Chapter 04 Medium Access Control

Associate Prof. Hong Zheng （郑宏）
Computer School
Beijing Institute of Technology

Key Points

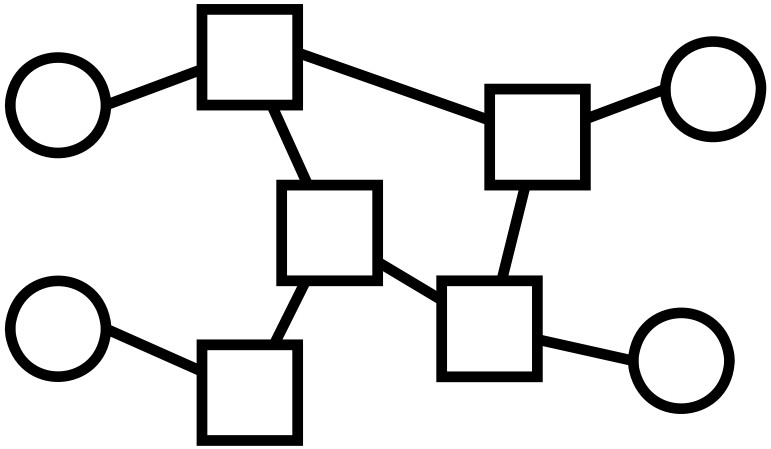
MAC	LAN model & Topology	熟练掌握
	CSMA/CD & analysis, CSMA/CA	
Ethernet	Frame format, MAC Address	熟练掌握
	CSMA/CD & Backoff Algorithm	
	Standards	掌握
Inter-connection	repeater/hub, bridge/switch, router, gateway	掌握
LAN Switching	Methods	掌握
	Learning, Filtering & Forwarding	熟练掌握
	Spanning Tree Protocol	掌握
VLAN	Benefits,Types, Trunk, 802.1Q tagged frame	熟练掌握

Questions to be answered

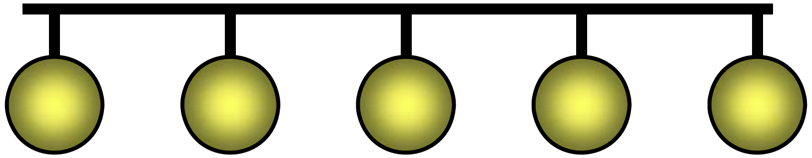
- In broadcast networks, how is the channel divided between competing users?
- What is Medium Access Control (MAC)?
- What protocols are used for allocating a multiple access channel?

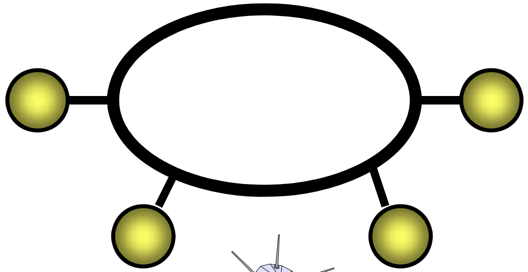
Chapter 4: Roadmap

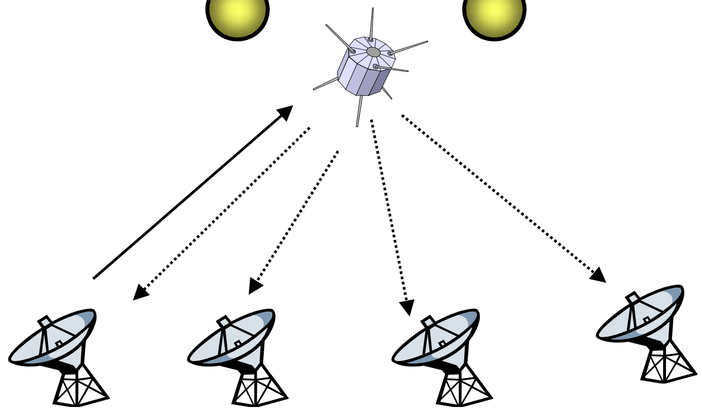
- Medium Access Control
- Local Area Networks (LANs) and IEEE 802
- Ethernet
- Wireless LAN
- LAN Interconnection
- LAN Switching
- VLAN



Broadcast channel







Point-point link:

- Error and flow control.

Broadcast link:

- Media access control.
- Scalability.

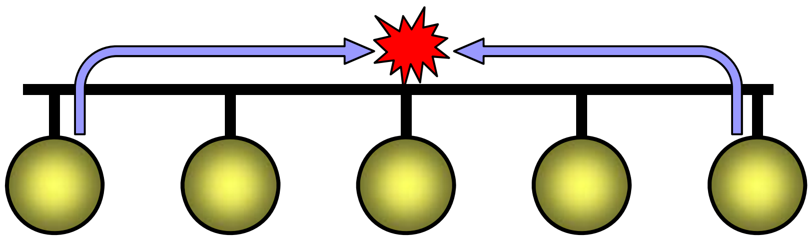
Multiple Access

- **What are the multiple access?**

  Multiple hosts sharing the same medium. If one station sends, all the others get to hear it.

- **What are the new problems?**

  When two or more nodes transmit at the same time, their frames will collide and the link bandwidth is wasted during collision



Multiple Access

- **For Broadcast network and shared channel, the key issue is:**

  How to determine who gets to use the channel when there is competition for it?

Solution

Allocate the channel to one of the competing stations.

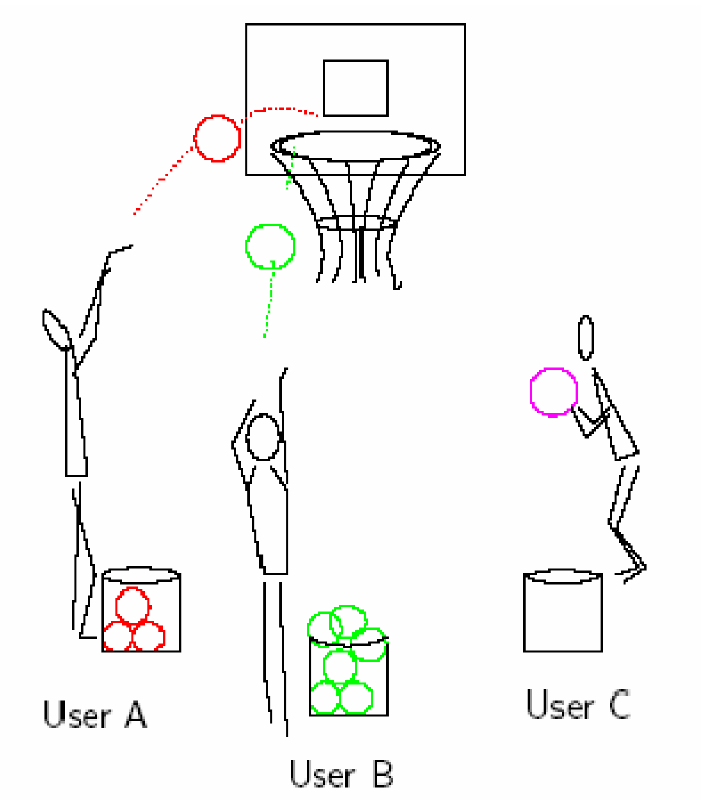
The Channel Allocation Problem

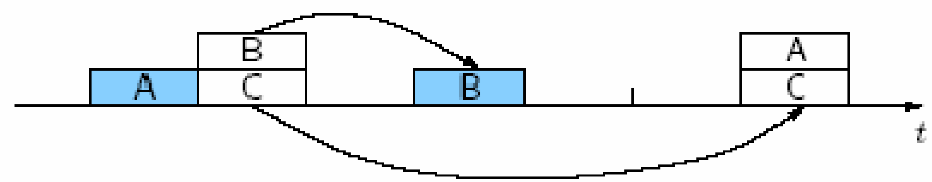
- **Requirements:**
  - ***efficiently*** ***i.e.*** **maximize message throughput**
  - ***fairly***
  - ***minimize*** **mean waiting time**
- **Two schemes to allocate a single channel:**
  - **Static Channel Allocation**
  - **Dynamic Channel Allocation**

Static Channel Allocation

- **Each user is statically allocated the bandwidth or time slots.**
  - **FDM and TDM**
  - **No interference between users.**
- **inefficient**
- **Poor performance**

Dynamic Channel Allocation





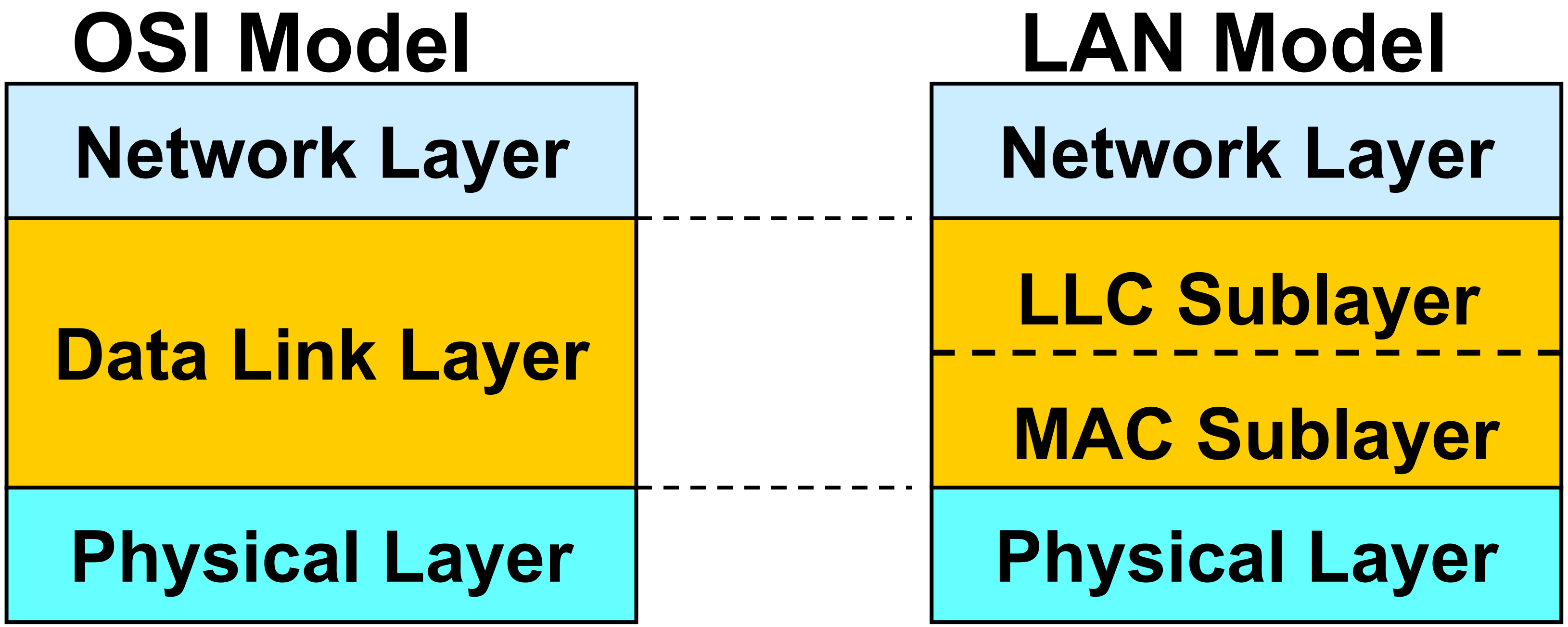
Objectives

- Small delay in light traffic.
- **Bounded** delay for a large (possibly infinite) number of users.

Collision may occur

What is MAC?

- **MAC: medium access control.**
- **MAC is a sublayer of the Data-link layer.**



Data link layer divided into two functionality-oriented sublayers

MAC Protocols: a taxonomy

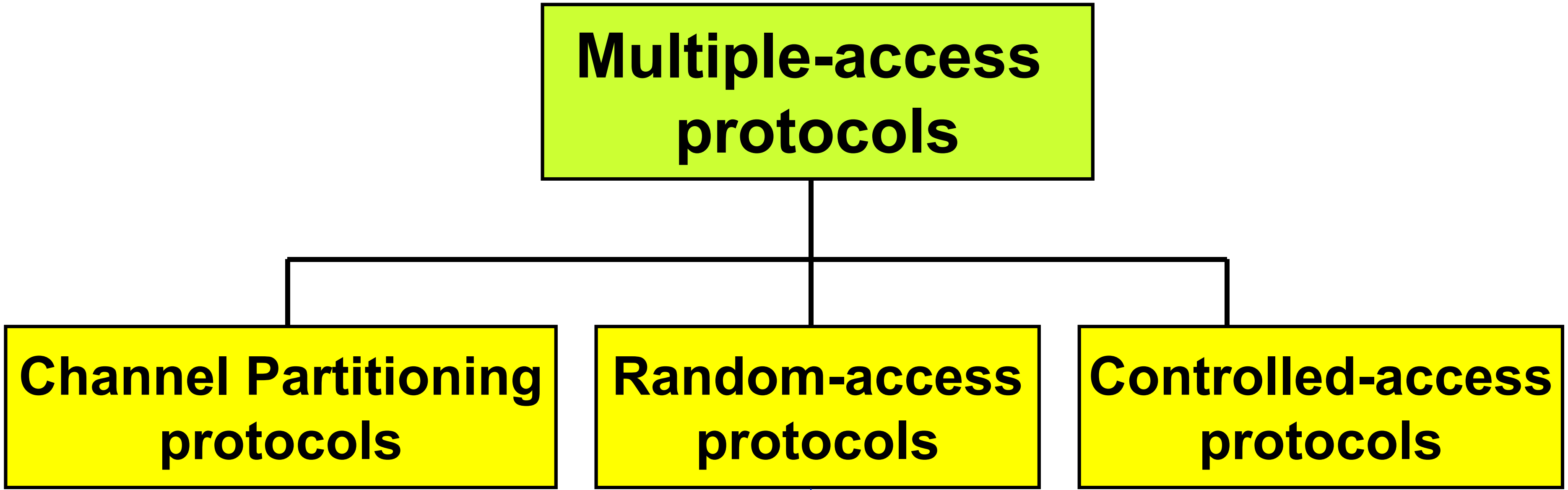
- The **MAC protocols** used to determine who goes next on a multi-access channel belongs to a MAC sublayer.
- Three broad classes:
  - Channel Partitioning
  - Random-access
  - Controlled-access protocols

MAC Protocols: a taxonomy

- Multiple-access protocols
  - Channel Partitioning protocols
  - Random-access protocols
  - Controlled-access protocols

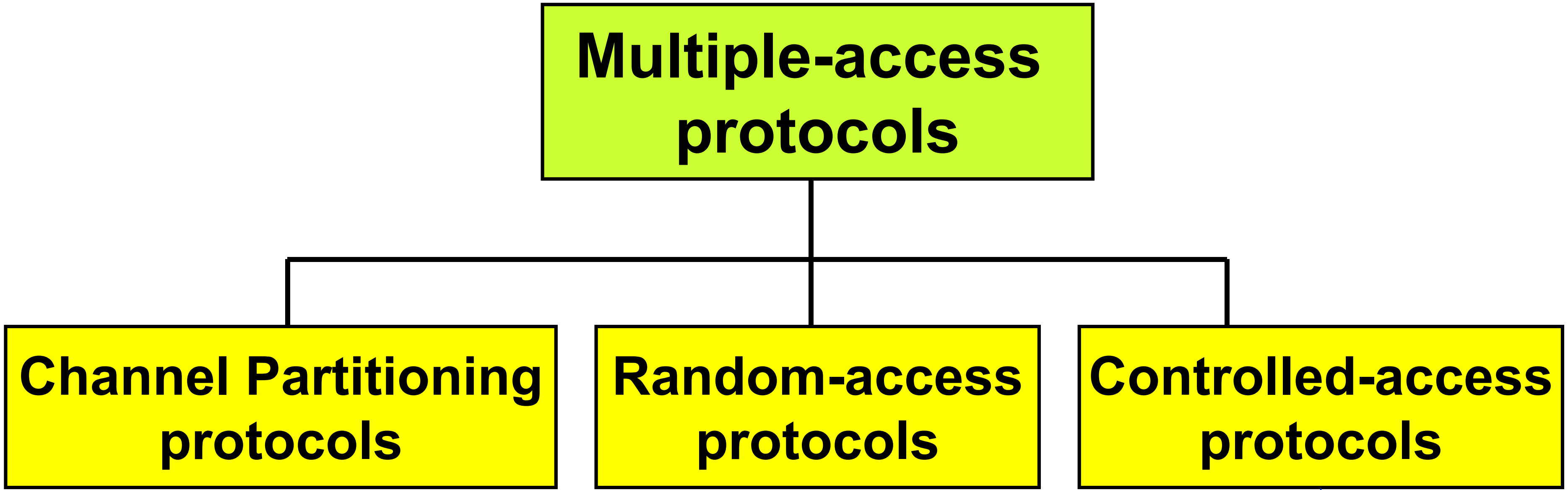
•Divide channel into smaller "pieces" (time slots, frequency, code)
•Allocate piece to node for exclusive use

MAC Protocols: a taxonomy



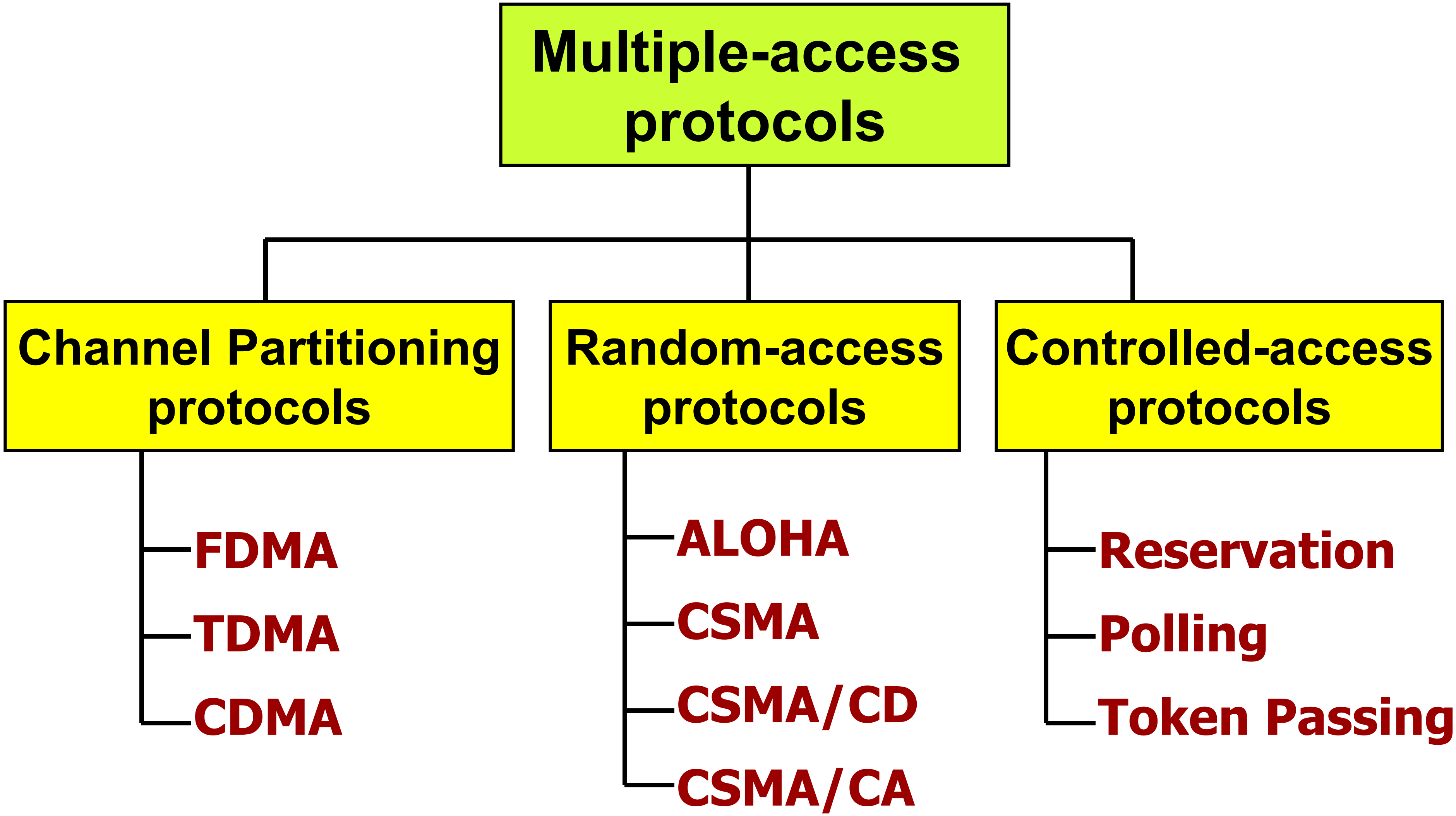
- Channel not divided, allow collisions
- “Recover” from collisions

MAC Protocols: a taxonomy



•Nodes take turns to shared medium so that every station has chance to transfer (fair protocol).

MAC Protocols: a taxonomy



Random Access

When node has packet to send:

- **transmit at full channel data rate.**
- **no *a priori* coordination among nodes.**

Random Access

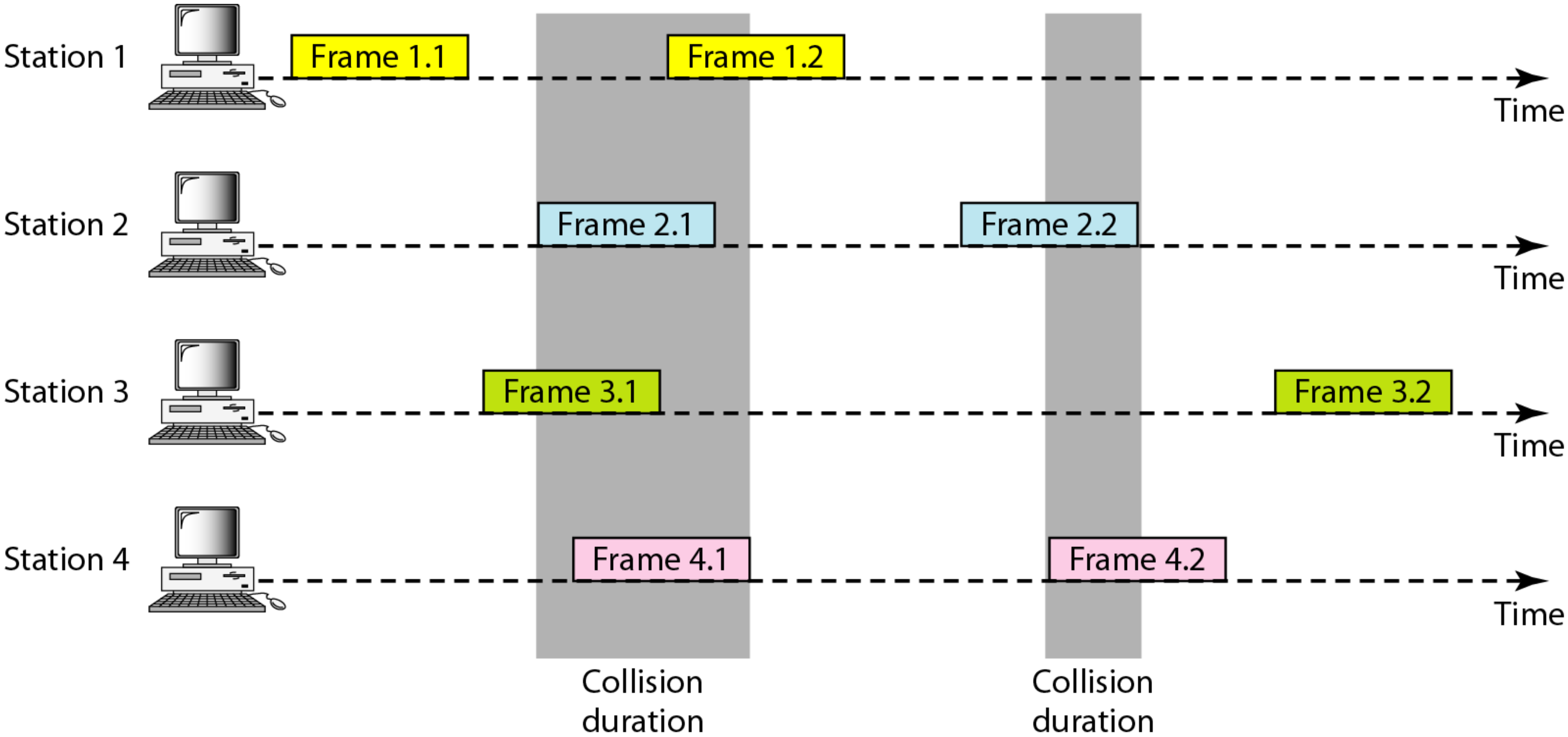
When two or more transmitting nodes ➜ "collision"

- Random access MAC protocol specifies:
  - how to detect collisions
  - how to recover from collisions (e.g., via delayed retransmissions)
- Examples:
  - ALOHA: pure ALOHA, slotted ALOHA
  - CSMA, CSMA/CD, CSMA/CA

Pure ALOHA

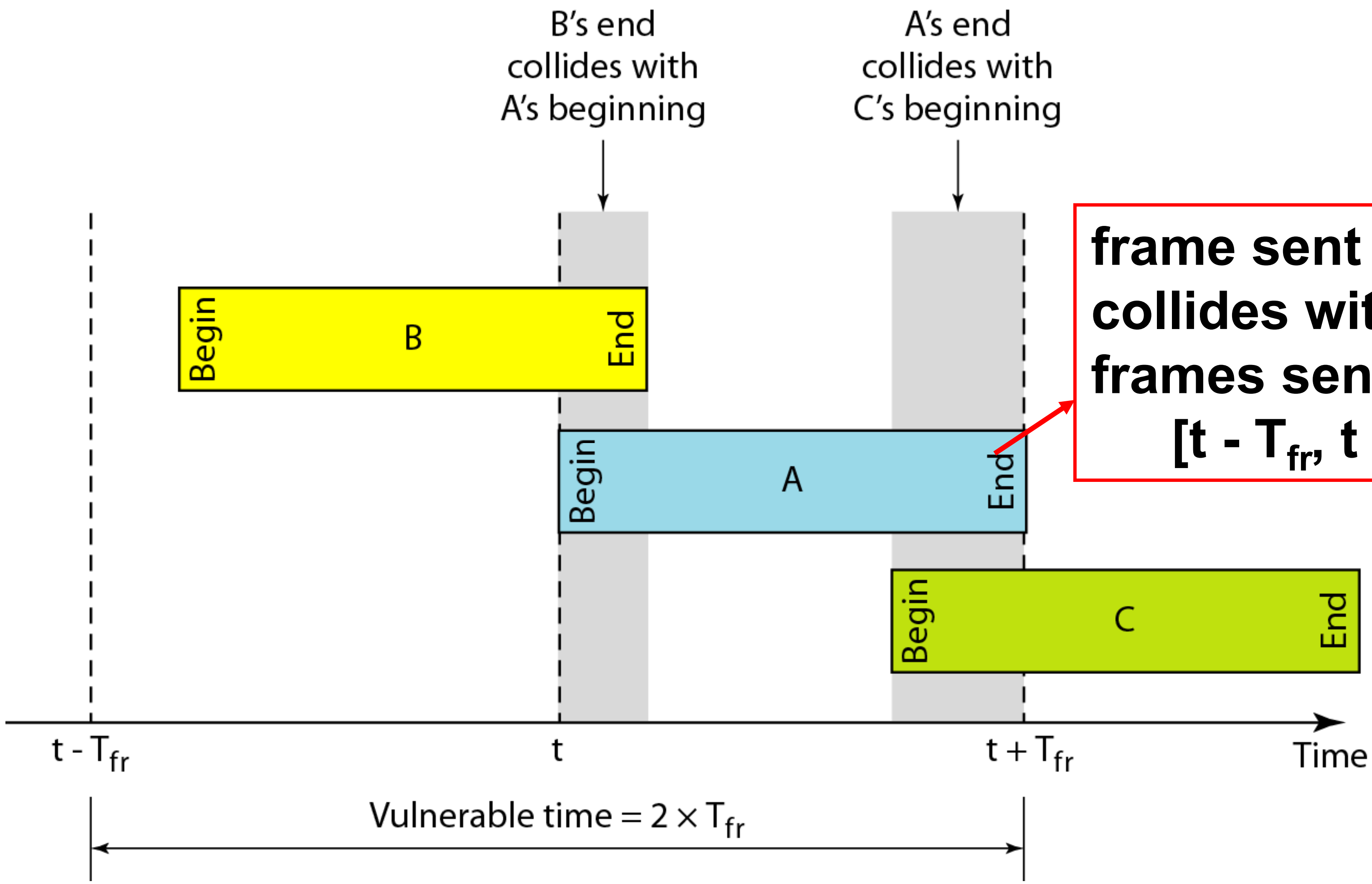
- 1, Let users transmit whenever they have data to be sent.
- 2, Expected collisions will occur.
- 3, The collided frames will be destroyed.
- 4, Using a feedback mechanism to know about the status of frame.
- 5, Retransmit the destroyed frame

Pure ALOHA



Frames in a pure ALOHA network

Pure ALOHA



frame sent at t collides with other frames sent in $[t - T_{fr}, t + T_{fr}]$.

Vulnerable time for pure ALOHA protocol

Pure ALOHA

Note

The throughput for pure ALOHA is
$S = G \times e^{-2G}$.
The maximum throughput
$S_{max} = 0.184$ when G= (1/2).

Pure ALOHA

Example

A pure ALOHA network transmits 200-bit frames on a shared channel of 200 kbps. What is the requirement to make this frame collision-free?

Solution

Average frame transmission time T_{fr} is 200 bits/200 kbps or 1 ms. The vulnerable time is 2×1 ms = 2 ms. This means no station should send later than 1 ms before this station starts transmission and no station should start sending during the one 1-ms period that this station is sending.

Pure ALOHA

Example

A pure ALOHA network transmits 200-bit frames on a shared channel of 200 kbps. What is the throughput if the system (all stations together) produces

a. 1000 frames per second

b. 500 frames per second

c. 250 frames per second.

Pure ALOHA

Solution

The frame transmission time is 200/200 kbps or 1 ms.

a.

If the system creates 1000 frames per second, this is 1 frame per millisecond. The load is 1. In this case $S = G \times e^{-2G}$ ***or*** $S = 0.135$ ***(13.5 percent). This means that the throughput is*** $1000 \times 0.135 = 135$ ***frames. Only 135 frames out of 1000 will probably survive.***

Pure ALOHA

Solution

b.

If the system creates 500 frames per second, this is (1/2) frame per millisecond. The load is (1/2). In this case $S = G \times e^{-2G}$ or $S = 0.184$ (18.4 percent). This means that the throughput is $500 \times 0.184 = 92$ and that only 92 frames out of 500 will probably survive.

Note that this is the maximum throughput case.

Pure ALOHA

Solution

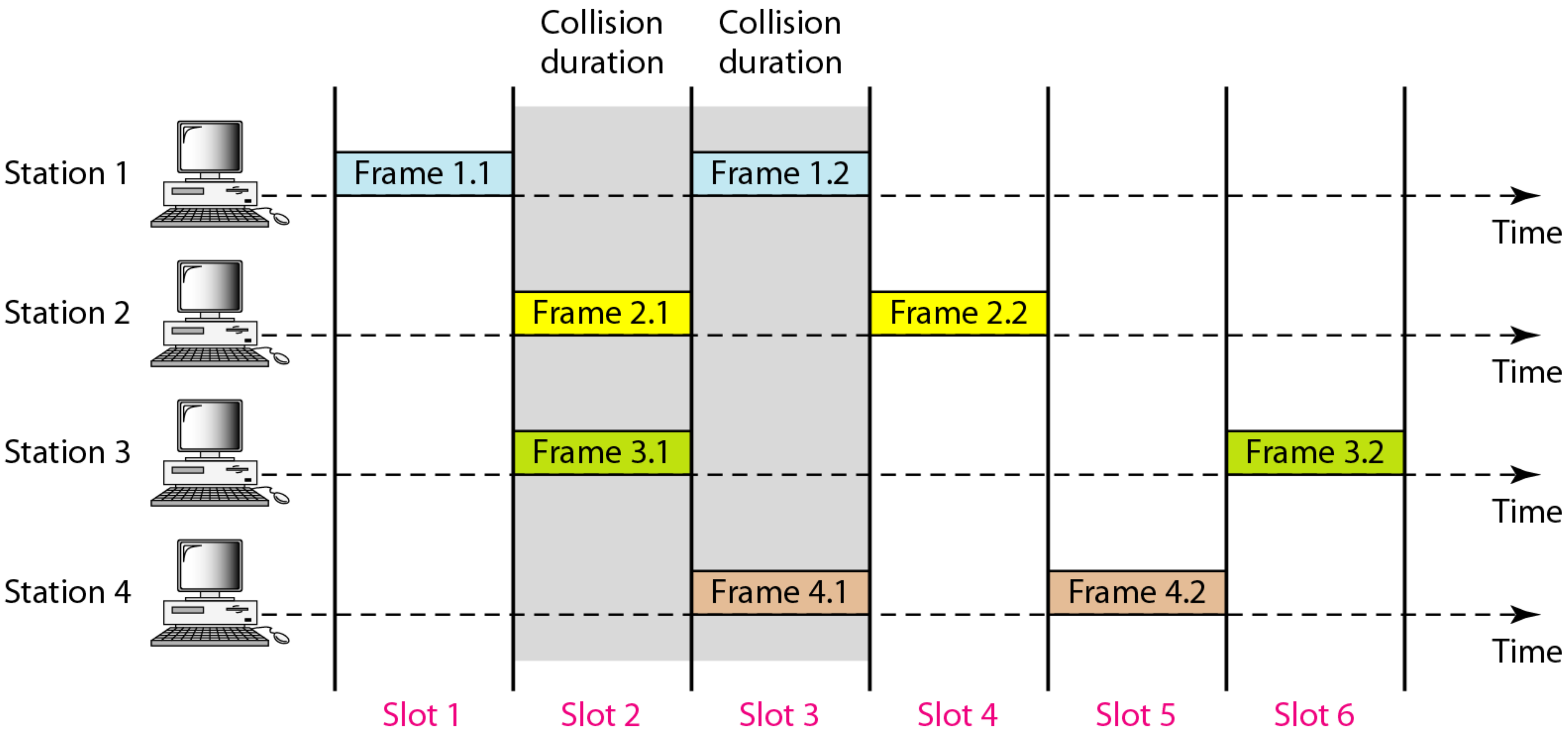
c.

If the system creates 250 frames per second, this is (1/4) frame per millisecond. The load is (1/4). In this case $S = G \times e^{-2G}$ or $S = 0.152$ (15.2 percent). This means that the throughput is $250 \times 0.152 = 38$. Only 38 frames out of 250 will probably survive.

Slotted ALOHA

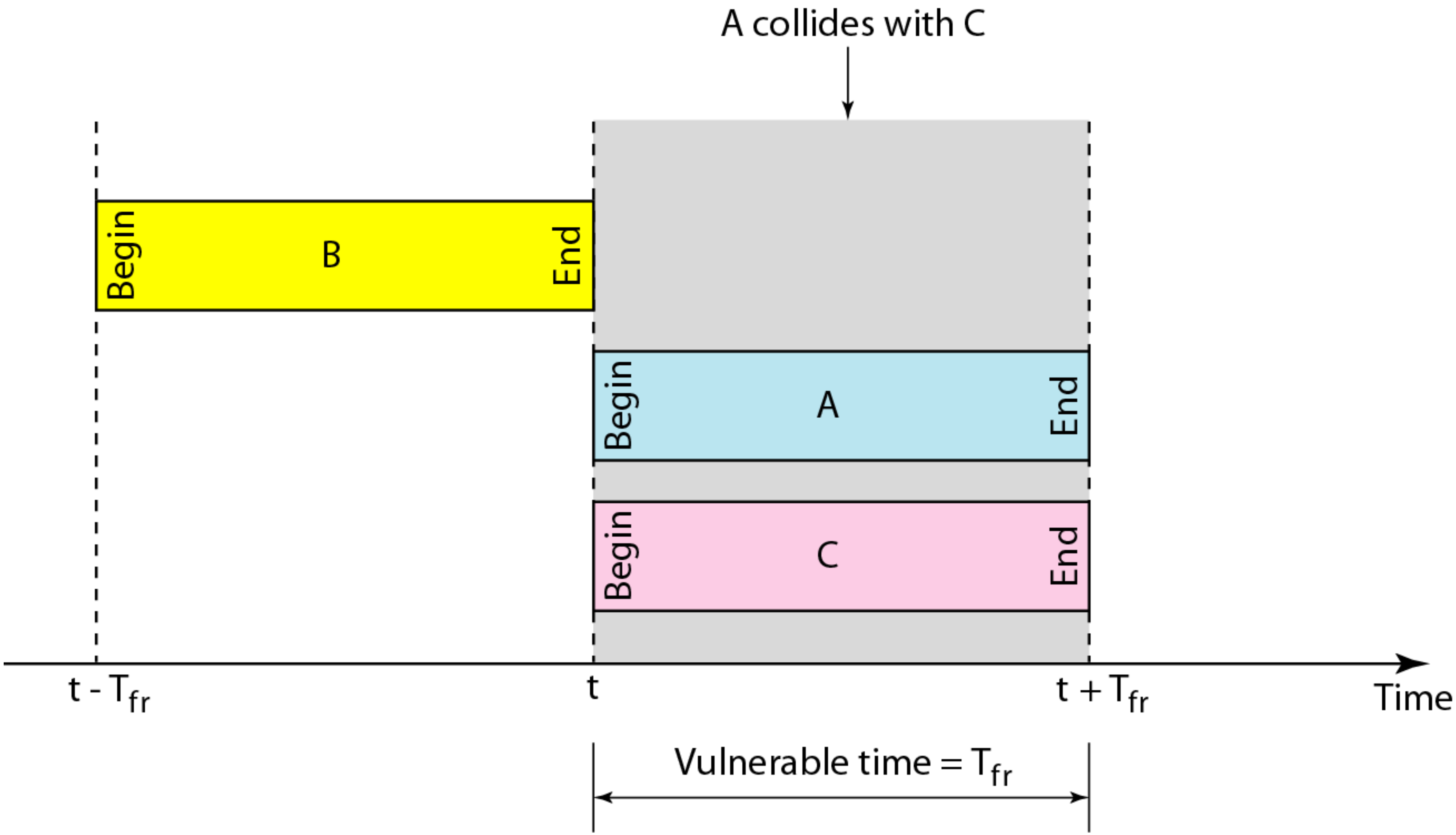
- Split time into pieces (slots), each slot equals to frame transmission time.
- A station can transmit at the beginning of a slot only.
- If a station misses the beginning of a slot, it has to wait until the beginning of the next time slot.
- A central clock or station informs all stations about the start of a each slot

Slotted ALOHA



Frames in a slotted ALOHA network

Slotted ALOHA



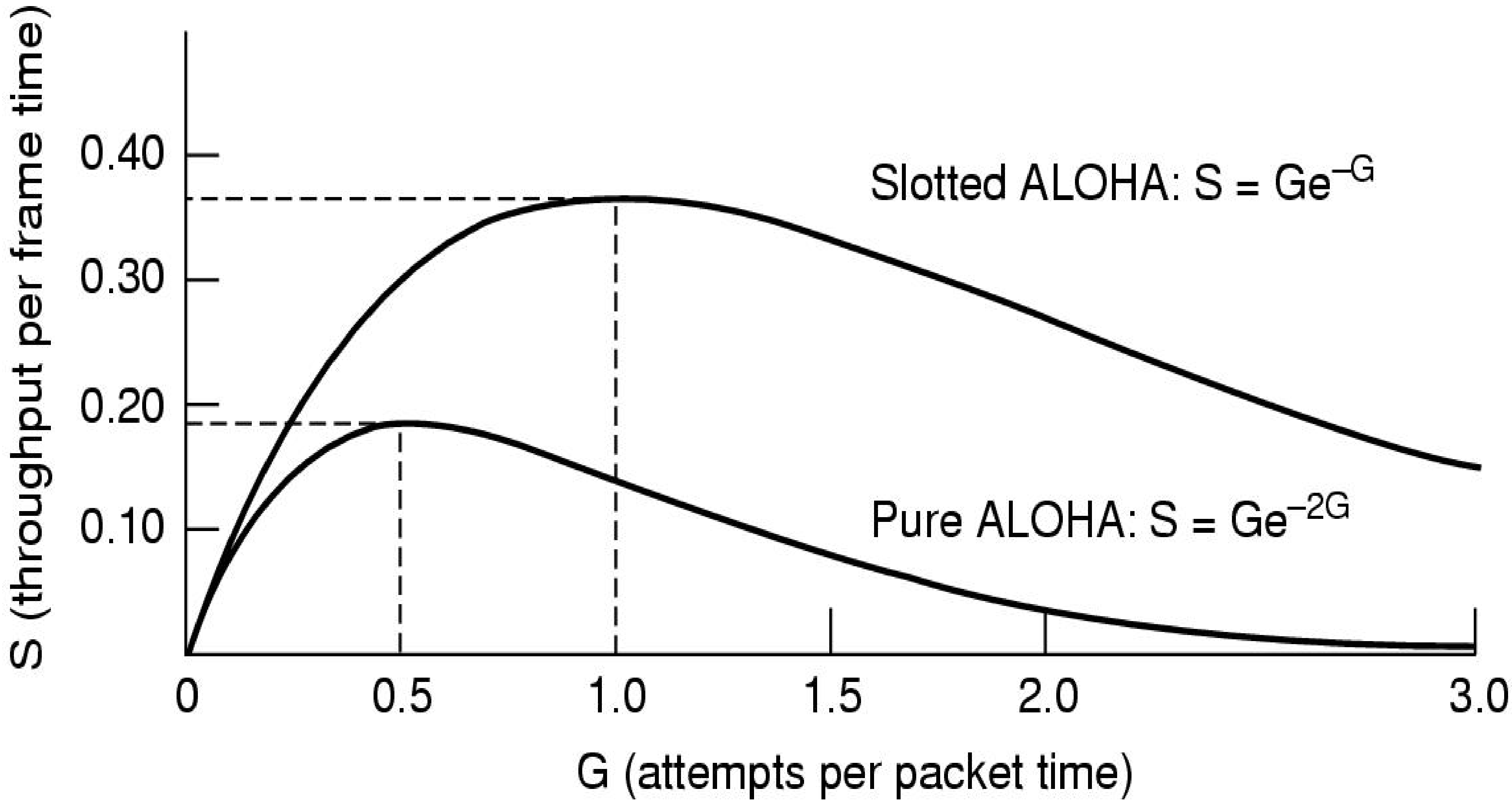
Vulnerable time for slotted ALOHA protocol

Slotted ALOHA

Note

The throughput for slotted ALOHA is
$S = G \times e^{-G}$.
The maximum throughput
$S_{max} = 0.368$ when G = 1.

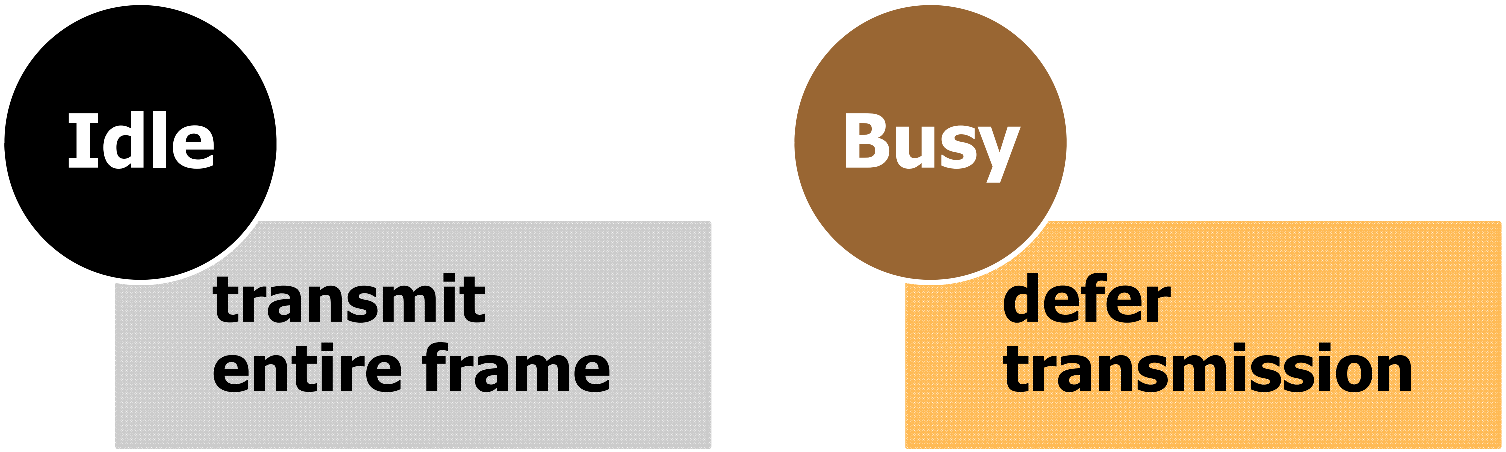
Efficiency of Aloha



Throughput versus offered traffic for ALOHA systems

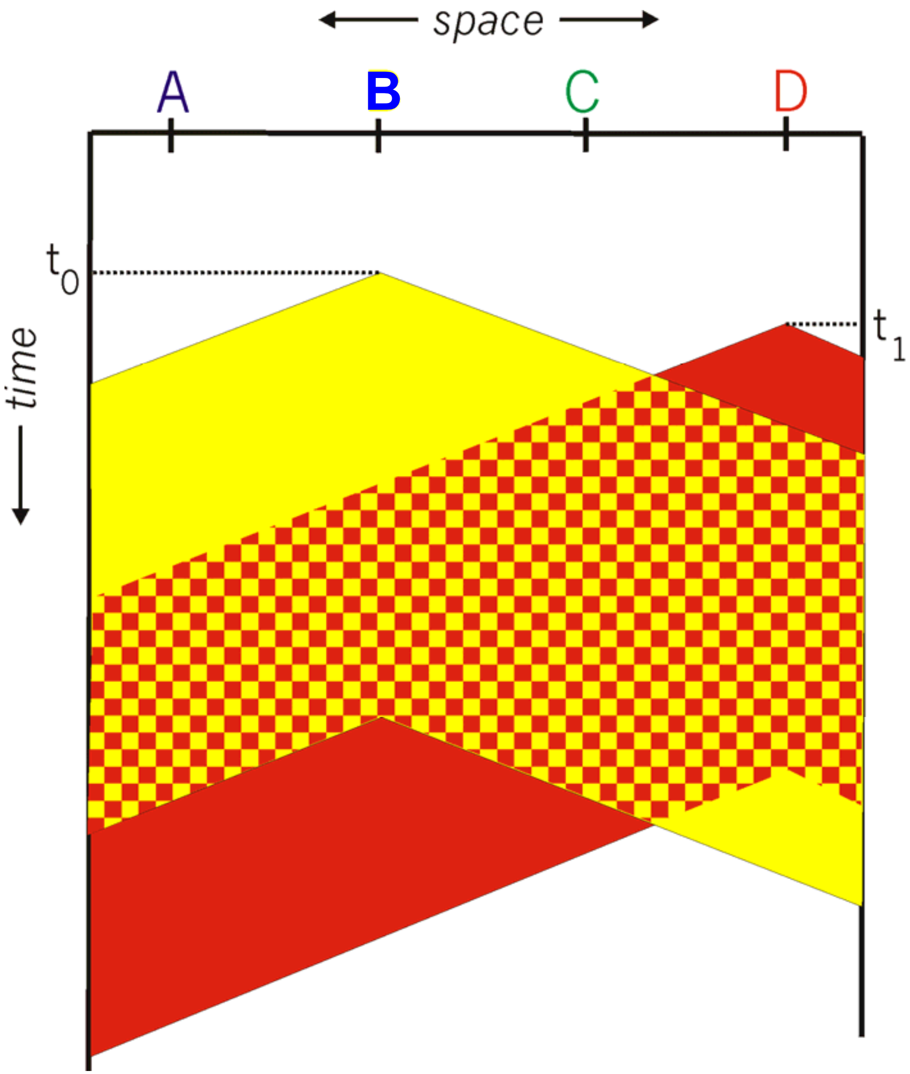
CSMA

- **Carrier Sense Multiple Access**
- **Monitor the channel before transmission.**



CSMA collisions

spatial layout of nodes



collisions can still occur:

propagation delay means two nodes may not hear each other's transmission

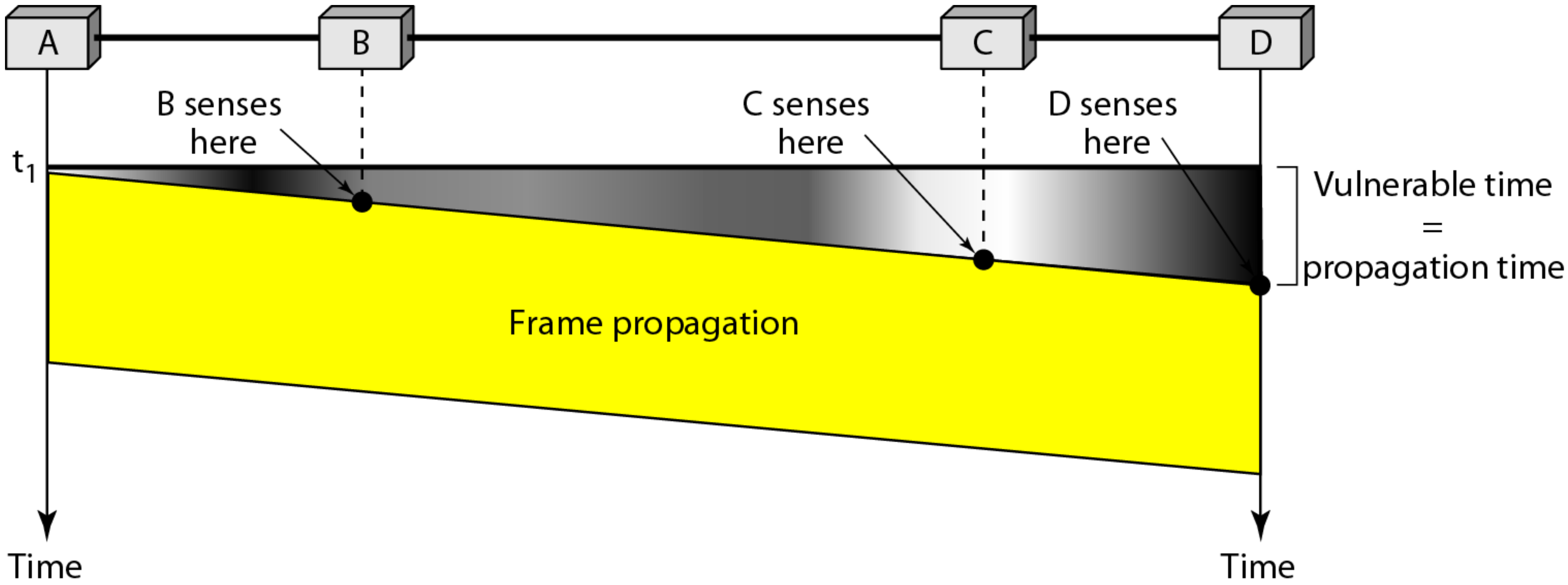
collision:

entire transmission time wasted

CSMA

- Vulnerable time for CSMA is **the maximum propagation time, t_{prop}**
- The **longer** the propagation delay, the **worse** the performance of the protocol because of the above case.

CSMA



Vulnerable time in CSMA

Types of CSMA

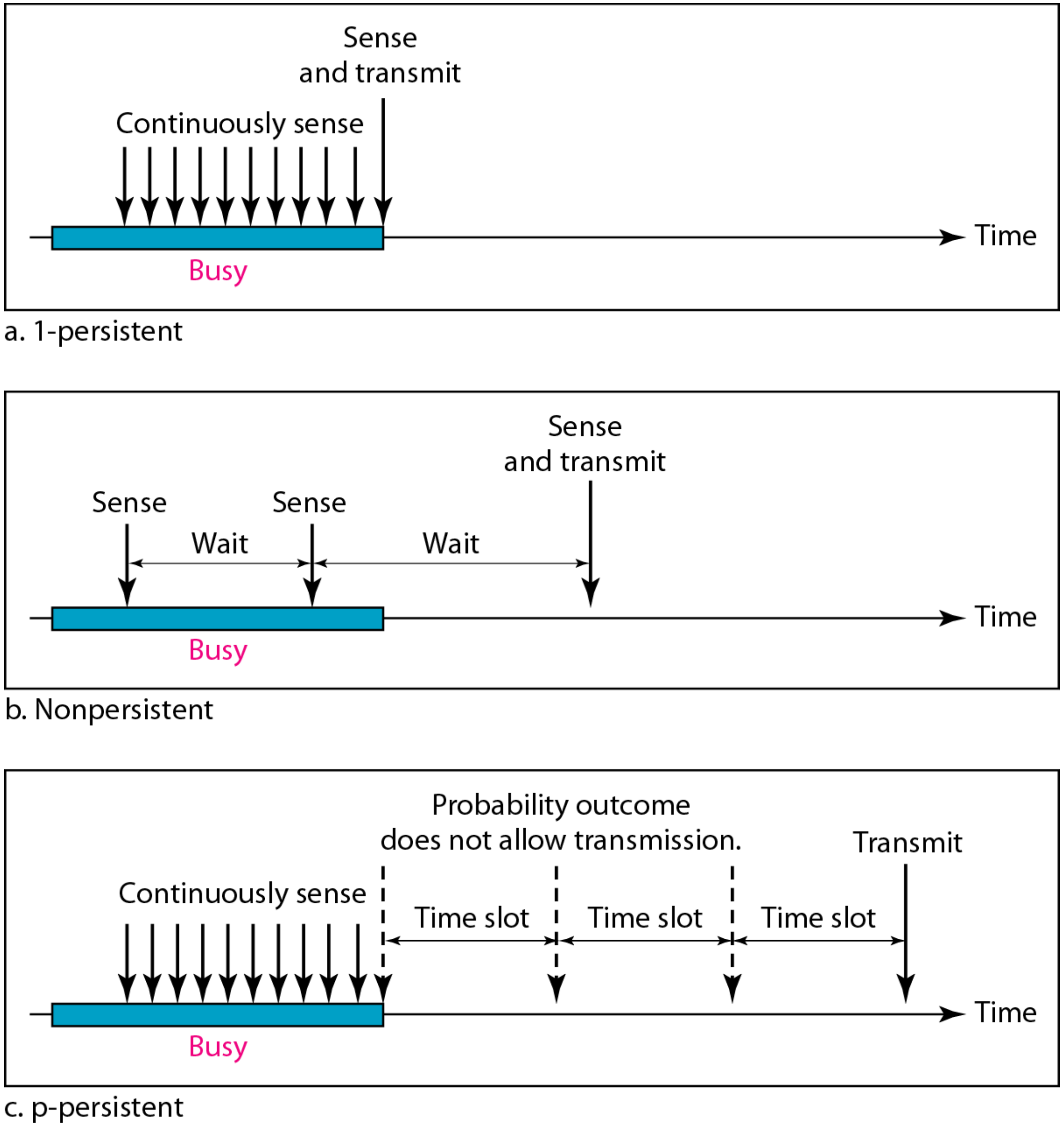
- **Different CSMA protocols that determine:**
  - **What a station should do when the medium is idle?**
  - **What a station should do when the medium is busy?**
- **Different techniques**
  - **Non-Persistent CSMA**
  - **1-Persistent CSMA**
  - **p-Persistent CSMA**

Types of CSMA

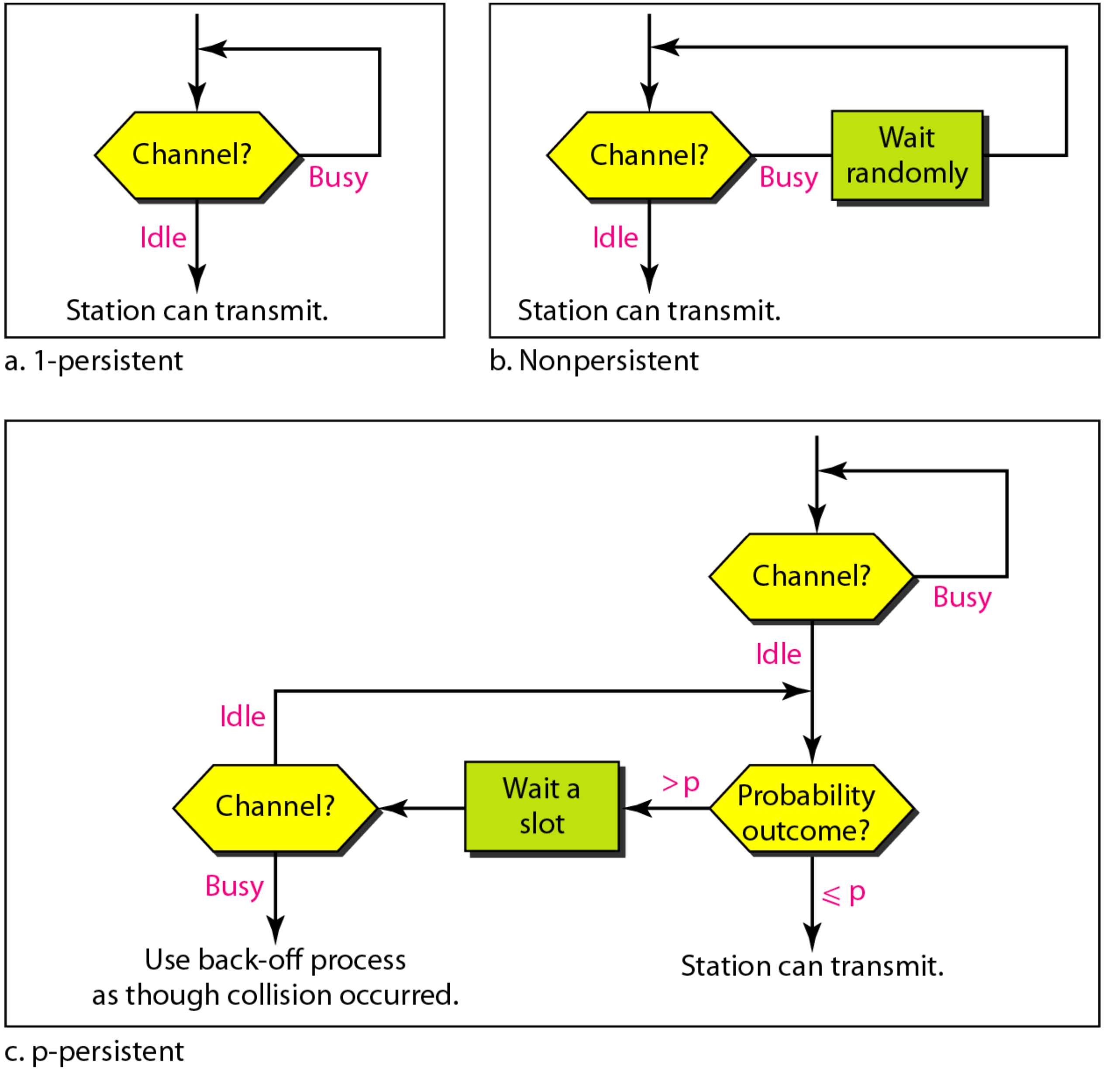
- **1-persistent:**
  - if busy, constantly sense channel
  - if idle, send immediately
  - if collision is detected, wait a random amount of time before retransmitting
- **Non-persistent:**
  - if busy, wait a random amount of time before sensing again;
  - if idle, transmit as soon as it is idle
  - collisions reduced because sensing is not immediately rescheduled
  - drawback: more delay

Types of CSMA

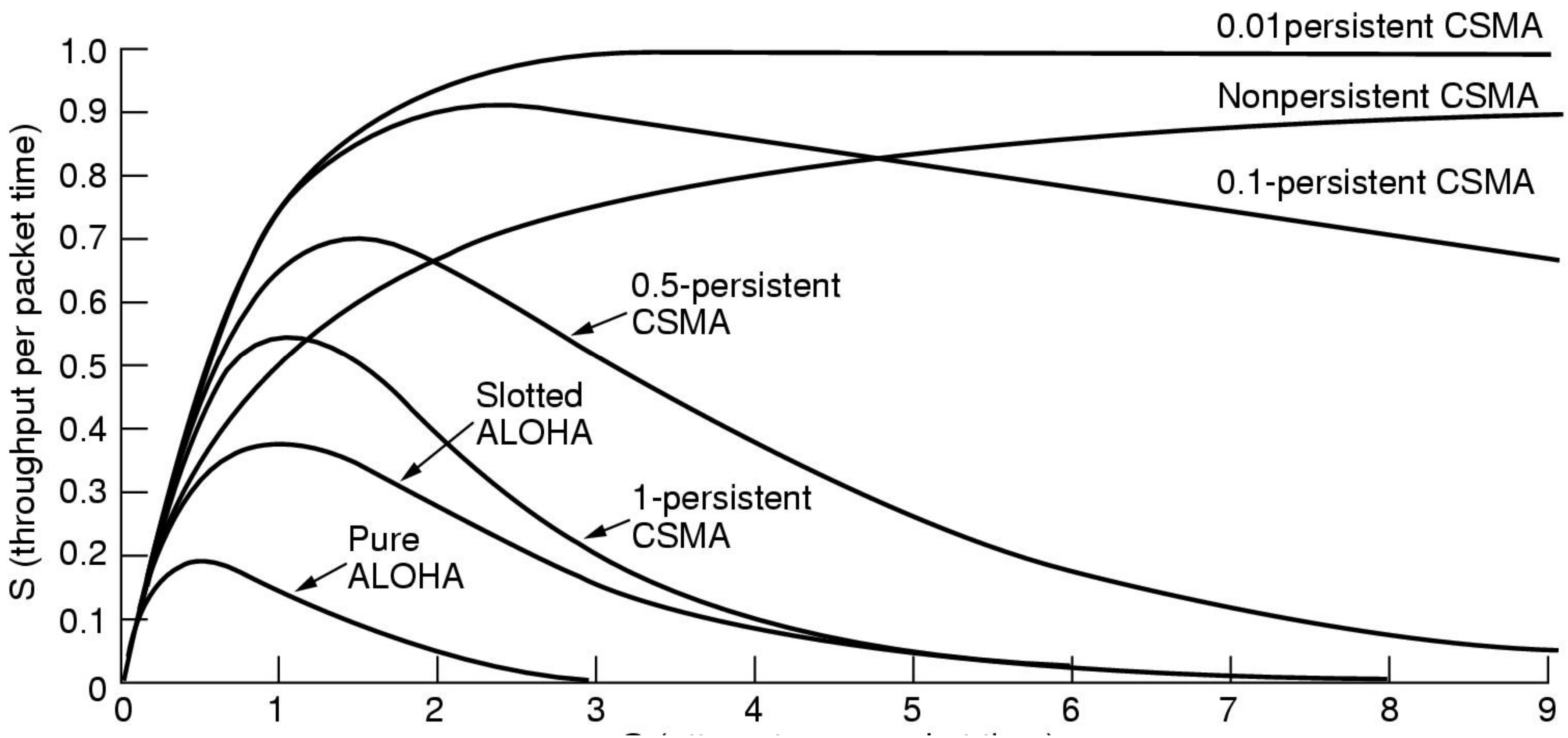
- **p-persistent:** combines 1-persistent goal of reduced idle channel time with the non-persistent goal of reduced collisions.
  - sense constantly if busy
  - if the channel is idle, transmit packet with probability *p*
  - with probability 1-*p* station waits an additional t_{prop} before sensing again
  - *p*=1 is not really good, *p*=0 makes you *really* polite



Behavior of three persistence methods



Flow diagram for three persistence methods



Comparison of the channel utilization versus load for various random

Question: What are we actually displaying here? Should the conclusion be that p-persistent protocols are really good with p is close to 0?

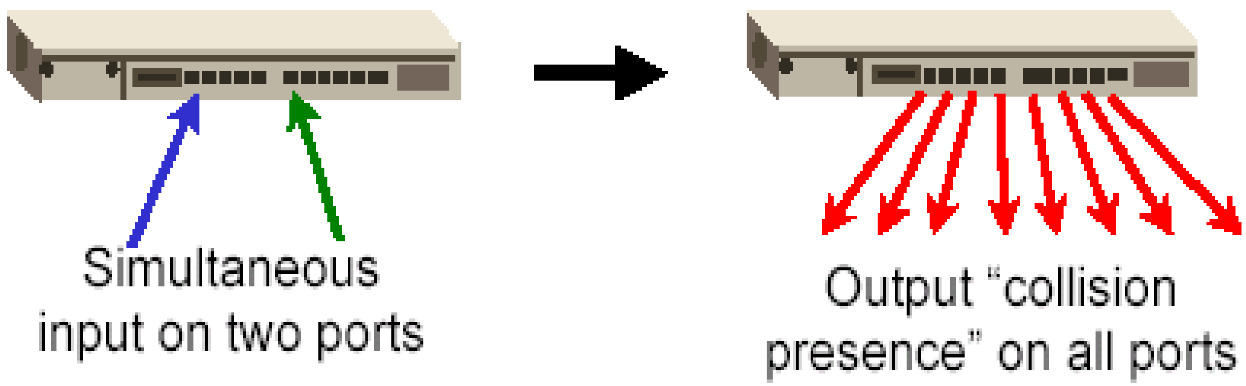
CSMA/CD (Collision Detection)

- CSMA has an inefficiency.
- CSMA/CD (Carrier Sense Multiple Access with Collision Detection) overcomes this as follows:
  - While transmitting, the sender is listening to medium for collisions. (Listening while talking)
  - Sender stops transmission if collision has occurred reducing channel wastage.
- CSMA/CD is Widely used for bus topology LANs (IEEE 802.3, Ethernet).

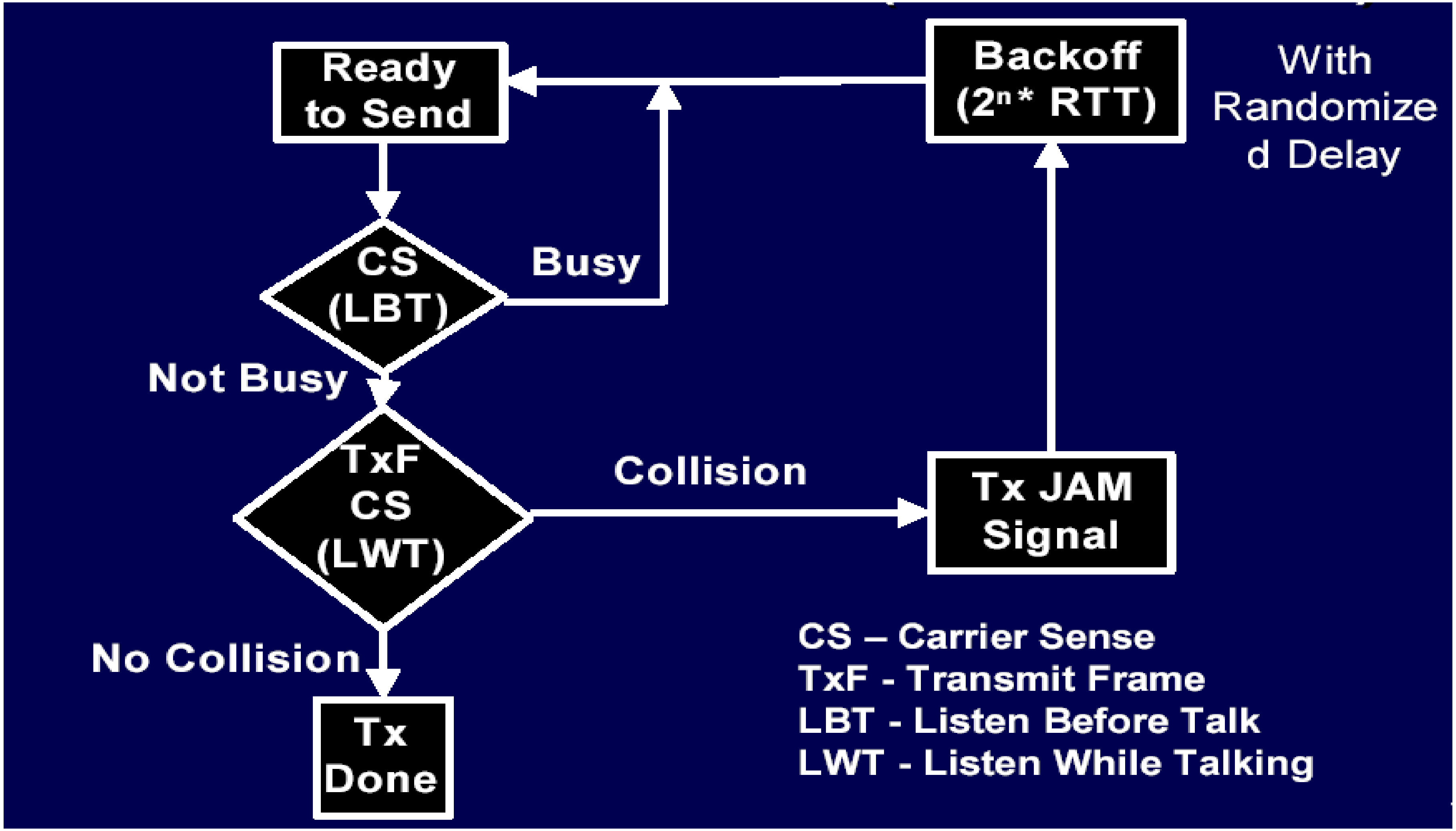
How to detect collision?

Transceiver: **A node monitors the media while transmitting. If the observed power is more than transmitted powe**of its own signal, it means collision occurred

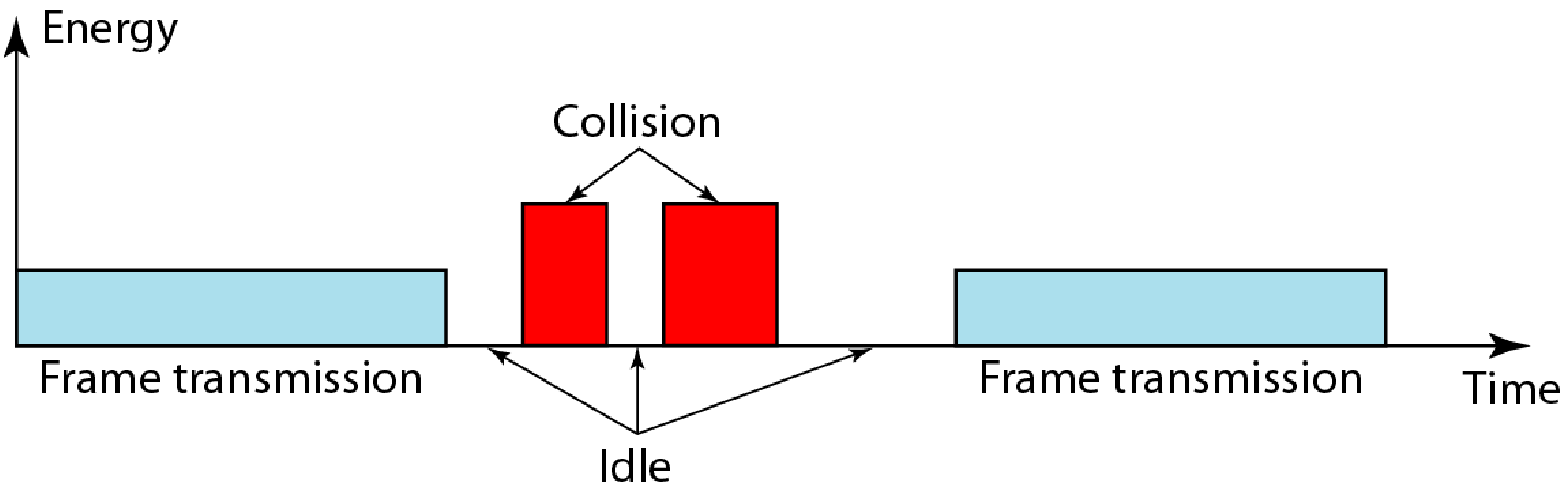
Hub: **if input occurs simultaneously on two ports, it indicates a collision. Hub sends a collision presence signal on all ports.**



CSMA/CD Protocol



CSMA/CD Protocol



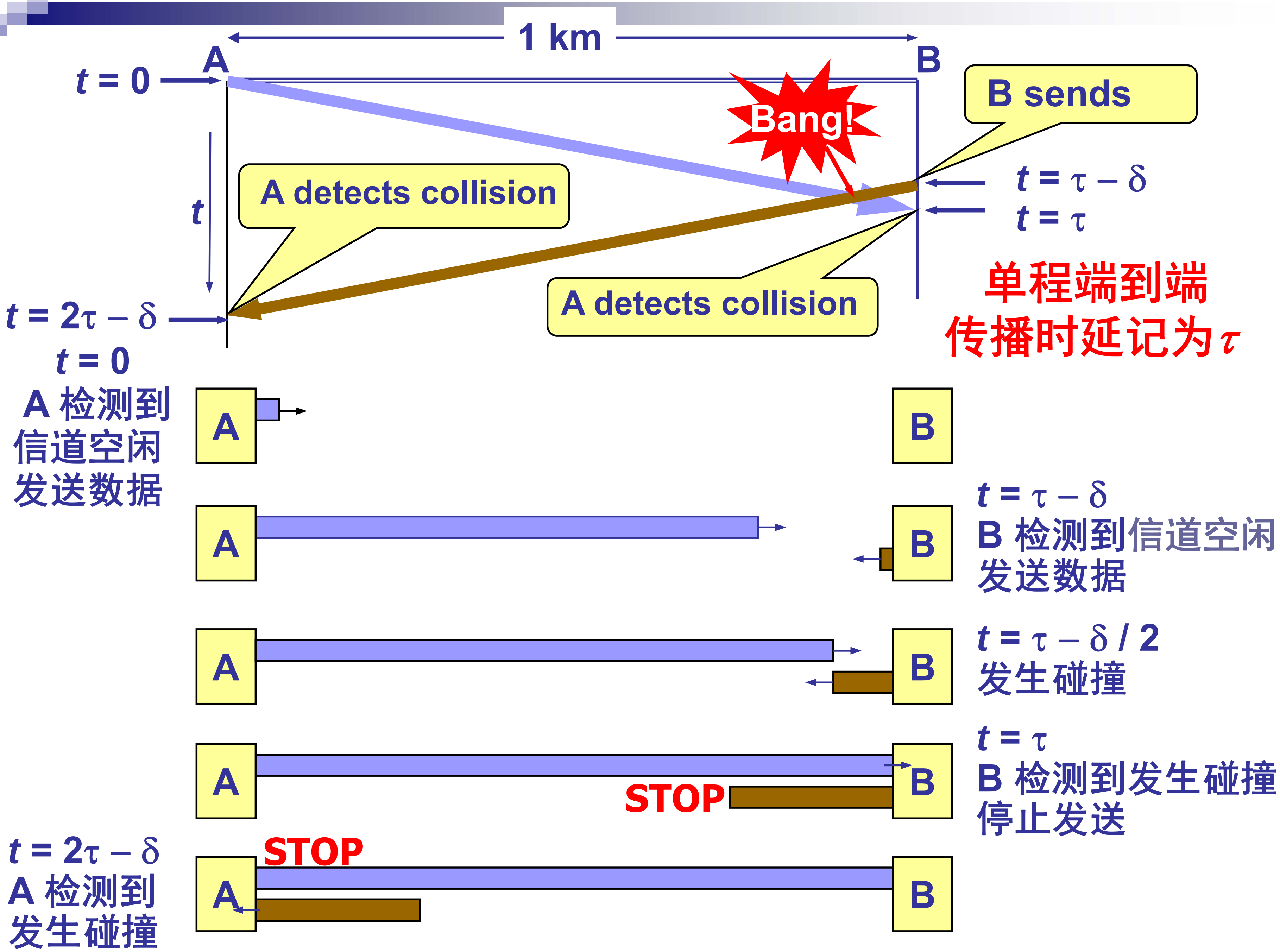
Energy level during transmission, idleness, or collision

Contention Period

- ***Question:*** **How long does it take to detect a collision?**
- ***Answer:*** ***In the worst case,*** **twice the maximum propagation delay of the medium.**

Contention slot must be 2 ح.

ح is maximum propagation delay on a channel.


1 km
A
B
t = 0
Bang!
B sends
A detects collision
t
t = τ – δ
t = τ
A detects collision
t = 2τ – δ
单程端到端
传播时延记为τ
t = 0
A 检测到
信道空闲
发送数据
A
B
t = τ – δ
B 检测到信道空闲
发送数据
t = τ – δ / 2
发生碰撞
t = τ
B 检测到发生碰撞
停止发送
STOP
t = 2τ – δ
A 检测到
发生碰撞
STOP

Contention Period

- **Restrictions of CSMA / CD:**

  Frame transmission time should be at least as long as the time needed to detect a collision (2 * maximum propagation delay + *jam sequence* transmission time).

  Otherwise, CSMA/CD does not have an advantage over CSMA.

TimeSlot >=

$2 * T_{prop}$ + jam sequence transmission time

Contention Period

- **Minimum frame length**

$$L_{\min} = R \cdot a = 2R\left(S/0.7C + T_{phy}\right)$$

where :

a - Contention period

R – Data rate

T_{phy} – Physical layer delay

Note: Process delay must be taken account into whole contention slot.

Example

A network using CSMA/CD has a bandwidth of 10 Mbps. If the maximum propagation time (including the delays in the devices and ignoring the time needed to send a jamming signal) is 25.6 μs, what is the minimum size of the frame?

Example

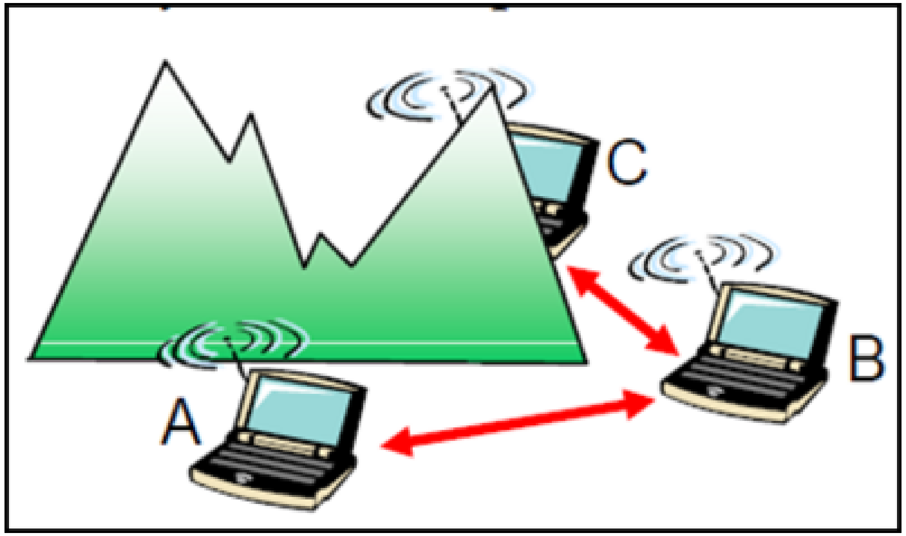
Solution

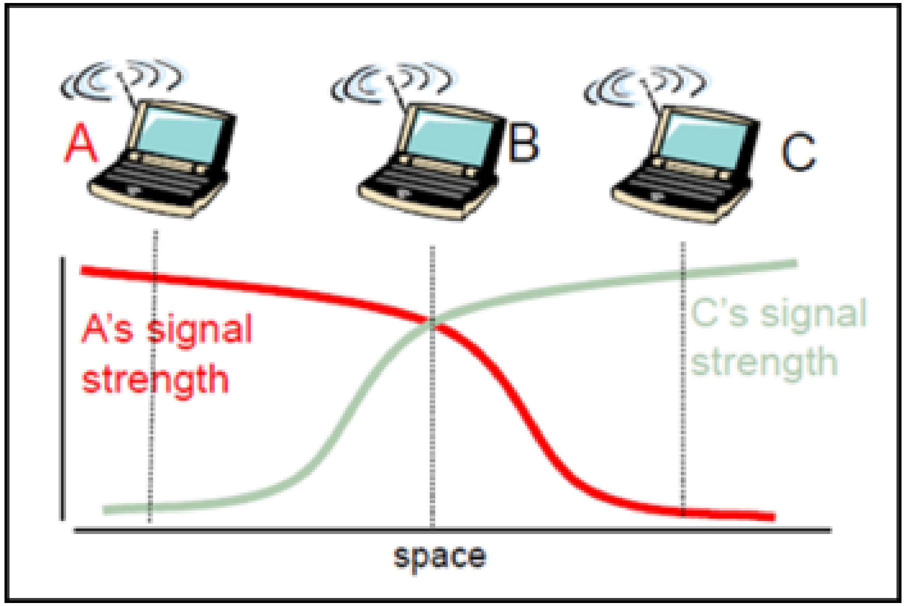
- The frame transmission time $T_{fr} = 2 \times T_p = 51.2\ \mu s$. This means, in the worst case, a station needs to transmit for a period of 51.2 μs to detect the collision.
- The minimum size of the frame is 10 Mbps × 51.2 μs = 512 bits or 64 bytes.

CSMA/CA

- Used often in wireless networking.
- CA: collision avoidance
  - Collision avoidance BEFORE transmission
- CSMA/CD does not work in wireless LAN.
- Three reasons:
  - Station must be able to send and receive data at the same time.
  - Collision may not be detected because of the hidden terminal problem.
  - Distance between stations in wireless LANs can be great. Signal fading could prevent a station at one end from hearing a collision at other end.

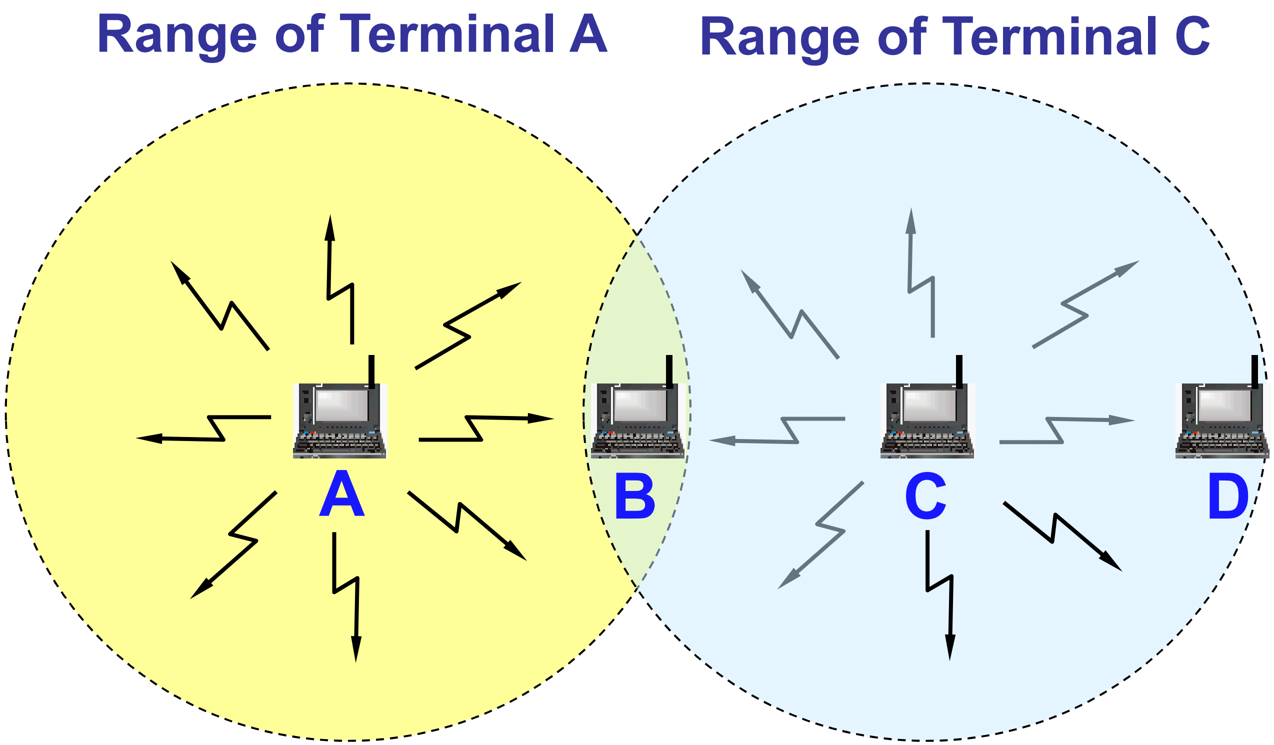
Hidden terminal problem





The hidden terminal problem occurs when station A is visible to station B while not visible to station C, and station C is visible to station B

Hidden terminal problem

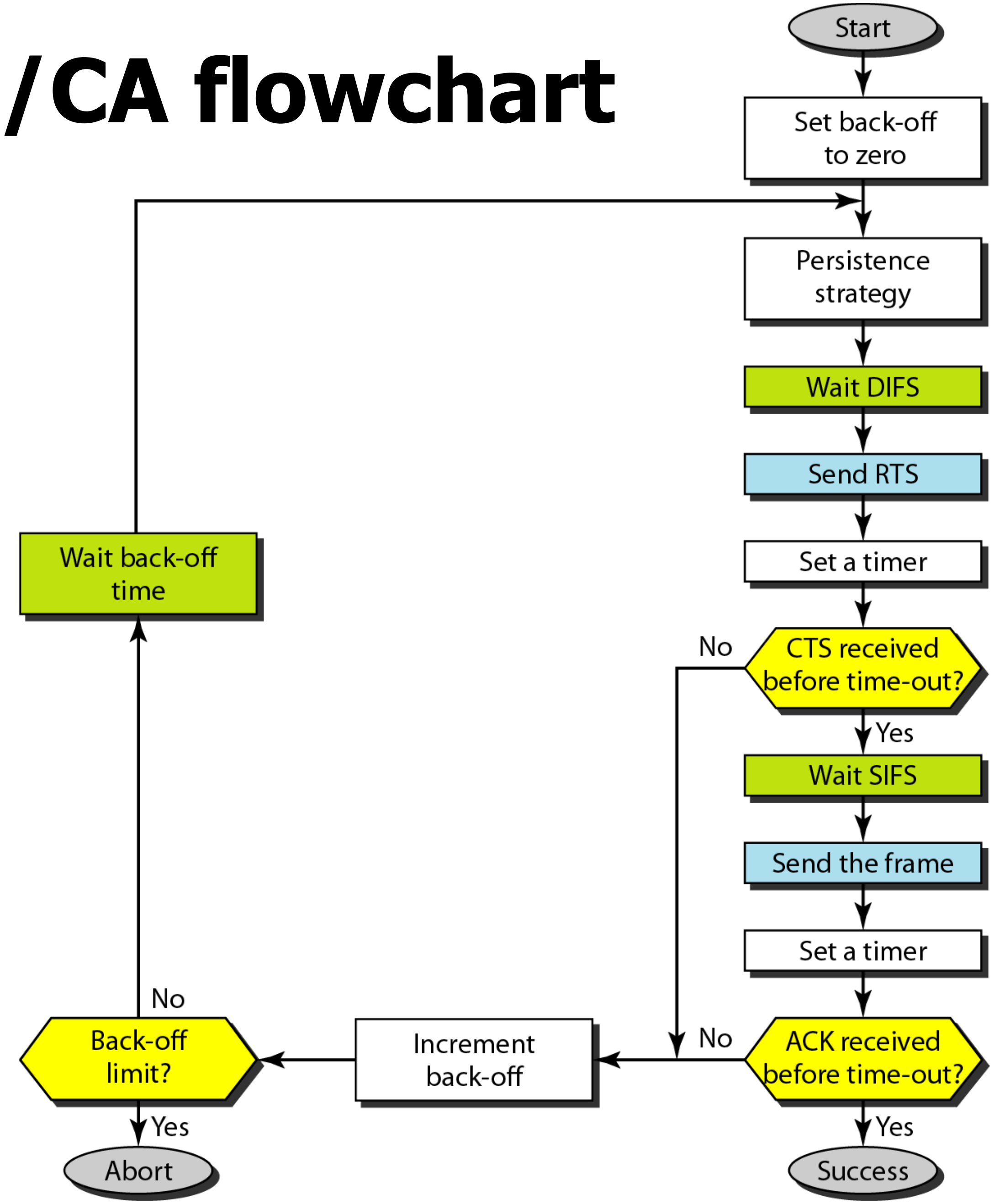


collision occurs when station A starts a transmission to station B while simultaneously, station C (it doesn't hear station A) starts another transmission to station B.

CSMA/CA

- Before sending a frame, source senses the medium by checking the energy level at the carrier frequency.
  - Backoff until the channel is idle.
  - After the channel is found idle, the station waits for a period of time called the Distributed interframe space (DIFS); then the station sends a control frame called request to send (RTS).
- After receiving RTS, the destination waits for a period called Short interframe space (SIFS), the destination station sends a control frame, called Clear to Send (CTS) to source. This control frame indicates that the destination station is ready to receive data.
- Source sends data after waiting for SITS
- Destination sends ACK after waiting for SITS.

CSMA/CA flowchart



Performance of Random Access Protocols

- **Simple and easy to implement.**
- **Decentralized.**
- **In low-traffic, frame transfer has low-delay**
- **However, limited throughput and in heavier traffic, frame delay has no limit.**
- **In some cases, a station may never have a chance to transfer its frame (unfair protocol).**

Performance of Random Access Protocols

- A node that has frames to be transmitted can transmit continuously at the full rate of channel (R) if it is the only node with frames.
- If (M) nodes want to transmit, many collisions can occur and the rate for each node will not be on average R/M.

Others

- Collision-Free Protocols
- Limited-Contention Protocols
- The Adaptive Tree Walk Protocol
- Wavelength Division Multiple Access Protocols
- ……
- Study by yourself !!!

Controlled Access or Scheduling

- Provides in order access to shared medium so that every station has chance to transfer (fair protocol)
- Eliminates collision completely
- Three methods for controlled access:
  - Reservation
  - Polling
  - Token Passing
- Study by yourself !!!

Chapter 4: Roadmap

- Medium Access Control
- Local Area Networks (LANs) and IEEE 802
- Ethernet
- Wireless LAN
- LAN Interconnection
- LAN Switching
- VLAN

Local Area Networks (LANs)

- Privately-owned
- Small area
- High speed
- High reliability
- Easy management

LAN Applications (1)

- **Personal computer LANs**
  - Low cost
  - Limited data rate
- **Back end networks**
  - Interconnecting large systems
    - High data rate
    - High speed interface
    - Distributed access
    - Limited distance
    - Limited number of devices

LAN Applications (2)

- **Storage Area Networks**
  - Separate network handling storage needs
  - Detaches storage tasks from specific servers
  - Shared storage facility across high-speed network
  - Improved client-server storage access
  - Direct storage to storage communication for backup

LAN Applications (2)

- **High speed office networks**
  - Desktop image processing
  - High capacity local storage
- **Backbone LANs**
  - Interconnect low speed local LANs
  - Reliability
  - Capacity
  - Cost

Various Local Area Networks

- Ethernet (Fast Ethernet, Gigabit Ethernet, 10G Ethernet)
- Wireless LAN
- Token ring
- FDDI (Fiber Distributed Data Interface)
- ATM LAN
- ……

LAN

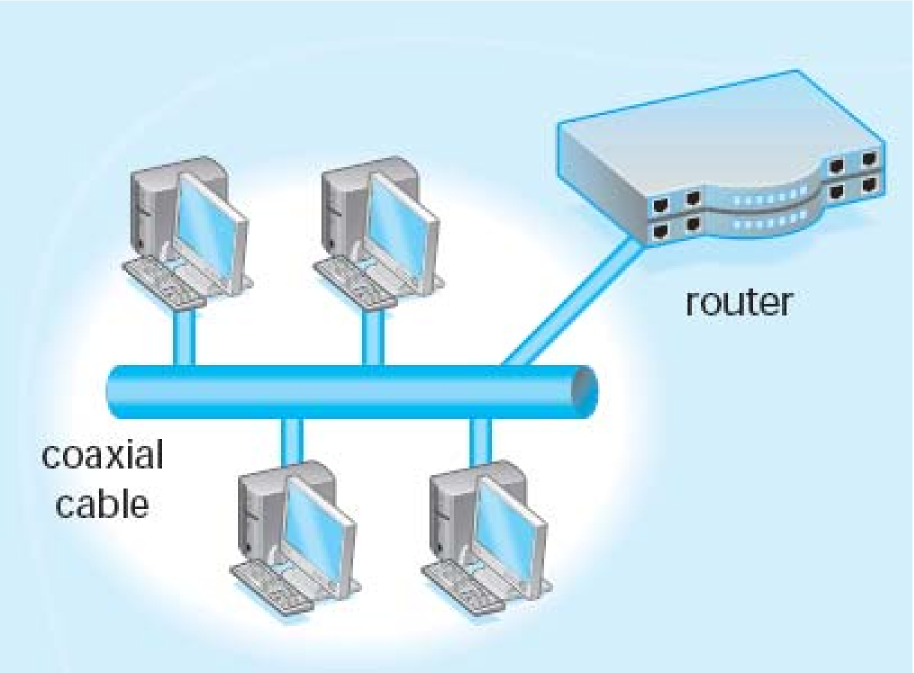
- LAN characteristics are determined by:
  - Topologies
  - MAC (Medium Access Control)
  - Transmission media
  - Size of coverage

LAN Topologies

- **Physical topology** is the actual location and arrangement of physical connections between devices on the network.
- **Logical topology** is the path that a given datagram travels between two devices. Often there is more than one way to get from one host to another.

Bus Topology

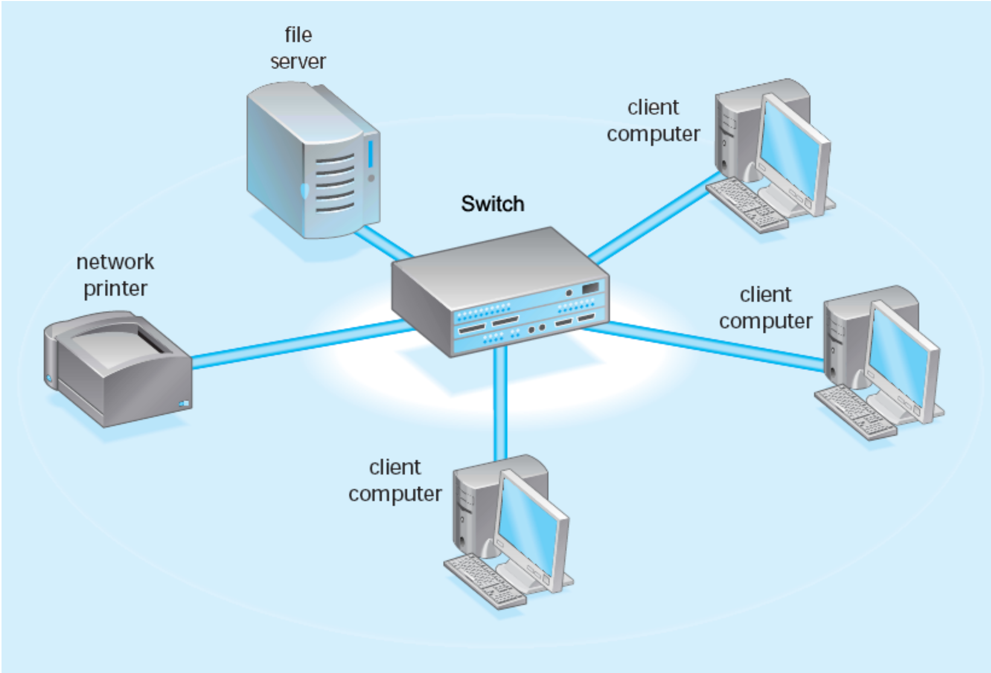
- All network devices connected to a common cable in logical linear fashion.
- Transmissions are sent along the length of the bus segment.



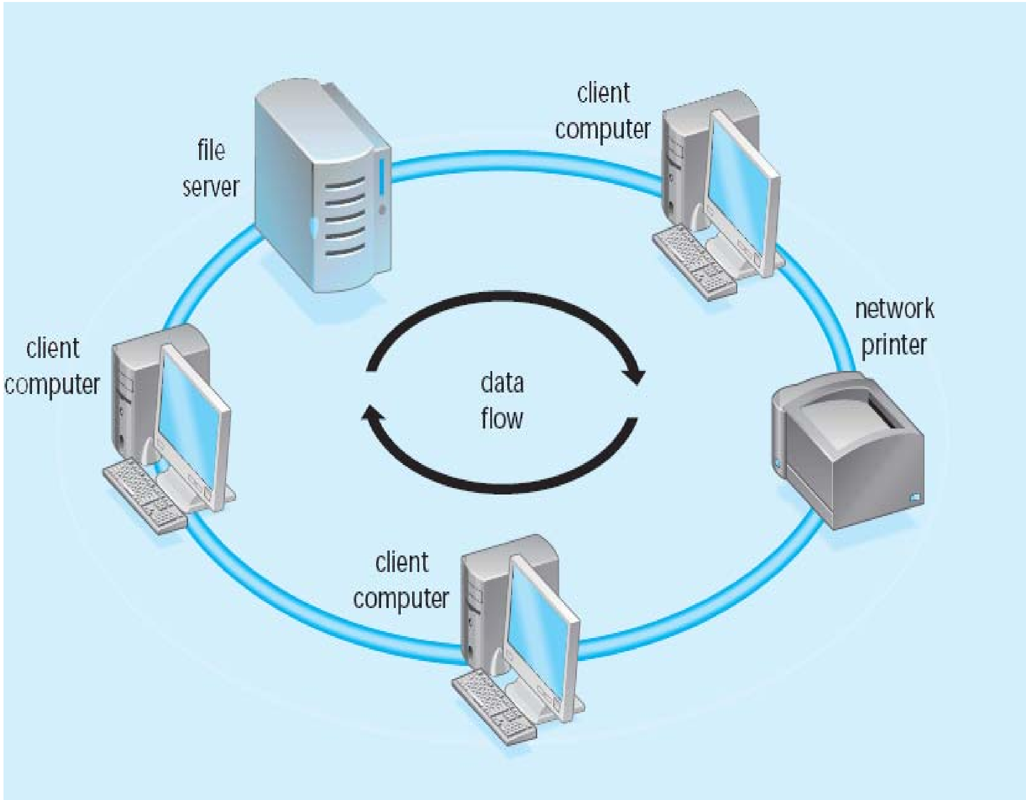
- Adding hosts to the network requires breaking the network.
- Failure of one host can cause failure of network.

Star Topology

- Connection from each device to a central location, usually a switch.
- Most commonly used physical topology.
- Failure of one cable does not bring down network.

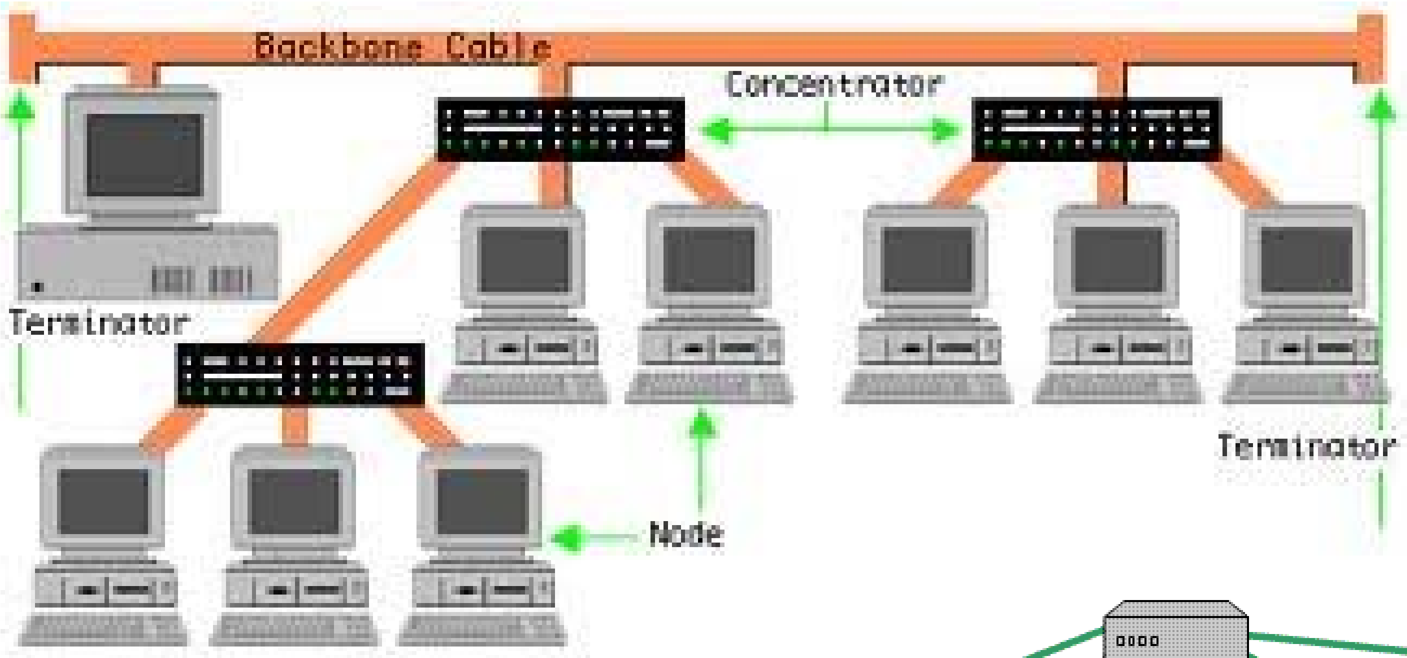


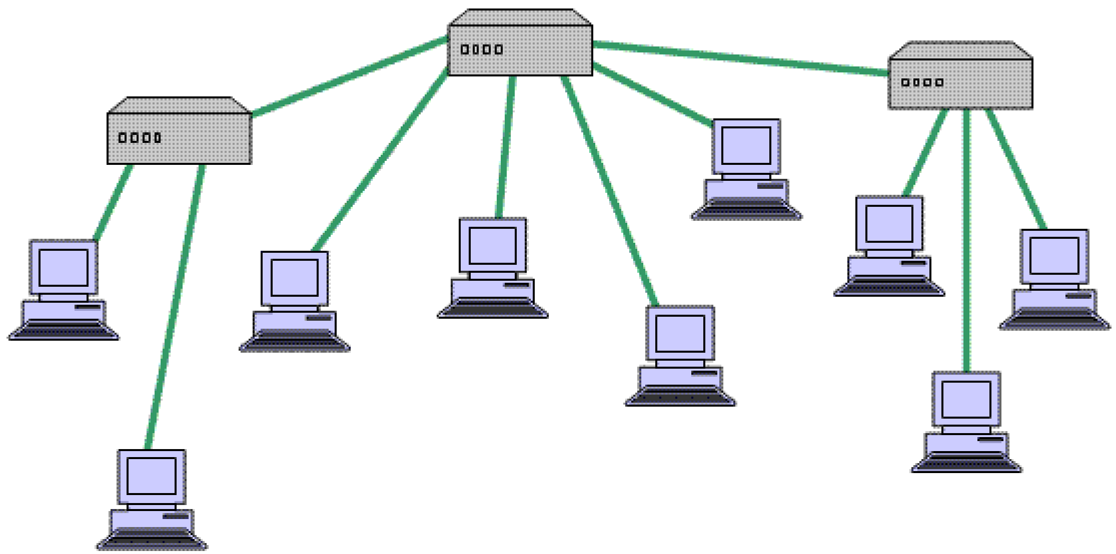
Ring Topology



- Network is connected in an endless loop.
- No termination required.
- Uncommon topology today, more common in 1980s.

Tree Topology

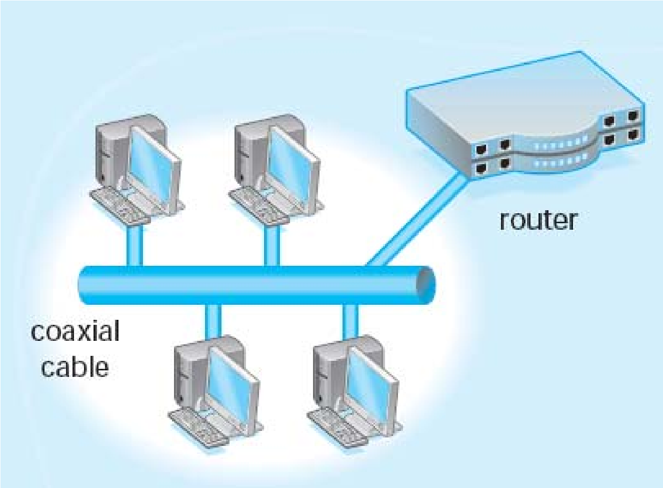


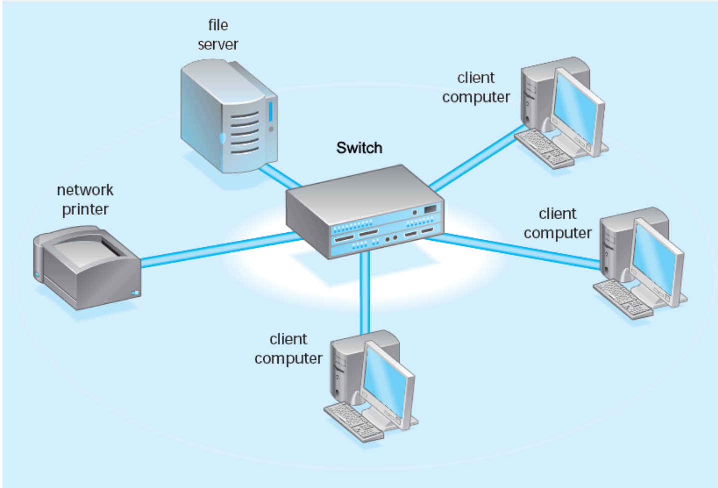


Choice of Topology

- Reliability
- Expandability
- Performance
- Needs considering in context of:
  - Medium
  - Wiring layout
  - Access control

MAC (Medium Access Control)





In a broadcast LAN, transmitted information will be received by all stations simultaneously. The medium access schemes are random access such as CSMA/CD and controlled access such as token-passing.

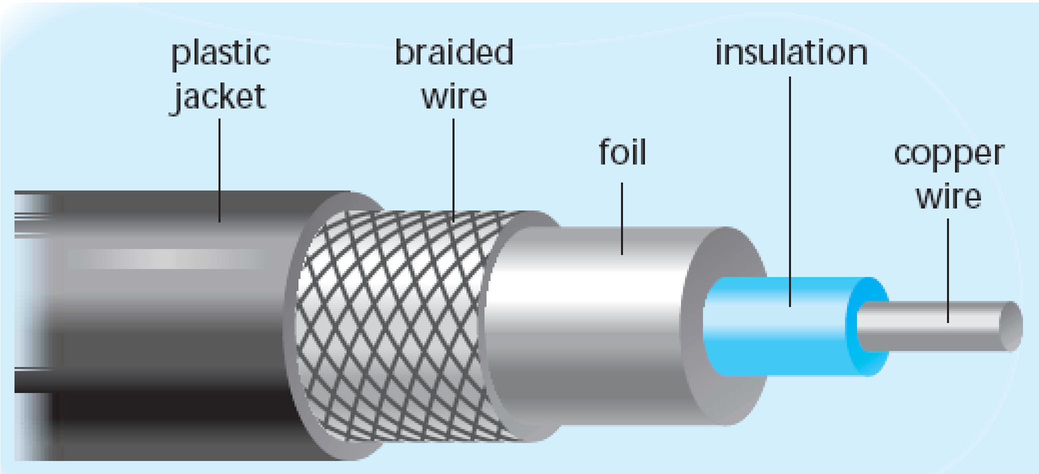
In a switched architecture, a switch forward data packets to their destinations that may be a single user station or another LAN segment.

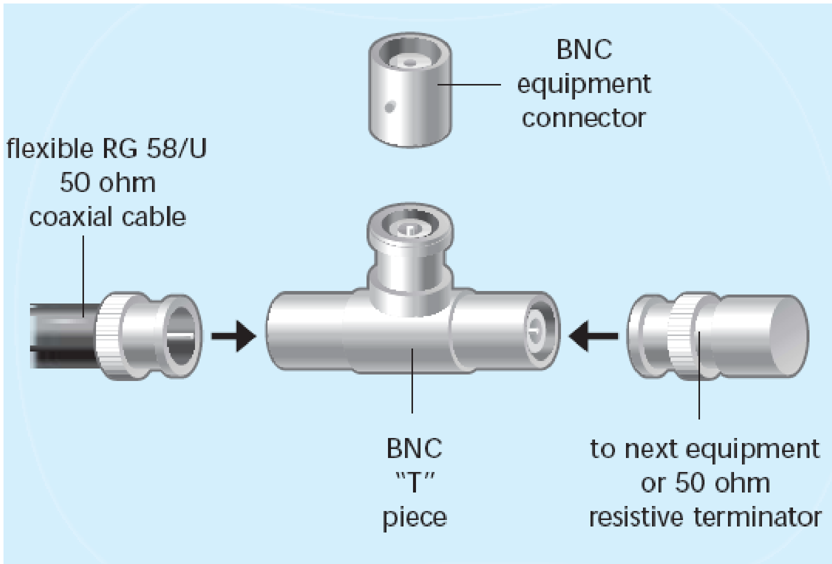
Transmission media

- Physical cabling is also known as **bounded media**
  - Transmissions are bound to the physical media.
  - To communicate, hosts *must* be physically connected to that media
- Wireless network is known as **unbounded media**
  - Transmissions are not bound to a physical cable.
  - To communicate, hosts *do not need* to be physically connected

Coaxial Cable

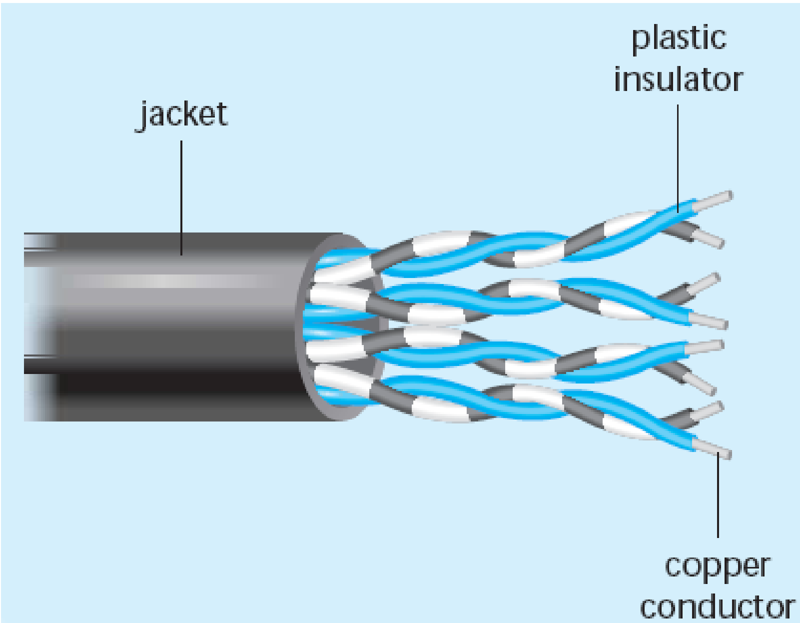
- Coaxial cable is often used in older LANs.
- Known as **RG58**, **Thinnet**, and **10Base2**.
- Maximum bandwidth of 10 Mbps.
- Maximum segment length of 185 meters.
- Maximum of 30 hosts per segment.

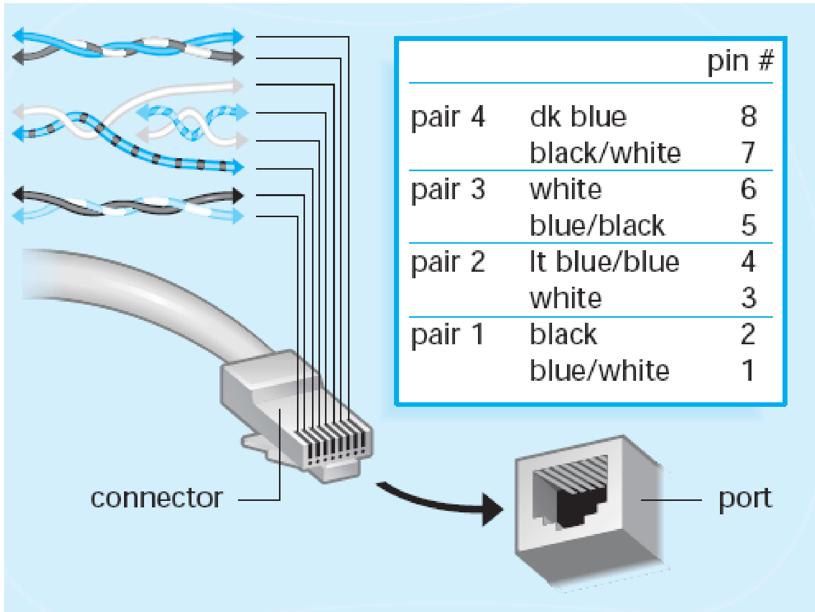




Twisted Pair Cable

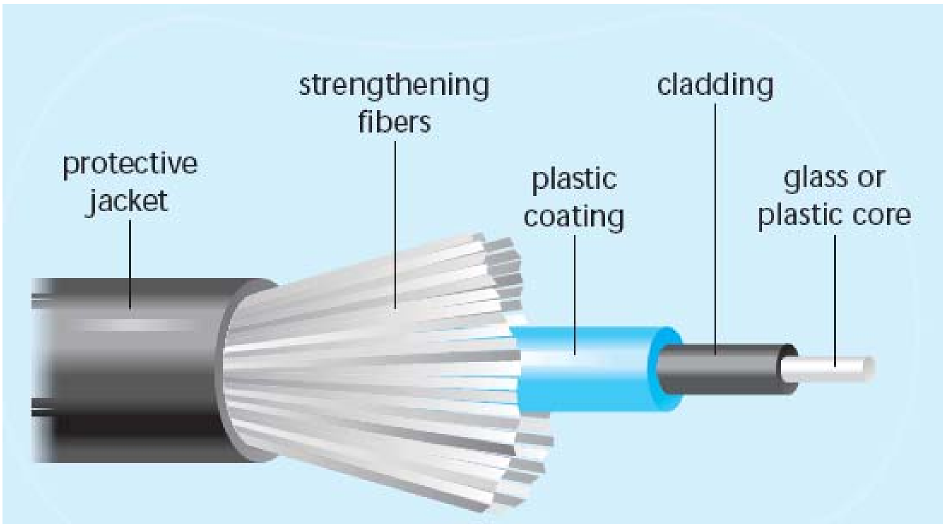
- The most common cabling technology in use today.
- Twists are used because they reduce interference.
- Maximum length: 100 meters.
- Maximum bandwidth: 1000 Mbps.





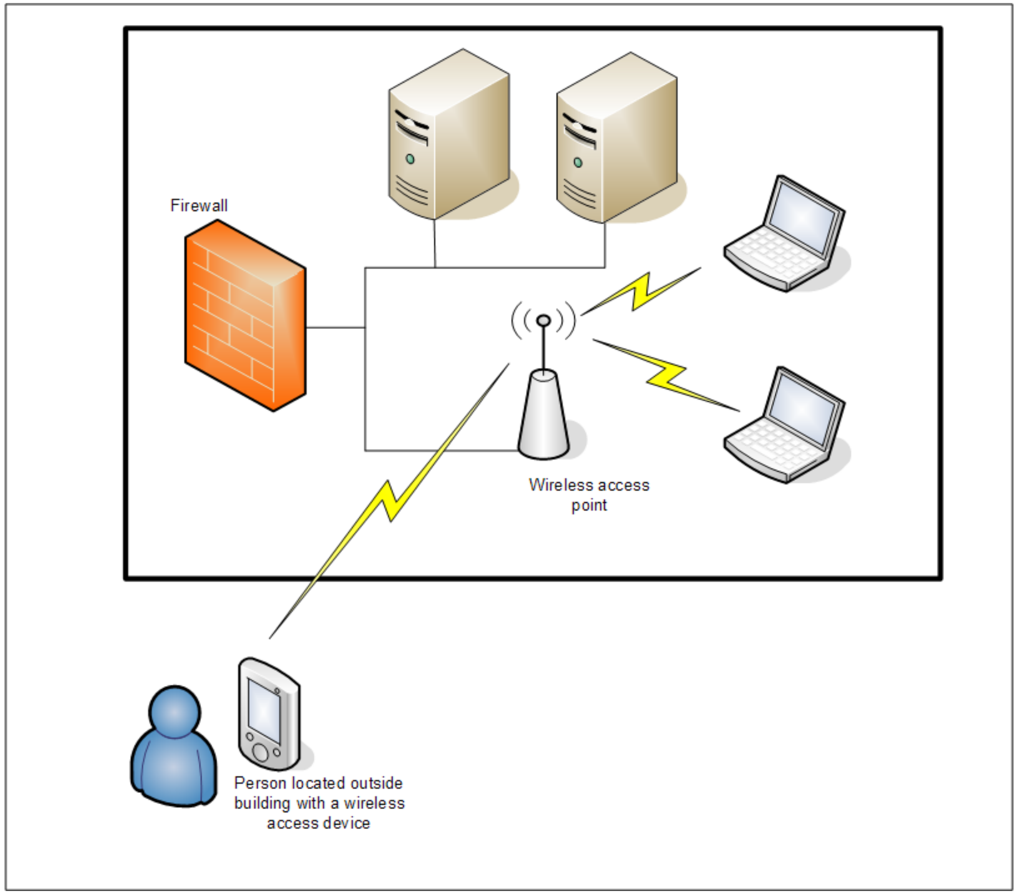
Fiber Optic Cable

- **Fiber optic cable has better data security than twisted pair or RG58. You can't intercept the signals without breaking the cable.**
- **Fiber optic cable is immune to electromagnetic interference.**
- **Fiber optic cable is mostly use as a backbone to connect LANs together, rather than connecting hosts together on a LAN.**

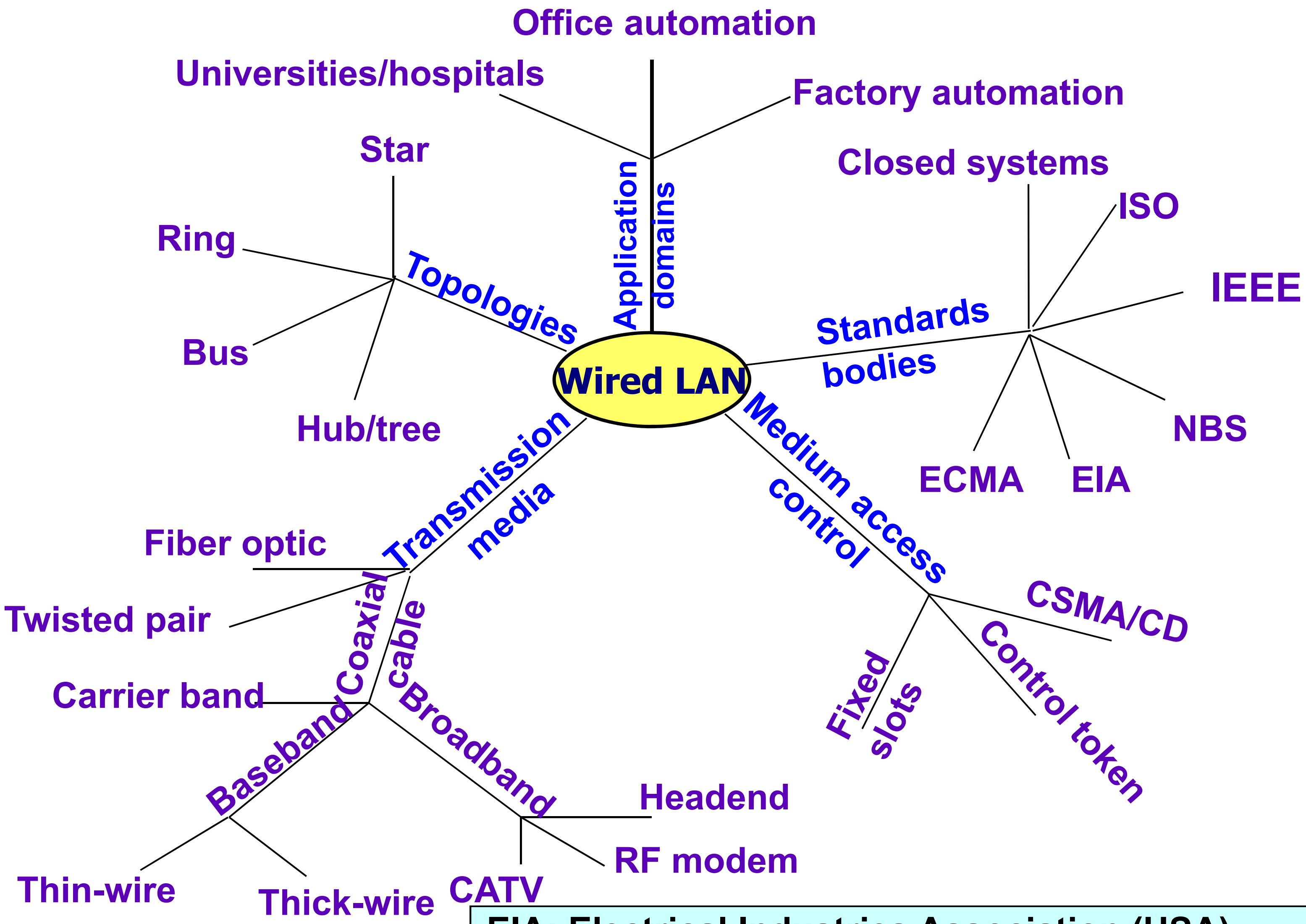


Wireless

- Wireless networks do not require physical infrastructure like cables.
- Wireless networks have short range.
- Wireless networks have limited bandwidth.
- Transmissions can be intercepted easily by a person outside building with a wireless access device.

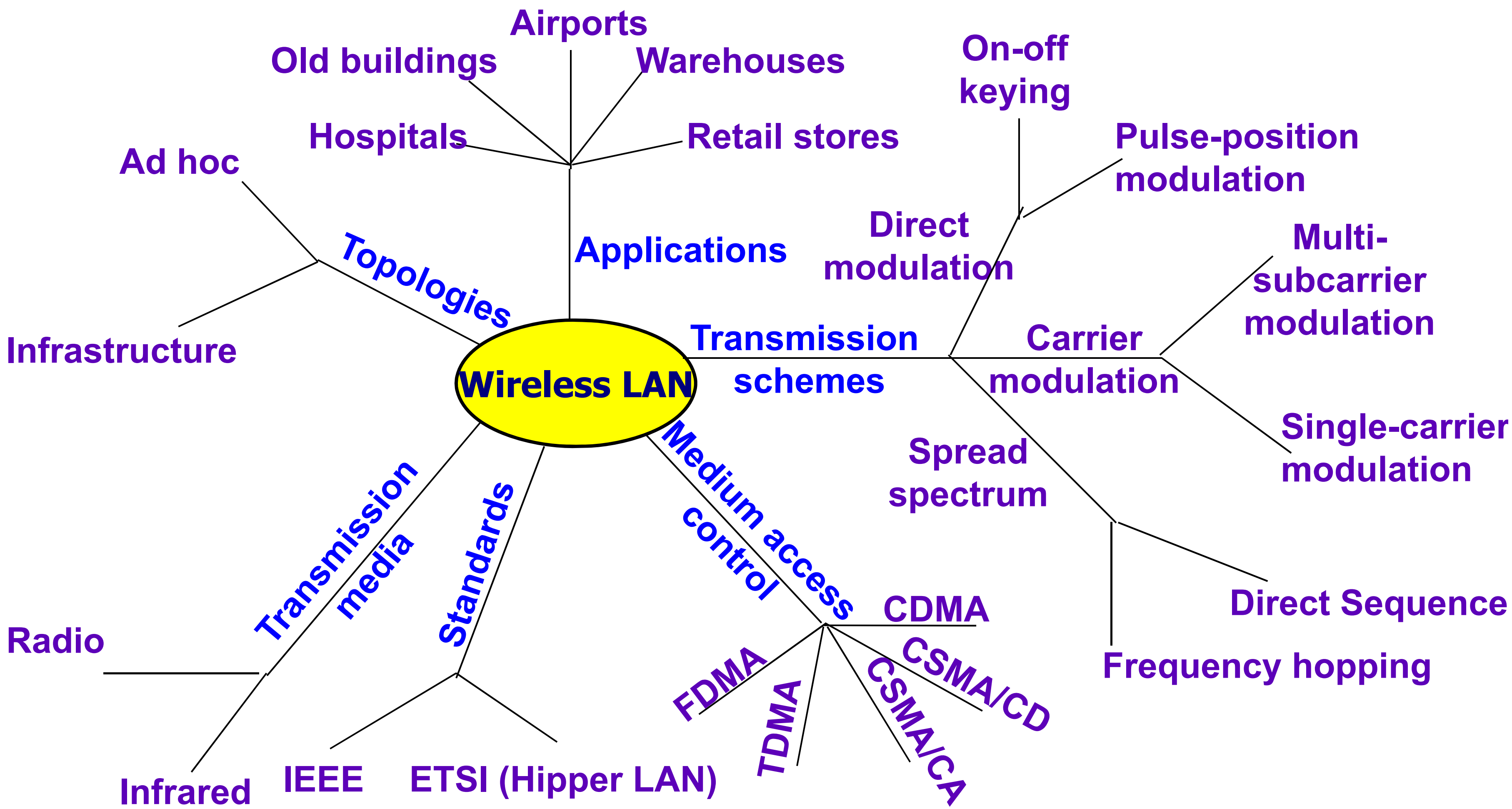


LAN Selections - Wired



EIA: Electrical Industries Association (USA)
ECMA: European Computer Manufacturers Association
NBS: National Bureau of Standards

LAN Selections - Wireless

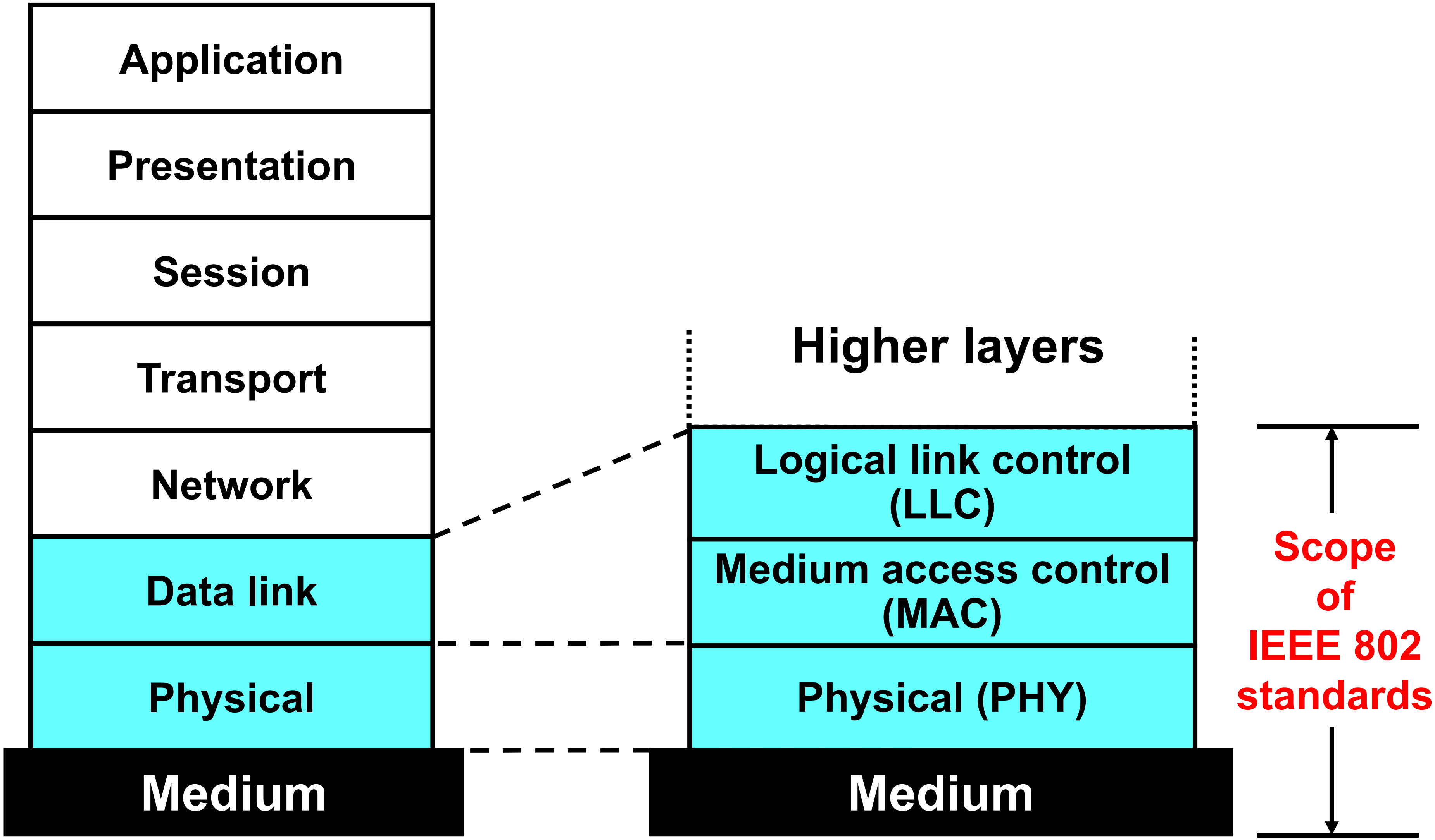


CDMA: Code Division Multiple Access
CSMA/CD: CSMA with Collision Detection
CSMA/CA: CSMA with Collision Avoidance
ETSI: European Telecom. Standards Institute
FDMA: Frequency Division Multiple Access
TDMA: Time Division Multiple Access

IEEE 802 Reference Model

- IEEE 802 committee developed, revises, and extends standards
- Three-layer protocol hierarchy:
  - physical
  - medium access control (MAC)
  - logical link control (LLC)

IEEE 802 Reference Model



Physical Layer

- **Encoding/decoding of signals and bit transmission/reception**
- **Specification of the transmission medium.**
- **The choice of transmission medium is critical in LAN design, and so a specification of the medium is included**

Logical Link Control

- Specifies method of addressing and controls exchange of data
- Independent of topology, medium, and medium access control
- Services:
  - Unacknowledged connectionless service (higher layers handle error/flow control, or simple apps)
  - Acknowledged connectionless service (no prior connection necessary)
  - Connection-mode service (devices without higher-level software)

Media Access Control

- Assembly of data into frame with MAC control, address and error detection fields
- Disassembly of frame
  - Address recognition
  - Error detection
- Govern access to transmission medium
- For the same LLC, several MAC options may be available

IEEE LAN Standards

802 Executive Committee

802.10 LAN Security

802.1 Higher Layer LAN Protocols

802.2 Logical Link Control

802.3 MAC	802.4 MAC	802.5 MAC	802.6 MAC	802.9 MAC	802.11 MAC	802.12 MAC	802.15 MAC	802.16 MAC	802.17 MAC
CSMA/CD	Token Bus	Token Ring	DQDB	Isoc. LAN	WLAN	100VG	PAN	Broad-band Wireless Access	RPR

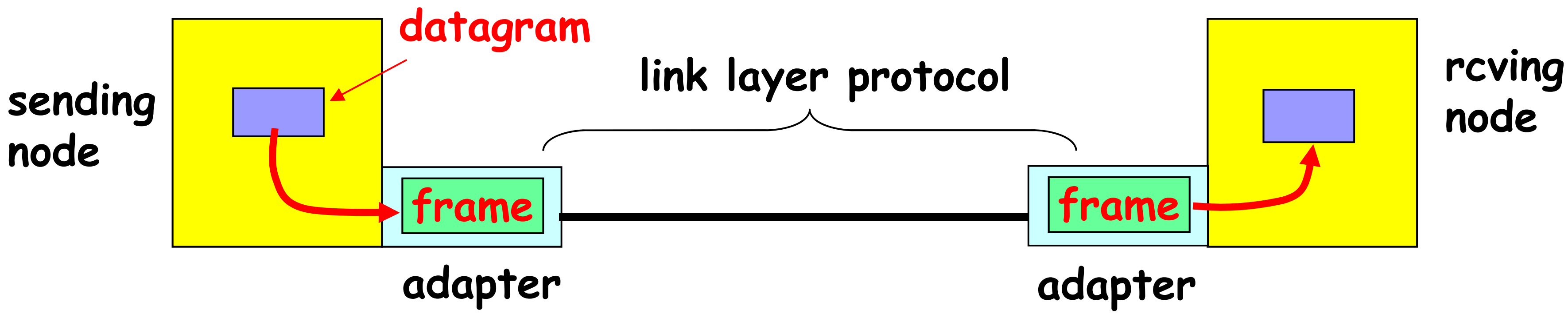
Data Link

Phy-sical

IEEE LAN Standards

- 802.1 Higher LAN Protocols
- 802.2 Logical link control (LLC) (No Activity)
- **802.3 CSMA/CD (Ethernet)**
- 802.4 Token Bus (No Activity)
- 802.5 Token Ring (No Activity)
- 802.6 Metropolitan area network (No Activity)
- 802.7 Broadband technical advisory (No Activity)
- 802.8 Fiber optic technical advisory (Obsolete)
- 802.9 Integrated services LAN (No Activity)
- 802.10 Interoperable LAN Security (No Activity)
- **802.11 Wireless LAN**
- 802.12 100 VG-AnyLAN (No Activity)
- 802.14 Cable-TV based broadband (Obsolete)
- 802.15 Wireless Personal Area Network
- 802.16 Broadband Wireless Access (WiMAX)
- 802.17 Resilient Packet Ring (RPR)

IEEE LAN Standards



- **link layer implemented in "adaptor" (aka NIC)**
  - Ethernet card, PCMCI card, 802.11 card
- **sending side:**
  - encapsulates datagram in a frame
  - adds error checking bits, flow control, etc.
- **receiving side**
  - looks for errors, flow control, etc
  - extracts datagram, passes to rcving node
- **adapter is semi-autonomous link & physical layers**

Discussion Questions

- What is the difference between a physical and a logical topology?
- What is the difference between a bus and a star topology?
- Which media access method sends an intent to transmit signal?
- What are the benefits of using twisted pair over RG58?

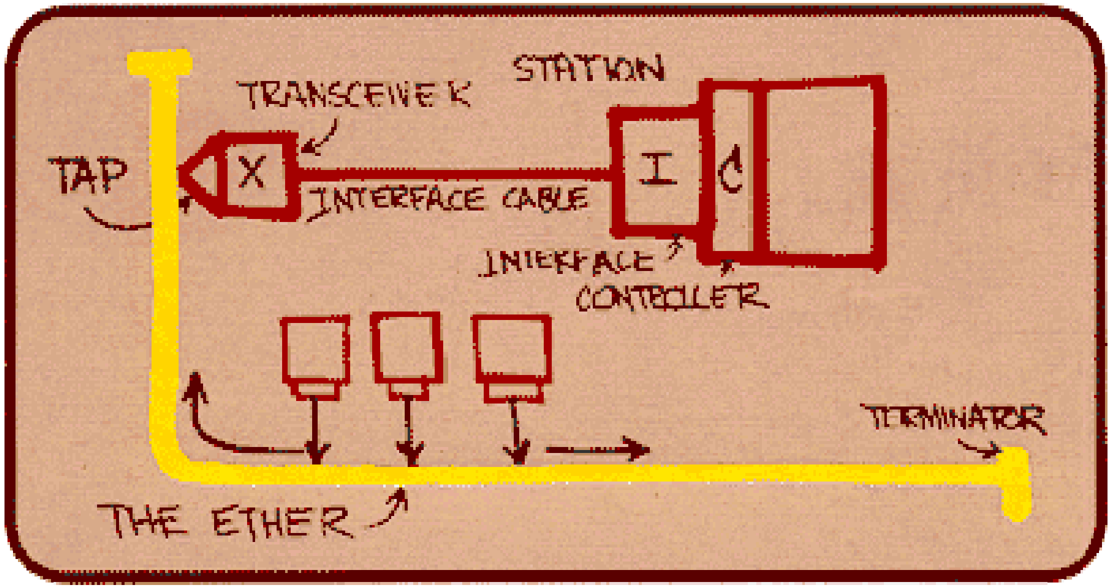
Chapter 4: Roadmap

- Medium Access Control
- Local Area Networks (LANs) and IEEE 802
- Ethernet
- Wireless LAN
- LAN Interconnection
- LAN Switching
- VLAN

Ethernet

"Dominant" wired LAN technology

- First widely used LAN technology
- Simpler, cheaper than token LANs and ATM
- Kept up with speed race: 10 Mbps – 10 Gbps



Metcalfe's Ethernet Sketch, 1972

Origin of Ethernet

- Developed by Xerox Palo Alto Research Center (PARC) in late 1972
- Original designed as a 2.94 Mbps system to connect 100 computers on a 1 km cable
- Later, Xerox, Intel and DEC drew up a standard support 10 Mbps
- Basis for the IEEE's 802.3 specification

Ethernet Basics

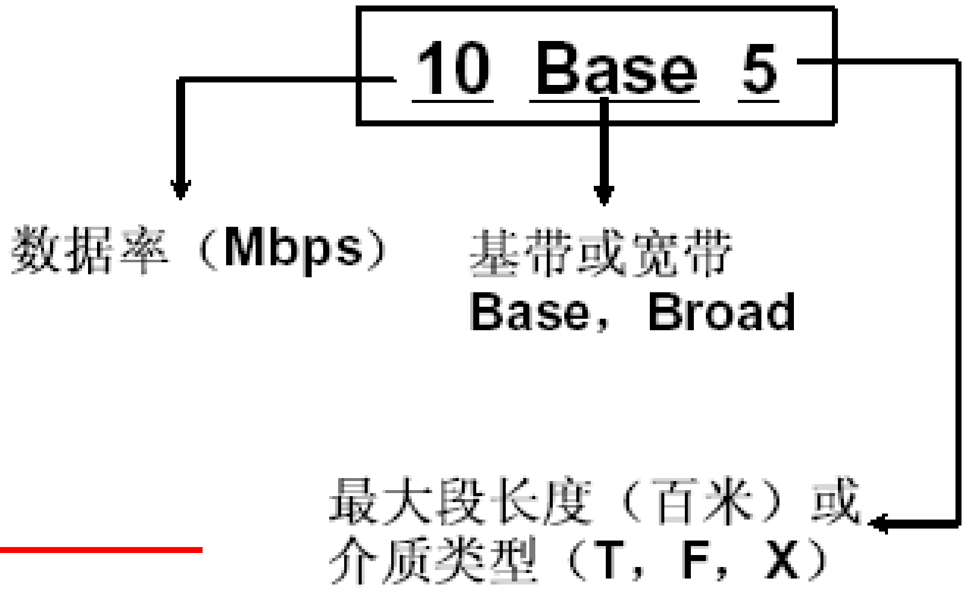
- **Topologies:** Linear bus, Star, Tree
- **Signaling:** Mainly baseband (digital)
- **Access method:** CSMA/CD
- **Specifications:** IEEE 802.3
- **Transfer speed:** 10 Mbps, 100 Mbps, or above
- **Cable types:** Coaxial cables, UTP

Ethernet Basics

<table>
<tr><td colspan="9">Logical Link Control Sublayer</td></tr>
<tr><td colspan="9">802.3 Media Access Control</td></tr>
<tr><td>Physical Signaling Sublayer</td><td rowspan="2">10BASE5 (500m) 50 Ohm Coax N-Style</td><td rowspan="2">10BASE2 (185m) 50 Ohm Coax BNC</td><td rowspan="2">10BASE-T (100m) 100 Ohm UTP RJ-45</td><td rowspan="2">100BASE-TX (100m) 100 Ohm UTP RJ-45</td><td rowspan="2">1000BASE-CX (25m) 150 Ohm STP mini-DB-9</td><td rowspan="2">1000BASE-T (100m) 100 Ohm UTP RJ-45</td><td rowspan="2">1000BASE-SX (220-550m) MM Fiber SC</td><td rowspan="2">1000BASE-LX (550-5000m) MM or SM Fiber SC</td></tr>
<tr><td>Physical Medium</td></tr>
</table>

802.3 Cabling

- 1Base5 双绞线
- 10Broad36 CATV
- 10Base5 粗同轴
- 10Base2 细同轴
- 10BaseT UTP
- 10BaseF MMF
- 100BaseT UTP
- 100BaseF MMF/SMF
- 1000BaseX STP/MMF/SMF
- 1000BaseT UTP



Ethernet Frame Structure

Preamble	Dest. Address	Source Address	Type	Data	PAD	CRC
8 Bytes	6 Bytes	6 Bytes	2 Bytes	0 - 1500 Bytes	0 - 46 Bytes	4 Bytes

Preamble:

7 bytes with pattern 10101010 followed by one byte with pattern 10101011, used to synchronize receiver, sender clock rates.

Addresses: 6 bytes (48 bits)

if adapter receives frame with matching destination address, or with broadcast address, it passes data in frame to net-layer protocol otherwise, adapter discards frame.

Ethernet Frame Structure

Preamble	Dest. Address	Source Address	Type	Data	PAD	CRC
8 Bytes	6 Bytes	6 Bytes	2 Bytes	0 - 1500 Bytes	0 - 46 Bytes	4 Bytes

Type:
indicates the higher layer protocol (mostly IP but others may be supported such as Novell IPX and AppleTalk).
CRC:
checked at receiver, if error is detected, the frame is simply dropped.
Frame length: Min. =64B, Max. = 1518B

Ethernet Frame Structure

Bytes	8		6	6	2	0-1500	0-46	4
(a)	Preamble		Destination address	Source address	Type	Data	Pad	Check-sum
(b)	Preamble	S o F	Destination address	Source address	Length	Data	Pad	Check-sum

(a) Ethernet frame (b) IEEE 802.3 frame

Ethernet Address

- 48 bits long: 00 00 E2 15 1A CA
- Governed by IEEE and are usually imprinted on Ethernet cards when the cards are manufactured ➔ physical address or hardware address.
- Type:
  - Single address: one station
  - Group address: a group of stations
  - Broadcast address (all '1'): all stations

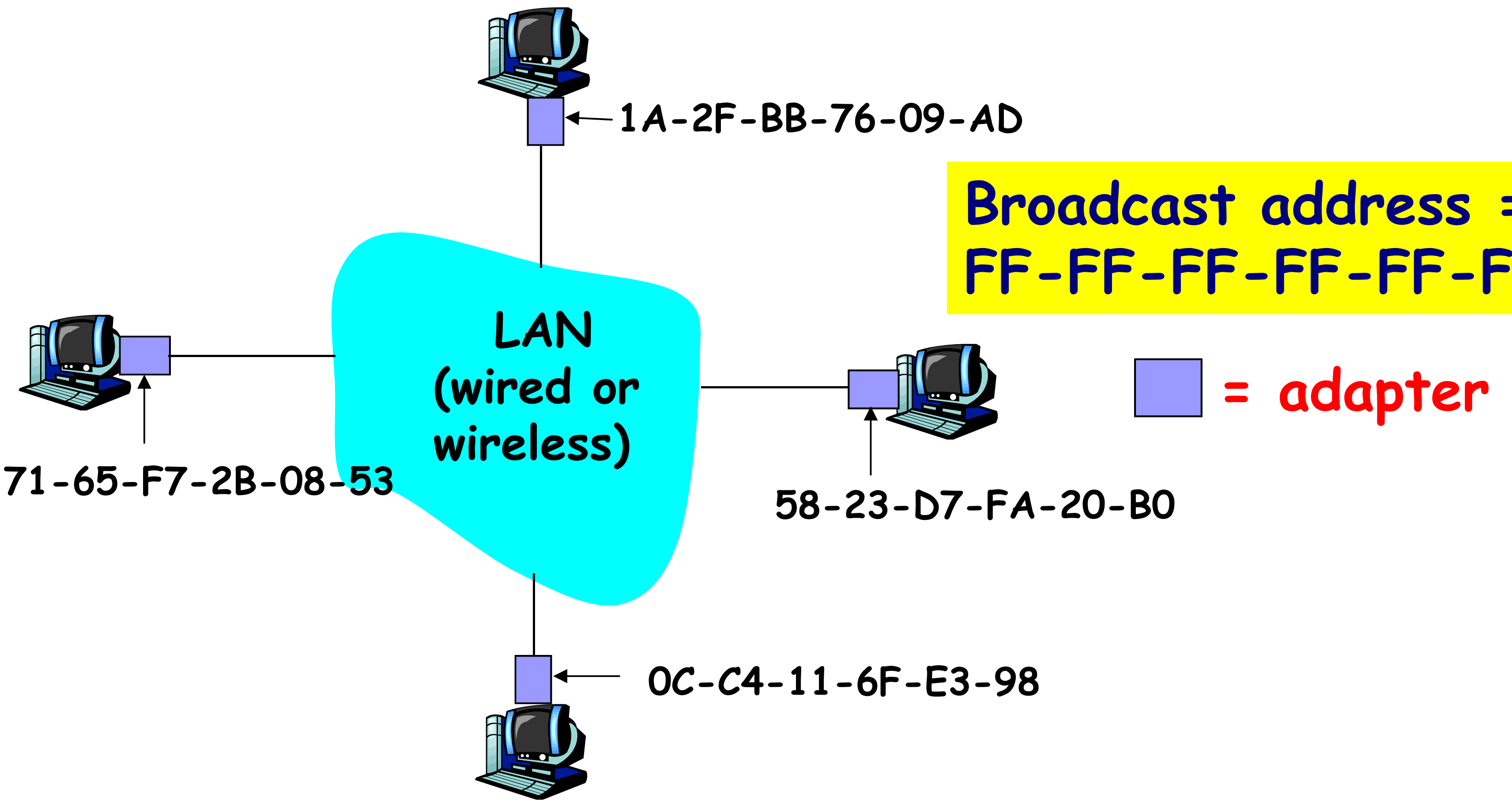
Ethernet Address

Examples of Manufacturer IDs

Cisco	**: 00-00-0C-**	**3Com**	**: 00-60-8C-**
	: 00-60-09-		**: 00-60-08-**
Sun	**: 08-00-20-**	**IBM**	**: 08-00-5A-**
Nokia	**: 00-40-43-**		

Ethernet Address

Each adapter on LAN has unique address



Broadcast address = FF-FF-FF-FF-FF-FF

Ethernet uses CSMA/CD

- **1-persistent CSMA/CD**
  - If line is idle (no carrier sensed)
    - Send immediately
    - Send maximum of 1500B data (1518B frame)
    - Wait 9.6 μs before sending again
  - If line is busy (carrier sensed)
    - Wait until line becomes idle
      - called *1-persistent* sending
  - If collision detected
    - Stop sending and send jam signal
    - Try again later

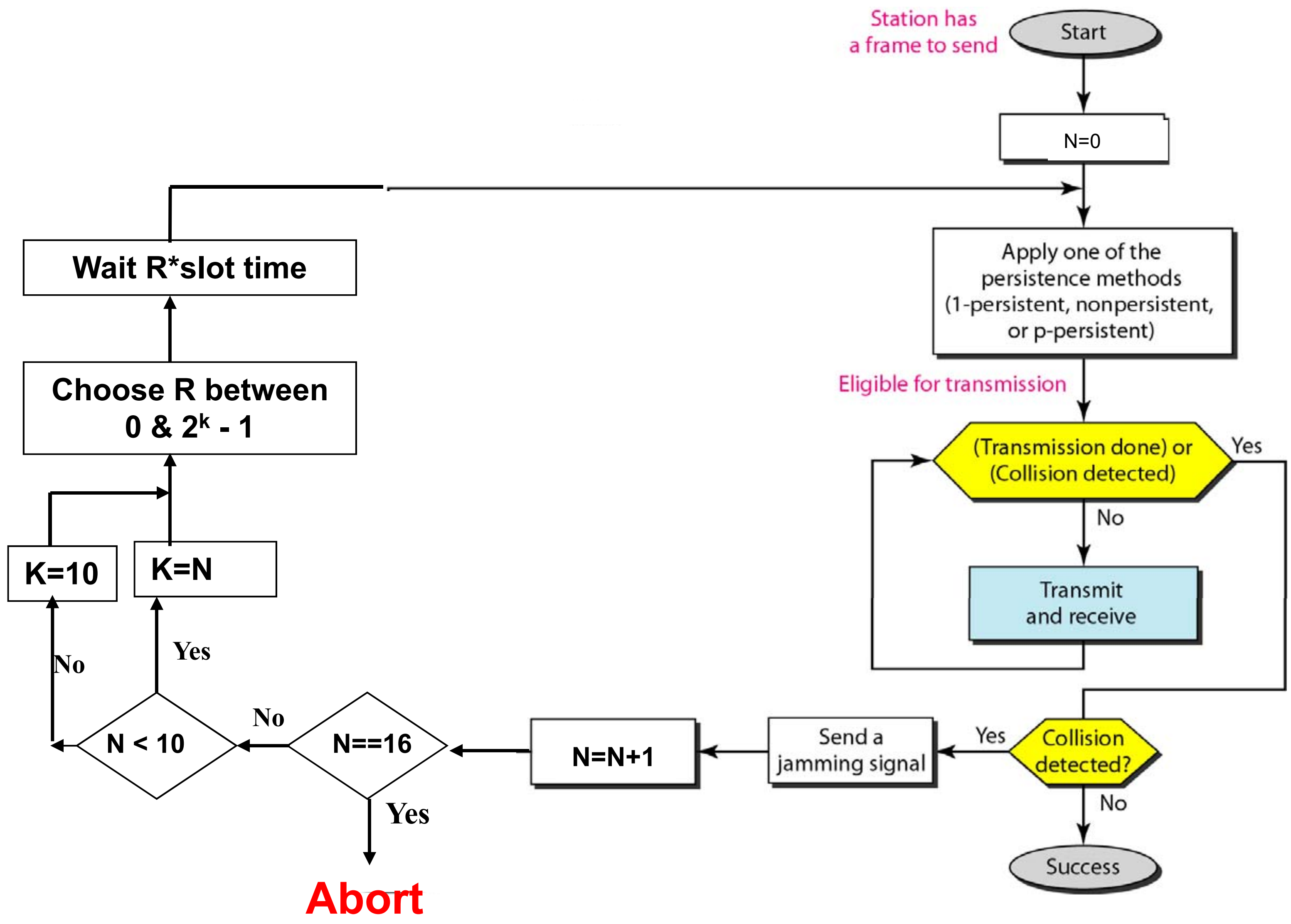
Exponential Backoff Algorithm

- **If a station is involved in a collision, it waits a random amount of time before attempting a retransmission.**
- **The random time is determined by the Exponential Backoff Algorithm**

Exponential Backoff Algorithm

$$\begin{cases} R = random[0, 2^{k-1}] \\ WaitingTime = R \bullet SlotTime \end{cases}$$

- **K = Min[# of retransmission, 10]**
- **SlotTime = 2*maximum propagation delay + Jam sequence transmission time (= 51.2 usec for Ethernet 10-Mbps LAN)**
- **Give up after 16 unsuccessful attempts and report failure to higher layers.**



Flow diagram for the CSMA/CD

Questions

- **How comes the minimum frame size of 64 bytes?**
  - **IEEE 802.3 specifies max value of slot to be 51.2us.This relates to maximum distance of 2500m between hosts (propagation delay)**
  - **At 10Mbps it takes 51.2us to send 512 bits (64B)**
  - **So, Ethernet frames must be at least 64B long**
    - **14B header, 46B data, 4B CRC**
    - **Padding is used if data is less than 46B**
  - **The minimum frame size is also called slottime**
- **Why we need minimum size?**
  - **Detecting frame collision**
  - **Distinguish good frame from damaged ones**

Questions

- **Q: If we keep the minimum frame size of 64 bytes for compatibility reason, what is the contention time for 100M and 1000M Ethernet?**
  - **5.12 μs for 100M, network span is 204m**
  - **0.512 μs for 1000M, network span is 20m ???**
- **1000M Ethernet contention time is 4.096 μs, remain the network span 204m**

CSMA/CD Maximum efficiency

- **When every nodes send in turn without collision, max. throughput achieved**

$$T = \frac{L}{t_p + t_{trans}} = \frac{L}{d/v + L/R}$$

Where

L – frame length

t_p – propagation delay

t_{trans} – frame transmission delay

R – Data rate

d – distance

v – signal speed

CSMA/CD Maximum efficiency

- **Maximum efficiency：**

$$U = \frac{T}{R} = \frac{L/R}{d/v + L/R}$$

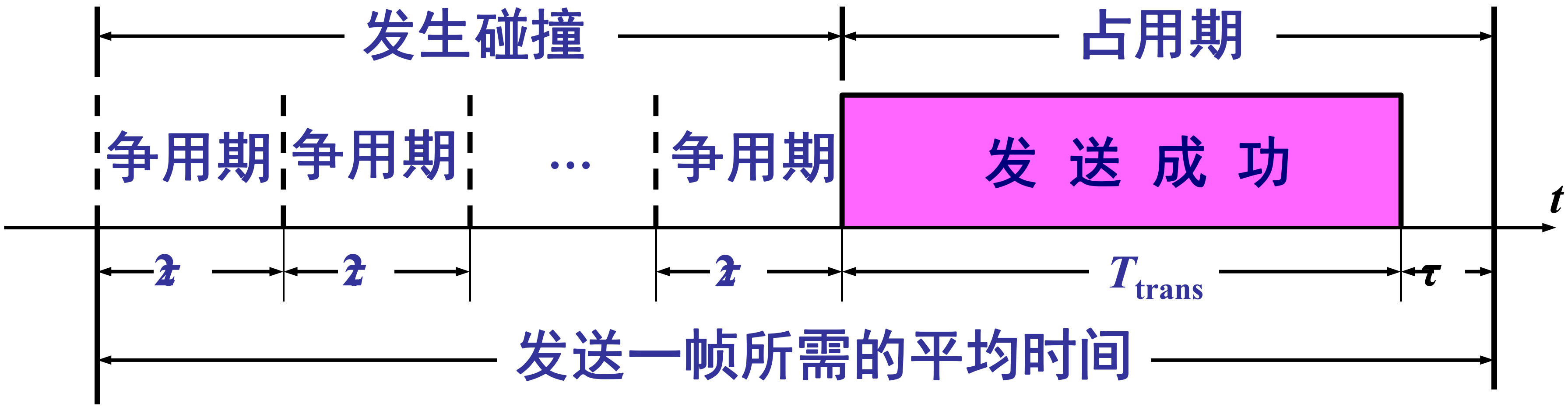
- **Let a=t_p/t_{trans}**

$$U = \frac{1}{\alpha + 1}$$

- **a = (d/v) / (L/R) = Rd/vL**
  a越大（R与d乘积越大）信道利用率越低
  其中

CSMA/CD efficiency

- 一个帧从开始发送，经可能发生的碰撞后，将再重传数次，到发送成功且信道转为空闲(即再经过时间 τ 使得信道上无信号在传播)时为止，是发送一帧所需的平均时间。



CSMA/CD efficiency

- t_{prop} = max prop between 2 nodes in LAN
- t_{trans} = time to transmit max-size frame

$$\text{efficiency} = \frac{1}{1 + 5t_{prop} / t_{trans}} = \frac{1}{1 + 5\alpha}$$

Where

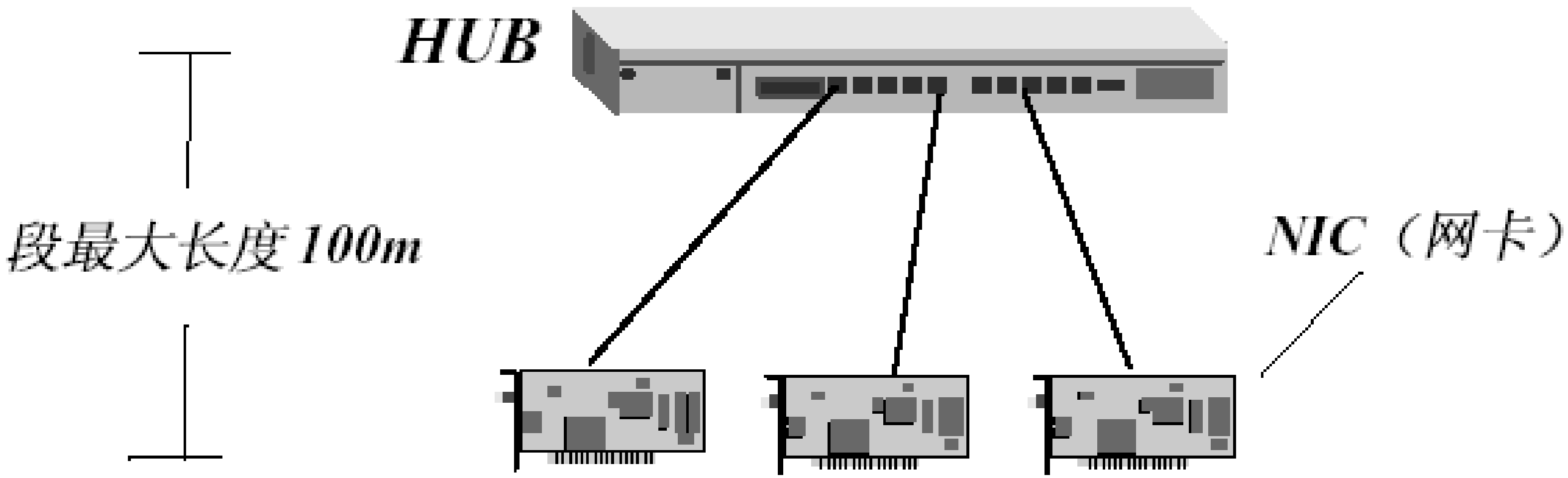
$\alpha = t_{prop}/t_{trans}$ = T/(R/L) = RT/L

T – Max. prop. time between two nodes

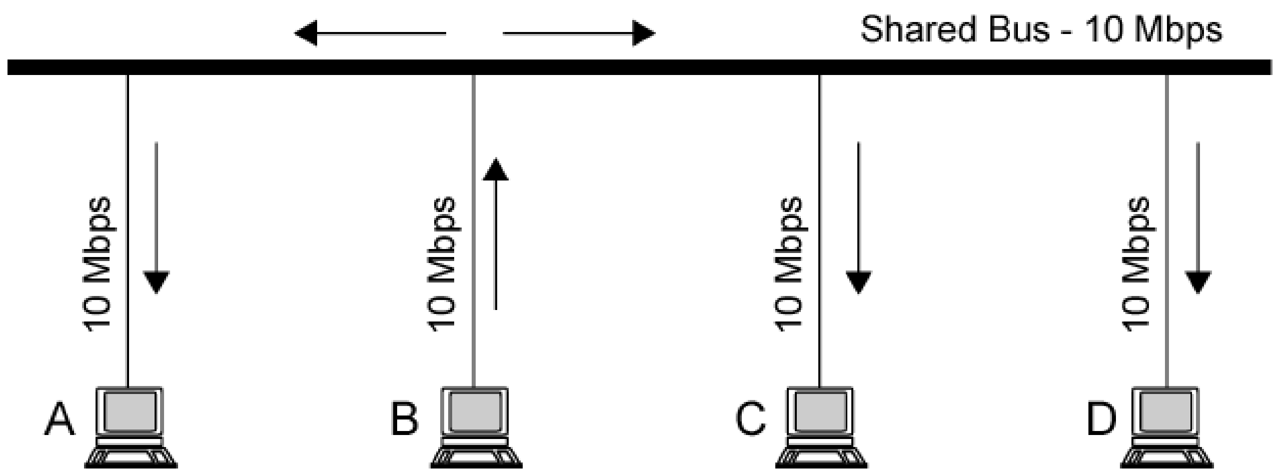
α small → early collision detection, efficiency
α large → late collision detection, inefficiency

Ethernet: 10Base_T

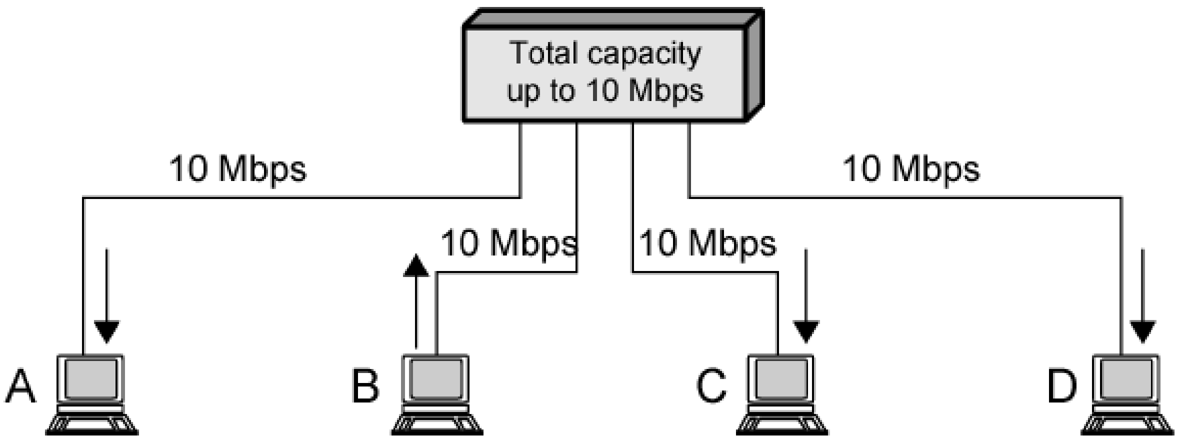
- 10Mbps rate
- T stands for Twisted Pair
- Nodes connect to a hub: "star topology"; 100m max distance between nodes and hub



Shared Medium Bus and Hub



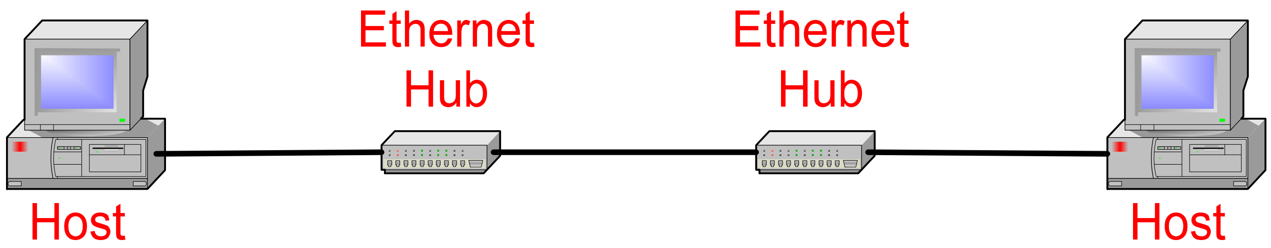
(a) Shared medium bus



(b) Shared medium hub

Ethernet Hub

- **Used to connect hosts to Ethernet LAN and to connect multiple Ethernet LANs**
- **Collisions are propagated.**



Hub:

•Essentially physical-layer repeater

•Same bandwidth shared by all nodes

•Single Collision Domain

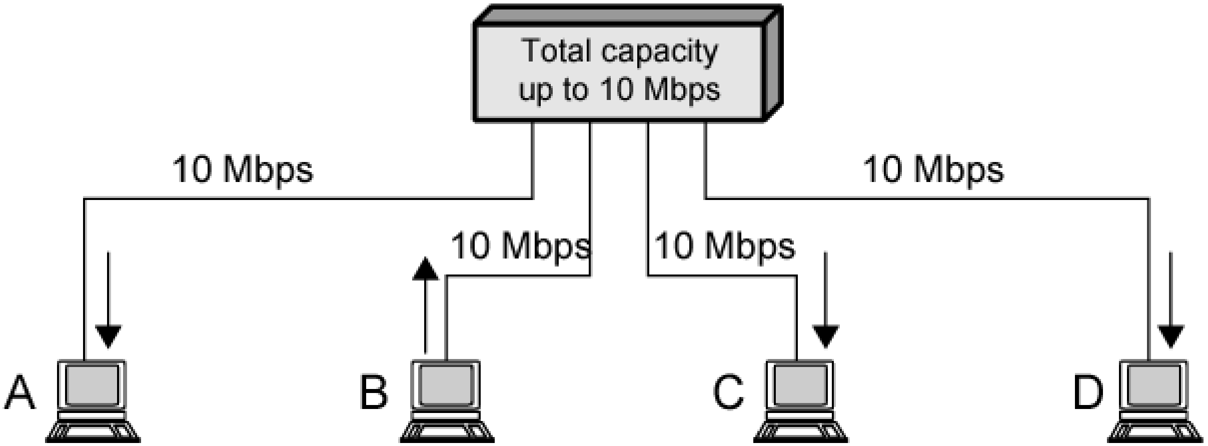
Switched Ethernet

- **Shared Ethernet Problem:**

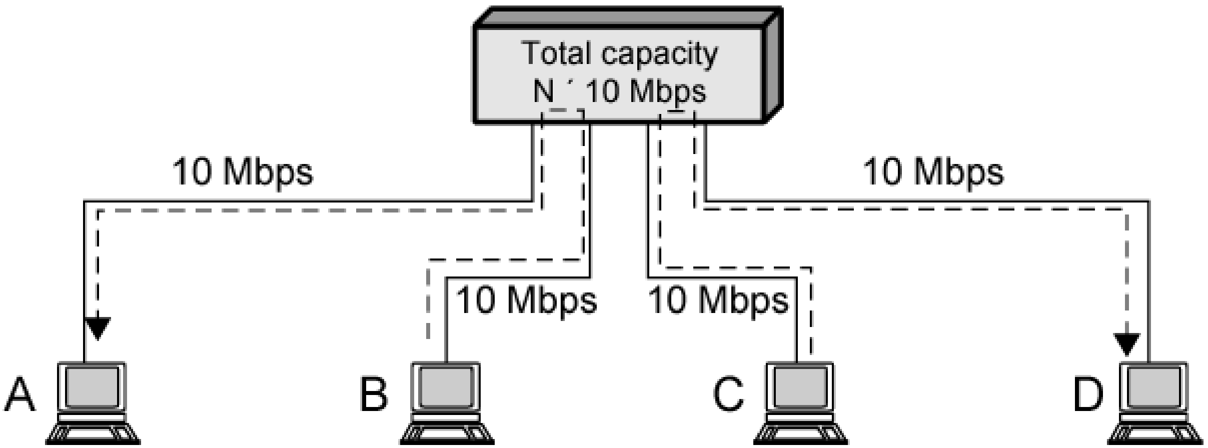
  As more stations are added, traffic will go up, and so will the possibility of collisions, then, the network will saturate.

- **Solution: Divide the network into separate sub-LANs and connect them through a high-speed switch.**

Switched Ethernet



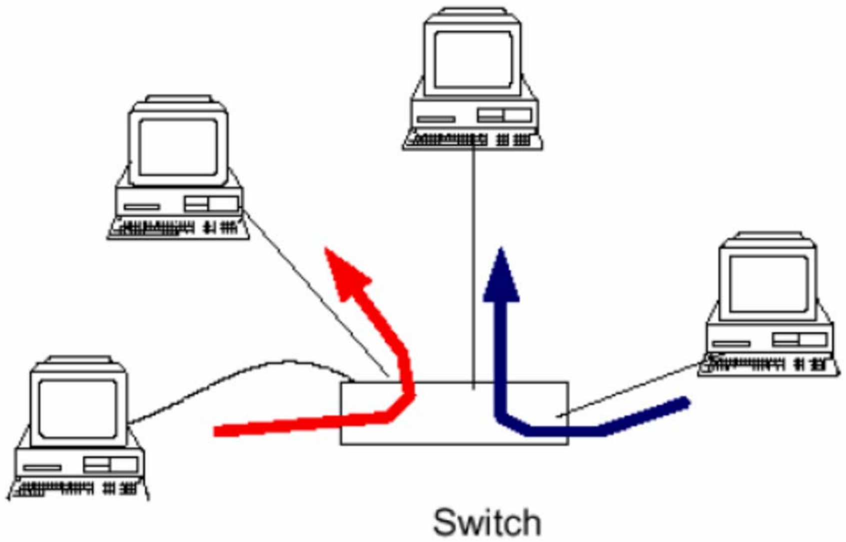
(b) Shared medium hub



(c) Layer 2 switch

Switched Ethernet

- Multiple transmissions are possible
- Switch stores frames that wait for same output

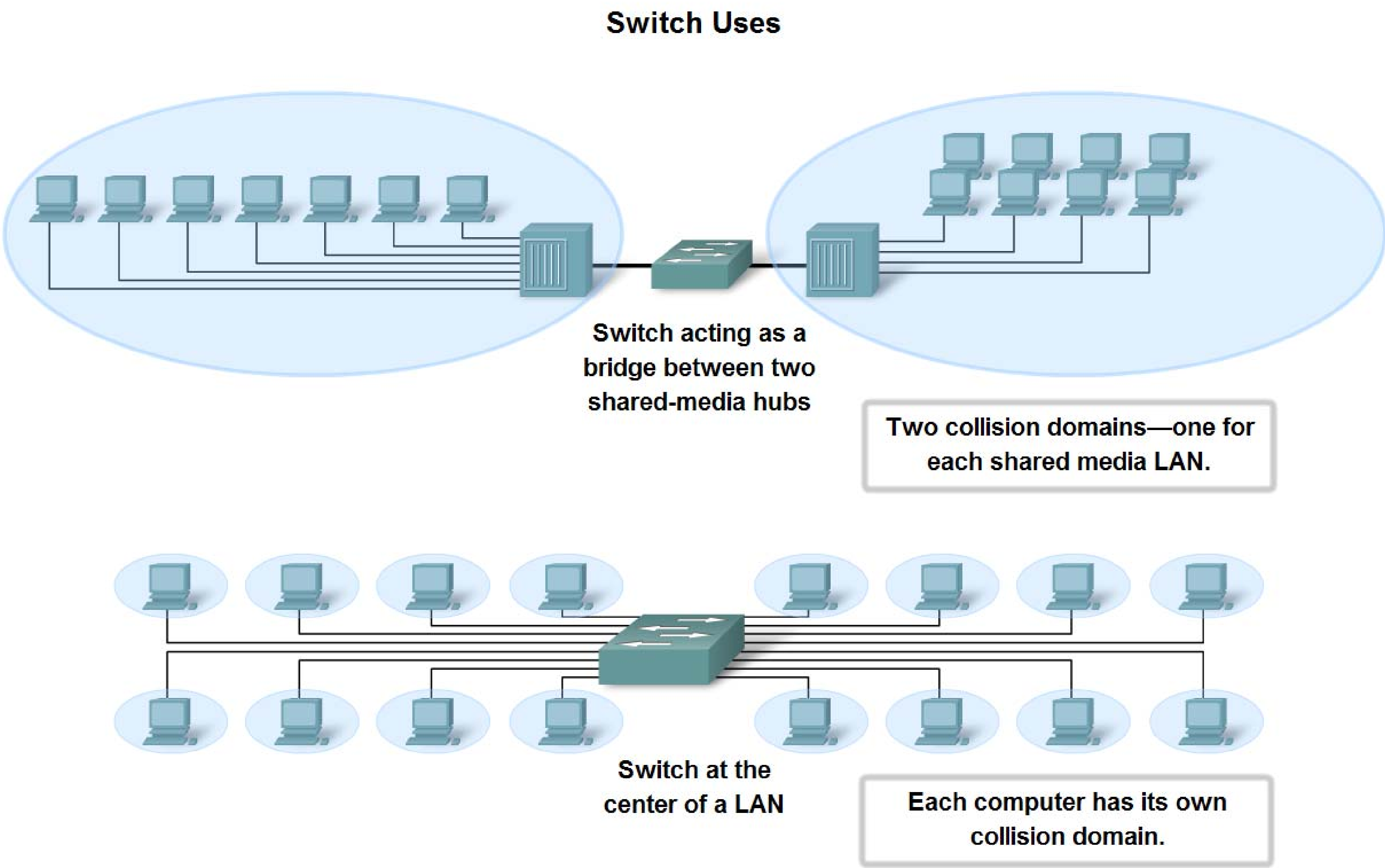


Switch:

•Dedicated bandwith

•Each switch port is a collision domain

Switched Ethernet



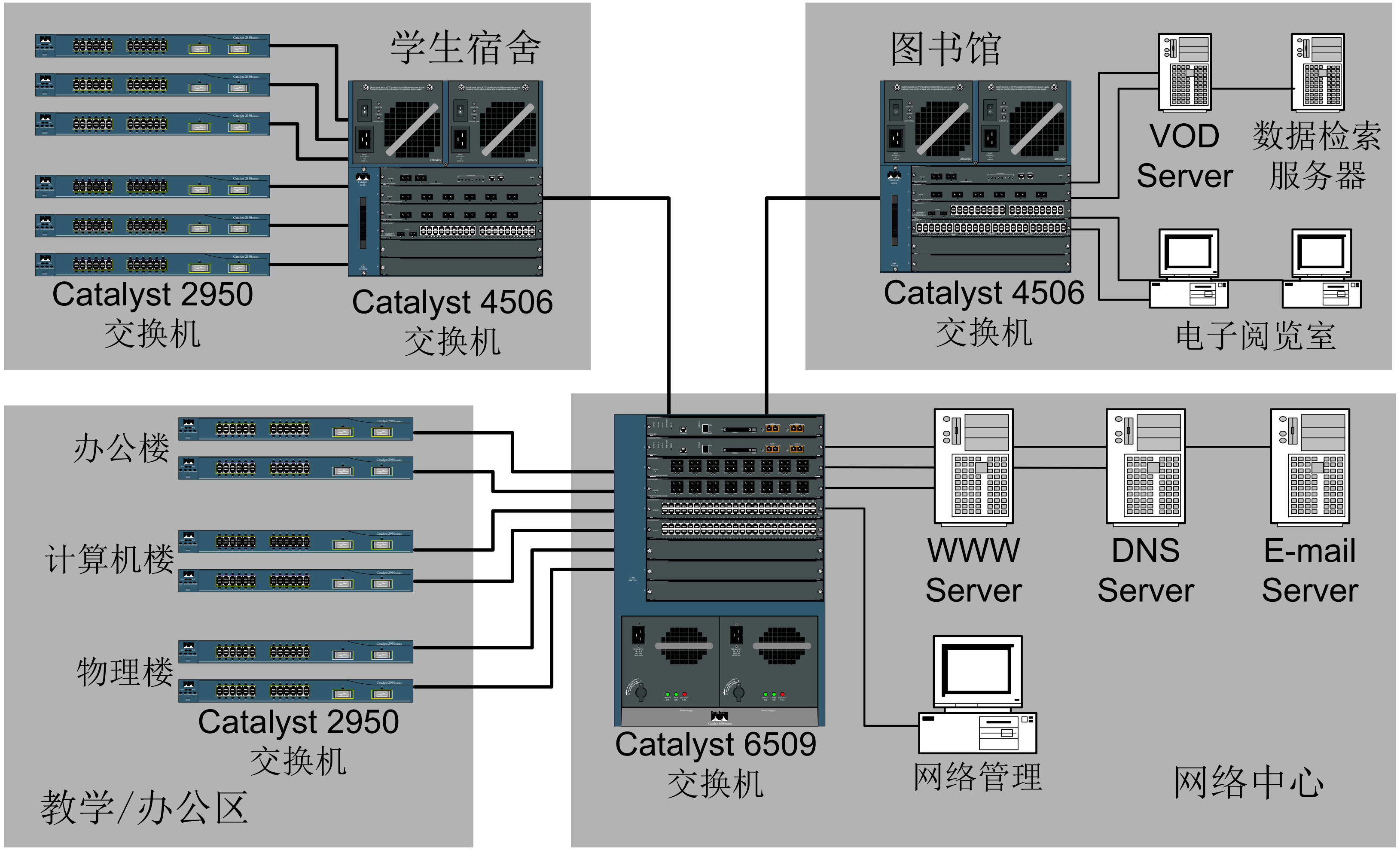
以太网交换机

- **Cisco Catalyst 6500**系列交换机

以太网交换机

- **Cisco Catalyst 3750系列交换机——**堆叠实例

交换机组网实例

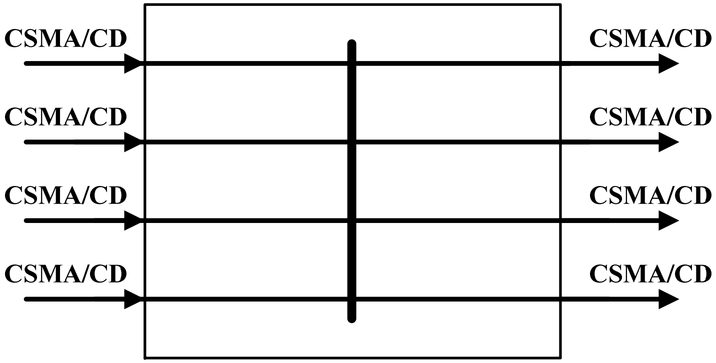


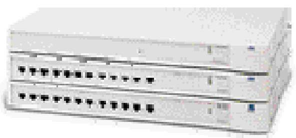
学生宿舍
Catalyst 2950
交换机
Catalyst 4506
交换机
图书馆
VOD
Server
数据检索
服务器
Catalyst 4506
交换机
电子阅览室
办公楼
计算机楼
物理楼
Catalyst 2950
交换机
教学/办公区
Catalyst 6509
交换机
WWW
Server
DNS
Server
E-mail
Server
网络管理
网络中心

Ethernet Hubs vs. Ethernet Switches

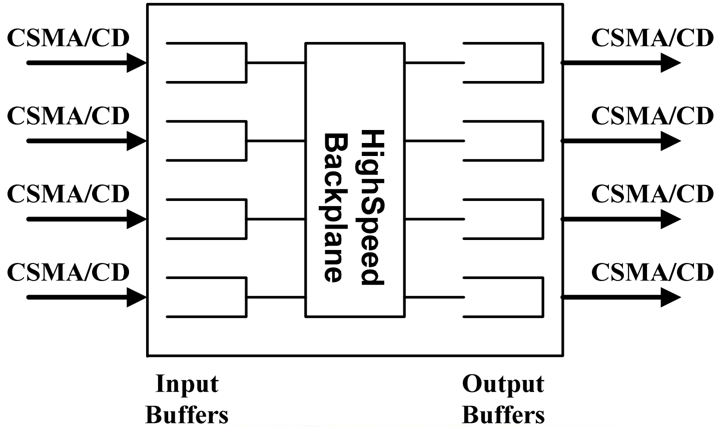
- An **Ethernet switch** is a switch for Ethernet frames
  - **Buffering** of frames prevents collisions.
  - Each port is isolated and builds its own **collision domain**
- An **Ethernet Hub** does not perform buffering:
  - Collisions occur if two frames arrive at the same time.

Hub





Switch

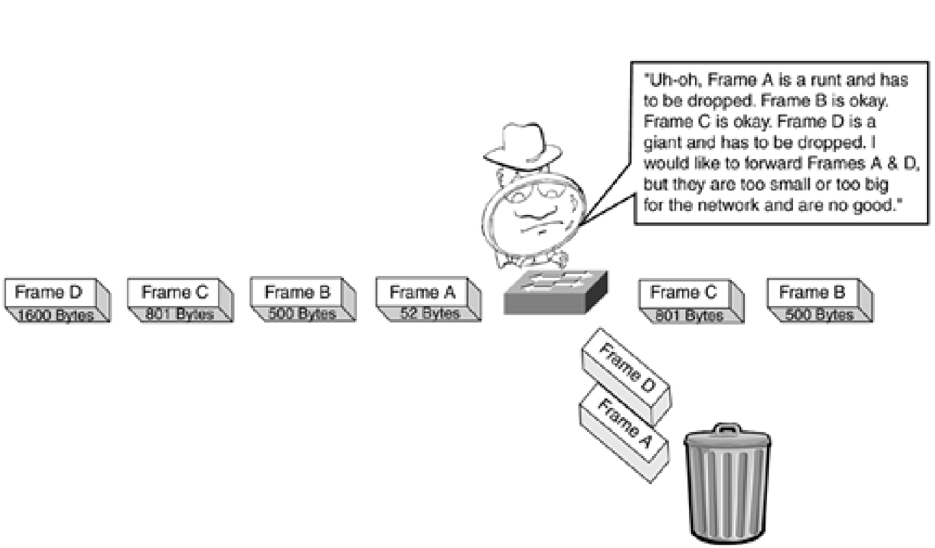


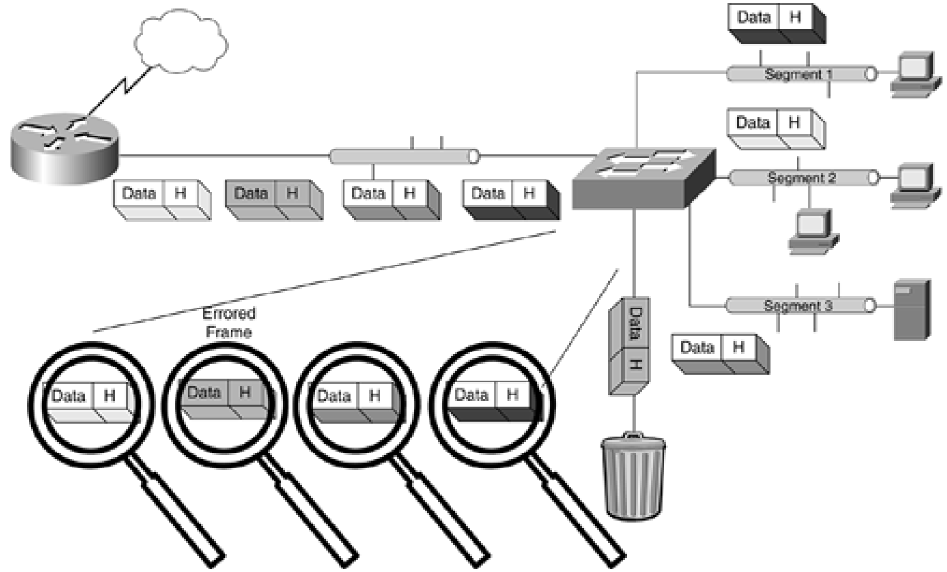
Layer 2 Switches

- Central hub acts as switch
- Incoming frame from particular station switched to appropriate output line
- Unused lines can switch other traffic
- More than one station transmitting at a time
- Multiplying capacity of LAN

Layer 2 Switching Methods

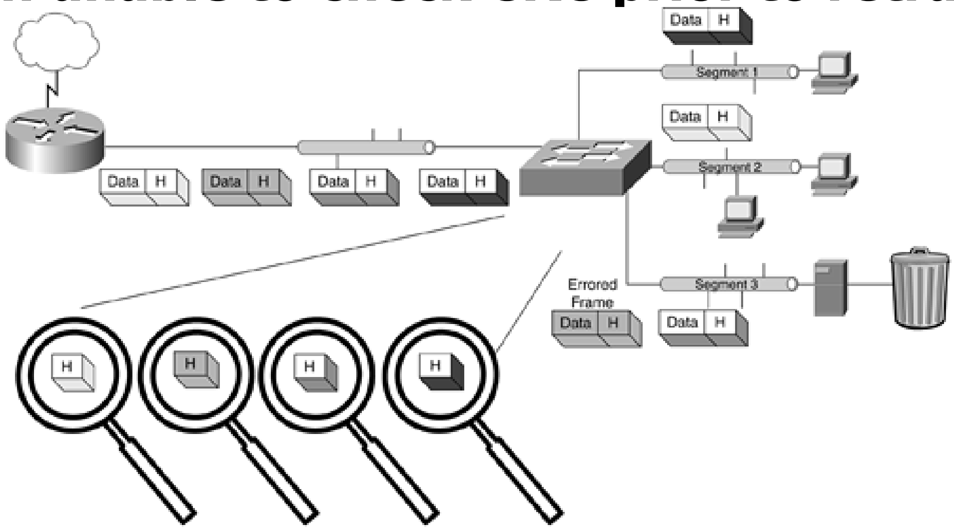
- **Store-and-forward switching**
  - Accepts frame on input line
  - Buffers it briefly,
  - Then routes it to appropriate output line
  - Delay between sender and receiver
  - Boosts integrity of network





Layer 2 Switching Methods

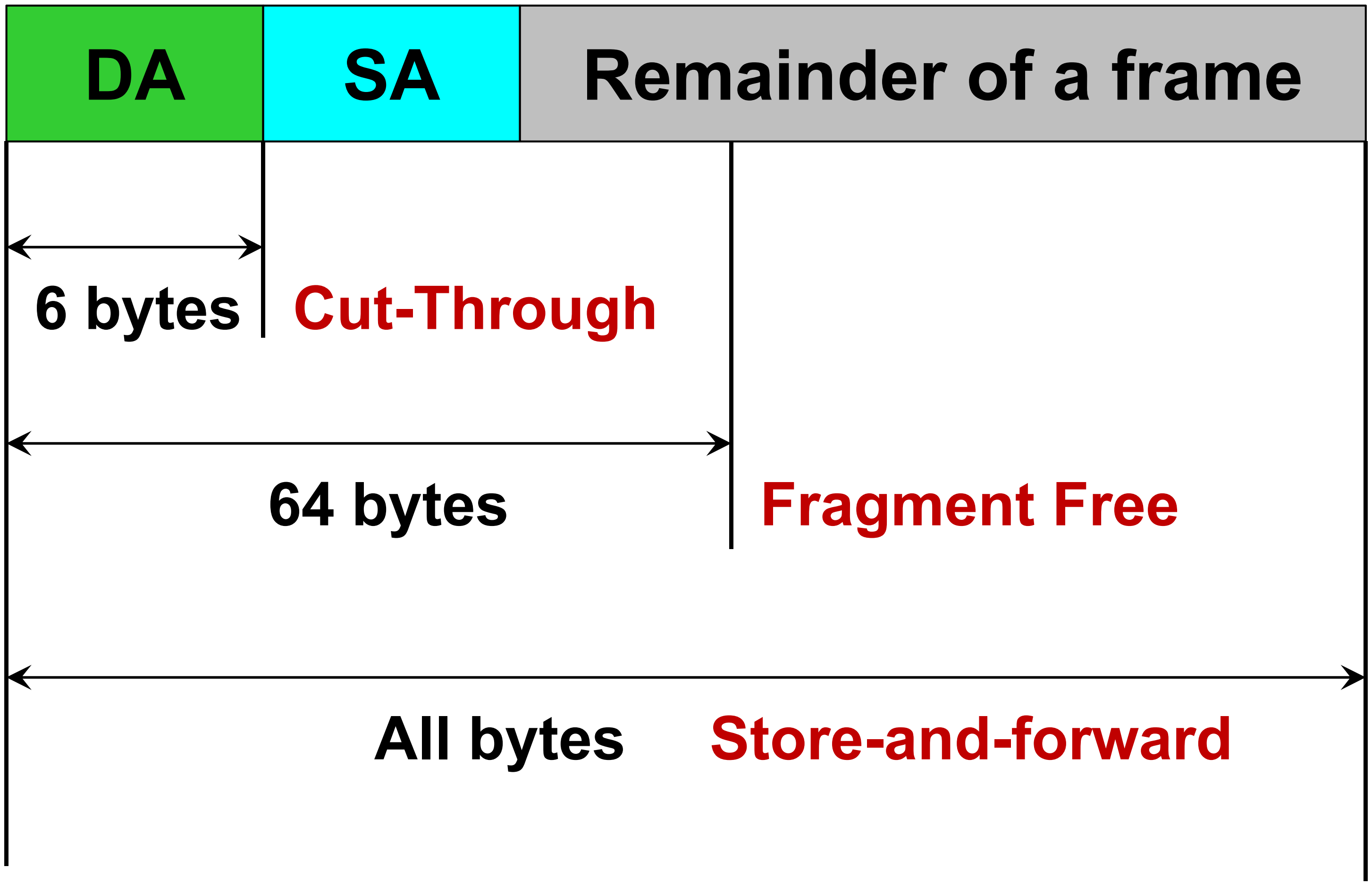
- **Cut-through switching**
  - Takes advantage of destination address appearing at beginning of frame
  - Switch begins repeating frame onto output line as soon as it recognizes destination address
  - Highest possible throughput
  - Risk of propagating bad frames
    - Switch unable to check CRC prior to retransmission



Layer 2 Switching Methods

- **Fragment Free switching**
  - A hybrid version of Store and Forward and Cut- Through.
  - It stores and checks the first 64 bytes of the frame before forwarding. It processes only those frames that have first 64bytes valid.
  - Any frame less than 64 bytes is known as runt. Runt is an invalid frame type.
  - This method filters runt while maintaining the speed.

Layer 2 Switching Methods



Layer 2 Switch

- Challenge
  - Learning which frames to copy across links
  - Avoiding forwarding loops

WHY and HOW ?

Fast Ethernet

- IEEE 802.3u
- 10x speed increase (100m max cable length retains min 64 byte frames)
- Replace Manchester with 4B/5B
- Full-duplex operation using switches
- Speed & duplex auto-negotiation

•在半双工方式下，仍使用 IEEE 802.3 的CSMA/CD 协议。
•可在全双工方式下工作而无冲突发生。此时不使用 CSMA/CD 协议
•MAC 帧格式仍然是 802.3 标准规定的

Fast Ethernet

- 三种不同的物理层标准
  - **100BASE-TX**
    - 使用 2 对 UTP 5 类线或屏蔽双绞线 STP
  - **100BASE-FX**
    - 使用 2 对光纤
  - **100BASE-T4**
    - 使用 4 对 UTP 3 类线或 5 类线

Name	Cable	Max. segment	Advantages
100Base-T4	Twisted pair	100 m	Uses category 3 UTP
100Base-TX	Twisted pair	100 m	Full duplex at 100 Mbps
100Base-FX	Fiber optics	2000 m	Full duplex at 100 Mbps; long runs

Gigabit Ethernet

- IEEE 802.3{z, ab}
- uses standard Ethernet frame format
- allows for point-to-point links and shared broadcast channels
- in shared mode, CSMA/CD is used; short distances between nodes required for efficiency
- Full-Duplex at 1 Gbps for point-to-point links

Gigabit Ethernet Specifications

- **Two different modes of operation**
  - **Full-duplex mode:** allows traffic in both direction at the same time, point-to-point communication, no contention, **CSMA/CD is not used**.
  - **Half-duplex mode:** connected to a hub, collisions are possible, **CSMA/CD is required**, the maximum distance is 100 times less, or 25 meters for 64-byte short frame, to maintain the essential properties of Ethernet.

Gigabit Ethernet Specifications

- **Two features to the standard to increase the radius**
  - **Carrier extension:** tells the hardware to add its own padding after the normal frame to extend the frame to 512 bytes, has a line efficiency of 9%(46/512).
  - **Frame bursting:** allows a sender to transmit a concatenated sequence of multiple frames in a single transmission, if the total burst is less than 512 bytes, the hardware pads it again.

Gigabit Ethernet

- 不同的物理层
  - 1000BASE-X：基于光纤通道的物理层
    - 1000BASE-SX SX表示短波长
    - 1000BASE-LX LX表示长波长
    - 1000BASE-CX CX表示铜线
  - 1000BASE-T
    - 使用 4对 5 类线 UTP

Name	Cable	Max. segment	Advantages
1000Base-SX	Fiber optics	550 m	Multimode fiber (50, 62.5 microns)
1000Base-LX	Fiber optics	5000 m	Single (10 μ) or multimode (50, 62.5 μ)
1000Base-CX	2 Pairs of STP	25 m	Shielded twisted pair
1000Base-T	4 Pairs of UTP	100 m	Standard category 5 UTP

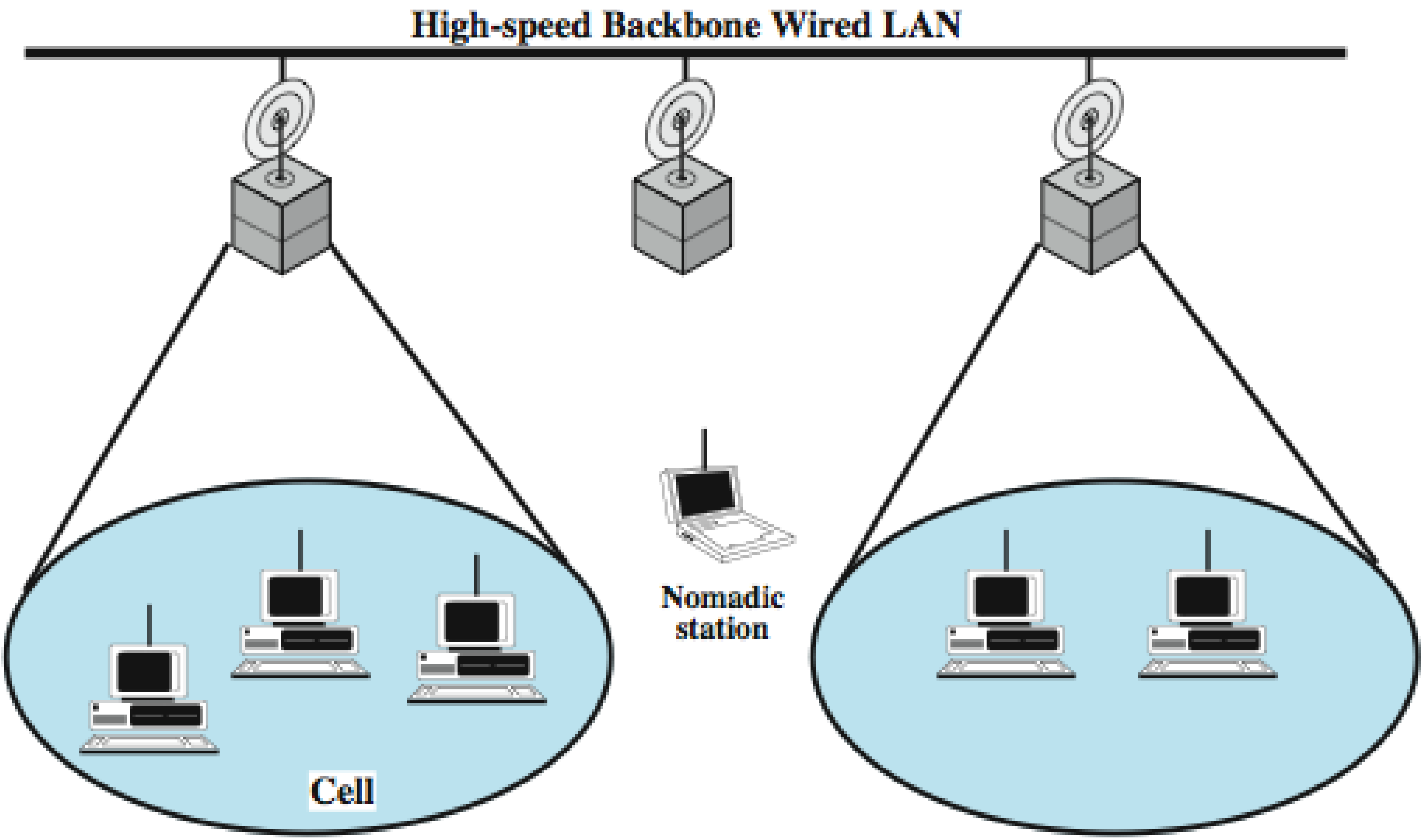
Chapter 4: roadmap

- Medium Access Control
- Local Area Networks (LANs) and IEEE 802
- Ethernet
- Wireless LAN
- LAN Interconnection
- LAN Switching
- VLAN

Wireless LAN

- A wireless LAN or WLAN is a wireless local area network that uses radio waves as its carrier.
- Key application areas:
  - LAN extension
  - cross-building interconnect
  - nomadic access
  - ad hoc networking

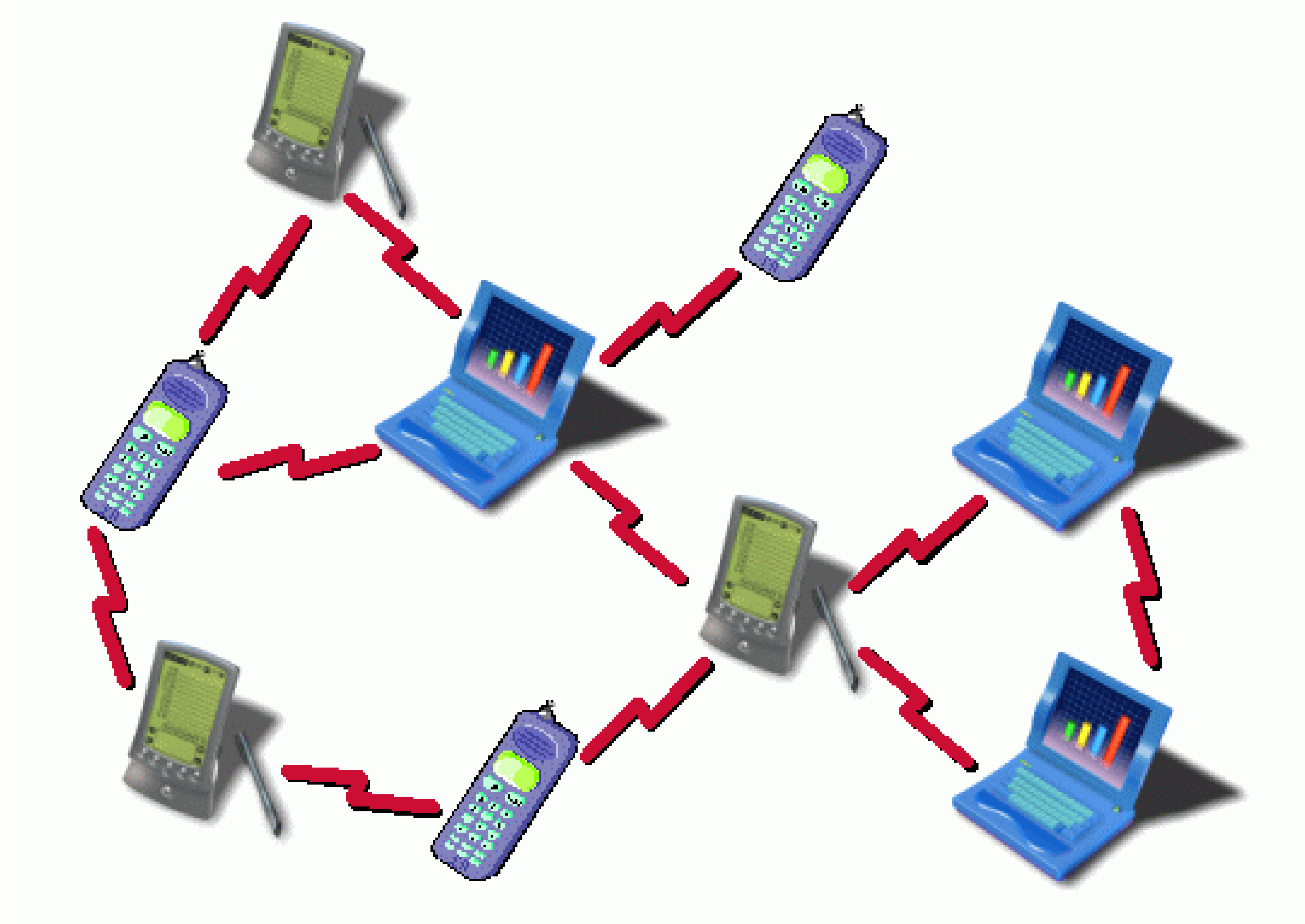
Infrastructure Wireless LAN



(a) Infrastructure Wireless LAN

Ad Hoc Networking

temporary peer-to-peer network (no infrastructure)



Wireless LAN Requirements

THROUGHPUT – should make efficient use of medium	**NUMBER OF NODES-** hundreds of nodes across multiple cells	**CONNECTION TO BACKBONE LAN –** use of control modules
SERVICE AREA – coverage area of 100 to 300m	**BATTERY POWER CONSUMPTION –** reduce power consumption while not in use	**TRANSMISSION ROBUST AND SECURITY–** reliability and privacy/security
COLLOCATED NETWORK OPERATION – possible interference between LANs	**LICENSE-FREE OPERATION –** not having to secure a license for the frequency band used by the LAN	**HANDOFF/ROAMING–** enable stations to move from one cell to another
DYNAMIC CONFIGURATION- addition, deletion, relocation of end systems without disruption		

Wireless LAN Technologies

spread spectrum LANs

- mostly operate in ISM (industrial, scientific, and medical) bands
- no Federal Communications Commission (FCC) licensing is required in USA

OFDM LANs

- orthogonal frequency division multiplexing
- superior to spread spectrum
- operate in 2.4 GHz or 5 GHz band

infrared (IR) LANs

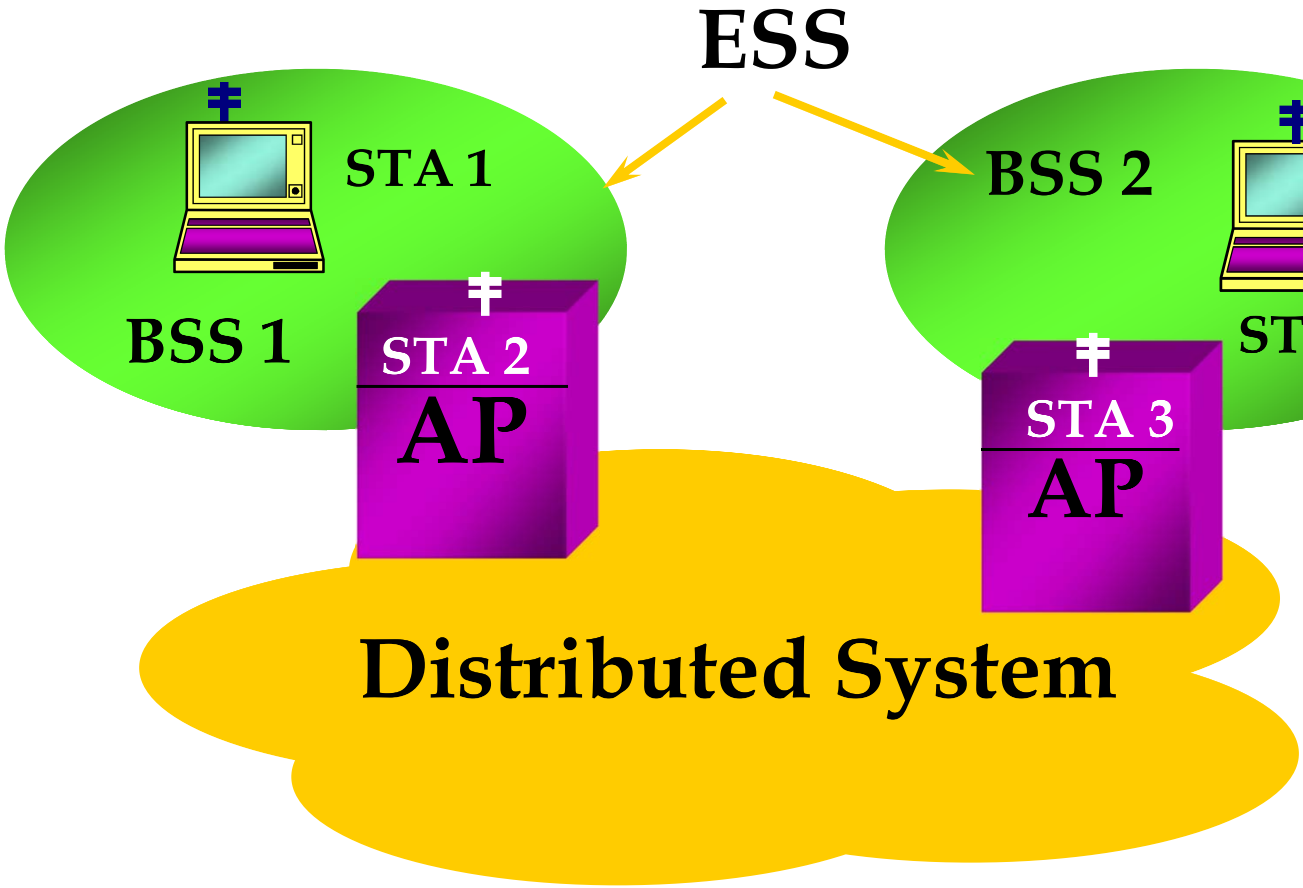
- individual cell of IR LAN limited to single room
- IR light does not penetrate opaque walls

Standard	Scope
IEEE 802.11	Medium access control (MAC): One common MAC for WLAN applications
	Physical layer: Infrared at 1 and 2 Mbps
	Physical layer: 2.4-GHz FHSS at 1 and 2 Mbps
	Physical layer: 2.4-GHz DSSS at 1 and 2 Mbps
IEEE 802.11a	Physical layer: 5-GHz OFDM at rates from 6 to 54 Mbps
IEEE 802.11b	Physical layer: 2.4-GHz DSSS at 5.5 and 11 Mbps
IEEE 802.11c	Bridge operation at 802.11 MAC layer
IEEE 802.11d	Physical layer: Extend operation of 802.11 WLANs to new regulatory domains (countries)
IEEE 802.11e	MAC: Enhance to improve quality of service and enhance security mechanisms
IEEE 802.11f	Recommended practices for multivendor access point interoperability
IEEE 802.11g	Physical layer: Extend 802.11b to data rates >20 Mbps
IEEE 802.11h	Physical/MAC: Enhance IEEE 802.11a to add indoor and outdoor channel selection and to improve spectrum and transmit power management
IEEE 802.11i	MAC: Enhance security and authentication mechanisms
IEEE 802.11j	Physical: Enhance IEEE 802.11a to conform to Japanese requirements
IEEE 802.11k	Radio resource measurement enhancements to provide interface to higher layers for radio and network measurements
IEEE 802.11m	Maintenance of IEEE 802.11-1999 standard with technical and editorial corrections
IEEE 802.11n	Physical/MAC: Enhancements to enable higher throughput
IEEE 802.11p	Physical/MAC: Wireless access in vehicular environments
IEEE 802.11r	Physical/MAC: Fast roaming (fast BSS transition)
IEEE 802.11s	Physical/MAC: ESS mesh networking
IEEE 802.11,2	Recommended practice for the Evaluation of 802.11 wireless performance
IEEE 802.11u	Physical/MAC: Interworking with external networks

IEEE 802.11 Standards

IEEE 802.11 only standardizes the physical and medium access control layers.

802.11 Architecture Components

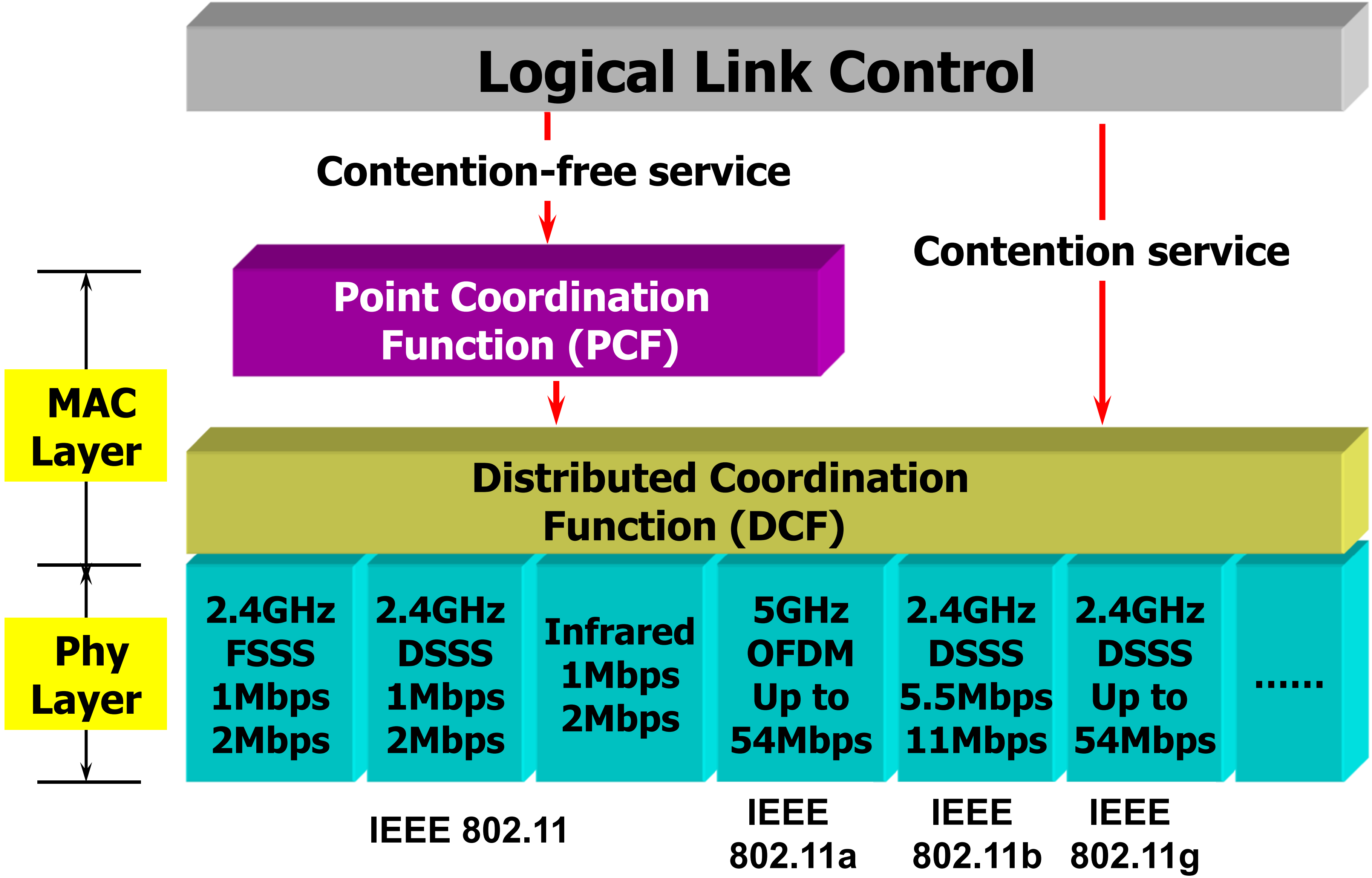


AP: Access Point

IEEE 802.11 Terminology

- **Access point (AP):** A station that provides access to the DS.
- **Basic service set (BSS):** A set of stations controlled by a single AP.
- **Distribution system (DS):** A system used to interconnect a set of BSSs to create an ESS.
  - DS is implementation-independent. It can be a wired 802.3 Ethernet LAN, or another 802.11 medium.
- **Extended service set (ESS):** Two or more BSS interconnected by DS
- **Portal:** Logical entity where 802.11 network integrates with a non 802.11 network.

Medium Access Control



Medium Access Control

- MAC layer covers three functional areas:
  - Reliable data delivery
  - Access control
  - Security

Reliable Data Delivery

- Loss of frames due to noise, interference, and propagation effects
- Frame exchange protocol
  - Source station transmits data
  - Destination responds with acknowledgment (ACK)
  - If source doesn't receive ACK, it retransmits frame
- Four frame exchange for enhanced reliability
  - Source issues request to send (RTS)
  - Destination responds with clear to send (CTS)
  - Source transmits data
  - Destination responds with ACK

Access Control

- **Distributed Coordination Function (DCF)**
  - **Distributed access protocol**
  - **Contention-Based**
  - **Makes use of CSMA/CA rather than CSMA/CD**
  - **Suited for ad hoc network and ordinary asynchronous traffic**
- **Point Coordination Function (PCF)**
  - **Alternative access method on top of DCF**
  - **Centralized access protocol**
  - **Contention-Free**
  - **Works like polling**
  - **Suited for time bound services like voice or multimedia**

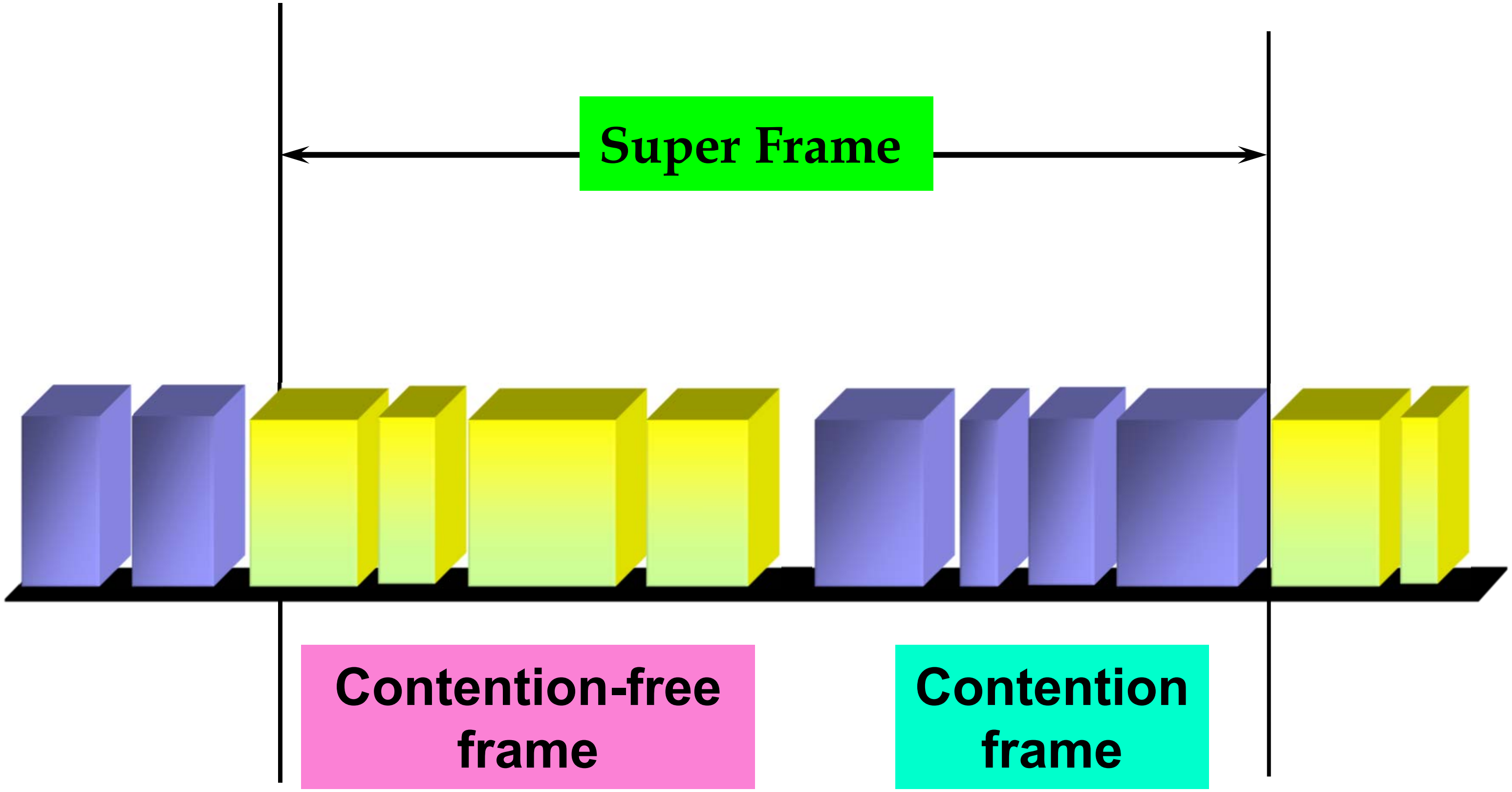
CSMA/CD vs. CSMA/CA

- **CSMA/CD – CSMA/Collision detection**
  - For wire communication
  - No control BEFORE transmission
  - Generates collisions
  - Collision Detection - How?
- **CSMA/CA – CSMA/Collision Avoidance**
  - For wireless communication
  - Collision avoidance BEFORE transmission
  - Why avoidance on wireless?
  - Difference in energy/power for transmit & receive
  - Difficult to distinguish between incoming weak signals, noise, and effects of own transmission

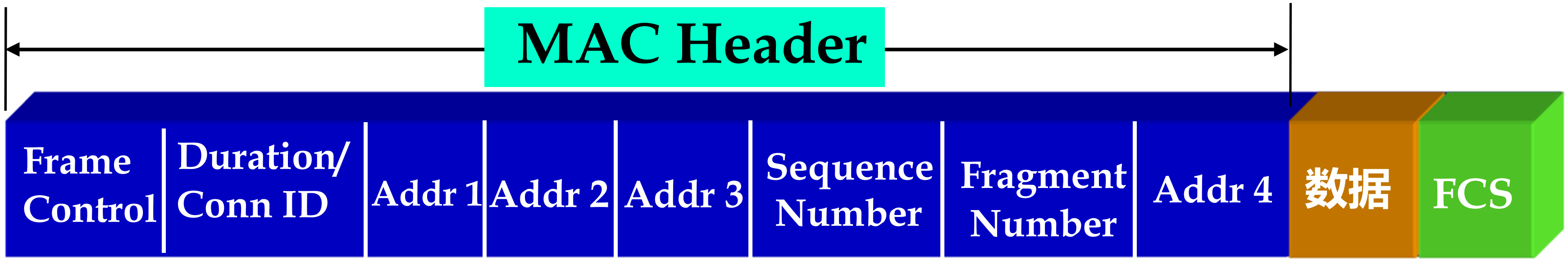
Access Control

- **Coexistence of DCF and PCF**
  - **Both the DCF and PCF shall coexist without interference.**
  - **They are integrated in a superframe in which a contention-free burst occurs at the beginning, followed by a contention period.**

Access Control



Data Frames



To DS	From DS	Addr 1	Addr 2	Addr 3	Addr 4
0	0	DA	SA	BSSID	N/A
0	1	DA	BSSID	SA	N/A
1	0	BSSID	SA	DA	N/A
1	1	RA	TA	DA	SA

BSSID： **The AP address, if the station is an AP or associated with an AP. BSS ID of the ad hoc LAN, if the station is a member of an ad hoc LAN**

802.11 Services

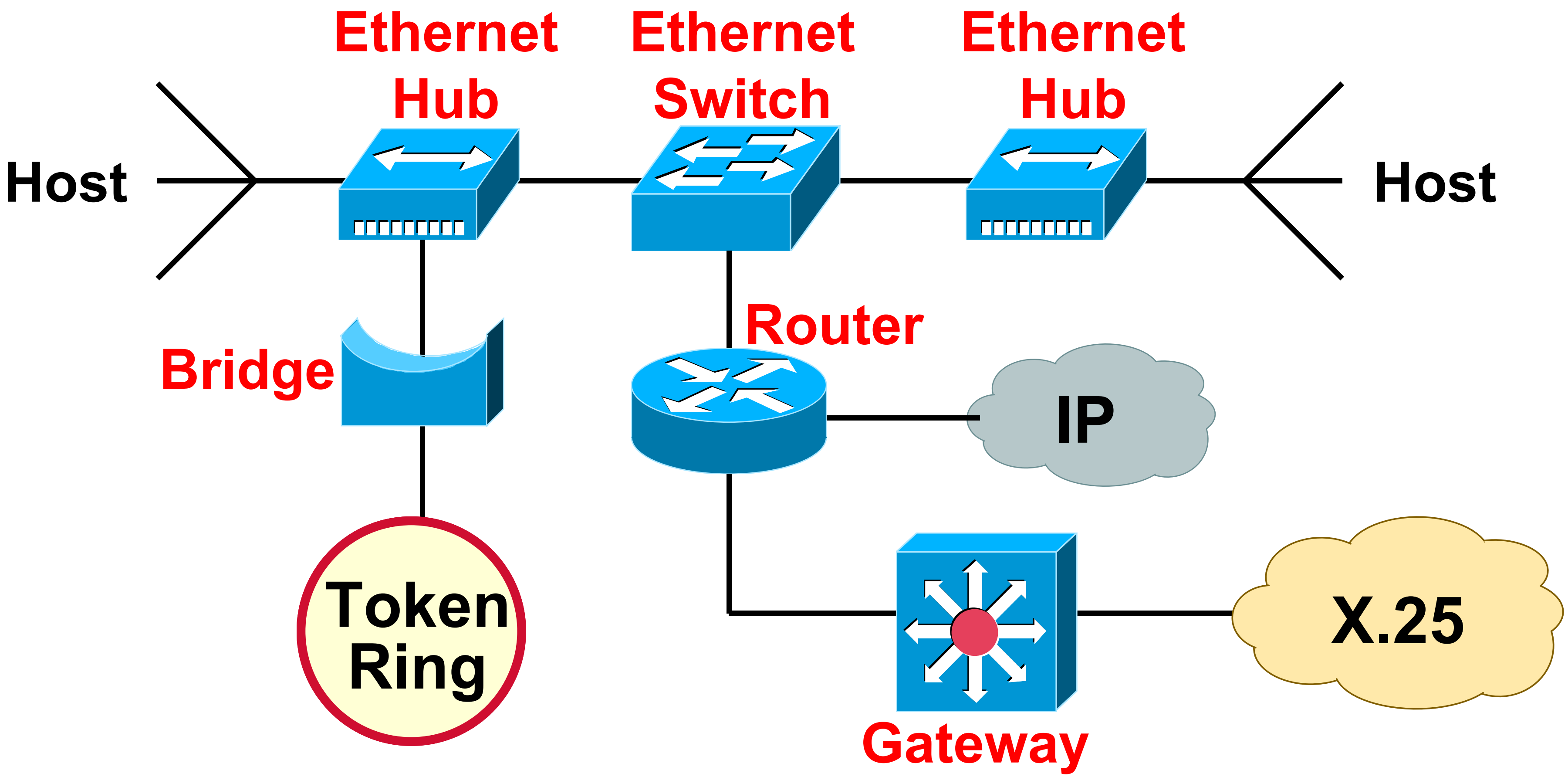
- Distribution Services
  - Association
  - Disassociation/Reassociation
  - Distribution
  - Integration
- Intracellular Services
  - Authentication / Deauthentication
  - Privacy
  - Data Delivery

Chapter 4: Roadmap

- Medium Access Control
- Local Area Networks (LANs) and IEEE 802
- Ethernet
- Wireless LAN
- LAN Interconnection
- LAN Switching
- VLAN

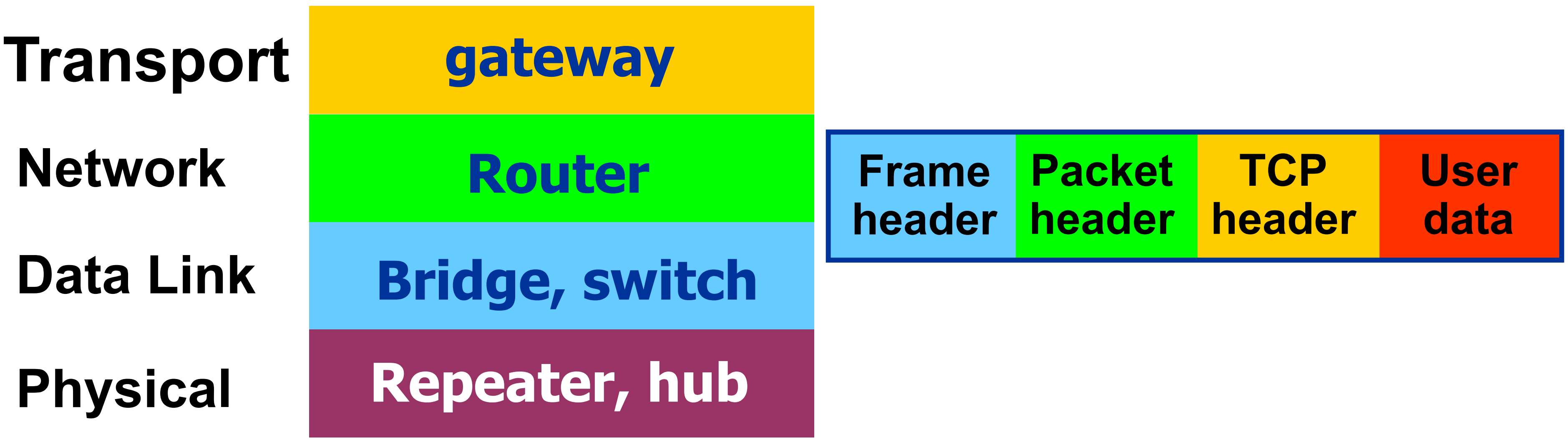
LAN Interconnection

- There are many different devices for interconnecting networks



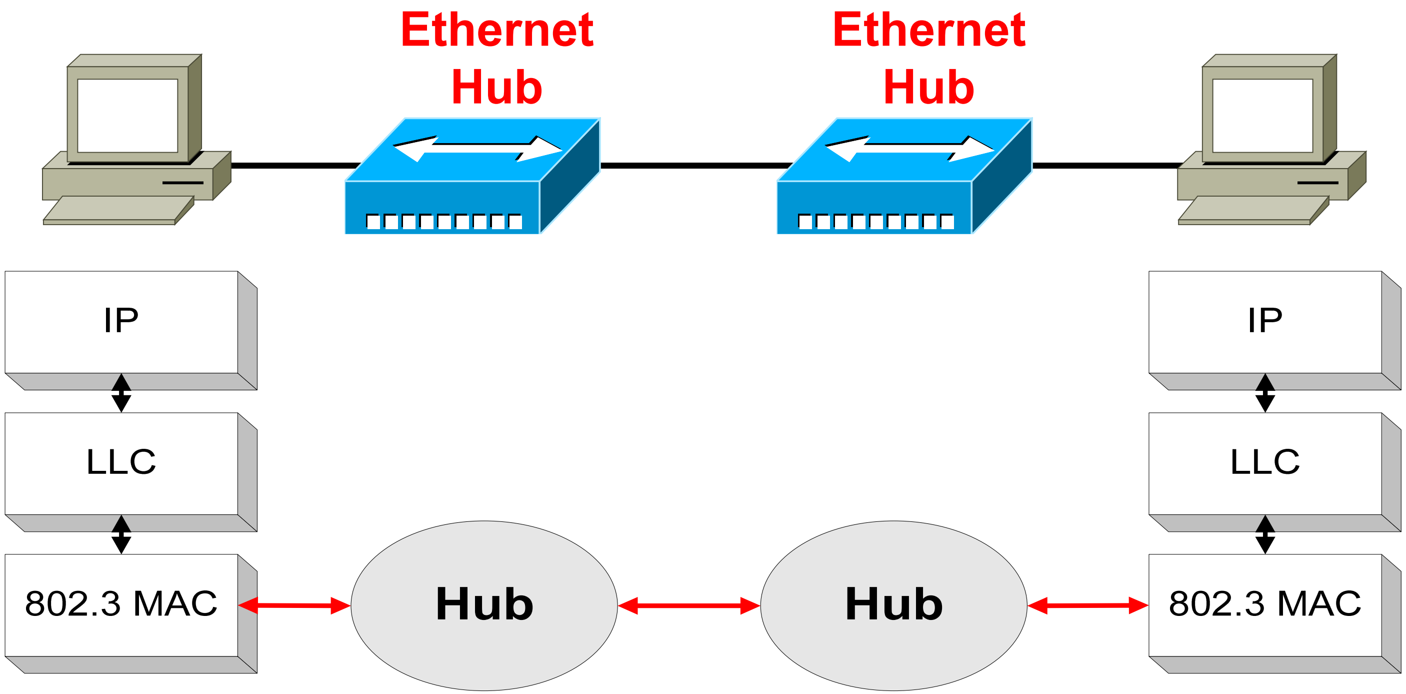
LAN Interconnection

- **Different devices switch different things**
  - **Physical layer:** electrical signals (repeaters and hubs)
  - **Link layer:** frames (bridges and switches)
  - **Network layer:** packets (routers)
  - **Transport and above layers:** message (gateways)

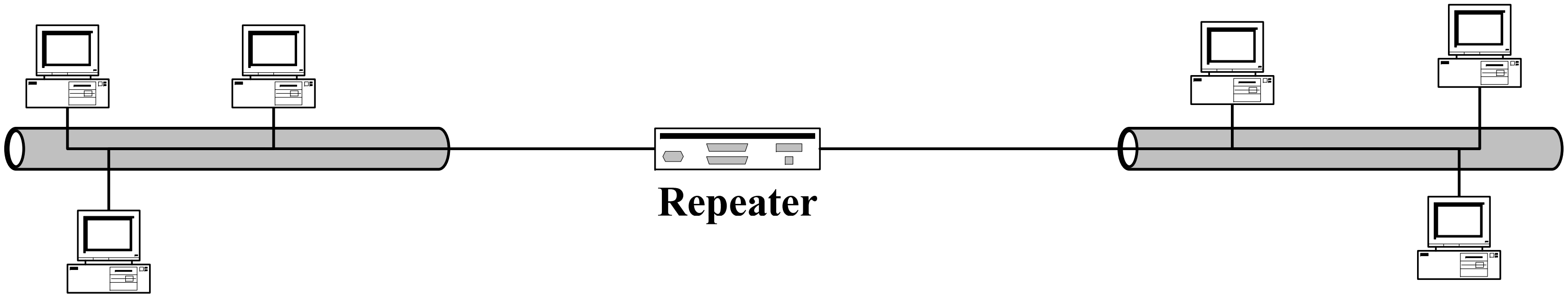


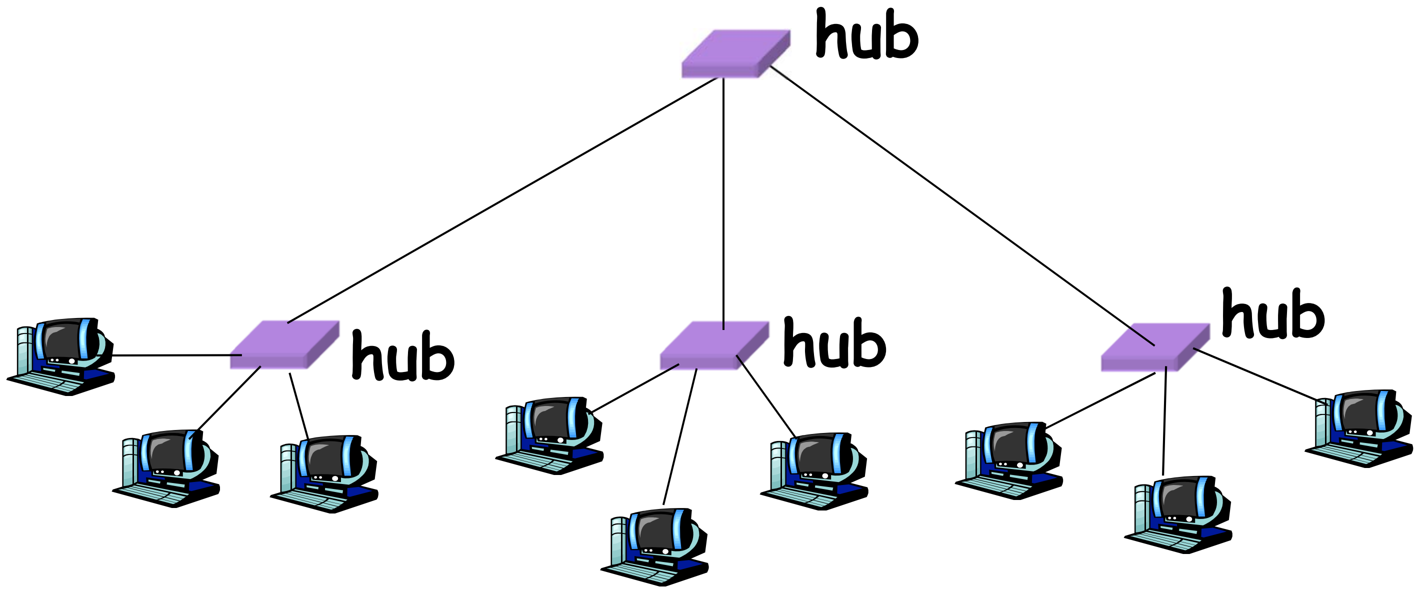
Hub/Repeater

- Operate at physical layer
- Used to connect hosts to Ethernet LAN and to connect multiple Ethernet LANs
- Collisions are propagated

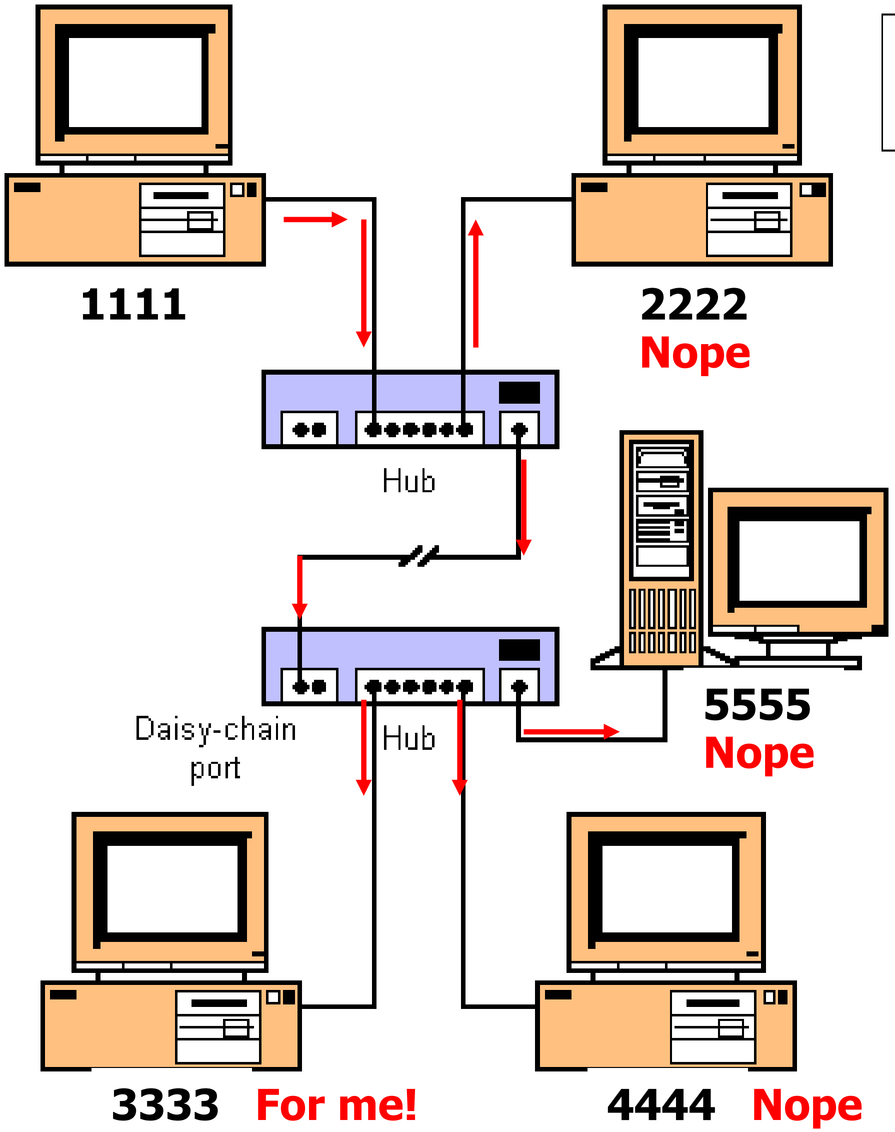


Interconnecting with Hub/Repeater





Interconnecting with Hub/Repeater



Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				

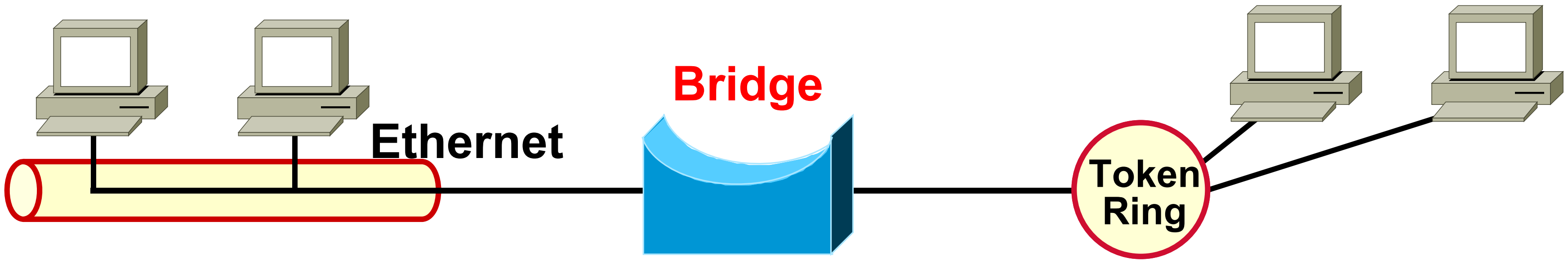
- The hub will **flood** it out all ports except for the incoming port.
- A hub or series of hubs is a single collision domain.

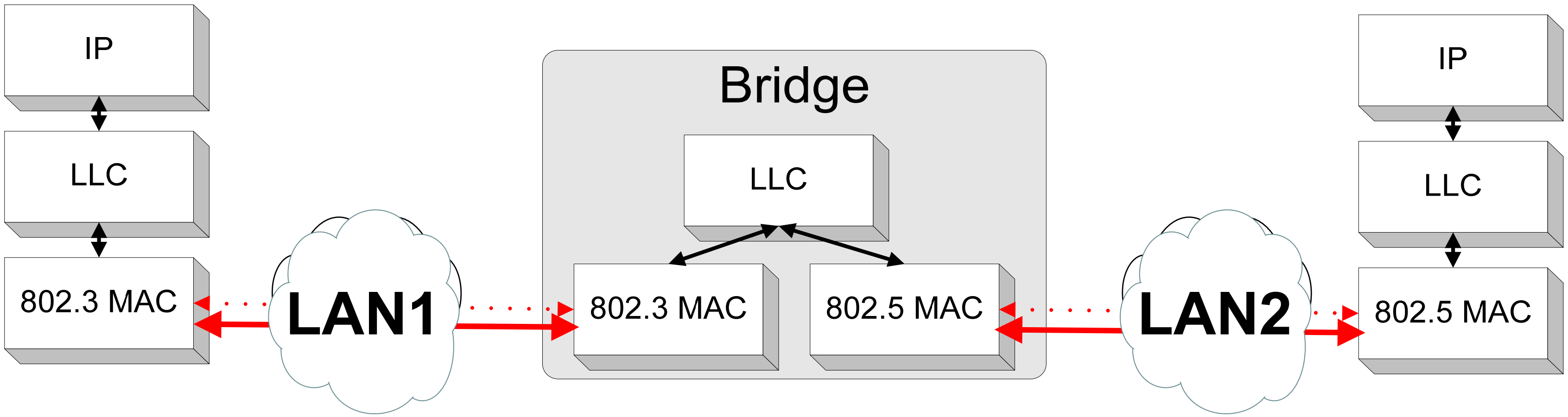
Limitations of Repeaters and Hubs

- **One large collision domain**
  - Every bit is sent everywhere
  - So, aggregate throughput is limited
  - E.g., three departments each get 10 Mbps independently
    - … and then connect via a hub and must share 10 Mbps
- **Cannot support different LAN technologies**
  - Does not buffer or interpret frames
  - So, can't interconnect between different rates or formats
    - E.g., 10 Mbps Ethernet and 100 Mbps Ethernet
- **Limitations on maximum nodes and distances**
  - Does not circumvent the limitations of shared media
    - E.g., still cannot go beyond 2500 meters on Ethernet

Bridges

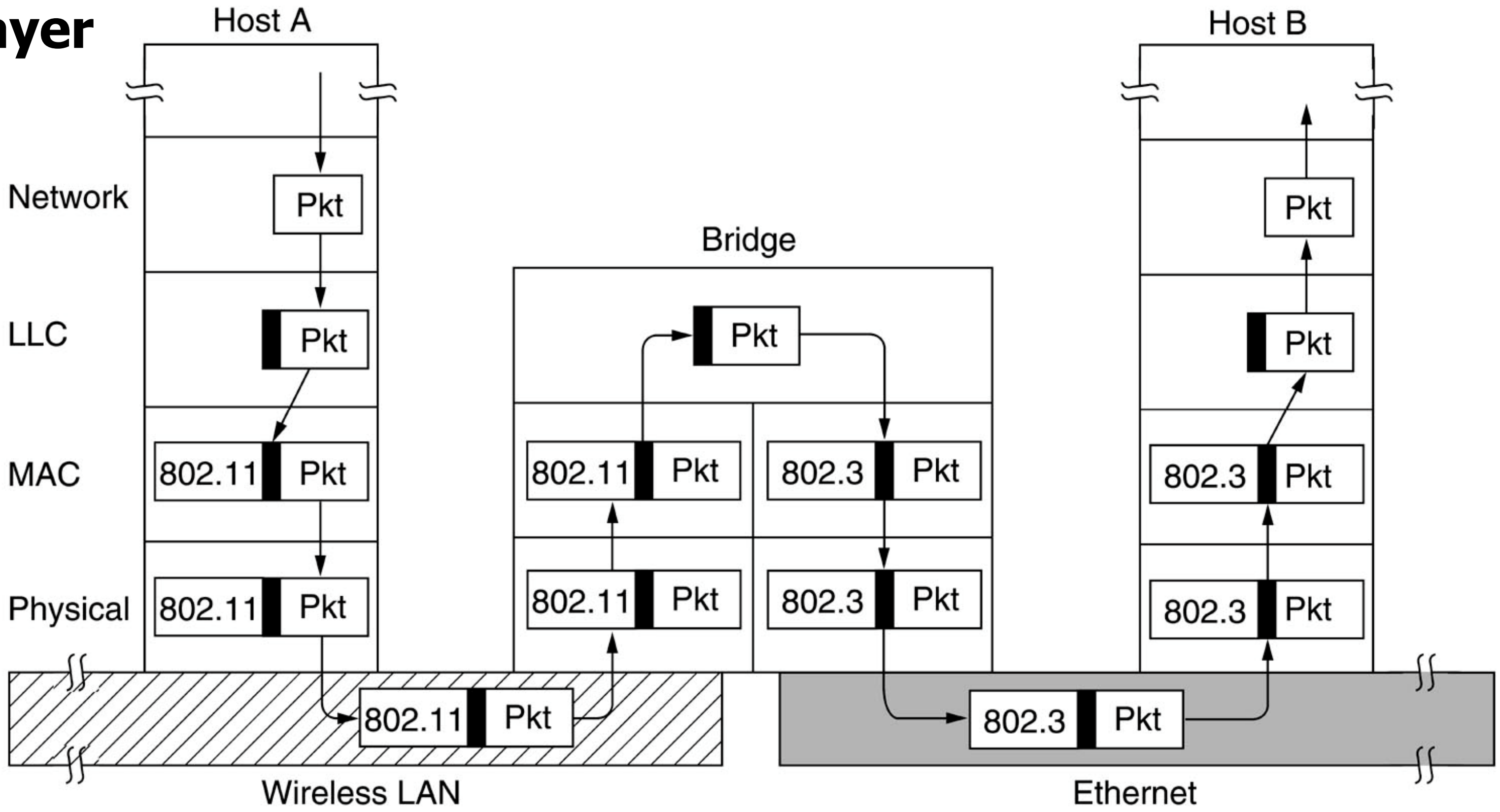
- **Operate at Data-Link layer (Layer 2)**
- **Interconnects two or more Local Area Networks (LANs) and forwards frames between these networks.**





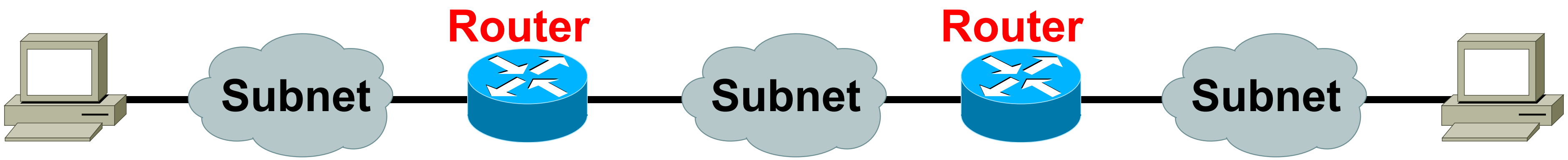
Bridges from 802.x to 802.y

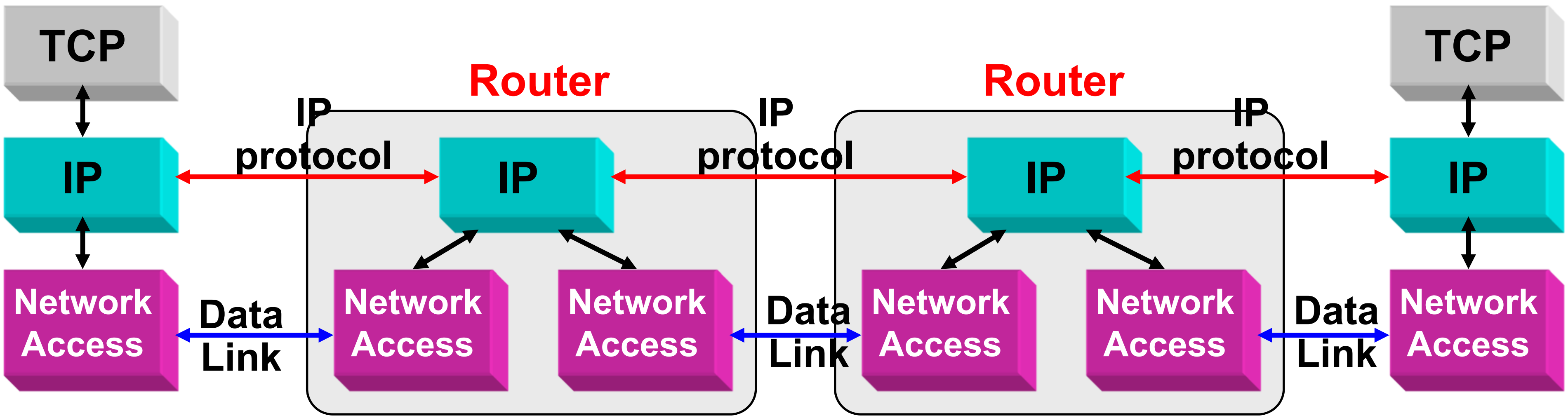
- **Principle Operations**
  - A packet is passed to the data link layer (LLC part)
  - It is then passed to the MAC layer (specific access strategy)
  - A bridge converts the stuff above the MAC layer, in the LLC layer



Routers

- **Operate at the Network Layer (Layer 3)**
- **Interconnect different subnetworks**





Routers

- Packet forwarding
- Packet filtering
- Packet switching (Routing)
- Traffic management
- QoS
- ...

Not transparent to hosts !

Gateways

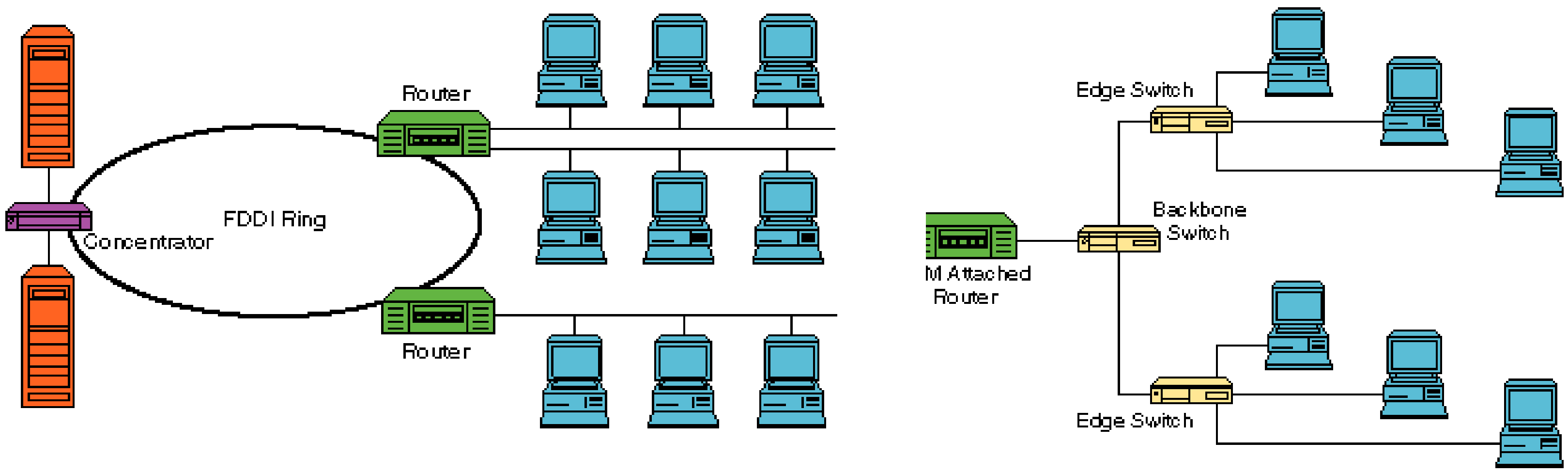
- Different meanings in different contexts:
  - a generic term for **routers** (Level 3)
  - also used for a device that interconnects different Layer 3 networks and which performs translation of protocols ("**Multi-protocol router**")

Gateways

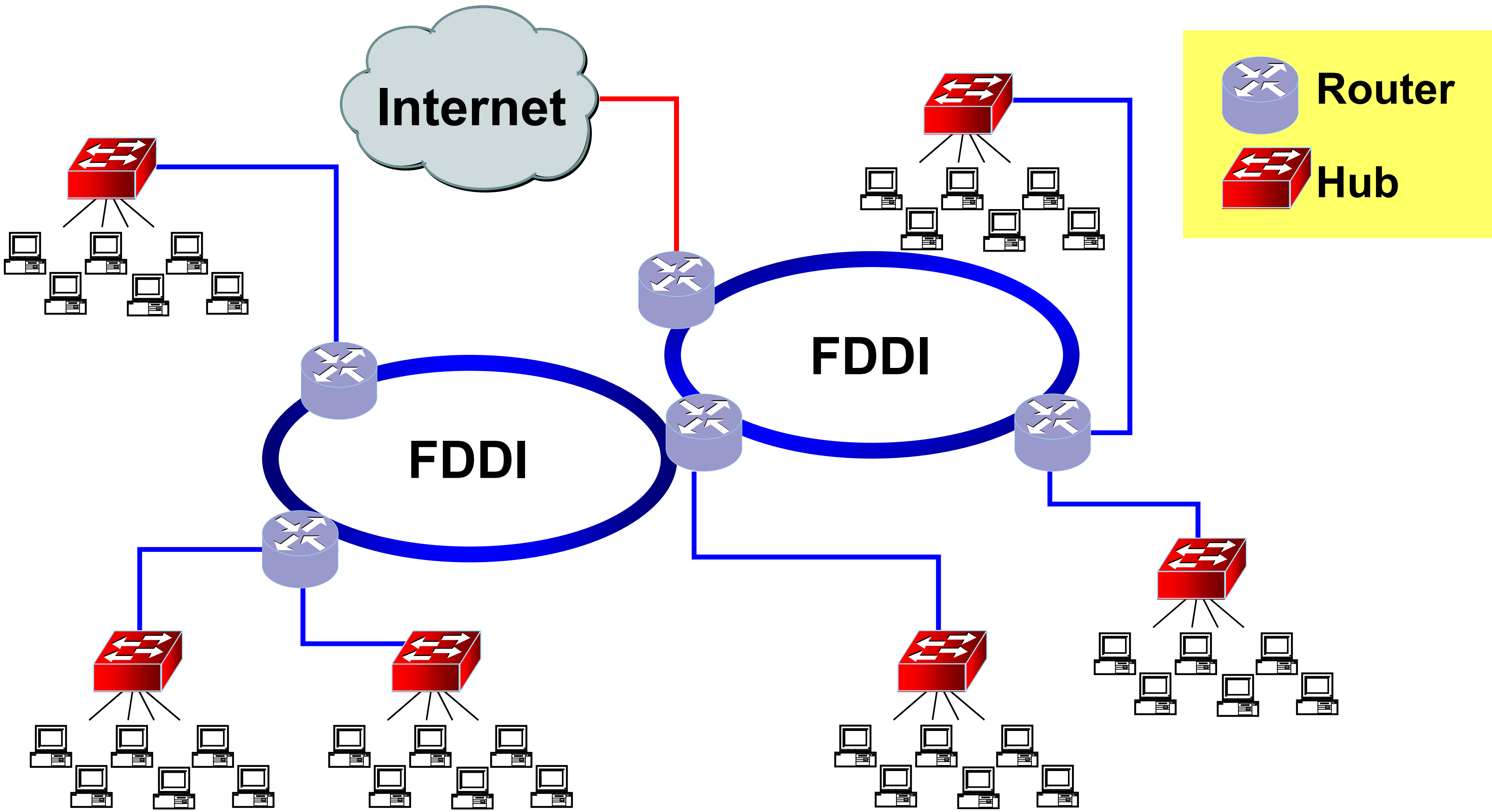
- 功能：
  - 报文格式转换
  - 地址映射
  - 网络协议转换
  - 原语连接转换
  - 连接不同体系结构的网络

Bridges/Switches versus Routers

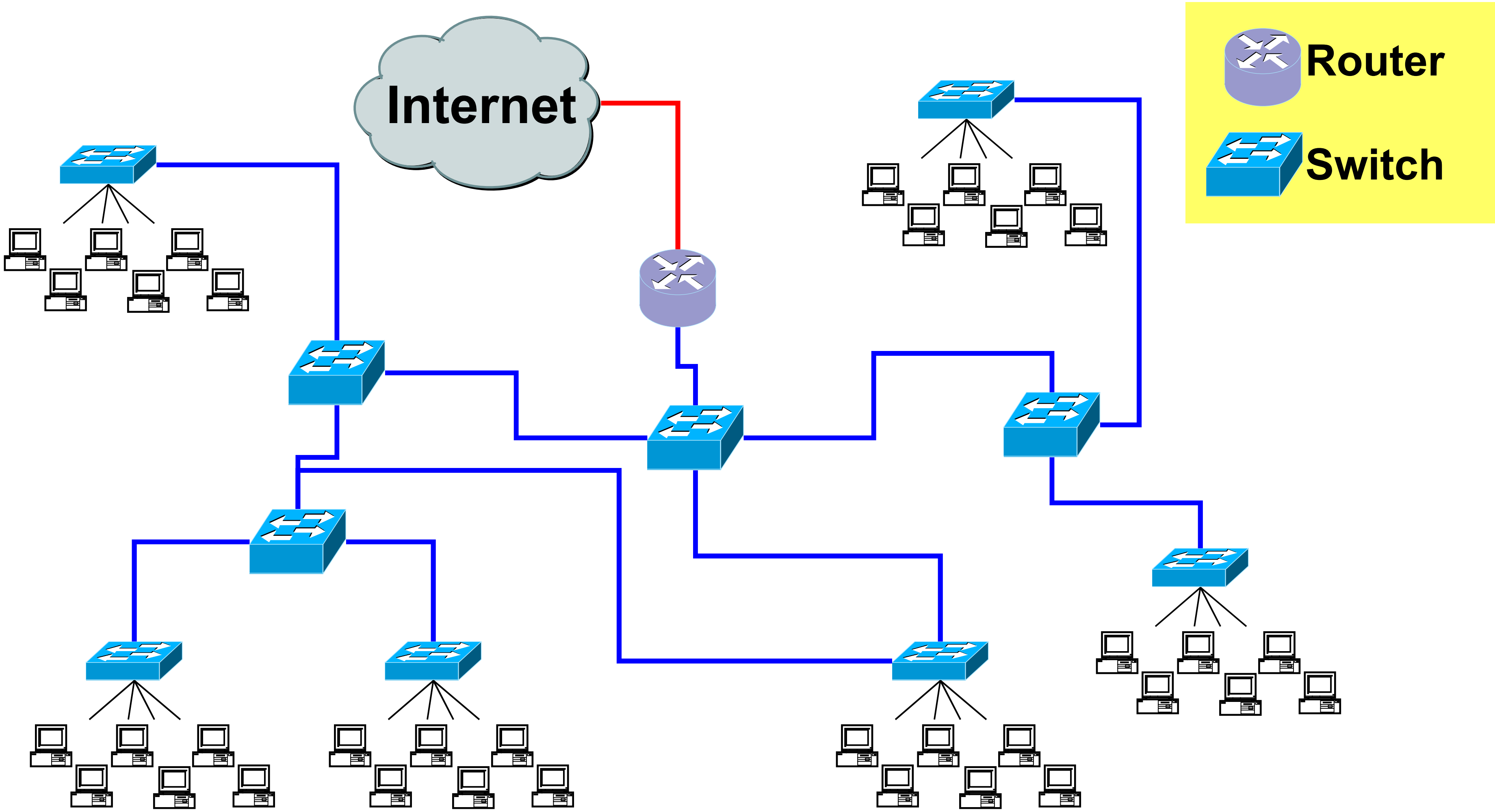
- An enterprise network (e.g., university network) with a large number of local area networks (LANs) can use routers or bridges
- Until early 1990s: most LANs were interconnected by routers
- Since mid1990s: LAN switches replace most routers



A Routed Enterprise Network



A Switched Enterprise Network



Bridges/Switches versus Routers

Routers

- Each host's IP address must be configured
- If network is reconfigured, IP addresses may need to be reassigned
- Routing done via RIP or OSPF
- Each router manipulates packet header (e.g., reduces TTL field)

Bridges/Switches

- MAC addresses are hardwired
- No network configuration needed
- No routing protocol needed (sort of)
  - learning bridge algorithm
  - spanning tree algorithm
- Bridges do not manipulate frames

Chapter 4: roadmap

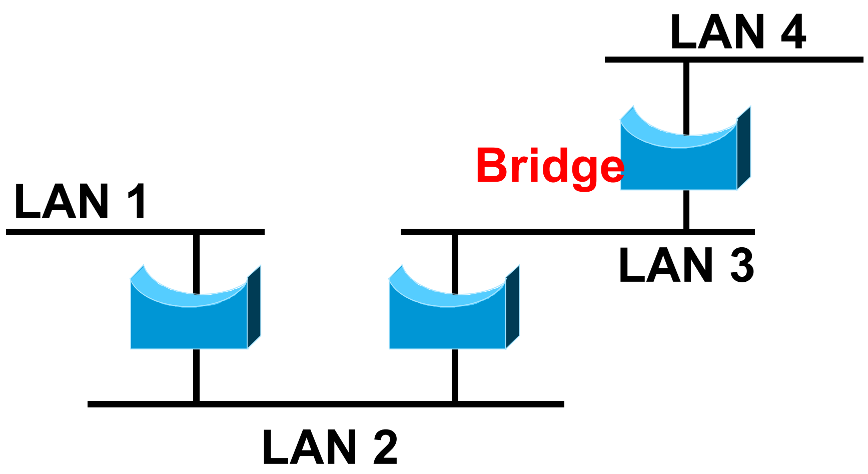
- Medium Access Control
- Local Area Networks (LANs) and IEEE 802
- Ethernet
- Wireless LAN
- LAN Interconnection
- LAN Switching
- VLAN

LAN Switching

- Traditional LAN
  - Shared medium (e.g., Ethernet)
  - Cheap, easy to administer
  - Supports broadcast traffic
- Problem
  - Scale
    - Larger geographic area (> O(1 km))
    - More hosts (> O(100))
  - But retain LAN-like functionality
- Solution
  - Bridges/Switches

Bridges

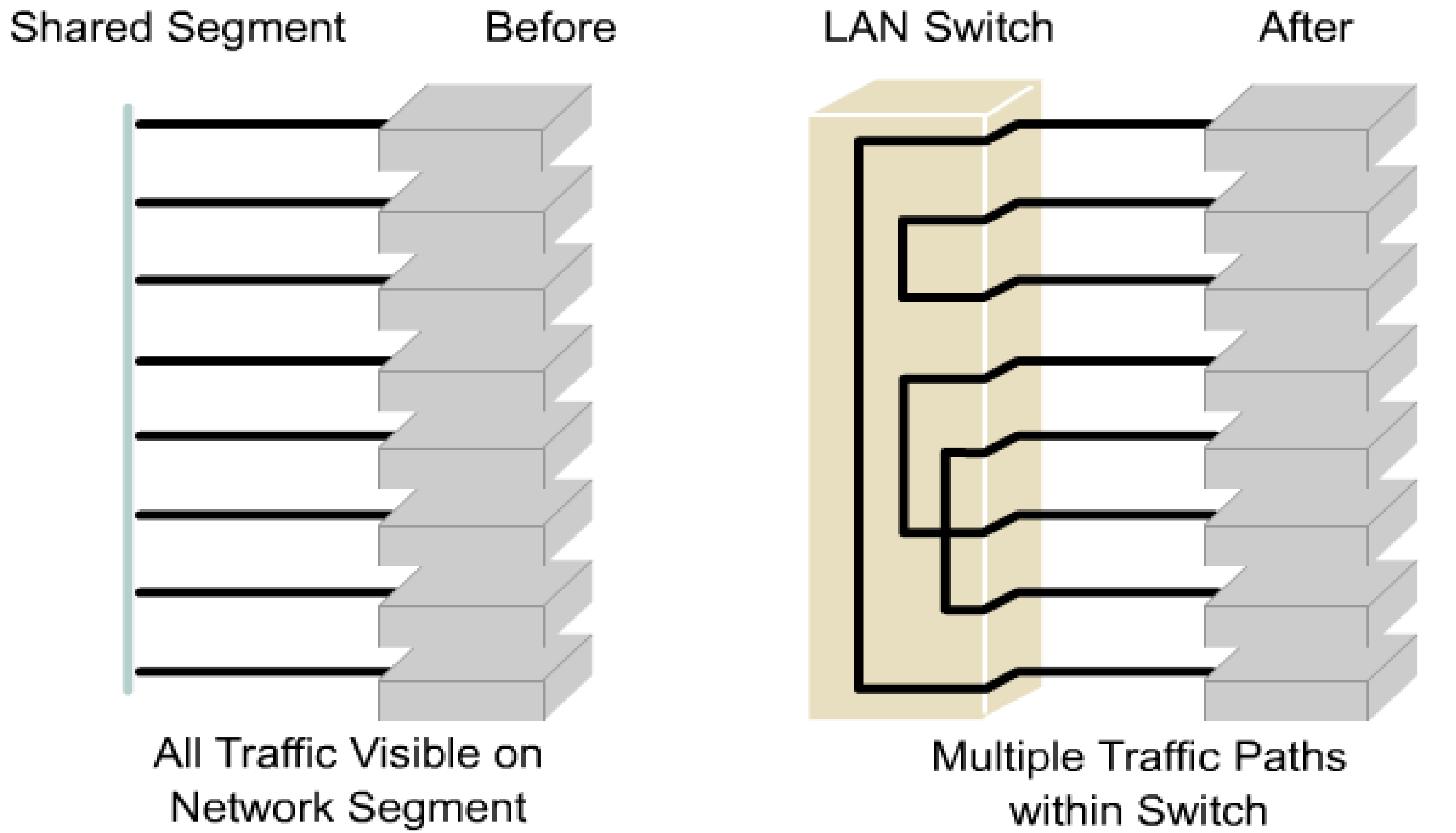
- Connect two or more LANs
  - Accept and forward
  - Level 2 connection (no extra packet header)
  - Each LAN is its own collision domain
- A collection of LANs connected by bridges is called an extended LAN



Bridges vs. Switches

- **Bridge**
  - Connect shared media
  - All ports bidirectional
  - Limited scalability
  - Slow and Expensive
- **Link layer switch**
  - Connects hosts and/or shared media
  - Many interfaces
  - Hardware-based switching fabric

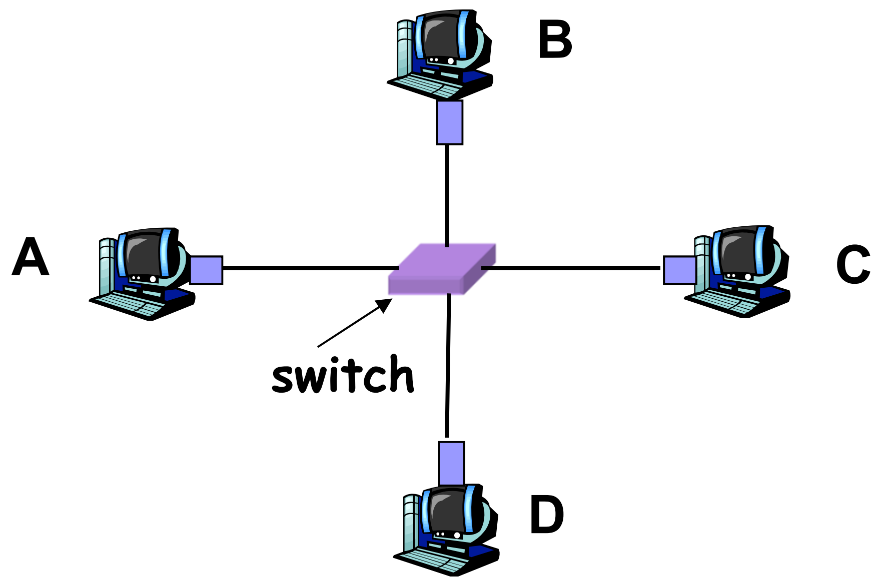
Switched Fabric



Switches: dedicated access

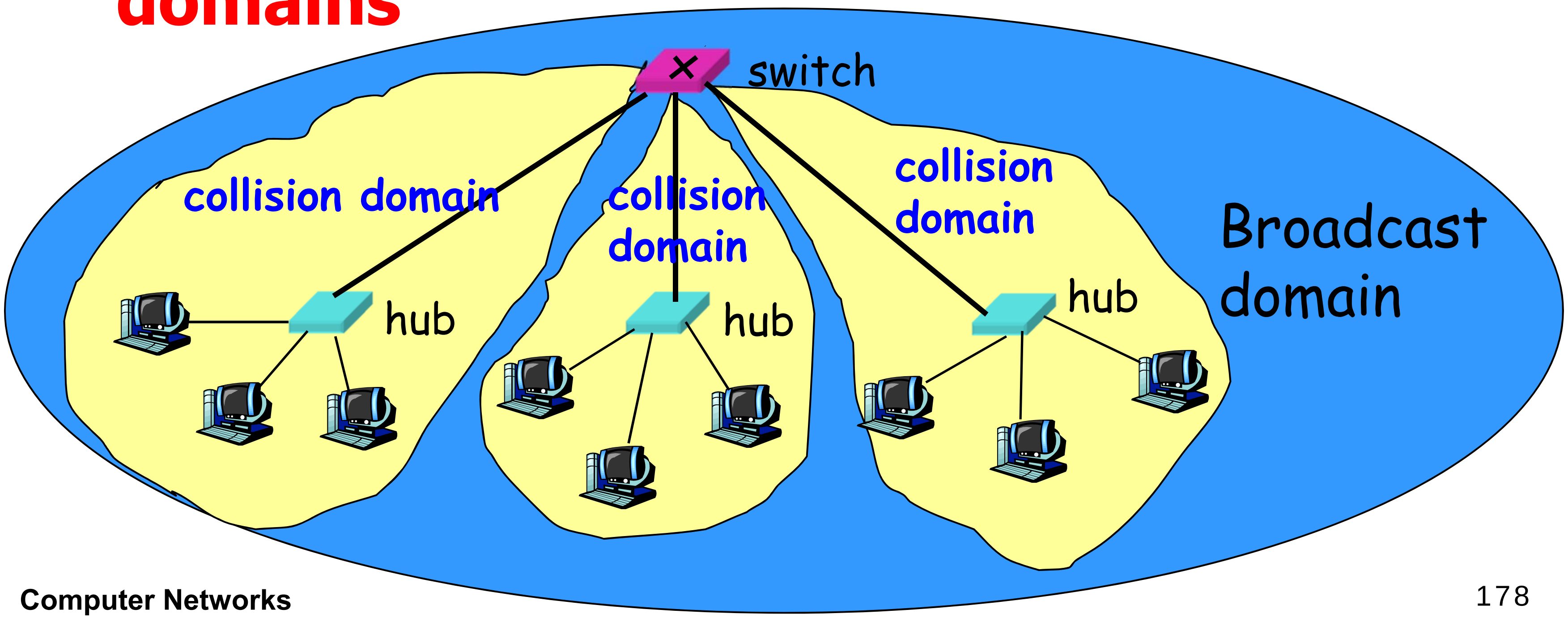
- With many interfaces
- Hosts have direct connection to switch
- full duplex
- No collisions;

Switching:
A-to-C and B-to-D simultaneously, no collisions.



Switch: traffic isolation

- Switch breaks subnet into LAN segments
- Switch filters frames:
  - same-LAN-segment frames not usually forwarded onto other LAN segments
  - segments become separate collision domains



Advantages Over Hubs/Repeaters

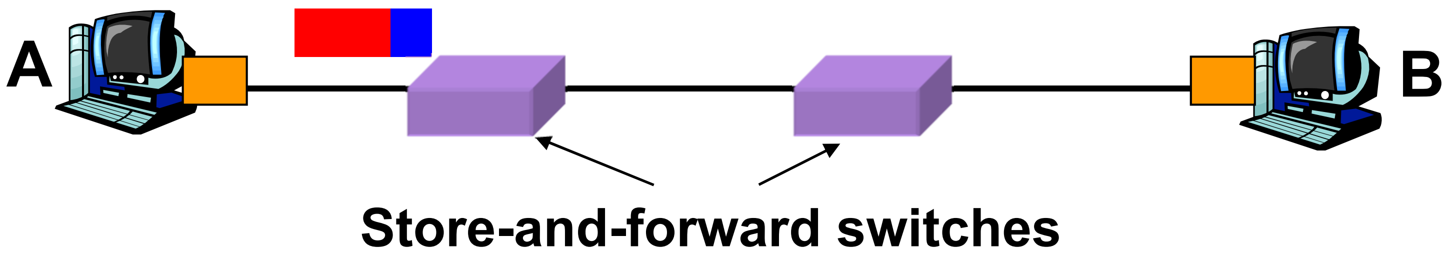
- **Only forwards frames as needed**
  - Filters frames to avoid unnecessary load on segments
  - Sends frames only to segments that need to see them
- **Extends the geographic span of the network**
  - Separate collision domains allow longer distances
- **Improves privacy by limiting scope of frames**
  - Hosts can "snoop" the traffic traversing their segment
  - … but not all the rest of the traffic
- **Applies carrier sense and collision detection**
  - Does not transmit when the link is busy
  - Applies exponential back-off after a collision
- **Joins segments using different technologies**

Disadvantages Over Hubs/Repeaters

- **Delay in forwarding frames**
  - Switch must receive and parse the frame
  - … and perform a look-up to decide where to forward
  - Storing and forwarding the packet introduces delay
  - **Solution: cut-through switching**
- **Need to learn where to forward frames**
  - Switch needs to construct a forwarding table
  - Ideally, without intervention from network administrators
  - Solution: self-learning
- **Higher cost**
  - More complicated devices that cost more money

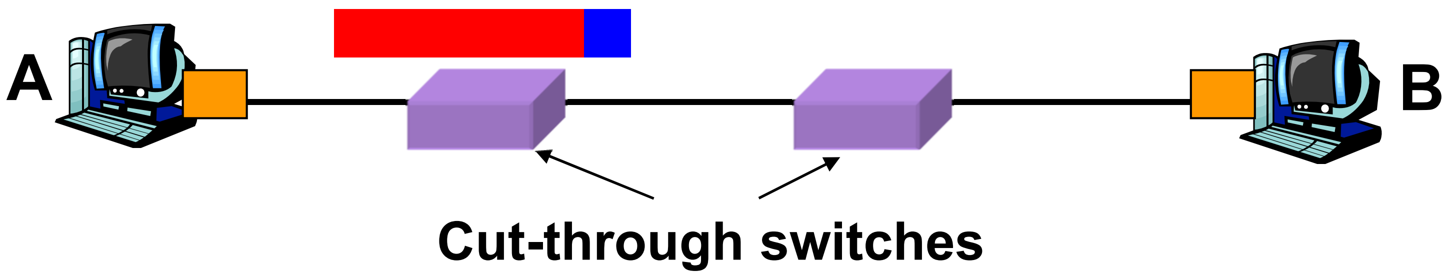
Motivation For Cut-Through Switching

- **Buffering a frame takes time**
  - **Suppose L is the length of the frame, and R is the transmission rate of the links**
  - **Then, receiving the frame takes L/R time units**
- **Buffering delay can be a high fraction of total delay**
  - **Propagation delay is small over short distances**

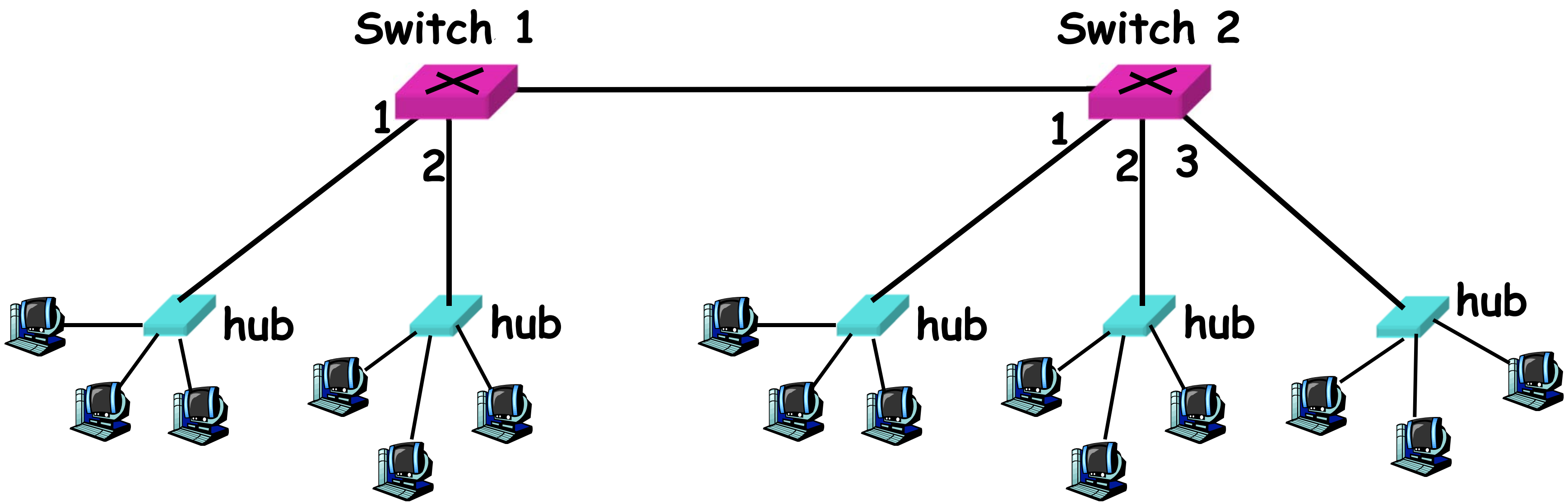


Cut-Through Switching

- Start forward transmission as soon as possible
  - Inspect the frame header and do the look-up
  - If outgoing link is idle, start forwarding the frame
- Overlapping transmissions
  - Transmit the head of the frame via the outgoing link,
  - … while still receiving the tail via the incoming link
  - Delay: head of the frame



Forwarding



- **Question:**
  How do determine onto which LAN segment to forward frame?

Self learning

- A switch has a switch table (Forwarding database, Forwarding table, MAC table)
- Entry in switch table:

  (MAC Address, Port, Age)

  MAC address: host name or group address

  port: port number of switch / bridge

  age: aging time of entry (stale entries in table dropped)

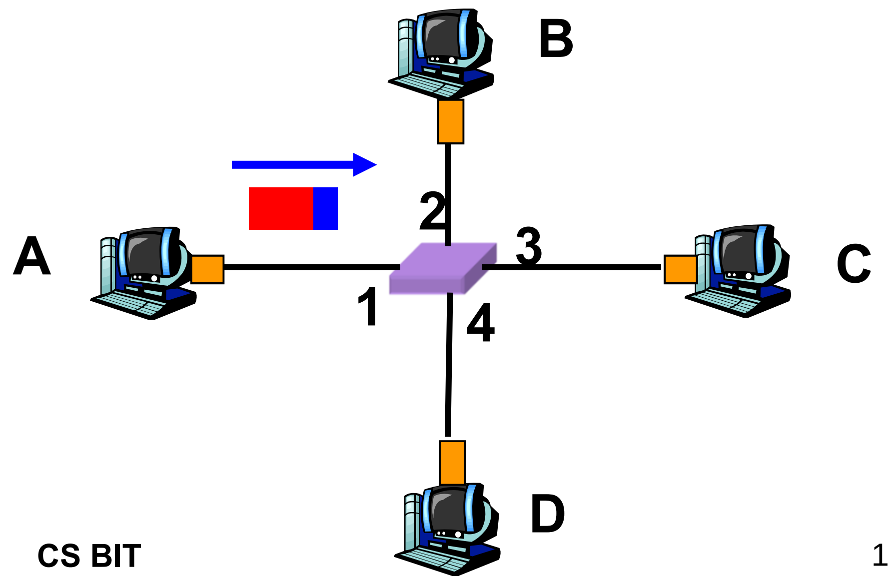
Interpretation:

a machine with MAC address lies in direction of the port number from the bridge. The entry is age time units old.

Self Learning: Building the Table

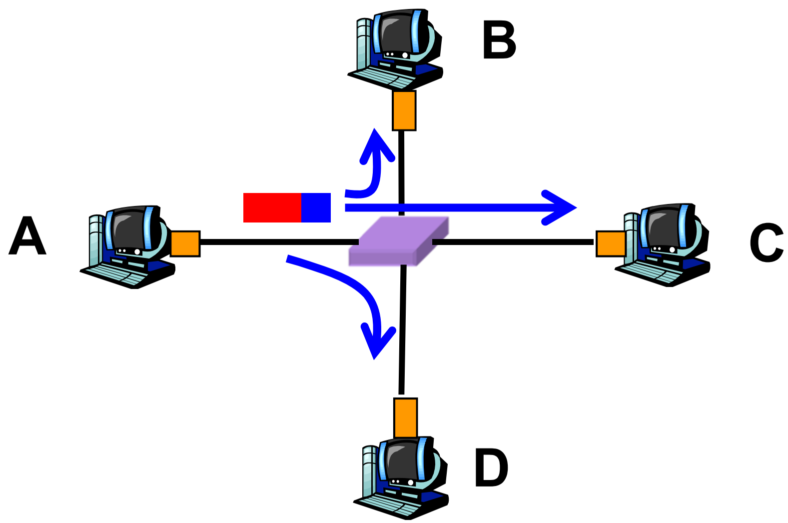
- **When a frame arrives**
  - Inspect the ***source*** **MAC address**
  - Associate the address with the ***incoming*** **interface**
  - Store the mapping in the switch table
  - Use a time-to-live field to eventually forget the mapping

Switch learns how to reach A.



Self Learning: Handling Misses

- **Miss:** output port to destination is not in switch table
- When frame arrives with unfamiliar destination, forward the frame out all of the interfaces
  - except for the one where the frame arrived



Filtering/Forwarding

When switch receives a frame:

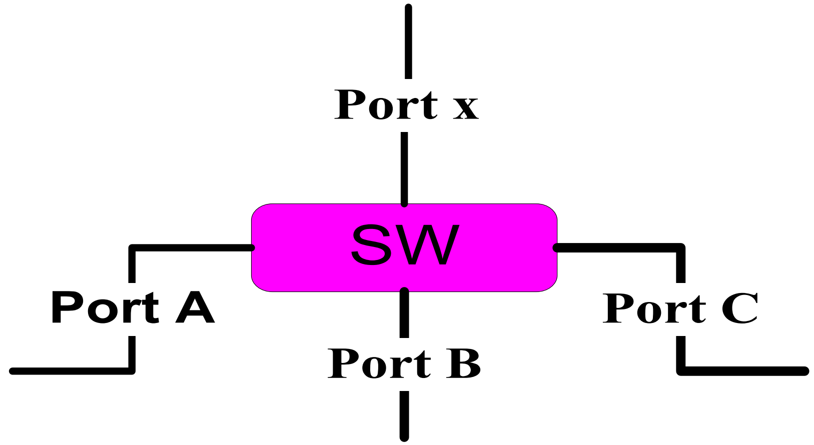
index switch table using MAC dest address

if entry found for destination
then{
if dest on segment from which frame arrived
then drop the frame
else forward the frame on interface indicated
}
else flood

forward on all but the interface on which the frame arrived

Filtering/Forwarding

- **Assume a MAC frame arrives on port x.**



Is MAC address of destination in forwarding database for ports A, B, or C ?

Found? → **Forward the frame on the appropriate port**

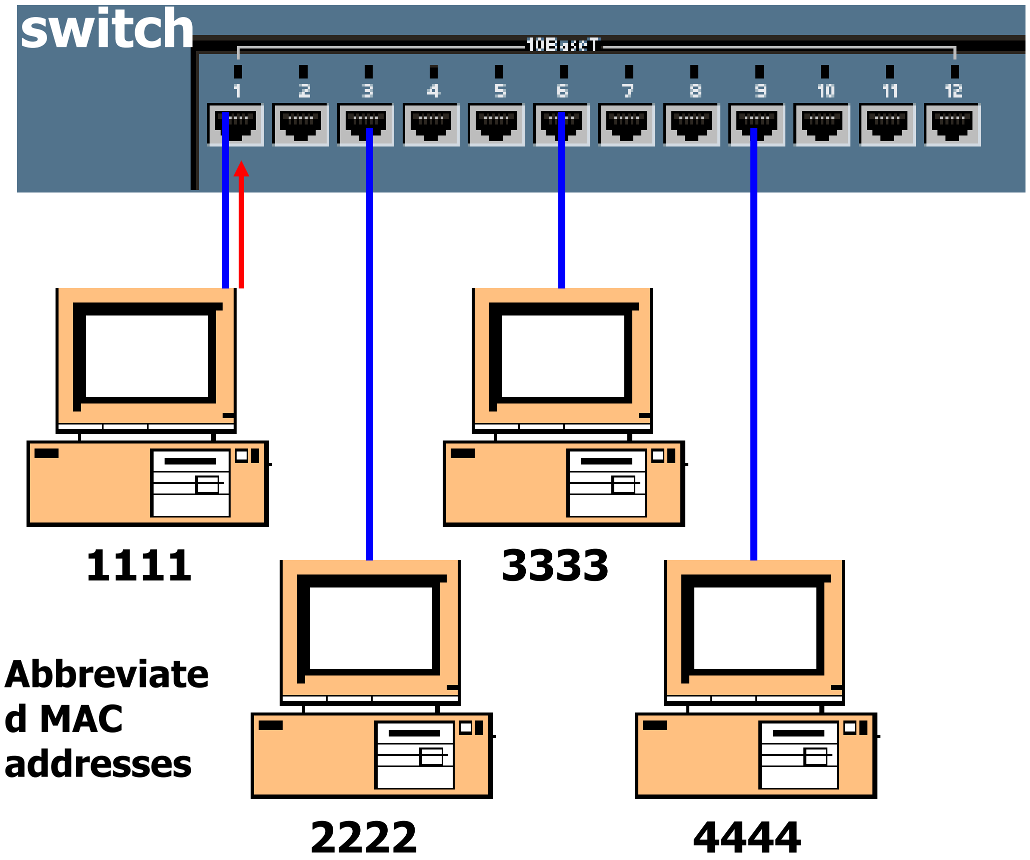
Not found ? → **Flood the frame, i.e., send the frame on all ports except port x.**

Sending and receiving Ethernet frames via a switch

Source Address Table

Port	Src. MAC Add.	Port	Src. MAC Add.

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				



Abbreviated MAC addresses

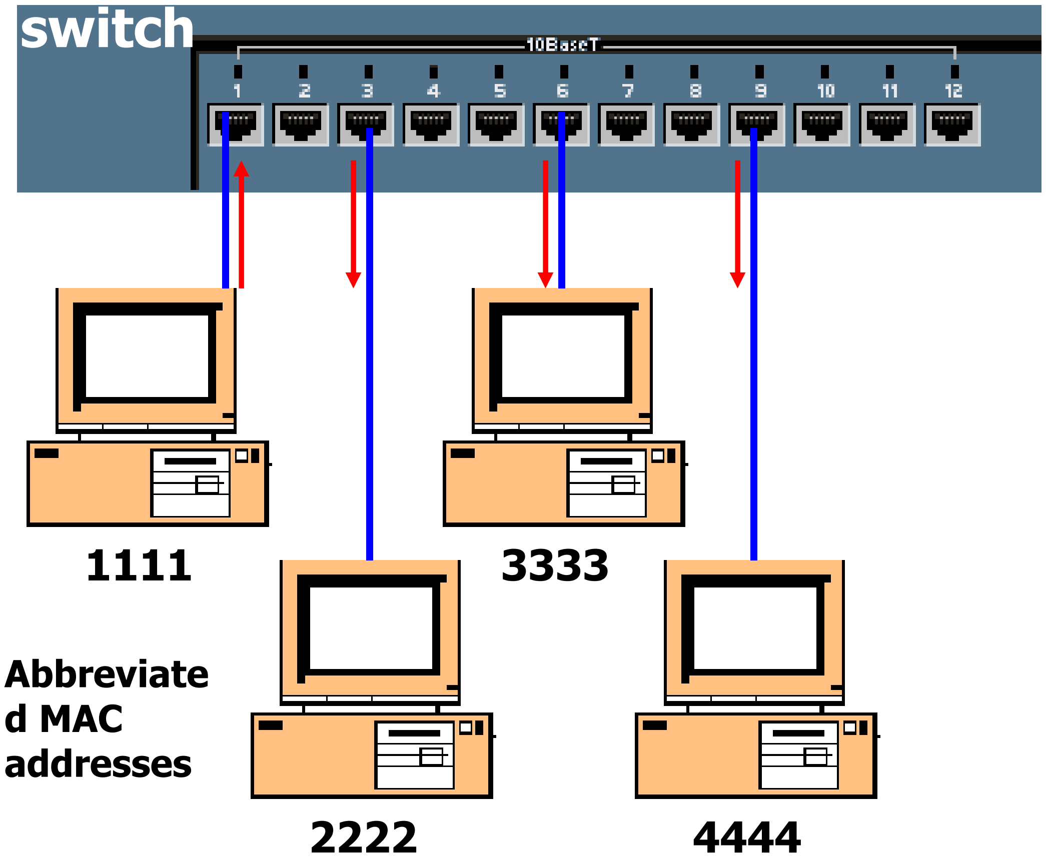
- Switches are also known as **learning bridges** or learning switches.
- A switch receives an Ethernet frame it searches the source address table for the Destination MAC address.

No Destination Address in table, Flood

Source Address Table

Port	Src. MAC Add.	Port	Src. MAC Add.
1	1111		

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				



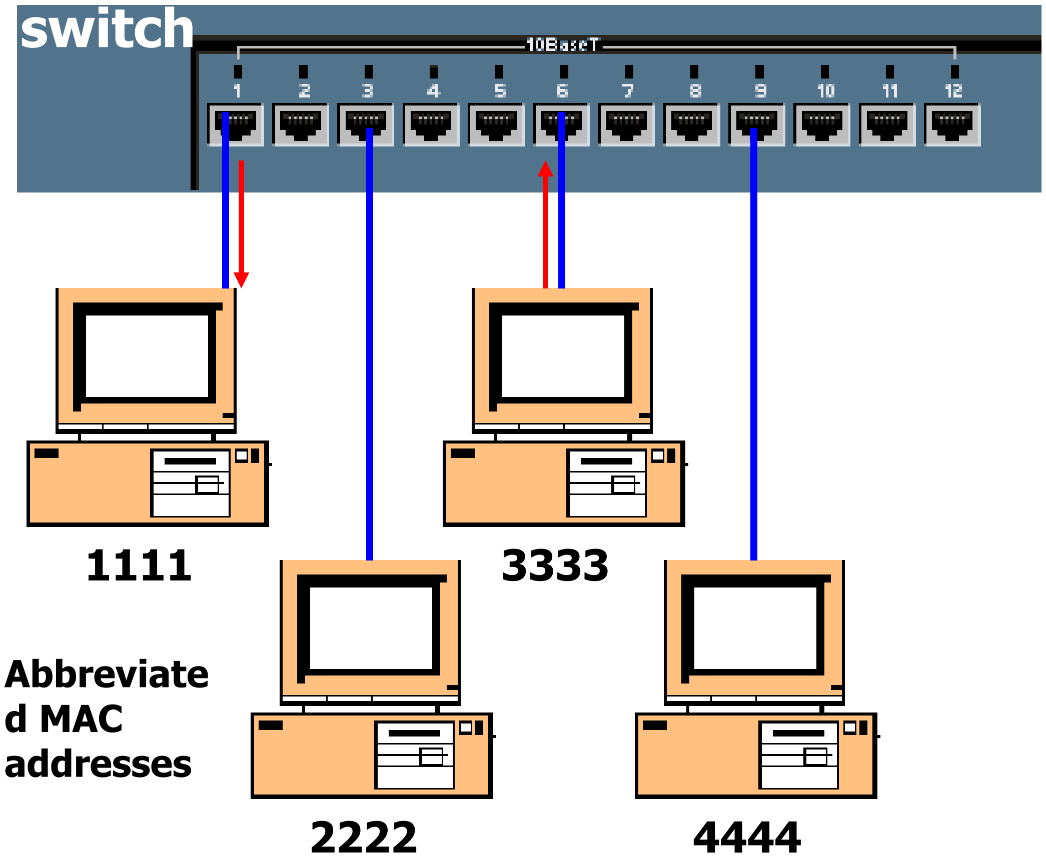
- If the SA (1111) is in it's table, it resets the timer (more in a moment).
- If it is NOT in the table it adds it, with the port number.
- Next, the switch will flood the frame out all other ports, because the DA is not in the source address table.

Destination Address in table, Filter

Source Address Table

Port	Src. MAC Add.	Port	Src. MAC Add.
1	1111	6	3333

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	1111	3333				



Abbreviated MAC addresses

- Now 3333 sends data back to 1111.
- The switch sees if it has the SA stored.
- It does NOT so it adds it.
- Next, it checks the DA and in our case it can filter the frame, by sending it only out port 1.

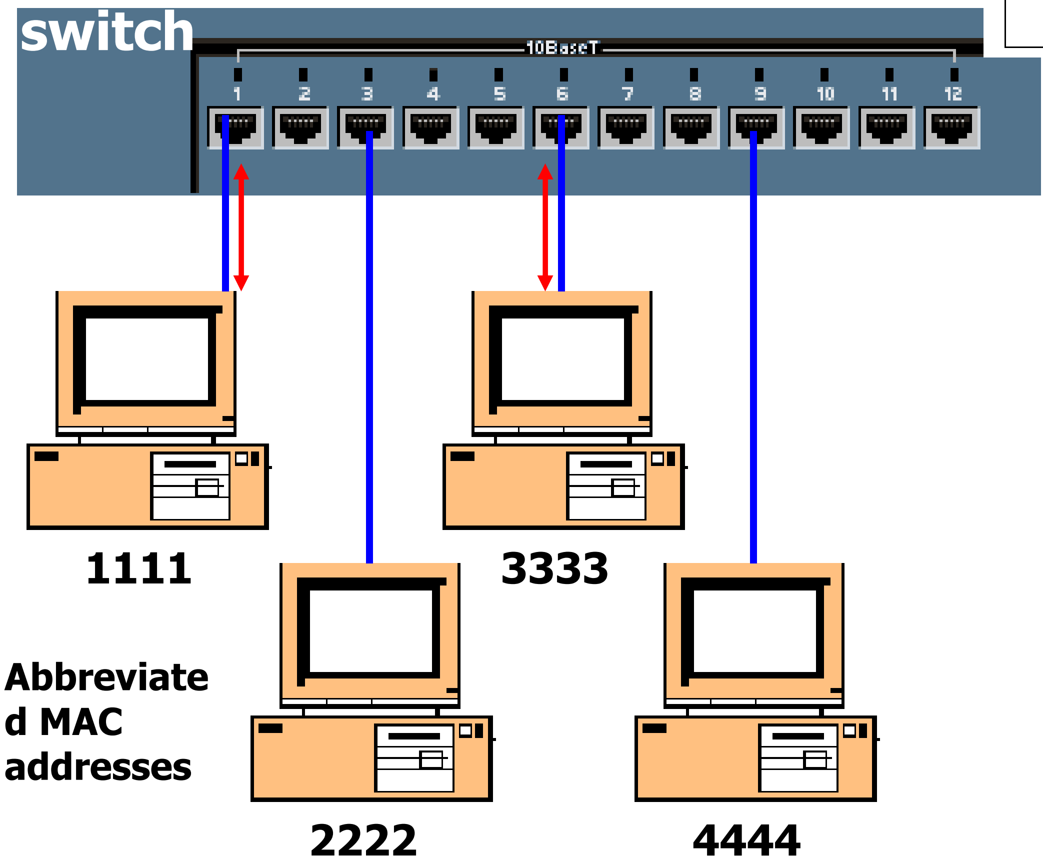
Destination Address in table, Filter

Source Address Table

Port	Src. MAC Add.	Port	Src. MAC Add.
1	1111	6	3333

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	1111	3333				



Abbreviated MAC addresses

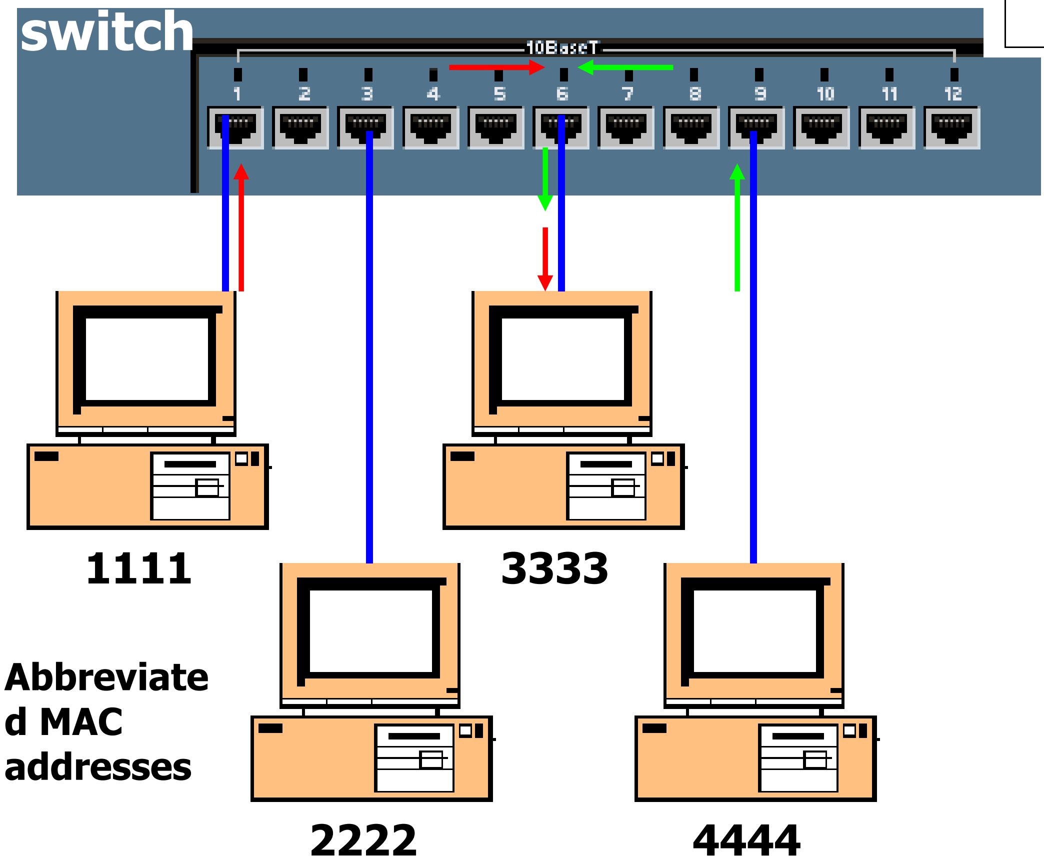
- **Question:**
- **What happens when two devices send to same destination?**
- **What if this was a hub?**
- **Where is (are) the collision domain(s) in this example?**

No Collisions in Switch, Buffering

Source Address Table			
Port	Src. MAC Add.	Port	Src. MAC Add.
1	1111	6	3333
9	4444		

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	4444				



Abbreviated MAC addresses

- Unlike a hub, a collision does NOT occur, which would cause the two PCs to have to retransmit the frames.
- Instead the switch buffers the frames and sends them out port #6 one at a time.

Collision Domains: Half Duplex vs. Full Duplex

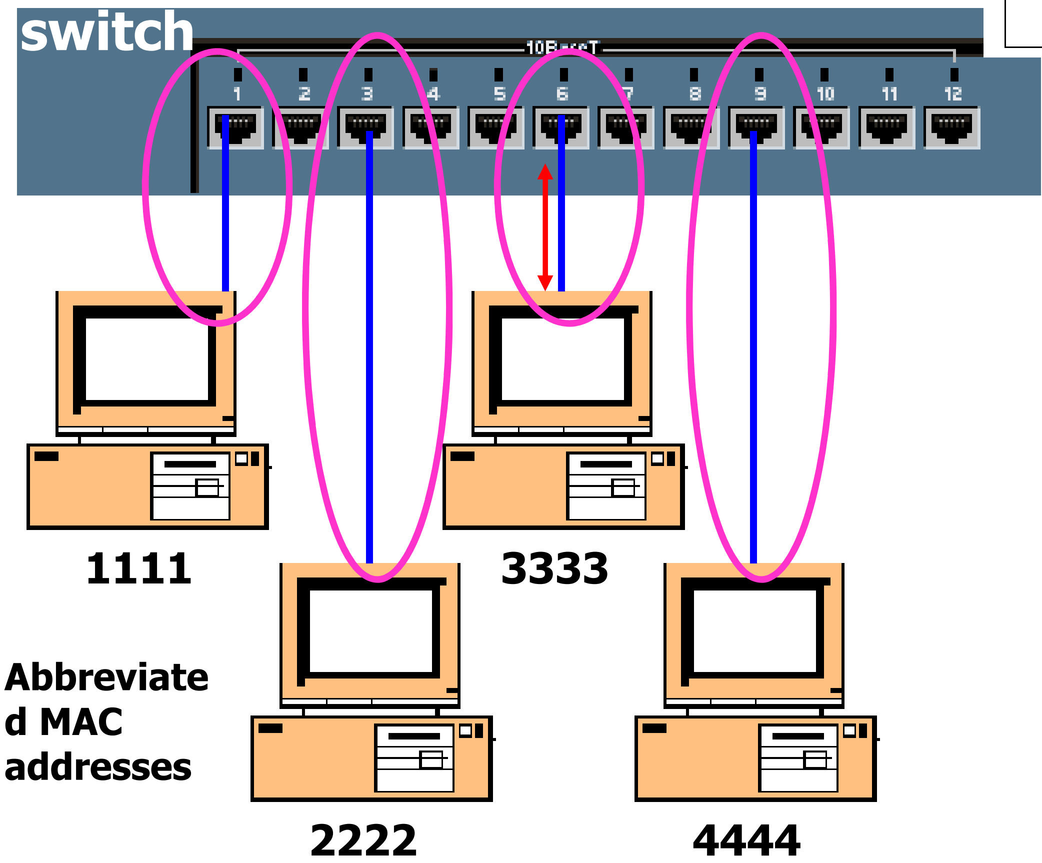
Source Address Table

Port	Src. MAC Add.	Port	Src. MAC Add.
1	1111	6	3333
9	4444		

Collision Domains

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	1111				

Preamble	Destination Address	Source Address	Type	Data	Pad	CRC
	3333	4444				

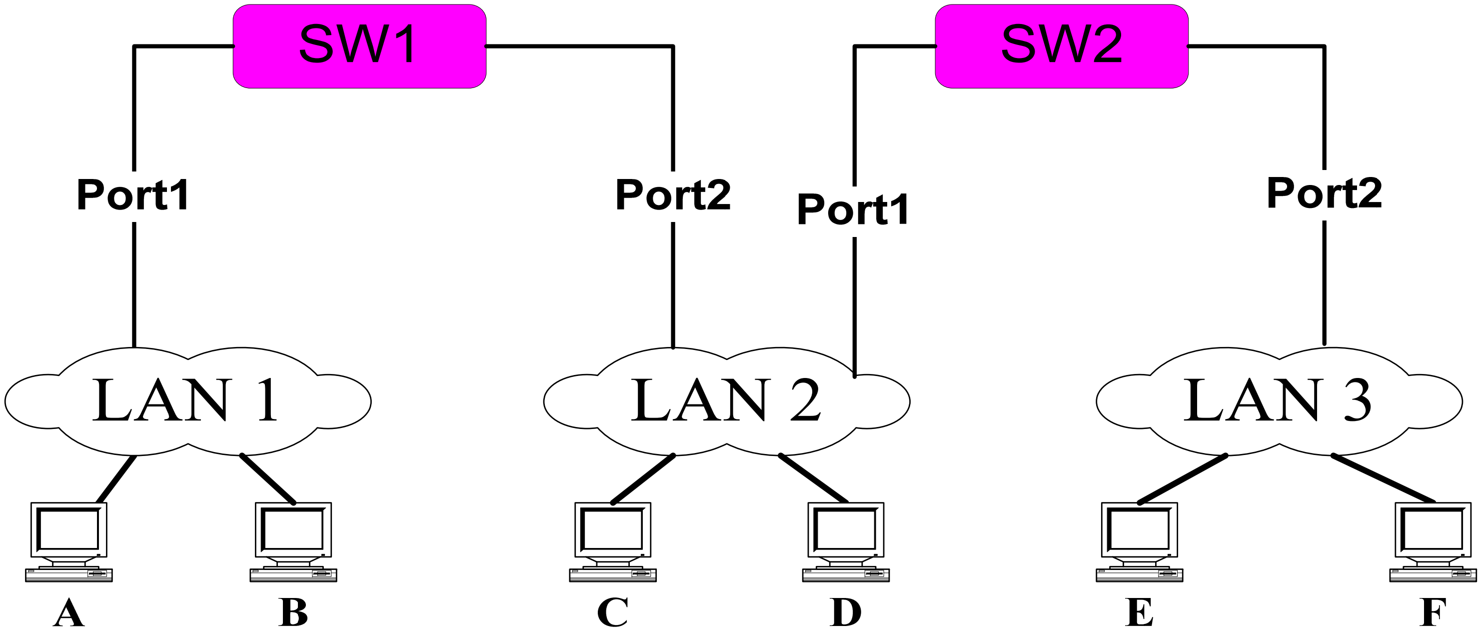


Abbreviated MAC addresses

- In **half-duplex** mode, the collision domain is only between the PC and the switch.
- With a **full-duplex** PC and switch port, there will be no collision.

Example

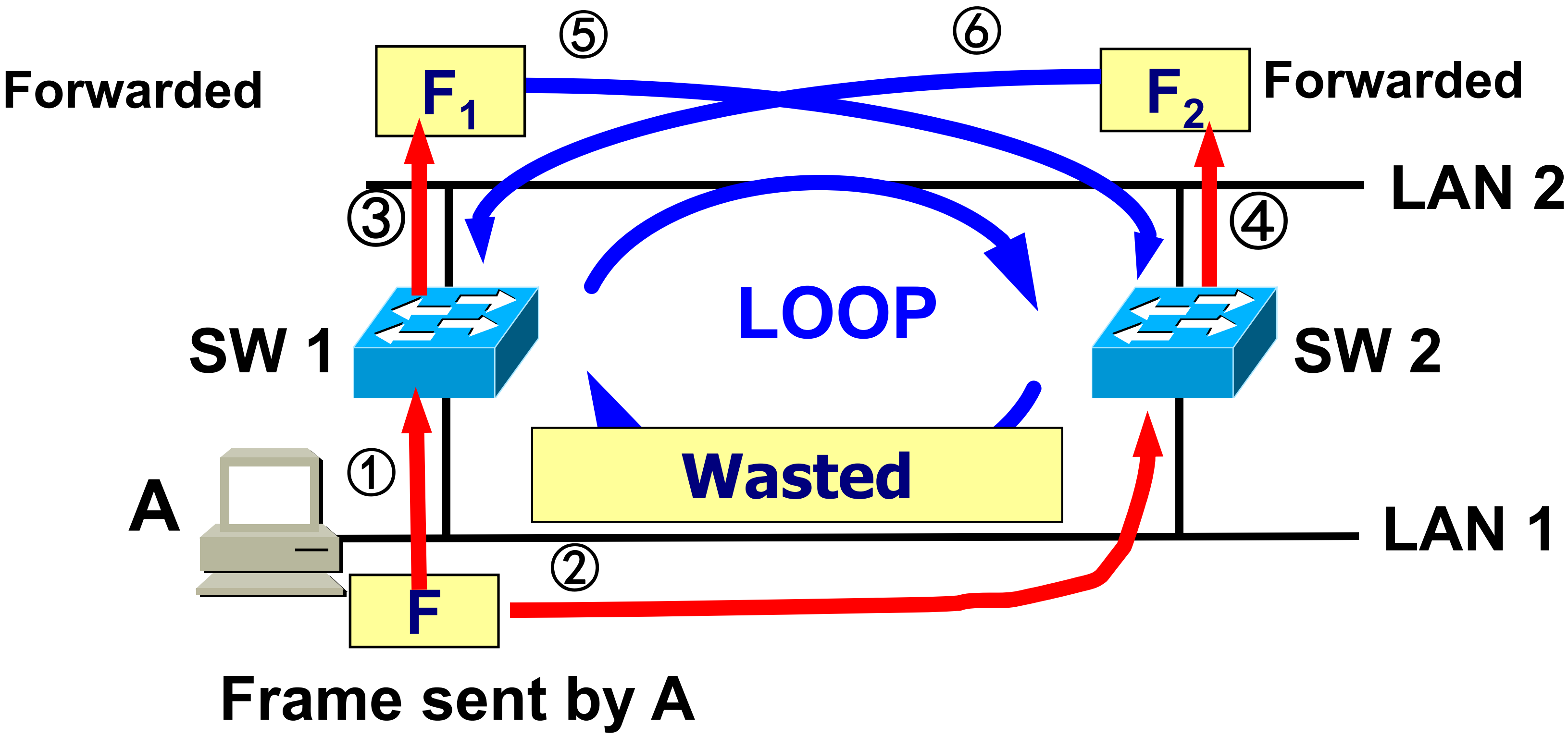
- Consider the following packets:
  - (Src=A, Dest=F)
  - (Src=C, Dest=A)
  - (Src=E, Dest=C)
- What have the switches learned?



Flooding Can Lead to Loops

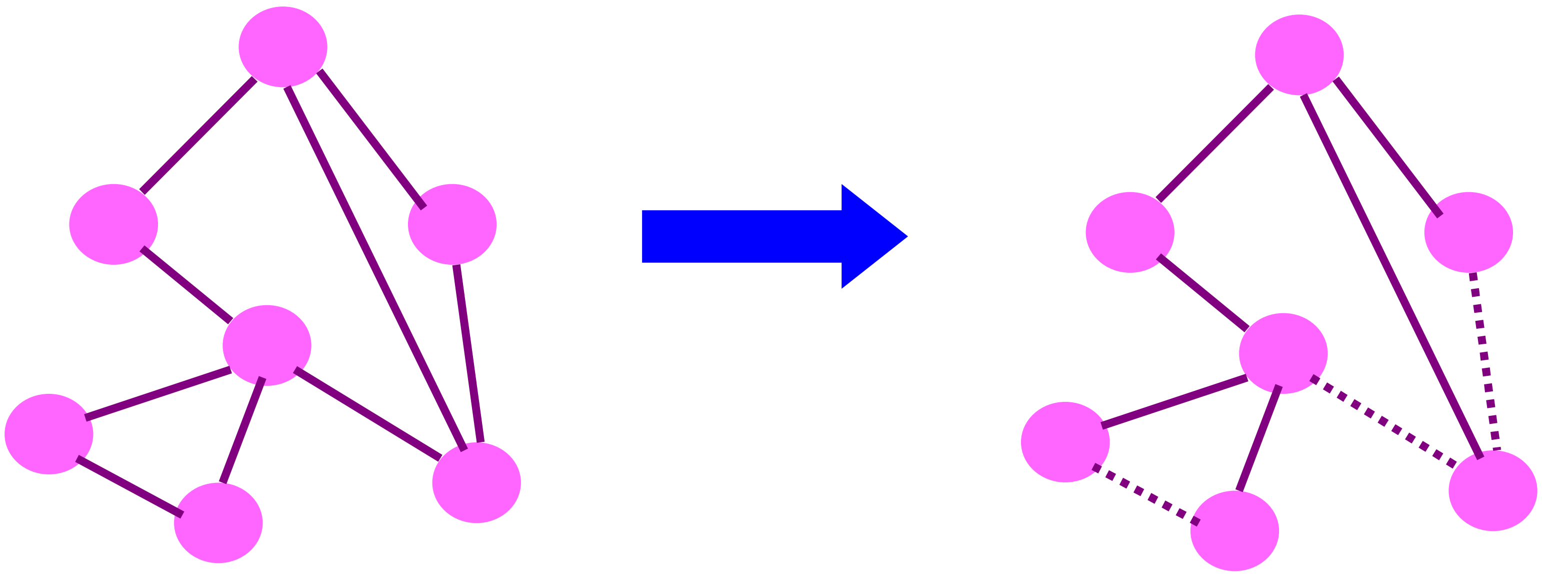
- Switches sometimes need to broadcast frames
  - Upon receiving a frame with an ***unfamiliar destination***
  - Upon receiving a frame sent to the ***broadcast address***
- Broadcasting is implemented by flooding
  - Transmitting frame out every interface
  - … except the one where the frame arrived
- Flooding can lead to forwarding loops
  - E.g., if the network contains a cycle of switches
  - Either accidentally, or by design for higher reliability

Forwarding loop



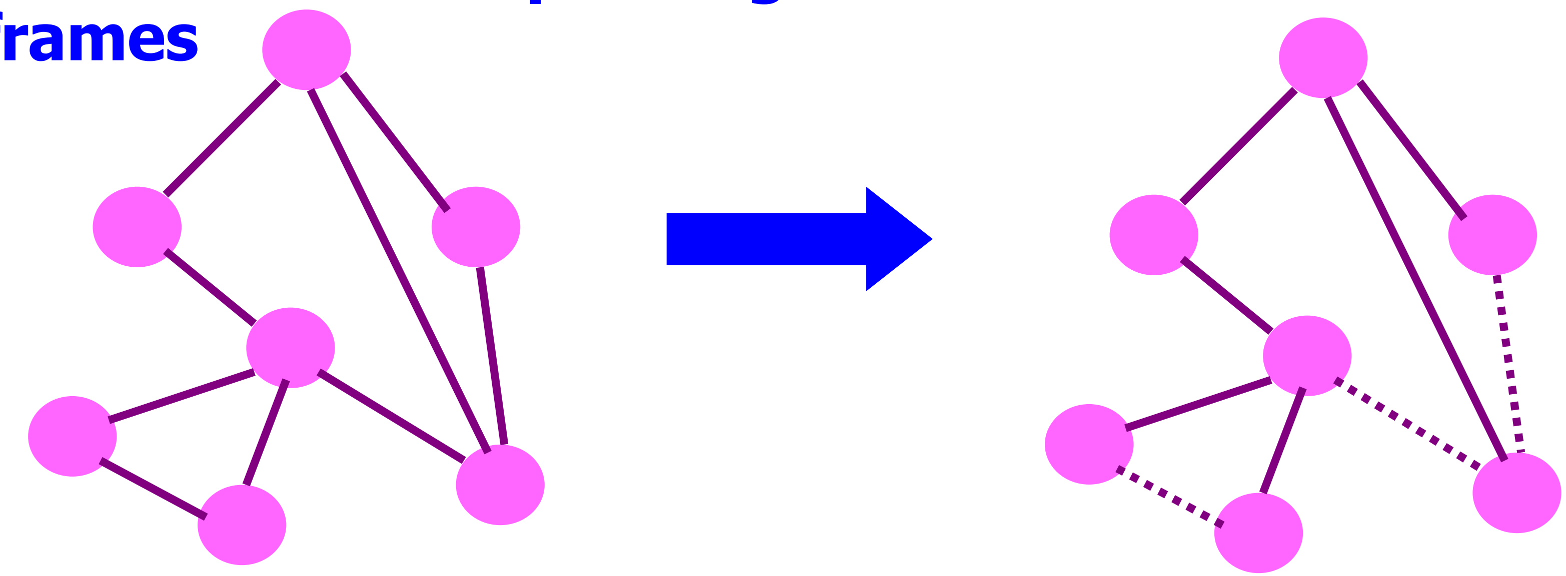
Solution: Spanning Trees

- Ensure the topology has no loops
  - Avoid using some of the links when flooding
  - … to avoid forming a loop
- Spanning tree
  - Sub-graph that covers all vertices but contains no cycles

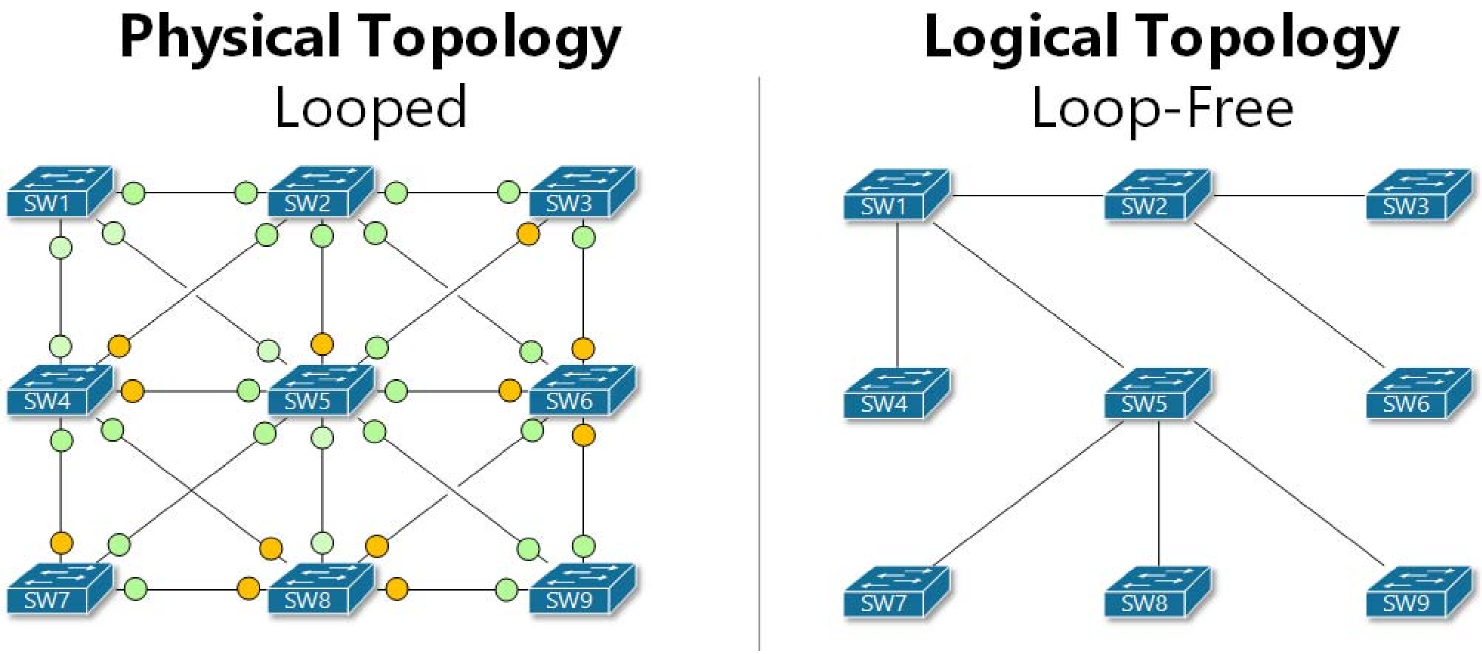


Solution: Spanning Trees

- **Ensure the topology has no loops**
  - Avoid using some of the links when flooding
  - … to avoid forming a loop
- **Spanning tree**
  - Sub-graph that covers all vertices but contains no cycles
  - Links not in the spanning tree do not forward frames



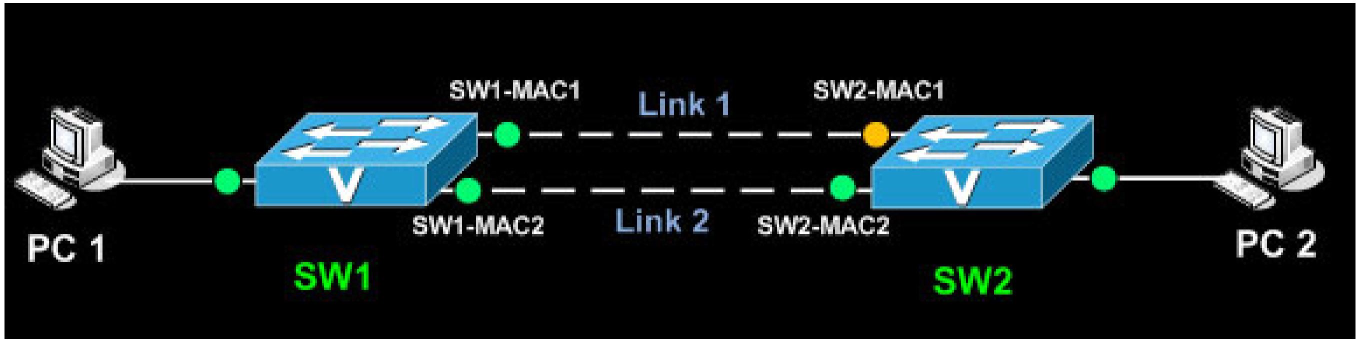
Redundant topology and spanning tree



- It is a spanning tree because all devices in the network are reachable or spanned.
- The algorithm used to create this loop free logical topology is the **spanning-tree algorithm**.

Spanning Tree Protocol

- **IEEE 802.1d** has an algorithm that builds and maintains a **spanning tree** in a dynamic environment
- Bridges/Switches that run 802.1d are called **transparent bridges**
- Bridges exchange BPDU (**Bridge Protocol Data Unit**) to configure the bridge to build the tree.



What do the BPDUs do?

With the help of the BPDUs, bridges can:

- Elect a single bridge as the **root bridge**.
- Calculate the distance of the **shortest path** to the root bridge
- Each bridge can determine a **root port**, the port that gives the best path to the root.
- Each LAN can determine a **designated bridge**, which is the bridge closest to the root bridge. **The designated bridge will forward packets towards the root bridge.**
- Select ports to be included in the spanning tree.

Transparent Bridges

Bridges that execute the spanning tree algorithm are called **transparent bridges**

Overall design goal: **Complete transparency**

- "Plug-and-play"
- Self-configuring without hardware or software changes
- Bridges should not impact operation of existing LANs

Three parts to transparent bridges:

(1) Forwarding of Frames

(2) Learning of Addresses

(3) Spanning Tree Algorithm

BPDU (Bridge Protocol Data Unit)

- **Send to STP multicast address 01:80:C2:00:00:00 every 2 seconds by default.**

Field	Byte	Value (default)
Protocol Identifier	2	0x0000
Protocol version ID	1	0x00
BPDU Type	1	0x00 or 0x80
Flags	1	0000 0000
Root BID	8	0x8000(32768)+MAC
Path Cost to Root Bridge	4	4(1G) / 19(100M) / 100(10M)
Sender BID	8	0x8000(32768)+MAC
Port ID	2	0x80 + Port number
Message Age	2	
Max Age	2	20s
Hello time	2	2s
Forward Delay	2	15s

key concepts

- **Bridge ID (BID)**
  - **to identify each bridge/switch;**

Bridge Priority	MAC Address
2 bytes	6 bytes

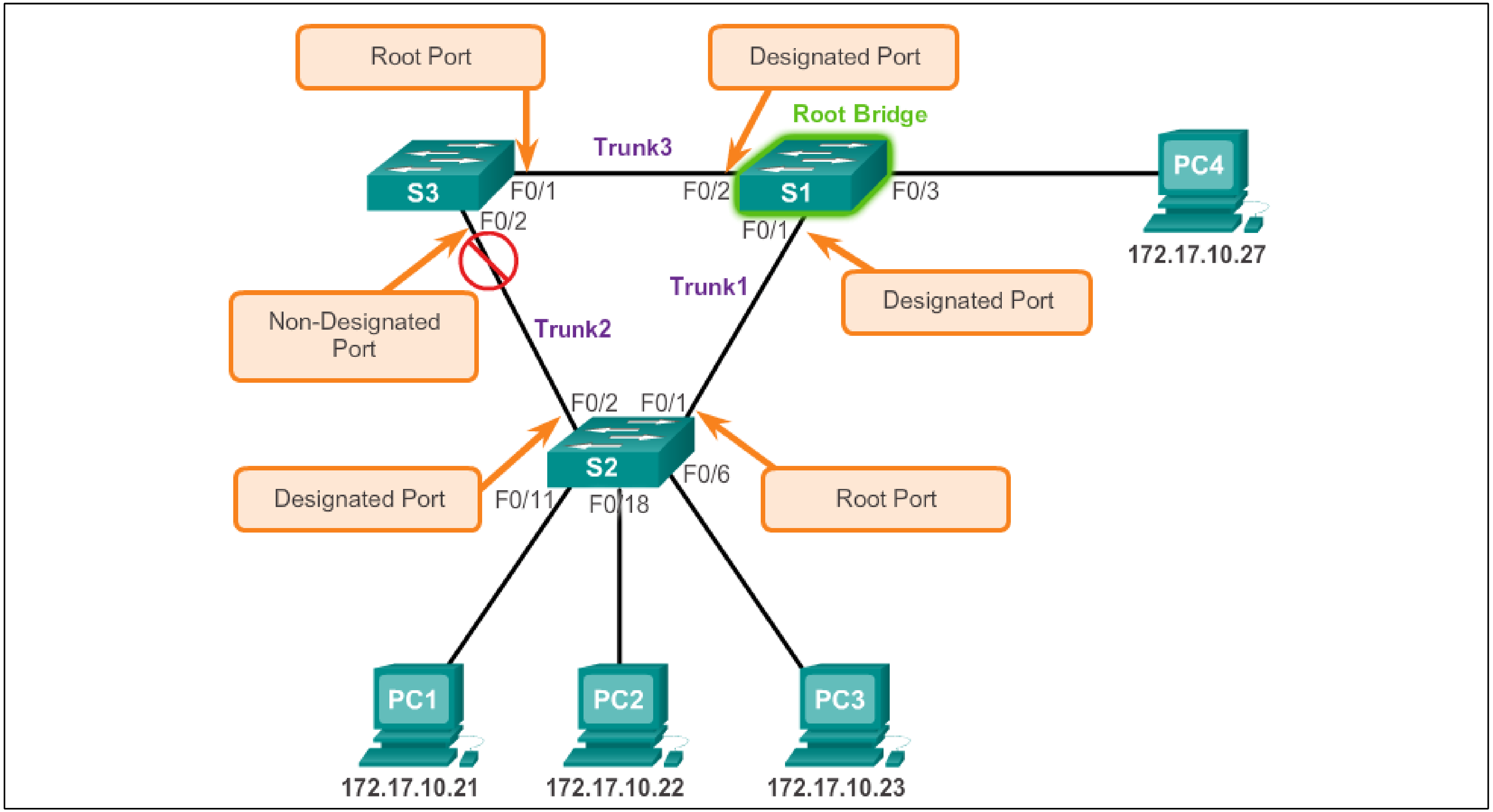
Configurable
Range: 0 - 65535
Default: 32758

Smaller value, more chances to be selected.

key concepts

- Root Bridge/SW
  - The bridge/SW with the lowest bridge ID.
- Designated Bridge/SW
  - A bridge/SW closest to the root on each LAN segment.
  - Root port
    - Lowest cost port to the Root Bridge/SW
  - Designated port
    - Lowest cost port on segment

Port Roles



Root Port
Designated Port
Root Bridge
Trunk3
S3
F0/1
F0/2
S1
F0/3
PC4
172.17.10.27
F0/2
F0/1
Trunk1
Designated Port
Non-Designated Port
Trunk2
F0/2
F0/1
S2
F0/6
Designated Port
F0/11
F0/18
Root Port
PC1
PC2
PC3
172.17.10.21
172.17.10.22
172.17.10.23

key concepts

- **Path Cost**
  - A path cost value is given to each port.
  - The lower the cost, the closer the switch is to the root.
- **Path cost to root bridge**
  - Accumulated port cost to the root bridge.

Bandwidth	Cost
4 Mbps	250
10 Mbps	100
16 Mbps	62
45 Mbps	39
100 Mbps	19
155 Mbps	14
622 Mbps	6
1 Gbps	4
10 Gbps	2

key concepts

- **Port ID**
  - **16 bits long**

Priority (6 bits)	Port Number (10 bits)

Configurable
Range: 0 - 255
Default: 128
Smaller value, more chances to be selected.

Spanning-Tree Operation

- When the network has stabilized, it has *converged* and there is ***one spanning tree per network***
- For every switched network the following elements exist:
  - One root bridge per network
  - One root port per non root bridge
  - One designated port per segment
  - Unused, non-designated ports
- Root ports and designated ports **forward** data traffic.
- Non-designated ports **discard** data traffic
  - These ports are called blocking or discarding ports

Three Steps of STP Convergence

- **Step 1 Elect one Root Bridge**
- **Step 2 Elect Root Ports**
- **Step 3 Elect Designated Ports**

Four-Step decision Sequence:

Step 1 - Lowest BID

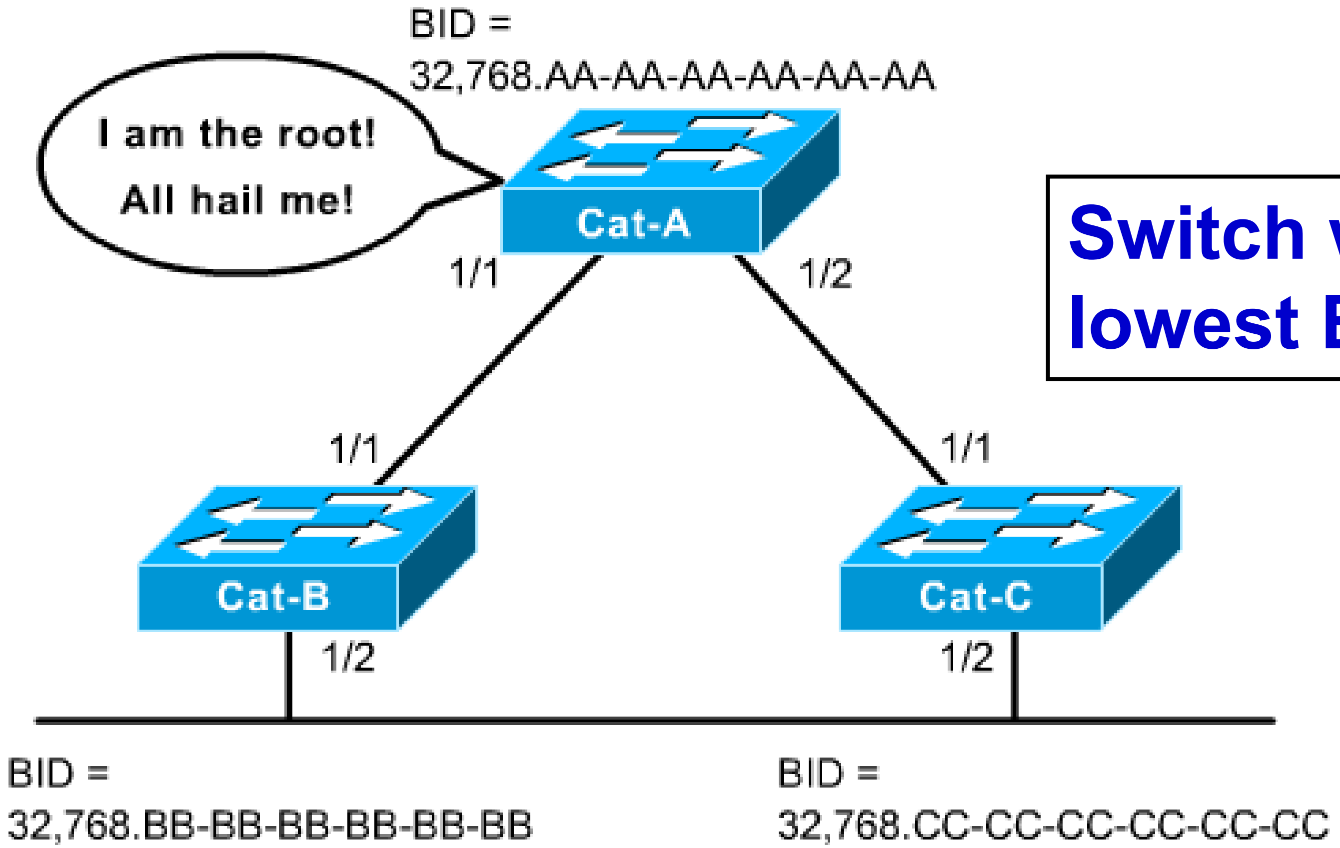
Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 - Lowest Port ID

Step 1 Elect one Root Bridge

Cat-A has the lowest Bridge MAC Address, so it wins the Root War!



Switch with the lowest BID wins

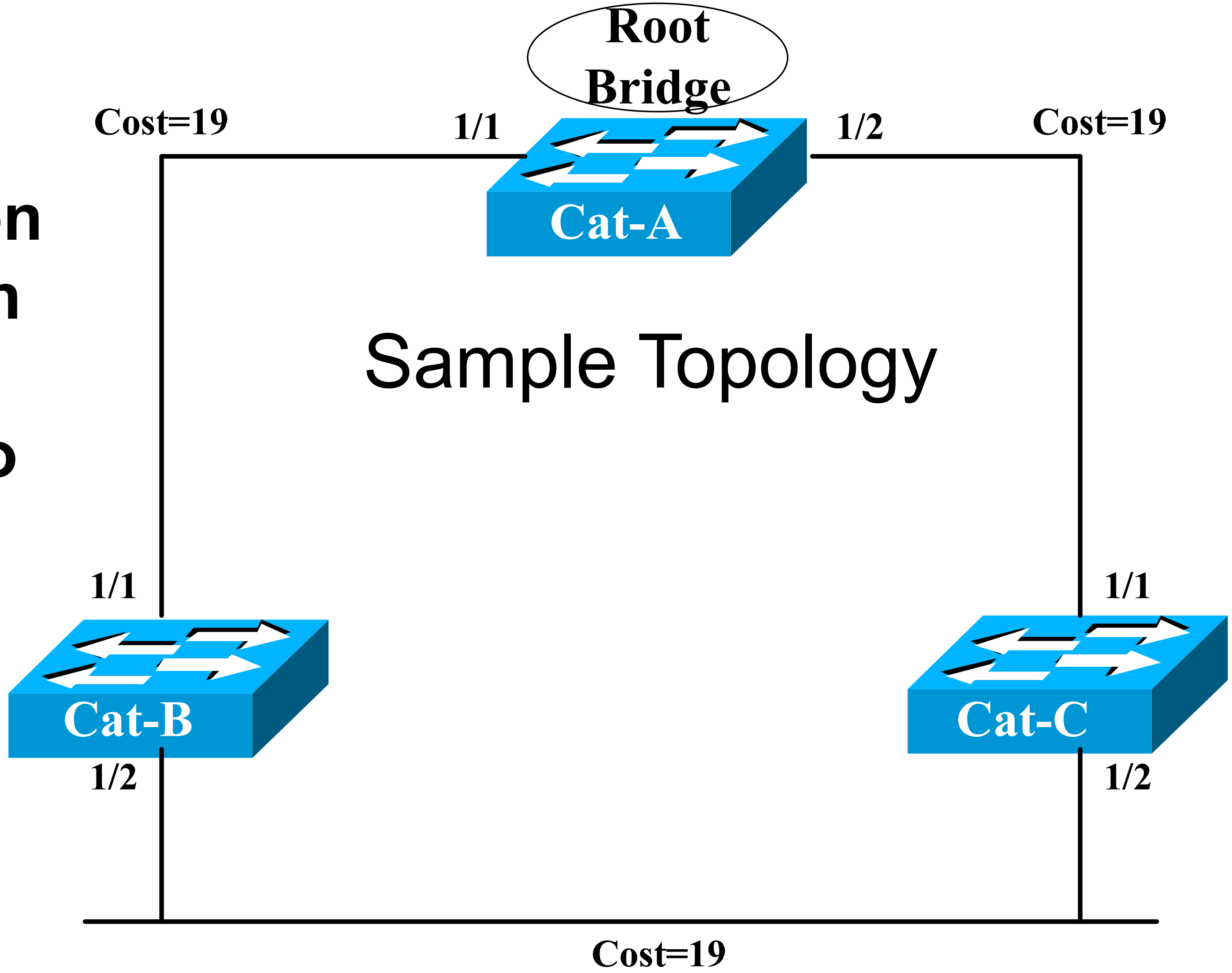
All 3 switches have the same default Bridge Priority value of 32,768

Step 2 Elect Root Ports

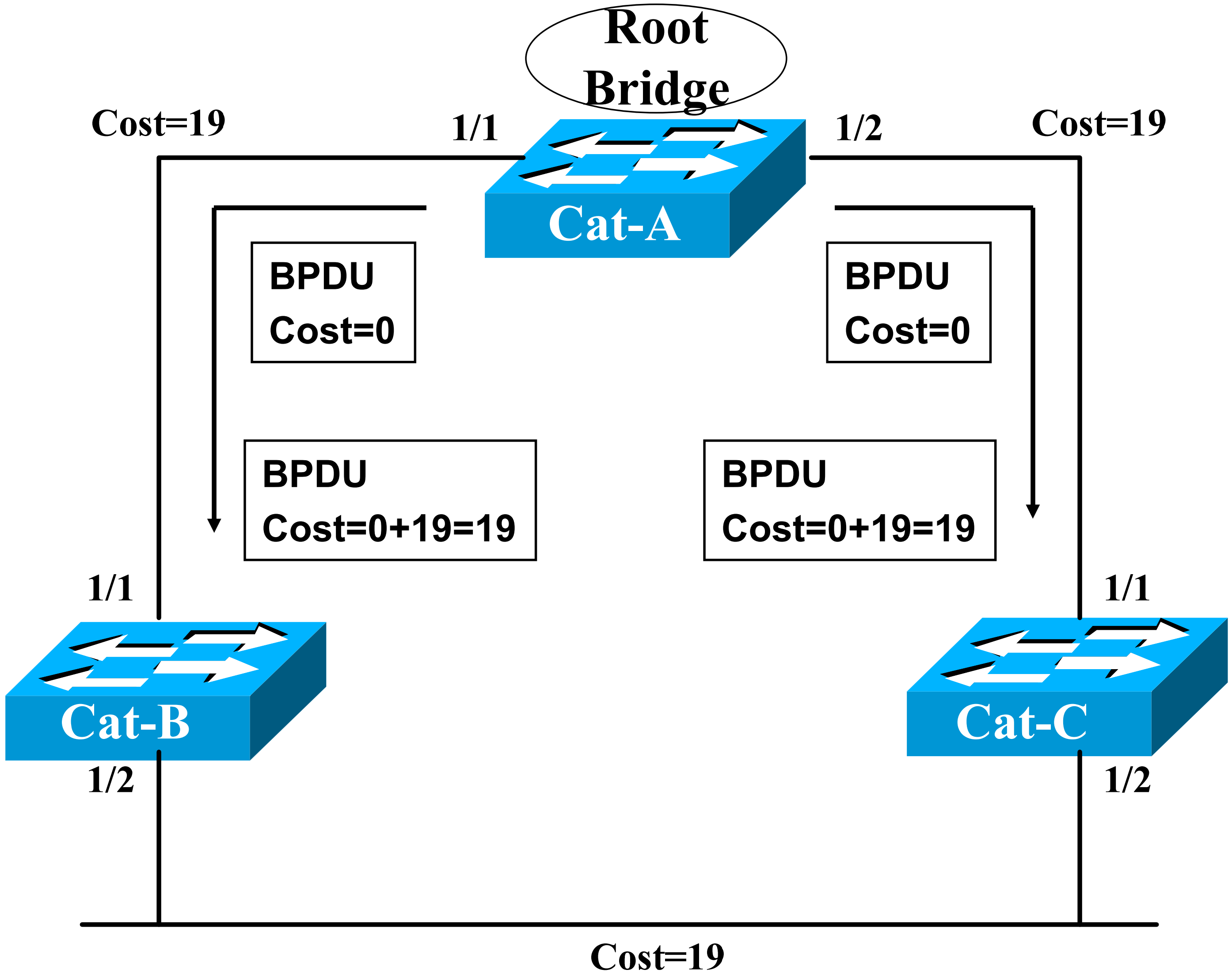
- **The switch looks at three components of the BPDU:**
  - **Lowest path cost to root bridge**
  - **Lowest sender Bridge ID**
  - **Lowest port priority/port ID**

Step 2 Elect Root Ports

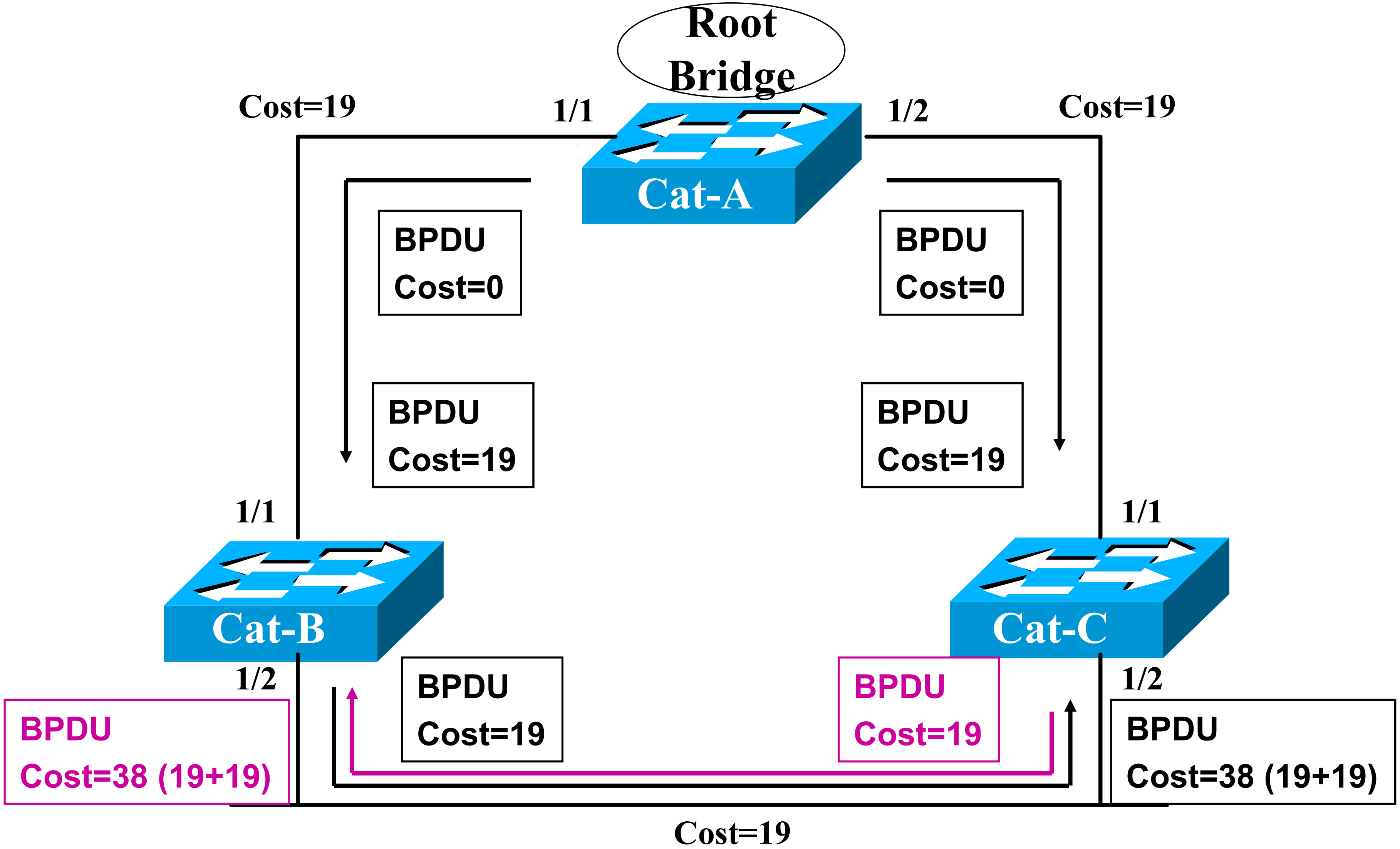
Root port on each switch will be the one used to connect to the root switch.



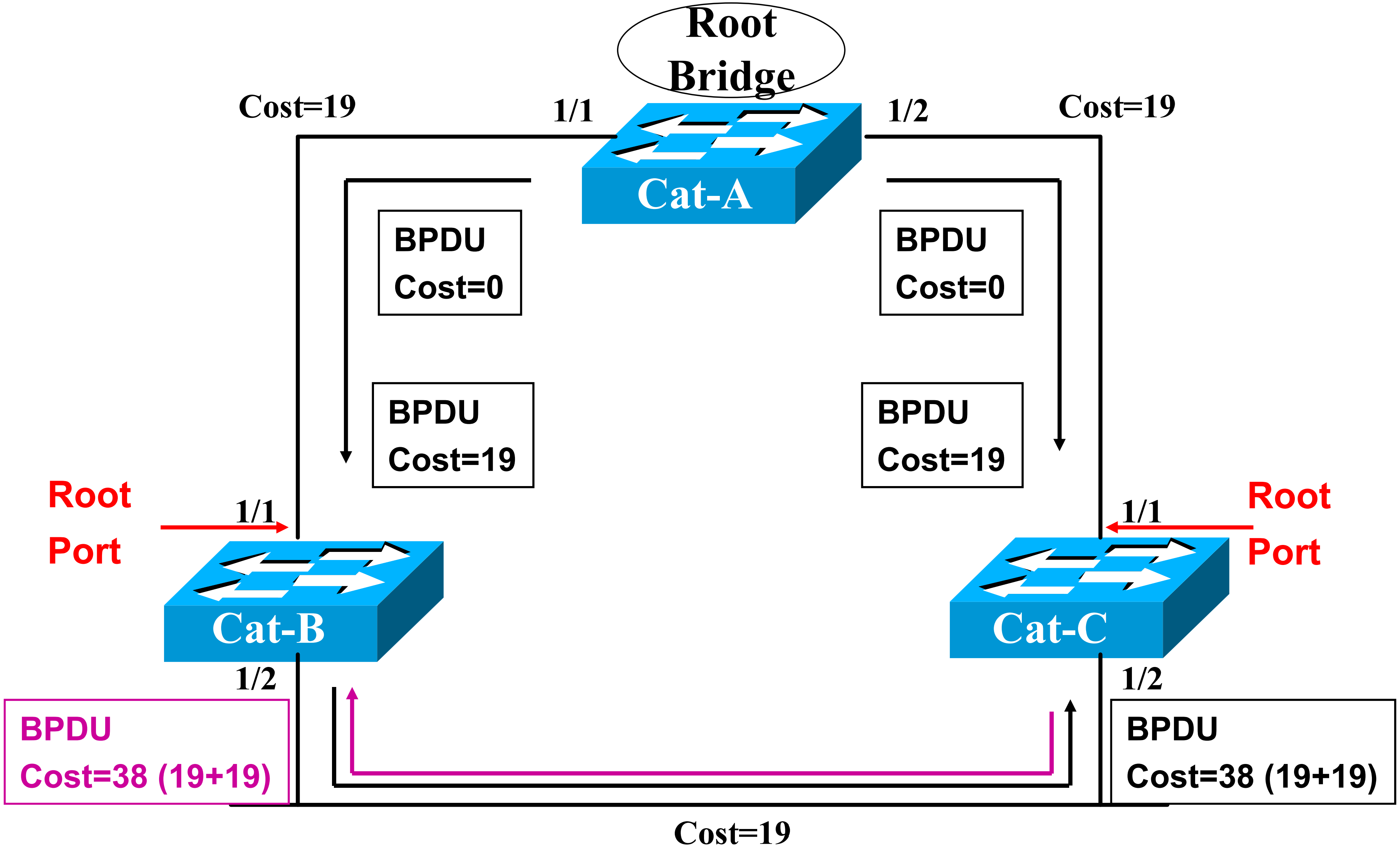
Step 2 Elect Root Ports



Step 2 Elect Root Ports



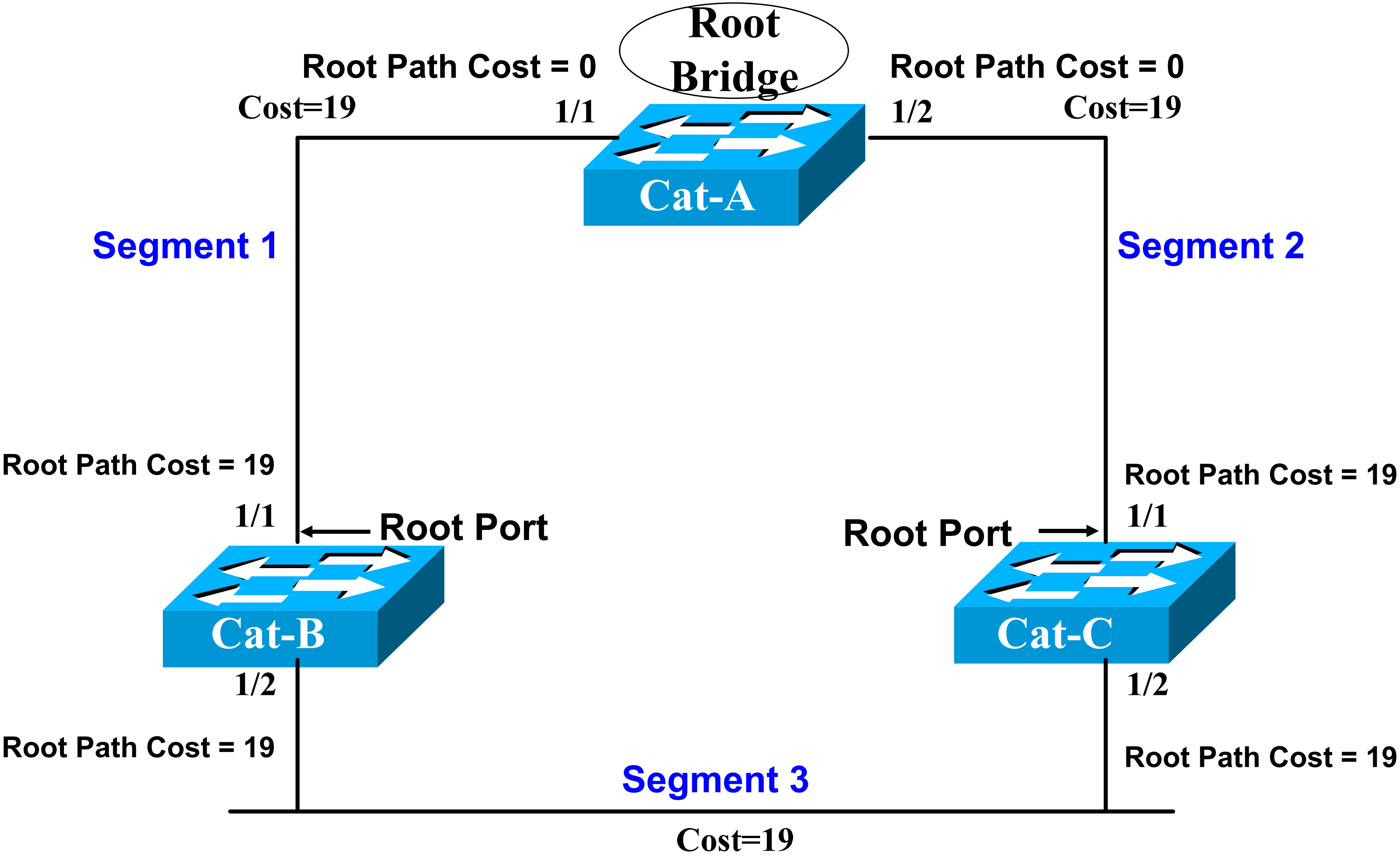
Step 2 Elect Root Ports



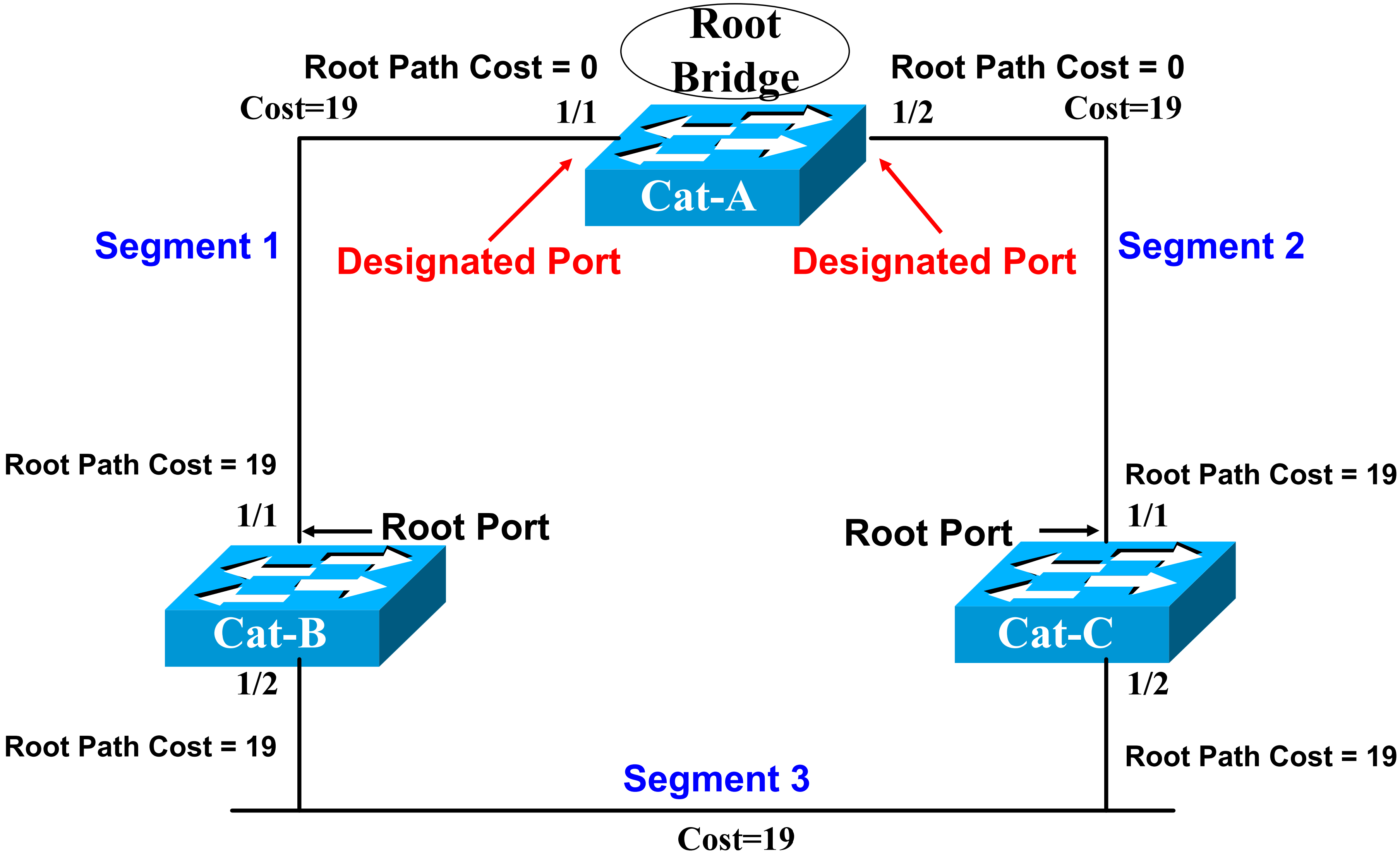
Step 2 Elect Root Ports

- On a tie, choose the neighboring switch with the lowest bridge ID.
- If a tie for the ID, select port with the lowest priority.
- If a tie, select the lowest port number.

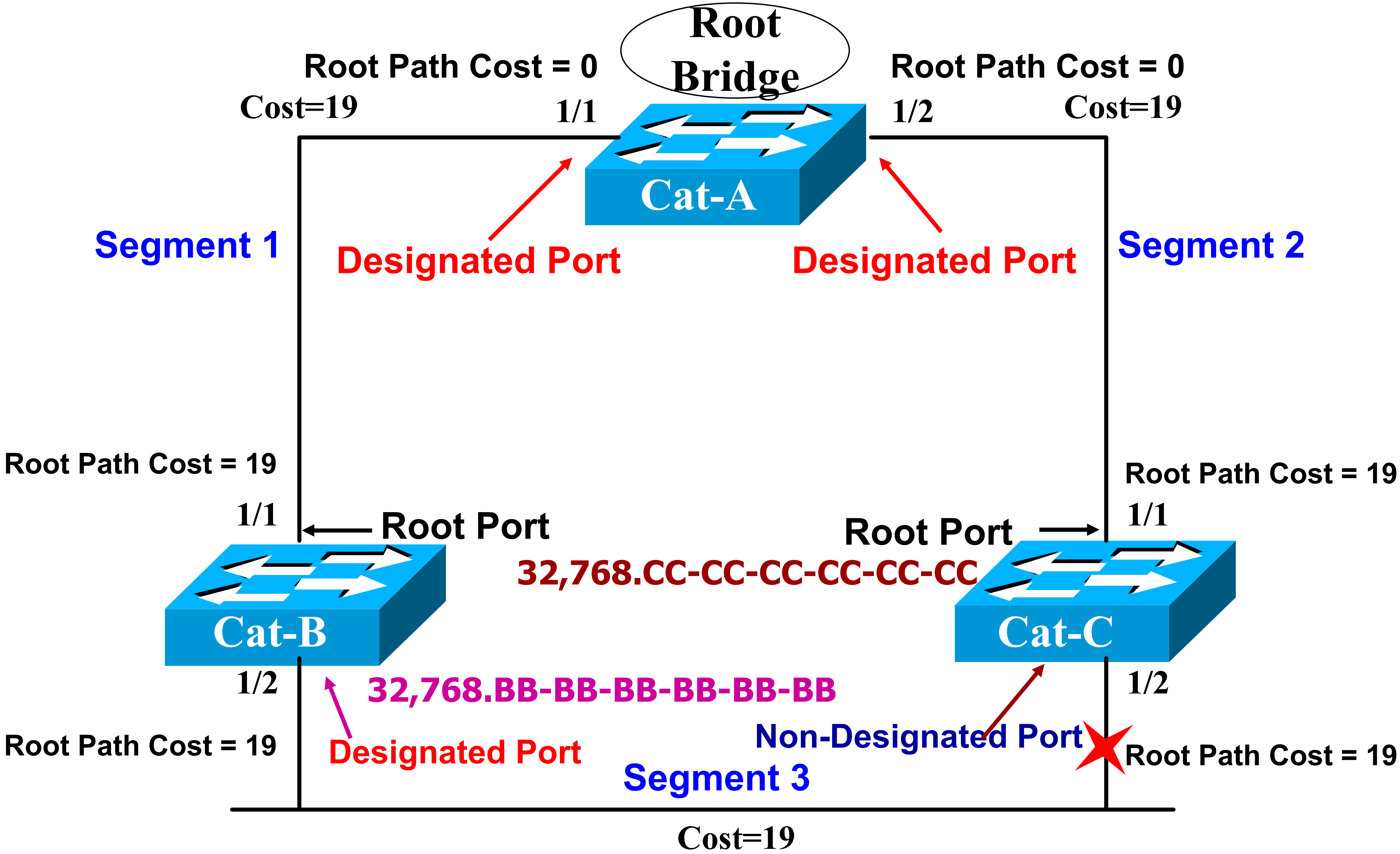
Step 3 Elect Designated Ports



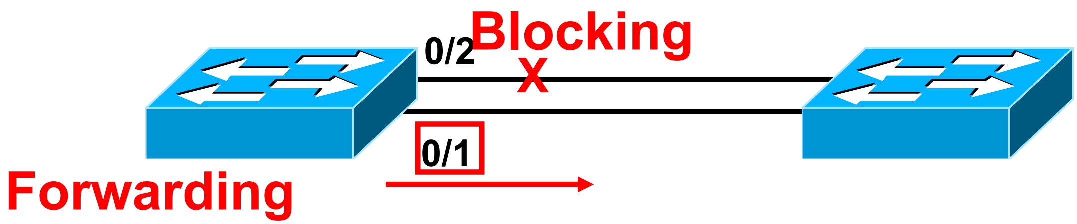
Step 3 Elect Designated Ports



Step 3 Elect Designated Ports



Step 3 Elect Designated Ports



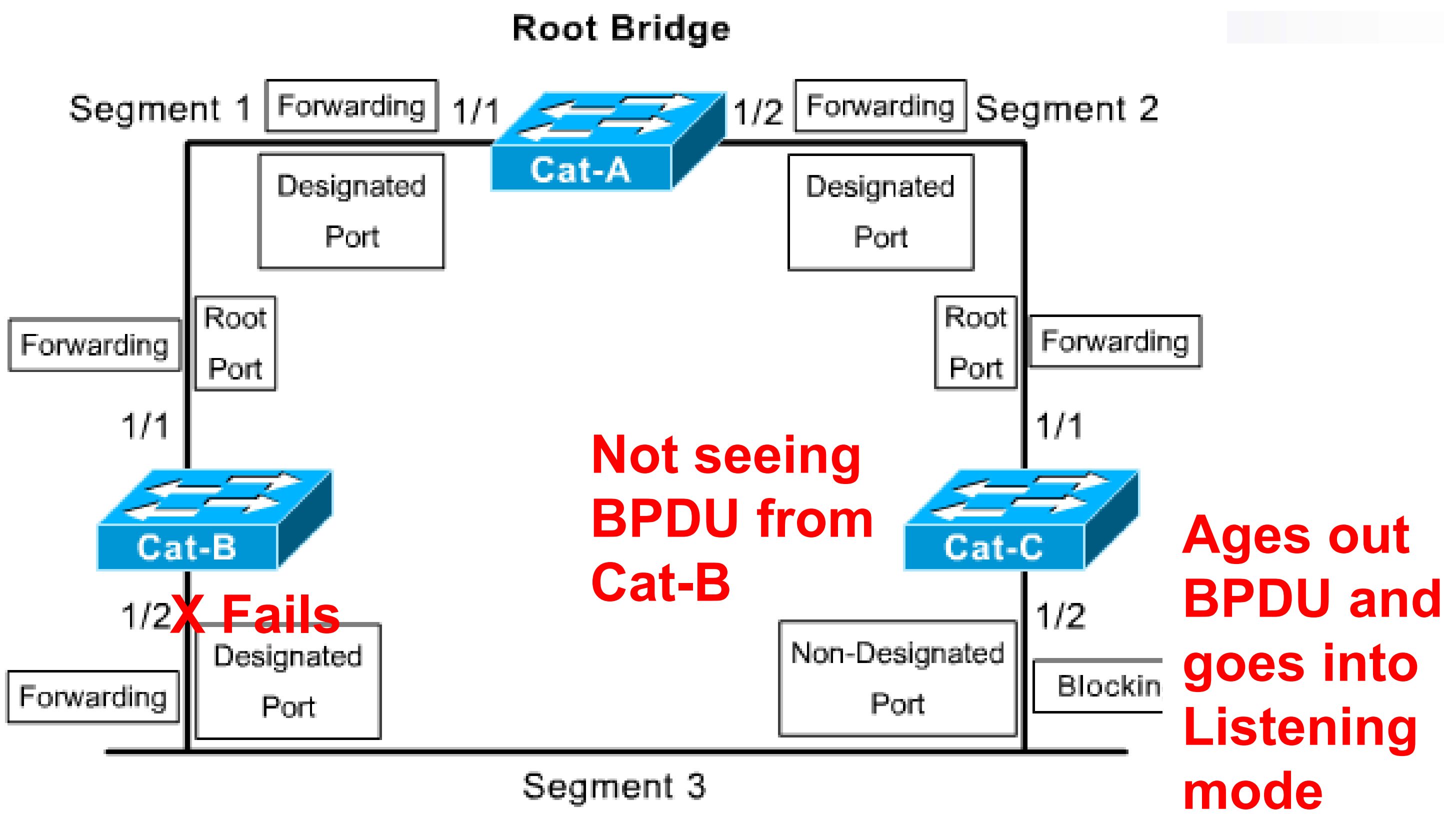
- If the path cost and bridge IDs are equal (as in the case of parallel links), the switch goes to the port priority as a tiebreaker.
- Lowest port priority wins (all ports set to 32).
- You can set the priority from 0 – 63.
- If all ports have the same priority, the port with the lowest port number forwards frames.

Port States

State	Forwards Data Frames?	Learns MACs based on Received Frames?	Transitory or Stable State?
Blocking	No. only receive BPDUs.	No	Stable
Listening	No. BPDUs processed.	No	Transitory
Learning	No BPDUs processed.	Yes	Transitory
Forwarding	Yes BPDUs processed.	Yes	Stable
Disabled	No	No	Stable

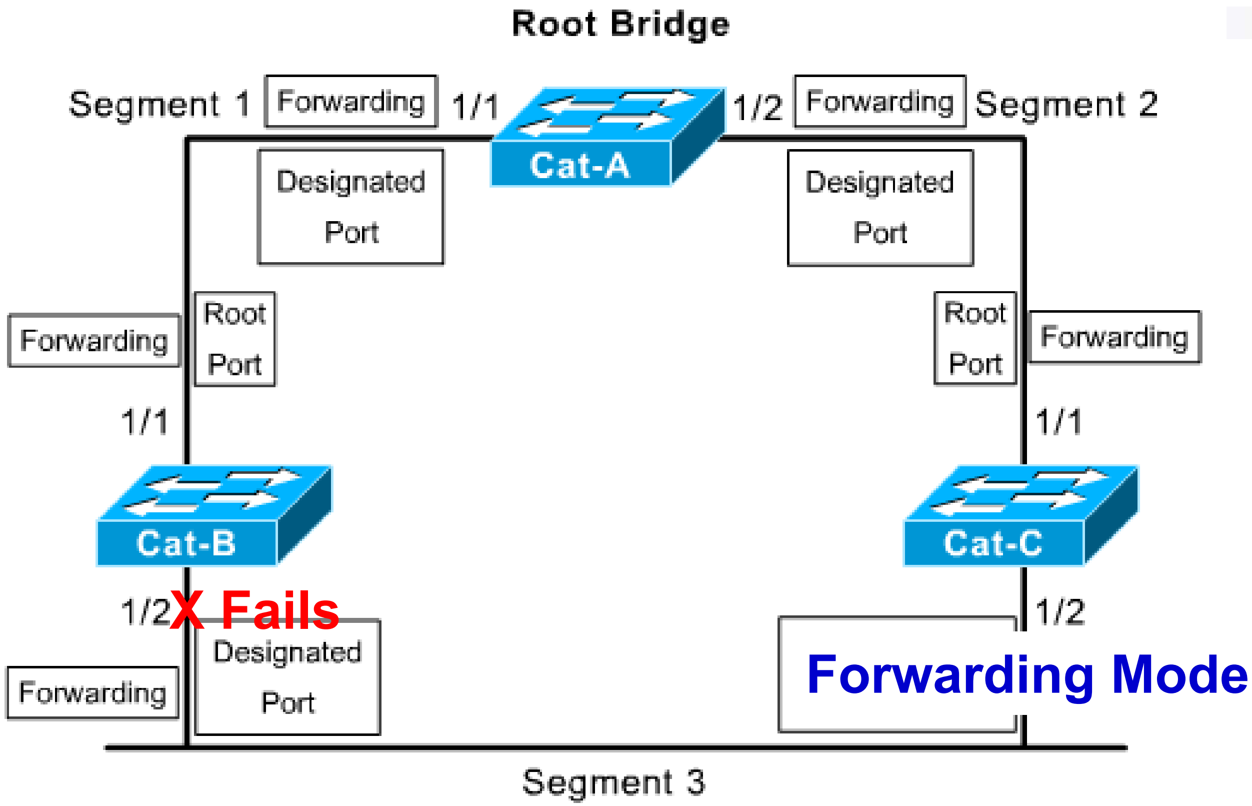
Forwarding and Blocking

Characterization of Port	STP State	Description
All the root switch's ports	Forwarding	The root switch is always the designated switch on all connected segments.
Each non-root switch's root port	Forwarding	The port through which the switch has the least cost to reach the root switch.
Each LAN's designated port	Forwarding	The switch forwarding the lowest-cost BPDU onto the segment is the designated switch for that segment.
All other working ports	Blocking	The port is not used for forwarding frames, nor are any frames received on these interfaces considered for forwarding.



If Cat-B: 1/2 fails:

- **Cat-C notices it is not receiving BPDUs from Cat-B.**
- **20 seconds (max age) after the failure, Cat-C ages out the BPDU that lists Cat-B as having the DP for segment 3.**
- **This causes Cat-C:1/2 to transition into the Listing state (15 seconds) in an effort to become the DP.**



- **Cat-C:1/2 now offers the most attractive access from the Root Bridge to this link, it eventually transitions to Learning State (15 seconds), then into Forwarding mode.**
- **In practice this will take 50 seconds (20 max age + 15 Listening + 15 Learning) for Cat-C:1/2 to take over after the failure of Cat-B:1/2.**

Using Hubs / Repeater

- Layer 1 devices
- Inexpensive
- In one port, out the others
- One collision domain
- One broadcast domain

Using Switches / Bridges

- Layer 2 devices
- Layer 2 filtering based on Destination MAC addresses and Source Address Table
- One collision domain per port
- One broadcast domain across all switches

Chapter 4: Roadmap

- Medium Access Control
- Local Area Networks (LANs) and IEEE 802
- Ethernet
- Wireless LAN
- LAN Interconnection
- LAN Switching
- VLAN

Evolution Toward Virtual LANs

- In the olden days…
  - Thick cables snaked through cable ducts in buildings
  - Every computer they passed was plugged in
  - All people in adjacent offices were put on the same LAN
  - Independent of whether they belonged together or not
- More recently…
  - Hubs and switches changed all that
  - Every office connected to central wiring closets
  - Often multiple LANs (*k* hubs) connected by switches
  - Flexibility in mapping offices to different LANs

Group users based on organizational structure, rather than the physical layout of the building.

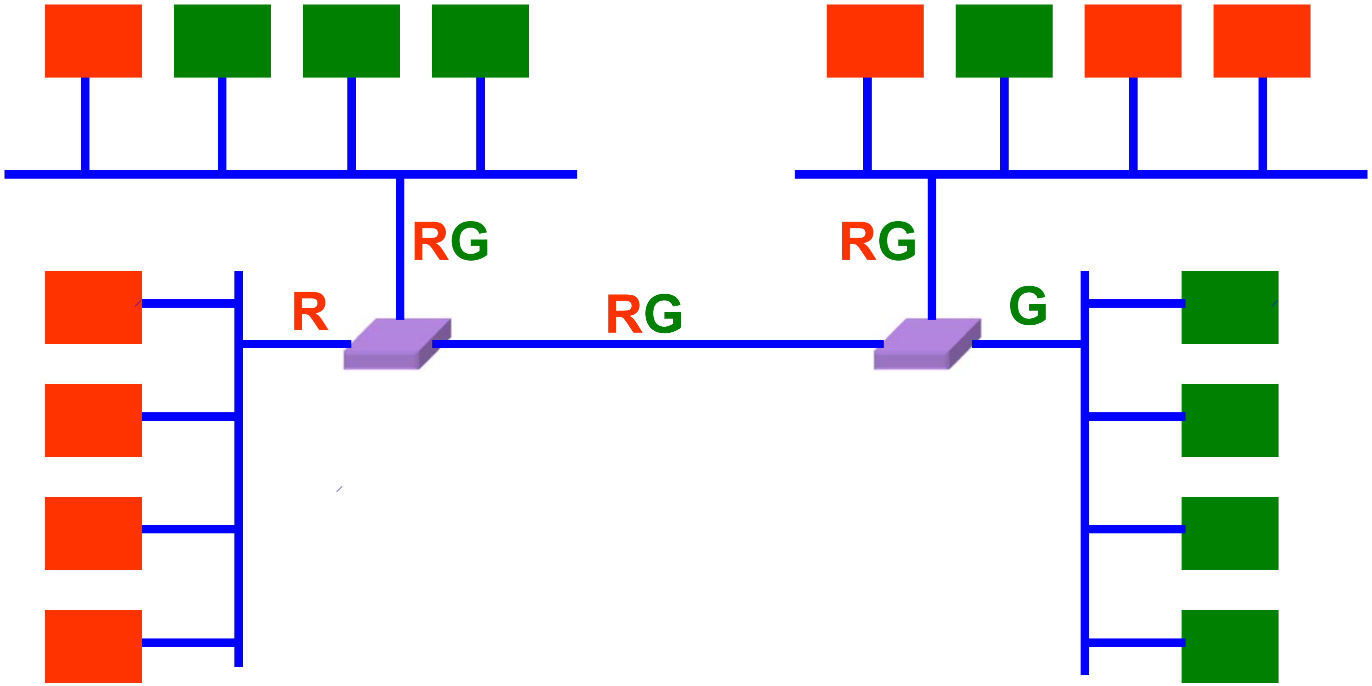
Why Group by Organizational Structure?

- **Security**
  - Ethernet is a shared media. Any interface card can be put into "promiscuous" mode, and get a copy of all of the traffic (e.g., midterm exam)
  - Isolating traffic on separate LANs improves security
- **Load**
  - Some LAN segments are more heavily used than others, can saturate their own segment and not the others
  - Plus, there may be natural locality of communication

People Move, and Roles Change

- Organizational changes are frequent
  - E.g., faculty office becomes a grad-student office
  - E.g., graduate student becomes a faculty member
- Physical rewiring is a major pain
  - Requires unplugging the cable from one port
  - … and plugging it into another
  - … and hoping the cable is long enough to reach
  - … and hoping you don't make a mistake
- Would like to "rewire" the building in software
  - The resulting concept is a Virtual LAN (VLAN)

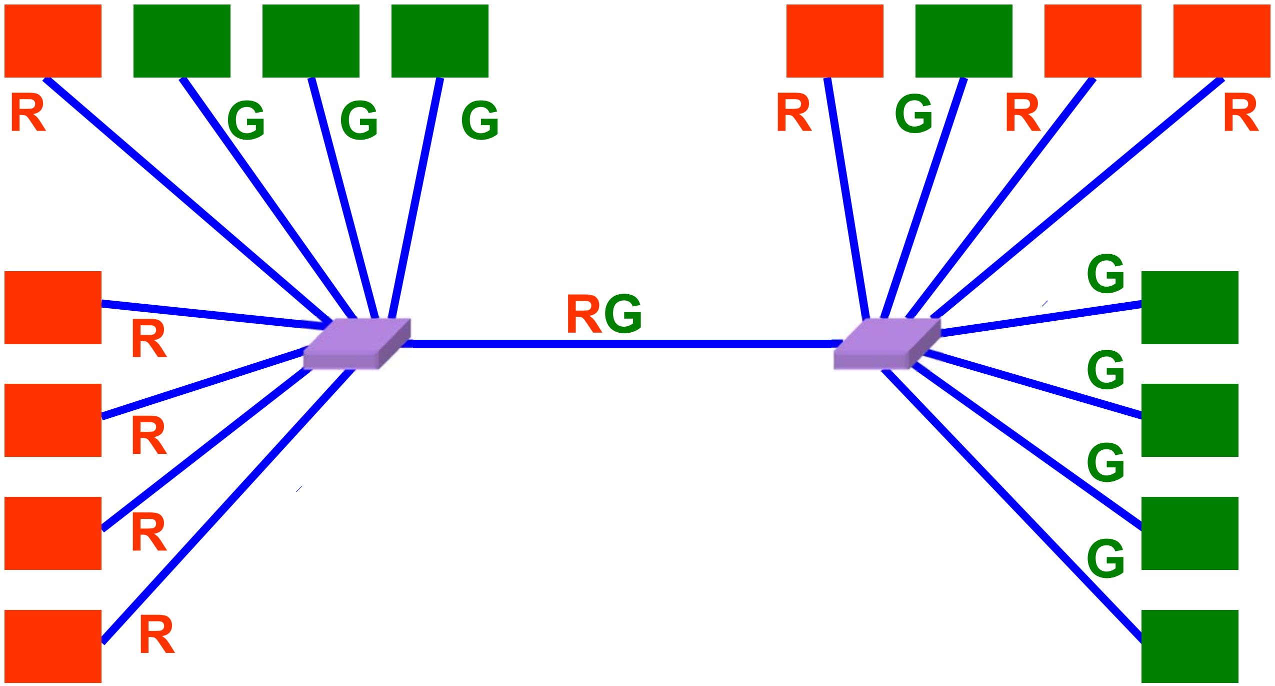
Example: No Virtual LANs



Red workgroup and Green workgroup

Bridges/Switches forward traffic to all

Example: Two Virtual LANs



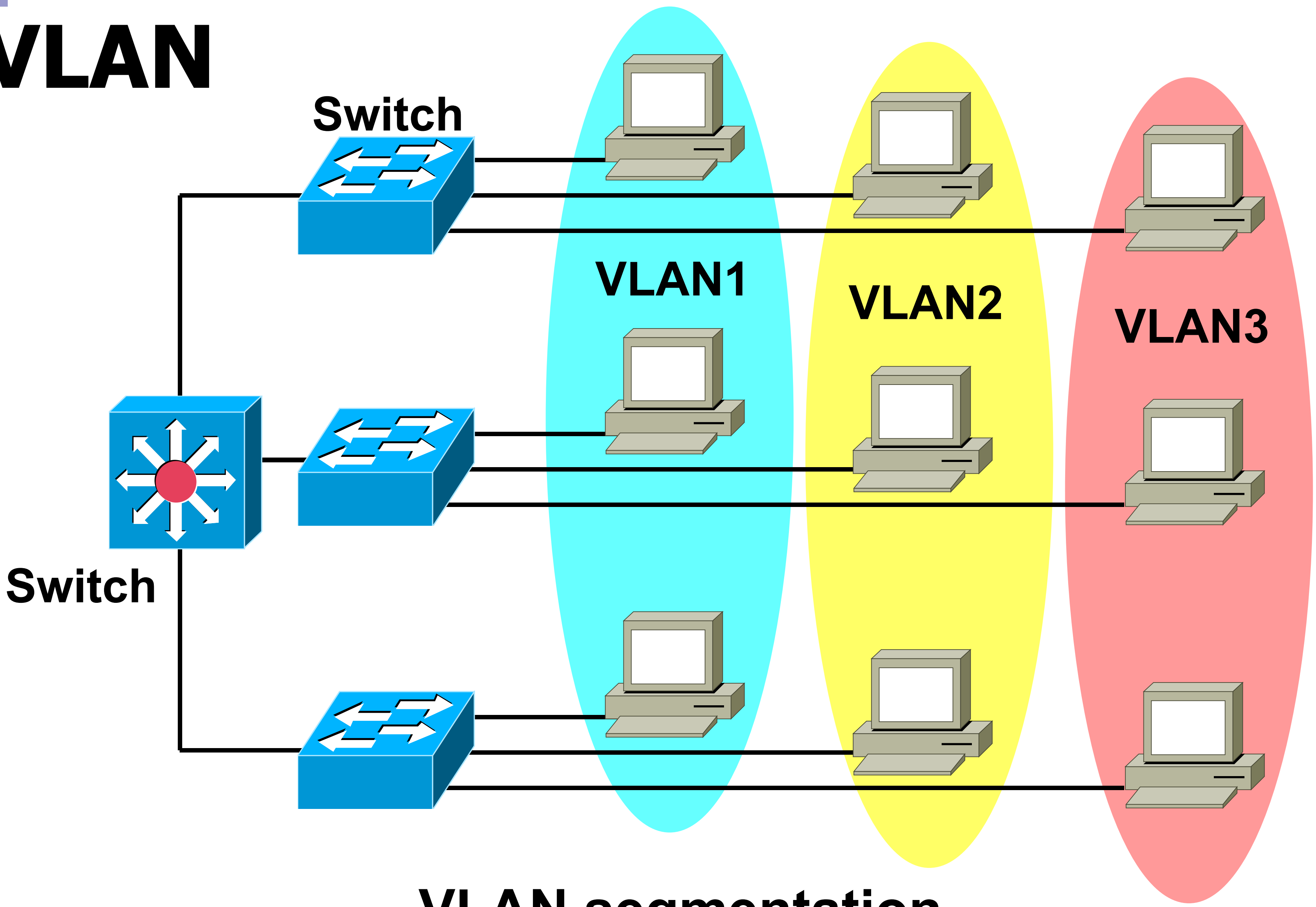
Red VLAN and Green VLAN

Switches forward traffic as needed

VLAN

- VLAN stands for Virtual Local Area Network.
- Can be seen as a group of end hosts, perhaps on multiple physical LAN segments, that are not constrained by their physical location and can communicate as if they were on a common LAN.
- Configured through software rather than hardware.

VLAN

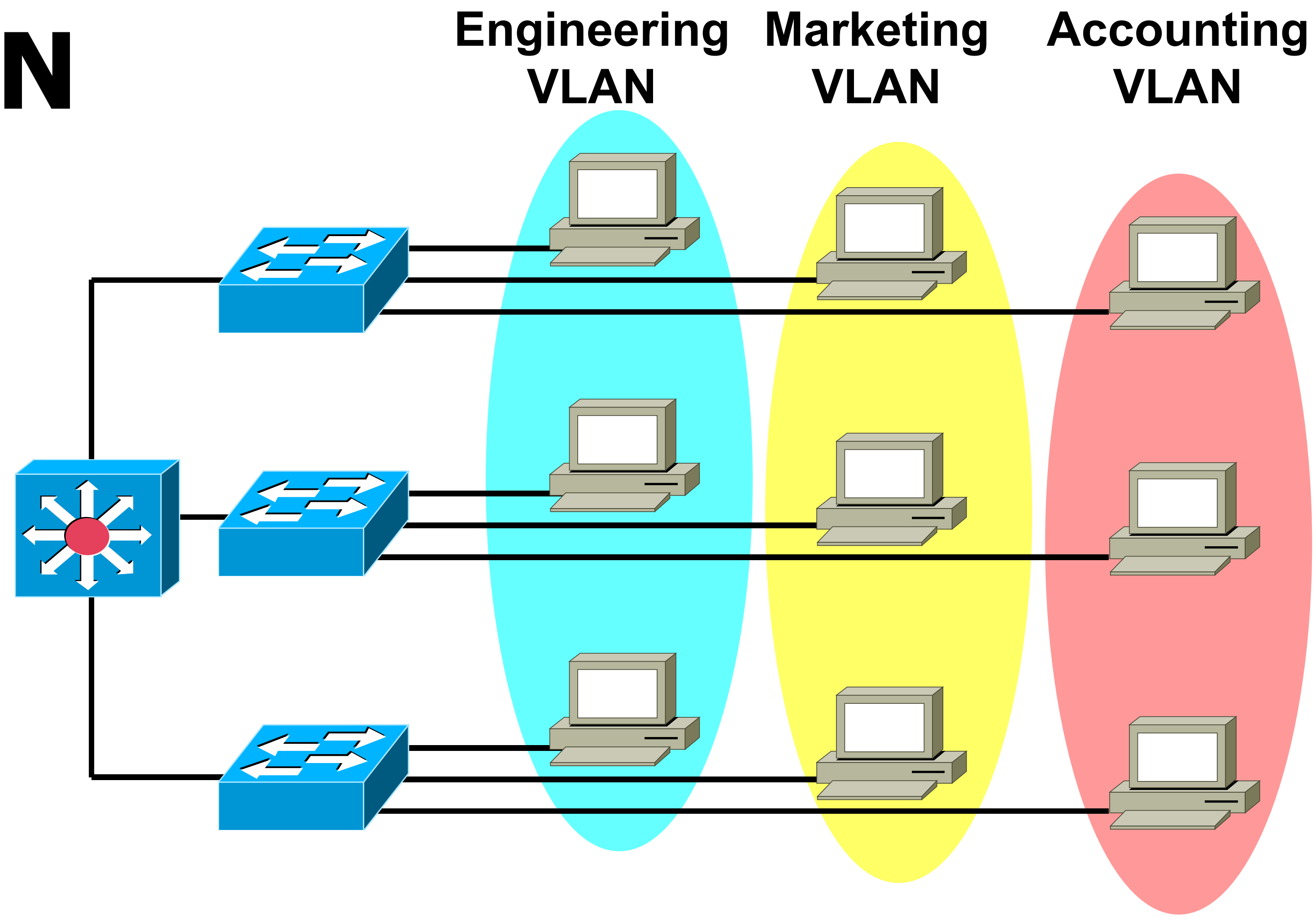


VLAN segmentation

Why Using VLAN?

- **Separate broadcast domains:** a group of end hosts will not be bothered by the broadcast traffic generated by another group of end hosts.
- **Achieve higher security:** now a host cannot snoop on the traffic of another group of hosts.
- **Ease management:**
  - do not need to change a host's IP address when it moves.
  - VLANs can be assigned and managed dynamically without physical limitations.
  - VLAN can be used to balance bandwidth allotment per group

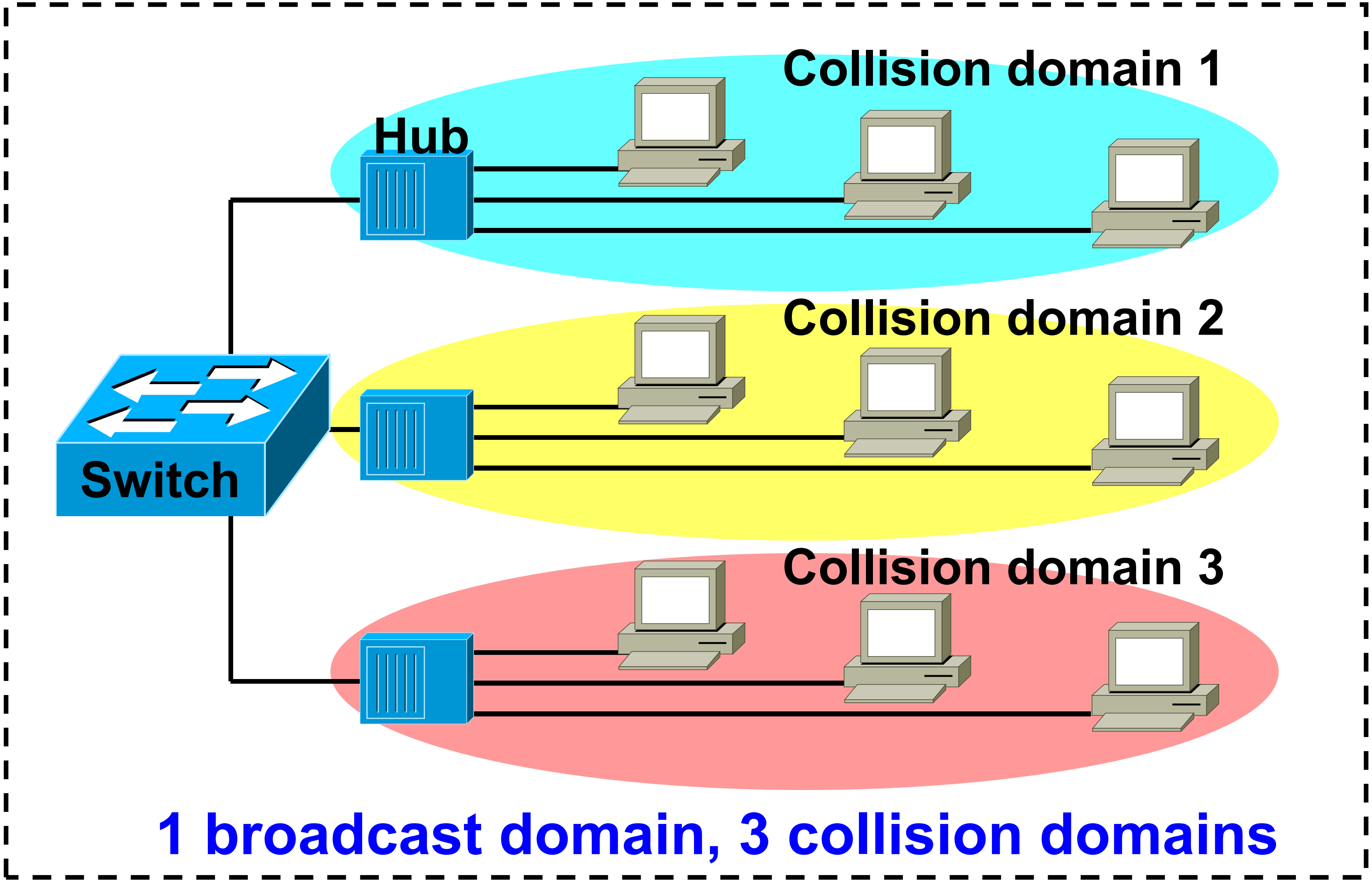
VLAN



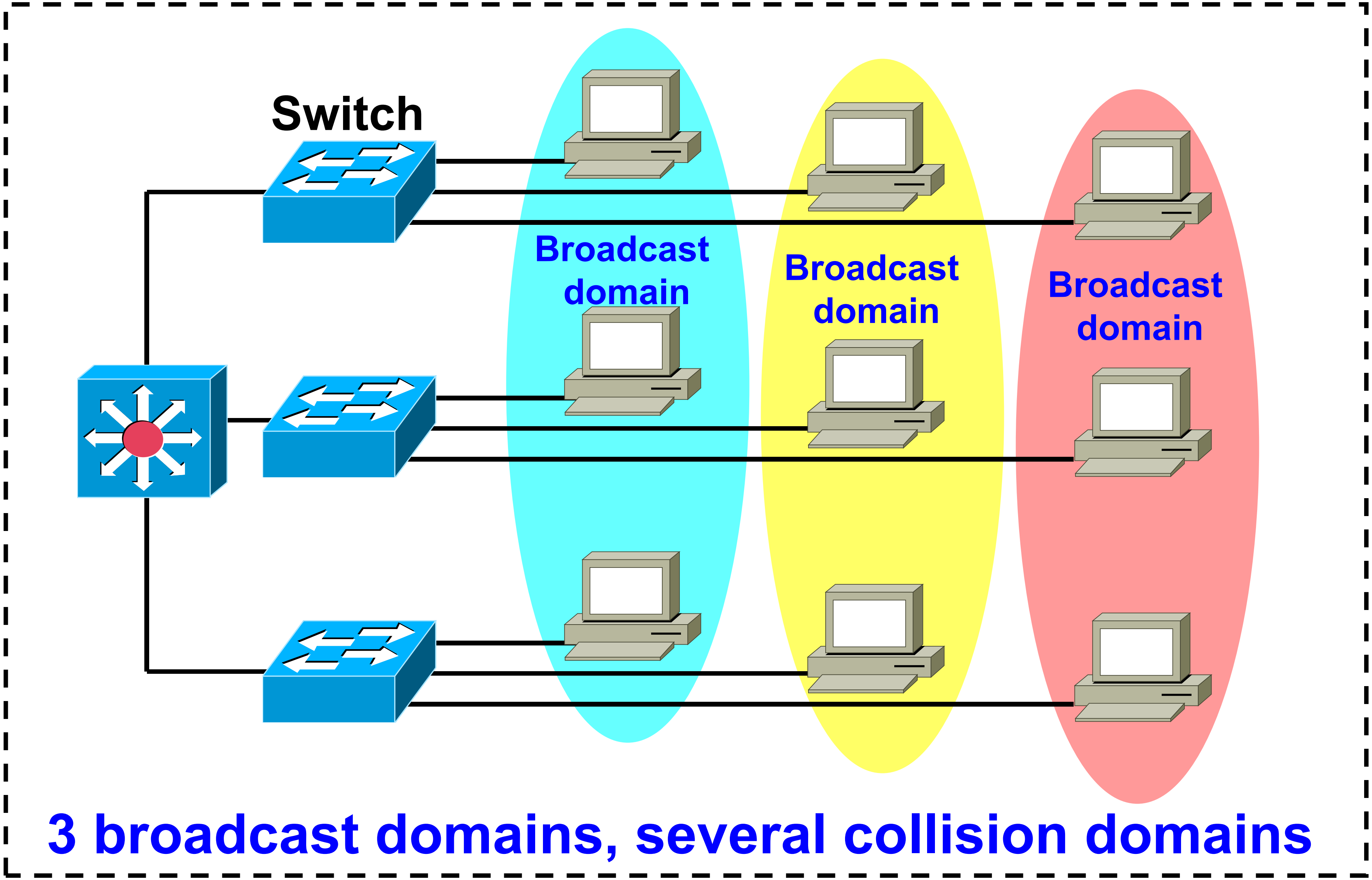
VLANs logically segment switched networks based on the functions, project teams, or applications of the organization regardless of the physical location or connections to the network.

VLAN

- **VLANs provide segmentation based on broadcast domains.**
- **Each VLAN is a broadcast domain created by one or more switches.**



Traditional LAN segmentation

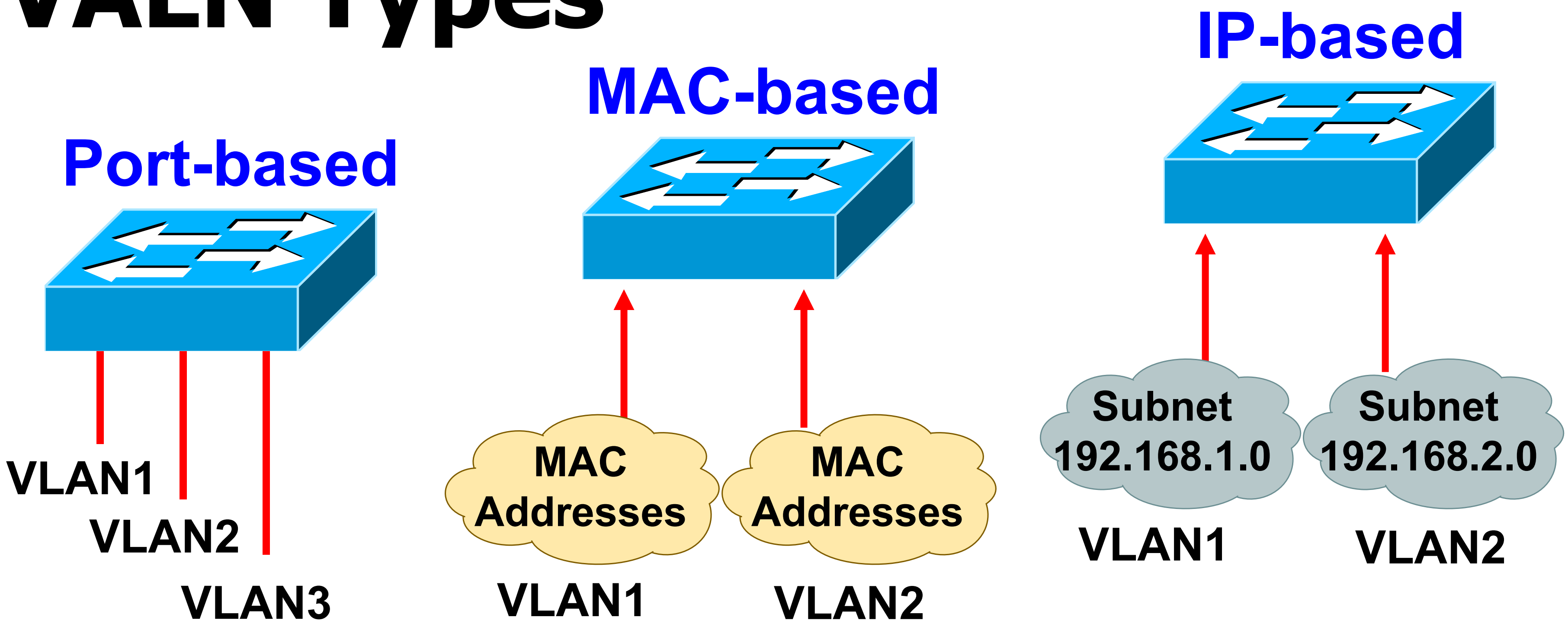


VLAN segmentation

VALN Types

- Port-based
  - Most common configuration method
- Protocol-based
- MAC-layer grouping
- Network-layer grouping
- Multicast grouping
- Application grouping
- Policy grouping

VALN Types



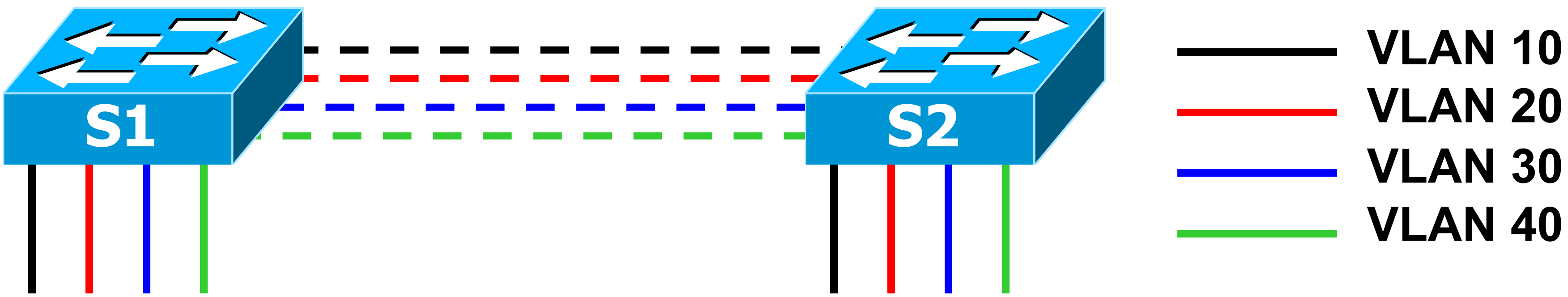
Port-based VLAN

- most common configuration method.
- Port assigned individually, in group or across more switches.
- Simple to use.

VLAN Trunk

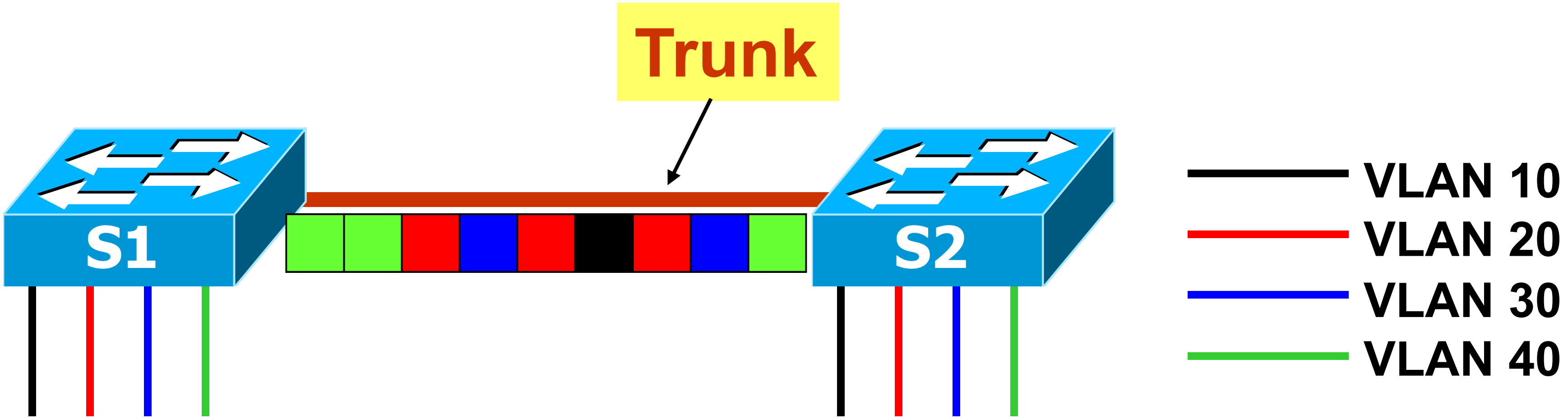
- A trunk is a point-to-point link between one or more Ethernet switch interfaces and another networking device, such as a router or a switch.
- Ethernet trunks carry the traffic of multiple VLANs over a single link.
- A VLAN trunk allows you to extend the VLANs across an entire network.

VLAN Trunk



When not use trunk,
4 switch ports needed, one for each VLAN

VLAN Trunk



When use trunk,

1 switch ports for 4 VLANs.

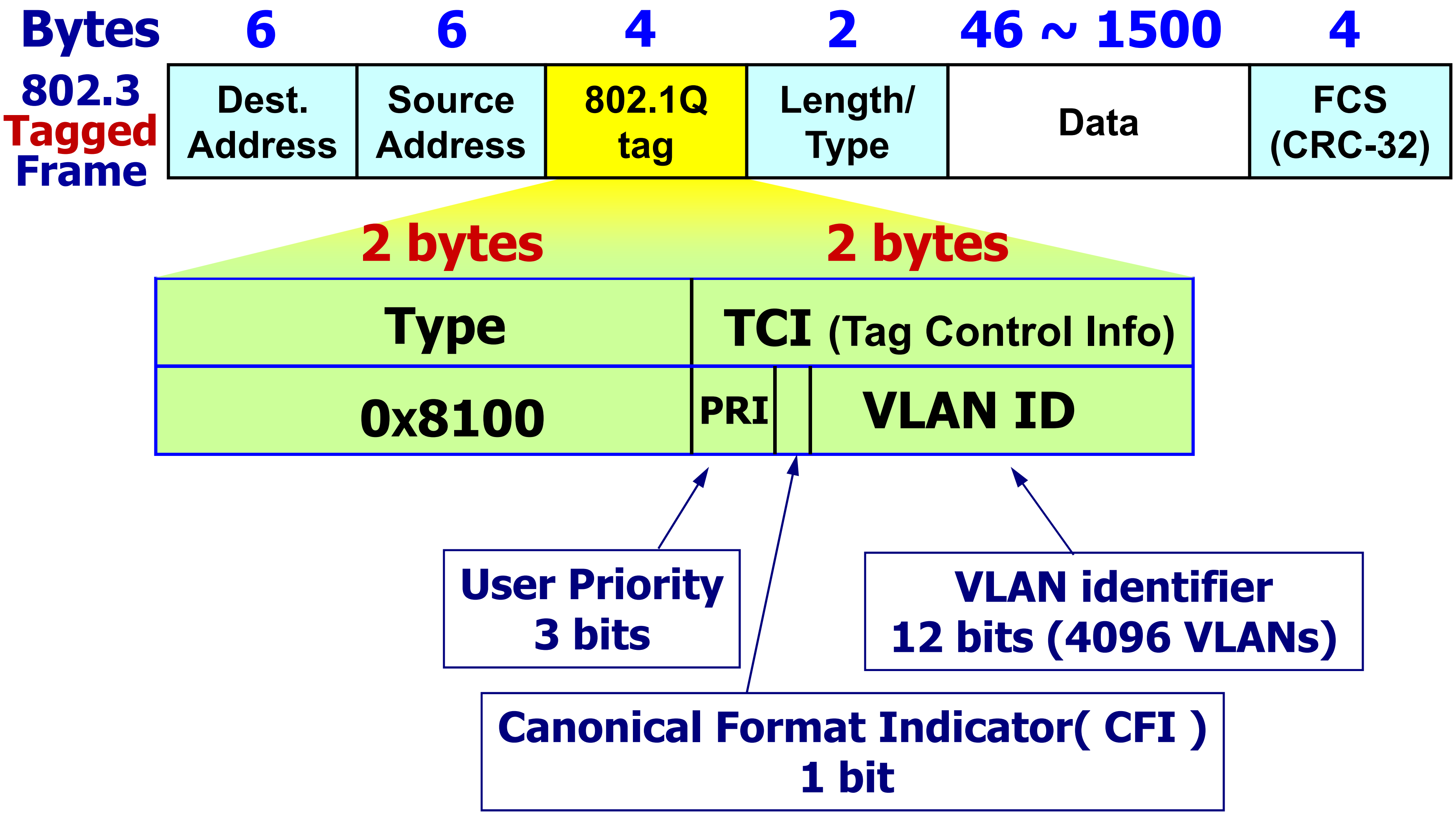
Problem: how can S1 and S2 know which VLAN the traffic is in and intended for?

VLAN Trunk - 802.1Q Frame tagging

- **VLAN Tagging is used when a link needs to carry traffic for more than one VLAN.**

There are two major methods of frame tagging, Cisco proprietary Inter-Switch Link (ISL) and IEEE 802.1Q.

VLAN Trunk - 802.1Q Frame tagging



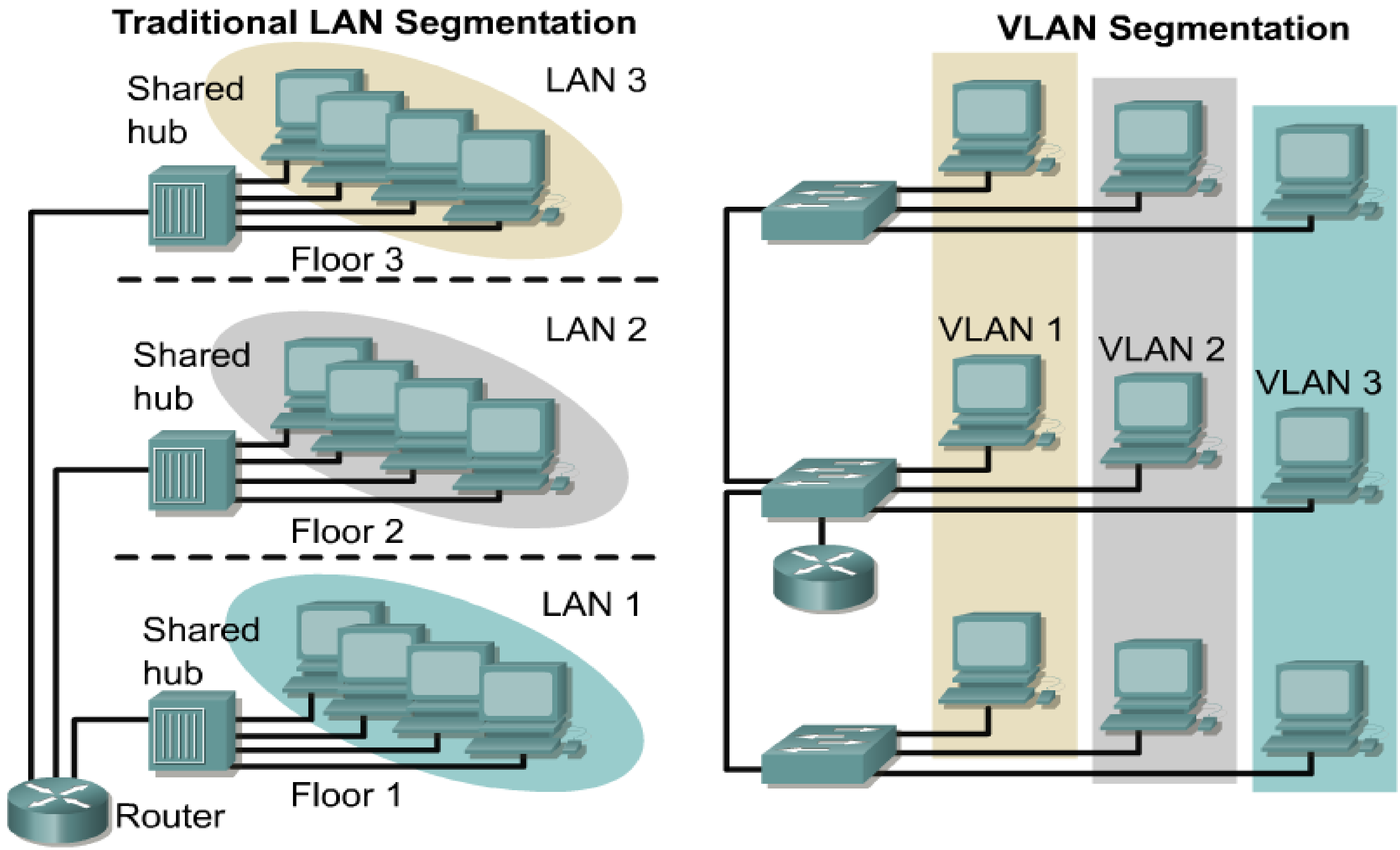
VLAN Trunk - 802.1Q Frame tagging

- The tag is automatically inserted into a frame by the switch when the frame needs to be forwarded to another switch.
- Because a host does not know anything about VLAN, the VLAN tag must be removed by a switch before the frame is forwarded to a host.

VTP

- VLAN Trunking Protocol
- VTP reduces the complexity of managing and monitoring VLAN networks
- VTP maintains VLAN configuration consistency across a common network administration domain
- VTP allows VLANs to be trunked over mixed media
- VTP provides for accurate tracking and monitoring of VLANs
- VTP provides “Plug-and-Play” configuration when adding new VLANs

Review of VLAN Basics



VLANs allow to group devices together, regardless of their physical location

VLAN Review

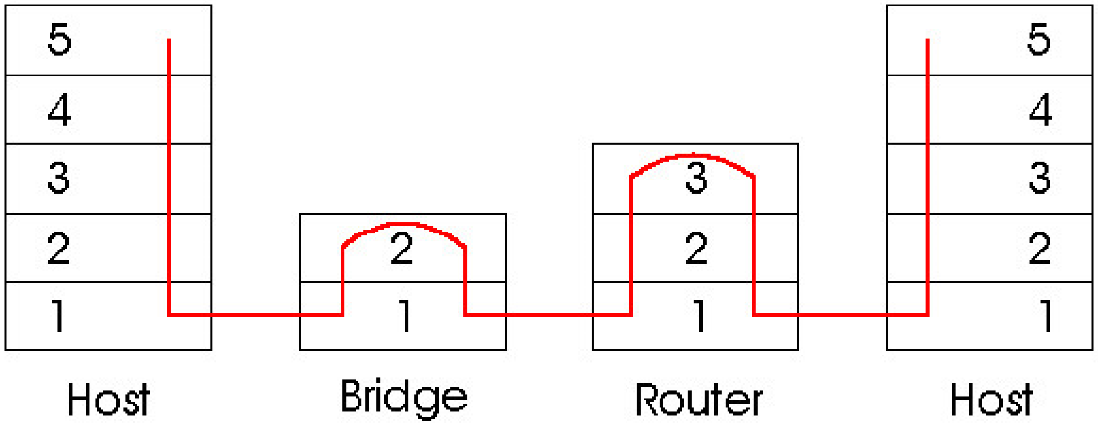
- **A VLAN is a *logical grouping* of devices or users that can be grouped by function, department, or application regardless of their physical location.**
- **VLANs are configured at the switch through *software*.**
- **VLANs can span single building infrastructures or interconnected buildings.**

More VLAN Review…

- **Network devices in different VLANs cannot directly communicate without the intervention of a Layer 3 routing device.**
- **A router is necessary to route the traffic between VLANs**
  - **Without the routing device, inter-VLAN traffic would not be possible**
  - **Put another way…when a host on one VLAN wants to communicate with a host on another, *a router must be involved***

Switches vs. Routers

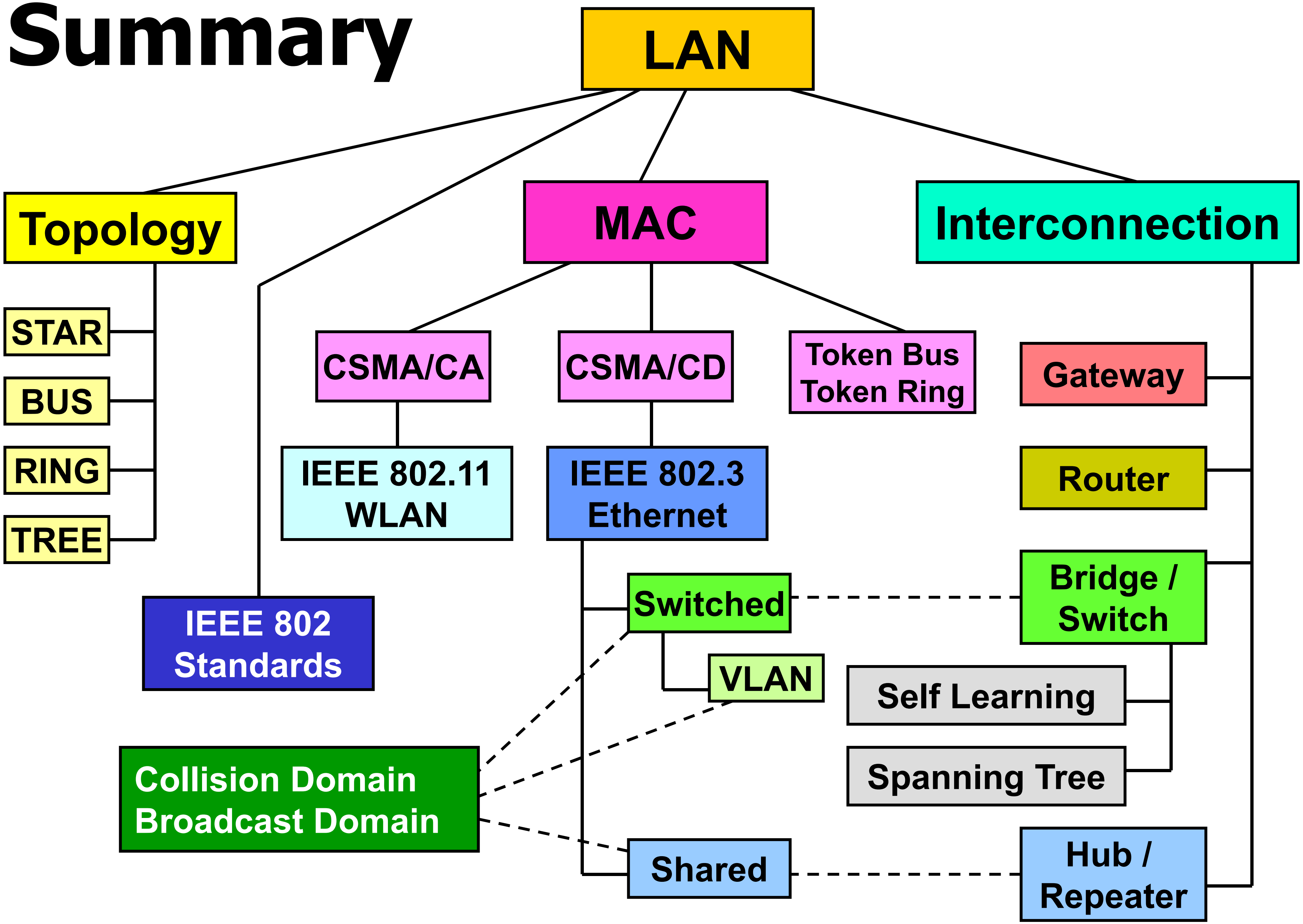
- both store-and-forward devices
  - routers: network layer devices (examine network layer headers)
  - switches are link layer devices
- routers maintain routing tables, implement routing algorithms
- switches maintain switch tables, implement filtering, learning algorithms



Hub, Switch and Router

	hubs	routers	switches
traffic isolation	no	yes	yes
plug & play	yes	no	yes
optimal routing	no	yes	no
cut through	yes	no	yes

Summary


LAN
Topology
STAR
BUS
RING
TREE
MAC
CSMA/CA
CSMA/CD
Token Bus
Token Ring
IEEE 802.11
WLAN
IEEE 802.3
Ethernet
Interconnection
Gateway
Router
Bridge / Switch
Hub / Repeater
IEEE 802 Standards
Switched
VLAN
Shared
Self Learning
Spanning Tree
Collision Domain
Broadcast Domain