Ethernet LANs

This chapter describes the ubiquitous Local Area Network (LAN) standard called Ethernet which is defined in IEEE standard 802.3.
After this chapter you should be able to: compute the values of the fields of an 802.3 Ethernet frame; identify the type of 802.3 LAN PHY according to the bit rate, number of pairs used, grade of cable and line code; specify the PHY parameters that would be chosen by autonegotiation between two Ethernet PHYs; specify the port(s) on which a frame will leave a learning bridge.

802.3 Frame Format

The following diagram[1] shows the contents of an Ethernet frame:

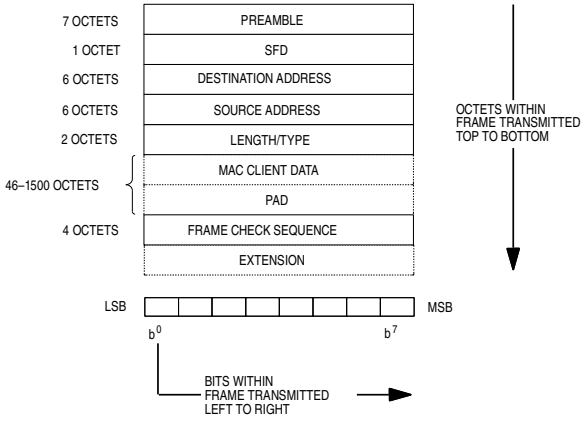


Figure 3–1—MAC frame format

Following a preamble used for synchronization consisting of 7 bytes of 0x55 and a start frame delimiter (SFD) byte of value 0xd5, each frame consists of:

Destination Address (6 bytes) - this is the MAC or "hardware" address of the destination.

Source Address (6 bytes) - the MAC address of the source interface.

Length/Type (2 bytes) - in some older Ethernet versions this was the length of the frame but it is now used to identify the protocol used in the Data field. The type field is necessary because an Ethernet frame can carry payload for different types of higher-layer protocols (e.g. IPv4, IPv6, ARP). The actual length of the frame is determined by the PHY layer (hardware) as described below.

[1]Non-hand-drawn diagrams in this lecture are taken from IEEE 802.3.

Data (up to 1500 bytes, or 9000 for non-standard "Jumbo" frames) - the payload. If necessary, padding must be added to reach a PHY-dependent minimum frame duration.

FCS (4 bytes) - a 32-bit CRC.

Note that some of these fields change slightly depending on the specific physical layer used.

There must be a minimum gap between frames (e.g. 96 bit times). Short pulses (Normal Link Pulses or NLP) are transmitted periodically in-between frames to allow devices to detect that they're connected.

Note that bits in each byte are transmitted least-significant bit (lsb) first while the length/type field is transmitted most-significant byte (MSB) first ("big endian" order).

MAC Addresses

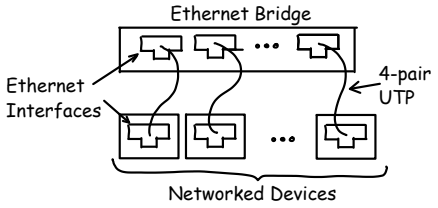
Each Ethernet interface has a unique MAC address. The first three bytes are called an Organizationally Unique Identifier (OUI) and is assigned by the IEEE to a specific manufacturer. The second three bytes are assigned by the manufacturer to a specific device and are programmed into the interface hardware. The first transmitted bit of the OUI (the LS bits of the first destination address byte) indicates a broadcast (or multicast) address which is meant to be received by all (or many) stations.

Exercise 1: How many possible OUI's are there? How many devices can be manufactured for each OUI?

Bridged Ethernet

Modern Ethernet systems use a direct connection from each networked device to a central "bridge" thus creating a "star" architecture. The bridge has

one interface (or port) for each device connected to it. The bridge receives a complete frame from a device then retransmits it on the appropriate port(s).



Ethernet PHYs

The IEEE 802.3 standard defines many physical interfaces (PHYs) all of which share a common frame format. Only three PHYs are in common use today: 10BASE-T, 100BASE-TX and 1000BASE-T operating at 10, 100 and 1000 Mb/s respectively[2]. These are defined in IEEE 802.3 "clauses" (chapters) 14, 25 and 40 respectively. Clauses 3 and 4 describe the frame format and the medium sharing protocol (Media Access Control or MAC) respectively which apply to all PHYs.

These standards use unshielded twisted pair (UTP) 24-gauge four-pair cables with a 100 m maximum cable length. Each end of the cable has an "RJ-45" 8-position modular connector. The links usually operate in full-duplex mode although half-duplex operation is also possible.

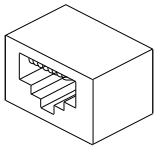


Figure 40–29—MDI connector

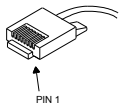


Figure 40–30—Balanced cabling connector

The MDI (Media Dependent Interface) connector pins can be wired as either "MDI" (typically devices) or "MDI-X" (MDI crossover, typically on a bridge). For example pins 1 and 2 are outputs on a 100 Mb/s MDI device but inputs on an MDI-X device. However, most interfaces today have "automatic MDI-X" which detects link pulses on either pair and switches the pin assignments as necessary.

[2] 10, 100 and 1000 are the bit rates, 'BASE' refers to operation at baseband, T/TX refer to use over twisted pair (F/FX refer to fiber).

10BASE-T

This PHY was designed for operation over typical telephone twisted pairs, typically 24 gauge, up to 100m long. It used one pair in each direction and allowed either full- or half-duplex operation. As with all twisted pair Ethernet PHYs, the signalling is differential and the peak voltages at the transmitter are approximately ±1 V.

The 10BASE-T PHY line code is Manchester at a 10 MHz symbol rate (100 ns per symbol, 20 MHz baud rate). The transmit signal is filtered to avoid ISI. The transmit filter is defined by defining a time-domain "mask" or template that the signal has to pass within when transmitting a symbol that is a high-to-low transition:

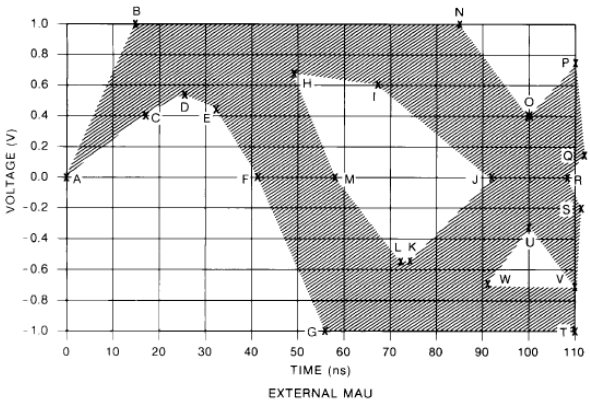


Figure 14–10—Voltage template

The start of a frame is indicated by a preamble while the end of the frame is indicated by a coding violation (the absence of a transition following the last bit of the frame).

100BASE-TX

The 100 Mb/s PHY operates over two 100 ohm pairs, one in each direction, with a maximum length of 100 m. The cable is required to meet certain requirements ("Category 5" defined in an EIA standard). The same connector and pin assignments are used as in 10BASE-T.

The second pair of pairs in each cable often carries analog phone signals or provides power to remote devices (called PoE for "Power over Ethernet").

The symbol rate is 125 MHz. The 4B5B line code is used to convert each set of 4 bits to 5 bits. MLT-3 is

then used to generate a three-level signal with lower bandwidth than if NRZ were used. The nominal voltages are +1, 0, -1V although, as with 10BASE-T, the waveform is smoothed out due to equalization filtering.

The start of frame, end of the frame and line-idle conditions are indicated by the use of specific 4B5B symbols (see line coding lecture).

1000BASE-T

The 1 Gb/s PHY uses four pairs, with each pair used in both directions simultaneously:

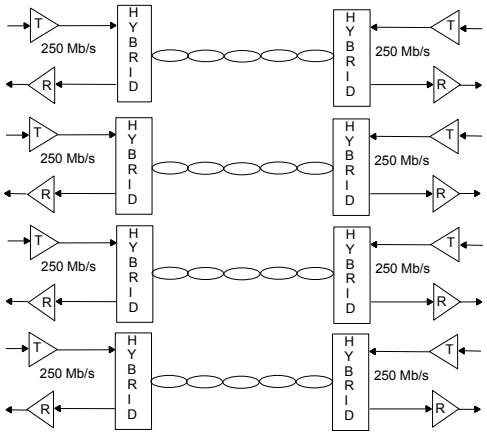


Figure 40–2—1000BASE-T topology

The symbol rate is 125 MHz with two bits per symbol. The overall bit rate is thus $125 \times 10^6 \times 2$ bits/symbol $\times$ 4 pairs = 1 Gb/s.

This PHY uses the same connectors and Cat 5 cable as 100BASE-TX.

Since both sides transmit over the same pair simultaneously in both directions, the hardware must be able to subtract out the transmitted signal, echoes, and crosstalk from the received signal. This is implemented by digitizing the four signals and using Digital Signal Processing (DSP).

The line code is called 4D-PAM5 for four-dimensions (four pairs) with five levels[3]. The input to the line coder, known as the PCS (Physical Coding Sublayer), is two bits per symbol per pair (8 bits per symbol). The output is four voltages ("dimensions") of five different levels with peak voltages of $\pm 1\ V$. There are thus $5^4 = 625$ different symbols available to encode $2^8 = 256$ possible inputs. As with 4B5B some symbols are not used because they have undesirable properties and others are used to signal special conditions such as the start and end of frame.

Independent scramblers are used for the two directions (called Master and Slave) and the scrambling sequences on the different pairs are offset from each other. In addition to scrambling, convolutional coding is used on the data. This helps meet the error rate requirement of 10^{-10}.

As with 100BASE-T, special sequences from the 8B10B block line code are used to mark the start and end of the frame.

[3]PAM is Pulse Amplitude Modulation.

Autonegotiation

All three PHYs above share the same connector and cabling type. In addition, most 100 Mb/s devices support 10 Mb/s and most Gigabit Ethernet devices support both slower speeds. Interfaces can additionally be used in full- or half-duplex mode. To avoid misconfiguration of interfaces the 802.3 standard defines an autonegotiation signalling scheme, using sequences of link pulses, that can be used by devices to discover and configure the best (fastest) possible mutually-compatible interface configuration.

Hubs, Bridges, Switches, Routers

A hub is the simplest way to connect multiple Ethernet devices. Similar to a co-ax-based Ethernet, a hub creates the logical OR of all devices' outputs and uses this signal to drive all of the devices' inputs. Although hubs are very inexpensive, they are rarely used today.

A bridge contains two independent, typically full-duplex, ethernet ports. Frames received on one port are transmitted on the other port. A learning bridge keeps track of the Ethernet addresses on each side of the bridge and only repeats frames that are destined for the other side of the bridge.

A switch is a learning bridge that contains more than two ports. Often only one device is connected to each port.

A learning bridge stores the address seen arriving on each port in a table and uses this table when deciding which port to forward a frame on. If the destination address is not known (such as when a device is first connected), the bridge must transmit the frame over all of the interfaces except the one the frame was received on. This is called "flooding".

A router is similar to a switch but uses the layer 3 (IP) address to select the outgoing port. Being a layer 3 device, a router is not limited to Ethernet and routers often have interfaces for other layer 2 protocols such as wireless (e.g. 802.11) or optical (e.g. SONET).

Exercise 2: Classify each of a hub, learning bridge and switch according to the following: can operate in full-duplex mode, can have independent PHY rates, collisions can happen, can receive from multiple ports simultaneously.