CO7219: Internet and Cloud Computing

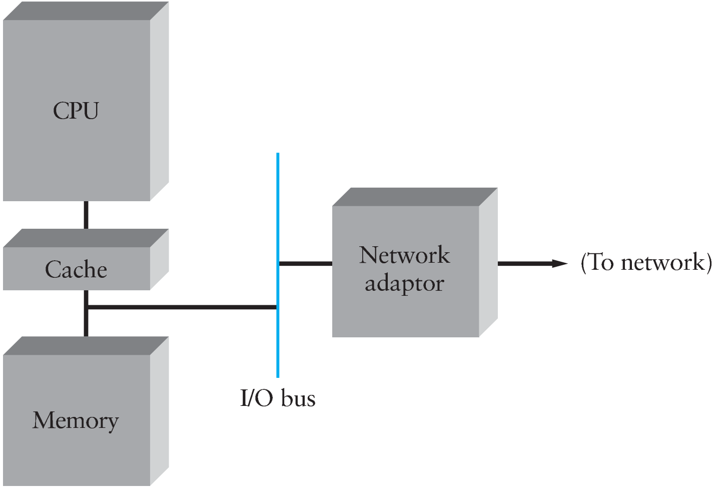
2. Sliding Window, Ethernet and WiFi

Direct Link Issues

- Simplest network: Two or more hosts directly connected by some physical medium.
- Five main issues:
  - **Encoding** bits so that they can be understood by a receiver.
  - **Framing**: delineating the sequence of bits into complete messages that can be delivered to the end node.
  - **Error detection**
  - Make the link appear **reliable** in spite of occasional corruption of frames.
  - **Media access control**: coordinate access to the link among multiple hosts

Hardware Building Blocks

- Typical node: general-purpose computer



Reliable Transmission

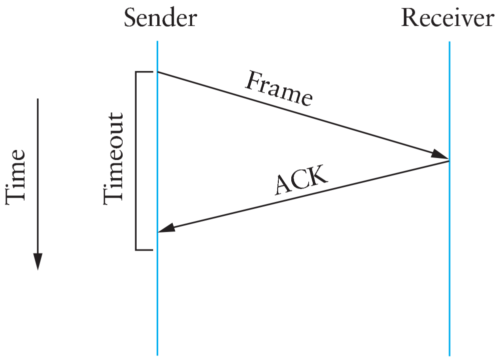
- Frames are sometimes corrupted (and this is detected using CRC or other error-detecting codes).
- Even when error-correcting codes are used, some errors are too severe to be corrected.
- A link-level protocol that wants to deliver frames reliably must somehow recover from lost frames.
- Two fundamental mechanisms:
  - **Acknowledgment** (ACK): Receiver confirms that a frame has arrived. The ACK is either a control frame (only header) or piggybacked on a data frame.
  - **Timeout**: If no ACK is received within a certain time, the original frame is retransmitted.

Approach is called **automatic repeat request** (ARQ).

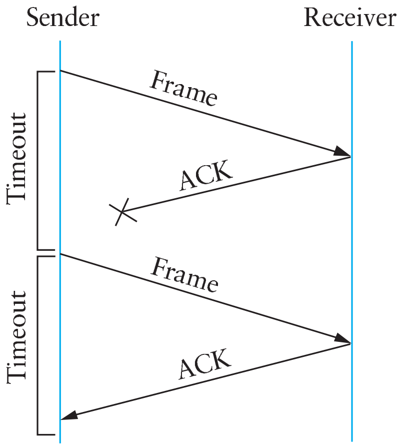
Stop-and-Wait

- Stop-and-wait is the simplest ARQ scheme.
- After transmitting one frame, the sender waits for an acknowledgment before transmitting the next frame.
- If the acknowledgment does not arrive after a certain time, the sender times out and retransmits the original frame.

Stop-and-Wait Scenarios

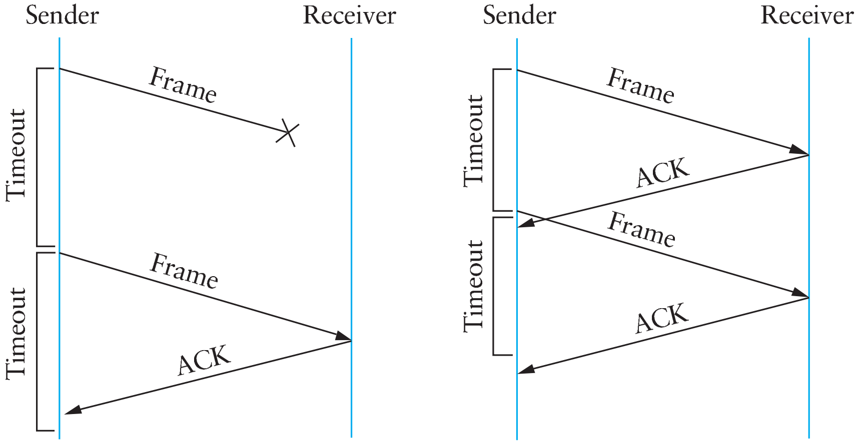


(a)



(c)

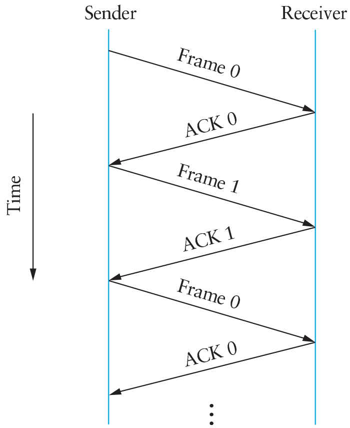
Stop-and-Wait Scenarios II



(b)   (d)

Sequence Numbers

- In scenarios (c) and (d), the receiver gets two copies of the same frame.
- In order to recognize such duplicates, 1-bit sequence numbers are used:



Shortcoming of Stop-and-Wait

- Stop-and-wait allows the sender to have only one outstanding frame on the link at any time.
- For example, on a 1.5-Mbps link with RTT 45 ms, a frame size of 1 KB implies a maximum sending rate of

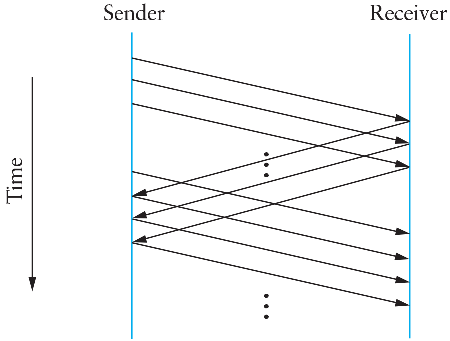
$$\frac{\text{BitsPerFrame}}{\text{TimePerFrame}} = \frac{1024 \times 8 \text{ bits}}{0.045\text{s}} \approx 182\text{Kbps}$$

  This is only one-eighth of the link capacity!

  Note that the delay $\times$ bandwidth product is 45 ms $\times$ 1.5 Mbps $=$ 67500 bits $\approx$ 8 KB for this link.

Sliding Window

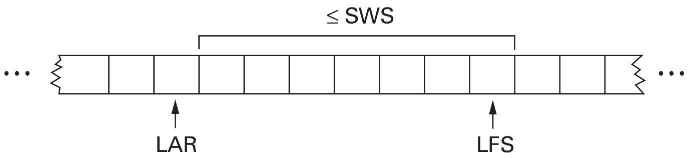
- Consider again the link with delay $\times$ bandwidth product of 8 KB.
- We would like to "keep the pipe full" and send the ninth frame roughly when the first acknowledgment arrives.



Sliding Window Algorithm

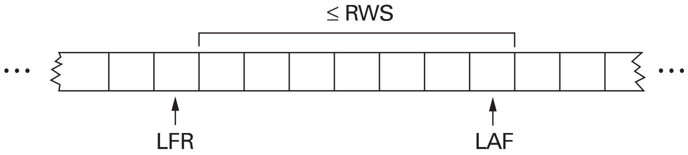
- Assign sequence number SeqNum to each frame.
- Sender maintains:
  - SWS (send window size): upper bound on number of outstanding (unacknowledged) frames that the sender can transmit
  - LAR (last acknowledgment received): sequence number of the last frame for which an ACK was received
  - LFS (last frame sent): largest sequence number of a frame that has been sent.
- Invariant: LFS – LAR $\leq$ SWS

Sliding Window on Sender



- When an ACK arrives, LAR moves to the right and the sender can send another frame.
- Sender associates a timer with each frame it transmits, and a frame is retransmitted if the timer expires before an ACK is received.
- Sender must be willing to buffer up to SWS frames, as they may need to be retransmitted.

Sliding Window on Receiver



- Receiver maintains:
  - RWS (receive window size): upper bound on number of out-of-order frames the receiver is willing to accept.
  - LAF (largest acceptable frame)
  - LFR (last frame received)
- Invariant: LAF – LFR $\leq$ RWS

Receiver Behaviour

- When a frame with sequence number SeqNum arrives:
  - If SeqNum $\leq$ LFR or SeqNum $>$ LAF, discard frame.
  - If LFR $<$ SeqNum $\leq$ LAF, the frame is accepted.

    Let SeqNumToAck be the largest sequence number that has the property that all frames with sequence number up to SeqNumToAck have been received.

    If frame SeqNumToAck has not yet been acknowledged, the receiver sends an ACK for sequence number SeqNumToAck (even if higher-numbered frames have been received) and sets LFR $=$ SeqNumToAck and LAF $=$ LFR $+$ RWS.
- The ACKs are called **cumulative** here, as an ACK for frame SeqNum acknowledges also all earlier frames.

Sliding Window Example

- Assume LFR $= 5$ (i.e. last ACK was sent for SeqNum 5) and RWS $= 4$, thus LAF $= 9$.
- The receiver accepts frames with SeqNum 6 to 9.
- If frames 7 and 8 arrive, the receiver stores them, but does not need to send an ACK. (Technically, it may also resend an ACK for frame 5.)
- If, after that, frame 6 arrives, the receiver stores it and sends an ACK for SeqNumToAck $= 8$. LAF increases to 12.
- In this scenario, frame 6 was probably lost during the first transmission and resent after the timeout.

Remarks on Sliding Window

- When a frame is lost and a timeout occurs, the amount of data in transit decreases, as the sender can advance its window only when the lost frame is acknowledged.
- Variations of sliding window:
  - Can use **duplicate acknowledgments** to detect frame losses.
  - **Negative Acknowledgment** (NACK): When frame 5 and then frame 7 arrives, receiver sends a NACK for frame 6.
  - **Selective Acknowledgment**: Receiver acknowledges exactly those frames it has received, not just the highest-numbered frame received in order.

Setting the Window Size

- The send window size determines the maximum number of outstanding frames on the link.
- SWS can be computed from the delay $\times$ bandwidth product of the link (if known!).
- RWS can be anything between 1 and SWS.
  - RWS $=$ 1: Receiver does not buffer any out-of-order frames.
  - RWS $=$ SWS: Receiver can buffer **any** frame that sender transmits.

Finite Sequence Numbers

- In practice, sequence numbers cannot grow infinitely large.
- Instead, they are stored in a header field of some finite size.
- For example, a 3-bit header field for sequence numbers allows only numbers 0,1,. . . ,7.
- Sequence numbers must be reused: they "wrap around."
- Receiver must be able to distinguish between different incarnations of the same sequence number.
- In stop-and-wait, only one frame could be outstanding at any time, and two different sequence numbers (0, 1) were enough.

How Many Sequence Numbers?

- Assume that sequence numbers range from 0 to MaxSeqNum $-$ 1.
- If RWS $=$ 1, then MaxSeqNum $\geq$ SWS$+$1 is sufficient.
- If RWS $=$ SWS, we need MaxSeqNum $\geq 2 \cdot$ SWS.
  - Consider SWS $=$ RWS $=$ 7 and sequence numbers 0,...,7. This does **not** work:
    - Sender transmits 0,...,6
    - Receiver acknowledges, but ACKs are lost
    - Sender retransmits 0,...,6, but receiver expects frames 7,0,1,2,3,4,5, so the retransmitted frames will be mistaken for new frames.
- All this assumes that frames are not reordered in transit!

Roles of Sliding Window Protocol

- It delivers frames reliably across an unreliable link.
- It can preserve the order in which frames are transmitted, so that they are delivered to the higher layer protocols at the receiving end in the same order as they were sent.
- It can support **flow control**: a feedback mechanism by which the receiver (or the network) is able to throttle the sender (by adapting the send window size).

Some of these roles are more relevant for end-to-end connections than for individual links.

Therefore, the sliding window algorithm is also used in some end-to-end protocols.

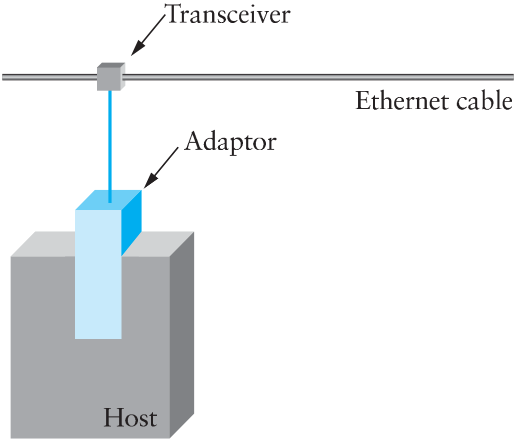
Concurrent Logical Channels

- Simple alternative to sliding window algorithm used in ARPANET.
- Keep the pipe full while still using the simple stop-and-wait algorithm.
- Idea: Multiplex several logical channels onto a single point-to-point link, and run stop-and-wait on each of these channels.
- When a node has a frame to send, it uses the lowest idle channel.
- Note: Reordering of frames can happen.
- In ARPANET: 8 logical channels over ground links, 16 logical channels over satellite links.

Ethernet (802.3)

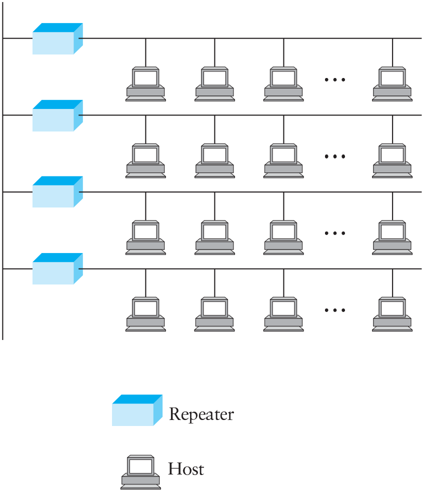
- Most successful local area networking technology of the last 20 years.
- Developed in the mid-1970s at Xerox Palo Alto Research Center (PARC).
- Based on **Carrier Sense Multiple Access with Collision Detect (CSMA/CD)** approach.
  - Carrier Sense: Nodes can distinguish between idle and busy link.
  - Collision Detect: Node listens as it transmits and can detect when a frame it is transmitting has interfered with another transmission.
- 10 Mbps Ethernet standard (Xerox, DEC, Intel) in 1978, basis of IEEE 802.3 (also: 100 Mbps – Fast Ethernet, 1000 Mbps – Gigabit Ethernet).

Transceiver and Adaptor



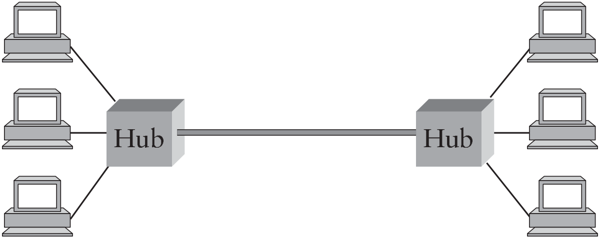
- Ethernet segment: up to 500 m in length (coaxial cable), taps at least 2.5 m apart.

Ethernet Repeater



- Repeaters forward the signal on all outgoing segments.

Ethernet Hub



- With 10BaseT cable (twisted pair, typically Category 5), it is common to have several point-to-point segments coming out of a multiway repeater (called **hub**).
- Even if an Ethernet consists of several segments connected by repeaters or hubs, data transmitted by any one host reaches **all other hosts**.

Ethernet Frame Format

- Digital–Intel–Xerox frame format:

64	48	48	16		32
Preamble	Dest addr	Src addr	Type	Body	CRC

  - 48-bit addresses (MAC addresses, MAC = media access control), usually burnt into ROM of adaptor.
  - **Type** field serves as demultiplexing key.
  - **Body**: 46 to 1500 bytes of data.
- Minor difference in IEEE 802.3: Length instead of Type, Type at beginning of Body.
- An adaptor receives frames addressed to it (**unicast**), to the **broadcast address** (all 1s), or to a **multicast address** that it has been instructed to listen to.
- In **promiscuous mode**, it receives all frames.

Transmitter Algorithm

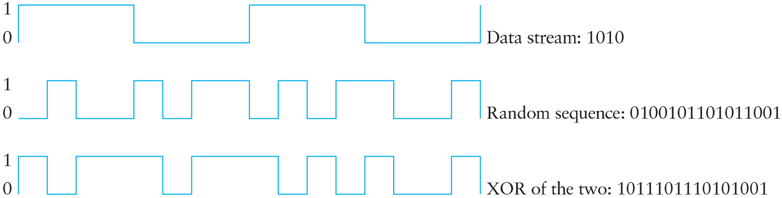
- If the adaptor has a frame to send:
  - If the line is idle, send immediately.
  - Otherwise, wait until current transmission finishes, and then transmit immediately.
- When the sender detects a **collision** (someone else was sending at the same time), it sends a **jamming sequence** (32 bits) and stops transmitting.
- In Ethernet, the maximum latency is at most $51.2\mu s$ and the minimum frame length is 512 bits, so every collision is detected. (It takes $51.2\mu s$ to transmit 512 bits over a 10 Mbps link.)

Exponential Backoff

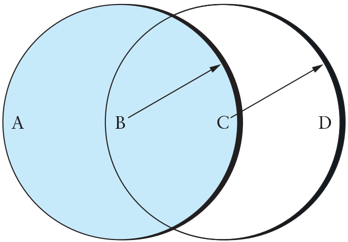
- After a collision, the sender repeatedly waits a certain amount of time and tries again:
  - After n collisions, it waits for $k \times 51.2\mu\text{s}$, where $k \in \{0, 1, \ldots, 2^n - 1\}$ is chosen **randomly**.
- This means:
  - After 1st collision, wait $k \times 51.2\mu\text{s}$ for $k \in \{0, 1\}$.
  - After 2nd collision, wait $k \times 51.2\mu\text{s}$ for $k \in \{0, 1, 2, 3\}$.
  - After 3rd collision, wait $k \times 51.2\mu\text{s}$ for $k \in \{0, \ldots, 7\}$.
  - After 4th collision, wait $k \times 51.2\mu\text{s}$ for $k \in \{0, \ldots, 15\}$.
  - ...and so on
- Adaptors typically retry up to 16 times (but cap n at 10).
- Ethernet works best under lightly loaded conditions (utilization below 30%).

Wireless (802.11)

- IEEE 802.11 Wireless LAN
- Designed for limited geographical area (homes, office buildings, campuses).
- Radio-based versions run at 11 Mbps or 54 Mbps.
- Spread spectrum radio: signals are spread over a wider frequency band than normal, so as to minimise the impact of interference from other devices:



Collisions

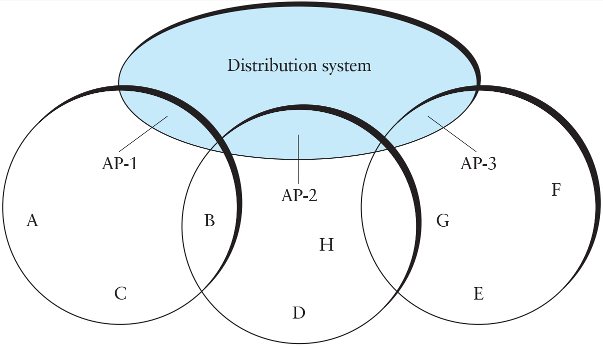


- Collision detection as in Ethernet is not always possible.
- If C and A send to B at the same time, the collision cannot be detected by C or A (**hidden node** problem).
- Simultaneous transmission from B to A and from C to D is possible, but when C hears B's transmission, it may conclude that it should not send (**exposed node** problem).

Collision Avoidance

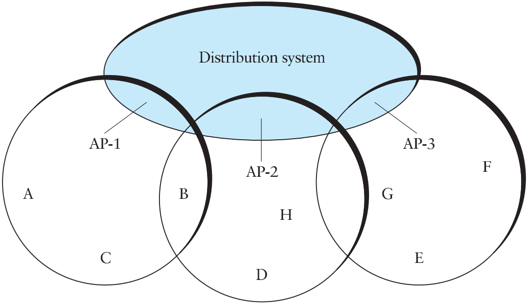
- **MACA**: Multiple Access with Collision Avoidance
  - When A wants to send a data packet to B, A sends an RTS (request to send) control frame that includes the length of the data packet.
  - When B receives RTS, it responds with a CTS (clear to send) control frame (with same length field).
  - When A receives the CTS, it sends the data packet.
  - Nodes that hear the CTS do not send during the packet transmission.
  - Nodes that hear the RTS but not the CTS are free to transmit.
  - If two RTS frames collide, the sender does not receive a CTS and waits a random time before trying again (exponential backoff).

Distribution System



- In practice, most IEEE 802.11 networks use **access points** (AP) that are connected by a **distribution system** (layer 2, e.g. Ethernet).
- Instead of nodes talking directly to each other, each node associates itself with an access point and sends/receives frames only to/from that AP.

Example



- If A wants to communicate with E:
  - A sends a frame to its access point AP-1.
  - AP-1 forwards the frame to AP-3 via the distribution system.
  - AP-3 sends the frame to E.

Selecting an AP: Scanning

- When a node joins the network or is unhappy with its current AP, it uses **active scanning**:
  1. The node sends a **Probe** frame.
  2. All APs within reach respond with a **Probe Response** frame.
  3. The node selects an AP and sends it an **Association Request** frame.
  4. The AP replies with an **Association Response** frame.
- If a node moves and changes its AP, the new AP notifies the old AP of the change.
- APs also periodically broadcast a **Beacon** frame that advertises its capabilities, and a node can change to this AP by sending it an **Association Request** frame (**passive scanning**).

802.11 Frame Format

16	16	48	48	48	16	48	0–18,496	32
Control	Duration	Addr1	Addr2	Addr3	SeqCtrl	Addr4	Payload	CRC

- **Control**:
  - Type: data, RTS, CTS, or scanning-related
  - Two 1-bit fields: ToDS and FromDS
- If one node sends directly to another, ToDS/FromDS are both 0, Addr2 is the source and Addr1 is the destination.
- If the frame was sent by a node to its AP, forwarded to another AP, and is now being sent to the destination:
  - ToDS and FromDS are 1
  - Addr4 is source node, Addr3 is first AP
  - Addr2 is second AP, Addr1 is destination node