


Strategic Directions **Moving** the **Decimal Point: An Introduction** to **10 Gigabit Ethernet**

by: Bruce Tolley



The next step for scaling the performance and functionality of enterprise and service provider networks is to combine multigigabit bandwidth and intelligent services in order to achieve scaled, intelligent, multigigabit networks with network links that range in speed from 10 Mbps to 10,000 Mbps. Since March 1999 the Ethernet industry has been working on increasing the speed of Ethernet from one to ten gigabits per second. This technology is very significant because not only will Ethernet run at 10 Gigabits per second and serve as a local-area network (LAN) connection, but it will also work in metropolitan area networks (MANs) and wide-area networks (WANs). With 10 Gigabit Ethernet, network managers will be able to build LANs, MANs, and WANs using Ethernet as the end-to-end Layer 2 transport.

Introduction

Managers of enterprise and service provider networks have many choices when they design networks. They have multiple media, technologies, and interfaces to choose from to build campus and metro connections, including OC-12 (622 Mbps), OC-48 (2.488 Mbps), synchronous optical network (SONET) and its European and international equivalent, synchronous digital hierarchy (SDH), packet over SONET (POS), dynamic packet transport (DPT), and spatial reuse protocol (SRP). By early 2001, customers will also be able to choose prestandard versions of 10 Gigabit Ethernet, the latest IEEE 802.3 Ethernet interface.¹

With this new Ethernet technology, bandwidth can be scaled from one to ten gigabits per second without sacrificing any of the intelligent network services such as multiprotocol label switching (MPLS), Layer 3 switching, quality of service (QoS), caching, server load balancing, security, and policy-based networking. These services can be delivered at line rates over the Ethernet network and supported over all network physical infrastructures in the LAN, MAN, and WAN.

10 Gigabit Ethernet is Ethernet. 10 Gigabit Ethernet uses the IEEE 802.3 Ethernet media access control (MAC) protocol, the IEEE 802.3 Ethernet frame format, and the IEEE 802.3 frame size. 10 Gigabit Ethernet is full duplex, just like full-duplex Fast Ethernet and Gigabit Ethernet; therefore, it has no inherent distance limitations. Because 10 Gigabit Ethernet is still Ethernet, it minimizes the user's learning curve by maintaining the same management tools and architecture.

1. For the purposes of this white paper SONET and SDH will be considered synonymous.



The IEEE 802.3ae 10 Gigabit Ethernet Task Force is chartered with developing the 10 Gigabit Ethernet specifications. This group is a subcommittee of the larger 802.3 Ethernet Working Group. While some details are still being worked out, the Task Force has achieved consensus not only on scope and purpose and but also on the core technical content. In contrast to previous Ethernet standards, 10 Gigabit Ethernet targets three application spaces: the LANs, MANs, and WANs.

While customers will see prestandard 10 Gigabit Ethernet products in 2001, the Task Force does not expect the standard to be completed until March 2002. A key milestone for the standards effort was the creation of a first draft in September 2000.

This white paper presents an overview of 10 Gigabit Ethernet technology and communicates Cisco's strategic direction regarding IEEE 802.3ae Ethernet by addressing the following topics:

- Technology and application trends
- 10 Gigabit Ethernet and other higher-layer services
- Expected LAN applications
- Expected MAN applications
- Expected WAN access applications
- 10 Gigabit Ethernet technology and standard
- Goals of the IEEE 802.3ae Task Force
- Task force milestones
- Optical transceivers and media supported
- Fiber infrastructure issues
- Recommendations for new fiber installation

Technology and Application Trends

The last incarnation of Ethernet, 1000BASE-T, extended Gigabit Ethernet operation over the installed Category 5 copper infrastructure and delivered Gigabit bandwidth and intelligence to enterprise and service provider data centers and server farms. The newest version of Ethernet, 10 Gigabit Ethernet, moves the decimal point from one to ten gigabits per second, enabling Ethernet to match the speed of the fastest technology on the WAN backbone, OC-192, which runs at approximately 9.5 Gbps.

The compelling reasons to develop 10 Gigabit Ethernet technology (also known as 10GBASE-X) are that it achieves the following:

- Supports all services (packetized voice and video, data) and Layer 3 through Layer 7 intelligence
- Easily scales enterprise and service provider networks
- Leverages the 300 million plus installed base of Ethernet switch ports
- Supports LANs, MANs, and WANs
- Matches the MAN/WAN backbone speed of OC-192

Not only does 10 Gigabit Ethernet match the data rate of SDH or SONET OC-192, but the IEEE 802.3ae 10 Gigabit Ethernet Task Force has specified an optional interface that delivers Ethernet data at a rate compatible with SONET/SDH. This WAN physical layer (PHY) interface enables the attachment of packet-based IP/Ethernet switches to SONET/SDH access equipment.



Support for All Other Services

Because 10 Gigabit Ethernet is Ethernet and requires no change to the Ethernet media access control (MAC) protocol or packet format, 10 Gigabit Ethernet supports all upper-layer services. These include all of the Cisco Intelligent Network Services that operate at Layer 2 and Layer 3 and higher of the OSI seven-layer model: high availability, multiprotocol label switching (MPLS), quality of service (QoS) including voice over IP (voIP) services, security and policy enforcement, server load balancing (SLB) and Web caching. 10 Gigabit Ethernet will also support all standard Layer 2 functions: 802.1p, 802.1Q VLANs, EtherChannel™, and spanning tree.

Simple to Scale

Because almost all network traffic today starts out as Ethernet and Internet protocol (IP) traffic, building Ethernet networks with the next step up in speed will be the easiest way to scale enterprise and service-provider networks. A fundamental rule of building switched networks is that a faster technology is always needed to aggregate multiple, lower-speed segments. As the density and the number of 100-Mbps segments at the edge of the network increase, 1000BASE-X and 1000BASE-T will become the uplink technology from the wiring closet to the core of the network. At the close of 2000, the Ethernet industry is shipping over 250 thousand Gigabit ports a month. 10 Gigabit is needed to aggregate these Gigabit segments.

In short, for enterprise LAN applications, 10 Gigabit Ethernet will enable network managers to scale their Ethernet networks from 10 Mbps, 100 Mbps, or 1000 Mbps to 10,000 Mbps, while leveraging their investments in Ethernet as they increase their network performance. For service provider metropolitan and wide-area applications, 10 Gigabit Ethernet will provide high-performance, cost-effective links that are easily managed with Ethernet tools.

Leverages the Installed Base of Ethernet: 10/100/10000

Over 300 million switched Ethernet ports have been installed worldwide. The development of 10 Gigabit Ethernet will not only enable network managers to take advantage of their investments in the installed equipment, network management and analysis tools, and IT staff expertise, it will also enable the Ethernet industry to take advantage of merchant silicon and optical components.

10 Gigabit Ethernet will be a cost-effective means of building 10-Gbps links. The reasons for this are two-fold. First, there is the design philosophy of the Ethernet industry, which assumes high-volume manufacturing and low-cost design. In fact, the Task Force adopted as a design goal the objective to develop 10 Gigabit Ethernet interfaces that would offer ten times the performance at three to four times the cost of the previous generation of Ethernet.

Second, because the whole industry, from chipset vendors and optical components manufacturers up the value chain to systems vendors, participates in the standards process, the 802.3 Ethernet interfaces are very well defined and can be implemented with available technology. This process enables intense market competition at every stage of the value chain, which lowers the costs of components and subsystems and lowers the costs of the systems that are available to the IT managers who are the final end customers. This process also means that most of the complexity of the encoding schemes and electronic circuitry becomes embedded in merchant silicon, which lowers costs and increases competition among vendors. The contribution of Ethernet to technical innovation, competition, and increasingly lower costs to final end users is evident in the traditional Ethernet cost curves witnessed with Ethernet, Fast Ethernet, and, most recently, Gigabit Ethernet. Finally, in contrast to a 10-Gbps telecommunications laser, 10 Gigabit Ethernet short links (less than 40 km on single-mode fiber) will use low-cost, uncooled optics and, in the future, vertical cavity surface emitting lasers (VCSELs), which are very low cost.



Network Applications of 10 Gigabit Ethernet

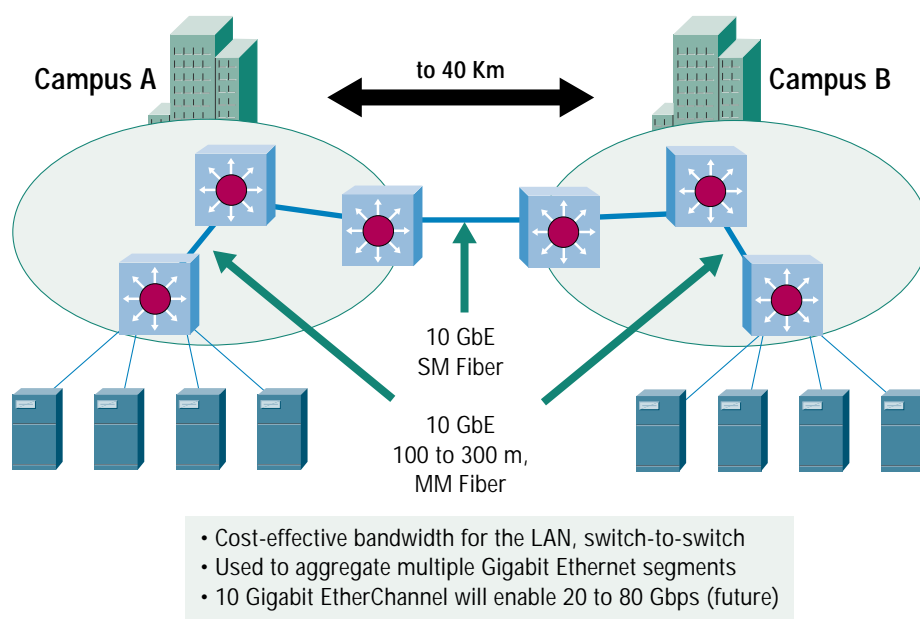
LAN Applications

10 Gigabit Ethernet has many potential applications for both service provider and enterprise networks. Figure 1 shows the standard LAN applications for 10 Gigabit Ethernet, which include the following:

- Server interconnect for clusters of servers
- Aggregation of multiple 1000BASE-X or 1000BASE-T segments into 10 Gigabit Ethernet downlinks
- Switch-to-switch links for very high-speed connections between switches in the same data center, in an enterprise backbone, or in different buildings

The building cabling media supported for LAN applications will be single-mode fiber, FDDI-grade multimode fiber, and possibly the new, higher-bandwidth multimode fiber.

Figure 1 10 Gigabit in the LAN



10 Gigabit Ethernet is used to link the core switches in both Campus A and Campus B. In fact, a primary use of 10 Gigabit Ethernet in the LAN will be to aggregate multiple Gigabit Ethernet segments. By 2003, the Dell'Oro Group forecasts shipments of 24 million Gigabit Ethernet ports. Most of these Gigabit ports will require 10 Gigabit Ethernet as an aggregation technology.

Many LAN customers expect to need 10 Gbps bandwidth soon. For example, network planners at Lawrence Berkeley National Laboratory have estimated their network bandwidth needs for 2001 to range between 5 and 40 Gbps. The applications expected to drive this demand for bandwidth include computations, real-time steering, remote visualization, and remote I/O. (M. Bennett, "A User's Perspective for Ten Gigabit Ethernet," Presentation to IEEE 802.3 Higher Speed Study Group, June 1, 1999.)

10 Gigabit Ethernet will be used in service provider and enterprise data centers and LANs. Initially, network managers will use 10 Gigabit Ethernet to provide high-speed interconnection between large-capacity switches inside the data center or computer room, or between buildings. As the need for bandwidth increases, 10 Gigabit Ethernet will be deployed throughout the entire LAN, including switch to server and MAN and WAN access applications.



Metropolitan Area Network (MAN) Applications

Dark Fiber Metro Applications

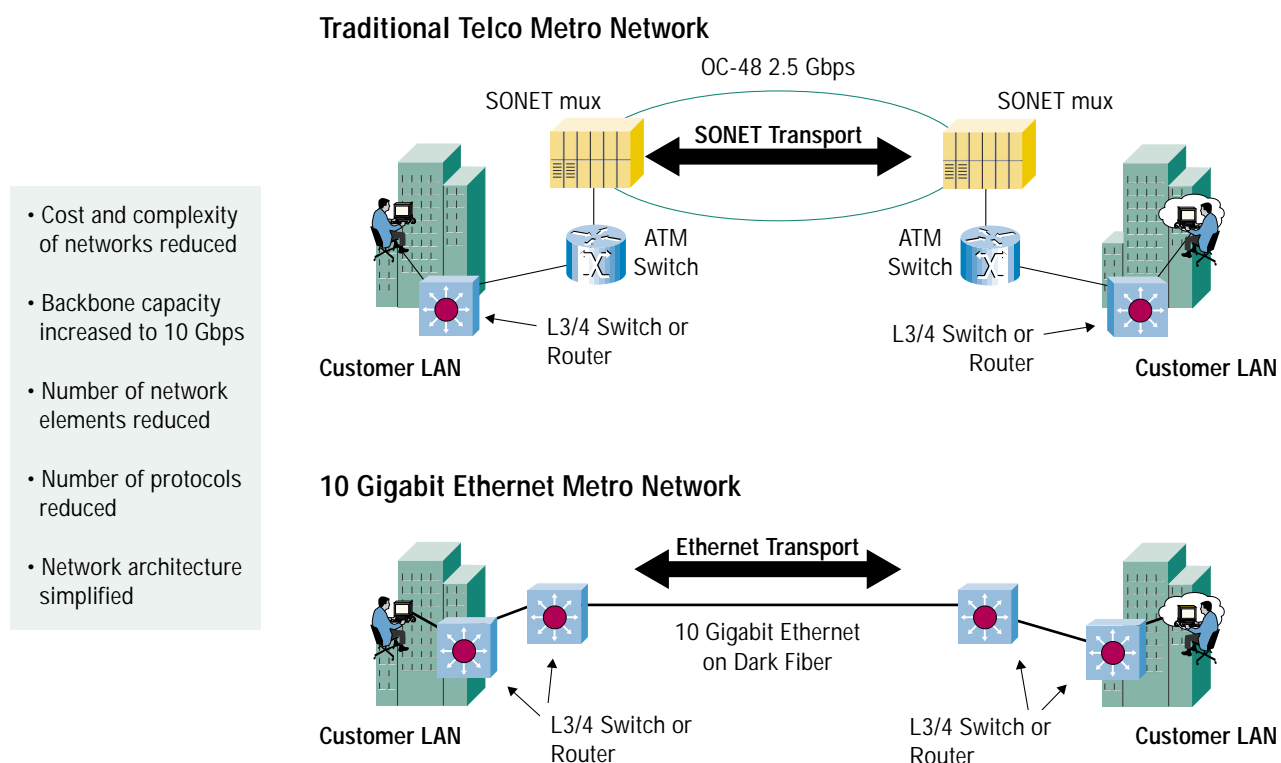
One of the most exciting innovations in the Gigabit space has been the growth in the deployment of long distance Gigabit Ethernet using long wavelength optics (1000BASE-LX and 1000BASE-ZX) on dark fiber to build network links that reach metropolitan distances. Cisco led the way with long distance Gigabit Ethernet, helping customers to build links ranging up to 100 kilometers with the use of pluggable transceivers called *gigabit interface converters* or GBICs.

10 Gigabit Ethernet will be deployed in MAN applications over *dark fiber*, over dark wavelengths, and as a fundamental transport for facility services. The term *dark fiber* refers to unused single-mode fiber capacity from fiber that has been installed for long distance applications that usually reach up to 100 kilometers without amplifiers or optical repeaters. This fiber is not currently “lit,” meaning that it is not carrying traffic and is not terminated.

Service provider and local exchange carriers (LECs) are evaluating the use of 10 Gigabit Ethernet as the backbone technology to build out highly cost-effective, large-scale Layer 2 and Layer 3 infrastructures over which a service could be delivered to customers.

It is important to point out that the use of the term *MAN* in this context refers to the type of distances reached by the whole network and the type of service delivered to the customer. The actual physical and logical topology is that of a very large LAN because the facility provider is not building a traditional SONET/ATM MAN, but instead is using 10 Gigabit Ethernet to build point-to-point links between Layer 3 through 4 switches. TCP/IP or other services can be delivered over the Ethernet network.

Figure 2 10 Gigabit Ethernet Metropolitan Network



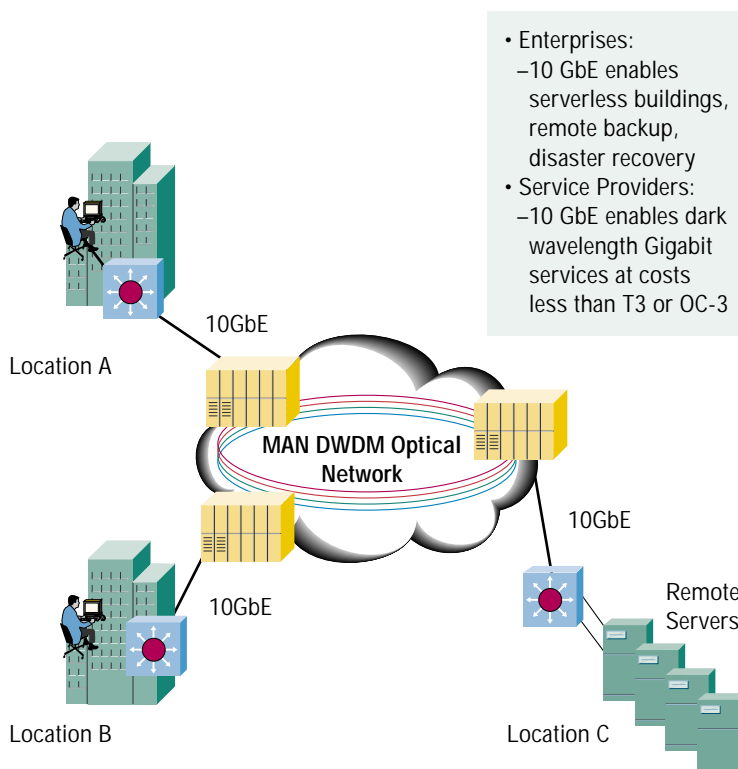


As shown in Figure 2, 10 Gigabit Ethernet metropolitan networks will enable service providers to reduce the cost and complexity of their networks while increasing backbone capacity to 10 Gbps by eliminating the need to build out an infrastructure that contains not only several network elements required to run TCP/IP and data traffic but also the network elements and protocols originally designed to transport voice. Reduction in the number of network elements and network layers lowers equipment costs, lowers operational costs, and simplifies the network architecture. With 10 Gigabit Ethernet backbone networks, service providers will be able to offer native 10/100/1000 Mbps Ethernet as a public service to customers, namely offering the customer twice the bandwidth of the fastest public MAN services OC-3 (155 Mbps) or OC-12 (622 Mbps) with no need for the added complexity of SONET or ATM and no need for protocol conversion.

Dark Wavelength Metro Applications with DWDM

As Figure 3 shows, 10 Gigabit Ethernet will be a natural fit with dense wave division multiplexing (DWDM) equipment as more of the latter is deployed in metropolitan applications. For enterprises access to 10 Gigabit Ethernet services over DWDM will enable serverless buildings, remote backup and disaster recovery. For service providers 10 Gigabit Ethernet in the MAN will enable the provisioning of dark wavelength gigabit services at very competitive costs.

Figure 3 10 Gigabit Ethernet in the MAN over DWDM



The term *dark wavelength* or *dark lambda* refers to unused capacity available on a DWDM system. WDM is a long established technology in the WAN backbone that enables multiple data streams to be transformed into multiple independent wavelengths. DWDM refers to systems that apply the tight wavelength spacing specified by the International Telecommunications Union (ITU), which is normally less than a nanometer (nm). Coarse or wide wavelength division multiplexing (CWDM or WWDM) refers to much cheaper systems that use wider spacing between wavelengths, which is normally about 10 nm. The WDM device then multiplexes the multiple (16, 32, and 64) streams into one stream of “white light” across one fiber pair, increasing the bandwidth capacity of the link by 16-, 32-, or 64. At the opposite end, the multiple wavelengths are demultiplexed into the original data streams. Many MANs and much of the WAN backbone today contain installed DWDM equipment that has unused capacity or dark wavelengths.

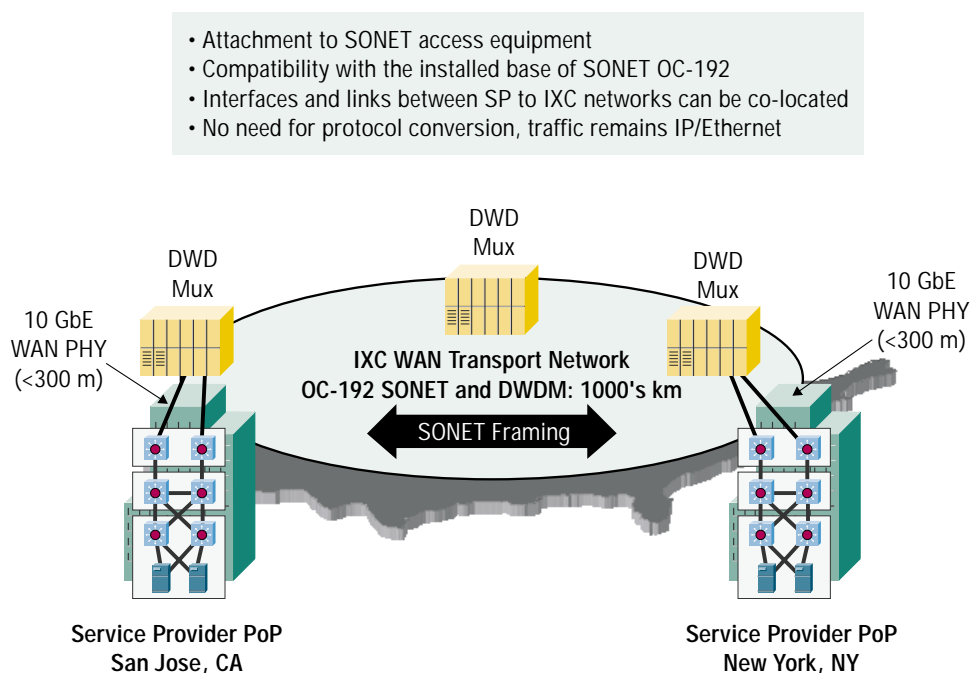


10 Gigabit Ethernet WAN Applications

WAN applications for 10 Gigabit Ethernet look very similar to MAN applications: Dark fiber, dark wavelength, and support for SONET infrastructure. Figure 4 shows multilayer switches and terabit routers attaching via 10 Gigabit Ethernet to the SONET optical network, which includes add drop multiplexers (ADMs) and DWDM devices. When dark wavelengths are available, 10 Gigabit Ethernet can be transmitted directly across the optical infrastructure reaching distances from 70 to 100 km.

SONET is the dominant transport protocol in the WAN backbone today and most MAN public services are offered as SONET OC-3 (155 Mbps) or OC-12 (622 Mbps).

Figure 4 10 Gigabit Ethernet in the WAN



Because most of today's installed optical infrastructure is built out with a specific architecture and specific timing requirements to support OC-192 SONET, the IEEE 802.3ae Task Force specified an Ethernet interface that attaches to the SONET-based TDM access equipment at a data rate compatible with the payload rate of OC-192c/SDH VC-4-64c. This is accomplished by means of a physical layer link based on the WAN PHY between Gigabit or Terabit switches and Ethernet line-terminating equipment (LTE) which is attached to the SONET network. The WAN PHY interface does not attach directly to a SONET OC-192 interface. This WAN PHY interface will allow the construction of MANs and WANs that connect geographically dispersed LANs between campuses or POPs through the SONET transport network. In other words, 10 Gigabit Ethernet interfaces that are compatible with SONET OC-192 payload rate facilitate the transport of native Ethernet packets across the WAN transport network with no need for protocol conversion at Layer 3 at the WAN edge. The conversion or encapsulation happens at Layer 1 in the PHY.

As stated above, reducing the need for protocol conversion increases the performance of the network and makes it simpler, easier to manage, and less costly because protocol conversion is CPU-intensive and adds complexity and additional elements to the network. Figure 4 shows that this WAN PHY interface will enable Internet service providers (ISPs) and network service providers (NSPs) to create very high-speed links at very low cost between co-located, carrier-class switches and routers. Note that in certain applications, this WAN PHY link might well be less than 300 meters.



The flexibility, compatibility, and simplicity of 10 Gigabit Ethernet cannot be overemphasized. It will be deployed in LANs and MANs. It will be used with OC-48 and OC-192. It will also be used over dark fiber without OC-192 or SONET.

10 Gigabit Ethernet: Technology and Standard

The Ubiquity of Ethernet

Without question, 10/100/1000 Mbps Ethernet is the world's most ubiquitous networking technology. Over 90 percent of today's network-attached desktops are attached with Ethernet. Since 1985, over 300 million 10, 100, and 1000 Mbps Ethernet ports have shipped with a market value exceeding billions of U.S. dollars. Once volume shipments begin to ramp, 10 Gigabit Ethernet interfaces are expected to be a very cost-effective means of deploying 10-Gbps bandwidth. The ubiquity of Ethernet makes it easy to find human and other resources to manage Ethernet networks. Because 10 Gigabit Ethernet is still Ethernet, it minimizes the IT manager's learning curve by maintaining the same management tools and architecture.

10 Gigabit Ethernet is the next stage in the natural evolution of Ethernet. All Gigabit Ethernet ports shipped today are full duplex and switched; Gigabit Ethernet is no longer a shared domain, half-duplex technology. 10 Gigabit Ethernet will also be a full-duplex, switched technology, maintaining compatibility with the 802.3 Ethernet MAC protocol and the Ethernet frame format.

Roles of the IEEE 802.3ae Task Force and 10 Gigabit Ethernet Alliance

In March 1999 representatives from Cisco Systems helped start the IEEE 10 Gigabit Ethernet standards effort by participating in the creation of a study group called the Higher Speed Study Group (HSSG). Since then, approximately 70 companies and 150 engineers have been working on 10 Gigabit Ethernet in the IEEE 802.3 HSSG. This study group concluded its work in November 1999 and in February 2000 became the IEEE 802.3ae 10 Gigabit Ethernet Task Force. The goal of the Task Force is to write the specifications for a 10 Gigabit Ethernet standard to be ratified by March 2002.

In February 2000, Cisco cofounded with six other companies the 10 Gigabit Ethernet Alliance (10GEA), an industry consortium with three primary aims:

- Promote industry awareness, acceptance, and advancement of the 10 Gigabit Ethernet standard
- Provide a forum for technical discussion and provide technical contributions to the IEEE 802.3ae Task Force; thus accelerating the standards effort
- Provide resources to establish and demonstrate multivendor interoperability of 10 Gigabit Ethernet products

The IEEE 802.3 Working Group, with its subcommittees such as the IEEE 802.3ae Task Force, is a standards body where individuals not companies participate and vote. In comparison, the Alliance is an industry consortium whose membership consists of companies. The Alliance does not write the standard, but supports the IEEE 802.3 standards effort. Cisco helps lead the activities of the Alliance by serving on the Alliance Board of Directors. A representative from Cisco is currently Vice President of the Alliance. By September 2000, membership of the 10 GEA reached over 80 members including all the major networking and telecommunications equipment vendors and many startup companies.

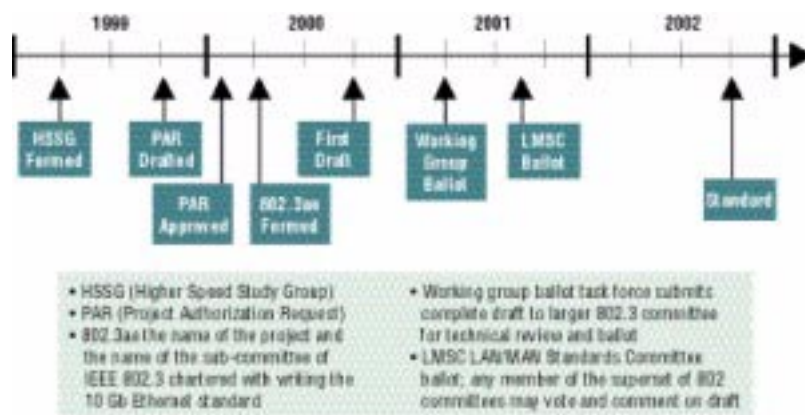
Task Force Milestones

In March 1999, the Task Force adopted the schedule shown in Figure 5. The Task Force created its First Draft in September 2000 and is on track for completion of the standard by March 2002. This schedule is very aggressive. In the cases of Fast Ethernet (100BASE-T) and Gigabit Ethernet (1000BASE-X), the Ethernet industry borrowed technology from the ANSI FDDI PHY and the ANSI Fibre Channel PHY, respectively. In contrast, the 802.3ae effort has had to define one or more new PHYs and four PMDs. In March 2001, the Task Force and IEEE 802.3 Working Group both voted overwhelmingly to forward the Draft specification to Working Group ballot.



Figure 5 Task Force Schedule

Standards-based products will not be available until the standard is complete in the first half of 2002. Prestandard products will be available in 2001.



Scope of the Standards Effort

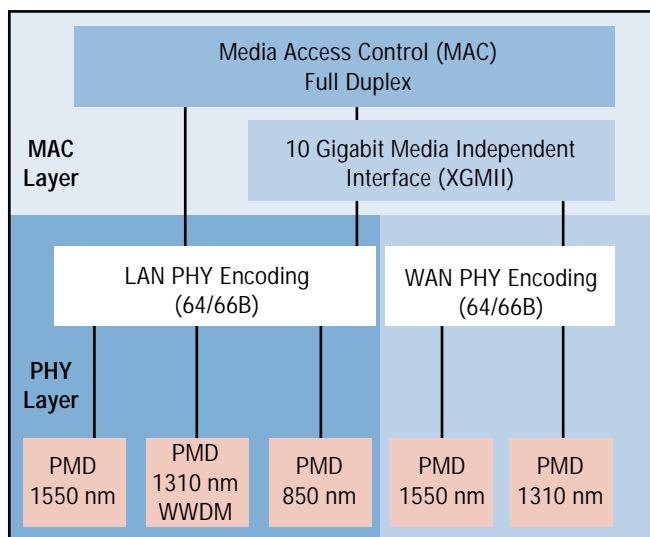
Like all previous Ethernet technologies, 10 Gigabit Ethernet is a layered protocol (see Figure 6). The standard will define protocols and interfaces at and below the MAC layer, including additions to the MAC layer, management objects, and encoding schemes for transmission and signaling over supported media. This means that the scope of the Task Force's work includes defining 10 Gigabit Ethernet at Layer 2, and at Layer 1 defining the physical layer devices (PHYs) and defining the optical transceivers or PMDs.

During 1999 and 2000, after months of intensive investigation, consensus and critical mass of support developed for the technical approach, scope, and goals of the standards effort represented in its formal goals. The objectives of the 802.3ae Task Force include the following:

- Preserve the 802.3 Ethernet frame format at the MAC client service interface
- Preserve the minimum and maximum frame size of the current IEEE 802.3 Standard
- Support full-duplex operation only
- Support star-wired LANs using point-to-point and structured cabling topologies
- Specify an optional media-independent interface
- Support the IEEE 802.3ad standard for link aggregation
- Support a speed of 10,000 Gbps at the service interface between the MAC and the physical layer signaling sublayer (PLS)
- Define two families of physical layers (PHYs):
 - A LAN PHY operating at a data rate of 10,000 Gbps
 - A WAN PHY operating at a data rate compatible with the payload rate of OC-192c/SDH VC-4-64c
- Define a mechanism to adapt the MAC/PLS data rate to the data rate of the WAN PHY
- Support fiber media selected from the second edition of the international and European building wiring standards, ISO/IEC 11801



Figure 6 10 Gigabit Ethernet Layered Protocol



In terms of the classical OSI layer model, Ethernet is a Layer 2 protocol. Just like previous generations of Ethernet such as 1000BASE-X and 1000BASE-T remained Ethernet, 10 Gigabit Ethernet remains Ethernet. 10 Gigabit Ethernet uses the IEEE 802.3 Ethernet MAC protocol and the same Ethernet frames as 10, 100, and 1000 Mbps Ethernet.

Although the IEEE 802.3 standard for Gigabit Ethernet supported both full and half duplex, customers selected only products that provided full-duplex operation. For this reason, the IEEE 802.3ae Task Force decided unanimously that 10 Gigabit Ethernet would be full duplex only. Because 10 Gigabit Ethernet is full duplex only, it is unlimited in reach. Only the physics of transmission and the physical media limit the distance of the link. In other words, because there are no packet collisions in a full-duplex link, the link distances are determined by optics and not by the diameter of an Ethernet collision domain.

Ethernet remains an asynchronous link protocol, so that timing or synchronization is maintained within each character in the bit stream of data and the receiving router, switch, or hub has the opportunity to retime the data and resynchronize at the beginning of the character. In contrast, each device in a SONET network must share the same clock in order to avoid timing drift between transmission and reception.

Layer 1: Physical Layer Devices

In order to address the LAN, MAN, and WAN applications, most of the work of the Task Force involves defining appropriate PHYs. The Ethernet PHY at Layer 1 of the OSI model connects the media to the MAC (Layer 2) and defines electrical and optical signaling, line states, clocking guidelines, data encoding, and circuitry needed for data transmission and reception. Contained within the PHY are several sublayers that perform these functions, including the physical coding sublayer (PCS) and the optical transceiver or physical media dependent (PMD) sublayer for fiber media. The PCS is made up of coding (for example, 8b/10b) and serializer or multiplexing functions.

The IEEE 802.3ae 10 Gigabit Ethernet specification defines two PHY types: the LAN PHY and the WAN PHY. The latter is defined by the specification as an optional WAN PHY that can interface with the existing WAN SONET infrastructure. The goal of the Task Force was to define a SONET-compatible PHY that will operate at a data rate that is compatible with the payload rate of OC-192c/SDH VC-4-64c. This SONET-friendly interface allows 10 Gigabit Ethernet switches and routers to attach to SONET access equipment and to use the SONET infrastructure for Layer 1 transport. The WAN PHY is simply an optional extended operating feature that is added to a LAN PHY. The WAN PHY interface is not SONET compliant or



SONET compatible and it does not attach directly to a SONET interface. Finally, it should be pointed out that 10 Gigabit Ethernet PMDs are needed at both ends of the 10 Gigabit Ethernet link between the gigabit switch or router and the LTE (see Table 1).

Table 1 The WAN PHY

| SONET Friendly | Not SONET Compliant |
|---|---|
| Enables use of SONET infrastructure for Layer 1 transport: <ul style="list-style-type: none">• SONET ADMs, DWDM Transponders, optical regenerators | Connects to SONET access devices but not directly to SONET infrastructure |
| Requires some SONET features: <ul style="list-style-type: none">• OC-192 link speed• SONET framing• MinimalPath/Section/Line overheard processing | Avoids most costly aspects of SONET: <ul style="list-style-type: none">• No TDM support• Concatenated OC-192c only• Does not require meeting SONET grid laser specifications, jitter requirements, stratum clocking• Minimal operations, administration, maintenance, and provisioning (OAM&P) |

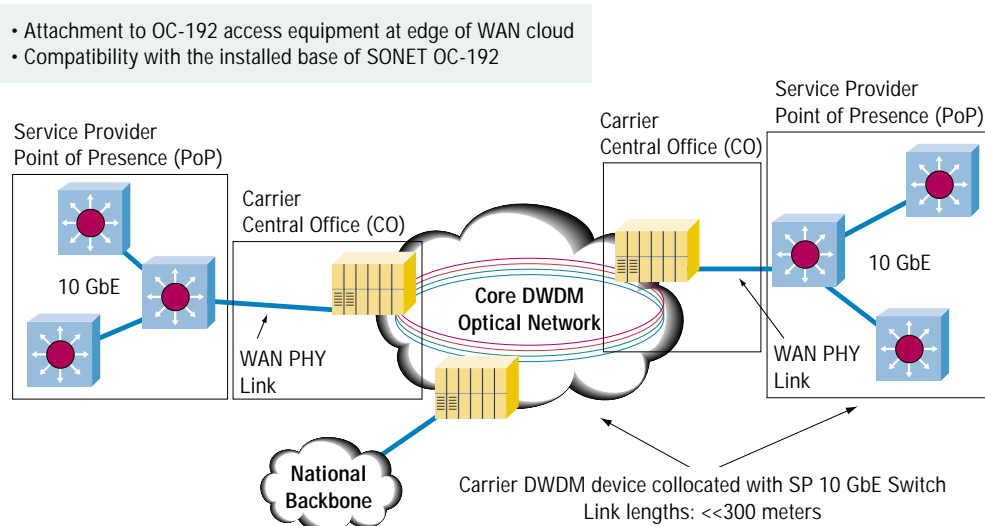
Both the LAN and WAN PHY will support each physical medium-dependent (PMD) sublayer and, therefore, support the same distances. These PHYs are distinguished solely by the PCS. In other words, the LAN and WAN PHYs are distinguished solely by the PCS encoding sublayer and both will be able to use the same optical transceivers.

The WAN PHY differs from the LAN PHY by the inclusion of a simplified SONET framer. Because the line rate of SONET OC-192 is very close to 10 Gbps, it is simple to implement a MAC that can operate either with a LAN PHY at 10 Gbps or with a WAN PHY at the SONET OC-192 payload rate. To enable low-cost WAN PHY implementations, the Task Force specifically rejected conformance to SONET jitter, stratum clock, and other SONET optical specifications.

The WAN PHY is a PHY that does some SONET framing and supports the SONET/SDH payload rate of 9.58 Gbps. Early in the standards process representatives from Cisco Systems proposed a unified PHY that would support both LAN and WAN applications by including the SONET framing functionality in a WAN interface sublayer (WIS). This solution has been instantiated by makers of chipsets that include both LAN and WAN PHY functions. Whether WAN PHY or unified PHY with the WIS enabled, such a PHY will most likely be used to build short links between collocated switches and routers and optical line terminating equipment (LTEs) that are attached to the SONET-based optical cloud (see Figure 7).



Figure 7 Use of the WAN PHY



To conclude the discussion of PHYs, 10 Gigabit defines a LAN PHY that, with simple encoding, will transmit Ethernet packets on dark fiber and dark wavelengths. The LAN PHY is intended to support the existing Ethernet applications at ten times the bandwidth with the most cost-effective solution. 10 Gigabit Ethernet also defines an optional WAN PHY that is used to attach data equipment, like switches or routers, to SONET/SDH access equipment. This allows the creation of Ethernet links that are extended over the SONET/SDH network. It is of course possible to build optical networks without SONET/SDH. Over time it is expected that the LAN PHY will be used in pure optical switching environments for MANs and WANs.

The Operation of the WAN PHY

The “S” in SONET stands for synchronous. SONET/SDH networks use high-accuracy stratum clocks to form a synchronous clock hierarchy. Each SONET/SDH interface within the network is synchronized via this clock hierarchy to a common, central, network clock (typically stratum 1). Having all of the interfaces synchronized in this way enables signals to be transported across the network with a minimal amount of buffering and retiming required at intermediate SONET/SDH nodes. In contrast, the 10 Gigabit Ethernet WAN PHY operates like any other asynchronous network interface. A store and forward buffer device like a router, bridge, or repeater separates each link from the clock domain of the next link.

The functions of the WAN PHY are performed inside an optional sublayer defined by the 10 Gigabit Ethernet specification called the WAN interface sublayer (WIS). The WAN PHY is distinguished from the LAN PHY by its PCS. While the PCS functions of frame delineation and 64b/66b encoding stay the same for both LAN and WAN PHY, in the WAN PHY the whole PCS data stream (including idles) is combined with the SONET path overhead (mostly fixed values) in the correct order and is mapped into the standard STS-192c payload envelope. In the next step, the STS-192c payload envelope is combined with the SONET line and section overhead (mostly fixed values) and is mapped into the WIS frame, which is equivalent to a standard STS-192c frame. This WIS frame is SONET scrambled (X^7+X^6+1) and then sent to the PMD through the 16-bit XBI interface. Note that everything but the first row of the section overhead is scrambled, which is standard SONET practice.

In the receive direction, the reverse happens. First, the transport overhead is stripped off, then the path overhead is stripped off, and the consequent PCS data stream can be delineated and decoded by the PCS. Ethernet frames are mapped into the SONET payload envelope until the latter is full and, should a packet not be fully mapped by the time the envelope is full, the remainder will simply be mapped into the beginning of the next envelope. While packaging the Ethernet bit stream into the SONET payload, the WAN PHY provides some SONET/SDH management information that enables the SONET network manager to view the Ethernet WAN PHY links as though they are SONET/SDH links. It is then possible to perform performance monitoring and fault isolation on the entire network including the 10 Gigabit Ethernet WAN PHY links from the SONET/SDH management station. The WIS that includes the SONET framer also provides the SONET/SDH management information. The WIS operates between the PCS and PMD layers that are common to the LAN PHY.

To sum up, the WAN PHY is not a SONET/SDH interface. It is asynchronous, not synchronous. It does not support complex features of SONET/SDH networks such as protection switching. The WAN PHY provides only a minimal set of SONET/SDH management information so the link can be managed as a SONET/SDH link. The WAN PHY is simply a low-cost link that uses SONET framing and standard Ethernet PMDs for attachment to SONET access equipment.



10 Gigabit Ethernet and Fiber Infrastructure Issues

Optical Transceivers (PMDs) and Supported Media

The 802.3ae Task Force investigated, in some detail, media and distance goals for 10 Gigabit Ethernet which are listed in Table 2.

Table 2 10 Gigabit Ethernet Link Distance and Media Goals

| |
|--|
| At least 65 meters over multimode fiber |
| At least 300 meters over installed multimode fiber |
| At least 2 km over single-mode fiber |
| At least 10 km over single-mode fiber |
| At least 40 km over single-mode fiber |

Note: As approved by the 802.3ae Task Force and the 802.3 Working Group

Optical transceivers or PMDs are defined for specific media and distances. The PMD sublayer transforms the incoming bit stream of changing voltages into light pulses so that the data can be sent across the fiber optic media. By late 2000, the IEEE 802.3ae Task Force had selected four optical transceivers:

- 1310 nm serial PMD to support at least 10 kilometers for single-mode fiber
- 1550 nm serial PMD to support at least 40 kilometers on single-mode fiber
- 850 nm serial PMD to support at least 65 meters on 500 MHz*km 50/125 micron multimode fiber
- 1310 nm WWDM PMD to support at least 300 meters on installed 160 MHz*km 62.5/125 micron multimode fiber (or better)

Support of the 40 km PMD is an acknowledgement that Gigabit Ethernet is already being successfully deployed in metropolitan and private, long-distance applications. The 1550 nm serial laser PMD reaches over 40 km and possibly up to 80 km over existing single-mode fiber. With such 10 Gigabit Ethernet technology, service providers will be able to build simple Ethernet networks with Layer 3 through Layer 4 switches over dark fiber without SONET or ATM and provision high-speed 10/100/1000 Mbps services at very low cost.

The 1310 nm serial PMD is likely to be the PMD of choice when building links to the optical SONET infrastructure.

The 1310 WWDM PMD, based on wavelength division multiplexing, supports installed multimode fiber backbones by transmitting, multiplexing, and demultiplexing upon reception four optical wavelengths across one 62.5/125 micron multimode fiber pair up to 300 meters. This PMD supports the installed base of 62.5/125 micron multimode fiber with 160/500 MHz*km bandwidth up to distances of 300 meters. This type of fiber is commonly known as FDDI-grade fiber. This PMD will also reach up to 10 km on single-mode fiber.

The Task Force identified the need to support low-cost, short-reach optics for 50/125 micron multimode fiber links reaching at least 65 meters for data centers and other short reach applications. The Task Force selected the 850 nm serial PMD for this application. See Table 3 for a summary of the optical transceivers and the media and distances supported.



Table 3 Optical Transceivers for 10 Gigabit Ethernet

| PMD (Optical Transceiver) | Fiber Supported | Diameter (microns) | Bandwidth (MHz*km) | Minimum Distance (meters) |
|----------------------------------|------------------------|---------------------------|---------------------------|----------------------------------|
| 850 nm Serial | Multimode ¹ | 50.0 | 500 ² | 65 |
| 1310 nm WWDM ³ | Multimode | 62.5 | 160 | 300 |
| 1310 nm WWDM | Single Mode | 9.0 ⁴ | N.A. | 10,000 |
| 1310 nm Serial | Single Mode | 9.0 | N.A. | 10,000 |
| 1550 nm Serial | Single Mode | 9.0 | N.A. | 40,000 |

1. The 802.3ae Task Force is still investigating the merits of supporting high-bandwidth 50 micron multimode fiber

2. 500 MHz*km is the OFL bandwidth value. Rated laser bandwidth might change when the TIA FO.2 work is complete

3. WWDM: Wide wavelength division multiplexing

4. ANSI/TIA/EIA-568-A specifies that the nominal "mode field diameter" shall be 8.7 to 10.0 microns with a tolerance of +/- 0.5 micron at 1310 nm

N.A. Not applicable

To summarize, the IEEE 802.3ae Task Force has selected four optical transceivers: 1310 nm serial for single-mode fiber, 1550 nm serial for single-mode fiber, 850 nm serial for 50/125 micron multimode fiber, 1310 nm wide wavelength division multiplexing (WWDM) for installed 62.5/125 or FDDI-grade multimode fiber and single-mode fiber. The 802.3 Ethernet Working Group ratified the selection of these four PMDs and the refinement of the distance objectives.

Recommendations for Fiber Optic Media in New Installations

Various fiber optic and cabling vendors have been marketing 50 micron diameter multimode fiber cable products with enhanced bandwidth specifications. This fiber type is not yet fully standardized. The second edition of the ISO 11801 (the European and International standard for building cabling) is in the midst of a ballot process. This document includes multimode fiber types with bandwidth specifications up to 2000 MHz*km. The Telecommunications Industry Association (TIA), the U.S. standards body for building cabling, will conduct a ballot process on these specifications starting in November 2000 as an Addendum to TIA/EIA-568-B.3, the Optical Fiber Cabling Components Standard. More important, the optical component and the fiber manufacturers are still working in the TIA Fiber Optic-2.2.1 Task Group on ways both to characterize the launch of lasers into this media and to measure the higher bandwidth. This work is still in progress. At the end of 2000, the IEEE 802.3ae Task Group had not explicitly specified operation of 10 Gigabit Ethernet on this media.

To support the deployment of 10 Gigabit Ethernet, the best choice is to deploy single-mode fiber such as Corning SMF-28 or Mohawk/CDT TruLite™ or equivalent. When support for legacy 100BASE-FX or 1000BASE-SX ports is required, customers should consider a hybrid single-mode and multimode product that contains as many pairs of single-mode fiber as possible. By use of such a product, network managers achieve backward compatibility with 100BASE-FX and backward compatibility with 1000BASE-SX ports, which are cheaper than 1000BASE-LX ports. Conventional 100BASE-FX uses light-emitting diodes (LEDs) and will not run on single-mode fiber.

It is worth noting that some Cisco customers have chosen to cable not only their networks but also their data centers completely with single-mode fiber. For building backbone distances over 300 meters, single-mode fiber will be the only specified solution. Although the termination and connectors as well as the optical transceivers (switch ports) for single-mode fiber are more expensive than for multimode fiber, it cannot be overemphasized that single-mode fiber is the only media that without exception will support not only 10 Gigabit Ethernet in 2001 and 2002, but also the next generation of Ethernet (100 Gbps) in 2005 and 2006.



The bottom line is that for fiber optic premises cabling inside the building network managers and planners should deploy a hybrid single-mode and multimode product that contains as many pairs of single-mode fiber as possible, the multimode fiber for legacy support, and the mode for future multigigabit applications. Outside the building and for any link over 300 meters, single-mode fiber is the right choice.

Single-Mode Fiber and Multimode Fiber

Single-mode fiber cable reduces the diameter of the fiber cable core to between 8 and 10 microns, which is very close to the order of a wavelength, so that one single angle or “mode” of light energy passes along the fiber. Because there is a single transmission path with single-mode fiber, the light can travel long distances with little loss of signal. Single-mode fiber is commonly used between buildings, for metro access, and other telecommunications applications.

In contrast, multimode fiber has wider core diameters. Light is reflected along the core at multiple angles or paths or modes. Because the light is propagated along multiple paths, each path has a different length and hence takes a different time to traverse the fiber. These multiple angles or modes cause the signal elements to spread out in time. Consequently, distortions occur that limit the distance over which the integrity of the light signal can be maintained. Multimode fiber is the predominant type of LAN fiber that is installed within buildings and usually requires less expensive connectors and less expensive active electronics, namely, the optical transceivers on the switch ports.

Because of the various ways that single-mode fiber is marketed and sold, there is occasionally confusion about the difference between 8.3, 9.0, and 10.0 micron diameter single-mode fiber. The difference is fairly esoteric. Industry convention refers to fiber optic cable by the diameter of the core, which is the region where light propagates. Corning, the largest maker of fiber optic cable, prefers to specify its single-mode fiber in terms of the “mode field diameter.” This also refers to the region through which light propagates. So Corning single-mode fiber that has a mode field diameter of roughly 9 microns at 1310 nm, has a core diameter of 8.3 microns. Mode field diameter is dependent on the specific wavelength of the light propagating down the fiber. At 1550 nm, the same Corning single-mode fiber has a mode field diameter of roughly 10 microns. Table 2 above lists 9.0 microns as the diameter of single-mode fiber. ANSI/TIA/EIA-568-A specifies that the nominal mode field diameter shall be 8.7 to 10.0 microns with a tolerance of plus or minus 0.5 microns at 1310 nm.

Conclusion and Summary

To conclude, 10 Gigabit Ethernet might well become the technology of choice for enterprise, metropolitan, and wide-area networks. In terms of physical media, 10 Gigabit Ethernet will support distances to 300 meters on installed multimode fiber and over 40 km (or more) on single-mode fiber. With 10 Gigabit Ethernet, network managers will be able to build LANs, MANs, and WANs using Ethernet as the end-to-end Layer 2 transport. Long-distance reach on single-mode fiber enables network managers to build simple, low-cost metropolitan-sized networks with Layer 3 through Layer 4 switches and 10 Gigabit Ethernet backbones. In addition, 10 Gigabit Ethernet will support an optional SONET-friendly PHY to enable transmission of Ethernet over the SONET transport infrastructure.

Web References:

For IEEE presentations, meeting minutes, and tutorials that are in the public domain, refer to the following URL:
<http://grouper.ieee.org/groups/802/3/ae/index.html>

For an IT manager’s perspective on 10 Gigabit Ethernet:
http://grouper.ieee.org/groups/802/3/10G_study/public/june99/bennett_1_0699.pdf

For general and technical information on 10 Gigabit Ethernet:
www.10gea.org

For information on cabling issues:
<http://www.bicsi.com/>
<http://www.mohawk-cdt.com/main.html>



Glossary

CWDM—Coarse wavelength division multiplexing. A relatively low-cost WDM technology that uses wide spacing (about 10 nm) between wavelengths and uncooled optics.

DWDM—Dense wavelength division multiplexing. A carrier-class WDM technology that uses expensive, cooled optics and tight spacing between wavelengths of less than a nanometer based on the specifications of the International Telecommunications Union (ITU) DWDM wavelength grid.

HSSG—Higher speed study group. The group commissioned by the IEEE 802.3 Ethernet working group with drafting the technical scope and objectives of the standards effort.

IEEE 802.3 Working Group: The permanent standing Ethernet committee of the IEEE 802 standards body.

IEEE 802.3ae Task Force: The group chartered by the IEEE 802.3 working group with writing the specification for 10 Gigabit Ethernet. IEEE 802.3ae is also the name of the document produced by the Task Force.

IEC/ISO 11801: The international and European standard for building cabling.

MAC: Media access control protocol.

PHY: Physical layer device. The Ethernet PHY at Layer 1 of the OSI model defines the electrical and optical signaling, line states, clocking guidelines, data encoding, and circuitry needed for data transmission and reception. Contained within the PHY are several sublayers that perform these functions including the physical coding sublayer (PCS) and the optical transceiver or physical media dependent (PMD) sublayer for fiber media. The Ethernet PHY connects the media to the MAC (Layer 2).

PAR: Project authorization request. The document drafted by the HSSG requesting that a standards effort be started.

PCS: Physical coding sublayer.

PMD: Physical media dependent sublayer, usually an optical transceiver.

TIA 568: The U.S. standard for building cabling.

VCSEL: Vertical cavity surface emitting laser. A laser technology based on the methods of semiconductor manufacturing and packaging that enable high-volume manufacturing, lower-cost, and greater efficiencies.

WAN PHY: The optional PHY that includes a simple SONET framer and operates at a data rate compatible with the payload rate of OC-192c/SDH VC-4-64c.

WDM: Wavelength division multiplexing; A means of data transmission that uses optical multiplexing to enable two or more wavelengths each with its own data source to share a common fiber optic medium.

WWDM: Wide wavelength division multiplexing. A relatively low-cost WDM technology that uses wide spacing and uncooled optics.



Corporate Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters

Cisco Systems Europe
11, Rue Camille Desmoulins
92782 Issy-les-Moulineaux
Cedex 9
France

www.cisco.com
Tel: 33 1 58 04 60 00
Fax: 33 1 58 04 61 00

Americas Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters

Cisco Systems Australia, Pty., Ltd
Level 17, 99 Walker Street
North Sydney
NSW 2059 Australia

www.cisco.com
Tel: +61 2 8448 7100
Fax: +61 2 9957 4350

Cisco Systems has more than 190 offices in the following countries. Addresses, phone numbers, and fax numbers are listed on the

Cisco.com Web site at www.cisco.com/go/offices.

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China • Colombia • Costa Rica • Croatia • Czech Republic • Denmark • Dubai, UAE
Finland • France • Germany • Greece • Hong Kong • Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The
Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia
Slovenia • South Africa • Spain • Sweden • Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Copyright © 2000 Cisco Systems, Inc. All rights reserved. EtherChannel is a trademark, and Cisco, Cisco IOS, Cisco Systems, and the Cisco Systems logo are registered trademarks of Cisco Systems, Inc. or its affiliates in the U.S. and certain other countries. All other brands, names, or trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0008R)