Link Layer Technologies

Ethernet

→ CSMA/CD

→ copper, fiber

Evolution:

- 10Base5 (ThickNet): coax, segment length 500 m, 100 nodes/segment
- 10Base2 (ThinNet): coax, segment length 200 m, 30 nodes/segment
- 10Base-T: category 3, 4, 5 UTP, segment length 100 m, 1024 nodes/segment
- 100Base-T (Fast Ethernet): cat 5 UTP; fiber
- 1000Base-T (GigE): copper

  → IEEE 802.3ab: cat 5, up to 100 m

- IEEE 802.3z

  → 1000Base-SX: multi-mode fiber

  → 1000Base-LX: single-mode fiber

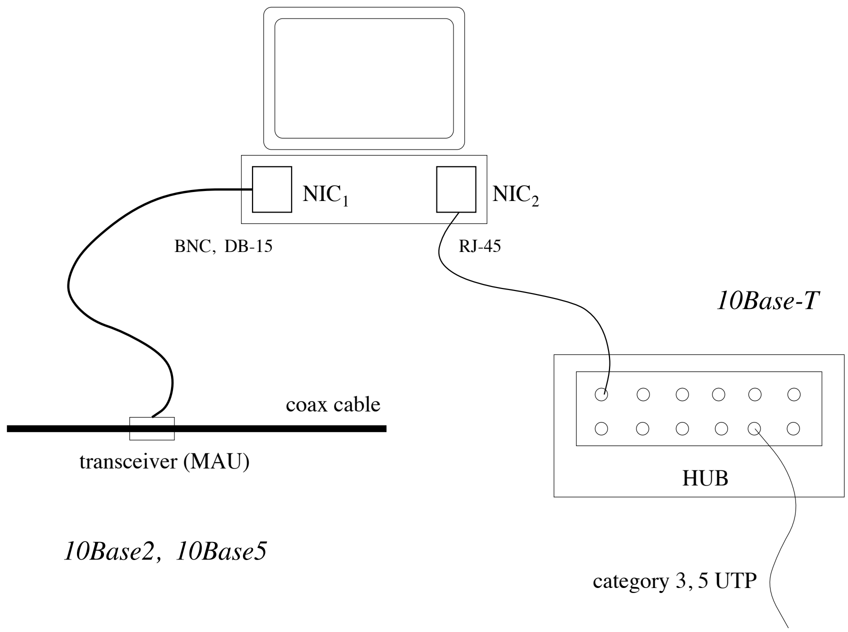
- 10, 40, 100, 400 Gbps Ethernet: fiber; cat 6, 7 copper

  → 100, 200, 400 Gbps: backbone trunk link, data center links

- IEEE 802.3ck: copper
- IEEE 802.3db: multi-mode fiber
- IEEE 802.3cs: single-mode fiber, subscriber network with reach up to 50 km
- IEEE 802.3dj: 200, 400, 800 Gbps, 1.6 Tbps

Active, on-going work, trend.

Ethernet technology evolution:



- single-homed vs. multi-homed
- unique 48-bit Ethernet address per NIC

$\longrightarrow$ ancient stuff but . . .

Technology reused in powerline networks: IEEE 1901, HomePlug industry group

High-speed Ethernets have shorter network diameter:

- 2500 m for 10 Mbps 10Base5 ThickNet
- 925 m for 10 Mbps 10Base2 ThinNet

  → distance limitations: due to Ethernet CSMA/CD protocol

Switched Ethernet: 100 m segment length for 10Base-T, FastEthernet, 1000Base-T GigE

→ full-duplex fiber, copper links

→ signal degradation main limiting factor

Addressing:

- 48 bit unique address

  → called hardware or MAC address

- broadcast address: all 1's

Sender: adds "from" and "to" address

→ source and destination

Receiver: Ethernet NIC accepts frames with matching destination address.

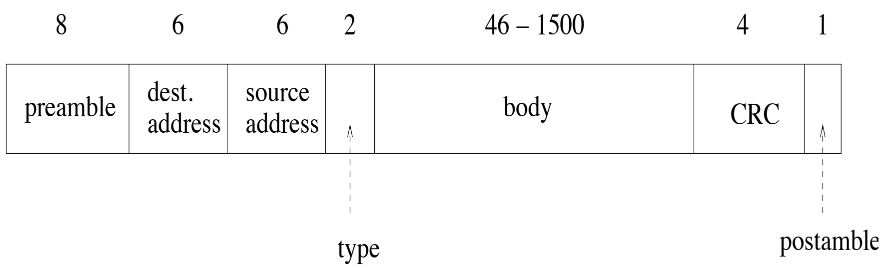
- default
- can accept all frames

  → promiscuous mode

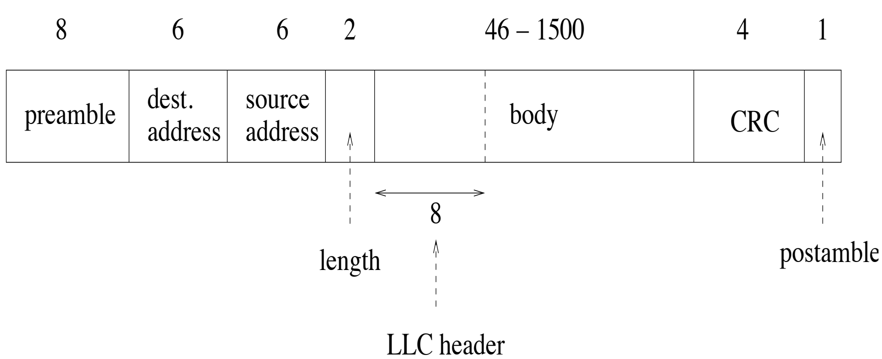
  → requires root privilege

  → useful for traffic monitoring/sniffing

DIX Ethernet frame:



IEEE 802.3 Ethernet frame:



→ type: e.g., 0x0800 for IPv4

→ DIX dominant: incorporated as IEEE standard

Modulation: baseband (square waves)

Cat 4, 10Base-T, Manchester encoding

→ bit rate is 1/2 of symbol rate (baud)

→ use modulation to facilitate clock synchronization

Cat 5, FastEthernet (IEEE 802.3u), 4B5B encoding

→ 4 data bits replaced by 5 code bits

→ 0000 → 11110, 1111 → 11101

→ use coding to facilitate clock synchronization

→ data rate 100 Mbps, symbol rate 125 MHz

→ full-duplex, 2 pairs

Cat 5/5e, GigE (IEEE 802.3ab), 5-level PAM

→ 4 levels for data: 2 bits

→ 5th level for error correction

→ mix of signal modulation and coding

→ data rate 1000 Mbps

→ all (4) pairs, 250 Mbps over each pair

→ full-duplex achieved using hybrid circuits

→ technology from telephony

Today: 10, 40, 100, 400, 800 Gbps Ethernets over fiber and copper.

Ethernet MAC protocol: CSMA/CD

- MA (Multiple Access): multiple nodes are allowed simultaneous access

  → just send

- CS (Carrier Sense): can detect if some other node is using the link

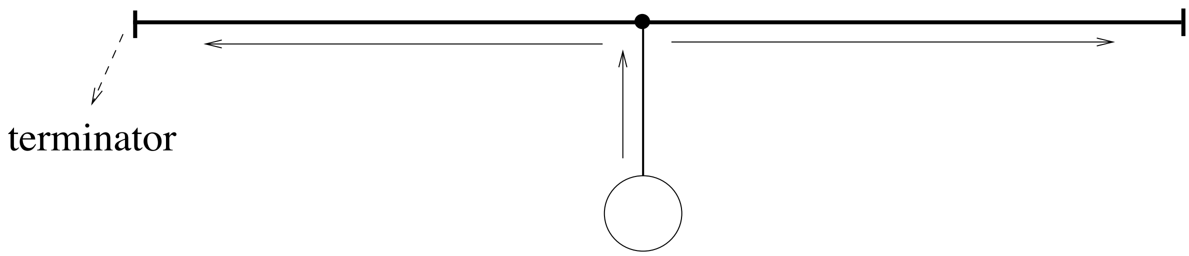
  → rule: if busy, wait until channel is not busy

- CD (Collision Detection): can detect frame collision stemming from simultaneous transmissions

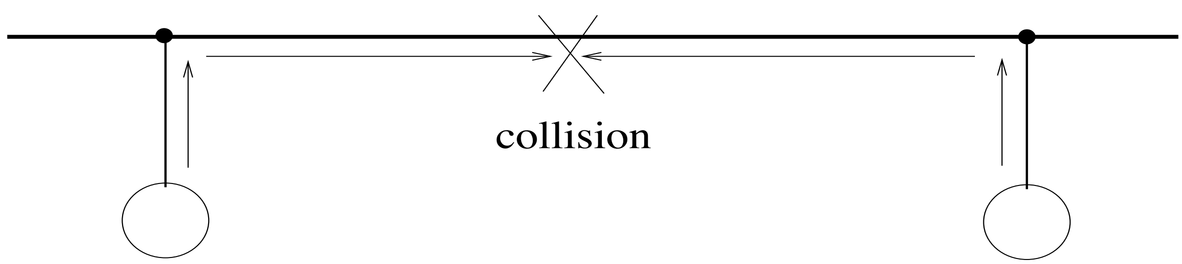
  → rule: if collision, try later

Collision detection mechanism:

→ broadcast signal

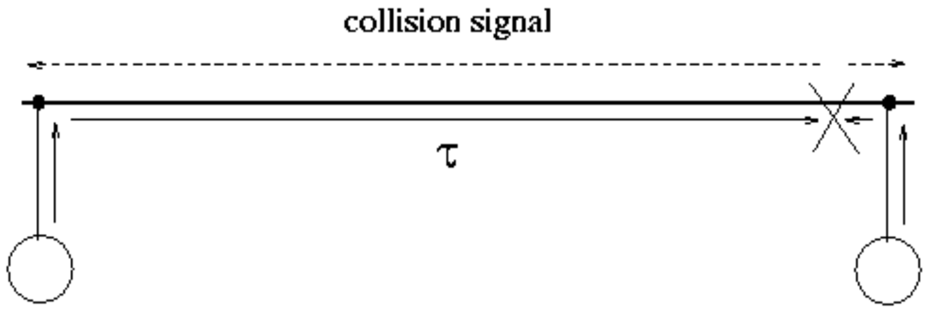


Collision scenario: best-case



→ meet in the middle

Collision scenario: worst-case



- τ: one-way propagation delay
- sender needs to wait 2τ sec before detecting collision

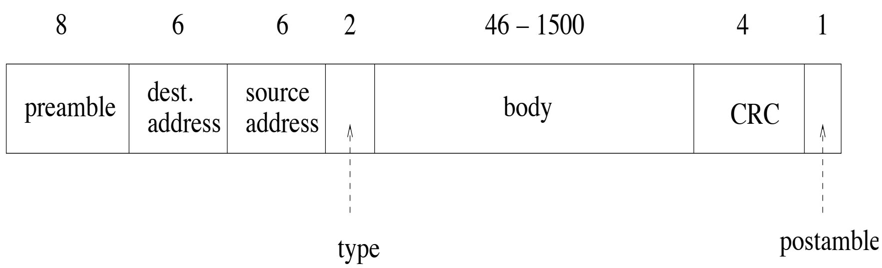
  $\rightarrow$ time for echo to bounce back
- for 2500 m length, 51.2 μs round-trip time (2τ)
- enforce 51.2 μs slot time
- at 10 Mbps, 512 bits: minimum frame size

  $\rightarrow$ assures collision detection

Transmit at least 512 bits for CD:

$\rightarrow 6 + 6 + 2 + 46 + 4 = 64$ B $= 512$ bits

$\rightarrow$ minimum payload size of Ethernet frame



To achieve collision detection (CD) in 100 Mbps Ethernet, what must happen?

1 Gbps, 10 Gbps?

Upon collision: when attempt retransmission?

$\rightarrow$ stop-and-wait with collision signal as negative ACK

Retransmission protocol: exponential backoff

1. Wait for random $0 \leq X \leq 51.2$ μs before 1st retry
2. Two consecutive collisions: wait for random $0 \leq X \leq 102.4$ μs before 2nd retry
3. Three consecutive collisions: wait for random $0 \leq X \leq 204.8$ μs before 3rd retry
4. i consecutive collisions: wait for $0 \leq X \leq 2^{i-1}\,51.2$ μs before next attempt
5. Give up if $i > 16$

$\rightarrow$ why exponential backoff?

$\rightarrow$ how good is throughput of CSMA/CD?

Today: switched Ethernet with full-duplex links

- not shared bus anymore

  → every device connected by point-to-point link to switch

  → sender/receiver cannot collide

  → switch: a computer

  → with special hardware support to speed up packet handling

- arriving Ethernet frames subject to scheduling

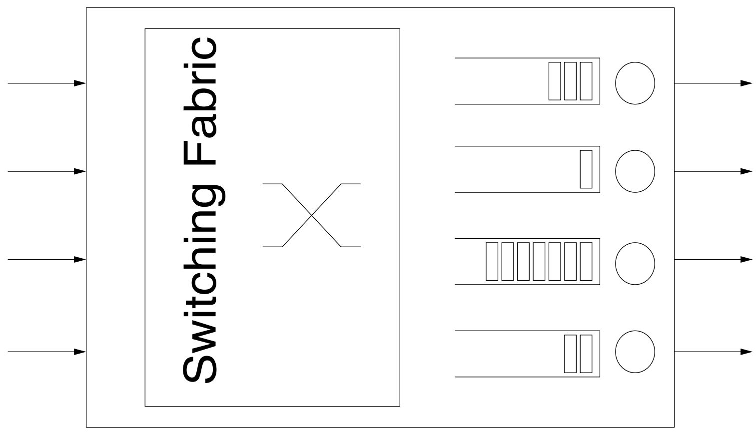
  → e.g., FIFO, priority, fair queueing

  → finite buffers: who is dropped?

  → frame losses occur due to buffer overflow

  → not collision

Example diagram of 4-port switch:



→ output buffered switch

→ switches: both input and output buffers

→ switching fabric: hardware

→ switching: pure hardware, firmware, processes in OS

→ e.g., Cisco's router OS: IOS (Internet OS)

Real-world importance of backward compatibility:

- legacy Ethernet NICs speak CSMA/CD
- switched Ethernet interoperate with legacy NICs

  → constraint of new networking technologies

Links between high-speed Ethernet switches:

→ less complications

→ turn off CSMA/CD

→ Ethernet in short-haul data center, long-haul backbone links

→ making inroad in automotive networks

Issue of congestion control

→ intrinsic

→ subject we will study separately