Gigabit Ethernet: From 1 to 100 Gbps and Beyond

by William Stallings

Since its first introduction in the early 1980s, Ethernet has been the dominant technology for implementing *Local-Area Networks* (LANs) in office environments. Over the years, the data-rate demands placed on LANs have grown at a rapid pace. Fortunately, Ethernet technology has adapted to provide ever-higher capacity to meet these needs. We are now in the era of the Gigabit Ethernet.

Ethernet began as an experimental bus-based 2.94-Mbps system[1] using coaxial cable. In a shared bus system, all stations attach to a common cable, with only one station able to successfully transmit at a time. A *Medium Access Control* (MAC) protocol based on collision detection arbitrates the use of the bus. In essence, each station is free to transmit MAC frames on the bus. If a station detects a collision during transmission, it backs off a certain amount of time and tries again.

The first commercially available Ethernet products were bus-based systems operating at 10 Mbps[2]. This introduction coincided with the standardization of Ethernet by the IEEE 802.3 committee. With no change to the MAC protocol or MAC frame format, Ethernet could also be configured in a star topology, with traffic going through a central hub, again with transmission limited to a single station at a time through the hub. To enable an increase in the data rate, a switch replaces the hub, allowing full-duplex operation. With the switch, the same MAC format and protocol are used, although collision detection is no longer needed. As the demand has evolved and the data rate requirement increased, some enhancements to the MAC layer have been added, such as provision for larger frame sizes.

Currently, Ethernet systems are available at speeds up to 100 Gbps. Table 1 summarizes the successive generations of IEEE 802.3 standardization.

Table 1: IEEE 802.3 Physical Layer Standards

Year Introduced	1983	1995	1998	2003	2010
Maximum data transfer speed	10 Mbps	100 Mbps	1 Gbps	10 Gbps	40 Gbps, 100 Gbps
Transmission media	Coax cable, unshielded twisted pair, optical fiber	Unshielded twisted pair, shielded twisted pair, optical fiber	Unshielded twisted pair, optical fiber, shielded twisted pair	Optical fiber	Optical fiber, backplane

Ethernet quickly achieved widespread acceptance and soon became the dominant technology for LANs. Its dominance has since spread to *Metropolitan-Area Networks* (MANs) and a wide range of applications and environments.

The huge success of Ethernet is due to its extraordinary adaptability. The same MAC protocol and frame format are used at all data rates. The differences are at the physical layer, in the definition of signaling technique and transmission medium.

In the remainder of this article, we look at characteristics of Ethernet in the gigabit range.

1-Gbps Ethernet

For many years the initial standard of Ethernet, at 10 Mbps, was adequate for most office environments. By the early 1990s, it was clear that higher data rates were needed to support the growing traffic load on the typical LAN. Key drivers in this evolution include:

- *Centralized Server Farms:* In many multimedia applications, there is a need for the client system to be able to draw huge amounts of data from multiple, centralized servers, called *Server Farms*. As the performance of the servers has increased, the network has become the bottleneck.
- *Power Workgroups:* These groups typically consist of a small number of cooperating users who need to exchange massive data files across the network. Example applications are software development and computer-aided design.
- *High-speed Local Backbone:* As processing demand grows, enterprises develop a configuration of multiple LANs interconnected with a high-speed backbone network.

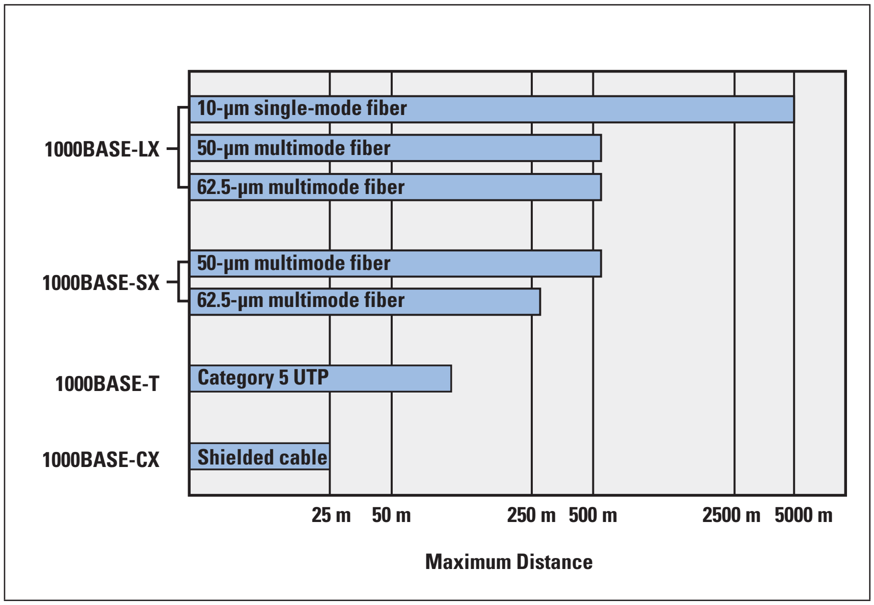
To meet such needs, the IEEE 802.3 committee developed a set of specifications for Ethernet at 100 Mbps, followed a few years later by a 1-Gbps family of standards. In each case, the new specifications defined transmission media and transmission encoding schemes built on the basic Ethernet framework, making the transition easier than if a completely new specification were issued.

The 1-Gbps standard includes a variety of transmission medium options[3, 4] (Figure 1):

- *+1000BASE-SX:* This short-wavelength option supports duplex links of up to 275 m using 62.5-µm multimode or up to 550 m using 50-µm multimode fiber. Wavelengths are in the range of 770 to 860 nm.
- *1000BASE-LX:* This long-wavelength option supports duplex links of up to 550 m of 62.5-µm or 50-µm multimode fiber or 5 km of 10-µm single-mode fiber. Wavelengths are in the range of 1270 to 1355 nm.

- *1000BASE-CX:* This option supports 1-Gbps links among devices located within a single room or equipment rack, using copper jumpers (specialized shielded twisted-pair cable that spans no more than 25 m). Each link is composed of a separate shielded twisted pair running in each direction.
- *1000BASE-T:* This option uses four pairs of Category 5 unshielded twisted pair to support devices over a range of up to 100 m, transmitting and receiving on all four pairs at the same time, with echo cancellation circuitry.

Figure 1: 1-Gbps Ethernet Medium Options (log scale)



The signal encoding scheme used for the first three Gigabit Ethernet options just listed is 8B/10B. With 8B/10B, each 8 bits of data is converted into 10 bits for transmission[3]. The extra bits serve two purposes. First, the resulting signal stream has more transitions between logical 1 and 0 than an uncoded stream; it avoids the possibility of long strings of 1s or 0s that make synchronization between transmitter and receiver more difficult. Second, the code is designed in such a way as to provide a useful error-detection capability.

The signal-encoding scheme used for 1000BASE-T is 4D-PAM5, a complex scheme whose description is beyond our scope.

In a typical application of Gigabit Ethernet, a 1-Gbps LAN switch provides backbone connectivity for central servers and high-speed workgroup Ethernet switches. Each workgroup LAN switch supports both 1-Gbps links, to connect to the backbone LAN switch and to support high-performance workgroup servers, and 100-Mbps links, to support high-performance workstations, servers, and 100-Mbps LAN switches.

10-Gbps Ethernet

Even as the ink was drying on the 1-Gbps specification, the continuing increase in local traffic made this specification inadequate for needs in the short-term future. Accordingly, the IEEE 802.3 committee soon issued a standard for 10-Gbps Ethernet. The principal requirement for 10-Gbps Ethernet was the increase in intranet (local interconnected networks) and Internet traffic. Numerous factors contribute to the explosive growth in both Internet and intranet traffic:

- An increase in the number of network connections
- An increase in the connection speed of each end station (for example, 10-Mbps users moving to 100 Mbps, and analog 56k users moving to DSL and cable modems)
- An increase in the deployment of bandwidth-intensive applications such as high-quality video
- An increase in Web hosting and application hosting traffic

Initially, network managers used 10-Gbps Ethernet to provide high-speed, local backbone interconnection between large-capacity switches. As the demand for bandwidth increased, 10-Gbps Ethernet began to be deployed throughout the entire network, to include server farm, backbone, and campus-wide connectivity. This technology enables *Internet Service Providers* (ISPs) and *Network Service Providers* (NSPs) to create very-high-speed links at a very low cost, between co-located, carrier-class switches, and routers[5].

The technology also allows the construction of MANs and *Wide-Area Networks* (WANs) that connect geographically dispersed LANs between campuses or *Points of Presence* (PoPs). Thus, Ethernet begins to compete with *Asynchronous Transfer Mode* (ATM) and other wide-area transmission and networking technologies. In most cases where the customer requirement is data and TCP/IP transport, 10-Gbps Ethernet provides substantial value over ATM transport for both network end users and service providers:

- No expensive, bandwidth-consuming conversion between Ethernet packets and ATM cells is required; the network is Ethernet, end to end.
- The combination of IP and Ethernet offers *Quality of Service* (QoS) and traffic policing capabilities that approach those provided by ATM, so that advanced traffic-engineering technologies are available to users and providers.
- A wide variety of standard optical interfaces (wavelengths and link distances) have been specified for 10 Gigabit Ethernet, optimizing its operation and cost for LAN, MAN, or WAN applications.

The goal for maximum link distances cover a range of applications is from 300 m to 40 km. The links operate in full-duplex mode only, using a variety of optical fiber physical media.

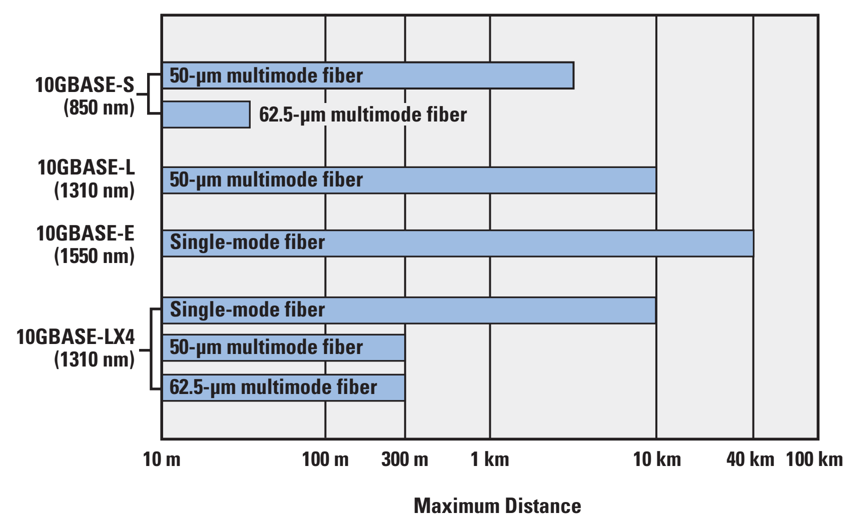


Figure 2: 10-Gbps Ethernet Distance Options (log scale)

Four physical layer options are defined for 10-Gbps Ethernet (Figure 2). The first three have two suboptions: an "R" suboption and a "W" suboption. The R designation refers to a family of physical layer implementations that use a signal encoding technique known as 64B/66B. With 64B/66B, each 64 bits of data is converted into 66 bits for transmission, resulting in substantially less overhead then 8B/10B but providing the same types of benefits. The R implementations are designed for use over *dark fiber* meaning a fiber-optic cable that is not in use and that is not connected to any other equipment. The W designation refers to a family of physical layer implementations that also use 64B/66B signaling but are then encapsulated to connect to SONET equipment.

The four physical layer options are as follows:

- *10GBASE-S (short):* Designed for 850-nm transmission on multimode fiber. This medium can achieve distances up to 300 m. There are 10GBASE-SR and 10GBASE-SW versions.
- *10GBASE-L (long):* Designed for 1310-nm transmission on single-mode fiber. This medium can achieve distances up to 10 km. There are 10GBASE-LR and 10GBASE-LW versions.
- *10GBASE-E (extended):* Designed for 1550-nm transmission on single-mode fiber. This medium can achieve distances up to 40 km. There are 10GBASE-ER and 10GBASE-EW versions.
- *10GBASE-LX4:* Designed for 1310-nm transmission on single-mode or multimode fiber. This medium can achieve distances up to 10 km. This medium uses *Wavelength-Division Multiplexing* (WDM) to multiplex the bit stream across four light waves.

40-/100-Gbps Ethernet

Ethernet is widely deployed and is the preferred technology for wired local-area networking. Ethernet dominates enterprise LANs, broadband access, and data center networking, and it has also become popular for communication across MANs and even WANs. Further, it is now the preferred carrier wire-line vehicle for bridging wireless technologies, such as Wi-Fi and WiMAX, into local Ethernet networks.

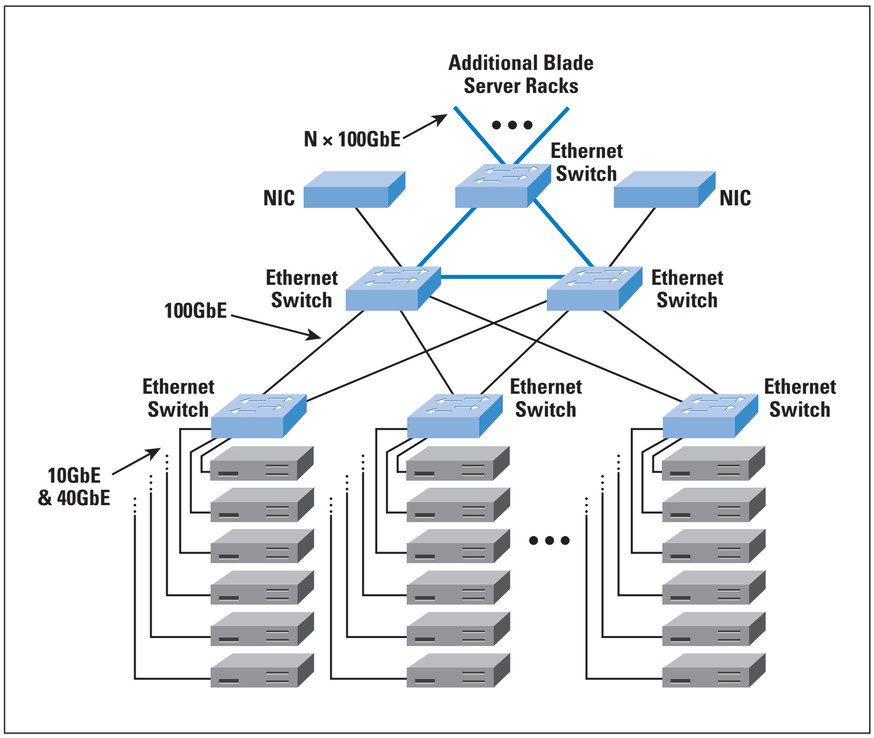
This popularity of Ethernet technology is due to the availability of cost-effective, reliable, and interoperable networking products from a variety of vendors. The development of converged and unified communications, the evolution of massive server farms, and the continuing expansion of *Voice over IP* (VoIP), *Television over IP* (TVoIP), and Web 2.0 applications have accelerated the need for ever-faster Ethernet switches. The following are market drivers for 100-Gbps Ethernet:

- *Data center/Internet Media Providers:* To support the growth of Internet multimedia content and Web applications, content providers have been expanding data centers, pushing 10-Gbps Ethernet to its limits. These providers are likely to be high-volume early adopters of 100-Gbps Ethernet.
- *Metro-Video/Service Providers:* Video on demand has been leading a new generation of 10-Gbps Ethernet metropolitan/core network build-outs. These providers are likely to be high-volume adopters in the medium term.
- *Enterprise LANs:* Continuing growth in convergence of voice/video/data and in unified communications is accelerating network switch demands. However, most enterprises still rely on 1-Gbps or a mix of 1-Gbps and 10-Gbps Ethernet, and adoption of 100-Gbps Ethernet is likely to be slow.
- *Internet exchanges/ISP Core Routing:* With the massive amount of traffic flowing through these nodes, these installations are likely to be early adopters of 100-Gbps Ethernet.

In 2007, the IEEE 802.3 working group authorized the *IEEE P802.3ba 40-Gbps and 100-Gbps Ethernet Task Force*. The 802.3ba project authorization request cited numerous examples of applications that require greater data-rate capacity than 10-Gbps Ethernet offers, including Internet exchanges, high-performance computing, and video-on-demand delivery. The authorization request justified the need for two different data rates in the new standard (40 Gbps and 100 Gbps) by recognizing that aggregate network requirements and end-station requirements are increasing at different rates.

An example of the application of 100-Gbps Ethernet is shown in Figure 3. The trend at large data centers, with substantial banks of blade servers, is the deployment of 10-Gbps ports on individual servers to handle the massive multimedia traffic provided by these servers. Typically, a single blade-server rack will contain multiple servers and one or two 10-Gbps Ethernet switches to interconnect all the servers and provide connectivity to the rest of the facility. The switches are often mounted in the rack and referred to as *Top-of-Rack* (ToR) switches. The term ToR has become synonymous with a server access switch, even if it is not located "top of rack." For very large data centers, such as cloud providers, the interconnection of multiple blade-server racks with additional 10-Gbps switches is increasingly inadequate. To handle the increased traffic load, switches operating at greater than 10 Gbps are needed to support the interconnection of server racks and to provide adequate capacity for connecting off-site through *Network Interface Controllers* (NICs).

Figure 3: Example 100-Gbps Ethernet Configuration for Massive Blade-Server Cloud Site



The first products in this category appeared in 2009, and the IEEE 802.3ba standard was finalized in 2010. Initially, many enterprises are deploying 40-Gbps switches, but both 40- and 100-Gbps switches are projected to enjoy increased market penetration in the next few years[6, 7, 8].

IEEE 802.3ba specifies three types of transmission media as shown in Table 2: copper backplane, twin axial (a type of cable similar to coaxial cable), and optical fiber. For copper media, four separate physical lanes are specified. For optical fiber, either 4 or 10 wavelength lanes are specified, depending on data rate and distance[9, 10, 11].

Table 2: Media Options for 40- and 100-Gbps Ethernet

	40 Gbps	100 Gbps
1-m backplane	40GBASE-KR4	
10-m copper	40GBASE-CR4	1000GBASE-CR10
100-m multimode fiber	40GBASE-SR4	1000GBASE-SR10
10-km single-mode fiber	40GBASE-LR4	1000GBASE-LR4
40-km single-mode fiber		1000GBASE-ER4

Naming nomenclature:
Copper: K = Backplane; C = Cable assembly
Optical: S = Short reach (100 m); L - Long reach (10 km);
E = Extended long reach (40 km)
Coding scheme: R = 64B/66B block coding
Final number: Number of lanes (copper wires or fiber wavelengths)

Multilane Distribution

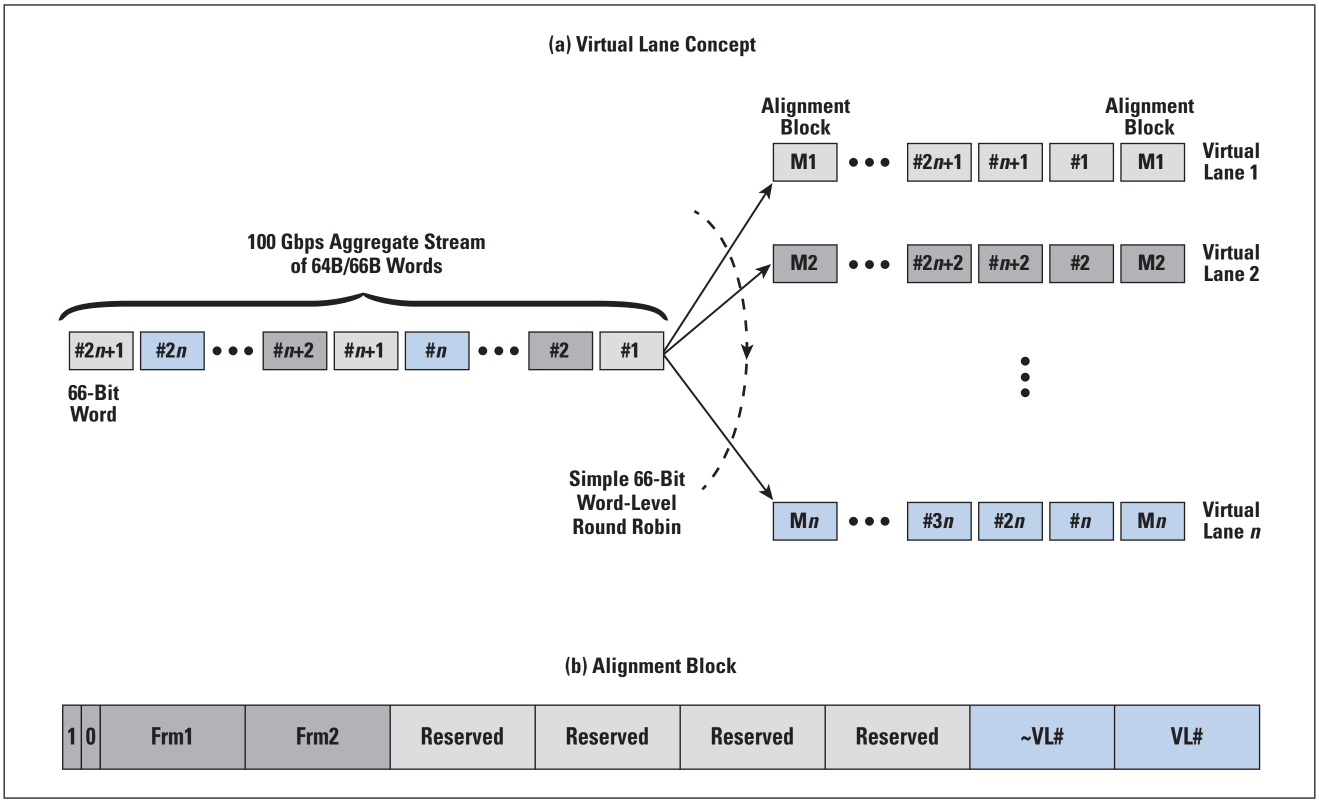
The 802.3ba standard uses a technique known as multilane distribution to achieve the required data rates. Two separate concepts need to be addressed: multilane distribution and virtual lanes.

The general idea of *multilane distribution* is that, in order to accommodate the very high data rates of 40 and 100 Gbps, the physical link between an end station and an Ethernet switch or the physical link between two switches may be implemented as multiple parallel channels. These parallel channels could be separate physical wires, such as four parallel twisted-pair links between nodes. Alternatively, the parallel channels could be separate frequency channels, such as provided by WDM over a single optical fiber link.

For simplicity and manufacturing ease, we would like to specify a specific multiple-lane structure in the electrical physical sublayer of the device, known as the *Physical Medium Attachment* (PMA) sublayer. The lanes produced are referred to as *virtual lanes*. If a different number of lanes are actually in use in the electrical or optical link, then the virtual lanes are distributed into the appropriate number of physical lanes in the *Physical Medium Dependent* (PMD) sublayer. This is a form of inverse multiplexing.

Figure 4a shows the virtual lane scheme at the transmitter. The user data stream is encoded using the 64B/66B, which is also used in 10-Gbps Ethernet. Data is distributed to the virtual lanes one 66-bit word at a time using a simple round robin scheme (first word to first lane, second word to second lane, etc.). A unique 66-bit alignment block is added to each virtual lane periodically. The alignment blocks are used to identify and reorder the virtual lanes and thus reconstruct the aggregate data stream.

Figure 4: Multilane Distribution for 100-Gbps Ethernet



The virtual lanes are then transmitted over physical lanes. If the number of physical lanes is smaller then the number of virtual lanes, then bit-level multiplexing is used to transmit the virtual lane traffic. The number of virtual lanes must be an integer multiple (1 or more) of the number of physical lanes.

Figure 4b shows the format of the alignment block. The block consists of 8 single-byte fields preceded by the 2-bit synchronization field, which has the value 10. The Frm fields contain a fixed framing pattern common to all virtual lanes and used by the receiver to locate the alignment blocks. The VL# fields contain a pattern unique to the virtual lane: one of the fields is the binary inverse of the other.

25-/50-Gbps Ethernet

One of the options for implementing 100 Gbps is as four 25-Gbps physical lanes. Thus, it would be relatively easy to develop standards for 25- and 50-Gbps Ethernet, using one or two lanes, respectively. Having these two lower-speed alternatives, based on the 100-Gbps technology, would give users more flexibility in meeting existing and near-term demands with a solution that would scale easily to higher data rates.

Such considerations have led to the formation of the *25 Gigabit Ethernet Consortium* by numerous leading cloud networking providers, including Google and Microsoft. The objective of the consortium is to support an industry-standard, interoperable Ethernet specification that boosts the performance and slashes the interconnect cost per Gbps between the NIC and ToR switch. The specification adopted by the consortium prescribes a single-lane 25-Gbps Ethernet and dual-lane 50-Gbps Ethernet link protocol, enabling up to 2.5 times higher performance per physical lane on twinax copper wire between the rack endpoint and switch compared to current 10- and 40-Gbps Ethernet links. The IEEE 802.3 committee is presently developing the needed standards for 25 Gbps, and it may include 50 Gbps[12, 13].

It is too early to say how these various options (25, 40, 50, and 100 Gbps) will play out in the marketplace. In the intermediate term, the 100-Gbps switch is likely to predominate at large sites, but the availability of these slower and cheaper alternatives gives enterprises numerous paths for scaling up to meet increasing demand.

400-Gbps Ethernet

The growth in demand never lets up. IEEE 802.3 is currently exploring technology options for producing a 400-Gbps Ethernet standard, although no timetable is yet in place[14, 15, 16, 17]. Looking beyond that milestone, there is widespread acknowledgment that a 1-Tbps (terabit per second, trillion bits per second) standard will eventually be produced[18].

2.5-/5-Gbps Ethernet

As a testament to the versatility and ubiquity of Ethernet, and at the same time the fact that ever-higher data rates are being standardized, consensus is developing to standardize two lower rates: 2.5 and 5 Gbps[19, 20]. These relatively low speeds are also known as *Multirate Gigabit BASE-T* (MGBASE-T). Currently, the *MGBASE-T Alliance* is overseeing the development of these standards outside of IEEE. It is likely that the IEEE 802.3 committee will ultimately issue standards based on these industry efforts.

These new data rates are intended mainly to support IEEE 802.11ac wireless traffic into a wired network. IEEE 802.11ac is a 3.2-Gbps Wi-Fi standard that is gaining acceptance where more than 1 Gbps of throughput is needed, such as to support mobile users in the office environment[21]. This new wireless standard overruns 1-Gbps Ethernet link support but may not require the next step up, which is 10 Gbps. Assuming that 2.5 and 5 Gbps can be made to work over the same cable that supports 1 Gbps, then this standard would provide a much-needed uplink speed improvement for access points supporting 802.11ac radios with their high-bandwidth capabilities.

Conclusion

Ethernet is widely deployed and is the preferred technology for wired local-area networking. Ethernet dominates enterprise LANs, broadband access, and data center networking, and it has also become popular for communication across MANs and even WANs. Further, it is now the preferred carrier wire-line vehicle for bridging wireless technologies, such as Wi-Fi and WiMAX, into local Ethernet networks. Further, the Ethernet marketplace is now large enough to accelerate the development of speeds for specific use cases, such as 25/50 Gbps for data center ToR designs and 2.5/5 Gbps for wireless infrastructure backhaul. The availability of a wide variety of standardized Ethernet data rates allows the network manager to customize a solution to optimize performance, cost, and energy consumption goals[22].

This popularity of Ethernet technology is due to the availability of cost-effective, reliable, and interoperable networking products from a variety of vendors. The development of converged and unified communications, the evolution of massive server farms, and the continuing expansion of VoIP, TVoIP, and Web 2.0 applications have accelerated the need for ever-faster Ethernet switches.

The success of Gigabit Ethernet and 10-Gbps Ethernet highlights the importance of network-management concerns in choosing a network technology. The 40- and 100-Gbps Ethernet specifications offer compatibility with existing installed LANs, network-management software, and applications. This compatibility has accounted for the survival of 30-year-old technology in today's fast-evolving network environment.

References

[1] Metcalfe, R., and Boggs, D., "Ethernet: Distributed Packet Switching for Local Computer Networks," *Communications of the ACM,* July 1976.

[2] Shoch, J., Dalal, Y., Redell, D., and Crane, R., "Evolution of the Ethernet Local Computer Network," *Computer,* August 1982.

[3] Stallings, W., "Gigabit Ethernet," *The Internet Protocol Journal,* Volume 2, No. 3, September 1999.

[4] Frazier, H., and Johnson, H. "Gigabit Ethernet: From 100 to 1,000 Mbps," *IEEE Internet Computing,* January/February 1999.

[5] GadelRab, S., "10-Gigabit Ethernet Connectivity for Computer Servers," *IEEE Micro,* May-June 2007.

[6] Mcgillicuddy, S., "40 Gigabit Ethernet: The Migration Begins," *Network Evolution E-Zine,* December 2012.

[7] Chanda, G., and Yang, Y., "40 GbE: What, Why & Its Market Potential," Ethernet Alliance White Paper, November 2010.

[8] Nowell, M., Vusirikala, V., and Hays, R., "Overview of Requirements and Applications for 40 Gigabit and 100 Gigabit Ethernet," Ethernet Alliance White Paper, August 2007.

[9] D'Ambrosia, J., Law, D., and Nowell, M., "40 Gigabit Ethernet and 100 Gigabit Ethernet Technology Overview," Ethernet Alliance White Paper, November 2008.

[10] Toyoda, H., Ono, G., and Nishimura, S., "100 GbE PHY and MAC Layer Implementation," *IEEE Communications Magazine,* March 2010.

[11] Rabinovich, R., and Lucent, A., "40 Gb/s and 100 Gb/s Ethernet Short Reach Optical and Copper Host Board Channel Design," *IEEE Communications Magazine,* April 2012.

[12] Morgan, T., "IEEE Gets Behind 25G Ethernet Effort," *Enterprise Tech,* July 27, 2014.

[13] Merritt, R., "50G Ethernet Debate Brewing," *EE Times,* September 3, 2014.

[14] Nolle, T., "Will We Ever Need 400 Gigabit Ethernet Enterprise Networks?" *Network Evolution E-Zine,* December 2012.

[15] D'Ambrosia, J., Mooney, P., and Nowell, M., "400 Gb/s Ethernet: Why Now?" Ethernet Alliance White Paper, April 2013.

[16] Hardy, S., "400 Gigabit Ethernet Task Force Ready to Get to Work," *Lightwave,* March 28, 2014.

[17] D'Ambrosia, J., "400GbE and High Performance Computing," *Scientific Computing Blog,* April 18, 2014.

[18] Duffy, J., "Bridge to Terabit Ethernet," *Network World,* April 20, 2009.

[19] D'Ambrosia, J., "TEF 2014: The Rate Debate," *Ethernet Alliance Blog,* June 23, 2014.

[20] Kipp, S., "5 New Speeds – 2.5, 5, 25, 50 and 400 GbE," *Ethernet Alliance Blog,* August 8, 2014.

[21] Stallings, W., "Gigabit Wi-Fi," *The Internet Protocol Journal,* Volume 17, No. 1, September 2014.

[22] Chalupsky, D., and Healey, A., "Datacenter Ethernet: Know Your Options," *Network Computing,* March 28, 2014.

WILLIAM STALLINGS is an independent consultant and author of numerous books on security, computer networking, and computer architecture. His latest book is *Data and Computer Communications* (Pearson, 2014). He maintains a computer science resource site for computer science students and professionals at **`ComputerScienceStudent.com`**. He has a Ph.D. in computer science from M.I.T. He can be reached at **`ws@shore.net`**