

Ethernet Technologies

Background

The term *Ethernet* refers to the family of local area network (LAN) implementations that includes three principal categories.

- Ethernet and IEEE 802.3—LAN specifications that operate at 10 Mbps over coaxial cable.
- 100-Mbps Ethernet—A single LAN specification, also known as Fast Ethernet, that operates at 100 Mbps over twisted-pair cable.
- 1000-Mbps Ethernet—A single LAN specification, also known as Gigabit Ethernet, that operates at 1000 Mbps (1 Gbps) over fiber and twisted-pair cables.

This chapter provides a high-level overview of each technology variant.

Ethernet has survived as an essential media technology because of its tremendous flexibility and its relative simplicity to implement and understand. Although other technologies have been touted as likely replacements, network managers have turned to Ethernet and its derivatives as effective solutions for a range of campus implementation requirements. To resolve Ethernet's limitations, innovators (and standards bodies) have created progressively larger Ethernet pipes. Critics might dismiss Ethernet as a technology that cannot scale, but its underlying transmission scheme continues to be one of the principal means of transporting data for contemporary campus applications. This chapter outlines the various Ethernet technologies that have evolved to date.

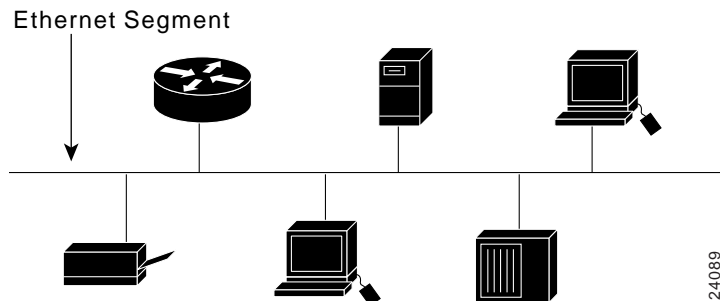
Ethernet and IEEE 802.3

Ethernet is a baseband LAN specification invented by Xerox Corporation that operates at 10 Mbps using carrier sense multiple access collision detect (CSMA/CD) to run over coaxial cable. Ethernet was created by Xerox in the 1970s, but the term is now often used to refer to all CSMA/CD LANs. Ethernet was designed to serve in networks with sporadic, occasionally heavy traffic requirements, and the IEEE 802.3 specification was developed in 1980 based on the original Ethernet technology. Ethernet Version 2.0 was jointly developed by Digital Equipment Corporation, Intel Corporation, and Xerox Corporation. It is compatible with IEEE 802.3. Figure 7-1 illustrates an Ethernet network.

Ethernet and IEEE 802.3 are usually implemented in either an interface card or in circuitry on a primary circuit board. Ethernet cabling conventions specify the use of a transceiver to attach a cable to the physical network medium. The transceiver performs many of the physical-layer functions, including collision detection. The transceiver cable connects end stations to a transceiver.

IEEE 802.3 provides for a variety of cabling options, one of which is a specification referred to as 10Base5. This specification is the closest to Ethernet. The connecting cable is referred to as an attachment unit interface (AUI), and the network attachment device is called a media attachment unit (MAU), instead of a transceiver.

Figure 7-1 An Ethernet network runs CSMA/CD over coaxial cable.



Ethernet and IEEE 802.3 Operation

In Ethernet's broadcast-based environment, all stations see all frames placed on the network. Following any transmission, each station must examine every frame to determine whether that station is a destination. Frames identified as intended for a given station are passed to a higher-layer protocol.

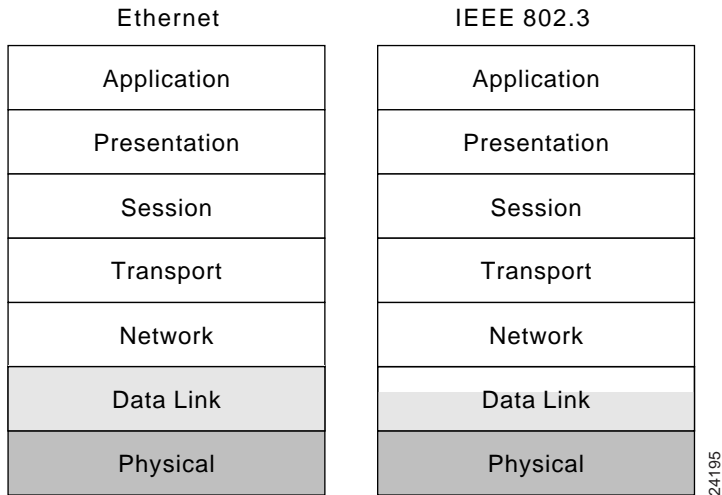
Under the Ethernet CSMA/CD media-access process, any station on a CSMA/CD LAN can access the network at any time. Before sending data, CSMA/CD stations listen for traffic on the network. A station wanting to send data waits until it detects no traffic before it transmits.

As a contention-based environment, Ethernet allows any station on the network to transmit whenever the network is quiet. A collision occurs when two stations listen for traffic, hear none, and then transmit simultaneously. In this situation, both transmissions are damaged, and the stations must retransmit at some later time. Back-off algorithms determine when the colliding stations should retransmit.

Ethernet and IEEE 802.3 Service Differences

Although Ethernet and IEEE 802.3 are quite similar in many respects, certain service differences distinguish the two specifications. Ethernet provides services corresponding to Layers 1 and 2 of the OSI reference model, and IEEE 802.3 specifies the physical layer (Layer 1) and the channel-access portion of the link layer (Layer 2). In addition, IEEE 802.3 does not define a logical link-control protocol but does specify several different physical layers, whereas Ethernet defines only one. Figure 7-2 illustrates the relationship of Ethernet and IEEE 802.3 to the general OSI reference model.

Figure 7-2 Ethernet and the IEEE 802.3 OSI reference model.



Each IEEE 802.3 physical-layer protocol has a three-part name that summarizes its characteristics. The components specified in the naming convention correspond to LAN speed, signaling method, and physical media type. Figure 7-3 illustrates how the naming convention is used to depict these components.

Figure 7-3 IEEE 802.3 components are named according to conventions.

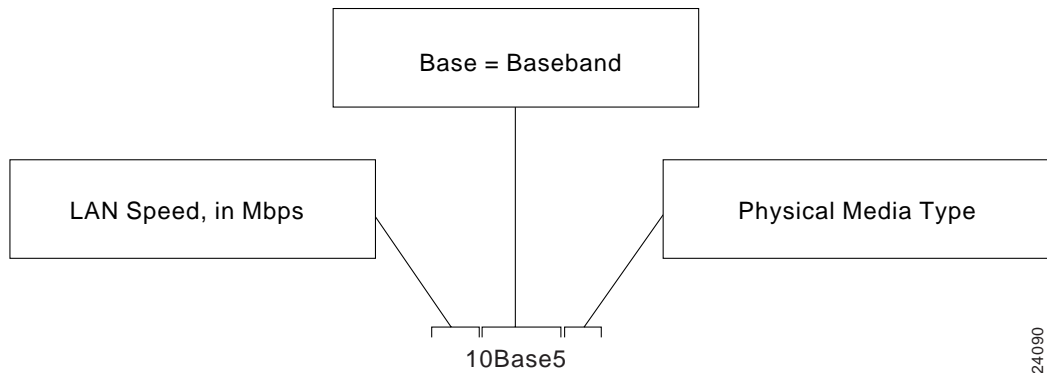


Table 7-1 summarizes the differences between Ethernet and IEEE 802.3, as well as the differences between the various IEEE 802.3 physical-layer specifications.

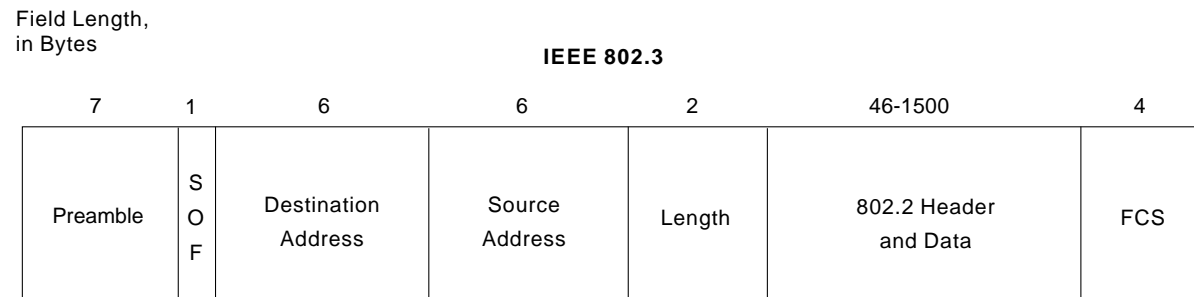
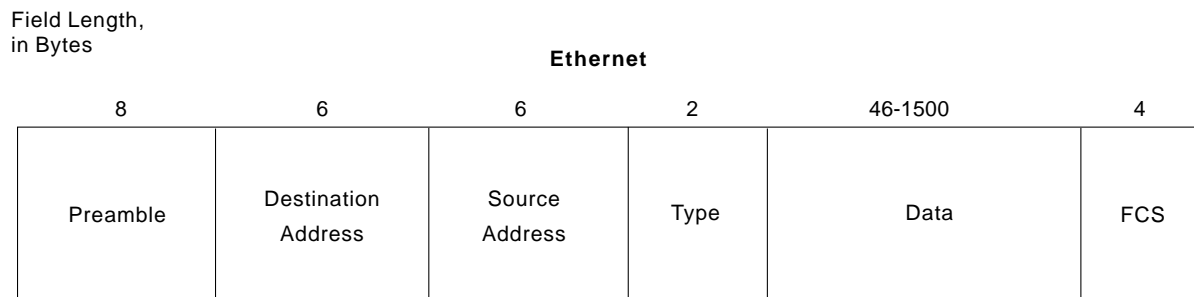
Table 7-1 Comparison of Various IEEE 802.3 Physical-Layer Specifications

Characteristic	Ethernet Value	IEEE 802.3 Values				
		10Base5	10Base2	10BaseT	10BaseFL	100BaseT
Data rate (Mbps)	10	10	10	10	10	100
Signaling method	Baseband	Baseband	Baseband	Baseband	Baseband	Baseband
Maximum segment length (m)	500	500	185	100	2,000	100
Media	50-ohm coax (thick)	50-ohm coax (thick)	50-ohm coax (thin)	Unshielded twisted-pair cable	Fiber-optic	Unshielded twisted-pair cable
Topology	Bus	Bus	Bus	Star	Point-to-point	Bus

Ethernet and IEEE 802.3 Frame Formats

Figure 7-4 illustrates the frame fields associated with both Ethernet and IEEE 802.3 frames.

Figure 7-4 Various frame fields exist for both Ethernet and IEEE 802.3.



SOF = Start-of-Frame Delimiter
 FCS = Frame Check Sequence

ith0706

The Ethernet and IEEE 802.3 frame fields illustrated in Figure 7-4 are as follows.

- *Preamble*—The alternating pattern of ones and zeros tells receiving stations that a frame is coming (Ethernet or IEEE 802.3). The Ethernet frame includes an additional byte that is the equivalent of the Start-of-Frame field specified in the IEEE 802.3 frame.

- *Start-of-Frame (SOF)*—The IEEE 802.3 delimiter byte ends with two consecutive 1 bits, which serve to synchronize the frame-reception portions of all stations on the LAN. SOF is explicitly specified in Ethernet.
- *Destination and Source Addresses*—The first 3 bytes of the addresses are specified by the IEEE on a vendor-dependent basis. The last 3 bytes are specified by the Ethernet or IEEE 802.3 vendor. The source address is always a unicast (single-node) address. The destination address can be unicast, multicast (group), or broadcast (all nodes).
- *Type (Ethernet)*—The type specifies the upper-layer protocol to receive the data after Ethernet processing is completed.
- *Length (IEEE 802.3)*—The length indicates the number of bytes of data that follows this field.
- *Data (Ethernet)*—After physical-layer and link-layer processing is complete, the data contained in the frame is sent to an upper-layer protocol, which is identified in the Type field. Although Ethernet Version 2 does not specify any padding (in contrast to IEEE 802.3), Ethernet expects at least 46 bytes of data.
- *Data (IEEE 802.3)*—After physical-layer and link-layer processing is complete, the data is sent to an upper-layer protocol, which must be defined within the data portion of the frame, if at all. If data in the frame is insufficient to fill the frame to its minimum 64-byte size, padding bytes are inserted to ensure at least a 64-byte frame.
- *Frame Check Sequence (FCS)*—This sequence contains a 4-byte cyclic redundancy check (CRC) value, which is created by the sending device and is recalculated by the receiving device to check for damaged frames.

100-Mbps Ethernet

100-Mbps Ethernet is a high-speed LAN technology that offers increased bandwidth to desktop users in the wiring center, as well as to servers and server clusters (sometimes called server farms) in data centers.

The IEEE Higher Speed Ethernet Study Group was formed to assess the feasibility of running Ethernet at speeds of 100 Mbps. The Study Group established several objectives for this new higher-speed Ethernet but disagreed on the access method. At issue was whether this new faster Ethernet would support CSMA/CD to access the network medium or some other access method.

The study group divided into two camps over this access-method disagreement: the Fast Ethernet Alliance and the 100VG-AnyLAN Forum. Each group produced a specification for running Ethernet (and Token Ring for the latter specification) at higher speeds: 100BaseT and 100VG-AnyLAN, respectively.

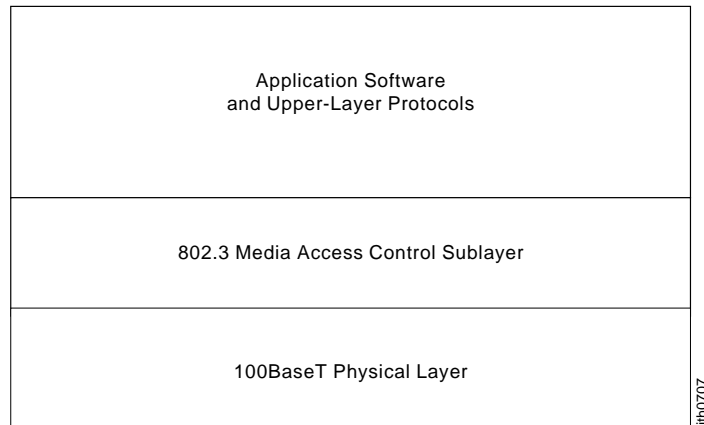
100BaseT is the IEEE specification for the 100-Mbps Ethernet implementation over unshielded twisted-pair (UTP) and shielded twisted-pair (STP) cabling. The Media Access Control (MAC) layer is compatible with the IEEE 802.3 MAC layer. Grand Junction, now a part of Cisco Systems Workgroup Business Unit (WBU), developed Fast Ethernet, which was standardized by the IEEE in the 802.3u specification.

100VG-AnyLAN is an IEEE specification for 100-Mbps Token Ring and Ethernet implementations over 4-pair UTP. The MAC layer is *not* compatible with the IEEE 802.3 MAC layer. 100VG-AnyLAN was developed by Hewlett-Packard (HP) to support newer time-sensitive applications, such as multimedia. A version of HP's implementation is standardized in the IEEE 802.12 specification.

100BaseT Overview

100BaseT uses the existing IEEE 802.3 CSMA/CD specification. As a result, 100BaseT retains the IEEE 802.3 frame format, size, and error-detection mechanism. In addition, it supports all applications and networking software currently running on 802.3 networks. 100BaseT supports dual speeds of 10 and 100 Mbps using 100BaseT fast link pulses (FLPs). 100BaseT hubs must detect dual speeds much like Token Ring 4/16 hubs, but adapter cards can support 10 Mbps, 100 Mbps, or both. Figure 7-5 illustrates how the 802.3 MAC sublayer and higher layers run unchanged on 100BaseT.

Figure 7-5 802.3 MAC and higher-layer protocols operate over 100BaseT.



100BaseT Signaling

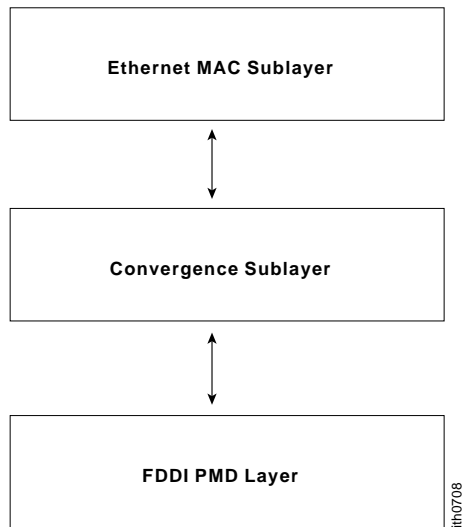
100BaseT supports two signaling types:

- 100BaseX
- 4T+

Both signaling types are interoperable at the station and hub levels. The media-independent interface (MII), an AUI-like interface, provides interoperability at the station level. The hub provides interoperability at the hub level.

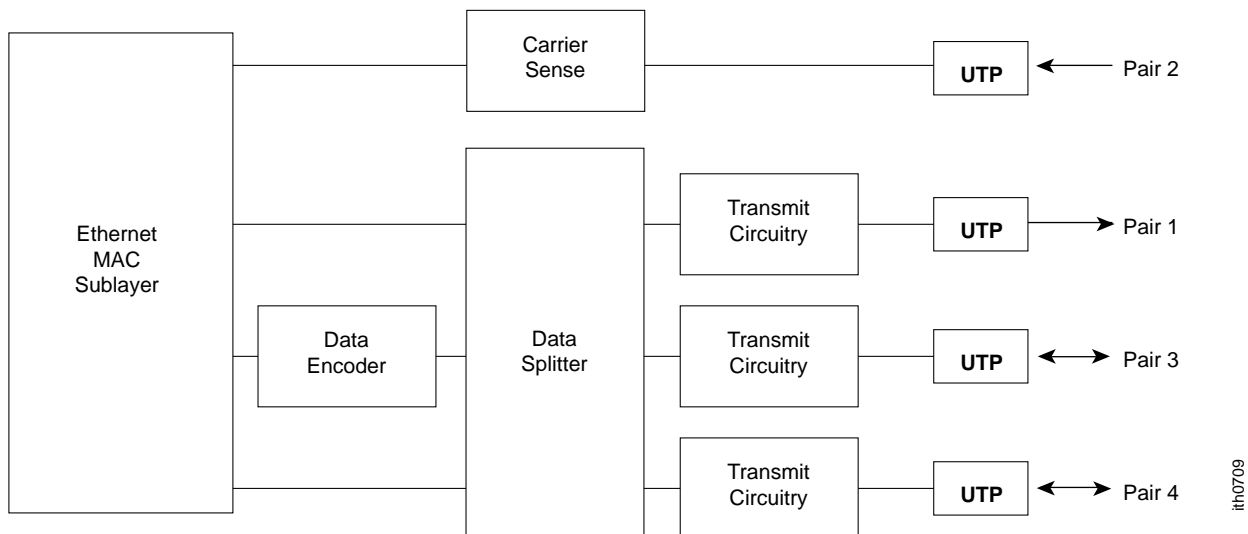
The 100BaseX signaling scheme has a convergence sublayer that adapts the full-duplex continuous signaling mechanism of the FDDI physical medium dependent (PMD) layer to the half-duplex, start-stop signaling of the Ethernet MAC sublayer. 100BaseTX's use of the existing FDDI specification has allowed quick delivery of products to market. 100BaseX is the signaling scheme used in the 100BaseTX and the 100BaseFX media types. Figure 7-6 illustrates how the 100BaseX convergence sublayer interfaces between the two signaling schemes.

Figure 7-6 The 100BaseX convergence sublayer interfaces two signaling schemes.



The 4T+ signaling scheme uses one pair of wires for collision detection and the other three pairs to transmit data. It allows 100BaseT to run over existing Category 3 cabling if all four pairs are installed to the desktop. 4T+ is the signaling scheme used in the 100BaseT4 media type, and it supports half-duplex operation only. Figure 7-7 shows how 4T+ signaling requires all four UTP pairs.

Figure 7-7 4T+ requires four UTP pairs.



100BaseT Hardware

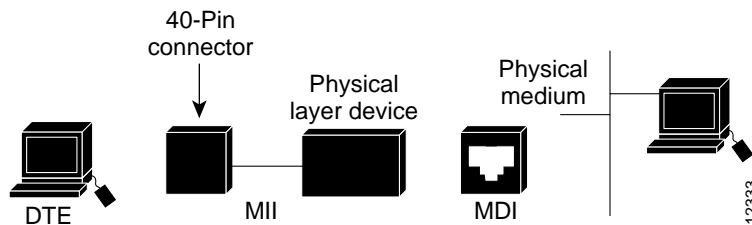
Components used for a 100BaseT physical connection include the following:

- *Physical Medium*—This device carries signals between computers and can be one of three 100BaseT media types:
 - 100BaseTX

- 100BaseFX
- 100BaseT4
- *Medium-Dependent Interface (MDI)*—The MDI is a mechanical and electrical interface between the transmission medium and the physical-layer device.
- *Physical-Layer Device (PHY)*—The PHY provides either 10- or 100-Mbps operation and can be a set of integrated circuits (or a daughter board) on an Ethernet port, or an external device supplied with an MII cable that plugs into an MII port on a 100BaseT device (similar to a 10-Mbps Ethernet transceiver).
- *Media-Independent Interface (MII)*—The MII is used with a 100-Mbps external transceiver to connect a 100-Mbps Ethernet device to any of the three media types. The MII has a 40-pin plug and cable that stretches up to 0.5 meters.

Figure 7-8 depicts the 100BaseT hardware components.

Figure 7-8 100BaseT requires several hardware components.



100BaseT Operation

100BaseT and 10BaseT use the same IEEE 802.3 MAC access and collision detection methods, and they also have the same frame format and length requirements. The main difference between 100BaseT and 10BaseT (other than the obvious speed differential) is the network diameter. The 100BaseT maximum network diameter is 205 meters, which is approximately 10 times less than 10-Mbps Ethernet.

Reducing the 100BaseT network diameter is necessary because 100BaseT uses the same collision-detection mechanism as 10BaseT. With 10BaseT, distance limitations are defined so that a station knows while transmitting the smallest legal frame size (64 bytes) that a collision has taken place with another sending station that is located at the farthest point of the domain.

To achieve the increased throughput of 100BaseT, the size of the collision domain had to shrink. This is because the propagation speed of the medium has not changed, so a station transmitting 10 times faster must have a maximum distance that is 10 times less. As a result, any station knows within the first 64 bytes whether a collision has occurred with any other station.

100BaseT FLPs

100BaseT uses pulses, called FLPs, to check the link integrity between the hub and the 100BaseT device. FLPs are backward-compatible with 10BaseT normal-link pulses (NLPs). But FLPs contain more information than NLPs and are used in the autonegotiation process between a hub and a device on a 100BaseT network.

100BaseT Autonegotiation Option

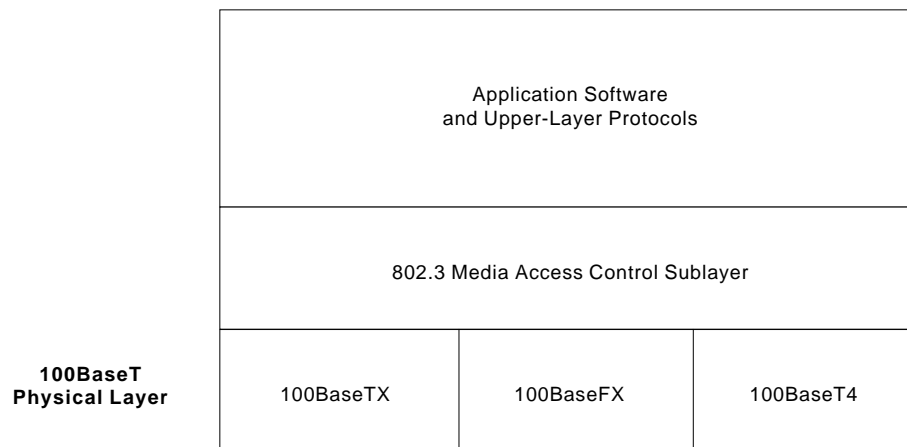
100BaseT networks support an optional feature, called autonegotiation, that enables a device and a hub to exchange information (using 100BaseT FLPs) about their capabilities, thereby creating an optimal communications environment.

Autonegotiation supports a number of capabilities, including speed matching for devices that support both 10-and 100-Mbps operation, full-duplex mode of operation for devices that support such communications, and an automatic signaling configuration for 100BaseT4 and 100BaseTX stations.

100BaseT Media Types

100BaseT supports three media types at the OSI physical layer (Layer 1): 100BaseTX, 100BaseFX, and 100BaseT4. The three media types, which all interface with the IEEE 802.3 MAC layer, are shown in Figure 7-9. Table 7-2 compares key characteristics of the three 100BaseT media types.

Figure 7-9 Three 100BaseT media types exist at the physical layer.



100BaseTX

100BaseTX is based on the American National Standards Institutes (ANSI) Twisted Pair-Physical Medium Dependent (TP-PMD) specification. The ANSI TP-PMD supports UTP and STP cabling. 100BaseTX uses the 100BaseX signaling scheme over 2-pair Category 5 UTP or STP.

Table 7-2 Characteristics of 100BaseT Media Types

Characteristics	100BaseTX	100BaseFX	100BaseT4
Cable	Category 5 UTP, or Type 1 and 2 STP	62.5/125 micron multi-mode fiber	Category 3, 4, or 5 UTP
Number of pairs or strands	2 pairs	2 strands	4 pairs
Connector	ISO 8877 (RJ-45) connector	Duplex SCmedia-interface connector (MIC) ST	ISO 8877 (RJ-45) connector
Maximum segment length	100 meters	400 meters	100 meters
Maximum network diameter	200 meters	400 meters	200 meters

The IEEE 802.3u specification for 100BaseTX networks allows a maximum of two repeater (hub) networks and a total network diameter of approximately 200 meters. A link segment, which is defined as a point-to-point connection between two Medium Independent Interface (MII) devices, can be up to 100 meters. Figure 7-10 illustrates these configuration guidelines.

100BaseFX

100BaseFX is based on the ANSI TP-PMD X3T9.5 specification for FDDI LANs. 100BaseFX uses the 100BaseX signaling scheme over two-strand multimode fiber-optic (MMF) cable. The IEEE 802.3u specification for 100BaseFX networks allows data terminal equipment (DTE)-to-DTE links of approximately 400 meters, or one repeater network of approximately 300 meters in length. Figure 7-11 illustrates these configuration guidelines.

Figure 7-10 The 100BaseTX is limited to a link distance of 100 meters.

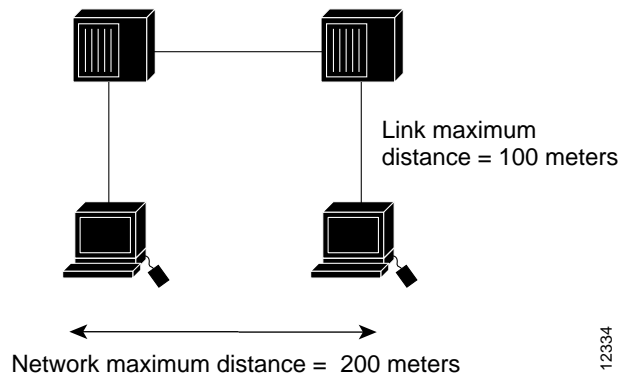
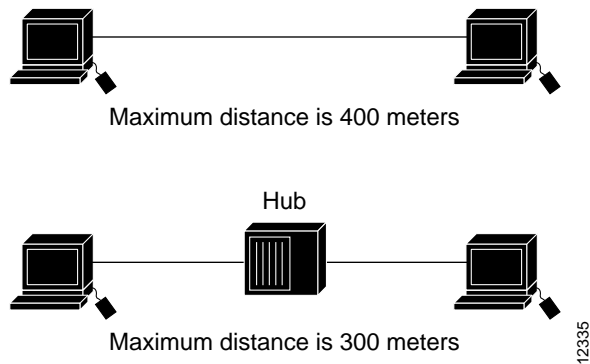


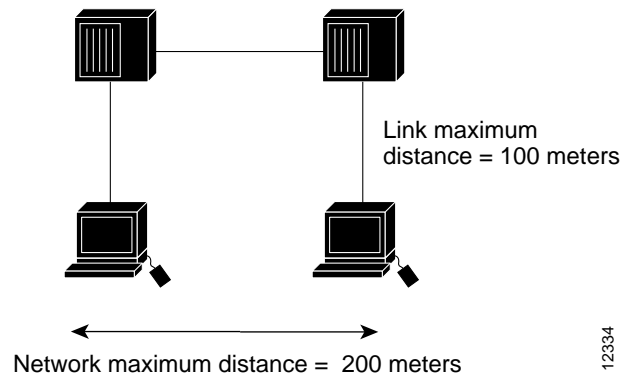
Figure 7-11 The 100BaseFX DTE-to-DTE limit is 400 meters.



100BaseT4

100BaseT4 allows 100BaseT to run over existing Category 3 wiring, provided that all four pairs of cabling are installed to the desktop. 100BaseT4 uses the half-duplex 4T+ signaling scheme. The IEEE 802.3u specification for 100BaseT4 networks allows a maximum of two repeater (hub) networks and a total network diameter of approximately 200 meters. A link segment, which is defined as a point-to-point connection between two MII devices, can be up to 100 meters. Figure 7-12 illustrates these configuration guidelines.

Figure 7-12 The 100BaseT4 supports a maximum link distance of 100 meters.



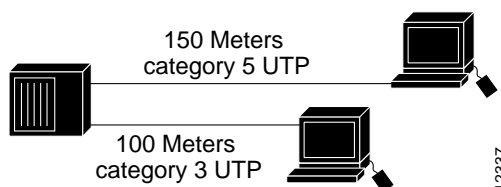
100VG-AnyLAN

100VG-AnyLAN was developed by HP as an alternative to CSMA/CD for newer time-sensitive applications, such as multimedia. The access method is based on station demand and was designed as an upgrade path from Ethernet and 16-Mbps Token Ring. 100VG-AnyLAN supports the following cable types:

- 4-pair Category 3UTP
- 2-pair Category 4 or 5 UTP
- STP
- Fiber optic

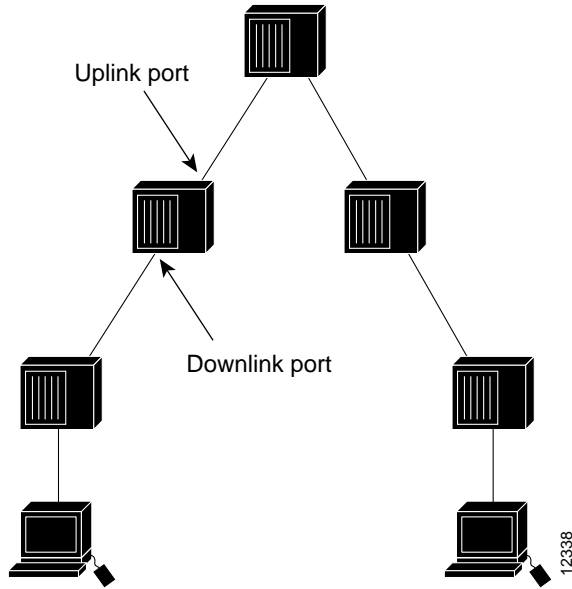
The IEEE 802.12 100VG-AnyLAN standard specifies the link-distance limitations, hub-configuration limitations, and maximum network-distance limitations. Link distances from node to hub are 100 meters (Category 3 UTP) or 150 meters (Category 5 UTP). Figure 7-13 illustrates the 100VG-AnyLAN link distance limitations.

Figure 7-13 100VG-AnyLAN link-distance limitations differ for Category 3 and 5 UTP links.



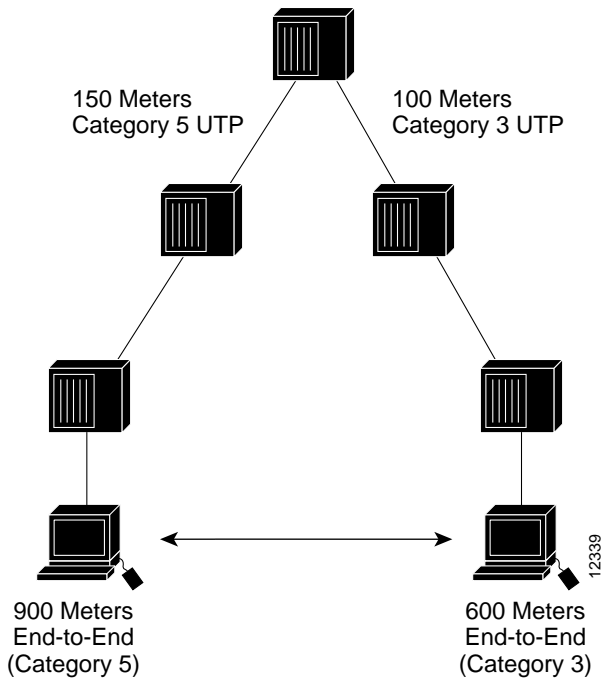
100VG-Any LAN hubs are arranged in a hierarchical fashion. Each hub has at least one uplink port, and every other port can be a downlink port. Hubs can be cascaded three-deep if uplinked to other hubs, and cascaded hubs can be 100 meters apart (Category 3 UTP) or 150 meters apart (Category 5 UTP). Figure 7-14 shows the 100VG-AnyLAN hub configuration.

Figure 7-14 100VG-AnyLAN hubs are arranged hierarchically.



End-to-end network-distance limitations are 600 meters (Category 3 UTP) or 900 meters (Category 5 UTP). If hubs are located in the same wiring closet, end-to-end distances shrink to 200 meters (Category 3 UTP) and 300 meters (Category 5 UTP). Figure 7-15 shows the 100VG-AnyLAN maximum network distance limitations.

Figure 7-15 End-to-end distance limitations differ for 100VG-AnyLAN implementations.



100VG-AnyLAN Operation

100VG-AnyLAN uses a demand-priority access method that eliminates collisions and can be more heavily loaded than 100BaseT. The demand-priority access method is more deterministic than CSMA/CD because the hub controls access to the network.

The 100VG-AnyLAN standard calls for a level-one hub, or repeater, that acts as the root. This root repeater controls the operation of the priority domain. Hubs can be cascaded three-deep in a star topology. Interconnected hubs act as a single large repeater, with the root repeater polling each port in port order.

In general, under 100VG-AnyLAN demand-priority operation, a node wanting to transmit signals its request to the hub (or switch). If the network is idle, the hub immediately acknowledges the request and the node begins transmitting a packet to the hub. If more than one request is received at the same time, the hub uses a round-robin technique to acknowledge each request in turn. High-priority requests, such as time-sensitive videoconferencing applications, are serviced ahead of normal-priority requests. To ensure fairness to all stations, a hub does not grant priority access to a port more than twice in a row.

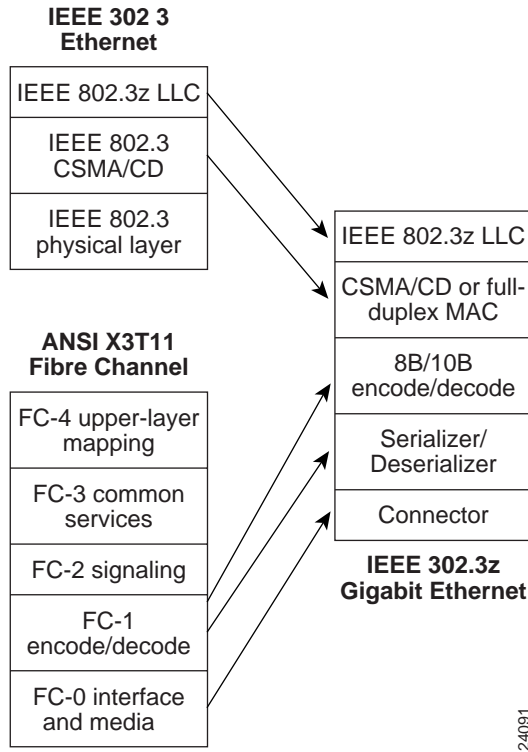
Gigabit Ethernet

Gigabit Ethernet is an extension of the IEEE 802.3 Ethernet standard. Gigabit Ethernet builds on the Ethernet protocol but increases speed tenfold over Fast Ethernet, to 1000 Mbps, or 1 Gbps. This MAC and PHY standard promises to be a dominant player in high-speed LAN backbones and server connectivity. Because Gigabit Ethernet significantly leverages on Ethernet, network managers will be able to leverage their existing knowledge base to manage and maintain Gigabit Ethernet networks.

Gigabit Ethernet Protocol Architecture

To accelerate speeds from 100-Mbps Fast Ethernet to 1 Gbps, several changes need to be made to the physical interface. It has been decided that Gigabit Ethernet will look identical to Ethernet from the data link layer upward. The challenges involved in accelerating to 1 Gbps have been resolved by merging two technologies: IEEE 802.3 Ethernet and ANSI X3T11 Fibre Channel. Figure 7-16 shows how key components from each technology have been leveraged to form Gigabit Ethernet.

Figure 7-16 The Gigabit Ethernet protocol stack was developed from a combination of the Fibre Channel and IEEE 802.3 protocol stacks.

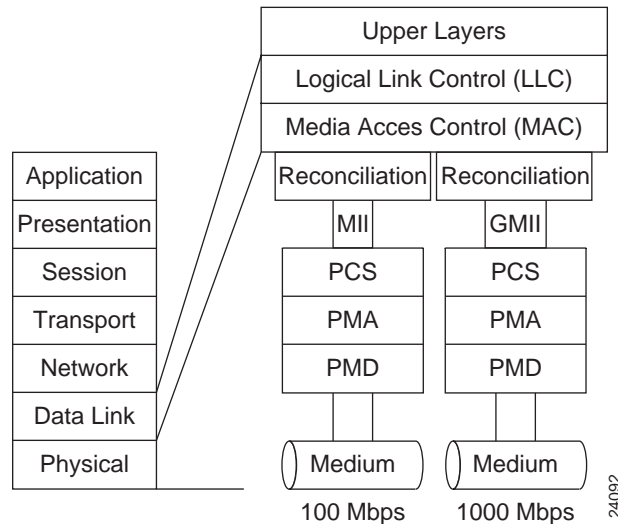


Leveraging these two technologies means that the standard can take advantage of the existing high-speed physical interface technology of Fibre Channel while maintaining the IEEE 802.3 Ethernet frame format, backward compatibility for installed media, and use of full-or half-duplex (via CSMA/CD).

A model of Gigabit Ethernet is shown in Figure 7-17.

24091

Figure 7-17 This diagram shows the architectural model of IEEE 802.3z Gigabit Ethernet. (Source: IEEE Media Access Control Parameters, Physical Layers, Repeater, and Management Parameters for 1000 Mbps Operation.)



The Physical Layer

The Gigabit Ethernet specification addresses three forms of transmission media: long-wave (LW) laser over single-mode and multimode fiber (to be known as 1000BaseLX), short-wave (SW) laser over multimode fiber (to be known as 1000BaseSX), and the 1000BaseCX medium, which allows for transmission over balanced shielded 150-ohm copper cable. The IEEE 802.3ab committee is examining the use of UTP cable for Gigabit Ethernet transmission (1000BaseT); that standard is expected sometime in 1999. The 1000BaseT draft standard will enable Gigabit Ethernet to extend to distances up to 100 meters over Category 5 UTP copper wiring, which constitutes the majority of the cabling inside buildings.

The Fibre Channel PMD specification currently allows for 1.062 gigabaud signaling in full-duplex. Gigabit Ethernet will increase this signaling rate to 1.25 Gbps. The 8B/10B encoding (to be discussed later) allows a data transmission rate of 1000 Mbps. The current connector type for Fibre Channel, and therefore for Gigabit Ethernet, is the SC connector for both single-mode and multimode fiber. The Gigabit Ethernet specification calls for media support for multimode fiber-optic cable, single-mode fiber-optic cable, and a special balanced shielded 150-ohm copper cable.

Long-Wave and Short-Wave Lasers over Fiber-Optic Media

Two standards of laser will be supported over fiber: 1000BaseSX (short-wave laser) and 1000BaseLX (long-wave laser). Short-wave and long-wave lasers will be supported over multimode fiber. There are two available types of multimode fiber: 62.5-millimeter and 50-millimeter diameter fibers. Long-wave lasers will be used for single-mode fiber because this fiber is optimized for long-wave laser transmission. There is no support for short-wave laser over single-mode fiber.

The key differences between the use of long-wave and short-wave laser technologies are cost and distance. Lasers over fiber-optic cable take advantage of variations in attenuation in a cable. At different wavelengths, "dips" in attenuation will be found over the cable. Short-wave and long-wave lasers take advantage of those dips and illuminate the cable at different wavelengths. Short-wave lasers are readily available because variations of these lasers are used in compact disc technology.

Long-wave lasers take advantage of attenuation dips at longer wavelengths in the cable. The net result is that short-wave lasers will cost less, but transverse a shorter distance. In contrast, long-wave lasers will be more expensive but will transverse longer distances.

Single-mode fiber has traditionally been used in networking cable plants to achieve long distances. In Ethernet, for example, single-mode cable ranges reach up to 10 kilometers. Single-mode fiber, using a 9-micron core and 1300-nanometer laser, demonstrate the highest-distance technology. The small core and lower-energy laser elongate the wavelength of the laser and allow it to transverse greater distances. This enables single-mode fiber to reach the greatest distances of all media with the least reduction in noise.

Gigabit Ethernet will be supported over two types of multimode fiber: 62.5-micron and 50-micron diameter fibers. The 62.5-micron fiber is typically seen in vertical campus and building cable plants and has been used for Ethernet, Fast Ethernet, and FDDI backbone traffic. This type of fiber, however, has a lower modal bandwidth (the ability of the cable to transmit light), especially with short-wave lasers. This means that short-wave lasers over 62.5-micron fibers will be able to transverse shorter distances than long-wave lasers. The 50-micron fiber has significantly better modal bandwidth characteristics and will be able to transverse longer distances with short-wave lasers relative to 62.5-micron fiber.

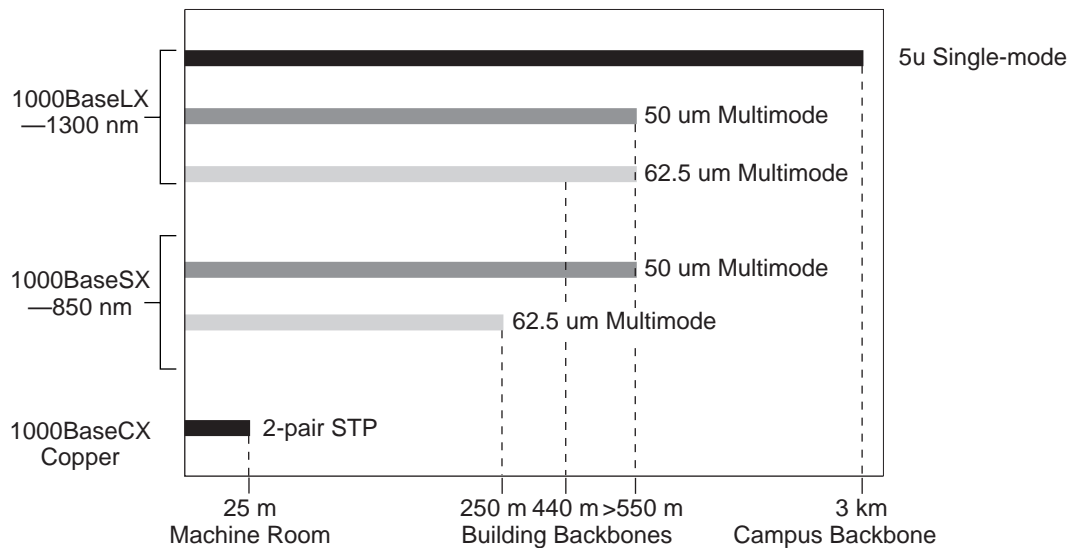
150-Ohm Balanced Shielded Copper Cable (1000BaseCX)

For shorter cable runs (of 25 meters or less), Gigabit Ethernet will allow transmission over a special balanced 150-ohm cable. This is a new type of shielded cable; it is not UTP or IBM Type I or II. In order to minimize safety and interference concerns caused by voltage differences, transmitters and receivers will share a common ground. The return loss for each connector is limited to 20 dB to minimize transmission distortions. The connector type for 1000BaseCX will be a DB-9 connector. A new connector is being developed by Aero-Marine Products called the HSSDC (High-Speed Serial Data Connector), which will be included in the next revision of the draft.

The application for this type of cabling will be short-haul data-center interconnections and inter-or intrarack connections. Because of the distance limitation of 25 meters, this cable will not work for interconnecting data centers to riser closets.

The distances for the media supported under the IEEE 802.3z standard are shown in Figure 7-18.

Figure 7-18 The Gigabit Ethernet draft specifies these distance specifications for Gigabit Ethernet.



NB: Distances are draft targets; may change in final specification.

24093

The Serializer/Deserializer

The physical media attachment (PMA) sublayer for Gigabit Ethernet is identical to the PMA for Fibre Channel. The serializer/deserializer is responsible for supporting multiple encoding schemes and allowing presentation of those encoding schemes to the upper layers. Data entering the PHY will enter through the PMD and will need to support the encoding scheme appropriate to that medium. The encoding scheme for Fibre Channel is 8B/10B, designed specifically for fiber-optic cable transmission. Gigabit Ethernet will use a similar encoding scheme. The difference between Fibre Channel and Gigabit Ethernet, however, is that Fibre Channel utilizes a 1.062 gigabaud signaling, whereas Gigabit Ethernet will utilize 1.25 gigabaud signaling. A different encoding scheme will be required for transmission over UTP. This encoding will be performed by the UTP or 1000BaseT PHY.

8B/10B Encoding

The Fibre Channel FC1 layer describes the synchronization and the 8B/10B encoding scheme. FC1 defines the transmission protocol, including serial encoding and decoding to and from the physical layer, special characters, and error control. Gigabit Ethernet will use the same encoding/decoding as specified in the FC1 layer of Fibre Channel. The scheme used is the 8B/10B encoding. This is similar to the 4B/5B encoding used in FDDI; however, 4B/5B encoding was rejected for Fibre Channel because it lacks DC balance. The lack of DC balance can potentially result in data-dependent heating of lasers due to a transmitter sending more 1s than 0s, resulting in higher error rates.

Encoding data transmitted at high speeds provides some advantages:

- Encoding limits the effective transmission characteristics, such as ratio of 1s to 0s, on the error rate.
- Bit-level clock recovery of the receiver can be greatly improved by using data encoding.

- Encoding increases the possibility that the receiving station can detect and correct transmission or reception errors.
- Encoding can help distinguish data bits from control bits.

All these features have been incorporated into the Fibre Channel FC1 specification.

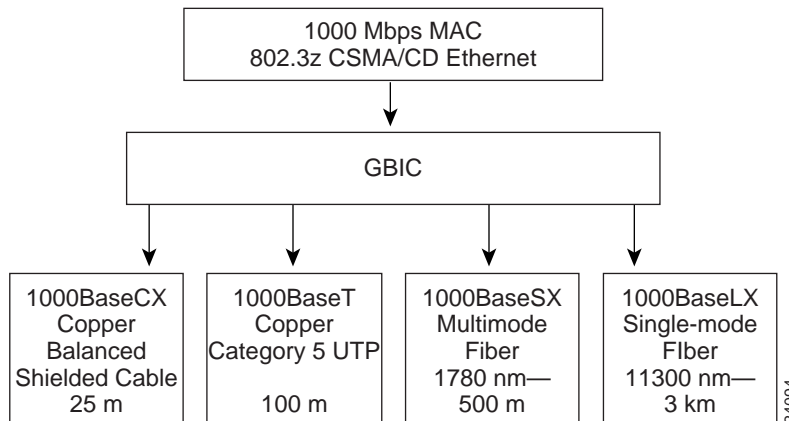
In Gigabit Ethernet, the FC1 layer will take decoded data from the FC2 layer, 8 bits at a time from the reconciliation sublayer (RS), which “bridges” the Fibre Channel physical interface to the IEEE 802.3 Ethernet upper layers. Encoding takes place via an 8-bit to 10-bit character mapping. Decoded data comprises 8 bits with a control variable. This information is, in turn, encoded into a 10-bit transmission character.

Encoding is accomplished by providing each transmission character with a name, denoted as Zxx.y. Z is the control variable that can have two values: D for *data* and K for *special character*. The xx designation is the decimal value of the binary number composed of a subset of the decoded bits. The y designation is the decimal value of the binary number of remaining decoded bits. This implies that there are 256 possibilities for data (D designation) and 256 possibilities for special characters (K designation). However, only 12 Kxx.y values are valid transmission characters in Fibre Channel. When data is received, the transmission character is decoded into one of the 256 8-bit combinations.

Gigabit Ethernet Interface Carrier (GBIC)

The GBIC interface allows network managers to configure each Gigabit port on a port-by-port basis for short-wave and long-wave lasers, as well as for copper physical interfaces. This configuration allows switch vendors to build a single physical switch or switch module that the customer can configure for the required laser/fiber topology. As stated earlier, Gigabit Ethernet initially supports three key media: short-wave laser, long-wave laser, and short copper. In addition, fiber-optic cable comes in three types: multimode (62.5 um), multimode (50 um) and single-mode. A diagram for the GBIC function is provided in Figure 7-19.

Figure 7-19 This diagram displays the function of the GBIC interface.



In contrast, Gigabit Ethernet switches without GBICs either cannot support other lasers or need to be ordered customized to the laser types required. Note that the IEEE 802.3z committee provides the only GBIC specification. The 802.3ab committee may provide for GBICs as well.

The MAC Layer

The MAC layer of Gigabit Ethernet is similar to those of standard Ethernet and Fast Ethernet. The MAC layer of Gigabit Ethernet will support both full-duplex and half-duplex transmission. The characteristics of Ethernet, such as collision detection, maximum network diameter, repeater rules, and so forth, will be the same for Gigabit Ethernet. Support for half-duplex Ethernet adds frame bursting and carrier extension, two functions not found in Ethernet and Fast Ethernet.

Half-Duplex Transmission

For half-duplex transmission, CSMA/CD will be utilized to ensure that stations can communicate over a single wire and that collision recovery can take place. Implementation of CSMA/CD for Gigabit Ethernet will be the same as for Ethernet and Fast Ethernet and will allow the creation of shared Gigabit Ethernet via hubs or half-duplex point-to-point connections.

Because the CSMA/CD protocol is delay sensitive, a bit-budget per-collision domain must be created. Note that delay sensitivity is of concern only when CSMA/CD is utilized; full-duplex operation has no such concerns. A collision domain is defined by the time of a valid minimum-length frame transmission. This transmission, in turn, governs the maximum separation between two end stations on a shared segment. As the speed of network operation increases, the minimum frame transmission time decreases, as does the maximum diameter of a collision domain. The bit budget of a collision domain is made up of the maximum signal delay time of the various networking components, such as repeaters, the MAC layer of the station, and the medium itself.

Acceleration of Ethernet to Gigabit speeds has created some challenges in terms of the implementation of CSMA/CD. At speeds greater than 100 Mbps, smaller packet sizes are smaller than the length of the slot-time in bits. (*Slot-time* is defined as the unit of time for Ethernet MAC to handle collisions.) To remedy the slot-time problem, carrier extension has been added to the Ethernet specification. Carrier extension adds bits to the frame until the frame meets the minimum slot-time required. In this way, the smaller packet sizes can coincide with the minimum slot-time and allow seamless operation with current Ethernet CSMA/CD.

Another change to the Ethernet specification is the addition of frame bursting. Frame bursting is an optional feature in which, in a CSMA/CD environment, an end station can transmit a burst of frames over the wire without having to relinquish control. Other stations on the wire defer to the burst transmission as long as there is no idle time on the wire. The transmitting station that is bursting onto the wire fills the interframe interval with extension bits such that the wire never appears free to any other end station.

It is important to point out that the issues surrounding half-duplex Gigabit Ethernet, such as frame size inefficiency (which in turn drives the need for carrier extension) as well as the signal round-trip time at Gigabit speeds, indicate that, in reality, half-duplex is not effective for Gigabit Ethernet.

IEEE 802.3x Full-Duplex Transmission

Full-duplex provides the means of transmitting and receiving simultaneously on a single wire. Full-duplex is typically used between two endpoints, such as between switches, between switches and servers, between switches and routers, and so on. Full-duplex has allowed bandwidth on Ethernet and Fast Ethernet networks to be easily and cost-effectively doubled from 10 Mbps to 20 Mbps and 100 Mbps to 200 Mbps, respectively. By using features such as Fast EtherChannel, “bundles” of Fast Ethernet connections can be grouped together to increase bandwidth up to 400%.

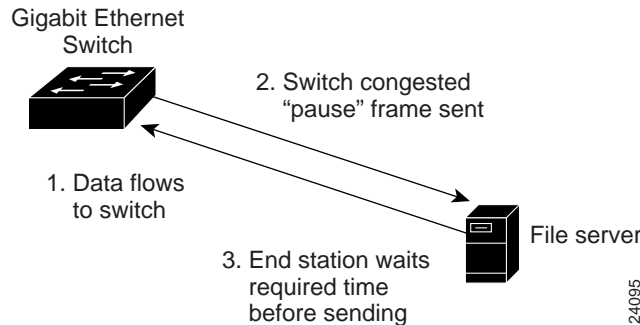
Full-duplex transmission will be utilized in Gigabit Ethernet to increase aggregate bandwidth from 1 Gbps to 2 Gbps for point-to-point links as well as to increase the distances possible for the particular media. Additionally, Gigabit EtherChannel “bundles” will allow creation of 8 Gbps connecting between switches. The use of full-duplex Ethernet eliminates collisions on the wire;

therefore, CSMA/CD need not be utilized as a flow control or access medium. However, a full-duplex flow control method has been put forward in the standards committee with flow control as an optional clause. That standard is referred to as IEEE 802.3x; it formalizes full-duplex technology and is expected to be supported in future Gigabit Ethernet products. Because of the volume of full-duplex 100-Mbps network interface cards (NICs), it is unlikely that this standard will realistically apply to Fast Ethernet.

Optional 802.3x Flow Control

The optional flow control mechanism is set up between the two stations on the point-to-point link. If the receiving station at the end becomes congested, it can send back a frame called a *pause frame* to the source at the opposite end of the connection; the pause frame instructs that station to stop sending packets for a specific period of time. The sending station waits the requested time before sending more data. The receiving station can also send a frame back to the source with a time-to-wait of zero and instruct the source to begin sending data again. Figure 7-20 shows how IEEE 802.3x will work.

Figure 7-20 This figure presents an overview of the operation of the IEEE 802.3 flow control process.



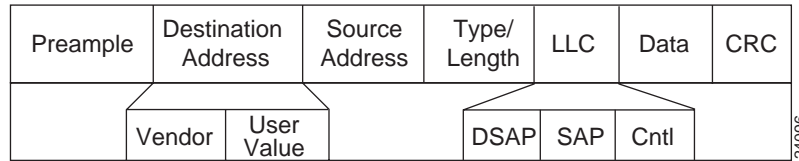
This flow control mechanism was developed to match the sending and receiving device throughput. For example, a server can transmit to a client at a rate of 3000 pps. The client, however, may not be able to accept packets at that rate because of CPU interrupts, excessive network broadcasts, or multitasking within the system. In this example, the client would send a pause frame and request that the server hold transmission for a certain period. This mechanism, although separate from the IEEE 802.3z work, will complement Gigabit Ethernet by allowing Gigabit devices to participate in this flow-control mechanism.

The Logical Link Layer

Gigabit Ethernet has been designed to adhere to the standard Ethernet frame format, which maintains compatibility with the installed base of Ethernet and Fast Ethernet products and requires no frame translation. Figure 7-21 describes the IEEE 802.3/Ethernet frame format.

The original Xerox specification identified a Type field, which was utilized for protocol identification. The IEEE 802.3 specification eliminated the Type field, replacing it with the Length field. The Length field is used to identify the length in bytes of the data field. The protocol type in 802.3 frames are left to the data portion of the packet. The LLC is responsible for providing services to the network layer regardless of media type, such as FDDI, Ethernet, Token Ring, and so on.

Figure 7-21 This figure shows the fields of the IEEE 802.3/Ethernet frame format.



In order to communicate between the MAC layer and the upper layers of the protocol stack, the Logical Link Control (LLC) layer of LLC protocol data units (or PDUs) makes use of three variable addresses to determine access into the upper layers via the LLC/PDU. Those addresses are the destination service access point (DSAP), source service access point (SSAP), and control variable. The DSAP address specifies a unique identifier within the station that provides protocol information for the upper layer. The SSAP provides the same information for the source address.

The LLC defines service access for protocols that conform to the Open System Interconnection (OSI) model for network protocols. Unfortunately, many protocols do not obey the rules for those layers. Therefore, additional information must be added to the LLC to provide information regarding those protocols. Protocols falling into this category include Internet Protocol (IP) and Internetwork Packet Exchange (IPX).

The method used to provide this additional protocol information is called a Subnetwork Access Protocol (SNAP) frame. A SNAP encapsulation is indicated by the SSAP and DSAP addresses being set to 0xAA. This address indicates that a SNAP header follows. The SNAP header is 5 bytes long: The first 3 bytes consist of the organization code, which is assigned by the IEEE; the second 2 bytes use the Type value set from the original Ethernet specifications.

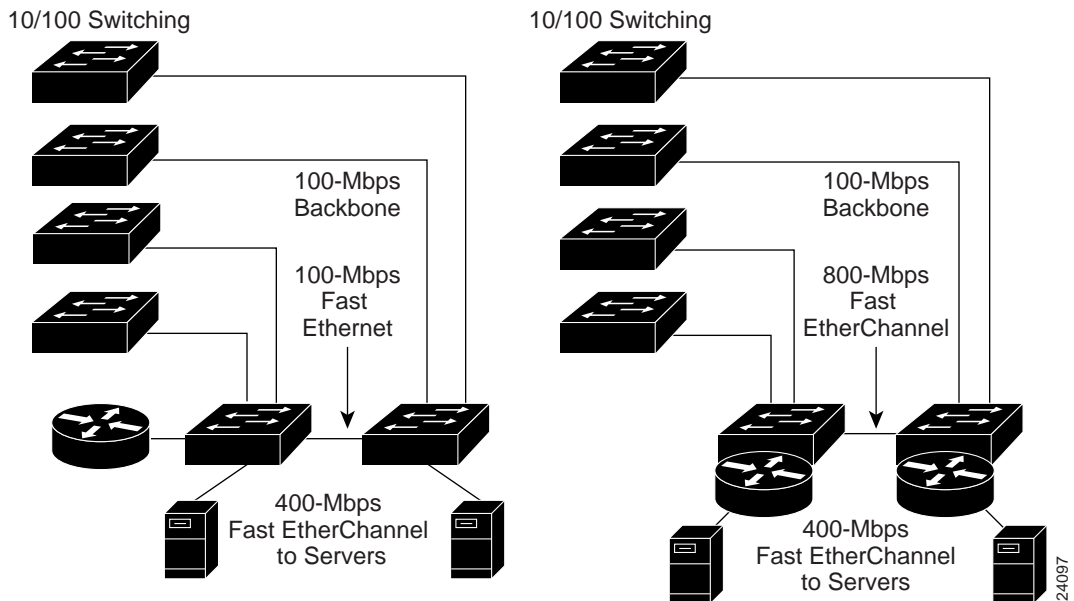
Migration to Gigabit Ethernet

Several means can be used to deploy Gigabit Ethernet to increase bandwidth and capacity within the network. First, Gigabit Ethernet can be used to improve Layer 2 performance. Here, the throughput of Gigabit Ethernet is used to eliminate Layer 2 bottlenecks.

Scaling Bandwidth with Fast EtherChannel and Gigabit EtherChannel

Bandwidth requirements within the network core and between the network core and the wiring closet have placed significant demands on the network. Fast EtherChannel allows multiple Fast Ethernet ports to be bundled together and seen logically by the switches as a fat pipe. Fast EtherChannel allows the bundling of up to four ports, for an aggregate bandwidth of 800 Mbps. With support from NIC manufacturers such as Sun Microsystems, Intel, SGI, Compaq, and Adaptec, Fast EtherChannel can now be provided directly to high-end file servers. Figure 7-22 provides a possible Fast EtherChannel topology.

Figure 7-22 EtherChannel allows the bundling of up to four ports, for an aggregate bandwidth of 800 Mbps.

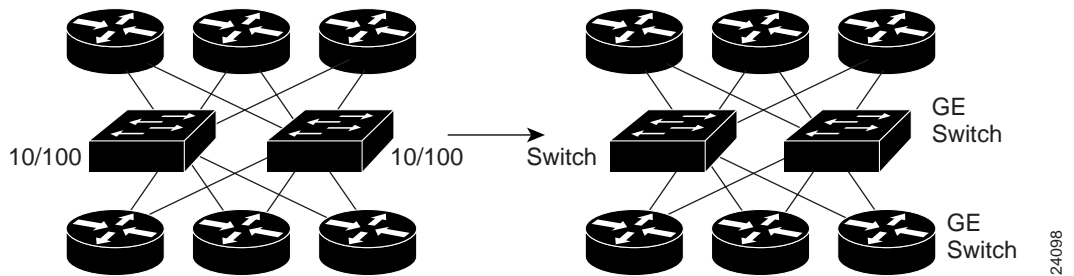


Scaling Router Backbones

Many large-scale networks use a meshed core of routers to form a redundant network backbone. This backbone typically consists of FDDI, Fast Ethernet, or ATM. However, as newer network designs heavily utilize switching with 100-Mbps links to these routers, a potential design bottleneck can be created. Although this is not currently a problem, the migration of services away from the workgroup and toward the enterprise can potentially lead to slower network performance.

The solution demonstrated in Figure 7-23 uses Gigabit Ethernet switches that provide aggregation between routers in a routed backbone. Gigabit Ethernet and Gigabit switching are used to improve speed and capacity between the routers. Gigabit Ethernet switches are placed between the routers for improved throughput performance. By implementing this design, a fast Layer 2 aggregation is utilized, creating a high-speed core.

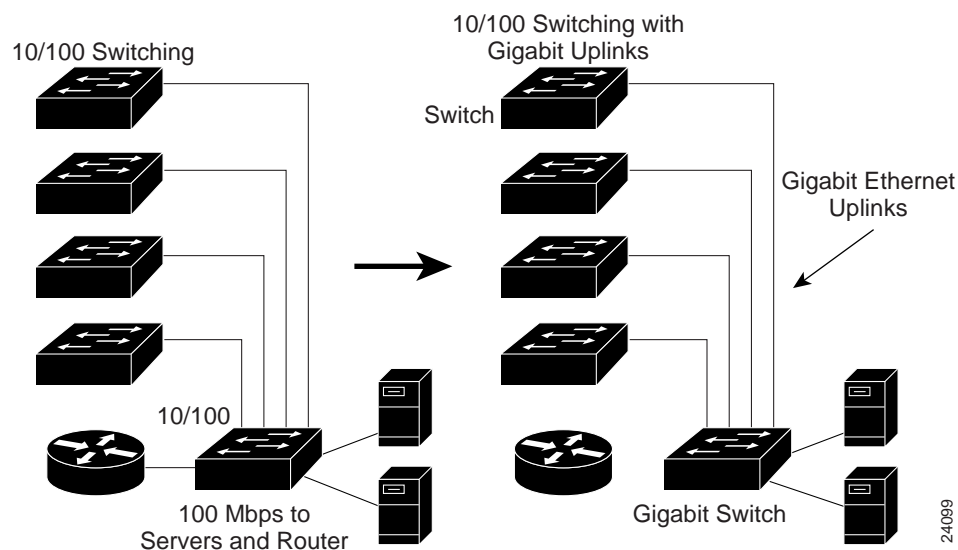
Figure 7-23 This design provides a scalable switching solution that increases throughput in a router backbone.



Scaling Wiring Closets

Gigabit Ethernet can also be used to aggregate traffic from wiring closets to the network core (see Figure 7-24). Gigabit Ethernet and Gigabit switching are used to aggregate traffic from multiple low-speed switches as a front end to the router. Low-speed switches can be connected either via Fast Ethernet or by a Gigabit Ethernet uplink while the switches provide dedicated 10-Mbps switching or group switching to individual users. The file servers are connected via Gigabit Ethernet for improved throughput performance. Keep in mind that as bandwidth requirements to the core or within the core increase, Gigabit EtherChannel can produce a fourfold increase in performance.

Figure 7-24 This design demonstrates the use of Gigabit Ethernet switching to improve data center applications.

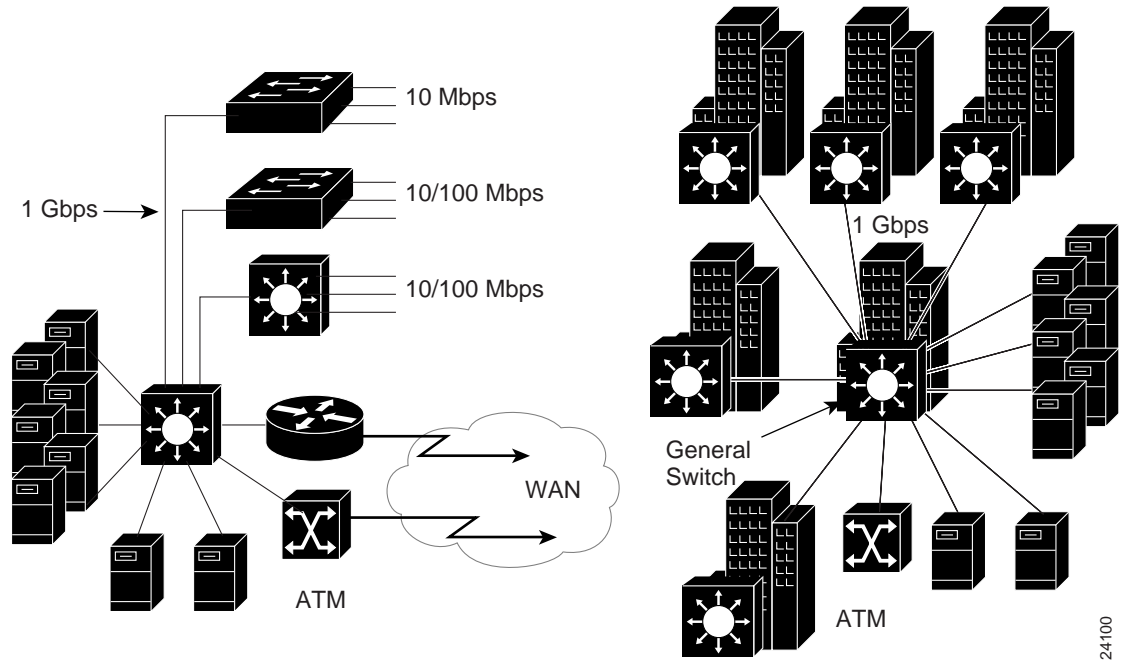


Gigabit Ethernet can also improve Layer 3 performance. This essentially means coupling Layer 2 performance with the benefits of Layer 3 routing. By using the switching paradigm as a road map, Gigabit switching and distributed Layer 3 services can improve the scalability and performance of campus intranets.

Gigabit Ethernet Campus Applications

The key application of Gigabit Ethernet is expected to be use in the building backbone for interconnection of wiring closets. A Gigabit multilayer switch in the building data center aggregates the building's traffic and provides connection to servers via Gigabit Ethernet or Fast Ethernet. WAN connectivity can be provided by traditional routers or via ATM switching. Gigabit Ethernet can also be used for connecting buildings on the campus to a central multilayer Gigabit switch located at the campus data center. Servers located at the campus data center are also connected to the Gigabit multilayer switch that provides connectivity to the entire campus. Once again, Gigabit EtherChannel can be utilized to significantly increase the bandwidth available within the campus backbone, to high-end wiring closets, or to high-end routers. Figure 7-25 illustrates potential multilayer Gigabit switching designs.

Figure 7-25 This design provides an example of a multilayer Gigabit switching environment.



24100