Caught In The Ethernet

The venerable networking technology continues to reinvent itself, going from LANs to MANs and WANs.

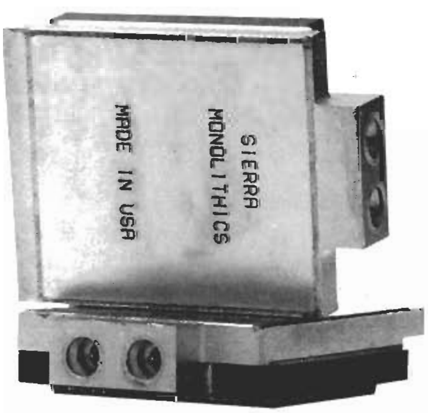
Upon entering its fourth decade, you'd think Ethernet would show signs of age, being pushed aside for newer, state-of-the-art networking technologies. Well, ummm, no.

Rather, Ethernet has evolved in step with changing requirements and standards. In the beginning, it involved PC ports on a coax bus local-area network (LAN). Over time, the technology has morphed into many other forms of networking, from simple I/O ports on embedded controllers to metro-area-network (MAN) and wide-area-network (WAN) implementations including backbone and backhaul—not to mention wireless *(see "Ethernet: A History," p. 35)*.

THE 10-GBIT/S ROLLOUT

The 10-Gbit/s version of Ethernet has been around for more than five years in several forms. But it has taken time for prices to moderate and companies to adopt these higher-speed versions. Usually, 10/100 Ethernet works for most small and medium businesses, and 1 Gbit/s is fast enough for an aggregation backbone.

Yet with ever-larger networks, more wireless extensions, and larger data flows, the advent of 1 Gbit/s to the desktop and 10-Gbit/s backbones has become necessary. With server virtualization, the increased growth of data centers and server farms with blade servers, the growing number of Internet searches, video downloads, and streaming video, the need for 10-Gbit/s networks is critical.



1. Sierra Monolithics' SMI4027 multiplexer/clock multiplier unit and the SMI4037 clock and data-recovery/demuxiplexer modules target 40-Gbit/s optical transport systems.

Today, 10-Gbit/s equipment is more common and becoming the main link between servers in data centers. While most 10-Gbit/s LANs are fiber, the newer copper unshielded twisted-pair (UTP) version is slowly being adopted in data centers and other short-distance (<30 to 100 m) applications.

The 10-Gbit/s copper version (10GBaseT) is an interesting technical case because it virtually defies the laws of physics. To achieve that seemingly impossible rate, the standard calls for 16 levels of pulse-amplitude modulation (PAM) simultaneously occurring at an 800-Msymbol rate over each of the four pairs in a special low-loss CAT6a UTP. A low-density parity check (LDPC) coding scheme helps overcome data errors due to cable losses and noise.

On top of that, all sorts of fancy equalization, echo, and crosstalk cancellation for near-end crosstalk (NEXT), far-end crosstalk (FEXT), and alien (adjacent cables) crosstalk are implemented with DSP. Lots of circuitry (i.e., 800-MHz 10-bit analog-to-digital converters, DSP logic) is used to achieve these connections, making most implementations power hogs. Several companies have brought the power consumption per port down to 3 to 5 W, which is barely acceptable.

Yet 10GBaseT is still significantly lower cost per port than fiber versions. While 10-Gbit/s copper network interface cards (NICs) are now available, 10GBaseT has never really taken off, primarily because of the rivalry and squabbling between the competing companies over intellectual property (IP) and other issues. This small group of companies has never come together to build a market despite the fact it did actually agree on a standard (802.3an).

2. The RAD ETX-202A Carrier Ethernet Demarcation device, which is owned and operated by the service provider and installed at the customer site, has two Ethernet network ports and up to four Ethernet subscriber ports that use either copper or SFP-based interfaces.

A variety of Ethernet optical modules is now available in X2, XFP, SFP, and other multisource agreement (MSA) formats. With the new SFP+ optical transceiver modules for 10 Gbits/s now priced competitively, there may be fewer reasons to keep the copper version. The market will no doubt let us know.

THE 40/100-GBIT/S STANDARD

With Ethernet steadily increasing in speed, there have been debates on its use as a network backbone and in core MANs and WANs. It's highly desirable due to its fully standardized nature, which includes full interoperability and backward-compatibility between the components of many different vendors. Its low cost is also attractive, as is the fact that it's well-known among IT and network personnel. The forthcoming 40- and 100-Gbit/s versions will make Ethernet a real contender in the MAN and WAN space.

Right now, Sonet/SDH is still the leading technology in these core, edge, and access networks. Sonet/SDH networks of the OC-48 (2.5 Gbits/s) and OC-192 (10 Gbits/s) varieties see wide usage. Also, Sonet/SDH defines standards to OC-768 or 40 Gbits/s. The 40-Gbit/s versions aren't common, but are now getting more play. Holding back their widespread adoption is the high price.

Recent Significant Ethernet Standards

IEEE standard	Ratification date	Explanation
802.3ac	1998	Frame size increased to 1522 bytes to allow tagging for virtual LANs and establishing data priority
802.3ae	2003	10-Gbit/s standard; multiple versions over fiber
802.3af	2003	Power over Ethernet (PoE); dc power distribution over the UTP medium to power wireless access points up to 15 W
802.3an	2006	10 Gbits/s over UTP
802.3ap	2008	1 and 10 Gbits/s over printed-circuit-board backplanes
802.3at/au	2009	PoE enhancements and isolation; power boost to the 30- to 45-W range
802.3av	2009	10-Gbit/s Ethernet passive optical network (EPON)
802.3ba	2010	40/100 Gbits/s over fiber

Thanks to developments in 10-Gbit/s Ethernet, semiconductors, and optical processing, data rates are reaching 100 Gbits/s. In 2006, the IEEE established the High Speed Study Group (HSSG) to investigate the development of a 100-Gbit/s standard. Then in 2007, the group agreed to add a 40-Gbit/s version as an interim technology for data centers and other markets.

An organization called the Road to 100G Alliance was formed in 2008 to study the problem and make recommendations. That Alliance was rolled into the Ethernet Alliance in late 2008. In early 2008, the IEEE established a Task Force to create the new 40/100-Gbit/s standard designated 802.3ba. A final standard isn't expected until 2010.

Additionally, the International Telecommunications Union-Telecommunications (ITU-T) has a 100-Gbit/s standard study group that's looking into this technology (designated as Recommendation G.709). Overall, several versions are possible, and most of them involve parallel data paths on fiber-optic cable.

The 40-Gbit/s standard would simply put four 10-Gbit/s data streams on four parallel fibers. One 100-Gbit/s version would require eight or ten 10-Gbit/s data paths on parallel fibers. These two versions would be used primarily in short runs, for example, in data centers. For longer runs, coarse or dense wavelength-division multiplexing (DWM) using four or eight/ten wavelengths (λ) of light on a single fiber is an option.

Whatever the technique, the industry's ultimate goal is a 100-Gbit/s rate for up to 40 km (25 miles) and up to 10 km (6 miles) on single-mode fiber (SMF), as well as up to 100 m (328 ft) on optical multimode 3 (OM3) multimode fiber (MMF). Possibilities include the use of SNAP-4 or quad small form-factor pluggable (QSFP) cables and connectors for 40 Gbits/s, as well as SMAP-12 cables and connectors for 100 Gbits/s.

The Optical Internetworking Forum (OIF) is working with the IEEE group on a version that would transmit 25 to 28 Gbits/s over four parallel paths. A single serial version over SMF isn't expected because of the attenuation and dispersion problems at such high data rates. A copper version with a reach of 5 to 10 m or a backplane version may be included in the standard.

A variety of modulation schemes such as differential phase-shift keying (DPSK) and differential quadrature phase-shift keying (DQPSK), dispersion compensation, and forward-error-correction (FEC) coding methods is being considered. Further goals include the support of full duplex operation and maintaining the frame format and media-access-controller (MAC) interface, as well as having a bit error rate (BER) of better than 10^{-12} at the MAC interface.

Designers can use the Sierra Monolithic SMI4027 multiplexer/clock multiplier unit (MUX/CMU) to implement 40-Gbit/s systems *(Fig. 1)*. It and its companion SMI4037 clock and data recovery unit/demultiplexer (CDR/DEMUX) implement the physical layer (PHY), operate from 39.8 to 44.6 Gbits/s, and incorporate a fully SFI-5 client-side interface. Furthermore, they are fully usable with Sonet/SDH or with Ethernet.

Ethernet Prepositions

Ethernet, still going strong, after 36 years, can be associated with these varied prepositions: up, into, across, over, and down.

It's going *up* in local-area-network (LAN) speed to 40 and 100 Gbits/s. Recall that we started Ethernet in 1973 at 2.94 Mbits/s.

It's going *into* the core of our long-haul networks, replacing Sonet at 40 Gbits/s today over long-haul dense wavelength-division multiplexing (DWDM). Terabit Ethernet is on the horizon.

It's going *across* the carrier-induced "telechasm," replacing T1 and other old gear, with burgeoning Carrier Ethernet services.

It's going *over* the airwaves as Wi-Fi and WiMAX.

And, it's going *down* in forms, including ZigBee, the wireless networking of an increasing fraction of each year's 10 billion new embedded microcontrollers.

ED ONLINE 20901

BOB METCALFE of Polaris Venture Partners received the National Medal of Technology in 2005 for his leadership in the invention, standardization, and commercialization of Ethernet, which added 350 million new switch ports last year, not counting Wi-Fi, according to IDC.

CARRIER ETHERNET

Ethernet has always been a "best effort" networking service, meaning not all packets may be received. Packets could be dropped because of noise, collisions, packet payload violations (too long, too short), or other reasons. Its non-deterministic characteristics caused by latencies and variable delays also diminish its usefulness in some applications.

These flaws make Ethernet questionable for serious data services such as financial and healthcare transactions, video streaming, and other mission-critical applications where quality of service (QoS) is mandatory. Sonet/SDH is a guaranteed service, and that's the reason for its continued use and development despite costs and complexity. Now an alternative known as Carrier Ethernet is turning Ethernet into a service with similar qualities.

Created by the Metro Ethernet Forum (MEF) and its members, Carrier Ethernet defines a set of standards and implementation agreements for a carrier-class service and network that distinguishes it from LAN Ethernet. Also, it defines five basic attributes that make Ethernet function at the carrier-grade level in the access, metro, and core networks:

- Standard services such as E-Line, E-LAN, and transparent private line, virtual private line, and LAN service; it's suitable for converged voice, video, and data networks with choice and granularity of bandwidth and QoS options
- Scalability across access, metro, and wider area networks
- Reliability, implying the ability to detect and recover from incidents without impacting users
- QoS with service level agreements (SLAs)
- Service management, meaning the ability to monitor, diagnose, and centrally manage the network; carrier-class operation, administration, and management (OAM); and rapid provisioning

The MEF provides testing procedures and certification to ensure that the standards are met and interoperability is achieved. Carrier Ethernet is implemented by equipment like RAD Data Communications' demarcation device *(Fig. 2)*.

STORAGE-AREA NETWORKS

Storing large quantities of data is common in the enterprise, government, and other organizations. RAID (redundant array of independent disks) and JBOD (just a bunch of disks) storage systems are a growing market, and the need to connect them to LANs and the Internet is critical. That's why special storage-area networks (SANs) were developed.

Ethernet: A History

Robert Metcalfe and his associates invented Ethernet in 1972 *(see "Ethernet Prepositions," p. 34)*. The original Ethernet was a coax bus topology with a bit rate of 2.94 Mbits/s. During the next 10 years, it evolved into a 10-Mbit/s coax bus, and a frame format and protocol emerged.

In 1983, the IEEE standardized this 10-Mbit/s version as 802.3 and designated it as 10Base5. It defined the physical-layer (PHY) and media-access-control (MAC) layers of the network. The medium was RG-8/U coax, generally called thicknet. In 1985, a version using RG-58/U coax (thinnet) was defined.

During the early years of its use as a local-area network (LAN), Ethernet competed with several other technologies, most prominently ARCNET, a token coax bus, and IBM's Token Ring of twisted pair. But Ethernet's proponents kept making improvements and additions that sustained its popularity for a wide range of networking applications.

The big change came with the 1987 version, 802.3j, which specified low-cost unshielded twisted pair (UTP) as the medium. The designation was 10BaseT. Unshielded twisted pair (UTP) with its RJ-45 connectors was cheaper and easier to work with than coax, so 10BaseT soon became the LAN of choice, gradually edging out the competition.

Another big breakthrough came in 1995 with the 802.3u standard, which defined three versions of a 100-Mbit/s data-rate technology. The 100BaseTX version soon became the mainstay of enterprise LANs everywhere. It used CAT5 UTP, and this so-called 10/100 version of Ethernet is still the anchor of the technology.

But Ethernet wasn't through yet. Faster, smaller semiconductors and other technologies soon made it possible to go way beyond 10/100. A 1-Gbit/s version was developed and standardized in 1998 as 802.3z. This fiber version was supplemented by a UTP version in 1999 as 802.3ab and dubbed 1000BaseT.

Since then, the technology has improved at a furious pace. The latest versions run at 10 Gbits/s. There are multiple fiber and UTP versions as well as coax and backplane versions. Other developments have kept Ethernet at the head of the pack, too. Currently, it's without a doubt the only remaining LAN technology in use. And due to its high speed and other improvements, its use is rapidly extending into the metropolitan and wide-area networks that were once implemented with ATM over T1 and DS3 lines as well as Sonet/SDH fiber networks.

Ethernet uses an access method called carrier-sense multiple-access/collision detection (CSMA/CD). Bus nodes use this arbitration method to access the network. Each node listens to the bus and then transmits if no one is using the bus. If another carrier (transmission) is sensed, the node waits until the bus is free. Collisions occur when two or more nodes try to access the bus simultaneously. The nodes then back off and retransmit at random intervals.

PRE	SFD	DA	SA	Length type	Data payload	FCS
7	1	6	6	2	←46 to 1522→	4

Field length in bytes

PRE = preamble (alternating zeros and ones)
SFD = start-of-frame delimiter (alternating ones and zeros ending in 11)
DA = destination address
SA = source address
Length type = number of bytes in data payload
FCS = frame check sequence (CRC)

The packet frame format for Ethernet is preceded by a sync bit stream (preamble and start of frame delimiter) of 64 bits and then begins with source and destination addressed, each 6 bytes long. This is followed by a 2-byte sequence designating the number of bytes in the data payload. The data payload follows and can be any byte length from 46 minimum to 1522 bytes maximum. A 32-bit cyclic redundancy code (CRC) for error detection concludes the packet.

This means the bus, being a shared medium, causes the maximum data rate to be divided or shared depending on how many users are trying to access the bus. The throughput is lower by that factor. Data is transmitted in frames like those shown in the figure.

The table summarizes some of the latest standards of interest *(see "Recent Significant Ethernet Standards," p. 34)*. The most exciting include the next speed increments to 40 and ultimately 100 Gbits/s.

ED ONLINE 20900

The most widely used SAN, Fibre Channel (FC), is an optical-fiber network that connects the hard-disk arrays to the servers and the network through host bus adapter (HBA) cards. The medium is fiber-optical cable, and the T11 organization has set standards for data rates of 1, 2, 4, and 8 Gbits/s. Also, a 10-Gbit/s standard has been defined.

Though FC works well and is widely used, it's expensive to expand and maintain. Nonetheless, most users want to continue it in some way even though lower-cost alternatives like the Internet Small Computer Systems Interface (iSCSI) are out there.

The iSCSI option is a serial version of the widely used older SCSI parallel connections and protocol used with hard drives. It uses Ethernet as the networking medium to tunnel SCSI commands and data. As a result, it costs much less and can run in parallel with existing Ethernet networks. But, except for small and medium businesses, iSCSI's adoption has been limited.

A forthcoming alternative that combines the best of FC and Ethernet is the Fibre Channel over Ethernet (FCoE) standard being developed by the T11 organization. It packages FC packets and tunnels them over standard Ethernet LANs.

BUT THAT'S NOT ALL...

Many other enhancements, additions, and applications surround Ethernet, such as industrial Ethernet, the LXI instrumentation system, and Ethernet passive optical networks (EPONs).

Industrial Ethernet uses standard Ethernet in place of other special networks developed for the harsh and critical nature of industrial applications. These networks have to operate with high levels of noise and in extreme environments of temperature variations and corrosive atmospheres. Special shielded and protected cables and connectors help in these cases.

In addition, many industrial applications require determinism. That is, they must be able to accurately time data transmissions and related control and measurement actions.

Ethernet is fraught with latencies, delays, and timing inaccuracies common to a best effort service. One solution is to apply the IEEE's 1588 Precision Time Protocol (PTP). It provides accurate synchronization of nodes on a network by using hardware-generated time stamps. Precision in the nanosecond range is possible. Employing 1588 PTP on Ethernet suits this combination for even the most time-critical industrial applications. The IEEE 1588 PTP is also used in Carrier Ethernet applications.

LXI (LAN eXtensions for Instrumentation) is the relatively new instrumentation communications standard designed to replace the general-purpose instrumentation bus (GPIB) so widely used in test systems over the years. It connects test instruments, PCs, and the Internet for any combination of testing, measuring, recording, storing, and accessing data. Many instruments already incorporate LXI, and more are forthcoming. LXI is based on the use of Ethernet because of its low cost and wide availability.

Finally, work is underway on a 10-Gbit/s version of Ethernet designed for PONs in metro networks that deliver digital TV and Internet access services. PONs use low-cost fiber to distribute video and other services to homes and businesses without the need for costly repeaters or other active intermediary hardware in the field. PONs are widely deployed in cable TV and telecom infrastructures to deliver higher speeds and quality video.

The original PON standards, which were set by the ITU-T, are designated as APON, BPON, and the latest, GPON. There's also a low-speed Ethernet PON known as EPON or GEPON for gigabit service. Designated 802.3ah, it is widely used in Asia but not in the U.S., where GPON is the dominant standard. The IEEE is currently working on a 10-Gbit/s version designated 802.3av that will utilize different wavelengths of light for 1- and 10-Gbit/s services on a single fiber. A final version of 802.3av is expected later in 2009.

For More Information

Information on the various IEEE standards groups can be found online at *grouper.ieee.org/groups/802/3/*. Though the site isn't in a friendly format, it's a good source of information. Published documents on works in process are hard to access and interpret, but it's still worth a look. The Ethernet Alliance *(www.ethernetalliance.com)* was started in 2005 to help foster the development of Ethernet and disseminate relevant information. It's a great source of news, status, and other information. You should read it daily. Brad Booth, EA chairman, was a major contributor to this article. Finally, the Metro Ethernet Forum *(www.metroethernetforum.org)* is the ultimate goldmine source of Carrier Ethernet information.

ED ONLINE 20899