Chapter 1. IEEE 802.3. Ethernet

1.1 Overview

IEEE Std 802.3 was first published in 1985. Since the initial publication, many projects have added functionality or provided maintenance updates to the specifications and text included in this standard. Each IEEE 802.3 project/amendment is identified with a suffix (e.g., IEEE Std 802.3ba™-2010).

The Media Access Control (MAC) protocol specified in IEEE Std 802.3 is Carrier Sense Multiple Access with Collision Detection (CSMA/CD). This MAC protocol was included in the experimental Ethernet developed at Xerox Palo Alto Research Center. While the experimental Ethernet had a 2.94 Mb/s data rate, IEEE Std 802.3-1985 specified operation at 10 Mb/s. Since 1985, new media options, new speeds of operation, and new capabilities have been added to IEEE Std 802.3.

Back in the 1970s at the Xerox Palo Alto Research Center, Dr. Robert M. Metcalf developed a network standard that enabled the sharing of printers to personal workstations. This original system, entitled the “Alto Aloha Network” (later re-named “Ethernet”), was able to transmit data at a rate of 3 Mb/s between all connected computers and printers. Later, in 1980 a multi-vendor consortium consisting of DEC, Intel, and Xerox released the DIX Standard for Ethernet. It was through this effort that Ethernet was able to become an open standard for network operations.

At the same time, the Institute of Electrical and Electronic Engineers (IEEE) created a group designated the 802 Working group to standardize network technologies. This group created standards that they would later number 802.x, where x was the subcommittee developing the particular standard. The subcommittee that developed the standards for the CSMA/CD, functionally very similar to the DIX Ethernet system, was 802.3. Later in 1985, the official standards were released for the IEEE 802.3. The standards were for Carrier Sensing Multiple Access with Collision Detection access method.

It is important at this point to distinguish between original DIX Ethernet and what is termed Ethernet today. Ethernet today refers to the IEEE 802.3 standards. Although functionally similar, there are subtle differences between the two standards. The main differences lay with the frame structure, which will be discussed later, and the data transfer rates. For now, when the term Ethernet is used, it refers to the IEEE 802.3 standards.

This is an international standard for Local and Metropolitan Area Networks (LANs and MANs), employing CSMA/CD as the shared media access method and the IEEE 802.3 (Ethernet) protocol and frame format for data communication. This international standard is intended to encompass several media types and techniques for a variety of MAC data rates.

1.1.1 Scope

This standard defines Ethernet local area, access and metropolitan area networks. Ethernet is specified at selected speeds of operation; and uses a common media access control (MAC) specification and management information base (MIB). The Carrier Sense Multiple Access with Collision Detection (CSMA/CD) MAC protocol specifies shared medium (half duplex) operation, as well as full duplex operation. Speed specific Media Independent Interfaces (MIIs) provide an architectural and optional implementation interface to selected Physical Layer entities (PHY). The Physical Layer encodes frames for transmission and decodes received frames with the modulation specified for the speed of operation, transmission medium and supported link length. Other specified capabilities include: control and management protocols, and the provision of power over selected twisted pair PHY types.

1.1.2 Basic concepts

The Ethernet system relays on the CSMA/CD standard. CSMA/CD simply means that the computers all have access to the transmission medium, and can send and receive data whenever the network is idle. The benefit of Ethernet is that it has the ability to sense collisions on the network. A collision occurs when two or more machines (nodes) try to send data at the same time. There are sophisticated techniques used to keep this from occurring on a regular basis. These techniques will be discussed later.

When a node on an Ethernet network wishes to send information to another node, it first listens to the network to see if there is network traffic. If the station detects no traffic, it will begin sending the frames of data. These frames will be transmitted throughout the network and ALL nodes on the particular Ethernet segment will receive the frames. However, only the node for which it was intended will be able to view the contents of the frame. This is done through source and destination addressing, which will be described later. If, however, more and more nodes become active on the network the probability of multiple nodes trying to send information at the same time increases. If two or more nodes send data at the same time a collision will occur. When this happens, the sending station will send out a jam sequence alerting all other nodes that there has been a collision and that any data received should be discarded. The node then waits a period of time and re-sends the frame. A mathematical algorithm termed “Truncated Binary Exponential Backoff “determines the amount of time the node waits.

The CSMA/CD standard can be broken down into its individual parts and applied to the description above. The Carrier Sensing (CS) is the ability of the computers to listen to the network and determine if there is activity. Multiple Access (MA) refers to the fact that all nodes on the network have access to the transmission medium at all times, and finally, the Collision Detection (CD) process was explained above.

This standard provides for two distinct modes of operation: half duplex and full duplex. A given IEEE 802.3 instantiation operates in either half or full duplex mode at any one time. The term “CSMA/CD MAC” is used throughout this standard synonymously with “802.3 MAC,” and may represent an instance of either a half duplex or full duplex mode data terminal equipment (DTE), even though full duplex mode DTEs do not implement the CSMA/CD algorithms traditionally used to arbitrate access to shared-media LANs.

1.2 Carrier Sensing Multiple Access with Collision Detection (CSMA/CD)

The following section goes into greater detail explaining how the CSMA/CD standard functions and how it enables the network to transmit data efficiently and with very few errors.

As was explained earlier, CSMA/CD allows machines to send and receive data any time it senses that the network is inactive. This method allows for a much more efficient use of the network resources and transmission medium. First it is necessary to explain how the Ethernet network is set up.

A network can be made up of two or more machines connected together with a transmission medium. These nodes that are connected together form an Ethernet Segment or a Collision Domain. It is called a collision domain because all of the nodes will receive every other nodes traffic. This being the case, the transmission medium is truly shared, therefore collision predisposed to. Machines can be on the same network, but not on the same Collision Domain. This is done through the use of bridges and switches.

Each machine or node on the network has a unique MAC (Medium Access Control) address. This MAC address is permanently imprinted on the NIC (Network Interface Card) in the form of a ROM (Read Only Memory) chip. The MAC address has to be unique within a Local Area Network. It is this address that distinguishes a node from other machines on the LAN.

1.2.1 Data Transmission

On any collision domain of an Ethernet network any information that is sent out over the network propagates in both directions in order to reach all nodes. All nodes receive every frame that is sent over the network, whether it is intended for that particular node or not. Only if the frame is addressed to that particular node, is the node allowed to accept it.

When a particular node on the network is ready to send information it goes through a series of steps which are outlined below:

1. The node listens to the network to see whether any other node or machine is transmitting. The node is able to listen by sensing the carrier signals present on the network transmission medium. If there is activity, the node continues to wait.
2. When no signal is detected, the node starts transmission of the message of frame.
3. While the node is transmitting, it also listens to the network. The node compares the received message with what was transmitted. If they are the same, the node continues to transfer, putting a 9.6 µs gap between frames.
4. If what is received is not what was sent, the node assumes it was a collision and stops transmitting.
5. The node transmits a Jam sequence which tells other nodes that a collision has been detected
6. The node waits a random amount of time and then begins again

This is the basic process that each node goes through when it transmits a packet. Figure 1.1 illustrates this process in detail. It seems very simple, but the truth is that there are a lot of behind-the-scenes processes that enable these simple six steps to take place.

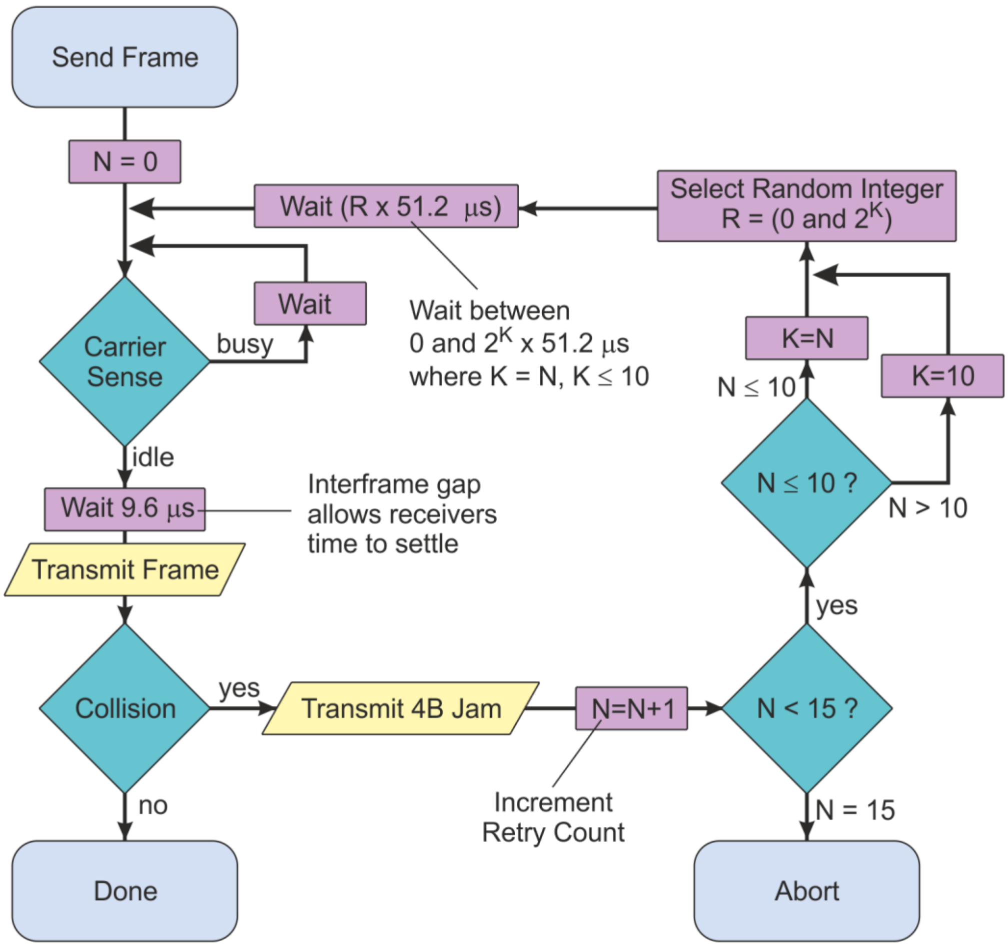


Figure 1.1. IEEE 802.3 Transmission Algorithm

The node “listens” to the transmission medium by use of a transceiver. This transceiver monitors the current flow along the cable. When the transceiver picks up current flow that translates to a bit flow (about 18-20 mA), it says the cable is busy and does not transmit any data. If the transceiver senses no activity, i.e. no current flow, it can begin transmission of data.

1.2.2 Frame Collisions

While the node is transmitting, it also continues to “listen”. It monitors all data that it has sent over the network. When it senses a collision on the network it halts transmission. It is able to sense a collision using the same transceiver. When the transceiver detects excess current on the line, it stops transmission of data and transmits a 32-bit jam sequence. This sequence is to let any node that may be receiving the damaged frame, to discard it. The receiving machine knows this because the 32-bit jam sequence is designed to take the place of the 32-bit CRC (Cyclic Redundancy Check) error-checking portion of the data frame. When the receiving node gets this jam sequence, it checks it against its CRC and determines it is an error and discards the frame. Following a detection of a collision, the node(s) will wait a random amount of time before transmitting the frame again. This process is known as “Truncated Binary Exponential Backoff”.

Simply put, after the collision and jam sequence has been sent and received, each node involved in the collision can either transmit immediately (following the 9.6 µs gap), or it can wait one window period. A window period is defined as the time it takes one frame to propagate the round-trip length of the network. The standard window time has been set to 51.2 µs. This is calculated by the fact that a standard frame is at least 512 bits in length. Since the transmission speed is 10 Mb (10,000,000 bits) / second, it would take 51.2 µs to send one frame.

After a collision, the node will select a multiple of the base window time (51.2 µs) to wait before sending again. This multiple comes from a set of numbers generated by the node with each successive send attempt. For example, if the node encounters a collision, it will either send immediately, or wait one window (51.2 µs). If it encounters another collision, it will then select a multiple of the window from the set of (0, 1, 2, or 3).

There are four options; therefore there is only a 25% chance that both nodes will choose the same time interval. The set of multiples continues to increase for each repeated attempt.

The formula for the set of numbers is simply 2^K where K is the number of attempted resends. So for the first attempt it would be 2^1 for a set of (0 or 1) times the window value. If it was the second attempt, it would be 2^2 for a set of (0, 1, 2, or 3) times the window and so on. This will continue up through fifteen re-sends. At this point, the node will stop attempting to send, and look to higher OSI-level software to decide what to do next.

Incidentally, the total of fifteen re-sends corresponds to a set of windows ranging from 0 to 1023, which corresponds directly with the 1024 maximum number of nodes allowable on any collision domain.

1.2.3 Half duplex operation

In half duplex mode, the CSMA/CD media access method is the means by which two or more stations share a common transmission medium. To transmit, a station waits (defers) for a quiet period on the medium (that is, no other station is transmitting) and then sends the intended message in bit-serial form. If, after initiating a transmission, the message collides with that of another station, then each transmitting station intentionally transmits for an additional predefined period to ensure propagation of the collision throughout the system. The station remains silent for a random amount of time (backoff) before attempting to transmit again. Each aspect of this access method process is specified in detail in subsequent clauses of this standard. Half duplex operation can be used with certain media types and configurations as defined by this standard.

1.2.4 Full duplex operation

Full duplex operation allows simultaneous communication between a pair of stations using point-to-point media (dedicated channel). Full duplex operation does not require that transmitters defer, nor do they monitor or react to receive activity, as there is no contention for a shared medium in this mode. Full duplex mode can only be used when all of the following are true:

a) The physical medium is capable of supporting simultaneous transmission and reception without interference.
b) There are exactly two stations connected with a full duplex point-to-point link. Since there is no contention for use of a shared medium, the multiple access (i.e., CSMA/CD) algorithms are unnecessary.
c) Both stations on the LAN are capable of, and have been configured to use, full duplex operation.

The most common configuration envisioned for full duplex operation consists of a central bridge (also known as a switch) with a dedicated LAN connecting each bridge port to a single device. Repeaters as defined in this standard are outside the scope of full duplex operation. Full duplex operation constitutes a proper subset of the MAC functionality required for half duplex operation.

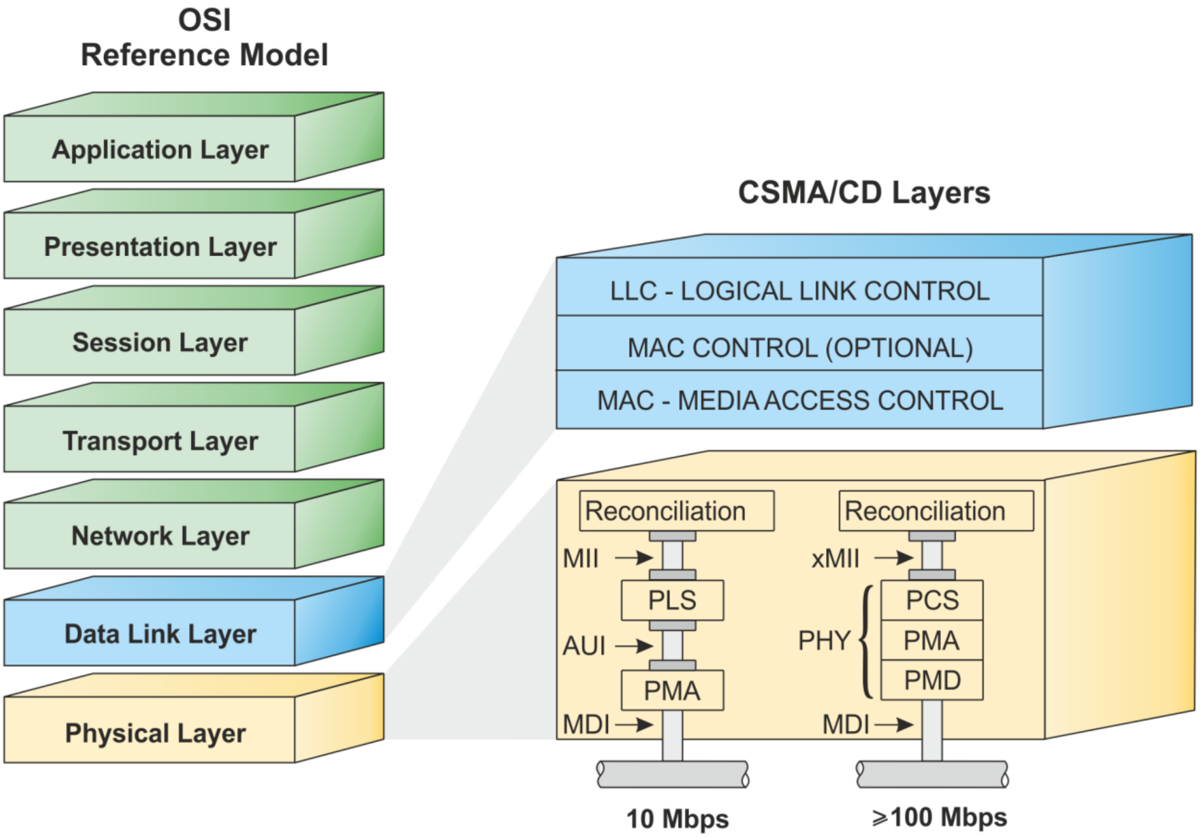
1.3 Architectural perspectives

There are two important ways to view network design corresponding to the following:

a) Architecture. Emphasizing the logical divisions of the system and how they fit together.
b) Implementation. Emphasizing actual components, their packaging, and interconnection.

This standard is organized along architectural lines, emphasizing the large-scale separation of the system into two parts: the Media Access Control (MAC) sublayer of the Data Link Layer and the Physical Layer. These layers are intended to correspond closely to the lowest layers of the ISO/IEC Model for Open Systems Interconnection (see Figure 1.2). The Logical

Link Control (LLC) sublayer and MAC sublayer together encompass the functions intended for the Data Link Layer as defined in the OSI model.



AUI = Attachment Unit Interface
MAU = Medium Attachment Unit
MDI = Medium Dependent Interface
MII = Media Independent Interface

PCS = Physical Coding Sublayer
PHY = Physical Layer Device
PLS = Physical Layer Signaling
PMA = Physical Medium Attachment
PMD = Physical Medium Dependent

NOTE—In this figure, the xMII is used as a generic term for the Media Independent Interfaces for implementations of 100 Mb/s and above. For example: for 100 Mb/s implementations this interface is called MII; for 1 Gb/s implementations it is called GMII; for 10 Gb/s implementations it is called XGMII; etc.

Figure 1.2—IEEE 802.3 standard relationship to the
ISO/IEC Open Systems Interconnection (OSI) reference model

1.3.1 Architectural rationale

An architectural organization of the standard has two main advantages:

a) Clarity. A clean overall division of the design along architectural lines makes the standard clearer.

b) Flexibility. Segregation of medium-dependent aspects in the Physical Layer allows the LLC and MAC sublayers to apply to a family of transmission media. Partitioning the Data Link Layer allows various media access methods within the family of LAN standards.

The architectural model is based on a set of interfaces that may be different from those emphasized in implementations. One critical aspect of the design, however, shall be addressed largely in terms of the implementation interfaces: compatibility.

1.3.2 Compatibility interfaces

The following important compatibility interfaces are defined within what is architecturally the Physical Layer:

a) Medium Dependent Interfaces (MDI). To communicate in a compatible manner, all stations shall adhere rigidly to the exact specification of physical media signals defined in the appropriate clauses in this standard, and to the procedures that define correct behavior of a station. The medium-independent aspects of the LLC sublayer and the MAC sublayer should not be taken as detracting from this point; communication in an Ethernet Local Area Network requires complete compatibility at the Physical Medium interface (that is, the physical cable interface).
b) Attachment Unit Interface (AUI). Some DTEs are located some distance from their connection to the physical cable. A small amount of circuitry will exist in the Medium Attachment Unit (MAU) directly adjacent to the physical cable, while the majority of the hardware and all of the software will be placed within the DTE. The AUI is defined as a second compatibility interface. While conformance with this interface is not strictly necessary to ensure communication, it is recommended, since it allows maximum flexibility in intermixing MAUs and DTEs. The AUI may be optional or not specified for some implementations of this standard that are expected to be connected directly to the medium and so do not use a separate MAU or its interconnecting AUI cable. The PLS and PMA are then part of a single unit, and no explicit AUI implementation is required.
c) Media Independent Interface (MII). It is anticipated that some DTEs will be connected to a remote PHY, and/or to different medium dependent PHYs. The MII is defined as a third compatibility interface. While conformance with implementation of this interface is not strictly necessary to ensure communication, it is recommended, since it allows maximum flexibility in intermixing PHYs and DTEs. The MII is optional.
d) Gigabit Media Independent Interface (GMII). The GMII is designed to connect a 1 Gb/s capable MAC or repeater unit to a 1 Gb/s PHY. While conformance with implementation of this interface is not strictly necessary to ensure communication, it is recommended, since it allows maximum flexibility in intermixing PHYs and DTEs at 1 Gb/s speeds. The GMII is intended for use as a chip-to-chip interface. No mechanical connector is specified for use with the GMII. The GMII is optional.
e) Ten-bit Interface (TBI). The TBI is provided by the 1000BASE-X PMA sublayer as a physical instantiation of the PMA service interface. The TBI is recommended for 1000BASE-X systems, since it provides a convenient partition between the high-

frequency circuitry associated with the PMA sublayer and the logic functions associated with the PCS and MAC sublayers. The TBI is intended for use as a chip-to-chip interface. No mechanical connector is specified for use with the TBI. The TBI is optional.

f) 10 Gigabit Media Independent Interface (XGMII). The XGMII is designed to connect a 10 Gb/s capable MAC to a 10 Gb/s PHY. While conformance with implementation of this interface is not necessary to ensure communication, it allows maximum flexibility in intermixing PHYs and DTEs at 10 Gb/s speeds. The XGMII is intended for use as a chip-to-chip interface. No mechanical connector is specified for use with the XGMII. The XGMII is optional.

g) 10 Gigabit Attachment Unit Interface (XAUI). The XAUI is designed to extend the connection between a 10 Gb/s capable MAC and a 10 Gb/s PHY. While conformance with implementation of this interface is not necessary to ensure communication, it is recommended, since it allows maximum flexibility in intermixing PHYs and DTEs at 10 Gb/s speeds. The XAUI is intended for use as a chip-to-chip interface. No mechanical connector is specified for use with the XAUI. The XAUI is optional.

h) 10 Gigabit Sixteen-Bit Interface (XSBI). The XSBI is provided as a physical instantiation of the PMA service interface for 10GBASE-R and 10GBASE-W PHYs. While conformance with implementation of this interface is not necessary to ensure communication, it provides a convenient partition between the high-frequency circuitry associated with the PMA sublayer and the logic functions associated with the PCS and MAC sublayers. No mechanical connector is specified for use with the XSBI. The XSBI is optional.

i) 40 Gigabit Media Independent Interface (XLGMII). The XLGMII is designed to connect a 40 Gb/s capable MAC to a 40 Gb/s PHY. While conformance with implementation of this interface is not necessary to ensure communication, it allows flexibility in intermixing PHYs and DTEs at 40 Gb/s speeds. The XLGMII is a logical interconnection intended for use as an intra-chip interface. No mechanical connector is specified for use with the XLGMII. The XLGMII is optional.

j) 40 Gigabit Attachment Unit Interface (XLAUI). The XLAUI is a physical instantiation of the PMA service interface to extend the connection between 40 Gb/s capable PMAs. While conformance with implementation of this interface is not necessary to ensure communication, it is recommended, since it allows maximum flexibility in intermixing PHYs and DTEs at 40 Gb/s speeds. The XLAUI is intended for use as a chip-to-chip or a chip-to-module interface. No mechanical connector is specified for use with the XLAUI. The XLAUI is optional.

k) 40 Gigabit Parallel Physical Interface (XLPPI). The XLPPI is provided as a physical instantiation of the PMD service interface for 40GBASE-SR4 and 40GBASE-LR4 PMDs. The XLPPI has four lanes. While conformance with implementation of this interface is not necessary to ensure communication, it allows flexibility in connecting the 40GBASE-SR4 or 40GBASE-LR4 PMDs The XLPPI is intended for use as a chip-to-module interface. No mechanical connector is specified for use with the XLPPI. The XLPPI is optional.

l) 100 Gigabit Media Independent Interface (CGMII). The CGMII is designed to connect a 100 Gb/s capable MAC to a 100 Gb/s PHY. While conformance with implementation of this interface is not necessary to ensure communication, it allows flexibility in intermixing PHYs and DTEs at 100 Gb/s speeds. The CGMII is a logical interconnection intended for use as an intra-chip interface. No mechanical connector is specified for use with the CGMII. The CGMII is optional.

m) 100 Gigabit Attachment Unit Interface (CAUI). The CAUI is a physical instantiation of the PMA service interface to extend the connection between 100 Gb/s capable PMAs. While conformance with implementation of this interface is not necessary to ensure communication, it is recommended, since it allows maximum flexibility in intermixing PHYs and DTEs at 100 Gb/s speeds. The CAUI is intended for use as a chip-to-chip or a chip-to-module interface. No mechanical connector is specified for use with the CAUI. The CAUI is optional.
n) 100 Gigabit Parallel Physical Interface (CPPI). The CPPI is provided as a physical instantiation of the PMD service interface for 100GBASE-SR10 PMDs. The CPPI has ten lanes. While conformance with implementation of this interface is not necessary to ensure communication, it allows flexibility in connecting the 100GBASE-SR10 PMDs. The CPPI is intended for use as a chip-to-module interface. No mechanical connector is specified for use with the CPPI. The CPPI is optional.

1.4 Media Access Control (MAC) service specification

This clause specifies the services provided by the Media Access Control (MAC) sublayer to the client of the MAC (see Figure 1.2). MAC clients may include the Logical Link Control (LLC) sublayer, Bridge Relay Entity, or other users of ISO/IEC LAN International Standard MAC services (see Figure 1.3). The services are described in an abstract way and do not imply any particular implementation or any exposed interface. Other clauses in this standard may add optional protocol sublayers directly above the MAC that preserve the service interface to the MAC client.

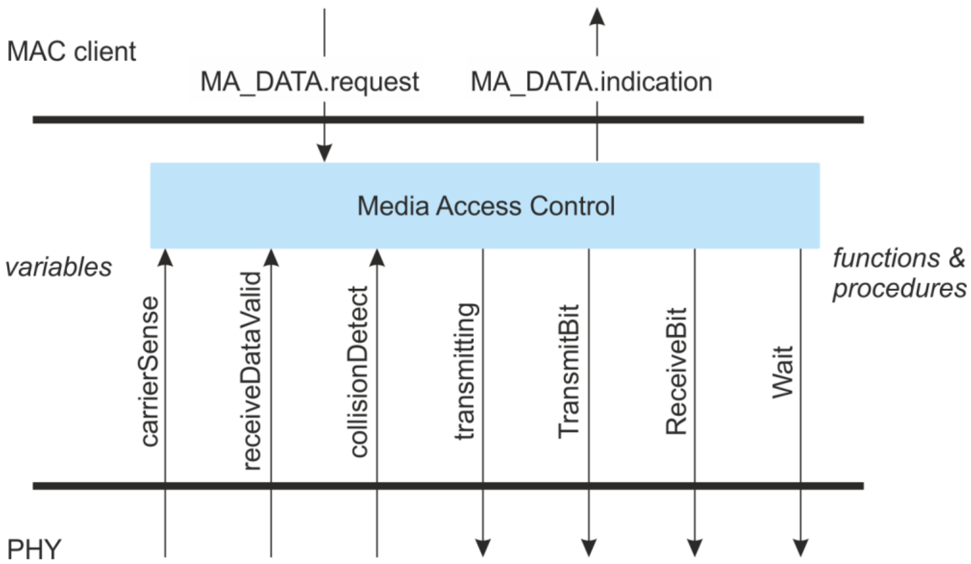


Figure 1.3—Service specification primitive relationships

1.5 Media Access Control (MAC) frame and packet specifications

1.5.1 Overview

This clause defines the mapping between MAC service interface primitives and Ethernet packets, including the syntax and semantics of the various fields of MAC frames and the fields used to form those MAC frames into packets.

During Ethernet's history, capabilities have been added to allow data link layer (layer 2) protocol encapsulations within the MAC Client Data field. As a result, there are now more than one type of MAC frame.

The frame format specified in this clause includes the following three types of MAC frames:

a) A basic frame
b) A Q-tagged frame
c) An envelope frame

All three frame types use the same Ethernet frame format.

1.5.2 Packet format

Figure 1.4 shows the fields of a packet: the Preamble, Start Frame Delimiter (SFD), the addresses of the MAC frame's destination and source, a length or type field to indicate the length or protocol type of the following field that contains the MAC client data, a field that contains padding if required, and the Frame Check Sequence (FCS) field containing a cyclic redundancy check value to detect errors in a received MAC frame.

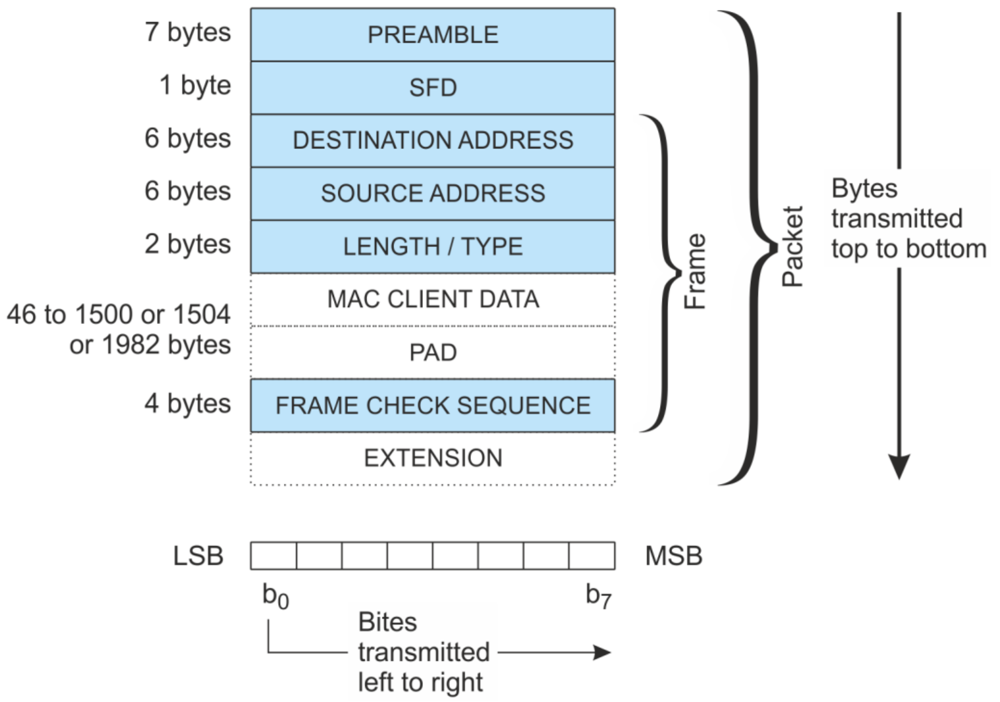


Figure 1.4—Packet format

An Extension field is added, if required (for 1000 Mb/s half duplex operation only). Of these fields, all are of fixed size except for the MAC Client Data, Pad and Extension fields, which may contain an integer number of octets between the minimum and maximum values that are determined by the specific implementation of the MAC.

The minimum and maximum MAC frame size limits refer to that portion of the packet from the Destination Address field through the Frame Check Sequence field, inclusive (i.e., the MAC frame).

Relative to Figure 1.4, the octets of a packet are transmitted from top to bottom, and the bits of each octet are transmitted from left to right.

The IEEE 802.3 data frame consists of seven different fields. These fields are put together to form a single data frame. Figure 1.4 illustrates the seven following fields: Preamble, Start-of-Frame delimiter, Destination Address, Source Address, Length, Data, and Frame Check Sequence. Each is discussed below.

Preamble
The preamble is a field that tells the receiving node that a data frame is coming. This field is simply a 56-bit (7 byte) alternating pattern of 1s and 0s.

Start of Frame Delimiter
The start-of-frame delimiter is used in conjunction with the preamble to synchronize the receiving clock with the transmitting clock though on lock of the Digital Phase Lock Loop (DPLL). The Digital Phase Lock Loop circuit is used to lock onto the phase timing of the frame which is imbedded in the Manchester Encoding of the data.

Along with the recognized bit pattern supplied by the preamble, the DPLL circuit is able, through the use of shift registers, to lock the receiving clock on to the timing of the sending clock.

Destination Address
The destination address is the MAC address of the machine to which the particular frame is to be delivered. As explained earlier, each NIC (Network Interface Card) has a unique MAC address. It is this address that is part of the destination address field, which is a 6 byte or 48-bit address.

The destination address can be one of three types: unicast, multicast, or broadcast. A unicast address is addressed to a single node on the network. This is the MAC address of the machine. Multicast is where a single frame can be sent to a number of nodes in a particular group. This is done by programming individual nodes to listen for specific multicast addresses. If one of these addresses is present in the destination address, any node that is set up to receive that address will retrieve that data. The third type of destination address is a broadcast address. When the destination address contains a broadcast address, every node on the network will be able to retrieve the data in that frame. The standard broadcast address is a 48-bit number of all 1s. This addressing scheme allows the network to be very flexible in transmission of data.

Source Address
The source address, like the destination address, is a 48-bit field. However, this value is always a unicast address and always reflects the MAC address of the sending node.

Length
The length field consisting of 16 bits contains the total number of bits of information contained in the following Data field.

Data
The data field contains the actual data to be processed by upper level protocols of the recipient node. The length of the data must be between 46 – 1500 bytes. The 46-byte minimum is to ensure that the entire length of the data frame is at least 64 bytes in length. The 64 bytes equates to 512 bits. This is the minimum size a data frame must be for nodes on either end of the network to be able to detect collisions. This is related to the propagation time of the data frame through the 10 Mb/s network.

Frame Check Sequence
The frame check sequence is a 4 byte, 32 bit Cyclic Redundancy Check (CRC) value. This value is calculated by the transmitting node and appended to the frame. On the receiving end, the receiving node also calculates this value. If the values do not match, there has been a transfer error and the frame is discarded.

Cyclic Redundancy Check
The main error checking method for frames transferred over a CSMA/CD network is the Cyclic Redundancy Check (CRC). This is a 32 bit value that is appended to the end of the data frame as explained above. The CRC is calculated by the transmitting node and then again by the receiving node. If they do not match, i.e. if the receiving node does not calculate the same CRC number as the one in the data frame, there has been a transmission error, and the frame is discarded.

1.6 Signal characteristics

1.6.1 Signal encoding

Two different signal encoding mechanisms may be used by the AUI. One of the mechanisms is used to encode data, the other to encode control.

1.6.1.1 Data encoding

Encoding is the process of adding the correct transitions to the message signal in relation to the data that is to be sent over the communication system. The first step is to establish the data rate that is going to be used. Once this is fixed, then the mid-bit time can be determined as ½ of the data rate period.

Manchester encoding is used for the transmission of data across the AUI. Manchester encoding is a binary signaling mechanism that combines data and clock into “bit-symbols.” Each bit-symbol is split into two halves with the second half containing the binary inverse of the first half; a transition always occurs in the middle of each bit-symbol. During the first half of the bit-symbol, the encoded signal is the logical complement of the bit value being encoded. During the second half of the bit-symbol, the encoded signal is the uncomplemented value of the bit being encoded. Thus, a bit 0 is encoded as a bit-symbol in which the first half is HI and the second half is LO. A bit 1 is encoded as a bit-symbol in which the first half is LO and the second half is HI. An example of Manchester waveform is shown in Figure 1.5.

The line condition IDL is also used as an encoded signal. An IDL always starts with a HI signal level. Since IDL always starts with a HI signal, an additional transition will be added to the data stream if the last bit sent was a zero. This transition cannot be confused with clocked data (CD0 or CD1) since the transition will occur at the start of a bit cell. There will be no transition in the middle of the bit cell. The IDL condition, as sent by a driver, shall be maintained for a minimum of 2 bit times. The IDL condition shall be detected within l.6 bit times at the receiving device.

a) System jitter considerations make detection of IDL (etd, end transmission delimiter) earlier than 1.3 bit times impractical. The specific implementation of the phase-locked loop or equivalent clock recovery mechanism determines the lower bound on the actual IDL detection time. Adequate margin between lower bound and 1.6 bit times should be considered.
b) Recovery of timing implicit in the data is easily accomplished at the receiving side of the interface because of the wealth of binary transitions guaranteed to be in the encoded waveform, independent of the data sequence. A phase-locked loop or equivalent mechanism maintains continuous tracking of the phase of the information on the Data circuit.

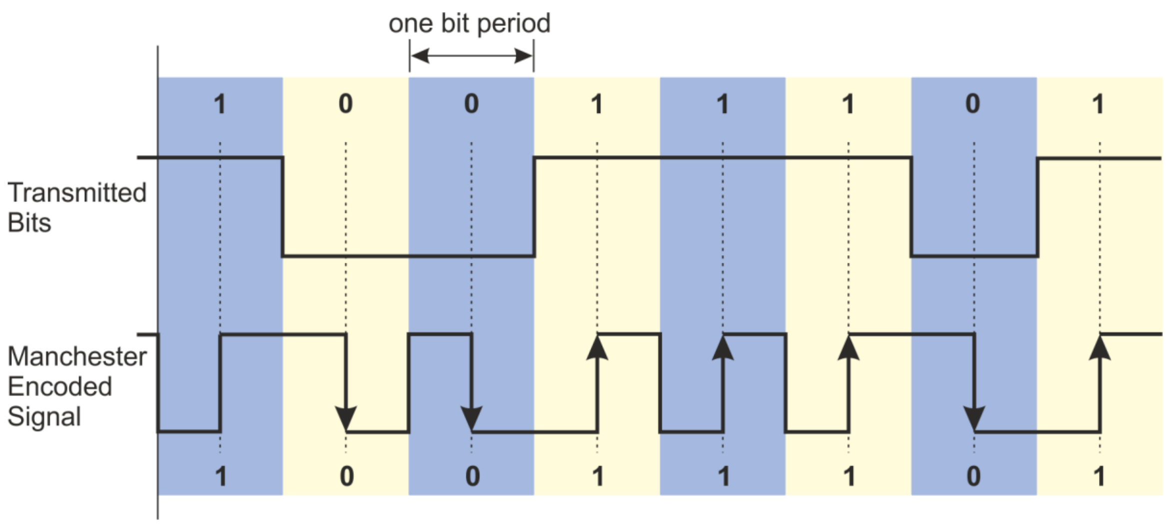


Figure 1.5. Example of Manchester encoded signal

1.6.1.2 Control encoding

A simpler encoding mechanism is used for control signaling than for data signaling. The encoded symbols used in this signaling mechanism are CS0, CS1, and IDL. The CS0 signal is a signal stream of frequency equal to the bit rate (BR). The CS1 signal is a signal stream of frequency equal to half of the bit rate (BR/2). If the interface supports more than one bit rate, the bit rate in use on the data circuits is the one to which the control signals are referenced. The IDL signal used on the control circuits is the same as the IDL signal defined for the data circuits. The Control Out circuit is optional (O) as is one message on Control In.

1.7 Ethernet Media

A big part of designing and installing an Ethernet system is using the right medium. A large selection of cables is a result of Ethernet evolution over the years and Ethernet flexibility

The three major types of media in use today are:

- Coax
  - Thick wire 10Base-5
  - Thin wire 10Base-2
- Twisted pair
  - 10Base-T
  - 100Base-T (twisted pair and fiber-optic cable)
  - 1000Base-T
- Fiber optic
  - 10Base-F
  - 1000Base-X

The first Ethernet systems used 10Base-5 coax cable. The name identifies different functions and limitations of the cabling:

- 10—Mb/s bandwidth of the signal
- Base—Cable carries only Ethernet
- 5—500 meters is the longest distance a cable segment can carry the signal

The expensive thick wire cable (10 mm) was soon replaced by less expensive, but less efficient, 10Base-2 thin wire cable. The maximum range for thin wire is 185 meters (that rounded up to 2 for the name of the cable).

The most popular Ethernet wiring is UTP cable that comes in a wide variety of grades (categories), performance levels, and prices. The UTP cable is similar to telephone plug cable, except it has eight connectors that are either straight-through or crossover. Straight-through connections are used in most Ethernet networks in order to connect computers to a switch, while crossover cables can connect two computers directly to each other without the use of a hub or switch.

Cat. 3 to Cat. 5 UTP cable is used for 10Base-T and 100Base-T electrical signals. Cat. 5 or better is used for 1000Base-T.

Fiber optic cables are gaining more in popularity each year. While it is more expensive than electrical cable, it is unaffected by many environmental conditions that can be problems for coax or UTP. Additionally, SMF optic cable provides practically unlimited bandwidth expandability and the ability to transmit Ethernet signals considerably longer distances.

The newer networking configuration is 1000Base-X or Gigabit Ethernet, with 10 Gig Ethernet looming on the horizon (IEEE 802.3ae was released in June 2002). Short wavelength (1000Base-SX) and long wavelength (1000Base-LX) are based on an 8B/10B block encoding scheme. Note: 8B/10B stands for 8 bits of data transmitted in a 10-bit sequence where the last two additional bits are used for signal and control functions.

The proliferation of media not only exemplifies Ethernet's growth and flexibility, it is also an example of Ethernet's complexity.

1.8 Fast Ethernet: 100BaseT

The idea of Fast Ethernet was first proposed in 1992. In August 1993, a group of vendors came together to form the Fast Ethernet Alliance (FEA). The goal of the FEA was to speed Fast Ethernet through the Institute of Electrical and Electronic Engineers (IEEE) 802.3 body, the committee that controls the standards for Ethernet. Fast Ethernet and the FEA succeeded, and in June 1995, the technology passed a full review and was formally assigned the name 802.3u.

The IEEE's name for Fast Ethernet is 100BaseT, and the reason for this name is simple: 100BaseT is an extension of the 10BaseT standard, designed to raise the data transmission capacity of 10BaseT from 10Mbits/sec to 100Mbits/sec. An important strategy incorporated by 100BaseT is its use of the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol which is the same protocol that 10BaseT uses because of its ability to work with several different types of cable, including basic twisted-pair wiring. Both of these features play an important role in business considerations, and they make 100BaseT an attractive migration path for those networks based on 10BaseT.

The basic business argument for 100BaseT resides in the fact that Fast Ethernet is a legacy technology. Because it uses the same transmission protocol as older versions of Ethernet and is compatible with the same types of cable, less capital investment will be needed to convert an Ethernet-based network to Fast Ethernet than to other forms of high-speed networking. Also, because 100BaseT is a continuation of the old Ethernet standard, many of the same network analysis tools, procedures, and applications that run over the old Ethernet network work with 100BaseT. Consequently, managers experienced at running an Ethernet network should have found the 100BaseT environment familiar, meaning less time and money must be spent by the company on training.

1.8.1 Protocol Preservation

Perhaps the shrewdest strategy taken with Fast Ethernet was the decision to leave the transmission protocol intact. The transmission protocol, in this case CSMA/CD, is the method a network uses to transmit data from one node to another, over the cable. In the OSI model, CSMA/CD is part of the Media Access Control, or MAC, layer. The MAC layer specifies how information is formatted for transmission and the way in which a network device gains access to, or control of, a network for transmission.

1.8.2 Types of Fast Ethernet

Another important consideration, along with the adoption of the CSMA/CD protocol, was the decision to design 100BaseT so it uses many basic forms of cabling those used by older versions of Ethernet and newer forms of cabling, as well. To accommodate the different types of cables, Fast Ethernet comes in three forms: 100BaseTX, 100BaseT4, and 100BaseFX. Both 100BaseTX and 100BaseT4 work with twisted-pair cabling standards, while 100BaseFX was created to work with fiber optic cabling.

The 100BaseTX standard is compatible with two pairs of UTP or STP. One pair is designated for reception and the other for transmission. The two basic cabling standards that meet this requirement are EIA/TIA-568 Category 5 UTP and IBM's Type 1 STP. The attractiveness of 100BaseTX lies in its ability to provide full-duplex performance with network servers and the fact that it uses only two of the four pairs of wiring, leaving two pairs free for future enhancements to your network.

However, if you plan on using Category 5 cabling with 100BaseTX, be forewarned of the drawbacks related to Category 5. The cable is more expensive than other types of four-pair cabling, such as Category 3, and it requires the installation of punch down blocks, connectors, and patch panels that are all Category 5-compliant.

The 100BaseT4 standard requires a less sophisticated cable than Category 5. The reason is that 100BaseT4 uses four pairs of wiring: one for transmission, one for reception, and two that can either transmit or receive data. Therefore, 100BaseT4 has the use of three pairs of wiring to either transmit or receive data. By dividing up the 100Mbit/sec data signal among the three pairs of wiring, 100BaseT4 reduces the average frequency of signals on the cable, allowing lower-quality cable to handle the signal successfully. Categories 3 and 4 UTP cabling, as well as Category 5 UTP and Type 1 STP, can all work in 100BaseT4 implementations.

The advantage of 100BaseT4 is its flexible cable requirements. Category 3 and 4 cabling was once more prevalent in existing networks, and if they aren't already being used in your network, they cost less than Category 5 cabling. The downside is that 100BaseT4 uses all four pairs of wiring, and it does not support full-duplex operation.

Fast Ethernet also offers a standard for operation over multimode fiber with a 62.5 micron core and 125 micron cladding. The 100BaseFX standard is designed mainly for backbone use, connecting Fast Ethernet repeaters scattered about the building. The traditional benefits of fiber optic cabling are still valid with 100BaseFX: protection from electromagnetic noise, increased security, and longer distances allowed between network devices.

1.8.3 Segment length

Although Fast Ethernet is a continuation of the Ethernet standard, the migration from a 10BaseT network to a 100BaseT network isn't a straight, one-to-one conversion of hardware some changes to the network topology may be required.

Theoretically, Fast Ethernet limits the end-to-end network diameter or the network segment diameter to 250 meters; only 10 percent of the 2,500-meter maximum theoretical size of Ethernet. Fast Ethernet's restriction is based on the speed of 100Mb/s transmission and the nature of the CSMA/CD protocol. A 100BaseT network can't be longer than 250 meters.

For Ethernet to work, a workstation transmitting data must listen long enough to make sure the data has reached its destination safely. In a 10Mb/s Ethernet network, such as 10Base5, the length of time a workstation listens for a collision is equivalent to how far a 512-bit frame (the frame size is specified in the Ethernet standard) travels before the workstation is finished processing it. In a 10Mb/s Ethernet network, that distance is 2,500 meters.

However, a 512-bit frame (the 802.3u standard specifies the same frame size, 512 bit, as the 802.3 standard) being transmitted by a workstation in the faster 100Mb/s Ethernet network travels only about 250 meters before the workstation is finished processing it. If the receiving workstation was located farther than 250 meters from the transmitting workstation, the frame may collide with another frame down the line, and the transmitting workstation, having finished processing the transmission already, would not be listening for the collision. For this reason, the maximum network diameter for a 100BaseT network is 250 meters.

To take advantage of the 250 meters, though, you will need to install two repeaters to connect all of the nodes. And, a node cannot be located farther than 100 meters away from a repeater Fast Ethernet adopted the 10BaseT rule that determines 100 meters to be the farthest allowable distance a workstation can be from a hub. Due to latency introduced by connection devices such as repeaters, the actual operational distances between nodes will probably prove to be less than those stated. So, it would be prudent to measure distances on the short side.

To incorporate longer runs in a network, it has to be invested in fiber cabling. For example, it can be used 100BaseFX in half-duplex mode to connect a switch to either another switch or an end station located up to 450 meters away. A full-duplex 100BaseFX installation will allow two network devices up to two kilometers apart to communicate.

1.9 Gigabit Ethernet

Gigabit Ethernet allows network transfers up to 1.000 Mbps using standard Cat 5 UTP (unshielded twisted pair) cabling. Below is explained how this can be accomplished as long Cat 5 cables can run only up to 100 Mbps.

Ethernet Cat 5 cables have eight wires (four pairs), but under 10BaseT and 100BaseT standards (10 Mbps and 100 Mbps, respectively) only four (two pairs) of these wires are actually used. One pair is used for transmitting data and the other pair is used for receiving data.

Ethernet standard uses a technique against electromagnetic noise called cancellation. As electrical current is applied to a wire, it generates an electromagnetic field around the wire. If this field is strong enough, it can create electrical interference on the wires right next to it, corrupting the data that were being transmitted there. This problem is called crosstalk.

What cancellation does is to transmit the same signal twice, with the second signal "mirrored" (inverted polarity) compared to the first one, as you can see in. So when receiving the two signals, the receiving device can compare the two signals, which must be equal but "mirrored". The difference between the two signals is noise, making it very simple to the receiving device to know what is noise and to discard it. "+TD" wire standards for "Transmitting Data" and "+RD" wire standards for "Receiving Data". "-TD" and "-RD" are the "mirrored" versions of the same signal being transmitted on "+TD" and "+RD", respectively.

On 10BaseT standard each bit that the computer wants to transmit is physically coded into a single transmitting bit, i.e., for a group of eight bits being transmitted, eight signals will be generated on the wire. Its 10 Mbps transfer speed means that its clock is of 10 MHz, but just because each clock cycle a single bit is transmitted. On other standards this is different.

100BaseT uses a coding scheme called 8B/10B, where each group of eight bits is coded into a 10-bit signal. So, differently from 10BaseT, each bit does not directly represent a signal on the wire. If you make the proper math, with a 100 Mbps data transfer rate, the clock rate of 100BaseT is of 125 MHz (10/8 x 100).

So, Cat 5 cables are certified to have a transmission speed of up to 125 MHz.

What Gigabit Ethernet does is to change the coding. Instead of making each bit to be coded into a single signal like 10BaseT or to code each 8-bit group into a 10-bit signal, it codes two bits per signal. So, a signal over a Gigabit Ethernet cable represents two bits, instead of a single bit. In order words, instead of just using two voltages on a signal representing merely “0” or “1”, it uses four different voltages, representing “00”, “01”, “10” and “11”.

Also, instead of using just four wires of the cable, Gigabit Ethernet uses all wires.
On top of this, all pairs are used in a bi-directional fashion. Both 10BaseT and 100BaseT use different pairs for transmission and reception; on 1000BaseT, as Gigabit Ethernet cabling is also called, the same pairs are used for both data transmission and reception.

The beauty of Gigabit Ethernet is that it still uses the 100BaseT/Cat 5 clock rate of 125 MHz rate, but since more data is transmitted per time, the transfer rate is higher. The math is quite simple: 125 MHz x 2 bits per signal (i.e., per wire pair) x 4 signals per time = 1.000 Mb/s.

This modulation technique is called 4D-PAM5 and it actually uses five voltages (the fifth voltage is used for its error-correction mechanism).

So it is a mistake to say that Gigabit Ethernet runs at 1.000 MHz. It doesn't. It runs at 125 MHz just like Fast Ethernet (100BaseT), but it achieves a 1.000 Mbps because it transmits two bits per time and uses the four pairs of the cable.

Gigabit Ethernet works much the same way as 10 Mb/s and 100 Mb/s Ethernet, only faster. It uses the same IEEE 802.3 frame format, full duplex, and flow control methods. Additionally, it takes advantage of CSMA/CD when in half-duplex mode, and it supports simple network management protocol (SNMP) tools. Gigabit Ethernet takes advantage of jumbo frames to reduce the frame rate to the end host. Standard Ethernet frame sizes are between 64 and 1518 bytes. Jumbo frames are between 64 and 9215 bytes. Because larger frames translate to lower frame rates, using jumbo frames on Gigabit Ethernet links greatly reduces the number of packets (from more than 80000 to less than 15000 per second) that are received and processed by the end host.

Gigabit Ethernet can be transmitted over CAT 5 cable and optical fiber such as the following:

- 1000Base-CX—Short distance transport (copper)
- 1000Base-SX—850 nm wavelength (fiber optics)
- 1000Base-LX—1300 nm wavelength (fiber optics)

10 Gigabit Ethernet

The operation of 10 Gigabit Ethernet is similar to that of lower speed Ethernets. It maintains the IEEE 802.3 Ethernet frame size and format that preserves layer 3 and greater protocols. However, 10 Gigabit Ethernet only operates over point-to-point links in full-duplex mode. Additionally, it uses only multimode and single mode optical fiber for transporting Ethernet frames.

Note: Operation in full-duplex mode eliminates the need for CSMA/CD.
The 10 Gigabit Ethernet standard (IEEE 802.3ae) defines two broad physical layer network applications:

- Local area network (LAN) PHY
- Wide area network (WAN) PHY