HEWLETT-PACKARD

Journal

FEBRUARY 1998

Technical Information from the Hewlett-Packard Company

HEWLETT-PACKARD Journal

FEBRUARY 1998

New Directions

This issue of the Journal completes the series of articles on wireless communications that we started in the December issue. There are also articles on HP CaLan, a cable system tester, and on the pager testing capability of the HP 8648A signal generator.

So, while this issue embodies continuity, there are also many changes taking place at the Journal. You may have noticed the graphic redesign that we rolled out with the December issue. Our goals are to make the Hewlett-Packard Journal easier to read and more contemporary looking. We're pretty sure we've achieved the second goal, and we hope we've succeeded with the first as well.

Other changes aren't quite as visible. In 1998, the Journal will become a quarterly publication. We'll publish in May, August and November. We're doing this for a couple of reasons. First, we want to devote more time and resources to our online edition of the Journal. If you would like to be notified, via e-mail, about new articles or features appearing at our site, use ***E-Mail Notification*** *at our website (see URL on next page) to register your e-mail address. We'll do the rest.*

THE COVER

A reflective look at communications appliances used in the past contrasted with those used today.

Volume 49 • Number 1

The Hewlett-Packard Journal is published bimonthly by the Hewlett-Packard Company to recognize technical contributions made by Hewlett-Packard personnel.

WHAT'S AHEAD

In May we will have eight articles on HP's implementation of OpenGL (a graphics API) and associated products, including one article on the Kayak PC worstation. We will also have articles on enterprise link software, a software tool for customer support, and a study of errors in forecasting the demand for components used in products.

The other reason for becoming a quarterly is that we want to spend more time to make the Journal even more reflective of what's going on at HP. The company is changing rapidly—new technologies, new markets, and new businesses are taking shape in every part of the company, and we want to bring our readers the stories of, and behind, HP's many new directions.

Finally, there's been a change on the Journal's editorial staff. Dick Dolan has retired after 32 years with HP, including 17 years as editor of the Journal. His high standards and wise stewardship have been crucial to the Journal's long-term success. I'd like to thank Dick for a job very well done.

Dick's successor is Steve Siano. Steve has been working for the Kobe (Japan) Instrument Division within HP's Test and Measurement Organization, and he brings a lot of fresh ideas and energy to the staff.

Steve Beitler
Executive Editor

The Hewlett-Packard Journal Online

The Hewlett-Packard Journal is available online at:

http://www.hp.com/hpj/journal.html

A quick tour of our site:

Current Issue—see it before it reaches your mailbox.

Past Issues—review back to February 1994, complete with links to other HP sites.

Index—search back to our first issue in 1949, complete with an **Order** button for issues or articles you'd like for your library.

Subscription Information—guidelines for U.S. and international subscriptions and a form you can fill out to receive an e-mail message about upcoming issues and changes to our Website.

Previews—contains an early look at future articles.

WWW

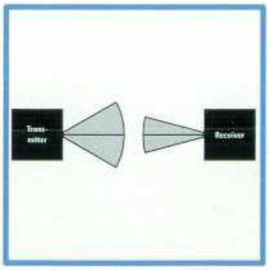
Articles

William J. McFarland

The tremendous growth in wireless communication products has created a parallel growth in research and development for high-performance components for these products.

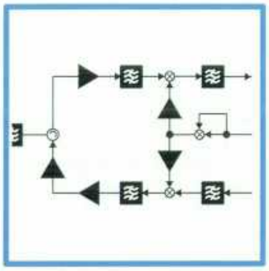
Iain Millar, Martin Beale, Bryan J. Donoghue, Kirk W. Lindstrom, and Stuart Williams

A core set of industry standards has been created to ensure interoperability among infrared-capable products.



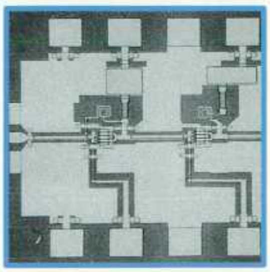
Kevin J. Negus, Bryan T. Ingram, John D. Waters, and William J. McFarland

Current RF wireless system standards and applications are placing demands on RF semiconductor manufacturers to produce highly specific and optimized RFIC solutions.



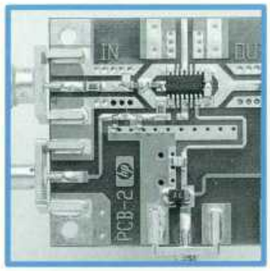
The MODFET Design and Fabrication Team

The authors describe the component characteristics that must be optimized to produce the high-performance MODFETs (modulation-doped FETs) used in HP's high-speed communications applications.



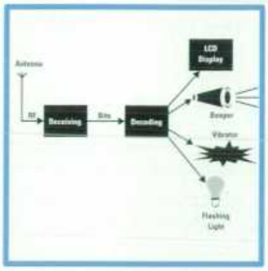
Der-Woei Wu, John S. Wei, Chung-Yi Su, Ray M. Parkhurst, Shyh-Liang Fu, Shih-Shun Chang, and Richard B. Levitsky

A 3-volt enhancement-mode pseudomorphic high-electron-mobility transistor (E-PHEMT) has been developed for the single-supply output power amplifiers used in PCS handsets.



Shunichiro Yajima and Antoni C. Niedzwiecki

The authors provide an overview of direct broadcast satellite (DBS) service and describe HP's efforts to develop a GaAs transistor used in the receiver portion of a DBS system.



The Hewlett-Packard Journal Online

http://www.hp.com/hpj/journal.html

What's new?

- The **Previews** section contains the following articles:

Linking Enterprise Business Systems to the Factory Floor

Knowledge Harvesting, Articulation, and Delivery

A Theoretical Derivation of Relationships between Forecast Errors

Strengthening Software Quality Assurance

A 150-MHz-Bandwidth Membrane Hydrophone for Acoustic Field Characterization

Reliability Enhancement of Surface Mount Light-Emitting Diodes for Automotive Applications

Engineering Surfaces in Ceramic Pin Grid Array Packaging to Inhibit Epoxy Bleeding

A Low-Complexity, Fixed-Rate Compression Scheme for Color Images and Documents

E-Mail Registration

- Use **E-Mail Notification** to register your e-mail address so that we can notify you when new articles are published.

WWW

Wireless Communications: A Spectrum of Opportunities

William J. McFarland

The tremendous growth in the consumer market for wireless communications products, such as cellular and cordless telephones, has created a parallel growth in research and development for higher-performance components for these products.

In the past five years consumer and manufacturer interest in wireless communication have exploded. Fueled by the tremendous acceptance of cellular telephones, cordless telephones, pagers, satellite delivered video, and simple infrared communications, the market for wireless communication hardware has grown to well over U.S. $10 billion per year worldwide. These market advances were enabled by tremendous reductions in cost and power consumption of wireless communications technologies.

Beyond the success of these established technologies, a wide range of new wireless communications services are being developed. Many of these services are substantial enhancements to existing systems, such as two-way paging, digital cellular service with short messaging capability, and sophisticated local communication using infrared such as the IrDA-NG (Infrared Data Association-Next Generation) standard. Other fast-growing systems will provide totally new capabilities. For example, the Global Positioning System for determining location (based on satellite communication) may eventually be used in all automobiles. Wireless multimedia networks and infrared-based control networks may someday be in every home.

The desire to communicate without wires appears natural. However, the value proposition for most wireless technologies is more diverse and subtle than the simple image of untethered information access would portray.

William J. McFarland

William McFarland is a project manager for radio communication circuits at HP Laboratories. Bill received a BSEE degree in 1983 from Stanford University and an MSEE degree in 1985 from the University of California at Berkeley. He joined HP Laboratories in 1985. Bill was born in Milwaukee, Wisconsin. He is married and has one daughter and he enjoys running, guitar, and cooking.

Advantages

In each application, the use of wireless communications provides the following advantages in varying degrees:

- Eliminates wiring
- Provides wide area coverage
- Allows mobility while using the service.

It is interesting to examine which of these advantages are fundamental to popular services. Cellular telephones provide mobility and wide area coverage, but since most homes and workplaces still get wired telephone service, cellular telephones have not reduced the amount of telephone wiring. Cordless telephones provide mobility within the home, but do not provide wide area coverage, nor have they eliminated much wiring (most homes have a telephone jack in many different rooms). On the other hand, direct broadcast satellite (DBS) systems do not provide mobility. What DBS provides is a way to compete with cable companies without having to string cable to every home. DBS is described in greater detail in the article on page 52.

Depending on the value of these wireless advantages, wireless systems must still compete with wired solutions on the basis of cost and performance. Despite the tremendous improvements in technology in the last few years, wireless communications circuits are generally more expensive and provide lower-bandwidth communication than wired solutions.

Wired (or fibered) solutions will always have an advantage in terms of bandwidth. Because signals on different wires interact very little, each new wire can use the same part of the electromagnetic spectrum as all others. Unfortunately, with wireless communication there is only one spectrum that must be shared among all users that are within range (coverage) of each other. In the case of services with nationwide coverage, such as cellular telephones and satellite systems, the spectrum allocated for those services typically can only be used for that one application. This situation limits the amount of bandwidth such a service could fairly use and has created government regulation to prevent interference.

Although the electromagnetic spectrum is infinite, not all of it is of equal value. The propagation properties, the cost of the technologies required, the regulatory issues, and available bandwidth all determine the value of a given piece of the spectrum.

Propagation

Frequencies between 10 and 30 MHz are able to propagate around the world with remarkably low transmit powers. Above 100 MHz, signals do not propagate over the horizon. Signals from 100 MHz to approximately 1 GHz can still penetrate most buildings, trees, and people fairly well, and can diffract easily around objects they cannot penetrate. From 1 GHz to approximately 100 GHz, signals progressively lose their ability to penetrate and diffract. Above 100 GHz, propagation is line-of-sight and has properties similar to light.

Technology

Technology is advancing rapidly, so any statement about relative costs holds only for a short time. Presently, there is an increase in cost somewhere between 1 and 5 GHz, with further increases somewhere between 12 and 30 GHz. An exception to the higher-frequency, higher-cost rule is infrared communications. The LEDs and photodetectors used to transmit and receive infrared signals (~500 THz) have become very inexpensive. However, these devices are noncoherent and must compete with the noise of sunlight. Therefore, infrared communication presently requires a great deal more transmitted power to reach a given range of communication. One potential advantage of higher frequencies (above 10 GHz, including infrared) is the small size and simplicity of highly directional antennas.

Regulation

Regulation is also changing (although much less quickly than technology). Presently, frequencies above 300 GHz are not regulated. Frequencies below 300 GHz fall into two regulation categories, licensed and unlicensed. To coordinate users and avoid potential interference, government agencies throughout the world require users to obtain a license for the use of most frequencies. While a licensed service has the advantage of protection from interference, the cost and administrative difficulties of obtaining a license for each transmitter are often prohibitive. Government agencies have set up some unlicensed frequencies. Transmitters in these bands do not need individual licenses. Instead, they must abide by certain rules, which are verified through type certification.* **Table I** lists some of the most important unlicensed bands and their properties. **Table I** also shows a general property that holds true for licensed and unlicensed frequencies—the higher the frequency,

* Type certification is a process by which the United States Federal Communication Commission (FCC) approves a product for sale but does not test each unit sold.

Table I

Selected Unlicensed Frequency Bands

Frequency	Bandwidth	Geographic Availability	Power Level
900 MHz	26 MHz	North America	Up to 1W
2.4 GHz	84 MHz	United States	Up to 1W
2.4 GHz	84 MHz	Europe	Up to 100 mW
2.4 GHz	26 MHz	Japan	Up to 10 mW
5.2 GHz	200 MHz	Available in U.S. and Europe for HIPERLAN only	50 or 250 mW
5.8 GHz	125 MHz	North America	Up to 1W
24 GHz	259 MHz	North America	Up to 25 mW
60 GHz	5 GHz	North America now and hopefully Europe and Japan soon	500 mW
>300 GHz	Theoretically infinite but for practical use <50 MHz	Worldwide	Limited by eye safety rules

the more bandwidth available for the service.

HP's Involvement

The combination of desired propagation, technology cost, and regulatory issues decides for any given application what frequencies to use. For each frequency range and application, a whole set of technologies are required. These go from the basic device technology, through packaging and system integration, to protocols and networking standards, and finally applications. HP is involved at all levels of this hierarchy over a wide range of frequencies, as the other wireless communications articles in this issue demonstrate.

The article on page 10 describes the applications and most widely accepted standards for infrared communications. The standards provide interoperability between devices for mobile professionals and consumers. The IrDA standards provide short-range, walk-up, point-and-shoot-type communications. Infrared provides a very low-cost implementation and worldwide freedom from regulation.

RF technology trade-offs for wireless communications is the subject of the article on page 27. The article concentrates on the frequency range from 900 MHz to 5 GHz. This frequency range has the most activity presently for local and wide area data networks. The article describes different cost and performance trade-offs, and the effects of the continually increasing performance and scale of integration in silicon technologies.

As the carrier frequency becomes higher, silicon alone cannot provide sufficient performance. One solution to this problem is a multitechnology chipset for 12-GHz direct broadcast satellite (DBS) receivers, which is described in the article on page 52. The delivery of video via a direct broadcast satellite has had the fastest initial growth rate of any consumer electronics product in history. HP has played a significant role in this success by providing very high-performance circuits at consumer prices.

An extremely advanced technology is described in the article on page 37, which describes a 0.1-μm MODFET (modulation doped FET) for use in high-speed wireless communications. This technology provides transistors with such tremendous performance that applications such as 60-GHz wireless LANs and 70-GHz collision avoidance radar become feasible. In addition, the technology can be used at 12 GHz, in which case the devices provide lower noise and higher power efficiency for DBS systems. This article includes a discussion of packaging issues as well, no simple matter for such high frequencies.

The article on page 39 takes a similar tack of using a very advanced technology (enhancement-mode PHEMT*) to provide remarkable performance in a commonly used frequency range. In this case, the frequency is the 850-MHz cellular telephone band, and the performance benefit is output power (two watts of RF power from a single IC operating from a 3V power supply) and excellent efficiency (50%).

* PHEMT = pseudomorphic high-electron-mobility transistor.

Conclusion

The incredible growth of wireless communications is bound to continue as consumers demand the convenience of untethered mobile access. The articles mentioned above are only a sampling of the work in wireless communications going on at HP. HP is involved in nearly every wireless communications market supplying everything from discrete devices, integrated circuits, and packaging to complete communication modules, networking, and application software. This breadth and depth of involvement positions HP to take full advantage of this rapidly growing market.

The IrDA Standards for High-Speed Infrared Communications

Iain Millar

Martin Beale

Bryan J. Donoghue

Kirk W. Lindstrom

Stuart Williams

As more data communications products, such as printers and laptop PCs, are released with infrared capability, support for a core set of IrDA standards has strong support from many manufacturers because, among other things, they want to ensure that their products will interoperate in a transparent and user-friendly manner.

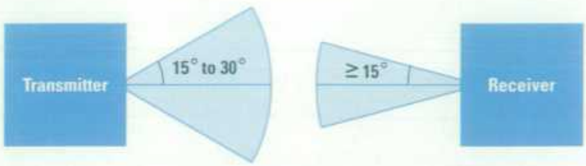
The use of infrared techniques for data communications has been around for several years, and by 1993 several commercial products were available with this capability. However, each company has tended to have its own infrared standard, and although devices from the same manufacturer could communicate with each other, competing systems tended not to be interoperable. Examples of such proprietary infrared systems include Hewlett-Packard's HP SIR (serial infrared), Sharp's ASK systems, and General Magic's MagicBeam. The resulting confusion in the marketplace meant that users viewed infrared as having only limited utility.

On June 28, 1993, the Infrared Data Association (IrDA) had its first meeting with the purpose of establishing a ubiquitous, low-cost, point-to-point serial infrared standard. Some 50 representatives from 20 interested companies were expected, but over 120 people representing more than 50 companies actually attended. It was clear that the industry was interested in developing a standard that would allow the true value and utility of infrared to be realized. At the culmination of this process—and due in no small part to the enthusiasm and spirit of cooperation of the participating companies—the first IrDA standards were published, just one year and two days after the initial meeting.

To date, IrDA has specified the physical and protocol layers necessary for any two devices that conform to the IrDA standards to detect each other and exchange data. The initial IrDA 1.0 specification detailed a serial, half-duplex, asynchronous system with transfer rates of 2400 bits/s to 115,200 bits/s at a

Figure 1

Viewing angle specified in IrDA specification 1.0.



range of up to one meter with a viewing half-angle of between 15 and 30 degrees (see **Figure 1**). More recently, IrDA has extended the physical layer specification to allow data communications at transfer rates up to 4 Mbits/s.

This paper presents details of the individual IrDA specifications, focusing specifically on the high-speed extensions that allow data communications at up to 4 Mbits/s. The first section gives details of the objectives that resulted in the series of IrDA specifications. The specifics of the user model and the technical requirements of the specification are also presented. Next the IrDA architecture is introduced, highlighting how the IrDA specifications together provide overall functionality. The infrared physical-layer specification with particular emphasis on modulation format, packet framing, transceiver design, and clock recovery is discussed in the next section. The transceiver design for the HP HDSL-1100 IrDA transceiver is also described in this section. The last section covers the protocol layers of the IrDA specifications. Finally, IrDA's current status is summarized.

IrDA Objectives

When IrDA was established, it set for itself the following objective:

"To create an interoperable, low-cost infrared data interconnection standard that supports a walk-up, point-to-point user model* that is adaptable to a broad range of mobile appliances that need to connect to peripheral devices and hosts."[1]

IrDA chose the short-range, walk-up, point-and-shoot directed infrared communications model for two main reasons. First, it was perceived that the initial target market for IrDA-enabled devices would be the mobile

* The phrase "walk-up, point-to-point user model" refers to the fact that to ensure data transfer between devices with infrared capabilities, they must be placed close together (< 2 m) with their infrared transceivers pointed at one another.

Glossary

Cell. A symbol in PPM.

Chip. A pulse within a symbol (cell) in PPM.

ENDEC. The encoder-decoder used in the IrDA physical layer.

HTTP. Hypertext Transfer Protocol.

HDLC. A bit-oriented, synchronous High-level Data Link Control protocol that applies to the message-passing (data link) layer of the Open Systems Interconnect (OSI) model for computer-to-computer communications.

IAS. The information access service maintains information about the services available on the host device and provides services that allow access to information on remote devices.

IrCOMM. IrDA specification for the emulation of serial and parallel port communications.

IrLAN. IrDA specification for accessing a LAN over an infrared medium.

IrLAP. IrDA specification for Link Access Protocol. This document specifies an HDLC-based protocol for controlling access to the infrared medium.

IrLMP. IrDA specification for Link Management Protocol. This protocol provides the LM-MUX and LM-IAS services.

IrOBEX. IrDA specification that defines the protocol for generic object exchange in an IrDA-enabled device.

IrPHY. The specification that describes the physical layer properties of the IrDA standard.

LM-IAS. The Link Management Information Access Service allows a pair of IrDA devices to interrogate each other to determine the services available on each device.

LM-MUX. The Link Management Multiplexer allows any pair of IrDA devices to simultaneously and independently use a single IrDA connection between themselves.

LSAP. Link Service Access Ports are address fields that uniquely identify applications on the source and destination devices.

LSAP-SEL. Link Service Access Port Selector.

PPM. Pulse position modulation.

SIR. Serial infrared.

Tiny TP. Lightweight transport protocol specification.

professional who is also a computer user. The environment for the use of such devices would be in a typical working environment in which the majority of stationary devices, such as printers or computers, would be located within the user's own reach space, on the desktop or in the immediate vicinity. Typical use of such devices would consist of short, conscious interactions such as file transfer or printing. Such use scenarios do not require the devices to be continually connected to each other, and a directed model of communications was adopted in which the user consciously points the infrared device at the target.

Previously, mobile professionals might connect their laptops to various peripherals using parallel or serial cables. Connecting such devices using LAN connections might also be possible if cost were not an issue. However, a problem arises when the user becomes mobile—for example, when visiting customers in their office. Setting up a laptop at the customer office to achieve even simple tasks, such as printing or file transfer, would more than likely require significant reconfiguration. IrDA aimed to change this by providing standards for ubiquitous access to such devices.

Second, IrDA chose this communications model to minimize cost. The use of a single LED and photodiode in the transceiver enables an extremely low-cost implementation. The model simplifies the protocol software by restricting the number of visible devices, hence limiting the contention and interference between IrDA devices. The limited range also allows reuse of the infrared medium, allowing multiple pairs of devices to communicate at the same time.

IrDA aimed to allow its standards to support a wide class of computing devices and peripherals that might be used by mobile professionals. These devices would range from very sophisticated, high-power notebook or laptop personal computers, through palmtop computers and personal digital assistants, to simple single-function devices like electronic business cards or phone dialers. Target peripheral devices would include conventional computer-oriented devices like printers and modems, as well as automatic teller machines and public and mobile telephones. It was also envisaged that IrDA would enable new classes of devices such as information access points.

To target such a broad range of devices, a set of general requirements was placed on any prospective standard. These requirements included:

- Low cost
- Industry standard
- Compact, lightweight, low-power
- Intuitive and easy to use
- Noninterfering.

Using these requirements, the IrDA committee developed a series of standards aimed at providing ubiquitous, low-cost, directed infrared communications for all classes of mobile computing devices. In IrDA's vision of the world, the user of such devices would be able to roam across international boundaries using IrDA communications to access information, computing, and communications services in a uniform and transparent manner. The days of the mobile computer user travelling the globe with a multitude of modem, serial, and parallel cables, including adapters, will be gone.

The remainder of this paper presents details of the standard IrDA has put in place to achieve this vision.

The IrDA Architecture

After the initial marketing requirements had been specified, the technical committee within IrDA moved quickly towards the development of the initial standards. In April 1994, the first IrDA standard was published covering the physical layer properties. This document, the Infrared Physical Layer (IrPHY) specification,[2] describes an infrared transmission system based on a UART modulation strategy. The document specifies the necessary parameters to provide an asynchronous half-duplex serial communications link over distances of at least one meter at data rates between 2400 bits/s and 115.2 kbits/s. The cone half-angle of the infrared transmission is specified as being at least 15 degrees, but no more than 30 degrees. The IrPHY specification was quickly followed with the publication of the Infrared Link Access Protocol (IrLAP) in June 1994.[3] IrLAP specifies an HDLC-based protocol for controlling access to the infrared medium and providing the basic link-level connection between a pair of devices.

During the development of IrPHY and IrLAP, it was realized that some additional functionality was required in addition to the ability to provide a single connection between a pair of devices. The Infrared Link Management (IrLMP) layer was conceived.[4] This layer has two primary functions.

First, it provides the mechanism by which multiple entities within any pair of IrDA devices can simultaneously and independently use the single IrLAP connection between those devices. This function is called the link management multiplexer (LM-MUX).

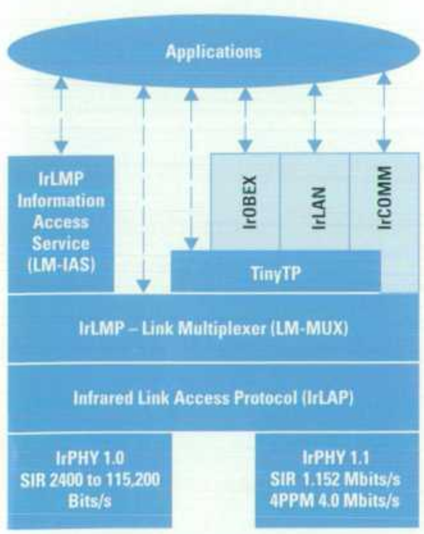
Second, it provides a way for entities using the IrDA services to discover what services are offered by a peer device and to register available services within the local device. This link management information access service (LM-IAS) considerably benefits the ease of use of portable devices, allowing pairs of devices to interrogate each other to discover information about the applications within each device.

These three standards—IrPHY, IrLAP, and IrLMP—form the core of the IrDA architecture, and all are required for a device to be IrDA-compliant. Since the core documents were published, several extensions have been added. The current complete IrDA architecture is shown in **Figure 2**.

In October 1995, optional extensions to the physical layer, adding data transmission speeds of up to 4 Mbits/s, were accepted by the IrDA committee. These changes resulted in the IrDA IrPHY 1.1 specification.[5] The IrLAP and IrLMP documents have also recently been updated to version 1.1 to incorporate various improvements that resulted from practical experience in implementing and using the IrDA protocols.[6,7]

Figure 2

The IrDA architecture.



In addition to the base standards, IrDA has specified a protocol called Tiny TP.[8] This protocol is an extremely lightweight transport protocol designed to provide application-level flow control as well as segmentation and reassembly of application data units. This protocol has proved to be useful and is now implemented by most applications that support the IrDA architecture.

To complement the functionality of the main components of the IrDA architecture, several application-level protocols have been and are in the process of being developed. These protocols are aimed at providing convenient and uniform interfaces to the functionality of the IrDA protocols for both old and new applications.

The original target for IrDA was cable replacement. The need for a protocol to support the redirection of serial and parallel cable traffic resulted in the IrCOMM serial and parallel port emulation protocol specification.[9] This protocol enabled the redirection of conventional serial and parallel ports over the infrared medium, allowing many existing applications to operate unchanged over an IrDA link. Another area seen as a suitable application of IrDA, particularly as a result of the high-speed extensions, is wireless access to local area networks. The protocol IrLAN was developed to allow an IrDA-enabled device to access a LAN over the infrared medium.[10] The protocol, in combination with an IrLAN-compatible LAN access device, provides the IrDA device with the equivalent functionality of a LAN card and the advantages of wireless connectivity.

Both IrCOMM and IrLAN address legacy-style applications. However, it is envisioned that many new applications will be enabled by the IrDA standards. Using IrDA on low-end devices gives rise to the need for a flexible, lightweight information exchange protocol suitable for devices with varying resource capabilities. A protocol for generic object exchange, IrOBEX, is currently under development within IrDA.[11] This protocol is based on HTTP (Hypertext Transfer Protocol) but is more compact. When completed, IrOBEX will provide a device independent method for exchanging arbitrary units of data between IrDA-enabled devices.

The IrDA Physical Layer

The IrDA physical layer is split into three distinct data rate ranges: 2400 to 115,200 bits/s, 1.152 Mbits/s, and 4 Mbits/s. Initial protocol negotiation takes place at 9600 bits/s, making this data rate compulsory. All other rates are optional and can be added if a device requires a higher data rate. The links are designed to be used in a line-of-sight, point-and-shoot manner and hence have a modest minimum coverage of one meter, with a ±15° viewing angle. This modest coverage is advantageous, since it allows a low-cost, high-data-rate link to be produced in a small package.

2400-to-115,200-bit/s Link

This is based on the HP-SIR link developed for HP calculators.[12] All IrDA-compliant devices implement this type of link since initial protocol negotiation takes place at 9600 bits/s. The architecture of the link (**Figure 3**) is designed for easy implementation and low cost. Hardware costs can be kept to a minimum by implementing the protocol, packet framing, and CRC calculation in software on the host processor. Bytes of data from the processor are converted to a serial data stream by a UART (universal asynchronous receiver-transmitter). Since many systems already include a UART for RS-232 communications, this places no extra cost burden on the system. Only the ENDEC (encoder-decoder) and transceiver represent an additional hardware cost for the system.

Infrared receivers contain a high-pass filter to remove background daylight. This high-pass filter forces the use of encoding on the link to ensure that long strings of zeros or ones are not lost in transmission. The encoding used on this link is return-to-zero (RZ). Zeros are represented by a pulse of 3/16-bit duration, and ones by the absence of a pulse (**Figure 4**). For example, 3/16 of a pulse width at 115,200 bits/s is 1.6 μs. The code is power-efficient since infrared light is only transmitted for zeros. The tall narrow pulse has better signal-to-noise ratio performance than a short wide pulse of the same energy.

Figure 4

The coding on a 2400-to-115,200-bit/s link.

The 1.152-Mbit/s Link

At speeds above 115,200 bits/s, packet framing and CRC generation and checking become a significant burden to the host processor. At 1.152 Mbits/s, these tasks are performed in hardware by a packet framer (see **Figure 5**). The packet format is slightly different from that used in the 2,400-to-115,200-bit/s link, but the line code remains similar.[5] Higher-level protocols are less processor intensive than packet framing or CRC generation and are still implemented in software on the host processor.

The 4-Mbit/s Link

The 4-Mbit/s link architecture is shown in **Figure 6**. As in the 1.152-Mbit/s link, packet framing and CRC generation and checking are performed in hardware to relieve the burden on the host processor, while higher-level protocols are implemented in software on the host processor. The link uses a new encoding scheme (described below) and a new, more robust packet structure. A phase-locked loop replaces edge detection as the means of recovering the sampling clock from the received signal. The packet

Figure 3

The 2400-to-115,200-bit/s link architecture.

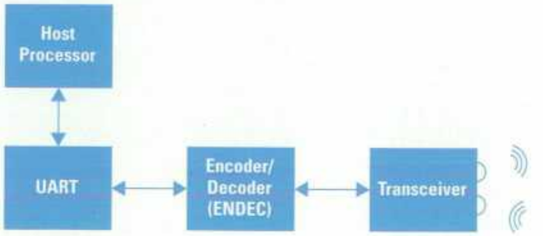


Figure 5

The 1.152-Mbit/s link architecture.

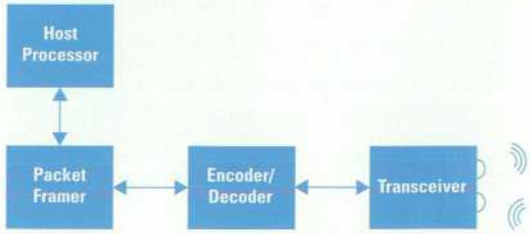


Figure 6

4-Mbit/s link architecture.

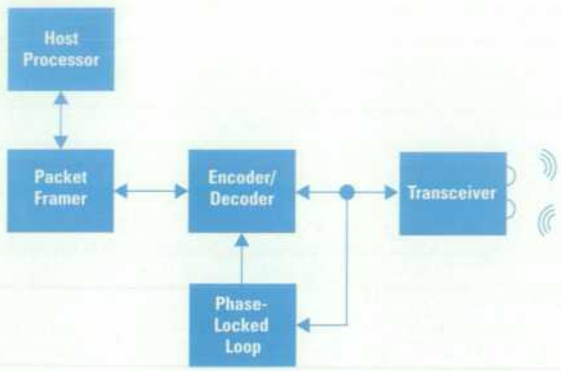
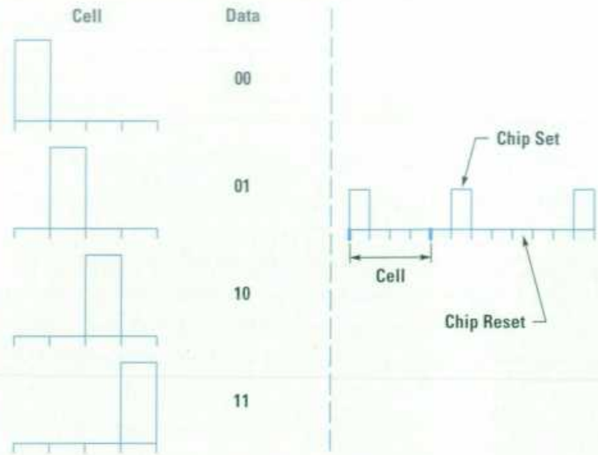


Figure 7

4PPM message encoding.



framer, ENDEC, and phase-locked loop are more complex than the UART and ENDEC in the 2400-to-115,200 bit/s link. However, this added complexity need not be expensive. The components are specified in a hardware description language and can be added quickly and inexpensively to one of the host system's ASICs. PC chipsets including the 4-Mbit/s hardware are already available from leading semiconductor manufacturers.

Coding and Packet Format. Pulse position modulation (PPM) was chosen as the line code for the 4-Mbit/s link. Data is transmitted within a PPM signal by varying the position of a pulse (referred to here as a chip) within a symbol (referred to here as a cell). The PPM modulation for the 4-Mbit/s link allows one chip to be set in one of four possible positions; thus it is known as 4PPM. Since a chip can be set in one of four possible positions, four different messages can be sent within one cell, allowing two bits of data to be encoded per cell. **Figure 7** shows the four possible messages that can be transmitted by 4PPM.

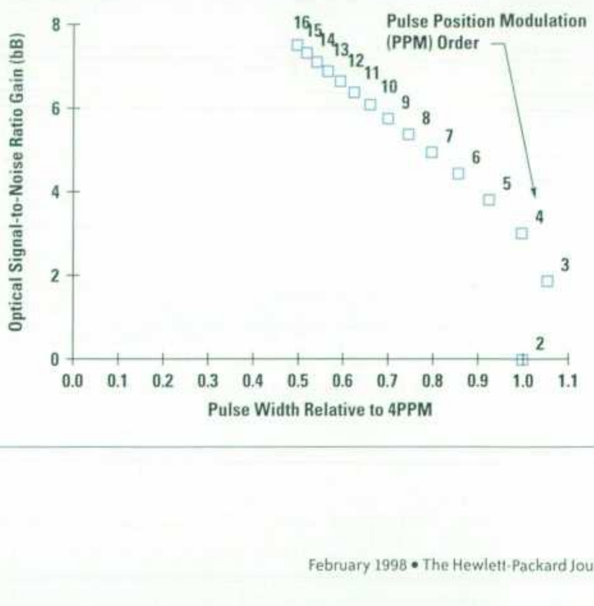
Pulse position modulation has many properties that make it attractive for use on the free-space optical channel. One of the main properties is the sparseness of the code. Sparse code allows high peak powers to be employed for set chips while maintaining a reasonable average power. The eye-safety rules stipulate a maximum average optical power, and LEDs tend to be average-power-limited at moderate duty cycles.

Pulse position modulation also contains significant and regular timing content, which facilitates synchronous clock recovery using a phase-locked loop. It is a modulation format that has very little dc content and can be high-pass filtered at 100 kHz, avoiding interference generated by fluorescent lighting without adversely affecting the receiver's eye diagram. A particularly interesting feature of PPM—one that had important ramifications in the choice of end delimiters—is its ability to detect line code errors.

Higher orders of PPM give lower duty cycles and theoretically greater signal-to-noise ratio gains on the infrared medium. **Figure 8** illustrates the interesting relationship between signal-to-noise ratio gain achievable with various

Figure 8

The signal-to-noise ratio gain and pulse width trade-off.



orders of PPM and the required pulse width. It is interesting to note that the optimum order of PPM from a bandwidth efficiency perspective would be 3PPM. This result might be of theoretical interest, but is fairly useless in a practical system. Since the fastest bright LEDs have a rise time of around 40 ns, and the rise time of an LED is proportional to the pulse width, the use of high-order PPMs at 4 Mbits/s becomes impractical. The decision to adopt the order four for the PPM was motivated by knowledge of the range of duty cycles over which LEDs are peak-power-limited, the rise and fall time of available LEDs, and the frequent timing content provided at order four.

Packet Format. The 4-Mbit/s physical layer packet has distinct features that perform a useful and well-defined role (see **Figure 9**). A preamble allows dc balance to be attained, and more important, permits the phase-locked loop to achieve chip-level synchronization. The length of the preamble was considered carefully such that the preceding two goals could be achieved without a significant impact on efficiency. The start and stop delimiters provide cell and frame synchronization and were chosen so as not to compromise overall packet robustness or adversely affect the receiver eye diagram. To distinguish the preamble and the end delimiters from the frame body, these fields contain code violations. The body of the packet is 4PPM-coded and has a 32-bit cyclic redundancy check (CRC) field appended to it. The choice of a 32-bit CRC provides a guaranteed level of robustness to undetected data errors over the range of error rates expected on a free-space infrared channel. The CRC is performed on the data bits rather than on the PPM-encoded chips.

Error Detection and Delimiters. A decoder may choose to exploit the error detection capabilities of 4PPM. The only portions of the packet allowed to contain violations are the preamble and the frame delimiters. If a decoder finds code violations within the frame body or CRC portion of the packet, it can flag that packet as being corrupted. In

Figure 9

The 4-Mbit/s packet format.

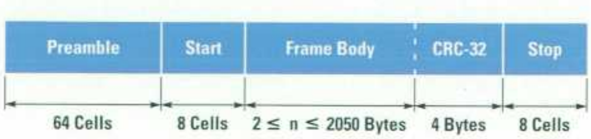
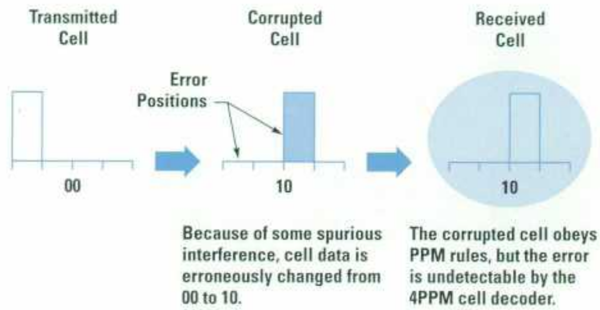


Figure 10

An undetectable 4PPM error.



the same way that a sufficient number of carefully positioned errors can produce a correct-looking CRC for a corrupted packet, there are some error patterns that a 4PPM decoder cannot detect. An example is shown in **Figure 10**.

The role of the CRC is to detect those error patterns that the PPM cell decoder cannot detect. Owing to the combined distance structure of the CRC and the pulse position modulation, the packet can be made very robust to withstand either random or burst errors at any signal-to-noise ratio.

A more worrisome error mechanism that had to be considered was the possibility of the corruption of the frame delimiters. The frame delimiters are not in themselves protected by the CRC. If the situation arose whereby a false Stop delimiter appeared in a valid position within the data and CRC portion of the packet, the packet would be protected solely by the scrambling effect of the CRC. In this case, a corrupted packet would be flagged as correct with a probability of $(0.5)^{32}$. Thus, it is important to ensure the unlikelihood of either random or burst errors causing a false delimiter to appear within the data portion of the packet. This is achieved by choosing delimiters with a large Hamming distance from the data (or shifted versions of the data, to ensure serial uniqueness) and with a sufficient number of chips such that "bursty" channel error models can be tolerated. A further constraint on the delimiter choice is that delimiters must not adversely affect the eye diagram of the complete packet. The lack of long strings of contiguous set or reset chips within the 4-Mbit/s delimiters allows this goal to be attained. The delimiters chosen ensure packet robustness at any signal-to-noise ratio, for any length of packet, over random and burst-error models—all without affecting the receiver eye diagram.

Clock Recovery. The UART-style clock recovery of the 2400-to-115,200-bit/s link uses a single signal edge to set the phase of the recovered sampling clock. This inevitably gives rise to phase jitter on the recovered clock and a consequent signal-to-noise ratio penalty. The phase-locked loop used by the 4-Mbit/s link generates a sampling clock with much less jitter because it uses timing information from many signal edges to set the phase of the clock. An analog phase-locked loop could have been used for clock recovery and might have achieved a low phase jitter, but it would have been unable to achieve the rapid phase lock of a digital phase-locked loop. Rapid phase lock is important in a packetized data system, because it determines the length of the training sequence, or preamble, required at the start of every packet to allow the phase-locked loop to lock.

The lock time is dictated by the accuracy with which the nominal frequency of the phase-locked loop's variable oscillator can be set. The nominal frequency of the variable oscillator in an analog phase-locked loop is highly variable, since it is determined by the (usually poor) tolerance of the resistors and capacitors. By contrast, the nominal frequency of the variable oscillator in a digital phase-locked loop can be locked to a crystal reference with a tolerance of less than 100 ppm. Implementations of digital phase-locked loops have the additional advantage that they can be quickly and easily ported between ASIC designs. The architecture of a typical digital phase-locked loop for the 4-Mbit/s link is shown in **Figure 11**.

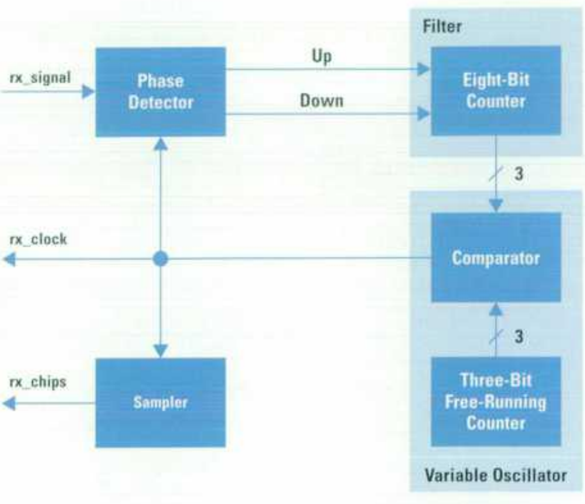
The phase detector is a state machine that compares the edges in the received signal (rx_signal) with those of the recovered clock (rx_clock). Rising edges only occur in rx_signal at PPM chip boundaries. Rising edges of rx_clock should occur halfway between chip cell boundaries. If rx_signal is earlier than expected, then the phase detector produces a Down signal, thereby advancing the phase of rx_clock. If rx_signal is later than expected, then the phase detector produces an Up signal.

The three most significant bits of the 8-bit counter set the phase of rx_clock. The five least significant bits ensure that the counter acts as a low-pass filter, since many Up and Down signals are required to change the phase of rx_clock. The three-bit free-running counter and the comparator together act as a variable phase oscillator. All blocks within the phase-locked loop are clocked by the same system clock. The system clock can be either 40, 48, 56 or 64 MHz, the choice being set by the rollover point of the three most-significant bits of the 8- and 3-bit counters (100, 101, 110, or 111). A 40-MHz system clock means that rx_clock should be very granular, with only five possible phase steps within a chip period. The effective number of phase steps is, however, doubled by making use of both the positive and negative edges of the system clock in the phase detector and sampler. The choice of whether to use positive or negative edges can be made by examining the fourth most-significant bit of the 8-bit counter.

The fast lock of the digital phase-locked loop is further aided by using a dual control loop within the digital phase-locked loop. A lock state machine within the phase detector decides whether the digital phase-locked loop is in or out of lock by examining the average deviation of the rx_clock edges from the rx_signal edges. If the digital phase-locked loop is out of lock, then multiple Up or Down pulses are generated for each edge in rx_signal to ensure rapid lock. Once locked, only single Up or Down pulses are generated since multiple pulses would increase phase jitter on rx_clock.

Figure 11

4-Mbit/s digital phase-locked loop.



The Hewlett-Packard HSDL-1100 IrDA Transceiver

The HP HDSL-1100 from HP's Communication Semiconductor Solutions Division is the world's first fully IrDA-compliant transceiver capable of operating at all IrDA data rates from 2400 bits/s to 4 Mbits/s. The HSDL-1100 fits within the same small package as its predecessor, the HSDL-1000, which operated at data rates from 2400 bits/s to 115,200 bits/s. The small package size available for pins, IC, passive components, and heat dissipation imposed design constraints on the complexity of the transceiver. The IC uses a low-density bipolar in-house process, which is low in cost and allows quick turn times on wafers for IC development.

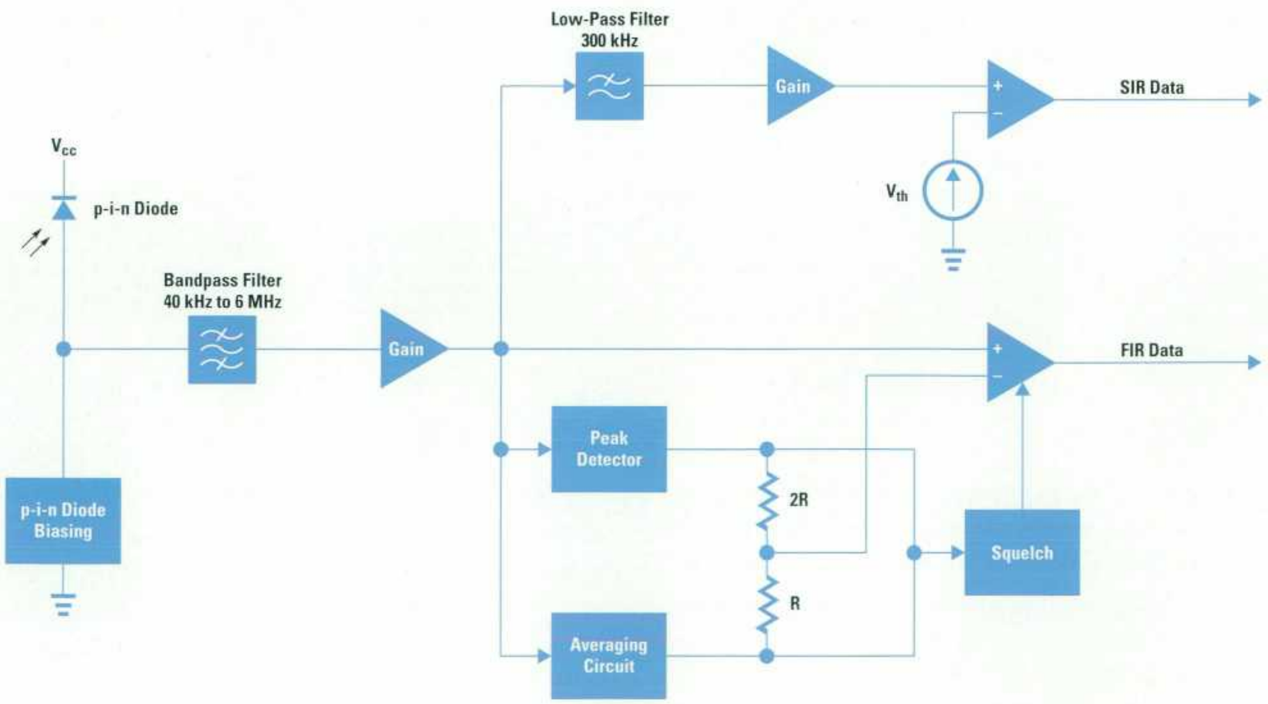
Transmitter design was straightforward. However, the multiple data rates, line codes, and large dynamic range made receiver design much more challenging. The receiver's dual-channel architecture is shown in **Figure 12**. A shared p-i-n diode detects all infrared signals with a modulation frequency between 40 kHz and 6 MHz. An amplifier boosts this signal before it is split into separate receiver channels. IrDA signals at 2400 to 115,200 bits/s pass through the serial infrared (SIR) channel*, while 1.152-to-4-Mbit/s signals pass through the fast infrared (FIR) channel. The lower bandwidth of the SIR channel (40 to 300 kHz) means lower noise and allows the SIR channel to meet the IrDA 4 $\mu W/cm^2$ sensitivity requirement. The higher-bandwidth (40 kHz to 6 MHz) FIR channel has higher noise, but still meets the 10 $\mu W/cm^2$ sensitivity requirement for 1.152-to-4-Mbit/s IrDA links. Since the different data rate IrDA links overlap in their modulation spectra, the received signal will appear on both channels. The ENDEC relies on information provided by the protocol to ensure that it listens on the correct channel.

The receiver converts signals from an analog to a digital form by comparing them with a threshold voltage. The two channels have different threshold detection circuits to meet the different requirements for the signals. The SIR channel has a fixed threshold set at the level of the weakest received signals. Although the fixed threshold

* At low rates, such as 2400 or 9600 baud, only the leading edge of the signal passes through the 40-kHz to 6-MHz bandpass filter. The signal is still correctly decoded since the ENDEC is able to tolerate received SIR pulses as short as 1 to 4 μs.

Figure 12

The HDSL-1100 receiver architecture.



tends to extend the duration of high-level pulses, the line code for the 2400-to-115,200-bit/s ENDEC is tolerant of pulses that extend to five times their nominal width. The 4-Mbit/s ENDEC is far less tolerant of pulse extension, so a dynamic threshold is required on the FIR channel. The dynamic threshold tracks the 50% level between the peak extensions of the 4PPM signal. A peak detector tracks the 100% level of the signal and an average circuit tracks the 25% level. The 50% threshold level is derived from a 2R-R voltage divider connected to these levels. Between packets, the dynamic threshold drops to zero. This would allow the FIR_Data output to "chatter" on noise or on feedback between the output pin and the p-i-n diode. The 1.152-Mbits/s ENDEC is intolerant of the extra pulses produced by such chatter, so a squelch circuit was added to switch off the FIR_Data output at low signal levels. The dynamic threshold also takes time to settle at the start of a packet, which causes some of the packet's initial infrared pulses to be lost or distorted. While this would be disastrous for the 2400-to-115,200-bit/s link, the 1.152- and 4-Mbit/s packets include a preamble to allow the receiver to settle before decoding data.

Another challenge for receiver design was the dynamic range of infrared signals. IrDA specifications allow received signal strength to vary between 4 $\mu W/cm^2$ and 500 mW/cm^2. This is a dynamic range of 51 dB. Since the p-i-n diode is a square law detector, this dynamic range doubles to 102 dB within the receiver. The receiver achieves this dynamic range by allowing the signal to be clipped while maintaining the timing of the signal. The impedance of the p-i-n diode biasing circuit decreases with signal level, reducing the signal voltage and the receiver amplifier's limit without saturating. The p-i-n diode has also been carefully designed to ensure that the induced signal decays rapidly once an infrared pulse disappears.

The IrDA Protocol Layers

The Infrared Link Access Protocol

IrLAP is the IrDA protocol that provides the basic link layer connection between a pair of IrDA devices. It is based on the HDLC protocol providing functions like connection establishment, data transfer, and flow control.[13,14] However, IrLAP has significant additional features as a result of the specific properties of the infrared medium.

The infrared medium over which IrLAP is required to operate is a point-to-point, half-duplex medium. While the narrow cone angle of IrPHY limits the number of other devices that can be seen, it does increase the probability of hidden devices. In such a situation, one device may see many other devices. However, it does not follow that those devices will see each other. This can result in collisions where transmissions from devices hidden from each other may overlap, resulting in the inability of the receiving device to decode those frames correctly. The characteristics of the infrared medium also result in there being no reliable way to detect transmission collisions. Conventional carrier sensing with collision-detection protocols would therefore be unsuitable, and IrLAP provides a mechanism for ensuring contention-free access to the medium, at least during data transfer.

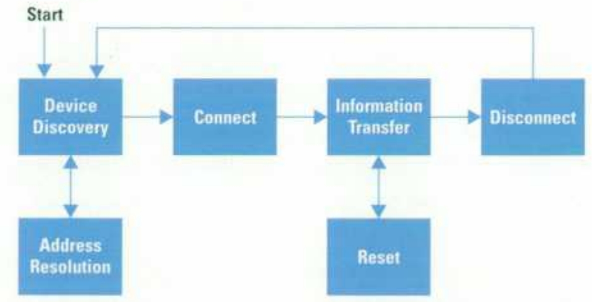
The IrLAP has three distinct phases of operation: link initialization, nonoperational mode, and operational mode. Nonoperational and operational modes are distinguished by the absence or presence of a connection with another device. During link initialization, the IrLAP layer chooses a random 32-bit device address. This address is randomly chosen to negate the need to select and maintain fixed device addresses for all IrDA devices. Although it is unlikely that two or more devices within range of each other will choose the same address, procedures are defined to detect and resolve address collisions. After the link is initialized, the IrLAP layer enters nonoperational mode.

The nonoperational mode is derived from HDLC's normal disconnect mode (NDM). In this mode, all devices contend for the medium. To do this, each device must check that the medium is not busy before transmission. This is achieved by listening for activity—that is, listening for physical layer transitions for at least 500 ms. Transmissions in the normal disconnect mode use link parameters that can be supported by all IrDA devices at a rate of 9600 bits/s. In this mode, the device will initiate device discovery, address resolution (if required), and connection establishment.

Once the connection has been established, the IrLAP layer moves into the operational or, in HDLC terms, normal response mode. This mode is an unbalanced mode of operation in which one device assumes the role of primary station and the other assumes a secondary role. This is the phase in which information is exchanged under

Figure 13

The IrLAP procedure flow.



control of the primary station. The link parameters are negotiated during the connection setup procedure and remain constant during the connection. During this phase, all other devices within range of either the primary or secondary stations remain idle in the normal disconnect mode. The two communicating devices therefore have unrestricted access to the medium for the duration of the connection. Once the information has been transferred, the link is disconnected and the device returns to the normal disconnect mode. The flow of procedures for the IrLAP layer is shown in **Figure 13**.

Device Discovery and Address Resolution. The discovery procedure is the process an IrDA device uses to determine whether or not there are any devices within communications range. In doing so, the device discovers the address of any device within range, the version number of the IrLAP protocol operating in each device, and some discovery information specified by the IrLMP layer in each device. The discovery procedure is controlled by the initiating device, which divides the discovery process into equal periods or time slots. The slotted nature of the discover procedure minimizes the likelihood of collisions when there are multiple devices within range.

After waiting for a period of 500 ms (normal disconnect mode rules), the initiating device starts the discovery procedure and broadcasts frames marking the beginning of each slot. On hearing the initial discovery slot (which also details the number of slots in the discovery process: 1, 6, 8 or 16), a device randomly selects one of the slots in which it will respond. When the device receives the frame marking its chosen slot, it transmits a discovery response frame. All frames in the discovery procedure use the HDLC unnumbered format of type XID (exchange identification).* An example of the discovery process is shown in **Figure 14**.

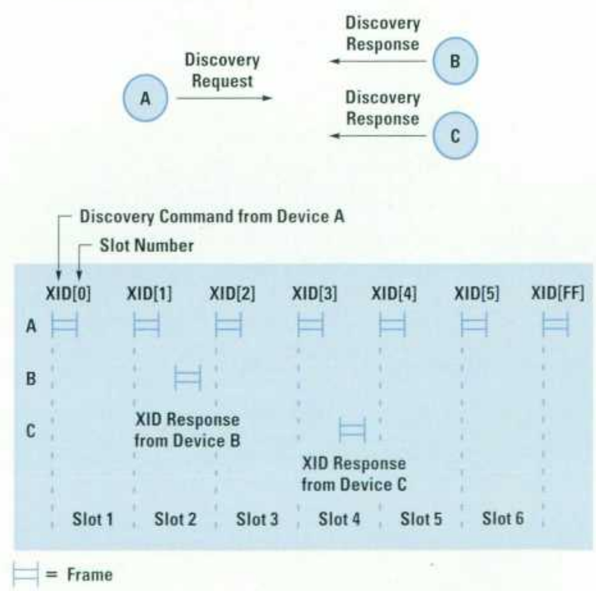
Figure 14 shows a three-device scenario in which device A is within range of devices B and C. Device A initiates the discovery process by transmitting a discovery XID command frame which, in this case, indicates that this is a six-slot discovery process and that this is the initial slot. Device A continues to transmit discovery command XID frames indicating the appropriate slot number. The final frame, after slot 6, is indicated by a slot number 0xFF. The final slot also contains information about the initiating device.

When the initial discovery XID command frame is received, devices B and C randomly choose slots in which to respond—in this example, slots 2 and 4. Device B then waits until it hears the discovery XID command indicating slot 2, and responds with a discovery XID response frame containing information about itself. Similarly, device C transmits a response during slot 4. Once the discovery process is over, all devices have the address and other

* In this context XID is a type of HDLC frame as specified in the ISO standard.

Figure 14

The discovery procedure.



information of all the devices within range: that is, device A has information about devices B and C, while devices B and C each have knowledge of device A. However, devices B and C are mutually hidden and as a result have no information about each other. This discovery information is passed to the upper layers whose responsibility it is to determine if there are any address collisions that need to be dealt with.

Should any of the devices that participated in the discovery process have duplicate addresses, then an address resolution process can be initiated. Address resolution follows a procedure similar to the discovery process, except that the device detecting the address conflict initiates the procedure, and resolution involves only the devices that have conflicting addresses. In this case, the initiating device transmits an address resolution XID command directed at the conflicting address. Devices with this address select another random address and a slot in which to respond. The initiator transmits the slot markers as before, and the previously conflicting devices respond in the appropriate slot. Once the process is over, each device should have a unique address. In the unlikely event that an address conflict still exists, the procedure can be repeated.

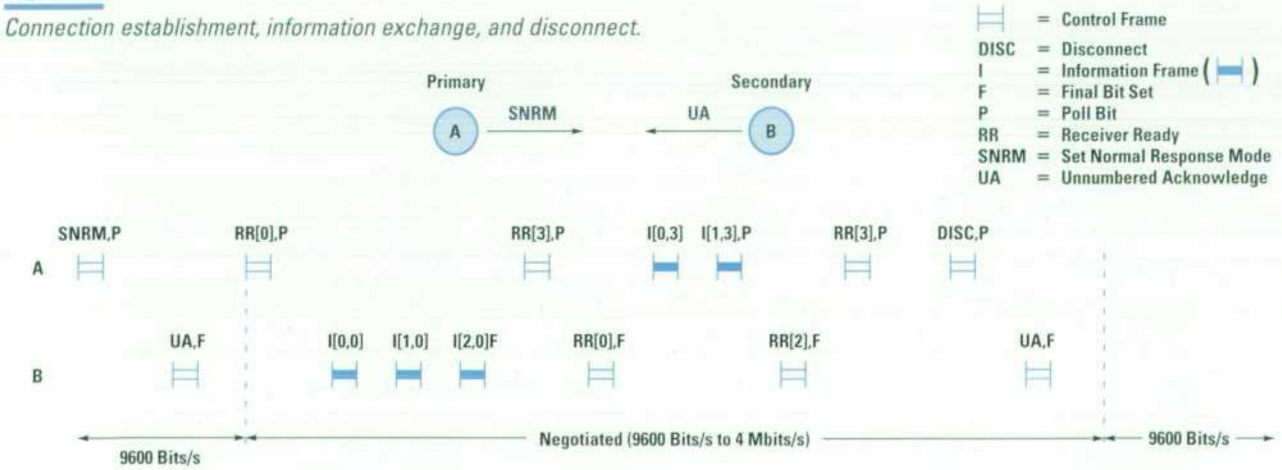
Connection Establishment. Once the discovery and address resolution processes are complete, the application layer may decide that it wishes to connect to one of the discovered devices. To connect, the application layer will issue a connection request which will ultimately result in the appropriate IrLAP service primitive being invoked.

The IrLAP layer connects to the remote device by transmitting a set normal response mode (SNRM) command frame with the poll bit set. This command informs the remote device that the source wishes to initiate the connection and the poll bit indicates that a response is required. Assuming the remote device can accept the connection, it responds with an unnumbered acknowledge (UA) response frame with the final bit set. This indicates that the connection has been accepted. Under normal circumstances, the device that initiates the connection (transmits the set normal response mode) will become the master, or primary, device, and the other device will become the slave, or secondary device. An example of connection establishment is shown in **Figure 15**.The notation used in the frames in **Figure 15** has the general form I(x,y) and RR(y), where x is the sequence number of the information frame and y is the sequence number of the next frame the source device expects to receive from the destination device.

The connection establishment takes place in normal disconnect mode (9600 bits/s), and once this is completed, the two devices will be in normal response mode. While in normal response mode, the devices can exchange data at any IrDA defined rate. However, not all IrDA devices will support all IrDA data rates or link parameters. It is therefore necessary for the devices to negotiate the parameters for normal response mode during connection setup. IrDA has defined several link parameters that can be negotiated:

Figure 15

Connection establishment, information exchange, and disconnect.



- Data rate
- Maximum turnaround time
- Data size
- Window size
- Number of additional start of frame symbols (BOFs)
- Minimum turnaround time
- Link disconnect threshold time.

Data rate defines the data transfer rate during normal response mode (9600 bits/s to 4 Mbits/s), while maximum turnaround time defines the length of time either device may transmit before giving the other device a chance to transmit (50, 100, 250, or 500 ms). Data size determines the maximum length of the data field in an information frame (64 to 2048 bytes), and, in combination with the retransmission window size, which defines the number of outstanding frames that may be unacknowledged, allows devices with only limited resources to restrict the rate at which they will receive data. Number of additional BOFs and minimum turnaround time relate to physical layer restrictions, while link disconnect threshold time determines how long a device will wait without receiving a response from another device before assuming the link has failed and informing the upper layer that the link has disconnected.

Well-defined rules exist that ensure that after the set normal response mode-UA exchange has been completed, both devices will know the negotiated normal response mode parameters. Once both devices are in normal response mode, the primary device polls the secondary device by transmitting a receiver ready (RR) frame with the poll bit set, thereby initiating the information exchange phase.

Information Exchange and Link Reset. The information exchange procedure operates in a master-slave mode in which the primary device controls the secondary device's access to the medium. The primary device issues command frames to the secondary device which responds with response frames. To ensure that only one device can transmit frames at any one time, a permission-to-transmit token is exchanged between the primary and secondary devices. The primary device passes the permission-to-transmit token to the secondary by sending a command frame with the poll bit set. The secondary device returns the token by transmitting a response frame with the final bit set. The secondary device can only retain the token while it is transmitting data, and it must return it to the primary device if it has no data to transmit or if it reaches the maximum turnaround time. The primary device, however, within the limits imposed by the maximum turnaround time, can hold the token even if it has no data to transmit.

Although the physical layer has been designed to provide a low bit error rate channel, the dynamic nature of the infrared connection results in a possibility that frames may be lost in transit because of corruption by noise. To cope with this, the IrLAP protocol uses a sequenced information exchange scheme with acknowledgments. Should a frame be corrupted by noise, the CRC will highlight this error and the frame will be discarded. At the IrLAP layer, this error will be detected by virtue of the noncontiguous sequence numbers on the information frames. The IrLAP protocol implements an automatic repeat request strategy in the same manner as HDLC with options of using stop and wait, go back to N, and selective reject retransmission schemes.[13] This strategy allows the IrLAP layer to provide an error-free, reliable link to the IrLMP layer. An example of an error-free information exchange between two devices is shown in **Figure 15**.

Under exceptional circumstances, however, it may not be possible for the IrLAP entities in each device to recover from an error condition while maintaining the sequenced delivery of error-free information (I) frames. In this case, the IrLAP entity is allowed to reset the link. This reset involves discarding any undelivered information and reinitializing the sequence numbers and timers for the link. Although this may result in the loss of data, which the higher-level layer must deal with, it does allow the link to recover without the need for a total disconnection.

Connection Termination. Once the data exchange has taken place, the IrLAP link may be disconnected by either the primary or secondary devices. Should the primary wish to disconnect, it sends a disconnect command to the secondary device with the poll bit set. The secondary responds by returning an unnumbered acknowledge frame with the final bit set. Both devices will now be in normal disconnect mode, and the default normal disconnect mode parameters (9600 bits/s data rate) will apply. If the secondary wishes to disconnect, it transmits a request disconnect response with the final bit set when it is polled by the primary. The primary will then respond by transmitting a disconnect command, and both devices will be in normal disconnect mode. An example of a primary-initiated disconnection is shown in **Figure 15**. Once the two devices are in normal disconnect mode, the medium is free for any other device to initiate the discovery, address resolution, or connection procedures.

The Infrared Link Management Protocol

The Link Management Protocol (IrLMP) is layered on top of IrLAP, and has two main functions: application and service discovery and multiplexing of application level connections over the single IrLAP connections. The IrLMP layer allows individual service users (applications) to connect and exchange information with similar entities in the peer device, independent of any other service users that may be using the IrLAP connection. The IrLMP layer provides multiple independent channels to the IrLMP layer in the remote device. The IrLMP layer also provides a service with which applications can locally register themselves and some significant parameters in an information base. Services are also provided that enable those applications to access equivalent information in the information base of remote devices. Using this service, an application does not need prior knowledge of the applications in a remote device. This is an extremely useful feature for the kind of ad-hoc interactions typical of IrDA devices.

The two main functions provided by IrLMP are split between two sublayers. The Link Management Multiplexer (LM-MUX) provides the facilities for multiplexing application level connections over a an IrLAP connection between a pair of devices. The Link Management Information Access Service (LM-IAS) provides the services necessary to allow applications to discover devices and access the information in the information base of a remote device.

The Link Management Multiplexer. The LM-MUX adds two bytes of overhead to the IrLAP information frame, which are primarily used for addressing the individual multiplexed connections. The address fields uniquely identify the link service access points (LSAPs) in both the source and destination devices. Each LSAP is addressed by a seven-bit selector (LSAP-SEL), and LSAP-SELs within the range 0x01 to 0x6F can be used by applications. LSAP-SELs 0x00 and 0x70 are reserved for the information access service server and the connectionless data service respectively. The remaining LSAP-SEL values, 0x71 to 0x7F, are reserved for future use. Connections between IrLMP service users are called LSAP connections, and although an LSAP may terminate other LSAP connections, there is only one LSAP connection between any pair of LSAPs. All LSAP connections use the single IrLAP connection between the pair of devices.

Information Access Service. The information access service maintains information about the services provided by the host device and provides services that allow access to the information base on remote devices. The information access service allows devices to discover which services are available on the host device and provides the configuration information necessary to access those services. As an example, the most common piece of information required is the LSAP-SEL value, which tells where a particular service is located.

The information stored in the information base consists of a number of objects. Each object belongs to a specific class, and there may be several objects of the same class in the information base. The class defines the attributes that are present in each object, and these attributes can be assigned a particular value. The attributes of a class can be of type user string, octet sequence, signed integer, or missing. **Figure 16** shows an example of the information access service database for a device offering three unique services.

The example shows a device with three individual applications: e-mail, calendar, and IrOBEX (file transfer application). The information base contains three objects associated with these applications. The required Object 0 is always present within the information access service database, and it provides information about the device name and the version of IrLMP the device supports. All other devices can address Object 0 to get this information. Objects in the information base typically detail information about the services provided—for example, the LSAP-SEL where these services can be accessed. In the case of the calendar application, this service can be accessed using the Tiny TP flow-control mechanism on LSAP-SEL 3, or directly on LSAP-SEL 2. The difference is encoded in the attribute name.

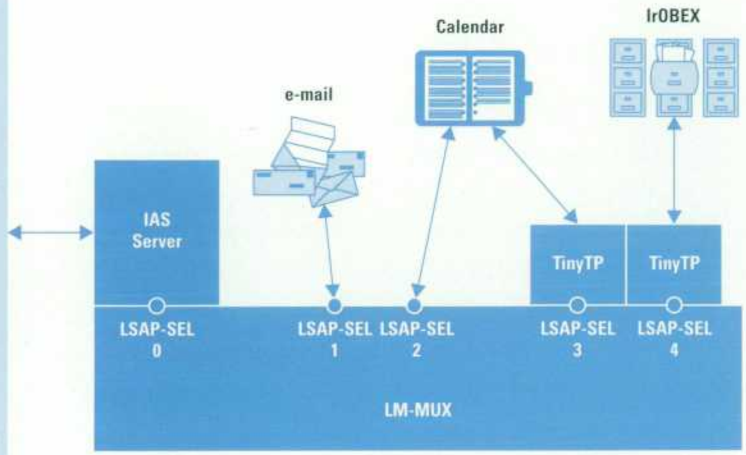
The IrLMP layer provides several service primitives to access information access service data. However, the only mandatory service is GetValueByClass. This service requires the service user to provide the class and attribute names of the service it is interested in. The service returns a list of

Figure 16

Example information access service database.

Information Base

```
Object 0:Class Device {
  DeviceName="My Device"
  IrMPSupport=0x01
}
Object 1:Class Email {
  IrDA:IrLMP:LSapSel=0x01
  IrDA:IrLMP:InstanceName="Mail"
}
Object 5:Class Calendar {
  IrDA:IrMP:LSapSel=0x02
  IrDA:TinyTP:LSapSel=0x03
  IrDA:IrLMP:InstanceName="CalTool"
}
Object 8:Class IrOBEX {
  IrDA:TinyTP:LSapSel=0x04
  IrDA:IrLMP:InstanceName="Launcher"
}
```



object identifiers and attribute values for all objects in the information base with the requested class and attribute name. Referring to the example in **Figure 16**, if a peer device issued a GetValueByClass with parameter Calendar for the class name and IrDA:IrLMP:LSapSel for the attribute name, the service would return a single element list with the entry containing object identifier 5 and attribute value 2.

Tiny TP Flow Control Mechanism

Although the IrLAP layer does have provisions for flow control, its use can result in deadlock situations, particularly where more than one IrLMP connection is operating. Such a deadlock situation can occur if an application in one device is waiting for its peer application to send it some data before releasing its buffer space. However, another connection may use up the remaining buffer space, causing the IrLAP layer to flow-control the link until buffer space becomes available. If both connections are waiting for data from the remote device before freeing the buffers, then clearly a deadlock has occurred that cannot be resolved without some form of higher-level intervention such as a system reset.

To overcome this problem, IrDA provides the lightweight transport protocol called Tiny TP.[8] Tiny TP adds a single byte of overhead to each frame and provides a per-LSAP-connection credit-based flow control mechanism with the possible segmentation and reassembly of service data units of up to 4 Gbytes in size. When a Tiny TP connection is initiated, the maximum service data unit size is negotiated and some initial credit is extended to each connection endpoint. Sending data causes the credit to be decreased by one, and periodically the receiver issues more credit. Without credit, the transmitter cannot send any data. It must wait until such times as the receiver extends it some more credit. Using Tiny TP, a device can ensure that credit is distributed among its applications, ensuring that the applications can communicate without reducing the buffer space to such a degree that IrLAP flow control must be used.

Conclusion

IrDA has completed the core standards necessary to enable any mobile computing platform with ad-hoc, point-and-shoot infrared communications from 2400 bits/s to 4 Mbits/s. Support for the IrDA platform from a wide variety of manufacturers is now becoming apparent, as many products—ranging from printers to laptop PCs and PDAs to mobile phones—are being released with IrDA capability. All these devices will have the ability to interoperate with one another should that be required. With over 130 companies actively maintaining membership in IrDA, currently released IrDA-enabled products represent only the tip of the iceberg. In the coming months and years, it is expected that more and more computing and other devices will be released with built-in IrDA capability.

However, providing the hardware platform to support IrDA is only half the story. Current activity within IrDA is directed at finishing off the IrDA series of standards to enable application-level developers to access the IrDA features in a uniform and efficient manner. The needs of legacy serial/parallel applications have been addressed with the IrCOMM standard. Legacy networking applications will be able to use the IrDA features implemented in the forthcoming IrLAN protocol. However, it is expected that a new class of applications will be developed with the express purpose of using the unique features of IrDA-enabled devices. The IrOBEX protocol, when completed, will provide application programmers with a generic method by which data can be exchanged with other applications without having to know the details of the destination application. As an example, transferring a graphic to another PDA (which will display it) or to a printer (which will print it) will be no different from the source application's point of view. Alternately, a more flexible approach to accessing the IrDA communications facilities will be to directly access them through the operating system's application programming interface. An example of this is the WinSock-style API to IrDA, called IrSock,[15] currently being developed for the Microsoft® Windows 95 operating system.

In conclusion, the future for infrared is bright. With cross-industry support, IrDA is fast becoming the ubiquitous infrared communications system for portable and peripheral devices. Although legacy support for other infrared systems will persist for some time to come, the IrDA standard is now used on so many platforms that it is unlikely any new systems will be anything other than IrDA-enabled.

References

1. *IrDA Marketing Requirements—Basis for the IrDA Technical Standards*, Version 3.2, The Infrared Data Association, November 23, 1993.

2. *Serial Infrared (SIR) Physical Layer Link Specification*, Version 1.0, The Infrared Data Association, April 27, 1994.

3. *Serial Infrared Link Access Protocol (IrLAP)*, Version 1.0, The Infrared Data Association, June 23, 1994.

4. *Link Management Protocol (IrLMP)*, Version 1.0, The Infrared Data Association, August 12, 1994.

5. *Serial Infrared Physical Layer Link Specification*, Version 1.1, The Infrared Data Association, October 17, 1995.

6. *Link Management Protocol (IrLMP)*, Version 1.1, The Infrared Data Association, January 23, 1996.

7. *Serial Infrared Link Access Protocol (IrLAP)*, Version 1.1, The Infrared Data Association, June 16, 1996.

8. *Tiny TP: A Flow-Control Mechanism for use with IrLMP*, Version 1.0, The Infrared Data Association, October 25, 1995.

9. *IrCOMM: Serial and Parallel Port Emulation over IR (Wire Replacement)*, Version 1.0, The Infrared Data Association, November 7, 1995.

10. *LAN Access Extensions for Link Management Protocol (IrLAN)*, Proposal, Version 1.0, The Infrared Data Association, January 1, 1996.

11. *Object Exchange Protocol (IrOBEX)*, Draft, Version 0.5f, The Infrared Data Association, July 24, 1996.

12. S. L. Harper and R. S. Worsley, "The HP 48SX Calculator Input/Output System, *Hewlett-Packard Journal*, Vol. 42, no. 3, June 1991.

13. *Information technology—Telecommunications and information exchange between systems—High-level data link control (HDLC) procedures—Elements of procedures*, ISO/IEC 4335, International Organization for Standardization, December 15, 1993.

14. *Information technology—Telecommunications and information exchange between systems—High-level data link control (HDLC) procedures—Classes of procedures*, ISO/IEC 7809, International Organization for Standardization, December 15, 1993.

15. *Infrared Sockets (IrSockets) Specification: Functional Specification*, Revision 1.0, The Microsoft Corporation, July 10, 1996.

Iain Millar

As a member of the technical staff at HP Laboratories in Bristol, England, Iain Millar is involved in the development of protocols for the next generation of IrDA systems. He attended the University of Aberdeen in Scotland where he received a BSc (Eng.) degree (1988) in electrical and electronic engineering and a PhD (1995) in the area of fault tolerant protocols for LANs.

Martin Beale

Martin Beale is a member of the technical staff at Hewlett-Packard Laboratories in Bristol, England. He is working on the physical layer for the next generation IrDA systems. He earned a PhD degree (1994) in reduced complexity decoding of convolutional codes from the University of Cambridge. Outside of work he enjoys rock climbing, walking, cycling, skiing.

Bryan J. Donoghue

Bryan Donoghue is a member of the technical staff at HP Laboratories in Palo Alto where he is working on the digital system design for high-speed wireless radio LANs. He received a MEng degree (1991) in electrical and electronic engineering from Loughborough University in England. Bryan was born in Llanelli, Wales and outside work he enjoys traveling, learning foreign languages, backpacking, and cycling.

Kirk W. Lindstrom

A member of the technical staff at HP's Communication Semiconductor Solutions Division, Kirk Lindstrom is responsible for the design of ICs for infrared products. He joined HP in 1978 and since that time he has worked on the design of optocouplers, fiber-optic modules, and infrared transceivers. He holds a BS degree in EECS (1979) from the University of California at Berkeley. Besides being an ardent windsurfer, he also has an interest in doing and writing about investing.

Stuart Williams

A project manager at HP Laboratories in Bristol, England, Stuart Williams is responsible for the infrared communications group. He has worked on various infrared protocol-related projects since he joined HP in 1992. He has a PhD degree (1986) from the University of Bath. Stuart was born in Rugby, Warwickshire, England, is married and has two sons. Biking and sailing occupy his free time.

RF Technology Trade-offs for Wireless Data Applications

Kevin J. Negus

Bryan T. Ingram

John D. Waters

William J. McFarland

Rapidly evolving wireless system standards and applications are placing demands on RF semiconductor manufacturers to produce highly specific and optimized RFIC solutions for specific growth segments including wireless data terminals.

Current RF wireless connectivity standards used for LANs and WANs include cellular and PCS (Personal Communications Services) protocols such as GSM (Global System for Mobile Communications) and AMPS (Advanced Mobile Phone System), trunk radio systems such as RAM and Ardis (proprietary systems), and ISM (industrial-scientific-medical) systems. In the future, satellite-based standards and dedicated wireless data systems such as HIPERLAN (European 5-GHz LAN) and U-NII (Unlicensed National Information Infrastructure) will be more commonplace. From a wireless data user standpoint, the radios based upon these standards can be implemented in a variety of physical form factors including removable PC cards that contain an entire radio, radios that are built into a dedicated data collection terminal, or cellular and PCS phones that connect to a laptop modem via a cable.

The applications listed above are made possible through high-performance RF semiconductor components. These applications range in frequency from several hundred megahertz to 6 GHz and above and require several watts of power output in some cases. These requirements need to be satisfied along with aggressive cost goals, achieved by means of low-cost, surface mount plastic packaging. The RF semiconductor technologies on which these components are based currently include gallium arsenide (GaAs) and silicon bipolar and will increasingly include CMOS in the future. Some specific examples of end-product benefits that RF component technologies can affect on a first-order basis include the cost, size, weight, and battery life of the wireless data terminal.

Lower cost is now being achieved in end-user applications as a result of the tremendous growth of the RF semiconductor industry and the corresponding economies of scale that are created. Innovative RF architectures and approaches to high-level integration can also lead to much lower cost.

Smaller wireless data terminals are made possible by highly integrated components that allow multiple functions and previously external support elements to be incorporated onto a single RFIC. This approach is usually only justified for high-volume standards and applications and the resulting IC does not have the flexibility to be used for a wide variety of standards. For the digital cellular standard GSM (Global System for Mobile Communications), a highly integrated approach is justified because of the large production volumes and relative homogeneity of the technical approaches used by various customers. However, the relatively small wireless data market currently consists of a fragmented group of existing and emerging standards. Therefore, a highly integrated RFIC approach is risky to the component manufacturer and limits the number of these RFICs that are being developed for the market. An alternative way of accomplishing the small size objective is to use very small, flexible, high-performance building-block RFICs that contain a few key functions per product. These RFICs can be used in a variety of systems and with a variety of customers. This wide range of use allows these RFICs to be produced at a very low cost.

Substantially longer battery operation has been achieved in RF communications equipment partly by RF components that have lower supply current requirements (higher efficiency). Within the radio system, the power amplifier usually requires a large percentage of the supply current budget relative to the rest of the RF components. As a consequence, many of the component developments geared for improved efficiency are targeted toward the power amplifier.

Lower-weight systems have emerged because RF component manufacturers have been able to reduce their supply voltages, thus enabling operation from fewer battery cells. The battery is the heaviest component in many battery-powered communication devices. Despite its high cost, lithium-ion is a common battery technology that is in favor for its improved battery life, reduced memory effects, and longer storage times.

Target Applications and Available Technologies

RF system specifications have a great impact upon the ultimate cost, size, weight, and battery life of any radio implementation. System-level trade-offs that include data rate, bit error rate (BER), system capacity, and range ultimately drive the radio specifications. These specifications directly affect the choice of RF components and technologies used in the radio. The major RF component drivers include operating frequency, RF output power, and modulation technique. Key system specifications of some examples of wireless data systems are summarized in **Table I**.

Table I

Examples of Wireless Data Systems

System	Frequency Band (MHz)	Data Rate (kbits/s)	Channel Spacing (MHz)	Modulation	Transmit Power (W)	Range (m)	Comments
CDPD	824-849 Tx 869-894 Rx	9.6	0.03	GMSK	0.6	10000	Uses empty analog cellular channel
U-PCS	1910-1930	1152	1.0	$\pi/4$ DQPSK	0.1	200	Unlicensed part of U.S. PCS bands
IEEE 802.11	2400-2483	1000 2000	1.0	2-GFSK 4-GFSK	0.1, 1	100	Subject to significant interference
HIPERLAN	5150-5300	23500	30	GMSK	1	50	

CDPD = Cellular Digital Packet Data.
GMSK = Gaussian minimum shift keying.
2,4-GFSK = 2-level or 4-level Gaussian frequency shift keying.
IEEE 802.11 = 2.4-GHz wireless LAN standard.
$\pi/4$ DQPSK = $\pi/4$ differential quadrature phase shift keying.

The operating frequency is a key specification because it affects many RF component decisions including whether GaAs or lower-cost silicon bipolar components can be used. Other considerations include whether the convenience of easy-to-use RFICs can be exploited or if discrete components are the best way to achieve the required noise figure and power efficiency. Development time and available RF expertise are also factors in choosing between RFICs and discrete components.

The modulation technique employed in a radio is a critical specification because it drives the type of modulator used in the system. A simple FSK (frequency-shift keying) technique can be implemented with minimal cost by direct modulation of the phase-locked loop. FSK-based systems do not require a linear transmit path, so lower-cost and lower-current components can be used. Techniques based upon QPSK (quadrature phase-shift keying) modulation require an I-Q (in-phase and quadrature) modulator in the transmit path, which adds cost. In addition, linear power amplifiers, which consume larger currents, are required for modulation techniques that have envelope fluctuations, such as QPSK. The benefit of QPSK systems is greater spectral efficiency or user density, but this is often sacrificed in unlicensed wireless data applications where no revenue stream is available to subsidize the terminal cost.

The power output levels of wireless data terminals typically range from milliwatts to watts. On the high end of the range, a GSM data terminal (two watts out of the antenna) might use a hybrid power module. A wireless LAN output stage with power output of 10 mW can use a simple bipolar transistor or RFIC that may cost only one twentieth as much as the GSM solution. The required power levels, along with the modulation technique, help determine the semiconductor technology that will be used. For a 2.4-GHz wireless LAN system that requires 1W output power and uses QPSK modulation, a GaAs-based output amplifier is essential for the higher gain and linearity provided by this technology. On the other extreme, a 0.5-mW average remote control device operating at 900 MHz with an on/off modulation technique may favor the use of a low-cost silicon bipolar transistor or RFIC.

Table II summarizes the key HP RF semiconductor technologies currently used to develop high-volume wireless RFICs, as well as potential future candidate technologies. GaAs RFICs based on PHEMTs (pseudomorphic high-electron-mobility transistors) are typically restricted to fairly low-levels of integration in front-end components such as low-noise amplifiers, transmit/receive switches, or power amplifiers. PHEMTs can also be used to make excellent upconverters and downconverters, but the cost penalty relative to silicon bipolar often overcomes any slight performance advantage. HP's ISOSAT silicon bipolar technology features good high-speed n-p-n transistors combined with excellent passive elements including high-Q inductors, precision high-Q capacitors, integrated varactor diodes, and integrated 100-GHz Schottky diodes. This makes ISOSAT ideal for most frequency converter applications including I-Q vector modulators and high-dynamic-range downconverters. ISOSAT passive elements can also be used to improve narrowband amplifier performance and even build fully monolithic voltage-controlled oscillators (VCOs).

Table II
RF Device Technologies

Process Technology	f_T (GHz)	f_{max} (GHz)	$Q_{inductor}$ (4 nH at 2 GHz)	Manufacturable Die Size (mm^2)	Cost
ISOSAT	14	30	10	2-3	Medium
HP-25	25	30	5	10-20	Low
HP PHEMT	60-90	N/A	20	< 1	High
Next-Generation RF Bipolar	30	> 50	15	10-20	Low

The HP-25 process listed in **Table II** is an example of a state-of-the-art, high-speed, highly manufacturable silicon bipolar process that can provide high levels of integration even for RF applications. HP-25 is based on a mainstream CMOS process and is built in the same high-volume fabrication facility with the same tools as CMOS. Next-generation RF bipolar processes are expected to have even higher-performance n-p-n transistors while using passive elements improved upon from ISOSAT. By continuing the HP-25 philosophy of using CMOS-based unit processes, these next-generation RF bipolar processes should be highly manufacturable and low in cost. However, CMOS technologies are also improving dramatically over time, and next-generation CMOS will be capable of implementing many RF functions, especially those used in frequency synthesis and IF demodulation.

Example 1: Low-Cost 900-MHz Remote Control Data Device

Consider an extreme example of a very simple wireless data device implemented at the lowest possible cost. This type of one-way device could be used for keyless entry (with a data-encrypted security code), meter reading, or remote control of digital equipment. Key attributes are low cost and current consumption at the expense of short range, very low data rate, and somewhat unreliable operation.

A possible circuit schematic is given in **Figure 1**. Conceptually, the transmitter consists of a crystal oscillator, frequency tripler, frequency doubler, and output amplifier to provide an on/off keyed data stream in the 902-to-928-MHz ISM band under U.S FCC Part 15.249 low-power rules. The RF section would have the crystal oscillator and tripler turned on several milliseconds before transmission. The on/off keying could be as crude as simply enabling V_{CC2} by a depletion-mode FET controlled by a CMOS bit from a microcontroller. This can work quite well for low data rates of 10 kbits/s or less (100-μs-duration pulses). Under Part 15.249 rules, the average transmit power in any 100-ms period must be less than 0.5 mW. This would allow, for example, 5-mW transmit power at the antenna in **Figure 1** so long as the bursts are only 10 ms long (or 100 bits at 10 kbits/s) and spaced 100 ms apart.

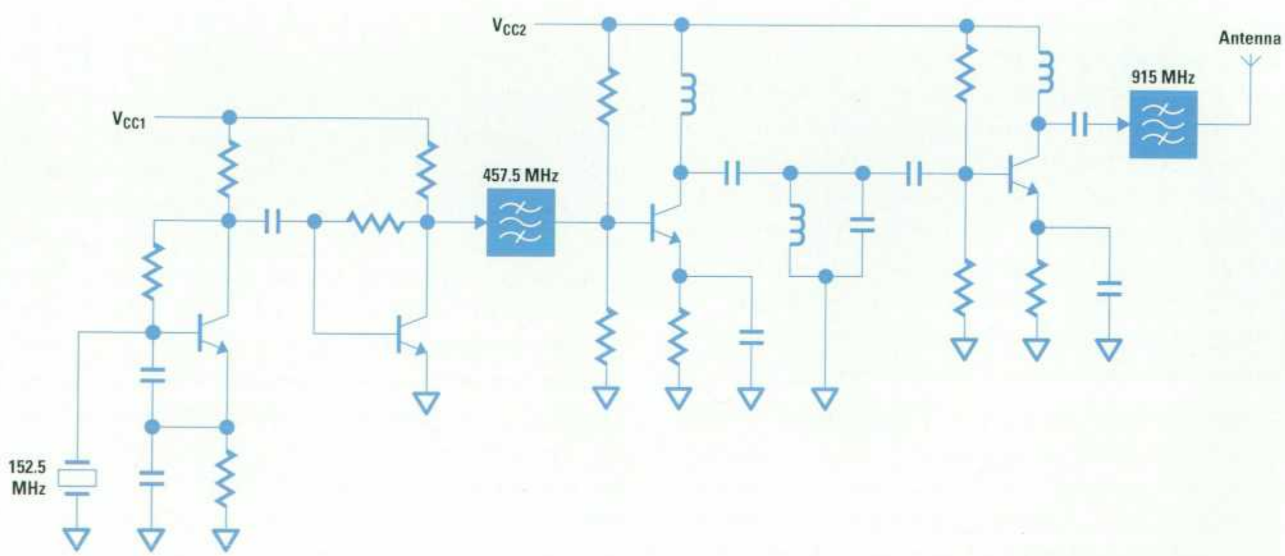
As illustrated in **Figure 1**, the RF semiconductor portion of the transmitter is only four discrete bipolar transistors, which can be identical, ultralow-cost AT41533s. The 457-MHz bandpass filter can be realized easily by printed coupled lines, so the 152.5-MHz crystal and the 915-MHz bandpass filter represent the only significant passive elements in terms of cost. In the U.S.A., in sufficiently high volume, the RF transistors are much less than one dollar for all four and the total cost of this solution is only two to three U.S. dollars, making it competitive with an infrared solution.

Example 2: Highly Integrated 2.4-GHz Wireless LAN

In the 2.4-GHz wireless LAN market, meeting the small size requirements of the popular PC-card form factor demands a high level of integration, both in the RF and digital circuits of the device. The digital part of the device is now routinely accommodated in only a few components, typically an application-specific controller IC and associated ROM and RAM, and as such can easily be accommodated within the PC-card form factor. However,

Figure 1

Low-cost transmitter.



the RF sections are not so easily integrated, and therefore careful consideration must be given to the radio architecture to ensure an optimum trade-off of size, power consumption, and component cost. One such architecture for a frequency-hopping wireless LAN, with the majority of functionality resident in only a few ICs, is illustrated in **Figure 2**.

The radio consists of a double-conversion superheterodyne receiver and a single-conversion transmitter. This arrangement is helpful in achieving the 100-MHz operating width of the 2.4-GHz ISM band while providing adequate suppression of transmitted spurious and received image responses. Some care must be exercised in choosing the first intermediate frequency (IF) to ensure that it is not affected by spurious products generated elsewhere (e.g., harmonics of the crystal reference) and that its own harmonics will not be a problem. In this example, 236 MHz is used for both the transmitter and the receiver since this allows small, low-cost, three-pole ceramic filters to be used in the transmit chain, meets the limits for spurious emissions dictated by the regulatory bodies, and is not affected by the 16-MHz reference.

Translation between the first IF and 2.4 GHz is performed by the combination of the HPMX-5001 upconverter/downconverter IC and the HPLL-8001 high-speed CMOS synthesizer. The synthesizer, operated in a dual-modulus configuration with the HPMX-5001's 32/33 dual-modulus prescaler is used to define the channel of operation. Driven from a 16-MHz reference—an integer multiple of the system channel spacing—the synthesizer can hop channels easily within the 224 μs required by the IEEE 802.11 standard. Transmit/receive turnaround time is less than 19 μs, the fast turnaround dictated by the RTS/CTS protocol normally used in a wireless LAN system. This is achieved by continuously running the modulator synthesizer at twice the second IF so as not to interfere during receive periods, and then connecting via a divide-by-2 circuit to the HPMX-5001 only during transmit periods.

The HPMX-3003 provides the major front-end functions of low-noise amplifier, switch, and power amplifier. The transmitted signal from the HPMX-5001 followed by the AT31011 driver is amplified to a final output power level at the antenna in excess of 200 mW by the HPMX-3003. Note that 100% duty cycle operation is possible with this device, which is an important consideration in asynchronous wireless LAN applications, in which the transmit duty cycle is determined by traffic loading to a large extent and is not usually limited to any significant degree. Transmit/receive antenna switching is provided by the HPMX-3003's low-loss switch. In the receive path the HPMX-3003's high gain and low noise figure ensure a good overall sensitivity for the receiver.

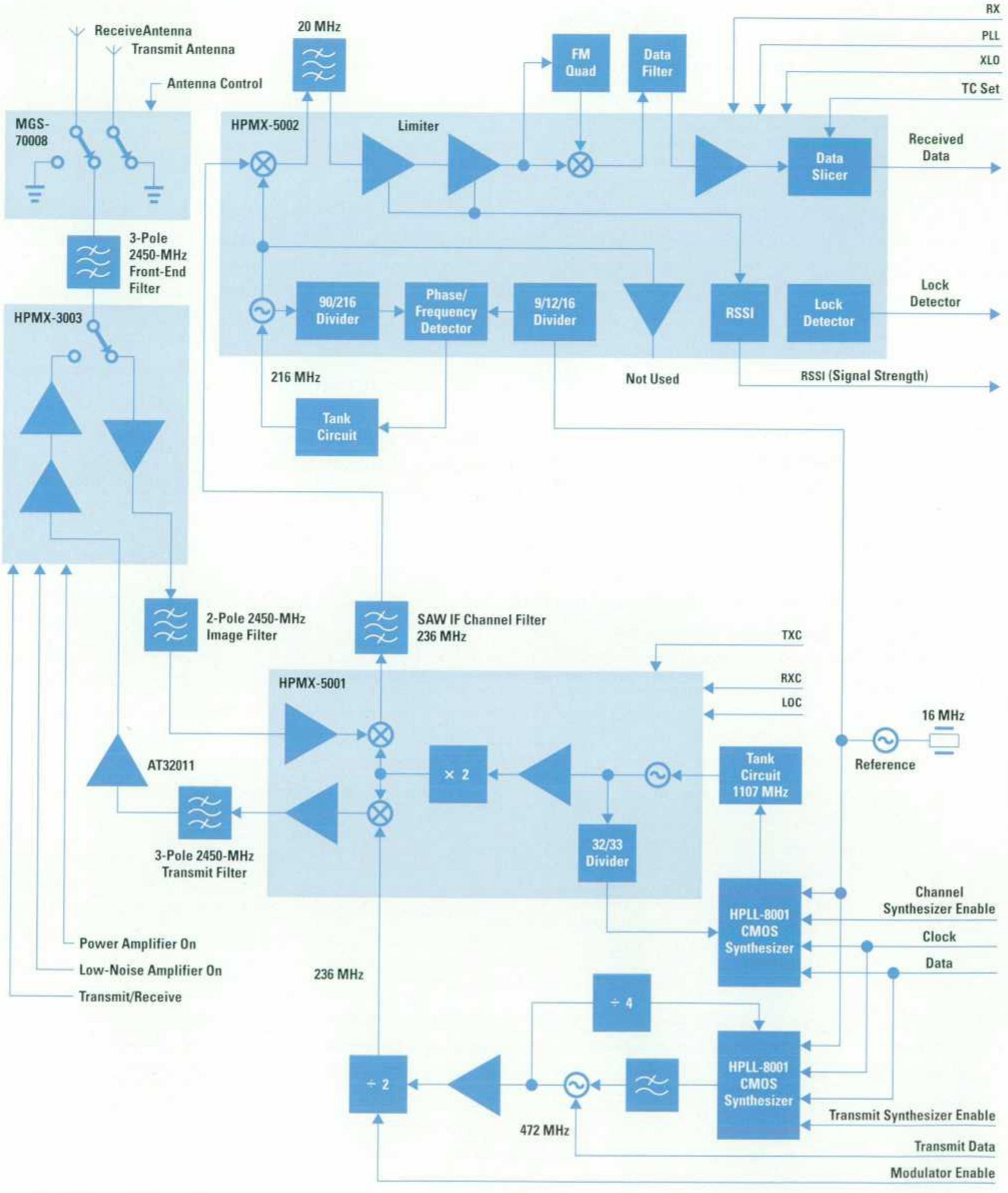
To improve the overall system performance and data throughput it is often desirable to include some form of diversity in the transmission channel. Most popular in low-cost systems is receiver antenna diversity, that is, a dual-antenna receiver, which can offer up to 10-dB improvement in system error performance. In **Figure 2** we illustrate this with two antennas spaced nominally $\lambda/2$ apart and selected by an MGS-70008 GaAs MMIC switch.

The HPMX-5002 provides all the necessary functionality for the demodulation of a downconverted 2-FSK signal. The HPMX-5002 contains the receiver's second mixer, limiting IF amplifier chain, discriminator, data slicer, lock detector, and RSSI (receive signal strength indicator) circuits. Also included are the necessary active components and dividers to generate the receiver's second LO. In **Figure 2** we use a second LO frequency of 216 MHz derived from the 16-MHz reference to create a 20-MHz second IF. The second IF is chosen low enough to ensure that the quadrature discriminator can be designed with low-tolerance components to minimize cost and avoid production adjustments while maintaining a low fractional bandwidth. Output from the discriminator's Gilbert-cell mixer is fed via an external data filter to the data slicer. The sliced receive data is then passed to the radio controller where it is processed and decoded according to the MAC (media access control) protocol being used (e.g., IEEE 802.11).

The complete RF section of this highly integrated wireless LAN consisting of RFICs, filters, and associated passive components is accommodated in less than 15 cm^2 of printed circuit board surface area, making it ideally suited to PC-card applications. The use of low-voltage operation in the devices combined with their power management facilities creates a low-power design particularly suited to portable applications. Interfacing to the radio controller is straightforward, with the majority of signals interfacing directly at normal CMOS levels. Only a minimum of interface circuitry is required to implement the full level of control required for the current generation of wireless

Figure 2

Frequency-hopping 2.4-GHz wireless LAN.



LAN systems. Many of the key performance parameters are summarized in **Table III**.

Table III

Summary of 2.4-GHz Wireless LAN RF Section Performance

Supply Voltage	3V
Receive Mode Current	100 mA
Transmit Mode Current	500 mA
Frequency Band	2.4 to 2.48 GHz
Transmit Power	200 mW
Data Rate	1 Mbit/s
Sensitivity	– 80 dBm
Range (Loss of Signal)	> 100 m
Board Area	15 cm^2

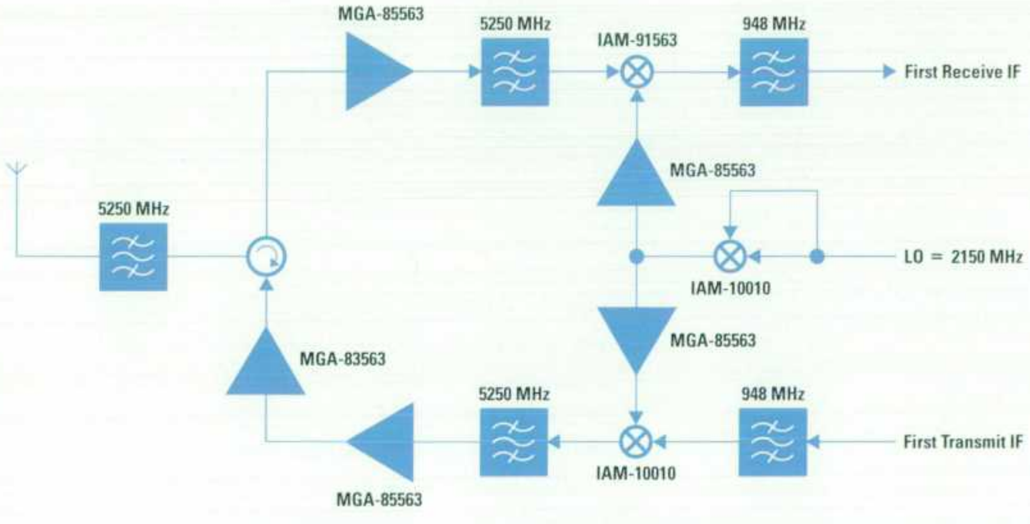
Example 3: Rapid Development of a 5-GHz Wireless LAN

Recently, considerable interest has arisen in building wireless LANs at 5 GHz to obtain much higher data rates (of the order of 10 Mbits/s) in unlicensed operation without the burdensome rules and problematic interference of the 2.4-GHz ISM band. An example of this is HIPERLAN in Europe. In the U.S.A., the FCC is considering allocation of the band from 5.15 to 5.35 GHz for similar types of high-data-rate wireless LANs. However, wireless terminal manufacturers wanting to build real products in this band will not have highly integrated chipsets as shown in Example 2 for the near future because of the immaturity of the applications and the standard, as well as the difficulty of RF integration at 5 GHz. While discrete RF devices using circuit techniques similar to Example 1 are theoretically possible at 5 GHz, the development time and engineering resources required would be huge for a complex system such as a high-data-rate wireless LAN. The best alternative by far is to use RFIC building blocks. HP manufactures many different building-block RFICs using both ISOSAT (silicon bipolar) and PHEMT (gallium arsenide) technologies. In general, these products are available in tiny surface mount plastic packages such as SOT-363, MSOP-10 and SSOP-16. The building blocks are usually one-function or two-function devices that have a wide range of uses and can be easily matched without stability problems.

An example of a 5-GHz wireless LAN RF front end running completely from a single supply voltage of 3.0V (a single lithium-ion cell) is shown in **Figure 3**. This block diagram can be implemented completely using low-cost, single-function RF building blocks as denoted by the HP part numbers in **Figure 3**. The choice of 948 MHz as a first IF is not accidental. Most proposals for this 5.15-to-5.35-GHz band have wide channels ranging from 10 to 25 MHz. At 948 MHz, two very high-volume off-the-shelf SAW (surface acoustic wave) filters that could be used as channel filters for this application are the 935-to-960-MHz GSM receive filter or the 940-to-956-MHz PDC transmit filter. There are

Figure 3

Building-block up/downconverter for a 5-GHz wireless LAN.



a number of highly integrated GSM receiver RFICs available, and these could be leveraged to minimize design effort. Also, a variety of transmit solutions are possible, including, for example, the HPMX-2007 vector modulator and upconverter.

In **Figure 3**, the nominally 4300-MHz first LO is generated by frequency doubling a 2150-MHz VCO. The doubler is an ISOSAT silicon bipolar RF building block, the IAM-10010, which can also be used as an upconverter at 5 GHz in the transmit chain. The channel synthesizer for the nominally 2150-MHz LO can be implemented in many ways, including, for example, a low-cost prescaler and the HPLL-8001. The output power amplifier in this example is the MGA-83563, which is a fully monolithic, two-stage PHEMT power RFIC usable to 6 GHz with 200-mW output power matched into 50 ohms. On the receive side, the MGA-85563 (another PHEMT RFIC) is used as the low-noise amplifier for its <2-dB noise figure and excellent matching even at these high frequencies. The IAM-91563 downconverter is also a PHEMT RFIC and is ideal for this application because of its 10-dB SSB noise figure and high IF bandwidth including conversion gain into 50 ohms. The versatile MGA-85563, although nominally a low-noise amplifier, is also used in three other blocks in **Figure 3** as an LO buffer and transmit driver.

At first glance, the use of eight separate RF building blocks at 5 GHz in **Figure 3** to provide the functionality of two highly integrated RFICs at 2.4 GHz in **Figure 2** may seem to increase the size of the solution astronomically. However, it is important to note the extremely small physical size of these RF building blocks. The IAM-10010 is housed in a tiny MSOP-10 package, which requires less than 50% of the area of the industry-standard SOIC-8. The MGA-83563, MGA-85563, and IAM-91563 are all housed in the ultraminiature SOT-363 package (or 6-lead SC-70), which is actually 30% smaller than the industry-standard SOT-23 transistor package. In addition, separate RF building blocks often provide much better board layout flexibility for filter placement than do highly integrated RFICs. A summary of some key performance parameters for the block diagram of **Figure 3** is given in **Table IV**.

Table IV

Summary of 5-GHz Wireless LAN Up/Downconverter Performance

Supply Voltage	3V
Receive Mode Current	70 mA
Transmit Mode Current	250 mA
Frequency Band	5.15 to 5.35 GHz
Transmit Power	100 mW
Receiver Noise Figure	<10 dB
Receiver Input Third-Order Intercept (IIP3)	−10 dBm
Board Area	20 cm^2

Future Technology Trends

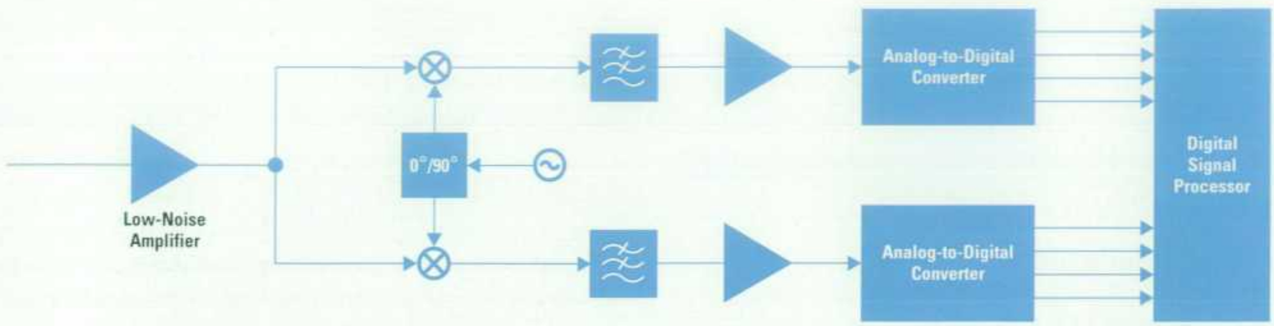
A key trend is the growing proliferation of incompatible data communications services. Data communications is being provided by paging, short-messaging services, data connections through cellular systems, wireless in-building LANs, and fixed satellite systems. Services that are just becoming available include mobile satellite systems, community wireless networks, and wireless multimedia networks. This proliferation of communications standards is occurring in the U.S. cellular industry as well. The new Personal Communications Services (PCS) bands are being populated with at least four different digital communications standards, and compatibility with the existing analog standard will be desired as well. A technical challenge, but one many consumers request, is to build a single radio that is flexible enough to be used for many of these multiple applications.

Another major trend arises from the demand for higher bit rates, which require the allocation of larger blocks of spectrum. Such expansive swaths of bandwidth are only available at higher frequencies. For example, the U.S. FCC has recently reallocated 300 MHz of bandwidth to unlicensed data communications at 5 GHz, and has already approved 5 GHz of spectrum at 60 GHz for unlicensed data communications. One advantage of using higher frequencies is a reduction of the size of an antenna. However, creating circuits that operate at such high frequencies requires improved design techniques and higher-performance devices.

Presently, cellular telephones are made from hundreds of components assembled on high-quality printed circuit boards. While the integration level of microprocessors has grown exponentially, the very high performance requirements of radio communications has slowed efforts toward integration. However, as low-cost technologies with high integration capabilities gain greater performance, the ability to integrate radios increases. An obvious example is the HPMX-5001 from the wireless LAN

Figure 4

Direct-conversion homodyne receiver architecture.



of Example 2 at 2.4 GHz. This one RFIC replaces several building-block-level RFICs or dozens of discrete components. However, CMOS, with its tremendous scale of integration capabilities and rapidly improving performance, offers the potential for substantial increases in radio integration. Research at universities hypothesizes that it may someday be possible to put a complete radio on a single IC with very few external components, at least for fairly simple systems such as ISM-band wireless LANs. The size reduction, along with the ease of embedding such a one-chip radio, should enable the use of radios in palm-tops, lap-tops, and other computing appliances. Also, much of the cost of a radio is presently in the discrete passive components, assembly, and test.

In Examples 2 and 3, **Figures 2 and 3** showed the block diagram of a traditional superheterodyne radio. This architecture has been used in virtually every commercial radio for the past 50 years. The great advantage of this technique is that the filtering to select the desired channel can be done in stages at conveniently located fixed frequencies.

Unfortunately, the filters shown in the heterodyne radio architecture are very difficult to implement on an integrated circuit. Precise high-performance filtering can be performed on an IC, but only at low frequencies. A number of techniques, which have been known for years but rarely used, are being reexamined to solve this problem. **Figure 4** shows an implementation of a direct-conversion (or homodyne) radio. In this architecture the carrier frequency is immediately converted to a very low frequency where it can be filtered and digitized with great precision by on-chip circuitry. There are several serious drawbacks to this approach, many related to the very high dynamic range required by radios in a naturally interference-prone environment. Solving these drawbacks will require further research, but the new design possibilities available in monolithic integration could provide possible solutions.

Figure 4 also raises the issue of digital versus analog processing. Presently, most radios convert the signal into the digital domain only after extensive processing in the analog domain. All filtering, gain control, and demodulation are performed in the analog domain. This approach has been mandated by the great precision and relatively high speed required for these functions.

However, as CMOS technology has progressed, the ability to do digital computations has grown nearly exponentially. It is now possible to perform digital operations to accomplish what has been done with analog filters at comparable power dissipation levels. The advantages of digital processing are many. Digital circuits can be designed more rapidly and verified more accurately. Digital circuits operate extremely reliably. They are less sensitive to temperature, power supply, and processing variations, and do not need tuning or adjustment.

Perhaps most important, the move to more digital processing simplifies attempts to provide flexible radios for multiple applications. For example, if the final channel filtering is performed in the digital domain, it is much simpler to adapt to the different channel bandwidths that are required for various applications. Similarly, different modulation formats can be decoded by a simple change in the programming of the digital part. Such wide-ranging flexibility (consider switching from a 30-kHz-wide channel for data on U.S. cellular bands to 1-MHz channels for data on the U.S. ISM bands) is very difficult for analog filters.

In the long term, the ability of high-performance CMOS to alter the boundary between traditional analog circuitry and digital processing will have a profound effect on RF products independent of the tremendous potential for CMOS as an RF analog device.

Conclusion

Rapidly evolving system standards and applications are now placing even more demands upon the RF semiconductor manufacturer. As standards stabilize, more semiconductor manufacturers will produce highly specific and optimized RFIC solutions for specific growth segments including wireless data terminals.

Customers increasingly demand easy-to-use RF solutions that are available from a reduced set of vendors. This is especially true in the area of wireless data as non-traditional RF customers add connectivity to their end-products to increase the utility of these products. These customers desire component solutions that allow quick entry into the market with minimum risk.

Active RF components will continue to absorb functionality that was previously implemented with passive RF components. This provides benefits in the areas of cost, size, and weight. HP has already introduced silicon bipolar products that include on-chip bandpass filtering, on-chip Schottky diodes in a mixer configuration, and reactive elements for output stage matching.

Despite its relatively small market size in 1997, increased attention and resources are being devoted to the wireless data market by RF component manufacturers, who are betting on the emergence of widely used applications, adoption of interoperability standards, and installation of infrastructure so that this market can grow in accordance with explosive forecast projections.

Kevin J. Negus

Kevin Negus is an R&D section manager at the HP Communications Semiconductor Solutions Division. He is responsible for silicon and gallium arsenide RFIC design, RF systems design, and new product pathfinding. He has been with HP (formerly Avantek) since 1988. A native of Fredericton, New Brunswick, Canada, he received his PhD degree in mechanical engineering from the University of Waterloo in 1988. In his free time he enjoys water and snow skiing, ice hockey, and hunting.

Bryan T. Ingram

Bryan Ingram is the strategic marketing manager for RF semiconductor components at the HP Communications Semiconductor Solutions Division, responsible for business strategy and planning, new product initiatives, and strategic alliances. He received his MSEE degree from Johns Hopkins University in 1990 and joined HP in 1991. Born in Carbondale, Illinois, he is married, has two children, and plays golf and basketball.

John D. Waters

John Waters is a project manager for wireless data networks in the home communications department of HP Laboratories Bristol, England. He received his BSc degree in electrical engineering from Bath University in 1976 and joined HP in 1988. Born in London, he is married and has two children. His interests include cabinet making and classic car renovation.

William J. McFarland

Author's biography appears on page 6.

0.1-μm Gate-Length AlInAs/GaInAs/GaAs MODFET MMIC Process for Applications in High-Speed Wireless Communications

Hans Rohdin
Avelina Nagy
Virginia Robbins
Chung-Yi Su
Arlene S. Wakita
Judith Seeger
Tony Hwang
Patrick Chye
Paul E. Gregory
Sandeep R. Bahl
Forrest G. Kellert
Lawrence G. Studebaker
Donald C. D'Avanzo
Sigurd Johnsen

To ensure high performance of MODFETs used in HP's high-speed communications applications, their high-frequency signal, noise, and power characteristics must be optimized.

Hewlett-Packard has developed an ultrafast III-V (AlInAs/GaInAs/GaAs) modulation-doped field-effect transistor (MODFET) technology for use in high-speed monolithic microwave integrated circuits (MMICs). Applications for these devices include:

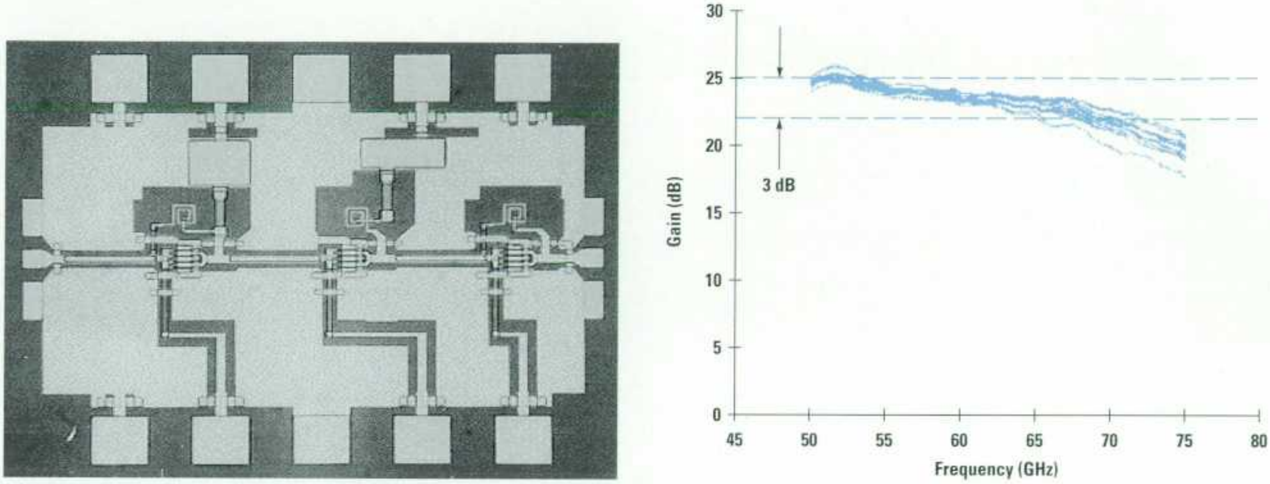
- Wireless millimeter-wave communications
- Fiber-radio personal communications systems
- Automobile collision-avoidance radar
- Optical fiber and low-noise direct broadcast satellite (DBS) communications receivers.

This article provides a summary of the device physics necessary to optimize the high-frequency signal and noise characteristics, semiconductor material, and output-signal power of MODFETs. It also discusses the material, processes, devices, and circuits developed and optimized at HP Laboratories for performance and manufacturability. We conclude with a discussion on process developments at HP's Communication Semiconductor Solutions Division and Microwave Technology Division.

Figures 1 and **2** are examples of circuits that have been fabricated with the MMIC process.

Figure 1

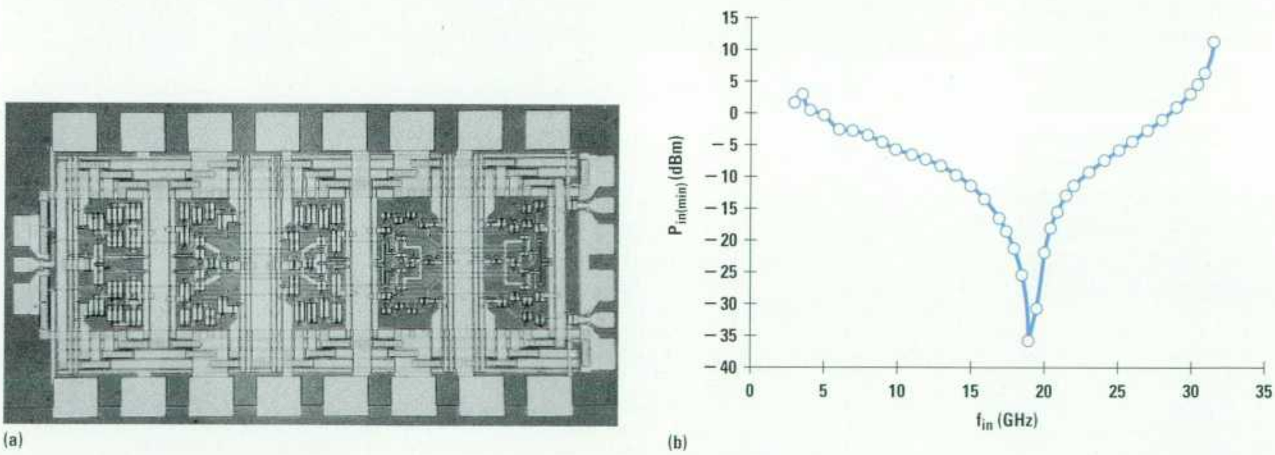
(a) Three-stage feedback MMIC amplifier, and (b) its V-band frequency response. The traces correspond to different circuit locations on the 2-inch wafer.



(a) (b)

Figure 2

(a) Static divide-by-4 circuit, and (b) its input sensitivity curve up to 31 GHz (measured on-wafer).



(a) (b)

The Hewlett-Packard Journal Online

This complete article can be found at:
http://www.hp.com/hpj/98feb/feb98a4.htm

An Enhancement-Mode PHEMT for Single-Supply Power Amplifiers

Der-Woei Wu

John S. Wei

Chung-Yi Su

Ray M. Parkhurst

Shyh-Liang Fu

Shih-Shun Chang

Richard B. Levitsky

To address the growing handset power amplifier needs for the emerging Personal Communications Services (PCS) markets, a 3-volt, single-supply, enhancement-mode pseudomorphic high-electron-mobility transistor (E-PHEMT) has been developed. The device exhibits +33-dBm output power and 65% drain efficiency at 1.88 GHz.

PCS telephone handset manufacturers have had to face some tough choices to find the most suitable technology for their output power amplifiers. Most manufacturers would prefer to use an amplifier that operates on the 3.0V to 3.6V provided by three nickel-cadmium battery cells or one lithium-ion battery cell.

However, GaAs MESFETs (metal-semiconductor field effect transistors) or PHEMTs (pseudomorphic high-electron-mobility transistors) capable of meeting this need have also required an additional negative voltage supply for proper biasing. Alternatively, a manufacturer could choose a single-supply amplifier using silicon bipolar junction transistors, GaAs heterojunction bipolar transistors, or silicon MOSFETs (metal-oxide-semiconductor field effect transistors). These devices typically require a supply voltage around 4.8V, and to use them the manufacturer would have to raise the battery voltage or use a dc-to-dc converter. All of these approaches have efficiency, size, complexity, or cost trade-offs that manufacturers would prefer not to accept if a better device technology were available.

To address this need, an enhancement-mode PHEMT (E-PHEMT) has been developed that provides excellent performance using a single 3V supply. In this paper, the original enhancement-mode technology development by HP Laboratories is presented, followed by device development work performed at the HP Communication Semiconductor Solutions Division (CSSD). RF performance results of both a process control monitor device and a 12-mm gate periphery device are discussed, and a comparison of E-PHEMT performance with other device technologies highlights the merits of this technology. Preliminary device reliability and qualification test results are presented, and test results from a low-cost, plastic packaged, 12-mm gate periphery transistor are discussed.

Development at HP Laboratories

HP Laboratories has been developing an enhancement-mode field effect transistor (EFET) to improve the speed of driver devices in enhancement/depletion-type digital circuitry. This type of device offers two special benefits in handset power amplifier circuitry. Firstly, an EFET only requires positive bias voltage, and can be turned off with zero volts on the gate, but conducts current with positive voltage applied to the gate. Since the drain of a PHEMT is also biased with positive voltage, eliminating the need for negative gate voltage (required for the conventional depletion-mode FET, or DFET) would simplify the operation of the PHEMT.

Secondly, because of the requirement to deliver maximum current over a small gate voltage swing ($\sim$1 volt), a correctly designed EFET inherently has higher transconductance (g_m), and, more important for class-B power FET operation, better g_m linearity than a conventional DFET. The HP Laboratories' approach employs molecular beam expitaxy (MBE) to place a highly doped, thin electron donor layer close to the gate, with the electron donors on both sides of the $In_{0.2}Ga_{0.8}As$ channel. Highly selective reactive ion etching is used to define the vertical position of the Schottky gate. **Figure 1** is a schematic diagram of the MBE layer structure of the enhancement mode PHEMT.[1]

Figure 1

Schematic diagram of the MBE layer structure of the enhancement-mode PHEMT.

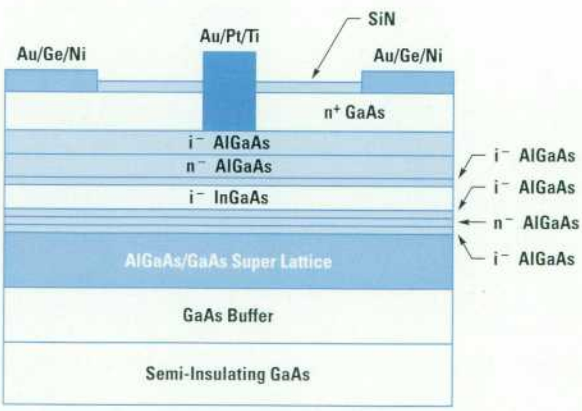


Table I summarizes the key characteristics of HP Laboratories' enhancement-mode self-aligned contact (SAC) PHEMT. An important point is that the maximum current achieved is no less than that of similarly fabricated depletion-mode PHEMTs, and is far higher than that of typical MESFETs.

Table I

Key Characteristics of HP Laboratories' E-PHEMT

Parameter	Unit	Conditions	Typical Wafer Mean	Typical Wafer σ
Threshold Voltage	V	$V_d = 2V$	0.06	0.04
Maximum Transconductance	mS/mm	$V_d = 2V$	708	57
Maximum Drain Current	mA/mm	$V_d = 2V$, $I_g < 100$ mA/mm	515	36
f_T	GHz	optimum	60	
f_{max}	GHz	optimum	120	

To explore the potential of HP Laboratories' enhancement-mode SAC PHEMT as a handset power device, a large-signal model was extracted from 100-μm gate width devices. An extensive HP Microwave Design System simulation was performed for both analog and digital handset applications. **Table II** summarizes the simulated results for both two-tone input (digital modulation) and one-tone input (analog modulation).

Table II

Power Performance of E-Mode SAC PHEMT

Parameter*	Unit	Specification	Simulation Results
2-Tone Power Out	dBm	28.5	28.5
Power-Added Efficiency @ 28.5 dBm	%	>50	53
Gain @ 28.5 dBm	dB	>15	19.6
IM3 @ 28.5 dBm**	dBc	< −26	−40
IM5 @ 28.5 dBm**	dBc	< −36	−37
IM7 @ 28.5 dBm**	dBc	< −45	−45
1-Tone Power Out	dBm	31.5	31.5
Power-Added Efficiency @ 31.5 dBm	%	>60	71.6
Gain @ 31.5 dBm	dB	>15	19.5
Quiescent Current	mA		145
Gate Width	mm		8.6

* Test Condition: 900 MHz and $V_d = 3V$

** Intermodulation (IM) distortions have been calculated for two tones spaced by 5 kHz at 900 MHz.

The very encouraging simulated results prompted the HP Communication Semiconductor Solutions Division (CSSD) to further examine HP Laboratories' enhancement-mode SAC PHEMT. Load-pull measurements at 2 GHz performed on 100-μm gate width devices indicated 13.4-dBm (219-mW/mm) power, 19.4-dB gain and 75.3% power-added efficiency at $V_d = 3$ volts. Since these results exceeded the anticipated product specifications by significant margins, a strategy was formulated to make available a first-generation product based on a simpler, nonself-aligned process.

Figure 2 shows the power performance at 2 GHz of 200-μm gate width DFETs and EFETs fabricated at CSSD with a nonself-aligned process. Although the performance was not as good as HP Laboratories' self-aligned FETs, the 223-mW/mm saturated power and 66% power-added efficiency exhibited by these devices at $V_d = 3$ volts are still state-of-the-art. **Figure 2** also clearly establishes the superior gain and efficiency of EFETs compared to DFETs at low input power.

The power measurements taken on EFETs and DFETs, which are processed the same way, also dispell another concern of employing EFETs for power amplification: that EFETs will draw unusually large gate leakage current in power saturation. From the load-pull measurements done for the data shown in **Figure** 2, the quiescent gate leakage current at 3-dB gain compression is positive and < 1 mA for the EFET, and is negative and < 1 mA for the DFET.

With the groundwork at HP Laboratories clarifying the potential of an enhancement-mode power PHEMT, a project was launched at CSSD and the results from that project are described in the remainder of this article.

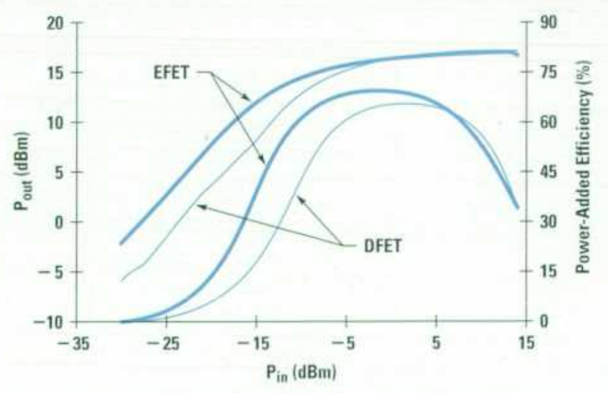
Development at CSSD

For the 3-volt, enhancement-mode, high-gain power transistor, one has to investigate candidate devices to determine their potential for low knee (saturation) voltage, high transconductance at low quiescent current, and high current supply capability. High transconductance at low current gives better linearity when the device is operated at low currents in high-linearity applications, and high current capability leads to better power output. The 3-volt operation favors PHEMTs because the PHEMT drain current saturates at very low voltages (low knee voltage). This allows greater dynamic voltage swing. In addition, the high electron mobility of PHEMTs provides high gain and high current. Pseudomorphic HEMTs are fabricated on GaAs substrates using MBE.

Discrete devices were fabricated with several different MBE material profiles. To evaluate these devices, on-wafer small-signal data and load-pull data was collected

Figure 2

Comparison of the power performance of EFETs and DFETs. Gates are 200 μm wide and 0.2 μm long. Gates and contacts are not self-aligned to each other.



on small devices, while RF circuit evaluation was done on large devices of 12-mm total gate periphery. The quick turnaround of the on-wafer measurements speeded the optimization of MBE material structures. Circuit evaluation and load-pull data on 12-mm devices further validated the small-signal results. CSSD's implementation of the material structure from HP Laboratories (the enhancement-mode PHEMT originally for digital applications) initially showed poor performance. After material growth optimization with the help of HP Laboratories, performance improved dramatically and a final structure was adopted for the project.

This demonstrates the problems that can occur in technology transfer, since the material growth and processing conditions used at HP Laboratories needed to be adapted and modified to run on the different equipment and processes used at CSSD. Through close collaboration with HP Laboratories, high-quality E-PHEMT epitaxial material is now routinely grown at CSSD. Wafer processing was refined and adapted to the standard PHEMT process used in the CSSD GaAs fabrication facility.

Device Layout

Key factors that determine the best device layout are total current capability, frequency response, cost, thermal dissipation, and specific performance requirements. A power transistor consists of an array of interdigitated source and drain pads separated by gate fingers, essentially a combination of small unit cells. The total current available is determined by the product of the number of fingers and the width of each finger. The current capabiity of a PHEMT also varies with the material structure, and the goal is to have the current necessary for a specified power output in the smallest device possible. To evaluate device layout trade-offs, a mask was generated with eight devices having different geometries.

For better response at higher frequencies, the width of the individual fingers of the transistor needs to be smaller. Finger widths range from 200 to 400 μm for the eight devices. The wider the fingers, the more symmetric the device aspect ratio, the better the real estate utilization, hence the lower the device cost, a typical trade-off of cost versus performance. On a microwave transistor, the bonding pads often occupy half the device. On this mask set, the cost of some devices was reduced by having separate narrow source bonding pads instead of a large conventional one-piece source pad surrounding all the gate pads.

Within each device, in addition to the gate finger width, the gate length (the narrow dimension of the gate) sets the upper limit of the frequency of operation. For the PCS/PCN frequency bands, 0.5-μm gate length is suitable, and has been adopted in the present project.

Fabrication Process

The wafer processing uses the standard CSSD PHEMT process. Fabrication of an enhancement-mode PHEMT requires some precaution. The gate metal needs to be placed just on top of the upper undoped AlGaAs layer to achieve zero-volt threshold operation. Since the active channel of the device lies very close to the exposed top surface, inadvertent exposure to some chemical solutions during processing can erode the channel, resulting in nonuniformity and current loss. With conventional chemical wet etching of 10-nanometer-thick layers, nonuniformity is inevitable. Fortunately, the composition of the different layers lends itself to a selective plasma reactive ion etching process that is self-stopping at the top AlGaAs layer.

After isolating the active area of the transistor with proton implantation, ohmic contacts to the source and drain pads are deposited. Then the gate opening is delineated in photoresist and the wafer placed in a reactive ion etcher to remove the top GaAs layer from the gate opening region. Immediately afterwards, multilayer metal is evaporated onto the wafer to form the gate electrode. Excess metal is lifted off by removing the photoresist in solvent. Wafer passivation is achieved using plasma silicon nitride. A final gold plating step forms the "air-bridge" interconnect metal (second metal), which links the individual fingers of the transistor array through bridges arching over bus bars underneath.

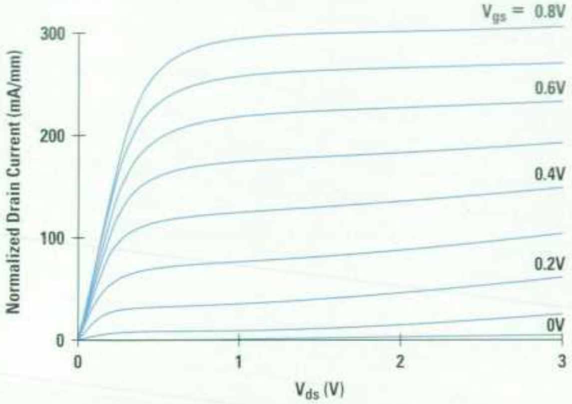
Process Control Monitor Device Performance

The current-voltage (I-V) characteristics of a 60-μm ($L_g = 0.5$ μm), double-doped E-PHEMT process control monitor device are shown in **Figure 3**.

Typical dc results are as follows. The drain-source saturation current density (I_{dss}) measured at $V_{ds} = 3$ volts is less than 10 mA/mm. The maximum drain current density (I_{max}), measured at $V_{gs} = 0.8$ volts and $V_{ds} = 3$ volts is greater than 300 mA/mm. The maximum transconductance (g_m) is greater than 370 mS/mm. The gate-to-drain breakdown voltage (BV_{gdo}) measured at a gate current of 0.5 mA/mm is more than 10 volts. The pinch-off voltage (V_p) measured at $V_{ds} = 2$ volts and drain current density of

Figure 3

The I-V characteristics of a 60-μm process control monitor device.



2 mA/mm, is 0 volts. The knee voltage, defined as the drain-source bias when the drain-source current density becomes 200 mA/mm with V_{gs} equal to 0.8 volt, is 0.3 volt. The on-resistance, measured at $V_{gs}=0.8$ volt, is 1.5 ohms-mm.

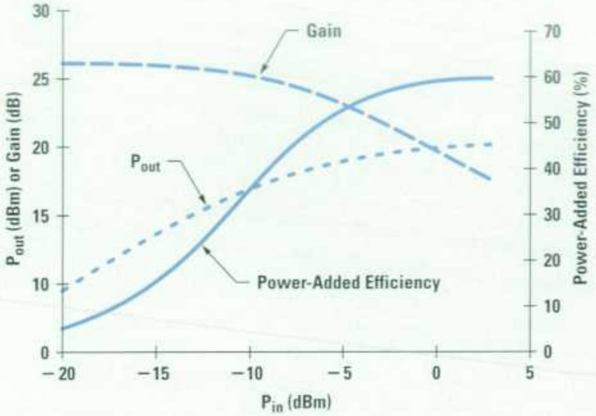
The small-signal parameters of the 60-μm devices were measured on-wafer using an HP 8510C automatic network analyzer. Based on the s-parameter results, the unity-current-gain frequency (f_T) for $V_{ds}=3$ volts and $I_{ds}=20$ mA is calculated to be 36 GHz. The maximum available gain/maximum stable gain at 2 GHz is 22 dB under the same bias condition.

The large-signal performance of the process control monitor devices was also monitored on-wafer using an automatic load-pull system. The large-signal characteristics were measured at $V_{ds}=3$ volts and $I_{ds}=20$ mA. With fixed input power, the system sets the input impedance to provide maximum small-signal gain, and then tunes the load impedance for maximum gain, maximum power, and maximum efficiency.

The power saturation characteristics for a 300-μm device under the maximum output power tuning are shown in **Figure 4**. The device exhibited an output power of 16.4 dBm at the 3-dB gain compression point with 3-volt bias at 2 GHz, corresponding to a power density of 140 mW/mm. In addition, the device demonstrated a saturated output power of 18 dBm, power-added efficiency of 60%, and power gain of 13 dB, which corresponds to a power density of 210 mW/mm.

Figure 4

The power saturation characteristics of the 300-μm process control monitor device at 3 volts and 2 GHz.



12-mm Device Performance

The next step in the E-PHEMT development project was to measure the performance of a large device that could actually be used as the output stage in a PCS telephone. The power device used for this evaluation was a 0.5-μm gate length and 12-mm (330-μm × 36-finger) gate periphery E-PHEMT. The layout of the device is shown in **Figure 5**.

The power performance of the 12-mm E-PHEMT was measured with a drain bias of 3 volts and a quiescent drain current of 100 mA (V_{gs} = +23 mV) at 1.8 GHz. To

Figure 5

The layout of a 12-mm power device.

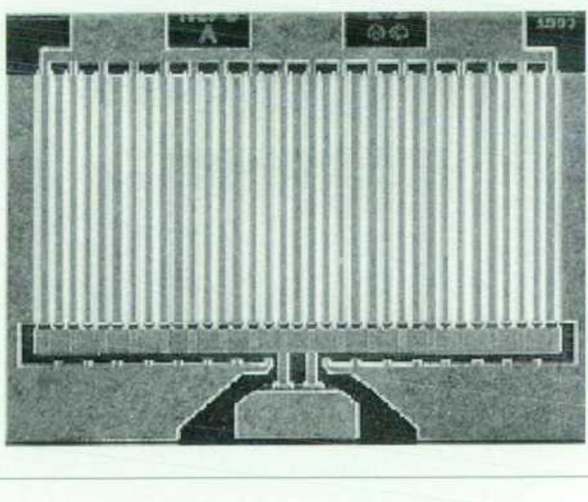


Figure 6

A 12-mm device mounted on a microprobing test fixture used in the on-wafer active load-pull system.

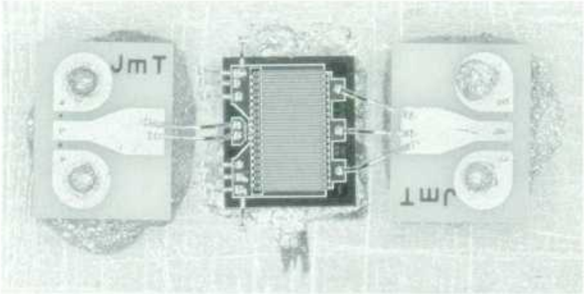
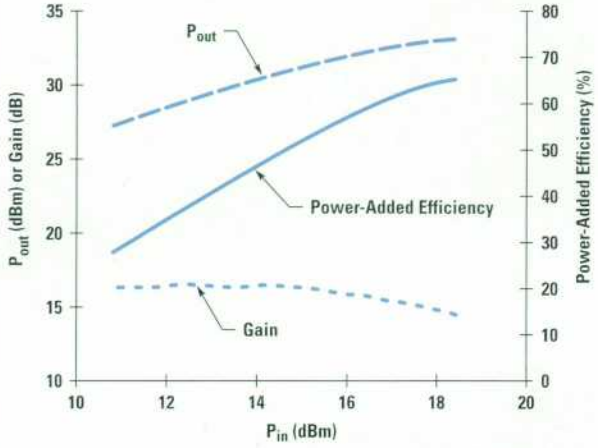


Figure 7

Power saturation characteristics for the 12-mm device at 3 volts and 1.8 GHz.



improve yield and reduce cost, the wafer was lapped to a thickness of 0.004 inch with no via grounding. The device was then eutectically mounted onto a metal carrier, and two ceramic probe adapters, which provide the coplanar-to-microstrip transition, were placed at the input and output of the device (**Figure 6**). Bond wires were used to connect the gate and drain metallization to the adapter microstrip. Several bond wires were also bonded down from the source pads to the ground metallization to ensure low ground inductance.

The power characteristics were measured using a vector-corrected, active load-pull system.[2] The use of an active load-pull system ensures that sufficiently low impedance is presented to the device for optimal performance. In addition, it also provides the capability to study the effects of harmonic tuning on power performance. **Figure 7** shows the power saturation characteristics for the 12-mm device under maximum-power tuning. The results were obtained with both fundamental and second-harmonic tuning. The improvement in power-added efficiency by terminating the second harmonics is about 4 to 6 percent. The 12-mm device achieved +33-dBm output power (corresponding to a power density of 167 mW/mm), 14.7-dB power gain (2-dB gain compression), and 65.4% power-added efficiency at 3 volts and 1.8 GHz, which is the highest combined power performance ever reported at such a low bias voltage.

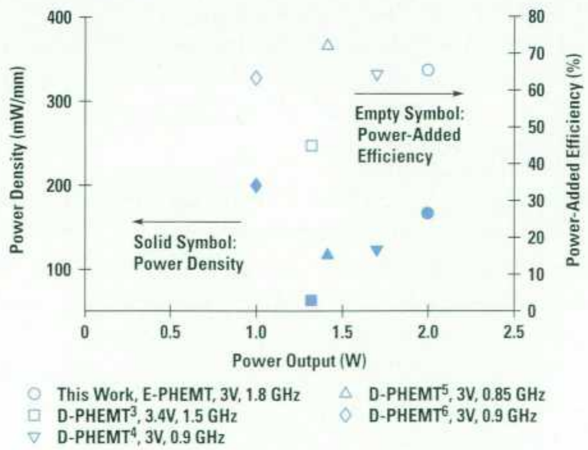
State-of-the-Art Performance Comparison

To benchmark the performance, the results obtained were compared with several competing technologies used in similar applications. The power performance (in terms of the power density) of the E-PHEMT was compared with several state-of-the-art, depletion-mode PHEMT devices operating in the 3-volt range (**Figure 8**). Several results were reported in the 900 MHz range. Also, most of the depletion-mode devices require dual supplies, as opposed to a single supply in our case.

Table III compares HP's 12-mm E-PHEMT device performance with high-performance MESFETs, GaAs/AlGaAs

Figure 8

Power performance comparison of state-of-the-art PHEMTs in the 3-volt range.



heterojunction bipolar transistors (HBTs), high-performance silicon bipolar junction transistors (BJTs), Si lateral-diffused metal-oxide-semiconductor (LDMOS) devices, and SiGe HBTs. The ability to operate from a single power supply and demonstrated excellent power performance (battery-efficient) at low voltages make the E-PHEMT very suitable for power generation in portable wireless applications.

Table III

Power Performance Comparison for Competing Technologies in Portable Wireless Applications

Device Technology	Frequency (GHz)	P_{out} (dBm)	Power-Added Efficiency (%)	V_{ds} (V)
This Work	1.8	33	65.4	3
SiGe HBT[7]	1.9	30	44	4.7
Double-Poly Si BJT[8]	1.8	24	60	3.5
Hi-Lo MESFET[9]	0.9	31.3	68	2.3
MESFET, $2f_o$ Tuning[10]	0.93	32.8	71	3.5
GaAs/AlGaAs HBT[11]	1.9	22.6	69	3
GaAs/AlGaAs HBT[12]	1.88	33	70	5
BPLDD SAGFET*[13]	1.9	24.7	54	3.3
Delta-Doped MESFET[14]	1.5	30.4	48	3.5
Si LDMOS[15]	0.85	31.8	65	5.8

* BPLDD SAGFET = buried p-layer lightly doped drain self-aligned gate field effect transistor.

Reliability

Reliability studies have been performed on devices from five separate wafer runs. The initial purpose of these studies was to determine whether a 2-layer or 3-layer passivation process should be applied. Stress-induced changes in I_d ($V_g = 0.75V$) and BV_{gdo} ($I_g = 500$ mA/mm) were examined. Since this is an enhancement-mode device, it is impractical to use changes in I_{DSS} as a failure parameter, so I_d ($V_g = 0.75V$) was selected.

In a 3-layer SiN_x passivated E-PHEMT, the first layer of SiNx is plasma deposited immediately after MBE. It is believed that this step protects the surface of the wafer from contamination during processing. Any contamination can create surface states that provide a leakage path between the gate and drain fingers, thus increasing sheet resistance and eventually degrading BV_{gdo}. The second layer of SiN_x is deposited after the formation of the ohmic contact metal on the source and drain. The final SiN_x passivation is performed after gate metal liftoff. The 2-layer passivated E-PHEMTs skip the first step.

Table IV shows the change in BV_{gdo} after greater than 1000 hours of stress. All devices were electrically stressed at $I_{ds} = 50$ mA/mm, $V_{ds} = 3V$. In addition to a less severe change in BV_{gdo}, the 3-layer devices also show less degradation at I_d ($V_{gs} = 0.6V$) and $g_{m\,max}$, as shown in **Figures 9a and 9b**.

Since this data indicates superior performance for a 3-layer process, all subsequent reliability tests have been performed using devices with a 3-layer passivation.

Table IV

Degradation of BV_{gdo} in 2- and 3-Layer Passivated E-PHEMTs

2-Layer SiN_x Passivation

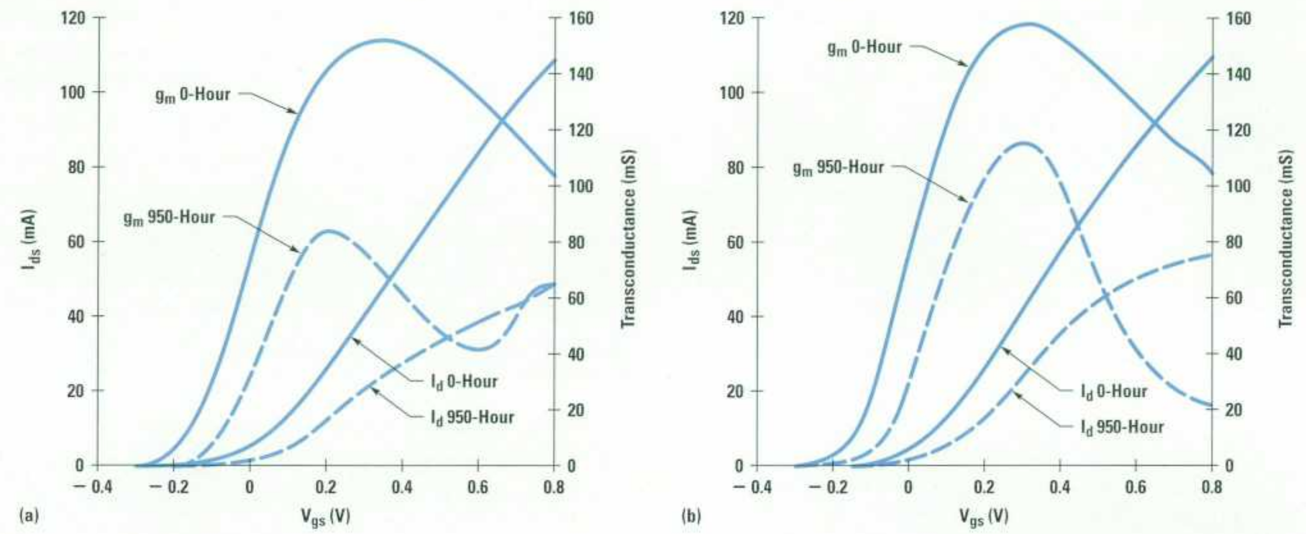
$T_{channel}$ (°C)	BV_{gdo} (V) 0-hr	BV_{gdo} (V) 1085-hr	Percent Change
175	8.72	7.71	−11.58
200	9.52	7.92	−16.80
225	9.31	6.54	−29.75

3-Layer SiN_x Passivation

$T_{channel}$ (°C)	BV_{gdo} (V) 0-hr	BV_{gdo} (V) 1181-hr	Percent Change
175	9.66	10.22	5.79
200	N/A	8.97	
225	9.1	9.12	0.21

Figure 9

(a) I-V characteristics after 950 hours of stress for a device with 2-layer passivation. Solid lines are for 0 hours. I_{ds} (0.6V) and $g_{m\ max}$ degraded by 54% and 49%, respectively. (b) I-V characteristics after 950 hours of stress for a device with 3-layer passivation. Solid lines are for 0 hours. I_{ds} (0.6) and $g_{m\ max}$ degraded by 42% and 27%, respectively.



The change in s-parameters during dc burn-in has also been examined, as shown in **Figure 10**. The filled data points are for 0 hours, while the open data points are after 1709-hour burn-in. The results from this test are extremely encouraging, in that s_{21} shows virtually no change. The biggest change (average 3-dB reduction) was seen in s_{12}. However, the sensitivity of this measurement could account for this change.

To date, the reliability tests have been preliminary. For the technology qualification, three recent wafers, with 30 devices from each wafer, will be dc stressed at 175°C, 200°C, and 225°C. The parameters that will be monitored are I_d ($V_g = 0.4$) and BV_{gdo} ($I_g = 500$ μA/mm). RF burn-in has also been planned, and additional tests to compare results between 2-layer and 3-layer passivated devices will be done.

Packaged Device Evaluation

The low-cost plastic package chosen to evaluate the performance of the E-PHEMT was an HP micro-small-outline package (MSOP) with grounded backside. This package features an exposed die attach paddle that can be soldered directly to the printed circuit board for thermal and electrical grounding (**Figure 11**).

The die attach paddle is approximately 0.040 inch wide by 0.090 inch in length, which is large enough to incorporate the first LC sections of the RF matching structures. These matching networks consist of silicon MOS capacitors attached to the paddle, resonated with the device drain and gate bond wires (see **Figure 12** for a photograph of the structure and **Figure 13** for a schematic diagram). Placing the capacitors inside the package on the grounded die attach paddle provides the shortest possible ground return path and minimizes parasitic inductive loss elements.

Evaluation Printed Circuit Board

The matching networks contained within the package provide only partial matching for the device gate and drain. Additional matching networks are required external to the package to match the input to 50 ohms and the output to the optimum impedance for maximum power, linearity, or efficiency. The evaluation printed circuit board

Figure 10

Zero and 1709-hour s-parameter data at 1.8 GHz, showing the effect of burn-in. The channel temperatures were 175° C, 200° C, and 225° C for devices 46-52, 53-59, and 60-67 respectively. Devices 68 and 69 were control devices.

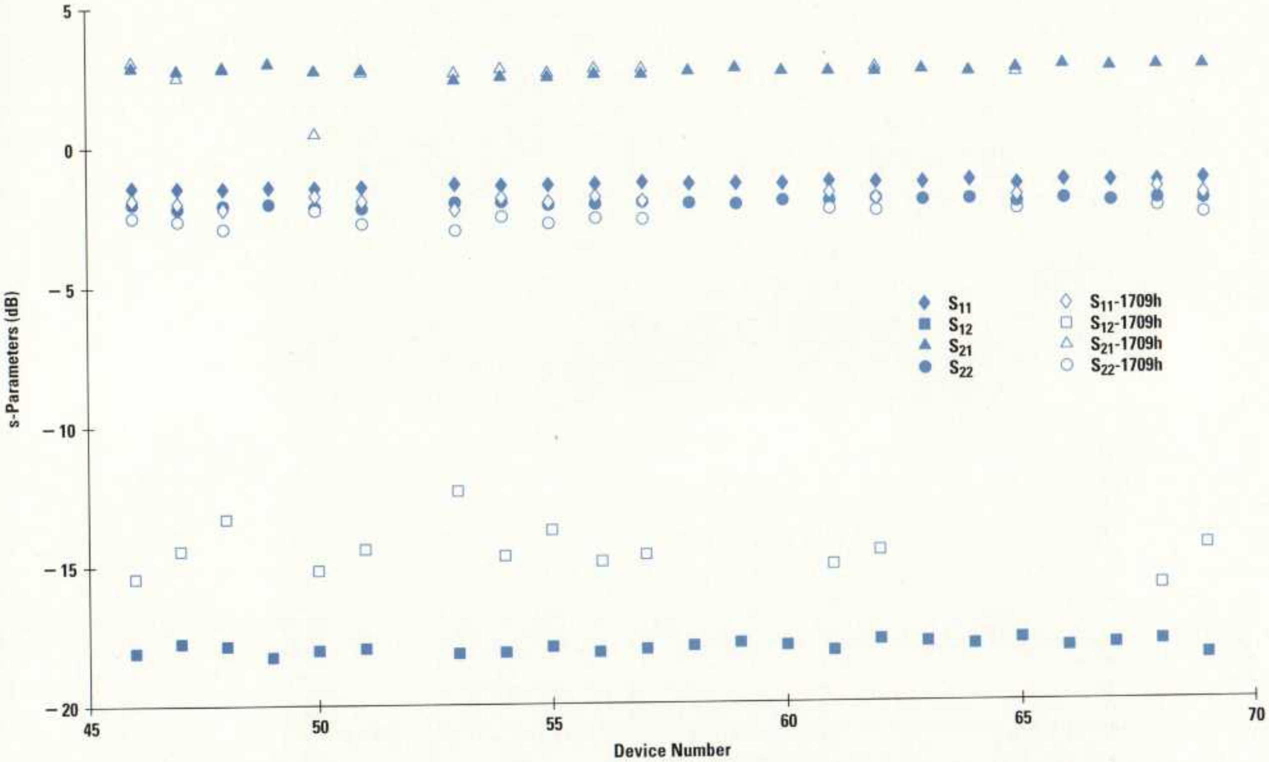


Figure 11

HP MSOP package top and bottom views.

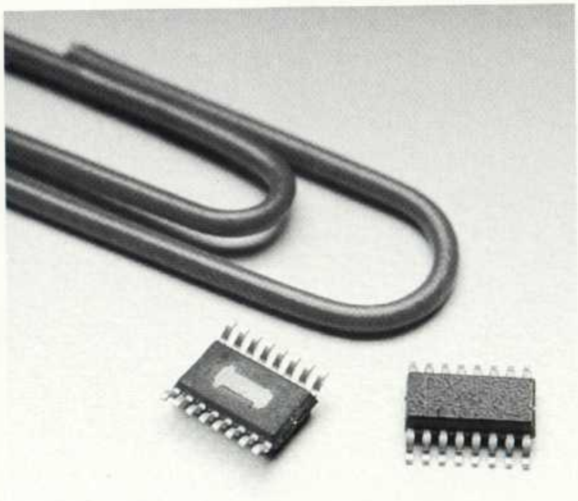


Figure 12

Internal bonding configuration of the E-PHEMT in the MSOP package.

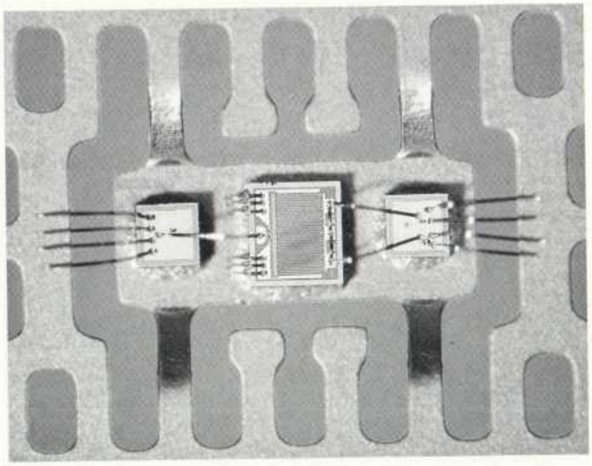
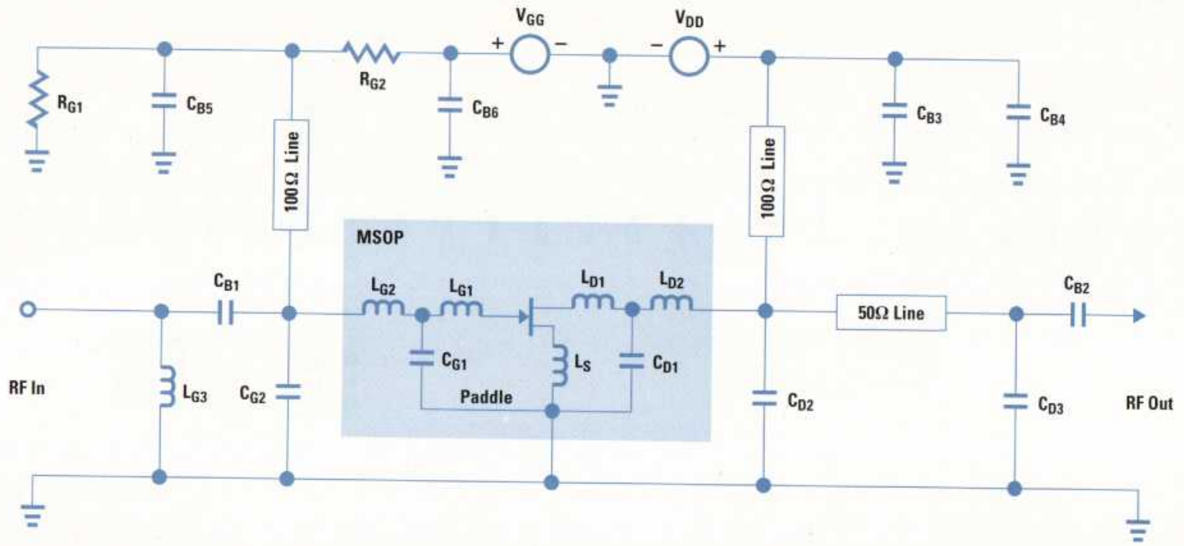


Figure 13

Schematic diagram of the packaged E-PHEMT evaluation circuit.

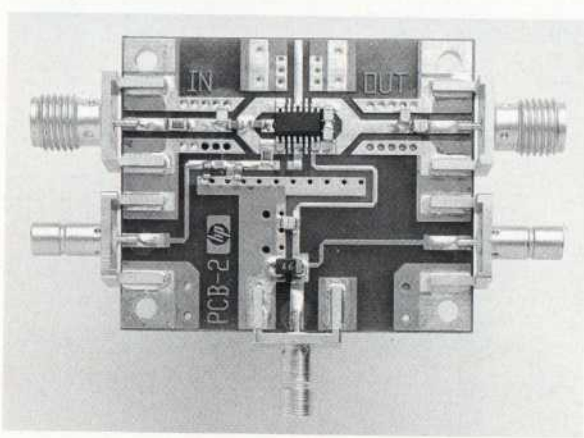


(see **Figure 14**) provides pi-section matching networks on both the input and output to achieve these goals, along with quarter-wavelength transmission line networks for gate and drain biasing.

The complete schematic of the packaged device, external matching networks, and biasing networks is shown in **Figure 13**.

Figure 14

Printed circuit board used to evaluate the packaged E-PHEMT.



Performance

The following test conditions were used to characterize the basic gain, power, and efficiency performance of the packaged E-PHEMT:

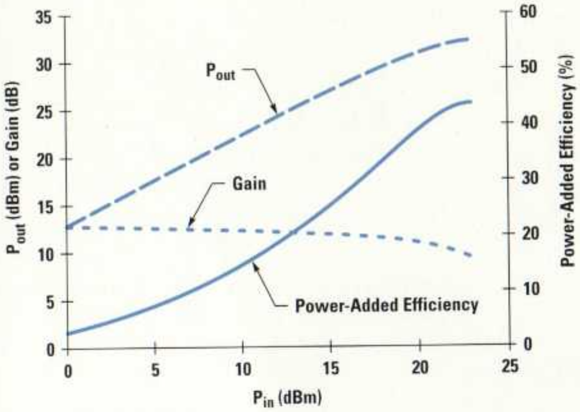
- Frequency = 1880 MHz
- V_{dd} = 3.0 volts
- Duty Cycle = 100% (CW)
- I_{ds} (no signal) = 200 mA (set by adjusting a positive bias voltage at the gate).

The measured small-signal gain was 12 dB. Measured output power was +31.0 dBm at 1-dB compression and better than +32.6 dBm at 3-dB compression. At 32.6 dBm, the power-added efficiency was 42 percent. The performance is summarized in **Figure 15**. Note that for the power and efficiency measurements, the matching network losses were not removed.

In addition to the basic single-tone performance described above, PCS systems also require the handset power amplifier to meet certain linearity requirements to maintain low bit error rates in the transmitted signal. The requirements vary with the modulation technique employed. **Table V** summarizes the performance of the packaged E-PHEMT compared with the requirements for typical CDMA and TDMA PCS systems. Test conditions for

Figure 15

Performance of the packaged E-PHEMT in the evaluation circuit.



these measurements were the same as for the preceding gain, power, and efficiency tests. A 0.8-dB correction has been applied to account for the measured output circuit matching loss.

Table V

Performance of Packaged HP E-PHEMT versus PCS System Requirements

Parameter	PCS CDMA	HP E-PHEMT	PCS TDMA	HP E-PHEMT
Output Power (dBm)	+28.5	+28.4	+28.5	+28.5
Efficiency (%)	30	33.6	40	43
Adjacent Channel Power (dBc)	−29	−31	−26	−27

Conclusion

Leveraging initial work at HP Laboratories, an enhancement-mode PHEMT device technology has been developed that offers excellent gain, power, and power-added efficiency using a 3V bias. Fundamental elements of this new device technology include careful selection of MBE material structure, adaptation of HP Laboratories' fabrication process to mesh with CSSD's existing PHEMT process, and device layout optimization to achieve the best overall measured performance. Devices built using this new technology have demonstrated +33-dBm output power with 65% power-added efficiency and 15-dB power gain when operated from a 3V supply at 1.8 GHz. This compares very favorably with the performance that can be achieved using other device technologies, which require additional components like dc-to-dc converters or higher supply voltages.

HP's new E-PHEMT technology is very well-suited for use in power amplifiers for PCS telephones. Using this technology, PCS power amplifiers that can operate from a 3V single supply with excellent performance are now a reality.

Acknowledgments

This project's success has been due to the contributions of the whole E-PHEMT project team and support from many other individuals both at CSSD and HP Laboratories. In particular, Patrick Chye from GaAs process R&D, Wes Fields from GaAs product engineering, Paul Gregory, Jeff Miller (HP Laboratories), and Virginia Robbins (HP Laboratories) from MBE materials have all provided invaluable contributions. We would also like to thank Steve Kofol, Hans Rohdin, and Avelina Nagy at HP Laboratories for their initial work on enhancement-mode power devices. We are grateful to Alex Cognata an the HP Microwave Technology Division for his assistance in performing active load-pull measurements. Finally, Dennis Moy from CSSD marketing was a great help in coordinating the initial draft and submission of this paper.

References

1. H. Rohdin, et al, "0.1-μm Gate-Length AlInAs/GaInAs/GaAs MODFET MMIC Process for Applications in High-Speed Wireless Communication," *Hewlett-Packard Journal Online*, February 1997, URL:

   http://www.hp.com/hpj/journal.html

2. B. Hughes, A. Ferrero, and A. Cognata, "Accurate On-Wafer Power and Harmonic Measurements of mm-Wave Amplifiers and Devices," *IEEE Microwave Theory and Techniques Symposium Digest*, 1992, pp. 1019-1022.

3. Y. Hasegawa, Y. Ogata, I. Nagasako, K. Inosako, N. Iwata, M. Kanamori, and T. Itoh, "3.4V Operation Power Amplifier Multichip ICs for Digital Cellular Phones," *IEEE GaAs IC Symposium Digest*, 1995, pp. 63-66.

4. K. Inosako, K. Matsunaga, Y. Okamoto, and M. Kuzuhara, "Highly Efficient Double-Doped Heterjunction FETs for Battery-Operated Portable Power Applications," *IEEE Electron Devices Letters*, Vol. 15, 1994, pp. 248-250.

5. V. Nair, S. Tehrani, D. Halchin, E. Glass, E. Fisk, and M. Majerus, "High-Efficiency Power HFETs for Low-Power Wireless Applications," *IEEE Microwave and Millimeter-Wave Monolithic Circuits Symposium Digest*, 1996, pp. 17-20.

6. Y.-L. Lai, E. Chang, C.-Y. Chang, T.K. Chen, T.H. Liu, S.P. Wang, T.H. Chen, and C.T. Lee, "5-mm High-Power-Density Dual-Delta-Doped Power HEMTs for 3V L-Band Applications," *IEEE Electron Devices Letters*, Vol. 17, 1996, pp. 229-231.

7. A. Schöppen, S. Gerlach, H. Dietrich, D. Wandrei, U. Seiler, and U. König, "1W SiGe Power HBTs for Mobile Communication," *IEEE Microwave and Guided Wave Letters*, 1996, pp. 341-343.

8. M. Versleijen, R. Dekker, W.v.d.Einde, and A. Pruijmboom, "Low-Voltage High-Performance Silicon RF Power Transistors," *IEEE Microwave Theory and Techniques Symposium Digest*, 1995, pp. 563-566.

9. J.-L. Lee, J.K. Mun, H. Kim, J.-J. Lee, and H.-M. Park, "A 68% PAE, GaAs Power MESFET Operating at 2.3V Drain Bias for Low-Distortion Power Applications," *IEEE Transactions on Electron Devices*, Vol. 43, no. 4, 1996, pp. 519-526.

10. M. Maeda, H. Takehara, M. Nakamura, Y. Ota, and O. Ishikawa, "A High-Power and High-Efficiency Amplifier with Controlled Second-Harmonic Source Impedance," *IEEE Microwave Theory and Techniques Symposium Digest*, 1995, pp. 579-582.

11. Y. Matsuoka, S. Yamahata, M. Nakatsugawa, M. Muraguchi, and T. Ishibashi, "High-Efficiency Operation of AlGaAs/GaAs Power Heterojunction Bipolar Transistors at Low Collector Supply Voltage," *Electronics Letters*, 1993, pp. 982-984.

12. P. Walters, P. Lau, K. Buehring, J. Penney, C. Farley, P. McDade, and K. Weller, "A Low-Cost Linear AlGaAs/GaAs HBT Power Amplifier with Active Bias Sensing for PCS Applications," *IEEE GaAs IC Symposium Digest*, 1995, pp. 67-70.

13. N. Kasai, M. Noda, K. Ito, K. Yamamoto, K. Maemura, Y. Ohta, T. Ishikawa, Y. Yoshii, M. Nakayama, H. Takano, and O. Ishihara, "A High-Power and High-Efficiency GaAs BPLDD SAGFET with WSi/W Double-Layer Gate for Mobile Communication Systems," *IEEE GaAs IC Symposium Digest*, 1995, pp. 59-62.

14. S. Makioka, S. Enomoto, H. Furukawa, K. Tateoka, N. Yoshikawa, and K. Kanazawa, "A Miniaturized GaAs Power Amplifier for 1.5-GHz Digital Cellular Phones," *IEEE MMWC Symposium Digest*, 1996, pp. 13-16.

15. D. Ngo, C. Dragon, J. Costa, D. Lamey, E. Spears, and W. Burger, "RF Silicon MOS Integrated Power Amplifier for Analog Cellular Applications," *IEEE Microwave Theory and Techniques Symposium Digest*, 1996, pp. 559-562.

Der-Woei Wu

Dave Wu is a project leader for CDMA/AMPS power amplifier modules at the HP Communication Semiconductor Solutions Division. Before joining HP in 1996 he was a senior engineer at M/A-COM. He received his PhDEE degree from Pennsylvania State University in 1993, specializing in microwave heterojunction bipolar transistors. A native of Taiwan, he is married, has a son, and enjoys Chinese calligraphy.

John S. Wei

John Wei received his PhD degree in physics from the University of Illinois Urbana-Champaign in 1974. With HP since 1992, he is the GaAs fabrication unit process manager at the HP Communication Semiconductor Solutions Division and is involved with developing manufacturing processes for GaAs devices. Born in Hankow, China, he is married and has two children.

Chung-Yi Su

Chung-Yi Su joined HP Laboratories in 1981 after receiving his PhDEE degree from Stanford University. He is now GaAs fabrication manager at the HP Communication Semiconductor Solutions Division. He has published 67 papers in the areas of semiconductor devices, solid surfaces, and interfaces. Born in Taiwan, he is married and has three children.

Ray M. Parkhurst

Ray Parkhurst is an R&D project manager for handset power modules at the HP Communication Semiconductor Solutions Division. His BS degree in electronic engineering is from the California Polytechnic State University at San Luis Obispo (1988). A native of Sacramento, California, he is married and enjoys woodworking and audio.

Shih-Shun Chang

After a year in reliability engineering at the HP Communication Semiconductor Solutions Division, Mark Chang recently joined the technology access and devices group as a device engineer. He received his PhDEE degree from the University of Colorado in 1996. His research topic was GaN/SiC heterojunction bipolar transistors for high-temperature applications.

Richard B. Levitsky

Rich Levitsky is the R&D section manager for integrated products and programs at the HP Communication Semiconductor Solutions Division. He holds an MSEE degree from the University of California at Davis (1988) and an MBA degree from Saint Mary's College of California (1991).

Shyh-Liang Fu

Shyh-Liang Fu received his PhD degree from the University of California at San Diego in 1996. Since joining HP in 1996 he has been engaged in research and development for GaAs field-effect transistors. His current work involves the design, fabrication, and characterization of AlGaAs/InGaAs/GaAs pseudomorphic high-electron- mobility transistors (PHEMTs).

Direct Broadcast Satellite Applications

Shunichiro Yajima

Antoni C. Niedzwiecki

One of the main reasons for the popularity of direct broadcast satellite (DBS) service is the small size of the parabolic dish antenna. The key to the small-size dish is a low-noise GaAs transistor used in the low-noise block of the DBS receiver system. One of HP's efforts in this area has been to develop an AlInAs/GaInAs device fabricated on a conventional GaAs substrate that has a lower noise figure, higher gain, and lower cost.

The first digital direct broadcast satellite (DBS) service started in June 1994. Today, DBS technology enables consumers to receive digitally modulated television signals directly from satellites. Consumers who bought digital satellite receivers are enjoying excellent high-quality pictures and more channel access. The dynamics of the DBS market are fueled by heavy competition between multiple DBS operators, who are motivated to provide better services and lower prices. This same competition also motivates receiver suppliers to reduce hardware costs.

DBS systems are now spreading at a rapid rate all over the world. Analog DBS systems have been very popular in Europe and Japan, with installations in 25 million homes since 1984. Both of these areas of the world are now switching to digital systems. In the United States, five million subscribers have already installed digital small-dish satellite systems, and the market is still growing. DBS services are just starting to be implemented in Asia and South America.

Shunichiro Yajima

Shun Yajima is a strategic marketing engineer at HP's Communication Semiconductor Solutions Division. Shun joined HP in 1984 after graduating from Shinshu University with a BSEE degree with an emphasis in semiconductors. He was born in Toyohashi, Aichi Prefecture, Japan. He is married and enjoys windsurfing, playing soccer, and skiing.

Antoni C. Niedzwiecki

An R&D section manager at HP's Communication Semiconductor Solutions Division, Tony Niedzwiecki is responsible for device characterization and packaging. Tony joined HP in 1991. He received an MSEE degree from the University of Technology, in Warsaw, Poland in 1971. He was born in Warsaw, Poland and is married and has two children.

Figure 1

A DBS receiver system.

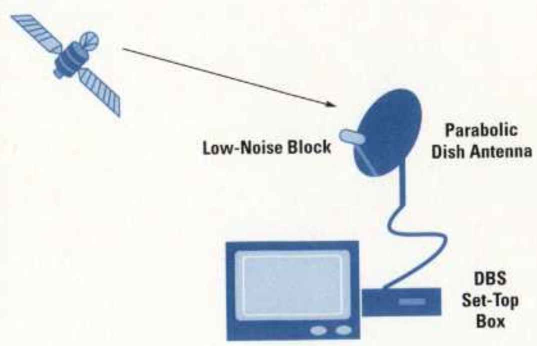
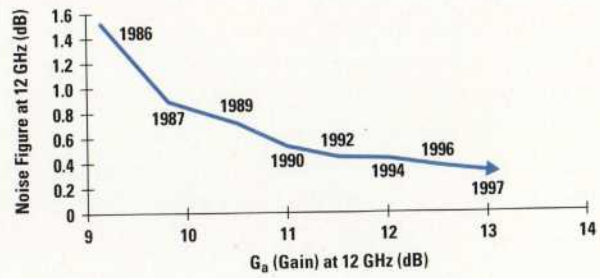


Figure 2

GaAs transistor performance for DBS applications.



Low-Noise Block

A DBS receiver system consists of two main components: a parabolic dish antenna with a low-noise block (LNB) downconverter and a digital set-top box (see **Figure 1**). The system receives digitally modulated video programs from high-powered Ku-band* transponders situated on satellites in a geosynchronous orbit (where the satellite is always in the same position relative to a position on earth). Using Ku-band carriers, the DBS system requires a much smaller dish antenna (18 inches or smaller) than C-band systems.

A key device that enables the use of a small-size dish is a low-noise GaAs transistor in the low-noise block. Lowering the noise figure improves the carrier-to-noise ratio for a given size of dish antenna, while still maintaining the signal quality. Therefore, the history of developing low-noise block downconverters is the same as improving the noise figure of GaAs transistors.

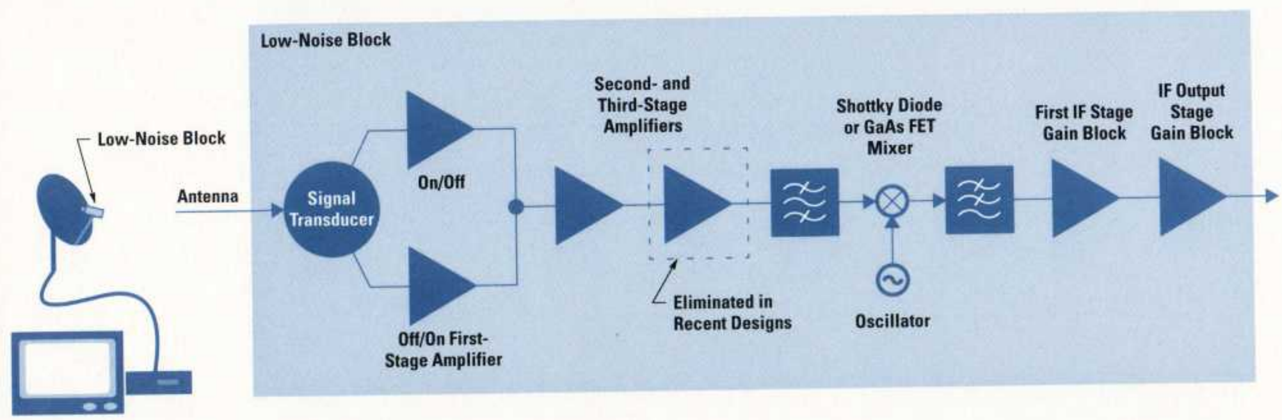
Typical GaAs FETs offered a noise figure of 1.2 dB to 1.5 dB for the European and Japanese DBS systems in 1985. Now, GaAs PHEMT (pseudomorphic high-electron-mobility transistor) technology has been pushing the noise figure of commercial PHEMTs down to the 0.4 dB to 0.5 dB level at 12 GHz. **Figure 2** shows these performance trends.

A typical low-noise block consists of a two- or three-stage low-noise amplifier, a downconverter mixer, a local oscillator, and a two-stage IF amplifier (see **Figure 3**). The first-stage low-noise amplifier sets the system noise figure. Therefore, the transistor with the lowest noise figure

* Ku-band broadcasts are between 10.95 GHz and 12.75 GHz.

Figure 3

A DBS low-noise block downconverter.



should be selected. The second and third stages should provide an appropriate gain as well as a suitable noise figure to meet the system specifications. Because of strong competitive pricing pressures, low-noise block manufacturers are forced to reduce costs each year, while at the same time improve performance. More recently, a radio architecture has eliminated the third stage low-noise amplifier and using an active mixer has become the optimal solution for lowering cost. However, the noise figure and gain per stage in the low-noise amplifier chain have become more critical.

Noise Figure and Packaging

GaAs transistor suppliers have been improving the noise figures every year by modifying the device structure and fine-tuning the lithograph technology. However, GaAs material technology appears to be approaching its limits. Although InP-based devices might be required for further noise-figure improvements, this technology has higher costs. Conventional packages for these low-noise transistors are ceramic-based microstrip packages, and the cost of these ceramic packages appears to be approaching bottom. The obvious choice is a plastic package such as the SC70 (SOT-343). However, plastic material degrades the noise figure and the G_a (associated gain) at high frequencies. The plastic package must be designed to optimize microwave performance and keep the cost low by using standard plastic package manufacturing processes. Manufacturing capacity is an important consideration to ensure that the peak demand rate can be met.

HP's solution to this manufacturing concern is based on two research-and-development projects that have the following objectives:

- Develop an AlInAs/GaInAs device fabricated on a conventional GaAs wafer to achieve a lower noise figure and much higher gain at lower cost
- Develop a surface mount plastic package optimized for microwave frequency, which is to be manufactured by standard molding processes and tested by high-volume testers (see "Packaging" next page).

A transistor on the wafer made with the LGLTBL (linearly graded low-temperature buffer layer) technology typically provides a 0.10- to 0.15-dB lower noise figure and a 2-dB higher associated gain than the best GaAs PHEMT. We have developed a device, which is packaged in a plastic surface-mount package, that shows a noise figure of 0.5 dB and a G_a of 13.5 dB. This noise figure is close to the top level, and the G_a is the highest on the performance mapping.

A higher G_a effectively contributes to reducing the total system noise figure, especially for a two-stage low-noise amplifier system (assuming that the third-stage amplifier is eliminated to keep costs down). In this regard, HP provides better performance at a lower cost for the emerging DBS market. Furthermore, the high-volume production will offer a low-cost manufacturing platform for future Ka-band and millimeter-wave applications.

* Ka-band includes frequencies between 18 and 31 GHz.

Packaging

At microwave frequencies, even a very well-designed and well-built package can have a detrimental effect on the performance of the semiconductor device it houses. On the other hand, a package makes the semiconductor device accessible and easier to handle, and provides some degree of mechanical and environmental protection.

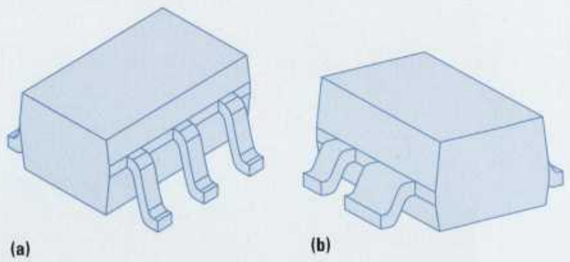
In the wireless market, the choice of a package is very important in terms of the cost and performance trade-offs that must be made for a specific semiconductor product. Ceramic packages, which have been used in industry for years, are prohibitively expensive. There has been an increasing number of plastic packages used at RF and microwave frequencies over the past few years. Considerable effort is being invested in developing electrical and thermal models of plastic packages.

When choosing or designing a package for a semiconductor device, one must consider a number of factors. Package size affects not only the amount of required circuit-board space, but also parasitics that are associated with wire bonds and lead lengths. The shape of the package body and the arrangement of its leads determines how quickly and at what cost the finished product can be tested and packed for shipment. Finally, the package affects assembly choices customers must make to use the product.

Hewlett-Packard's low-noise transistors that are destined for direct broadcast satellites (DBS) and wireless communications markets are packaged in plastic packages belonging to the SC-70 family, which has body dimensions of $2 \times 1.3 \times 0.9$ mm. Six- and four-lead SC-70 configurations designated as SOT-363 and SOT-343 are currently used for 12-GHz low-noise devices (See Fig. 1). The SOT-363 package was built first and represents a standard approach to the SC-70 type package design.

Figure 1

(a) SOT-363 package. (b) SOT-343 package.



This package is used for a wide range of RF products. The SOT-343 package was developed specifically for high-frequency, low-noise transistors. The leadframe design, lead-forming profile, and molding-compound selection were optimized to improve the high-frequency performance of the package. The noise contribution of the SOT-343 package at 12 GHz is only half that of the SOT-363 package.

Low-noise transistors in the SOT-363 and SOT-343 packages are tested on fully automated test systems. Each transistor is placed in a test circuit, and its input and output are loaded with specific impedances at the test frequency of 12 GHz. Noise figure, gain, and gate current are measured for each device. Special compliant contacts are used to achieve repeatable connections between transistors and test boards without causing package-lead distortions.

Pager Testing with a Specially Equipped Signal Generator

Matthew W. Bellis

This paper reviews current trends in the paging industry, describes typical pager designs, presents the test requirements of modern pagers, and discusses the contribution to pager testing of the HP 8648A signal generator with Option 1EP, the pager signaling option.

Today there are over 88 million subscribers of paging services throughout the world. By the year 2000, the number of subscribers is expected to grow to over 140 million. To meet this demand, pager manufacturers will need to produce over 30 million pagers per year. This makes pagers one of the leading RF devices in production today.

The design and testing of pagers from concept to production requires sophisiticated test equipment. In addition to the typical RF measurements, a paging signal is required to test the finished product. This article will review modern paging formats, typical pager designs, and methods for testing pagers.

Paging Review

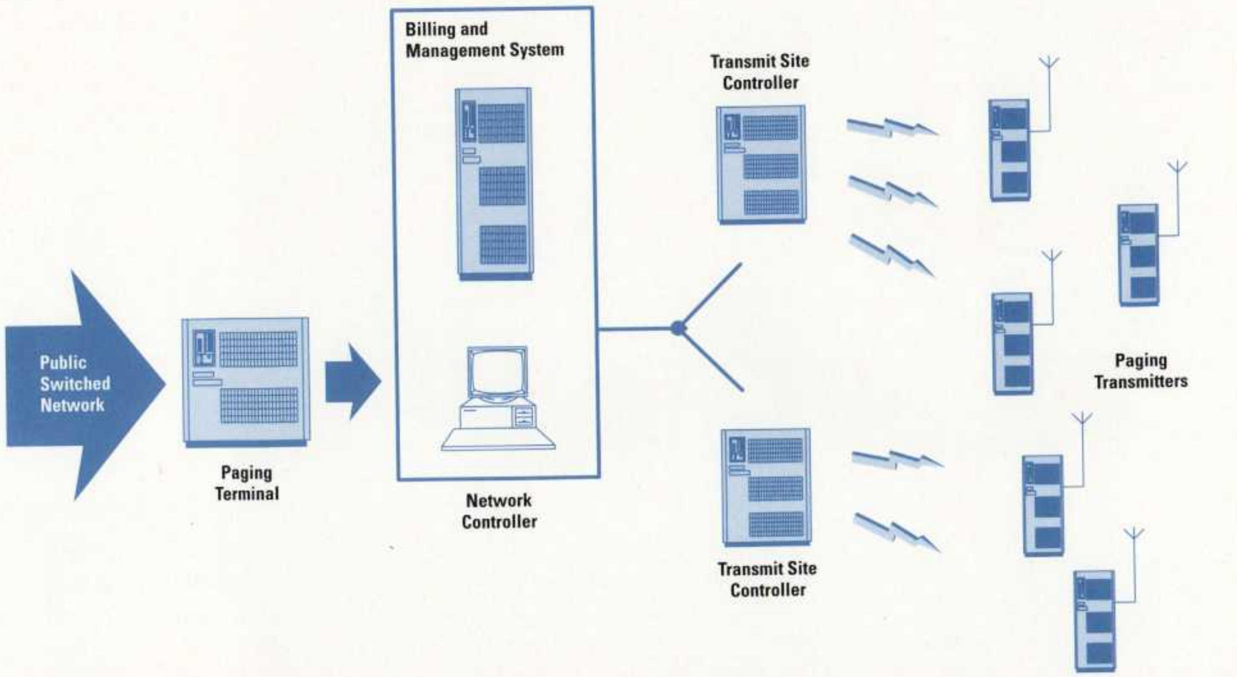
A paging network (see **Figure 1**) begins at the connection to the public switched network or telephone lines. The caller can initiate a page through voice mail or a paging operator, or can leave a message by entering touchtones from a telephone. Pages are assembled in the paging terminal and sent to the network controller, where they are combined into batches based on their final destination. Billing and management are also controlled at this point.

Many paging companies cover more than just one geographical area. For example, the company may serve an entire state or even country. The network controller specifies the site controllers for which the batched messages are intended and sends them out. Each site, covering a particular geographical location, can contain one or more paging transmitters. Once the site controller receives the batch of pages, it uplinks them to the paging transmitters, which

Matthew W. Bellis
Matt Bellis is the U.S. marketing manager for the HP Kobe Instrument Division. He has been with HP for five years, during which he has focused primarily on wireless system and wireless component testing. Before coming to HP he worked as an R&D engineer in the aerospace industry. He has an MSEE degree from the University of California at Santa Barbara and BA degree in physics from Northwestern University.

Figure 1

A paging network.



then transmit the batch of pages at the same time on the same frequency using a simulcast technique. Simulcast means that two or more transmitters are used to transmit identical information at the same time. This allows the system to provide seamless coverage on a single frequency.

Many different types of paging formats are currently in use. The most common and only worldwide standard at this time is POCSAG (Post Office Code Standardization Advisory Group), also known as RPC1. The digital formats currently under development are FLEX (a Motorola trademark), ERMES (European Radio Message System), and FLEX-TD (RCR-43). FLEX is receiving worldwide attention and has been implemented in North America, China, Indonesia, Singapore, and Thailand. ERMES is intended mainly for use in Europe. Japan has developed its own paging format, FLEX-TD, which is based on FLEX.

Table I shows a comparison of some of the main characteristics of current paging formats. A bit more detail regarding the protocol of each format is given in the following paragraphs. Probably the most important reason for the creation of the newer digital formats is the ability to send more information, faster, to more people.

POCSAG

POCSAG, Post Office Code Standardization Advisory Group, is a digital paging scheme that was developed by British Telecom to provide a standard signaling format for the United Kingdom. POCSAG is the most common paging format in use today.

POCSAG is an asynchronous paging scheme with a preamble (see **Figure 2**). Paging base stations transmit pages simultaneously within a service area. Messages are sent in batches and each pager address is located in a specific frame within a batch. Within the assigned frame, the pager will look for a message. Messages can exceed one frame, in which case the message is continued in the following frame. Either five numeric or 2.8 alpha characters can be sent in a message codeword. This means that for a numeric pager to receive a caller's 10-digit phone number, two message codewords must be sent following the pager's address codeword.

Table I

Characteristics of Current Paging Formats

Protocol	Data Rate (bits/s, forward/reverse)	Channel (kHz, forward/reverse)	Modulation	Timing	Applications
POCSAG	2400	25	2-FSK	Asynchronous	One-way
FLEX	6400	25	2- or 4-FSK	Synchronous	One-way
ReFLEX-25	6400/9600	25/12.5	2- or 4-FSK	Synchronous	Two-way
ReFLEX-50	6400/9600	50/12.5	2- or 4-FSK	Synchronous	Two-way
InFLEXion	6400/9600	50/12.5	2- or 4-FSK	Synchronous	Two-way and voice messaging
FLEX-TD	6400	25	2- or 4-FSK	Synchronous	One-way, Japanese version of FLEX
ERMES	6250		4-FSK	Synchronous	One-way, European standard
pACT				Synchronous	Two-way

The pager, after detecting the preamble and synchronizing to the frame sync word, turns off its receiver circuits until the proper frame appears—this increases the life of the pager battery. POCSAG pagers support tone, numeric, and alphanumeric messages. Voice is not supported.

POCSAG uses frequency-shift keying (FSK), a simple form of digital modulation. FSK offers good noise immunity and is inexpensive to generate. However, FSK is not suitable for high-data-rate transmission.

Figure 2

POCSAG paging format.

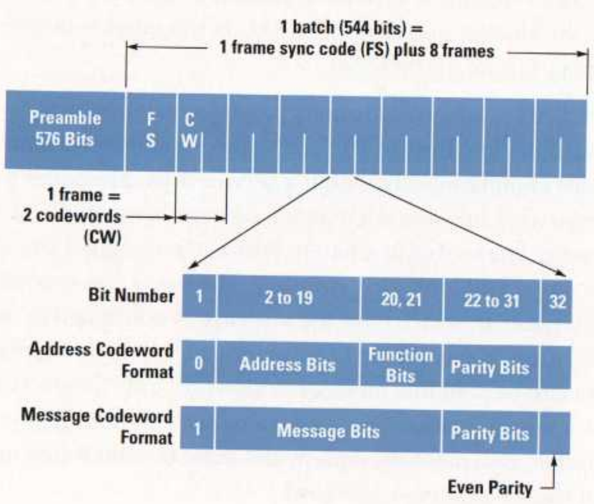


Table II shows typical specifications for a POCSAG pager. These specifications are warranted by the manufacturer and many are verified during production of the pager. Several of these specifications are indirectly verified or are verified by design.

Table II

Typical Specifications for a POCSAG Pager

Code Format	POCSAG
Channel Spacing	25 kHz
Frequency Deviation	$\pm$4.5 kHz
Paging Sensitivity	5 μV/m for 1200 bits/s
Spurious and Image Rejection	60 dBc
EIA Selectivity	65 dB at $\pm$25 kHz
Frequency Stability	$\pm$0.002% of reference
Power Consumption	3.75 mA (0.383 mA standby) for 1200 bits/s
Battery Life	>2100 hours (1200 bits/s)

FLEX Family

In many areas there is growing demand for increased capacity because existing systems are reaching their maximum capacity. There is also a demand for more sophisticated messaging capability. To meet this demand, Motorola introduced the FLEX family of paging protocols. The FLEX family consists of an enhanced one-way paging protocol, FLEX, and three two-way paging protocols, ReFLEX-25, ReFLEX-50, and InFLEXion. In addition to two-way messaging, InFLEXion also supports voice messaging.

FLEX is a synchronous protocol that assigns pagers to frames (see **Figure 3**). These frames occur at a rate of 128 every four minutes. One FLEX cycle equals 128 frames. There are 15 FLEX cycles per hour. Each pager is assigned to look at a particular frame (or frames) within a FLEX cycle. Upon power-on, the pager synchronizes with the FLEX network. During synchronization, network timing information is passed to the pager. The synchronization information is used to determine the pager's assigned frames. Thereafter, the pager will only turn on its receiver circuitry during the assigned frames. This feature of FLEX results in improved pager battery life since it eliminates the need for preamble detection. The frame assignment for a pager is configured in the pager or determined from information within the paging signal.

Each frame (**Figure 4**) is 1.875 seconds long and consists of a sync field and 11 blocks. The sync information takes up 115 ms and is composed of three parts: sync 1, frame information, and sync 2. The first two parts are always transmitted at 1600 bits/s using 2-level FSK. Sync 1 provides timing information and an indication of the speed of the rest of the frame. The frame information word contains the frame number (0 to 127), the cycle number (0 to 14), and other information. Sync 2 provides synchronization at the frame's block speed so the remainder of the frame can be properly decoded.

The information that is transmitted within a FLEX frame is shown in **Figure 4**. The fields are not constrained by block boundaries. The fields that are transmitted in a frame are a block information field, an address field, a vector field, a message field, and idle blocks.

The block information field is typically one word long and contains information that indicates where in the frame the address and vector fields start, frame information from the system as discussed above, and the number of priority addresses placed at the beginning of the address field. Priority addresses are assigned to pages with an urgent priority.

The system collapse value is also contained in the block information field. The collapse concept enables service providers to better manage their paging networks. The pager's assigned frame is represented by a seven-bit word (any number from zero to 127 can be represented by a seven-bit binary word). The system collapse value, or collapse cycle, instructs pagers to mask one or more of the address bits. For example, if the collapse cycle is set to three, the pager will mask all but the three least-significant bits of its assigned frame and the current frame being transmitted. When the collapse cycle is three, the pager will monitor one out of eight frames instead of one out of 128 frames. Thus, the 128-frame cycle (2^7) is collapsed to an eight-frame cycle (2^3). In addition to the system collapse cycle, the pager has a default collapse cycle. During operation, the collapse cycle in effect will

Figure 3

FLEX paging format cycle.

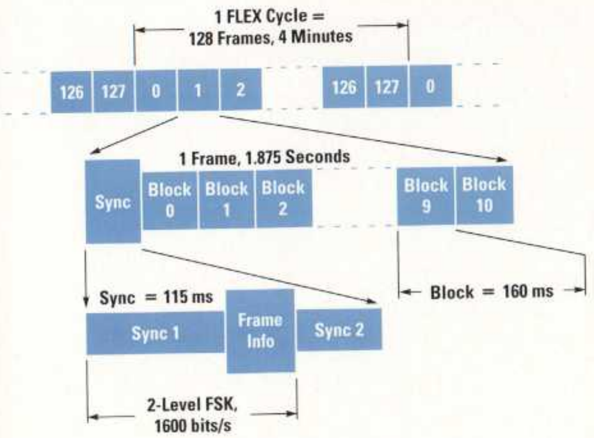
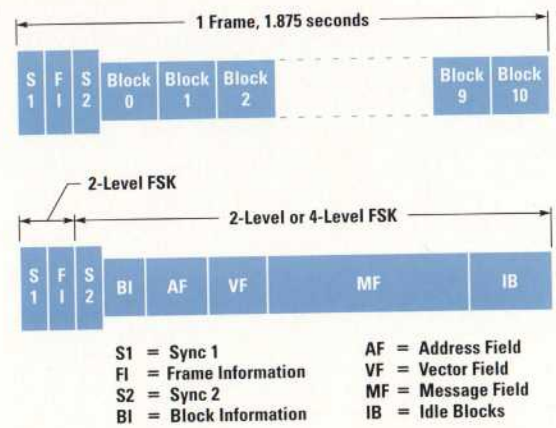


Figure 4

FLEX frame.



be the lesser of the system collapse cycle and the pager collapse cycle.

The address field contains addresses of the pagers that are being paged during this frame time. The vector field has a one-to-one relationship with the address field, so the pager knows where to look for the vector. The vector points to the start word of the message and indicates the length of the message, again so the pager knows where its message occurs. The vector field also indicates the type of message. FLEX supports several types of messages (see **Table III**). The type of message indicates how the bits within the message field should be grouped and decoded to recover the information being transmitted.

The message field follows the vector field. Any unused blocks in a frame are filled with bits representing idle: alternating 1s and 0s at 1600 bits/s. At higher speeds, the symbols, adjusted for the speed and modulation format, must represent that same pattern.

The FLEX protocol can be transmitted at three different speeds: 1600, 3200, or 6400 bits per second. FLEX pagers can operate at any of the three speeds. FLEX pagers automatically decode the correct signaling speed based on information from the paging signal (located in the sync word).

FLEX uses two-level and four-level FSK to achieve the various data rates. **Figure 5** indicates the frequency deviations and bit assignments. Four-level FSK, a more complex modulation format, enables FLEX to operate at higher data rates.

Some typical test specifications are listed in **Table IV**. Many of these specifications are verified during production of the pagers or are verified by design.

Table III

Common FLEX Message Types

Vector Type	Description
Numeric Vector	The received message should be displayed as a number.
Numeric Vector with Format (Special)	Special formatting should be applied to the message. For example, the parentheses and dashes normally displayed in a telephone number are added by the pager and not sent over the channel. This saves one message word.
Alphanumeric Vector	The received message should be displayed as alphanumeric.
Hex/Binary Vector	Starting with the third word in the message field, each four-bit field represents one of sixteen combinations. This vector type is selected when transmitting Chinese characters. Several four-bit fields are combined to represent a character.
Numeric Vector with Message Number	In addition to receiving a message, a number is assigned to the transmitted message. This number can be used by the pager (or subscriber) to identify missed messages.

Table IV

Typical Specifications for a FLEX Pager

Code Format	FLEX
Channel Spacing	25 kHz
Frequency Deviation	± 4.8 kHz
Paging Sensitivity	20 μV/m for 6400 bits/s
Spurious and Image Rejection	50 dBc
EIA Selectivity	55 dB at ± 15 kHz
Frequency Stability	± 0.02 ppm of reference
Battery Life	> 10,000 hours (6400 bits/s)

Typical Pager Design

The top-level block diagram of a pager is fairly simple (see **Figure 6**). Pagers have receiving circuitry, digital decoding circuitry, and the devices that alert the subscriber and display a message.

The receiving section demodulates the data from the RF carrier and passes the data to the decoding section. The decoder then looks at the data and decides whether the information being received is for the subscriber. If so, that information is displayed on the pager and the subscriber is alerted by the selected means: beep, vibration, or flashing light.

Figure 5

Two-level and four-level frequency-shift keying (FSK) in the FLEX protocol.

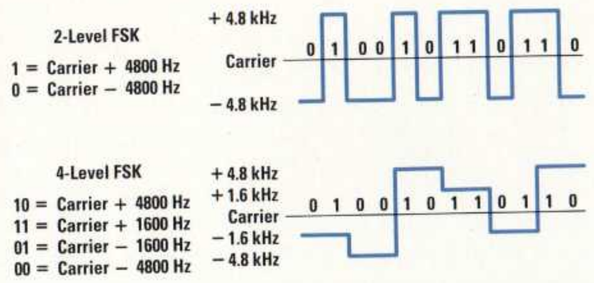
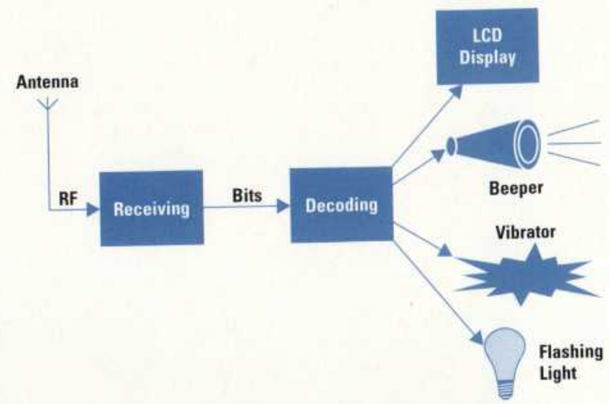


Figure 6

Top-level block diagram of a pager.



Throughout the design process, testing of each section of the pager is required. In addition, designs must be able to withstand reasonable manufacturing tolerances and still meet specifications. During production, testing is done primarily to ensure functional operation and compliance with sensitivity specifications.

Most pager receivers are double-downconversion superheterodyne receivers, as shown in **Figure 7**. A loop antenna is used to couple the incoming RF. The coupling is primarily magnetic. The incoming RF signal is filtered to remove unwanted out-of-band signals. Two IF filters are employed, but the second IF filter often determines the selectivity of the receiver. After the second IF filter, an FM detector extracts the FSK signal. An analog-to-digital converter (ADC) is used to convert the analog waveform to a series of digital words that are processed to recover the information. Often, the RF antenna, the preselector, the first mixer, and the first local oscillator (LO) contain variable components that must be adjusted during production. Components used after the first mixer are generally fixed (e.g., ceramic filters, demodulators) and do not require adjustment. Most manufacturing testing is done because of the adjustable components that are in the pager receiver section.

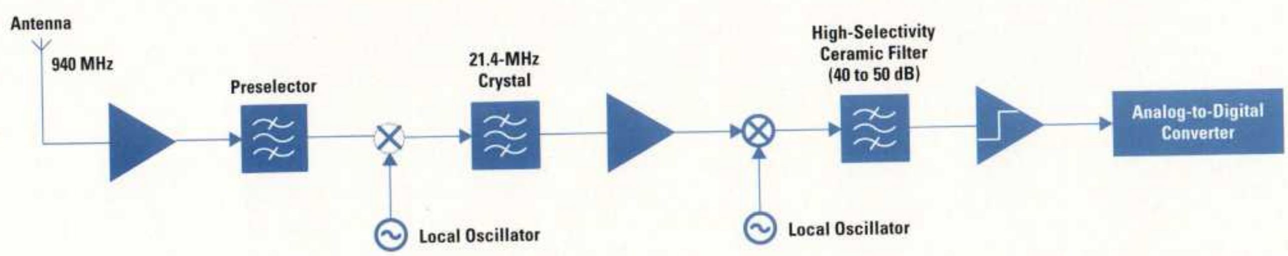
This type of design produces highly sensitive pagers whose specifications can be maintained during high-volume production. However, high levels of integration are difficult to achieve with the above design. Recently, receiver architectures using image-reject mixers and direct downconversion have been introduced with performance equal to double-downconversion receivers. Because these designs use fewer filters, oscillators, and mixers, higher levels of integration are more easily achieved.

Pager Testing with the HP 8648A Signal Generator

The HP 8648A Option 1EP synthesized signal generator incorporates the necessary protocol to test POCSAG,

Figure 7

Double-downconversion superheterodyne pager receiver.



FLEX, and FLEX-TD pagers. In addition, the HP 8648A includes an arbitrary message that can be used to create user-defined POCSAG bursts or FLEX and FLEX-TD frames or cycles. The use of the arbitrary message allows sophisticated users access to all of the features supported by POCSAG, FLEX, and FLEX-TD.

Pager Tests

A variety of testing is done during the design and production of pagers. The following list, although not exhaustive, covers much of the testing:

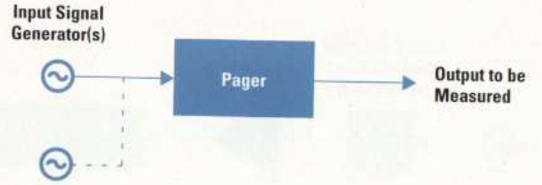
- Preselector alignment
- Oscillator tuning
- Sensitivity
- Adjacent channel selectivity
- Functional.

The first four tests are primarily RF tests. Functional testing of the pager is a system-level test. All testing is done with the same basic test setup (see **Figure 8**). A pager is placed in some type of RF isolation, one or more inputs signals are applied, and the parameter of interest is measured.

Preselector Alignment. Many pager designs use adjustable components in the preselector section of the receiver. The receiver of a popular POCSAG pager, for example, has one tunable capacitor (2 pF to 10 pF) and three tunable inductors (two 3-turn and one 17-turn). These components are adjusted during testing. For example, if the pager is designed to operate at 929.1125 MHz with a 25-kHz channel width, the components in the preselector will be adjusted to filter out frequencies outside of this channel.

Figure 8

Generic pager test setup.



For this test, a signal generator outputs a paging signal (for example, a four-level FSK signal for a FLEX pager) at the channel frequency (for example, 929.1125 MHz). The adjustable components—the variable capacitor and inductors—are adjusted to maximize the received signal strength after the amplifier. Many pager designs have a test point where a voltage level that is proportional to the received signal strength can be measured. For these designs a voltmeter can be used to verify that the preselector has been properly aligned.

Often during design a spectrum analyzer is used to ensure that the spectrum of the signal is properly passed. The received signal strength can be viewed directly on the spectrum analyzer by adjusting the center frequency of the spectrum analyzer to the desired channel (for example, 929.1125 MHz) and setting the span to at least the channel width (for example, 25 kHz). The spectrum analyzer also displays any asymmetry in the preselector filter shape. The amplitude levels of the FSK sidelobes should be of equal height. Sidelobes that are of unequal height indicate that the preselector filter shape needs to be adjusted.

Oscillator Tuning. After adjusting the preselector to the appropriate channel, the local oscillator (LO) needs to be adjusted to downconvert the RF signal to the correct intermediate frequency (IF). Many pager designs incorporate a variable inductor into the oscillator design. A simple method of tuning is to connect a frequency counter at the IF output of the first mixer and adjust the oscillator value until the counter reads the proper IF frequency. The signal generator must provide a sine wave at the required channel frequency (for example, 929.1125 MHz).

This is an inexpensive solution. However, the counter requires a filter to remove unwanted mixing products and generally requires a preamplifier to raise the signal level to the detection threshold. For designs that use a fixed IF filter, the measurement can be made at the IF filter output. This eliminates the need for an external filter at the counter input. A spectrum analyzer can be used instead of a frequency counter. The spectrum analyzer is frequency selective and has greater sensitivity. This eliminates the need for an external filter and a preamplifier.

Some manufacturers use laser trimming to adjust the LO frequency. In laser trimming, a YAG laser is aimed at a laser trimmable component (usually a capacitor). The capacitor is generally a multilayer ceramic capacitor. The energy from the YAG laser removes metal and ceramic material from the capacitor. The removal of material changes the value of the capacitor and thus changes the LO frequency.

Laser trimming systems require an automated system that shuts off the laser at the proper time. In an automated system, an input signal is applied and the LO is tuned while the received signal strength at the first IF is monitored. When the received signal strength reaches the specified level, the YAG laser is shut down.

Sensitivity. Once the pager has been properly aligned and tuned, the sensitivity of the pager can be measured. Perhaps the most important specification for a pager is the receiver sensitivity. The receiver sensitivity determines the ability of the pager to receive low-level signals properly. A pager with poor sensitivity will not detect messages reliably, resulting in unhappy customers.

Sensitivity specifications are usually in microvolts per meter (μV/m). A sensitivity measurement must be made with a calibrated, known field strength. To achieve this, the pager is placed in an RF isolation enclosure, typically a TEM cell,* screen room, or isolation chamber. A signal generator is attached to the enclosure. The field strength generated inside the enclosure should be uniform. Bit error rate (BER) is usually the performance measure for receivers of digitally modulated signals. The sensitivity of a pager is formally defined as the minimum signal level that produces a specified BER. In practice, however, less cumbersome measurements are often substituted for BER.

The 9-of-10 method is a common technique for measuring sensitivity. In this method, the signal generator output is set to the sensitivity of the pager. Ten pages are sent. To pass, at least nine of the ten pages should be received. When using a TEM cell, the pager should be oriented to achieve the maximum sensitivity reading. Other techniques involve varying the orientation of the pager with respect to the incoming paging signal.

A second technique, known as the 3/20 method, is more involved. In the 3/20 method, the pager is placed inside a TEM cell. The measurement begins with the pager placed in the upright position inside the TEM cell. A total of eight measurements are made, with the pager rotated 45 degrees for each measurement. The sensitivity of the pager, in μV/m, is recorded at each orientation.

To obtain the pager sensitivity at each orientation, three consecutive pages are sent. If the pager responds to all three, the output power is reduced by 1 dB and three more pages are sent. The output power is gradually reduced to the lowest level that triggers a response on each of the three successive pages. The output field strength is recorded as E_{3p}.

The output power is then further reduced by 1 dB, and twenty pages are sent. If three successive pages are received, the output power is reduced by 1 dB. If no pages are received, the output power is increased by 1 dB. Otherwise, the output power remains the same. The output field strength is recorded as E_{20p}.

The pager sensitivity at each orientation is the average of E3p and E20p:

$$E_n = \frac{E_{3p} + E_{20p}}{2},$$

where n = 1, 2, ..., 8 identifies the orientation. The overall pager sensitivity is calculated as:

$$E_{sens} = \sqrt{\frac{8}{\sum_{n=1}^{8} 1/E_n^2}}.$$

Adjacent Channel Selectivity. Most pager designs have tight requirements, on the order of 65 dB, for adjacent channel selectivity. The adjacent channel selectivity of a pager receiver is a measure of the receiver's ability to receive a modulated input signal on its assigned channel frequency in the presence of a second modulated signal in the adjacent channel (±25 kHz). Again, the standard performance measure is BER. Two signal generators are needed to measure adjacent channel selectivity accurately. One signal generator provides the in-channel paging signal, while the second provides the interfering signal. The level of the paging signal is generally set to a level above the pager's sensitivity (usually 3 dB) and the level of the interfering signal is increased until the specified BER is achieved (usually the same BER specified for the sensitivity test). The level difference between the two signals is then recorded as the adjacent channel selectivity.

* A TEM cell provides electromagnetic isolation. Radiation propagating within the cell is largely confined to the transverse electromagnetic (TEM) mode.

In practice, BER measurements are generally not done. The following test method is often used. For a pager receiver that has been tuned to receive 929.1125 MHz, the signal generator carrier frequency is set to 929.1125 MHz. The output of the signal generator is set to produce a paging signal. The amplitude of the signal generator is set to 3 dB above the sensitivity of the pager. The frequency of the interfering signal generator is set to an adjacent channel, 929.1375 MHz. The interfering signal is FSK modulated with the proper deviation. The strength of the interfering signal is set below the adjacent channel selectivity specification and is increased until the (modified) sensitivity of the pager (under the current conditions) is reached. The difference between the paging signal level and the interfering signal level is the adjacent channel selectivity.

The phase noise performance of the interfering signal can greatly affect the measurement. The following equation is used to determine the acceptable level of phase noise:

$$\text{Signal Generator Single-Sideband Phase Noise} < -(\text{Adjacent Channel Selectivity}) - (\text{Noise/Hz}) - \text{Margin}.$$

For a specification of 65 dB, a channel width of 25 kHz, and a margin of 10 dB, the single-sideband phase noise at the channel offset of 25 kHz must be less than or equal to $-65\text{ dB} - 10\log(25\text{ kHz}/1\text{ Hz}) - 10\text{ dB} = -119\text{ dBc/Hz}$.

The HP 8648A signal generator's phase noise at a 20-kHz offset for a 1-GHz carrier is -116 dBc/Hz. The HP 8657A/B signal generator's phase noise at 20-kHz offset for a 1-GHz carrier is approximately -130 dBc/Hz. For adjacent channel selectivity measurements on pagers, the HP 8657A/B provides a better phase noise margin. However, the HP 8648A provides a cleaner solution for generating an FSK signal.

In the previous calculation, a 10-dB margin was used to determine the acceptable level of phase noise for a signal source. With a 10-dB margin, the single-sideband phase noise of the signal generator will add 0.4 dB of error to the adjacent channel selectivity measurement. The following table indicates the error contribution for other margins.

Margin (dB)	0	1	2	3	4	5	10
Error contribution (dB)	3	2.5	2.1	1.8	1.5	1.2	0.4

Functional Test. Functional test requires that an actual page be sent. To do this, the test signal must contain the proper protocol. Because of the need for pager testing, the HP 8648A option 1EP incorporates much of the protocol for the POCSAG, FLEX, and FLEX-TD paging formats.

In general, protocol is not tested at the system level. However, when the protocol affects the RF performance, the interaction of the protocol with the RF section should be tested and verified. The following tests verify this interaction:

- Receiving a message
- Reset
- Resynchronization
- Roaming
- Receiving a message.

During production, the functional test of a pager is simple: Does the pager respond? The requirements for production testing of POCSAG and FLEX pagers are generally few. The pager should respond to a message, all of the alert methods (sound, light, vibration) should work, and the entire display should work.

Sending a message sounds like a relatively simple task. However, messages can be formatted in numerous ways. Most manufacturers of pagers produce a variety of pagers for a variety of markets. To ensure that all needs are met, the HP 8648A option 1EP supports the following:

- POCSAG
  - Numeric
  - Alphanumeric
  - Hex/Binary
  - 14-Bit and 16-Bit Chinese Characters
- FLEX, FLEX-TD
  - Numeric
  - Numeric with Format (Special)
  - Alphanumeric
  - Hex/Binary (Used for Chinese Characters)
  - Numeric with Message Numbering.

When testing FLEX pagers, synchronization must be provided by the test signal. In addition, after the test signal is removed, it is often desirable to reset the pager to remove the timing synchronization. The HP 8648A can be configured to reset the pager.

On occasion, the entire FLEX network may lose timing synchronization. When this happens, all of the pagers on the network must be resynchronized. A resynchroniztion signal is sent by the network to resynchronize the pagers. During design, the ability of the pager to do this should be tested.

The ability to support roaming is a major contribution of FLEX. The FLEX paging protocol defines two methods for roaming: simulcast system identification (SSID) and network identification (NID).

With SSID, a list of simulcast areas is programmed into the pager. When the pager encounters one of these areas, the pager is instructed when to look (which frame) and where to look (which channel) to receive pages. The SSID list contains codes that identify the simulcast areas. In addition, a channel scan list with the information required to find and identify each simulcast system is stored. For a pager to fully support SSID roaming, the pager must have the ability to change channels.

NID roaming is a superset of SSID roaming. NID is supported for subscribers that may roam across national and possibly global regions. In such large areas, storage of all SSID information is impractical.

Roaming is an important functional parameter to test because roaming connects a feature of the protocol to the RF control of the receiver. Although within the firmware implementation of the protocol the proper bits may be set, the functional aspect should be tested because the change in the RF channel should be verified.

RF Isolation

RF isolation is needed when making any type of measurement on a pager. Without isolation, stray signals will be coupled into the pager by the antenna and even the printed circuit traces. Most specifications, however, are written in terms of field strengths. Therefore, when using an isolation chamber, the power levels of the test signals must be converted to field strengths:

$$E = \sqrt{P} \times R/d,$$

where E is the field strength in volts per meter, P is the power from the signal generator in watts, R is the output resistance of the signal generator, and d is the distance over which the power is radiated. In the more common units of dB above one microvolt per meter (dBμV/m):

$$E_{\mu} = 20 \log(E \times 10^6),$$

where E_{μ} is the field strength in dBμV/m.

The HP 8648A comes with an optional TEM cell that provides RF isolation and a calibrated field strength. For this TEM cell, the conversion from power level to field strength is shown in **Table V**.

Table V
HP 8648A TEM Cell Conversion Table

Signal Generator Power (dBm)	TEM Cell Field Strength (μV/m)	TEM Cell Field Strength (dBμV/m)
−120	2.91	9.29
−115	5.18	14.29
−110	9.21	19.29
−105	16.38	24.29
−100	29.14	29.29

Conclusion

This article has reviewed the current trends in the paging industry, typical pager designs, and the test requirements of modern pagers. In addition, the contribution to pager testing of the HP 8648A signal generator with Option 1EP, the pager signaling option, was discussed.

HP CaLan: A Cable System Tester that Is Accurate Even in the Presence of Ingress

Daniel D. Van Winkle

Today, cable system operators have to deal with bidirectional traffic from sources such as pay-per-view television, high-speed Internet access, and two-way telephony. A cable testing system is described that can handle bidirectional traffic even with RF noise (ingress) on the return path.

The deregulation of the telecommunications industry has resulted in an unprecedented surge of effort to implement new telecommunications services. The cable television industry, in particular, is pushing hard to provide customers with two-way telephony, fast Internet service, and interactive video programming. As they prepare to implement and maintain two-way communications over their cable networks, cable industry engineers and providers of cable TV test equipment are facing a whole new set of challenges.

This article describes the background and design of the new HP CaLan 3010H and 3010R sweep/ingress analyzer (see **Figure 1**), which enables cable television providers to do return path alignment in two-way cable systems and to troubleshoot the system quickly regardless of the presence of RF noise (ingress).

Daniel D. Van Winkle

Dan Van Winkle is a hardware design engineer at HP's Microwave Instruments Division and is currently working on the HP Broadband Series Test System. Since joining HP in 1990 he has worked as a production engineer in the modular measurement system and microelectronic departments. He received a BSEL (electronics engineering) degree in 1984 from the California Polytechnic State University in San Luis Obispo, California, and an MSEE degree in 1988 from Santa Clara University. Born in Greenbrae, California, Dan is married and has two children. In his free time he enjoys outdoors sports such as hiking, mountain biking, and skiing.

Figure 1

The sweep/ingress analyzer consists of the HP 3010H headend unit and the HP 3010R field unit.

Background

Cable systems have historically been one-directional. Signals from sources such as satellite feeds, community access programming from local studios, tape machines with prerecorded programming and commercial insertion, and various antennas for both local and distant broadcast television signals are received at a central cable office (headend). Here they are appropriately combined and rebroadcast to various branches over the cable system until they finally reach the subscriber's home.

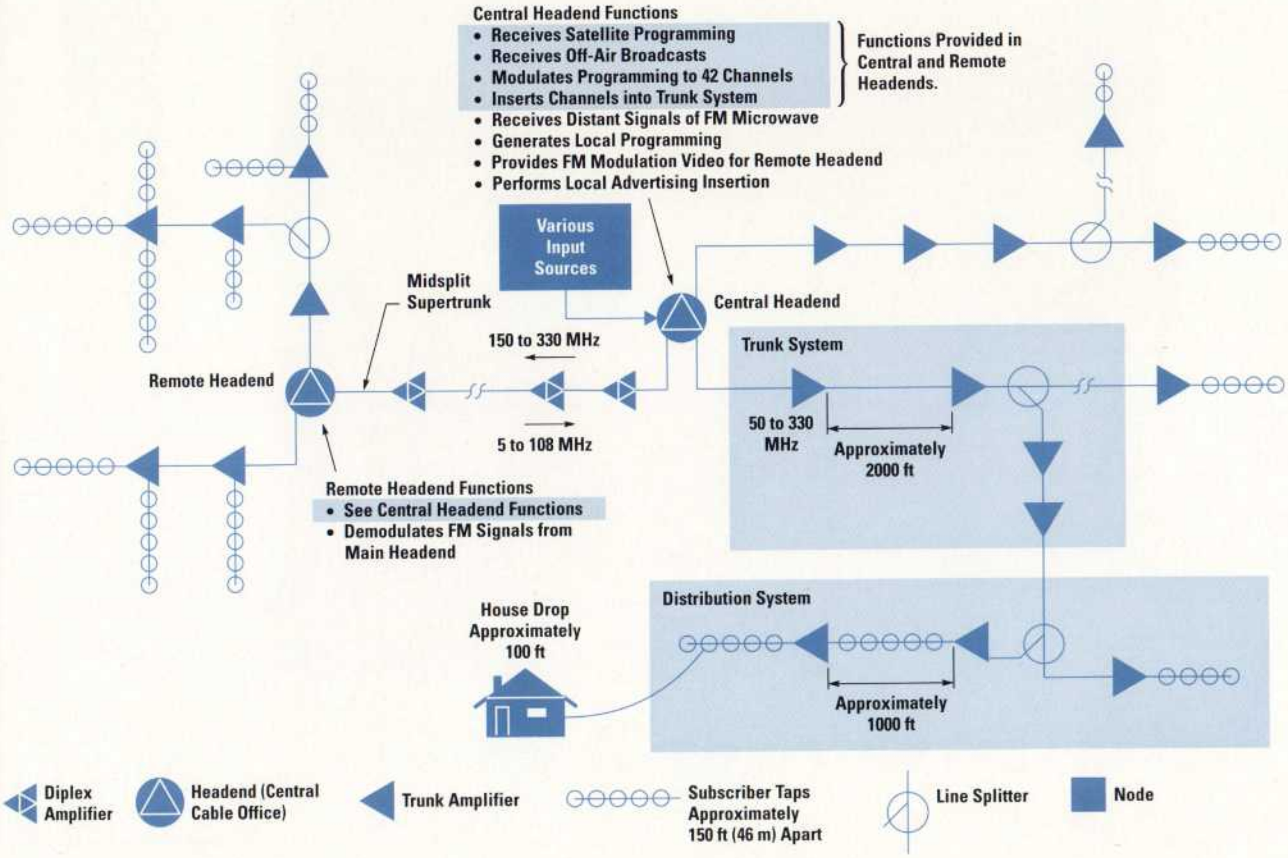
A typical cable system has trunk lines that feed bridge amplifiers. The amplifiers feed distribution lines, which in turn feed the flexible cable drops to the subscriber's home. Approximately 40% of the system's cable footage is in the distribution system and 45% is in the flexible drops to the home.[1] A cable system can be as simple as a small headend with a single trunk line and only a couple of distribution lines, or it can have multiple trunk lines, each with multiple distribution lines and possibly a remote headend connected to the central headend through a midsplit super trunk. **Figure 2** shows a portion of a typical conventional cable system.

Many cable systems are upgrading their plants* from low-frequency coaxial cable systems to hybrid fiber/coaxial cable distribution systems. In hybrid fiber/coaxial systems (also called fiber backbone systems) the system is divided into several smaller cable systems with amplifier cascades limited to four to six amplifiers. Each of the smaller cable systems is fed with a fiber link to the headend (see **Figure 3**).

* The entire cable system is referred to as the "cable plant."

Figure 2

Conventional low-frequency coaxial cable system.



Glossary

Ingress. Any undesired signals present on the cable system. These signals usually come from inside the home, but can also come from ham radio operators, car ignitions, and other sources.

Signal Level Meter (SLM). A signal level meter is used to measure signal levels in a cable system. It is typically a tuned receiver that is calibrated to measure a fairly wide dynamic range.

Forward and Return Pilot. The forward pilot is a modulated fixed-frequency carrier (set by the cable operator) that is used for forward communication in a cable system. Forward communications travel from the headend to the remote site. The return pilot is a modulated fixed-frequency carrier (set by the cable operator) that is used for return communications in a cable system. Return communications are from the remote site.

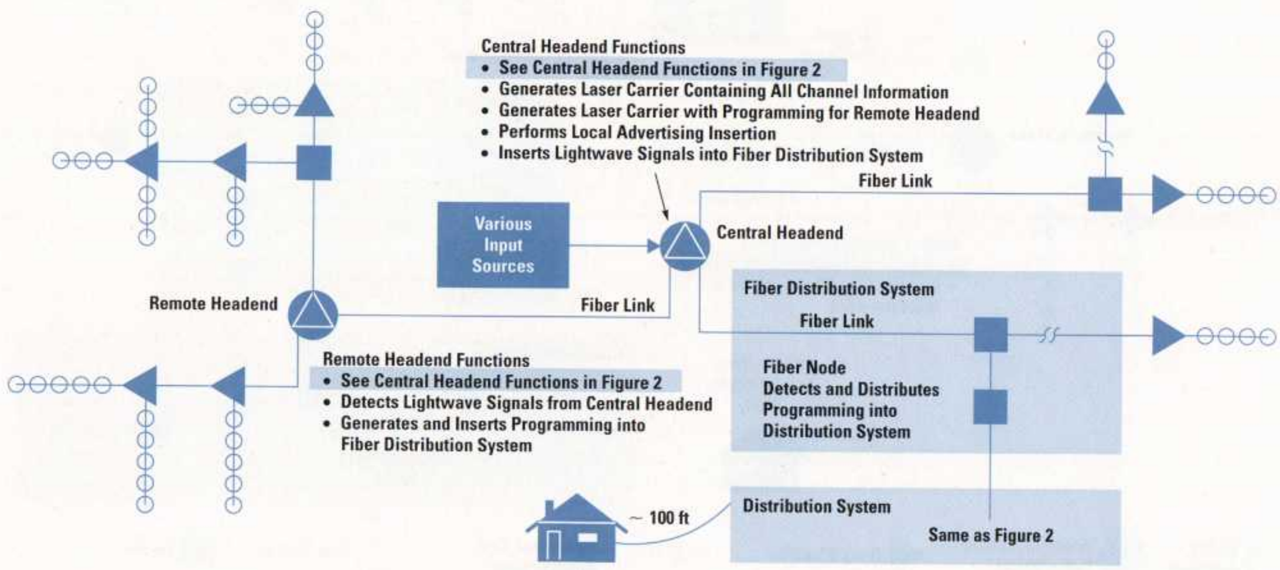
Headend. The headend in a cable system is the location where signals are gathered from various sources (off-air, community access, tape machines, and so on). The headend is also the central distribution point for the entire cable system. In upgraded two-way systems, the headend will likely contain some type of file server connected to the Internet through a T1 line.

Remote Unit. The remote unit as defined for the HP 3010H and HP 3010R sweep system is the unit that is taken into the field to do forward and return sweeping. It is handheld and battery operated.

Return Sweep. The return sweep is the magnitude response versus frequency of the entire (or partial) cascade of all the amplifiers and cable lengths in a cable system in the return direction. Return sweep (and forward sweep) are used as alignment tools for cable systems.

Figure 3

Hybrid fiber/coaxial cable distribution system.



The advantages of using hybrid fiber/coaxial systems include significantly lower vulnerability to amplifier outages, reduced bandwidth restrictions, lower noise buildup (because fewer amplifiers are in series), and greatly reduced ingress. The greatly reduced ingress makes a two-way system practical. Another major benefit of the fiber backbone approach is that implementation cost is relatively low.

Many cable operators are also upgrading their plants to be bidirectional. Not only will the system transmit signals from the headend, but signals from the subscribers' homes will also be transmitted from various in-home devices back to the headend. **Figure 4** shows a generic return path system.

The return path allows new services, such as audio and video telephony, video-on-demand, and Internet access, to be offered. As the cable operators bring up these systems, they are finding that the return path (from the home to the headend) is causing some new and interesting challenges. One-way cable systems have a tree-like structure, in that information (modulated video carriers) is fed from the root up through the trunk, into individual branches, and finally into the leaves (subscribers' homes). On the other hand, two-way systems have a river-like structure, in that every house drop is like a small creek that feeds into a larger river and then pours into a lake (headend). Any debris (noise) that comes from small creeks will also end up being dumped into the lake (headend). In a large system, this noise power can be enough to cause significant interference. Since most cable operators plan on using the low-frequency (5 to 40 MHz) end of the cable system, and since most of the noise generated in homes is in this range, careful attention must be paid to establishing and maintaining the return path. As mentioned above, a hybrid fiber/coaxial distribution system can help significantly in reducing return-path ingress.

The Development Process

The development process for the HP CaLan 3010H and 3010R sweep/ingress analyzer proceeded in three phases. First, we developed a proposal and model of an "ideal" cable testing system. Second, we created the design concept for the new analyzer by making a list of hardware and software tasks and performing a system analysis. Finally, using the output from the previous two phases, we developed the product.

The Proposed System

The HP CaLan 2010/3010 product line has historically been used by cable operators to sweep and align their forward path systems. The HP Calan 2010A was essentially a signal level meter (SLM) that allowed cable operators to measure carrier levels and perform some FCC testing. The HP 3010A (the first instrument in the HP 3010 family) added the capability to perform a forward sweep like a scalar network analyzer. For example, to perform a forward sweep, an HP 1777 integrated sweep transmitter was required in the headend (see **Figure 5a**). This system worked by stepping the HP 1777 source across the 5-MHz-to-1-GHz band at a power level below the visual and audio carriers. (Guardbanding was used to prevent the sweep source from causing unwanted interference.) The HP 1777 was pulsed for approximately 5 to 10 μs at each frequency step. The HP 3010A then widened its receiver time window to catch the signal from the HP 1777. Careful timing of the source (HP 1777) and the receiver (HP 3010R) was required to maintain synchronization. Also, because the signal from the HP 1777 passed through the entire cable system, the resultant signal level received by the HP 3010A showed the swept amplitude response of the entire cable plant from the headend to the remote site where the system was being tested.

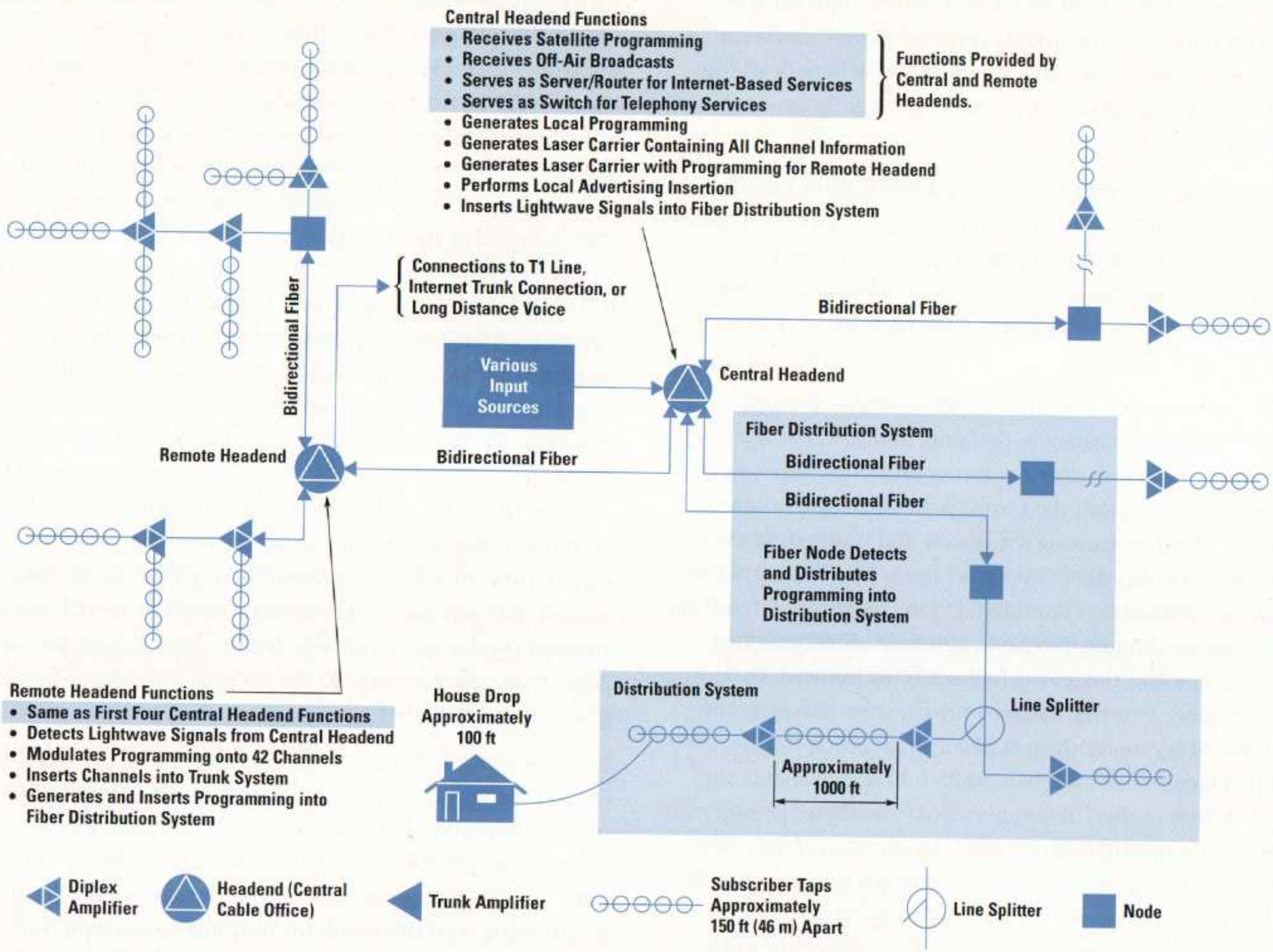
Because cable companies are working rapidly to upgrade their return path systems, an urgent measurement opportunity has emerged: *return sweep*. To perform a return sweep, the remote unit has to contain a frequency source with enough frequency range to cover the entire return bandwidth, and the headend unit has to contain a receiver. Because the existing products (HP 3010A and HP 1777) did not contain the necessary hardware, the design team proposed a scheme in which a sweep source would be added to the remote unit and a new headend unit would be created. The headend unit would have the same functionality as the remote unit but would be capable of acting as a central control point for the system by coordinating the sweeps of multiple remote units. A block diagram of this scheme is shown in **Figure 5b**.

The Design Concept

After evaluating the system shown in **Figure 5b**, the design team decided that to build these new instruments they needed to:

Figure 4

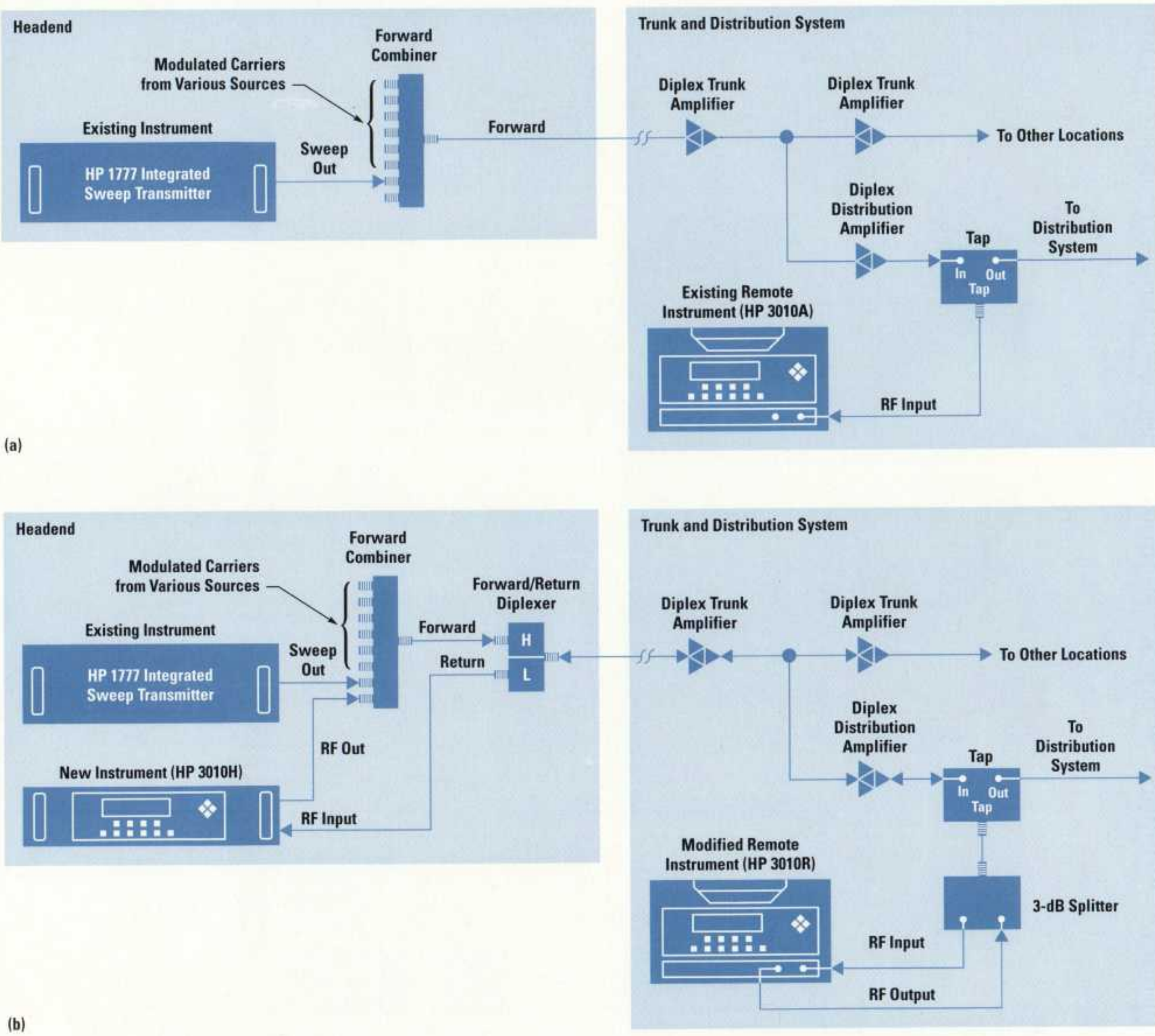
Modern two-way cable system.



- Create a sweep/ingress analyzer that includes a sweep source and maintains the functionality of the original HP 3010A
- Move the design to a process that is compatible with surface mount technology and use preferred parts where possible
- Move the design to a less expensive substrate process (The original HP 3010A RF board was Teflon laminated onto FR4.)
- Develop firmware changes to create a communications protocol that would allow communication between a remote unit and a headend unit over the system being measured.
- Correct any of the current RF board yield issues, including:
  - The tuning range issues in the first LO caused the 4.4-GHz-to-5.6-GHz oscillator to have chronic problems with dropping out at the low end of the frequency range at higher temperatures.
  - The first mixer input flatness was very dependent on the positioning of the dual-diode mixer package.
  - The 20-dB preamp was very dependent on transistor f_T (cutoff frequency), and since devices tend to vary from lot to lot, the preamp often required hand-selected parts.
  - The 20-dB input match required tuning.

Figure 5

Setup for a forward sweep (a) and a return sweep (b).



At the beginning of the project, the required hardware changes to the original HP 3010A block diagram (see **Figure 6**) were not readily obvious. We considered several schemes for the sweep source portion of the new remote unit. We finally settled on an approach that is based upon a tracking generator concept used in most spectrum analyzers. In a spectrum analyzer, a tracking generator is implemented by operating the RF conversion in reverse. That is, instead of converting a broad input frequency range into a fixed IF using superheterodyne techniques, a broad output range is created by operating a superheterodyne receiver in reverse (see **Figure 7**).

The tracking generator approach appeared to be really promising. However, to continue to use the unit as a receiver, we added switches around the RF converter section to switch between transmit and receive. The final configuration is shown in **Figure 8**.

Figure 6

Original HP 3010A block diagram.

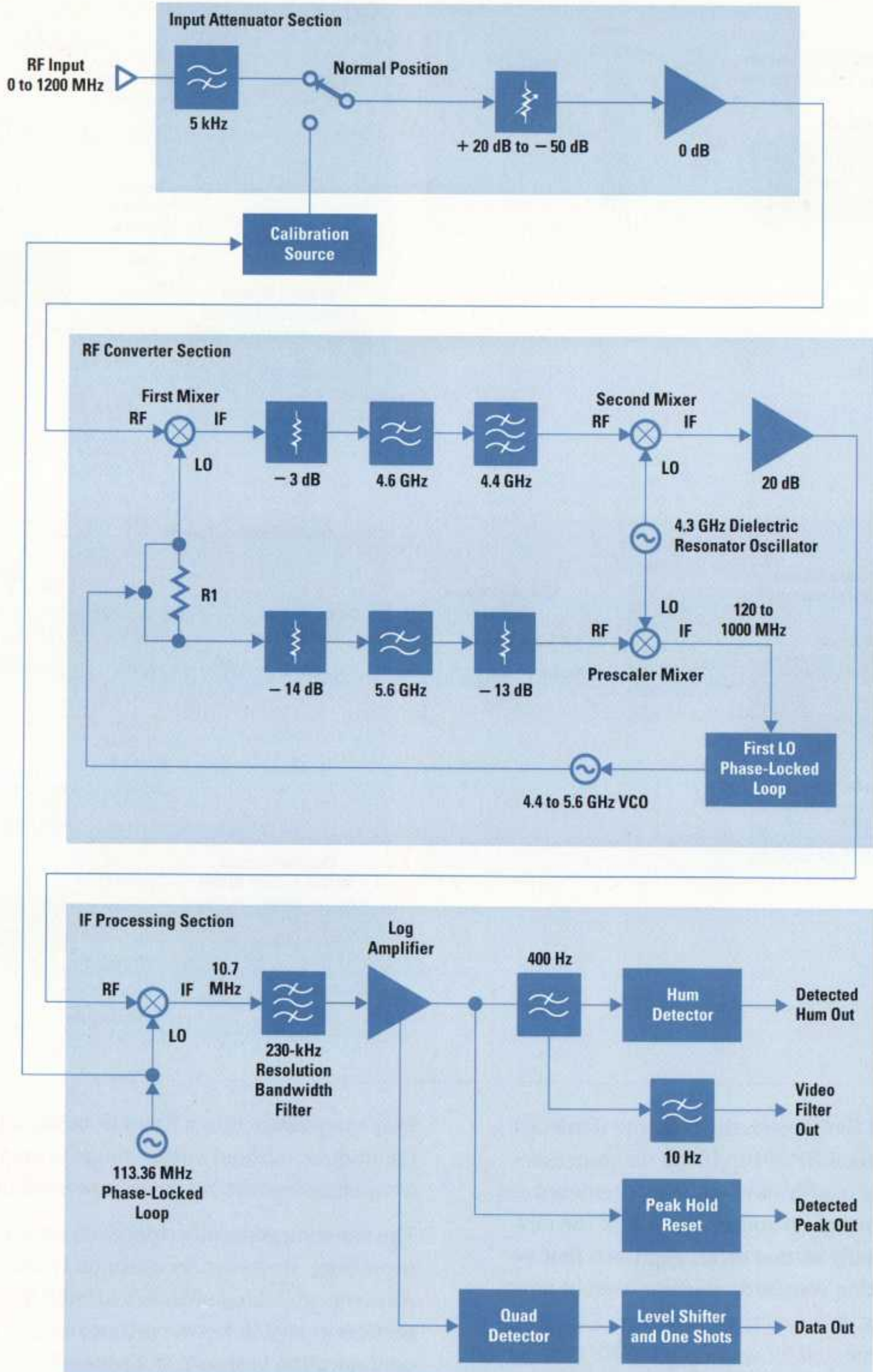
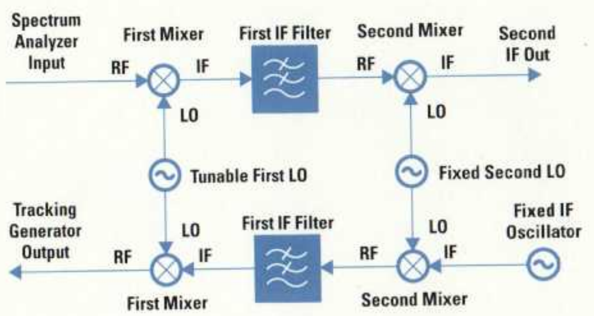


Figure 7

Generic spectrum analyzer tracking generator.



This approach operates as follows. In the receive mode, the signal enters through the RF input and the attenuator section, is passed through the first transmit/receive switch, and is injected into the first mixer. Based upon the reference level, the attenuators are set to keep the input power at the first mixer at approximately −30 dBm to −40 dBm. The first mixer upconverts the signal to a fixed IF of 4.454 GHz, which is filtered and passed on to the second mixer where it is downconverted to 124.0568 MHz. The 124.0568-MHz signal is passed through the second transmit/receive switch to the IF section where it is amplified and mixed a second time with 113.3568 MHz and downconverted to the final IF of 10.7 MHz. The signal is amplified again, passed through the resolution bandwidth filter, and applied to the log amplifier. The log amplifier output can either be filtered through a 400-Hz low-pass filter, or it can go directly to the peak detector. The output of the peak detector is sent to the main processor board, which contains an analog-to-digital converter (ADC).

In the transmit mode, the two transmit/receive switches are reversed, and the 124.0568 MHz phase-locked loop oscillator is turned on. The 124.0568 MHz oscillator signal is passed into the second transmit/receive switch and injected into the second mixer where it is upconverted to 4.454 GHz, filtered, and injected into the first mixer. Based upon the frequency of the first LO, the first mixer, operating in reverse, downconverts this fixed IF to a signal in the range of 5 MHz to 1 GHz. The resulting signal is then applied through the first transmit/receive switch to the output ALC (automatic level control) amplifier, which provides approximately 20 dB of ALC range to allow for slope correction and vernier 1-dB adjustments of the output power. The signal is then passed through an output attenuator which can be set to 0, 10, 20 or 30 dB, giving the unit an overall output power control range from 10 dBmV to 50 dBmV.

Overall Design

With the concept defined, detailed hardware development could be started. Five main hardware tasks had to be resolved. From the concept definition, two more design tasks were created: a 124.0568-MHz phase-locked loop oscillator and a leveled sweep amplifier (which would be adjustable from 10 dBmV to 50 dBmV).

Since one of the main priorities of this project was time to market, breadboarding was critical to start firmware development and identify unforeseen implementation problems early. The breadboards ended up being very solid and, in fact, are still being used by the firmware engineers to develop further enhancements to the product.

The next step in the hardware design was to convert the board from the fairly expensive low-loss material to a fairly inexpensive GeTek* material. The circuits that were most likely to cause problems were those that caused problems in the original design: the first mixer and the first local oscillator.

First Mixer

A symbolic schematic of the original first mixer is shown in **Figure 9**. The main function of the first mixer is to take the input signal (5 MHz to 1 GHz) and upconvert it to the fixed first IF of 4.454 GHz by multiplying it by the first LO frequency. For an upconversion at 5 MHz, the first LO must be 4.459 GHz, and for 1 GHz, it must be 5.454 GHz.

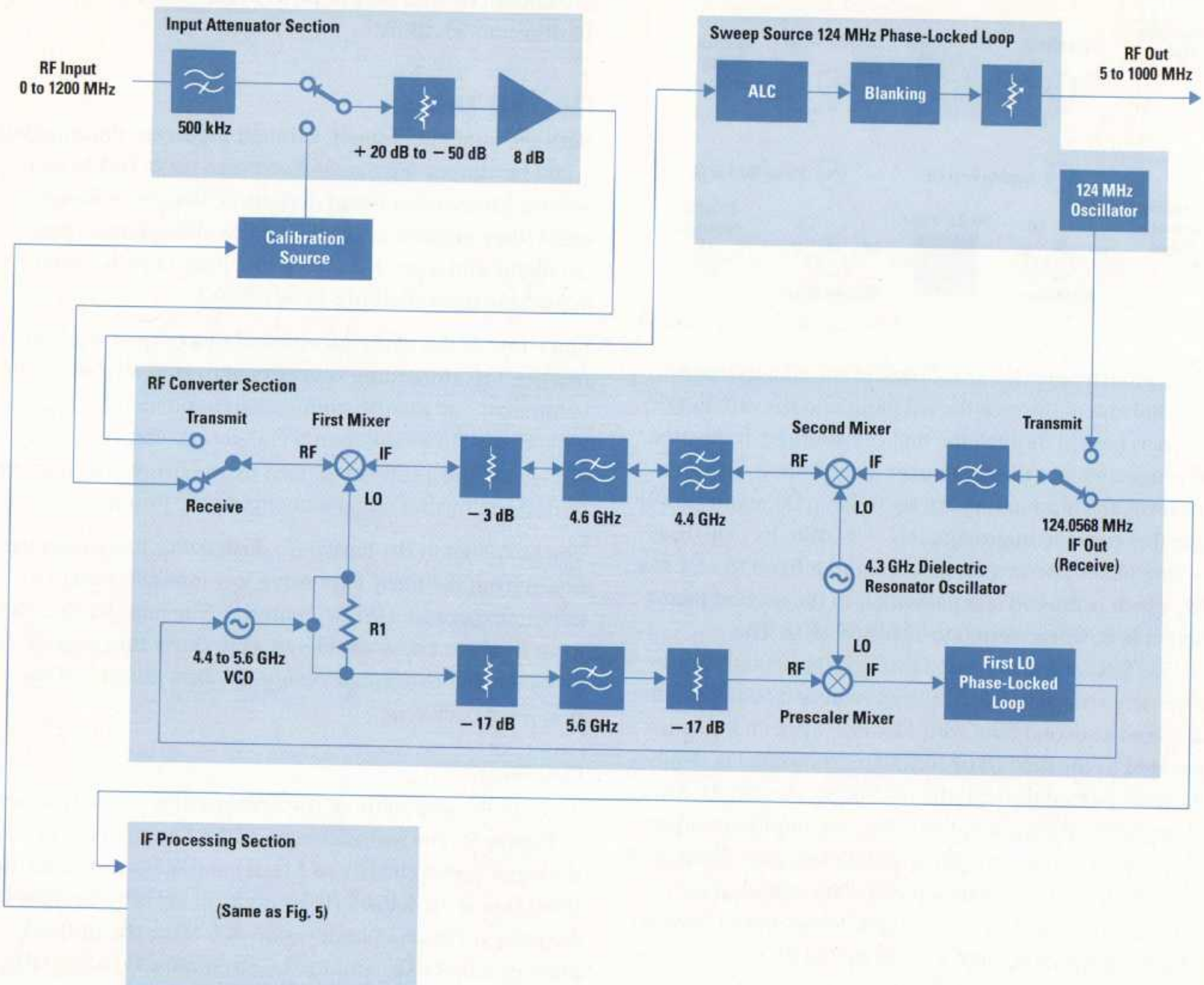
The mixer in **Figure 9** works as follows:

- During the first half of the LO cycle, D1 is forward biased and D2 is reverse biased. A forward-biased Schottky diode looks simply like a resistor, so the RF signal flows through T1 into the microstrip hybrid, emerging at the T3/T4 junction of the hybrid shifted by 270 degrees. Because a reverse-biased Schottky diode

* GeTek is a fairly low-loss, woven-glass type of material that is much easier to process than Teflon-based material laminated to FR4. Because GeTek is easier to fabricate, the raw board is significantly less expensive than one made with Teflon.

Figure 8

Final block diagram for the new HP 3010.



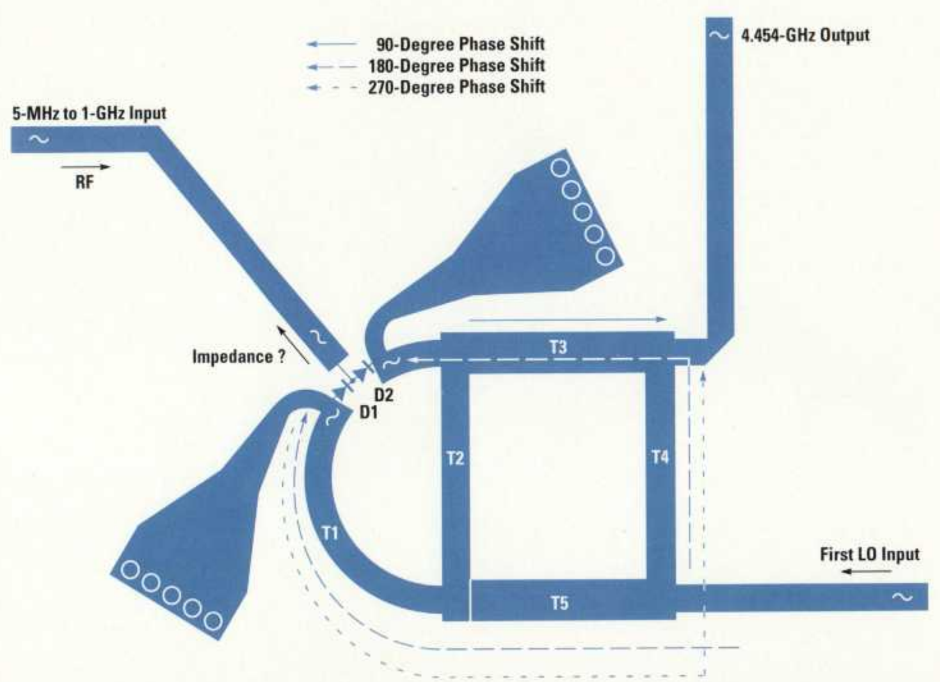
looks like an extremely small capacitor, an insignificant amount of the RF power flows into the T2/T3 junction.

- During the second half of the LO cycle, D2 is forward biased and D1 is reverse biased so that now the RF signal flows into the hybrid at the T2/T3 junction, emerging at the T3/T4 junction shifted by 90 degrees.

The net signal coming out of the T3/T4 junction can be seen to be a multiplication of the LO signal and the RF signal. That is, the phase reversal of 90 and 270 degrees of the RF signal is essentially the same as multiplying the RF signal by a unit LO pulse train with an amplitude of ± 1. **Figure 10a** shows an overlay of an RF input signal of 1 GHz and the first LO signal of 4.5 GHz. The solid trace in **Figure 10b** shows the voltage at the anode of D1. The dashed trace in **Figure 10b** shows the voltage at the cathode of D2. As these signals are phase shifted through the hybrid, the resultant trace at the T3/T4 junction is shown in **Figure 10c**. This is equivalent to multiplying the two traces shown in **Figure 10d**.

Figure 9

The original configuration of the first mixer used in the HP 3010A.



This multiplication performed in the time domain results in a translation in the frequency domain and the net result is the sum and difference of the RF and first LO frequencies. More specifically, if the input frequency is f_{rf} and the first LO frequency is f_{lo}, mixing produces two new frequencies: $f_{hi} = (f_{lo} + f_{rf})$ and $f_{low} = (f_{lo} - f_{rf})$. In our case we are interested in the f_{hi} so we simply filter the f_{low} with the first IF bandpass filter.

It was not clear, at first, what was causing the conversion loss to be so sensitive to the position of the mixer diode pair. A closer look at the layout shown in **Figure 9** and the functional description reveals that when D1 is forward biased the impedance looking into the common leg is extremely important. In fact, it would be best if it were a short circuit at the LO frequency, and conversely, with an open circuit, the mixer would hardly function. Looking out of the mixer RF input leg, it was extremely hard to tell what that impedance would be at 4.4 GHz or 5.6 GHz, and it was even harder to control. That leg is fed by a low-frequency (5-MHz-to-1-GHz) 8-dB preamp and its output impedance is not easily defined or controlled at 4.4 GHz. What was needed at that point was an ac short over the LO range (4.4 to 5.5 GHz) that would pass a signal in the 5-MHz-to-1-GHz RF input range. This was a prime application for a radial stub. After conducting simulation we decided to add a pair of radial stubs to the mixer, resulting in the layout shown in **Figure 11**.

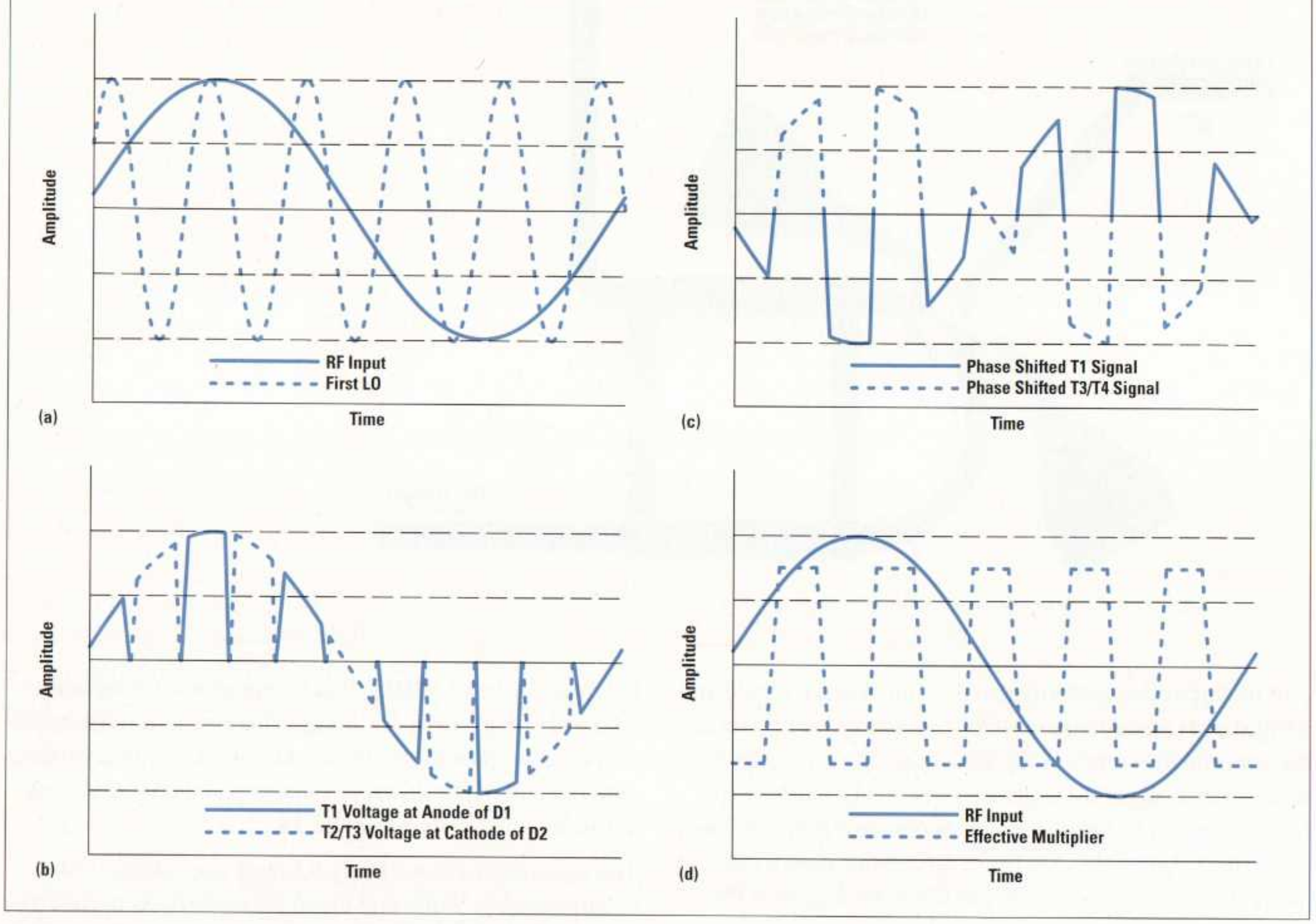
The new mixer shows a much better conversion loss (approximately 2 dB) and much better flatness across the band. We also found that the new mixer is less susceptible to LO power variations, and when operating in the tracking generator mode, shows approximately a 10-dB improvement in harmonic performance. A breadboard of the mixer was built and tested on the new GeTek material and was found to perform quite well, with the measured results closely matching those of the simulation.

First Local Oscillator

The first local oscillator, which had been a chronic problem circuit on the original board material, needed to cover the 4.4 to 5.6 GHz range. This was perhaps the single most risky circuit to try to fabricate on a lossy dielectric material. Because oscillators require fairly high Q resonators so that the negative impedance can

Figure 10

Waveforms appearing at the junctions of the first mixer shown in Figure 9. (a) An overlay of an RF input signal of 1 GHz and the first LO signal of 4.5 GHz. (b) The voltage at the T1 junction (anode of D1) and the cathode of D2 (T2/T3 junction). (c) The phase-shifted signals at the junction of T3/T4. (d) The two signals that could be multiplied together to produce the trace shown in Figure. 10c.



overcome the lossy portion of the resonator, going to a higher loss material seemed like it could only make the problems worse.

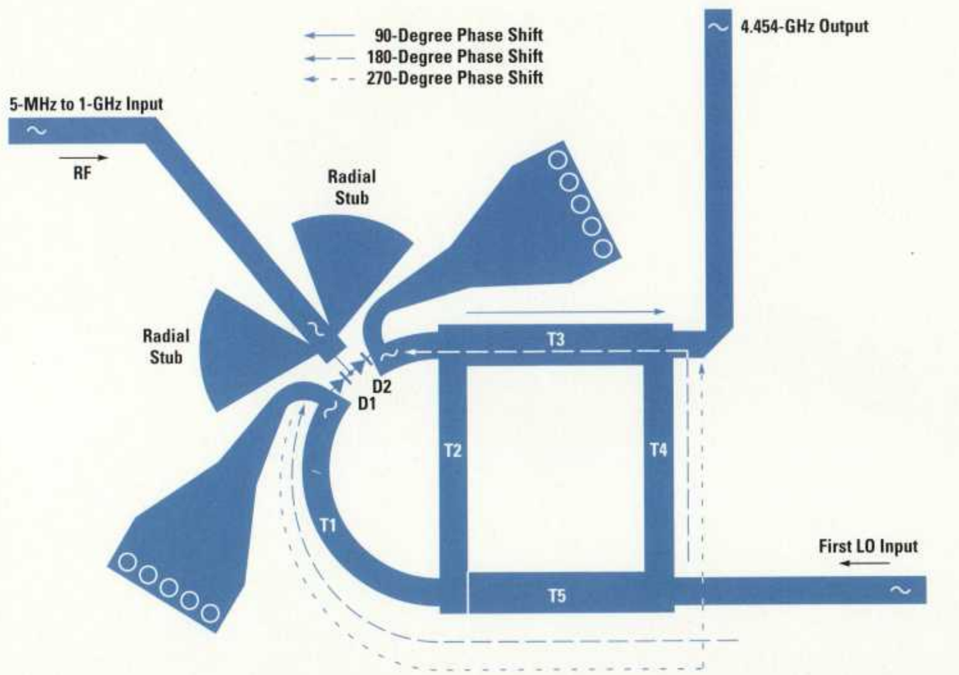
We simulated the oscillator in its original configuration to understand the nature of the original circuit's problem with the low end of the frequency range dropping out of oscillation when subjected to heat. Also, the oscillator had been susceptible to *moding* (multiple oscillation modes) in the past. The oscillator was a negative-resistance oscillator with a microstrip stub and a varactor as its resonator. The simulation revealed several things about the oscillator that gave insight into the cause of the recurring problem at the low end of the frequency range. The absolute value of the negative resistance was barely enough to overcome the losses of the resonator, and as the active device's parameters moved with increased temperature, the oscillator often dropped out of oscillation. To rectify this problem, the circuit was modified by adding a small capacitor on the collector of the oscillator transistor. The added capacitance had the effect of moving the negative resistance to a higher absolute value, thereby enabling the oscillator to overcome the losses of the resonator. Furthermore, the resonator was adjusted to more closely center on the negative resistance of the active device.

Second Mixer and Dielectric Resonator Filter

The design team originally thought that the second mixer and dielectric resonator filter would not need many

Figure 11

A modification to the configuration shown in Figure 9. This is the configuration used in the new HP 3010H.



changes other than scaling the line widths and lengths. However, because the second LO is loaded by the second mixer and dielectric resonator, the line lengths become critical (see **Figure 12**).

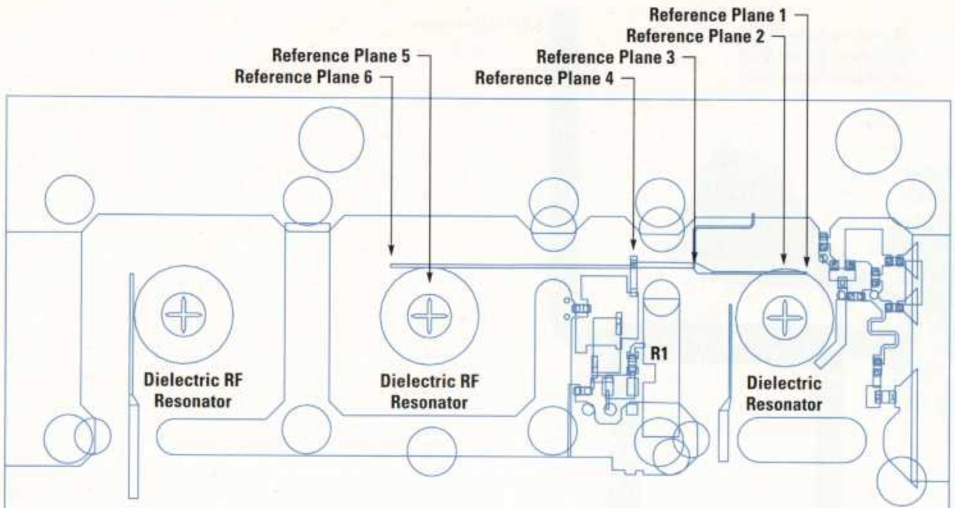
To prevent loading the dielectric resonator oscillator (DRO) at its center frequency (4.33 GHz), the combined electrical length from reference plane 2 to reference plane 6 should be a multiple of 180 electrical degrees (or half wavelengths). In addition, it is best that the electrical length from reference plane 2 to the mixer diode reference plane 4 be a multiple of 180 degrees at the IF frequency of 4.454 GHz. This is to prevent loading the mixer diode and to create a maximum voltage swing across it. Finally, the electrical length from reference plane 1 to reference plane 5 had to be a multiple of 180 degrees at 4.454 GHz (the first IF frequency) to prevent loading the dielectric RF resonator (DRF) by the DRO. The result is that the overall length of the long interconnecting microstrip line between the mixer and oscillator has to be some multiple of half wavelengths (such as 0.5, 1.0, or 1.5).

The original Teflon-based design called for this circuit to have approximately one wavelength from the DRF reference plane (reference 5) to the DRO reference plane (reference 2). This is made up of 0.5 wavelength from reference plane 5 to reference plane 4 and 0.5 wavelength from reference plane 2 to reference plane 4. Since GeTek has a higher dielectric constant than the Teflon substrate used in the old design, and since higher dielectric constants result in shorter wavelengths, it was physically impossible to leave the length of the line at one wavelength. Unfortunately, we didn't realize this until after the first board turn.

Much of the health of this circuit is based on the amount of mixer bias voltage coming from the average dc current induced in the mixer diode by the second LO. In the first turn of the new board, the mixer bias voltages across R1 were approximately 30 to 40 mV. In comparison, the original board had a mixer bias voltage of approximately 100 mV. After this mistake was realized, and an extra 180 degrees of electrical length from reference plane 4 to reference 5 was added to the line, the mixer bias voltage was actually increased to an average of 110 mV.

Figure 12

The second mixer.



124.0568-MHz Oscillator

The design of the 124.0568-MHz oscillator was fairly straightforward. It was implemented as a classical phase-locked loop design using a Motorola MC145150 phase-locked loop IC, having a reference frequency of 2.7648 MHz. Since all the data exchanged between the headend and the remotes employs phase shift keying (PSK) modulation of the carrier, this capability was included in the 124.0568-MHz oscillator design. Since frequency equals the derivative of phase with respect to time, a step in phase is equivalent to an impulse in frequency. This phase shift keying modulation was implemented by impulsing the tuning line of the oscillator at a frequency outside the loop bandwidth of the oscillator. Since the VCO's output frequency is a function of voltage, we can accomplish a step in phase by applying an impulse of voltage through a high-pass differentiator to the tuning input of the oscillator (see **Figure 13**).

Output ALC Amplifier

The output ALC amplifier was implemented using classical design techniques. This included a variable gain amplifier (HP IVA05208) on the input to provide the variable gain required for leveling. The signal was then passed through two fixed-gain amplifiers and detected using a classical diode detector with a reference diode for temperature stability.

Input Attenuator Match

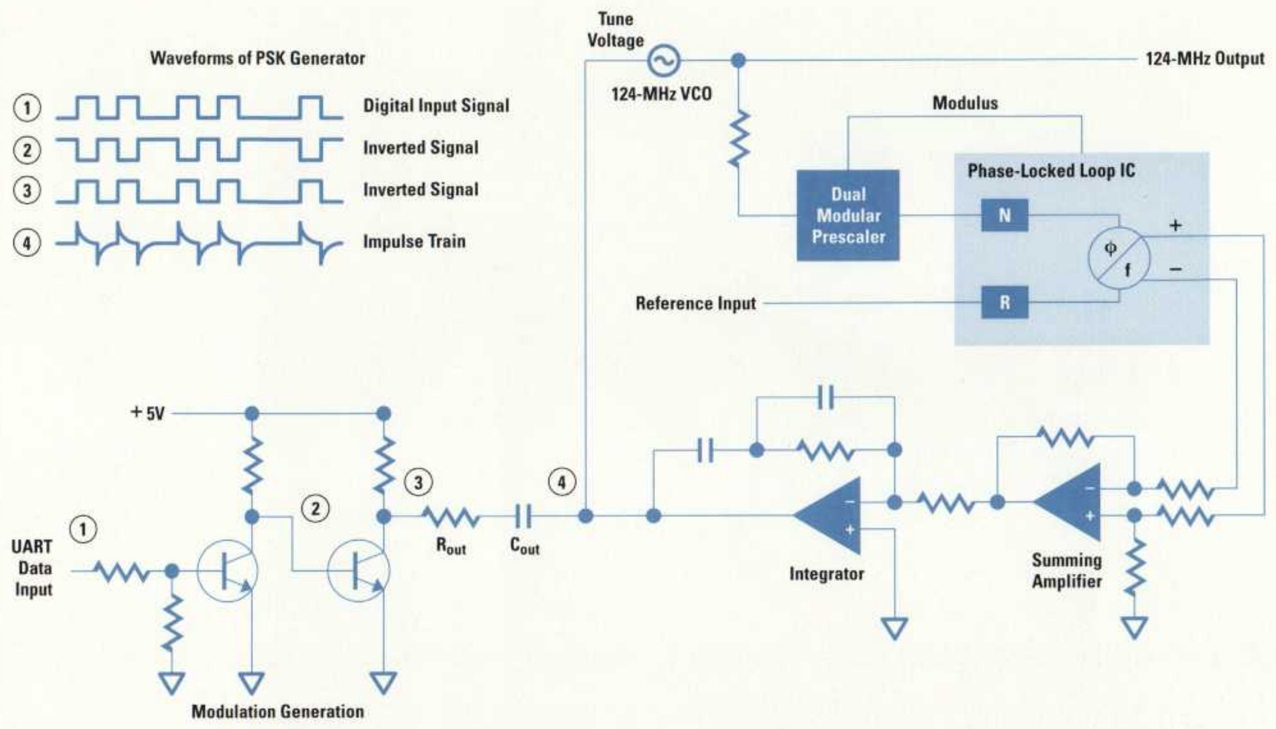
Since the HP 3010H and HP 3010R sweep/ingress analyzer is used as a signal level meter, a large part of its accuracy depends on the input match. The specification for input return loss of the instrument is 18 dB. The RF module has a specification of 23 dB.* The previous design had some difficulty meeting this specification, and often the modules had to be manually adjusted to reach the desired performance. One goal of this project was to eliminate manual tuning to achieve the input match. This instrument is required for use in cable systems having a 75-ohm characteristic impedance. All of the switches and relays that were available for the instrument design were designed for 50-ohm systems (75-ohm relays and GaAs switches do not exist).

To use these parts in a 75-ohm system required some unconventional matching techniques. In a 75-ohm system, a 50-ohm transmission line (if it is short enough) tends to

* The RF module has a tighter specification than the instrument because there is a cable and two connectors between the module output and the instrument output. These cables and connectors degrade the return loss of the module.

Figure 13

PSK modulation generation.



function like a parallel capacitor. For a superheterodyne receiver, like the HP 3010, it is desirable to have a low-pass filter in the input. By using 50-ohm parts as shunt capacitors in a classical inductor and capacitor low-pass filter structure and high-impedance transmission lines as inductors, we were able to achieve good input return loss and proper filtering. At the initial production release, the input return loss did not quite meet specifications. However, with some minor repeatable modifications, the new RF assembly aligned much faster than the original RF assembly. The complete RF converter board is shown in **Figure 14**.

Communication Protocol Development

Another major part of this project was the development of firmware. Three major features were added to the existing firmware:

- Reverse sweep and communications protocol
- Ingress detection
- Digital-channel power measurement.

Reverse Sweep

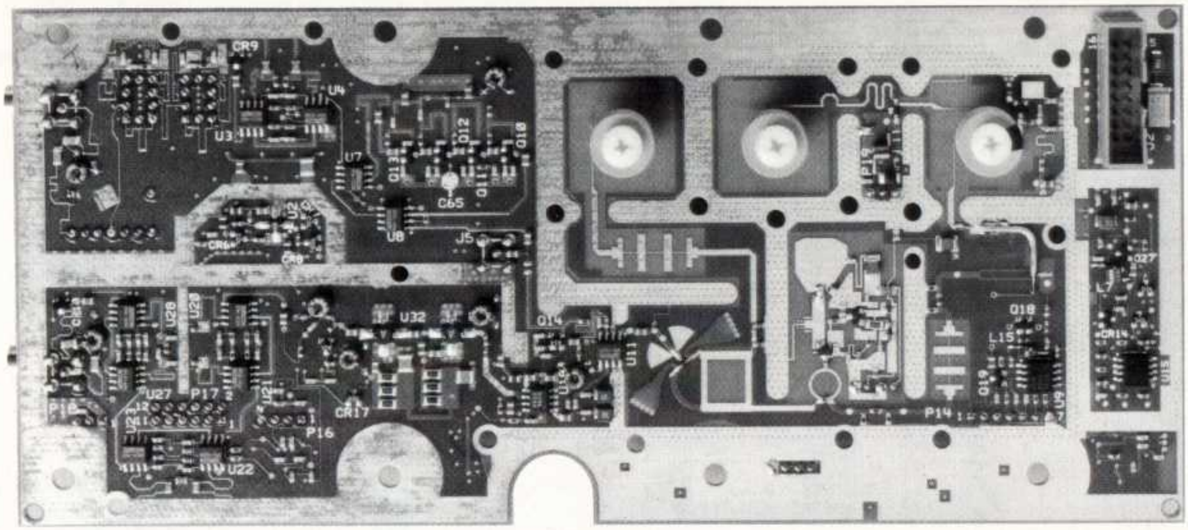
The primary effort of the firmware development was to create an algorithm to perform a return sweep. Return sweep is used to generate a sweep response plot of the return (subscriber's-home-to-headend) position of the cable system (see **Figure 15**). With this sweep response, technicians can align their amplifiers for the best possible response.

On initial inspection, the problem of communication between a remote unit and a headend unit seemed like a fairly simple proposition, with the communication going something like:

- Headend to Remote: "Hey, are you there?"
- Remote to Headend: "Yes, I'm here."
- Headend to Remote: "OK, what do you want to do?"
- Remote to Headend: "I want to sweep."

Figure 14

RF converter board.



- Headend to Remote: "OK, go sweep."
- Remote to Headend: "OK, sync now."

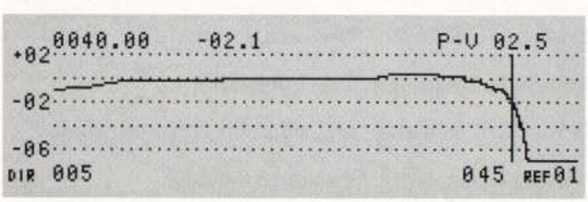
The headend and remote would be synchronized, and the remote would then send out pulses at various frequencies until the end of sweep was reached. At that point, the headend could simply process the data and send it back to the remote. It is easy to see that a simple one-remote system would be easy to program.

Trying to handle multiple users (remotes) with this technique might result in the following communication dialogue:

- Headend to Remotes: "Hey, are you there?"
- Remote 1 to Headend: "Yes, I'm here."

Figure 15

Typical return sweep response.



- Remote 2 to Headend: "Yes, I'm here."
- Remote 3 to Headend: "Yes, I'm here."
- Remote 4 to Headend: "Yes, I'm here."
- Remote 5 to Headend: "Yes, I'm here."
- Remote 6 to Headend: "Yes, I'm here."
- Remote 7 to Headend: "Yes, I'm here."
- Headend to Remotes: "Wait! Wait! Wait! I can't understand you all at once."

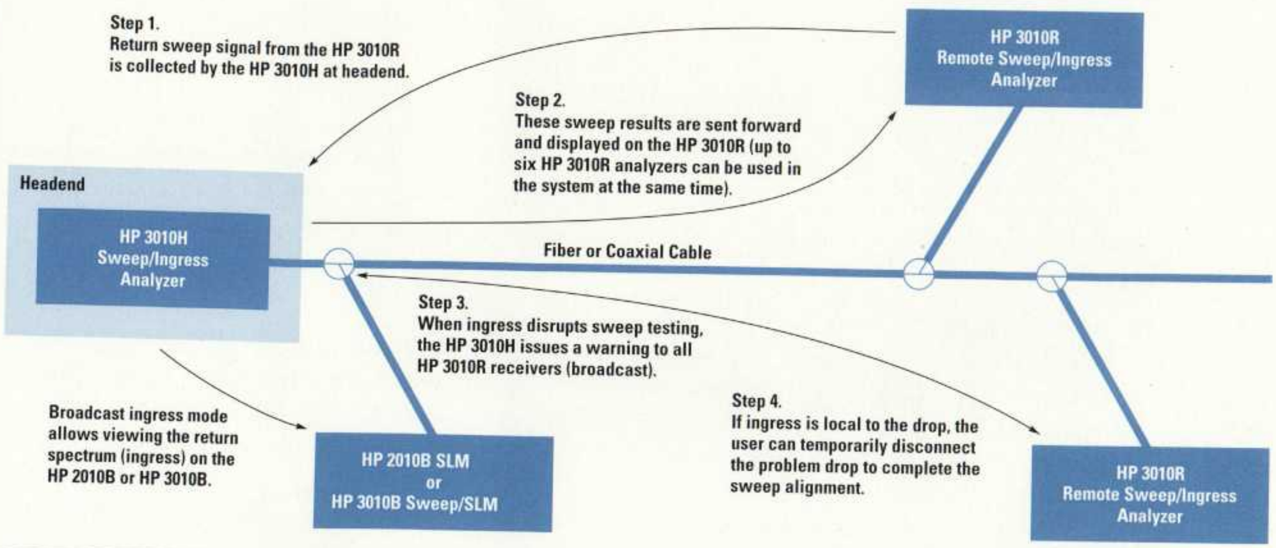
At this point, communication would break down.

In our development, we chose to take a more civilized approach. Each remote unit is assigned a serial number to enable the headend to identify and communicate with it. In addition, we decided there would be two states for the headend and remote units: *connected* and *unconnected.* In the connected state the headend and remote units are in sweep loops. The headend can be connected to many remotes. It keeps a list that is used to control when each remote is run through the sweep loop. The unconnected state would indicate for the headend that it is not connected to any remotes, and for any remote that it is not connected to the headend.

New users (remotes) would be polled by repeating messages broadcast on the forward pilot from the headend.

Figure 16

Ingress detection.



The forward pilot is a frequency that is set aside in the cable spectrum for communication. The HP 3010R uses a forward and return pilot to complete the communication loop.

When new users are first connected to the cable system, they respond to the new user poll by returning (over the return pilot) their serial number and a cyclic redundancy checksum (CRC) of the sweep table stored in a remote's memory.* The sweep table checksum response is important because if it does not match the headend's sweep table checksum when the unit begins to sweep, the headend and remote units will not be synchronized, resulting in bad sweep data. Once a new user has responded to a new user poll and received a sweep message from the controlling headend, it is then considered connected and is added to the headend's queue.

Since it is possible for more than one unit to respond to a new user poll simultaneously, the headend can use four allotted time slots to listen for responses. To create random time slot selection, each remote in the not-connected state that attempts to respond will choose a time slot based on a pseudorandom number generated from its internal clock. Since the likelihood of any two remotes being set to the same time is very slim, a fairly equal probability exists for a given remote to respond in any one of the four time slots. Because the headend can only respond to one time slot at a time, the first remote unit's time slot that contains a message will be the next unit to be connected. To conclude the connection, any responding remote that receives a sweep message from the headend will be considered connected.

Ingress Detection

Since the influx of noise (ingress) is a problem, we wanted to add a feature that would allow remote field technicians to see the level of ingress being received at the headend (or at a hub site). In addition, it is important for field technicians to be able to see this ingress even if the remote unit can no longer communicate with the headend. To accomplish this, we decided to have two modes of ingress operation: *return spectrum* and *broadcast ingress* (see **Figure 16**).

Return spectrum, as the name implies, is nothing more than a snapshot of the return spectrum taken by the headend and sent to the remote unit over the forward path. The return spectrum measurement, like the return sweep measurement, is a demand measurement (the remote unit

* Sweep tables are used in the HP 3010 SLM system to keep the sweep signals out of the way of visual, aural, and, in some cases, digital carriers.

demands the measurement). It is very similar to the return sweep measurement, except that the remote unit does not send sweep pulses. As a consequence, no synchronization pulse is required. When the remote unit demands a return spectrum measurement (during the synchronization message) it waits in the listen mode. The headend then performs its usual new user poll, but instead of listening for the sweep pulse from the remote unit, it measures the noise ingress. Upon completing the return spectrum measurement, the headend sends a data message to the remote unit that requested the return spectrum measurement.

Broadcast ingress, on the other hand, is not a demand measurement. Broadcast ingress measurements can only occur if the headend is in continuous mode, or if it receives noise above a certain level while trying to receive a message from the remote unit. When this occurs, the headend sends a broadcast message containing the return spectrum data to all remotes. When the remote units receive this message, they add a softkey to their display and store the data in an array. The user can then view the data at any time by simply pressing the softkey. When the headend is set to continuous ingress mode, it sends a broadcast ingress message every other sweep cycle.

Digital Channel Power Measurement Algorithm

Because many cable operators are now beginning to offer digital services, and since most digital services are broadcast using some sort of digital modulation scheme (such as QAM, QPSK, and so on) that results in a broadband noise-like signal, it is no longer an easy matter to measure channel power. Digital signals tend to look like noise pedestals in the frequency domain (see **Figure 17**). A new algorithm has been added to the HP 2010 and HP 3010 firmware to enable the user to measure digital channel power easily and accurately.

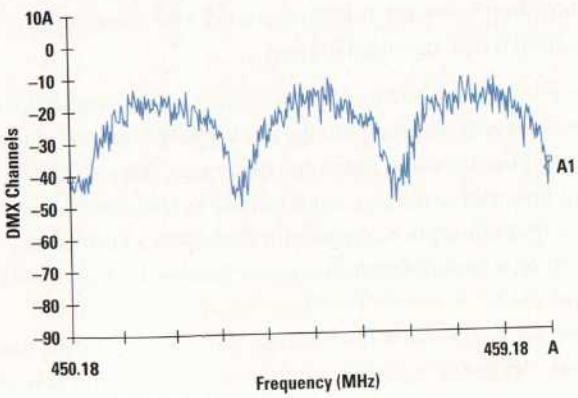
An algorithm developed for the HP 8591C spectrum analyzer has been leveraged into the HP CaLan 2010 and 3010 signal level meters. This algorithm is based on the fact that, at any given center frequency, the power level detected at the output of the log amplifier is a time sample of the total power contained within the resolution bandwidth of the receiver. That is, it equals the bandwidth multiplied by the noise power density in that band (watts/Hz). As an example, with a noise power density of -100 dBm/Hz, the total power contained within a 230-kHz bandwidth would be -46.383 dBm. (This calculation must be performed with linear power. To do the conversion in dB, $-100 + 10\log(230 \times 10^3) = 46.383$.) Since the resolution bandwidth of the HP 2010 and HP 3010 modules is approximately 230 kHz, to measure an 8-MHz channel requires at least 35 steps (8 MHz/0.230 MHz). To calculate the total power in a channel, the receiver is tuned to the low end of the channel and then stepped across the entire band in steps narrower than the bandwidth of the filter. At each frequency step, the power is converted to linear, averaged five times, weighted by the overlap in the step size, and added to the other frequency data. After the final frequency is reached, the total power is converted back to dBmV for display.

The accuracy of this measurement depends on two major factors: resolution bandwidth accuracy and the ability of the detector to operate in sample mode. Since the shape of the resolution bandwidth filter is not easy to model mathematically, it is necessary to define the filter's noise equivalent bandwidth. This can be defined as the effective ideal bandwidth of the filter. It can be calculated by finding the area under the normalized curve of the filter. For example, an ideal brick wall filter would have a normalized amplitude response of one and a bandwidth of $\Delta\omega$. Thus, the brick wall filter has a noise equivalent bandwidth of $\Delta\omega$.

Since the resolution bandwidth filter is not consistent from unit to unit, we devised a way to calibrate the noise equivalent bandwidth of the HP CaLan 2010 and 3010 modules. This is accomplished by turning on the internal

Figure 17

Digital channels.



calibrator and stepping the first LO in fine increments around the center frequency. The responses are then summed in a trapezoidal integration to give the resultant noise equivalent bandwidth.

For accurate measurement of a noise-like signal, it is necessary for the detector to be in sample mode. This is very important because it is impossible to know what the peak to average power level ratio is for any given type of modulation scheme. Fortunately, it is possible to run the detector in a sample mode. The final algorithm gives results that correlate well with an rms power meter and with the HP 8591C digital channel power downloadable program.

Other firmware changes included upgrades for printer support and improvements to the user interface to accommodate the newly added features.

Conclusion

The new HP CaLan 2010 and 3010R and 3010H sweep/ingress analyzers were completed quickly to fill the urgent needs of cable system operators to sweep and align the forward and return paths of their cable systems. The final product is easy and intuitive to use and includes many system improvements. This was achieved by setting and maintaining clear priorities from the very beginning. In the order of precedence, these priorities were schedule, quality and reliability, performance, and cost. By focusing on the priorities, we were able to introduce a new product within a very short time.

Acknowledgments

Many people were instrumental in getting this product to market in what may be record time for the Microwave Instruments Division. The lab team made up of Bill Morgan (overall project leader), Brenda Coomes (firmware project leader), Bill Scales (digital design), Cinda Craven (mechanical hardware design), Marv Milholland (remote firmware design), Chin-Min Wang (headend and digital power firmware design), and John Mullan (user interface and printer support firmware design) were all instrumental in the creation of this product. The documentation team made up of Jerry Green and Al Barnes did an excellent job of creating a comprehensive set of manuals. The manufacturing team including Dave Wolbeck (hardware support and RF board test development), Jim Stock (hardware support and RF board test development), Dennis Nishi (production engineering RF board), Ken Silk (production engineering instrument test development), Phil Melman (production engineering support for everything else), Adam Pirog (instrument line technician and all-around troubleshooter), Kim Holliday (industrial design), Charlie Wolfe (production engineering mechanical engineer), and Patty Hicks (line supervisor) were instrumental in reaching our shipment and pilot run goals. The new product introduction group consisting of Vic Sallee (new product introduction purchasing), James Ellison (design drafter), Steve Comins (SMTC test support), Linda Morrison (new product introduction coordinator), and Julie Silk (SMTC liaison) were also quite helpful in pulling things together.

Reference

1. W. S. Ciciora, *Cable Television in the United States—An Overview*, Cable Television Labs, Inc. 1995 (http://www.cablelabs.com/C-IT/ciciora.html).

Online Information

For more information about the product described in this article, take a look at the information located at the following URLs:

- http://www.tmo.hp.com/tmo/datasheets/English/HPCaLan.html
- http://www.tmo.hp.com/tmo/datasheets/English/HP3010H.html
- http://www.tmo.hp.com/tmo/datasheets/English/HP3010R.html
- http://www.tmo.hp.com/tmo/datasheets/English/HP3010B.html
- http://www.tmo.hp.com/tmo/datasheets/English/HP2010B.html

The Hewlett-Packard Journal

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.

To receive an **HP employee subscription** send an e-mail message indicating your HP entity, employee number, and mailstop to: **ldc_litpro@hp0000.hp.com**

Qualified non-HP individuals, libraries, and educational institutions in the **U.S.** can request a subscription by going to our website and following the directions to subscribe.

To request an **International subscription** locate your nearest country representative listed on our website and contact them directly for a subscription. Free subscriptions may not be available in all countries.

Back issues of the Hewlett-Packard Journal can be ordered through our website.

Our Website

Current and recent issues are available online at **http://www.hp.com/hpj/journal.html**

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. Before doing any work on an article, please contact the editor by sending an e-mail message to: **hp_journal@hp.com**

Copyright

Inquiries

Please address inquiries, submissions, and requests to:

Editor
Hewlett-Packard Journal
3000 Hanover Street, Mail Stop 20BH
Palo Alto, CA 94304-1185 U.S.A.

HEWLETT-PACKARD Journal

FEBRUARY 1998 • Volume 49, Number 1

Technical Information from the Hewlett-Packard Company

5966-2129E