Ethernet LANs

This lecture describes the ubiquitous Local Area Network (LAN) called Ethernet which is defined in IEEE standard 802.3. After this lecture you should be able to: decide if a network is a LAN or PAN; compute the values of the fields of an 802.3 Ethernet frame; decide whether a bus or star LAN topology best meets specific requirements; 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 with and without VLANs configured; and choose disabled ports on interconnected bridges that will result in a spanning tree.

LAN, PAN, WAN

Local Area Networks (LAN) are networks that transfer data between devices that are with about 100 m of each other. Typically all of these devices are within one building and all are owned and controlled by one entity such as one company.

By analogy, the term Personal Area Network (PAN) refers to networks with a more limited range often just a few meters around a person. PANs are often used to connect peripherals such as mice to computers or headsets to phones. The most common example is Bluetooth, a wireless PAN.

A Wide Area Network (WAN) is a network with wider span than a LAN. Data networks provided by cellular service providers are considered wireless WANs. Other -AN terms you may hear occasionally are MAN (Metropolitan), BAN (Body). The latter terms are not well-defined or widely used (although the networks may be).

802.3 Frame Format

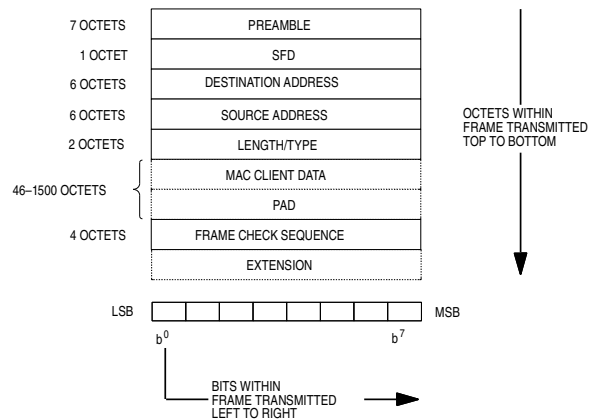The following diagram[1] shows the contents of an Ethernet frame:

Figure 3–1—MAC frame format

Following a preamble used for synchronization and a start frame delimiter (SFD) byte, 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 is now used to identify the type of frame and is used to interpret the contents of the Data field. The type field is necessary because the Data payload can be formatted in different ways depending on the Layer 3 (network) protocol in use. The actual length of the frame is determined by the PHY layer (hardware).

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.

_____

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

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

MAC Addresses

Each Ethernet interface has a unique MAC address. The first three bytes is 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 it. The two most significant bits of the OUI are used to indicate broadcast (or multicast) addresses which are 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?

Early Ethernet

The first Ethernet LANs used a bus topology with a single co-ax cable that each device in a network "segment" connected to by "vampire" tapping to the center conductor through a hole drilled in the shield. The CSMA/CD protocol was used to recover from collisions.

This "bus" style of network architecture allowed simple deployment and flexibility. Only one cable needed to be installed and new devices could be added at any point along this cable.
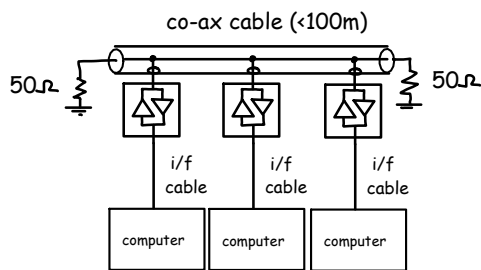

However, a bus architecture has a number of problems:

- the capacity of the bus (cable) must be shared between all users
- there is no privacy since all users can hear each others' transmissions
- a problem with one node (e.g. a short or open circuit) affects all users on the segment

This version of Ethernet is obsolete.

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).
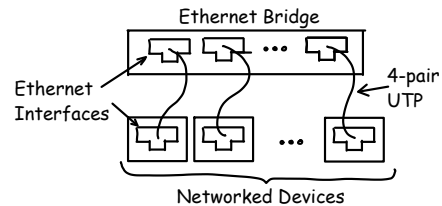
This has a number of advantages compared to the bus architecture:

- each port can operate at full speed and ports can even operate at different speeds
- there are no collisions caused by two stations deciding to transmit at the same time
- with some exceptions, frames are only forwarded to the desired port and this improves privacy and security
- electrical problems with one link do not affect other users

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.

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.

_____

[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).
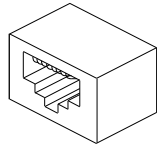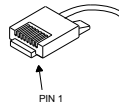
Figure 40–29—MDI connector

PIN 1

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 an MDI device but inputs on an MDI-X device. However, most interfaces today have "automatic MDI-X" which automatically switches the pinout as necessary.

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 ± 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:
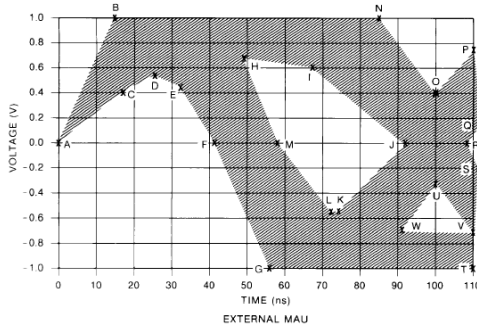

Figure 14–10—Voltage template

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.

1000BASE-T

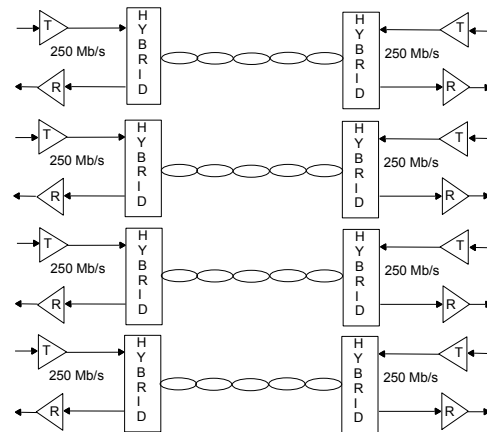The 1 Gb/s PHY uses four pairs, with each pair used in both directions simultaneously:

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, 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

_____
[3]PAM is Pulse Amplitude Modulation.

3

symbol). The output is four voltages ("dimensions") of five different levels with peak voltages of ± 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}.

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 that can be used by devices to discover and configure the best (fastest) possible mutually-compatible interface configuration.

Learning Bridges

A bridge looks at the source address field of each Ethernet frame to learn the MAC address(es) of the device(s) attached to each port. These addresses are stored in a table and are used 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".

Spanning Tree Protocol

In order to extend the capacity and range of an Ethernet network, bridges can be connected to each other as shown below:
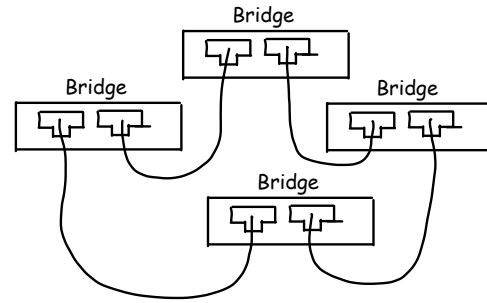
In the example above there are also redundant links. These links increase the reliability of the network.

However, these links can result in loops. In this case broadcast frames will be forwarded around the loop indefinitely since there is no way for a bridge to detect that it has already seen a particular frame. The resulting "broadcast storm" quickly saturates the network making it unusable.

To avoid this possibility, bridges use a distributed algorithm known as the "Spanning Tree Protocol" to eliminate loops by disabling certain ports.

VLANs

Virtual LANs (VLANs) are a way to separate devices into groups so that each group of devices appears to be on its own LAN. This is done by restricting broadcast and flooded frames to devices that are in the same VLAN. This improves network security and capacity. For example at a school one VLAN could be used for the students and another for the administration.

VLAN membership is typically configured by network administrators. The VLAN can be determined by source MAC address or by the port on the bridge a frame arrives on.

To preserve this information as frames traverse the LAN, IEEE standard 802.1q defines a modified 802.3 frame format that is used to 'tag' frames with their VLAN ID. The 802.1q VLAN tag is a 16-bit value inserted into the frame between the source address and the length field. It is stripped off unless the frame is being sent to another VLAN-capable bridge.