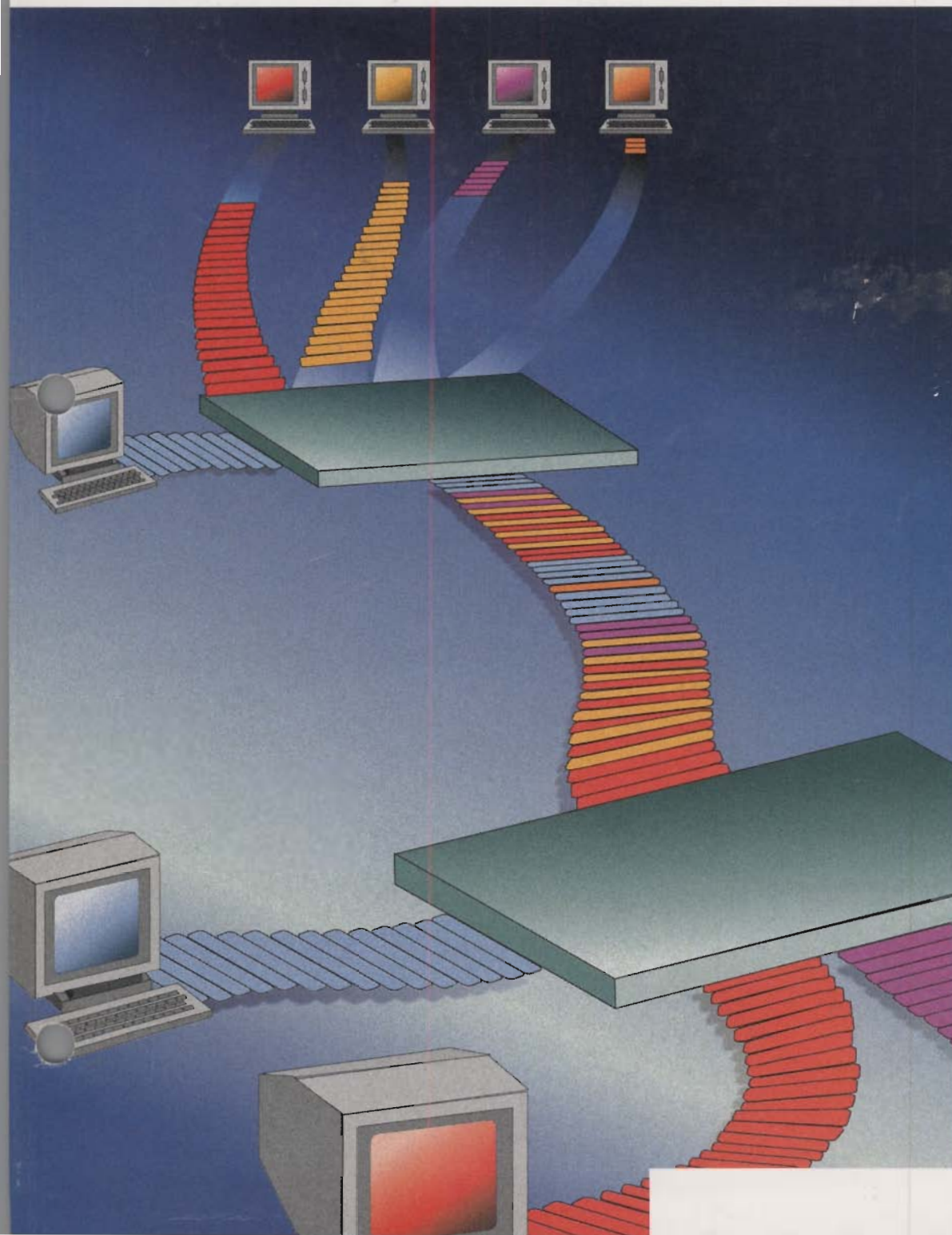


# HEWLETT-PACKARD JOURNAL

August 1995



## Articles

- 
- 6 **Introduction to 100VG-AnyLAN and the IEEE 802.12 Local Area Network Standard**, by Alan R. Albrecht and Patricia A. Thaler
- 7 **Cable Types**
- 10 **Other Network Technologies**
- 
- 13 **Demand Priority Protocol**, by Alan R. Albrecht, Michael P. Spratt, Patricia A. Thaler, and Gregory C.A. Watson
- 15 **Network Protocol Layers**
- 
- 18 **Physical Signaling in 100VG-AnyLAN**, by Alistair N. Coles, David G. Cunningham, Joseph A. Curcio, Jr., Daniel J. Dove, and Steven G. Methley
- 19 **Cross Talk in Unshielded Twisted-Pair Cables**
- 21 **Multilevel Signaling**
- 22 **Cross Talk Analysis**
- 26 **Optical-Fiber Links for 100VG-AnyLAN**
- 
- 27 **Coding in 100VG-AnyLAN**, by Simon E.C. Crouch and Jonathan Jedwab
- 30 **IEEE 802.3 and 802.5 Frame Formats**
- 31 **Polynomial Arithmetic and Cyclic Redundancy Checks**
- 
- 33 **Multimedia Applications and 100VG-AnyLAN**, by John R. Grinham and Michael P. Spratt
- 35 **Remote Bridge Example**
- 36 **Higher-Level Protocols**
- 38 **Related Projects**
- 

**Editor**, Richard P. Dolan • **Associate Editor**, Charles L. Leath • **Publication Production Manager**, Susan E. Wright • **Illustration**, Renée D. Pighini

**Advisory Board**, Rajeev Badyal, *Integrated Circuit Business Division, Fort Collins, Colorado* • Thomas Beecher, *Open Systems Software Division, Chelmsford, Massachusetts* • Steven Brittenham, *Disk Memory Division, Boise, Idaho* • William W. Brown, *Integrated Circuit Business Division, Santa Clara, California* • Rajesh Desai, *Commercial Systems Division, Cupertino, California* • Kevin G. Ewert, *Integrated Systems Division, Sunnyvale, California* • Bernhard Fischer, *Böblingen Medical Division, Böblingen, Germany* • Douglas Gennetten, *Greeley Hardcopy Division, Greeley, Colorado* • Gary Gordon, *HP Laboratories, Palo Alto, California* • Mark Gorzynski, *Inkjet Supplies Business Unit, Corvallis, Oregon* • Matt J. Harline, *Systems Technology Division, Roseville, California* • Kiyoyasu Hiwada, *Hachioji Semiconductor Test Division, Tokyo, Japan* • Bryan Hoog, *Lake Stevens Instrument Division, Everett, Washington* • C. Steven Joiner, *Optical Communication Division, San Jose, California* • Roger L. Jungeman, *Microwave Technology Division, Santa Rosa, California* • Forrest Kellert, *Microwave Technology Division, Santa Rosa, California* • Ruby B. Lee, *Networked Systems Group, Cupertino, California* • Alfred Maute, *Waldbromm Analytical Division, Waldbromm, Germany* • Andrew McLean, *Enterprise Messaging Operation, Pinewood, England* • Dona L. Miller, *Worldwide Customer Support Division, Mountain View, California* • Mitchell Mlinar, *HP-EEsof Division, Westlake Village, California* • Michael P. Moore, *VXI Systems Division, Loveland, Colorado* • Shelley I. Moore, *San Diego Printer Division, San Diego, California* • M. Shahid Mujtaba, *HP Laboratories, Palo Alto, California* • Steven J. Narciso, *VXI Systems Division, Loveland, Colorado* • Danny J. Oldfield, *Electronic Measurements Division, Colorado Springs, Colorado* • Garry Orsolini, *Software Technology Division, Roseville, California* • Han Tian Phua, *Asia Peripherals Division, Singapore* • Ken Poulton, *HP Laboratories, Palo Alto, California* • Günter Riebesell, *Böblingen Instruments Division, Böblingen, Germany* • Marc Sabatella, *Software Engineering Systems Division, Fort Collins, Colorado* • Michael B. Saunders, *Integrated Circuit Business Division, Corvallis, Oregon* • Philip Stenton, *HP Laboratories Bristol, Bristol, England* • Stephen R. Undy, *Systems Technology Division, Fort Collins, Colorado* • Jim Willits, *Network and System Management Division, Fort Collins, Colorado* • Koichi Yanagawa, *Kobe Instrument Division, Kobe, Japan* • Dennis C. York, *Corvallis Division, Corvallis, Oregon* • Barbara Zimmer, *Corporate Engineering, Palo Alto, California*

**HP Computer Museum**  
**[www.hpmuseum.net](http://www.hpmuseum.net)**

**For research and education purposes only.**

---

39	100VG-AnyLAN 15-Port Hub Design , by Lisa S. Brown
41	Invalid Packet Marker
43	HP AccuPage 2.0: A Toolkit for High-Quality Document Scanning , by Steven L. Webb, Steven G. Henry, Kevin S. Burke, and George Prokop
44	Glossary
51	An 11.8-in Flat Panel Display Monitor , by David J. Hodge, Bradly J. Foster, Steven J. Komrmusch, and Tom J. Searby
53	Liquid Crystal Display Technology
57	Product Design of the HP S1010A Flat Panel Display
59	A Note About VRAMs
61	Applying an Improved Economic Model to Software Buy-versus-Build Decisions , by Wesley H. Higaki
66	Benchmark Standards for ASIC Technology Evaluation , by Antonio A. Martinez, Aloke S. Bhandia, and Henry H.W. Lie

---

## Departments

---

4	In this Issue
5	Cover
5	What's Ahead
71	Authors

---

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

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

The Hewlett-Packard Journal is available online via the World-Wide Web (WWW) and can be viewed and printed with Mosaic. The uniform resource locator (URL) for the Hewlett-Packard Journal is <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. Please contact the Editor before submitting such articles. Also, the Hewlett-Packard Journal encourages technical discussions of the topics presented in recent articles and may publish letters expected to be of interest to readers. Letters should be brief, and are subject to editing by HP.

Copyright © 1995 Hewlett-Packard Company. All rights reserved. Permission to copy without fee all or part of this publication is hereby granted provided that 1) the copies are not made, used, displayed, or distributed for commercial advantage; 2) the Hewlett-Packard Company copyright notice and the title of the publication and date appear on the copies; and 3) a notice appears stating that the copying is by permission of the Hewlett-Packard Company.

Please address inquiries, submissions, and requests to: Editor, Hewlett-Packard Journal, 3000 Hanover Street, Palo Alto, CA 94304 U.S.A.

---

---

## In this Issue



Most local area networks (LANs) today have the same bandwidth, or maximum data transfer rate, as ten years ago—10 to 16 megabits per second. In the same time period the speed of computers and the size of files created by sophisticated applications have increased a hundredfold. As we might expect, the result is a rising level of user complaints about long response times, symptomatic of network congestion caused by too many users and a high volume of data-intensive applications such as database access, image analysis, desktop publishing, network printing, and computer-aided design. Multimedia applications demand even more bandwidth and will only make the situation worse unless the bandwidth of these LANs can be increased. An increase in the data rate to 100 mega-

bits per second would be a major improvement. Because existing local area networks already represent major investments in wiring, it would be ideal if the increased bandwidth were made available over existing twisted-pair building wiring, with fiber-optic cabling an option. The ideal technology would also be compatible with existing network protocols, topologies, and software, and would offer guarantees that such multimedia applications as real-time video would not experience unacceptable delays because of congestion. In 1992, HP Laboratories in Bristol, England and the HP Roseville Networks Division in California established teams to work on this challenge. Official approval of their solution, the proposed IEEE 802.12 demand priority local area network standard, is expected soon. This new networking technology, called 100VG-AnyLAN by HP, is introduced in the article on page 6. It offers an affordable upgrade path for congested LANs based on the IEEE 802.3 (Ethernet) and IEEE 802.5 (token ring) standards, as most existing LANs are. It provides a 100-megabit-per-second data rate with guaranteed bandwidth and bounded delay for time-critical applications such as interactive video. Perhaps best of all, a large proportion of network owners won't have to rewire their buildings to convert to it. The demand priority protocol, which is explained in the article on page 13, is implemented by intelligent hubs, or repeaters as they're called in the standard. The hubs use a round-robin technique to give all of the nodes in the network a fair opportunity to transmit data. Two priority levels make it possible to give priority service to time-critical applications. For transmission over common unshielded twisted-pair building wiring, the standard's quartet signaling technique takes advantage of the four-pair cable used in most buildings, transmitting at a rate of 25 megabits per second over each pair (see the article on page 18). Techniques for transmission over 25-pair cable, shielded twisted-pair cable, and fiber-optic cable are also specified in the standard. A spectrally efficient coding method called a 5-bit/6-bit block code (page 27) makes it possible to transmit at 25 megabits per second over unshielded cable without exceeding electromagnetic interference regulations. The coding scheme also provides the needed levels of error detection and correction. The article on page 33 tells how 100VG-AnyLAN meets the needs of multimedia applications, and the article on page 39 describes the design of an intelligent hub based on special repeater and transceiver chips.

Optical character recognition or OCR is the process of scanning a hard-copy page and turning it into computer editable text. It's harder than you'd think. Many documents a human can read easily simply can't be converted accurately by an ordinary OCR program. Small or closely spaced text, thin lines, multiple columns, and graphics can mean major trouble for an OCR program. HP AccuPage technology (page 43) provides image processing transforms that are used in OCR software to provide much more accurate text conversion and the ability to capture text and pictures at the same time. HP AccuPage techniques include adaptive threshold determination and upscaling of small text.

The HP S1010A flat panel display is a color liquid crystal display (LCD) monitor that's designed as a lighter, lower-power, smaller-footprint direct replacement for conventional computer monitors—you can unplug a conventional monitor and plug in the HP S1010A. The article on page 51 tells how the analog red-green-blue signals from the computer are converted to digital signals to drive the LCD, and how the 256-level colors from the computer are converted to noise-free 8-level colors for the LCD.

---

---

An application software package is typically composed of many components. Each component requires a design decision: to buy or to build. Experts at HP Laboratories who advise HP product divisions in such decisions have been recommending the use of third-party software whenever it makes sense economically. In the article on page 61, Wes Higaki of HP Laboratories proposes an improved economic model for buy-or-build decisions. The model extends an existing model with an extensive list of costs and benefits, net present value calculations, and estimation techniques for costs and benefits.

If you're a designer of an application-specific integrated circuit (ASIC), your choice of a supplier is critical. Choosing one with unrealistic performance claims can lead to failure, while choosing one with overly conservative claims leads to higher costs than necessary. For an objective way to evaluate ASIC suppliers' performance claims, HP procurement engineers have been using two benchmark circuits. As described in the article on page 66, they collect supplier sample data for the two circuits, calculate technology constants for various parameters, and plot the results to spot optimistic or conservative claims.

R.P. Dolan  
Editor

---

## Cover

The demand priority protocol of the IEEE 802.12 local area network standard ensures that all nodes in a network get a fair opportunity to transmit data, but gives higher priority to time-critical applications, represented here as brightly colored data packets as distinguished from the blue normal-priority data.

---

## What's Ahead

In the October issue, we'll have eight articles on the theory and operation of HP PE/SolidDesigner, a solid modeling software package for mechanical computer-aided design. Other subjects in the issue will be the HP 70910/11A wide bandwidth microwave receiver for extending spectrum analyzer operation and the use of the log weighted average for measuring printer throughput.



# Introduction to 100VG-AnyLAN and the IEEE 802.12 Local Area Network Standard

100VG-AnyLAN provides a 100-Mbit/s data rate with guaranteed bandwidth and maximum access delay for time-critical applications such as multimedia, using existing building wiring. It uses demand priority protocol. Developed by Hewlett-Packard and now supported by over 30 companies ranging from integrated circuit vendors to systems suppliers, demand priority is well on its way to becoming the IEEE 802.12 standard.

by Alan R. Albrecht and Patricia A. Thaler

100VG-AnyLAN is a new, high-speed addition to HP's AdvanceStack local area network (LAN) product group. It is an economically effective upgrade path for congested 10-Mbit/s 10Base-T Ethernet and 4/16-Mbit/s token ring networks. It provides a 100-Mbit/s data rate with guaranteed bandwidth and bounded access delay for time-critical applications, using existing building wiring. This provides high performance for traditional data transfer applications. It also provides emerging multimedia applications, such as interactive video, with the low delays they require. It delivers this performance over the most common networking medium, 4-pair unshielded twisted-pair (UTP) telephone wire.

100VG-AnyLAN uses the demand priority protocol, which was developed as a joint effort by Hewlett-Packard Laboratories in Bristol, England and the HP Roseville Networks Division in California. Now supported by over 30 companies ranging from integrated circuit vendors to systems suppliers, demand priority is well on its way to becoming the IEEE 802.12 standard.

The 100VG-AnyLAN articles in this issue look at the development of the demand priority protocol and the 100VG-AnyLAN product set.

## Local Area Network Technology

Before the initial development of local area networking in the late 1970s, the telephone system was the only generally available data communications option. Bandwidth (3 kHz on a voice-grade line) was clearly a problem and a number of different types of new data networks were proposed. Two, Ethernet<sup>1</sup> and token ring,<sup>2</sup> have emerged to dominate the local area networking market.

**Ethernet (IEEE 802.3).** Ethernet was developed in the late 1970s as a 10-Mbit/s answer to the limitations of the telephone system. It became an IEEE standard in 1985. All nodes were connected to a single central coax bus, which proved to be somewhat inflexible as users change locations or are added to the network.

The Ethernet access policy is CSMA/CD, which stands for carrier sense, multiple access with collision detection. It allows any node to transmit a packet (with up to 1500 bytes of data) anytime it detects silence (no signal) on the network. This can lead to packet collisions if two or more nodes need to transmit and detect silence at the same time. Each involved node is required to back off (cease transmitting) immediately after a collision is detected, but time is consumed and the available bandwidth is effectively reduced during high-traffic periods.

The protocol also requires each node to monitor the network traffic and to decode (filter) the destination address of each packet to determine whether it should be received by the node. Packets with the node's individual or group address are copied into memory and packets with nonmatching addresses are ignored.

The 10Base-T star topology was proposed by Hewlett-Packard in 1987 and became part of the standard in 1990. The center of the star is a network concentrator (hub) which is typically located in a wiring closet. Each node is connected to the hub by voice-grade twisted-pair cable. 10Base-T retains the basic features and access policy of the bus network and also adds a level of fault tolerance. Link faults at individual nodes are isolated by the hub and do not take down the entire network. 10Base-T has become the most common IEEE 802.3 network configuration.

**Token Ring (IEEE 802.5).** Token ring was proposed as a 4/16-Mbit/s solution to the Ethernet collision problem and became an IEEE standard in late 1984. The original network structure is a ring around which both tokens and information packets (up to 4500 data bytes) are passed. The network medium is IBM type 1 shielded twisted-pair (STP) cable. Token ring networks are also now commonly installed in star configurations.

The token ring access policy is designed to be both collision-free and priority-based. It prevents any node that does not currently "own" the token from transmitting a data packet,

and it provides eight priority levels to allow some classes of data to take precedence over other classes.

All data packets and tokens contain an access control field that allows the successive nodes on the network both to reserve the token and to indicate their reservation priority level. The node that currently owns the token transmits its data packet with the reservation bits in the access control field set to minimum priority. Each successive node forwards the packet as it is being received. It also interrogates the destination address field to determine whether it should copy the data frame and the access control field to determine the current reservation level. If the node needs to send a data packet and the reserved priority level is less than the node's level, the node indicates its need by changing the value of the reservation bits in the forwarded packet.

The sending node removes the packet from the network and transmits a new token with the priority bits of the access control field set to the priority level indicated in the returned packet. The token then circulates to the node that first reserved that priority. That node removes the token and transmits a data packet. The token circulates continuously at minimum priority in an idle network.

### The Local Area Networking Market

International Data Corporation (IDC), a market research firm, reports that the worldwide installed Ethernet base at the end of 1993 was 26,376,000 nodes, up 102% from 1992.<sup>3</sup> They predict an installed base of over 75,000,000 by the end of 1995, predominantly 10Base-T.

The token ring worldwide installed base was 6,744,000 at the end of 1993, up 115% from 1992.<sup>3</sup> IDC predicts that the token ring installed base will approach 14,900,000 by the end of 1995.

### Current Network Pressures

The last ten years has seen a hundredfold increase in the speed of computers and the size of files created by sophisticated applications. Meanwhile, the data transfer rate of most networks has remained constant at 10 Mbits/s to 16 Mbits/s.

The first signs of network strain are users complaining that performance is falling off and response times are rising. The cause is almost always congestion—too many users, too much information. Network-connected high-performance desktop systems are intended to give users instant access to any appropriate information, anywhere in the organization, at any time, and this generates high levels of traffic.

Sometimes, network bottlenecks arise from individual applications exceeding the bandwidth of the network. Data-intensive applications, such as database access, image analysis, desktop publishing, network printing, and CAD, require that very large amounts of information be transferred in a single burst. For example, a desktop publishing application might require 10 megabytes for a single page incorporating several typefaces, a bitmapped logo, and four-color graphics. On a typical Ethernet or token ring network, it could take as long as 20 seconds to retrieve that one page. A multipage document could take a minute or two to retrieve.

---

---

## Cable Types

Cables can be categorized in various ways: according to their physical structure, the material used for transmitting signals, and the uses for which they are suitable. Common types are listed below.

**UTP:** Unshielded twisted-pair, 100-ohm balanced cable. The lack of shielding makes UTP cable very low-cost, but introduces problems of cross talk when the pairs are in close proximity.

**Category 3:** Voice-grade cable, such as telephone wire, with 16-MHz bandwidth, used in 4-pair groups for each link. 25-pair bundles of Category 3 UTP are common in existing LANs. This is an important consideration when designing a network protocol.

**Category 4:** 20-MHz bandwidth, used in 4-pair groups for each link.

**Category 5:** Data-grade cable with 100-MHz bandwidth, used in 2-pair or 4-pair groups for each link.

**Optical-Fiber:** Cable consisting of a minimum of two strands of optical fiber running parallel within a protective jacket. Each fiber is usually composed of glass 125  $\mu\text{m}$  in diameter, and has a 62.5- $\mu\text{m}$  core. Transmission is by light beam at 850-nm or 1330-nm wavelength.

**STP:** 150-ohm balanced shielded twisted-pair cable. Usually used in 2-pair groups for each link.

---

---

The situation will rapidly worsen with the accelerated development of time-sensitive multimedia applications. Real-time audio and video for video conferencing and interactive video require that packets of data be transferred continuously with minimal delay. They cannot afford to have any packet delayed or dropped because of a collision or congestion on the network.

### Design Goals

The problem presented to HP's network design team contained several major goals and considerations:

- **Speed.** The current networks are clearly too slow. A major improvement would be to increase the speed of the network. 100 Mbits/s would allow ten times the amount of traffic.
- **Guaranteed Access.** While multimedia and other time-sensitive applications require the increased total bandwidth that 100 Mbits/s would provide, they also need guarantees that information will get through within a stated delay window, whatever the traffic on the network.
- **Cost.** Existing networks have already required major investments in the wiring structure. The new network should be able to operate over generic twisted-pair building wiring. Fiber-optic cabling should also be allowed.
- **Topology.** To be compatible with existing wiring, the new network must use a star topology. The allowed network diameter should be at least 2.5 km with three or more levels of hub cascading.
- **Software Compatibility.** The network should be compatible with both Ethernet and token ring frame formats and should preserve existing investments in network and applications software.



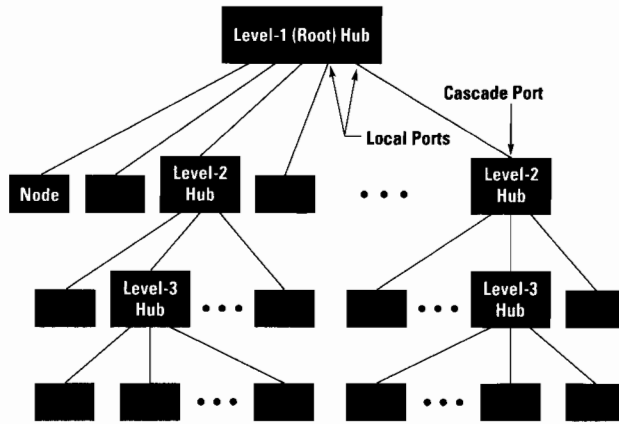


Fig. 1. A cascaded star network.

- **Error Susceptibility and Detection.** The implementation should have a physical layer bit error rate of less than  $10^{-8}$  and a coding scheme that will guarantee detection of errors in any three bits within the data frame without compromising the cyclic redundancy check (CRC) defined for the IEEE 802.3 and 802.5 frame formats.
- **Privacy.** Both Ethernet and token ring provide address filtering at the node level, effectively allowing any node to copy any data packet sent on the network. To enhance the privacy of data communications, address filtering of individually addressed packets should be provided as an option within the hub.
- **Robust Operation.** Continuous operation of the network is required. Physical connections should be tested before allowing nodes to enter the network and provision should be made to allow the identification and removal of disruptive nodes.
- **Network Management.** An optional network management capability should be provided to monitor network performance, isolate faults, and control network configuration.

### Demand Priority and 100VG-AnyLAN

The protocol for the new network is demand priority. It combines the best characteristics of both Ethernet (simple, fast access) and token ring (strong control, collision avoidance, and deterministic delay).

The HP product is 100VG-AnyLAN. It is based on chip technology developed by HP and AT&T that delivers an effective data rate of 100 Mbits/s over several different link configurations: 4-pair Category 3 unshielded twisted-pair (UTP) cable, 2-pair shielded twisted-pair (STP) cable, or 2-pair fiber-optic cable. UTP, STP, and fiber-optic cable can be intermixed on the same network. 100VG-AnyLAN can also operate with the 25-pair bundled cable that is used in many 10Base-T networks, as long as hubs are not connected through bundled cable.

### Network Topology

The basic topology used by 100VG-AnyLAN networks is the star configuration. Each hub has two or more local ports and can optionally have one cascade port for connection to a higher-level hub. Nodes can be user stations, bridges to other networks, LAN analyzers, or lower-level hubs.

A network can contain several levels of hubs interconnected in a cascade as shown in Fig. 1. The topmost hub is designated as the level 1 (root) hub. Hubs in lower levels of the cascade are designated by the number of links between them and the root hub. The level number of any particular hub can be determined by the equation:

$$\text{Hub Level} = 1 + (\text{number of link segments away from the root hub})$$

Hubs at the same cascade level have the same level number.

The maximum number of nodes that can be connected to the network is dependent on the level and frequency of traffic each node generates. The maximum topology diameters and the number of levels of cascading are limited by the allowable delay between the node and the root hub, and can be calculated for any proposed configuration. When there is only one intermediate hub between the root hub and the node, for example, the maximum distance between a node and the root hub is 4 km. Each additional intermediate hub reduces the hub-to-node distance by 1.0 km, resulting in a maximum of four intermediate hubs and a root-hub-to-node distance of 1 km.

Local hub-to-node distances depend on the type of media used for the link: 4-pair UTP links configured with Category 3 or Category 4 cable and links configured with STP cable should not exceed 100 m. Category-5 links can be up to 150 m long. Fiber-optic links can be even longer: 500 m with 850-nm transceivers and 2000 m with 1300-nm transceivers.

### Demand Priority Protocol

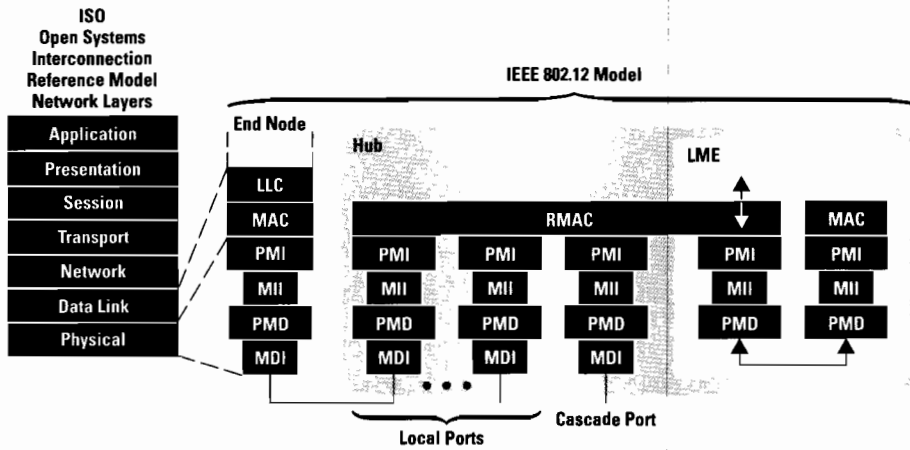
Control of a demand priority network is centered in the hubs and is based on a request/grant handshake between the hubs and their connected nodes. Access to the network is granted by the hub to requesting nodes in a cyclic round-robin sequence, based on the priority of the request. Within a priority level, selection of the next node to transmit is determined by its sequential location in the network rather than the time of its request. Data is encoded before transmission and is checked for errors at each intermediate hub and the receiving node. Either IEEE 802.3 or IEEE 802.5 frame format can be used.

### Architectural Model

The demand priority protocol contains four sublayers corresponding to the two lower layers of the ISO Open Systems Interconnection (OSI) reference model shown in Fig. 2.

The functions of the OSI data link layer are implemented in two sublayers: the LLC and demand priority MAC sublayers. The upper sublayer in a network node is the IEEE 802.2 logical link control (LLC) sublayer. The media access control (MAC) sublayer provides data formatting and control of packet transmission (or reception) in the transmitting (or receiving) node. The MAC also initiates outgoing control requests and acts on received control indications.

Each hub provides control of its connected star portion of the network. The RMAC sublayer provides a superset of the functions of the node's MAC sublayer (except frame formatting). It selects which node will next be granted permission



**Fig. 2.** The demand priority architectural model. LME is the layer management entity.

to send a packet, determines where the received packet will be sent, provides local control of packet reception and retransmission, and monitors each connected link for proper operation. The RMAC initiates outgoing control requests and acts on received control indications. There is no LLC sublayer in a hub.

The functions of the physical layer of the OSI model are also provided by two sublayers: the PMI and PMD sublayers. The physical medium independent (PMI) sublayer presents the same logical interface to both the MAC in a node and the RMAC in a hub. It provides the functions that are common to all link media: ciphering and encoding data and inserting stream headers and trailers before transmission, and removing stream headers and trailers and decoding and deciphering data during reception.

The physical medium dependent (PMD) sublayer contains the functions that are dependent on the particular link medium: bit encoding, signal conditioning, and multiplexing (if necessary) before transmission, and signal recovery, demultiplexing (if necessary), and bit decoding during reception. The PMD translates control requests and generates outgoing control signals. It also detects control signal transitions and generates the appropriate control indication for the MAC or RMAC.

The medium independent interface (MII) is defined in the draft standard as an optional, physically exposed connection to allow link configuration interchangeability (for example, changing from 4-pair UTP to a fiber link). The medium dependent interface (MDI) is the connector between the PMD and the link media.

### Standards Development

The demand priority protocol is currently in the process of becoming the IEEE 802.12 standard. The following milestones had been accomplished by May 1995:

- November 1992. A proposal was made for a demand priority development project.
- July 1993. IEEE established the 802.12 Demand Priority Working Group.
- November 1993. An initial draft document (D1) was submitted to the 802.12 Working Group for review.
- July 1994. The IEEE 802.12 draft standard (D4) was submitted to the Working Group for ballot.

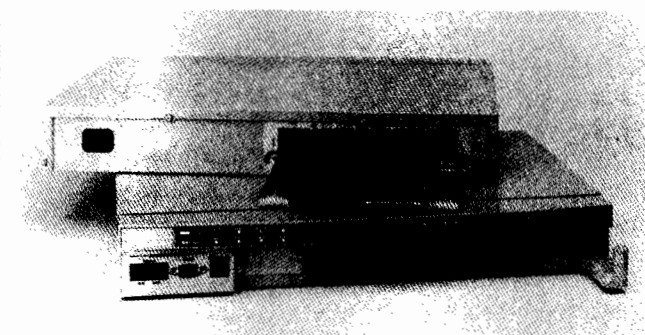
- January 1995. The IEEE 801.12 draft standard (D7) was submitted to the LAN/MAN Standards Committee for sponsor ballot.
- March 1995. Sponsor ballot was successfully completed.
- May 1995. The IEEE 802.12 draft standard (D8) was submitted to the Review Committee of the IEEE Standards Board. It is anticipated that the Review Committee and the Standards Board will vote on approving IEEE 802.12 in June.

As the base IEEE 802.12 standard approaches completion, the 802.12 Working Group is investigating future enhancements. At the March 1995 meeting, study groups were formed to investigate higher-speed operation (0.4 to 4 gigabits per second) and a PMD for 2-pair Category 5 UTP. Also under discussion are redundancy and full-duplex links.

### 100VG-AnyLAN Products

HP currently offers a variety of products for 100VG-AnyLAN within the HP AdvanceStack networking family:

- HP AdvanceStack 100VG Hub15 (HP J2410A) is a 15-port 100VG-AnyLAN hub (Fig. 3).
- HP AdvanceStack 100VG SNMP/bridge module (HP J2414A), when installed in the expansion slot of 100VG Hub15, adds SNMP network management and bridging to 10-Mbit/s Ethernet networks.



**Fig. 3.** The HP AdvanceStack 100VG Hub15 (HP J2410A) hub is a 15-port 100VG-AnyLAN hub.

## Other Network Technologies

**FDDI.** Fiber Distributed Data Interface (FDDI) networks run at a standard 100-Mbit/s transfer rate using multimode optical-fiber cabling. The purchase cost is high, partly because it necessitates new cabling for existing networks. FDDI is also available as a high-speed backbone network connecting LANs.

A twisted-pair version, TP-PMD (twisted-pair physical medium dependent) FDDI, is under development, using Category 5 UTP and STP copper cables. This version is sometimes called Copper Distributed Data Interface (CDDI).

**ATM.** Asynchronous transfer mode (ATM) is a new network technology particularly suitable for wide area networks and campus backbones. It is intended to allow seamless integration of campus LAN backbones into the wide area network.

ATM uses cell switching (53 bytes per cell) similar to high-speed telephone switching over existing UTP or optical-fiber cabling. It runs at 25 to 622 Mbits/s.

**Table I**  
Cabling and Topological Comparisons

	FDDI	ATM	100Base-T	Demand Priority
100-Mbit/s Category 3 cable supported?	No	50 Mbits/s with complex coding	Yes	Yes
Bundled cables supported?	No	No	No	Yes
Multiple cascades supported without bridging or routing?	Yes	Yes	No	Yes
Cost	High	High	Medium	Low

**Ethernet and Token Ring Switching.** Switching is also being introduced into Ethernet and token ring networks. To increase overall throughput, LANs are segmented. Hubs switch packets dynamically between connected segments allowing simultaneous transmissions among pairs of network segments. This increases bandwidth by two or more times that of individual segments.

**100Base-T.** 100Base-T is a scaling of CSMA/CD to 100 Mbits/s. There is no migration path or accommodation for existing token ring users. The technique cannot emulate 10Base-T topologies since the maximum topology is two repeaters.

**Table II**  
System Comparisons

	FDDI	ATM	100Base-T	Demand Priority
Supports multimedia with guaranteed delay and bandwidth?	Yes	Yes	No	Yes
End-node adapter card complexity	Node management is expensive	Segmentation or re-assembly of frames is expensive	Low	Low
Ethernet 802.3 networks can be upgraded without software changes?	No	No	Yes	Yes
Token ring 802.5 networks can be upgraded without software changes?	No	No	No	Yes

- HP 10/100VG selectable ISA, EISA, and PCI adapters (HP J2573A, J2577A, and J2585A) are PC LAN adapter cards with one RJ-45 connector for 10-Mbit/s 10Base-T and another for 100-Mbit/s 100VG-AnyLAN. The adapters automatically sense which network they are connected to and select the correct mode of operation.
- HP 100VG-AnyLAN/9000 (HP J2645AA, J2655AA) are adapters for HP 9000 Series 700 workstations. They support connection to either 10-Mbit/s 10Base-T or 100-Mbit/s 100VG-AnyLAN. The adapters automatically sense which network they are connected to and select the correct mode of operation.
- HP 100VG-AnyLAN development system (HP E2463A) is a development system for designers and operators of 100VG-AnyLAN network products. It finds the root cause of any IEEE 802.12 design and interoperability problems.

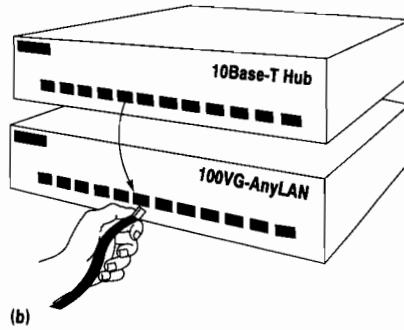
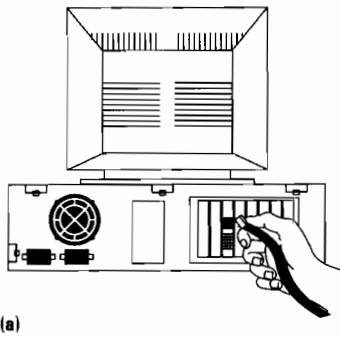
### Upgrading Existing Networks

The current 100VG-AnyLAN product set provides a smooth, step-by-step way for customers who wish to upgrade their existing 10Base-T Ethernet networks. The only required elements are new hubs for the network and new adapter cards for each node. Existing network management interfaces, operating systems, bridges, and routers can remain.

In most cases, not all of the network will have to be upgraded at the same time. Consider, for example, an existing 10Base-T network with a mix of high-traffic and low-traffic users that all need to access file servers and printers. The first phase is to identify the high-traffic users to determine how many nodes need to be upgraded. This will determine the number of 100VG-AnyLAN hub ports and network adapters that will be required (including network file servers and printers).

The second phase is to acquire the necessary 100VG-AnyLAN hubs and network adapter cards (one hub for each 15 nodes if HP AdvanceStack 100VG-AnyLAN hubs are used), and to install them in the network as depicted in Fig. 4. The link cable is disconnected from each node that has been identified for upgrading, a replacement PC LAN adapter card is installed, and the link cable is reconnected.

100VG-AnyLAN hubs are installed adjacent to the existing 10Base-T hubs that service high-traffic users. The network cables to each upgraded node are disconnected from the 10Base-T hub and connected to a 100VG-AnyLAN hub (some rearrangement of the 10Base-T node-to-hub connections



**Fig. 4.** Upgrading a 10Base-T network to 100VG-AnyLAN. (a) Upgrade adapter at the desktop. (b) Upgrade hub in the wiring closet.

may be advisable). A bridge module between the 100VG-AnyLAN and 10Base-T root hubs interconnects the two LANs.

The network is again ready for use with an upgraded topology as shown in Fig. 5. Additional 100VG-AnyLAN hubs and adapters can be added as needed to accommodate changing traffic levels in the remaining 10Base-T nodes.

### The Future

For 100VG-AnyLAN to become a pervasive network topology, the concept of the protocol must be adopted by three major groups: network product manufacturers, network system designers, and network users. Customers must be able to obtain complete sets of the products that are required to meet their individual needs, and the products they obtain should be compatible with each other even if they are obtained from different companies.

An early measure of 100VG-AnyLAN's acceptance can be seen by the number of organizations involved in the demand priority standards development, and in the number adopting the protocol. By January 1995, the following companies had already begun delivering or had announced intention to deliver 100VG-AnyLAN products:

- IC chipsets. AT&T, Motorola, Texas Instruments, Applied Micro Circuits, and Pericom Semiconductor.
- Hubs. HP, Thomas-Conrad, Chipcom, NEC, MultiMedia LANs, Compex, Alfa, Katron Technologies, Optical Data Systems, Anritsu, D-Link, and Ragula Systems.
- Node Adapters. HP, Thomas-Conrad, NEC, Compex, Alfa, Katron Technologies, Ragula Systems, Anritsu, D-Link, Interphase, Optical Data Systems, and Racore Computer Products.

- Multimedia Networking Software. Starlight Networks.
- Development Systems. HP and AT&T.
- Switches and other Internetworking Products. Cisco, Compex, Newbridge Networks, Optical Data Systems, and Plain Tree Systems.
- Desktop Systems. IBM, HP, Compaq, and NEC.

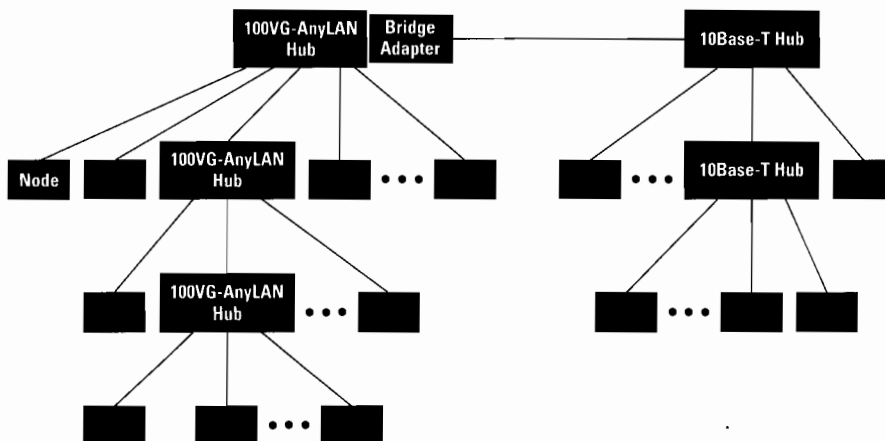
The current 100VG-AnyLAN products provide customers with ten times the speed and up to 16 times the throughput of a 10Base-T Ethernet network at about twice the price. Development of VLSI devices that provide greater integration of the physical layer and the MAC should lead to even lower-cost products.

Demand priority has been designed to be architecturally independent of any particular implementation technology. As such, future generations of demand priority networks may provide higher data rates.

### Other Articles in this Issue

The following 100VG-AnyLAN articles provide more detail for their respective areas:

- Demand Priority. The article on page 13 introduces the round-robin pointers, priorities, and bandwidth allocation capabilities of the protocol. A typical demand priority transmission is described in step-by-step fashion, and results of performance simulations are provided.
- Physical Signaling. The article on page 18 gives an expanded description of the physical sublayer. Several design decisions leading to the development of demand priority as a replacement for 10Base-T are described. Quartet signaling, cross talk avoidance, and control signal generation and detection are explained. Differences between the 4-pair UTP and the STP and fiber-optic PMD and link requirements are discussed.



**Fig. 5.** Hybrid network consisting of interconnected 100VG-AnyLAN and 10Base-T networks.

- **Coding.** The article on page 27 looks at the techniques of coding data to provide error detection and electrical balance while optimizing the efficiency of the network. Additional insight into the reasons behind the decisions to include both data ciphering and encoding are given along with an expanded discussion of the properties of the coding scheme.
- **Multimedia Requirements.** The article on page 33 examines the specific demands multimedia applications make on any network system and shows how 100VG-AnyLAN addresses them.

#### **Acknowledgments**

The development of 100VG-AnyLAN from public proposal to draft standard and product release has taken only two years. This is an incredibly short time, and has involved almost Herculean efforts of the Roseville and Bristol teams. 100VG-AnyLAN was able to progress from investigation to product in such a short time because of the efforts of a fantastic team. We would particularly like to acknowledge the contributions of our fellow inventors: Alistair Coles, Simon Crouch, David Cunningham, Joe Curcio, Dan Dove, Steve Goody, Jonathan Jedwab, Michael Spratt, and Greg Watson. Although the two halves of the team live 5000 miles apart, these engineers, mathematicians, and physicists collaborated closely to produce a high-performance network.

Bill Lane, a professor emeritus from the California State University at Chico, contributed his skills as an editor and writer to produce the text for the draft standard and assisted in the editing of several articles in this issue of the HP Journal.

Completing a major development in a short time requires more than individual commitment. Thus, we also thank our management on both sides of the Atlantic—David Dack, Mark Gasteen, Robert Gudz, Dave Harris, Bill Kind, Gary McAnally, Tim McShane, Carolyn Ticknor, and Steve Wright—for their trust and wholehearted support. They always ensured that we had what we needed to succeed.

#### **References**

1. *Information Technology—Local and Metropolitan Area Networks—Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Specifications*, ISO/IEC 8802-3:1993.
2. *Information Technology—Telecommunications and Information Exchange between Systems—Local and Metropolitan Area Networks—LAN/MAN Specific Requirements—Part 5: Token Ring Access Method and Physical Layer Specifications*, ISO/IEC DIS 8802-5:1992.
3. *Local Area Networks—Analysis of the Intelligent Hub Market*, International Data Corporation, May 1994.

# Demand Priority Protocol

In multiple-hub networks, demand priority ensures fairness of access for all nodes and guarantees access time for multimedia applications.

by Alan R. Albrecht, Michael P. Spratt, Patricia A. Thaler, and Gregory C. A. Watson

Demand priority is the media access control protocol defined in the IEEE 802.12 draft standard. It is called 100VG-AnyLAN by HP. Various techniques are used to ensure fairness of access for all nodes and to guarantee access time for individual nodes. Round-robin selection procedures are used to give each node an equal opportunity to transmit data. Two priority levels are provided so that time-critical traffic such as interactive video, audio, and multimedia can be given priority service with guaranteed low delay. A bandwidth allocator can be introduced to control the amount of bandwidth each application can use.

These techniques make demand priority able to accommodate multimedia applications with guaranteed performance levels, while allowing normal traffic to use the remaining bandwidth in a fair manner.

## Hubs

100VG-AnyLAN networks are centered around the concept of intelligent hubs (called repeaters in the IEEE 802.12 draft standard), each of which is at the center of a star configuration. A hub has several local ports, which are connected to individual nodes, which can be workstations, servers, bridges, routers, or lower-level hubs. A hub may also have one cascade port for connection to a higher-level hub.

## Network Control

Control in a demand priority network is based on a request/grant handshake between the node and the hub. Each node needing to send a frame indicates its need to transmit by sending a request signal to the hub and waiting for the hub to grant it permission to transmit the frame. The concept can be seen as a sophisticated and flexible successor to the token in token ring networks; it ensures that collisions cannot occur in the network.

Hubs use a round-robin node selection procedure to ensure a fair opportunity for nodes to transmit data. The purpose is to ensure that no node can send two successive normal-priority frames until all other nodes have also had an opportunity to send a frame at normal priority.

## Round-Robin Node Selection

The nodes are numbered in order around the entire network. The hub maintains a round-robin pointer that indicates which node is next to receive an opportunity to transmit. For example, if nodes 2, 3, 5, and 9 have indicated that they have packets to send and the round-robin pointer currently points to 3, node 3 gets to transmit. The pointer then moves on to 4. Assuming node 4 still has no data to send, node 5 will be selected next. The pointer will then move on

to 6. When the pointer reaches the last node, it cycles back to 1.

The normal-priority selection cycle will be temporarily suspended when any node lodges a high-priority request. A separate round-robin pointer is maintained for high-priority requests, so that they too are treated fairly. The decision as to which node is next is made immediately after a transmission has finished. The hub looks at all requests that have been received since it last made a selection. Earlier requests are not stored so that spurious requests caused by noise or previous requests from a node that has changed its mind do not cause a false selection.

## Typical Transaction

A typical transaction within a network using the demand priority protocol could run as follows. The example is for a network with a single hub, using 4-pair UTP cables. 4-pair UTP links require all four pairs to be used for data transmission in either direction, but dedicate two pairs in each direction for control when data is not being sent. Fig. 1 illustrates each step in the transaction.

While there is no data to be transmitted, all nodes send IDLE to the hub, and the hub returns IDLE to each node.

Suppose that node 1 has a frame to transmit to node 3 alone. It sends a request (REQ\_N for normal priority, REQ\_H for high priority) to the hub.

After referring to its round-robin pointer and checking for any other request, the hub decides to accept transmission from node 1. It indicates the selection by ceasing its IDLE transmission to node 1 (indicating a grant). This clears the outgoing signal from the hub and readies the link to receive data on all four pairs. Simultaneously, the hub also changes the control signal being sent to all other nodes from IDLE to INCOMING. (At this stage, the hub has no idea of the destination of the data packet.)

The other nodes respond to INCOMING by ceasing transmission to the hub. This clears their outgoing control signal so that they are prepared to receive the packet. The source node (node 1) starts transmitting its packet to the hub on all four pairs of wires as soon as it detects that the IDLE control signal has gone silent.

The hub decodes the beginning of the packet to determine the destination node or nodes. It finds that, in this case, the packet is unicast (destined for one node only) for node 3. It immediately starts sending the packet to node 3 and changes the control signal to all other nodes to IDLE.



Transmission continues until it is complete. Meanwhile, nodes not involved in the transfer can send REQ if they have packets to send. The hub notes such requests, but does not yet select the next node to transmit. When the source node (node 1) completes its transmission, it sends IDLE, or another REQ if it has another packet to send. After the hub has finished forwarding the frame, it selects the next node to transmit according to the round-robin selection procedure.

There is a possibility that a node that has received several consecutive packets and needs to make a transmission of its own might miss its turn in the round-robin cycle. This is because at the instant when the hub selects the next node to transmit, the receiving node is still receiving the end of the last packet. To overcome this, receiving nodes needing to transmit ignore the INCOMING control signal for a short period (called the *request window*) during which they submit their own request. The request window is shorter than the minimum gap between successive packets.

If the packet has a multicast address (it is destined for more than one node), the hub may need to wait until it has received the entire packet before transmitting it to the destination nodes. This is to avoid cross talk problems in 25-pair bundled cable. The article on page 18 provides more detail on 4-pair UTP links and the use of bundled cable.

In a network using 2-pair STP or fiber-optic cables, packets are transmitted down one fiber or pair of wires by multiplexing the four data streams. This means that the other fiber or wire pair is always available for lodging requests to the hub. See the articles on pages 18 and 27 for an expanded discussion of STP and fiber links.

### Networks with Cascaded Hubs

A network can have several hubs in a tree structure known as a cascade (see Fig. 1 on page 8). One hub is designated

the *root hub*. All other hubs will each have a connection to a next higher-level hub through their cascade port. They may also have lower-level hubs connected to some of their local ports.

Where there is more than one hub in a network, at any moment exactly one of them has *control* of the network activity. A hub is said to have control if it is selecting the next node to transmit. When the entire network is idle, the root hub has control.

If a network consists of a cascade of hubs, the round-robin pointers effectively treat all end nodes as if they were connected to a single hub, so that all nodes have an equal opportunity to transmit, however far they may be from the root hub. Fig. 2 shows the order of node numbering in a cascaded network.

A different control signal is needed when a hub receives a high-priority request while another hub is servicing normal-priority transactions. Suppose hub A is servicing normal-priority requests when hub B receives a high-priority request (see Fig. 3). Hub B sends the high-priority request up the cascade of hubs until it reaches hub X, which earlier passed control to hub A. Hub X then sends an ENABLE\_HIGH\_ONLY signal to hub A.

The ENABLE\_HIGH\_ONLY signal tells hub A to suspend its round-robin sequence at the end of the current transmission and return control to hub X. At the same time hub A tells hub X whether it had finished its normal-priority round-robin cycle. Hub X then passes control to hub B to service the high-priority request. When that is complete, assuming that hub A had not finished its portion of the round-robin normal-priority selection cycle, and assuming that no other high-priority requests have been received by hub X, control is passed back to hub A.

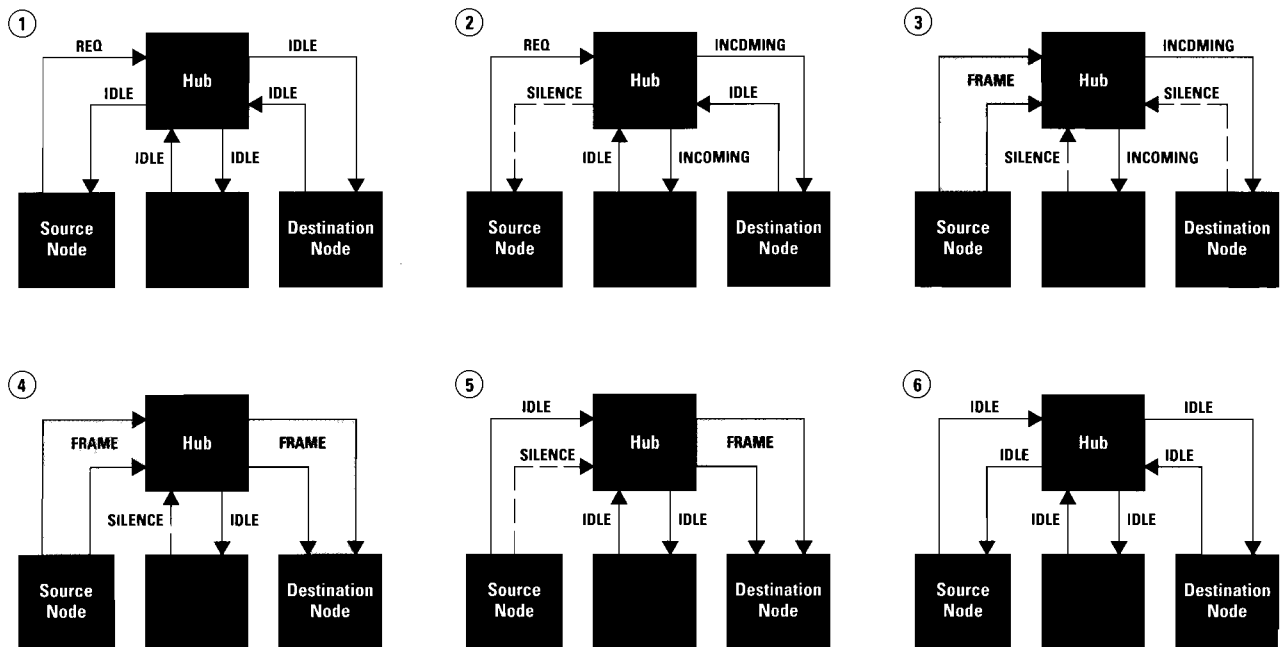


Fig. 1. The stages in a typical demand priority transaction.

## Network Protocol Layers

The Open Systems Interconnection (OSI) Reference Model of the International Organization for Standardization (ISO) describes the structure of networks. It defines the seven layers shown at the left side of Fig. 2 on page 9. The lower three layers are relevant to the demand priority protocol.

**Layer 3: Network Layer (NL).** The network layer is responsible for passing a packet of data through an internetwork, which can consist of many individual LANs and even wide area links.

**Layer 2: Data Link Layer (DLL).** The data link layer provides the transmission of data between two nodes on the same network. It receives a packet from Layer 3, the network layer, and adds source and destination addresses and other information to it. The DLL consists of two sub-layers: the logical link control (LLC) sub-layer and the media access control (MAC) sublayer.

The LLC sublayer links the network layer to the MAC. There are standard LLCs, enabling different protocols to link successfully. In 100VG-AnyLAN, the LLC can be either an IEEE 802.2 Class I LLC, supporting Type 1 unacknowledged, connectionless-mode transmission, or a Class II LLC, supporting Type 2, connection-mode transmission.

When a frame is ready for transmission in an end node, it is sent from the LLC sublayer to the MAC sublayer where the appropriate Ethernet or token ring MAC frame is built. The frame is then passed to the physical layer's PMI (physical medium independent) sublayer.

When a data packet is received from the physical layer, the MAC sublayer reassembles the MAC frame and performs various checks for errors in the received frame. Only valid frames are sent on to the LLC sublayer in an end node.

Frames received by the hub's RMAC sublayer are forwarded to the addressed destination (if it can be determined) and to all promiscuous ports regardless of

error condition (there is no LLC in a hub). Frames containing errors are marked with an invalid packet marker.

**Layer 1: Physical Layer (PHY).** The physical layer defines the procedures and protocols associated with the physical transmission of bits (such as cable interfaces, data signal encoding, and connector types and pinouts). It consists of two sublayers: the PMI sublayer and the physical medium dependent (PMD) sublayer.

The PMI sublayer includes provisions for quartet signaling, data ciphering, 5B/6B encoding, the addition of preambles, and start and end frame delimiters (SFD and EFD). See the articles on pages 18 and 27 for details.

During transmission, the PMI sublayer accepts data from the MAC sublayer and prepares the packet for transmission. It converts the octet data into quintets which are separated into four streams. In each stream, each quintet is ciphered and then encoded as a 5B/6B sextet. The PMI adds physical layer headers and trailers to each data stream.

When receiving data packets, the PMI removes physical layer headers and trailers and then passes the packet onto the MAC layer. It decodes each received 5B/6B sextet, decipheres the resulting quintet, merges the four deciphered quintet streams, and converts the result into a single octet stream for delivery to the MAC sublayer.

The PMI sublayer connects with the PMD sublayer through the medium independent interface (MII).

The PMD sublayer includes channel data packet multiplexing (for 2-pair STP and fiber-optic cabling only), NRZ encoding, link medium operation, and link status control. It connects with the physical medium (the cable) through the medium dependent interface (MDI).

### Link Initialization—Joining the Network

When a node first joins the network, a handshake training sequence occurs between it and the hub, during which the node sends the hub its 48-bit MAC address. It also sends other information, such as what type of frame it will use (IEEE 802.3 or IEEE 802.5) and if it wants to receive all packets of data whether addressed to it or not (promiscuous mode). The hub accepts or rejects this information. For example, the hub might be configured by the network administrator to reject promiscuous mode to preserve a high level of data privacy.

### Guaranteed Performance

The two priority levels make it possible to guarantee bandwidth to applications and to keep the access delay (the time for which nodes may have to wait before being allowed to transmit) within bounds.

Bandwidth and access delay depend on the size of the network and the way it is configured. Nodes using high-priority traffic may need to be configured so that the amount of bandwidth they can use is restricted, but nodes using only normal-priority traffic will operate totally unaware that a high-priority service is being provided to other nodes.

### Bandwidth

Without any other form of control, nodes wishing to send high-priority traffic will automatically be allowed to send an equal number of high-priority frames. Any bandwidth not used by this high-priority traffic is then automatically shared

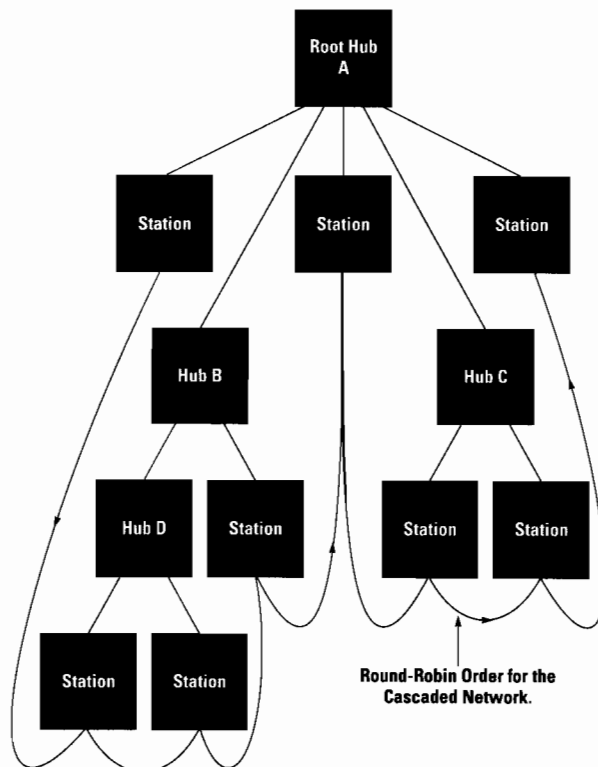


Fig. 2. The numbering of nodes in a cascade of hubs.

between all nodes wishing to send normal-priority traffic. Clearly, where there is excessive high-priority traffic, the bandwidth allowed for high-priority requests needs to be restricted in some way so that normal-priority traffic is never completely stopped. Ways of doing this are described in the article on page 33.

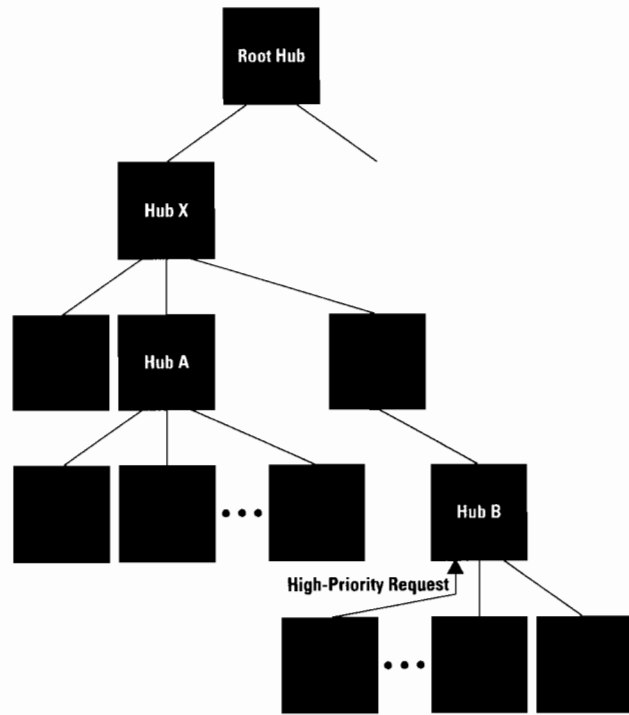
### Access Delay

The deterministic behavior of the round-robin selection procedure means that it is easy to establish a reasonably accurate estimate of the maximum access delay that a high-priority packet will experience. For a hub with  $n$  nodes using high-priority traffic, the worst-case delay will be  $nT$ , where  $T$  is the time it takes to transmit the largest possible frame. It is not  $(n - 1)T$  as might be expected, because the worst case occurs when a high-priority request is received from all nodes just at the moment when the hub starts to service a normal-priority request. For example, the worst-case delay for a hub with 32 nodes forwarding 1500-byte IEEE 802.3 frames is 4 ms, or for 4500-byte IEEE 802.5 frames, 12 ms.

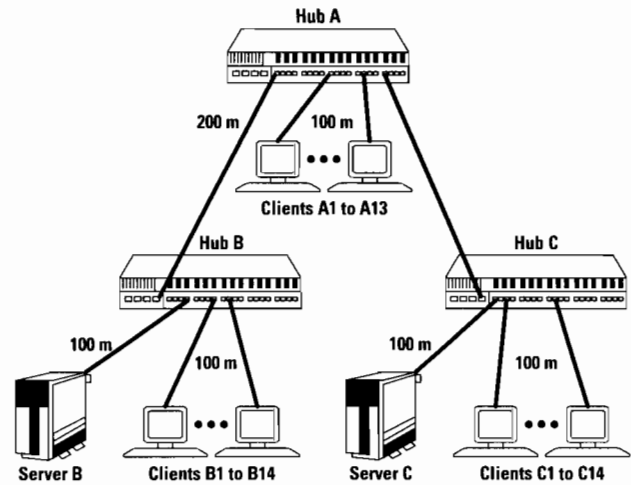
Access delay for a normal-priority request is more difficult to calculate, but the performance data presented in the next section shows what delays might be expected under heavy load.

### Performance in a Simulated Example

Various simulations have been run to examine the effect on high-priority traffic as normal-priority traffic increases. Each



**Fig. 3.** In a cascade of hubs, a hub (hub B) may receive a high-priority request while another hub (hub A) is servicing normal-priority transactions. Hub B sends the high-priority request up the cascade of hubs until it reaches hub X, which earlier passed control to hub A. Hub X then sends an ENABLE\_HIGH\_ONLY signal to hub A.



**Fig. 4.** Network configuration used in the simulated example.

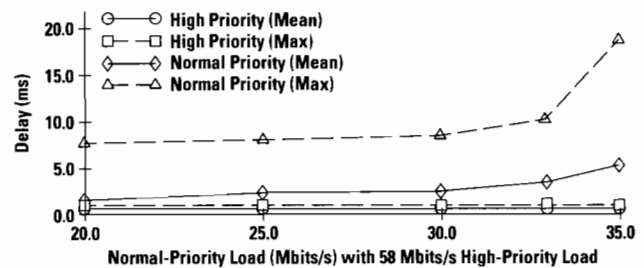
simulation ran for 100,000 frames, and the results were taken from the central portion of the run to avoid anomalies that occur at startup and shutdown.

The network configuration used for one of the simulations is shown in Fig. 4. All fifteen ports are used in each of the hubs. The hubs are connected to each other by 200-m links and the end nodes are connected to the hubs by 100-m links. Hubs B and C each have three nodes that are allocated high-priority bandwidth so that each node can send a block of eight maximum-sized frames every 10 ms. The bandwidth assigned to each node is equivalent to about 9.7 Mbits/s, and is substantially more than needed for MPEG-encoded video. The total high-priority traffic is therefore about 58 Mbits/s.

All stations, including the six with high-priority traffic, send some maximum-size frames at normal priority. The simulation measured the time for which nodes waited for access as the number of these normal-priority frames increased.

Fig. 5 shows the mean and maximum access delays that were observed for both high-priority and normal-priority traffic. Access delay is defined as the time a frame spends at the head of its node's transmission queue waiting for access to the network.

It is clear from the graph that the delay that high-priority traffic encountered was almost independent of the amount of normal-priority traffic. Even when the total load on the network was 93 Mbits/s (35 Mbits/s at normal priority in



**Fig. 5.** Access delays in the simulation.

addition to the 58 Mbits/s of high-priority load), high-priority traffic still had a mean access delay of less than 0.5 ms and a maximum access delay of less than 0.8 ms.

These delays are, of course, dependent on the number of stations that have high-priority traffic, but it is clear that a number of high-priority streams can be supported with a guaranteed low delay.

#### **Acknowledgments**

We would particularly like to acknowledge the contributions of our fellow team members from Roseville Networks Division: Joe Curcio, Dan Dove, and Steve Goody.

#### **Bibliography**

Additional information on the demand priority protocol is contained in the following papers:

1. A. Albrecht, J. Curcio, D. Dove, S. Goody, and M. Spratt, "An Overview of IEEE 802.12 Demand Priority," *Proceedings of GLOBECOM*, 1994.
2. G. Watson, A. Albrecht, J. Curcio, D. Dove, S. Goody, J. Grinham, M. Spratt, and P. Thaler, "The Demand Priority MAC Protocol," *IEEE Networks Magazine*, January 1995.

# Physical Signaling in 100VG-AnyLAN

A physical layer has been developed for demand priority local area networks that accommodates different cable types by means of different physical medium dependent (PMD) sublayers. The major goal was to provide 100-Mbit/s transmission on existing cables, including Category 3, 4, and 5 UTP, STP, and multimode optical fiber.

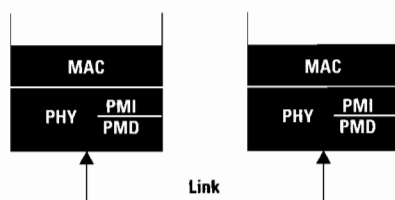
by **Alistair N. Coles, David G. Cunningham, Joseph A. Curcio, Jr., Daniel J. Dove, and Steven G. Methley**

The physical layer (PHY) of a 100VG-AnyLAN demand priority local area network (LAN) acts as an interface between the MAC (media access control) and the link (the cable), adding control signaling and data formatting to the MAC frame when necessary (see Fig. 1).

Several goals were identified in the early stages of the demand priority PHY development. First, the PHY should be as simple as possible, easy to implement, and above all, low in cost. Second, the PHY should provide robust data transfer. LAN performance deteriorates if multiple retransmissions of packets are necessary because of errors. Errors can occur if a PHY does not provide sufficient immunity against noise on the transmission medium (such as impulse noise on unshielded cable running close to switching gear). Typically a LAN is required to operate with less than one error in  $10^8$  bits.

Third, the PHY should support a range of existing media types. 10Base-T LANs operate over voice-grade, or Category 3, unshielded twisted-pair (UTP) wire. More recently, higher-quality UTP (Categories 4 and 5) has been specified and is now being used in new installations. Shielded twisted-pair (STP) has been used extensively in token ring LANs, although recently these too have been connected with UTP. Multimode optical fiber is also being used increasingly.

Fourth, the PHY should be capable of data transfer at 100 Mbits/s. Finally, when not transmitting data, the PHY should be capable of signaling five independent control states from one end of a link to another. These control signals are required for the operation of the demand priority protocol.

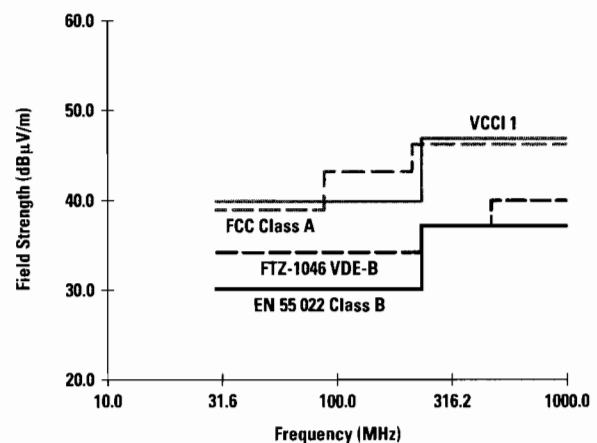


**Fig. 1.** Network layer model, showing the media access control (MAC) and the physical layer (PHY) with its physical medium dependent (PMD) and physical medium independent (PMI) sublayers.

The following constraints were also placed on the design of the demand priority PHY:

- **EMC.** The transmission techniques used for the PHY must not cause radiated emissions from the LAN equipment or cabling that would violate electromagnetic compatibility (EMC) regulations as shown in Fig. 2 (for example, FCC Class A in the U.S.A., EN 55022 in Europe). This constraint becomes most significant with Category 3 UTP.
- **Cable cross talk.** 25-pair bundles are often used for connections from a wiring closet to multiple wall outlets. As a result, several end nodes may be connected to a hub through a single 25-pair bundle. Simultaneous transmissions between the hub and more than one such end node may then cause cross talk within the cable (see "Cross Talk in Unshielded Twisted-Pair Cables" on page 19). The PHY must not be detrimentally affected by this cross talk.
- **Transformer Coupling.** Connections to twisted-pair (UTP and STP) are made through transformers so that dc currents cannot flow between devices with nonequipotential grounds. However, the transformers cause distortion of data signals with dc content, so the PHY must process the data to reduce the dc content of the signal. This often requires that some form of block coding be performed on the data before transmission.<sup>1</sup>

(continued on page 20)



**Fig. 2.** CISPR and FCC radiated emissions regulations. Radiated emissions must fall below the lines shown for each regulatory body.

## Cross Talk in Unshielded Twisted-Pair Cables

Cross talk in UTP cables is caused by capacitive coupling between pairs. Signals on pair A cause noise signals on pair B, and often the cross talk noise proves to be the limiting factor in the link performance. Cross talk occurs in two ways. Near-end cross talk (NEXT) happens when a signal from a transmitter at one end of a cable interferes with a receiver at the same end of the cable. Far-end cross talk (FEXT) occurs when a signal interferes with a receiver at the opposite end of the cable from the transmitter.

### Near-End Cross Talk (NEXT)

Near-end cross talk loss is defined as:

$$\text{NEXT} = -20 \log(V_n/V_i),$$

where  $V_n$  and  $V_i$  are shown in Fig. 1a. The minimum NEXT loss between pairs in a cable tends to follow a smooth curve, as shown in Fig. 1b, decreasing at a rate of 15 dB per decade. However, the actual NEXT between two particular pairs deviates significantly from this curve because of resonances in the twisted-pair. Typical measurements of the NEXT loss between some pairs in a 25-pair cable are also shown in Fig. 1b.

### Far-End Cross Talk (FEXT)

Far-end cross talk loss is defined as:

$$\text{FEXT} = -20 \log(V_f/V_i),$$

where  $V_f$  and  $V_i$  are shown in Fig. 2a. The minimum FEXT loss also decreases with frequency following a smooth curve, but at a rate of 20 dB per decade. As with

NEXT loss, the actual FEXT loss between two particular pairs deviates from this curve. Typical measurements of the FEXT loss between some pairs in a 25-pair cable are shown in Fig. 2b.

### Cross Talk Measurements

Our analysis of cross talk required a database of accurate and detailed measurements of cross talk between pairs in 25-pair cables. A measurement system was constructed to measure NEXT and FEXT losses of all pair combinations in 25-pair cables (see Fig. 3, next page).

Individual pairs were routed to the stimulus and response ports of a network analyzer via a computer-controlled switch. This allowed the automatic selection of 300 different pair combinations for NEXT measurements and 600 pair combinations for FEXT measurements. Any pair not being measured was terminated in 100 ohms via a balun and a 50-ohm termination internal to the switch. The network analyzer measured the cross talk loss (phase and magnitude) to 40 MHz, and this was downloaded to a computer database. Using this system, the NEXT and FEXT losses were measured for many thousands of pair combinations in a selection of 25-pair cables of varying manufacturer and age. The database was used to input NEXT and FEXT loss characteristics to the computation of cross talk noise described in "Cross Talk Analysis" on page 22.

(continued on next page)

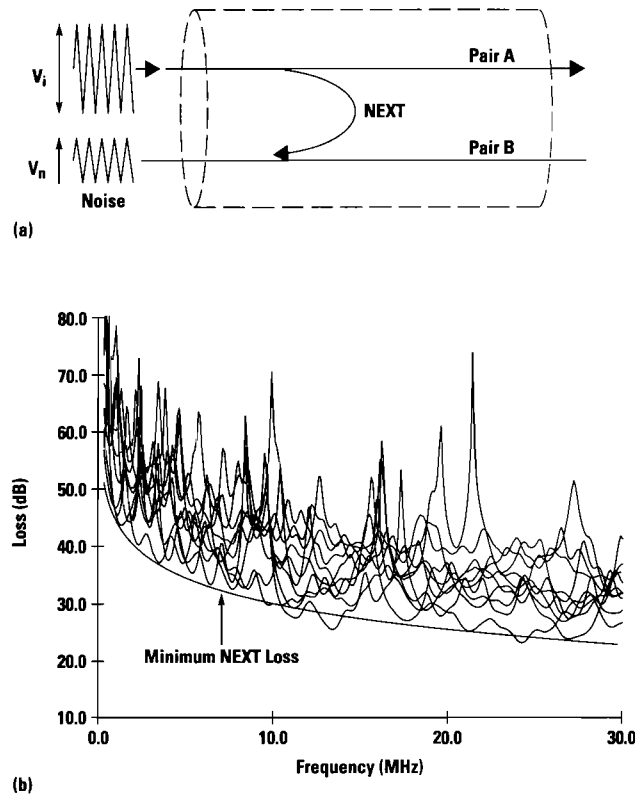


Fig. 1. (a) Near-end cross talk (NEXT). (b) Minimum theoretical NEXT loss and actual measurements.

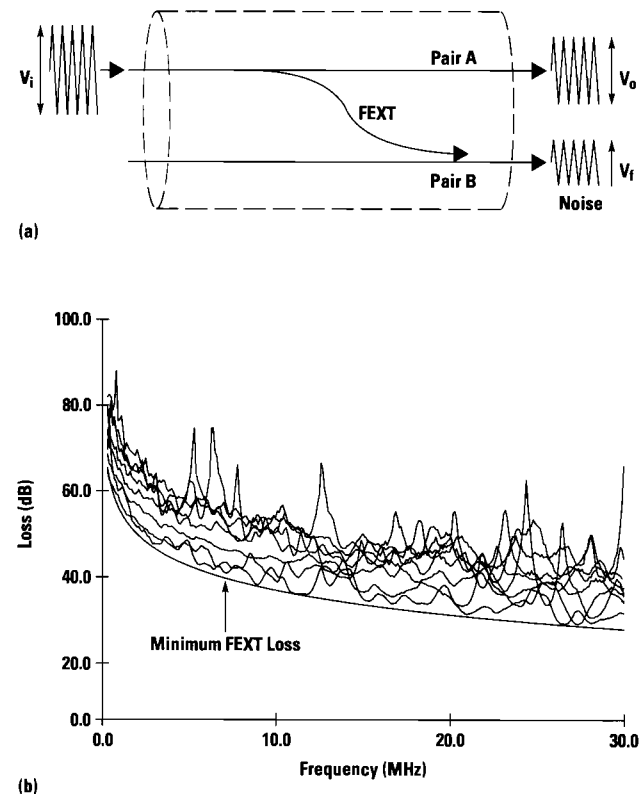


Fig. 2. (a) Far-end cross talk (FEXT). (b) Minimum theoretical FEXT loss and actual measurements.



(continued from page 19)

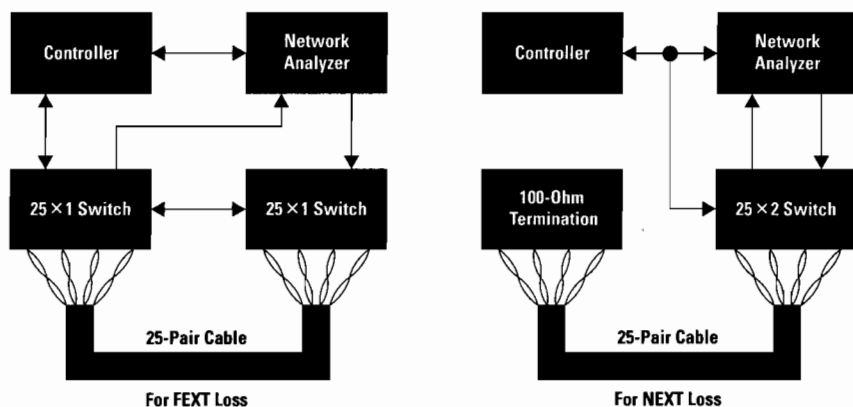


Fig. 3. Measurement system for NEXT and FEXT loss.

(continued from page 18)

Given these constraints, the design goals for the PHY were met by developing a physical medium dependent (PMD) sublayer for each media type (UTP, STP, and multimode optical fiber) and a physical medium independent (PMI) sublayer that contains all the functions common to all media types. These two sublayers together form the demand priority PHY.

In the remainder of this article we will discuss the design choices made for each of the three demand priority PMDs.

#### The Four-Pair UTP PMD

The first PMD to be developed was to support UTP cabling, since this addresses the large 10Base-T upgrade market.

10Base-T uses full-duplex signaling at 10 Mbit/s on UTP cabling.<sup>2</sup> One twisted pair is used to transmit and one to receive data. The dc content of the signal is minimized by Manchester coding the data before transmission on the twisted-pair channel. Manchester coding is a very simple form of block coding:

Bit	Code
0	10
1	01

This is a 1B/2B block code, and results in a 100% bandwidth expansion. Manchester coding is spectrally inefficient relative to other block codes, but does provide a guaranteed transition for every two symbols, has very low dc content, and is very simple to implement.

The transmission rate in 10Base-T is 20 megabaud, giving a data rate of 10 Mbits/s (1 baud is one symbol per second). This means that the transmitted signal can be low-pass filtered with a cutoff frequency somewhat less than 20 MHz. This helps minimize radiated emissions above 30 MHz, the lower bound of stringent EMC regulations.

#### From 10Base-T to 100VG-AnyLAN

The progression from 10Base-T to 100VG-AnyLAN was made in three simple steps.

First, it was recognized that since full-duplex transmission was not absolutely necessary in a hub-based network, both twisted pairs used for 10Base-T could be used simultaneously

for transmission in one direction or reception in the other. This immediately doubles the bit rate achieved on existing 10Base-T networks.

Second, the spectrally inefficient Manchester code was replaced with a more efficient 5B/6B block code (5 bits of data are coded to 6 transmitted binary symbols). This reduced the bandwidth overhead from 100% with Manchester coding to 20%. As a result, the data rate on each pair could be increased to 25 Mbits/s (that is, a symbol rate of 30 megabaud after 5B/6B coding) while the main lobe of the data spectrum remained below 30 MHz, the lower limit of EMC regulations. As with 10Base-T, low-pass filtering with a cutoff below 30 MHz can then be applied to minimize the risk of excessive radiated emissions. The 5B/6B code chosen also has very low dc content, which avoids distortion from the coupling transformers. See the article on page 27 for more details.

Third, the UTP PHY takes advantage of the two unused pairs available in every four-pair cable. Surveys of customer cable plants revealed that a large proportion of these customers adhered to structured cabling recommendations<sup>3</sup> when installing cable, and connected four-pair cable to each wall outlet. Two of these four pairs currently lie unused. By transmitting 5B/6B coded data at 30 megabaud on all four pairs, it is possible to provide a total signaling rate of 120 megabaud (100 Mbits/s) over UTP cable.

#### Quartet Signaling

The four-pair transmission scheme, called quartet signaling, uses binary transmission, that is, only two voltage levels are used as symbols. Other approaches to the UTP PMD were examined. One was multilevel (m-ary) signaling (see "Multilevel Signaling" on page 21). In a multilevel scheme, n data bits are mapped to one of  $m = 2^n$  symbols, and each symbol is a unique voltage level. For example, in a quaternary scheme, two data bits may be mapped to one of four voltage levels. In this way the number of symbols transmitted, and hence the transmission rate required, is reduced by a factor of n. However, for a fixed power supply voltage, the voltage separation between symbols is reduced by a factor of  $1/(m - 1)$  from the binary case. The binary quartet signaling scheme maximizes the voltage separation between symbols, which provides greater immunity to noise at the receiver.

## Multilevel Signaling

Multilevel signaling is often used as a means of compressing the bandwidth required to transmit data at a given bit rate. In a simple binary scheme, two single symbols, usually two voltage levels, are used to represent a 1 and a 0. The symbol rate is therefore equal to the bit rate. The principle of multilevel signaling is to use a larger alphabet of  $m$  symbols to represent data, so that each symbol can represent more than one bit of data. As a result, the number of symbols that needs to be transmitted is less than the number of bits (that is, the symbol rate is less than the bit rate), and hence the bandwidth is compressed. The alphabet of symbols may be constructed from a number of different voltage levels. Fig. 1 shows an example for a four-level scheme.

In the four-level scheme, groups of two data bits are mapped to one of four symbols. Only one symbol need be transmitted for each pair of data bits, so the symbol rate is half the bit rate. The drawback of the multilevel scheme is that symbols are separated by a smaller voltage than in the binary scheme. This means that when noise is added to the data signal (cross talk or impulse), the probability of the noise changing one symbol to another is increased. The symbol separation could be increased to that of the binary scheme by increasing the peak-to-peak transmitted voltage by a factor of  $(m - 1)$  for an  $m$ -level scheme, but this is generally not possible given fixed power supply voltages, and in any case it increases the power required for a transmitter.

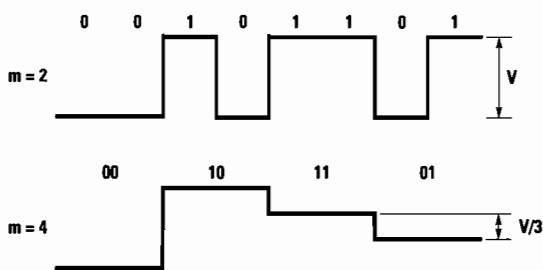


Fig. 1. Two-level and four-level signaling.

The susceptibility of a scheme to errors caused by noise is measured by the ratio of signal separation to noise. Fig. 2 shows the signal-to-NEXT-noise ratio plotted against the transmission bandwidth for several multilevel schemes and for multipair schemes for a bit rate of 100 Mbits/s. A 16-level scheme reduces the bandwidth to 25% of the bit rate, but the S/NEXT ratio is 13 dB (a factor of 4.5) worse than for a four-pair scheme with 25 Mbits/s per pair, which is the scheme used in the 100VG-AnyLAN standard.

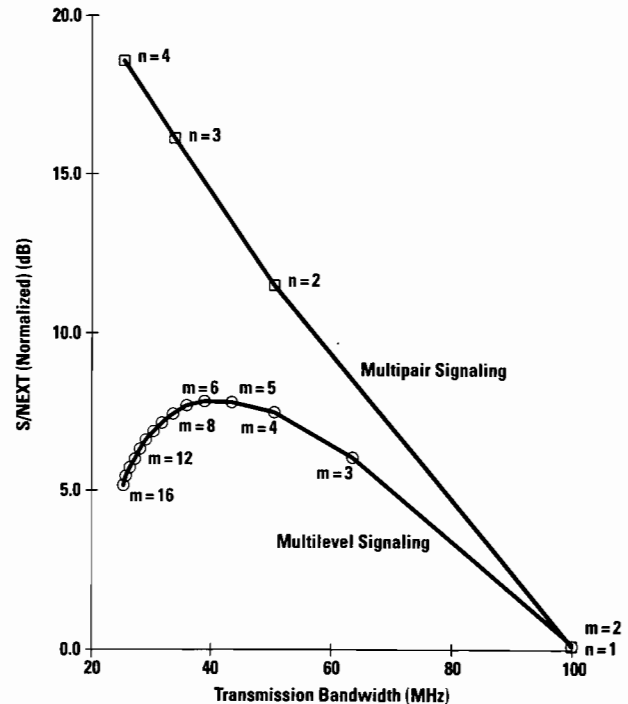


Fig. 2. Susceptibility to errors from noise.

Two noise sources are significant in UTP cabling: external noise and cross talk noise. External noise may be caused by electromagnetic radiation from radio stations (often referred to as continuous wave or CW noise) or impulses from switching equipment (impulse noise). Cross talk noise arises from capacitive coupling between twisted pairs within a cable. Cross talk affects links most significantly when the links use 25-pair bundles.

### Impulse Noise

Little data is available to describe the characteristics of impulse noise occurring on UTP cable plant. It was decided that the safest approach to impulse noise was to maintain the margin of 10Base-T, since the success of 10Base-T proves that its level of robustness is appropriate. The choice of binary signaling meant that the UTP PHY could provide the same immunity to impulse noise as 10Base-T.

### Cross Talk

The 100VG-AnyLAN cross talk environment is very different from that of 10Base-T. Since a four-pair UTP cable only carries a single network link in a network and the traffic on all four pairs is in the same direction, only far-end cross

talk (FEXT, see "Cross Talk in Unshielded Twisted-Pair Cables" on page 19) is a problem. FEXT is less severe than near-end cross talk (NEXT) in UTP cables, and can be disregarded in four-pair links.

However, when 25-pair cables are used to connect end nodes to hubs, up to six network links (each occupying four pairs) can populate one cable. In addition to FEXT, NEXT is also a problem at the hub end of a cable. When a packet is being received on one port, retransmission of that packet at other ports will result in NEXT at the receiving point. It is essential that the level of this NEXT be minimized to prevent errors at the receiving port. The protocol minimizes this NEXT in the following way.

When a packet is received that has a single destination address, the hub forwards the packet to that destination immediately. If the source and destination are attached to the hub by the same 25-pair cable, NEXT will occur from the retransmitted packet to the received packet. An extensive analysis of the cross talk noise generated in such a scenario verified that the received signal is robust to this level of cross talk (see "Cross Talk Analysis" on page 22).

(continued on page 23)

## Cross Talk Analysis

Many previous analyses of twisted-pair transmission systems have assumed that the distribution of cross talk noise is Gaussian. These have yielded reasonably accurate predictions of system bit error rates. However, in applying the error rate analysis, there is an implicit assumption that the cross talk noise is independent of the data on the disturbed system. This is often the case in telecommunication systems, but is not always the case in LANs, where the disturbing links are those on which the disturbed data is being retransmitted. For example, the NEXT interfering with data received at one port of a hub is a result of the retransmission of earlier bits of the same data on other ports. If the cross talk noise is of sufficient amplitude to cause an error in the received data, this error is extremely likely to be repeated every time the same data is transmitted. If a packet is errored by cross talk, that particular packet is likely always to be errored.

To guarantee the error-free transmission of any packet, the worst-case peak cross talk voltage must be found. The peak cross talk noise from multiple disturbers can be calculated directly from a knowledge of the cross talk channels and the disturbing data source. The NEXT and FEXT cross talk channel frequency responses can be calculated for any pair combination using measurements of the pair-to-pair NEXT or FEXT loss. Once the cross talk channel frequency response is known, it is possible to find the impulse response of this channel by inverse Fourier transform. We define the impulse response of the NEXT channel as:

$$g_n(t) = F^{-1}(H_n(f))$$

and the impulse response of the FEXT channel as:

$$g_f(t) = F^{-1}(H_f(f))$$

To find the cross talk noise voltage at the receiver decision point,  $n(t)$ , caused by any data pattern  $f(t)$ , the impulse response is convolved with  $f(t)$ :

$$n(t) = f(t) \star g(t), \quad (1)$$

where  $g(t)$  represents either  $g_f(t)$  or  $g_n(t)$  as appropriate. Our goal is to find the worst-case cross talk for any data pattern, so  $n(t)$  must be calculated for all values of  $f(t)$ . It is therefore useful to apply some limit to the duration of  $f(t)$  to shorten the computation time, and this can be done by taking into consideration the finite duration of the cross talk channel impulse response. A typical impulse response of a NEXT channel is shown in Fig. 1. The duration of the cross talk impulse response is typically less than 1400 ns, which is equivalent to 42 symbol periods for the 30-megabaud transmission rate used in quartet signaling. Therefore, the cross talk waveform during the last six symbols of a pattern  $f(t)$  can be predicted accurately if the duration of  $f(t)$  is restricted to 1600 ns.

The cross talk for any  $f(t)$  is calculated as follows. The disturbing data source is assumed to be the output of a 5B/6B block coding function and consists of eight sequential six-bit codewords, each chosen from an alphabet of 32 codewords. The total number of permutations that  $f(t)$  can take is therefore  $32^8$ . For each permutation,  $n(t)$  is computed according to equation 1. The peak cross talk noise voltage generated by any value of  $f(t)$  can then be found by searching each resultant  $n(t)$ .

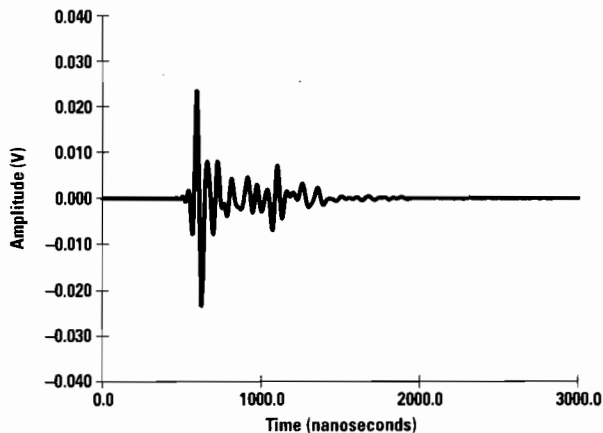


Fig. 1. Typical impulse response of a NEXT channel.

The peak cross talk noise voltage, which represents the maximum noise generated by a worst-case data pattern, can be calculated in this way for each pair combination in a 25-pair cable by repeating the search described above using the NEXT or FEXT loss particular to that pair combination. For each pair combination, the maximum cross talk noise voltage is recorded. The distribution for a typical cable is shown in Fig. 2 for NEXT and FEXT. (The cross talk noise voltage is normalized to the signal amplitude at the receiver decision point.) We denote the maximum NEXT noise voltage for pair  $i$  disturbing pair  $k$  as  $v_{pk,NEXT}^{i,k}$ , and the maximum FEXT noise voltage due to pair  $i$  disturbing pair  $k$  as  $v_{pk,FEXT}^{i,k}$ .

### Multiple-Disturber Cross Talk

The worst-case cross talk environment for UTP PMDs operating over 25-pair bundles consists of three far-end disturbers and four near-end disturbers. Modeling this environment is relatively straightforward given the distributions of  $v_{pk,NEXT}^{i,k}$  and  $v_{pk,FEXT}^{i,k}$ . For each choice of disturbed pair ( $k$ ), four NEXT disturbers ( $a,b,c,d$ ) and three FEXT disturbers ( $p,q,r$ ), the total noise voltage for multiple disturbers is:

$$v_{pk,total} = \sum_{i=a,b,c,d} v_{pk,NEXT}^{i,k} + \sum_{j=p,q,r} v_{pk,FEXT}^{j,k}$$

By calculating the multiple disturber noise voltage in this way we assume, pessimistically, that the maximum noise voltages for the worst-case disturbing patterns from each disturbing source occur at the same time and with the same polarity on the disturbed pair  $k$ . This is obviously a worst-case scenario.

A Monte-Carlo approach has been used to choose combinations ( $k,a,b,c,d,p,q,r$ ) randomly from the 25 pairs of a cable. For each choice,  $v_{pk,total}$  was calculated. The resulting distribution of  $v_{pk,total}$  is shown in Fig. 2. The maximum multiple-disturber cross talk noise expected on any pair of the cable for any choice of disturbing pairs can be estimated from the higher extreme of this distribution (such as the first percentile). This number represents the noise voltage for the worst-case choice of disturbing and disturbed pairs, with the maximum noise contributions from all disturbing pairs occurring simultaneously on the disturbed pair. For the example shown, the first percentile of the total peak noise distribution is 47% of the signal. This allows a substantial margin for error-free signal detection.

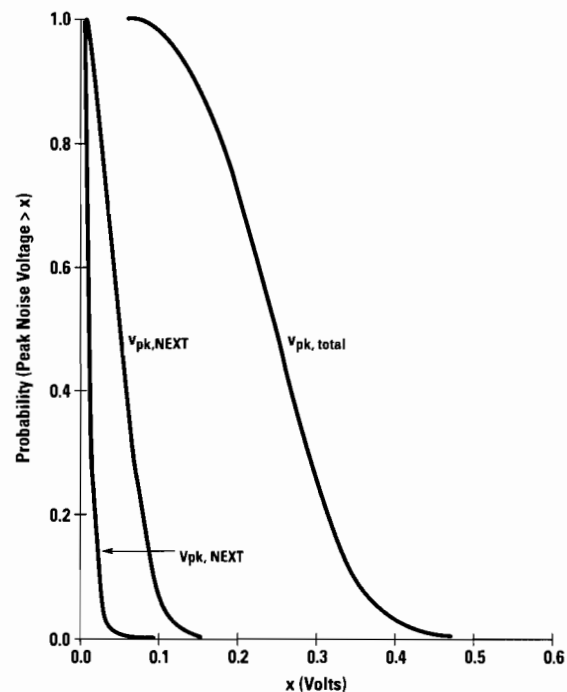


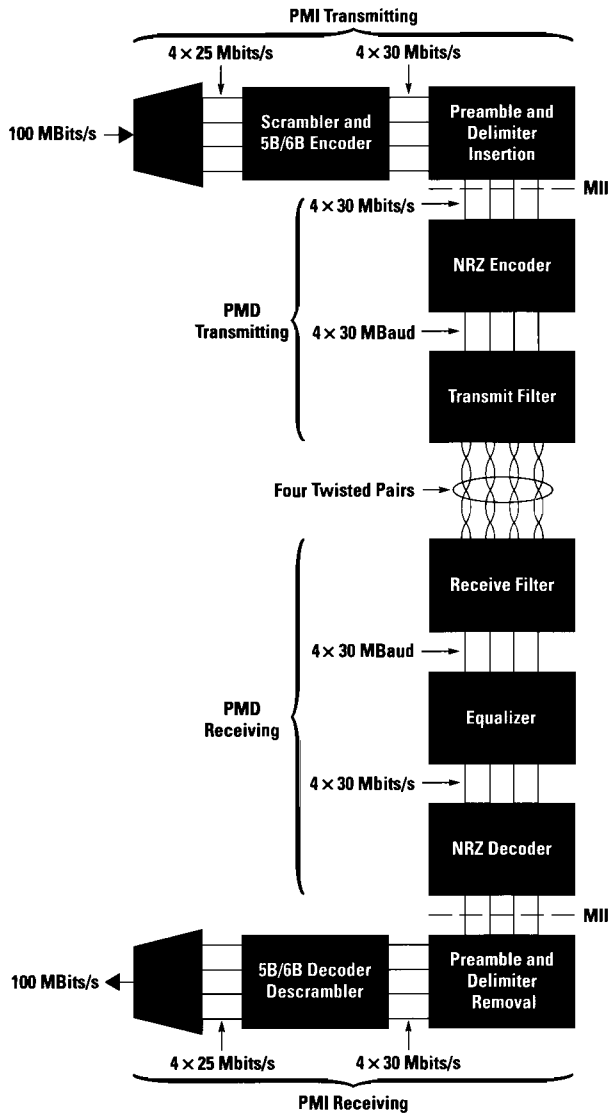
Fig. 2. Distribution of  $v_{pk,NEXT}^{i,k}$ ,  $v_{pk,FEXT}^{i,k}$ , and  $v_{pk,total}$  for a typical 25-pair cable.

When a broadcast or multicast packet is received, the hub does not immediately forward this packet. The NEXT caused by retransmitting the packet to several destinations sharing the same 25-pair cable as the source would result in erroneous reception. Rather, the packet is stored until reception is complete and is then forwarded to all destinations simultaneously. In this way, no NEXT occurs during reception. This store-and-forward technique is only implemented when 25-pair cables are attached to a hub. If only four-pair cables are used, all packets (single and multiple addresses) are forwarded immediately, since there is no NEXT between the individual four-pair cables.

### Implementation of Quartet Signaling

Fig. 3 shows a block diagram of the 100VG-AnyLAN implementation of quartet signaling. The PMI and PMD sublayer functions are considered separately when transmitting or receiving data.

**PMI Transmitting.** The PMI splits the data into four streams, each of which is scrambled. This removes patterns that



**Fig. 3.** Implementation block diagram. The MII is the medium independent interface.

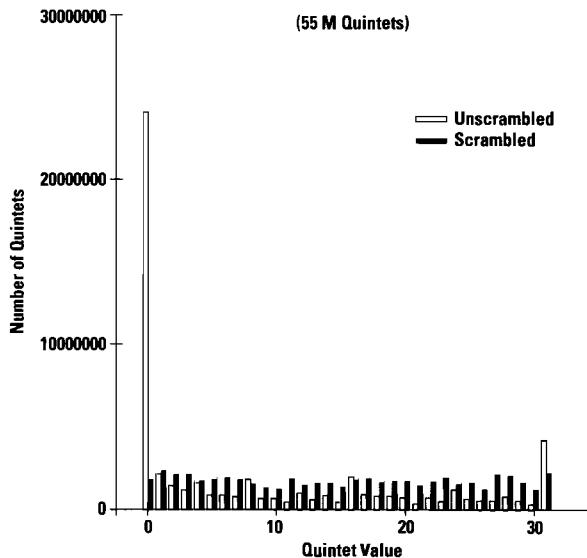
would result in a repetition of the same codeword in the output of the 5B/6B coder. This helps avoid spectral peaks that might violate EMC regulations when the coded data is transmitted.

Typically, LAN traffic contains data patterns that are simply repetitive 1s or 0s. If left unscrambled, when split into quintets (five bits) at the PMI, the distribution of quintets is heavily biased towards all 0s (quintet value of 0) or all 1s (quintet value of 31). This is confirmed by Fig. 4 which shows the distribution of quintets obtained from real LAN traffic. Scrambling removes this bias, providing a more random distribution of quintets at the input to the 5B/6B coder.

After scrambling, the PMI performs the 5B/6B coding. It then adds start and end delimiters to the four streams, and a preamble sequence (a 0101... pattern) to the start of each stream. The four parallel streams of coded data (30 Mbits/s per stream) are then passed to the PMD.

**PMD Transmitting.** The PMD converts the four parallel data streams to the binary signaling levels ( $\pm 2.5V$ ) on each of the four twisted pairs. The data is nonreturn-to-zero (NRZ) coded and low-pass filtered. The low-pass filter has a cutoff at 20 MHz, and is used to attenuate spectral components that would cause undesirable emissions above 30 MHz. A typical eye diagram at the output of the PMD is shown in Fig. 5.

**PMD Receiving.** When receiving data, the low-pass filter in the PMD rejects out-of-band noise on the twisted pair. The signals on the four channels are then equalized. This compensates for the attenuation of the cable and minimizes the intersymbol interference at the sampling point of the data. To perform this function for any cable length between 0 and 100 m the equalizer must be adaptive. The protocol provides a training sequence, during which the PMD equalizer trains its response to compensate for the length of cable present in the link. A typical eye diagram at the output of the receiver



**Fig. 4.** Distribution of quintets obtained from real LAN traffic.

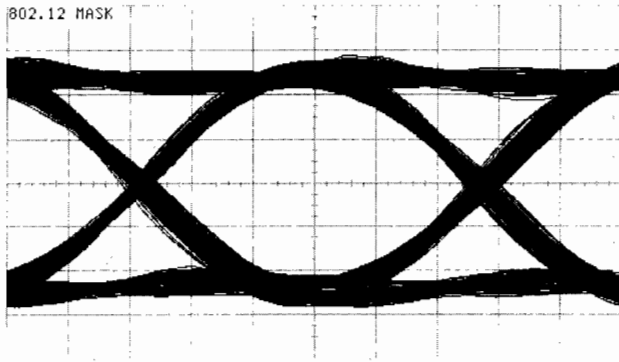


Fig. 5. Transmission eye diagram.

equalizer is shown in Fig. 6. In this case the cable length was 121 m, which approximates a worst-case 100-m cable.

After equalization the received signals are sampled. The PMD recovers a clock from the received signals by using the preamble pattern generated by the transmitting PMI. This clock is used to sample the received data. The four channels are often misaligned in time by up to two bit periods with respect to each other after traveling over a length of UTP cable. This is because the twist rate, and hence the propagation delay, varies from one twisted pair to another. The PMD realigns the received signals to a common clocking point before passing the four parallel streams of data to the PMI.

**PMI Receiving.** The PMI further realigns the received data to remove any skew between the start delimiters on each stream. The data on each stream is then decoded, unscrambled, and reformatted into a single data stream, which is passed to the receiving MAC. The PMI also performs a number of error-checking functions. These provide extra protection against errored packets being accepted as valid by the MAC, above the protection offered by the frame check sequence. The PMI first checks that the start delimiters on each stream are all valid and occur with the correct time relationship to each other. Then, while decoding, it checks that only valid 6-bit codewords are received. It signals to the MAC if an error is detected in the received data.

Quartet signaling meets all the design goals identified for the PHY apart from control signaling (described next). By retaining a binary signaling scheme the simplicity and robustness of

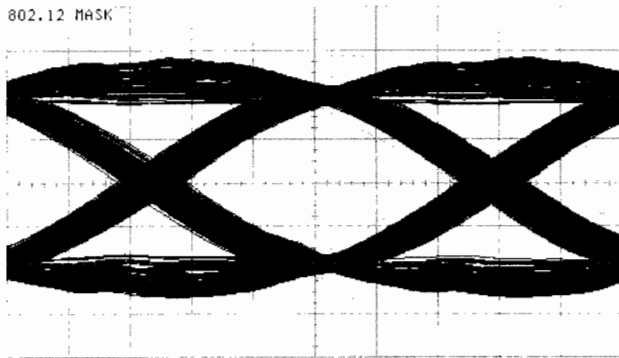


Fig. 6. Reception eye diagram.

10Base-T was maintained. Keeping the baud rate low makes it easy to implement in a standard CMOS process, thereby meeting the low-cost requirement. The increased data rate is mainly attributable to more efficient block coding and better use of the available twisted-pair channels.

### Control Signaling

The protocol uses control signals to transfer requests, acknowledgments, and training signals between the hub and the end nodes. These control signals must be continuously available on all network links whenever data is not being transferred. They must be full-duplex and distinct from any data pattern.

Since the control signals are required to operate continuously, signaling at the data rate of 30 megabaud was rejected because of the large levels of cross talk that might be generated when 25-pair cables were used. Instead, a low-frequency signaling scheme has been developed that allows all five control states to be transmitted using only signals with fundamental frequencies less than 2 MHz. No valid data pattern contains components with such low frequencies, providing distinct identity for the control signals. Full-duplex operation is achieved by using two pairs in each direction to carry control signals.

Two control signals are defined: CS1 is a 0.9375-MHz square wave formed by repeating a pattern of sixteen 0s followed by sixteen 1s, and CS2 is a 1.875-MHz square wave formed by repeating eight 0s followed by eight 1s. Four of the five control states are represented by combinations of CS1 and CS2 on two pairs as shown in Table I.†

The fifth control state is represented by silence (no energy) on the two pairs. Silence is used by a hub to indicate that a request to transmit data has been granted. On receiving silence, an end node can cease control signaling and begin transmission immediately, since the cable is already silent. This allows rapid turnaround of the half-duplex link.

The two square waves have been chosen to have a large separation in frequency so that the receiver is able to distinguish them without having a clock that is precisely phase-locked. The control signals are generated and recovered within the UTP PMD.

### The STP and Optical-Fiber PMDs

The STP PMD was developed to support existing token-ring network cabling. These consist of cables up to 100 m in length with two shielded twisted pairs. The optical-fiber PMD was developed to provide a means of connecting hubs and end nodes over longer distances than the 100 m provided by the STP and UTP PMDs. This extra distance is particularly important when cascaded networks are built and hubs are distributed over a campus area. The optical-fiber PMD allows LANs to have up to 4-km diameter.

† The control signals shown in Table I are for the 100VG-AnyLAN implementation of the IEEE 802.12 standard. The control signal definitions in the standard have different names.

**Table I**  
**Control Signaling in 100VG-AnyLAN**

Tone Wires		Transmitted from			Received by		
1	2	End Node	Root Hub	Another Hub	End Node	Root Hub	Another Hub
CS1	CS1	IDLE	IDLE	IDLE	IDLE	IDLE	IDLE
CS1	CS2	REQ_N	INCOMING	REQ_N	INCOMING	REQ_N	INCOMING
CS2	CS1	REQ_H	ENABLE_HIGH_ONLY	REQ_H	Reserved	REQ_H	ENABLE_HIGH_ONLY
CS2	CS2	REQ_T	REQ_T	REQ_T	REQ_T	REQ_T	REQ_T

REQ\_N = Normal-priority request

REQ\_H = High-priority request

REQ\_T = Link training request

INCOMING = A packet is about to be transmitted

ENABLE\_HIGH\_ONLY = Put normal round-robin sequence on hold while a high-priority request is serviced

STP has less attenuation and greater NEXT loss than even Category 5, data-grade UTP, and is already used for data transmission at 100 Mbits/s. For example, the SDDI specification<sup>4</sup> provides for transmission of FDDI traffic at 100Mbits/s over STP.

The STP physical layer uses the same modulation and coding techniques as SDDI, that is, binary signaling at a rate of 100 Mbits/s (before coding). The low attenuation of the cable makes it possible to transmit 100Mbits/s on a single pair, so one pair is dedicated to transmitting and one pair to receiving at each end of a link. The cable shield reduces radiated emissions to satisfactory levels.

Multimode optical-fiber links have also been in use for many years at data rates of over 100 Mbits/s (such as FDDI).<sup>5</sup> The optical-fiber PMD uses components that have recently been developed for low-cost FDDI implementations. Examples are the HFBR 5106 and HFBR 5107 developed by HP's Optical Communications Division. These use LED transmitters with wavelengths of either 850 nm (for 500-m links) or 1300 nm (for 2-km links). The transmitter forms part of a small module that also includes an optical receiver. This module forms a simple interface between the optical fiber and a transceiver chip, which can be identical to the STP transceiver (see "Optical-Fiber Links for 100VG-AnyLAN" on page 26).

For the demand priority STP and optical-fiber PHYs, the block code and NRZI coding specified for SDDI have been replaced with an alternative scheme based on the same 5B/6B block code that is used for the UTP PHY. This approach allows the 5B/6B block coder to be placed in the PMI sublayer. The amount of logic that is common to all physical layers is thereby increased, resulting in a lower-cost PHY device.

The PMI provides four 30-Mbit/s channels of scrambled and 5B/6B block coded data at the MII. These four channels are multiplexed, codeword by codeword, by the STP and fiber PMDs (see article, page 27). Even after codeword multiplexing, the 5B/6B code retains its advantageous properties. The serialized data stream is NRZ coded and transmitted at 120 megabaud on one pair using symbol levels of  $\pm 0.25V$  for STP media or passed to the optical module mentioned above for optical-fiber cables.

### Control Signaling on STP and Optical Fiber

The two-pair, two-tone control signaling developed for the UTP PMD is not suitable for the STP and optical-fiber PMD because only one pair per direction is available for control signaling. Because cross talk is not an issue with STP or optical-fiber links, we were able to explore the use of higher-frequency signals. Square waves were again attractive because they can easily be chosen to be distinct from valid data patterns.

One concern was that the control signals not produce harmonics that, when transmitted on STP, might cause radiated emissions that violate regulations. It was decided that the control signal spectra should always fall below the random data spectrum, since it was known that the data transmission met EMC regulations. Calculations showed that square waves with frequencies less than 4 MHz met this requirement. A lower bound on the control signal frequency was set by the need to avoid distortion of the square wave resulting from transmission through the transformers used to couple to twisted-pair.

The final choice of control signals is five square waves with frequencies between 1.875 and 3 MHz. Again, the control signals are separated in frequency sufficiently to allow detection without a phase-locked clock at the receiver.

### Summary

The physical layer developed for 100VG-AnyLAN local area networks accommodates different cable types by means of three different physical medium dependent (PMD) sublayers.

The major goal was to provide 100-Mbit/s transmission on the existing cable plant. The three PMDs allow LANs to operate across the vast majority of LAN media installed today: UTP (categories 3, 4, and 5), STP, and multimode fiber. As a result, customer investment in structured cabling is protected while at the same time an upgrade path to high-speed LANs is created. All three PMDs are designed to be simple and cost-effective. Customers can now benefit from a factor of up to eight improvement in the cost/performance ratio of their LANs without the significant cost penalty of replacing their cable plant.



---

## Optical-Fiber Links for 100VG-AnyLAN

As data rates increase, low-cost optical-fiber links play an increasingly significant role in LANs for extending the length of links beyond what can be achieved with copper media, while meeting the full range of electromagnetic emission and susceptibility requirements for networks.

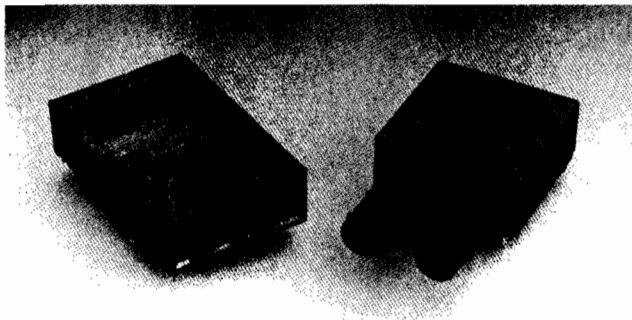


Fig. 1. HP optical transceiver.

The 100VG-AnyLAN standard defines a serialized interface with a 120-megabaud signaling rate for STP and multimode optical-fiber links. The standard defines two optical-fiber link length specifications, which allow the use of low-cost 850-nm technology for 500-m building backbones (this is the same technology used in the existing IEEE 802 standard CSMA/CD 10Base-F and 802.5J token-ring links) and 1300-nm technology for 2-km campus backbone links.

Fig. 1 shows the new Hewlett-Packard low-cost industry-standard optical transceiver package. This small package is 1 inch wide and 1.5 inches long and has a duplex SC optical connector on the front and a 1-by-9 row of electrical pins at the rear. HP transceivers HFBR 5106/5107 meet the two 100VG-AnyLAN link length standards and allow interchanging link technology in the same printed circuit board footprint. The transceivers are also available with AT&T ST optical connectors to address the large installed base having ST connectors for building and campus backbones.

Del Hanson  
Principal Engineer, Fiber-Optic  
Networks and Standards  
Optical Communications Division

---

### Acknowledgments

The authors were fortunate to be part of an enthusiastic and effective team based at HP Laboratories Bristol and the Roseville Networks Division. Major contributions to the physical layer developments were made by Pat Thaler of RND. Rod Chidzey of HP Laboratories was responsible for the construction of the cable cross talk measurement rig and for gathering cable data. Del Hanson of the Optical Communications Division provided valuable assistance in the specification of the fiber-optic PMD.

### References

1. E.A. Lee and D.G. Messerschmitt, *Digital Communication, Second Edition*, Kluwer Academic, 1994.

2. *Twisted-Pair Medium Attachment Unit (MAU) and Baseband Medium, Type 10Base-T*, IEEE Standard 802.3, section 14.4.4.1.
3. *The Commercial Building Telecommunications Wiring Standard*, EIA/TIA Standard 568, July 1991.
4. *SDDI: A Specification for FDDI on Shielded Twisted-Pair*, implementors' agreement published by AMD Inc., Chipcom Corp., IBM Corp., Madge Networks Inc., Motorola Inc., National Semiconductor Corp., Network Peripherals, Inc., Sumitomo Corp., Synoptics Communications Inc., SysKonnnect Inc., and Technitrol Inc.
5. *Fiber Distributed Data Interface (FDDI)—Token Ring Physical Layer Protocol (PHY)*, ANSI Standard X3.148-1988.

# Coding in 100VG-AnyLAN

A 5B/6B coding scheme in which five data bits are encoded into six-bit codewords is used in conjunction with offsetting the data on different channels by three bits in quartet signaling. It provides the level of error detection necessary, produces a signal balanced within narrow limits, and restricts strings of consecutive 0s or 1s to a maximum length of 6. It is also efficient.

by **Simon E. C. Crouch and Jonathan Jedwab**

Coding of data before transmission has three main purposes. First, it can ensure that the transmitted signals are dc balanced, that is, that there are equal numbers of 0s and 1s over an extended period of time. This is important in transformer-coupled systems, which are designed to prevent ground loops. Second, coding can aid clock synchronization by minimizing the length of strings of consecutive 0s or 1s. This ensures that there is a high density of signal transitions. Third, together with cyclic redundancy checks (CRCs), it can enable errors to be detected. There is an IEEE Project 802 functional requirement that any three bits in error within a data frame must be detected. There is also a need to detect burst errors; any coding system must not compromise the detection provided by the CRC.

The 100VG-AnyLAN team chose a 5B/6B coding scheme in which five data bits are encoded into six-bit codewords. This is used in conjunction with offsetting the data on different channels by three bits in quartet signaling. It provides the level of error detection necessary, produces a signal balanced within narrow limits, and restricts strings of consecutive 0s or 1s to a maximum length of 6. At the same time, it is efficient, adding only 20% to the data load, compared with the 100% produced by the 1B/2B Manchester coding used in 10Base-T. The 5B/6B code is also effective when the four data channels are multiplexed for transmission on STP or fiber-optic cables.

This article describes the design of the 5B/6B block code used in 100VG-AnyLAN and more generally in the IEEE 802.12 proposed local area network standard<sup>1</sup> and explains the reasons behind its design.

## Why Code?

The method used to transmit digital data has to take account of the physical constraints that the transmission medium imposes upon the transmission system. The physical media used by 100VG-AnyLAN include unshielded twisted pair (UTP) cable, shielded twisted pair (STP) cable, and optical fiber. The demands imposed by the use of UTP cable mean that to transmit at 100 Mbits/s (even when using quartet signaling), digital data has to be coded before it is transmitted. The design of the code has to take account of various constraints imposed by implementation considerations and by the need to be compatible with other data transmission systems. In particular, the 5B/6B 100VG-AnyLAN code scheme was designed to be compatible with the packet

structure and error detection capabilities defined by the IEEE 802.3 standard used in Ethernet local area networks and by the IEEE 802.5 standard used in token ring local area networks.

## Block Codes and Error Detection

Generally speaking, a code is a mapping from an alphabet of symbols to a set of sequences of symbols from some other alphabet. So, for example, we could define a code from the alphabet {a,b,c,...,z} to the set of sequences of zeros and ones by: a → 010, b → 0110, c → 01110, and so on. For another example, the representation of decimal numbers as their binary equivalents can be regarded as a code. This idea of a code should not be confused with the notion of a cipher, which is a code that is specifically designed to hide the meaning of transmitted data from those not meant to understand it. The codes that we are using are designed to be easy to decode and are capable of detecting errors in transmission. For the mathematical background of coding theory, see reference 2. For an engineering perspective, see reference 3.

A block code is a specific type of code in which every codeword has the same length. In a 5B/6B block code, each element of an alphabet of 32 different symbols (which itself can be represented by a 5-bit number) is encoded as a 6-bit codeword.

The major physical constraint that affects the use of a coding scheme is the need to maintain dc balance, that is, the spectral content of the transmitted signal should have no zero-frequency component. In mathematical terms, this translates into requiring that the difference between the number of ones and the number of zeros transmitted at any time must be kept as close to zero as possible.

Clearly, a simple way to get a balanced code is to represent a data character of 0 by the codeword 01 and a data 1 by the codeword 10. The number of 1s transmitted will then equal the number of 0s transmitted at the end of each codeword. Unfortunately, this is also a very inefficient way of achieving dc balance because it requires the physical transmission of two code bits for every data bit sent. The 100VG-AnyLAN team chose to implement a 5B/6B code because it gives a suitable trade-off between efficiency and cost of implementation.

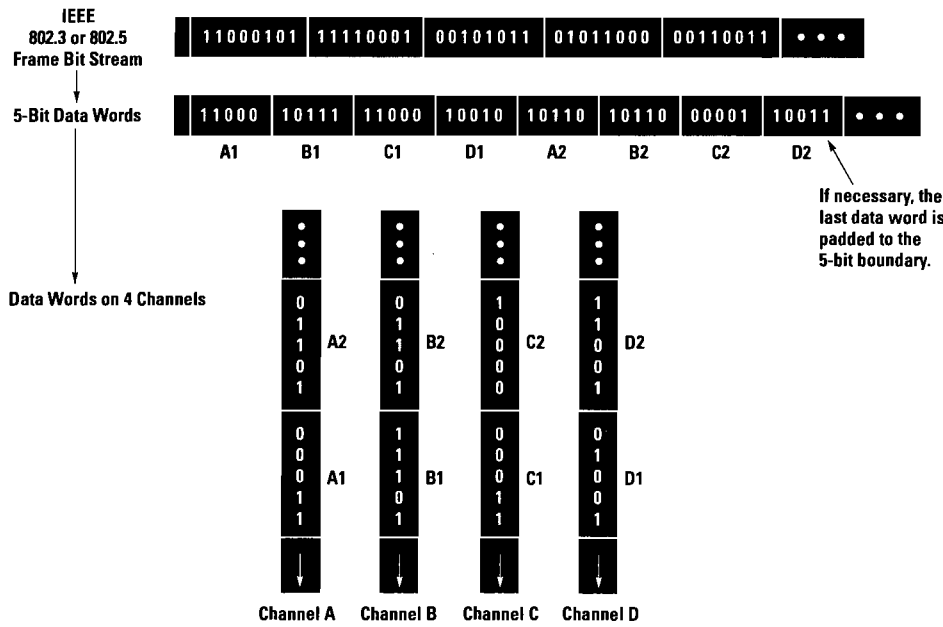


Fig. 1. Distribution of a data stream on four channels for quartet signaling.

The major compatibility constraint imposed upon the design of the 100VG-AnyLAN 5B/6B code was that it had to be compatible with the MAC frame formats defined by the IEEE 802.3 and 802.5 standards and in particular with the error-detection capabilities offered by the cyclic redundancy check (CRC) built into those formats. (See "Polynomial Arithmetic and Cyclic Redundancy Checks" on page 31). CRC-32, the CRC defined by both IEEE 802.3 and 802.5, detects up to three single-bit errors occurring anywhere in a frame, or a single burst error of up to 32 bits in the frame. The protocol will then discard that packet as flawed. (A data packet consists of the MAC frame delimited by a preamble and start and end delimiters.) The 100VG-AnyLAN team had to design the 5B/6B code so that similar error-detection characteristics will be maintained when 802.3 or 802.5

frames are coded and transmitted in parallel using quartet signaling. In particular, the requirement to detect up to three single-bit errors anywhere in a frame is a compulsory requirement of IEEE Project 802,<sup>4</sup> so the new IEEE 802.12 standard had to meet this demand.

In the next two sections we describe the process of coding the data frame and the error detection capabilities of 100VG-AnyLAN.

#### Coding a Data Packet

Data from the IEEE 802.3 or 802.5 frame is divided into 5-bit data blocks (with padding added at the end, if necessary) and is distributed between the four data streams transmitted using the quartet signaling scheme (see Fig. 1).

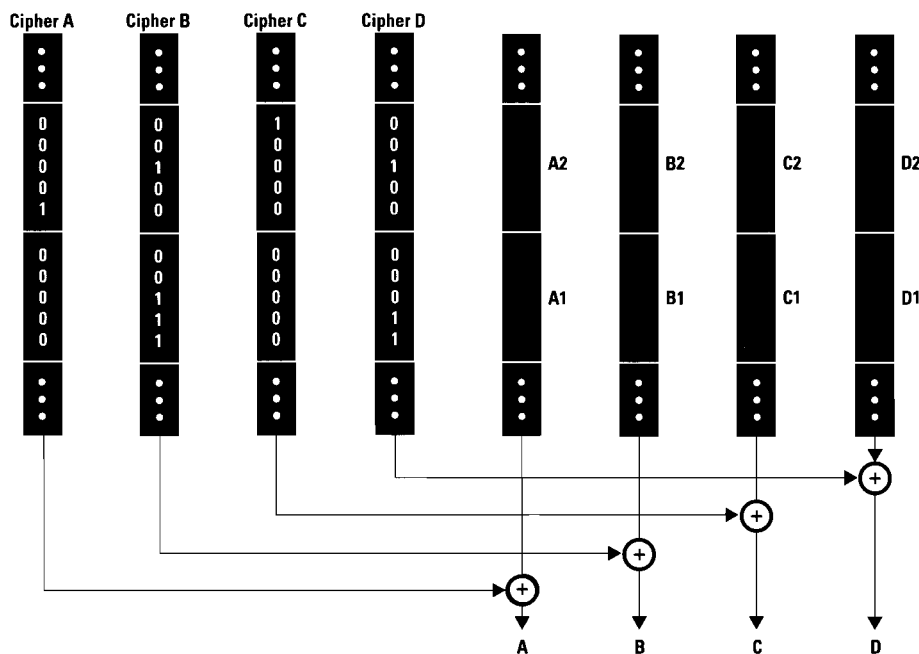


Fig. 2. The data words are XORed with pseudorandom stream ciphers.

Each data stream is exclusive-ORed with a stream of pseudo-random bits produced by stream ciphers (see Fig. 2). The ciphered data blocks on the four streams are then coded according to the following 5B/6B code table, as shown in Fig. 3.

Data	Code	Data	Code Pair
00001	101100	00000	001100 110011
00011	001101	00010	100010 101110
00101	010101	00100	001010 110101
00110	001110	01011	000110 111001
00111	001011	01100	101000 010111
01000	000111	01110	100100 011011
01001	100011	10000	000101 111010
01010	100110	10010	001001 110110
01101	011010	10101	011000 100111
01111	101001	10111	100001 011110
10001	100101	11010	010100 101011
10011	010110	11110	010010 101101
10100	111000		
10110	011001		
11000	110001		
11001	101010		
11011	110100		
11100	011100		
11101	010011		
11111	110010		

In the second column, all the 6-bit codewords are balanced, that is, there are equal numbers of 0s and 1s. There are not enough balanced 6-bit codewords to code all possible 5-bit data symbols, so twelve data symbols are coded by a choice of two different 6-bit codewords, one of weight two (two 1s and four 0s), and one of weight four (four 1s and two 0s). The two codewords are used in the following fashion, independently in each stream:

- For the first data symbol that codes to an unbalanced codeword, the weight-two codeword is chosen.
- When the next unbalanced codeword occurs, the weight-four codeword is used.

- Weight-two and weight-four codewords continue to alternate whenever an unbalanced codeword occurs.

The observant reader will notice that in all but one of the code pairs, the alternative unbalanced codewords are logical negatives of each other. This is not true of the second pair, for reasons having to do with the error properties of the code.

At the end of the stream, one of two end delimiters is used, independently in each stream (see Fig. 4). When the end of the stream is reached, if the next unbalanced codeword is due to have weight 2 (according to the rules above), end delimiter ED2 is used. If it is due to have weight 4, ED4 is used. The invalid packet marker (IPM) is used by repeaters to mark errored packets for disposal or further processing.

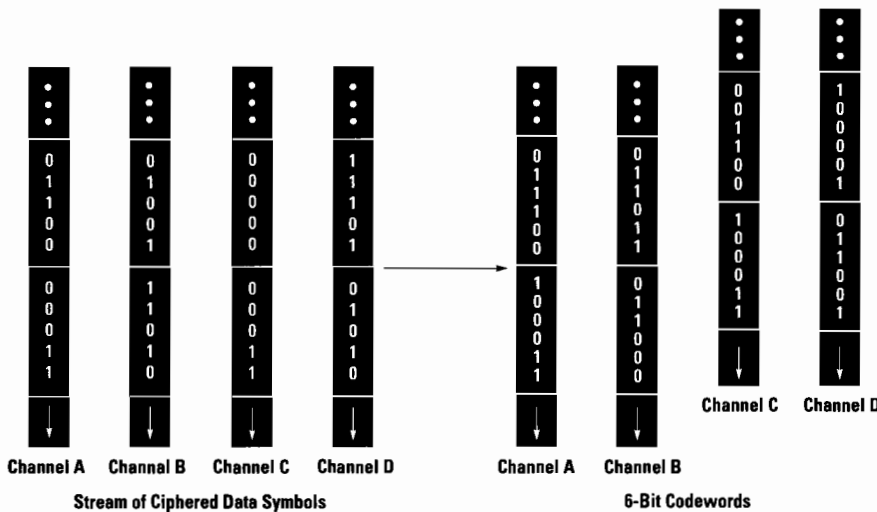
After coding, the data streams are offset by 3 bits with respect to each other before transmission (see Fig. 3). Again, the reason for this has to do with error-detection properties and will be explained below.

The actual transmitted bit sequence down each channel looks as shown in Fig. 4. The start and end of packet markers are designed to maintain error detection capabilities. The use of alternative end delimiters (effectively an extra parity check) is an essential part of the error detection scheme of 100VG-AnyLAN.

### Multiplexing

Another design constraint of the 5B/6B coding scheme was that it should behave well when the four data streams are multiplexed onto fewer channels for networks using STP or fiber-optic cables. As noted in the article on page 18, STP and fiber-optic PMDs pass four parallel streams of data through a multiplexer, which combines the four codeword streams into one stream, codeword by codeword (Fig. 5).

When combined in this way, the physical and error protection capabilities of the 5B/6B code are maintained, as explained in the next section.



**Fig. 3.** Conversion of 5-bit data symbols to 6-bit codewords. The three-bit offset of channels C and D is for error detection.

## IEEE 802.3 and 802.5 Frame Formats

The IEEE 802.3 frame (Fig. 1) consists of a 48-bit destination address, a 48-bit source address, a 16-bit length field, and then a data field ranging between 368 and 12,000 bits in length and consisting of data organized into octets. This is followed by a 32-bit cyclic redundancy check (CRC).

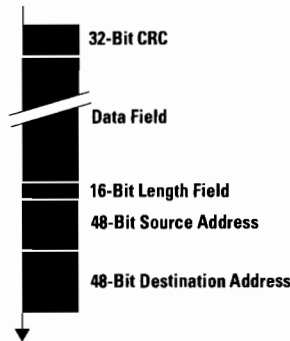


Fig. 1. IEEE 802.3 frame structure.

The IEEE 802.5 token-ring frame (Fig. 2) consists of an 8-bit access control field (not used in the 802.12 standard), an 8-bit frame control field, a 48-bit destination address, a 48-bit source address, between 0 and 240 bits of routing information, and then a data field ranging between 0 and 36,016 bits in length and consisting of data organized into octets. This data field is followed by a 32-bit CRC.

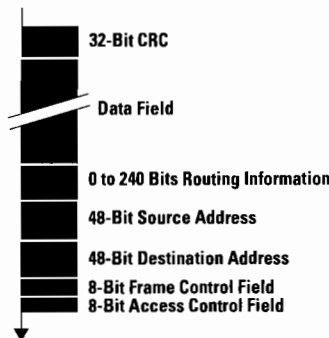


Fig. 2. IEEE 802.5 frame structure.

The CRC provides two types of protection for an IEEE 802.3 or 802.5 frame:

- Any three single-bit errors occurring anywhere in the frame are detected.
- Any burst of errors for which the distance between the first corrupted bit and the last corrupted bit is less than or equal to 32 bits is detected.

IEEE Project 802 requires that any transmission scheme developed under its aegis can perform the first type of error detection, that is, that at least three single-bit errors can always be detected.

### Properties of the Coding Scheme

The choice of 5B/6B code, together with the alternation of unbalanced codewords and the use of the alternative end delimiters, ensures that IEEE 802's requirement is met: up to three single-bit errors occurring anywhere on the four channels within a data frame are detected. This is not a trivial matter to confirm, because the use of a code means that a single-bit error in the transmitted bit stream may cause many more than one single bit to be in error in the data stream after decoding. A substantial portion of the code

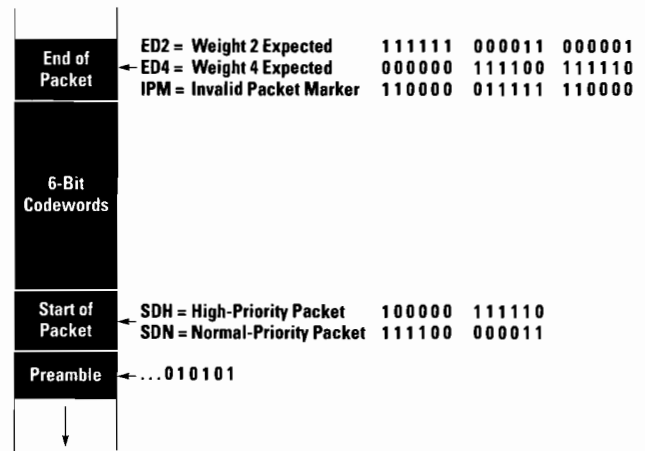


Fig. 4. Transmitted bit sequence on each channel.

design effort was dedicated to limiting the number of data bits affected by a single code bit in error.

It is not difficult to see how the use of the alternation rule for choosing unbalanced codewords together with the use of the alternative end delimiters leads to the detection of all triples of single-bit errors, when those errors occur in distinct codewords on a single channel. Fig. 6 shows several cases in which three single errors are detected by the alternation rules and the use of end delimiters.

A second major design thrust was to make sure that burst error-detection properties inherited from CRC-32 are not compromised in 100VG-AnyLAN. Because we are now talking about a system that can transmit over four parallel streams as well as a single stream, we have to be careful to define what we mean by a burst error. We decided that the clearest case could be made by considering a burst error to be any error caused by arbitrary corruption across all the streams for a certain number of bit periods (see Fig. 7). So, for example, for a burst error of seven code bit periods, there is a block seven code bits long by four code bits wide (corresponding to the four channels), within which any given code bit may or may not be inverted.

The first thing we realized was that if all the codewords were transmitted synchronously on all four channels, a burst error of length 2 (corrupting eight code bits) could cause trouble if it occurred at a codeword boundary (see Fig. 8). The error could corrupt eight code words and thus eight data blocks, corresponding to forty data bits. This is well above the 32 protected by the CRC. It would therefore require a remarkably good coding scheme to protect against even such a short error burst.

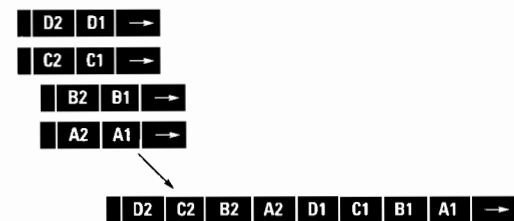


Fig. 5. Multiplexing four streams into one.

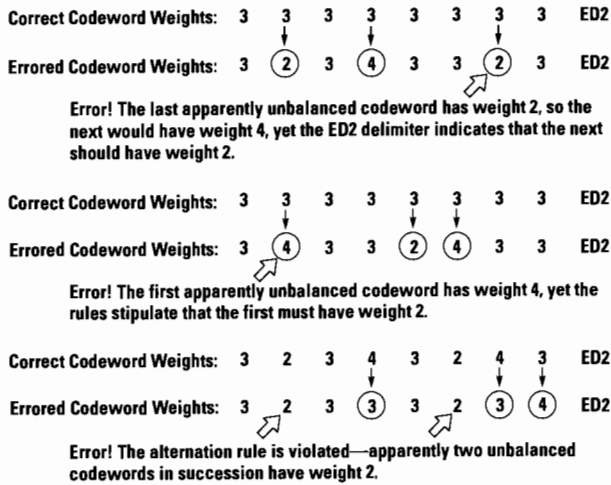


Fig. 6. Detection of errors through the alternation rule and end delimiters.

The solution to this problem is to offset the codewords on two of the channels relative to the other two. Fig. 9 clearly shows that with such an offset, even an error burst of length four (corrupting 16 code bits) can only corrupt the code bits of at most 6 codewords. This corresponds to 6 data blocks (30 data bits) and so will be detected automatically by the burst error properties of the CRC. Thus, even with an arbitrary code, any burst error of length four is detected.

Substantial theoretical work led us to design a code on top of this offset scheme that extends the burst error-detection capability to seven code bit periods (arbitrarily corrupting 28 code bits), while still meeting the requirement to detect three separate single-bit errors (see Fig. 10). The triumph of theory over trial and error should not be underestimated here—there are approximately  $10^{45}$  choices of 5B/6B code!

When multiplexed onto a single channel for STP or fiber-optic solutions, the error detection capabilities of the 5B/6B code are enhanced. The single-bit error protection remains at three while the burst error-detection capabilities are increased to 34 code bits.

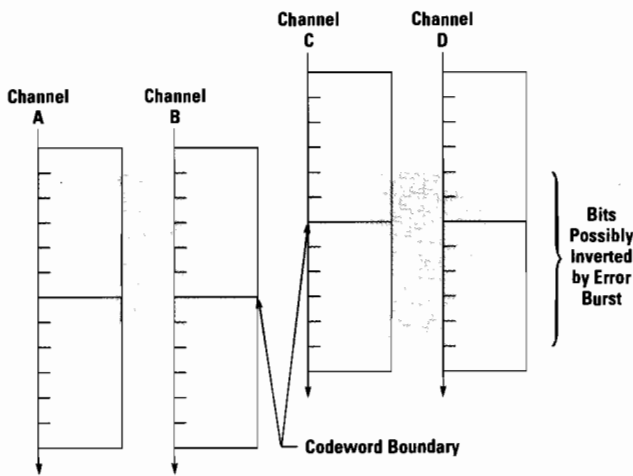


Fig. 7. Illustration of a burst error with a length of seven code bit periods.

## Polynomial Arithmetic and Cyclic Redundancy Checks

The calculation of cyclic redundancy checks (CRCs) depends upon the arithmetic of modulo 2 polynomials. A modulo 2 polynomial is an expression of the form  $a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ , where the coefficients  $a_0, a_1, \dots, a_n$  are integers modulo 2, that is, they can take values of 0 or 1. The coefficients obey regular modulo 2 arithmetic as follows:

Addition			Multiplication		
+	0	1	×	0	1
0	0	1	0	0	0
1	1	0	1	0	1

Electrical engineers will recognize the addition operation as the XOR operation defined on binary states.

The following are examples:

- $(1+x^2+x^4) + (1+x+x^4) = x+x^2$
- $(1+x^2+x^4)(1+x+x^4) = 1+x+x^2+x^3+x^5+x^6+x^8$

Division of modulo 2 polynomials is done in exactly the same way as it is for ordinary polynomials, remembering that the coefficients obey modulo 2 arithmetic as defined above.

A very important operation in the calculation of CRCs is the calculation of the remainder  $R(x)$  when one polynomial  $M(x)$  is divided by another  $G(x)$ .  $R(x)$  is uniquely defined by  $M(x) = G(x)K(x) + R(x)$ , where the degree of  $R(x)$  is less than the degree of  $G(x)$ . The degree of a polynomial is defined as the largest value of  $n$  for which the coefficient  $a_n$  is nonzero, so, for example,  $1+x^2+x^5$  has degree 5.

Although this sort of polynomial division may look formidable, there are very efficient means of calculation based on shift registers.

CRCs are calculated on a stream of data by assuming that the data represents the coefficients of some modulo 2 polynomial. So, given a stream of data  $n$  bits long, the first bit can be considered as the coefficient  $a_{n-1}$  of  $x^{n-1}$ , the second bit as the coefficient  $a_{n-2}$  of  $x^{n-2}$ , the  $(n-1)$ th bit as the coefficient  $a_1$  of  $x$  and the  $n$ th bit as the constant term  $a_0$ .

Roughly speaking, the 32 bits of the CRC are defined to be the polynomial remainder  $R(x)$  when the polynomial defined by the data  $M(x)$  is divided by a standard polynomial:

$$G(x) = 1+x+x^2+x^4+x^5+x^7+x^8+x^{10}+x^{11}+x^{12}+x^{16}+x^{22}+x^{23}+x^{26}+x^{32}$$

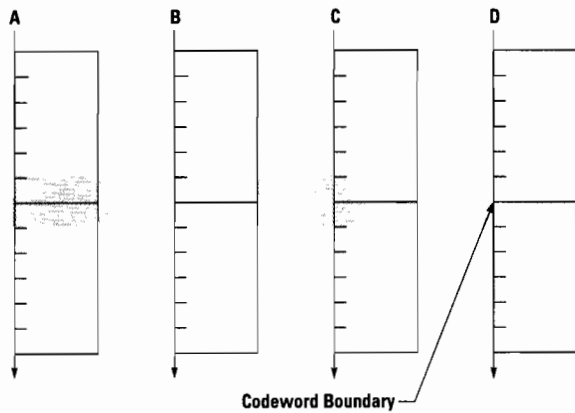
Slight modifications are made for implementation reasons, but as far as the error properties are concerned, this is what is calculated.

The CRC bits are appended onto the data. When the data is received, the CRC is calculated and compared with the received CRC. If there is a difference, the data is known to have been corrupted.

When a corruption occurs in transmission, a number of bits are inverted. Let the errored bits define the coefficients of a polynomial  $E(x)$ ; for example, if errors occur at bit positions  $p$  and  $q$ , the polynomial will be  $x^{n-p} + x^{n-q}$ . Since the operation of calculating CRCs is linear, the error is detected if and only if  $E(x)$  is not exactly divisible by  $G(x)$ . CRC-generating polynomials such as  $G(x)$  are chosen precisely to detect as many polynomials like  $E(x)$  as possible.

In addition to the error detection, the 5B/6B code has other properties that are highly pertinent to physical transmission. The first of these is run length—the maximum number of consecutive zeros or consecutive ones transmitted on any code stream. This is important in transmission systems where some form of clock information is recovered from the

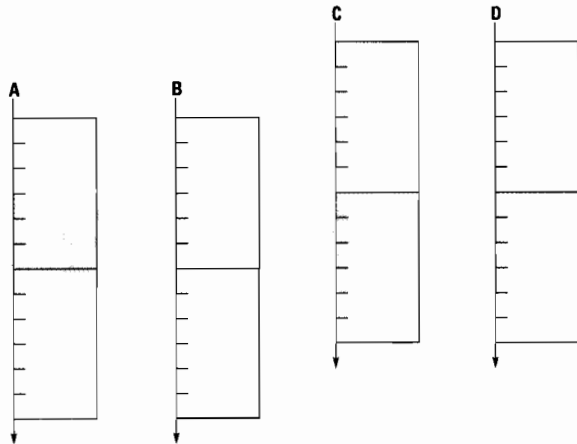




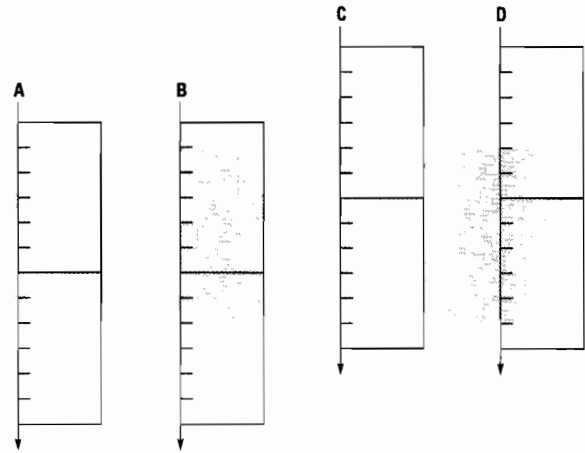
**Fig. 8.** A burst error across synchronous channels. A burst of length 2 corrupts up to eight code bits.

data stream, since clock recovery usually depends on receiving a reasonable density of signal transitions. For the 100VG-AnyLAN 5B/6B code this maximum run length is 6, whether on four channels or on one.

Another physical attribute of the 5B/6B code is the running digital sum (RDS)—the difference between the number of zeros and the number of ones in the transmitted bit stream since the code transmission began. A bounded RDS ensures bounded baseline wander. For the four-pair scheme, the RDS is bounded between  $-5$  and  $+3$ . For the single-stream scheme, the bounds are  $-11$  and  $+3$ . The bounds are not symmetrical because of the alternation rule—a weight-2



**Fig. 9.** Detection of a burst of length 4. A burst of length 4 will always be detected because of the offset between channels.



**Fig. 10.** Detection of a burst of length 7. A burst of length 7 is always detected because of the offset and the coding scheme.

codeword is always sent first on a channel when there is a choice of unbalanced codewords

In summary then, the choice of 5B/6B code, with the alternative end delimiters working on top of the offset transmission scheme, allows 100VG-AnyLAN to meet IEEE 802's requirements. It will detect three separate single-bit errors anywhere in the frame and is capable of detecting substantial burst errors for all the media choices while maintaining favorable physical transmission attributes.

#### Acknowledgments

The authors would like to thank all of the team members at HP Laboratories Bristol and the Roseville Networks Division for their individual contributions, which made the team effort so fulfilling. Special thanks should go to Pat Thaler and Dan Dove at RND and to David Cunningham, Alistair Coles, Miranda Mowbray, David Coker, and Michael Spratt at HP Laboratories, who all provided valuable advice in their own special ways.

#### References

1. *Information Technology—Local and Metropolitan Area Networks—Part 12: Demand Priority Access Method and Physical Layer Specification*, IEEE, 1994.
2. J.H. van Lint, *Introduction to Coding Theory*, GTM 86, Springer Verlag, 1982.
3. R.E. Blahut, *Theory and Practice of Error Correcting Codes*, Addison Wesley, 1983.
4. *Functional Requirements, IEEE Project 802, Local and Metropolitan Area Networks Standards Committee, Draft 6.10.*

# Multimedia Applications and 100VG-AnyLAN

Networks must guarantee bandwidth for multimedia traffic and must control end-to-end delay and delay jitter (fluctuation in the arrival time of packets). The new campus network, 100VG-AnyLAN, can meet these requirements in many circumstances through the basic operation of the protocol. More flexibility can be obtained through the use of bandwidth allocators and the target transmission time protocol. Until either the Broadband Integrated Services Digital Network (B-ISDN) or reliable Internet protocols become available, the use of dial-up remote bridges with existing WANs can accommodate multimedia traffic in the near term.

by John R. Grinham and Michael P. Spratt

The term *multimedia* covers a wide range of applications, ranging from simple image-in-text word processors through video conferencing. The qualifier "real-time" confines the applications to those that impose timing constraints on the underlying delivery system, such as:

- Conferencing systems using synchronized video, audio, and shared computer data
- Remote training and information databases containing digital video and audio data
- Remote control or manipulation of computer-generated models and robotic machines.

As desktop computers grow in power and decrease in price, and networks expand both in area and in data transfer rates, such applications are likely to have a major impact on all desktop computer systems and workgroup LANs in the mid to late 1990s. The factors that indicate a takeoff in multimedia can be divided into two groups: user demand and the development of appropriate technology.

## User Demand for Multimedia Applications

A major demand is for humans to interact with remote systems. This started in the fields of education and training. Computer-based training, which used to use analog video and audio, has become a mature market, and is moving rapidly towards using digital media. In the future, this data may need to be held centrally so that it can be updated easily. Training applications could be networked to this central source. This requires that control data be transmitted to a server with low delay, and that the server have the ability to transmit multimedia data back to the desktop, also with low delay.

Applications are growing in the scientific, engineering, and medical domains in which, for example, humans interact in real time with remote models and image data. These applications also place significant timing constraints on the underlying network.

Perhaps the most exciting growth area is in the support of real-time human conversation. Teams separated geographically need such systems for activities traditionally performed on a single site, such as meetings, training, marketing, and support. As an example, HP has already released a set of collaborative workgroup tools that includes imaging, digital audio, and shared data.

Systems such as these are a natural development from video conferencing systems. It is remarkable that over half of all Fortune 1000 companies use boardroom-based conferencing facilities, given that until recently they cost over \$20,000 per node and were relatively crude in operation. The possibility of low-cost desktop-based conferencing systems that integrate with existing computer applications has to be seen as extremely attractive to companies and teams distributed over several sites.

## Multimedia Technology

Demand for desktop multimedia applications connected over networks can only be satisfied if the technology exists to support it. Three areas are critical:

- The development of standards for video and audio coding that allow low-cost implementation and fast transmission
- The development of low-cost wide-area networks (WANs)
- The development of consistent interfaces to multimedia technology, both from WANs to appropriate LANs, such as 100VG-AnyLAN, and from LANs to applications.

**Video and Audio Coding Standards.** The ISO and the ITU-T (formerly CCITT) have produced several standards for audio and video coding algorithms, such as MPEG and H.261 for video conferencing and video telephony and JPEG for the compression of still images (now being extended to full-motion video). Since all these algorithms are based on discrete cosine transformations, low-cost chipsets that support all three are being developed by several manufacturers.

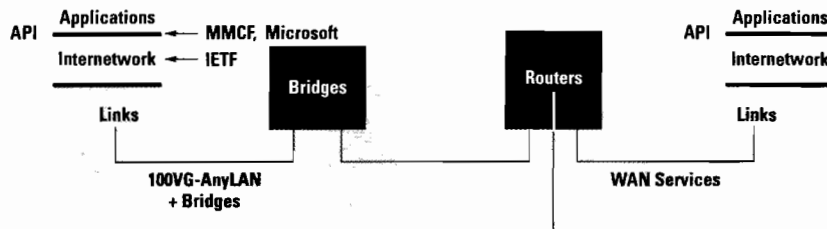


Fig. 1. Developments in different portions of a wide area network.

Audio CD-quality systems are already becoming standard features of workstations and PCs, and many manufacturers have released plans to incorporate hardware support for full-motion video. The increasing power of workstation and PC microprocessors is also making software compression and decompression of full-motion video feasible. HP's current workstations can already perform decompression of full-motion video. Real-time compression is likely to follow in a few years.

**Wide Area Networks (WANs).** The cost of WAN services has fallen dramatically in the last few years. For example, a long-distance video conferencing link that was around U.S.\$1500 per hour in 1984 now costs in the range of U.S. \$12 to \$30 per hour. In addition, there is increasingly widespread availability of a variety of dial-up WAN services.<sup>1</sup>

There is also a significant effort going into the development of wide area broadband services, perhaps leading to a Broadband Integrated Services Digital Network (B-ISDN) by the end of the decade.

**Consistent Interfaces.** For multimedia to work effectively over wide area networks and a variety of campus and local-area networks, there is a need for consistent interfaces between networks. 100VG-AnyLAN only covers a portion of a wide area networked system.

Various bodies are looking at other areas of network systems. Fig. 1 illustrates the situation schematically.

The Multimedia Communications Forum (MMCF), the Winsock 2.0 forum, and the Integrated Services Group of the Internet Engineering Task Force (IETF) are looking at network interface issues, while the IMA has been considering end-system services. Most major operating system and hardware manufacturers, including HP, have been involved in addressing these areas, and are developing standards to ensure consistency of interfaces.

### Demands of Multimedia on Networks

There are three areas in which real-time multimedia applications make demands on networks:

- Workgroup LAN delay must be minimized.
- It must be possible to determine delay bounds for all elements in the transfer path from source to destination so that it can be guaranteed that data will arrive at the destination within a specified time.
- Delay jitter (fluctuations in the arrival time of packets) must be controlled.

The following sections describe these three considerations in more detail and indicate how 100VG-AnyLAN addresses them.

**Control of End-to-End Delay.** The most important new requirement of real-time multimedia applications is for control of end-to-end delay.<sup>2,3,4</sup> The ITU-T's target is a maximum of

150 ms for low-quality voice services. Real-time 3D visualization transmissions are more critical. Even one-way delays of more than 45 ms have been found to cause unpleasant side effects. We decided to set a target of about 100 ms for WAN applications and 30 ms for local applications.

Table I illustrates the contribution to delay of various elements in the data path of a remote collaboration system. The system is shown in Fig. 2.

Table I  
Delays in a Remote Collaboration System

System Component	Typical Delay Value
Video Coder Delay	2 to 150 ms, depending on the CODEC used
Video Decoder Delay	2 to 150 ms, depending on the CODEC used
Packetization Delay	8 ms, using 1500-byte packets and a 1.5-Mbit/s stream
Workgroup LAN Delay	0.1 to over 200 ms, depending on the protocol used
Backbone Network Delay	10 ms typical for each backbone traversed
Router/Bridge Delay	5 to 10 ms, if priorities are implemented
Wide Area Network Delay	5 ms per 1000 km, or about 25 ms for the U.S.A., coast to coast

Of these delays, video coding and decoding have the greatest ranges. CODEC hardware is evolving extremely rapidly. Those using a JPEG processor such as the C-Cubed Microsystems CL550 have a delay of less than 10 scan lines (a few milliseconds). Currently H.261 and MPEG CODECs can impose a considerable delay. These should reduce as the technology progresses, but even so, it is possible for the

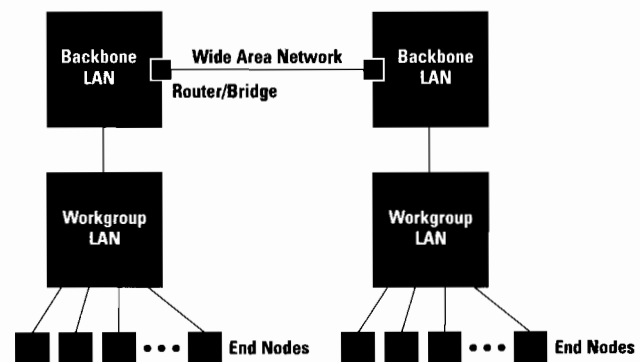


Fig. 2. Example of a remote collaboration network.

CODEC, packetization, and wide area network delays to account for all of the 100-ms delay budget. It follows then, that for effective transmission of real-time multimedia, the workgroup LAN delay must be minimized. To achieve our targets for wide area and local transmissions, we set a workgroup LAN delay target of about 10 ms.

In many cases the basic operation of the 100VG-AnyLAN protocol keeps packet delay across the LAN bounded below 10 ms. As network size and complexity increase, the high priority level can be used for multimedia traffic, possibly in conjunction with resource allocation schemes as described on page 37.

**Bandwidth Guarantees.** Application vendors need to be confident that their applications will work no matter what other traffic is on the network. The need for guaranteed high-bit-rate services to a single node or a few nodes extends across all elements within a network complex, from the LAN to bridges and routers and wide area networks. One simple method of connecting LANs to a WAN is through the use of a dial-up remote bridge (see "Remote Bridge Example" at right).

Within a LAN, 100VG-AnyLAN can guarantee bandwidth to multimedia applications either by exploiting the basic operation of the protocol or, if more flexibility is required, through the use of resource allocation schemes as described on page 37.

**Interarrival-Time Jitter.** Packet switching has been put forward as an effective technology for integrating stream-type data, such as video and audio, with conventional computer data. Packet switching can allow flexible allocation of bandwidth to individual calls, interfaces that are well-matched to computer systems, and efficient multiplexing of sources.

However, packet switching introduces a problem. Packets will be subject to variations in arrival time (known as interarrival-time jitter). How then can reconstruction of the continuous audio or video signal from a set of packets sent over a network be ensured?

Jitter is produced when a source transmits data packets at regular intervals but each packet is subject to a different delay before it arrives. Delay is caused partly by fixed elements, such as the propagation time and packetization delays, and partly by variable delays from queuing and waiting for access to the network. It is these last two that result in jitter.

Provided that any jitter is kept below a maximum value, it is possible to calculate a buffer size at the receiver that is sufficient to smooth out any variation in interarrival times. The

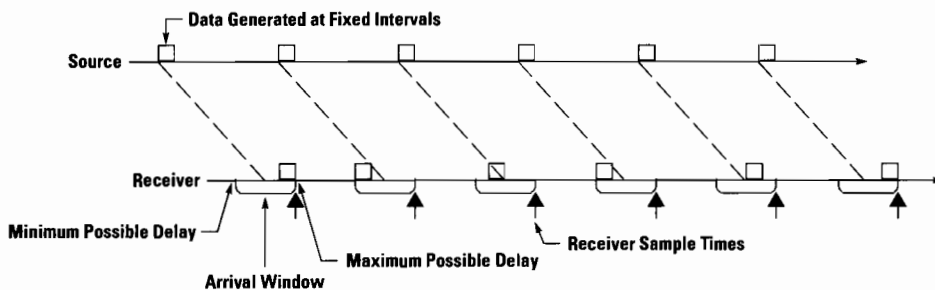


Fig. 3. Recovering jittered data.

## Remote Bridge Example

To illustrate the operation of a dial-up remote bridge, Fig. 1 shows two sites, each with a backbone FDDI network linking 100VG-AnyLAN hubs.

Users connected to end nodes 1 and 2 want to communicate by video and audio between the two sites. The backbone networks on each site are connected to the wide area network (WAN) through dial-up remote bridges specifically designed to carry multimedia traffic. We assume that site A and site B can also communicate through existing routes, such as some kind of corporate internet.

When the call between node 1 and node 2 is initialized, an isochronous link with the appropriate bit rate is set up between the two dial-up bridges. Bridge A is told the 48-bit MAC address of node 2. Any subsequent packets that appear on the backbone FDDI bearing that destination address will be forwarded across the WAN. Bridge B puts them on the backbone FDDI at site B, from which they reach node 2.

Any normal-priority data packets destined for node 2 will not be picked up by bridge A because they will carry the destination address of the router connecting site A to the corporate internet, not that of node 2.

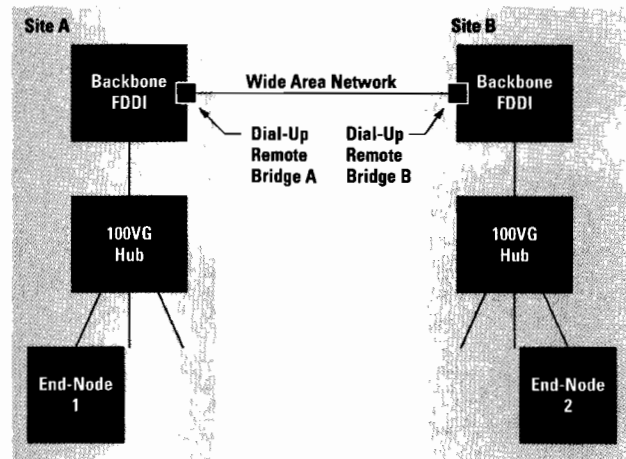


Fig. 1. Remote sites connected by dial-up remote bridges.

data can then be retrieved from the buffer in a steady stream for decoding and display. Fig. 3 shows how this is achieved when an arrival window can be determined.

The arrival window must be small, or the buffering delay will cause the total end-to-end delay of the transmitted data to exceed the 100 ms target.

## Introducing New Applications

Because interface standards and technology are still being developed for the whole wide area network structure, it is

## Higher-Level Protocols

The introduction of new LAN technologies (such as 100VG-AnyLAN) is only one element in the development of multimedia networking. Higher-level protocols are also needed that can control and transfer stream data over multiple network hops. The Internet Engineering Task Force (IETF) and the International Telecommunications Union (ITU-T) have been working in this area.

### IETF

The IETF has been developing protocols for a multiservice, packet-based network spanning the world (see Fig. 1), such as the Internet Stream Protocol (ST-II). This is a network-layer protocol roughly equivalent to Internet Protocol (IP), but specifically developed to support stream-based traffic. ST-II is being used in the Multimedia Teleservices section of the BERKOM-II program (see "Related Projects" on page 38).

The IETF has a number of groups looking at new protocols for multimedia that can be introduced into the worldwide Internet.

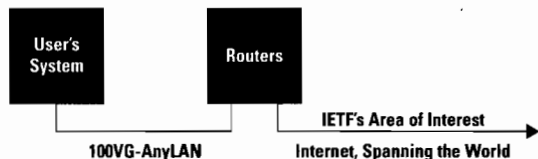


Fig. 1. IETF involvement in worldwide networks.

### ITU-T Narrowband ISDN Standards

The ITU-T works towards developing standards for data transmission over an enhanced telephone network, such as ISDN or other digital circuits with fixed bit rates up to 2 Mbits/s (Fig. 2). 100VG-AnyLAN could provide desktop connections to these services. ITU-T Recommendation H.320 includes the following:

- H.261: Video CODEC for audiovisual services at  $p \times 64$  kbits/s
- H.242: A system for establishing communication between audiovisual terminals
- H.221: A frame structure for a 64-to-1920-kbit/s channel in audiovisual teleservices
- Q.931/2: A D-channel signaling protocol.

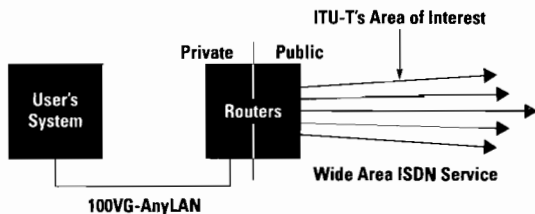


Fig. 2. ITU-T involvement in worldwide networks.

important that there be an easy growth path for the introduction of new multimedia applications. This also satisfies the need for users with existing networks in place to be able to build them up to cope with advanced multimedia traffic in stages, using the existing infrastructure and avoiding a large initial outlay. Unlike other high-speed campus LANs, it is possible with 100VG-AnyLAN to provide such a growth path for the introduction of multimedia applications.

The growth path has three phases:

- Phase 1. Run the application using a single priority level throughout the network.

- Phase 2. Move some nodes to high priority as multimedia traffic increases, with simple mechanisms for monitoring and controlling bandwidth.
- Phase 3. Introduce automatic bandwidth allocation mechanisms for all high-priority traffic.

### Phase 1: Multimedia Using a Single-Priority 100VG-AnyLAN.

In the first phase, since no new API for applications has been developed, the application is simply put on normal priority all the time. Two of the properties of the 100VG-AnyLAN demand priority protocol enable it to be used with multimedia applications even in this phase: bounded delay and bandwidth sharing.

**Bounded Delay.** The access delay for a data packet is guaranteed to be less than or equal to the maximum packet duration times the number of active ports on the hub. If Ethernet-length packets are used, the maximum packet length is 12,000 bits, giving a duration of 0.12 ms. With a 32-port hub, a queued packet will have to wait at most  $0.12 \text{ ms} \times 32 = 3.84 \text{ ms}$ . This will increase to 7.84 ms if all traffic is multicast and the system is using bundles.

These figures compare with  $32 \times \text{TTRT}$  (target token rotation time) for asynchronous FDDI (giving 320 ms for a 10-ms TTRT), and unbounded for CSMA/CD.

**Bandwidth Sharing.** Every active node has an equal share of transmission opportunities. In the worst case, when every node is actively transmitting full-length packets of data at the same priority, this will give a maximum bandwidth of about 3 Mbits/s per node on a 32-port hub. This reduces to 1.5 Mbits/s if all data is multicast in bundles. This guarantee is more than sufficient for many applications such as simple conferencing.

It follows that a 32-port hub can support simple applications using real-time data at 1.5 Mbits/s or less, and application vendors can be confident that the system will work no matter what other data is on the LAN.

### Phase 2: Use of High Priority without Allocation

If the second priority is used, the two properties listed in Phase 1 apply to high-priority traffic. If 32 nodes run at high priority in a large network, say of 100 nodes, the guarantees would be the same as in the previous example.

A malfunctioning high-priority node can be prevented from saturating the network by network monitoring (the hub can refuse access to a high-priority transmission), and by configuring the network so that normal-priority traffic is raised to high-priority after waiting a certain period of time.

This provides a simple way of introducing multimedia applications, but in the longer term, a mechanism for allocating bandwidth may be required. This will control the use of bandwidth by high-priority traffic automatically, and will ensure that delay is kept within an upper bound on large networks.

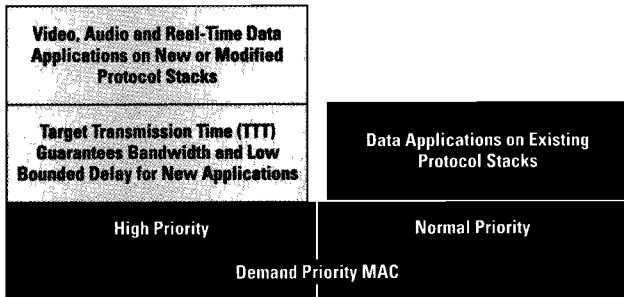


Fig. 4. 100VG-AnyLAN end node stacks.

### Phase 3: Resource Allocation in 100VG-AnyLAN

A third property of 100VG-AnyLAN facilitates the allocation of bandwidth on networks: spare bandwidth utilization. Any bandwidth not used by high-priority traffic is immediately available for normal-priority data.

This property is particularly relevant if high-priority traffic is bursty. It also ensures that normal-priority traffic does not need to participate in the allocation process.

Fig. 4 illustrates how the protocol stacks for high and normal priority are independent. This is important because resource allocation protocols need only be introduced on end systems that actually run multimedia applications. End systems that only use normal priority do not need any new software. Note that the use of a bandwidth allocator requires higher-level interfaces and services to be defined, such as those on which the IETF is working (see "Higher-Level Protocols" on page 36).

**Example of Bandwidth Allocators.** Consider the following simple scenario illustrating the use of a bandwidth allocator. Several 100VG-AnyLAN hubs are connected by an FDDI ring. The allocator ensures that the sum of the high-priority synchronous multimedia connections across the FDDI does not exceed 80 Mbits/s, and that the sum of the high-priority multimedia connections across any individual 100VG-AnyLAN network does not exceed 80 Mbits/s. This ensures that the 100VG-AnyLAN networks and the FDDI backbone are not overloaded with high-priority traffic.

Suppose that all the 100VG-AnyLAN hubs are 32-port or smaller. The protocol guarantees that in the event of all nodes simultaneously wishing to transmit normal-priority data, they will each have a bandwidth of a little over 3 Mbits/s. The allocator then only has to ensure that the sum of the high-priority synchronous multimedia traffic across the FDDI does not exceed 80 Mbits/s. A system such as this can typically support up to 80 video connections each with a peak bandwidth of 1 Mbits/s, or up to 200 364-kbit/s connections.

In 100VG-AnyLAN, the proposed mechanism for allocating bandwidth and setting delay bounds is the target transmission time.

**Example of Target Transmission Time Allocation.** A nominal time value, the target transmission time (TTT) is set, and nodes request permission from the allocator to transmit a maximum amount of data within a TTT period. Provided that the sum of all such requests is less than a preset maximum, say 80% of the network capacity, the allocator grants the request. The end node is then guaranteed that its data will be transmitted during the TTT period.

Fig. 5 shows a simple example of how this could work. In the example, there are four stations operating at high priority. Nodes 1 and 2 are each generating 30 Mbits/s, node 3 is generating 7.5 Mbits/s, and node 4 1.5 Mbits/s.

Node 4 generates one 1500-byte packet every 8 ms and requires this to be transmitted onto the network within 8 ms. If this is the lowest access delay, the TTT would be set at 8 ms. In this time period, the other nodes will generate 30kB (20 1500-byte packets), 30 kbytes, and 7.5 kbytes (5 1500-byte packets) of data, respectively.

Fig. 5 shows how the data in each of the end nodes' transmission queues is transmitted on the network.

**Bandwidth Monitoring.** High-priority traffic can be monitored centrally at the hub or through a separate device attached to a promiscuous port. From packet lengths and interpacket gap times, the monitor can easily calculate the aggregate

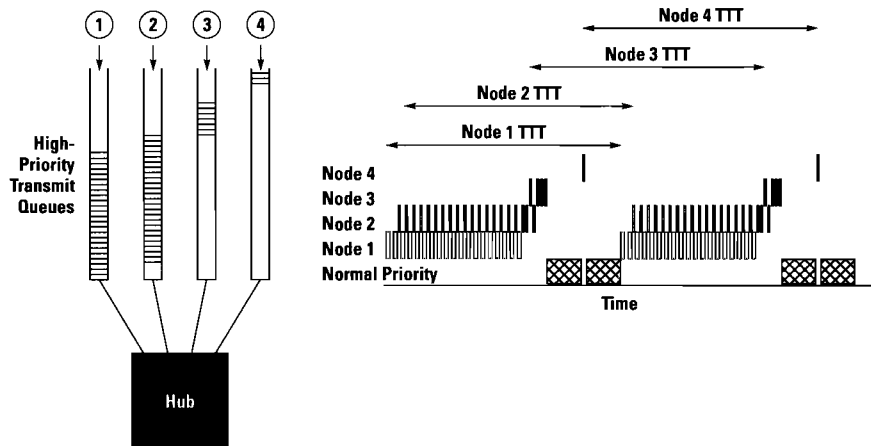


Fig. 5. Example of the use of the target transmission time (TTT).

---

---

## Related Projects

Two projects are under way that have a bearing on the development of networks to carry multimedia traffic.

### Teleservices Project

The Teleservices project is part of the Deutsches Bundespost (DB) Telekom BER-KOM technology and demonstrator program. The program covers wide area network and campus broadband communication studies, using broadband switches and local optical-fiber links installed in Berlin to form a large test bed. It now forms part of the German Broadband Integrated Services Digital Network (B-ISDN) asynchronous transfer mode (ATM) pilot network, with switching centers in Cologne and Hamburg.

The Teleservices Project is investigating the provision of multimedia services over wide area, broadband networks. HP is involved in two areas: multimedia collaboration (MMC), supporting conferencing and shared applications, and multimedia transport (MMT), developing high-speed networking protocols for multicast operation.

HP joined in April 1993, with collaboration from HP Germany, HP Laboratories in Bristol, and TUB (Technical University of Berlin, as an HP subcontractor). The main HP/TUB task is the development of high-speed network protocols and the overall implementation and integration of the complete collaboration application under the multimedia collaboration task for the demonstration of interoperability. Five HP 9000 Model 755 computers have been installed at the TUB for the multimedia implementation, together with a range of audiovisual equipment, communications interfaces, and associated software packages. Prototype HP ATM switches have also been installed.

The program will produce a limited HP multimedia communications and management product for use on HP 9000 Model 7x5 workstations in the German broadband trials. It will also provide measurement and performance evaluation results, including:

- Observations of the interactions between video and audio compression schemes, transport protocols, the different network technologies used, and their impact on perceived performance

---

---

throughput of the hub for each TTT. If the aggregate is higher than a certain threshold value, further action can be taken on individual streams. For example, the network manager could disable high-priority traffic at a particular port, or this could be done automatically through the network monitor. A technique such as this provides a high level of security to other nodes, even from malicious attempts to disrupt the high-priority service.

- Identification of particular issues and problems associated with multimedia applications related to HP products and commercial interests
- Performance measurement and behavior of a transport protocol designed for multimedia applications running over different network technologies.

A demonstration was given in August 1993 using the first version of the conference application with software video compression and relatively simple control options. In March 1994, an enhanced version was demonstrated at the CeBIT exhibition in Hannover. This had a range of management and control options, an audio server, and full-motion hardware compression (M-JPEG). It also included multivendor terminal interworking at the show over the ATM demonstration network and conferencing over ATM link connections back to a terminal in Berlin.

A demonstration of the Teleservices package between HP Laboratories in Bristol and the TUB, using ISDN at about 300 kbits/s showed that a minimum transfer rate of 1 Mbit/s is required for successful operation. Wideband operation over ISDN connections is also being evaluated.

### Distance Learning Seminar

In October 1994, a pan-European seminar on distance learning was held on a network linking three European universities (the Royal Institute of Technology, Stockholm, the Technical University of Berlin, and University College London) and various other European sites. The seminar was sponsored by HP and aimed to promote the use of distance learning and to explore the future of the technology and the role of multimedia.

The seminar was a practical demonstration of multimedia over a mix of transmission technologies. German Telekom provided B-ISDN (broadband ISDN ATM) links between Berlin and London and Berlin and Stockholm. Other links used inverse multiplexers at each site to provide contiguous bandwidth over ISDN of 384 kbits/s. This provided adequate bandwidth for 25-frames-per-second video using hardware compression based on the H.261 standard, plus good-quality audio.

## References

1. D. Blankenhorn, "MCI Announces Switched T1 and T3 Services," *Newsbytes*, 1991.
2. "General Characteristics of International Telephone Connections and Circuits, Recommendations G.100-G.181," *ITU-T Blue Book*, Vol. III, Fascicle III.1, 1988.
3. A. Shah, D. Staddon, I. Rubin, and A. Ratkovic, "Multimedia over FDDI," *LCN Conference Proceedings*, 1992, p. 110.
4. I. Corbett, "Videophony," *Audiovisual Telecommunications*, BT Telecomms Series, Chapman & Hall, 1992.



# 100VG-AnyLAN 15-Port Hub Design

Much of the intelligence and uniqueness of a 100VG-AnyLAN network is concentrated in the hub. Special repeater, transceiver, and end node chips implement the functionality of the HP J2410A AdvanceStack 100VG Hub 15.

by Lisa S. Brown

Hewlett-Packard has introduced several network products for 100VG-AnyLAN applications. Products available since mid-1994 include a 15-port hub, a 10Base-T/100VG bridge and SNMP management module, a 10Base-T/100VG ISA adapter card, and a 10Base-T/100VG EISA adapter card. These products were developed before completion of the IEEE 802.12 standard and are currently under modification to include changes that occurred in the standard between the time of product definition and the final version of the standard. Most notably, this includes changes to support token ring framing. Additional 100VG-AnyLAN products available in 1995 are a 10Base-T/100VG PCI adapter card, a multiport hub, a fiber hub, and a 10Base-T/100VG switch.

Much of the intelligence and uniqueness of a 100VG-AnyLAN network is concentrated in the hub. Therefore, this article

mainly focuses on the design of HP's 15-port hub, the HP J2410A AdvanceStack 100VG Hub 15 (see Fig. 3 on page 9). Some 100VG-AnyLAN-specific aspects of the other products are also explored in this article.

## Hub Architecture

The key components of a 100VG-AnyLAN hub are the repeater chips and transceivers. In the HP 15-port hub solution, a microprocessor subsystem provides configuration and management functions. Fig. 1 is a block diagram of the 15-port hub.

The repeater chip implements the demand priority protocol, the signaling to receive and transmit packets, and the interfaces to the system. The system interfaces include a port

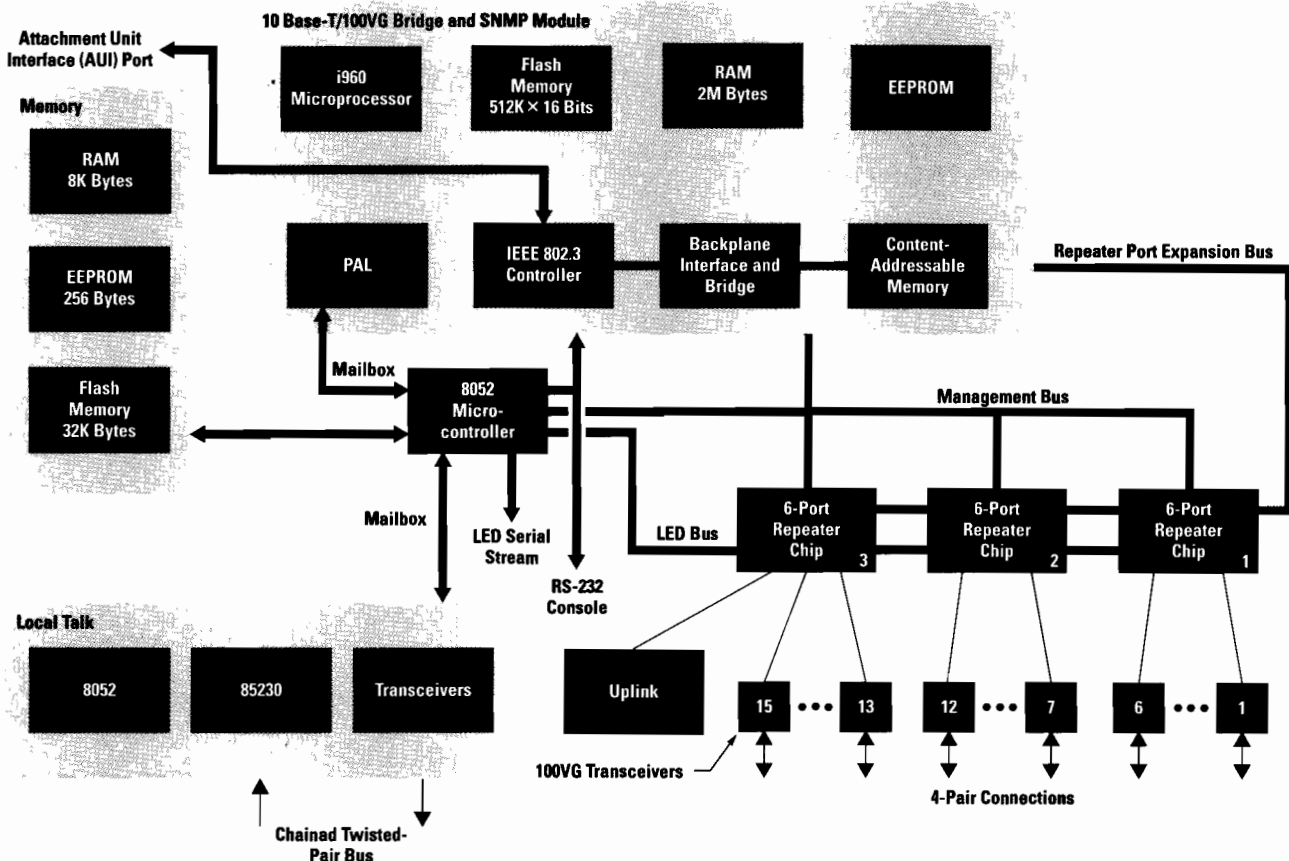


Fig. 1. Block diagram of the HP J2410A AdvanceStack 100VG Hub 15, a 15-port 100VG-AnyLAN hub.

expansion bus to connect several repeater chips into one logical repeater and an interface for network management functions, such as control and status information. The repeater chip is a six-port device that connects directly to six transceivers. Additionally, it has a single uplink port which also connects to a transceiver for use in a cascaded hub topology. For the 15-port hub, three repeater chips provide the 15 ports plus one downlink. The repeater chips act as one logical repeater through the use of the parallel expansion bus.

The transceiver chip implements the physical layer as described in the article on page 18. It is used in conjunction with a filter specially designed for use in 100VG-AnyLAN applications. The 15-port hub uses 16 transceiver chips, one for each of the 15 ports plus another for the uplink connection. The transceivers connect directly to the repeater chips.

### Repeater Chip

The repeater chip logic was developed by HP and AT&T engineers in conjunction with the IEEE 802.12 standard development. It is now commercially available through AT&T as the ATT2R01. It implements the following functions:

- Demand priority protocol
- 100VG-AnyLAN RMAC (repeater MAC)
- Port expansion bus to connect multiple repeater chips together
- Interface to six transceivers for downlink ports
- Interface to transceiver for uplink port
- Packet buffer RAM
- Six-port CAM (content-addressable memory)
- Management bus
- Control and status registers
- Management counters.

Fig. 2 is a block diagram of the repeater chip's internal modules.

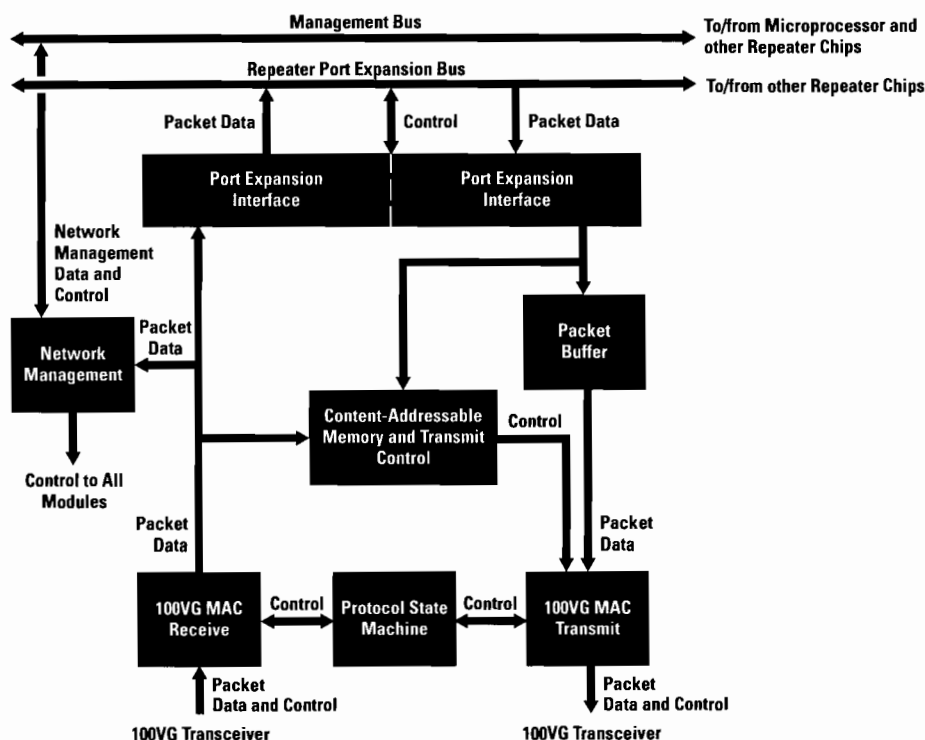


Fig. 2. Repeater chip block diagram.

A packet traverses through the hub in the following manner. The protocol state machine polls the ports as described in the article on page 13 and determines which port will receive a packet from its end node next. It initiates the proper tone sequence through the transceiver and readies the hub to receive a packet. When a packet is received, it first passes through the receive portion of the 100VG repeater MAC. The data is decoded as described in the article on page 27 and is passed through as bytes. Statistics for network management functions are gathered on the packet data and the packet is passed through to the port expansion bus. Each repeater chip on the port expansion bus loads the packet into its own internal packet buffer. The packet buffer is implemented as a RAM that is large enough to hold a maximum-length packet. The first six bytes of the packet are sent to the CAM. At this point, the destination address of the packet is decoded and compared with all entries in the CAM. The transmit control circuit decides which ports, if any, will transmit the packet to an end node. This is based on whether there was a match with a CAM entry, if the packet was a broadcast or multicast packet, or if any of the ports are configured as promiscuous. Finally, the packet passes through the transmit portion of the 100VG repeater MAC where it is encoded for transmission as described in the article on page 27.

The repeater chip also implements the training functions as described in the article on page 13. The protocol state machine controls the training sequence in a manner similar to its control of data transfers. As the training packets are received, the chip learns the status of the connected port and writes the source address into the internal CAM. It monitors the received packet, and if an error is detected the packet is not counted as a good training packet. Training continues until 24 sequential good packets are transmitted in each direction (see "Invalid Packet Marker" on page 41). If the

---

---

## Invalid Packet Marker

When a packet has been received with an error, the invalid packet marker is appended when the packet is retransmitted to an end node or another hub in a cascaded network. Invalid packet marker information is useful during training and during normal packet transmission.

During training, if a packet is received at the end node with an invalid packet marker attached, this implies that an error occurred on the link where the packet was transmitted from the end node to the hub. If a packet is received at the end node with no invalid packet marker but an error is detected on the link, this implies that an error occurred on the link where that packet was transmitted from the hub to the end node. In either case, the packet is not counted as a good training packet.

During normal packet transmission, the detection of the invalid packet marker indicates that an error occurred in a link upstream from where the invalid packet marker was detected. For example, in a cascaded repeater network, the combination of receive error and invalid packet marker detection can be used to determine which link in the cascade produced the error. Excessive errors on any one link could indicate a marginal physical connection. A network administrator might then decide to examine the link to improve network performance.

---

---

link is faulty, the node will not pass training and the protocol state machine will pause the training attempt, wait one to two seconds, and then allow training to be attempted again.

The management bus allows the processor subsystem to access the repeater chips. In the network management block of Fig. 2 there are many control registers, which network management applications can set as needed. These include functions such as disabling ports, allowing for promiscuous ports, setting the hub in a bundled cable configuration, security functions, and a host of others. The processor also has read access to the CAM, which can aid network managers in building a map of the network. Additionally, there are sixteen 24-bit counters per port and a number of status registers in the network management block that can be read via the management bus. The counters accumulate statistics on each port such as the number of bytes received, good and bad packets received, errors, high-priority and normal-priority packets received, and others. The LED control interface is a simple bus that provides port status information to the 8052 microcontroller for driving the hub LEDs.

### Transceiver Chip

The transceiver chip logic was developed by HP and AT&T engineers in conjunction with the IEEE 802.12 standard development. It is now commercially available through AT&T as the ATT2X01. It implements the following functions:

- 100VG-AnyLAN physical layer (PHY)
- Quartet signaling
- Adaptive equalization
- Timing recovery
- Data justifier
- Receive control signals (RCS) machine
- Transmit control signals (TCS) machine.

Further information on the functionality of the transceiver chip can be found in the article on page 18.

## Network Interface Card Architecture

The key components in the network interface cards, such as the 10Base-T/100VG ISA and EISA adapter cards, are the 10Base-T/100VG end node chip, the external buffer RAM, the 100VG transceiver, and the 10Base-T transceiver. The 10Base-T transceiver is available commercially through a number of vendors and is used to provide a 10-Mbit/s interface to Ethernet networks. The 100VG transceiver is the same device that is used in the hub, the ATT2X01, and is described above. The 10Base-T/100VG end node chip is a new device to accommodate both 10-Mbit/s and 100-Mbit/s networks.

The 10Base-T/100VG ISA and EISA adapter cards are implemented with the same end node chip, which includes both the ISA and EISA system buses and the 100VG and 10Base-T interfaces to transceivers. An EISA workstation adapter card is available for HP 9000 Series 700 servers. The end node chip was developed by HP and AT&T engineers and is now commercially available from AT&T as the ATT2MD01.

The end node chip implements ISA and EISA bus interfaces specifically tuned to handle 100 Mbits/s in the most efficient manner possible. The EISA interface provides bus mastering to allow 100VG servers to achieve LAN link rates with minimal CPU overhead. These bus interfaces comply with the industry-standard specifications for ISA and EISA machines.

Each end node device implements a 100VG module, which consists of a 100VG MAC and a protocol state machine. Through direct connection to a 100VG transceiver, this module generates and recognizes the appropriate tone sequences for training and data packet transmission. It converts internal chip bus data to 5B/6B encoded data and it decodes incoming 5B/6B data for the internal chip bus.

The end node chip also implements an interface to a 10Base-T network. It complies with all IEEE 802.3 specifications and connects directly to commercially available 10Base-T transceivers. This interface allows customers who may still be installing 10-Mbit/s networks to upgrade easily to 100VG-AnyLAN in the future without replacing adapter cards.

### Other System Components

HP has also implemented a combination 10Base-T/100VG bridge and SNMP module. The module provides a connection from a 100VG network to a 10-Mbit/s Ethernet legacy network. It also has enough memory to load and run SNMP management functions for Ethernet and token ring networks.

This module slides into the hub and interfaces to the management bus and the port expansion bus on the repeater chip. The module has an i960 microprocessor and memory subsystem. The i960 is fast enough to process the counter information from the repeater chips in real time. The large 2M-byte RAM provides enough memory for both SNMP and bridging functions to run on the card.

This module also has an IEEE 802.3 controller chip to interface between 10-Mbit/s and 100-Mbit/s systems. Between this controller chip and the bus attached to the repeater

chip is an ASIC which controls the data and management information flow through the module. The module has direct access to the repeater data bus so that packet data can flow between the IEEE 802.3 controller and the 100VG network.

#### **Acknowledgments**

The introduction of the products discussed in this article was the result of tremendous effort and dedication on the

part of the 100VG-AnyLAN team in Roseville and Bristol. Through the well-orchestrated efforts of the research, design, manufacturing, and marketing teams these products were successfully introduced into the marketplace, thus establishing 100VG-AnyLAN as a viable low-cost, high-speed network for the applications of today and tomorrow.

# Authors

August 1995

## 6 100VG-AnyLAN

### Alan R. Albrecht



Alan Albrecht graduated from the University of New Mexico with a BS degree in computer engineering in 1986. He joined HP's Roseville Networks Division in 1986 and is currently principal engineer in the high-speed workgroup, responsible for 100VG architecture, products, and standards, with a primary focus on demand priority protocol.

Lead engineer for 100VG-AnyLAN since its inception, his roles have included inventor, principal architect, and ASIC designer. His previous accomplishments at HP include work on the first two generations of Ethertwist 10Base-T hubs and ASIC design for a 10Base-T adapter. His work has produced several patents pending in the area of high-speed network technology. Alan is a member of the IEEE, has coauthored two papers on demand priority, and is interested in computer networks. He is married and has a three-year-old daughter.

### Patricia A. Thaler



Pat Thaler came to HP's Stanford Park Division in 1976 and is presently a principal engineer for LAN architecture and standards at the Roseville Networks Division. She was born in Utica, New York and attended the University of California where she received a BSEE degree in 1974. She continued her studies at Stanford University, receiving an MSEE in 1978. Previous accomplishments at HP include production engineering for the HP 8660 signal synthesizer, computer interface card hardware design, and Ethernet LAN product and standards engineering which included development of network interface cards, repeaters, and 10Base-T transceivers. She was a major contributor to the IEEE repeater standard and

chaired the IEEE 802.3 10Base-T task force. She is presently chair of IEEE 802.12 and is responsible for IEEE 802.12 enhancements and standards efforts that include gigabit speeds, full duplex, redundancy, and two-pair Category 5 PMD. She contributed to the overall architecture and system design of 100VG-AnyLAN and led the associated standards development. She is a senior member of the IEEE and the Society of Women Engineers and is interested in digital and analog circuit design and LAN technology including media access control methods and physical layer design. Her work has produced a patent related to 100VG-AnyLAN coding, and two more patents are pending. She is the author of a network systems tutorial for IEEE 802.3. Before joining HP, she was with IBM, where she developed machines for the manufacture of magnetic disks. Pat is married and has three sons. She participates in a radio amateur civil emergency service and enjoys pond gardening and koi, origami, and puzzles and games.

## 13 Demand Priority Protocol

### Alan R. Albrecht

Author's biography appears elsewhere in this section.

### Michael P. Spratt



Michael Spratt is a senior engineer with HP Laboratories at Bristol, England. He joined HP in 1989 and is currently investigating networking technologies for evolving data services to the home. He attended Bristol University for his BSc degree

in mathematics, then went on to Hatfield Polytechnic where he earned an MSc in digital systems in 1986. Previous work at HP includes the design of infrared multiple-access protocols, error propagation analysis for FDDI over copper twisted-pair cable, and a published analysis of nonunity ratio bandwidth sharing. Before joining HP, Michael was with British Telecom Research Laboratories, where he worked on switched-star cable-TV video-on-demand network development and on aspects of networking technology. He has been involved since the beginning of the project in the design and development of the 100VG-AnyLAN protocol, its extension to cascaded hubs, and the use of two priorities for multimedia traffic. His publications include ten journal and conference papers on new network protocols and performance analysis. His work has contributed to five patent applications in the areas of infrared multiple-access protocol and 100VG-AnyLAN technology. His interests include local area network design, LAN/WAN interface and interaction, and the development of new services and applications.

### Patricia A. Thaler

Author's biography appears elsewhere in this section.

### Gregory C.A. Watson



A native of Lilongwe, Malawi, Greg Watson attended the University of Manchester, receiving a BSc degree in computer engineering in 1984 and an MSc in systems design in 1985. He continued his studies at École Nationale Supérieure des Télé-

communications in Paris, France, where he received a PhD in 1992. With HP Laboratories at Bristol, England since 1985, he is presently a senior engineer developing products based on 100VG-AnyLAN technology at the Roseville Networks Division. He was responsible for the technical description of the MAC protocol within the IEEE 802.12 standard. His previous HP accomplishments include work on parallel architectures for declarative languages, FDDI, and gigabit-per-second LANs. He is interested in high-speed local area networks and LAN switches and has authored seven-teen journal articles about high-speed networks. Greg is a member of the IEEE and the ACM. He is married, has a son, and enjoys unicycling together with juggling.

## 18 Physical Signaling

### Alistair N. Coles



Born in Plymouth, United Kingdom, Alistair Coles attended the University of Bath, graduating in 1988 with a degree in electrical and electronic engineering. He has been a member of the technical staff of HP Laboratories at Bristol, Eng-

land since he joined HP in 1988 and has been involved in research on high-speed optical links, passive optical networks, optical amplifiers, and free-space optical links. He was involved with physical layer performance verification, specification of IEEE 802.12 physical layers, and physical layer standardization for the project reported in this issue. He is member of the IEEE and has written several papers in the areas of optical amplification and data communications. His work has produced a patent in the area of optical signal processing. He is married and enjoys playing the guitar and hiking.

### David G. Cunningham

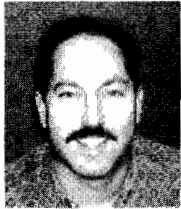


With HP Laboratories at Bristol, England since 1987, David Cunningham is a project manager for link technology in the networks and communications lab. A graduate of The Queens University at Belfast, he has a BSc degree in physics (1981) and

a PhD degree in laser physics (1985). He was the HP Laboratories project manager for 100VG-AnyLAN and served as standards coordinator/manager for the HP Laboratories IEEE 802.12 efforts. His previous HP projects include a gigabit network, the Serial HIPPI standard, and an FDDI physical network layer. His work has produced various patents in the areas of integrated optics, optical signal processing, coding,

and communications systems. He is interested in communications systems and laser physics and has published articles about atomic molecular physics, laser physics, integrated optics, optical signal processing, and communications systems. Before joining HP, David was a research assistant at The Queens University where he worked on high-powered lasers and spent two years at British Telecom Research Laboratories working with integrated optics. He was born in Belfast, Ireland, is married, has a son, and enjoys swimming, reading, movies, and most of all his family.

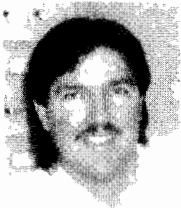
#### Joseph A. Curcio, Jr.



Joe Curcio is engineering manager at HP's Roseville Networks Division. He joined HP in 1984 after graduating from the University of California at Berkeley with a BSEE degree. On the 100VG-AnyLAN project, he was

involved with physical layer design and testing and hardware design of PC cards, and participated in the IEEE 802.12 standardization effort. He previously served as a development engineer and project lead for several HP networking products. He is a member of the IEEE and is interested in high-speed links. Joe was born in Oakland, California, is married, has four boys, and participates in youth sports activities as a soccer, baseball, and basketball coach. He's also a flight instructor.

#### Daniel J. Dove



Born in Torrance, California, Dan Dove attended California State University at Sacramento, receiving a BSEE degree in computer communications in 1988. A senior design engineer with the HP Roseville Networks Division, Dan has been with HP since

1980 and has worked on a wiretest instrument, Ethernet transceivers and repeaters, and 100VG-AnyLAN repeaters and transceivers. He began his career at HP as a production assembler after serving in the U.S. Navy, and he is currently developing high-performance link technology. Dan architected and co-designed the first 100VG transceiver, codeveloped the quartet signaling technology, and codeveloped the media-independent interface for 100VG-AnyLAN. He is a member of the IEEE and is interested in analog circuits, EMC design, and LAN physical layers. He is married, has three children, and enjoys trout fishing, camping, hiking, and acrobatic "skurfing" behind a ski boat.

#### Steven G. Methley



Steven Methley is a member of the technical staff at HP Laboratories at Bristol, England. He was born in Barnsley, Yorkshire, England, and attended Imperial College, University of London, receiving a BSc degree in engineering in 1979 and a PhD in

solid-state electronics in 1982. He has been with HP since 1989 and is presently a project leader for digital signal processing (DSP) for LANs. For 100VG-AnyLAN, Steve did electrical measurements and served as a subeditor for a chapter of the IEEE 802.12 standard. His previous work at HP includes receiver design for the HP 8146 OTDR, a low-cost optical receiver for Serial HIPPI, gigabit LED demonstration, ATM emissions work, and DSP investigation. His work has produced four patents related to optical receivers and one patent pending for active terminations. Before joining HP he was with British Telecom Research Laboratories where he worked on long-haul transmission systems. He is a member of the IEEE and is interested in low-noise design and signal processing. He has published six papers about optical communications and one paper on communication theory. Steve has a teenage son and enjoys off-road driving, mountain biking, swimming, classic car restoration, cinema, and dining out.

### 27 Coding

#### Simon E.C. Crouch



Project manager Simon Crouch leads the mathematics group at HP Laboratories at Bristol, England. Since joining HP in 1985 at the Pinewood Division, he has served as a performance engineer, senior quality engineer, and member of the

technical staff at HP Laboratories before becoming a project manager. Born in Oxford, England, he has an MA degree in mathematics from Cambridge University (1980) and a PhD in mathematics from Warwick University (1985). His work on 100VG-AnyLAN involved software coding and research into error detection issues, and resulted in several patent applications. He is a member of the American Mathematical Society and the London Mathematical Society and an associate fellow of the Institute of Mathematics and its Applications. His professional interest is the application of advanced mathematics to engineering. Before coming to HP, Simon was with the United Kingdom Atomic Energy Authority, working in the area of remote sensing. He is married and enjoys playing the flute, singing in his church choir, and listening to classical music. His artistic ambitions would be fulfilled by discovering some of the lost cantatas of J.S. Bach and by singing Siegfried at Bayreuth.

#### Jonathan Jedwab



A native of London, England, Jonathan Jedwab is a member of the technical staff of HP Laboratories at Bristol, England. He attended Cambridge University, receiving an MA degree in mathematics in 1986 and a diploma in mathematical statistics in

1987. He continued his studies at London University, where he earned a PhD in mathematics in 1991. He came to HP in 1988 and has contributed to a software implementation of the Rivest-Shamir-Adleman cryptosystem, the design of a bit error rate tester, and the analysis of sampling in HP embedded advanced sampling environment (EASE) products. He designed the 5B/6B code for 100VG-AnyLAN. His current responsibilities include coding techniques for high-speed data transmission. His work has produced two patents and five pending patents in the area of digital information communication. He has published twenty journal papers in the area of discrete mathematics. Jonathan is an associate fellow of the Institute of Mathematics and its Applications, a fellow of the Institute of Combinatorics and its Applications, and a member of the Mathematical Association of America. He was recently awarded a 1994 Kirkman Medal of the Institute of Combinatorics and its Applications. He is interested in combinatorial design theory and its application to digital signal design. He enjoys wildlife observation and photography and almost any form of dancing, including rock, jive, acrobatic, Israeli, Latin American, and ballroom.

### 33 Multimedia Applications

#### John R. Grinham



A project manager with HP Laboratories at Bristol, England, John Grinham is presently responsible for networked multimedia systems development. He was a member of the 100VG-AnyLAN team, specializing in LAN requirements for new

applications. His earlier work at HP has included FDDI implementation, simulation studies of multipath ATM switch architectures, and development of a networked video jukebox including end-system application software and network protocol design. He is a member of the IEEE and the Audio Engineering Society and is interested in real-time multimedia systems. His work has produced a patent for an RF spectrum analyzer based on a recirculating fiber-optic delay line and he has coauthored four papers on networked multimedia systems. John holds a BSc degree in physics (1984) and joined HP Laboratories in 1987. Before coming to HP he was with Cossor Electronics where he developed optical signal processing systems for crosscorrelation and spectral analysis of communication signals. He enjoys the study of human perception of audio energy and its impact on recording and listening to music.

#### Michael P. Spratt

Author's biography appears elsewhere in this section.

## 39 15-Port Hub Design

**Lisa S. Brown**



An R&D project manager with the Roseville Networks Division, Lisa Brown is presently responsible for 100VG-AnyLAN hub and core technology development. She served as a design engineer for the 100VG ASIC repeater chip used in the HP

AdvanceStack Hub 15 and managed the development of the Hub 15 and the 10Base-T/100VG bridge and SNMP module. She was born in New Hartford, New York, and attended Northwestern University at Evanston, Illinois, receiving her BSEE degree in 1987. Before coming to HP, Lisa was with Northern Telecom where she served as a senior hardware design engineer for PBX systems and at GTE Communications Systems where she developed ISDN hardware. She is a member of the IEEE and the Society of Women Engineers, and participates in the SWE outreach program to encourage young women to pursue engineering careers. She is married, has a new baby boy, and enjoys aerobics and piano.

## 43 HP AccuPage 2.0

**Steve Webb**



An R&D engineer with Greeley Hardcopy Division, Steve Webb joined HP in 1978. He received a BSME degree from the University of Florida in 1978 and a BSCS degree from Colorado State University in 1988. He is currently responsible for

firmware design and development for the next generation scanner. Previous HP accomplishments include scanner firmware design for the HP ScanJet IIc and the user interface design for scanner applications used with earlier HP scanners. Steve developed the interface code that lies between OCR applications and HP AccuPage technology. He also developed scanner control code and memory management code for HP AccuPage 2.0. His work has resulted in eight patents related to scanner performance. Born in Miami, Florida, he is married and has one daughter. He enjoys skiing, hunting, camping, sailing, and computer games.

**Steven G. Henry**



A project manager at the Greeley Hardcopy Division, Steve Henry came to HP in 1977 at the Fort Collins Division. He attended the University of California at Davis, from which he received a BSEE degree in 1977, and Colorado State University,

where he earned an MSEE in 1983. Past contributions at HP include software development for the HP 250 computer, materials engineering, hardware and software development, management of a project involving technology transfer from HP Labs, and work

on HP AccuPage 1.0. He is presently a project manager for the team working on third-party systems and connectivity. He provided project management for the HP AccuPage 2.0 project. Steve is named as coinventor in two patents, one related to a tape drive load sequence and a second concerning scanner color separation techniques. His publications include a paper on microprogram compaction using reduction operators, a coauthored article on tape backup, and a paper presented at HP's project management conference on technology transfer from HP Labs. Born in Heidelberg, Germany, he is married, has one daughter, and is a member of the Amateur Radio Relay League and Amnesty International. Hobbies include keeping up with his daughter and their lab retriever puppy, ham-radio operation (AA0VB), and church activities.

**Kevin S. Burke**



Born in Detroit, Michigan, Kevin Burke came to HP in 1989 at the Greeley Hardcopy Division. He attended Colorado State University, receiving a BSCS degree in 1989. Kevin was a lead engineer for HP AccuPage 2.0, contributing to the applica-

tion architecture, small text algorithm, and photo image processing algorithms. Previous accomplishments at HP include work on the autothresholding algorithm for HP AccuPage 1.0, OCR for HP AccuScan, and algorithms for reformatting text read by OCR. His work has resulted in patents related to autothresholding and small text enhancement. He is presently a software engineer with Formonix Inc. in Fort Collins, Colorado. Kevin is married and has three children.

**George Prokop**



George Prokop is an account manager at the Greeley Hardcopy Division. He came to HP in 1985 and has worked as a financial analyst, finance supervisor, sales development marketing engineer, and product support engineer. He is pres-

ently working as a third-party account manager where he coordinates comarketing, solution compatibility, and OEM activities with third-party vendors. He was the account manager for primary OCR vendors wishing to support HP AccuPage 2.0. George attended Purdue University from which he received a BS degree in industrial engineering in 1980 and Indiana University where he earned an MBA in product marketing and corporate finance in 1985. Before coming to HP he was with Rockwell International and Warner Electric Brake and Clutch as an industrial engineer. Born in Chicago, Illinois, he is married and has two daughters.

## 51 11.8-in Flat Panel Display

**David J. Hodge**



Dave Hodge came to HP at the Desktop Computer Division in 1979. He is currently a project manager in the graphics hardware lab at the Workstation Systems Division. As the lead engineer for the 11.8-in flat panel display project he was re-

sponsible for working with the flat panel manufacturer and designing some of the circuits on the HP S1010A interface board. He has also worked on many other HP graphics products including the HP 9845 desktop computer. Born in Boulder, Colorado, he attended the University of Colorado from which he received a BSEE degree in 1975 and is continuing his studies at National Technological University where he is pursuing a masters degree in electrical engineering. His work has resulted in a patent related to the memory mapping scheme used in the HP IGC-20 and HP IGC-10 and several pending patents for work related to the HP S1010A. Dave is a member of the IEEE and has coauthored an HP Journal article about graphics VLSI. Before joining HP, he spent three years working on the design of an electronic milking machine controller. He is married, has a five-year-old daughter and a son on the way, and enjoys model railroading and ham radio (KF0XD).

**Bradly J. Foster**



Born in Beaufort, South Carolina and raised in Cedar Rapids, Iowa, Brad Foster holds a BSEE degree (1988) from Iowa State University and is currently working towards an MSEE at Colorado State University. A member of the technical staff in the

graphics hardware lab at the Workstation Systems Division, he is presently doing graphics hardware ASIC design. He joined HP in 1988 at the Colorado Computer Manufacturing Operation where he held manufacturing responsibilities for HP Series 9000 products. In 1991 Brad joined the graphics hardware lab where he has worked on ASIC design and the design of the 11.8-in flat panel display. He co-designed, prototyped, and debugged the flat panel display breadboard and went on to design the flat panel display frame buffer, frame-buffer control, and display interface. He also designed the frame rate synchronization algorithm, color test-bar pattern circuitry, and production-test equipment to help verify the new display in manufacturing. His efforts related to the HP S1010A have resulted in several pending patents. Before joining HP, Brad worked summers at Rockwell International during his college years. He is married, has one child, and enjoys bicycling, soccer, softball, and basketball.



## Steven J. Komrusch



Steve Komrusch is a member of the technical staff at the Graphics Technology Division where he is responsible for simulation environment design for the next generation of graphics hardware. He joined HP in 1989 after working as a

research assistant for a real-time systems group at the Massachusetts Institute of Technology. He was also a summer hire with Tellabs Incorporated, in Lisle, Illinois working as a systems programmer on digital multiplexing and digital switching telecommunications products. Born in Chicago, Illinois, he attended the University of Illinois at Champaign-Urbana from which he received a BS degree in computer engineering in 1987. He continued his studies at the Massachusetts Institute of Technology, earning an MS in electrical engineering and computer science in 1989. Steve worked on the design of the digital-to-analog converter and the power supply control circuits for the 11.8-in flat panel display. Previous accomplishments at HP include microcode for the VRX and TurboVRX graphics subsystems, printed circuit board layout, timing analysis, and the design of the turn-on FPGA for the CRX48Z graphics subsystem. His work has resulted in pending patents for a computer system module and method for capping volume enclosing polyhedra after sectioning and for a computer system module and method for interference checking of polyhedra using capping polygons. He is interested in computer graphics and the simulation and test of complex Verilog designs. Steve is married and enjoys mountain biking, general relativity, international travel, and learning Japanese.

## Tom J. Searby



A product designer with HP's Systems Technology Division at the Fort Collins Systems Laboratory, Tom Searby is working on product design for the next generation of desktop workstations. His previous HP contributions include product

design on many of the HP 9000 Series 700 workstations and lead mechanical engineering on the flat panel display. He was responsible for the overall flat panel display mechanical concept and the detailed plastic and sheet-metal designs. Tom worked with an external design vendor on the industrial design of the HP S1010A flat panel display. Before joining HP's Workstation Systems Division in 1989, he was a coop student with Rockwell International where he worked on product design for global positioning systems. Tom's special interests include plastic design and ME tools administration. He is currently the ME CAD expert and consultant for the Fort Collins site. Tom was born in Centerville, Iowa, and attended the University of Iowa from which he received a BSME degree in 1989. He is married, has a 4-year-old son with another on the way and enjoys hiking, camping, running, weightlifting, and home improvement.

## 61 Software Buy-versus-Build Decisions

### Wesley H. Higaki



Wes Higaki is a project manager with HP Laboratories and presently leads a team of researchers in the area of test environments at the measurement systems department. He led the team of HP Laboratories engineers helping an HP division develop

the product described in his article, a project that included the development of a new comprehensive economic model for software buy-versus-build decisions. With HP since 1979, his previous accomplishments include software engineering with the HP Laboratories integrated circuits process lab and project management at the Scientific Instruments Division, the Information Networks Division, the Industrial Applications Center, and the Manufacturing Productivity Division. Wes is a member of the IEEE and has authored a previous HP Journal article. He is interested in integrating test information into the manufacturing enterprise. Born in Redwood City, California, he attended the University of California at Davis, receiving a BS degree in mathematics in 1977. He continued his studies at Santa Clara University where he earned an MS in computer science in 1980. He is married, has twin sons and two daughters, and enjoys running.

## 66 ASIC Technology Evaluation

### Antonio A. Martinez



Antonio Martinez is a scientist/engineer with HP Corporate Procurement. He came to HP in 1988 at the Integrated Circuits Business Division and is presently responsible for providing other HP divisions with critical consulting services in

ASIC technology, including supplier selection, technical risk management, and the achievement of long-term technology convergence between HP divisions and supplier capabilities. He recently developed useful benchmark standards for ASIC technology evaluation as reported in this issue. Previous contributions at HP include work on the CMOS26 standard-cell library, the CMOS14 standard-cell library, CMOS14 RAM, and timing modeling. His work has produced a patent for a self-timed dynamic programmable logic array. Born in Rio Piedras, Puerto Rico, he attended the California Institute of Technology, where he simultaneously earned BS and MS degrees in electrical engineering in 1978. He is an IEEE member and is interested in semiconductor circuit design. He has published thirteen papers related to ASIC design, standard-cell libraries, CMOS integrated circuits, timing evaluation, and physical topology level fault detection. Antonio is married, has two children, is a church daycare volunteer, and participates in period reenactment at the Ardenwood historic farm. He enjoys gardening, piano, rock climbing, and sky diving. He is fluent in Spanish and French and has had a lifelong interest in science and mathematics.

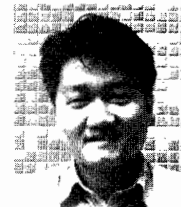
## Aloke S. Bhandia



Aloke Bhandia was born in Calcutta, India, and attended Syracuse University, graduating in 1981 with two MS degrees, one in solid-state science and technology, and another in chemical engineering. He continued his studies at Stanford University, earning an MS in engineering management in 1988. He came to HP in 1984 at the Integrated Circuit Business Division where he performed process integration and device design for high-speed integrated circuits used in HP instruments and microwave products. His contributions at HP also include management of VLSI ASIC technologies for HP's high-performance computers and providing HP RISC workstations with early access to fast-cache SRAM. Aloke is presently a manager with HP Corporate Procurement, where he heads the ASIC/ASSP technology center to facilitate technology convergence with HP suppliers. He led the effort to establish ASIC benchmark technology evaluation standards reported in this issue. Before joining HP, he was with National Semiconductor, where he worked with linear ICs and memory technology development. He is a member of the IEEE, has published six technical papers in the area of semiconductor technology development, and received an HP resident fellowship in 1986-1987 in recognition of outstanding contributions. He is married, enjoys photography and playing with his eight-month old son.

Henry H.W. Lie

### Henry H.W. Lie



A native of Hong Kong, Henry Lie attended the University of Western Ontario at Ontario, Canada, receiving a BSEE degree in 1986. He went on to attend Texas A&M University, earning an MSEE in 1988. He joined HP in 1988 at the Systems Technology Division where he worked as a hardware design engineer for development of a translation lookaside buffer control unit and a chip-level bus converter for HP 9000 technical and commercial multiuser systems. He is presently an ASIC/ASSP project leader with HP Corporate Procurement, providing consulting services to other HP divisions in the area of ASIC/ASSP supplier technology evaluation. His consulting services include technical risk management and the achievement of technology convergence between HP divisions and suppliers. Henry is a member of the IEEE and is interested in CMOS circuit designs, ASIC design methodology, CAD tools development, high-level system designs, and PC system architecture. He is married, is active in his church choir, and serves as an English/Chinese translator at church services. He enjoys basketball and hiking.