IEEE 802.3 Ethernet

- **1-persistent CSMA/CD LAN**
- **Ethernet Cabling**
  - **10base5**: Thicknet (good for backbones)
    - Thick cable, doesn’t bend well
    - vampire taps used to “tap” the network
    - max run is 500 meters
  - **10Base2**: Thin coax (cheapest system)
    - uses BNC “T” connectors
    - max run is 200 meters
  - **10baseT**: twisted pairs (easy to maintain - very popular now)
    - uses a central hub that simulates a shared cable
    - uses RJ-45 connectors
    - max run is 100 meters to hub
  - **10baseF**: fiber optics (best for between buildings)
    - expensive connectors and terminators
    - max run is 2000 meters
Manchester Encoding for 10Mbps Ethernet

- **Problem:** How to send a 0/1 bit?
  - Negative -5 volt for a 0 bit and 5 volts for a 1 bit?
  - How about a long string of 1’s and 0’s: the output is a constant volt over a long period of time
    - a drift of clock can result in a loss of synchronization

- **Answer:** Force a transition in the middle of a bit period
  - the mid-bit transition serves as a self-clocking mechanism
  - every bit is half low (L) and half high (H)
  - 1 is high then low (HL)
  - 0 is low then high (LH)
  - but this requires 2 signals per bit
    - Ex 4.18: What is the baud rate of 10Mbps 802.3 LAN?

- **Differential Manchester Encoding (used in token ring)**
  - 0 causes a transition at the start, 1 doesn’t
  - a change of voltage still occurs in the middle
  - better noise immunity (because more transitions are encoded)
  - e.g., if the initial voltage is L, then the next bit is a 1 if LH follows; 0 if HL
Ex. 4.22: Manchester encoding for 0001110101?
   – Ans: LH LH LH HL HL HL LH HL LH HL

Ex. 4.23: Differential Manchester encoding for the same bit string assuming the initial state is L
   – Ans: HL HL HL LH HL LH LH HL LH LH HL
Ethernet Frame Format

- **Preamble** (7 bytes of 10101010) is used to sync clock
- **Target Address** - 6 bytes
  - if it starts with a 0 it is for ordinary addresses (one to one)
  - if it starts with a 1 it is for group addresses (multicast)
  - if all 1 bits it is for all stations on the network (broadcast)
- **Length** - 0 to 1500 bytes
  - if less than 48 bytes then the pad field is used to fill the frame to the minimum frame size of 64 bytes (i.e., 26+48=64)
    - to make sure all frames must take more than 2τ to send so that the sender can have a chance to detect collision
- **4-byte Checksum field** to detect garbled data at link level
Frame Size vs. Data Rate in Ethernet

- **Why a minimum frame size 64 bytes in IEEE 802.3?**
  - The bottom line: a transmission of a frame cannot be completed before $2\tau$ time or the sender may be fooled
  - For a 10 Mbps Ethernet LAN with a maximum cable length of 2.5 km (with four repeaters and five 0.5km segments), $2\tau=51.2\ \mu$sec
  - So minimum frame size = $51.2/10\text{Mbps} = 512$ bits or 64 bytes

- **What if the data rate is 1Gbps?**
  - To cope with this higher data rate, we can
    - use a larger frame size while keeping the cable length the same
      $\Rightarrow$ Minimum frame size = $51.2/1\text{Gbps} = 6,400$ bytes, or
    - use a shorter cable while keeping the minimum frame size the same
      $\Rightarrow$ maximum cable length = 25 meters, or
    - use a combined approach: maximum cable length = 250 meters and minimum frame size = 640 bytes

- **Ex. 4.36: Fast Ethernet running at 100 Mbps has the same minimum frame size as Ethernet running at 10 Mbps, why?**
  - Ans: Cable length of Fast Ethernet has to be only 1/10 of Ethernet if the same minimum frame size is kept
CSMA/CD LAN Data Rate

Q: What is the effective data rate of a 10 Mbps CSMA/CD LAN (not 802.3) given the following data?

- A slotted ALOHA protocol is used for the channel contention interval with a slot time of $2\tau = 20 \mu \text{sec}$ (refer to Figure 4-5, p. 253)
- A station contends for a slot; if no collision occurs, it seizes the channel and sends one frame
- Data frames are 1000 bits including 24 overhead bits
- The receiver sends a 30-bit ACK frame (including the 24 overhead bits) to the sender after receiving a frame
- The first contention slot after a frame is sent is reserved for the receiver
- Assume no collisions

Ans: There are 4 time stages

- Time for the sender to seize the channel: $20 \mu \text{sec}$
- Time for the sender to transmit its frame: $1000/10\text{Mbps} = 100 \mu \text{sec}$
- Time for the receiver to seize the channel: $20 \mu \text{sec}$
- Time for the receiver to transmit the ACK frame: $30/10\text{Mbps} = 3 \mu \text{sec}$
- The effective data rate is only $(1000-24)/(20+100+20+3) = 6.82\text{Mbps}$
Collision Management

- **Binary Exponential Backoff**
  - time is divided into $2\tau$ slot times
  - after first collision, wait either 0 or 1 slot time
  - after second collision, wait either 0, 1, 2, or 3 slot times
  - in general, after $j$th collision, wait a r.n. between 0 and $2^j - 1$
  - limited to 1023 slots
  - after 16 collisions, data link layer gives up

- **Ex. 4-20**: what is the probability that the contention between two stations ends after attempt #k and what is the mean number of attempts per contention period?

- **Ans**: attempt $j$ (starting at 1) is distributed among $2^{j-1}$ slots
  - thus the probability of a collision on attempt $j$ is $1/(2^{j-1})$
  - the probability that the first (k-1) attempts fail, followed by a success on attempt $k$, denoted by $p(k)$, is $P(k) = (1 - 2^{-(k-1)}) \prod_{j=1}^{k-1} 2^{-(j-1)}$
  - The expected number of attempts is thus $\sum_{k=0}^{\infty} k \cdot p(k)$
Performance of Ethernet

- If each station wants to transmit with probability $p$ in a slot, then the probability that one of $k$ stations acquires the channel in a slot is
  
  - $A = k \left[ p^1 (1-p)^{k-1} \right]$
  
  - $A$ is maximized when $p = 1/k$, with $A \to 1/e$ as $k \to \infty$

- Probability that a contention interval consists of $j$ slots is therefore equal to $A(1-A)^{j-1}$

- So the mean number of slots per contention is
  
  $$\sum_{j=0}^{\infty} jA(1-A)^{j-1} = \frac{1}{A}$$

- And the mean contention time interval is $w = \frac{2\tau}{A}$
  
  - When $k \to \infty$, $w = 2\tau e$
  
  - How to estimate $k$ in general?
    
    - $k\lambda = 1/(P+w)$ where $\lambda$ is the arrival rate (frames/sec) of each station and $P$ is one frame transmission time
Ethernet Performance (cont.)

- Maximum Ethernet Channel efficiency is then:

\[
\frac{P}{P + 2\tau / A} = \frac{F/B}{F/B + 2 (L/c) e} = \frac{1}{1 + 2 BLe / cF}
\]

B = bandwidth  
L = cable length  
c = speed of signal propagation  
F = frame length

- When \(2BLe/cF\) is large, network efficiency is low
  - network efficiency decreases as bandwidth (B) or cable length (L) increases
    * this implies 802.3 is not designed for applications that require high bandwidth over long distances, e.g., fiber MANs
  - network efficiency increases as frame size (F) increases
Ethernet Performance (cont.)

- network efficiency increases as frame size (F) increases

Fig. 4-23. Efficiency of 802.3 at 10 Mbps with 512-bit slot times.
Variations on Ethernet

- **Traditional Ethernet is a bus**
  - limited to one host at a time

- **Switched Ethernet**
  - a switch connecting several Ethernet segments

- **Fast Ethernet**
  - 100 Mbps
  - uses hubs only
    - shared hub - all the stations are logically connected like 802.3, forming a single collision domain
    - switched hub - like in switched Ethernet where several plug-in cards with buffering capability are present within a hub, with each card forming a separate collision domain
  - the maximum cable length is reduced by a factor of 10 if the same minimum frame size is to be kept
Switched Ethernet

- a switch contains high-speed (1 Gbps) backplane and 4 to 32 plug-in cards, each containing up to 8 ports for 8 hosts
  - A plug-in card forms its own Ethernet segment (collision domain)
    - frames for the same segment are handled the same way
    - frames for other segment travel over backplane
    - hierarchy is possible: a hub can connect to one port of a card via an 802.3 LAN
  - individual stations retain the same 10BaseT card and cabling

A plug-in card which forms a collision domain
Fast Ethernet

- **Based on hubs**
  - limits cable length to 100 meters for twisted pairs to/from hubs
  - can use a shared hub (like 802.3) or a switched hub (with cards)

- **Cabling and Signaling (not using Manchester encoding)**
  - 100Base-T4 (to utilize existing telephone wiring)
    - four category-3 twisted pairs, each with 25Mhz bandwidth
    - during a single clock cycle, a wire can contain a 0, a 1, or a 2
    - three wires in the forward direction can cover 27 possible symbols, making it possible to send 4 bits per cycle (baud) and thus yielding effective data rate = 4b/cycle *25 Mhz= 100 Mbps
    - This encoding scheme is called 8B6T, meaning 6 signals can transmit 8 bits for each twisted pair, making each wire’s effective data rate 33Mbps
    - the remaining twisted pair can cover traffic in the reverse direction
  - 100Base-TX
    - two category-5 twisted pairs (with more twists per unit distance than category 3 twisted pairs) each with 125Mhz data rate
    - 4 bits out of 5 bits (4B5B) for data encoding - 4 bits sent in 5 clock periods
    - full duplex (to the hub and from the hub) both at 100Mbps
IEEE 802.4 Token Bus

- Ethernet is not suitable for real-time computation since the time to acquire the channel is not bounded
- token bus forms a logical ring out of a single bus
- by passing a token among logical stations and restrict the token holding time of each station, the channel access time can be bounded

Fig. 4-25. A token bus.
IEEE 802.5 Token Ring

- the bypass relay inside the wire center is controlled by current from the stations
  - can bypass ring breaks and failed stations automatically

- A ring consists of ring interfaces connected by point-to-point lines so there is a 1-bit copying/inspection delay at each station interface

- the ring must have a sufficient delay to contain a complete token to circulate. Why?

- Advantages:
  - point-to-point technology
  - fair with priority scheduling an option
  - bounded access time

Fig. 4-29. Four stations connected via a wire center.
Token Ring MAC Protocol

- A station must seize the token and convert the token (SD-AC-ED) into the start-of-frame sequence (SD-AC-FC) before transmitting frames.
- Bits that have completed the trip around the ring come back and removed by the sender; the sender can compare these bits too.
- A sender can only hold the token for a *token-holding time* (10 msec), after which it must regenerate the token onto the ring.
- Frame status (FS) can implement acknowledgement for each frame:
  - $A=0 \land C=0 \Rightarrow$ destination failed or not present
  - $A=1 \land C=0 \Rightarrow$ destination present but frame not accepted
  - $A=1 \land C=1 \Rightarrow$ destination present and frame copied

![Token Ring MAC Protocol Diagram](image)

**Fig. 4-30.** (a) Token format. (b) Data frame format.
Token Ring MAC Protocols (cont.)

- Q1: A 10Mbps token ring has a token-holding time of 10 msec. What is the longest frame that can be sent on this ring?
  - Ans1: 10 msec * 10 Mbps = 100K bits minus overhead bits

- Q2: What are the factors that should be considered in calculating the effective bandwidth of a heavily loaded token ring?
  - Ans2: in a heavily loaded token ring, one frame will be sent following another from adjacent stations. The following time intervals need to be considered:
    - transmission time for a data frame
    - token regeneration time (ignored) and token transmission time
    - token propagation time to the next station -- the longer the cable, the longer the token propagation delay
Priority Handling in IEEE 802.5 Token Ring

- the 1-byte AC field in the token and in each data frame can implement priority scheduling
  - each station can reserve the next token at a priority level (if a higher/equal one is not present) by modifying the reservation bits in the current data frame's AC field
  - when the current data frame is finished, the next token is generated by the sender at the reserved priority specified in the AC byte
  - when a station wants to transmit a priority $n$ frame, it must capture a token with priority less than or equal to $n$
  - the station raising the priority lowers the token priority when done
- It is possible a station with low priority frames may suffer starvation waiting for a low priority token to arrive
- However, priority handling makes Token Ring more suitable for real-time applications than Ethernet
IEEE 802.5 Token Ring Maintenance

- **Maintenance works include**
  - regenerating a new token if the token is lost
  - inserting artificial delay when a ring breaks or a station fails
  - cleaning garbled and orphan frames

- **Each token ring has a monitor station that oversees the ring and a contention protocol for electing the monitor**
  - checking lost tokens is implemented by a timer set to the longest possible token-less interval (the full token holding time) -- if this timer goes off, drain and issue a new token
  - checking garbled frames is done by seeing invalid format (e.g., HH or LL) or checksum -- drain the ring, and issue a new token
  - checking orphan frames is implemented by setting a *monitor* bit in the 1-byte AC field in the frame when an orphan frame passes through the monitor - if the same orphan frame passes twice, the monitor drains it

- **Problem -- a faulty monitor cannot be impeached**
IEEE 802.6 Distributed Queue Dual Bus

- For WANs since 802.3 (Ethernet), 802.4 (Token Bus) and 802.5 (Token Ring) are for LANs which have length limitation. Why?
- Two parallel, unidirectional buses snake through the city
- Each bus has a head-end which generates empty 53-byte cells downstream to be used by stations who are ready to send
- Stations transmit frames in global FIFO order without using a central queue
- Each station maintains two local variables
  - RC (Request Counter): counting # of downstream requests pending
    - RC is incremented when an upstream request is heard
  - CD: the station’s position in the global queue, e.g., CD=2 means the queue position is 2 now for the station, so it needs to pass two empty downstream cells and uses the third downstream cell to transmit
    - When a station is ready to send downstream, it sets CD=RC & RC=0
    - When a station passes by an empty downstream cell, it will decrement CD if CD > 0; otherwise it will decrement RC if RC>0
IEEE 802.6 DQDB (cont.)

- An example of achieving FIFO using two local variables RC and DC

Fig. 4-32. (a) Initially the MAN is idle. (b) After D makes a request. (c) After B makes a request. (d) After D transmits. (e) After B transmits.
IEEE 802.6 DQDB (cont.)

- Ex. 4.34: C, A, B request in succession and then send data in the same order. Show RC/DC values in each station

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
</tr>
</tbody>
</table>

- Initially: 0/0 0/0 0/0 0/0 0/0

- After C’s request:
  - 1/0 1/0 0/0 0/0 0/0

- After A’s request:
  - 0/1 1/0 0/0 0/0 0/0

- After B’s request
  - 1/1 0/1 0/0 0/0 0/0

- After the first frame sent by C
  - 1/0 0/0 0/0 0/0 0/0

- After the 2nd frame sent by A?

- After the 3rd frame sent by B?
Other IEEE 802 Standards

- **IEEE 802.2 Logical Link Control (LLC)**
  - To run on top of 802.x protocols to provide an error-controlled, flow-controlled data link protocol
  - LLC sits at the upper half of the data link layer, with the MAC sublayer below it
  - LLC provides interfaces to the network layer:
    - unreliable datagram services
    - acknowledged datagram services
    - reliable connection-oriented services

---

Fig. 4-33. (a) Position of LLC. (b) Protocol formats.