Direct Link Communication: Technology and Access Control

Point-to-point communication

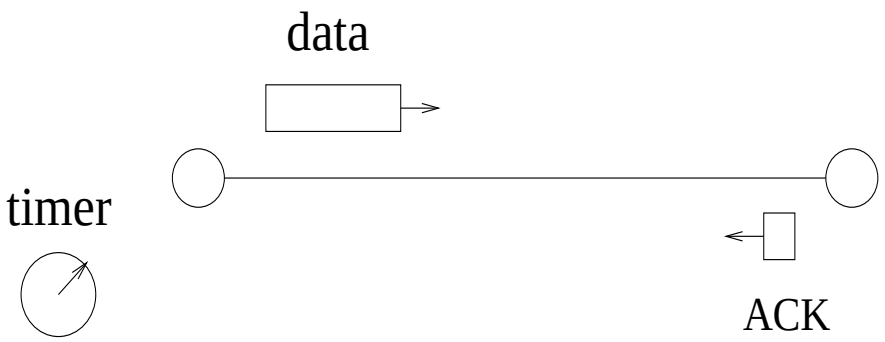
Already seen digital and analog transmission of digitial data including coding and error detection.

Reliable transmission

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

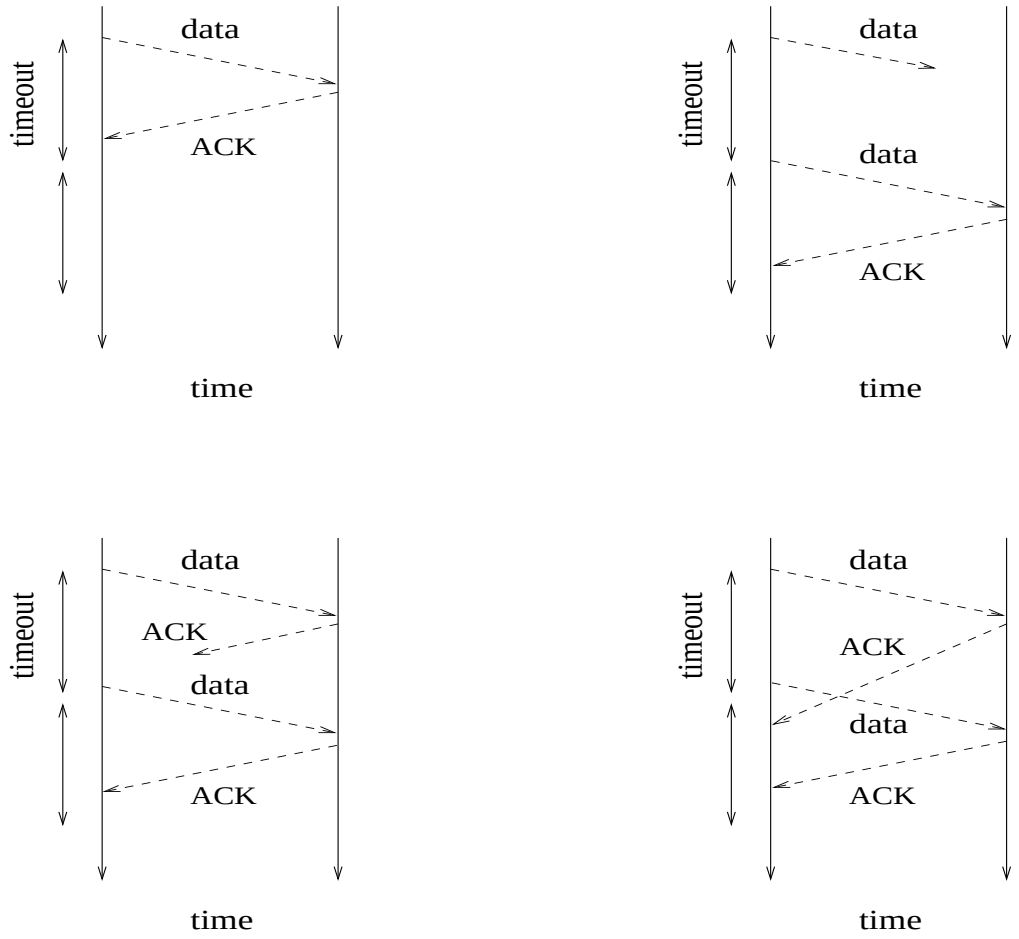
Three items:

- 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):

- Easier for single link than *path* in an internetwork.
- Still, a function of queueing effect at receiver.

More serious problem: Not keeping the pipe full.

$\longrightarrow$ bandwidth-delay product

Literally, volume of data in on the link.

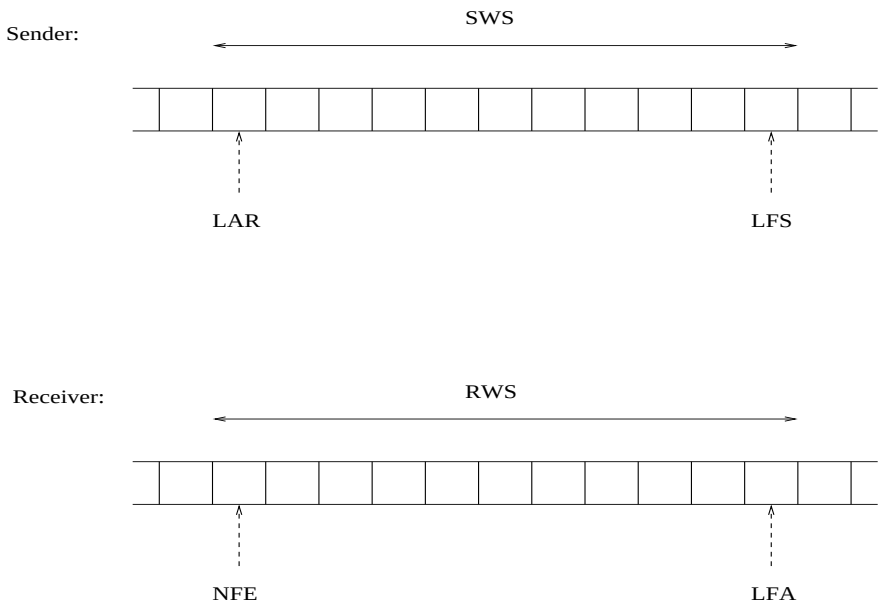
Example: Link BW 1.5Mbps, 45ms RTT; bandwidth-delay product = 1.5Mbps $\times$ 45ms = 67.5kb $\approx$ 8kB (using twice the delay).

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

Solution:

$\longrightarrow$ send blocks of data

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: Increment LAR, then LFS.

Receiver: Cumulative ACK; $SeqNumToAck$ is 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)

Credit-based link control revival (H. T. Kung). Achieve absolute reliability at link level.

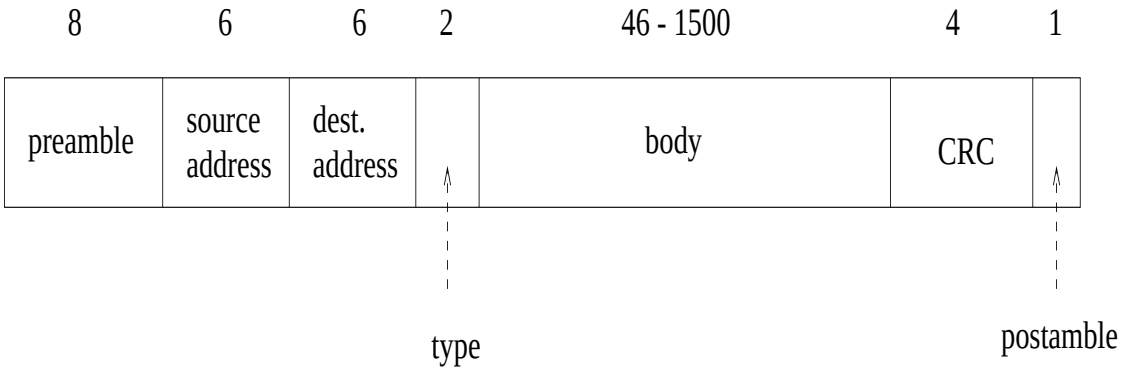
Multi-access communication

Ethernet and CSMA/CD

Types:

- 10Base5: coax, segment length 500m, 100 nodes/segment
- 10Base2: coax, segment length 200m, 30 nodes/segment
- 10Base-T: twisted pair, segment length 100m, 1024 nodes/segment
- 10Base-F: fiber, segment length 2000m, 1024 nodes/segment

Segments can be hooked up by repeaters, bridges, gateways.



Addressing:

- point-to-point
- broadcast
- multicast

Receiver side: Ethernet adaptor accepts frames with relevant address.

$\longrightarrow$ promiscuous mode

Sender side:

- *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 $0 < p \leq 1$.

51.2μs wait period after sending one frame.

Upon collision detection:

- Make sure to transmit at least 512 bits (jam time); on 1500m Ethernet, propagation delay 51.2μs.

  $\longrightarrow$ bandwidth-delay product: 512 bits

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

- Exponential backoff; wait for $0 \leq X \leq 51.2\mu$s before next attempt.
- On i'th collision, wait for $0 \leq X \leq (2^i - 1)51.2\mu$s before next attempt, $i \leq 16$.
- $X \leq 1023 \times 51.2\mu$s.
- $X = 0\mu\text{s}, 51.2\mu\text{s}, 2 \times 51.2\mu\text{s}, 3 \times 51.2\mu\text{s}, \ldots$.

  $\longrightarrow$ distributed bus arbitration mechanism

Performance (approximate) analysis of Ethernet 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 $\rightarrow$ 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 $2\tau/\eta \leq 2\tau e$ (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,

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

$\longrightarrow$ 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).

$\longrightarrow$ multiple collision domains

Fast Ethernet: 100Base-T (100m), 100Base-F (2000m)

$\longrightarrow$ clock rate problem

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

- Send frame (pure or slotted).
- Upon collision, retry after random waiting time.

Crucial missing component: carrier sense (CS)

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

FDDI (Fiber Distributed Data Interface)

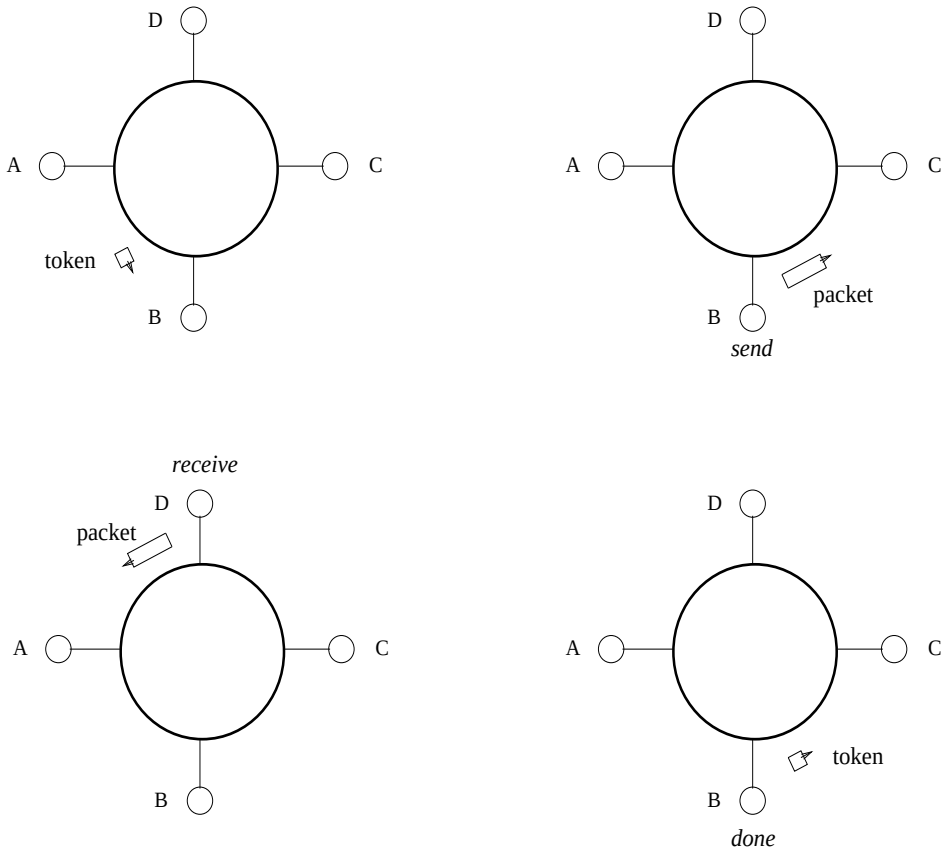
$\longrightarrow$ token ring architecture

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

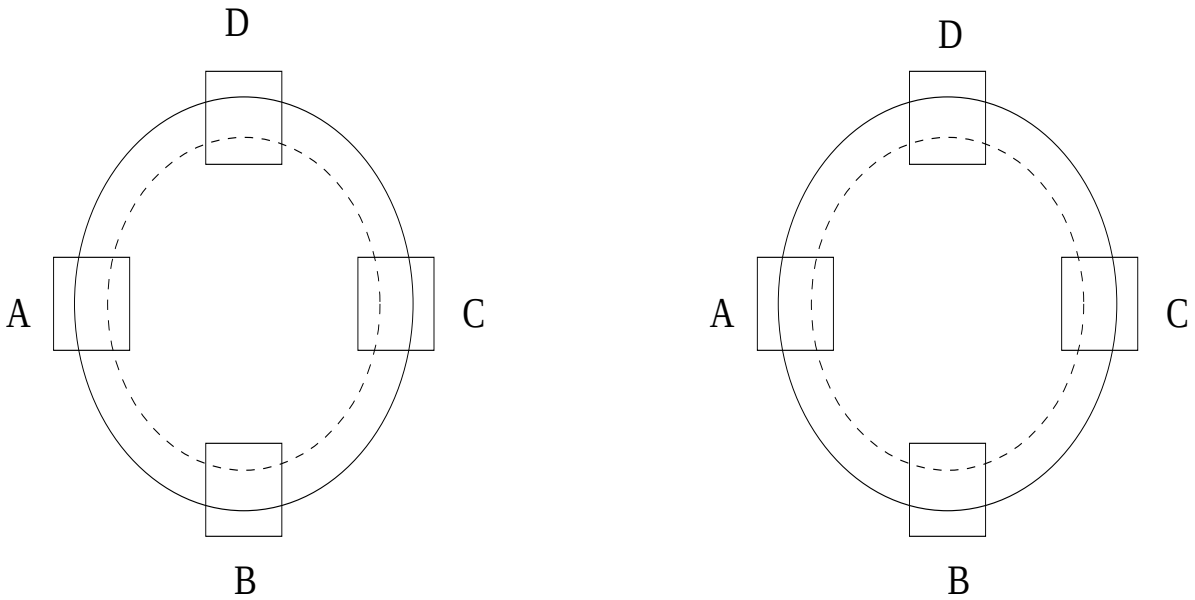
$\longrightarrow$ 100Mbps bandwidth

Mostly used as high-bandwidth LAN backbone.

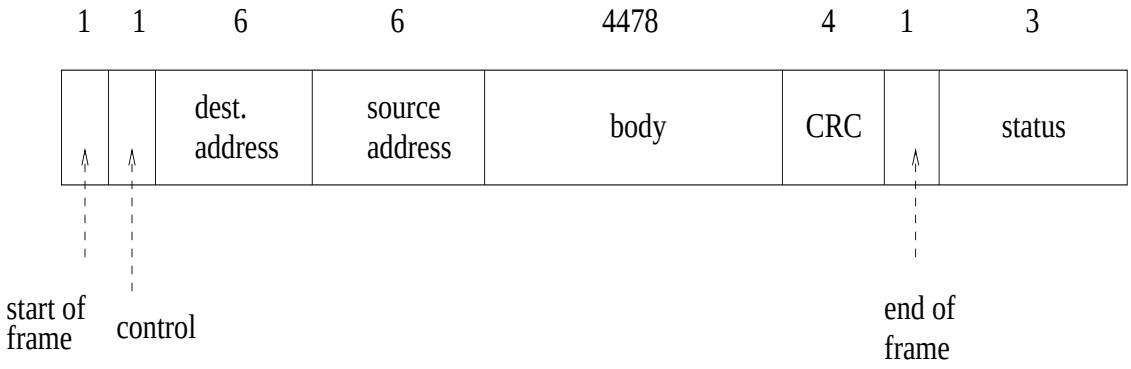
Basic operation



Fault-tolerance:



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



Performance issues: fairness and efficiency

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

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

THT determining factor:

- 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 } time.
- Synchronous frames get always sent.

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

$\longrightarrow$ real-time issues and QoS guarantees

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.
- Signal propagation speed c m/μsec, bit time c/x m.
- If $c = 300$, $x = 100$, then bit length 3 m.

Compare with IEEE 802.5 and IBM token ring; due to smaller bandwidth and medium length, wrap-around problem; use *draining*.

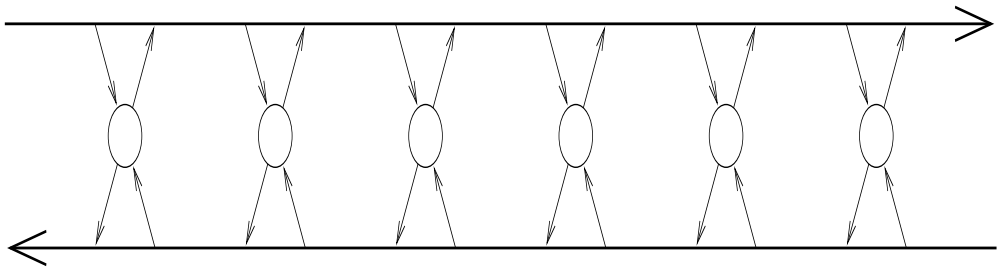
Token maintenance: *token claim process* determines TTRT.

- Smaller TTRT bid overrides higher TTRT bids.
- Winner: same bid value when token arrives.

IEEE 802.4 Token bus

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

Distributed queue dual bus (DQDB)



Specifications:

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

Performance: Under infinite source model, very efficient.

$\longrightarrow$ slotted system (cf. STDM)

Issue: fairness.

$\longrightarrow$ achieve FIFO queue scheduling

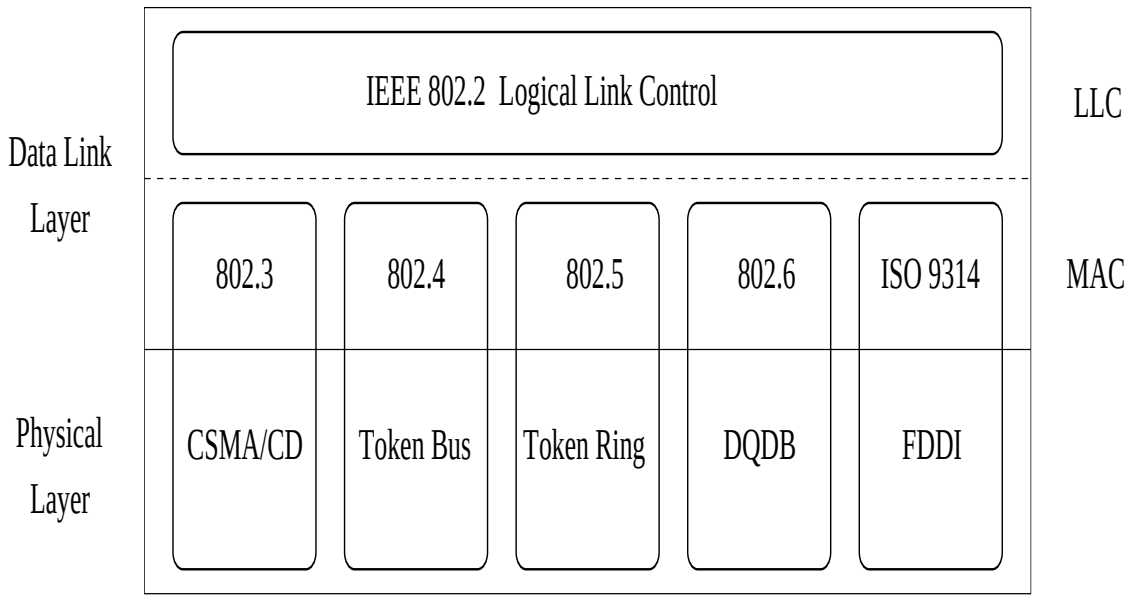
Method (assume sending to the right):

- 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 genetic mechanisms including reliability and flow control.

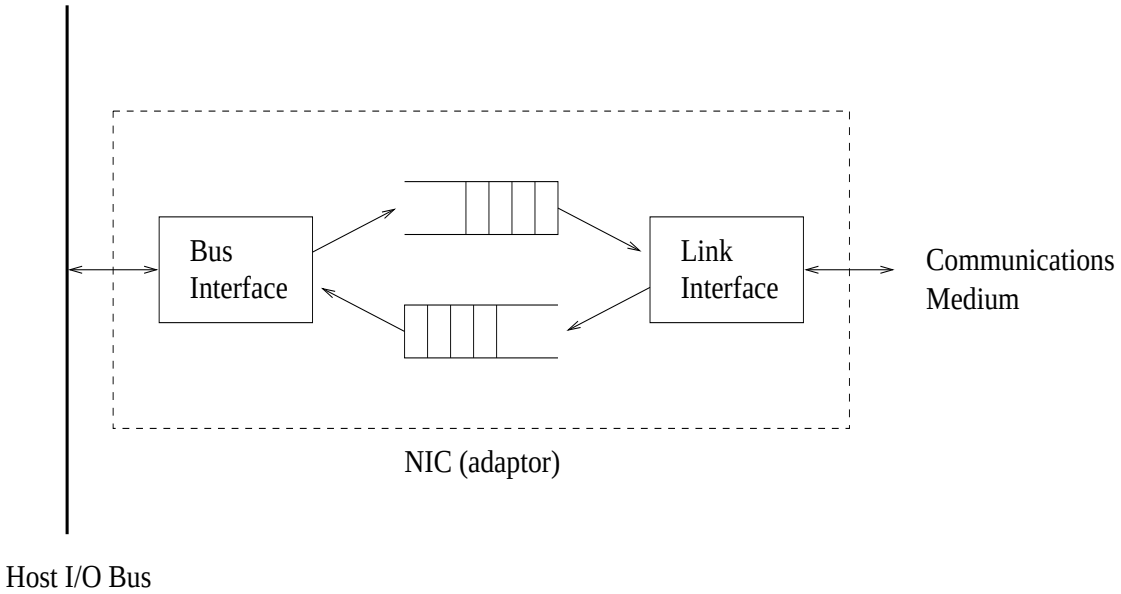
$\longrightarrow$ similar to HDLC (point-to-point)

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, Manchester encoding in Ethernet).
- Emphasis on ease of standardization, implementation, and cost over efficiency.
- Efficiency pressure is a recent phenomenon. Integrated performance measure—quality-of-service (QoS)—has become critical.
- 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).

Network adaptors



State of NIC is captured by *control status register* (CSR).

Polling vs. interrupts.

$\longrightarrow$ interrupt handler does minimal work

Direct memory access (DMA) vs. programmed I/O.

- DMA: Minimal buffering. Use buffer descriptor list to directly manipulate main memory.
- Programmed I/O: heavy burden on on-board buffer space.