10Mbps Ethernet MAC and PHY Standards

Ethernet is strictly based on a layered OSI model. As a result, you can easily combine the Ethernet MAC with different PHYs. This section discusses the 10Mbps Ethernet MAC standards, and the four major baseband PHY specifications.

The Ethernet CSMA/CD MAC

The Ethernet MAC technology is called *carrier-sense multiple access with collision detection* (CSMA/CD). Figure 3.2 illustrates the CSMA/CD flow.

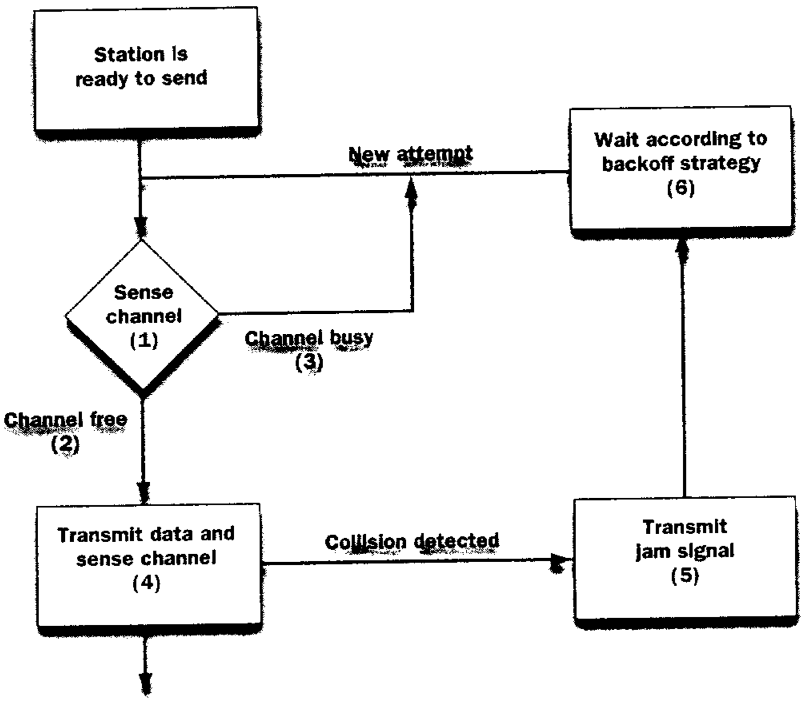

FIGURE 3.2 *A flow diagram illustrating the CSMA/CD MAC.*

CSMA/CD is the shorthand version for about seven different steps that make up an Ethernet transmission. Notice the analogy to human speech among multiple individuals:

1. A station wanting to transmit a frame of information has to ensure that no other nodes or stations are currently using the shared media, so the station listens to the channel first. (This is the *carrier sense* part, also known as "listen before talking.")

2. If the channel is quiet for a certain minimum period of time, called the interframe gap (IFG), the station may initiate a transmission ("talk if quiet").
3. If the channel is busy, it is monitored continuously until it becomes free for the minimum IFG time period. At this point, transmission begins (known as *multiple access*, or "wait for quiet before talking").
4. A collision (two stations transmitting on the cable at the same time) may occur when two or more stations listen while waiting to transmit, simultaneously determine that the channel is free, and begin transmitting at almost the same time. This event would lead to a collision and destroy both data frames. Ethernet continuously monitors the channel during transmission to detect collisions (*collision detection*, or "listen while talking").
5. If a station detects a collision during transmission, that transmission immediately stops. A jam signal is sent to the channel to guarantee that all other stations detect the collision and reject any corrupted data frame they may have been receiving (also part of the collision detection, or "one talker at a time").
6. After a waiting period (called a *backoff*) the stations that wish to transmit attempt to make a new transmission. A special random backoff algorithm (called binary exponential backoff, or BEB) determines a delay time that the different stations will have to wait before attempting to send their data again. Of course, another collision could occur after the first one, especially when many nodes are trying to obtain access at the same time. After 16 consecutive collisions for a given transmission attempt, the packet will be dropped. This can and does happen if the Ethernet channel is overutilized. This is also part of the multiple access method.
7. The sequence returns to step 1.

Ethernet uses frames of data to transmit the actual information, also known as *payload*, from source to destination. Like most other LANs in existence today, Ethernet transmits a frame of variable length. The length of the frame changes because the payload or data field can vary.

The original Ethernet frame specified by Digital, Intel, and Xerox is known as the DEC-Intel-Xerox Ethernet V2.0 frame, or just DIX or Ethernet II frame. The official IEEE Ethernet frame subsequently replaced it. The only difference is in the 2-byte frame type/length field. Figure 3.3 shows both frame types.

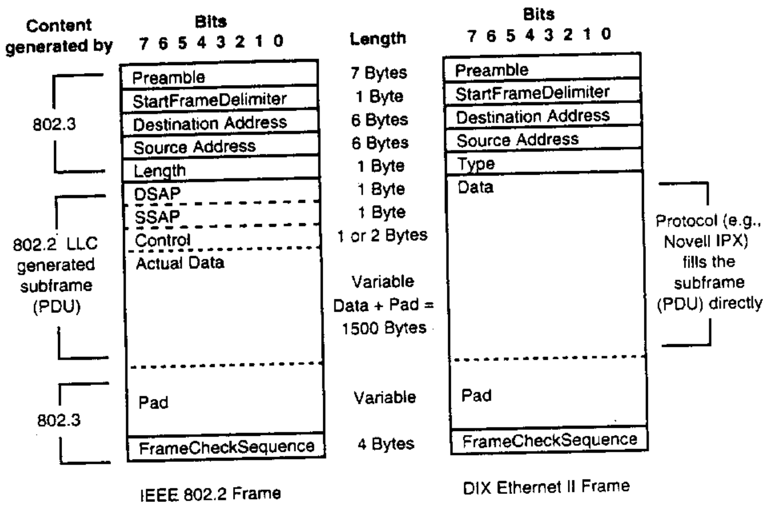

FIGURE 3.3 *The official IEEE and the older DIX Ethernet II frame differ in some respects.*

Let's look at the individual fields that comprise the Ethernet frame:

- The preamble is sent to allow the receiver to synchronize with the incoming transmission and locate the start of the frame. The preamble is a sequence of 01010101…, 7 bytes long.
- The start of frame delimiter (SFD) indicates that the MAC frame is about to commence. The SFD octet or byte is specified as 10101011.
- The source address denotes the sender. Each node has a unique address. The first three bytes of the address are called the block ID or organizationally unique identifier (OUI) and identify the manufacturer of the equipment. The IEEE assigns them. Intel, for example, is identified by the 00-AA-00 (hex) address, 3Com uses the 00-20-AF address, and Cisco uses 00-00-0C. The other three bytes are called the device ID and are assigned by each manufacturer. These are always unique. Three bytes of device ID allows for 16 million different and unique addresses: 2^24–1 or 16,777,215 to be exact. Some of the major Ethernet vendors have shipped more than 16 million Ethernet MAC devices and have started using new OUIs. Some of the references given in Appendix C, "Useful Web Links," show the complete list of vendor OUIs.
- The destination address specifies where the frame will be sent.

- The protocol type is a field in the original DIX V2.0 frame type. It specified which kind of Layer 3 protocol the data contained. A hexadecimal value of 08 00 indicates a TCP/IP packet, for example, and 81 37 indicates a Novell NetWare packet.
- The newer IEEE frame differs from the Ethernet II or DIX frame type in one key area. The IEEE frame is much more popular these days and no longer uses the protocol type field, where it has been replaced with the length field. This specifies the total length of the data that will be transmitted, which can vary from 0 to 1500 bytes. If the contents of this field are 0 to 1500 bytes, you can be sure that it is the length we are discussing. If the contents of the type length/type field are greater than 1500 bytes (for example, 8,137), we are talking about the old DIX frame type.
- The data field can vary from 0 to 1500 bytes in length. The data field is also known as the protocol data unit (PDU). We discuss the contents of the data field in more detail later.
- If the actual data is less than a minimum length required, the MAC will add a variable pad to maintain a minimum total frame size of 64 bytes. If the data is longer than the maximum frame allows, Layer 3 will typically split the packet into more than one frame.
- Finally, a frame check sequence (FCS) ensures accurate transmission. The cyclical redundancy check method (CRC) checks for invalid frames. This value is calculated from the rest of the packet's data and sent along in the FCS frame. The receiving station performs the same calculations and compares its results with the FCS transmitted with the packet. If the results are different, the packet is rejected.
- The maximum total frame size is 1518 bytes. (The frame officially starts with the source address.)

There are a total of four different frame types. The DIX Ethernet II frame discussed previously, plus three different versions of the IEEE frame type: the Ethernet 802.3 frame (also known as Novell Raw LLC), the 802.2 frame, and the 802.2 with SNAP frame. All look slightly different. Figure 3.4 shows the contents of the 1500-byte data field for the four different frame types. The following list details the three different versions of the IEEE frame type:

- The Novell Raw frame format is used only by Novell's NetWare operating system, and uses the 802.3 frame shell type without adding an IEEE 802.2 LLC header within the data field. In this case, NetWare adds its own IPX information. This type of frame is called "Novell Raw" because it encapsulates the IPX data in raw form without any 802.2 LLC information. Because the frame doesn't use the IEEE 802.2 LLC information, this frame type is actually Novell-proprietary and not 802.3 compliant.

- For the 802.2 frame type, the data field contains the 802.2 LLC-embedded information. The IEEE frame contains the protocol type information within the LLC subframe. Three fields are located at the beginning of the data field. The DSAP field, a 1-byte SSAP field, and a 1-byte Control field. The IEEE frame assigns Service Access Point numbers; among those currently defined are E0 for Novell, F0 for NetBIOS, 06 for TCP/IP, and AA for the Subnetwork Access Protocol (SNAP).
- The 802.2 with SNAP frame is almost identical to the 802.2 frame, except it supports more than 256 protocol types by adding a 5-byte Protocol Identification field. On any SNAP packet, both the DSAP and SSAP fields are set to AA, and the Control field is set to 03. The 5-byte protocol field then follows.

Table 3.2 lists all the relevant Ethernet MAC frame parameters. Most 10Mbps Ethernet/802.3 MAC parameters are listed in bit times. The Ethernet MAC is inherently scalable. All the parameters can be measured in terms of the time taken to transmit 1 bit of data, referred to as *bit times*. Note that the actual speed of Ethernet (10Mbps) is not mentioned in the MAC specification at all. This makes it very easy to run Ethernet at different speeds.

TABLE 3.2 KEY 10MBPS ETHERNET MAC AND FRAME PARAMETERS, SPECIFIED IN BIT TIMES

Parameter	Value (Bit Times)
`Time`	512 bit times
`MinInterFrameGap`	96 bit times (or 9.6 μs)
`AttemptLimit`	16 (tries)
`BackoffLimit`	10 (exponent)
`JamSize`	32 bits
`MaxFrameSize`	12144 bits (1518 bytes)
`MinFrameSize`	512 bits (64 bytes)
`AddressSize`	48 bits

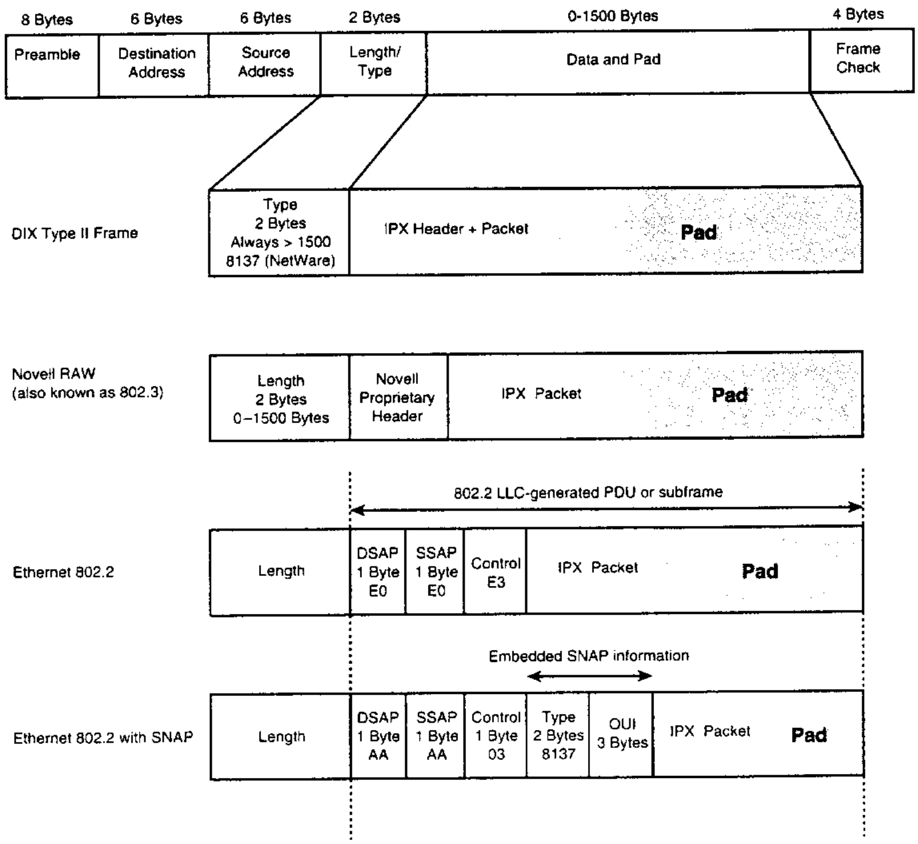

FIGURE 3.4 *The data field of the four different frame Ethernet types in use.*

Calculating the time to transmit 1 bit for 10Mbps Ethernet transmission becomes very easy:

$$1 \text{ bit-time} = \frac{1 \text{ bit}}{10\text{Mhz}} = 0.1\mu s \text{ or } 100\text{ns}$$

For 1Mbps Ethernet/StarLAN, the frame looks exactly the same. The only thing that changes is the bit time. For StarLAN, the bit time is 1/1MHz = 1μs, or 1000ns. Fast Ethernet works exactly the same way: The frame is identical again, but the bit time is reduced to 1/10, or 10ns.

Ethernet PHYs

This section looks at the different PHY implementations for 10Mbps Ethernet (see Figure 3.5). There are officially five ways to transmit 10Mbps Ethernet:

- 10BASE5 is the original thick Ethernet coaxial cable standard, dating back to the early 1970s.
- 10BASE2, also known as thin Ethernet, was added in the early 1980s and uses a thinner coaxial cable.
- In 1990, Ethernet over unshielded twisted-pair, known as 10BASE-T, was standardized.
- 10BASE-F, although less well known, is very important because it utilizes fiber cabling to carry Ethernet over extended distances. The physical layers mentioned so far all use baseband transmission methods, which means the entire frequency spectrum transmits the data.
- 10BROAD36 is different from all the other Ethernet PHY standards in that it uses broadband transmission technology to transmit. This allows different channels to communicate simultaneously on the same cable. 10BROAD36 is far less popular and no similar 100Mbps broadband PHY exists, so we will not discuss 10BROAD36 in this book.

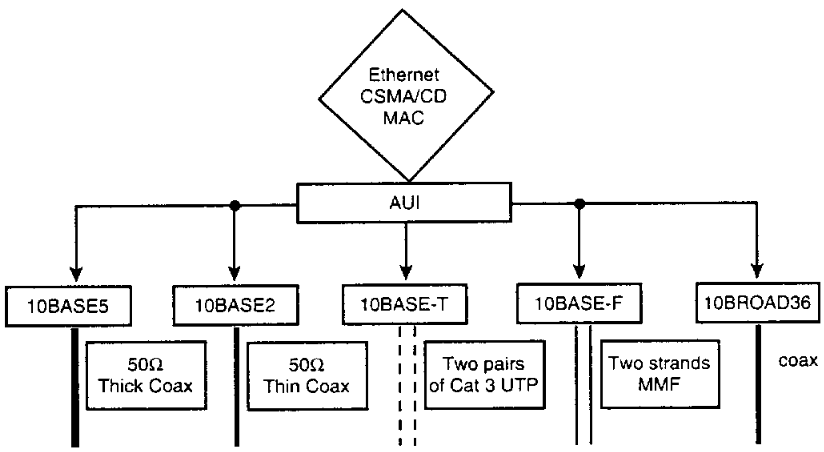

FIGURE 3.5 *The different Ethernet/802.3 PHYs.*

TABLE 3.5 SUMMARY OF DIFFERENT 10MBPS ETHERNET/802.3 PHY STANDARDS

Parameter	10BASE5	10BASE2	10BASE-F	10BASE-T
Maximum segment length	500m	185m	Varies from 400m to 2000m	100m
Topology	Bus	Bus	Star	Star
Medium	50-Ω thick coax	50-Ω thin coax	Multimode fiber (a few vendors provide for single-mode equipment, too)	100-Ω UTP
Connector	NICDB15	BNC	ST[1]	RJ-45
Medium attachment	MAU bolted to coax	External or on NIC	External or on NIC	External or on NIC
Stations/cable segment	100	30	N/A	2 (NIC and repeater)
Maximum segments	5	5	5	5

1 *ST is the official connector, but SC and MIC are actually more popular.*

Half- and Full-Duplex Ethernet

Regular Ethernet is a shared-media access method. All networks that utilize a MAC are based on a shared-media operation (hence the term *media access*, which implies that the transmission media are shared and access is negotiated and not always given).

By definition, a shared-media LAN transmission method also implies half-duplex operation. *Half-duplex* means that a station is either transmitting or receiving, but not both, at the same time. That's because in CSMA/CD, a station has to listen to see whether the channel is available, and only if it is can a station start transmitting. When one station is transmitting data, all others have to listen. Therefore, it is an either-or situation for all stations on the LAN. This method of operation was very efficient for the coaxial cable on which Ethernet grew up. Running a single coaxial cable throughout an entire office and providing everyone an access opportunity to the cable every few milliseconds was very efficient. Sharing the transmission media also brought with it collisions, something that has given Ethernet a bit of a bad name. The truth is that collisions are a very effective and efficient method of preventing overload.

Full-duplex, on the other hand, means that a station is simultaneously transmitting *and* receiving. In the early 1990s, many events occurred that have made Full-Duplex Ethernet a reality:

- The introduction of 10BASE-T wiring offered the capability for separate transmit and receive data paths. Before the arrival of 10BASE-T, coaxial cable didn't offer this capability; only one electrical (coax) wire made simultaneous transmission and reception impossible.

Note

Technically speaking, a single cable can carry separate transmit and receive data streams in two directions. This is known as dual-duplex and requires a specially designed physical layer. The electronic circuitry required to do this is very complex, and typically requires a sophisticated DSP. Until recently, the cost and complexity associated with DSP technology was considered prohibitive. 1000BASE-T, however, will utilize simultaneous transmit and receive on four pairs of UTP. With 10BASE-T, the existing physical layer could be kept for full-duplex transmission.

- The emergence of multiport Ethernet bridges or switches meant that the physical media were no longer being shared by multiple users but were increasingly being used to connect two switches or a switch and a NIC together in a point-to-point manner.

In 1992, Kalpana seized on this opportunity so that it could effectively double the speed of Ethernet through full-duplex transmission. Kalpana started working with several other industry vendors to establish a de facto industry standard for full-duplex Ethernet over UTP wire. Kalpana proposed the following scheme:

- One pair of UTP wire (or one fiber strand) would be used exclusively for transmission and one for reception of data.
- A new kind of multiport bridge or switch would be required at both ends of the wire.

The Ethernet MAC algorithm would be modified as follows:

- No carrier-sense (CRS) would be required, because a cable pair would now be dedicated for both transmission and reception.
- Similarly, no collision-detection (CDT) jam or exponential backoff are required because collisions are no longer present: They happen in a multiple-user segment only.

Essentially, *full-duplex* Ethernet means running Ethernet without the CSMA/CD MAC in operation (refer to Figure 3.2, without steps 1, 2, 3, 5, and 6). Ethernet frames are transmitted and received simultaneously on two pairs of UTP or fiber at any time. The only possible problem is one of data overflow. We can deal with this issue through flow control. Kalpana proposed that artificial collisions be generated by the receiving station if an overload condition exists.

Figure 3.7 illustrates the evolution of Ethernet from a shared-media coaxial-cable LAN with collisions to a dedicated media full-duplex technology.

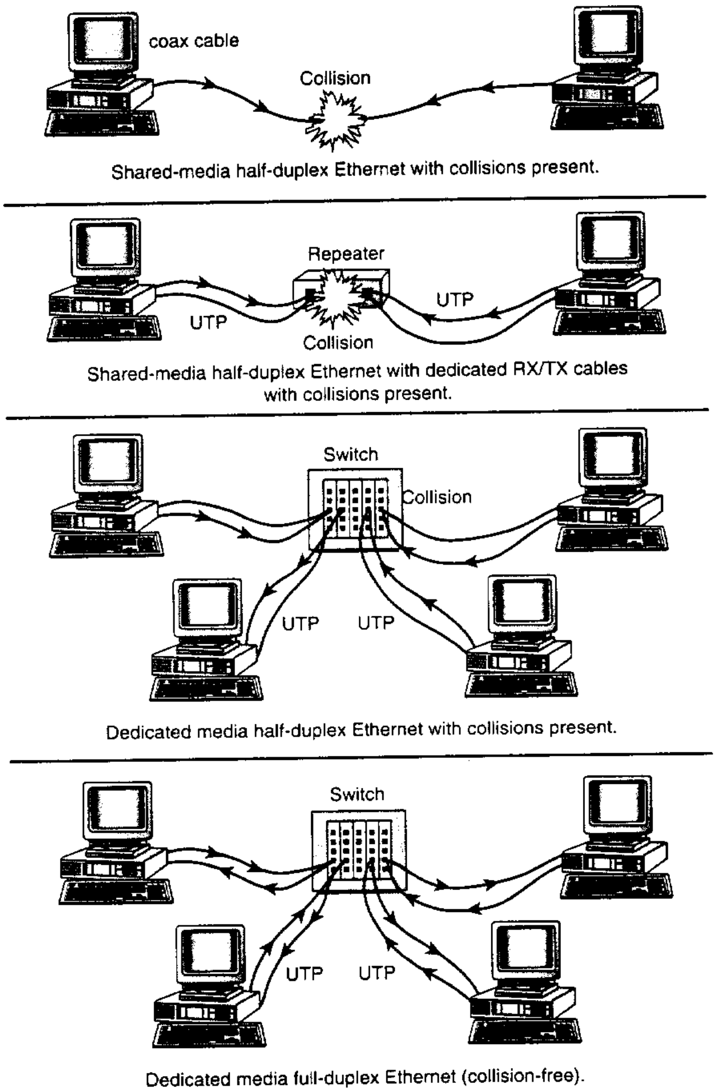

FIGURE 3.7 *The evolution of Ethernet from shared-media, half-duplex to dedicated media full-duplex.*

Note

A two-user switched connection is a prerequisite to Full-Duplex Ethernet because full-duplex requires a point-to-point connection with only two stations present. Note that Switched Ethernet does not automatically imply full-duplex operation.

Full-duplex has played an increasingly important role with the introduction of Fast and Gigabit Ethernet. In 1997, the 802.3X full–duplex/flow control standard was approved. (Chapter 4 covers this in more detail.) Some people in the IEEE even advocated making Gigabit Ethernet a full-duplex–only technology, which means dropping the all-too-familiar CSMA/CD MAC concept altogether. In the end, the half-duplex followers prevailed, and CSMA/CD operation was included in the standard. It is doubtful whether we will ever see any half-duplex shared Gigabit Ethernet equipment in the marketplace, however.

A trend is currently in place to move from repeated half-duplex to Switched full-duplex Ethernet. With the advent of Gigabit Ethernet in particular, it looks like CSMA/CD and collisions are heading for extinction. What will live on is the language of Ethernet, namely the 802.3 frame format. Our well-known Ethernet frames will be zooming along at 1, 10, 100, 1000, and someday 10000Mbps.

Gigabit Ethernet

In 1995, Fast Ethernet seemed like such an ingenious, yet simple, idea. 100BASE-T became an overnight success, and it was only a matter of time before the frequency would be increased another order of magnitude. Work started on Gigabit shortly after the Fast Ethernet standard had been ratified. Three years later, on June 25, 1998, the IEEE 802.3z standard was officially adopted.

If you have read this chapter's information on Fast Ethernet, we should just ask you to go back and read everything with an extra zero added everywhere. Gigabit Ethernet is that simple. Some differences arise in the physical layers, network design, and minimum frame size that we cover here and in Chapter 6. For the most part, however, Gigabit Ethernet is just supercharged Fast Ethernet.

The IEEE 802.3z standard includes the Gigabit Ethernet MAC, as well as three physical layers that use the 8B/10B encoding originally developed as part of the ANSI Fibre Channel technology.

Essentially, the IEEE engineers bolted the existing Fibre Channel PHY to the Ethernet MAC running at 10 times the speed of the Fast Ethernet MAC. The 802.3z standard encompasses two fiber physical layer standards, 1000BASE-LX and 1000BASE-SX as well as one copper PHY, 1000BASE-CX. The copper-based PHY was also inherited from the Fibre Channel standard and was included to enable cost-effective and quick cross-connects. A fourth PHY, called 1000BASE-T, is still under development.

Note

The ANSI X3T11 committee is responsible for the Fibre Channel technology. ANSI decided to spell Fibre with an re rather than an er to differentiate it from the fiber-optic cable on which it runs.

The IEEE has also defined a Gigabit MII (GMII), which is similar to the Fast Ethernet MII and connects Gigabit MAC and PHY. The GMII is only an electrical interface specification, and unlike the Fast Ethernet MII, the specification does not include a connector. The biggest benefit of the GMII is that it allows circuit designers to use existing Fibre Channel PHY chips and will allow for easy redesign of existing 1000BASE-LX equipment to accommodate future PHY transceivers.

Some people in the IEEE wanted to make Gigabit Ethernet a full-duplex–only technology; others wanted to preserve the classic CSMA/CD algorithm and the half-duplex shared operation. The reason for preserving the CSMA/CD part of

Ethernet was twofold. First, many vendors didn't want to redesign their Ethernet MAC chips completely. (Running it at 10 times the clock frequency didn't take a complete redesign.) Second, some IEEE members just wanted to preserve the 25-year heritage of CSMA/CD itself.

Figure 3.10 shows all the different Gigabit Ethernet components.

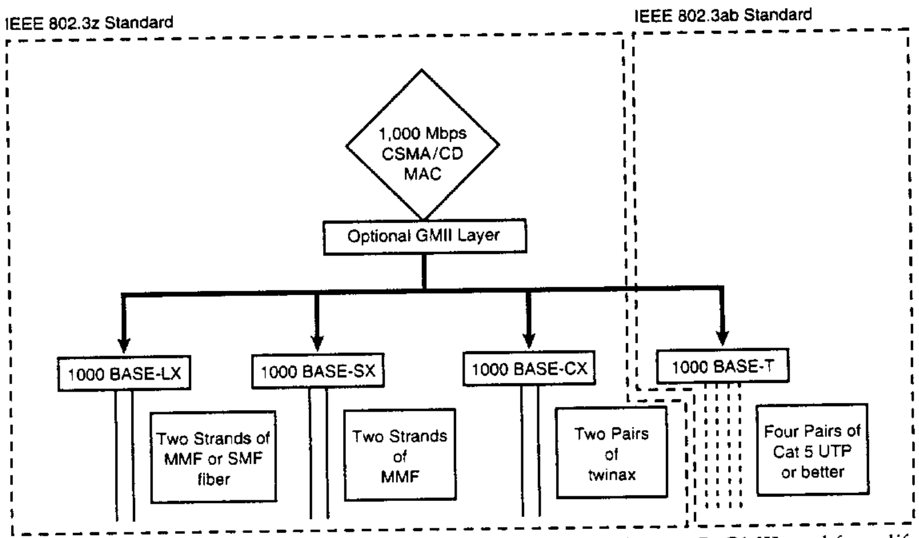

FIGURE 3.10 *A Gigabit Ethernet block diagram showing the MAC, GMII, and four different physical layers.*

100BASE-X refers to the original Fast Ethernet MAC and FX/TX PHYs. Similarly, 1000BASE-X refers to the 1000Mbps MAC and the LX/SX/CX transceiver technology.

The Gigabit Ethernet MAC

The main points of the Gigabit Ethernet MAC are the following:

- Gigabit Ethernet uses the official 802.3 frame format, identical to that of 10Mbps and 100Mbps Ethernet (refer to Figure 3.3).
- Like 10Mbps and 100Mbps Ethernet, Gigabit Ethernet can operate in both half- and full-duplex mode.
- The 1000Mbps MAC uses the original Ethernet MAC operating at 10 times the speed. Running the Ethernet MAC at Gigabit speeds has created some challenges in terms of the implementation of CSMA/CD. To make CSMA/CD work at 1GHz, a minor modification was required. The slot time has been increased to 512 bytes, as opposed to 64 bytes for 10

and 100Mbps Ethernet. The slot time, in essence, is the allocated time during which the complete frame needs to be transmitted. During the slot time, the transmitter retains control of the media. If the transmitted frame is smaller than 512 bytes, an extra carrier extension is added. The carrier extension concept is similar to the PAD that is added to the end of the data field within the frame. The carrier extension is added at the end of the completed frame to make it meet the new slot time of 512 bytes.

- IEEE 802.3z specification defines both half- and full-duplex operation. When very small frames are transmitted in a half-duplex environment, many carrier extension bits will be added. This makes Gigabit Ethernet very inefficient because large amounts of useless carrier extension bits are transmitted. Assume, for example, that we wanted to transmit only 64-byte minimum size frames. The carrier extension would add 438 bytes of carrier extension to meet the spec of a 512-byte slot time. To calculate the overall efficiency, we still need to add the interframe gap (IFG) overhead of 12 bytes: 64/(512+12) = 12% efficiency, or 122Mbps. This is only marginally better than 100BASE-T! Small frames are quite common, so this inefficiency for small frame sizes needed to be addressed. The IEEE 802.3z Gigabit MAC includes a feature called *burst mode*. In this case, a station may continuously transmit multiple smaller frames, up to a maximum of 8192 bytes worth of data. The interframe intervals will also be filled with carrier extension so that the wire never appears free to any other stations during the burst cycle.

Table 3.8 compares the 10, 100, and 1000Mbps Ethernet MAC parameters.

TABLE 3.8 A COMPARISON OF 10, 100, AND 1000MBPS ETHERNET MAC PARAMETERS

Parameter	Ethernet/802.3	Fast Ethernet/ 802.3u	Gigabit Ethernet/802.3z
SlotTime	512 bit times	Same[1]	512-byte times (4096 bit times)
`Minimum InterFrameGap`	96 bit times	Same	Same
`AttemptLimit`	16 (tries)	Same	Same
`BackoffLimit`	10 (exponential number)	Same	Same
`JamSize`	32 bits	Same	Same
`MaxFrameSize`	1518 bytes	Same	Same
`MinFrameSize`	64 bytes (512 bits)	Same	Same
`AddressSize`	48 bits	Same	Same

1 *Same means same value in bit times as 10 or 100Mbps Ethernet (for example, the value in the column to the left).*

Gigabit Ethernet PHYs

The 802.3z Gigabit Ethernet standard includes three PHYs: 1000BASE-SX and LX to support fiber-optic cable and 1000BASE-CX for shielded 150-Ohm copper cable. Gigabit Ethernet used the same proven concept that Fast Ethernet had already pioneered: It uses an off-the-shelf PHY standard. In this case, the ANSI X3T11 Fibre Channel PHY was chosen. Fibre Channel was a natural choice, as it was proven, and components were readily available. The only issue was speed. The Fibre Channel PMD runs at 1Gbaud and uses 8B/10B encoding. This translates to a data rate of only 800Mbps. The IEEE therefore increased the Fibre Channel PHY speed to 1.25Gbaud to obtain an actual data throughput of 1Gbps. This means the Fibre Channel components had to be retested or even redesigned for the 25% speed increase. See Figure 3.11 for an illustration.

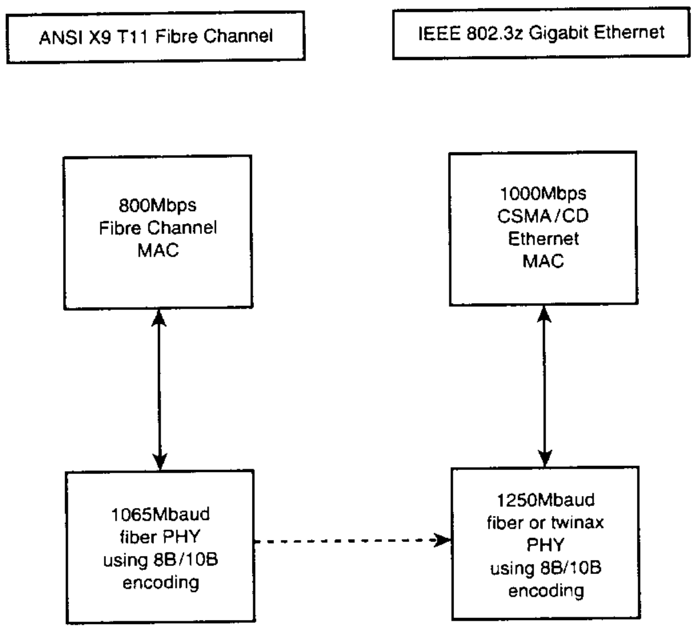

FIGURE 3.11 *Three of the Gigabit Ethernet PHYs are based on the ANSI Fibre Channel PHY running at 1.25Gbaud.*

Finally, the 802.3ab 1000BASE-T physical layer will enable Gigabit Ethernet over Category 5 UTP. This standard isn't complete at the time of writing this book, but should be approved in late 1998 or early 1999.

1000BASE-SX: Gigabit Ethernet for Horizontal Fiber

1000BASE-SX is targeted at cost-sensitive, shorter backbone, or horizontal connections (*S* as in "short wavelength or shorter cable runs"). 1000BASE-SX uses the same physical layer as LX and uses affordable 850nm short-wavelength optical diodes. 1000BASE-SX uses only multimode fiber (MMF). The distance supported varies from 220 meters to 550 meters, depending on the type of fiber cable used.

1000BASE-LX: Ethernet for Vertical or Campus Backbones

1000BASE-LX is targeted at longer backbone and vertical connections (*L* as in "long wavelength or longer runs"). LX can use either SMF or MMF. LX requires expensive 1300nm lasers. The IEEE has specified *a* segment length of 5000m for LX using SMF. For LX using MMF, the distance is 550m (full-duplex connections). Table 3.9 summarizes the different 1000BASE-SX and LX distance combinations. The IEEE specifies the SC connector for both 1000BASE-SX and 1000BASE-LX.

TABLE 3.9 1000BASE-SX AND LX MAXIMUM LENGTH VARIES DEPENDING ON TYPE AND QUALITY OF FIBER CABLING USED

Wave-length	Fiber Type	Fiber Size (µm)	Bandwidth	Attenuation	Maximum Distance[1]
			1000BASE-SX		
850	MMF	50/125µm[2]	400MHz/km	3.25	500m
			500MHz/km	3.43	550m
		62.5/125µm[3]	160MHz/km	2.33	220m
			200MHz/km	2.53	275m
			1000BASE-LX		
1300	MMF	50/125µm	400/500MHz/km	2.32	550m
		62.5/125µm	500MHz/km	2.32	550m
	SMF	10/125µm	Huge/infinite[4]	4.5	5000m

1 *All distances are for full-duplex. We don't know of a single vendor building half-duplex 1000BASE-X equipment. For the record, half-duplex distances are 316m for all 1000BASE-LX connections and between 220m to 316m for 1000BASE-SX.*

2 *The maximum length depends on the type of 50µm or 62.5µm fiber used. The 400MHz/km stuff can only accommodate 500m, but the 500MHz/km fiber is specified at 550m.*

3 *Again, the maximum length depends on the type of 62.5µm fiber used. The 160MHz/km stuff can only accommodate 220m, but the 200MHz/km fiber is specified at 275m. Most of the installed base of 62.5µm fiber has a bandwidth of 160MHz/km. The lower bandwidth 160 and 400MHz/km fiber is more prevalent in Europe than in the United States.*

4 *SMF has a modal bandwidth that exceeds the capabilities of today's electronics components. For practical purposes, it cannot be measured and is therefore infinite.*