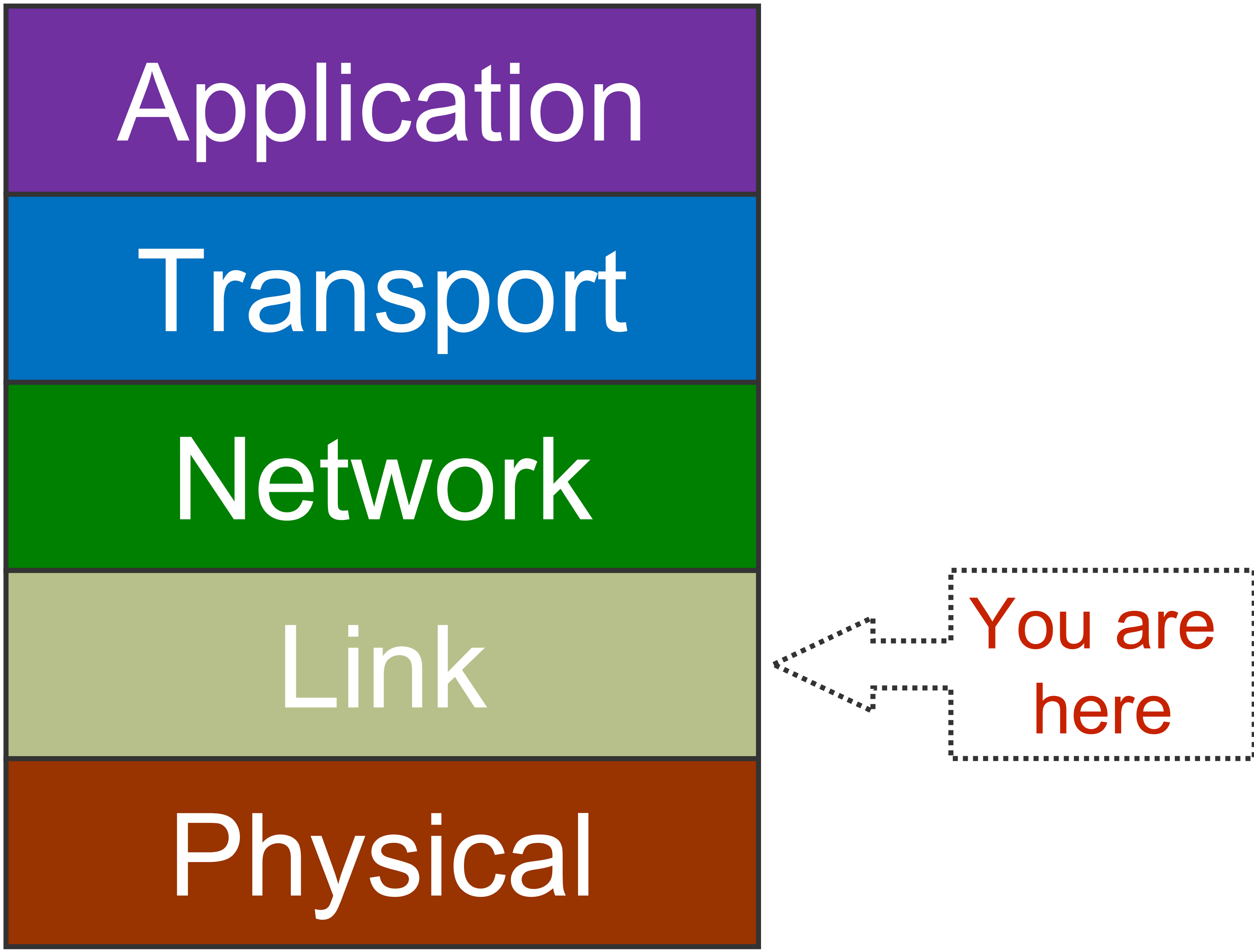
# CS2105

## An *Awesome* Introduction to Computer Networks

Lectures 9&10: The Link Layer

NUS National University of Singapore
Department of Computer Science
School of Computing


Application
Transport
Network
Link
Physical
You are here

# Lectures 9&10: The Link Layer

*After this class, you are expected to understand:*

- the role of link layer and the services it could provide.
- how parity and CRC scheme work.
- different methods for accessing shared medium.
- how ARP allows a host to discover the MAC addresses of other nodes in the same subnet.
- the role of switch in interconnecting subnets in a LAN.

# Lectures 9&10: Roadmap

6.1 Introduction to the Link Layer

6.2 Error Detection and Correction

6.3 Multiple Access Links and Protocols

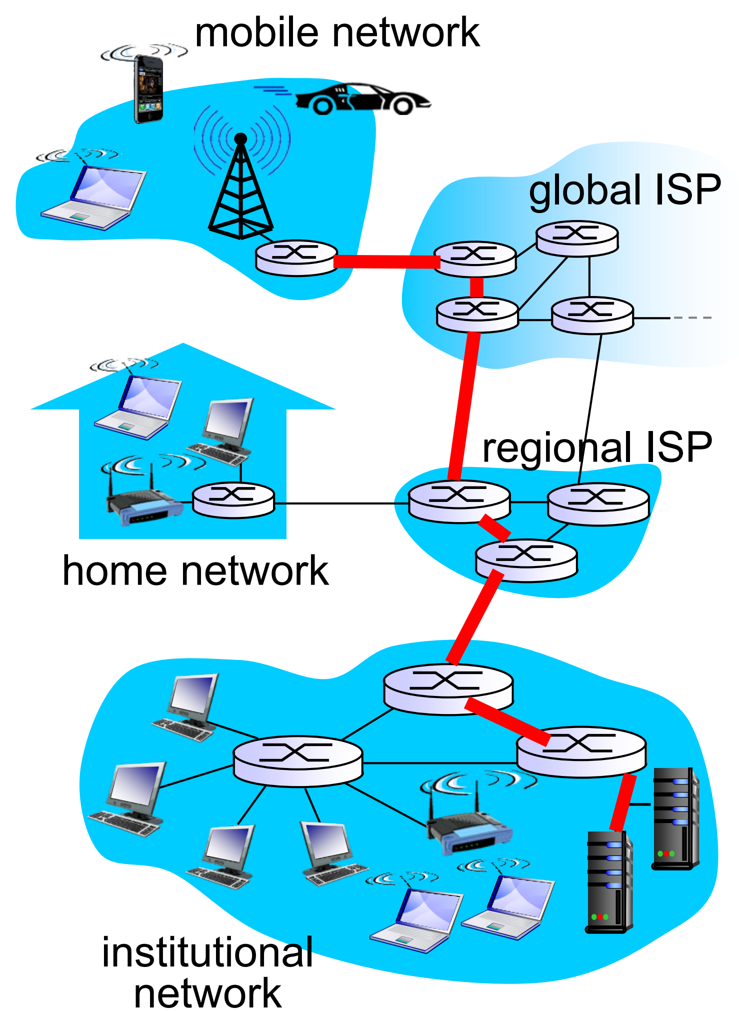
6.4 Switched Local Area Networks

Kurose Textbook, Chapter 6
(Some slides are taken from the book)

# Link Layer: Introduction (1/2)

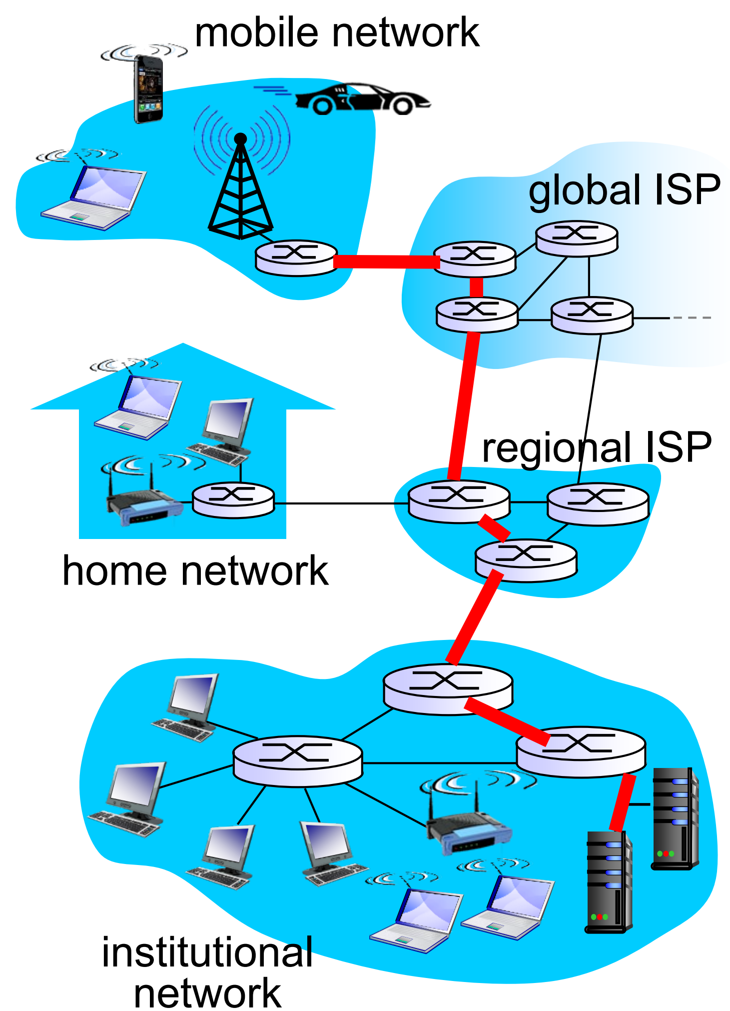
- Network layer provides communication service between any two hosts.

- An IP datagram may travel through multiple routers and links before it reaches destination.



# Link Layer: Introduction (2/2)

- Link layer sends datagram between adjacent nodes (hosts or routers) over a single link.
  - IP datagrams are encapsulated in link-layer frames for transmission.
  - Different link-layer protocols may be used on different links.
    - *each protocol may provide a different set of services.*



# Possible Link Layer Services (1/2)

- Framing
  - Encapsulate datagram into frame, adding header and trailer.

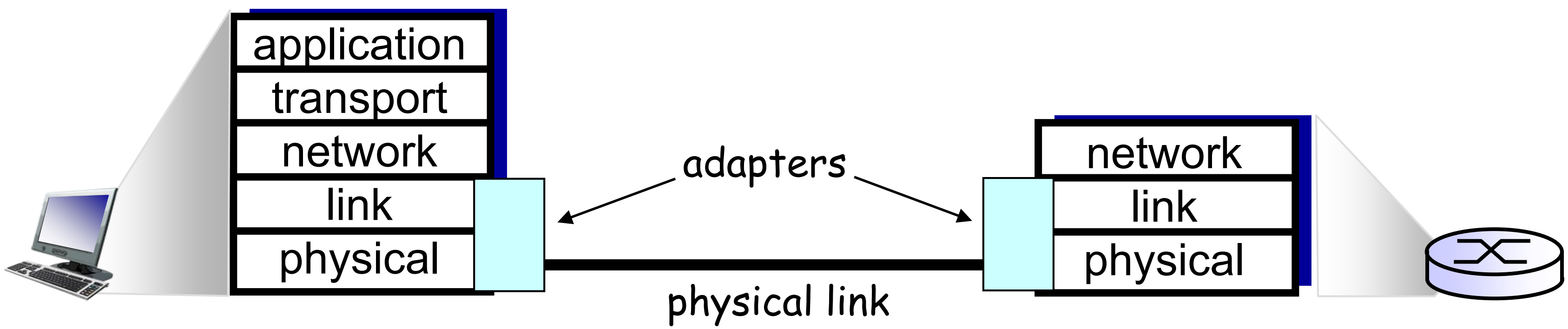

- Link access control
  - When multiple nodes *share* a single link, need to coordinate which nodes can send frames at a certain point of time.

humans at a cocktail party (shared air)

# Possible Link Layer Services (2/2)

- Reliable delivery
  - Seldom used on low bit-error link (e.g. fiber) but often used on error-prone links (e.g. wireless link).
- Error detection
  - Errors are usually caused by signal attenuation or noise.
  - Receiver detects presence of errors.
    - *may signal sender for retransmission or simply drops frame*
- Error correction
  - Receiver identifies and corrects bit error(s) without resorting to retransmission.

# Link Layer Implementation




- Link layer is implemented in "adapter" (aka NIC) or on a chip.
  - E.g., Ethernet card/chipset, 802.11 card
- Adapters are semi-autonomous, implementing both link & physical layers.

# Lectures 9&10: Roadmap

6.1 Introduction to the Link Layer

6.2 Error Detection and Correction

- 6.2.1 Parity Checks
- 6.2.3 Cyclic Redundancy Check (CRC)

6.3 Multiple Access Links and Protocols

6.4 Switched Local Area Networks

# Error Detection and Correction

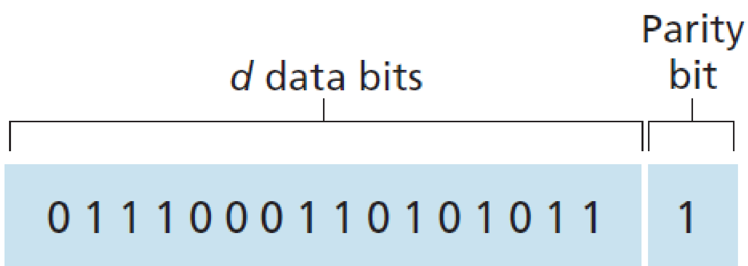
- Popular error detection schemes:
  - Checksum (used in TCP/UDP/IP)
  - Parity Checking
  - CRC (commonly used in link layer)

- Error detection schemes are not 100% reliable!
  - may miss some errors, but rarely.
  - larger error detection and correction (EDC) field yields better detection (and even correction).

# Parity Checking

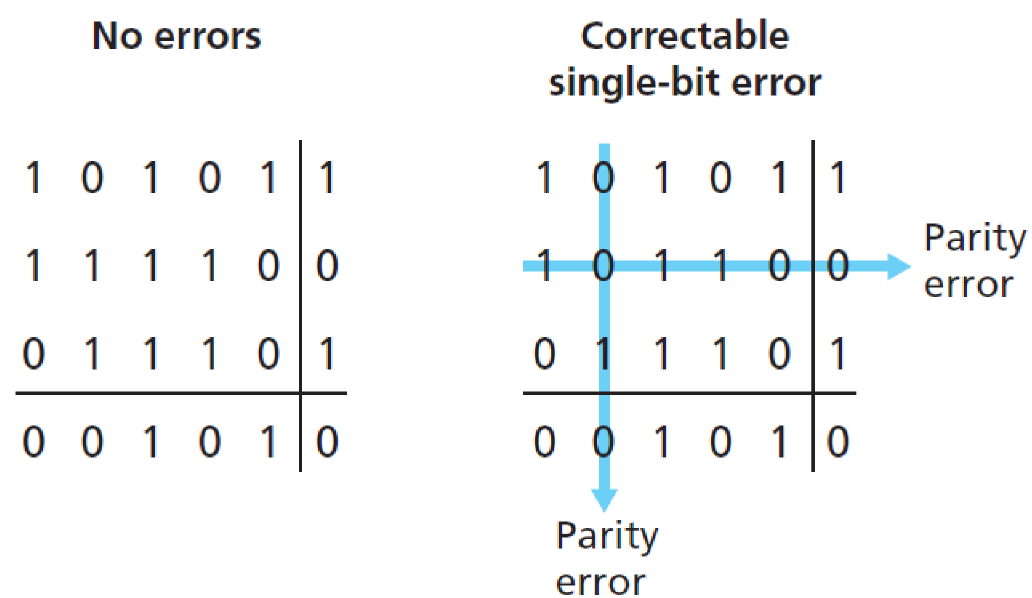
## Single bit parity

- can detect single bit errors in data.




## Two-dimensional bit parity

- can detect and correct single bit errors in data.
- can detect data any two bits errors in data.



# Cyclic Redundancy Check (CRC)

- Powerful error-detection coding that is widely used in practice (e.g., Ethernet, Wi-Fi)
  - $D$: data bits, viewed as a binary number.
  - $G$: generator of $r + 1$ bits, agreed by sender and receiver beforehand.
  - $R$: will generate CRC of $r$ bits.

Example: $r = 3$

```
1001 ) 101110000
       1001
       ----
        1010
        1001
        ----
         1100
         1001
         ----
          1010
          1001
          ----
           011
```

$G$

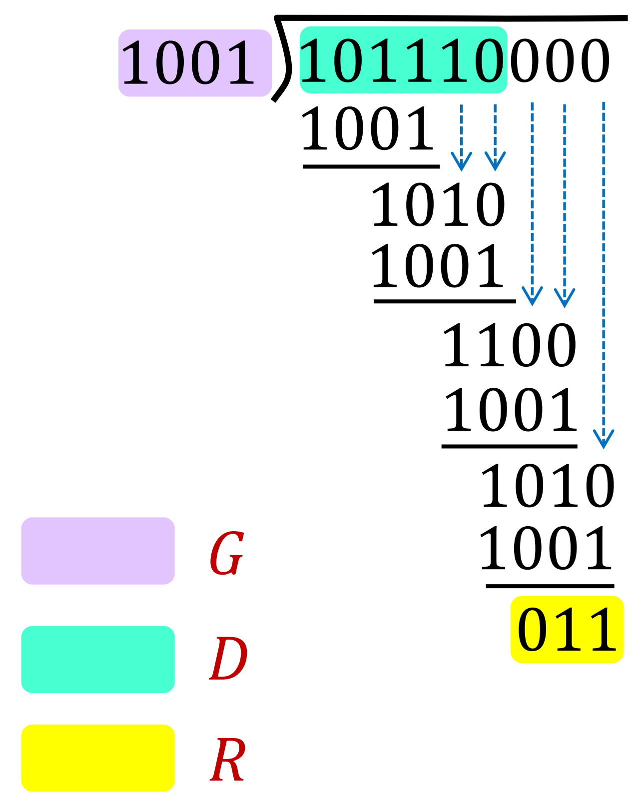
$D$

$R$

# Cyclic Redundancy Check (CRC)

- CRC calculation is done in bit-wise XOR operation without carry or borrow.
- Sender sends $(D, R)$
  - 101110011
- Receiver knows $G$, divides $(D, R)$ by $G$.
  - If non-zero remainder: error is detected!

Example: $r = 3$



# Lectures 9&10: Roadmap

6.1 Introduction to the Link Layer

6.2 Error Detection and Correction

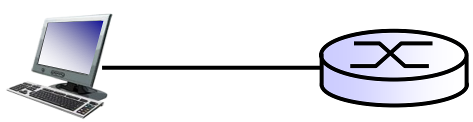
6.3 Multiple Access Links and Protocols

- 6.3.1 Channel Partitioning Protocols
- 6.3.2 Random Access Protocols
- 6.3.3 Taking-Turns Protocols

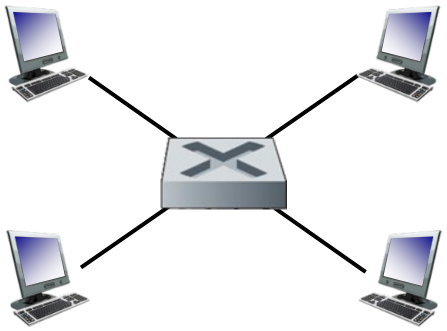
6.4 Switched Local Area Networks

# Two Types of Network Links

- **Type 1: point-to-point link**
  - A sender and a receiver connected by a dedicated link
  - Example protocols: Point-to-Point Protocol (PPP), Serial Line Internet Protocol (SLIP)
    - No need for multiple access control



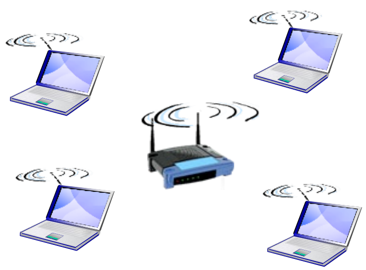
A host connects to router through a dedicated link



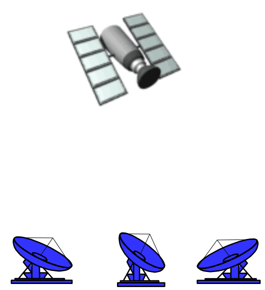
A point-to-point link between Ethernet switch and a host

# Two Types of Network Links

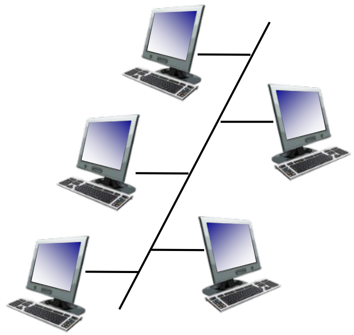
- **Type 2: broadcast link** (shared medium)
  - Multiple nodes connected to a shared broadcast channel.
  - When a node transmits a frame, the channel broadcasts the frame and each other node receives a copy.



802.11 Wi-Fi



Satellite



Ethernet with bus topology

# Multiple Access Protocols

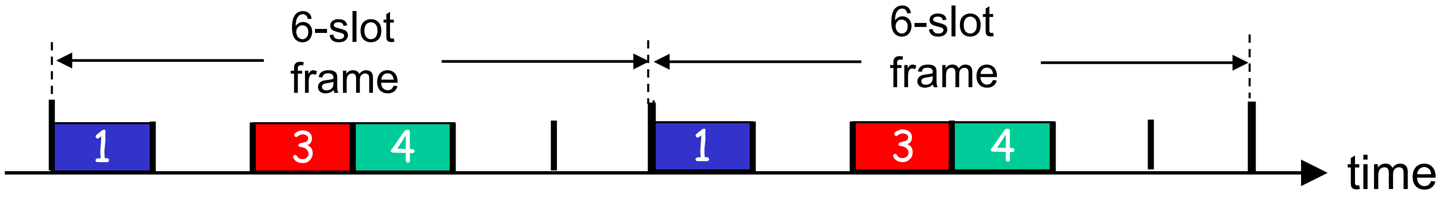
- In a broadcast channel, if two or more nodes transmit simultaneously
  - Every node receives multiple frames at the same time ➜ frames *collide* at nodes and none would be correctly read.
- Multiple Access Protocol
  - distributed algorithm that determines how nodes share channel, i.e. when a node can transmit.
  - However, coordination about channel sharing must use channel itself!
    - *no out-of-band channel signaling*

# Multiple Access Protocols

- Multiple access protocols can be categorized into three broad classes:
  - **Channel partitioning**
    - divide channel into fixed, smaller "pieces" (e.g., time slots, frequency).
    - allocate piece to node for exclusive use.
  - **"Taking turns"**
    - nodes take turns to transmit.
  - **Random Access**
    - channel is not divided, collisions are possible.
    - "recover" from collisions.

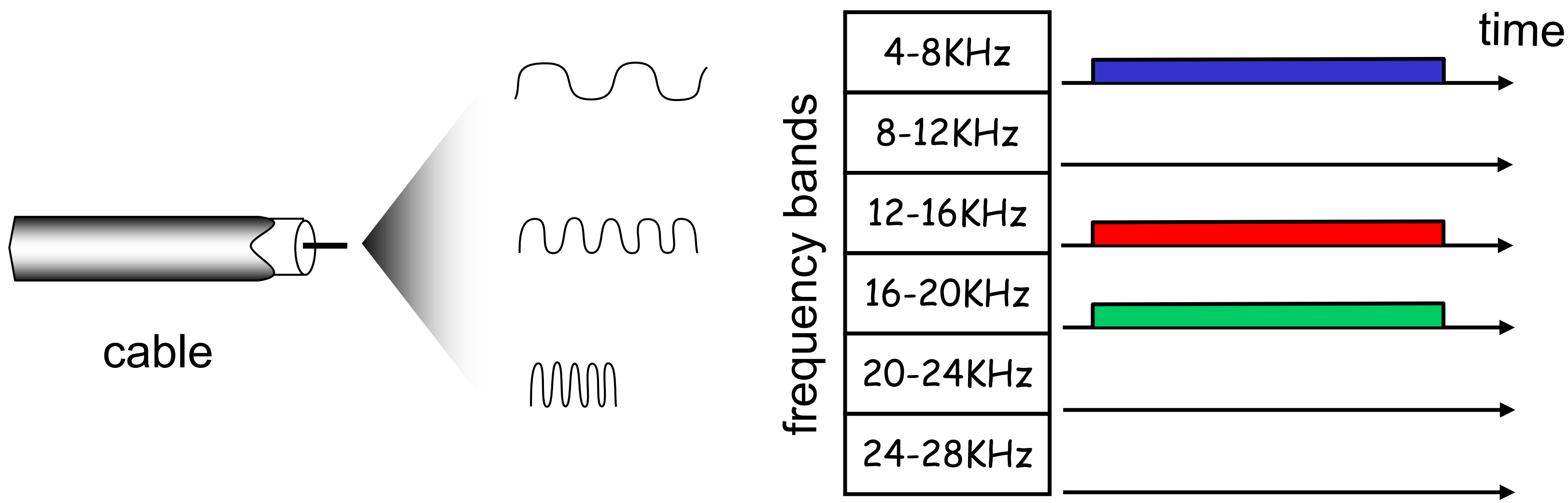
# Channel Partitioning Protocols

- TDMA (time division multiple access)
  - Access to channel in "rounds".
  - Each node gets fixed length slot (length = frame transmission time) in each round.
  - Unused slots go idle.
  - Example: 6 nodes sharing a link, 1, 3, 4 have frames, slots 2, 5, 6 are idle.



# Channel Partitioning Protocols

- FDMA (frequency division multiple access)
  - Channel spectrum is divided into frequency bands.
  - Each node is assigned a fixed frequency band.
  - Unused transmission time in frequency bands go idle.
  - Example: 6 nodes, 1, 3, 4 have frames, frequency bands 2, 5, 6 are idle.



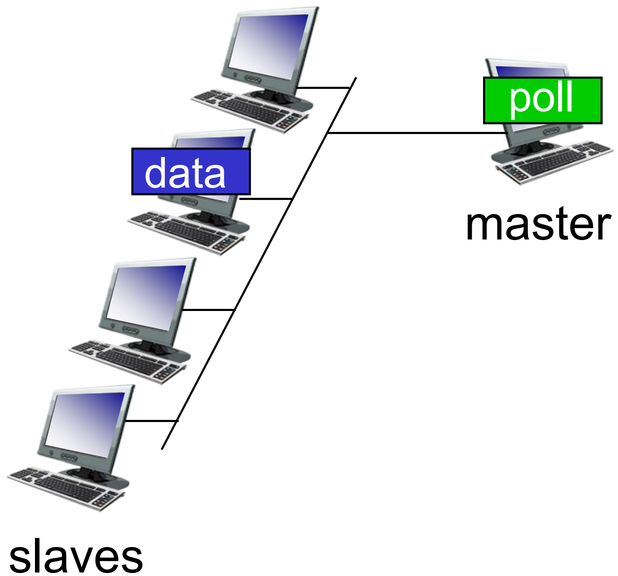
# Multiple Access Protocols

- Multiple access protocols can be categorized into three broad classes:
  - Channel partitioning
    - divide channel into smaller "pieces" (e.g., time slots, frequency).
    - allocate piece to node for exclusive use.
  - **"Taking turns"**
    - nodes take turns to transmit.
  - Random Access
    - channel is not divided, collisions are possible.
    - "recover" from collisions.

# "Taking Turns" Protocols

## Polling:

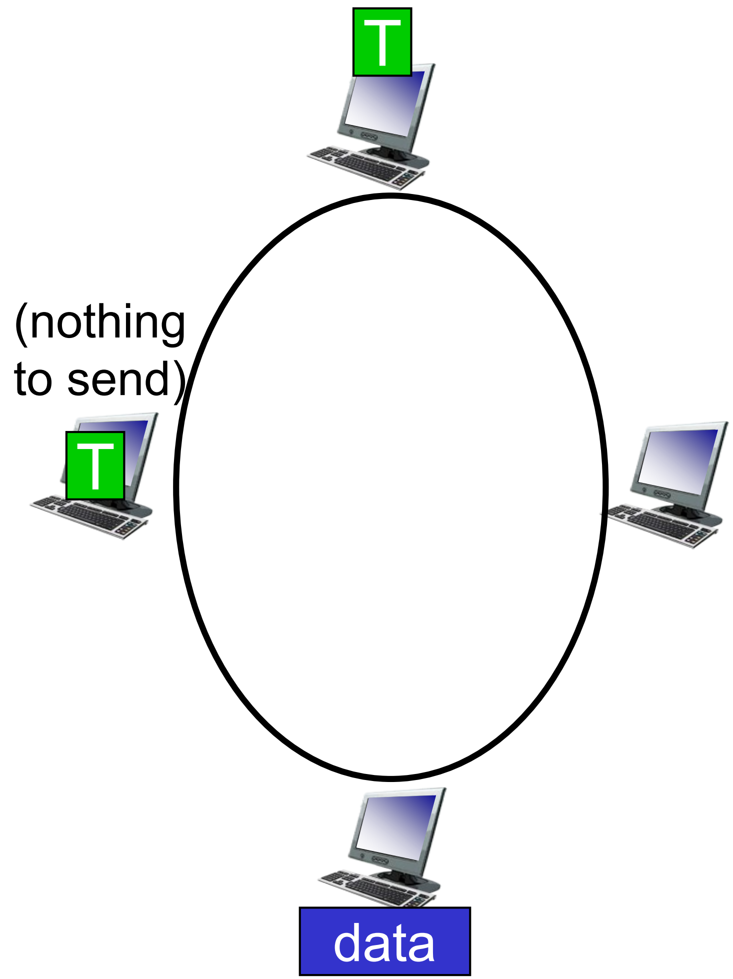
- master node "invites" slave nodes to transmit in turn.
- concerns:
  - polling overhead
  - single point of failure (master node)



# "Taking Turns" Protocols

## Token passing:

- control token is passed from one node to next sequentially.
- concerns:
  - token overhead
  - single point of failure (token)



# Multiple Access Protocols

- Multiple access protocols can be categorized into three broad classes:
  - Channel partitioning
    - divide channel into smaller "pieces" (e.g., time slots, frequency).
    - allocate piece to node for exclusive use.
  - "Taking turns"
    - nodes take turns to transmit.
  - **Random Access**
    - channel is not divided, collisions are possible.
    - "recover" from collisions.

# Random Access Protocols

- When node has packet to send
  - no *a priori* coordination among nodes
  - two or more transmitting nodes ➜ "collision"
- Random access protocols specify:
  - how to detect collisions
  - how to recover from collisions (e.g., via delayed retransmissions)
- We will skip the mathematical formulas on the efficiency of random access protocols.

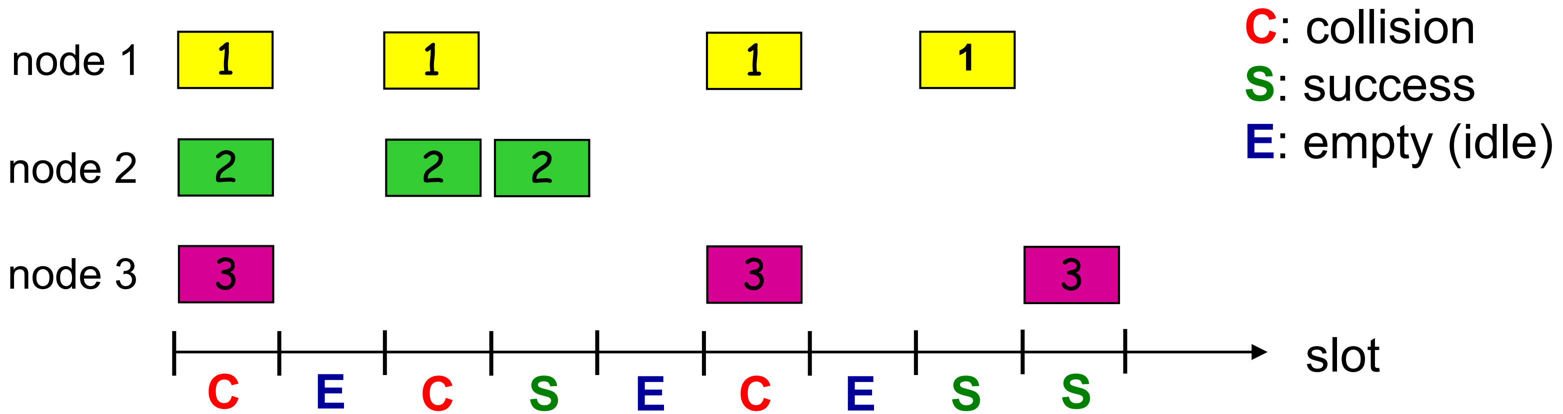
# Slotted ALOHA

**Assumptions:**

- All frames are of equal size.
- Time is divided into slots of equal length (length = time to transmit 1 frame).
- Nodes start to transmit only at the beginning of a slot.

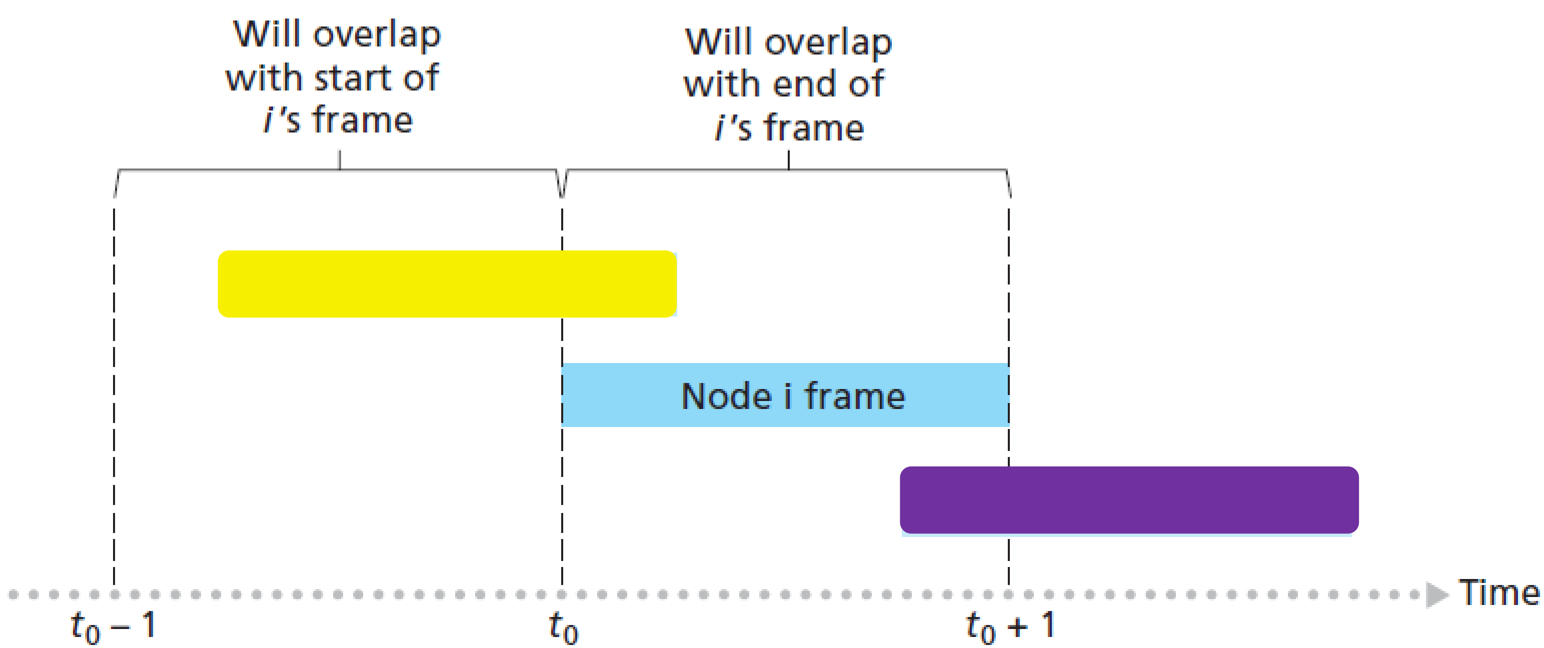
**Operations:**

- Listens to the channel while transmitting (**collision detection**).
- *if collision happens:* node retransmits a frame in each subsequent slot with probability $p$ until success.



# Pure (unslotted) ALOHA

- Even simpler: no slot, no synchronization
  - When there is a fresh frame: transmit immediately
  - Chance of collision increases:
    - frame sent at $t_0$ collides with other frames sent in $(t_0 - 1,\ t_0 + 1)$



# Carrier Sense Multiple Access

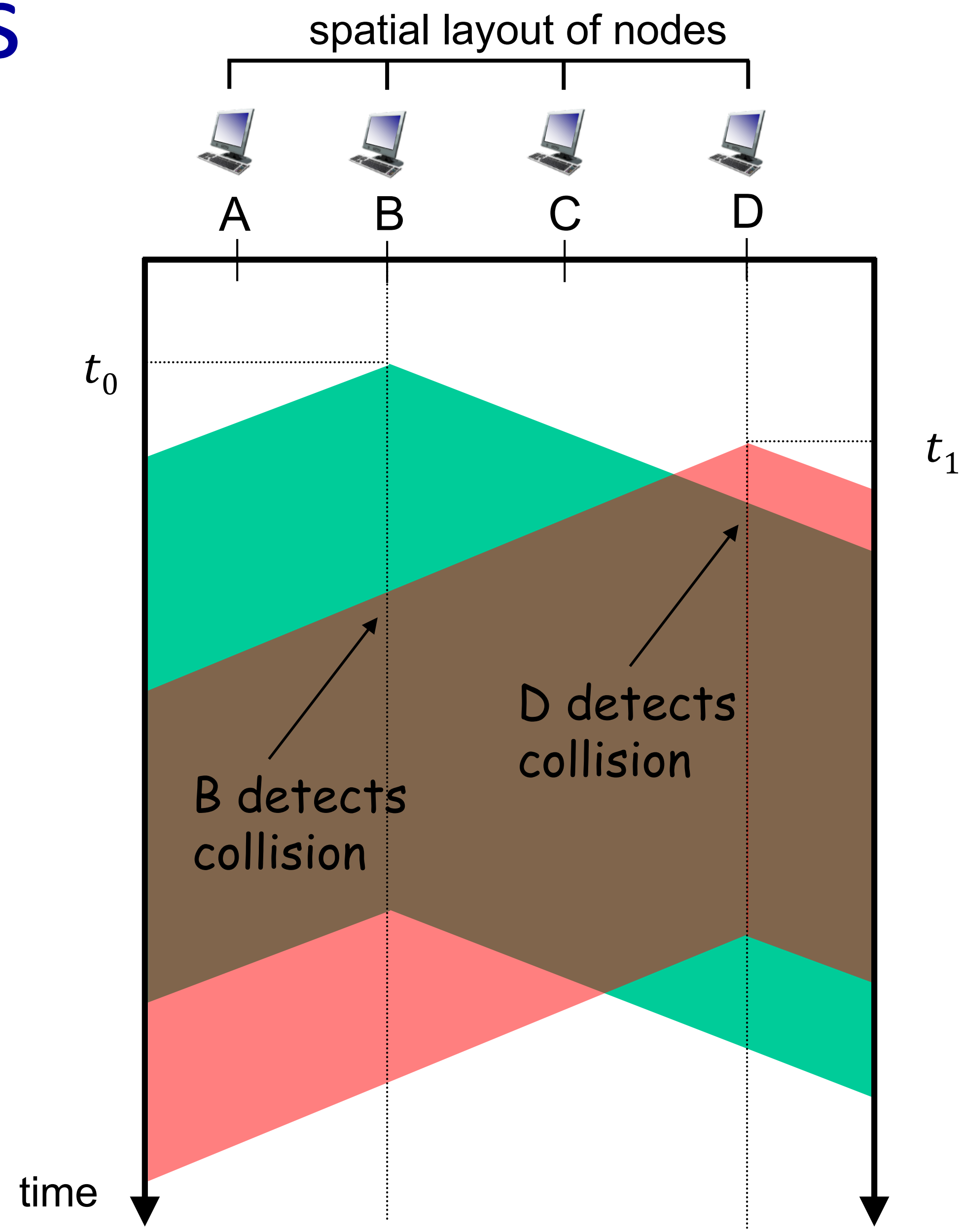
- CSMA (carrier sense multiple access)
  - Sense the channel before transmission:
    - *if channel is sensed idle,* transmit frame
    - *if channel sensed busy*, defer transmission

- Human analogy: don't interrupt others!

- Q: Will collision ever happen in CSMA?
  - collisions may still exist, e.g., when two nodes sense the channel idle at the same time and both start transmission.

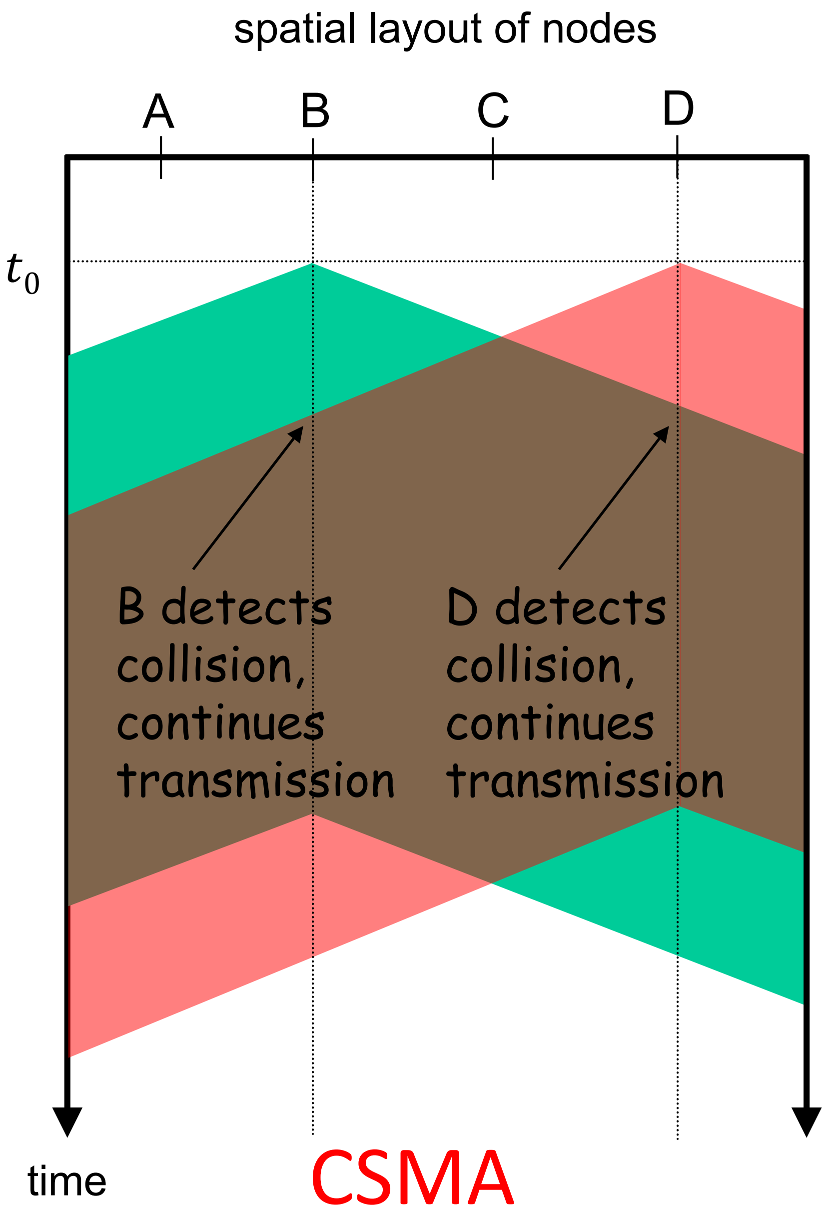
# CSMA Collisions

- Collisions can still occur:
  - propagation delay means two nodes may not hear each other's transmission immediately.

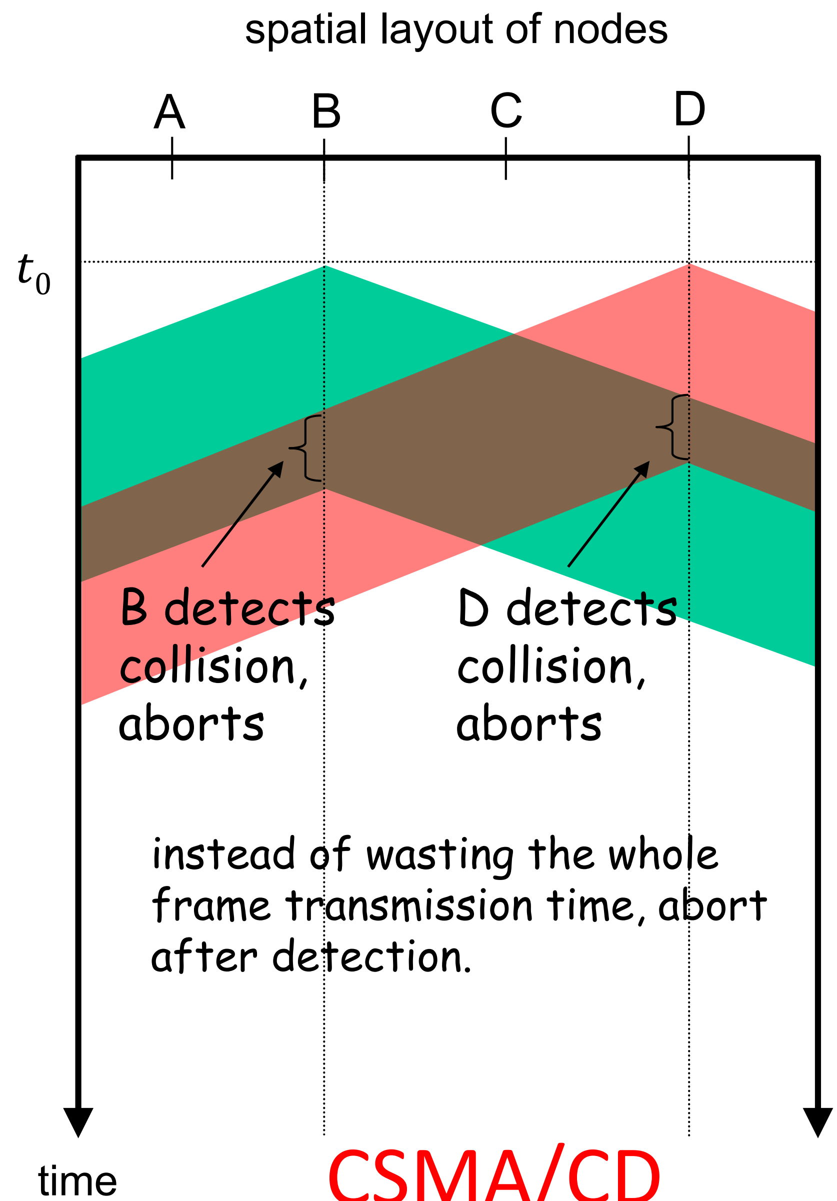


# CSMA/CD (Collision Detection)

- CSMA/CD
  - Carrier sensing & deferral as in CSMA
  - When collision is detected, transmission is aborted (reducing channel wastage).
  - Retransmit after a random amount of time.
    - *An example algorithm will be given in the next lecture*


spatial layout of nodes
A
B
C
D
$t_0$
B detects collision, continues transmission
D detects collision, continues transmission
time

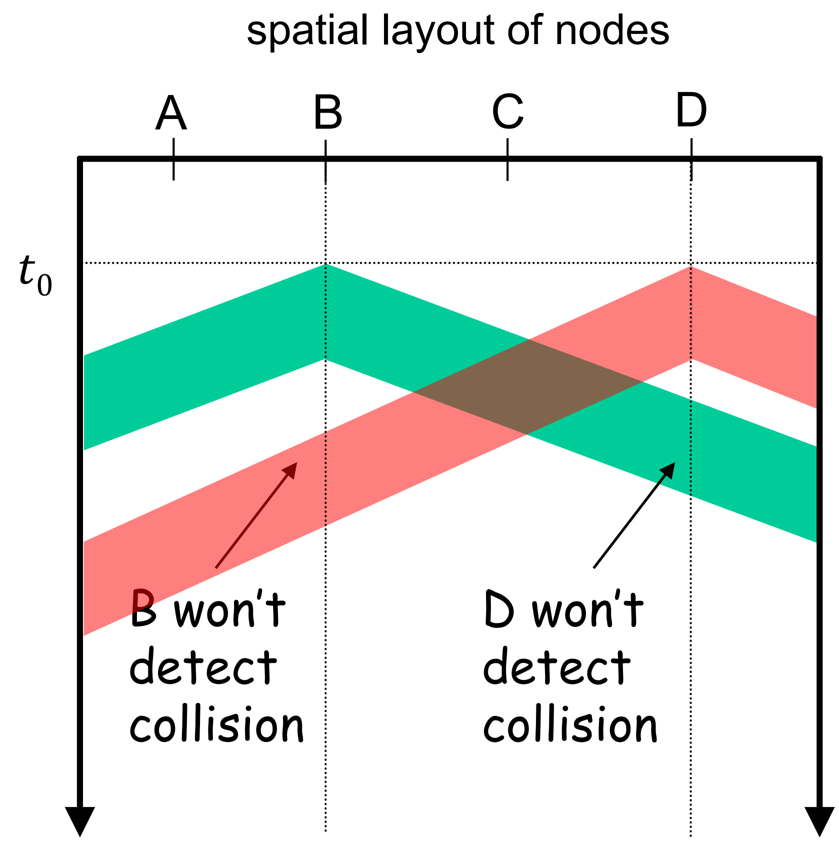
CSMA


spatial layout of nodes
A
B
C
D
$t_0$
B detects collision, aborts
D detects collision, aborts
instead of wasting the whole frame transmission time, abort after detection.
time

CSMA/CD

# Minimum Frame Size

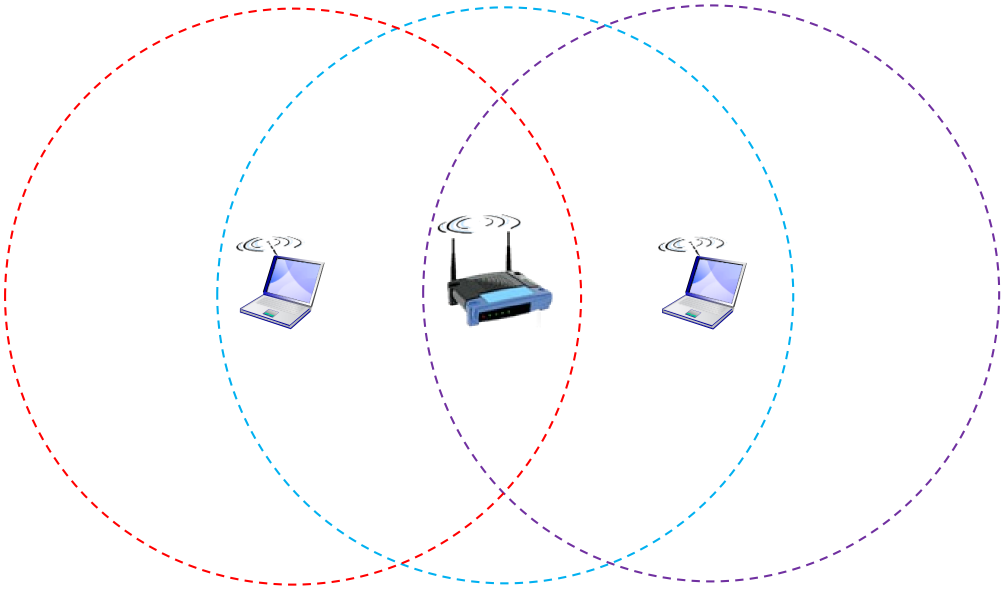
- What if the frame size is too small?
  - Collision happens but may not be detected by sending nodes.
    - No retransmission!




- For example, Ethernet requires a minimum frame size of 64 bytes.

# CSMA/CA (Collision Avoidance)

- Collision detection is easy in wired LANs, but difficult in wireless LANs. For example,



Hidden node problem (two laptops cannot detect each other)

- 802.11 (Wi-Fi) uses CSMA/CA protocol instead.
  - Receiver needs to return ACK if a frame is received OK.

# Lecture 9: Summary

- Channel partitioning
  - Divide channel by time, used in GSM
  - Divide channel by frequency, commonly used in radio, satellite systems
- Taking turns
  - polling from central site, used in Bluetooth
  - token passing, used in FDDI and token ring
- Random access
  - CSMA/CD used in Ethernet
  - CSMA/CA used in 802.11 Wi-Fi

# Lectures 9&10: Roadmap

6.1 Introduction to the Link Layer

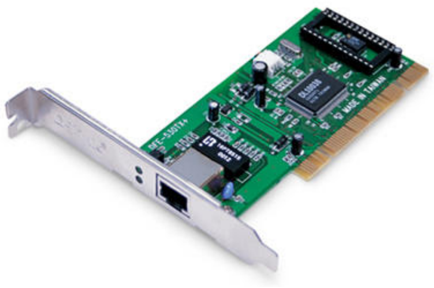
6.2 Error Detection and Correction

6.3 Multiple Access Links and Protocols

6.4 Switched Local Area Networks

- 6.4.1 Link Layer Addressing & ARP
- 6.4.2 Ethernet
- 6.4.3 Link-layer Switches

# MAC Address (1/2)



- Every adapter (NIC) has a MAC address (aka physical or LAN address).
  - Used to send and receive link layer frames.
  - When an adapter receives a frame, it checks if the destination MAC address of the frame matches its own MAC address.
    - If yes, adapter extracts the enclosed datagram and passes it to the protocol stack.
    - If no, adapter simply discards the frame without interrupting the host.

# MAC Address (2/2)

- MAC address is typically 48 bits, burned in NIC ROM (sometimes software settable).
  - Example: **5C-F9-DD-E8-E3-D2** hexadecimal (base 16) notation
  - MAC address allocation is administered by IEEE.
    - The first three bytes identifies the vendor of an adapter.
  - Several websites allow us to check the vendor given a MAC address, e.g.:

https://macvendors.com/

# IP Address vs. MAC Address

- IP address
  - 32 bits in length
  - network-layer address used to move datagram from source to dest.
  - Dynamically assigned; hierarchical (to facilitate routing)
  - Analogy: postal address
- MAC address
  - 48 bits in length
  - link-layer address used to move frame over every single link.
  - Permanent, to identify the hardware (adapter)
  - Analogy: NRIC number

# ARP: Address Resolution Protocol

- Question: How to know the MAC address of a receiving host, knowing its IP address?
  - Use ARP [RFC 826]

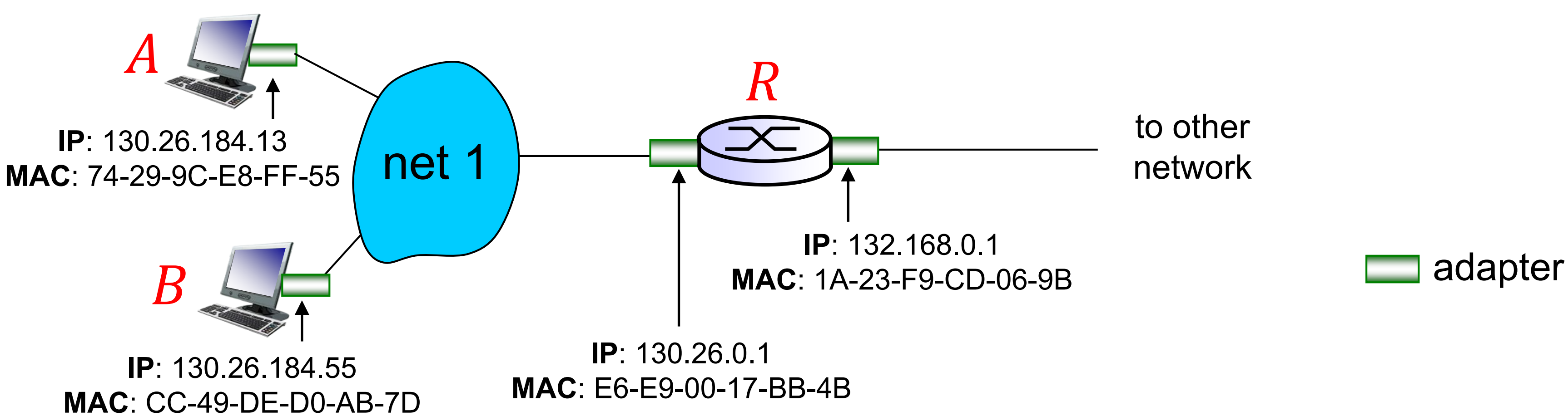
- Each IP node (host, router) has an ARP table.
  - Stores the mappings of IP address and MAC address of other nodes in the same subnet.

< IP address; MAC address; TTL >

time after which address mapping will be forgotten (typically a few minutes on Windows)

# Sending Frame in the Same Subnet

- Suppose $A$ wants to send data to $B$. They are in the same subnet.
  1. If $A$ knows $B$'s MAC address from its ARP table
     - create a frame with $B$'s MAC addresses and send it.
     - Only $B$ will process this frame.
     - Other nodes may receive but will ignore this frame.
  2. What if $A$ is not aware of $B$'s MAC address?



# Sending Frame in the Same Subnet

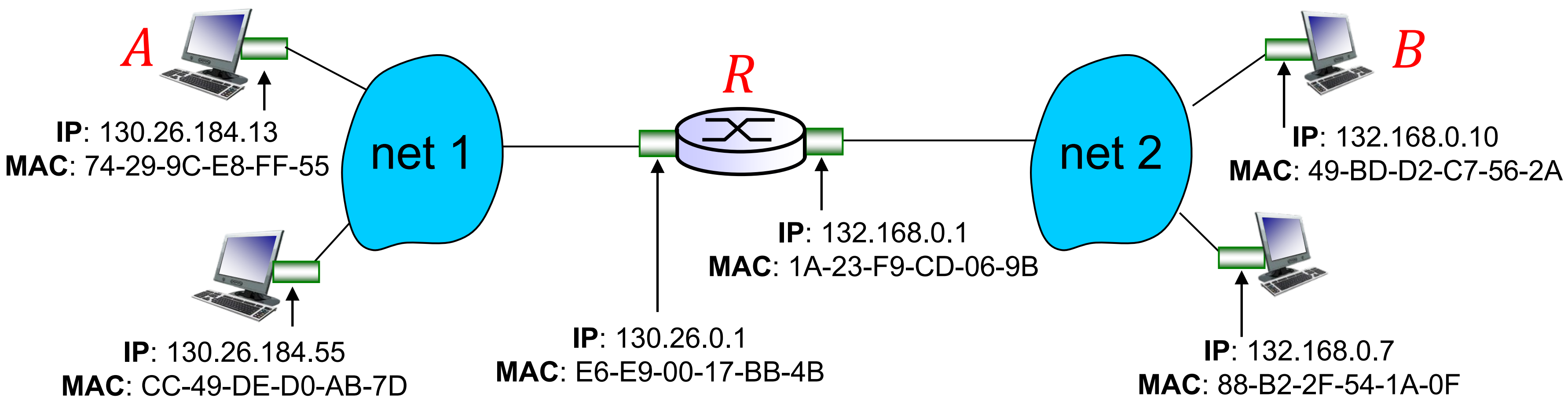
- What if $B$'s MAC address is not in $A$'s ARP table?

  1. $A$ broadcasts an ARP query packet, containing $B$'s IP address.
     - Dest MAC address set to FF-FF-FF-FF-FF-FF
     - All the other nodes in the same subnet will receive this ARP query packet, but only $B$ will reply it.
  2. $B$ replies to $A$ with its MAC address.
     - Reply frame is sent to $A$'s MAC address.
  3. $A$ caches $B$'s IP-to-MAC address mapping in its ARP table (until TTL expires).

Question: how to determine if $B$ is in the same subnet?

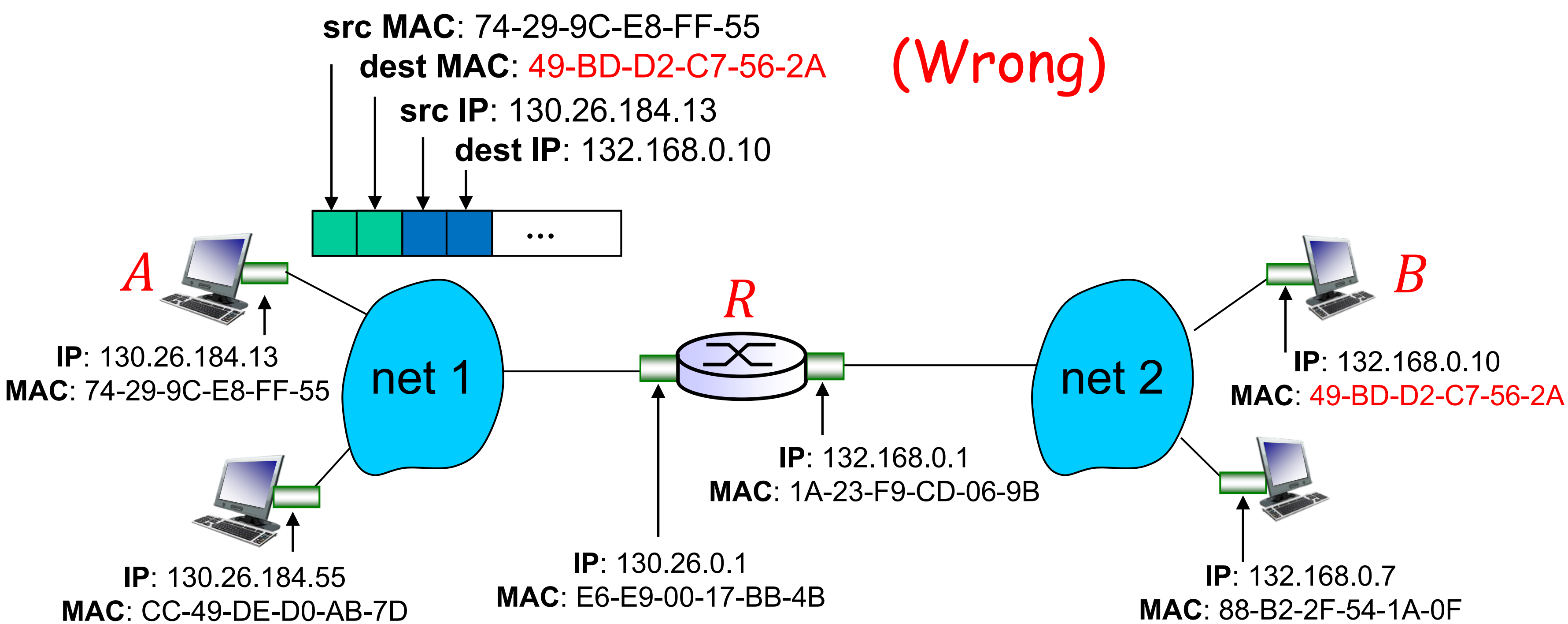
# Sending Frame to Another Subnet

- Question: What if we send data to a host in another subnet?
  - For example, $A$ sends datagram to $B$ in another subnet.



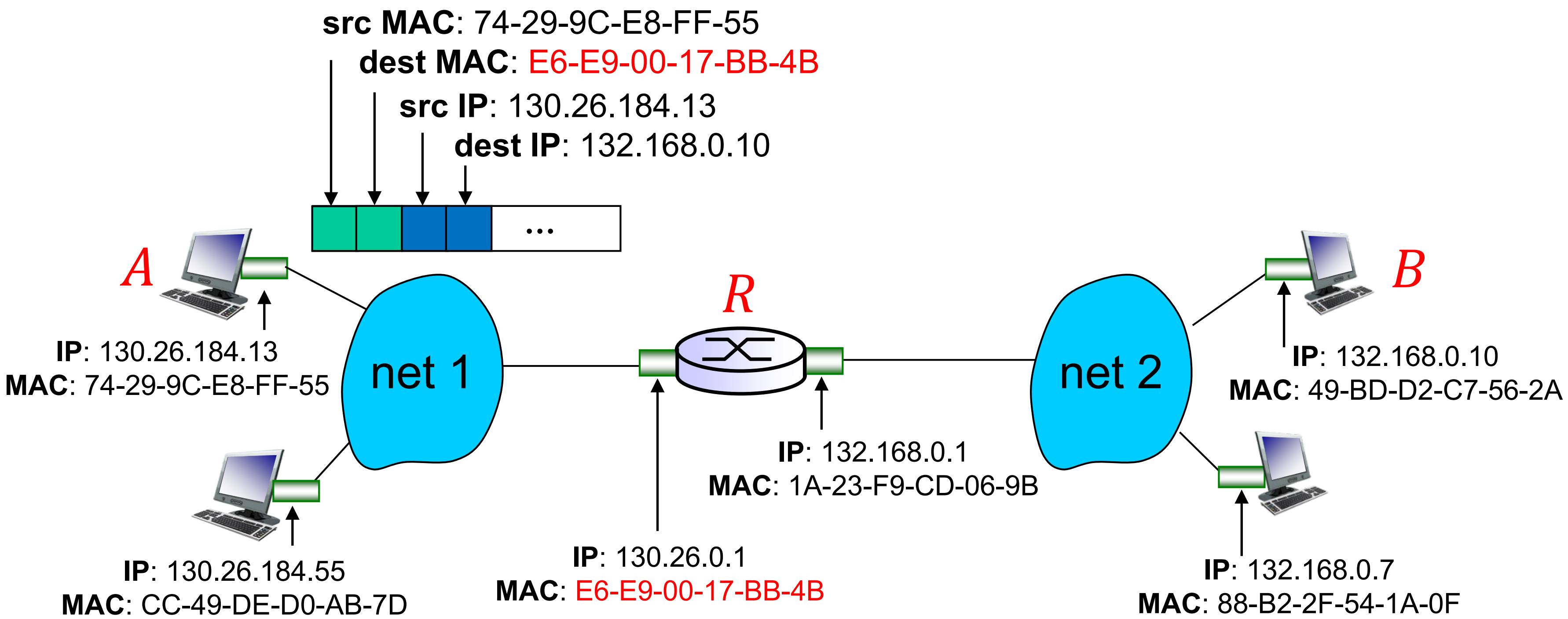
# Sending Frame to Another Subnet

- $A$ sends datagram to $B$ in another subnet.
  - Can $A$ create a frame as follows?
    - No. all adapters in net 1 will ignore this frame because of the mismatch of destination MAC address.



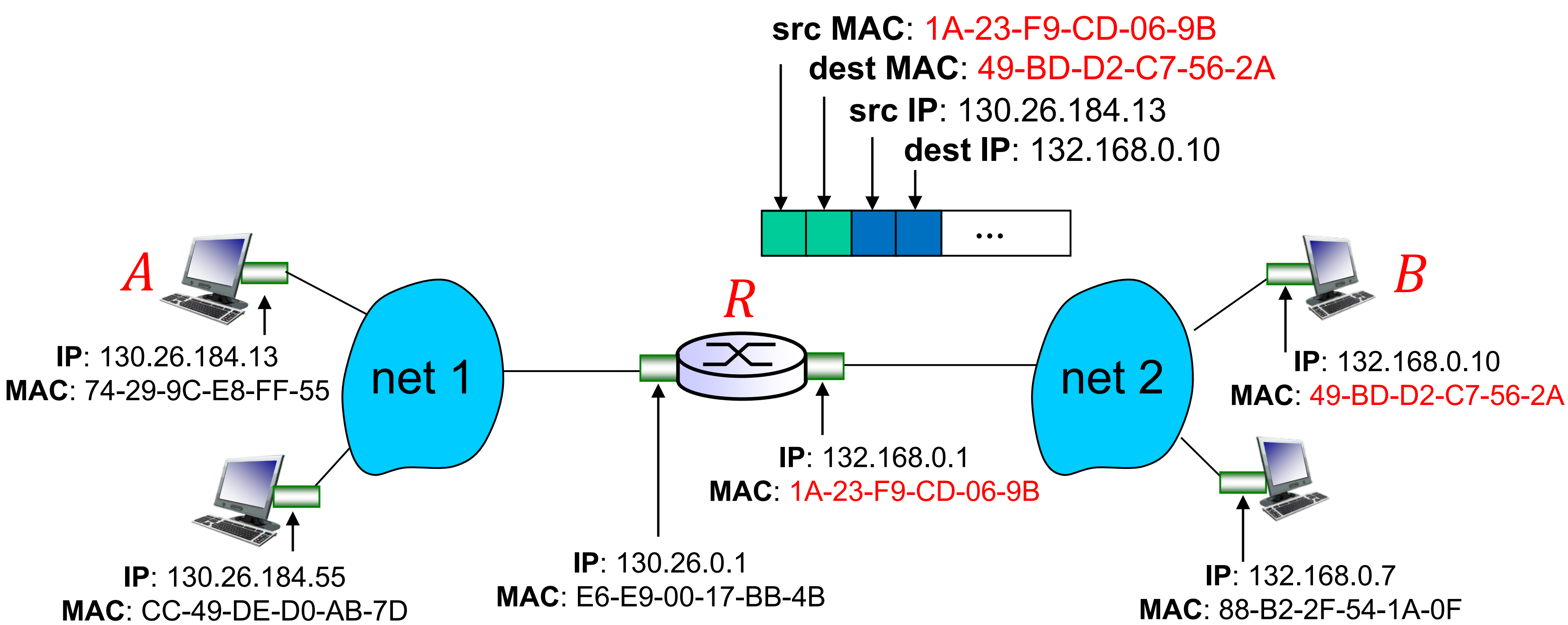
# Sending Frame to Another Subnet

- $A$ sends datagram to $B$ in another subnet.
  - $A$ should create a link-layer frame with (1) $R$'s MAC address (2) $B$'s IP address as destination.



# Sending Frame to Another Subnet

- $A$ sends datagram to $B$ in another subnet.
  - $R$ will move datagram to outgoing link and construct a new frame with $B$'s MAC address.



# Lectures 9&10: Roadmap

6.1 Introduction to the Link Layer

6.2 Error Detection and Correction

6.3 Multiple Access Links and Protocols

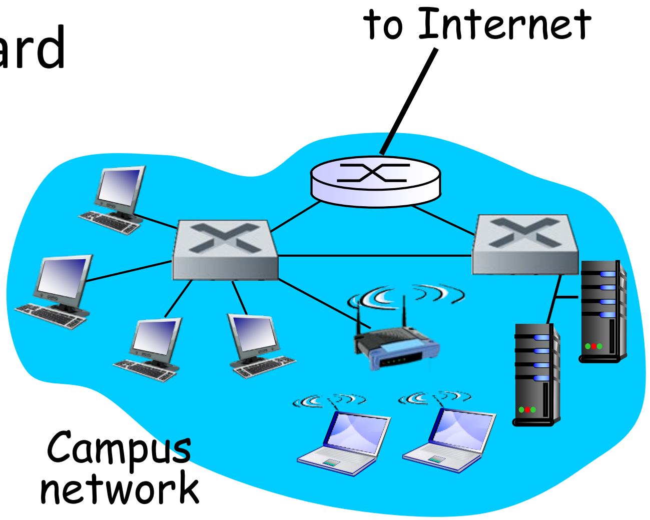
6.4 Switched Local Area Networks

- 6.4.1 Link Layer Addressing & ARP
- 6.4.2 Ethernet
- 6.4.3 Link-layer Switches

# Local Area Network (LAN)

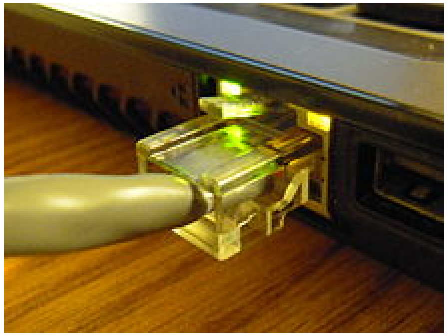
- LAN is a computer network that interconnects computers within a geographical area such as office building or university campus.

- LAN technologies:
  - IBM Token Ring: IEEE 802.5 standard
  - Ethernet: IEEE 802.3 standard
  - Wi-Fi: IEEE 802.11 standard
  - Others

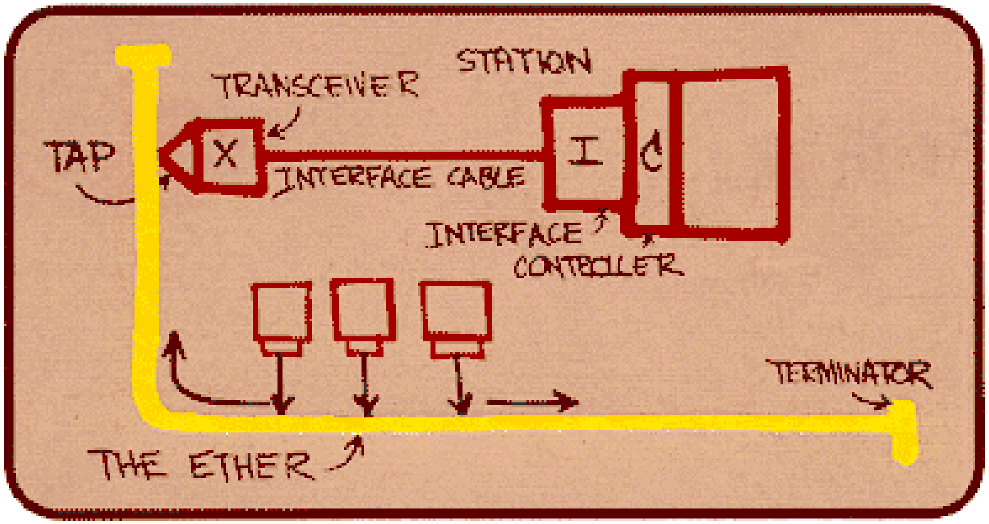


# Ethernet

❖ "dominant" wired LAN technology:
- Developed in mid 1970s
- Standardized by Xerox, DEC, and Intel in 1978
- Simpler and cheaper than token ring and ATM



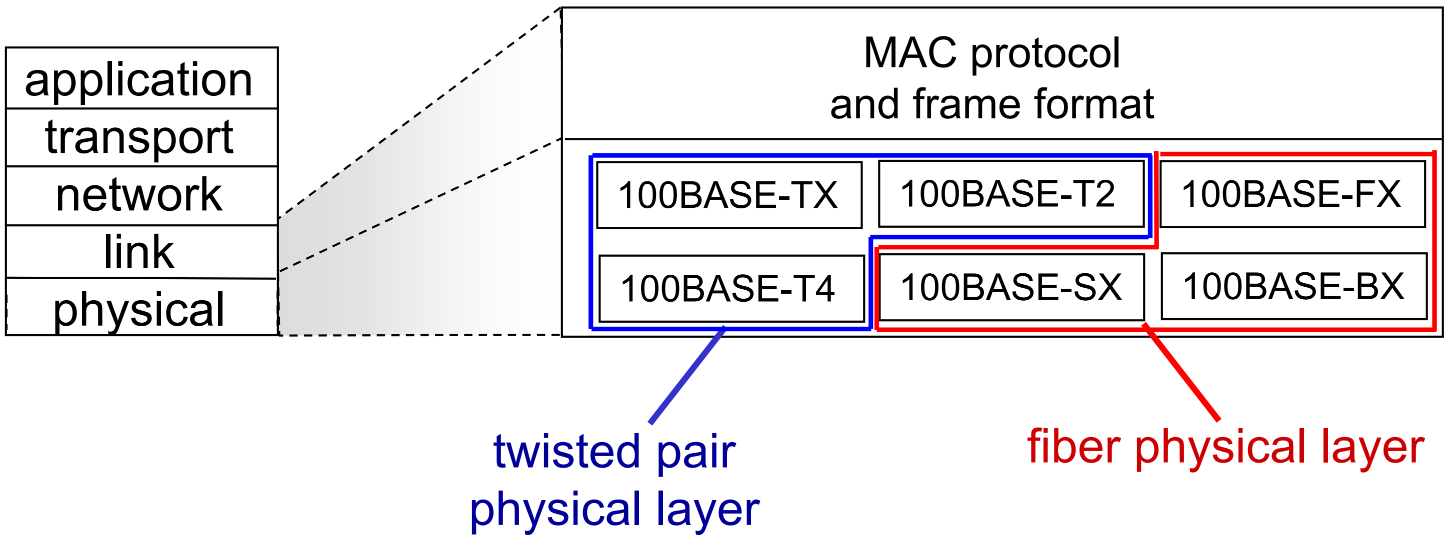
Ethernet connection
(Source: Wikipedia)




*Metcalfe's Ethernet sketch*

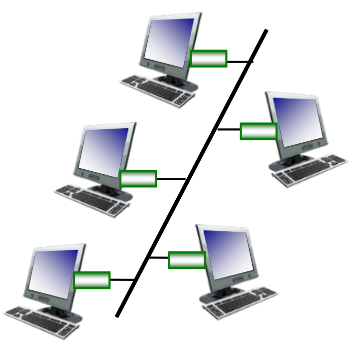
# 802.3 Ethernet Standards

- A series of Ethernet standards are developed over the years.
  - Different speeds: 2 Mbps, 10 Mbps, 100 Mbps, 1 Gbps, 10 Gbps, 100 Gbps
  - Different physical layer media: cable, fiber optics
  - MAC protocol and frame format remain unchanged

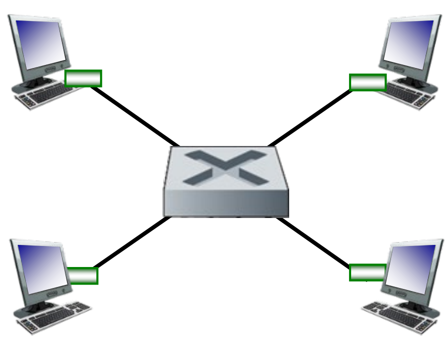


# Ethernet: Physical Topology

- Bus topology: popular in mid 90s
  - all nodes can collide with each other
- Star topology: prevails today
  - switch in center
  - nodes do not collide with each other



Ethernet with bus topology



Ethernet with star topology

# Ethernet Frame Structure (1/2)

- Sending NIC (adapter) encapsulates IP datagram in Ethernet frame.

| 8 bytes | 6 | 6 | 2 | 46 - 1500 | 4 |
|---|---|---|---|---|---|
| Preamble | Dest Addr | Src Addr | Type | Payload | CRC |

- *Preamble:*
  - 7 bytes with pattern **`10101010`** followed by 1 byte with pattern **`10101011`**.
  - used to synchronize receiver and sender clock rates.

# Ethernet Frame Structure (2/2)

| 8 bytes | 6 | 6 | 2 | 46 - 1500 | 4 |
|---|---|---|---|---|---|
| Preamble | Dest Addr | Src Addr | Type | Payload | CRC |

- *Source and dest MAC address:*
  - If NIC receives a frame with matching destination address, or with broadcast address, it passes data in the frame to network layer protocol.
  - Otherwise, NIC discards frame.
- *Type:* Indicates higher layer protocol (mostly IP).
- *CRC:* corrupted frame will be dropped.

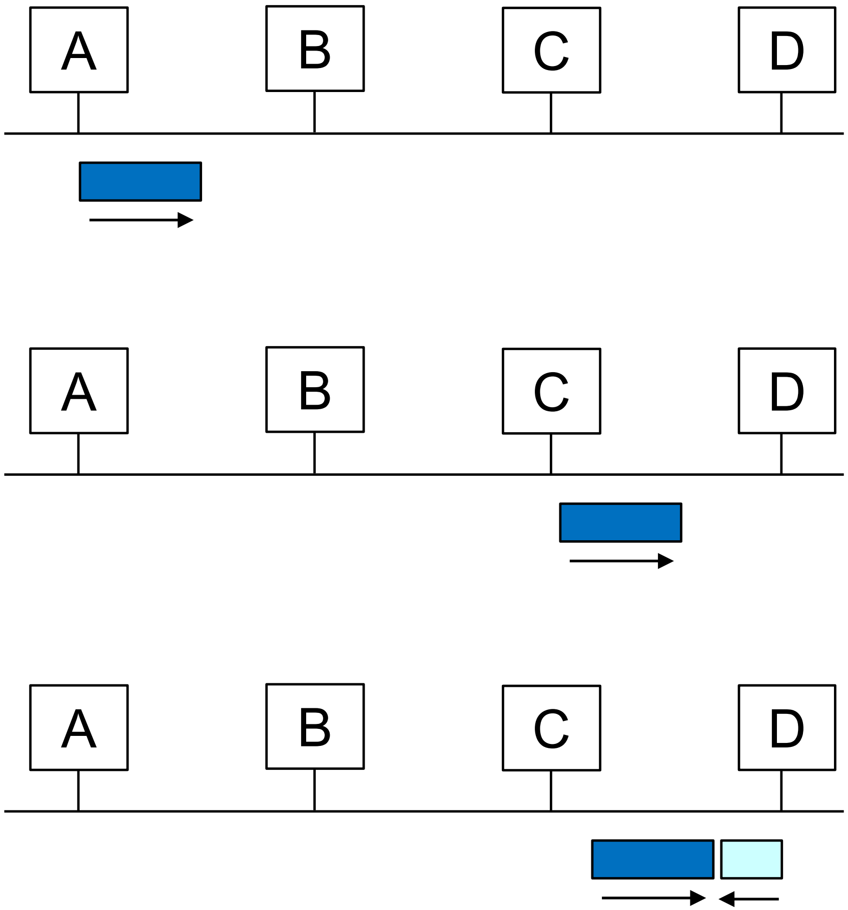
# Ethernet Data Delivery Service

- Connectionless: no handshaking between sending and receiving NICs.
- Unreliable: receiving NIC doesn't send ACK or NAK to sending NIC.
  - data in dropped frames will be recovered only if initial sender uses higher layer rdt (e.g. TCP); otherwise dropped data is lost.
- Ethernet's multiple access protocol: CSMA/CD with binary (exponential) backoff.

# Collisions in Bus Topology Ethernet

- Collision may happen in Ethernet of bus topology.

- For example:
  - $A$ sends a frame at time $t$.
  - $A$'s frame reaches $D$ at time $t + d$.
  - $D$ begins transmission at time $t + d - 1$ and collides with $A$'s frame.



# Ethernet CSMA/CD Algorithm

1. NIC receives datagram from network layer, creates frame.
2. If NIC senses channel idle, starts frame transmission. If NIC senses channel busy, waits until channel idle, then transmits.
3. If NIC transmits entire frame without detecting another transmission, NIC is done with frame!
4. If NIC detects another transmission while transmitting, aborts and sends jam signal.
5. After aborting, NIC enters binary back-off:
   - after $m^{th}$ collision, NIC chooses $K$ at random from {0, 1, 2, …, $2^m$-1}.
   - NIC waits $K$*512 bit times, returns to Step 2.

# Ethernet CSMA/CD Algorithm

## Exponential backoff:

- After 1st collision: choose $K$ at random from {0, 1}; wait $K$ * 512 bit transmission times before retransmission.
- After 2nd collision: choose $K$ from {0, 1,..., $2^2$-1}.

  ...
- After $m^{th}$ collision, choose $K$ at random from {0, 1, ..., $2^m$-1}

- *Goal*: adapt retransmission attempts to estimated current load
  - More collisions implies heavier load.
  - longer back-off interval with more collisions.

# Lectures 9&10: Roadmap

6.1 Introduction to the Link Layer

6.2 Error Detection and Correction

6.3 Multiple Access Links and Protocols

6.4 Switched Local Area Networks

- 6.4.1 Link Layer Addressing & ARP
- 6.4.2 Ethernet
- 6.4.3 Link-layer Switches

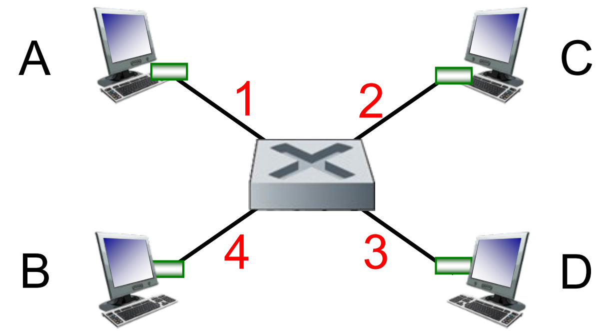
# Ethernet Switch

- A link-layer device used in LAN
  - Store and forward Ethernet frames
  - Examine incoming frame's MAC address, selectively forward frame to one-or-more outgoing links.
- Transparent to hosts
  - No IP address
  - Hosts are unaware of the presence of switches

a 50-port Ethernet switch
(Source: Wikipedia)

# Ethernet Switch

- In Ethernet of star topology, hosts have dedicated connection to switch.



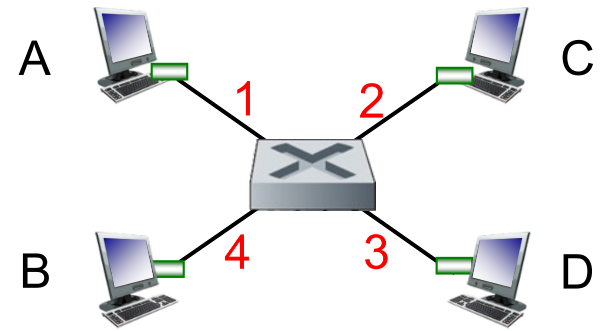

A switch with 4 interfaces (1, 2, 3, 4)

- Switch buffers frames and is full duplex.
  - A and D can send frames to each other simultaneously.

- Ethernet protocol is used on each link, but no collisions!

# Switch Forwarding Table

- Q: how does switch know A is reachable via interface 1, B is reachable via interface 4?
- A: each switch has a switch table.
  - Format of entry:



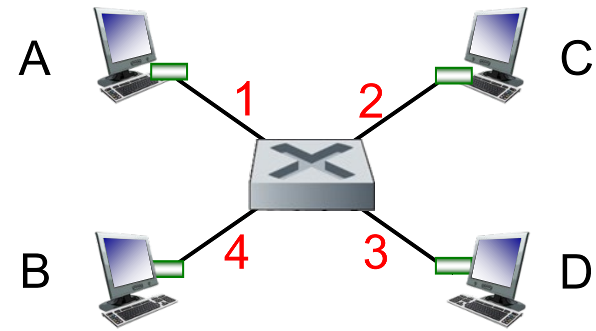

A switch with 4 interfaces (1, 2, 3, 4)

< *MAC address of host, interface to reach host, TTL* >

- Q: how are entries created and maintained in a switch table?

# Switch: Self-learning

- Switch *learns* which hosts can be reached through which interfaces.
  - When receiving a frame from $A$, note down the location of $A$ in switch table.
  - If destination $B$ is found in the table, forward the frame onto that link.
  - If destination $B$ is unknown, broadcast the frame to all outgoing links.



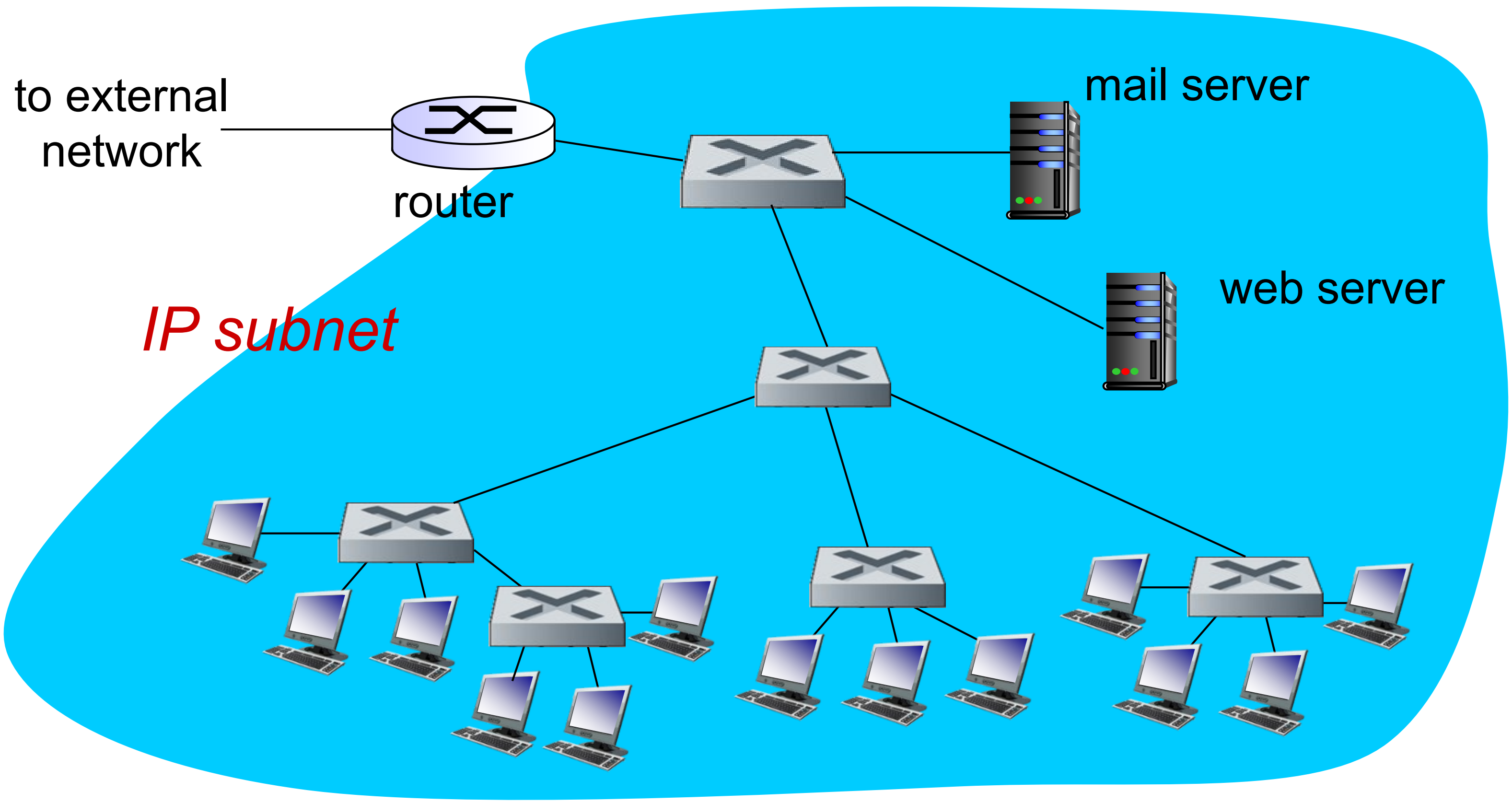

A switch with 4 interfaces (1, 2, 3, 4)

| MAC addr | Interface | TTL |
|---|---|---|
| A | 1 | 60 |
| | | |
| | | |
| | | |

Switch table (initially empty)

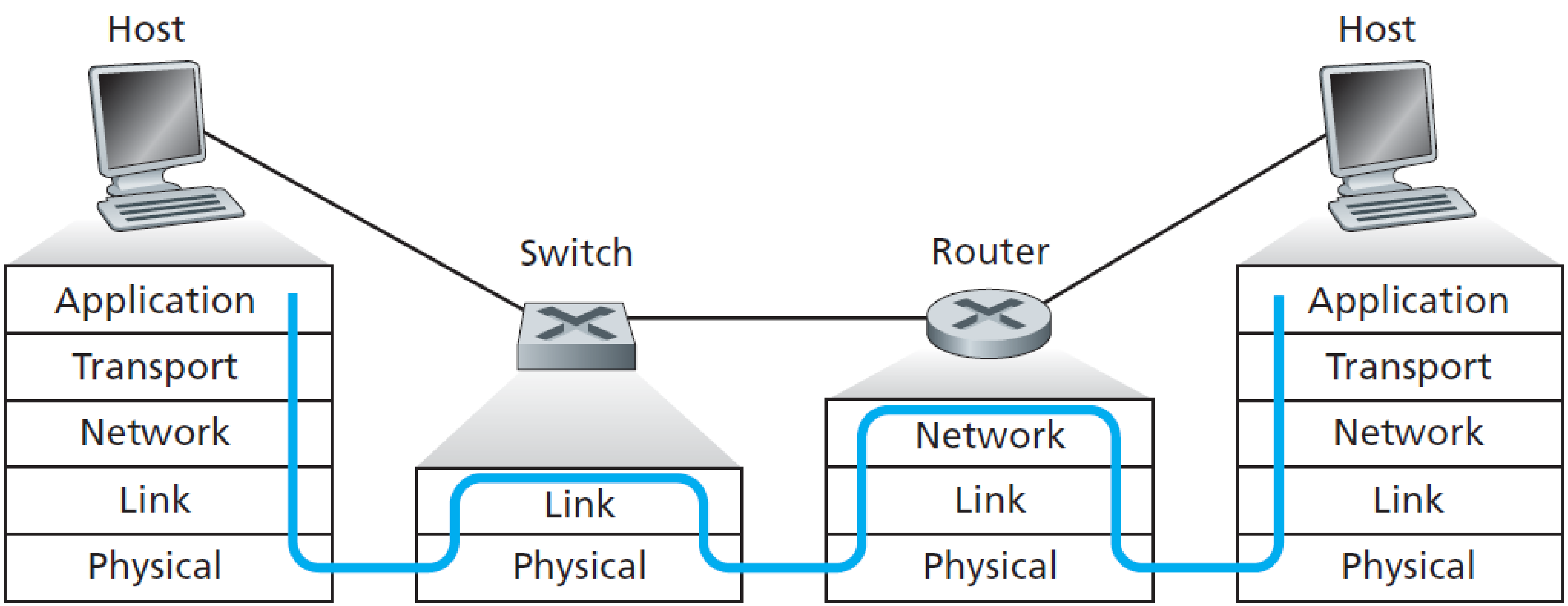
# Interconnecting Switches

- Switches can be connected in hierarchy.



# Switches vs. Routers

- **Routers**
  - Check IP address
  - Store-and-forward
  - Compute routes to destination
- **Switches**
  - Check MAC address
  - Store-and-forward
  - Forward frame to outgoing link or broadcast



# Lecture 10: Summary

- ARP [RFC 826] resolves the mapping from network layer (IP) address to link layer (MAC) address.

- Instantiation and implementation of link layer technologies.
  - Ethernet
  - CSMA/CD protocol with binary back-off
  - Ethernet switch and switch table