EECS 489
Computer Networks

Winter 2023

Z. Morley Mao

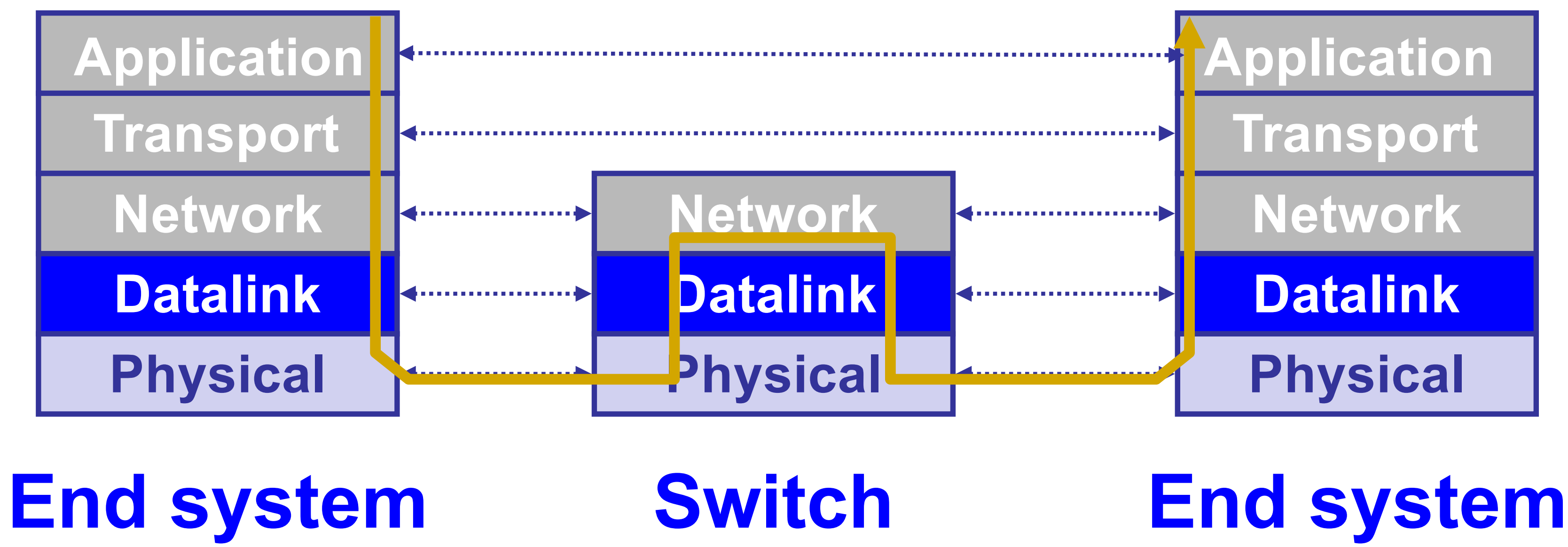
Material with thanks to Aditya Akella, Sugih Jamin, Philip Levis, Sylvia Ratnasamy, Peter Steenkiste, and many other colleagues.

Agenda

- Data link layer

Data link layer

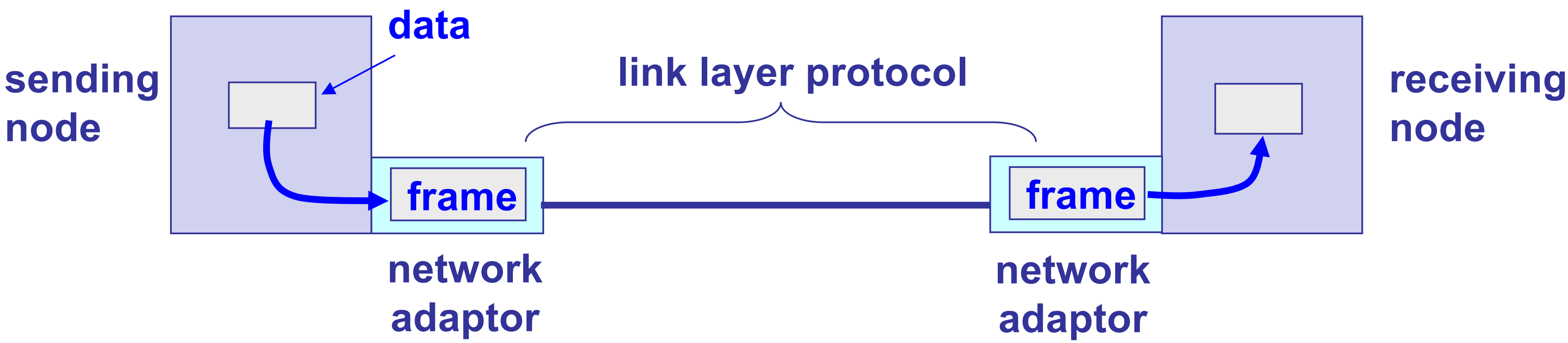
- Present everywhere
- Transfers data between adjacent nodes or between nodes on the same local area network



Data link layer

- Provides four primary services
  - Framing
    - Encapsulates network layer data
  - Link access
    - Medium access control (MAC) protocol defines when to transmit frames
  - Reliable delivery
    - Primarily for mediums with high error rates (e.g., wireless)
  - Error detection and correction

Packets are now "frames"



- Frames encapsulate network layer packets
- Link layer protocols are implemented in h/w
- Frame formats can change based on link layer protocol

Point-to-point vs. broadcast medium

- Point-to-point: dedicated pairwise communication
  - E.g., long-distance fiber link
  - E.g., Point-to-point link b/n Ethernet switch and host
- Broadcast: shared wire or medium
  - Traditional Ethernet (pre ~2000)
  - 802.11 wireless LAN

Multiple access algorithm

- Context: a shared broadcast channel
  - Must avoid having multiple nodes speaking at once
    - Otherwise, collisions lead to garbled data
  - Need distributed algorithm to determine which node can transmit
- Three classes of techniques
  - Channel partitioning: divide channel into pieces
  - Taking turns: scheme for deciding who transmits
  - Random access: allow collisions, and then recover
    - More in the Internet style!

Random access MAC protocols

- When node has packet to send
  - Transmit at full channel data rate **w/o** coordination
- Two or more transmitting nodes $\Rightarrow$ collision
  - Data lost
- Random access MAC protocol specifies
  - How to detect and recover from collisions
- Examples
  - ALOHA and Slotted ALOHA
  - CSMA, CSMA/CD, CSMA/CA (wireless)

Ethernet

- Invented as a broadcast technology
  - Hosts share channel
  - Each packet received by all attached hosts
  - CSMA/CD for media access control
- Modern Ethernets are "switched" (later)
  - Point-to-point links between switches and between a host and switch
  - No sharing $\Rightarrow$ no CSMA/CD
    - » Uses "self learning" and "spanning tree" algorithms for routing

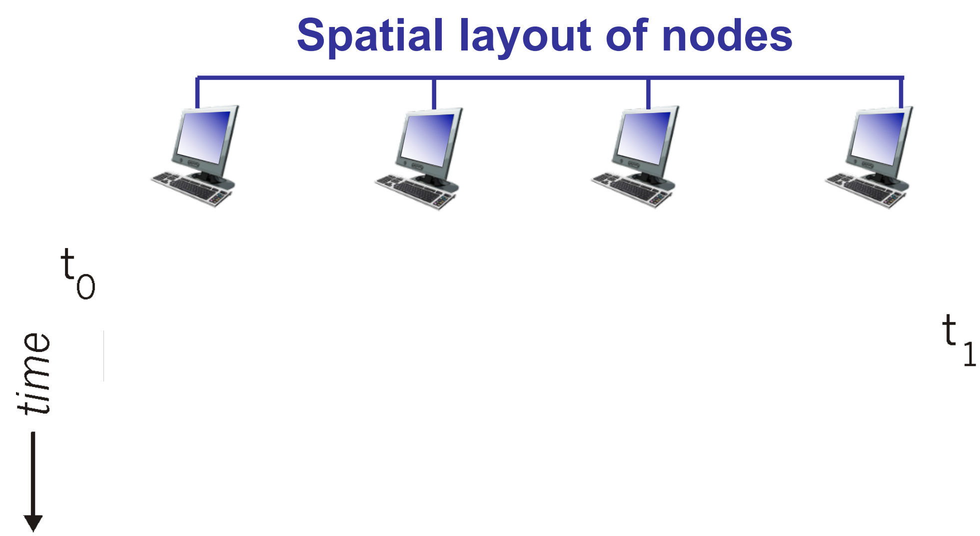
CSMA (Carrier Sense Multiple Access)

- CSMA: listen before transmit
  - If channel sensed idle: transmit entire frame
  - If channel sensed busy, defer transmission
- Human analogy: don't interrupt others!
- Does not eliminate all collisions
  - Why?

CSMA collisions

- Propagation delay: two nodes may not hear each other before sending

- CSMA reduces but does not eliminate collisions
- Collision: entire packet transmission time wasted
  - Distance and propagation delay affect collision probability

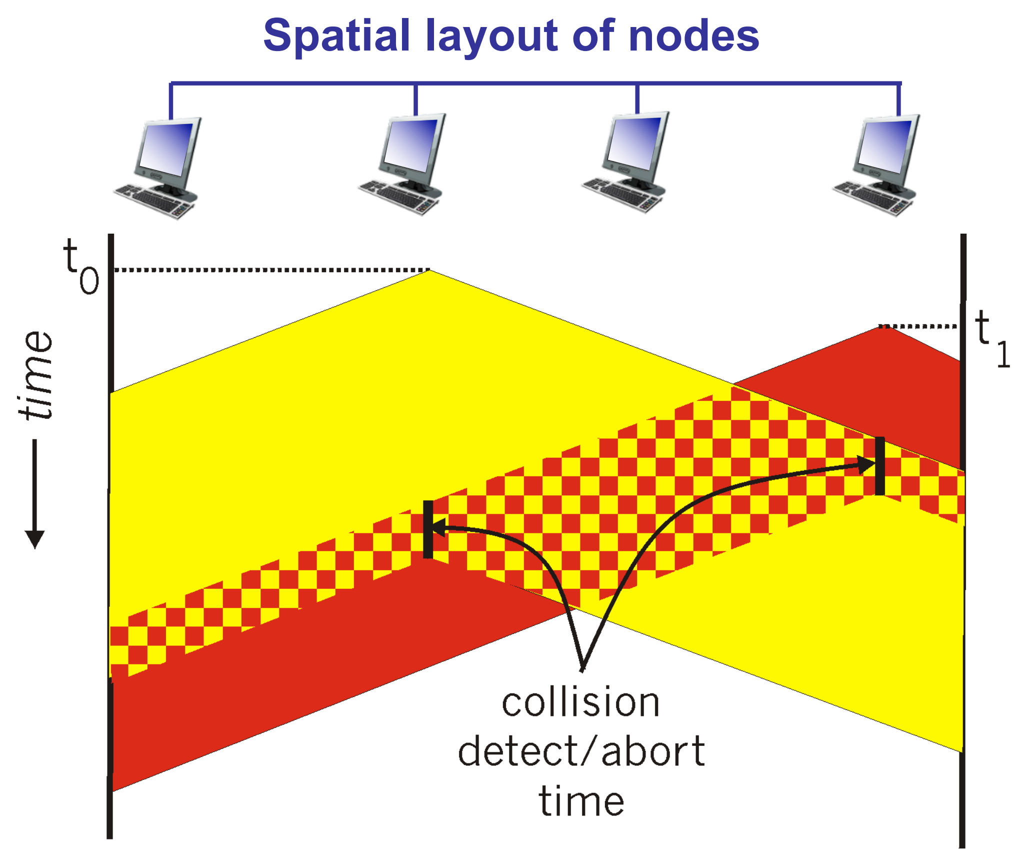


CSMA/CD (Collision Detection)

- CSMA/CD: carrier sensing, deferral as in CSMA
  - Collisions detected within short time
  - Colliding transmissions aborted, reducing wastage
- Collision detection easy in wired (broadcast) LANs
  - Compare transmitted, received signals
- Collision detection difficult in wireless LANs
  - Later!

CSMA/CD (Collision Detection)

- For this to work, need restrictions on minimum frame size and maximum distance
  - Why?



Limits on CSMA/CD network length

- Latency depends on physical length of link
  - Time to propagate a frame from one end to other
- Suppose A sends a frame at time **t**
  - And B sees an idle line at a time just before **t + d**
  - … so B happily starts transmitting a frame
- B detects a collision, and sends jamming signal
  - But A cannot see collision until **t + 2d**

Limits on CSMA/CD network length

- A needs to wait for time **2d** to detect collision
  - So, A should keep transmitting during this period
  - AND keep an eye out for a possible collision
- Imposes restrictions; e.g., for 10 Mbps Ethernet
  - Maximum length of the wire: 2,500 meters
  - Minimum length of a frame: 512 bits (64 bytes)

Three key ideas of random access

- Carrier sense
  - Listen before speaking and don’t interrupt
  - Checking if someone else is already sending data
  - … and waiting till the other node is done
- Collision detection
  - If someone else starts talking at the same time, stop
    - Make sure everyone knows there was a collision!
  - Realizing when two nodes are transmitting at once
  - …by detecting that the data on the wire is garbled

Three key ideas of random access

- Randomness
  - Don’t start talking again right away
  - Waiting for a random time before trying again

How long should you wait?

- Should it be immediate?
- Should it be a random number with a fixed distribution?

Ethernet: CSMA/CD Protocol

- Carrier sense: wait for link to be idle
- Collision detection: listen while transmitting
  - No collision: transmission is complete
  - Collision: abort transmission & send jam signal
- Random access: binary exponential back-off
  - After collision, wait a random time before retrying
  - After m^{th} collision, choose K randomly from $\{0, \ldots, 2^m-1\}$
    - Wait for K*512 bit times before trying again
    - If transmission occurring when ready to send, wait until end of transmission (CSMA)

Efficiency of CSMA/CD

- Efficiency is defined as the long-run fraction of time during which frames are being transmitted without collision
- d_{prop} = max propagation time between two adapters
- d_{trans} = time to transmit a max-sized frame

$$\text{Efficiency} \approx \frac{1}{1 + 5\, d_{prop} / d_{trans}}$$

Efficiency of CSMA/CD

- $d_{prop} \rightarrow 0$
  - Efficiency approaches 1
  - Colliding nodes abort immediately
- $d_{trans} \rightarrow \infty$
  - Efficiency approaches 1
  - Each frames uses the channel for a long time

$$\text{Efficiency} \approx \frac{d_{trans}}{d_{trans} + 5\, d_{prop}}$$

5-MINUTE BREAK!

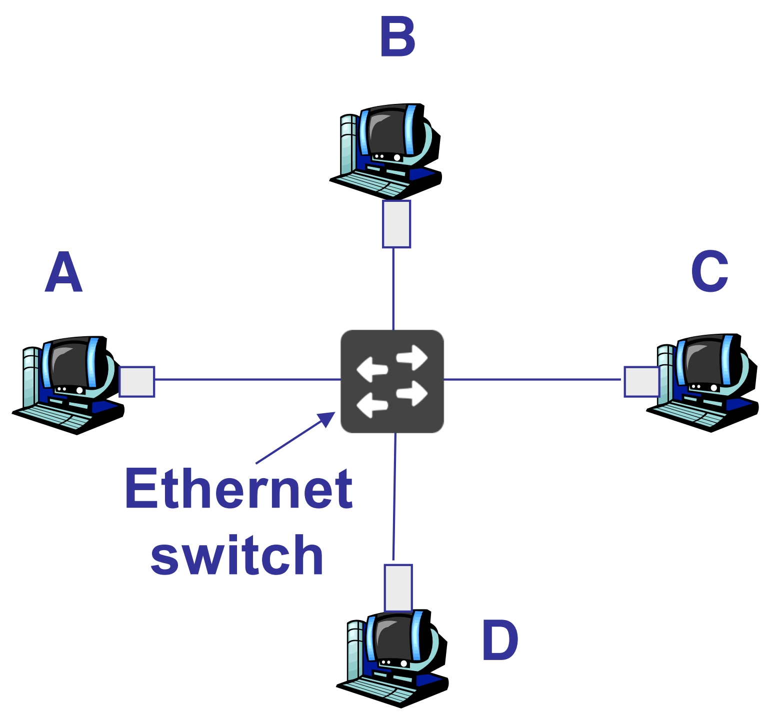
SWITCHED ETHERNET

Broadcast vs. switched Ethernet

- Invented as a broadcast technology
  - Hosts share channel
  - Each packet received by all attached hosts
  - CSMA/CD for media access control
- Modern Ethernets are "switched"
  - Point-to-point links between switches and between a host and switch
  - No sharing $\Rightarrow$ no CSMA/CD
    - » Uses "self learning" and "spanning tree" algorithms for routing

Why switched Ethernet?

- Enables concurrent communication
  - Host A can talk to C, while B talks to D
  - No collisions and no need for CSMA/CD
  - No constraints on link lengths, etc.



The evolution of Ethernet

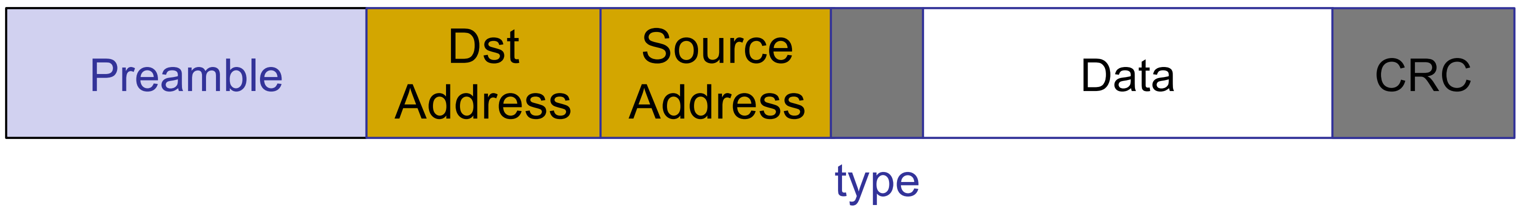
- Changed almost everything except the frame format
  - From the shared media coax cables to dedicated links
  - From 3 Mbit/s to 100 Gbit/s
  - From electrical signaling to optical
- Lesson: the right interface can accommodate many changes
  - Evolve the implementation while maintaining the interface (backward compatibility)

Topics

- Frames and framing
- Addressing
- Routing
- Forwarding
- Discovery

Ethernet “Frames”

- Encapsulates IP datagram



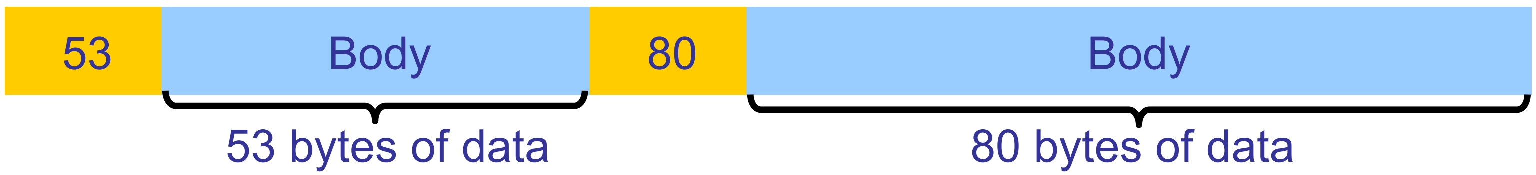
- Preamble: 7 bytes for clock synchronization and 1 byte to indicate start of frame
- Addresses: 6 bytes
- Type: 2 bytes, higher-layer protocol (e.g., IP)
- Data payload: max 1500 bytes, min 46 bytes
- CRC: 4 bytes for error detection

Framing frames

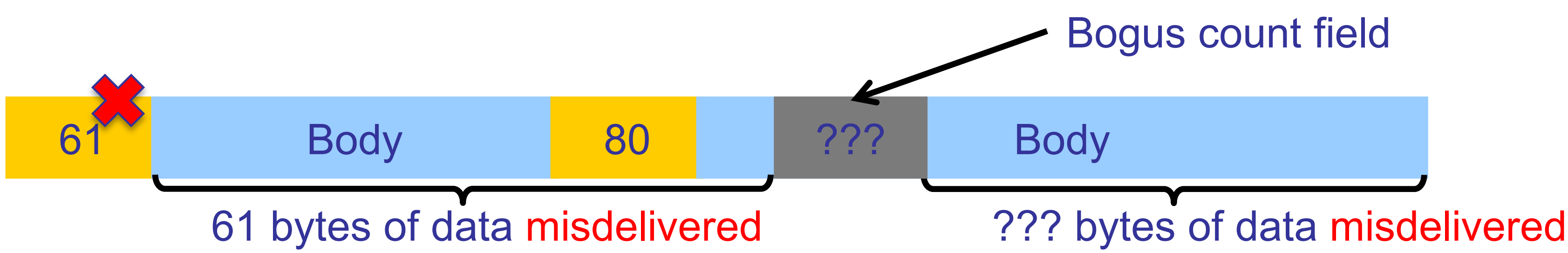
- Physical layer puts bits on a link
- But, two hosts connected on the same physical medium need to be able to exchange frames
  - Service provided by the link layer
  - Implemented by the network adaptor
- Framing problem: how does the link layer determine where each frame begins and ends?

Simple approach: Count bytes

- Sender includes number of bytes in header



  - Receiver extracts this number of bytes of body
- What if the Count field is corrupted?



  - L2 will frame the wrong bytes → a framing error
  - CRC tells you to discard this frame, but what about the next one?

Desynchronization

- Once framing on a link is desynchronized, it can stay that way
- Need a method to resynchronize

Framing with sentinel bits

- Delineate frame with special “sentinel” bit pattern
  - e.g., 01111110 $\Rightarrow$ start, 01111111 $\Rightarrow$ end

01111110	Frame contents	01111111

- What if sentinel occurs within frame?
- Solution: bit stuffing
  - Sender always inserts a 0 after five 1s in the frame contents
  - Receiver always removes a 0 appearing after five 1s

When receiver sees five 1s...

01111110	Frame content	01111111

- If next bit 0, remove it; begin counting again
  - Because this must be a stuffed bit; we can't be at beginning/end of frame (those had six or seven 1s)
- If next bit 1 (i.e., we've seen six 1s) then:
  - If following bit is 0, this is start of frame
    - Because the receiver has seen 01111110
  - If following bit is 1, this is end of frame
    - Because the receiver has seen 01111111

Example: sentinel bits

- Original data, including start/end of frame:
  - 01111110011111101111101111100101111111
- Sender rule: five 1s → insert a 0
  - After bit stuffing at the sender:
  - 01111110011111010111110011111000101111111
- Receiver rule: five 1s and next bit 0 → remove 0
  - 01111110011111101111101111100101111111

Topics

- Frames and framing
- Addressing
- Routing
- Forwarding
- Discovery

Medium Access Control (MAC) Address

- MAC address
  - Numerical address associated with a network adapter
  - Flat name space of 48 bits (e.g., 00-15-C5-49-04-A9 in HEX)
  - Unique, hard-coded in the adapter when it is built
- Hierarchical Allocation
  - Blocks: assigned to vendors (e.g., Dell) by the IEEE
    - »First 24 bits (e.g., 00-15-C5-**-**-**)
  - Adapter: assigned by the vendor from its block
    - »Last 24 bits

MAC address vs. IP address

MAC Addresses

- Hard-coded when adapter is built
- Flat name space of 48 bits (e.g., 00-0E-9B-6E-49-76)
- Like a social security number
- Portable, and can stay the same as the host moves
- Used to get packet between interfaces on same network

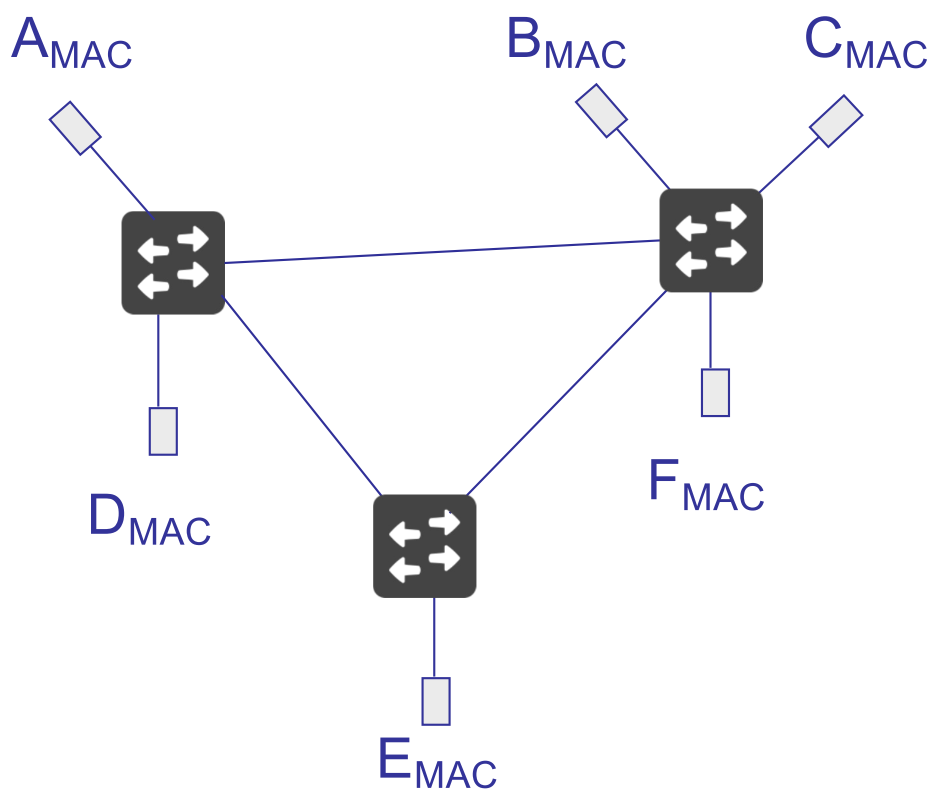
IP Addresses

- Configured, or learned dynamically
- Hierarchical name space of 32 bits (e.g., 12.178.66.9)
- Like a postal mailing address
- Not portable, and depends on where the host is attached
- Used to get a packet to destination IP subnet

Topics

- Frames and framing
- Addressing
- Routing
- Forwarding
- Discovery

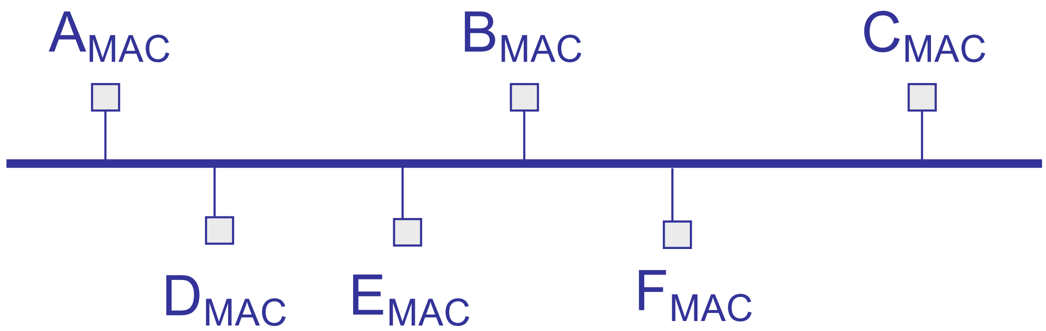
Routing with switched Ethernet?



Why does Ethernet not use LS/DV?

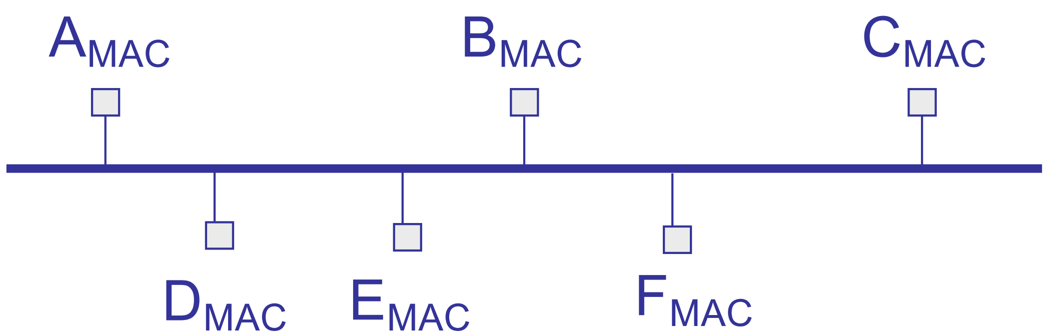
- Concerns over scalability
  - Flat MAC addresses cannot be aggregated like IP addresses
- Legacy

“Routing” with broadcast Ethernet



- Sender transmits frame onto broadcast link
- Each receiver’s link layer passes the frame to the network layer:
  - If destination address matches the receiver’s MAC address OR if the destination address is the broadcast MAC address (ff:ff:ff:ff:ff:ff)

"Routing" with broadcast Ethernet

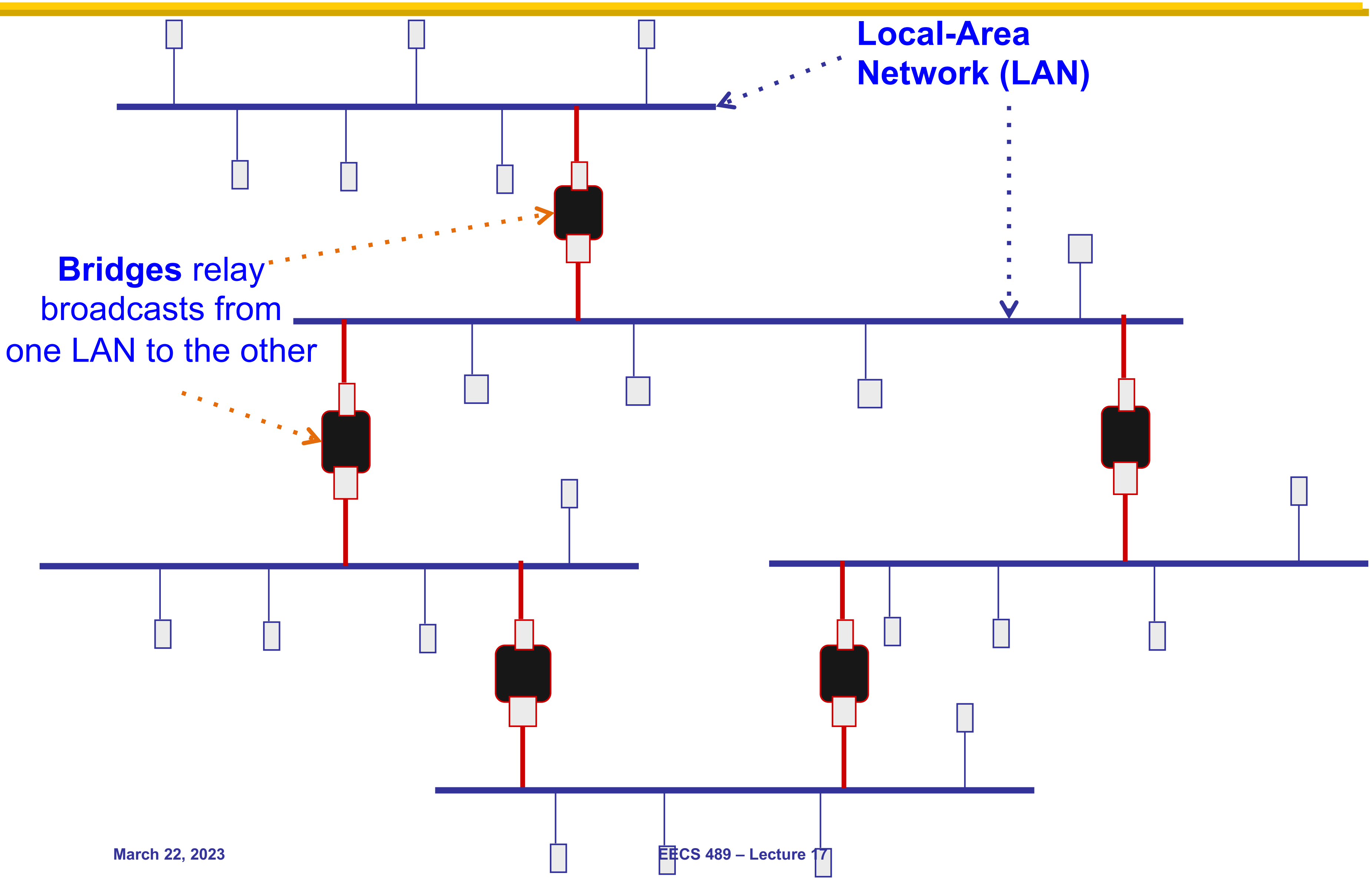


- Ethernet is "plug-n-play"
- A new host plugs into the Ethernet and is good to go
  - No configuration by users or network operators
  - Broadcast as a means of bootstrapping communication

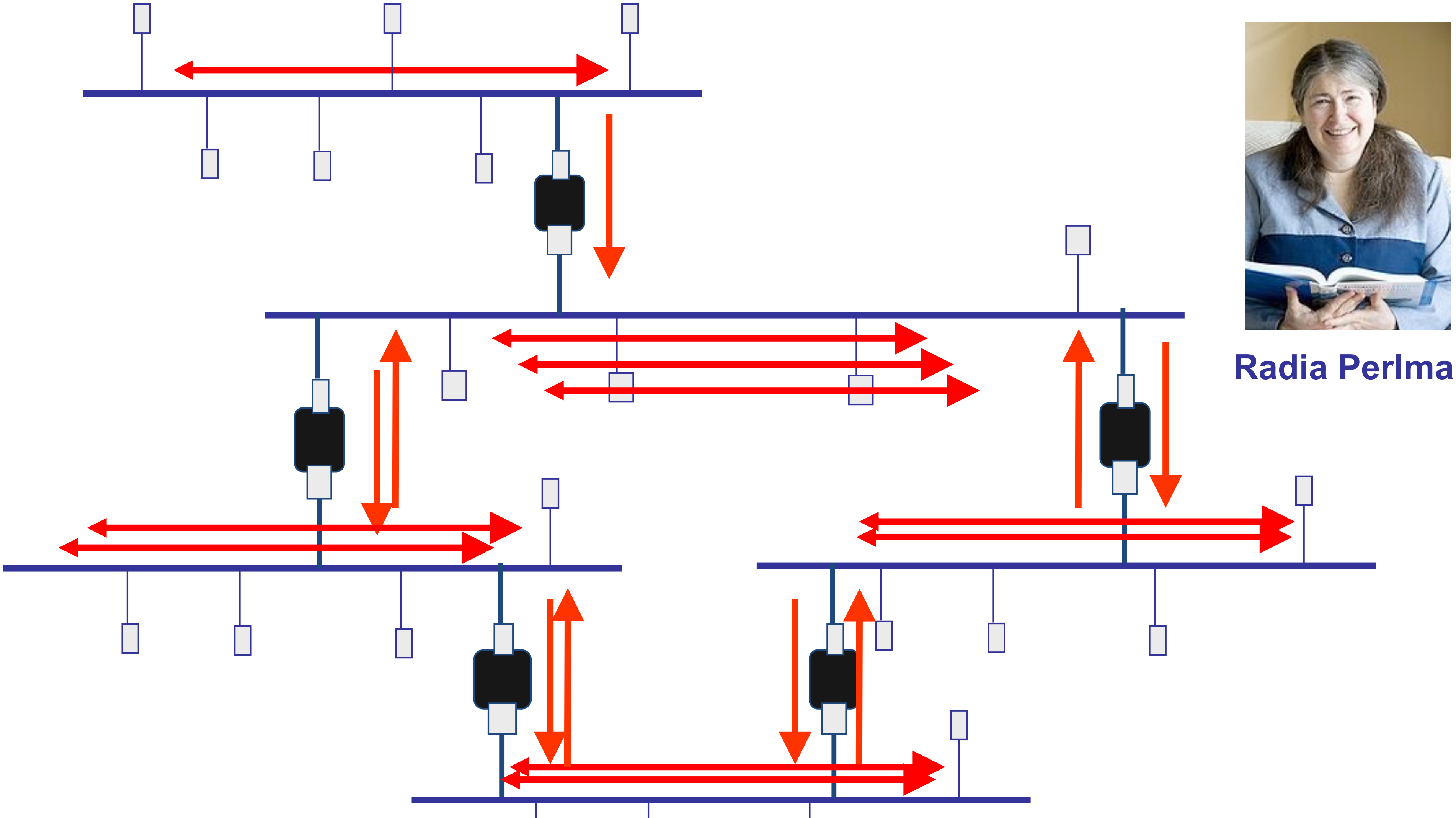
Why does Ethernet not use LS/DV?

- Concerns over scalability
  - Flat MAC addresses cannot be aggregated like IP addresses
- Legacy
  - Backward compatibility with broadcast Ethernet
  - Desire to maintain Ethernet's plug-n-play behavior
  - How broadcast Ethernet evolved

Routing in extended LANs



The “broadcast storm” problem



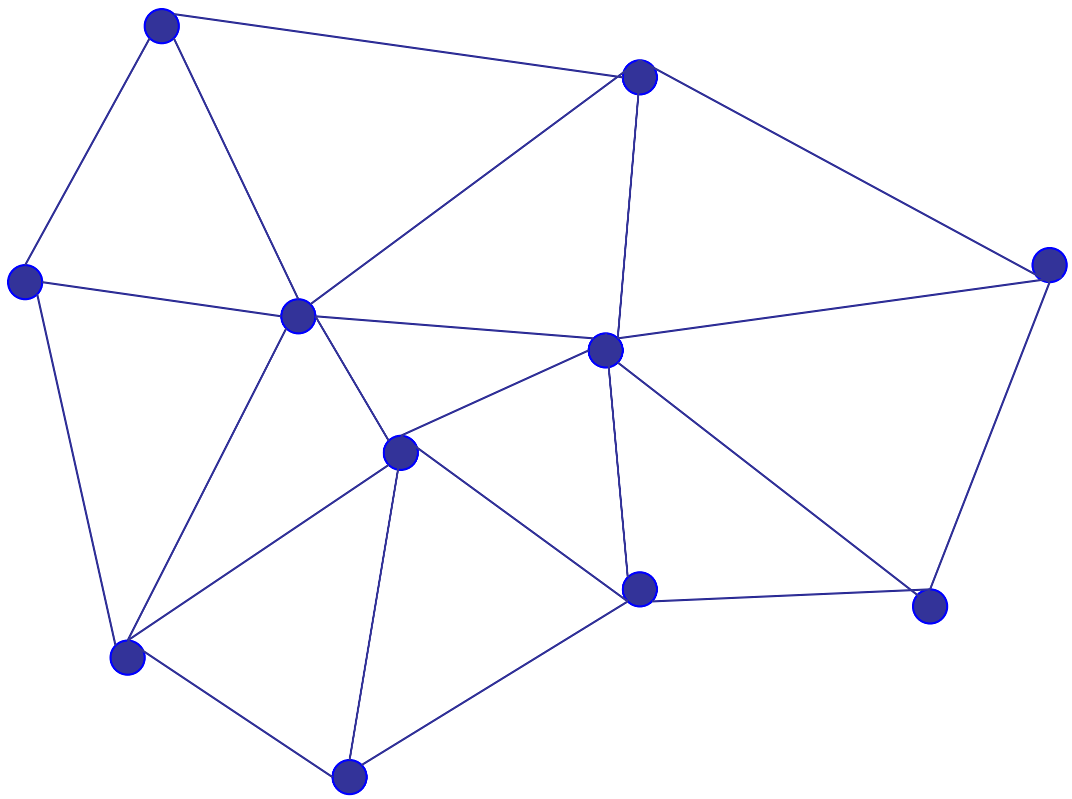
Radia Perlman

Perlman’s idea: eliminate loops in the topology

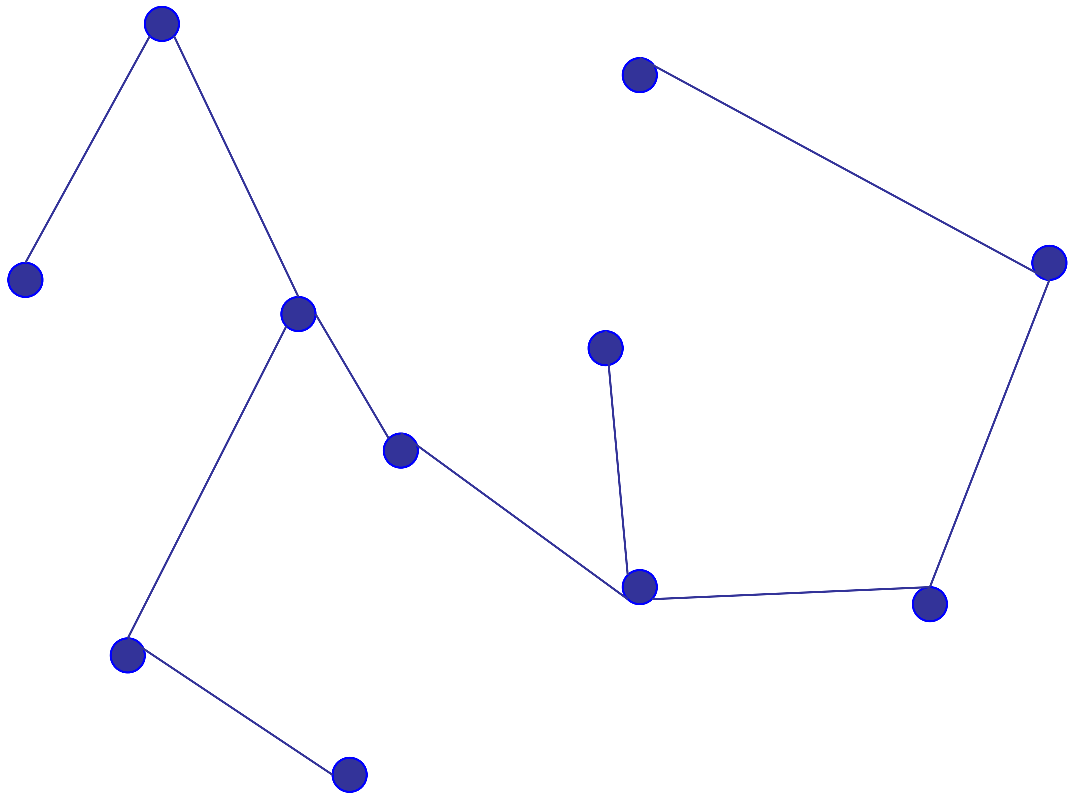
Easiest way to avoid loops

- Use a topology where loops are impossible!
- Take arbitrary topology and build a spanning tree
  - Sub-graph that includes all vertices but contains no cycles
  - Links not in the spanning tree are not used to forward frames

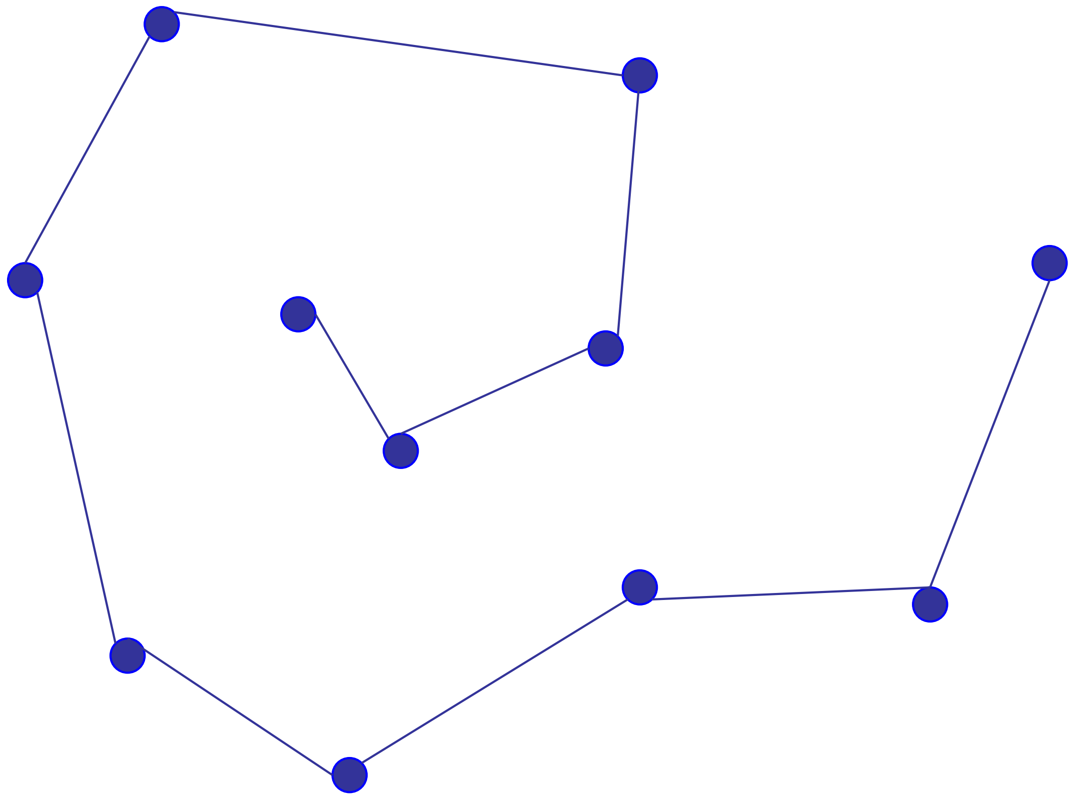
Consider a graph



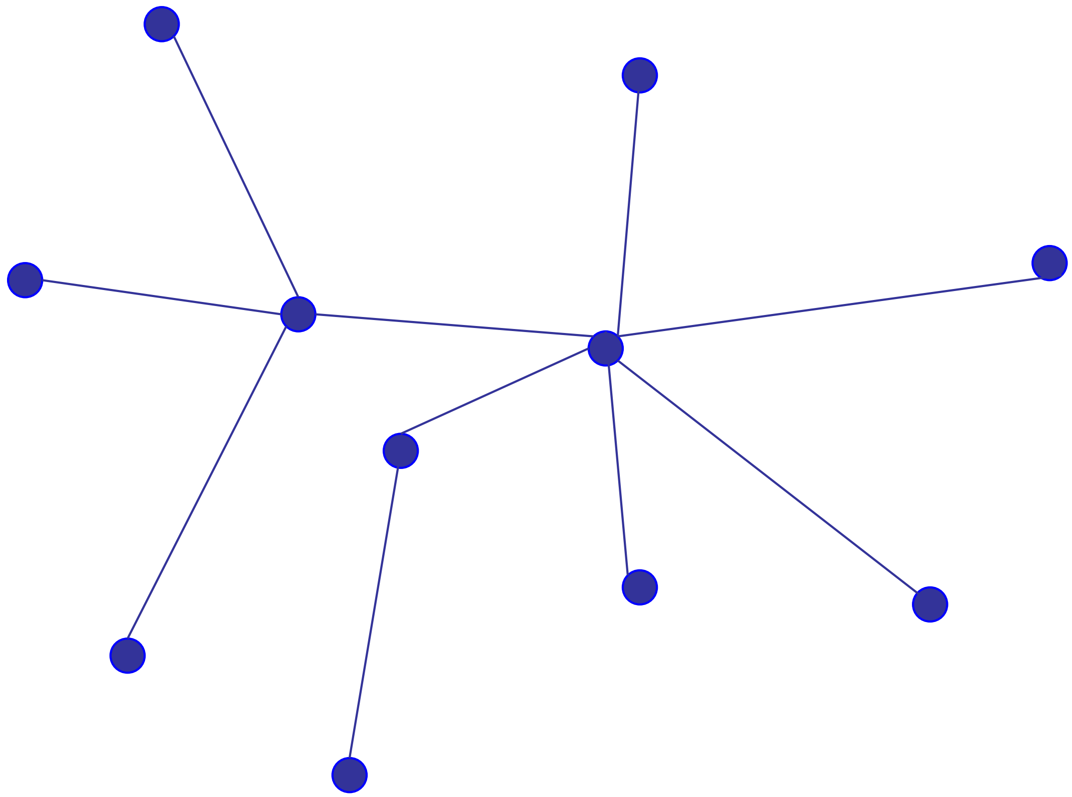
A spanning tree



Another spanning tree



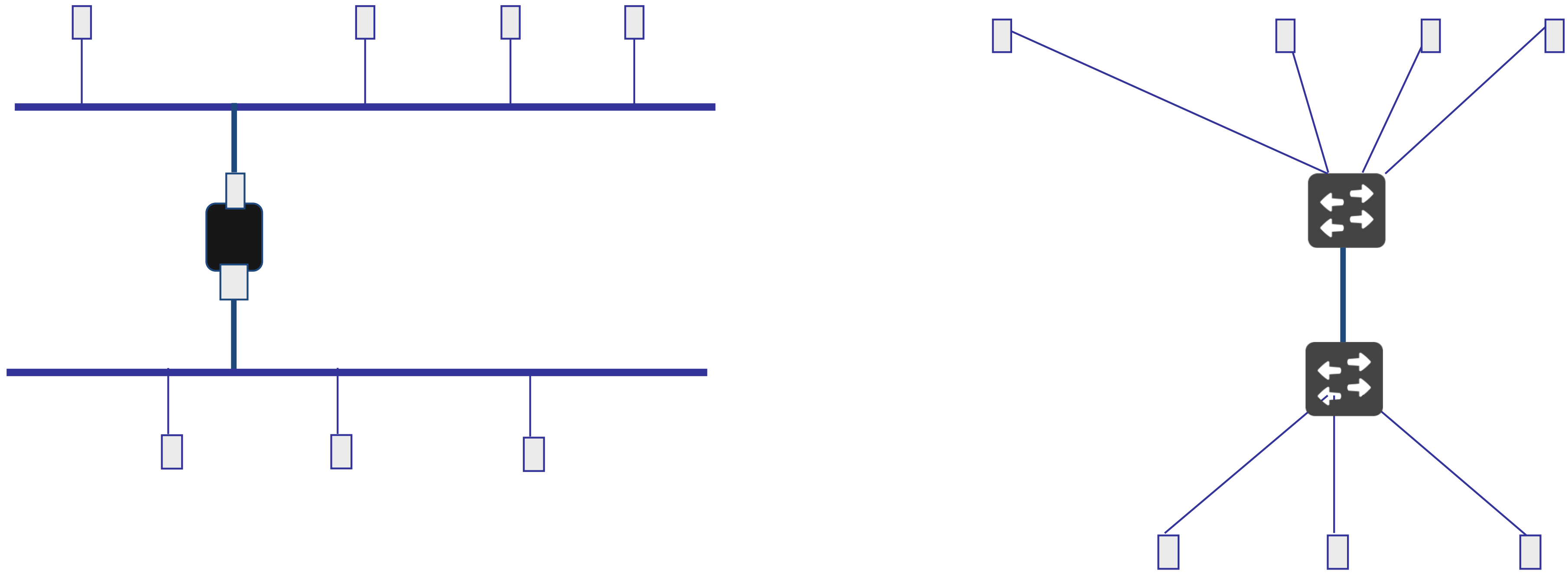
Yet another spanning tree



Spanning tree protocol (Perlman'85)

- Protocol by which bridges construct a spanning tree
- Nice properties
  - Zero configuration (by operators or users)
  - Self healing
- Still used today

From extended LANs to switched Ethernet



Switched Ethernet

- Constraints (for backward compatibility)
  - No changes to end-hosts
  - Maintain plug-n-play aspect
- Earlier Ethernet achieved plug-n-play by leveraging a broadcast medium
  - Can we do the same in a switched topology?

Summary

- Data link layer transfers data between adjacent nodes or nodes connected to the same switch
- Ethernet evolved from a broadcast medium to switched

- Next week: Link layer wrap up + putting everything together