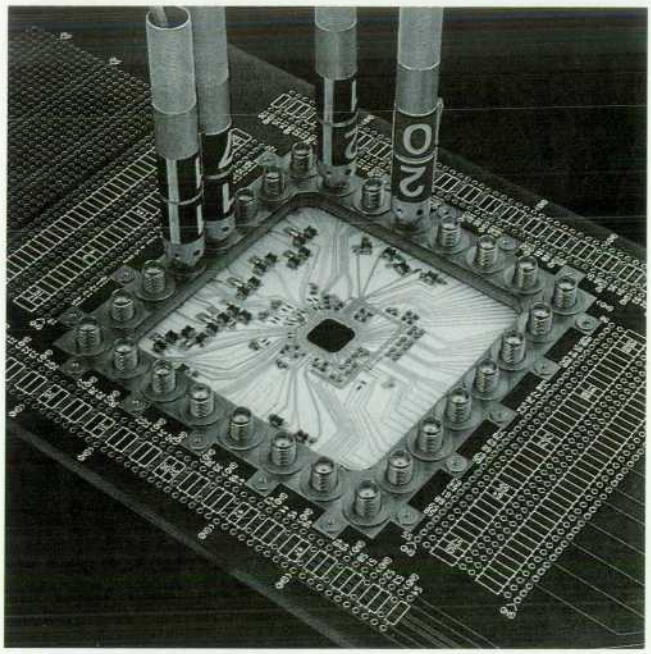


Fig. 2. Photograph of the new probe fixture without probes.

The most significant advantages and performance features of the new probe fixture are:

- 28 high-frequency, controlled-impedance signal paths up to 3.2 GHz
- 144 low-frequency, high-power signal paths up to 1A
- 150 maximum low-inductance W needle probes
- Fully customizable state-of-the-art thick-film interface circuitry close to the DUT
- Standard interfacing of all signals via the SMA connectors and printed circuit board ribbon cable connectors
- Use and mechanical performance comparable to conventional needle probe fixtures.

Sampler IC Wafer Testing

The performance of the probe fixture has been demonstrated by performing automatic wafer testing of the HP 54720/10 sampler IC (1DX4) in a typical manufacturing test environment. A similar probe fixture has been designed but not yet demonstrated for the other important HP 54720/10 data acquisition IC, the 1DX3 dual 7-bit ADC. Since the 1DX4 (along with two 1DX3s) is mounted in a costly hybrid assembly, at-speed wafer testing of this circuit is greatly desired. Results of the 1DX4 wafer test demonstration show that the probe fixture allows automated at-speed testing of a high-performance data acquisition circuit by providing convenient and low-loss interfacing of signals up to the circuit's full operating speed.

The 1DX4 is a fully custom, mixed-signal, 128-pad, 2-GHz-bandwidth bipolar integrated circuit. The 1DX4 is the front-end data sampler and the clock and timing generator of the high-speed data acquisition system incorporated in the HP 54720/10 oscilloscope. The HP 54720/10 data acquisition system is described in more detail in the article on page 11 and in reference 2.

To wafer test the 1DX4 at speed, the required high-speed signals provided via the probe fixture include a 2-GHz time base reference clock, a 100-MHz phase reference clock, and a 2-GHz-bandwidth analog input. The high-speed signals sensed via the probe fixture include a 100-MHz system clock, one of four 500-MHz sample clocks, and one of four 500-MHz postamplifier sample pulse outputs. Through automatic control and sensing of these signals, at-speed testing of the parameters shown in Table I is added to the conventional low-speed production test.

Table I
HP 54720/10 Oscilloscope Sampler At-Speed Test Parameters

Test Description	Test Specifications		Unit
	minimum	maximum	
VCO min frequency adjust		1000	MHz
VCO max frequency adjust	1000		MHz
100 MHz output frequency	95	105	MHz
DLL + delay range	460		ps
DLL – delay range		–265	ps
Programmed delay 1	115	135	ps
Programmed delay 2	240	260	ps
Programmed delay 3	365	385	ps
500-MHz output frequency	495	505	MHz
500-MHz output rise time		200	ps
500-MHz output fall time		200	ps
Sample clock phase adjust	40		ps
Sample maximum aperture	177		ps
Sample minimum aperture		163	ps
Sample-to-clock skew	455	760	ps
Input amplifier bandwidth	2		GHz
Postamplifier peak CMRR		–20	dB

VCO = voltage-controlled oscillator. DLL = delay-locked loop. CMRR = Common-mode rejection ratio.

Sampler Test Results

To provide a sense of the type of tests performed and their results, the 500-MHz output frequency and rise time, sample clock phase adjust, and sample maximum aperture parameter measurements from Table I will be discussed in more detail.

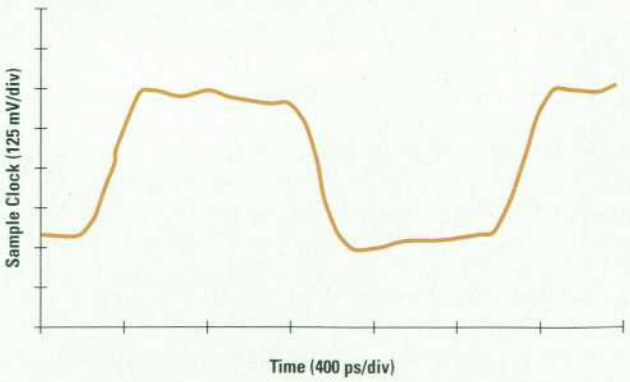


Fig. 3. 500-MHz sample clock output via the new probe fixture.

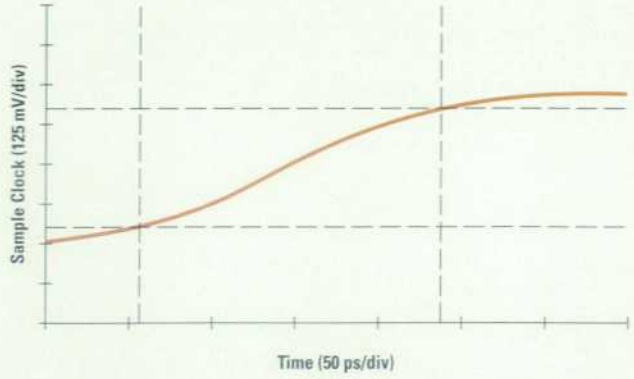


Fig. 4. 500-MHz sample clock output rising edge. t_r = 186 ps.

500-MHz Output Frequency and Rise Time. Fig. 3 displays the 500-MHz differential sample clock output waveform as seen by the ATE oscilloscope (HP 54120A) via the new probe fixture. During this test the 1DX4 is driven by the 2-GHz time base reference clock (not the VCO) so that the sample clock rates should measure almost exactly 500 MHz. When this waveform appears at the oscilloscope inputs, the test program asks the oscilloscope to measure the frequency and pass that value back to the test program for testing.

When the frequency measurement is complete, the test program asks the oscilloscope to find a rising edge of the waveform. This edge is then zoomed in on by decreasing the time scale to give the waveform shown in Fig. 4. Once zoomed in to increase measurement resolution, the test program asks the oscilloscope for the rise time measurement. For the waveform shown, the rise time is 186 ps. The same sequence is then executed for the fall time.

Sample Clock Phase Adjust. The sample clock phase adjust, with a nominal range of about 45 ps, is used to align the phases of the four sample clocks precisely. The adjustment range for a particular clock is determined from a low-pass filtered version of the exclusive-OR combination of the adjusted clock and one other fixed clock. Fig. 5 displays the histogram of 100 repeated measurements of the sample clock phase adjust of a typical 1DX4. As shown in the figure, the test exhibits an excellent resolution of 1.0 ps.

Sample Maximum Aperture. The sample aperture adjustment is used to set the sample aperture precisely to a desired

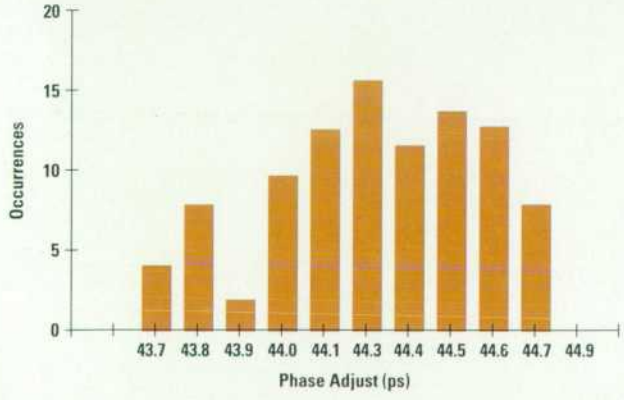


Fig. 5. Sample clock phase adjust measurement repeatability histogram.

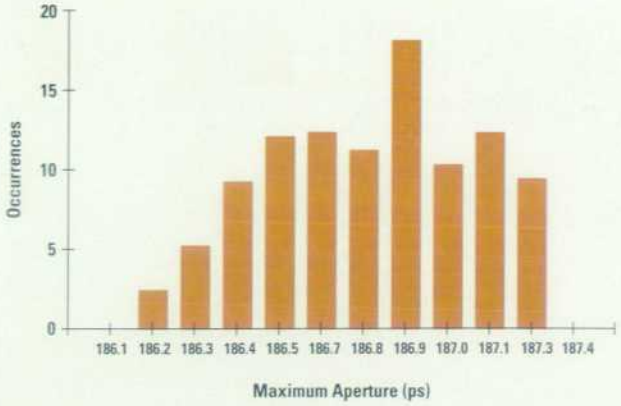


Fig. 6. Sample maximum aperture measurement repeatability histogram.

width, normally 170 ps. Since the aperture directly determines the bandwidth of the acquisition system, testing its adjustment range is crucial. With some on-chip circuitry to remove the gain from the sample pulse, the pulse width or aperture is calculated from the change in a low-pass filtered version of the sample pulse. Fig. 6 displays the histogram of 100 repeated measurements of the maximum sample aperture of a typical 1DX4. As shown in the figure, the test exhibits an excellent resolution of 1.1 ps.

Conclusion

A new probe fixture has been developed that allows at-speed wafer testing of high-performance data acquisition ICs. The probe fixture provides a wide bandwidth of 3.2 GHz, a maximum probe count of 150, a maximum high-frequency port count of 28, and mechanical performance comparable to conventional needle probes. The new probe fixture has been proved capable of performing at-speed wafer testing of the fully custom, 128-pad, 2-GHz-bandwidth 1DX4 sampler IC of the HP 54720/10 oscilloscope.

The new probe fixture provides a wide bandwidth, a high probe count, flexible interfacing, and low maintenance. The capabilities of the probe fixture are necessary to support the ongoing advances that Hewlett-Packard is making in data acquisition technology. With the testing capabilities that it allows, the new probe fixture can significantly benefit the design and manufacture of high-performance data acquisition ICs.

Acknowledgments

The author would like to thank Bruce Domen, Howard Iwane, Joe Garibaldi, Chris Schiller, Bart Jansen, Kathi Luiz, Buddy Yount, Dale Walz, and Bryan Gartner of the Hewlett-Packard Company as well as Jon Hicks and the engineers of Alphatronics Engineering Corporation for their contributions to this work.

References

1. D. Hamling, "A 3-GHz, 144-Point Probe Fixture for Automatic IC Wafer Testing," *Proceedings of the International Test Conference*, 1992.

2. C. Schiller and P. Byrne, "A 4-GHz 8-b ADC System," *IEEE Journal of Solid-State Circuits*, Vol. 26, no. 12, December 1991, pp. 1781-1789.

A High-Performance 1.8-GHz Vector Network and Spectrum Analyzer

Network and spectrum analyzers are frequently used together for RF component and circuit evaluation. The HP 4396A vector network and spectrum analyzer exploits this natural union by combining the two measurement modes into one instrument.

by Shigeru Kawabata and Akira Nukiyama

Frequently in RF component testing and circuit evaluation a network analyzer and a spectrum analyzer are used either together or alternately to make measurements. In many active component measurements, both instruments are necessary for measuring and analyzing frequency characteristics. A network analyzer measures the frequency responses of components, and a spectrum analyzer measures the power of signals.

For example, circuit designers design amplifiers to meet certain performance criteria such as input and output return losses, forward gain, reverse isolation, noise figure, harmonic or intermodulation distortion, and so on. These items are not independent, but are influenced by each other. If the gain is too high, the distortion performance may not be good. If the noise figure is very good, the input return loss may not be sufficient. Generally, many design iterations are necessary in these design processes, and since these measurements often involve both network measurements and spectrum measurements, designers have to exchange the two different analyzers frequently.

Because network and spectrum analyzers are frequently used together, it seemed very natural to us that the two analyzers should be combined together in one instrument. The HP 4396A network and spectrum analyzer is the result of our efforts in combining the two analyzers (see Fig. 1). The main features of the HP 4396A include:

Cost. The HP 4396A offers high-performance network and spectrum measurements at a reasonable price. It might be more expensive than an individual network analyzer or spectrum analyzer with performance characteristics similar to the HP 4396A, but it is much less expensive than the combined cost for two individual analyzers.

Space Efficiency. The size and the weight of the HP 4396A are smaller than the combined size and weight of an individual network analyzer and an individual spectrum analyzer. Many circuit blocks and components, such as, the CRT display, the power supply, the CPU, the synthesizer, and the receiver are used in common for both measurement modes in the HP 4396A.

Consistent User Interface. Since a network analyzer and a spectrum analyzer have similar measurement characteristics and are typically used together, it would seem to be convenient for users of these two measurement modes to have the same user interface. The user interface of the HP 4396A is designed to be consistent for both measurement modes.

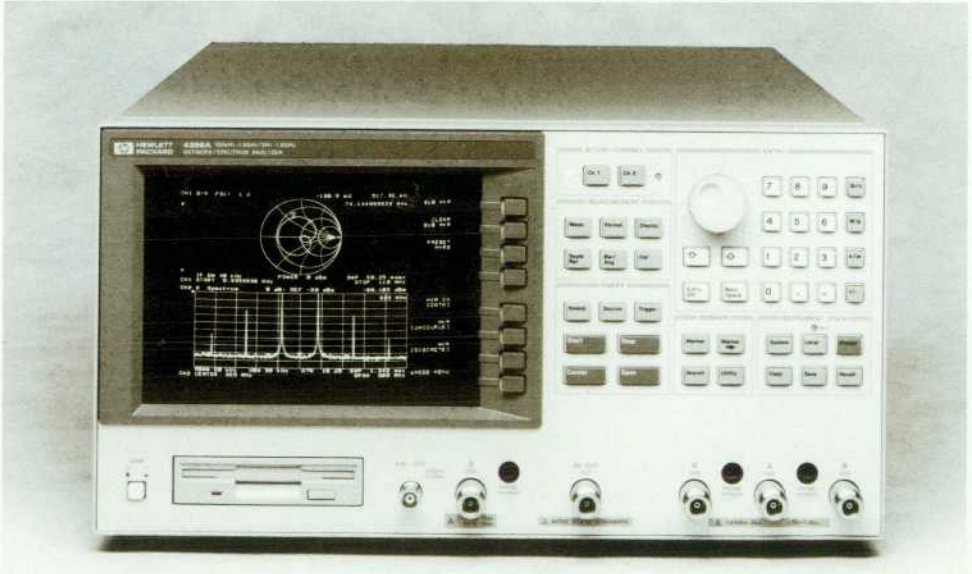


Fig. 1. The HP 4396A high-performance vector network and spectrum analyzer provides the capabilities of both types of analysis without compromising the performance of either. Options include burst-signal analysis and HP Instrument BASIC programming.

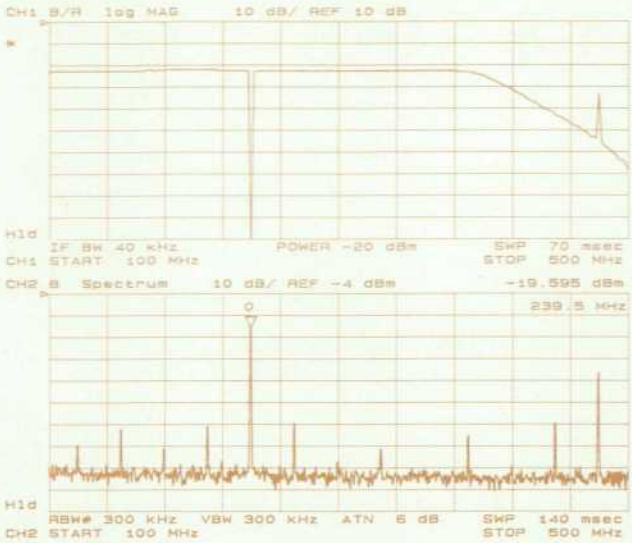


Fig. 2. The upper channel shows a network measurement in which the gain of an amplifier is being measured. The lower channel shows the spectrum of the signal being monitored at the output of the same amplifier. Without changing any connections, spurious signals that corrupt the amplifier gain measurement can be found by means of the spectrum monitor.

Common or similar functions are located under the same keys and have the same command names wherever possible. The HP 4396A's user interface is very similar to other conventional dedicated analyzers. The HP 4396A inherits its network measurement user interface from the HP 87xx Series of network analyzers, and its spectrum measurement user interface from the HP 85xx Series of spectrum analyzers.

Fewer Connection Changes. The R, A, and B ports of the HP 4396A are usually used as network measurement input ports. However, they can also be used as spectrum measurement input ports with some performance degradation. This type of measurement is called spectrum monitor. For instance, it is useful to check the self-oscillation of devices or spurious signals coming from outside devices under test during a network measurement. This can be done with the HP 4396A without changing connections; just turn off the RF output and check the spectrum using the other measurement channel. Fig. 2 shows an example in which a spurious signal is checked by channel 2 using the spectrum monitor while the network measurement of an amplifier is performed on channel 1.

Performance

The goal of the HP 4396A project was to develop a product that does not compromise performance in either network or spectrum measurements. In some cases, the HP 4396A has achieved better performance than some conventional network or spectrum analyzers. Wide dynamic range and high measurement throughput were considered to be two of the most important features of an excellent instrument. As a result, a great amount of development effort was put into providing these features in both measurement modes of the HP 4396A.

Network Measurement Performance

The HP 4396A can be used as a high-performance 100-kHz to 1.8-GHz network analyzer. Table I shows some of the HP 4396A performance values in the network measurement mode.

Wide Dynamic Range. The dynamic range and sweep rate of the HP 4396A in the network measurement mode are especially important features. More than a 110-dB dynamic range is guaranteed and typically more than a 120-dB dynamic range is achievable. These wide dynamic ranges can be obtained with a high sweep rate. Fig. 3 shows an example in which more than a 110-dB dynamic range is obtained in the measurement of a dielectric filter.

Fig. 4 is a plot of the dynamic range as a function of the sweep rate of the HP 4396A. More than a 95-dB dynamic range using 201 measurement points can be obtained within 0.2 second. The sweep rate is fast enough and the dynamic range is wide enough for most manual adjustments in a production line. If more dynamic range is required, the IF bandwidth can be narrowed at the cost of a slower sweep rate.

While the dynamic ranges at 10-kHz and 3-kHz IF bandwidths are almost the same, the sweep rate at 10-kHz bandwidth is more than two times faster. At 10-kHz and 40-kHz IF bandwidths, the design was tuned to get a wide dynamic

Table I
Typical Network Measurement Performance Values for the HP 4396A

Parameter	Typical	Conditions
Dynamic Range	> 120 dB	IF bandwidth = 10 Hz
Dynamic Accuracy		From –20 dB to –70 dB full scale at IF bandwidth = 10 Hz
Magnitude	< 0.02 dB	
Phase	< 0.15 degree	
Output Power Range	–60 dBm to +20 dBm (0.1-dB resolution)	
Power Sweep		
Range	20 dB (up to 30 dB depending on the stop power)	
Linearity	0.5 dB	50 MHz
Frequency	< 1-mHz resolution	
Calibration	Full one-port, full two-port, and one-path-two-port calibrations for type N, APC 7-mm or APC 3.5-mm connector type	

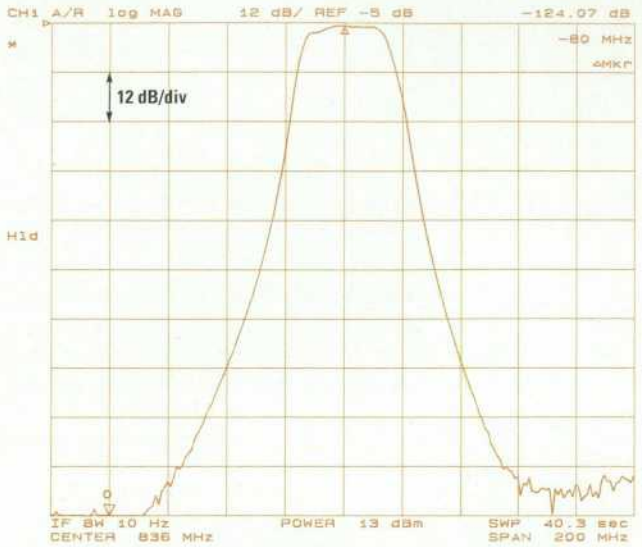


Fig. 3. The very wide dynamic range of the HP 4396A network measurement mode allows measurement of the 120-dB stop-band rejection of a dielectric filter.

range with a high sweep rate, whereas below 3-kHz IF bandwidth, image and alias responses are well-rejected for better high-Q device measurement.

Spectrum Measurement Performance

The HP 4396A is also a 2-Hz-to-1.8-GHz spectrum analyzer. Table II shows some of the key spectrum measurement performance values for the HP 4396A. The values in Table II show that the HP 4936A's spectrum measurement performance is as high as or even better than conventional spectrum analyzers.

Low System Noise. As shown in Table II, the average noise level is very low, especially at lower frequencies, thanks to the instrument's low conversion loss, double-balanced first mixer, and low-noise-figure first IF amplifier. These components make small-signal or low-noise measurements possible without attaching a low-noise preamplifier. Fig. 5 shows a measurement of about a –140-dBm signal using a 3-Hz resolution bandwidth.

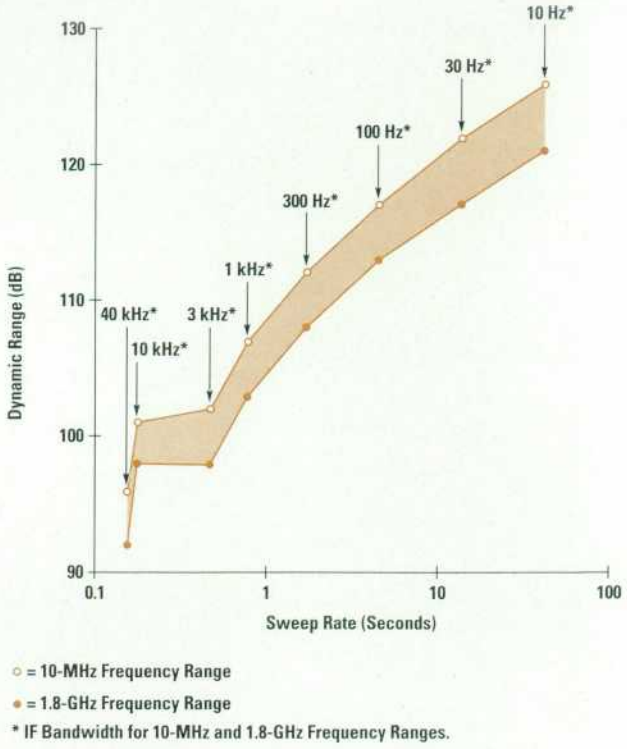


Fig. 4. The dynamic range of the network measurement mode is plotted as a function of the sweep rate and the IF bandwidth using 201 measurement points.

Good Scale Fidelity. The HP 4396A doesn't have an analog logarithmic amplifier like other conventional spectrum analyzers. Logarithms are calculated in firmware after analog-to-digital conversion. Thus, the HP 4396A has very good scale fidelity, which is almost the same as the dynamic accuracy in the network measurement mode.

Table II
Key Spectrum Measurement Performance Values

Parameter	Specification	Typical	Conditions
Average Noise Level	< –149 dBm/Hz	–155 dBm/Hz	10 MHz to 300 MHz, attenuator = 0 dB, reference level ≤ –40 dBm
Scale Fidelity	< 0.05 dB	< 0.02 dB	1-Hz to 3-kHz RBW, 0 to –30 dB from reference level, –50 dBm ≤ mixer level ≤ –20 dBm
RBW* Shape Factor	< 10 < 3	< 8.5 < 2.5	10-kHz to 3-MHz RBW 1-Hz to 3-kHz RBW
Third-Order IMD**	< –75 dBc	< –80 dBc	≥ 10 MHz, –30-dBm mixer level, two-tone with ≥ 20-kHz separation
Second Harmonic Distortion	< –70 dBc	< –80 dBc	≥ 10 MHz, –35-dBm mixer level
Residual Response	< –100 dBm	< –110 dBm	≥ 3 MHz, attenuator = 0 dB
Noise Sidebands	< –105 dBc/Hz	–115 dBc/Hz	1 GHz, 20-kHz offset
Frequency Response	< ±0.5 dB	< ±0.3 dB	≥ 10 MHz, attenuator = 10 dB, referenced at 20 MHz

* RBW = Resolution Bandwidth

** IMD = Intermodulation Distortion

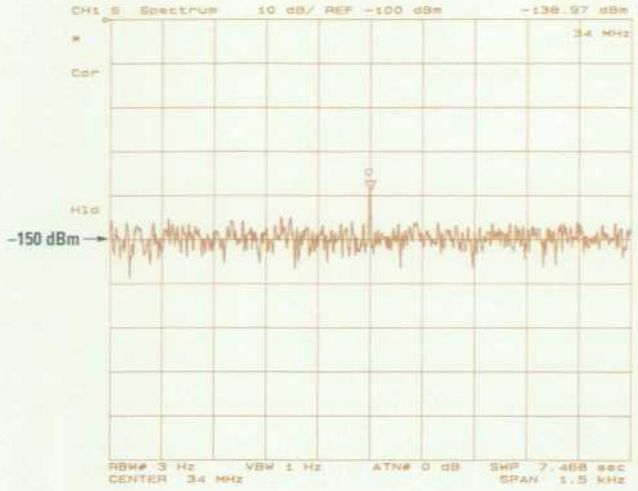


Fig. 5. Because of the high-sensitivity receiver design and the stepped FFT of the HP 4396A, a –140-dBm signal can be easily measured in 7.5 seconds.

High Sweep Speed at Narrow RBWs. Fig. 6 shows a plot of sweep time for various resolution bandwidths (RBW) of the HP 4396A compared to conventional, swept-type spectrum analyzers. When the RBW is narrow, the HP 4396A is much faster. Below 1-kHz RBW the sweep speed of the HP 4396A is 2 to 100 times faster than conventional swept-type spectrum analyzers. A stepped fast Fourier transform (FFT) technique produces this fast spectrum measurement. A signal of a certain frequency range is frequency converted and analog-to-digital converted to digital IF, and its spectrum is obtained simultaneously by using an FFT. The first local oscillator frequency is changed step-by-step to shift the input measurement frequency range to obtain the total spectrum over the required frequency span. See page 90 for more about the FFT techniques used in the HP 4396A.

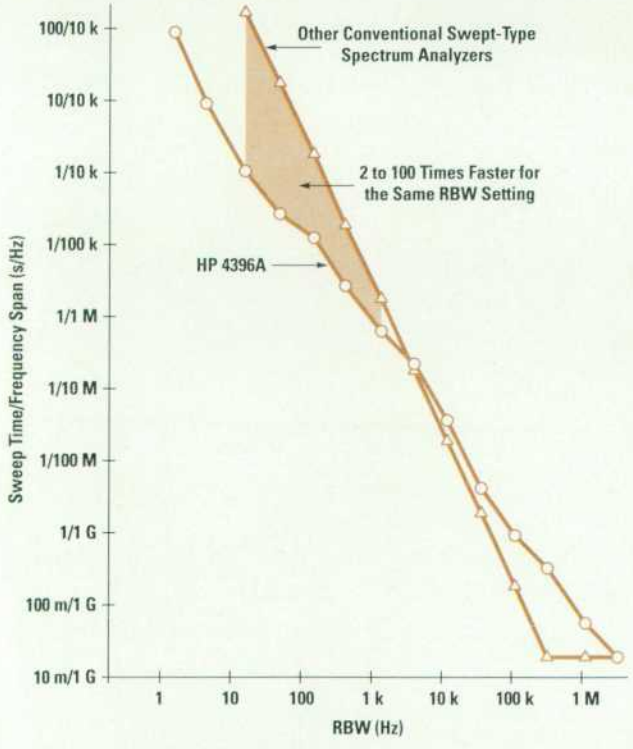


Fig. 6. The HP 4396A's sweep speed is compared with conventional swept-type spectrum analyzers.

In the measurement shown in Fig. 5, the sweep time of the HP 4396A is 7.5 seconds. If a conventional spectrum analyzer had been used at the same measurement settings, its sweep time would be more than 500 seconds.

Because of the low-noise design of the HP 4396A, a wider RBW can be used to realize the same sensitivity, which contributes to increasing the measurement throughput of the instrument. On the other hand, the HP 4396A is a bit slower than conventional swept-type spectrum analyzers at wider RBWs. However, the speed difference is not so noticeable because sweep time is much shorter at wider RBWs.

Gated Spectrum Measurement. The HP 4396A has a gated spectrum measurement option. This means that the instrument measures the spectrum of signals only while the measurement gate is turned on. The gate delay and the gate length can be set as low as 2 μs, which allows measurement of the spectrum of a very narrow-width burst signal.

Other Features

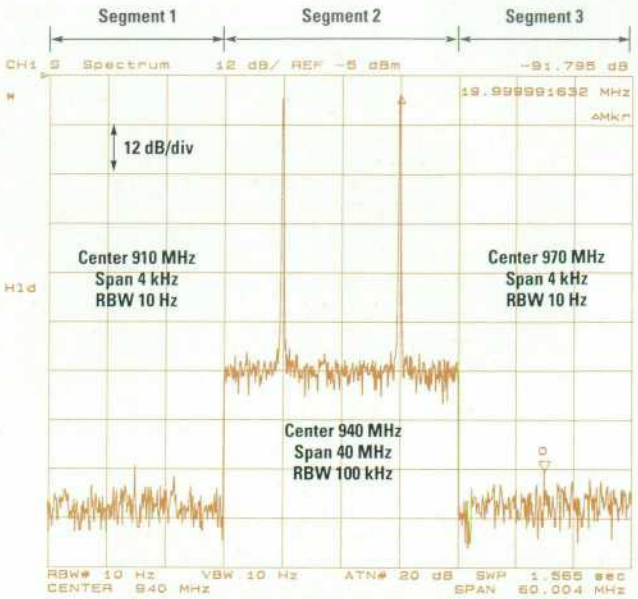
The following are features that are common to both the network and spectrum measurement modes of the HP 4396A.

List Sweep Capability. Many of the latest HP network analyzers have list sweep capability, which enables them to measure and display simultaneously several user-defined segments at different frequency ranges, numbers of points, IF bandwidths, and power levels. The HP 4396A has this capability not only in the network measurement mode but also in the spectrum measurement mode. List sweep is especially useful in the spectrum measurement mode.

As an example, consider a two-tone third-order intermodulation distortion (IMD) measurement. When the IMD level is very low, a narrow RBW should be selected to get enough sensitivity. If it is desired to measure widely separated two-tone fundamental signals and third-order IMD signals on the same screen, the sweep will become quite slow because of the narrow RBW necessary to detect the low-level IMD signals.

Fig. 7 shows this third-order IMD measurement using a list sweep of three segments. The fundamental two-tone signals are measured at the second segment which has a wide RBW, and the third-order IMD signals are measured at the first and third segments which have narrow RBWs and narrow frequency spans to shorten sweep time. The total sweep time is about 1.6 seconds. Without list sweep it would take more than two hours.

Measurement Channels. The HP 4396A has two measurement and display channels and each channel can be independently assigned to either a network measurement mode or a spectrum measurement mode. If channel 1 is assigned to a network measurement mode and channel 2 is assigned to a spectrum measurement mode, to change from one measurement mode to another it is only necessary to select the desired active channel key (**CH1** or **CH2**). Both measurements can be performed alternately by using the dual-channel mode. Since either channel works as an independent analyzer, the HP 4396A can be used as if it were two spectrum



(a)

Segment	Center	Span	Points	Power	RBW
>1	910 MHz	4 kHz	240	0 dBm	10 Hz
2	940 MHz	40 MHz	321	0 dBm	100 kHz
3	970 MHz	4 kHz	240	0 dBm	10 Hz
END					

(b)

Fig. 7. An example of the list sweep capability of the HP 4396A in which two-tone signals and their third-order IMDs are measured at the same time. (a) A list sweep display in the spectrum measurement mode. (b) List table.

analyzers, two network analyzers, or one spectrum analyzer and one network analyzer. The two channels can be associated with one another by using the cross-channel marker mode and the coupled-channel mode.

Versatile Marker Functions. The HP 4396A has many versatile, convenient marker functions like other HP spectrum and network analyzers. Among them, the cross-channel marker function is the most interesting. This function enables markers to be used to relate the upper and lower measurement channels on the display.

When a signal jumps to another frequency and close-in measurement of the signal is required, ordinary signal tracking will miss the signal because of the signal's large frequency jump and the analyzer's slow sweep speed. In such a case, cross-channel signal tracking is very useful. One channel is used to measure the signal with a wide frequency span to keep track of the signal. The other channel can be used to measure a close-in area of the signal, which is being tracked by the other channel.

The cross-channel marker zoom function can be another useful cross-channel function. One channel is used for the measurement of the zoomed-in area of a marker location while the other channel displays the original (before zooming) area. Fig. 8 illustrates use of the zoom function.

HP-IB Command Support. The HP 4396A supports two different types of HP-IB commands. One is the SCPI* command set,[1] and the other is the conventional HP-IB command set (IEEE 488.2). In most cases, the conventional command set is more compact and works faster than the SCPI command set. It may be easier for those who are familiar with the HP 87xx Series network analyzers or HP 85xx Series spectrum analyzers to use the conventional command set.

On the other hand, the SCPI command set was designed under industry-standard SCPI guidelines, and any SCPI-supported instruments have common HP-IB commands for the same functions.

Either SCPI or conventional HP-IB commands can be used to control the HP 4396A.

Controller Capabilities. To enable the HP 4396A to behave as an instrument controller, the instrument has HP Instrument BASIC (IBASIC), an external HP-HIL keyboard, and a general I/O port, which consists of IBASIC-controllable 4-bit inputs and 8-bit outputs. The combination of IBASIC programs and the HP-IB and general I/O ports can be used to build an instrument system consisting of the HP 4396A and other instruments without any need for an external computer. Programs can be easily written by means of the external keyboard and a command logging capability.

Storage Devices. The HP 4396A has a a built-in flexible disk drive and 448K bytes of internal volatile memory. The flexible disks can be used to store measurement data, instrument settings, and IBASIC programs. The flexible disk can read and write in DOS format for easy data communication with personal computers.

The volatile memory can be used for quick data storage and retrieval of data such as instrument settings. With a 448K-byte capacity, the memory can store up to six sets of 201 points of full two-port calibration data, which can be recalled instantly. The memory area is shared with IBASIC's work memory area, and the memory partitioning can be set by the user.

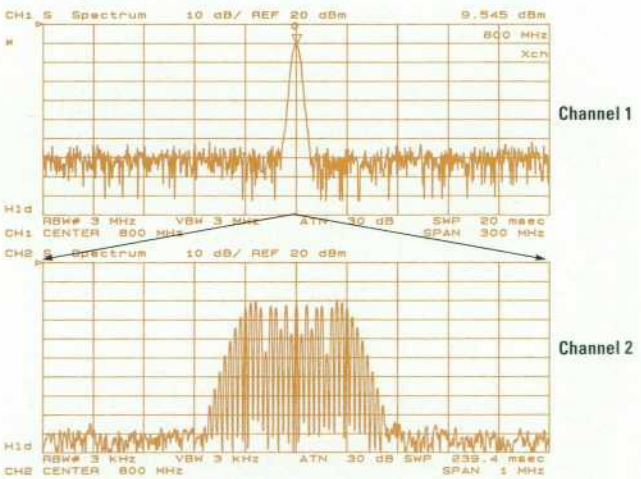


Fig. 8. An example of the cross-channel marker zoom function. In this case channel 1 is used for tracking a signal with a wide frequency span and channel 2 is used for zooming in to measure sidebands of the signal. Zoomed-in sidebands can be measured without missing a signal that is frequency hopping.

* SCPI means Standard Commands for Programmable Instruments.

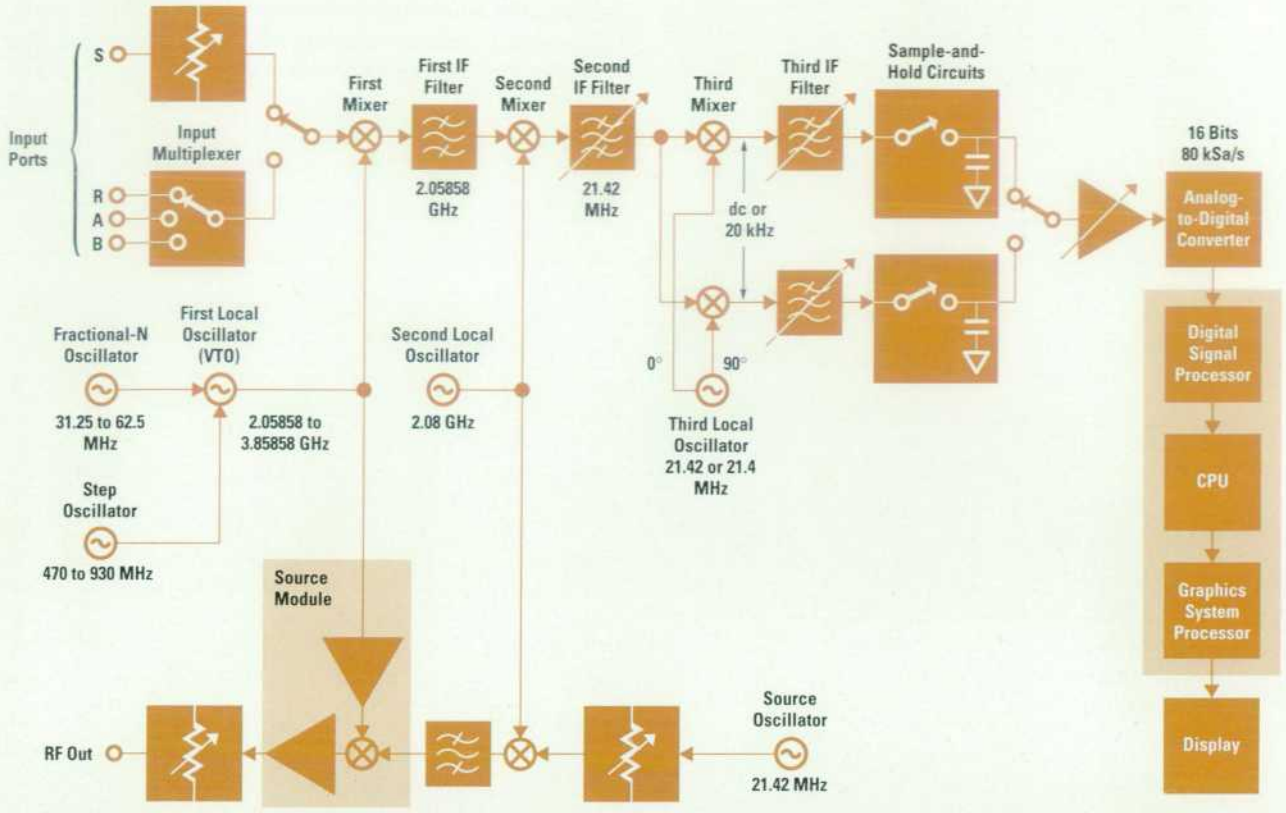


Fig. 9. Simplified block diagram of the HP 4396A.

Low-Cost Design

The main objective that guided the development of the HP 4396A was to provide a high-performance instrument but at a low cost. It is impossible to realize both of these contradictory themes in an ordinary system design. If a combined network and spectrum analyzer is built by just physically joining a stand-alone network analyzer and a stand-alone spectrum analyzer, the cost and the size would be prohibitive. Even sharing the cabinet, CRT display, the power supply, and the CPU would not decrease the cost of such a configuration.

To realize a very low-cost combination analyzer and at the same time keep the performance of network and spectrum measurements as high as possible required many new low-cost design techniques in the HP 4396A.

Receiver Design

Traditional vector network analyzers have three independent channel receivers for each of the three input measurement ports. While the cost for a network analyzer's receiver can be kept low, a spectrum analyzer's receiver is more expensive because it needs frequency upconversion and better distortion performance, making it too expensive to have three spectrum analyzer receivers.

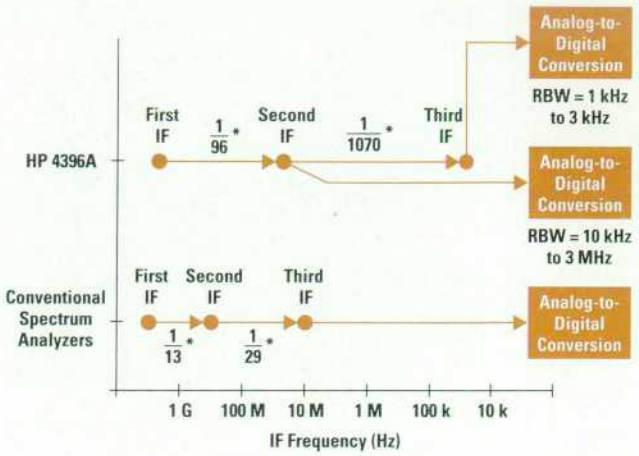
To keep receiver cost low, the HP 4396A has only one receiver, which is used for both vector network and spectrum measurements. To make vector network measurements possible, a three-channel multiplexer is used and measurements are performed using time division multiplexing. The three-channel multiplexer has a wide dynamic range and a high switching speed. The HP 4396A receiver is described in the article on page 85, and the three-channel multiplexer is described in the article on page 95.

Fig. 9 shows a simplified block diagram of the major components in the HP 4396A.

The HP 4396A doesn't use a logarithmic amplifier or narrow analog bandpass filters for RBW (resolution bandwidth) filters. RBW filtering below 3 kHz, video filtering, peak or sample detection, and logarithmic conversion are all computed by firmware after analog-to-digital conversion. This approach keeps the circuit size small and reduces the need for adjustments.

On the other hand, wide-RBW filters are realized with analog bandpass filters or low-pass filters with synchronous detectors in the HP 4396A. If a much higher-sample-rate analog-to-digital converter (ADC) were used, all the analog filters and the synchronous detection circuits could be eliminated and the total circuit size would be smaller and less adjustments would be needed. However, a very high-speed ADC with good linearity and the associated digital circuits needed to create the required RBW filters would be too expensive. Therefore, after considering the cost and performance balance, hybrid analog and digital RBW filters were chosen for the HP 4396A.

When a frequency down-conversion is performed by a frequency mixer, the ratio of the mixer output frequency to the input frequency usually ranges from 1/30 to 1/8. This ratio is mainly determined by the input filter's sharpness and the image-response rejection requirement.



$$\text{* IF Frequency Ratio} = \frac{\text{Output Frequency}}{\text{Input Frequency}}$$

RBW = Resolution Bandwidth

	First IF	Second IF	Third IF
HP 4396A	2.05858 GHz	21.42 MHz	20 kHz
Conventional Spectrum Analyzers	2 to 4 GHz	300 to 320 MHz	10 to 25 MHz

Fig. 10. IF frequency ratios of the HP 4396A compared to conventional spectrum analyzers.

The HP 4396A uses high-Q bandpass filters as IF filters and achieves quite small output/input frequency ratios. For the first-to-second-IF conversion, dielectric filters are used for the first IF filters and the frequency ratio is close to 1/100 (2.05858 GHz to 21.42 MHz). For the second-to-third-IF conversion, crystal filters are used for narrow-bandwidth measurements and the frequency ratio is below 1/1000 (21.42 MHz to 20 kHz). Both filters are small and have very high Qs and are adjustment free. Because of these filters some conversion stages and many adjustments have been eliminated, resulting in a low-cost, compact receiver IF section. Fig. 10 compares the frequency ratios for the IF sections of the HP 4396A and other conventional spectrum analyzers.

Synthesizer Design

For fine-resolution frequency generation, a digitally compensated fractional-N phase-locked loop is used. Its characteristics of low cost, small size, low noise, high speed, and no adjustments significantly contribute to making the HP 4396A production and service costs low.

A varactor-tuned voltage-controlled oscillator (VTO) is used for the first local oscillator instead of a YIG-tuned oscillator (YTO), which is typically used. One of the main reasons for using a VTO is cost. A VTO costs less than a YTO. The cost difference will be bigger in the future because a VTO has a much simpler structure. The driver circuits for the VTO can also be much simpler than those used for a YTO.

Digital Hardware Design

The Motorola MC68332 (16 MHz) and the Motorola DSP56001 (20 MHz) were selected as the CPU and the digital signal processor (DSP) for the HP 4396A. Since these chips have a lot of powerful on-chip peripherals as shown in Table III, they contributed to reducing the number of components, total component cost, board size, and hardware development costs for the HP 4396A. For controlling the CPU DRAMs, an off-the-shelf DRAM controller (74F1763A) was chosen, and for data communication between the CPU and the DSP, low-cost, dual-port IDT7134L SRAMs are used. The selection of these components helped reduce the number of parts and development costs and helped improve reliability.

Table III
Components on the MC68332 CPU Chip and DSP56001 DSP Chip

Component	Use
MC68332	
Time processor unit	Interrupt handler timer
System integration module	Address decoder
Quad serial module	Fractional-N chip interface
Serial communication interface	RS-232-C interface (for software debug)
DSP56001	
Host interface	Memory-mapped peripheral to interrupt handler
Synchronous serial interface	ADC serial interface
Address generation unit	Address decoder
Bootstrap ROM	Power-on bootstrap
Internal RAM	Fast-fetch program and data memory

Adoption of highly integrated and small packaged memory ICs, such as 4M-bit DRAMs in ZIP* packages and 1M-bit SRAMs in surface mount packages, have also reduced the number of parts and board area drastically and contributed to keeping the total cost low.

RF Shield Design

Typically, RF shielding tends to be expensive, especially molded shield blocks and dc feeding. A new low-cost shielding method was adopted for the HP 4396A RF modules.

Fig. 11 shows a representation of the components that make up an HP 4396A RF module. The module consists of two shielded RF printed circuit boards which are encased in one molded aluminum block to reduce the material cost of the module. The circuit side of a printed circuit board is attached to a piece of aluminum sheet metal and the component side is shielded and covered by one side of the molded block. Another printed circuit board is shielded in the same way but using the other side of the same molded block. Thus, the molded block is used as a shield cover for both printed circuit boards. One of the printed circuit boards has a 96-pin DIN connector on the bottom edge of the board which is connected to the motherboard. This connector is used for digital or low-frequency signals and dc power. RF connectors located on the upper side of the module are used for RF or other sensitive signals. Fig. 12 shows the two RF circuit boards and the aluminum block.

* Zigzag Inline Package.

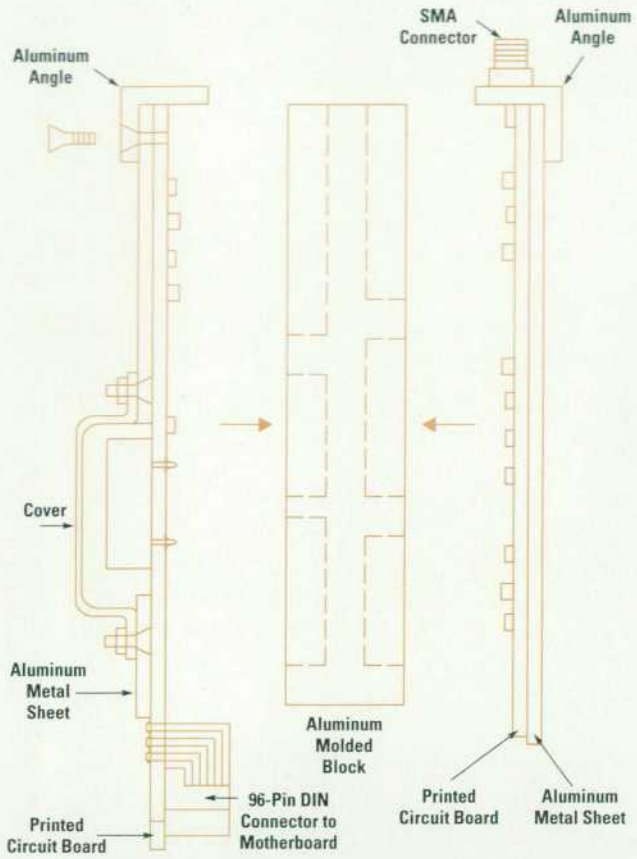


Fig. 11. A representation of the components that make up an HP 4396A RF module. One molded aluminum block is used for shielding between two RF circuit boards.

Surface mount type integrated low-pass filters are used as dc feedthroughs for HP 4396A RF modules. Fig. 13 shows the shape, equivalent circuit, and use of surface mount dc feedthrough filters, which are less expensive than a generally used screw-type dc feedthrough capacitor.

Isolation of the RF shielding structure with surface mount dc feedthrough filters was carefully investigated before it was adopted. Fig. 14 shows the experimental results from examining this RF isolation structure. No serious isolation problems have been observed in any HP 4396A RF modules.

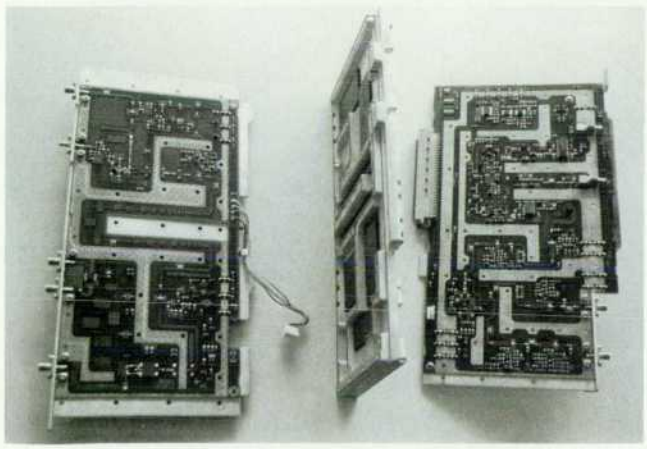


Fig. 12. Two RF circuit boards and the molded aluminum block.

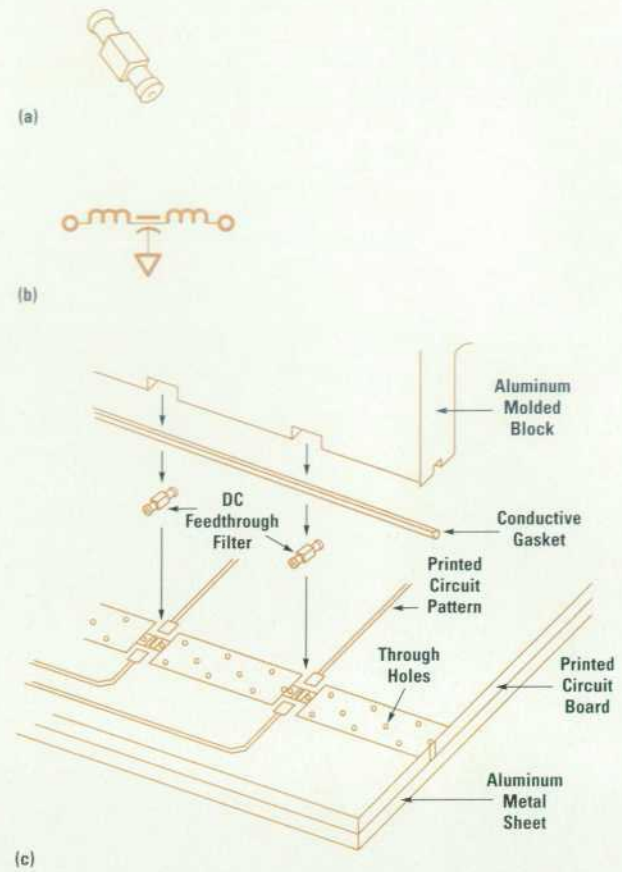


Fig. 13. Dc feedthrough on an HP 4396A RF module uses a surface mount dc feedthrough filter. This method contributes to cost reduction, productivity, and serviceability without degrading the shielding performance. (a) Feedthrough component. (b) Equivalent circuit. (c) Implementation.

A lot of cumbersome hand assembly operations can be eliminated by using this RF shielding structure. Surface mount dc feedthrough filters can be assembled using the standard surface mount assembly process, and printed circuit boards can be easily taken apart from the shield blocks without having to unsolder any parts.

Because of the implementation of these design considerations, the cost of the RF shield structure for the HP 4396A is very low.

Other Low-Cost Factors

More than 75% of the HP 4396A's electrical parts are surface mount devices, which reduces the board area and contributes to the total cost reduction. While the circuits in the HP 4396A deal with signal frequencies of almost 4 GHz, most of the RF circuits are built with surface mount and pattern components. Since microcircuits tend to be expensive, they are not used except for the source output module, which could not satisfy the performance requirements without using a microcircuit.

PPO (polyphenylene oxide) is used as the printed circuit board material for the HP 4396A's RF circuits. The dielectric loss of PPO lies somewhere between glass-epoxy and Teflon. PPO is more expensive than glass-epoxy but is much less expensive than Teflon. PPO printed circuit boards with

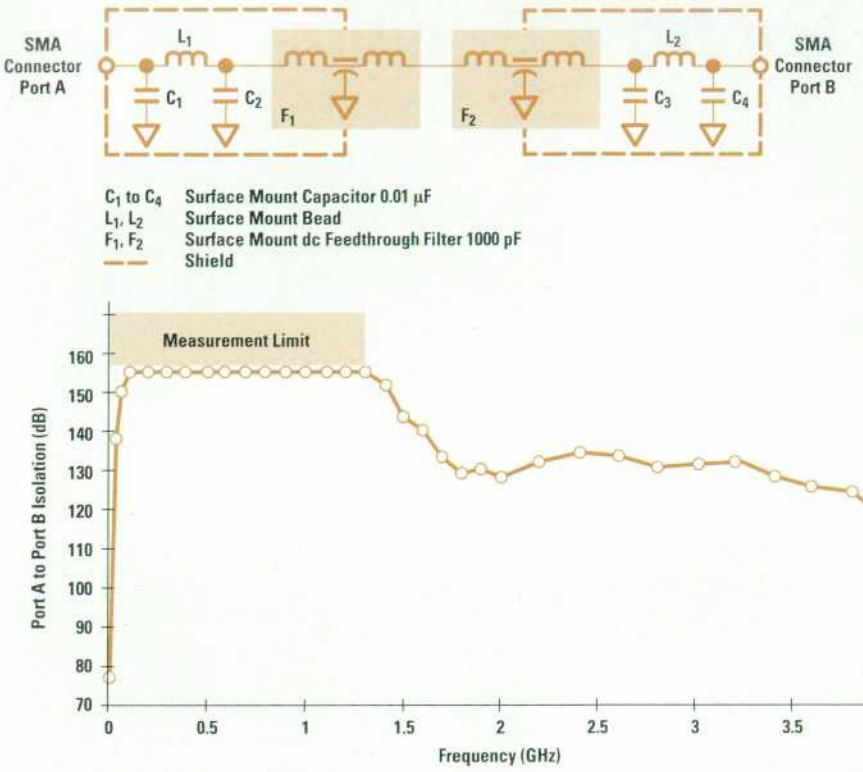


Fig. 14. Shield effectiveness of the RF structure using surface mount feedthrough filters.

through holes can be made using the same process used for glass-epoxy boards.

All interboard and module connectors and adjustment trimmers are accessible from the top, bottom, or side of the cabinet. No extender board is necessary for adjustments. This provides high productivity and serviceability, which contributes to cost reduction.

Acknowledgments

The HP 4396A design team members who deserve special recognition are Yoshiyuki Yanagimoto, who is the leader of the analog group, for the receiver RF section, Atsushi Ishihara for the receiver IF section, Koji Harada and Troy Morin for the synthesizer section, Tomokazu Furukawa for the signal source section, Tatsuo Furukawa for the power supply and the input multiplexer, Dean Nicholson and Kosuke Miyao for the source output module, which was developed at HP's Microwave Technology Division, Kazuhiro Matsui for the system test and environmental test, Akira Nukiyama, the leader of the digital hardware and firmware group, Yoshiharu Nakayama and Jun Shibata for the digital hardware and firmware, Masatoshi Yamamoto and Seishi Onishi for firmware, Norio Nakano for mechanical design, and Ken-ichi Katoh for industrial design. Special thanks are also due Kazuyuki Yagi, Tomio Wakasugi, Jun Kadowaki, and Hideyuki Takahashi for their intensive management. Many people from R&D, marketing, production, and product assurance contributed to making the HP 4396A successful.

Reference

1. S. Narciso and G. Hill, "The VXIbus from an Instrument Designer's Perspective," *Hewlett-Packard Journal*, Vol. 43, no. 2, April 1992, pp. 15-17.

Receiver Design for a Combined RF Network and Spectrum Analyzer

A low noise floor, reduction of distortion and local oscillator feedthrough, and removal of image frequencies and higher-order harmonic products were the main design objectives for the HP 4396A receiver.

by Yoshiyuki Yanagimoto

To keep receiver cost low, the HP 4396A 1.8-GHz vector network and spectrum analyzer uses only one receiver for both network and spectrum measurement modes. This article describes the design and implementation of some of the main components of the HP 4396A receiver.

A block diagram of the receiver portion of the HP 4396A is shown in Fig. 1. The receiver is located on the HP 4396A's RF converter board and IF board. A signal enters the RF converter board and hence the receiver via the GaAs FET switch either through the S (spectrum) input after the step attenuator or through the R, A, or B ports located at the input to the input multiplexer. The selected input signal enters the first mixer through the input low-pass filter. The signal is converted to the first IF frequency (2.05858 GHz) in the first mixer by the local oscillator (LO) whose frequency range is between 2.05858 GHz and 3.85858 GHz (tuned to correspond to RF input frequencies between dc and 1.8 GHz). A dielectric bandpass filter with a dummy load is connected to the first mixer to reject LO feedthrough and the unwanted products that might distort the first IF amplifier.

A 17-dB gain is provided by the first IF amplifier. The noise performance of the receiver is mostly decided at this point. The total sum of the conversion loss and the insertion loss of the circuits before the first IF amplifier is between 13 and 17 dB. The input noise of the IF amplifier gives a system noise floor of about –155 dBm/Hz to –151 dBm/Hz.

After going through a cascaded low-pass filter that rejects the higher harmonic products, the amplified signal is amplified again by an identical amplifier and then converted to the second IF frequency (21.42 MHz) by the second 2.08-GHz local oscillator in the second mixer.

The second IF is converted to either dc or 20 kHz, depending on the measurement mode, and then converted to a digital signal by the 16-bit, 80-kHz analog-to-digital converter (ADC). The digital signal is transferred to the digital signal processor (DSP) chip. The DSP calculates vector ratio, FFT, logarithm, average (digital filtering), and display decimation. The video filters are also implemented in the DSP chip.

First Mixer

The characteristics of the first mixer in the HP 4396A set the standard for almost all the distortion and noise performance of the instrument's network and spectrum measurements. For this reason a lot of time and effort was put into the design of the first mixer.

The input noise floor of the receiver is mostly determined by the first mixer. The conversion loss is the most significant factor affecting the input noise. Distortion is another very important factor. Harmonic distortion is generated in the first mixer, while third-order intermodulation distortion (IMD) is generated in all of the stages of the receiver.

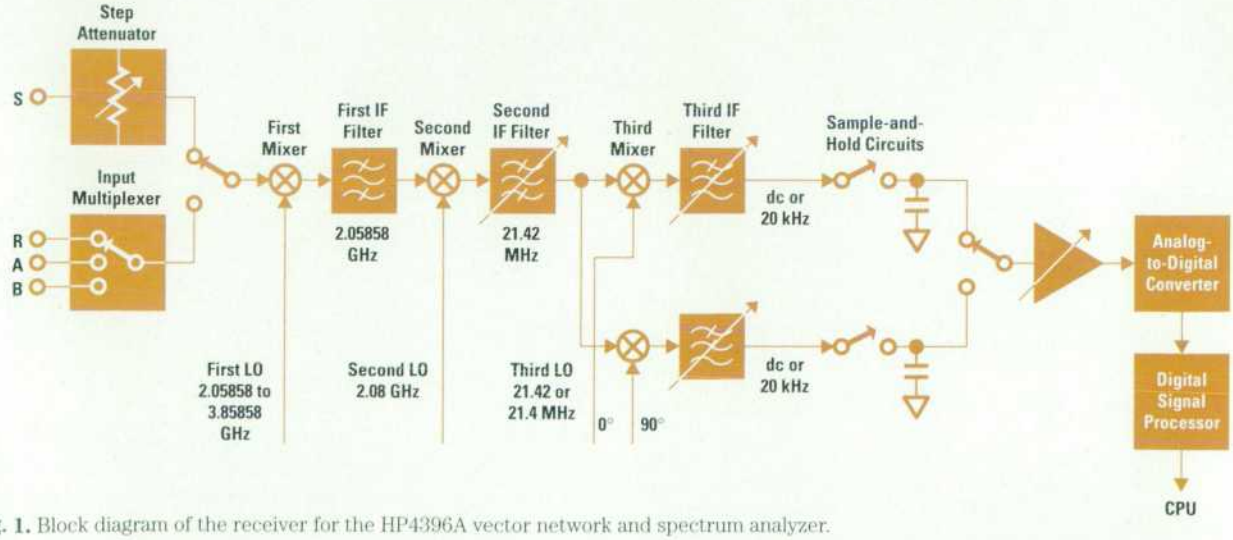


Fig. 1. Block diagram of the receiver for the HP4396A vector network and spectrum analyzer.

Gain compression is also caused by the nonlinearity of the mixer and the succeeding circuits. For a network analyzer, compression is a more frequently used concept than distortion, which is more commonly used in relation to a spectrum analyzer.

Dynamic range is defined as the ratio of the maximum input level to the equivalent input noise floor. The lower the input level, the smaller the distortion and the compression become. The maximum input level is defined so that the distortion or the compression is within the instrument's specifications. To meet compression specifications, the maximum input level for the HP 4396A's first mixer is –10 dBm. For two-tone input signals the maximum input level is –30 dBm to meet the third-order IMD specification of –80 dBc.

Fig. 2 shows the conversion loss of the first mixer and the input low-pass filter. The first mixer and other circuits, such as the input attenuator, the input switch, the low-pass filters, and the cables, make the total loss (sum of the insertion loss and the conversion loss) about 9 dB to 13 dB. The dielectric filter right after the mixer gives another 4-dB loss before the signal is amplified by the first IF amplifier. The noise figure of the first IF amplifier and the overall loss that accumulates before the amplifier sets the system noise floor for the HP 4396A at –155 dBm/Hz to –151 dBm/Hz.

Two approaches can be used to make the dynamic range wider. One is to increase the maximum input level. The other is to reduce the conversion loss, that is, lower the equivalent input noise. The approach chosen for the first mixer in the HP 4396A was to reduce the conversion loss.

It is well known that a double-balanced mixer has less conversion loss than a single-balanced mixer. However, HP had never made a double-balanced mixer in this frequency range. This is because the complexity of a double-balanced mixer makes the second-order harmonic distortion and the LO feedthrough worse than a single-balanced mixer, which has a much simpler configuration.

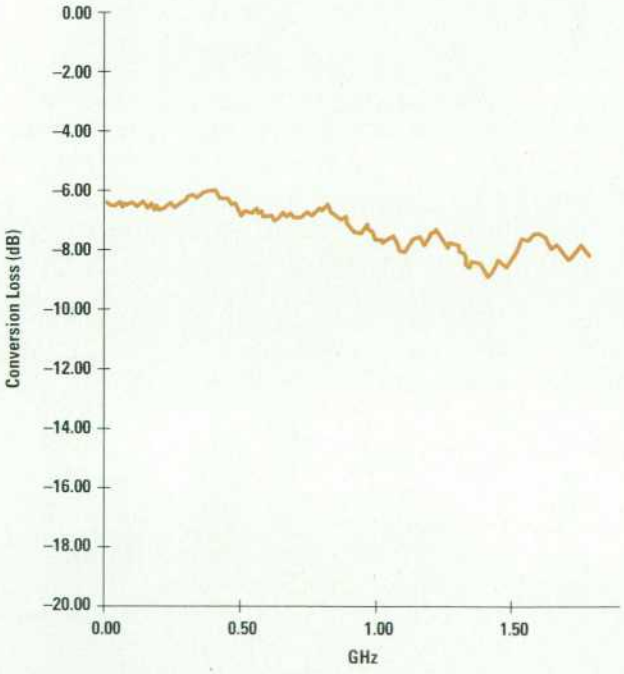


Fig. 2. Conversion loss of the double-balanced first mixer and the input low-pass filter. The double-balanced mixer design reduces the conversion loss and consequently increases the sensitivity (lowers the noise floor) of the HP 4396A.

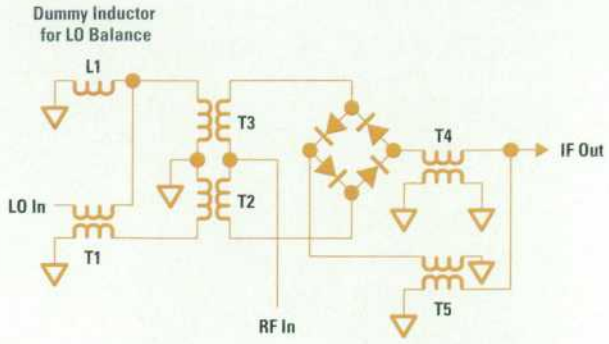


Fig. 3. Schematic diagram of the first mixer.

Implementation of the First Mixer

Regardless of the complications, we put our efforts into implementing a double-balanced mixer with low second-order harmonic distortion and low LO feedthrough. We designed the mixer so that it has complete electrical and physical symmetry.

Two points about second-order harmonic distortion and the LO feedthrough characteristics of double-balanced mixers are:

- The common-mode component of the LO signal going into the mixer diodes should be removed completely to reduce second-order distortion and LO feedthrough.
- Balance of the LO signal is obtained by the physical symmetry of the LO driving circuit if the even-order distortion of the LO signal is thoroughly removed by other circuits.

If the LO signal drives the mixer (diode ring) with perfect symmetry, which means there is absolutely no common-mode signal present, the residual cause of second-order distortion and LO feedthrough is a lack of balance in the four mixer diodes. Our experiments showed that the main cause of distortion and LO feedthrough is not the mixer diodes, but the physical asymmetry of the driving circuit and the consequent LO unbalance. Thus, our efforts focused on getting physical symmetry.

The schematic for the first mixer in the HP 4396A is shown in Fig. 3, and its layout on the printed circuit board is shown in Fig. 4. The input RF signal, after going through the input low-pass filter, is split into two 100-ohm microstrip lines to two nodes (B and D) of the four-node diode ring. The LO is balanced by the first balun (T1), and balanced LO signals are added to the RF input on the two 100-ohm microstrip lines by the other baluns (T2 and T3). The two signals being fed into the B and D nodes of the diode ring are $V_{in} + V_{LO}/2$ and $V_{in} - V_{LO}/2$, where V_{in} is the input signal voltage and V_{LO} is the LO voltage.

The LO signal turns on one side of the diode ring (say, D-A and A-B) and the output voltage at the center node A is V_{in} since $+V_{LO}/2$ and $-V_{LO}/2$ should be canceled at the center node according to the relationship above. When the other side of the ring (B-C and C-D) are turned on, the same V_{in}

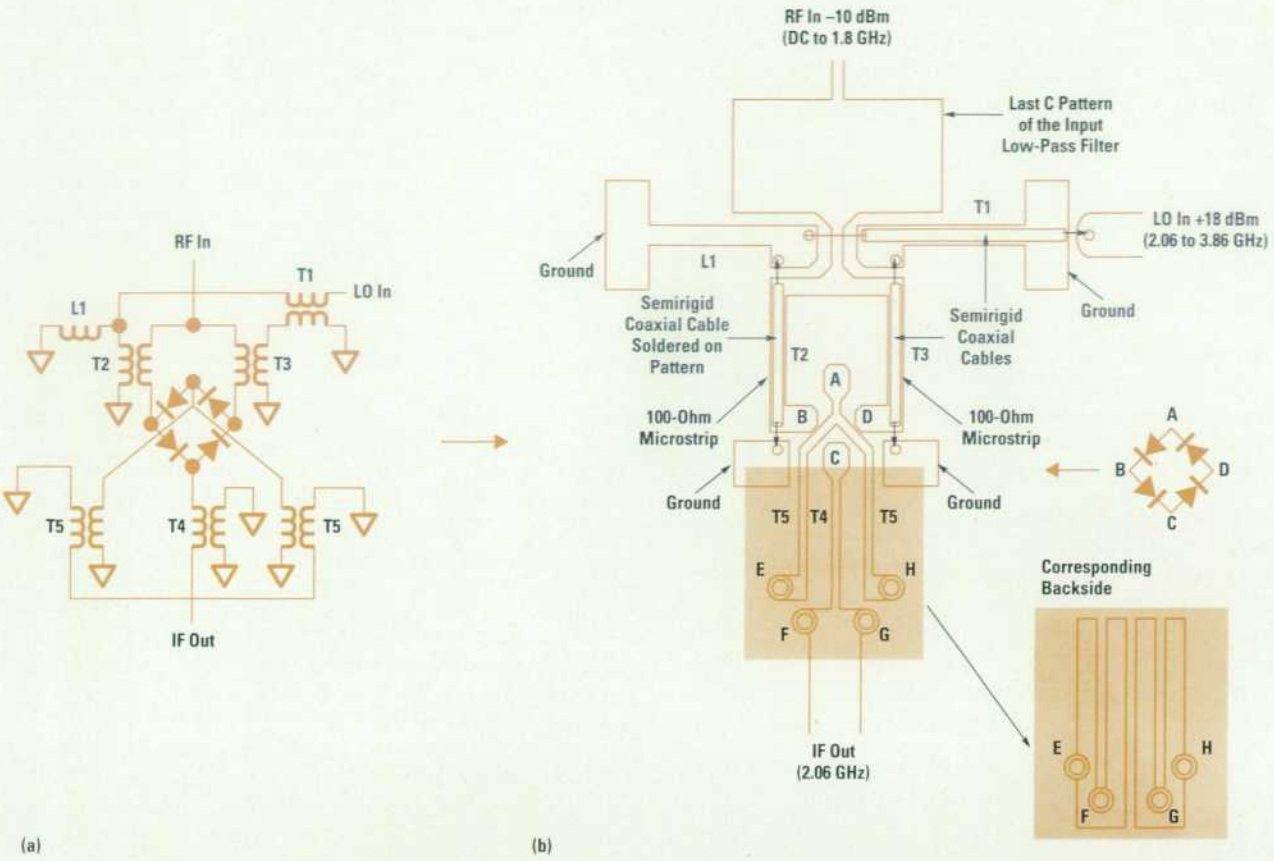


Fig. 4. (a) The layout of the first mixer based on the schematic shown in Fig. 3 and (b) the physical layout of the first mixer. The double balanced mixer is implemented on the printed circuit board in such a way that its layout has complete symmetry.

appears at the other center node C. Finally, one of the two outputs (C) is connected through T4 to the IF output node, and the other output (A) is inverted by T5 and connected to the IF output node.

The baluns T1, T2, and T3 are made with semirigid coaxial cable soldered on the pattern (T2 and T3 are soldered on the two 100-ohm microstrip lines). T4 is made with narrow patterns on both sides of the board facing each other. T5 consists of two baluns on the right and the left sides of T4 to maintain symmetry. The inverted outputs of T5 (ground lines promoted to signal lines) are connected to the IF output node.

The operation of this circuit is typical for a double-balanced mixer. Some important points to note about this circuit are:

- The balanced LO on the RF input signal side should be very well-balanced so that the output doesn't have the LO feedthrough component. An unbalanced LO will turn on one of the two diodes, causing second-order distortion.
- The output should not destroy the balance. One of the two output transformers (T5) is actually composed of two transformers in parallel to ensure good symmetry.

It was also found that gluing two pieces of microwave absorber on T4 and T5 on both sides of the board reduces distortion and LO feedthrough. This is probably because the LO signal coming into T4 and T5 through air, which would cause an unbalance of the driving signal, is absorbed by the absorber material.

These efforts reduce second-order harmonic distortion to less than –70 dBc at –30-dBm input level and LO feedthrough to less than –15 dB from the maximum input equivalent level.

First IF Amplifier Design

The first IF amplifier is a GaAs FET amplifier with the appropriate impedance matching. It has a 17-dB gain centered at 2.06 GHz.

The IF amplifier was designed using the HP Microwave Design System (MDS), which is a very powerful software tool for microwave design and simulation. Although the designers did not have a great deal of experience with the MDS, the IF amplifier was designed very quickly without any trouble. The short design time was accomplished with the prototyping system, which links the MDS data file with printed circuit board fabrication. This system allowed one-day delivery of the prototype board, or in other words, the simulated circuit became a printed circuit board the next day.

First IF Filter Design

IF filters are used for three main reasons. One is to reject the image frequency of the second IF. Usually, a bandpass filter is used for this purpose. Another is to remove all the higher harmonic products that are generated in the first mixer. A low-pass filter is used for this purpose. The third reason is to prevent LO feedthrough and unwanted products of the first mixer from distorting the IF amplifier. This is

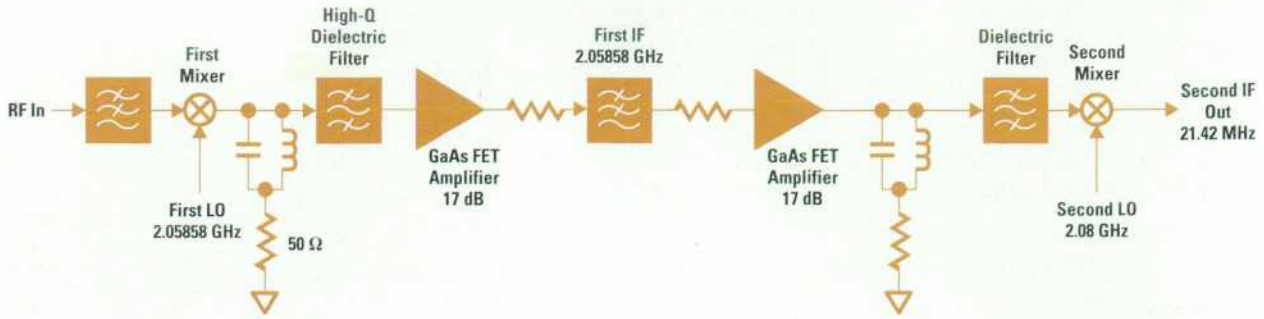


Fig. 5. Block diagram of the first IF section. All of these components are located on one board.

solved by locating the bandpass filter mentioned above between the first mixer and the IF amplifier. The first IF block in the HP 4396A is shown in Fig. 5.

The image-rejecting bandpass filter is essential in a spectrum analyzer. For example, the 2.05858-GHz first IF must be converted down to the second IF, 21.42 MHz, with the second LO at 2.08 GHz. The image frequency in this case is calculated to be 2.08 GHz + 21.42 MHz = 2.10142 GHz. If the image-rejecting filter were not in place, two signals would appear on the display 42.84 MHz away from each other. One would be real and the other would be an image. The 2.10142 GHz should be removed completely by the IF bandpass filter with 2.05858 GHz being passed through. The requirements for the IF filter are very strict and had it not been for the high-Q dielectric filters, the IF filter stage could not have been built. Actually, two dielectric filters in series are used to reject the image. A major advantage of this design is that one conversion stage was omitted that would otherwise be necessary. This contributed greatly to lowering the cost of the receiver design.

A harmonic of the first LO (2.05858 GHz to 3.85858 GHz) and a harmonic of the second LO (2.08 GHz fixed) can mix together to create a signal equal to the frequency of the second IF (21.42 MHz). This will cause a residual response. Fig. 6 shows an example of a residual response caused by mixing the fifth and seventh harmonics from two oscillators. If the instrument is tuned to 857.704 MHz, the first LO is set to 2916.284 MHz. The fifth harmonic of this is 14581.42 MHz. The seventh harmonic of the second LO is 14560 MHz. The difference of these two harmonics becomes 21.42 MHz, which would be detected as a fake signal. The fifth-seventh harmonic is just one of many harmonic combinations.

This higher harmonic mixing mechanism is one of the main causes of residual responses. Since the possible harmonic numbers are up to infinity, isolation between the first mixer and the second mixer is very important even at unused frequency ranges. Therefore, it is desirable that the first mixer and the second mixer be located on separate blocks. It is also necessary to insert a low-pass filter between the two mixers. Most spectrum analyzers have an independent low-pass filter between the first mixer block and the second mixer block.

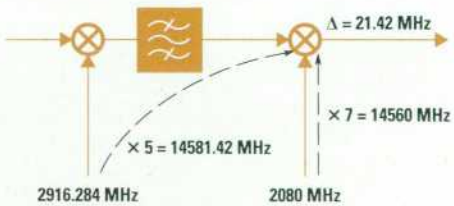


Fig. 6. One example of higher harmonic mixing causing a residual response. Mixing of the fifth and seventh harmonics generates a 21.42-MHz IF frequency when the first LO is 2916.284 MHz.

The HP 4396A includes all of the filters needed to reject all three of the undesired effects mentioned above on one board. A well-designed, multistage, low-pass filter and a low-cost, high-efficiency RF shielding method made this possible. The resulting residual response in the HP 4396A is –100 dBm for customer specifications and less than –110 dBm at production. The spurious response specification is –70 dBc.

IF Detection

The 21.42 MHz from the second IF filter is converted to lower frequencies so it can be handled by the 80-kHz, 16-bit ADC. Depending on the instrument settings, there are three modes for IF detection:

- Dc sampling mode, which is used in the spectrum measurement mode for wider resolution bandwidths (RBWs)
- FFT (fast Fourier transform) mode, which is used in the spectrum measurement mode for narrower RBWs
- Ac sampling mode, which is used in the network measurement mode.

DC Sampling Mode. The components involved in providing the dc sampling mode in the HP 4396A are shown in Fig. 7. This mode is used during spectrum analysis with resolution bandwidths in the range 10 kHz ≤ RBW ≤ 3 MHz. In this mode, the signal from the second IF filter is sent to the two mixers shown in Fig. 7. The LO frequency for this stage is the same 21.42 MHz passing through the second IF filter. The two local oscillators have in-phase and quadrature-phase relationships so that the two IFs in this section are the in-phase and quadrature components of the second IF.

The in-phase and quadrature components are sampled simultaneously by the two sample-and-hold circuits. The sum of the square of these two components is calculated by the digital signal processing (DSP) chip and then transferred to the CPU. The CPU detects the data as the input power. RBWs of 1 MHz and 3 MHz are shaped in the second IF stage (21.42 MHz) by bandpass filters, and RBWs of 10 kHz, 30 kHz, 100 kHz, and 300 kHz are shaped by the switchable low-pass filters in the third IF section.

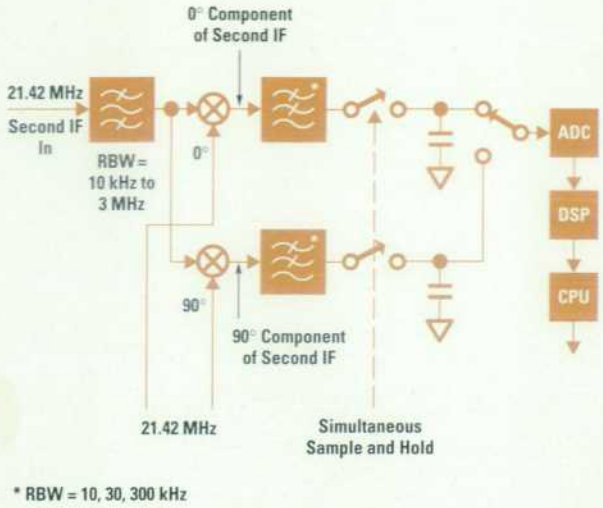


Fig. 7. Components involved in providing the dc sampling mode for spectrum analysis with wider RBWs. The in-phase and quadrature components are captured by the sample-and-hold circuits at the same time. The DSP chip calculates the power.

FFT Mode. The values and behavior of components in the FFT mode are shown in Fig. 8. This mode is used when the RBW is set between 1 Hz and 3 kHz. The FFT mode provides two great advantages. One is that the digital FFT filter replaces the analog filters for narrower RBWs. The analog filters would have required a very careful design and many precise adjustments resulting in a higher production cost. The other advantage of the digital FFT filter is the sweep speed. The sweep speed is very fast because the FFT algorithm can get information over a 10-kHz bandwidth at one time with the desired resolution, while conventional IF detection provides information at only one frequency at a time. The FFT method is described in more detail on page 90.

Only one of the two mixers in the third IF section is used in the FFT mode. The third LO is tuned to 21.4 MHz so that 20 kHz can be provided to the third IF section. This IF actually has a bandwidth of 10 kHz (15 kHz to 25 kHz) to provide the wide information bandwidth for the FFT mode. The 16-bit,

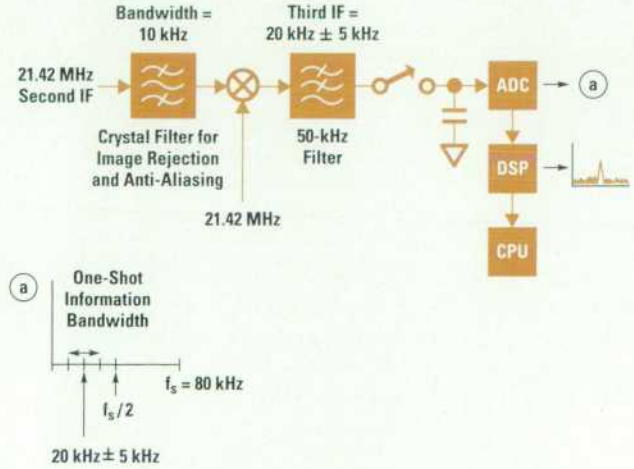


Fig. 8. Components involved in providing the FFT mode for spectrum analysis with narrower RBWs. The effective information bandwidth in this mode is 15 to 25 kHz. The DSP chip does the decimation and the FFT calculation.

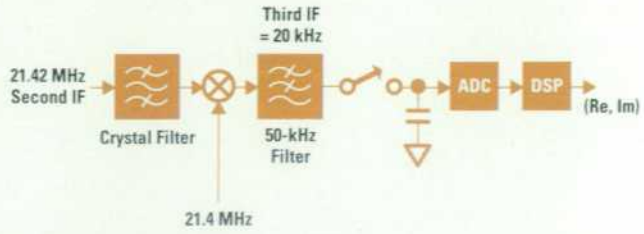


Fig. 9. Components involved in providing the ac sampling mode for network analysis. The 20-kHz IF is sampled and held at 80 kHz. The vector information is generated by the DSP chip, and the CPU receives the real and imaginary components of the IF signal.

80-kSa/s ADC samples the signal coming from the third IF filter, and the DSP chip calculates the FFT. The CPU treats the data from the DSP as the power spectrum of the input signal over the effective frequency range. To eliminate image and aliasing responses, a crystal filter is used at the second IF stage. Distortion products in the third IF stage are eliminated by the 50-kHz low-pass filter.

AC Sampling Mode. The ac sampling mode is shown in Fig. 9. When the instrument is configured as a network analyzer, the third IF section is set to 20 kHz using one of the two mixers in this section. With an IF bandwidth of 40 kHz (the fastest sweep) the ADC takes four data points to generate vector information as shown in Fig. 10. For narrower IF bandwidths, the vector data is averaged a number of times corresponding to the selected IF bandwidth. The CPU receives the real and imaginary components of the 20-kHz IF signal.

Unlike other spectrum analyzers, the HP 4396A doesn't have a logarithmic amplifier. As mentioned above, in-phase and quadrature detection or FFT methods are used, which eliminates the need for a costly logarithmic amplifier and various RBW bandpass filters. The logarithm is calculated by the CPU after linear detection by the ADC.

The traditional logarithmic amplifier method has the following disadvantages:

- The logarithmic amplifier's linearity, which is worse than the instrument ADC's, decides the system linearity of a spectrum analyzer.
- Dc offset voltage at the output of the rectifier limits the instrument's dynamic range.
- The logarithmic amplifier circuit is very complicated and consequently expensive.

(continued on page 92)

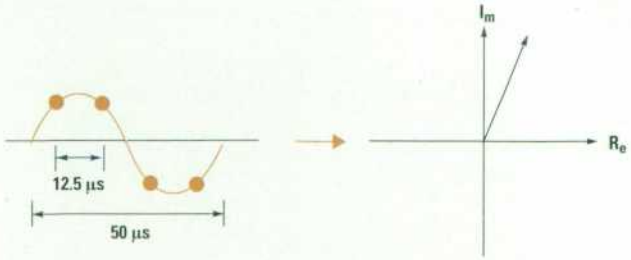


Fig. 10. Four data points are collected in the ac sampling mode to determine the real and imaginary parts of the 20-kHz signal. The sample rate is based on the 80-kHz (12.5-µs) ADC.

DSP Techniques for Digital IF

The Motorola 20-MHz DSP56001 is used for digital signal processing in the HP 4396A vector network and spectrum analyzer. It processes the data coming from the 16-bit, 80-kHz ADC located in the receiver section of the instrument. It performs one-to-three-step (1-Hz, 3-Hz, 10-Hz, ..., 3-kHz) RBW filtering without any specific hardware or large memory banks. For each RBW filter, the 56001 decimates* the input data from the ADC performs windowing, FFTs, square summing (power), video filtering, peak detection, and level calibration. The 56001 is not fast enough to do the required decimation in real time. On the other hand, huge amounts of memory would be required to batch process real-time data. In the HP 4396A, decimation is separated into two stages, with the first decimation stage processed in real time to decrease the memory volume for the data to be stored. To achieve this efficiently, several DSP techniques are used.

In designing a DSP mechanism it is important to consider the number of FFT points (N), the decimation factor (M), the FFT window, and the digital filter used for decimation. In the HP 4396A, the Remez exchange algorithm[1] is used in the design of the window function and the decimation filters.

Spectrum Resolution

Suppose the signal from the ADC has no images (all rejected), no aliases (anti-aliasing effectively applied), and the sampling frequency is f_s. An N-point FFT applied to the incoming signal to get its spectrum gives the following resolution bandwidth (RBW):

$$RBW = k \times f_s/N \quad (1)$$

where k is the 3-dB bandwidth factor of the window function. Table I lists some of these bandwidth factors for different window functions.

Table I
3-dB Bandwidth Factors

Window Functions	Bandwidth Factor
Rectangle	0.89
Hamming	1.30
Blackman	1.68

If f_s is constant and we want to ensure that the required RBW is wide enough, suitable values for k and N must be chosen since the RBW is a function of these two variables. In the HP 4396A, a suitable power of two is selected for N and then a window with a suitable k can be designed.

Necessity of Decimation

If a narrow RBW is required, according to equation 1 a larger N, a lower f_s, and a smaller k would make it possible. However, a larger N needs more memory and the FFT coherent quantization error would be an important factor. We found through simulation that N = 4096 is the maximum limit to avoid this error in the HP 4396A, and k cannot be less than four to meet the specification of passband ripple for the window functions.

Directly reducing f_s in hardware is not feasible for the HP 4396A. Therefore, we digitally convert the sampling rate of a signal from the given f_s to a lower value f_s/M, where M is an integer value called the decimation factor.[2] Thus,

$$RBW = k \times (f_s/M)/N \quad (2)$$

Note that once the signal is M-to-1 decimated, the information bandwidth of the result of one FFT is divided by M. Also, in the decimation process the appropriate anti-alias digital filtering should be performed before the M-to-1 sample rate reduction because the sample rate reduction generates aliasing M times. The execution time of a decimation is based on the filtering process.

* Decimation in digital filtering is the process of digitally converting the sample rate of a signal from a given rate f_s to a lower rate f_s', where $f_s' < f_s$.

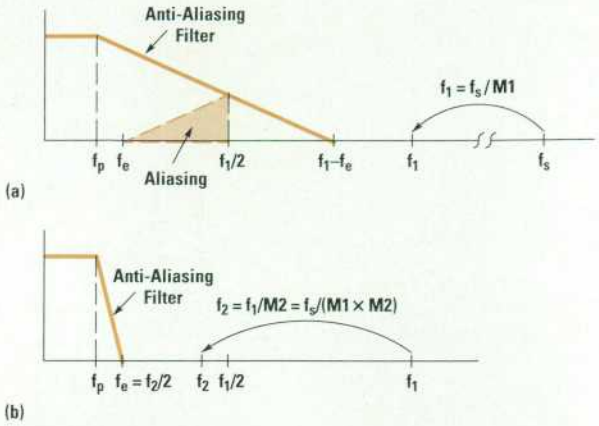


Fig. 1. Signal processing and filter specification for two-stage decimation. (a) First-stage decimation. (b) Second-stage decimation.

Two-Stage Decimation

For the case in which the whole computation of the resolution bandwidth is done in batch mode using equation 2, the amount of data required to be stored is:

$$M \times N + L$$

where L is the number of taps** of the digital anti-aliasing filter before decimation. This number is too large for hardware implementation.

Decimation processing in real time could reduce the amount of memory to the order of N, but the 56001 cannot the finish the whole M-to-1 decimation calculation in a sampling period (12.5 μs = 1/80 kHz). Therefore, we separate the decimation into two stages, that is,

$$M = M1 \times M2. \quad (3)$$

First, M1-to-1 decimation is performed and then M2-to-1 decimation is performed. On the condition that the necessary anti-aliasing is performed in the second decimation stage, some aliasing can be accepted in the first decimation stage. This means that the necessary amount of anti-alias filtering can be much smaller in the first decimation stage. Therefore, the first decimation is performed in real time since the amount of data required to be stored is reduced to:

$$M2 \times N + L \quad (4)$$

where $M2 < M$. Fig.1 illustrates the concept of this two-stage decimation process.

In the first decimation stage, the impulse response of the anti-aliasing filter is convolved with the signal, which is sampled with the rate f_s. Only every M1-th sample of the filtered output is saved. Consequently, the sample rate is converted from f_s down to f_1, which is equal to $f_s/M1$. Note that aliasing occurs between f_e and $f_1 - f_e$, but the effective information band (0 to f_p) is not affected.

In the second decimation stage, anti-alias filtering is performed and only every M2-th sample of the filtered output is saved. Consequently, the sample rate is converted from f_1 down to f_2, which is equal to $f_1/M2 = f_s/M1/M2$. In this stage, complete anti-aliasing is required so that the range from f_e to f_1 is completely eliminated by the filter.

** Taps refers to the number of coefficients used in the FIR filter.

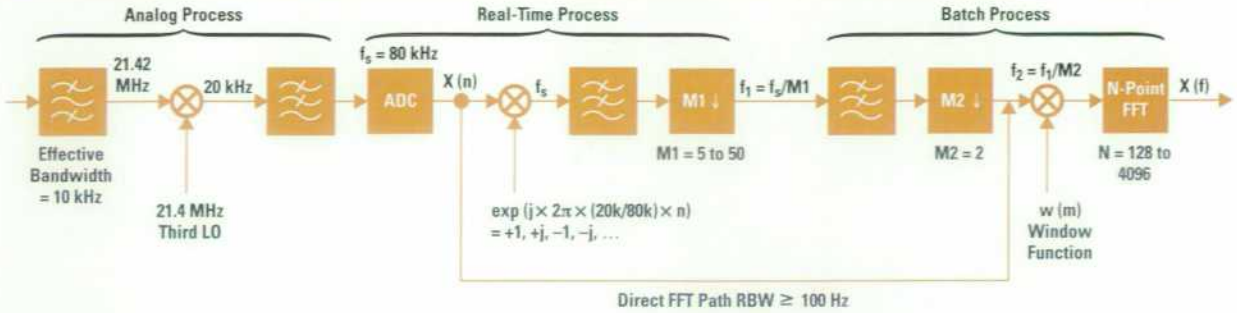


Fig. 2. HP 4396A signal processing block diagram.

The passband width ($2 \times f_p$) of the filter is the effective information bandwidth of the FFT and the step width of the stepped FFT.* The wider this is, the fewer the FFTs for the same span, but the more difficult it is to realize anti-alias filtering. In the HP 4396A, f_p is 75% of $f_2/2$.

According to equations 3 and 4, a larger value of M1 (smaller value for M2) can save more memory. In the HP 4396A, M2 is always two. Table II shows the parameters used in computing the RBW.

Table II
RBW Parameters

RBW (Hz)	1	3	10	30	100	300	1 k	3 k
N	4096	4096	4096	1024	4096	1024	256	128
M1	50	16	5	5				
M2	2	2	2	2				
f_s (kHz)	80	80	80	80	80	80	80	80
f_2 (kHz)	0.80	2.5	8	8				
EBW (kHz)	0.60	1.875	6	6	10	10	10	10

N = Number of FFT points

M1, M2 = First and second decimation factors

f_s = Sampling frequency

f_2 = Second decimated sampling frequency ($f_s/M1/M2$)

EBW = FFT effective information bandwidth
($75\% \times f_s/M1/M2$ for decimation, 10 kHz for direct FFT)

This two-stage decimation is performed in the signal processing section section of the HP 4396A, which is described below.

* See "High Sweep Speed at Narrow RBWs" on page 79 for a brief discussion about stepped FFT.

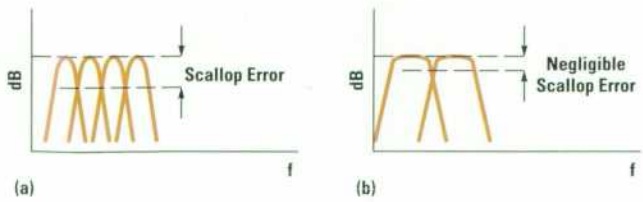


Fig. 3. A subtle side effect of using the Hamming window function is that if a component of the input signal is not centered in the filter's shape, a portion of the signal's amplitude will be attenuated. (a) Hamming passband filter shapes and the scallop error. (b) Using a window function with a flatter (flattop) passband reduces the scallop error.

Signal Processing Block Diagram

The signal processing block diagram in Fig. 2 shows the three blocks involved in signal processing in the HP 4396A. In the first block (analog process) the incoming signal is filtered and converted to digital format. In the second stage (real-time process) digital mixing and the first decimation are performed in real time. Finally, in the third block (batch process), second decimation, windowing, and FFT are performed. For cases in which the RBW is equal to 100 Hz, 300 Hz, 1 kHz, or 3 kHz no decimation is required, so only windowing and FFT are performed. This direct FFT process is also indicated in Fig. 2.

Crystal Bandpass Filter. This filter rejects the image responses and the aliases resulting from the frequency conversion that takes place between the second IF (21.42 MHz) and the third IF (20 kHz). The passband width of this filter is 10 kHz, which determines the step width of the stepped FFT when the RBW is 100 Hz, 300 Hz, 1 kHz, and 3 kHz.

ADC. The ADC used in the digital processing loop is an Asahi-Kasei Co. AK9202-VP (f_s = 80 kHz). The output of the ADC is transferred into the 56001 directly as 16-bit serial data. The 56001 has an on-chip serial interface and when data arrives from the ADC an interrupt occurs in the 56001. The data is read and the necessary processing is performed by the DSP.

Digital Mixer. Digital mixing is used only when decimation is required. Because $f_s = 4 \times$ (third IF), actual digital mixing is a simple operation. This mixing operation is complex mixing, so no unnecessary frequency shift is generated. Note that the mixed output signal is a complex signal.

Windowing. Because the passband width of a window function is designed to be the same width as the FFT resolution (f_s/N or $f_s/M/N$), the scallop error of the window does not affect the level flatness of the HP 4396A (see Fig. 3).

In the design of a window function, not only the flatness (passband ripple mentioned above), but also the elimination band ripple and the 3-dB bandwidth factor (k) are important.

Data simulation was repeated many times to satisfy the requirements for window functions.

FFT. In the HP 4396A, 128-, 256-, 1024-, and 4096-point FFTs are used.

Gain Adjustment. Because the 56001 is a fixed-point DSP chip, the coefficients of the window functions and the filters described have to be normalized to avoid a reduction of dynamic range. The gain of this normalization is adjusted after the FFT process.

References

1. L. R.Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice Hall, 1975.

2. R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice Hall, 1983.

Akira Nukiyama
R&D Development Engineer
Kobe Instrument Division

The in-phase and quadrature method has several advantages. A 16-bit ADC provides very good linearity and wider dynamic range at a lower cost. The drawback is that the detection speed is slow. This is because more ADC bits are necessary to get the same wide dynamic range provided in logarithm-based detection. This slow speed especially affects the zero-span spectrum measurement for a rapidly varying signal. The HP 4396A uses repetitive sampling for faster zero-span sweep to make up for this weakness.

System Performance

Fig. 11 shows that the system noise floor for the HP 4396A at 100 MHz is –155 dBm/Hz. The stepped FFT method for narrower RBWs produces 20 to 100 times faster sweep than that of conventional spectrum analyzers.

The scale fidelity (system linearity) of the HP 4396A spectrum analyzer is much better than logarithm-based spectrum analyzers. This is one of the great advantages of using a linear detection method. The most obvious difference is easily seen during a carrier-to-noise measurement. The HP 4396A shows more than 100-dB dynamic range while a logarithm-based spectrum analyzer has an obvious "bottom" because of the dc offset of the rectifier. Fig. 12 shows the difference for a spectrum measurement between the HP 4396A and a conventional spectrum analyzer.

The frequency response of the third-order intermodulation distortion of the HP 4396A is shown in Fig. 13. The low noise floor allows a measurement to be done easily and quickly. As shown in Fig. 14, intermodulation-distortion-free dynamic range is more than 100 dB.

It should be pointed out that the time to display intermodulation distortion at the same level is much shorter for the HP 4396A than other spectrum analyzers. This is because of the low noise floor and the FFT method with a dedicated DSP chip. Fig. 15 shows three different displays of –80-dB IMD taken from three different spectrum analyzers. The sweep

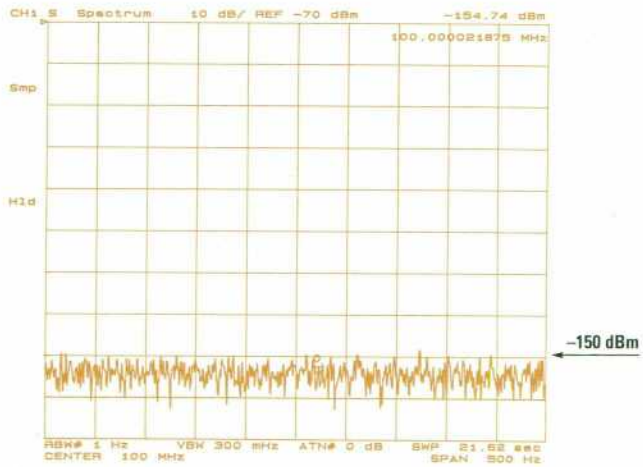
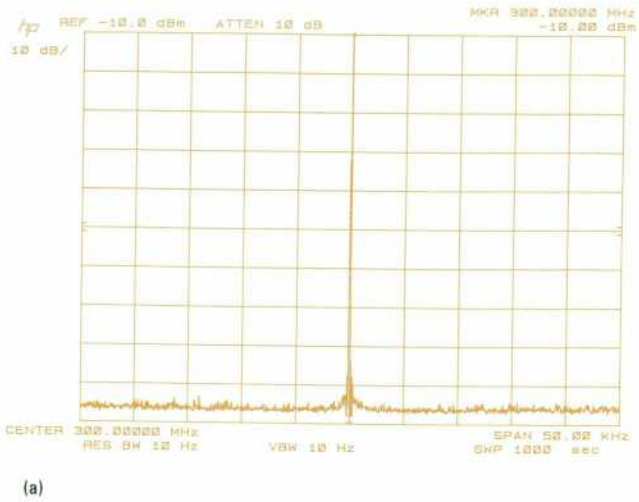
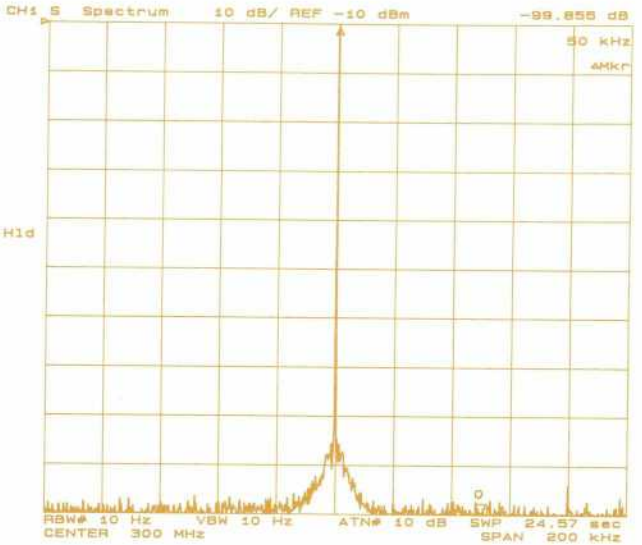


Fig. 11. The system noise floor for the HP 4396A is about –155 dBm/Hz at 100 MHz.



(a)



(b)

Fig. 12. The difference in scale fidelity between the HP 4396A and a conventional spectrum analyzer. (a) A dc offset of the rectifier is seen at about –95 dB in a conventional logarithmic-based spectrum analyzer's display. (b) No dc floor is seen in the display of the HP 4396A.

time of the HP 4396A is only 6.24 s (Fig. 15a) while a conventional spectrum analyzer like the HP 8568B takes 500 s (Fig. 15b), and even an FFT-based spectrum analyzer like the HP 8561E takes 53.3 s (Fig. 15c).

Time Gated Measurement

The HP 4396A has two advantages associated with time gated (pulsed RF) measurement. First, the synchronous sample-and-hold detection method can make a time gated measurement with 2-μs resolution. Second, the digitally implemented video filters can filter the displayed signal even though the input pulse width is shorter than the response time of the video filter.

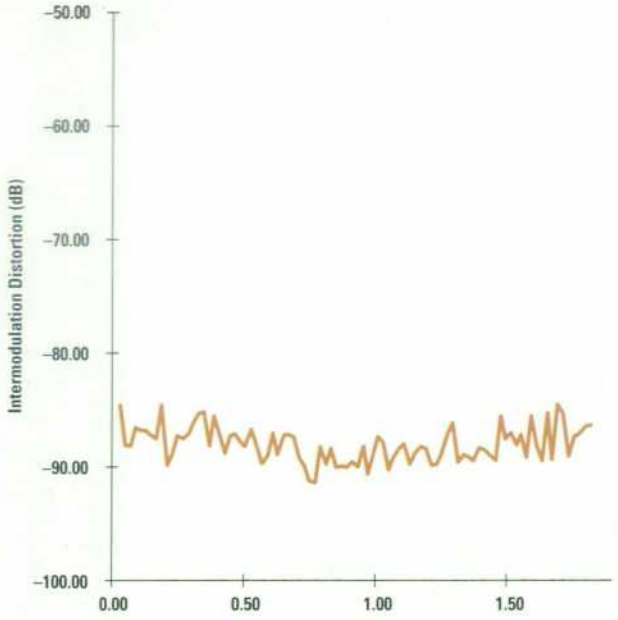


Fig. 13. Frequency response of the HP4396A's third-order IMD.

The 80-kHz ADC in the HP 4396A might seem to be unable to catch a 2-μs pulse width. However, synchronization of the sample-and-hold circuit and the ADC to an external trigger can make it possible. Some logarithm-based spectrum analyzers can catch a fairly narrow pulsed signal because they have an ADC with less resolution (fewer bits), which runs faster. However, they cannot provide level linearity because of the logarithmic amplifier.

The synchronization technique mentioned above provides narrow-pulse RF measurement with good linearity. Many applications of pulsed-RF measurements are carrier-to-noise measurements. If the signal-to-noise ratio is 60 dB, then the reliability of the value 60 dB depends on the linearity of the detector. The HP 4396A typically has only a 0.3-dB linearity error in this case.

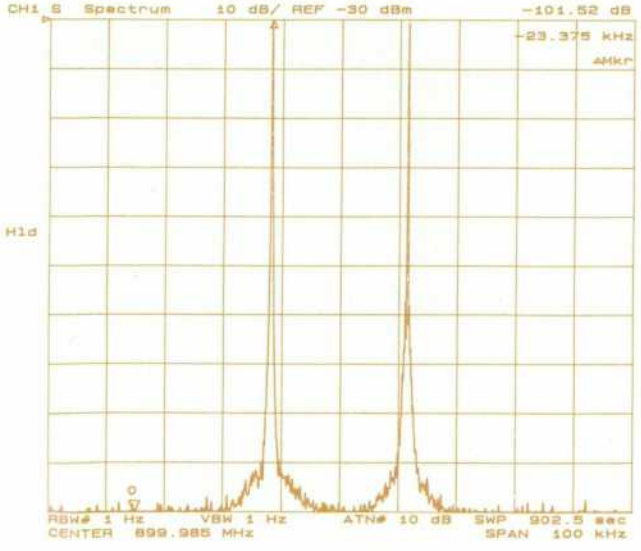


Fig. 14. Distortion free dynamic range is more than 100 dB for the HP 4396A.

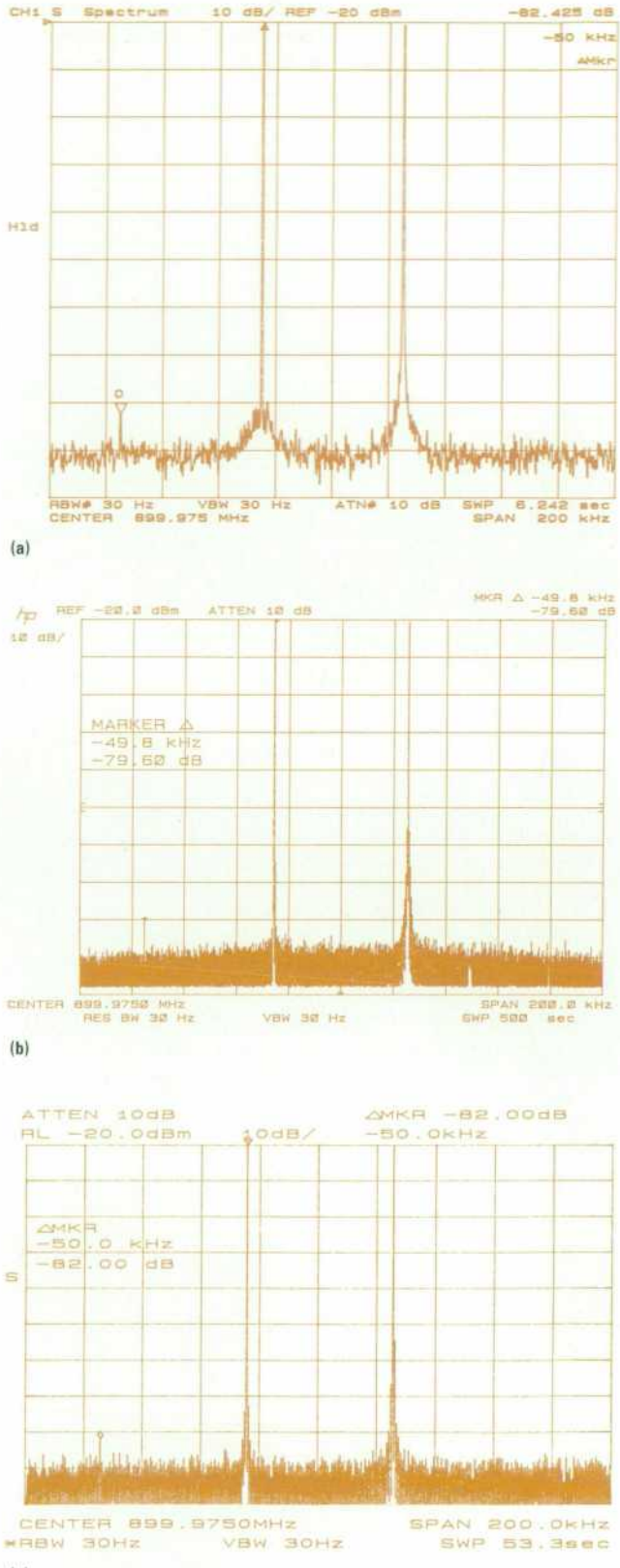
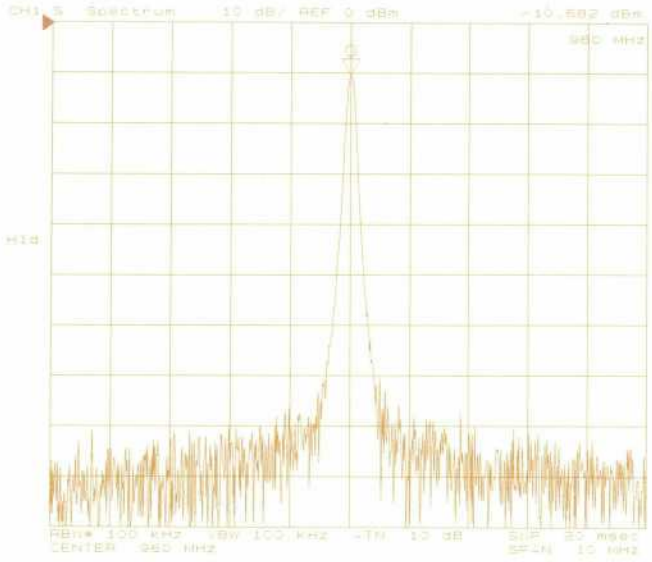
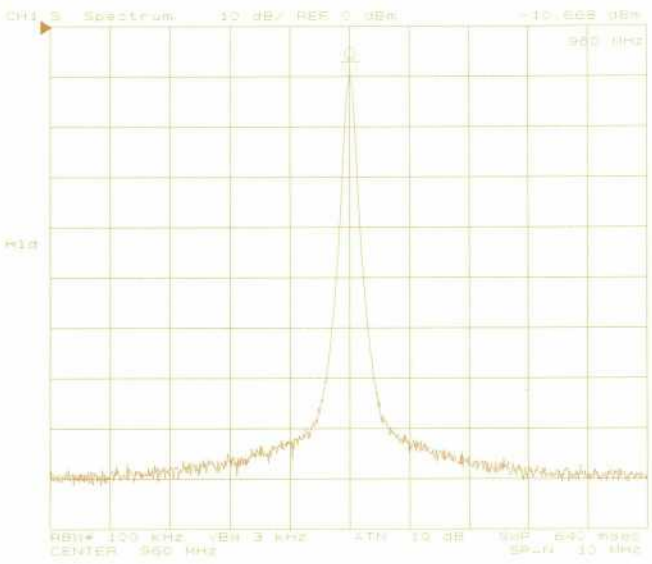


Fig. 15. Intermodulation distortion display times. (a) The HP 4396A shows –80-dB IMD in only 6.24 s thanks to the low-noise-floor design of the receiver and the dedicated DSP chip. (b) A conventional spectrum analyzer (HP 8568B) takes 500 s to show –80-dB IMD. (c) An FFT-based spectrum analyzer (HP 8561E) takes 53.3 s to show –80-dB IMD.



(a)



(b)

Fig. 16. An example of the improvement in measurement results provided by digital video filters. (a) A time-gated (pulsed-RF) measurement of a pulsed input signal with 80-μs pulse width and a 100-μs repetition interval without video filtering. (b) The 3-kHz video filter smooths the displayed signal even though the input pulse width is 80 μs.

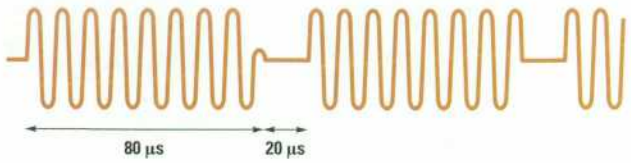


Fig. 17. The input signal for the results shown in Fig. 16.

The video filters implemented in the DSP chip created a new opportunity for time gated measurement. Conventional spectrum analyzers have analog video filters. An analog filter has a settling time. For a conventional spectrum analyzer to take the data on pulsed-RF signals with a video filter, the pulse width of the input signal should be longer than the settling time of the video filter.

The HP 4396A has digitally implemented video filters that don't require any settling time. Only the RBW filter needs some delay. Fig. 16 shows the measurement results for a pulsed-RF signal with a pulse width of about 80 μs with a repetition interval of 100 μs (Fig. 17). Fig. 16a shows a normal pulsed-RF measurement with the RBW filter set at 100 kHz without video filtering. Fig. 16b shows the situation in the HP 4396A in which a 3-kHz video filter is used with the same 100-kHz RBW filter.

A Fast-Switching, High-Isolation Multiplexer

A three-channel multiplexer with 140-dB isolation between channels, fast switching transient settling time, steady low return loss, and low noise and distortion provide the front end to the single receiver of the HP 4396A network and spectrum analyzer.

by Yoshiyuki Yanagimoto

To provide a three-input-port measurement capability to one receiver, a high-isolation, high-speed multiplexer is implemented at the front end of the HP 4396A network and spectrum analyzer. This solution gives the best cost/performance trade-off.

A network analyzer measures the transfer function or reflections from a DUT (device under test) as a function of frequency. Another way to define what a network analyzer does is to say that it measures the vector ratio of the transmitted or reflected signal relative to the source (incident) signal. In the setup shown in Fig. 1, since the internal source is used to stimulate the DUT, the receiver is easily tuned to the same source frequency. In this case, the first IF can be lower than the input frequency because there's no need to reject an image response. For example, the HP 8751A, which is a 5-Hz to 500-MHz network analyzer, uses about 1.5 MHz as its first IF.

Three, or at least two, measurement channels are necessary to get an accurate vector ratio (e.g., A/R or B/R, where A and B represent the measurement signal and R represents the reference signal). Fig. 2 shows an input configuration with three independent ports. Most existing network analyzers have this input configuration.

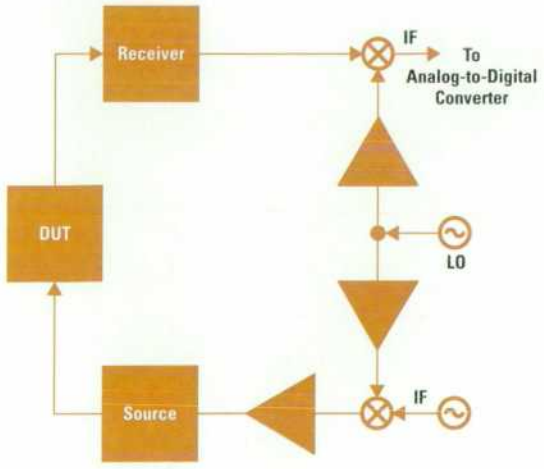


Fig. 1. A measurement setup in which the source to the DUT and the receiver are in the same network analyzer. The receiver of the network analyzer is easily tuned to the same frequency as the source.

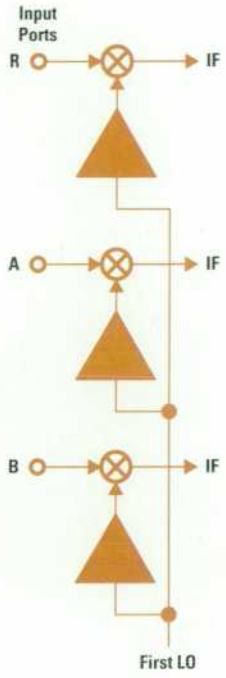


Fig. 2. The input port configuration for a typical network analyzer.

A spectrum analyzer has to be able to receive an unknown signal and identify the frequency of the signal. Most RF spectrum analyzers use an IF that is higher than the input frequency range to separate the input frequency from the image frequency. In general, a higher frequency requires more expensive parts and a better RF shield, and consequently, higher production costs.

To design a network and spectrum analyzer having three independent receivers with a spectrum analyzer configuration would not be the best solution because such an instrument would cost three times more.

Network Analyzer

Because the HP 4396A has only one receiver, a three-channel multiplexer at the front end of the instrument is necessary to make three-channel network analysis possible (Fig. 3). Time division multiplexing is used to measure signals from these three channels with a single receiver.

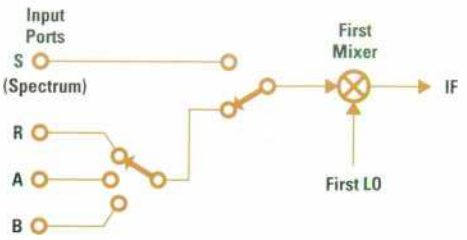


Fig. 3. Input multiplexer design to provide network and spectrum analysis in the HP 4396A.

Fig. 4 and the list below show an HP 4396A measurement sequence for the fastest vector ratio (A/R) measurement with an IF bandwidth of 40 kHz. For a narrower IF bandwidth, more time would be spent at each step in the sequence.

A time division network analysis measurement sequence includes the following internal steps:

1. Set the first local oscillator synthesizer to the measurement frequency.
2. Wait 100 μs for the synthesizer to settle.
3. Determine the IF gain for the R channel (this takes 50 μs).
4. Measure the R channel (50 μs).
5. Switch the input of the multiplexer to the A channel.
6. Wait 50 μs for the transient to settle.
7. Determine the IF gain for the A channel (50 μs).
8. Measure the A channel (50 μs).
9. Compute the A/R vector ratio in the DSP chip.
10. Repeat steps 1 through 9 for the next frequency.

The fast settling time of the synthesizer is essential for this type of measurement because any residual frequency error will become a measurement error. The multiplexer switching time is also very important. Any switching transient would cause a measurement error. Any leakage between channels is also transferred to a measurement error and limits the dynamic range.

From the sequence above it is clear that a fast-settling local oscillator and a high-isolation multiplexer are the keys to time division network analysis.

A frequency error is not a problem for a simultaneous network analyzer because all the input channels have the same frequency error, which is very easily canceled.

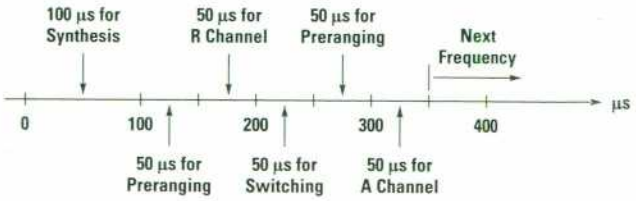


Fig. 4. The fastest A/R measurement sequence of the HP 4396A. The fastest measurement time is 350 μs per point.

Input Multiplexer Design

The important specifications of the input multiplexer for the HP 4396A include:

- 140-dB isolation between channels. This is directly reflected in the dynamic range of the overall instrument and is considered the most important specification.
- Less than 50-μs switching transient settling time for a –100-dB input. The dynamic range of the fastest measurement (IF bandwidth = 40 kHz) is about 95 dB and the waiting time is 50 μs. The effect of a transient should be negligible in 50 μs.
- Less than –40 dB for the return loss variation when the switch position is changed. The input return loss varies slightly because the switch is turned on and off during the measurement. This slight change has an effect on the dynamic accuracy and consequently the measurement accuracy after calibration. The effect should be within the specification of the dynamic accuracy (0.05 dB/0.3°).
- Low noise and low distortion. The full-scale signal-to-noise-floor ratio should be wider than that of the rest of the receiver. This was achieved without great difficulty.
- Low cost and high manufacturability. As is always true, the less expensive the better if the performance is kept the same. The multiplexer in the HP 4396A uses a surface mount process. This process kept our production costs low.

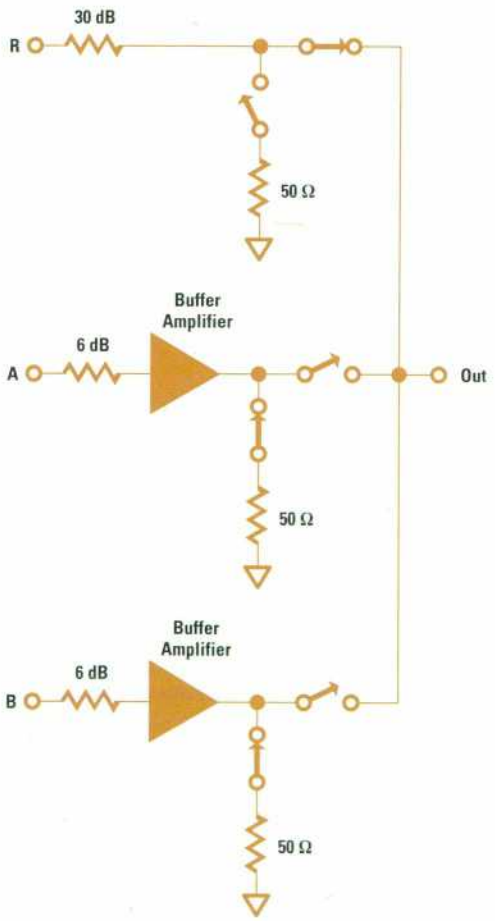


Fig. 5. Simplified diagram of the HP 4396A input multiplexer.

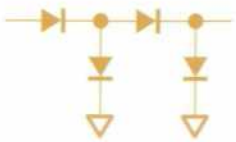


Fig. 6. Series-shunt type diode switch. This circuit provides less isolation efficiency.

Implementation of the Multiplexer. Fig. 5 shows a simplified schematic of the multiplexer. The switching device is made of surface mount p-i-n diodes manufactured by HP. To reduce production cost and increase manufacturability, a surface mount process was desirable. The on resistance of this diode is about 2.5 ohms and the off capacitance is about 0.3 pF. The diode has a bonding inductance of about 1 nH. A series-shunt type of switching was considered first (Fig. 6), but it turned out that the off efficiency (how many dB of isolation can be obtained per diode) was poorer than the finally chosen capacitive divider type (Fig. 7). In the HP 4396A there are six diodes per channel.

When a channel is on, all six diodes are turned on and the path forms a low-pass filter whose cutoff frequency is about 2.2 GHz. When a channel is off, the diodes are turned off and can be considered capacitors. Since the off capacitance of the diode is about 0.3 pF and the capacitance of the pattern is about 2.5 pF, an isolation of about 19.4 dB is obtained at each stage of the capacitive divider. The six stages give a total isolation of about 116 dB.

Although eight stages would have achieved the 140-dB requirement, a 24-dB pad was placed at the R (reference) channel input instead. This decision was based on cost considerations, board size, and customer needs. For most of the possible applications of the instrument, 140-dB isolation is really not needed. Applications in which 140-dB isolation is required are primarily in the area of filter measurements. Fig. 8 shows a case in which a filter with 140-dB isolation is measured. To make this measurement correctly, the leakage through the R channel into the receiver (which is measuring the A channel) should be at least 140 dB lower than the incident signal into the R channel. A 24-dB pad in the R channel added to the 116-dB switch isolation gives the required 140-dB isolation.

This method sacrifices high-gain amplifier measurements. For example, Fig. 9 shows the measurement of a 140-dB amplifier. The leakage through the A channel is 48 dB larger than the desired R channel signal. This means this amplifier cannot be measured correctly. Only up to 92-dB gain can be measured. Since the possibility of having such a high-gain amplifier as a DUT is far less than that of an isolation device such as a filter, the 24-dB pad at the R channel is considered a better solution than any other.

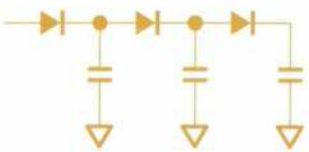


Fig. 7. Capacitive divider type diode switch. This circuit provides better isolation efficiency.

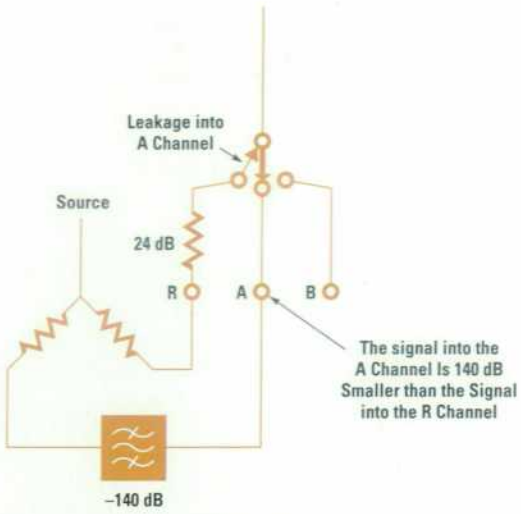


Fig. 8. An extra pad at the R channel makes 140-dB filter measurement possible.

Fig. 10 shows the HP 4396A multiplexer board. N-type connectors are directly connected to the body of the multiplexer. These connectors appear at the front panel of the instrument. This way, the production cost is lower compared to the case in which three cables are needed to connect the multiplexer to the front-panel connectors. This configuration also eliminates the degradation of return loss that would be caused by extra cables.

Low-Transient Design. Specifications require that any transient should be settled in 50 μs. The biggest cause of a transient spike is the switched current source that drives the p-i-n diode switches (Fig. 11). The circuit is carefully designed so that all the current sources have the same value and are switched at the same time. This way, during the switching process, the output port maintains a constant dc voltage (or at least has the smallest possible voltage change).

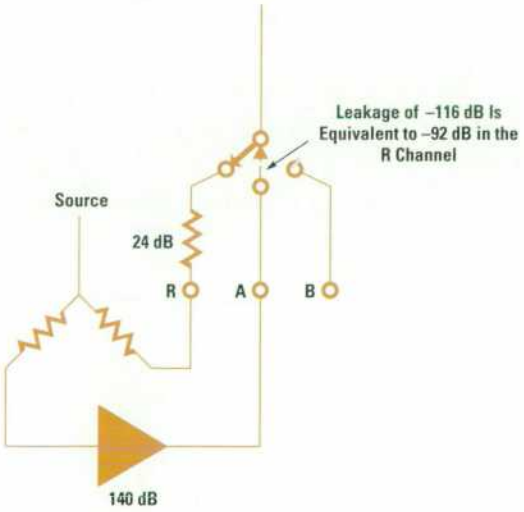


Fig. 9. Measurement of a high-gain amplifier (140 dB). The leakage through the A channel is 48 dB larger than the desired R-channel signal. Thus, the amplifier cannot be measured correctly because of the extra pad at the R channel.

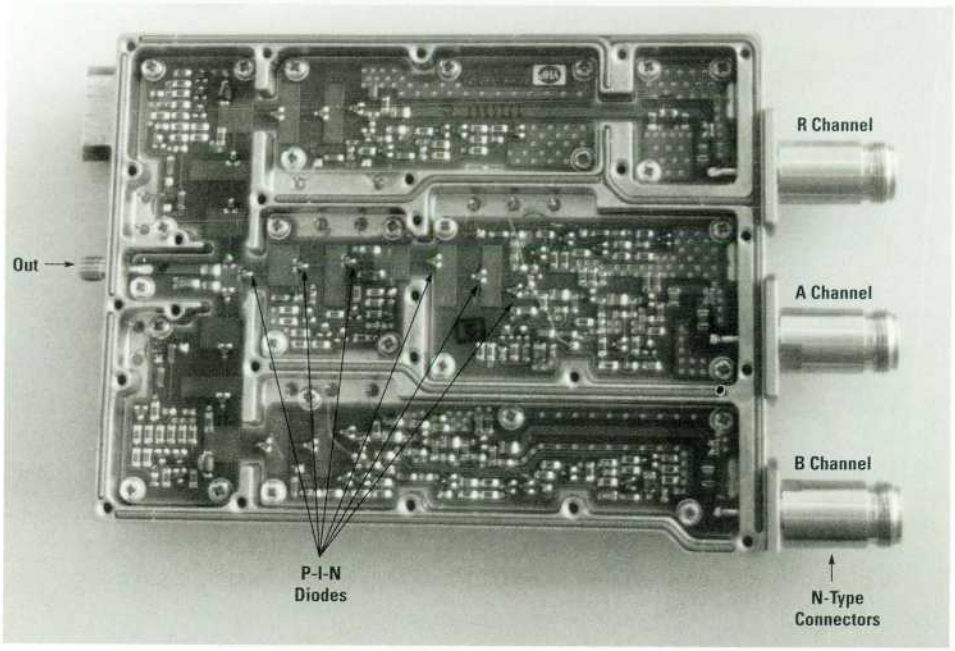


Fig. 10. The HP 4396A multiplexer board, showing the six diodes associated with one channel.

To prevent temperature drift, all the current sources are always active regardless of whether the channel is on or off. The switch only selects which node the current source should be connected to. This way, the temperature drift caused by the multiplexing sequence is kept around 0.005 dB.

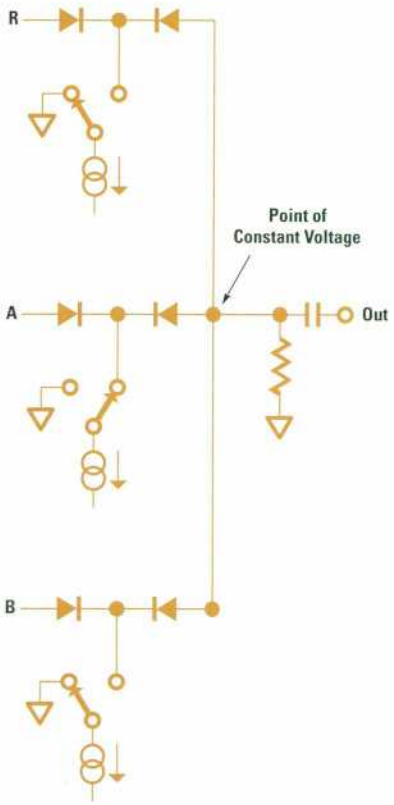


Fig. 11. The p-i-n diode switches of the multiplexer are driven by switched current sources. All the switches have been designed to switch simultaneously to reduce switching transients. Remember that there are six diodes per channel. Therefore, each diode in the schematic represents three diodes.

This drift is mainly caused by the diodes' on resistance when the diodes are heated.

High-Isolation Shielding Method. To get high isolation for low cost, an aluminum casting is used. This casting is solder-plated to prevent corrosion. Au, Sn, and Ni platings were also tested. None showed any degradation after a super-soak test, while bare aluminum showed excessive degradation. Solder was chosen because of its low cost.

All the dc signals, such as those coming from power supplies and control lines, are sent through feedthrough capacitors. All the other RF boards in the HP 4396A use surface mount filter capacitors, but the requirements for the multiplexer were too strict to allow a surface mount capacitor filter to be used.

The bottom of the board and the bottom of the shield casting must be flat to get such high isolation.

Input Return Loss Consistency. Any change of the input return loss will cause an error in the dynamic accuracy, which is specified to be less than 0.05 dB in amplitude and 0.3° in phase. The input return losses of normal network analyzers with three independent input channels are constant during measurement because there is no switch at all in the front end. However, since the HP 4396A has to switch input channels, there is a slight change of input return loss. This degrades the dynamic accuracy slightly, but the degradation is small enough for the dynamic accuracy specification to be the same as other network analyzers.

Fig. 5 shows the circuits that were implemented to reduce the return loss change. A 50-ohm dummy load is located at each input. This was not enough to eliminate the measurement error, so a 6-dB pad and a buffer amplifier are inserted between the input connector and the switch for the A and the B channels.

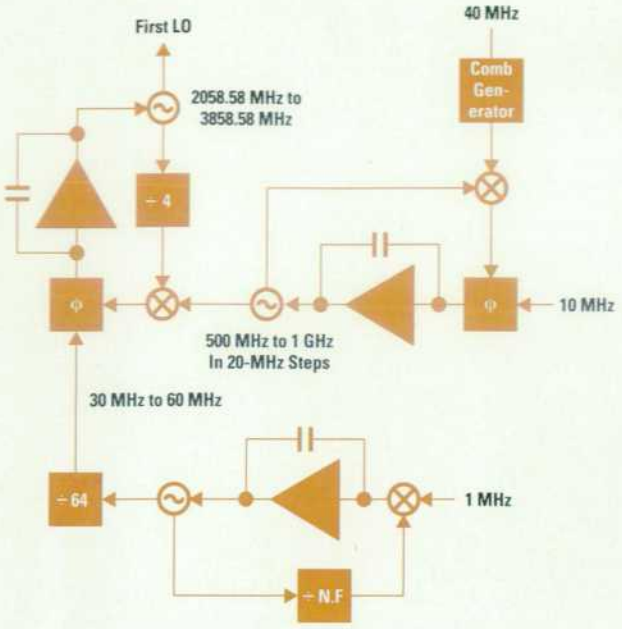


Fig. 12. Three phase-locked loops are used to reduce the phase noise of the VCO.

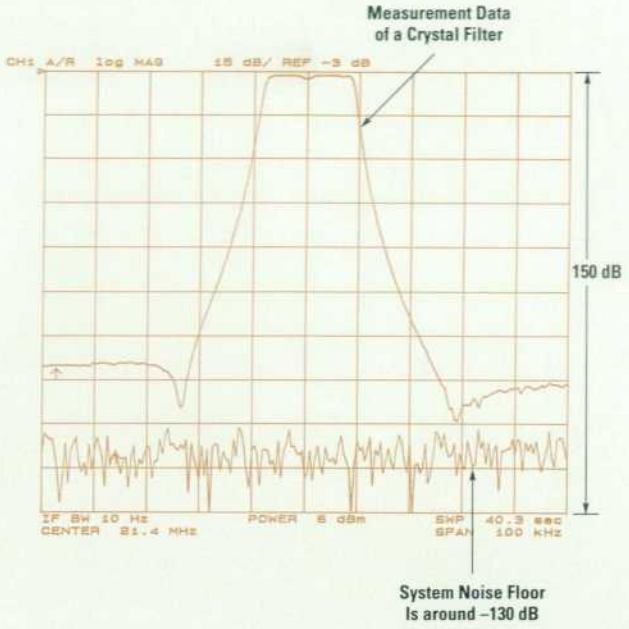


Fig. 14. The widest dynamic range is more than 120 dB.

Fast-Settling Local Oscillator

As was shown in Fig. 4, the synthesized local oscillator should be settled within 100 ms.

Most spectrum analyzers use a YIG oscillator as the first local oscillator. A YIG oscillator is tuned by a magnetic field induced by a large current flowing in a large inductance, on the order of 1 H. Therefore, driving a YIG oscillator and getting it settled within 100 ms is very difficult.

An alternative oscillator, a VCO (voltage-controlled oscillator), was chosen for the HP 4396A. A VCO is tuned by a voltage rather than a current. It can be tuned very quickly so that the main phase-locked loop bandwidth, which includes the VCO, can be set to 1 MHz. The major drawback of a VCO is its phase noise. This level of phase noise is negligible for network measurement but not for spectrum measurement. Most of our effort was put into reducing phase noise. Three phase-locked loops (Fig. 12) reduce the overall phase noise to a level equivalent to a YIG-based oscillator.

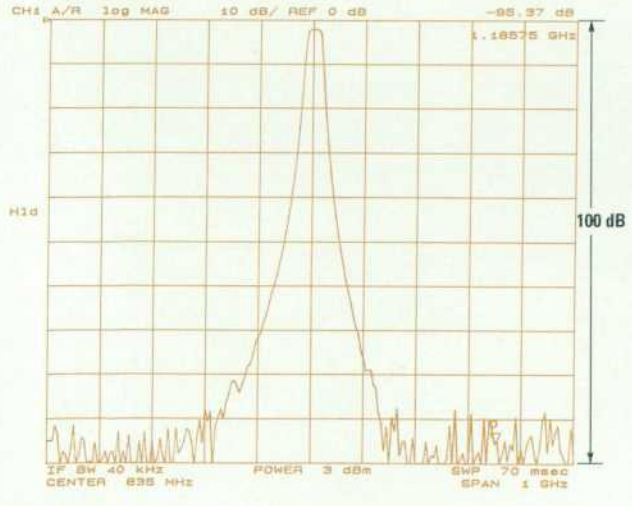


Fig. 13. Even the fastest sweep (350 µs/point) provides a 95-dB dynamic range.

To generate an arbitrary frequency, digitally-corrected fractional-N technology[1] is used in one of the loops. This loop has the narrowest bandwidth (10 kHz) and its time constant is about 15 µs. The settling time of the first local oscillator is determined by this loop (the main loop bandwidth is 1 MHz). We have found through experiments that ten times the time constant, or 150 ms (100 ms before the preranging sequence and 50 ms before the real measurement starts) provides adequate settling. The frequency error and the consequent phase-measurement error caused by the residual frequency error (the frequency not perfectly settled after 150 ms) has been verified to be within the instrument's specification. The performance of the HP 4396A is as good as other simultaneous network analyzers.

System Performance

The measurement of a dielectric filter for a cellular phone is shown in Fig. 13. The instrument was set to the fastest sweep with an IFBW (intermediate frequency bandwidth) of 40 kHz. The A/R measurement was made at 350 µs per point. The wide dynamic range of the receiver provides a 95-dB dynamic range even for such a fast measurement.

Fig. 14 shows the widest dynamic range. A 21.4-MHz crystal filter is the DUT. The displayed dynamic range is well over 120 dB and this measurement was done in only 40 s.

Reference

1. B. L. McJunkin, "Fractional-N Synthesis Module," *Hewlett-Packard Journal*, Vol. 40, no. 5, October 1989, p. 28.

A 10-Megasample-per-Second Analog-to-Digital Converter with Filter and Memory

In addition to analog-to-digital conversion, the HP E1430A addresses the problems of gain ranging, anti-aliasing protection, frequency band selection, triggering, data buffering, and multichannel synchronization.

by Howard E. Hilton

The Hewlett-Packard E1430A is a VXIbus-based analog-to-digital converter (ADC) module containing a high-dynamic-range, 23-bit-resolution, 10-MSa/s (megasample-per-second) ADC, a family of octave-spaced anti-aliasing filters, a complex frequency shifter, and a 4-MSa FIFO buffer memory. It is designed to provide maximum performance and flexibility for capturing a band-limited continuous analog signal in a format compatible with digital computers.

According to Nyquist's sampling theorem, any signal confined to a finite frequency bandwidth can be completely represented by a sequence of discrete samples taken at a rate of at least twice the signal bandwidth. If we are interested only in a finite time segment of the analog signal, all the necessary information is contained in a finite number of these samples taken from the appropriate segment of the sequence. In the absence of additive measurement noise we could theoretically represent the signal with infinite precision, although this would require infinite precision for each discrete sample. However, all analog measurements have some level of additive noise, which limits the amount of signal information available. Therefore, it is only necessary to store each sample with sufficient finite precision to retain the available signal information in the presence of the additive noise.

In other words, it is theoretically possible to completely determine a finite time segment of a bandlimited analog signal, to the extent allowed by additive measurement noise, by collecting a finite number of finite-precision samples of the signal. To maintain complete generality in capturing such a signal, the signal bandwidth, center frequency, start time, and time duration should all be independently specifiable. The HP E1430A offers a wide range of independent choices of all of these parameters while guaranteeing that the sample rate and data precision are sufficient to characterize the signal. The HP E1430A also minimizes the amount of additive measurement noise to preserve as much signal information as possible.

The HP E1430A is much more than an analog-to-digital converter. It also addresses the problems of gain ranging, anti-aliasing protection, frequency band selection, triggering, data buffering, and multichannel synchronization.

Module Description

The HP E1430A is implemented as a single-slot, C-size VXIbus module,[1,2] as shown in Fig. 1. The primary analog connections are the three BNC connectors on the front panel, which are for the analog input signal, an external clock, and an external trigger. The four SMB connectors on the front panel provide the capability of sending synchronizing signals from one VXIbus mainframe to another mainframe containing additional HP E1430A modules.

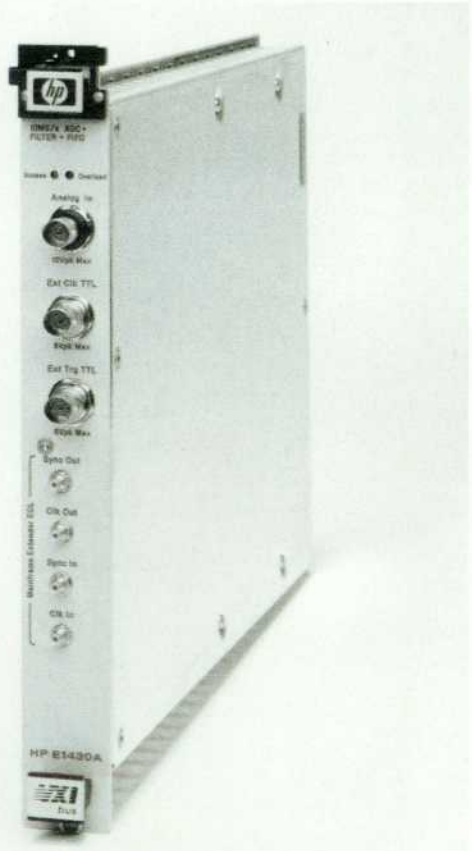


Fig. 1. HP E1430A analog-to-digital converter module.

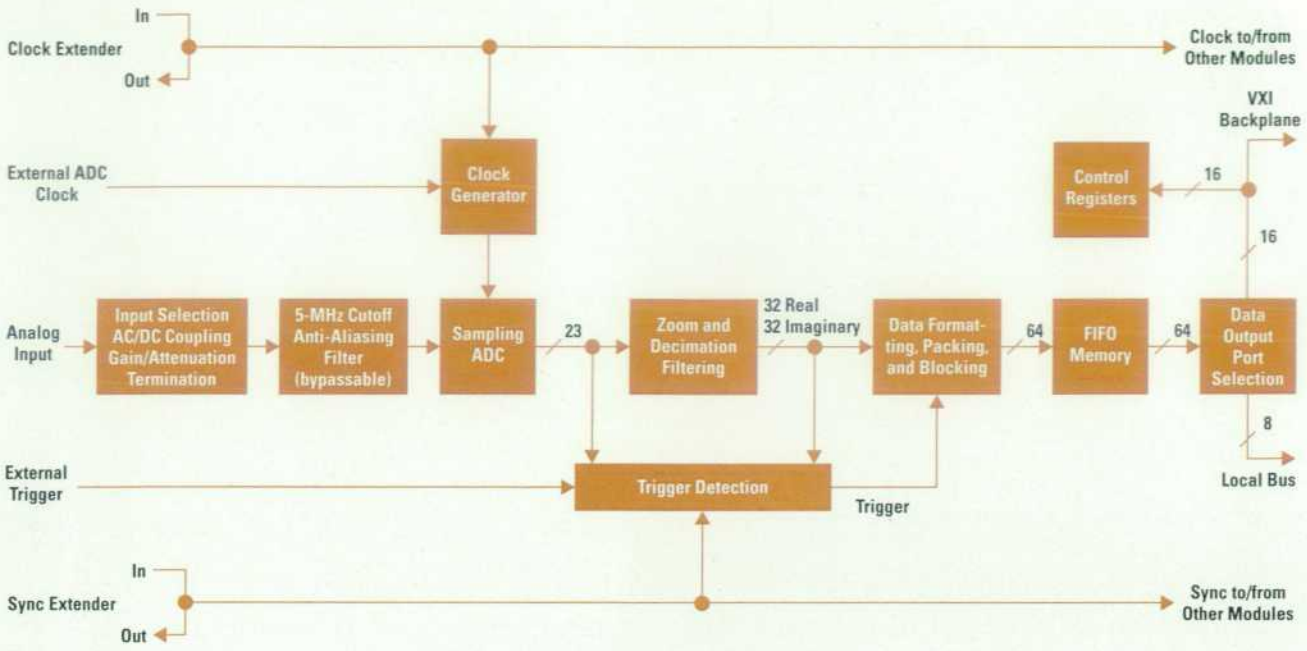


Fig. 2. Block diagram of the HP E1430A ADC module.

The rear panel contains the standard VXIbus connectors, which are used for programming and reading data from the module. The HP E1430A complies with the VXIbus register-based protocol. Status lights are provided to indicate when the module is being accessed via the VXIbus backplane or when the input range is exceeded, producing an overload in the ADC.

Fig. 2 shows a functional block diagram of the HP E1430A module.

Analog Signal Conditioning

It is common practice at audio frequencies to provide high-impedance balanced differential inputs for ADC modules. However, maintaining good frequency response to a bandwidth of 4 MHz requires the use of a terminated transmission line to drive the input. The HP E1430A implements a 50-ohm pseudofloating input as shown in Fig. 3. The cable ground is isolated from chassis ground by 50 ohms in parallel with a 0.04-μF capacitor. This is sufficient impedance to break up low-frequency ground loops, maintaining the key benefit of a differential input. At high frequencies where ground loops are no longer a problem, the 0.04-μF capacitor shorts out the common-mode signal, reducing the impact of common-mode feedthrough at high frequencies. The resistor damps out resonances of the input cable inductance with the cable-to-chassis capacitance.

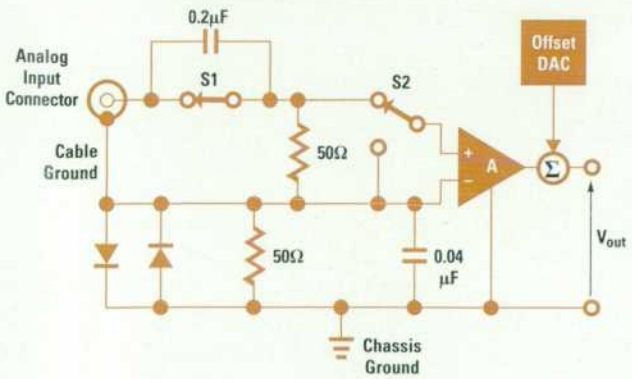


Fig. 3. Analog signal conditioning equivalent circuit.

Diodes are placed between the grounds to protect against damage and to satisfy safety concerns arising from high common-mode voltages. The result is an input termination that maintains good flatness to 4 MHz, suppresses low-frequency ground loop pickup, reduces high-frequency common-mode feedthrough, and eliminates unsafe high common-mode voltages.

Opening S1 under program control causes the input signal to be ac coupled through a 0.2-μF capacitor. This makes possible the measurement of low-level ac signals in the presence of a large dc offset. Programming S2 to the grounded position provides a 0-volt reference so that the offset DAC can be programmed to eliminate any dc offset in the input amplifier.

The gain or attenuation of the input amplifier is programmable in 6-dB steps so that sinusoidal input signals ranging from –32 dBm to +28 dBm can be scaled to produce a full-scale sine wave at the ADC. The noise added to the signal by the HP E1430A is –136 dB/Hz relative to full scale (dBfs/Hz) for the –14-dBm and higher ranges. It is –128 dBfs/Hz for the –20-dBm and lower ranges. This represents a 14-dB noise figure in the –32-dBm range. Most ADC modules have fixed, high-level input ranges requiring the user to provide low-noise external amplification.

Anti-Aliasing Filter

Since the normal ADC sample rate is 10 MHz, a complete representation of the input signal can be achieved only for bandwidths up to 5 MHz. To eliminate the possibility of higher-frequency components causing ambiguous results as a result of aliasing, all signal components above 5 MHz need to be removed before sampling occurs. The analog anti-aliasing filter in the HP E1430A is flat to 4 MHz and rejects signals above 6 MHz by at least 110 dB. Thus the 0-to-4-MHz

frequency range of the sampled signal will be alias-free. The analog filter transition band from 4 MHz to 6 MHz affects the flatness and allows some aliasing in the sampled signal frequency range of 4 MHz to 5 MHz. In some applications a complete, unambiguous representation of a continuous signal may not be necessary, or the user may have additional information about the signal to allow a valid interpretation of the aliased components. In those cases anti-aliasing filtering may not be necessary, and the analog filter may be bypassed. This programmable mode allows the user to take advantage of the full 20-MHz sampler bandwidth. The anti-aliasing filter bypass mode should be used with caution and is not recommended for normal operation.

Sampling ADC

The heart of the HP E1430A is a precision ADC that generates 23-bit outputs at sample rates up to 10.24 MHz. The amplitude resolution is far in excess of the converter's analog noise. Thus, the effects of finite quantization levels can be completely ignored, leaving the main error mechanisms, which are random white noise and linearity errors. For each sample the random error has a Gaussian amplitude distribution with an rms level of –70 dB relative to a full-scale sine wave. The random error for each sample is essentially uncorrelated with previous samples, meaning that the spectral energy of the noise is uniformly distributed across the 5-MHz Nyquist band. Therefore, the noise can be expressed as –137 dBfs/Hz. With the input amplifier noise included, the overall HP E1430A noise level is –136 dBfs/Hz (–128 dBfs/Hz for input ranges <-20 dBm). This low noise density is comparable to the best available ADCs at any sample rate.

In many applications, random errors can be filtered, averaged, or otherwise processed to reduce their impact on the final result. In these applications the deterministic signal-related errors—that is, distortion components—may limit the resulting accuracy unless they are significantly lower than the –70 dB broadband noise level. The HP E1430A achieves distortion errors of –80 dBfs to –110 dBfs depending on the level and dynamics of the applied signal. The graph shown in Fig. 4 shows the worst-case harmonic level for sinusoidal inputs of various levels and frequencies. This distortion performance is considerably better than traditional ADCs in the 10-MSa/s class.

A more complete discussion of ADC errors and how the HP E1430A minimizes them is given in the article on page 105.

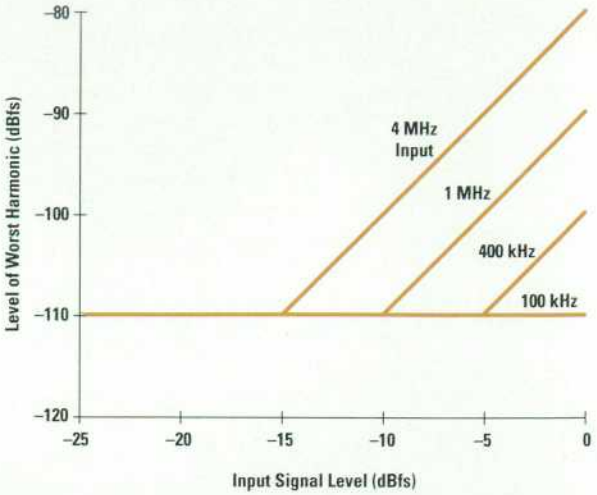


Fig. 4. Harmonic distortion as a function of input level and frequency.

Zoom and Decimation Filtering

For changing the signal bandwidth and center frequency, the HP E1430A provides a complex frequency shifter (zoom) and a complex low-pass filter. Both functions are implemented digitally with proprietary Hewlett-Packard high-speed ICs to achieve real-time operation. A block diagram of the digital signal processing is shown in Fig. 5.

The local oscillator generates cosine and sine waves with spurious components smaller than –110 dBc and frequency resolution better than 10 μHz. These are then multiplied by the incoming signal to produce the real and imaginary components of the down-converted complex baseband signal. The complex baseband signal is then filtered to the desired bandwidth by separately filtering the real and imaginary components.

Bandwidth choices are provided with a cascaded chain of digital low-pass filters, each of which reduces the bandwidth by a factor of two. With the ADC sample rate, f_s, set to the standard internal 10-MHz rate, the available bandwidth choices are ±5 MHz, ±2.5 MHz, ..., ±0.149 Hz around the programmed LO frequency. Each of the filters has ±0.35-dB amplitude flatness to 75% of its indicated corner frequency and has > 105-dB rejection for signals above 125% of its indicated corner frequency. Because of the sharp cutoff, the time-domain step response of the filters has approximately

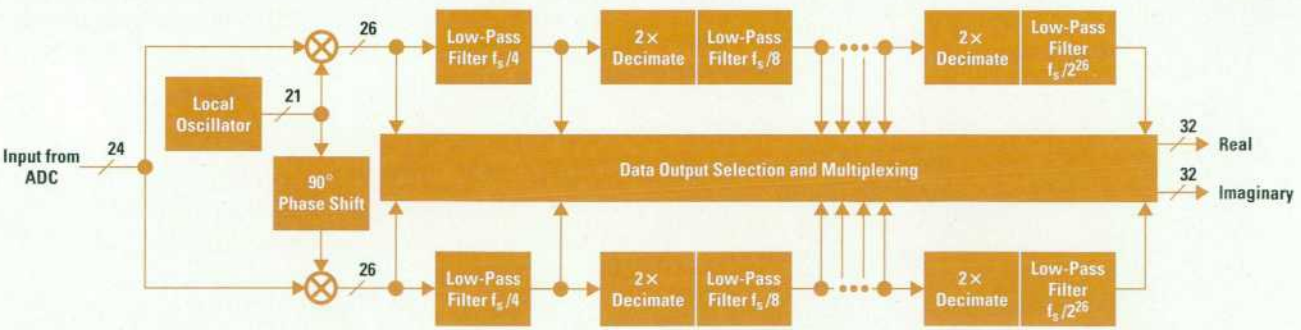


Fig. 5. Zoom and decimation filtering.

20% overshoot. Also, since the filters are not linear-phase, the time-domain impulse response is not symmetric. In time-domain applications where overshoot and/or impulse response symmetry are important the user can apply additional signal processing to achieve the desired filter response. Although the HP E1430A does not include this compensation filtering, all the necessary signal information is preserved to accomplish it within a host computer or signal processing module.

Once the signal bandwidth is reduced below $\pm f_s/4$ the sample rate is also reduced by a factor of two in each filter stage. Thus, each filter output is generated with a sample rate of four times the nominal cutoff frequency. This is sufficient to avoid any aliasing within the filter passband and transition band. The user can program an additional factor-of-two sample rate reduction to get an output sample rate of only two times the nominal filter cutoff. This is still sufficient to avoid aliasing within the passband, but the transition band will not be fully alias-free. This additional decimation is useful in applications such as FFT-based spectrum analysis, where the lower sample rate is beneficial but transition band aliasing is not of concern.

The data multiplexing block can be programmed to output only samples from a particular filter or to multiplex the outputs of all of the filters beyond a selected one. In the multiplexed filter mode each output sample is tagged with a number to indicate from which filter it came. This mode is useful in the implementation of 1/N-octave analysis algorithms.

The real and imaginary components are each computed to 32-bit precision to preserve the processing gain provided in the narrowband filters. Thus, each complex output sample contains 64 bits. Whether or not all these bits are stored in memory can be programmed in the data formatting block.

Data Formatting and FIFO Memory

The HP E1430A can be programmed to save only the real component of the signal or to save the complete complex signal. The data precision can be set to 16 bits or 32 bits. Thus, each sample occupies from two to eight bytes of memory. The data formatting block packs the selected data into 64-bit words, which are stored in the FIFO memory. Since the standard FIFO depth is 1M words (8M bytes), it is possible to hold up to 4M samples in memory at one time.

The memory can be configured either in block mode or in continuous mode. In block mode, data collection initiated by a trigger proceeds until a specified block length is captured. The measurement is then paused so that the data can be read out. Before a new block can be collected, the module must be rearmed and triggered again. This mode is useful in capturing single transient events or whenever the output data rate is too high to be read and processed in real time.

In the continuous mode, data collection is initiated by a trigger and continues as long as the FIFO memory does not overflow. Data can be read out of the memory while the measurement is in progress. If the reading of data is sufficiently fast then the memory will never overflow and the measurement will continue indefinitely. If the memory should ever overflow then the measurement will stop and wait until data is read out, the measurement is rearmed, and a new trigger occurs. This mode of operation is useful for real-time applications that employ a high-speed signal processor to read and operate on each sample of data. The deep FIFO memory allows the consumer to read the data in bursts to accommodate pauses for such things as disk access times or block mode computations.

The effective trigger time can be offset from the actual trigger event by programming a trigger timing offset. The pretrigger offset is limited to the physical depth of the FIFO memory. The post-trigger offset is limited to 2^{26} samples.

Data Output

The output data from the FIFO memory can be directed to a VXIbus register or a high-speed local bus. The VXIbus register can be read by any controller compatible with the VMEbus standard. The memory is unpacked from the 64-bit memory and sent to the 16-bit register as four separate words. Although this mode provides compatibility with a broad range of controllers, it limits the data flow to approximately 4 Mbytes/s. The local bus mode supports data transfers over a high-speed 8-bit ECL bus to an adjacent module (to the right) in the VXIbus mainframe. The HP E1430A can output data over the local bus at rates up to 80 Mbytes/s. This mode requires the use of a consumer module that supports Hewlett-Packard's ECL local bus protocol. The protocol accommodates multiple adjacent HP E1430A modules sending data to a single signal processor module such as the HP E1485A. In addition to the increased data rates, the local bus mode allows output data to flow concurrently with control traffic over the standard VMEbus backplane. This can simplify the design of real-time signal processing systems that require interactive control. In both of the data output modes the samples must be read out sequentially, beginning with the sample following the effective trigger. The HP E1430A does not support random access or memory-mapped access to the data.

Clock and Trigger Generation

Normally the ADC clock is produced by a 10-MHz crystal oscillator inside the clock generation block. However, for applications requiring a customer-supplied sample clock, the HP E1430A can accept an external TTL clock signal at a front-panel connector. The ADCs of multiple HP E1430A modules can be synchronized by programming them to use a common ECL clock line on the backplane. One of the modules can then be programmed as the clock master that drives this line. For systems involving more than one VXIbus mainframe, the backplane clock line can be extended to another mainframe by using the SMB connector on the front panel.

The trigger event used to start a measurement can be generated in four different ways: software trigger, external TTL, ADC threshold, and log magnitude. Any HP E1430A module can synchronously trigger multiple HP E1430A modules via a shared sync line on the VXIbus backplane. This line can be extended between mainframes in the same manner as the ADC clock described above. All modules in a synchronous system are triggered on exactly the same ADC sample. All triggering modes support slope selection. The ADC and log magnitude modes also allow user selection of a trigger threshold, with hysteresis to prevent noise-generated false triggers on the wrong slope. The log magnitude triggering is based on the magnitude of the complex signal after zooming

and filtering. The frequency selectivity of this mode is ideally suited to capturing low-level burst communication signals in the presence of larger interfering signals.

Control

All control of the HP E1430A module is accomplished by means of twenty-four writable and eighteen readable 16-bit registers mapped into the 16-bit VXIbus address space. The operating and service manual documents the function of each of these registers in detail. The module can be programmed from any VXIbus or VMEbus controller. The registers allow direct, high-speed access to all of the functions of the module.

To assist a programmer in using the HP E1430A effectively, the operating and service manual also includes documentation and a distribution disk or tape for a library of functions to facilitate programming the registers. These functions provide a C-language interface for setting up single modules and synchronous groups of modules spanning multiple VXIbus mainframes. Along with the low-level control functions, the library provides setup save and recall, autorange, autozero, and diagnostics. Because source code is included, the functions can be modified or translated to other languages. An executable program that invokes the diagnostic functions is included so that users with a supported controller can test the HP E1430A without writing any code.

For users who are accustomed to a high-level ASCII control interface, the distribution disk or tape includes software that will configure an HP E1405B command module to respond to ASCII commands from a supported external controller. The commands conform to the SCPI (Standard Commands for Programmable Instruments) protocol. The HP E1405B interprets each SCPI command and performs the appropriate register read/write operations on the HP E1430A. A driver is provided to support the HP ITG (Interactive Test Generator) and HP VEE-Test interactive environments. Either of these environments can use this driver and SCPI commands to provide a virtual front panel on the computer screen for control of the HP E1430A. The ITG and VEE software environments are sold separately.

Summary

The primary features that set the HP E1430A module apart from a typical ADC module are its high accuracy, high sample rate, selectable anti-aliasing filters, selectable center frequency, deep FIFO memory, analog signal conditioning, triggering, and fast data transfers. These are important considerations in modern communications receivers, radar and sonar processors, and transient capture equipment. Before digital signal processing algorithms can be applied effectively to signals, those signals must first be captured accurately in digital form. The HP E1430A provides all the necessary capabilities to do perform this function with a high degree of flexibility.

Acknowledgments

The key contributions to the HP E1430A were made by several people at the Lake Stevens Instrument Division. Dean Payne designed the analog portions of the module and managed schematic capture and printed circuit layout for the entire product. The digital design was done by Jerry Ringel with help from Hoang Nu. Software support included a function library provided by Doug Passey and test software written by Chris Sutton. The challenging printed circuit layout was accomplished by Lavonne Fogel and Allyson Riley. Rene Slocumb was the project coordinator. Marketing support and assistance in product definition came from Lee Meyer. As the first application programmer to use the HP E1430A, Mike Gribler provided valuable QA and definition feedback to the project. Ed Guppy generated the product documentation.

References

1. *VMEbus Extensions for Instrumentation System Specification*, Revision 1.3, July 14, 1989.

2. L.A. DesJardin, "VXIbus: A Standard for Test and Measurement System Architecture," *Hewlett-Packard Journal*, Vol. 43, no. 2, April 1992, pp. 6-14.

A 10-MHz Analog-to-Digital Converter with 110-dB Linearity

A classification outline is presented for the errors found in an analog-to-digital converter (ADC). A comparative analysis is done of errors caused by random noise, nonlinearities, and finite amplitude resolution (quantizing errors). An ADC implementation is presented that substantially reduces the nonlinearity errors and virtually eliminates the quantizing errors.

by Howard E. Hilton

An outline of the major error mechanisms in an analog-to-digital converter (ADC) is shown in Fig. 1. The quantizing error results from the use of a finite-resolution numeric representation to approximate the voltage of each input sample. Normally the outputs are expressed in binary form and the quantizing error is inferred by the number of bits used in the binary output. This number of "bits of resolution" is a common basis of comparison between ADCs. Unfortunately, this comparison is often misleading because of the effects of other error mechanisms.

The category labeled additive noise includes both random noise and spurious signals that are independent of the applied input signal. The random noise is a result of stochastic errors resulting from such analog sources as input thermal noise, 1/f or flicker noise, and shot noise. Typically it is dominated by Gaussian white thermal noise, meaning that the error of each ADC sample is independent of all other samples and the error amplitude has a Gaussian probability density. Spurious signals are nonrandom errors that are, in effect, added to the input signal. These typically result from unwanted pickup of such things as power line frequencies, digital clocks, and display monitor scanning circuits. A typical way to quantify the additive noise performance of an ADC is to specify the signal-to-noise ratio (SNR). The SNR is the rms level of a full-scale sine wave divided by the rms level of the total additive noise. It is usually expressed in decibels, 20 times the base-ten logarithm of the ratio.

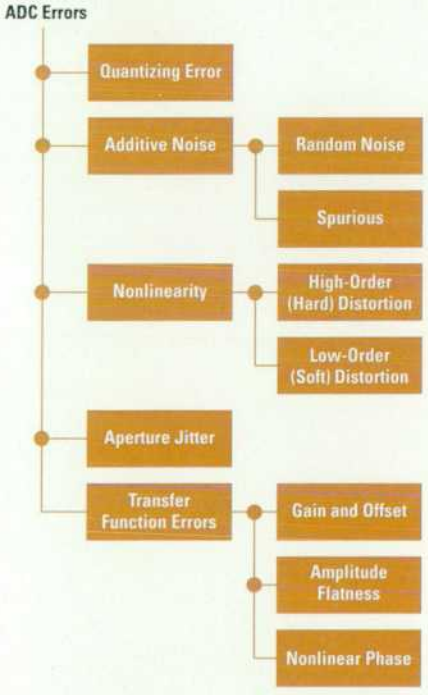


Fig. 1. Outline of analog-to-digital converter (ADC) errors.

The nonlinearities of an ADC result from a transfer function that is not a linear function of the input signal (or its time derivatives). If we consider only the static nonlinearities, those not involving time derivatives of the input signal, we can characterize the nonlinearities of an ADC with a graph of the output error as a function of input signal level. The worst-case deviation from a linear function is often called the integral nonlinearity. Step discontinuities in the error graph are often of more concern in many applications. The worst-case step discontinuity is usually called the differential nonlinearity. An alternative classification of nonlinearities can be defined by representing the error curve as a power series expansion of the input. The low-order terms of this expansion represent the "soft distortion" category while the high-order terms represent the "hard distortion." The importance of this distinction will be discussed in a later section.

Aperture jitter errors arise when the effective sample time of the ADC deviates from the desired sample time. If there is an input signal present that is rapidly changing, this timing error will result in an apparent amplitude error in the sampled ADC output. The amplitude error increases proportionally with input signal level and frequency. One way to specify the effect of aperture jitter is to apply a full-scale sinusoidal input and plot the ADC signal-to-noise ratio as a function of input frequency.

Another way to quantify aperture jitter is to specify the rms time deviation from the ideal sample time. If the aperture jitter is random and independent between samples, the resulting errors will appear as an increased level of white noise. If the aperture jitter is periodic then the resulting error

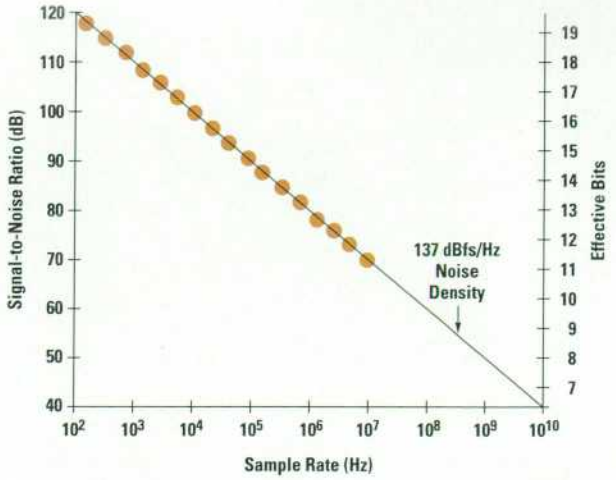


Fig. 2. Signal-to-noise ratio as a function of ADC sample rate.

will appear as sidebands on the applied signal. With < 4-MHz inputs the 10-MHz ADC described later in this paper has sufficiently low aperture jitter that its jitter can be ignored relative to the other ADC error mechanisms. Therefore, aperture jitter will not be discussed further in this paper.

The transfer function errors of an ADC are not as serious as the previously mentioned errors since they can be removed by the application of appropriate compensation to the ADC output sequence. The gain and offset errors can be removed by subtracting the offset and multiplying by the appropriate gain factor. The flatness and phase errors can be corrected with a compensation filter, assuming the use of a uniform ADC clock frequency at least twice the maximum frequency of the input signal. Although for uncompensated operation it is important to specify an ADC's transfer function errors, this paper will not discuss these error mechanisms further.

The 10-MHz calibrated analog-to-digital converter described in this article is the basis of the HP E1430A VXIbus ADC module (see article, page 100). The specifications are 10-MHz sample rate, 70-dB SNR, –110-dBfs hard distortion limit, and –80-dBc distortion at 4 MHz full scale. The product includes digital decimation filters to allow the user to take advantage of the exceptional linearity and better SNR at lower sample rates.

Random Noise

Regardless of the care taken in designing the circuitry in an ADC, the noise performance will at best be limited by the thermal noise of the sampler or track-and-hold amplifier. Since thermal noise is white (equal power per hertz of bandwidth), wideband samplers designed for high-frequency ADCs exhibit a higher level of noise than narrowband track-and-hold circuits used in low frequency ADCs. However, normalized to a 1-Hz bandwidth, the random noise is relatively independent of sampler bandwidth. Since the full-scale level of an ADC is practically limited by power supplies and by distortion considerations, the ratio between the full-scale signal and the noise power per hertz is nearly constant for state-of-the-art ADCs over a wide range of sample rates. Although better performance is theoretically possible, most ADCs are limited to a noise level of approximately –137 dB/Hz relative to a full-scale signal. By integrating this noise over the Nyquist bandwidth (half the sample frequency), it is possible to predict the SNR of high-quality converters optimized for various sample rates. Fig. 2 plots a line of SNR as a function of sample rate for converters with noise performance of –137 dBfs/Hz.

Fig. 2 shows that a 10-MHz ADC with 70-dB SNR has the same noise performance as a 1-kHz ADC with 110-dB SNR. This does not mean that the 10-MHz converter will give 110-dB noise performance merely by running it at the slower sample rate. It will not. The 1-kHz ADC must achieve its lower noise by limiting the effective input bandwidth accordingly. Normally this is accomplished by using a completely different ADC with a narrower-bandwidth sampler, amplifiers, comparators, and so on. However, the same effect can be obtained by using the 10-MHz ADC at its full sample rate, filtering its output with a 500-Hz low-pass digital filter, and resampling (decimating) to a 1-kHz sample rate. The filtering operation eliminates all the high-frequency noise (and signal) that would have been filtered out by the narrower-bandwidth components of the 1-kHz ADC. Thus, with the addition of digital decimation filters, a 10-MHz ADC with 70-dB SNR can be used to achieve state-of-the-art noise performance at any sample rate below 10 MHz. The points in Fig. 2 show the available SNR as a function of the sample rate choices available with the HP 1430A 10-MHz ADC, which includes a set of octave-spaced digital decimation filters.

High-Order (Hard) Distortion

The static nonlinearities of an ADC can be represented by an error graph as a function of input voltage. An example of such a graph is shown in Fig. 3. Although this curve looks noiselike, it represents a deterministic repeatable error in the expected value of the output for each input voltage. Typically such a curve has to be experimentally generated by stepping the input voltage and plotting the deviation of the averaged output from a straight line. Sufficient averaging must be used to eliminate the random noise discussed above.

Sharp discontinuities in the error graph are particularly undesirable since even a small input signal can generate the full range of output errors. This type of nonlinearity is called high-order distortion because a power series expansion of the error curve requires high-order terms to generate the step discontinuity. It is also referred to as hard distortion since the rms error level is relatively independent of input signal level. The errors form an unyielding, hard measurement limit below which amplitude information is questionable. This type of distortion characteristic is typical of most ADCs because of the differential nonlinearities between adjacent output codes.

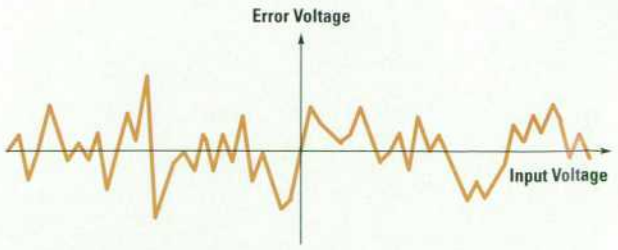


Fig. 3. Static nonlinearity (high-order).

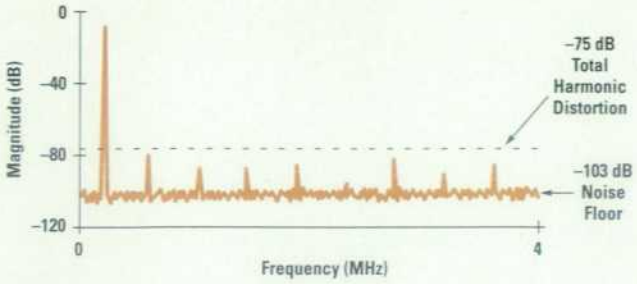


Fig. 4. Spectrum of the high-order nonlinearity.

ADC manufacturers are normally content when the high-order nonlinearity errors are comparable with, or slightly smaller than, the broadband noise level of the ADC, since the distortion then adds little to the overall rms error. Unfortunately, this eliminates the possibility of achieving a higher-precision, lower-sample-rate ADC through the use of digital filtering. The ADC merely becomes distortion-limited rather than noise-limited. Another difficulty with distortion comparable with the noise arises when FFT analysis is used on the data. The noise floor of the FFT result is determined by the equivalent noise bandwidth of each FFT bin, which is considerably narrower than the full Nyquist bandwidth of the ADC. Fig. 4 shows a typical FFT result from a 10-MHz ADC with 70-dB SNR and a hard distortion limit of –75 dBfs. The comb of harmonics becomes clearly visible over the noise because of the processing gain of the FFT. As the level of the input sine wave is decreased the harmonic level will remain near the hard distortion limit of –75 dBfs. The levels of individual harmonics will increase and decrease in an unpredictable manner. No signal information below the hard distortion limit can be trusted, even with low input signal levels.

Low-Order (Soft) Distortion

In addition to the high-order distortion caused by differential nonlinearity of the quantizer, most ADCs exhibit amplifier and sampler distortion, which is low-order. A typical static error curve for an amplifier is shown in Fig. 5. The resulting distortion is called low-order because the power series expansion for the error curve can be expressed with only low-order terms. This type of distortion is often much more acceptable than high-order distortion because of one key difference: the distortion level decreases with decreasing input levels. In fact, the second-order component drops twice as fast as the signal level, the third-order component drops three times as fast, and so on.

Since the low-order distortion mechanisms are usually found in the front-end amplifier and sampler portions of the ADC, these are also the mechanisms that are more likely to change when the input signal is dynamic rather than static. This is

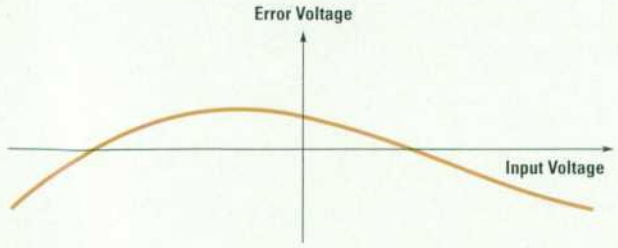


Fig. 5. Low-order nonlinearity error graph.

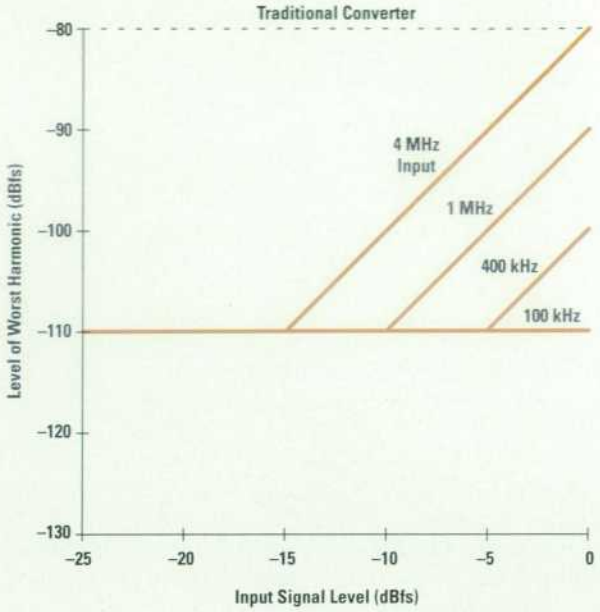


Fig. 6. Distortion performance of the HP 1430A 10-MHz ADC.

especially true in circuits that employ negative feedback to reduce distortion. Typically the decrease in loop gain at higher frequencies causes the feedback to be less effective, and the distortion increases with input signal frequency.

All the high-order, low-order static, and low-order dynamic distortion can be specified graphically by showing the worst-case harmonic level as a function of input signal level for a sine wave. Fig. 6 shows such a plot for the 10-MHz ADC described in this paper. The flat portion of the graph represents the hard distortion limit imposed by high-order differential linearity errors. The sloping lines show the effect of low-order dynamic distortion in the input amplifiers and track-and-hold circuit. The dotted line indicates the typical hard distortion limit of a high-performance, traditionally implemented, 10-MHz ADC with 70-dB SNR and 75-dB total harmonic distortion.

Quantizing Error

A perfect ADC should convert samples of the real-valued analog input into finite-resolution digital form with an ideal quantizer. A graph of the output codes as a function of input voltage from such a quantizer is shown in Fig. 7. The dotted line shows the output of a perfectly linear, infinite-resolution conversion, which is error-free.

The difference between the finite-resolution quantization and the dotted line in Fig. 7 is shown in Fig. 8. This difference is called the quantizing error because it is caused by the finite resolution of the quantizer. The worst-case error of an ideal quantizer is ±1/2 LSB (least-significant bit). In the absence of noise, the quantizing error represents a high-order nonlinearity in the amplitude transfer function. A perfect analog sinusoid will appear to have a comb of harmonics with a total rms error magnitude of approximately $1/\sqrt{12}$ LSB.

In some applications the analog input signal is stochastic and the quantizing error tends to appear as random additive noise with an rms magnitude of $1/\sqrt{12}$ LSB. It is common practice to model the quantizing error as random white noise and call it quantizing noise. In fact, it is also common

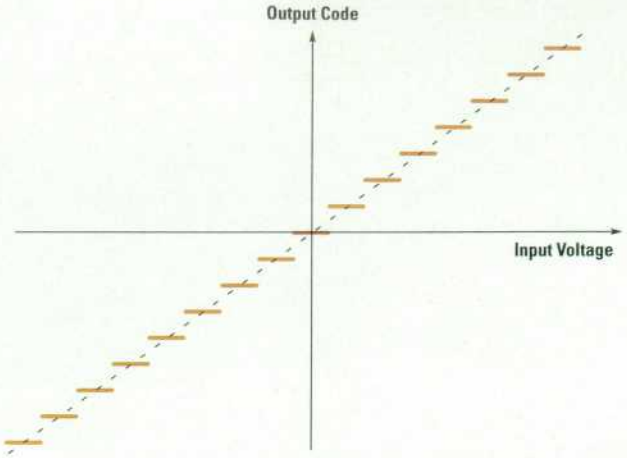


Fig. 7. Ideal quantizer transfer function.

practice to quantify all the noise of an ADC in terms of the number of bits of resolution needed for an ideal converter to produce the observed noise. The relationship between noise effective bits and SNR is given by:

$$\text{Effective Bits} = \frac{\text{SNR}-1.76}{6.02}, \qquad (1)$$

Fig. 2 shows the noise effective bits for a range of converter sample rates and SNR values.

The fallacy of the "quantizing noise" model can become painfully obvious when this model is used in the wrong circumstances. One way to ensure the validity of the quantizing noise model is to add random analog noise to the input signal before quantization as shown in Fig. 9. This ensures that the quantizing errors are in fact random. Because of the added random noise, the output code must be considered a random function of the input voltage. The expected value of the output error for a given input voltage can be computed by integrating the product of the quantizing error and the noise distribution as shown in Fig. 10. In general, the expected output error becomes the convolution of the quantizing error with the noise distribution. If the added noise has a Gaussian amplitude distribution with a standard deviation of σ, the peak expected error, E_p, is given in the formula below. Both σ and the error are expressed relative to one LSB.

$$E_p \approx \frac{e^{-2(\pi\sigma)^2}}{\pi}, \qquad (2)$$

With only 0.5 LSB rms of added analog Gaussian dither, the peak expected error drops to an insignificant level of 0.0023 LSB. This means that if the number of bits of resolution of an ideal converter is more than one plus the equivalent noise bits, the quantization noise model assumption will be valid. If the resolution is more than two plus the equivalent

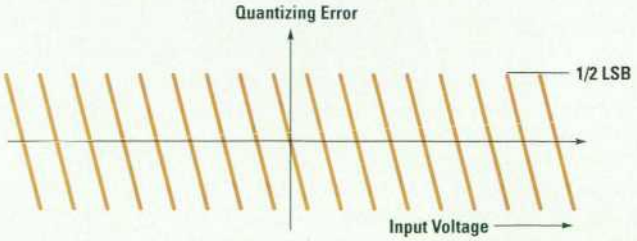


Fig. 8. Quantizing error for an ideal ADC.

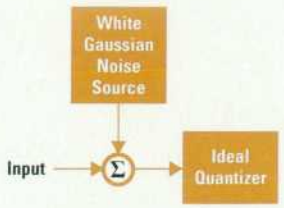


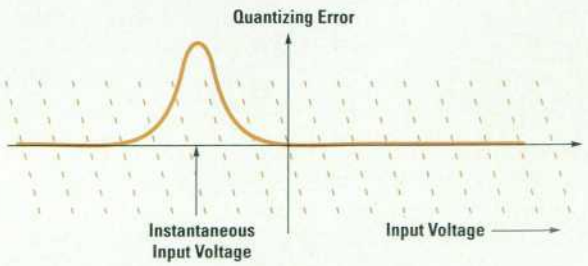
Fig. 9. Adding noise to randomize quantizing error.

noise bits, the quantization noise does not even contribute significantly to the ADC noise and its effects can be ignored. Because the 10-MHz ADC described in this paper has 23 bits of resolution, far in excess of the equivalent noise bits, we can ignore the quantization effects and concentrate on the important parameters: noise and distortion.

Nonlinearities in a Two-Pass ADC

The HP 1430A 10-MHz ADC is based on a variation of the standard two-pass architecture shown in Fig. 11. Each sample is first digitized by ADC1 to coarse resolution and accuracy (8 bits in the HP 1430A). The result of the ADC1 conversion is sent to a DAC, which subtracts that initial estimate from the input signal. The residual analog signal is small enough to be amplified with gain A before being digitized with high resolution by ADC2. The digitized residue is then divided by the gain A and added to the original estimate to produce the final output. The residue gain is normally chosen as a power of two because the divide operation can be accomplished simply by right-shifting the data by the appropriate number of bits. Thus, the digital portion of the circuit reduces to a simple addition. The attractiveness of this architecture is that the errors of ADC1 are not important to the final accuracy of the conversion. Notice that the output of ADC1 is both added to (via the DAC) and subtracted from (via the adder) the final output. Thus, errors in ADC1 are canceled out.

The standard two-pass quantizer is still susceptible to errors in the residue gain, the DAC, and ADC2. Making a quantizer with −110 dBfs hard distortion would require DAC linearity on the order of three parts per million and absolute residue gain accuracy of 0.08%. These are difficult numbers to achieve in analog circuitry at high speed. To achieve this goal,



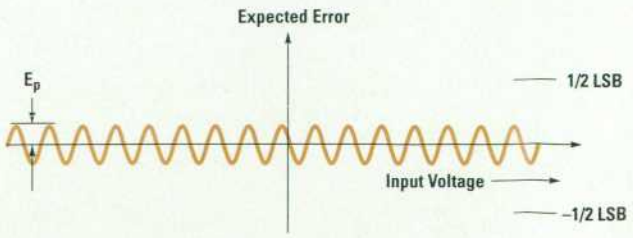


Fig. 10. Adding noise to reduce nonlinearity error.

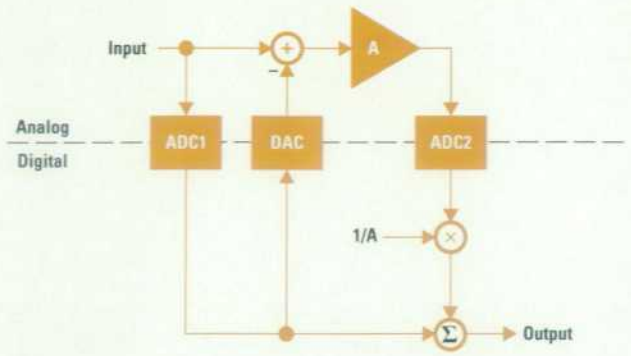


Fig. 11. Standard two-pass ADC.

the HP 1430A converter implements a calibration scheme to remove the residue gain errors and the DAC errors.[1] An abbreviated description of this scheme is is given below.

Residue Gain Calibration

The HP quantizer uses a statistically based continuous calibration scheme to adjust the residue gain. A block diagram of the calibrator is shown in Fig. 12. The added digital circuitry replaces the original multiplication by 1/A with multiplication by an adjustable gain Q.

For this analysis let us assume the DAC and ADC2 have no errors, so the variable gain value can be written as:

$$Q_{n+1} = Q_n + \frac{N_n}{k}(AQ_n(X_n - R_n - N_n) + N_n). \quad (3)$$

Since N is taken to be a random sequence of ±1 values, we can write the expected value of the variable gain as:

$$Q_{n+1} = Q_n + \frac{1}{k}(1 - AQ_n). \quad (4)$$

Note that the expected value of the gain is independent of the input signal since all terms involving X and R are multiplied by the noise sequence N, which has zero mean and is assumed to be uncorrelated with the input signal. The above recursion relation can be solved explicitly in terms of the initial conditions:

$$Q_n = \frac{1}{A} + \left(1 - \frac{A}{k}\right)^n \left(Q_0 - \frac{1}{A}\right). \quad (5)$$

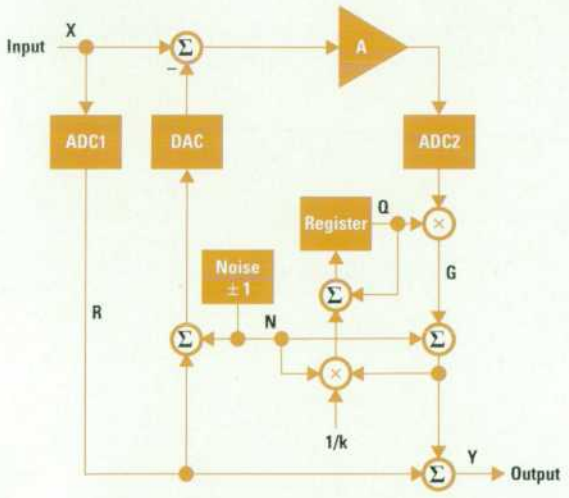


Fig. 12. Residue gain calibrator block diagram.

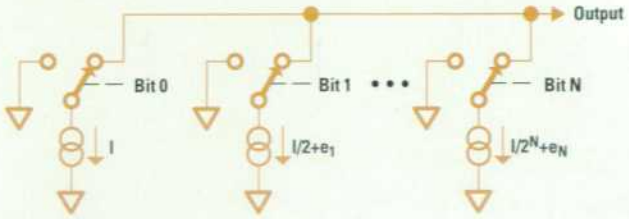


Fig. 13. DAC error model.

If the calibration algorithm runs long enough, the variable gain will approach exactly the desired value:

$$Q_\infty = \frac{1}{A}. \quad (6)$$

It should be kept in mind that although the expected value of the gain converges to the proper value, there will be a random component of the gain because of the discarded terms involving the product of the noise sequence and the residue voltage. This random gain fluctuation manifests itself as signal-related noise. The spectral content of the noise tends to cluster in a narrow frequency band around discrete signals and their harmonics. There is also a noise pedestal at dc that looks very much like 1/f noise. The amplitude of this noise can be made arbitrarily small by choosing a very long gain loop time constant. In the HP implementation, the time constant is 2^{24} samples. This provides sufficient averaging to ignore the random gain fluctuation.

DAC Calibration

To correct for the DAC errors in an efficient manner the following analysis assumes a model for the DAC as shown in Fig. 13, where $e_1, e_2, \ldots, e_N$ represent individual current source errors with respect to their ideal binary weighted values. This model assumes independence between bits in that the error of each is not a function of the states of the other bits. The DAC in the HP implementation was chosen based on conformance to this model.

Fig. 14 shows a block diagram of a modified two-pass quantizer that corrects the bit errors of the DAC. The DAC is replaced by a combination of three DACs with their outputs summed together. For the purposes of this analysis, assume that ADC2 is perfect and that the 1/A multiplication has been accomplished exactly so that the effective residue gain is one. Two independent noise sources have been added to ensure that the DAC bits are randomly exercised. Since the noise sources are both added to and subtracted from the signal path, their contribution to the residue will be zero provided that the DAC current weightings are exactly correct. It is this fact that allows the detection and correction of individual bit weights for all three DACs. The bold lines in Fig. 14 indicate vector quantities while the other lines indicate scalars. **B** is a vector of ±1 values that correspond to the states of all the bit lines driving the DACs. Representing the bit states as {−1,1} instead of the usual {0,1} simplifies the calculations of correlations. This is consistent with the normal coding of DACs and ADCs provided that a proper dc offset is applied. The most-significant bit of DAC1 is excluded from the vector **B** since that DAC bit is used as the reference for all the others and, by definition, does not require calibration. The vector **H** represents the corrections for the DAC bit errors.

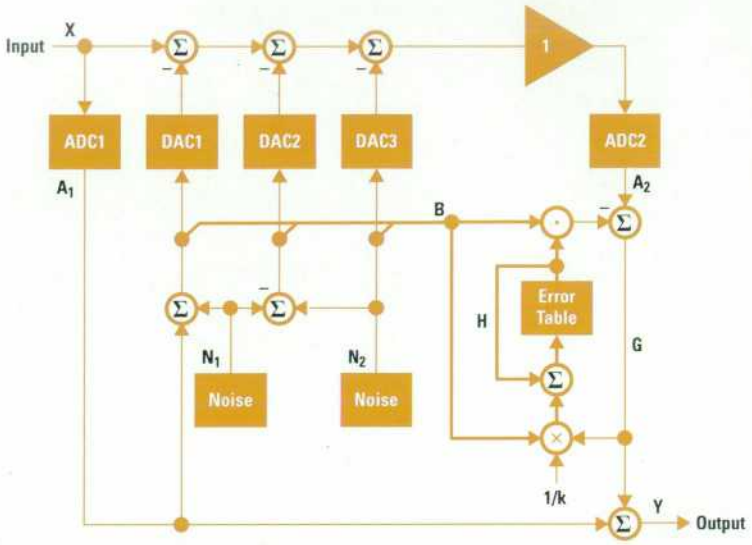


Fig. 14. DAC calibrator block diagram.

From the Fig. 14 we can write an equation for the nth sample of G in Fig. 14 as:

$$G_n = X_n - b_n - \mathbf{B}_n \cdot \mathbf{C} - \mathbf{B}_n \cdot \mathbf{H}_n. \quad (7)$$

In the second term b_n is the ±1 value corresponding to the most-significant bit of DAC1. This equation has been scaled so that the weight for this bit is 1. The third term contains the contribution from the remaining DAC bits, where **C** is a vector containing the DAC bit weights, including their errors. The fourth term contains the contribution of the error correction table. The idea, of course, is to cause the error correction table to converge to a vector such that $\mathbf{C}+\mathbf{H}_\infty$ is equal to the expected binary bit weights for each of the three DACs.

To check the convergence of the correction table we can form the following recursion relation, replacing dot products with products of transposed vectors:

$$\mathbf{H}_{n+1} = \mathbf{H}_n + \frac{\mathbf{B}_n}{k}\left(X_n - b_n - \mathbf{B}_n^T(\mathbf{C} + \mathbf{H}_n)\right). \quad (8)$$

The elements of $\mathbf{B}_n$ that go to DAC2 and DAC3 are derived from the noise sources only, have zero mean, and are independent of the input signal. It can be shown that, except for the most significant bit, the bits driving DAC1 have zero mean and have zero crosscorrelation with the input signal. Therefore, if we take the expected value of the recursion equation, adding **C** to both sides and defining $\mathbf{W}_n = E(\mathbf{C}+\mathbf{H}_n)$, the expected value of the equivalent DAC bit weights after correction, we get:

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \frac{1}{k}(\mathbf{D} + \mathbf{F}\mathbf{W}_n). \quad (9)$$

where $\mathbf{D} = E(\mathbf{B}_n b_n)$ is a vector, and $\mathbf{F} = E\left(\mathbf{B}_n \mathbf{B}_n^T\right)$ is a matrix. The computation of **D** and **F** is a tedious process of representing the DAC drive bits in terms of the noise bits and the ADC1 bits, multiplying out the indicated products, and eliminating all terms that have a zero expected value because of the independence of the noise bits. The derivation will not be shown here, but the results are:

$$\mathbf{D} = \begin{bmatrix} \mathbf{SA}-\mathbf{A} \\ \mathbf{A}^T\mathbf{SA}-\mathbf{A}^T\mathbf{A}-4^{-N} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} & -\mathbf{SA} & \mathbf{0} & \mathbf{0} \\ -\mathbf{A}^T\mathbf{S}^T & 1 & -\mathbf{A}^T & -\mathbf{A}^T \\ \mathbf{0} & -\mathbf{A} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{A} & \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (10)$$

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \cdots & \frac{1}{2^N} \end{bmatrix}^T$$

$$\mathbf{S} = \begin{bmatrix} 0 & s_{12} & s_{13} & \cdots & s_{1N} \\ 0 & 0 & s_{23} & \cdots & s_{2N} \\ 0 & 0 & 0 & \cdots & s_{3N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

In **C**, $\mathbf{W}_n$, **F**, and **D**, the indexing is such that the first component corresponds to the second most-significant bit of DAC1 down to the least-significant bit, followed by the MSB of DAC2 down to its LSB, followed by the MSB of DAC3 down to its LSB. The elements of **S** are defined as $s_{ij}=E(a_{in}a_{jn})$ where $a_{in} = \pm 1$ is the ith bit of the ADC1 output during the nth sample period.

The recursion equation for $\mathbf{W}_n$ can be solved explicitly in much the same way as the corresponding equation for the gain control:

$$\mathbf{W}_n = -\mathbf{F}^{-1}\mathbf{D} + \left(\mathbf{I}-\frac{1}{k}\mathbf{F}\right)^n\left(\mathbf{W}_0 + \mathbf{F}^{-1}\mathbf{D}\right). \quad (11)$$

The matrix that is raised to the nth power can be written as $\mathbf{\Phi}^T\mathbf{\Lambda}\mathbf{\Phi}$, where $\mathbf{\Phi}^T = \mathbf{\Phi}^{-1}$, and $\mathbf{\Lambda}$ is a diagonal matrix. Multiplying both sides of the equation by $\mathbf{\Phi}$, we get a decoupled set of equations:

$$\mathbf{\Phi}\mathbf{W}_n = -\mathbf{\Phi}\mathbf{F}^{-1}\mathbf{D} + \mathbf{\Lambda}^n\left(\mathbf{\Phi}\mathbf{W}_0 + \mathbf{\Phi}\mathbf{F}^{-1}\mathbf{D}\right). \quad (12)$$

Since $\mathbf{\Lambda}$ is diagonal, its nth power is obtained by merely taking the nth power of each of its diagonal elements. Thus if the magnitude of each of these diagonal elements is less than one, the effective bit weights will ultimately converge to:

$$\mathbf{W}_\infty = -\mathbf{F}^{-1}\mathbf{D} = \begin{bmatrix} \mathbf{A} \\ 1 \\ \mathbf{A} \\ \mathbf{A} \end{bmatrix}. \quad (13)$$

Notice that the final value of the effective bit weights is totally independent of the input signal. In fact, substituting this back into the equation for G_n gives:

$$G_n = X_n - A_{1n},$$

or (14)

$$Y_n = X_n,$$

where G, X, Y, and A_1 are signals shown in Fig. 14.

The speed with which the effective bit weights converge to the desired result is determined by the magnitudes of the diagonal elements of $\mathbf{\Lambda}$. The exponential time constant for convergence is $\leq k$ samples for all initial bit weight errors that are orthogonal to the vector:

$$\mathbf{\Phi}_1 = \left[\mathbf{SA} \quad \sqrt{2\mathbf{A}^T\mathbf{A} + \mathbf{A}^T\mathbf{S}^T\mathbf{S}\mathbf{A}} \quad \mathbf{A} \quad \mathbf{A}\right]^T. \quad (15)$$

For the component of the initial bit weight errors collinear with the above vector, the time constant can be stretched somewhat compared to k. The amount of stretching is a function of the number of bits in ADC1 and the statistics of the input signal. The longest stretch factor occurs when the input signal is a constant at midscale of ADC1, a condition that occurs when no signal is applied to the input. In this case the stretched time constant, in samples, is:

$$\tau = \frac{k}{1-\sqrt{1-2^{1-N} + N4^{-N}}}. \quad (16)$$

For an 8-bit ADC1 this worst-case stretched time constant is 260k. In the HP converter the normal DAC correction time constant is 2^{22} samples. Thus, with a 10-MHz sample rate and no input signal applied, the stretched time constant is 109 seconds.

Just as in the case of the automatic gain control, the DAC correction table will fluctuate randomly around its ideal expected value. In this case, however, the resulting noise at the output is white and will not impact the measurement accuracy provided that the averaging time constant is sufficiently large.

Although the DAC and residue gain errors are effectively eliminated with the HP converter, the errors of ADC2 are still present. With ADC1 resolution of 8 bits the 110-dB desired linearity goal still requires that the linearity of ADC2 must be at least 62 dB. This requirement is made easier to meet by the fact that the statistical calibration process inherently adds a substantial amount of random noise to the input of ADC2 compared to the residual signal level. The effect of convolving this noise distribution with the ADC2 nonlinearity error curve generates a smoother nonlinearity curve with smaller deviations. Thus, the required linearity can be achieved even with a low-cost monolithic ADC2.

Hardware Implementation

The digital portion of the gain and DAC calibration schemes described above have been implemented in a single digital ASIC. Fig. 15 shows how this ASIC is connected to the analog components to provide the complete ADC. The track-and-hold amplifier becomes the limiting factor for low-order dynamic nonlinearities at high frequencies. The adder that drives DAC1 is implemented with ECL logic to reduce the delay. The switch on the input of the residue amplifier prevents any nonlinear DAC switching transients from entering the residue amplifier. This allows the bandwidth of the residue amplifier to be reduced, for lower noise, without fear of "smearing" the transients into the sampling aperture of ADC2.

The timing generator uses analog delay lines to time the sequence of events for a single ADC conversion cycle. Relative to the clock edge that starts the conversion, at 0 ns, the track-and-hold amplifier switches from track to hold. At 22.5 ns, ADC1 converts and the ASIC generates new noise outputs. At 45 ns, the clamp releases the summing node. At 67.5 ns,

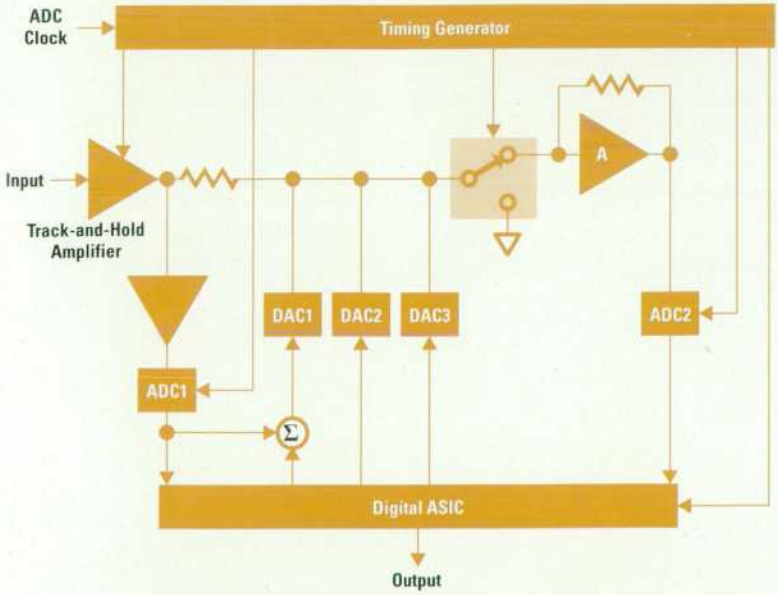


Fig. 15. Hardware block diagram of the 10-MHz ADC.

ADC2 converts. At 75 ns, the track-and-hold amplifier returns to track mode and the clamp grounds the summing node. At 90 ns, the timing circuit starts looking for another ADC clock edge.

Acknowledgments

The development of this precision ADC required the contributions of several people at the Lake Stevens Instrument Division. The critical and challenging design of the analog portions of the circuit was done by Dean Payne with help on the initial investigation of feasibility from Chuck Kingsford-Smith. Extensive simulation and verification of the calibration algorithm was done by Jerry Ringel, who also did the design of the calibration ASIC. The physical layout of the ASIC was done at the IC Business Division by Gary Petrie and Paul Nuber.

Reference

1. *Statistically Based Continuous Autocalibration Method and Apparatus*, U.S. Patent 4,996,530, United States Patent Office, February 1991.

Authors

October 1993

6 Modular Digitizing Oscilloscope

John A. Scharrer

An R&D section manager at the Colorado Springs Division, John Scharrer joined HP in 1965. He was born in Sheboygan, Wisconsin and attended the University of Wisconsin, from which he received a BSEE degree in 1963 and an MSEE degree in 1965. On past HP projects, he was a designer for the HP 1200 Series oscilloscopes and a designer and project manager for the HP 180 and 1700 Series oscilloscopes. He worked in a similar capacity on the HP 1615A logic analyzer, was project manager for the HP 64000 system logic and timing analyzers, and was section manager for the HP 64000 system emulators and the HP 54720/10 oscilloscopes. He currently manages development of high-performance digitizing oscilloscopes. His publications include two HP Journal articles and an EDN Magazine article related to amplifier design. John is married and has three daughters. His hobbies include tennis, hiking, and skiing.

11 8-GSa/s Data Acquisition System

Michael T. McTigue

Mike McTigue was born in Spokane, Washington and graduated from Arizona State University in 1979 with a BSEE degree. He joined the R&D laboratory at the Colorado Springs Division soon after that, and worked for several years on thick-film hybrid microcircuits for oscilloscopes. He also was part of the engineering team for the HP 54120T oscilloscope. For the HP 54720/10 oscilloscopes, he managed the group that developed the plug-in amplifiers in addition to working on the 1-GHz amplifier. Currently, he continues his involvement in high-frequency oscilloscope design. Mike's favorite pastimes include golf, skiing, and playing guitar.

Patrick J. Byrne

An R&D section manager at the Colorado Springs Division, Patrick Byrne has been with HP since 1982 when he joined the Circuit Technology Group. For several years, he was a device modeling and IC design engineer, working with various HP divisions. In 1985, he lectured at the University College of London, and then returned to the R&D lab, where he was a project manager for IC designs for workstations, peripherals, and test and measurement products. On his next assignment, he was a section manager for CAD tools, IC design, and product engineering. For the HP 54720 oscilloscope data acquisition system, he was project manager for the custom bipolar sampler and ADC ICs. Currently, he is a section manager for high-performance logic analysis systems. Patrick was born in Palo Alto, California and received a BSEE degree from the University of California at Berkeley in 1982 and an MSEE degree from Stanford University in 1988. His 1991 article on a 4-GHz 8-bit ADC system won a best paper award at the 1991 International Solid-State Circuits Conference. Patrick is married, has two children, and enjoys mountain biking and soccer.

24 Time Base and Trigger System

David D. Eskeldson

David Eskeldson is an electrical engineer at the Colorado Springs Division and joined HP in 1988 after completing work for a BSEE degree from Montana State University. He has contributed to the development of a high-speed trigger hybrid and worked on the time base board design and development for the HP 54720/10 oscilloscope. He's currently working on a high-speed logic analyzer.

Reginald Kellum

With HP since 1985, Reggie Kellum is an electrical development engineer at the Colorado Springs Division. He designed the one-picosecond-resolution time base interpolator for the HP 54720/10 oscilloscopes. Previously, he designed a time base and trigger system for the HP 54112 oscilloscope. Reggie received a BSEE degree from Stanford University in 1985 and an MSEE degree from the University of Arizona in 1991. He's now designing high-speed CMOS ICs.

Donald A. Whiteman

Born in Bexhill-on-Sea, Sussex, England, Don Whiteman received a degree in electrical engineering from Brighton College in 1956. After college, he developed radar systems and oscilloscopes for Marconi Instruments and designed nuclear radiographic products for Searle Radiographic. He joined the Colorado Springs Division in 1978. On his past HP projects, he designed oscilloscope CRTs, displays, and trigger systems. He developed the 2-GHz clock and the clock distribution system for the HP 54720/10 oscilloscopes. His present responsibilities involve the design of acquisition circuits for electro-optical instruments. In England, Don served for two years in the Royal Air Force and was named a chartered electrical engineer in 1968. He is married, has three children, and likes to ski.

31 2.5-GHz Active Probe

Thomas F. Uhling

Design engineer Tom Uhling came to HP's Colorado Springs Division in 1988 and specializes in the design of new probes and probing techniques. He was the electrical designer for the HP 54701A active probe. He is the author of three patents related to probes and the author or coauthor of two articles for trade magazines. Tom was born in Cottonwood, Idaho and attended the University of Idaho from which he received a BSEE degree in 1987. He's married and lists hunting, fishing, camping, target shooting, and bicycling as outside interests.

John R. Sterner

John Sterner is a hardware designer at the Corvallis Division and has been with HP since 1985, when he joined the Colorado Springs Division. A native of Winsted, Minnesota, he attended Coe College, from which he received a BA degree in mathematics in 1977, and the University of Idaho, where he completed work for a BSME degree in 1984. Initially, he was a process engineer and then moved to the R&D laboratory where he worked on the HP 54112 and HP 5412x oscilloscopes. He was responsible for the mechanical design of the HP 54701A probe. Before joining HP he designed dryers and evaporators for dairies. John is married and has a daughter. His hobbies include mountain biking, woodworking, and home brewing.

38 Interleaved ADC Systems

Allen Montijo

With HP since 1981, Allen Montijo is an R&D design engineer at the Colorado Springs Division. For the HP 54720A oscilloscope, he developed the hardware system, display board, acquisition board, and hybrid circuit logic. He also developed data acquisition and hybrid calibration software. His work on the HP 54111D involved similar design efforts. He is currently working on hardware design for future products and consulting on digital signal processing. He's named as an inventor in a patent on equivalent time sampling systems and is the author of an HP Journal article on HP 54111D signal processing. Born in Denver, Colorado, Allen received a BSE degree in electrical engineering and computer science from Princeton University in 1981 and an MSEE degree from Stanford University in 1985. He and his wife have two daughters and he enjoys photography, hiking, and skiing.

Kenneth Rush

Engineer/scientist Ken Rush has been with the Colorado Springs Division since 1979. He was the chief analog designer for the HP 54100A oscilloscope and the chief measurement system architect for the HP 54120A and HP 54720A oscilloscopes. His current responsibilities involve investigating digital design tools. Born in Winchester, Tennessee, he attended the University of Tennessee, from which he received a BS degree in aerospace engineering in 1973 and an MS degree in electronic engineering in 1975. He worked at NASA's Goddard Spaceflight Center and at Oak Ridge National Laboratories before joining HP. He's named as an inventor in two patents related to an oscilloscope active probe and an advanced ADC. He has also written numerous conference papers on analog-to-digital converters and oscilloscopes and received a best paper award from the International Solid-State Circuits Conference in 1991. He has published two articles in IEEE Spectrum magazine and is the author of an earlier HP Journal article. He serves on an IEEE committee that is writing a waveform measurement standard. Ken is married and has two children. During his leisure time, he breeds cocker spaniels for show and studies the stock market. He notes that the market is the ultimate signal-in-noise problem, so he is designing a trading system that takes advantage of his knowledge of signals and noise.

47 Pulse Parameter Accuracy

Kenneth Rush

Author's biography appears elsewhere in this section.

51 Oscilloscope Architecture

Dana L. Johnson

Dana Johnson was born in Edina, Minnesota and completed work for a BS degree in design science from the University of Minnesota in 1975. After leaving the university, he worked for Sheldahl, Inc. on the design of a solar collector control system. Currently a software engineer at the Colorado Springs Division, he joined HP in 1978 as a technical writer, later became an application engineer, and then joined the R&D laboratory to help develop the HP 19800 measurement library. Since then, he has worked on the HP 54200/201 and HP 54112 oscilloscopes. For the HP 54720/10, he did high-level architectural system design and designed and implemented the control and calibration of the plug-ins and acquisition system. He's currently working on enhancements to the HP 547xx oscilloscopes. His work on an oscilloscope probability density histogram display has resulted in one patent. Dana is married and has three sons. His outside interests include electronic music, guitar, duplicate bridge, gardening, and home brewing.

Christopher J. Magnuson

Chris Magnuson studied electrical engineering at Colorado State University, from which he received a BSEE degree in 1987. That same year he joined HP's Network Measurements Division, where he was a test engineer for the HP 8510 network analyzer. He later transferred to the Colorado Springs Division and is now developing software for digital oscilloscopes. He worked on the graphics processor system software for the HP 54720/10 oscilloscope. Born in Chicago, Illinois, Chris is married and has five children. Outside of software development, his interests include alpine mountaineering, ice climbing, and electronic music.

59 Firmware Development Processes

David W. Long

Project manager David Long is responsible for the development of system firmware and application software for the HP 547xx family of oscilloscopes. Earlier, he worked on the design of the computer display, system firmware, and user interface for the HP 54720/10. He was a software design engineer for a previous product, the HP 54120 oscilloscope. He's named as an inventor in a patent on a digitally synthesized gray scale for raster scan oscilloscope displays and is the author of a magazine article on oscilloscope color displays. David was born in Milwaukee, Wisconsin and completed work for a BSEE degree from Purdue University in 1984 before joining the Colorado Springs Division in 1985. His hobbies include watercolor painting, sketching, coin collecting, and science fiction.

Christopher P. Duff

R&D engineer Chris Duff has been at HP's Colorado Springs Division since 1988, the same year he completed work for a BS degree in computer science from Purdue University. He developed the postacquisition signal analysis firmware for the HP 547xx family of oscilloscopes and did research on algorithms for pulse parameter measurements and waveform mathematics techniques. He was also responsible for firmware development tool support. Currently, he's responsible for the acquisition and analysis firmware for the next-generation high-performance oscilloscope family.

66 Oscilloscope Mechanical Design

John W. Campbell

John Campbell was born in Washington, D.C. and grew up in Park Ridge, Illinois. He attended Bradley University, from which he received a BS degree in machine design engineering in 1967 and an MS degree in industrial and technical education in 1968. He was a captain and space systems analyst in the U.S. Air Force for four years before joining the Colorado Springs Division in 1972. He has been a design engineer or project manager for a number of products, including the HP 1740A, 1980A/B, and 54100A oscilloscopes, the HP 16500A logic analyzer, the HP 54720/10 oscilloscopes, and the HP 16517/18 timing analyzers. For the HP 54720/10 products, he developed the concept for the mechanical partitioning and architecture of the system and designed the cabinet and some of the subassemblies. He is named as an inventor in a patent on a position detector apparatus for a touchscreen system and is the author of two previous HP Journal articles. He's also a member of the American Society of Mechanical Engineers. John is married, has two sons, and is an avid golfer and model railroad enthusiast.

Kenneth W. Johnson

A California native, Kenny Johnson received a BS degree in mechanical engineering from Kansas State University in 1985 and came to HP's Colorado Springs Division the same year. He was a manufacturing development engineer responsible for automated printed circuit board assembly before transferring to R&D. He was responsible for the mechnical design of the HP 54711A, 54712A, 54713A, and 54721A plug-ins and for the plug-in slots of the mainframe design for the HP 54720/10 oscilloscope. Kenny is married and has two sons. His outside interests include running, mountain biking,

hiking, cooking, marine aquariums, and keeping exotic birds.

Wayne F. Helgoth

Design engineer Wayne Helgoth has been with the Colorado Springs Division since 1972. He held several engineering positions in the CRT department, where he developed the CRTs for the HP 1336A and HP 1345A displays and contributed to the development of the CRT for the HP 1727A oscilloscope. He continues to work on mechanical design for future products. He designed the cooling system and mechanics of the acquisition assembly for the HP 54720/10 oscilloscope. Born in Holyoke, Colorado, Wayne served in the U.S. Air Force for five years and received his BSME degree from the University of Colorado in 1971. He's married, has four children, and enjoys bicycling, hiking, travel, and working on home projects.

William H. Escovitz

Bill Escovitz is manufacturing engineering manager for new products at the Colorado Springs Division and has been with HP since 1979. Formerly an R&D project manager, he managed the mechanical design of the HP 54720/10 oscilloscope mainframe and plug-ins as well as all design for the HP 54701A active probe, mainframe CPU board, and display board. Previously, he was a development engineer or project manager for the HP 54003A high-impedance probe, the HP 10032A 100:1 probe, and the HP 54112A oscilloscope. A native of Pittsburgh, Pennsylvania, he studied physics at Dartmouth College (AB 1968) and at the University of Chicago (SM 1973 and PhD 1979). He taught high school science and mathematics and was a postdoctoral fellow at the University of Chicago before joining HP. Bill is a member of the American Physical Society and the author of about a dozen papers in physics, two in electronics, and one on project scheduling. He is married, has two sons, and enjoys skiing, running, and golf.

73 Wafer Test Probe Fixture

Daniel T. Hamling

A test development engineer at the Integrated Circuits Business Division, Dan Hamling developed the probe fixture used to test the data acquisition circuits for the HP 54720/10 oscilloscope. He has been with HP since 1987 and has developed numerous tests for components used in oscilloscopes, logic analyzers, pulse generators, and wireless communication products. He's named as the inventor in a pending patent on the probe fixture, and has written two other papers related to high-speed automatic wafer testing. Born in Bridgeport, Michigan, Dan attended the University of Michigan (BSEE 1987) and Stanford University (MSEE 1990). He and his wife have two daughters and he enjoys bicycling, basketball, baseball, and music.

76 Network and Spectrum Analyzer

Shigeru Kawabata

Shigeru Kawabata was born in Osaka, Japan and attended Kyoto University. He received his BSEE degree in 1980 and his MSEE degree in 1982, the same year he joined HP's Kobe Instrument Division. He was the R&D project manager for the HP 4396A vector network and spectrum analyzer, and earlier contributed to the development of the HP 4194A impedance/gain-phase analyzer and the HP 83595C sweeper plug-in. Shigeru is married and has two sons. His outside interests include gardening, collecting beautiful and precious stones, and painting and drawing.

Akira Nukiyama

Akira "Nuki" Nukiyama was born in Sendai, Japan and studied electrical engineering at Waseda University, from which he received a BSEE degree in 1981 and an MSEE degree in 1983. He is an R&D engineer at the Kobe Instrument Division, which he joined in 1983. He has contributed to the development of the HP 4194A impedance/gain-phase analyzer, the HP 54120A oscilloscope, and the HP 8751A network analyzer. He led the team that developed the firmware for the HP 4396A network and spectrum analyzer. He is currently working on software for PDC (personal digital cellular) testing. Nuki is married and likes skiing, bicycling, and camping with his wife and friends.

85 Receiver Design for the HP 4396A Analyzer

Yoshiyuki Yanagimoto

With HP since 1985, R&D development engineer Yoshi Yanagimoto managed the hardware development team and designed the receiver for the HP 4396A network and spectrum analyzer. His past projects include the HP 41941A/B impedance probe kits, the HP 4195A network and spectrum analyzer, and the HP 3588A spectrum and network analyzer. He is currently managing a project on synthesizer technology. He is named as an inventor in two patents related to mixer structure and a synchronized sampling method for gated sweep. Yoshi was born in Kyoto, Japan and received BS and MS degrees in physics from Kyoto University (1983 and 1985). He is married and has a young son. A former downhill ski racer, he still enjoys cross-country skiing as well as camping and skin diving. He says the thing he likes most is playing with his son.

95 Fast-Switching Multiplexer

Yoshiyuki Yanagimoto

Author's biography appears elsewhere in this section.

100 10-MSa/s ADC Module

Howard E. Hilton

Project manager Howard Hilton managed the development of the HP E1430A ADC module and is currently the project manager for the Lake Stevens Instrument Division Technology Group. He was born in St. Ignatius, Montana and studied electrical engineering at Montana State University, receiving a BSEE degree in 1973 and an MSEE degree in 1973. He joined HP's Loveland Instrument Division the same year. Products for which he has been a designer or manager include the HP 3571A tracking spectrum analyzer, the HP 3585A spectrum analyzer, the HP 239A low-distortion oscillator, and the HP 3565S multichannel measurement system. His work has resulted in three patents on digital IF, a self-identifying transducer, and a statistically calibrated ADC. Howard is married and has two children. His professional interests include digital signal processing algorithms, digital communications, surveillance, and ASIC development. He enjoys studying mathematics and physics when not at work.

105 Analog-to-Digital Converter

Howard E. Hilton

Author's biography appears elsewhere in this section.

Fr: Worldwide Roster/190 00106778 *
5731

To: LEWIS, KAREN
HP CORPORATE HEADQUARTERS
DDIV 0000 20BR

IDR #98234

HEWLETT-PACKARD
JOURNAL

October 1993 Volume 44 • Number 5

Technical Information from the Laboratories of
Hewlett-Packard Company

Hewlett-Packard Company, P.O. Box 51827
Palo Alto, California, 94303-0724 U.S.A.

Yokogawa-Hewlett-Packard Ltd., Suginami-Ku Tokyo 168 Japan

HEWLETT
PACKARD

5091-8890E