Direct Link Communication: Technology and Access Control

Point-to-point communication

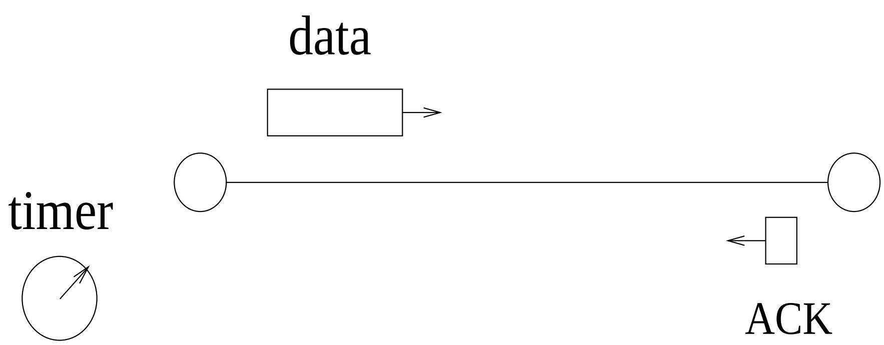
Already seen digital/analog transmission of digital data including coding and error detection.

Reliable transmission

Principal methodology: ARQ (Automatic Repeat reQuest) or PAR (Positive Acknowledgment with Retransmission) or backward error correction (BEC).

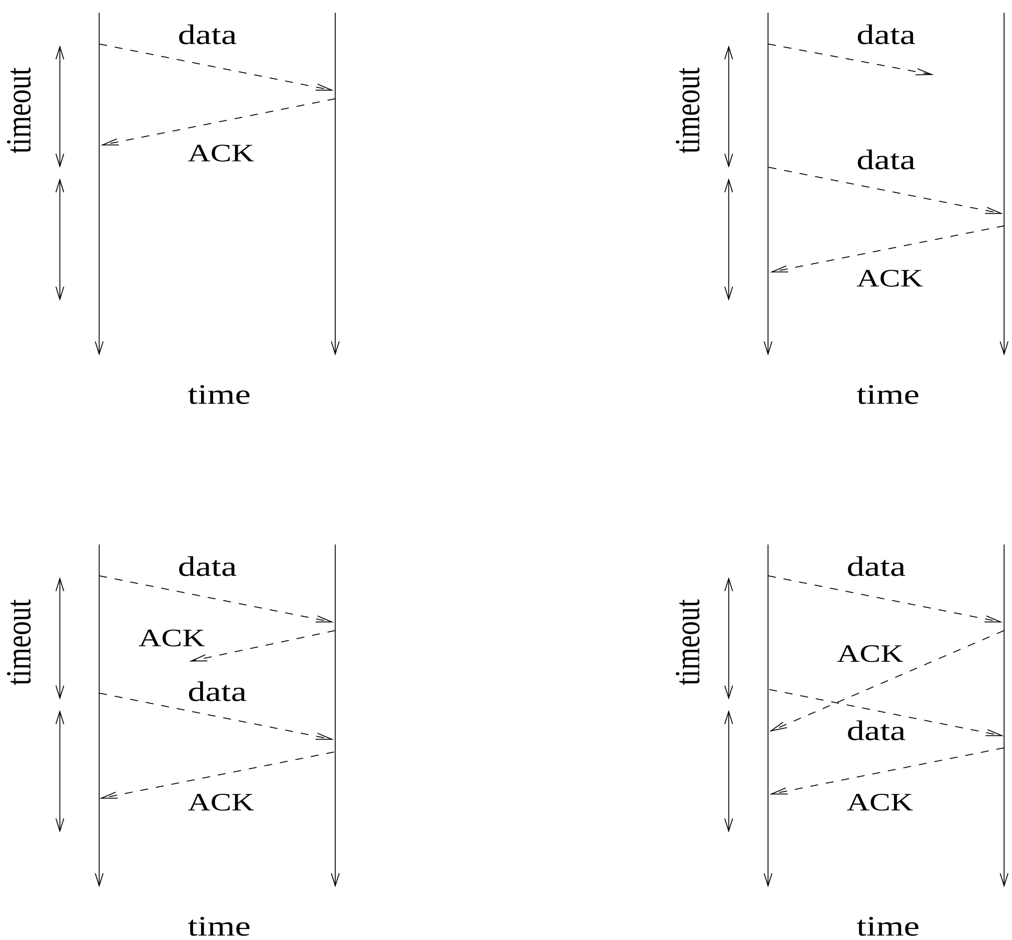
Three components:

- acknowledgment (ACK)
- timeout
- retransmit



Stop-and-wait

Assumption: Frame is "lost" due to corruption; discarded by NIC after error detection.



Issue of RTT (Round-Trip Time) and timer management:

- what is proper value of timer?
- RTT estimation
- easier for single link than *path* in an internetwork
- largely independent of queueing effect

More serious problem: Not keeping the pipe full.

$\longrightarrow$ bandwidth-delay product

Literally, volume of data in on the link.

To achieve high utilization, want to keep volume of traffic flowing close to the bandwidth limit.

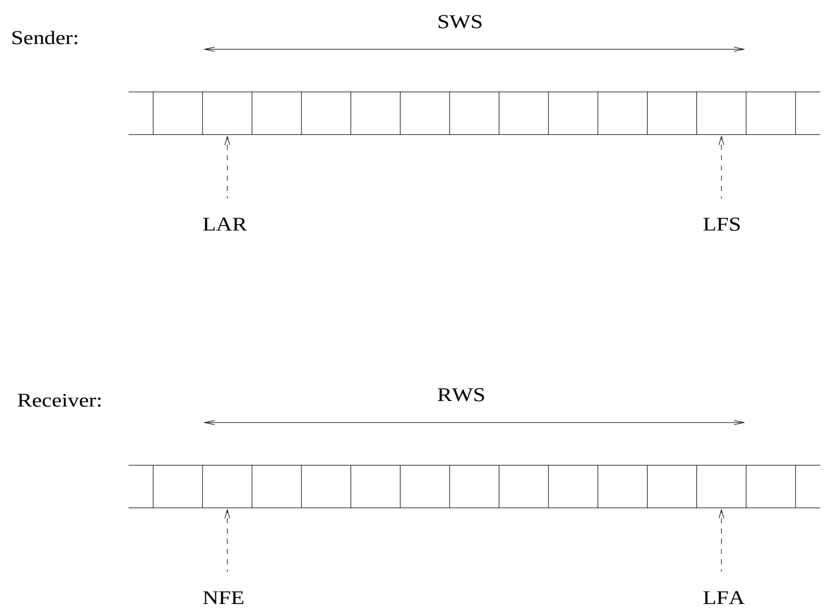
Example: Link BW 1.5 Mbps, 45 ms RTT; bandwidth-delay product = 1.5 Mbps $\times$ 45 ms = 67.5 kb $\approx$ 8 kB.

If frame size 1 kB, then (effective) throughput is $1024 \times 8/0.045 = 182$ kbps; utilization is only 0.125.

Solution: Other things being equal, must increase frame size.

- straightforward increase of frame size is problematic; why?
- send blocks of data, i.e., sequence of frames
- creates management problem

Sliding window protocol



- *SWS*: Send Window Size
- *RWS*: Receiver Window Size
- *LAR*: Last ACK Received
- *LFS*: Last Frame Sent
- *NFE*: Next Frame Expected
- *LFA*: Last Frame Acceptable

Assign sequence number (SeqNum) to individual frames.

Maintain invariants:

- $\text{LFS} - \text{LAR} + 1 \leq \text{SWS}$
- $\text{LFA} - \text{NFE} + 1 \leq \text{RWS}$

Sender: Update LAR, send more frames, then update LFS.

Receiver: Cumulative ACK; let SeqNumToAck denote the largest sequence number not yet acknowledged.

- $\text{NFE} \leftarrow \text{SeqNumToAck} + 1$
- $\text{LFA} \leftarrow \text{SeqNumToAck} + \text{RWS}$

ACK variants:

- piggyback
- negative ACKs
- selective ACKs

Sequence number wrap-around problem:

$$\text{SWS} < (\text{MaxSeqNum} + 1)/2.$$

$\longrightarrow$ similar to stop-and-wait (binary)

Further optimization/control variables in end-to-end case?

Why can packets still be lost given that link layer achieves reliability?

Link-based flow/congestion control revival (H. T. Kung). Achieve flow control/multiplexing (buffer sharing)/reliability at link level.

Multi-access communication

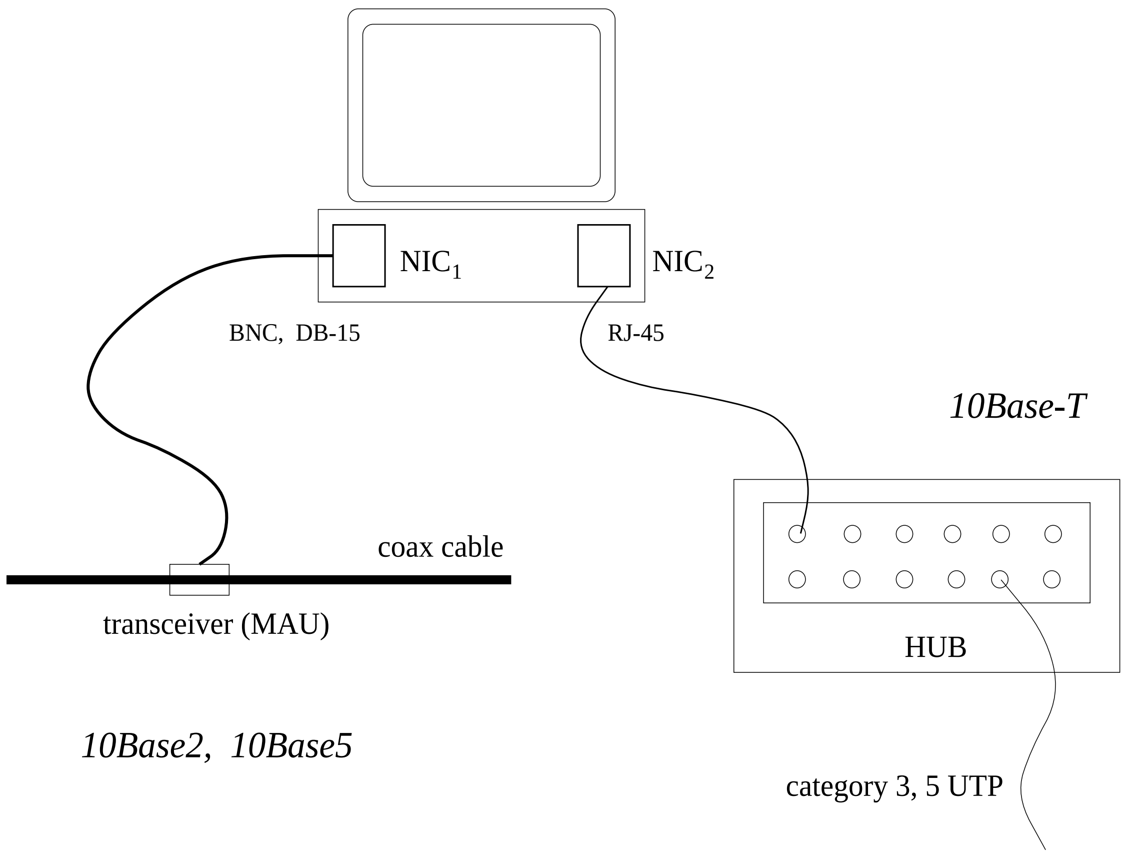
Ethernet and CSMA/CD

$\longrightarrow$ copper, optical fiber

Types:

- 10Base2 (ThinNet): coax, segment length 200 m, 30 nodes/segment
- 10Base5 (ThickNet): coax, segment length 500 m, 100 nodes/segment
- 10Base-T: twisted pair, segment length 100 m, 1024 nodes/segment
- 10Base-F: fiber, segment length 2000 m, 1024 nodes/segment
- 100Base-T (Fast Ethernet): category 5 UTP, fiber (also 100VG-AnyLAN)
- Gigabit Ethernet: fiber, category 5 UTP

Connectivity example:



⟶ bus/star configuration

⟶ multihomed/singlehomed

⟶ unique Ethernet address per NIC

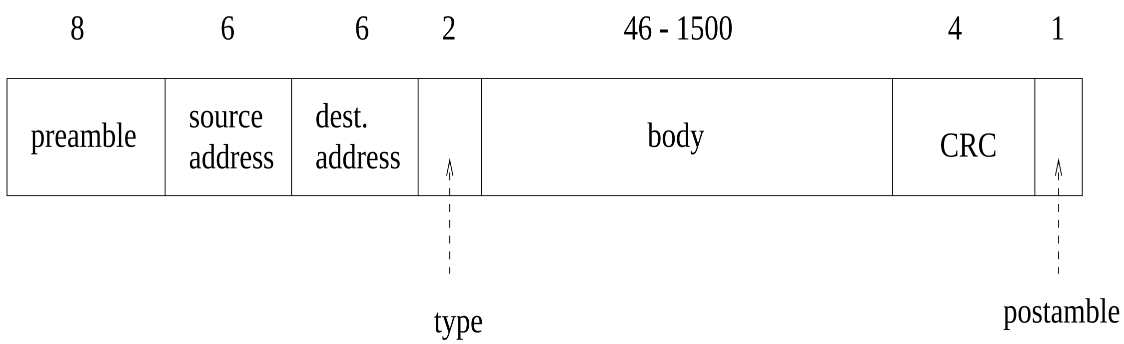
Segments can be hooked up by repeaters, bridges, gateways, (hub) switches.

- maximum of 2 (4 for IEEE 802.3) repeaters between two hosts; 1500 m
- for Fast Ethernet, 2 repeater hops

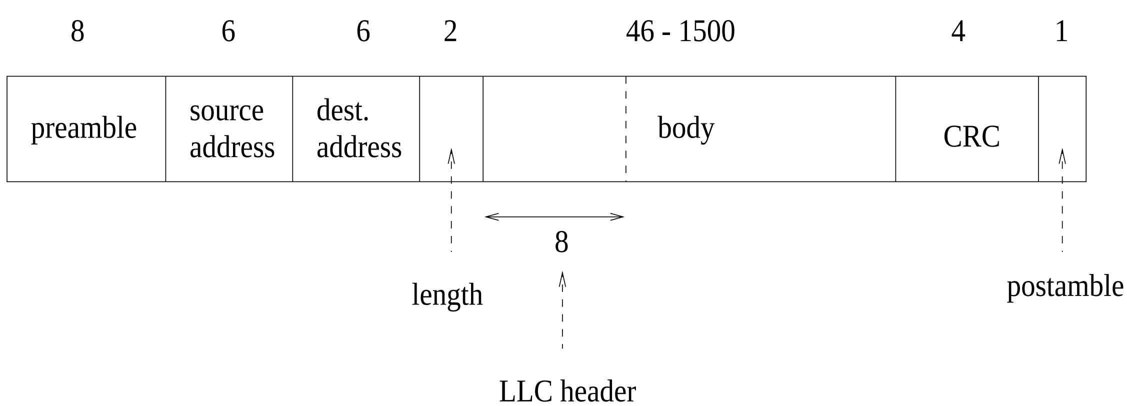
High-bandwidth Ethernets have *shorter* network diameter.

- about 2500 m for 10 Mbps Ethernet
- about 200 m for 100 Mbps Ethernet
- even shorter for 1 Gbps Ethernet

DIX Ethernet frame:



IEEE 802.3 Ethernet frame:



$\longrightarrow$ IEEE 802.2 LLC (Logical Link Control)

Encoding: Manchester

Addressing:

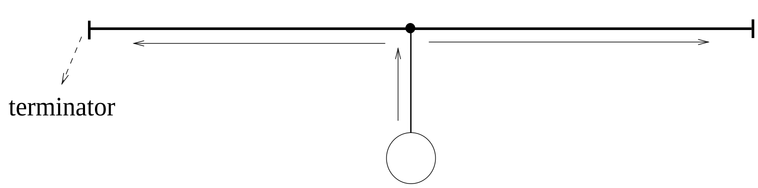
- 48 bit unique address
- point-to-point
- broadcast (all 1's)

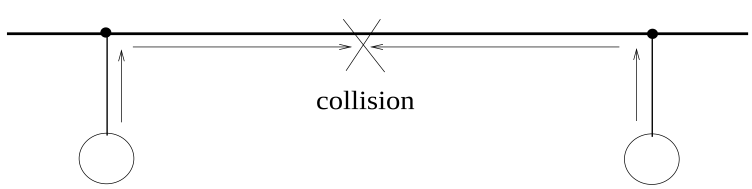
Receiver: Ethernet adaptor accepts frames with relevant address.

- accepts only own frame address
- accepts all frames (promiscuous mode)

MAC (Medium Access Control): CSMA/CD

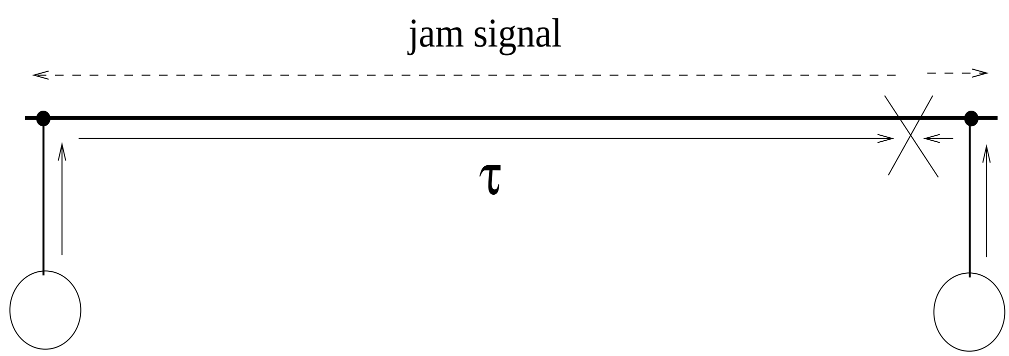
- *CS (Carrier Sense)*: Can detect if some other node is using the link.
- *MA (Multiple Access)*: Many nodes are allowed to simultaneously access the link.
- *CD (Collision Detection)*: Can detect if simultaneous access has occured (corrupted signal).





Ethernet is 1-*persistent* MA scheme; more generally, p-persistent where $0 < p \leq 1$.

Worst-case collision scenario:



- sender (worst case) needs to wait 2τ sec before detecting collision
- for 2500 m length, 51.2 μs round-trip time (2τ)
- enforce 51.2 μs slot (jam) time
- at 10 Mbps, 512 bits; i.e., minimum frame size

Hence, upon collision detection:

- Make sure to transmit at least 512 bits

  $\longrightarrow$ $2 \times$ bandwidth-delay product

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

- exponential backoff; wait for $0 \leq X \leq 51.2$ μs before next attempt
- on i'th collision, wait for $0 \leq X \leq (2^i - 1)51.2$ μs before next attempt, $i \leq 16$
- $X = 0\ \mu\text{s}, 51.2\ \mu\text{s}, 2 \times 51.2\ \mu\text{s}, 3 \times 51.2\ \mu\text{s}, \ldots$

$\longrightarrow$ distributed bus arbitration mechanism

Performance (approximate) analysis of CSMA/CD

Assumptions:

- *contention slot* (or mini slot) of duration 2τ (τ: end-to-end propagation delay)
- k hosts transmitting with constant probability p at every contention slot—slotted system

Fix contention slot; probability that a fixed host acquires slot successfully $p(1-p)^{k-1}$.

Hence, probability that some host acquires a given slot

$$\eta = kp(1-p)^{k-1}.$$

Fact: η is maximized at $p = 1/k$. Also,

$$\lim_{k \to \infty} \eta = 1/e.$$

Probability that contention time interval T (multiple of 2τ) lasts for exactly i contention slots $\eta(1-\eta)^{i-1}$.

Hence, mean contention duration (in units of 2τ)

$$\mathbf{E}(T) = \sum_{i=0}^{\infty} i\eta(1-\eta)^{i-1} = 1/\eta.$$

Mean contention duration $2\tau/\eta \approx 2\tau e$ (sec).

If mean frame transmission time is t_0 (sec), then

$$\rho = \frac{t_0}{t_0 + 2\tau e}$$

If mean frame length is F, bandwidth B, length of medium L, signal propagation speed c,

$$\rho = \frac{1}{1 + (2e/c)BL/F}$$

If B is increased, then either L must decrease and/or F must increase.

$\longrightarrow$ bad news for high-bandwidth media

In practice: debate on actual throughput of Ethernet (above 90% vs. much lower).

Current trend: use switching hubs (Ethernet switches).

- multiple collision domains ("multi-port bridge")
- switching (buffering at Ethernet switches)
- conceptually little distinction with full-fledged switching (e.g., ATM)
- realization of virtual LANs

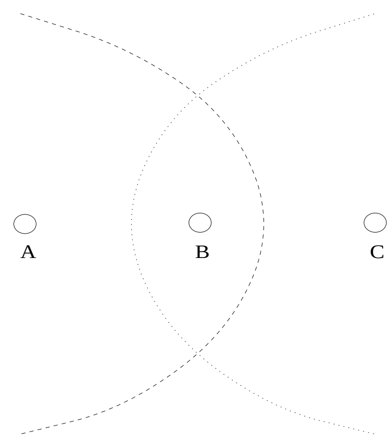
Gigabit and Fast Ethernet: use compendium of tricks. However, subject to instrinsic limitations.

- Fast Ethernet (100Base-T) uses various physical layer options including that of FDDI (emulate CSMA/CD)
- 100Base-T uses same frame size
- 100VG-AnyLAN (IEEE 802.12) uses priority scheduling (not CSMA/CD)
- gigabit Ethernets use broadband signalling
- maintaining consistent frame format & backward compatibility is an important factor

ALOHA: '70s Abramson et al. (Univ. of Hawaii); simplest form of multi-access system using radio waves.

- send frame (pure or slotted).
- upon collision, retry after random waiting time.
- missing component: carrier sense (CS)
- subject to perils of long propagation time (in msec)
- 37 % utilization limit (slotted)

Wireless LAN: short-range radio.



⟶ two problems with using simple CSMA

- hidden station problem
- exposed station problem

Hidden station problem

1. A transmits to B
2. C does not sense A; transmits to B
3. interference occurs at B

Exposed station problem

1. B transmits to A
2. C wants to transmits to D (to its right) but senses B
3. C refrains from transmitting although ok

MACA (Multiple Access with Collision Avoidance)

- sender sends short RTS (request to send) frame
- receiver sends short CTS (clear to send) frame to sender
- others sensing RTS or CTS wait for suitable duration
- collision can still occur on control messages
- exponential backoff

Next: *collision-free (reservation-based) protocols*

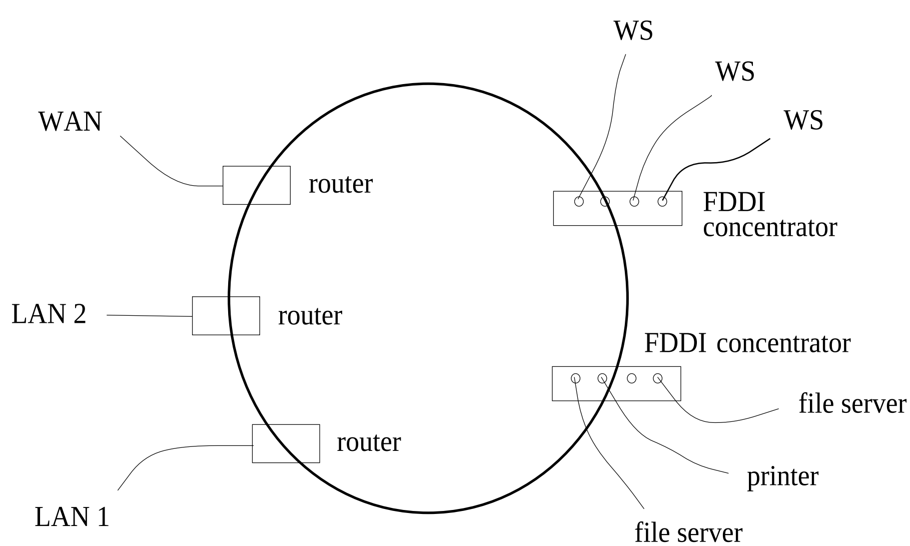
FDDI (Fiber Distributed Data Interface)

$\longrightarrow$ token ring architecture

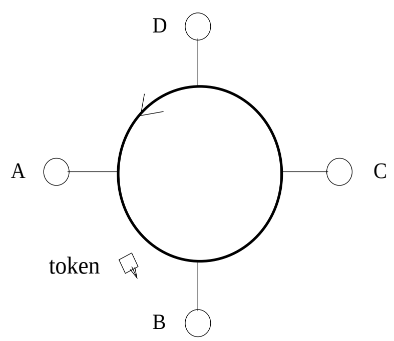
High-bandwidth extension of IBM 4 Mbps token ring and 16 Mbps IEEE 802.5 token ring standard.

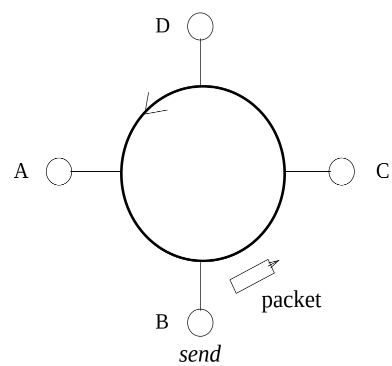
$\longrightarrow$ 100 Mbps bandwidth

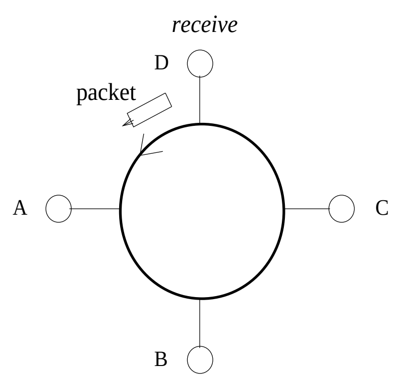
Mostly used as high-bandwidth LAN backbone.

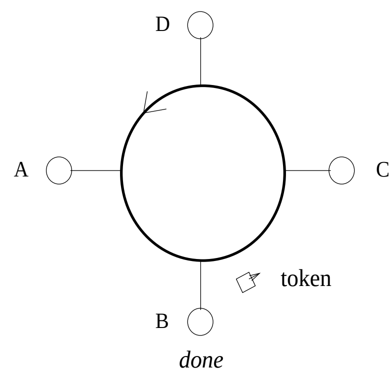


Basic operation:

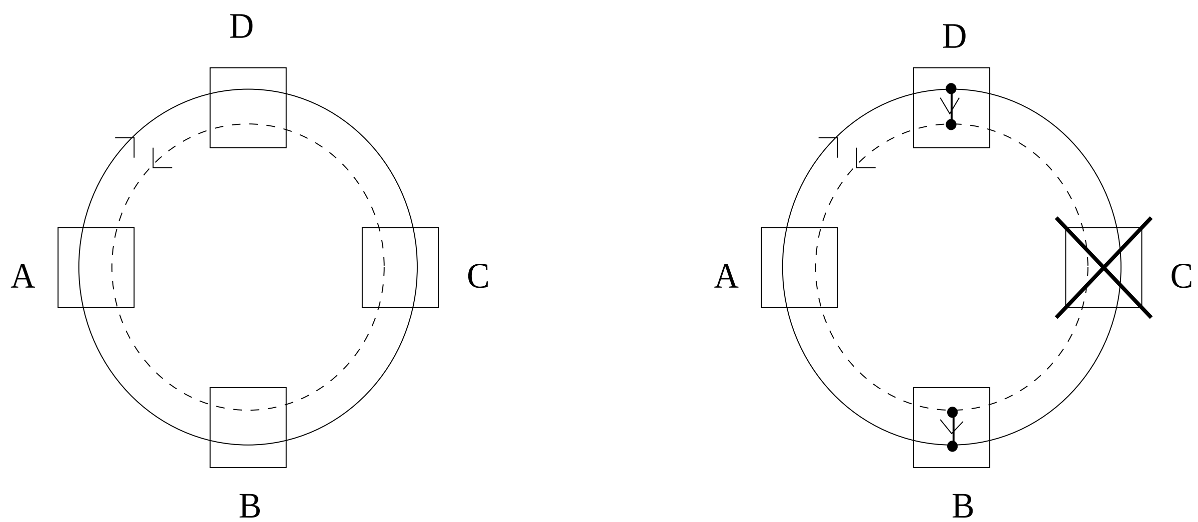




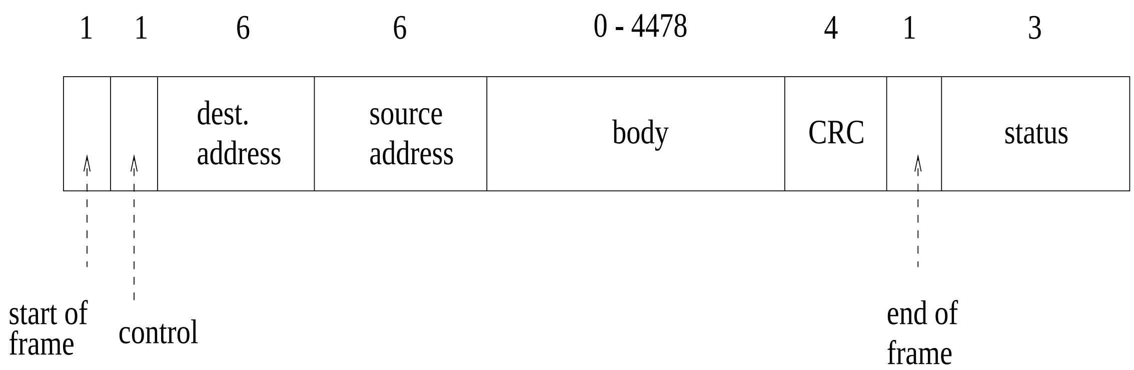




Fault-tolerance:



- DAS (dual attachment station)
- SAS (single attachment station)



- frame size < 4500 B
- 4B/5B encoding
- ANSI
- supports IEEE 802.2 LLC
- synchronous/asynchronous data
- 2 km inter-station distance
- 200 km diameter (multimode fiber); 100 km circumference

Performance issues: fairness and efficiency

- TRT (token rotation time)
- THT (token holding time)

$$\text{TRT} = \text{no. of nodes} \times \text{THT} + \text{link latency}$$

To increase efficiency: increase THT

$\longrightarrow$ let station send as much as it needs

$\longrightarrow$ $\text{THT} \uparrow \implies \rho \uparrow$

To increase fairness: limit THT

$\longrightarrow$ limit station's one-time sending of data

To facilitate fairness: introduce TTRT (target token rotation time).

THT determining factor (assume TTRT is given):

- If TRT $>$ TTRT, then late; don't send asynchronous data.
- If TRT $\leq$ TTRT, then early; send asynchronous data for max $\{$ TTRT $-$ TRT, single frame time $\}$ duration.
- Synchronous frames get always sent.

How to set TTRT?

$\longrightarrow$ token claim process

$\longrightarrow$ initiate when needed (e.g., start-up)

- Each station submits *claim frame* containing TTRT *bid*.
- Smaller TTRT bid overrides higher TTRT bids.
  - Compare claim frame bid against own desired TTRT.
  - If less, then reset own TTRT to lower value.
  - If larger, then put lower bid on claim frame and forward.
- Winner: same bid value when claim frame makes full circle.

At the end of the day, consistent TTRT value among all stations.

$\longrightarrow$ consensus problem

Last problem: When to reinsert token after sending data frame?

- immediate release (FDDI)
- delayed release (IEEE 802.5 token ring); problem of bit time and draining

"Easier" to implement quality of service (QoS).

$\longrightarrow$ real-time issues and QoS guarantees

$\longrightarrow$ partial solution only

Compare against Ethernet's CSMA/CD.

$\longrightarrow$ determinism vs. indeterminism

Throughput: under *infinite-source model*, as long as THT is not "too small," can approach close to 100% utilization.

$\longrightarrow$ what is "not too small"?

Bit length (duration of single bit on medium):

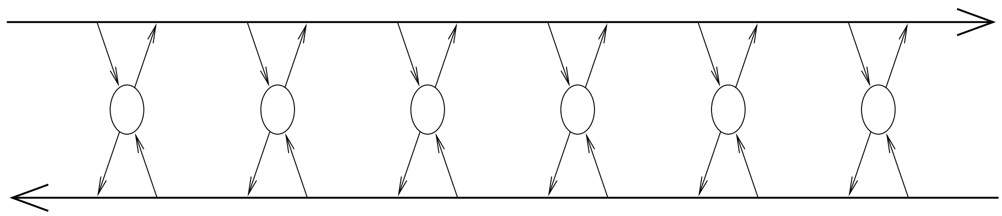
- Bandwidth x Mbps, a bit sent every $1/x$ μsec (bit time).
- Signal propagation speed c m/μsec, bit length c/x m.
- If $c = 300$, $x = 100$, then bit length 3 m.

What about infinite-source model in the case of Ethernet?

IEEE 802.4 Token bus

- Physical bus, logical ring.
- Token-based bus arbitration.

Distributed queue dual bus (DQDB)



Specifications:

- Two 150 Mbps unidirectional buses.
- 150 km.
- 53 B frames.
- IEEE 802.6.

Performance: Under infinite source model, very efficient.

$\longrightarrow$ slotted system (cf. STDM)

Issue: fairness.

$\longrightarrow$ achieve FIFO queue scheduling

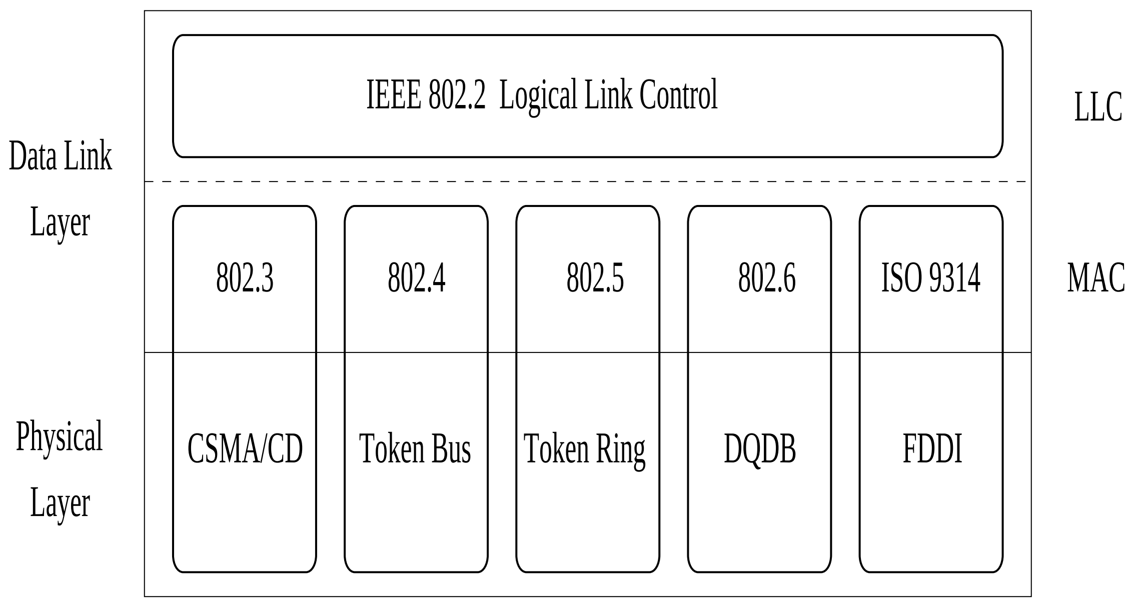
Method (assume sending to the right):

- Keep two counters RC (request counter), CD (count-down counter).
- To send to the right, mark request bit in frame of left bus.
- If request frame passes by on the left bus, increment RC.
- If request was made and CD $= 0$, then seize next free frame.
- If request was made but CD > 0, then decrement CD upon seeing next unoccupied frame.

Intuitive meaning of RC, CD?

So far, we have looked at architecture (hardware) and algorithms of direct-link media, i.e., data link layer.

Protocol specification of data link layer:



Logical link control: Isolation of common generic mechanisms including reliability and flow control.

$\longrightarrow$ similar to HDLC (point-to-point)

Medium access control: random access, deterministic access.

Lessons to be drawn:

- Although theory of direct-link communication exists, oftentimes, implementation is arbitrary and technology dependent (e.g., dual ring of FDDI, baseband/broadband & switched Ethernet).
- Emphasis on ease of standardization, implementation, and cost over efficiency.
- Efficiency pressure is a recent phenomenon. Integrated performance measure (QoS) has become important.
- Industry trend: increase bandwidth more and more; problematic fix to the problem.
- Computer networking tendency: shift away from data link issues (e.g., ATM over TCP/IP).