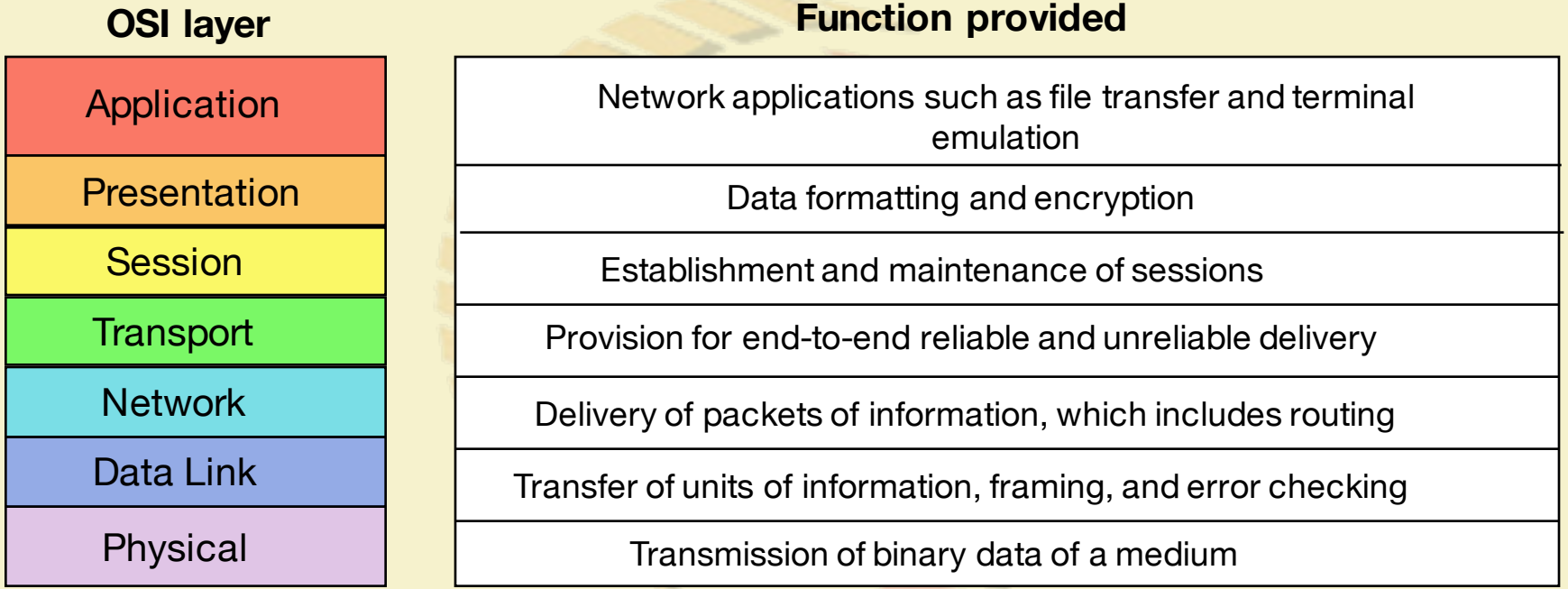
COMPUTER NETWORKS AND INTERNET PROTOCOLS

Data Link Layer - Overview

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

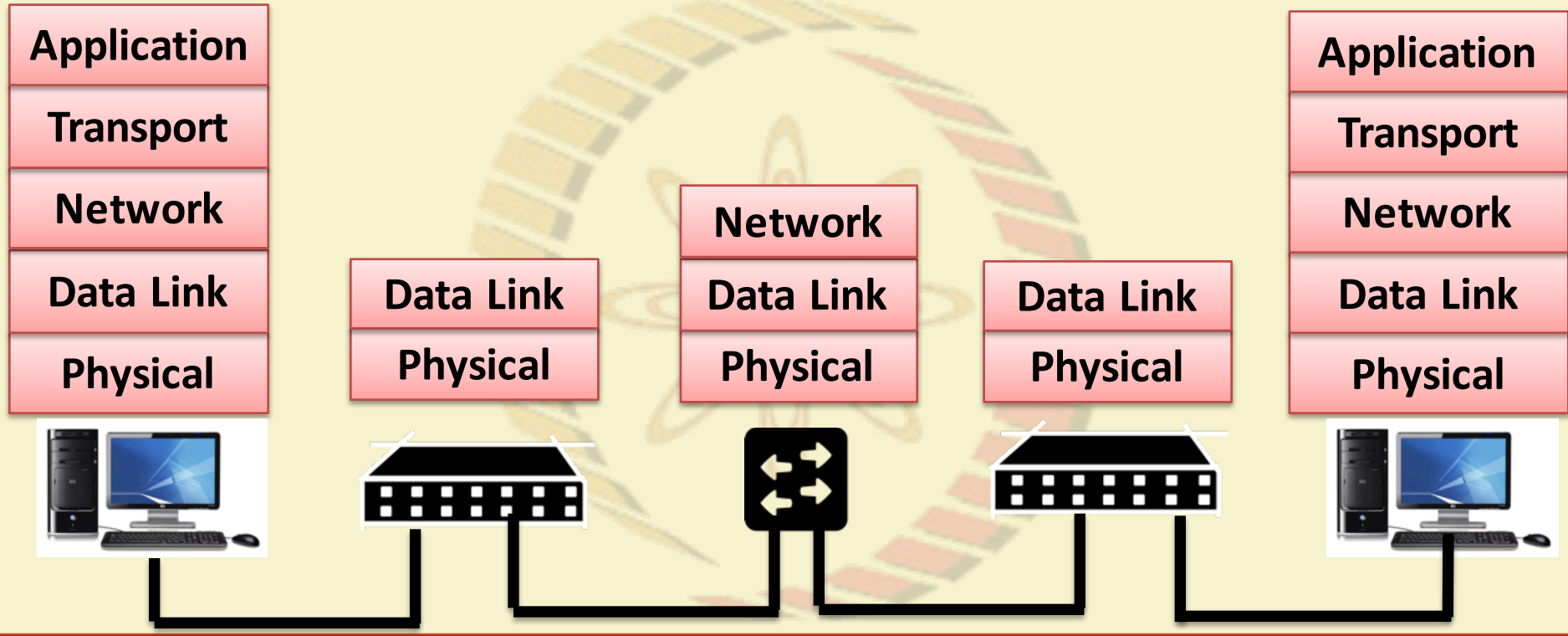
OSI Model Layers



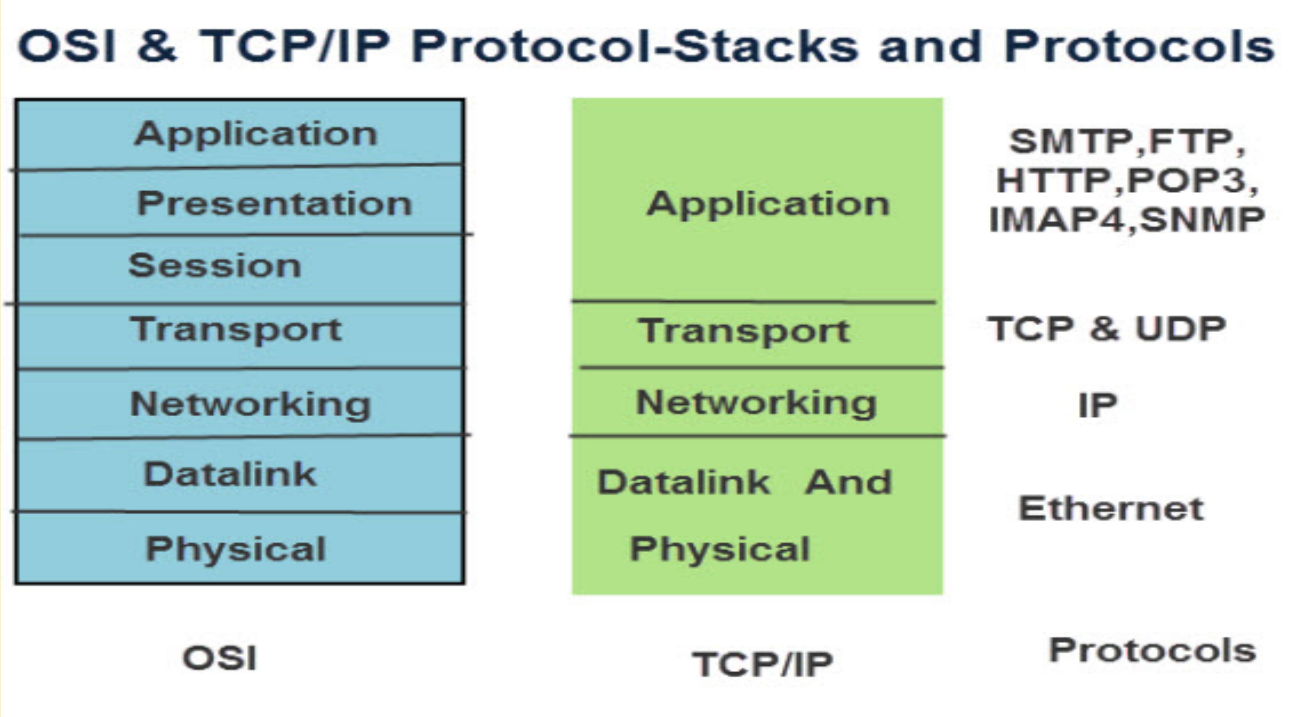
OSI layer	Function provided
Application	Network applications such as file transfer and terminal emulation
Presentation	Data formatting and encryption
Session	Establishment and maintenance of sessions
Transport	Provision for end-to-end reliable and unreliable delivery
Network	Delivery of packets of information, which includes routing
Data Link	Transfer of units of information, framing, and error checking
Physical	Transmission of binary data of a medium

Ref: Data and Computer Communications, W. Stallings; Local and Metropolitan Area Networks, W. Stallings; Data Communications and Networking, B.A. Forouzan; Local & Metropolitan Area Networks, L. Christofi; TCP/IP Tutorials, IBM Redbooks

Data Transfer between Two Remote Machines



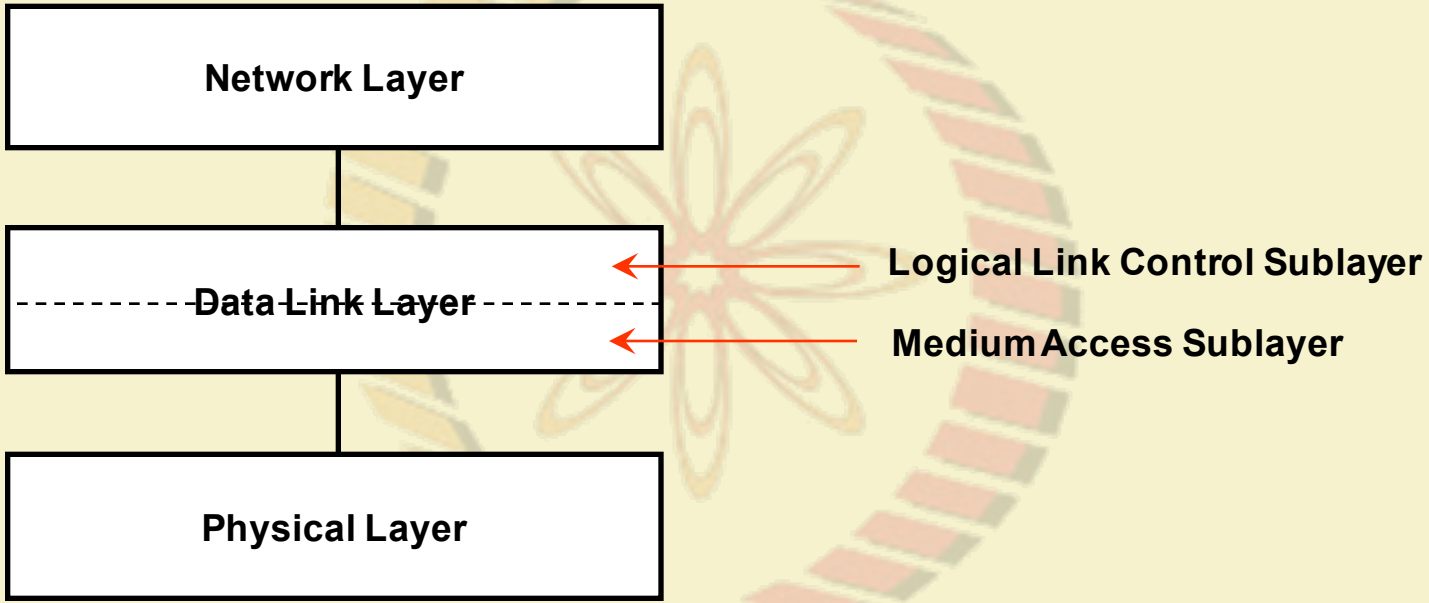
OSI and TCP/IP



Data Link Layer

- Network Layer
- **Data Link layer**
  - **Medium Access Control (MAC) layer**
  - **Logical Link Control (LLC) layer**
- Physical layer

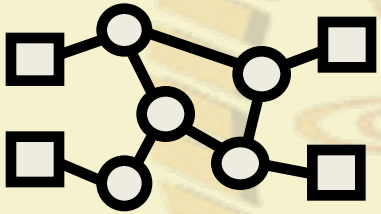
Data Link Layer



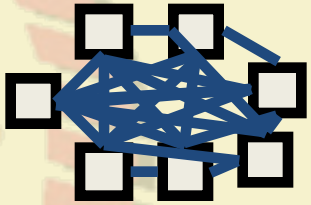
Issue 1 – Sharing a Wire

Learned how to connect hosts

- … But what if we want more hosts?

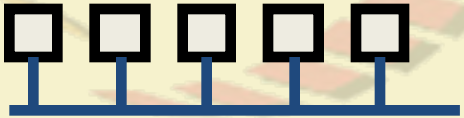


Switches

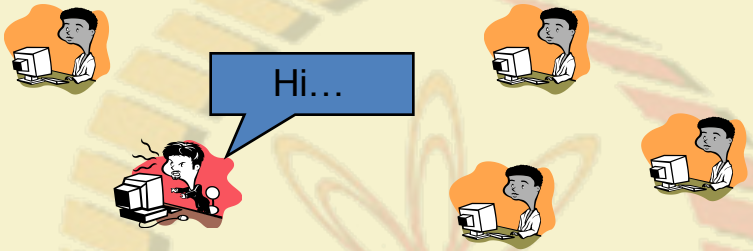


Wires for everybody!

- How to share a wire?

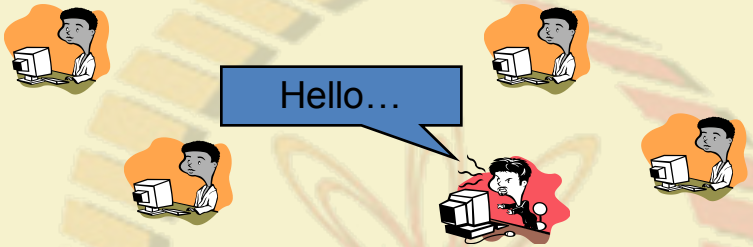


Issue 2 – Listen and Speak



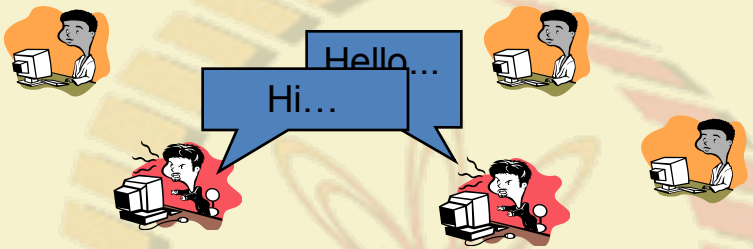
- Listen before you Speak…

Problem 2 – Listen and Speak



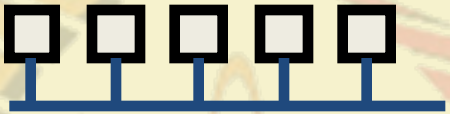
- Listen before you Speak…

Problem 2 – Listen and Speak



- Simultaneous Speaking!

Issue 3 – Recipient of the packet?

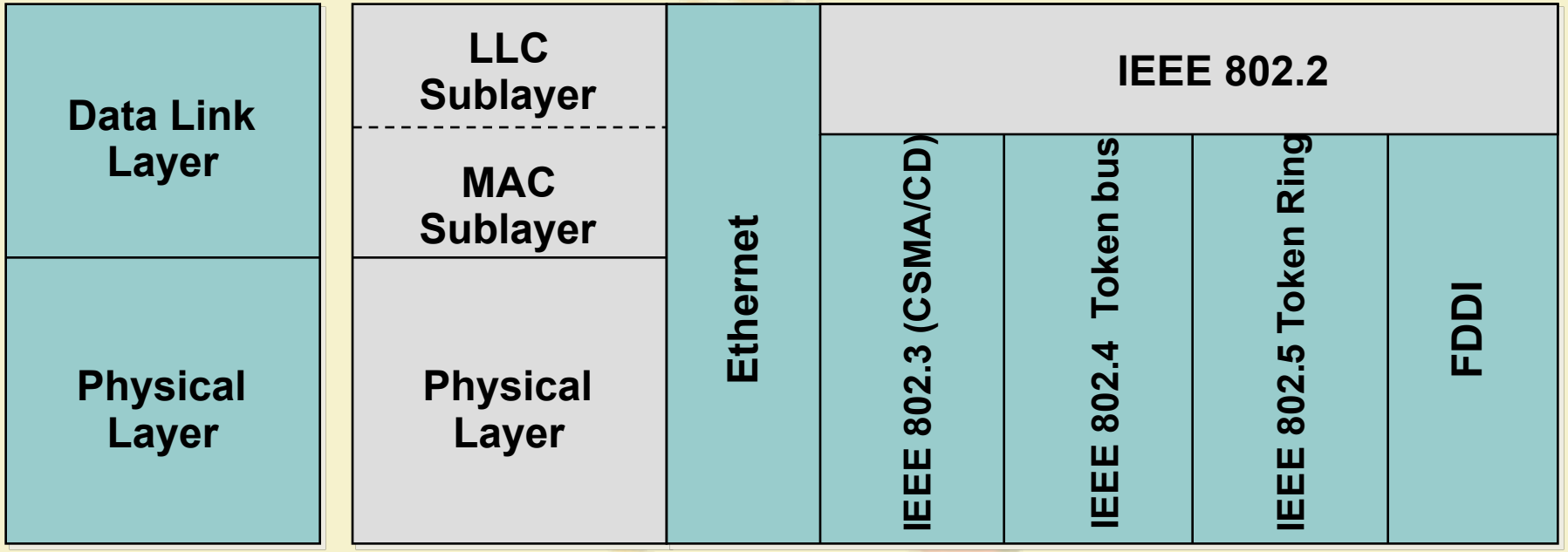


- Need to put an address on the packet
- Format?
- How do you know what address you want to send it to?

Medium Access Control

- Medium access (MAC) sub-layer is primarily used in broadcast or shared channel networks
- MAC protocols enable two stations (or nodes) using a shared communication resource to establish, maintain and terminate a connection.
- Examples: *Satellite, Ethernet, Cellular*

IEEE 802 Standard



OSI, TCP/IP Layers

LAN Specification

IEEE 802 Standard

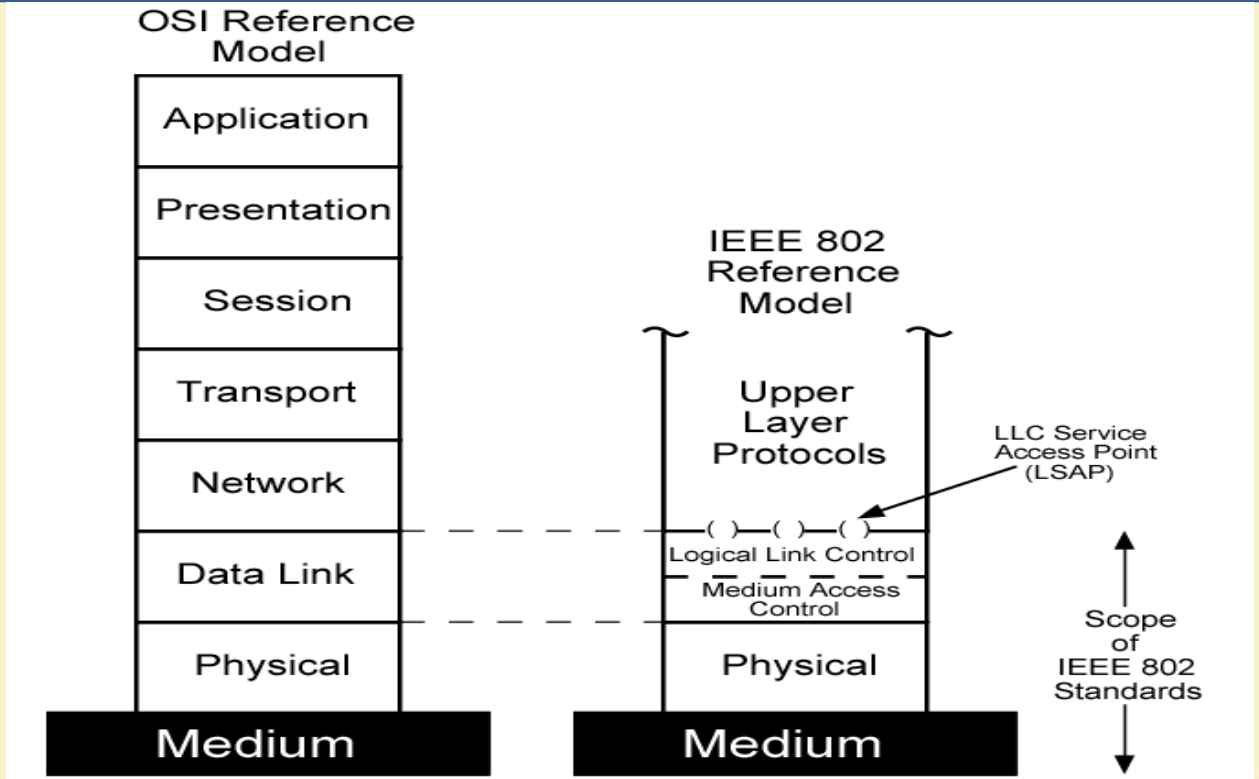
802.2: Logical Link Control (LLC)

802.3: CSMA/CD (Ethernet)

802.5: Token Ring

802.11: Wireless LANs

IEEE 802 Scope

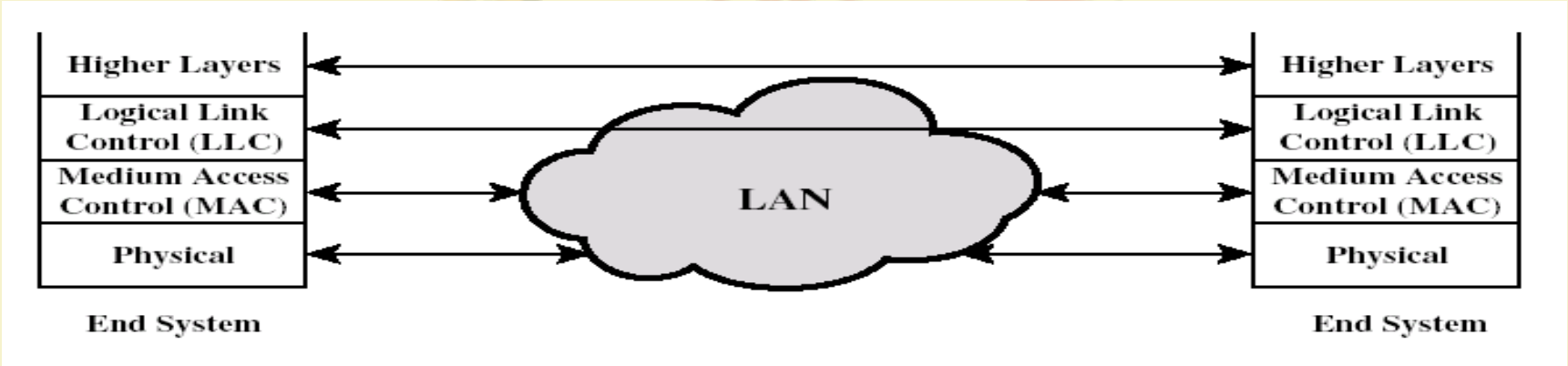


802 Layers functions

- **Physical**
  - Encoding/decoding
  - Preamble generation/removal
  - Bit transmission/reception
  - Transmission medium and topology
- **Logical Link Control**
  - Interface to higher levels
  - Flow and error control
- **Medium Access Control**
  - Data assembly and dismantle into frames
  - Govern access to LAN transmission medium

Scope of LAN protocols

- Two stations (End Systems) that communicate via a shared medium (LAN)
- Higher layers (above LLC) provide end-to-end services between the stations
- Below the LLC layer, the MAC provides the necessary logic for gaining access to the network



Logical Link Control (LLC)

- LLC layer for LANs is concerned with the transmission of a link-level protocol data unit (PDU) between two stations, without the necessity of an intermediate switching node
- Two characteristics:
  - It must support the multi-access, shared medium nature of the link
  - It is relieved from some details of link access by the MAC layer

LLC Services

- **Unacknowledged connectionless service**
  - Datagram-style
  - Does not involve any flow and error control mechanisms
  - Data delivery is not guaranteed
- **Connection mode service**
  - A logical connection is set up between two stations
  - Flow and Error control
- **Acknowledged connectionless service**
  - A cross between the previous services
  - Datagrams are to be acknowledged
  - No prior logical connection is set up

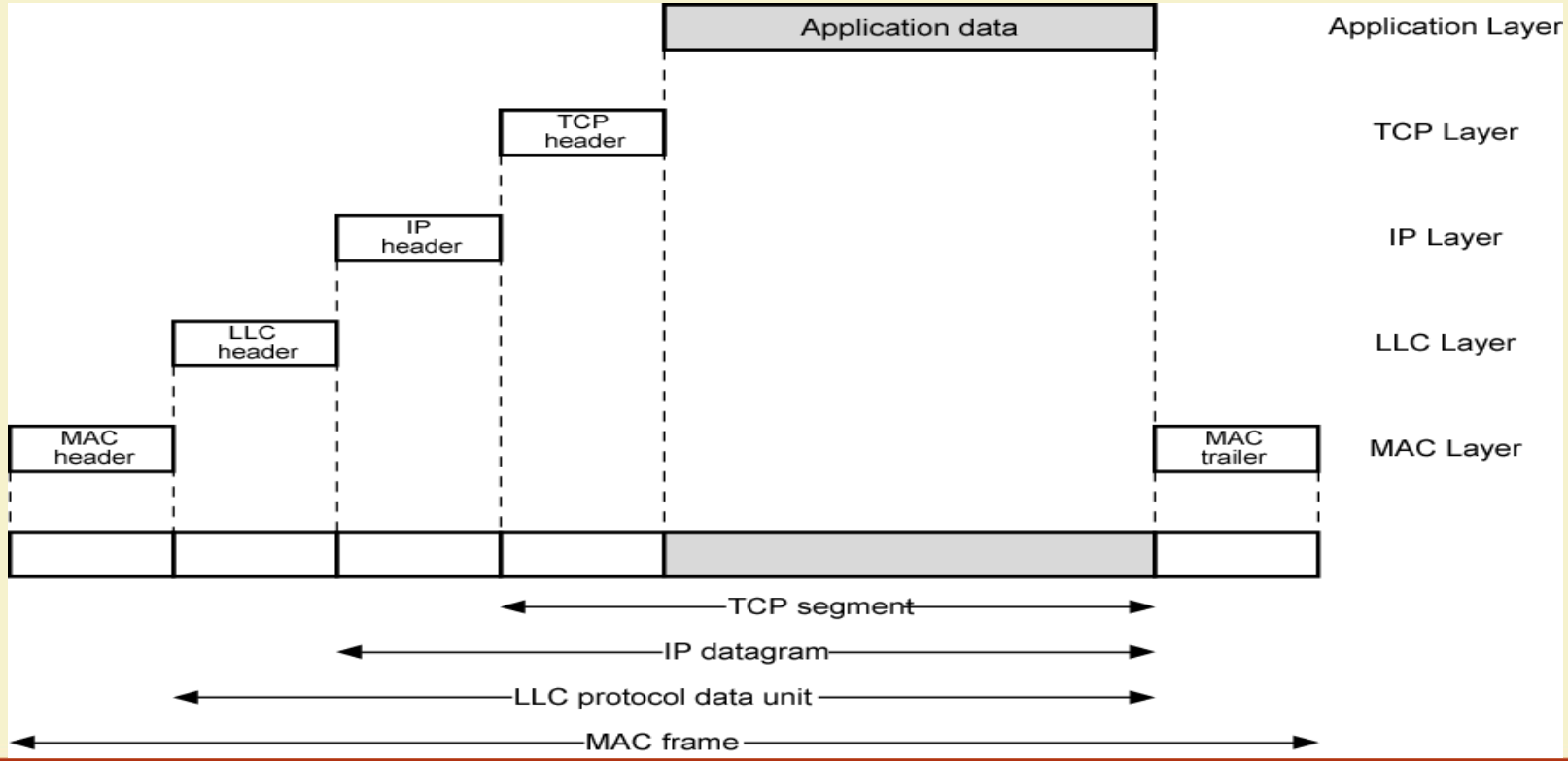
Medium Access Control (MAC)

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

Medium Access Control (MAC)

- All LANs consist of a collection of devices that must share the network's transmission capacity
- Function of a **Medium Access Control (MAC)** protocol: Some means of controlling access to the transmission medium is needed for efficient use of that capacity.
- Key parameters in any MAC technique are ***where*** and ***how***.
  - ***Where***, refers to whether control info is exercised in a centralized or distributed fashion.
    - **Centralized:** a controller has the authority to grant access to the network
    - **Distributed:** the stations collectively perform a MAC function to determine dynamically the order in which stations transmit
  - ***How***, is constrained by the topology and is a trade-off among competing factors, such as cost, performance and complexity

LAN Protocols – Messages




thank you!

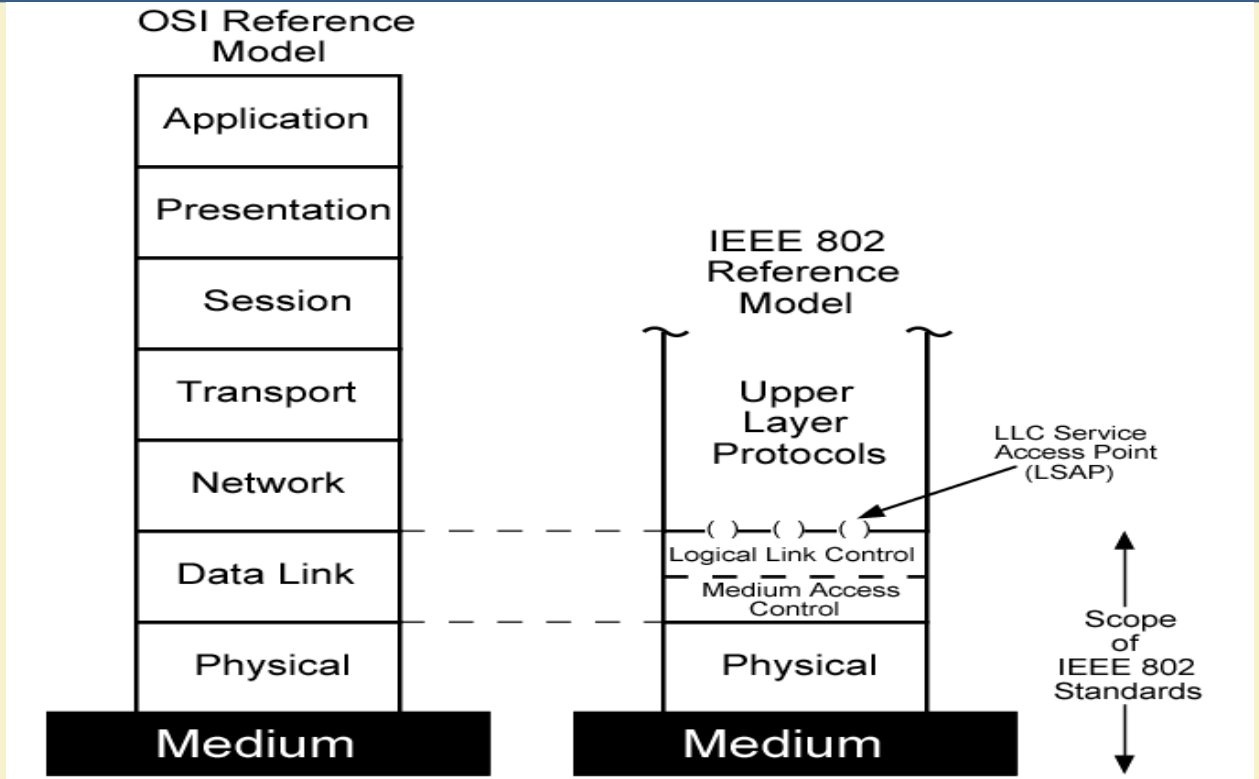
COMPUTER NETWORKS AND INTERNET PROTOCOLS

Data Link Layer - Overview

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

IEEE 802 Scope



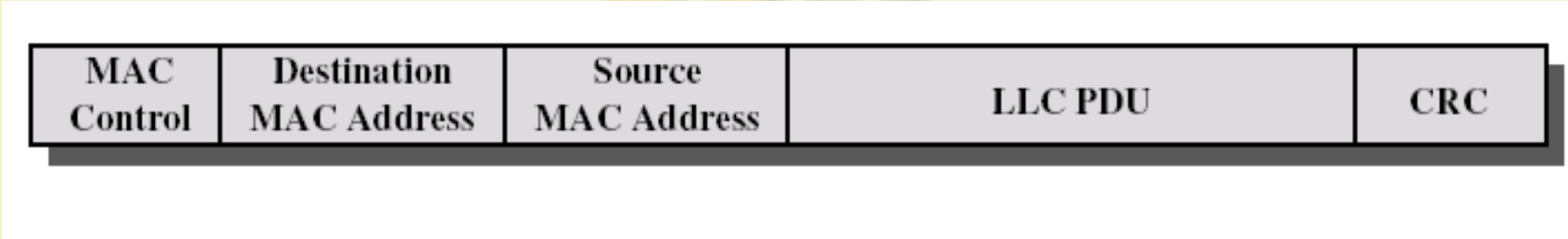
Logical Link Control (LLC)

- LLC layer for LANs is concerned with the transmission of a link-level protocol data unit (PDU) between two stations, without the necessity of an intermediate switching node
- Two characteristics:
  - It must support the multi-access, shared medium nature of the link
  - It is relieved from some details of link access by the MAC layer

Medium Access Control (MAC)

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

Generic MAC frame format



- **MAC control:** contains control info for the functioning of the MAC protocol, e.g. priority level
- **Destination MAC address:** the destination physical attachment point on the LAN for this frame
- **Source MAC address:** the source physical attachment point on the LAN for this frame
- **LLC:** The LLC data from the next higher layer
- **CRC:** Cyclic Redundancy Check field, used to check if a transmission error has occurred

MAC Techniques

- **Synchronous**
  - A specific capacity is dedicated to a connection
  - Same approach as in circuit-switching FDM or TDM, so not optimal for LANs/MANs because the needs of the stations are unpredictable
- **Asynchronous**
  - Capacity is allocated in a dynamic fashion, in response to demand
  - Subdivided into three categories
    - Round Robin
    - Reservation
    - Contention

Asynchronous MAC techniques

- **Round Robin**
  - Each station in turn is granted the right to transmit
  - After each station finishes transmitting, it passes the right to transmit to the next station in logical sequence
  - Efficient technique when many stations have data to transmit over an extended period of time
- **Reservation**
  - For stream traffic (voice, bulk file transfer etc)
  - Time on the medium is divided into slots, like synchronous TDM
  - A station whishing to transmit reserves slots for an extended period
- **Contention**
  - For bursty traffic (short, sporadic transmissions such as interactive terminal-host traffic)
  - No control is exercised to determine whose turn it is
  - Simple to implement and efficient for light loads

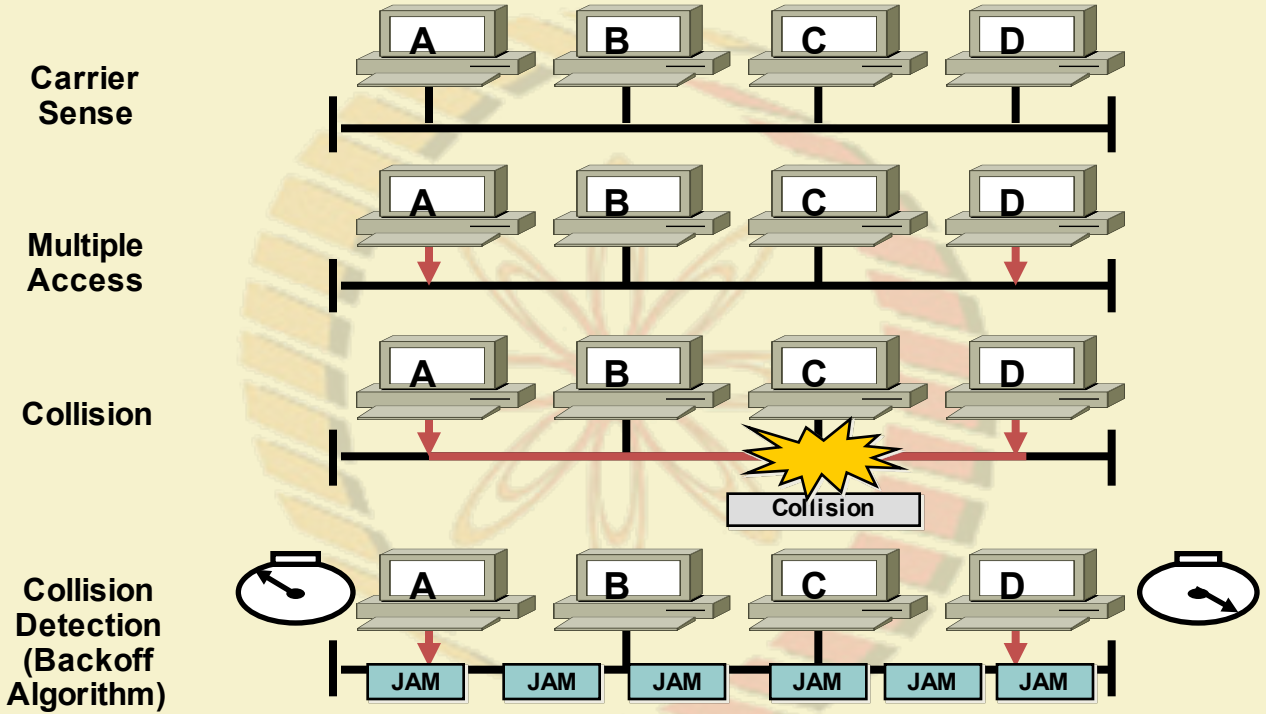
Medium Access Control methods

- Methods used for **Medium Access Control** are:
  - **Carrier-sense multiple-access with collision detection (CSMA/CD)** for ***bus*** topologies
  - **Control token** or **Token Passing** for ***bus*** and ***ring*** topologies

CSMA/CD

- CSMA/CD is used only in **bus** type networks, where a number of nodes share a common communication channel (wire) known as the bus.
- CSMA/CD is used in traditional **Ethernet**

CSMA/CD



Carrier sense multiple access collision detect (CSMA/CD)

CSMA/CD operation

Basic operation of CSMA/CD is as follows:

- To transmit data, the **source station** assembles a packet comprising of the destination address, the data and control info
- The source station **listens** to the cable to determine if the bus is currently in use. If so, it waits until the bus is free, else it transmits the packet. This operation is known as **carrier sensing.**
- During transmission, the source station continues to listen to the cable to detect if another station has also initiated a transmission thus causing a **collision**. This process is known as **collision detection**.
- If a collision is detected then, to ensure all stations are aware of the collision, the source station transmits a random bit pattern known as the **jam sequence**.
- Stations involved in a collision then **back off** for a random period before retrying for packet transmission.

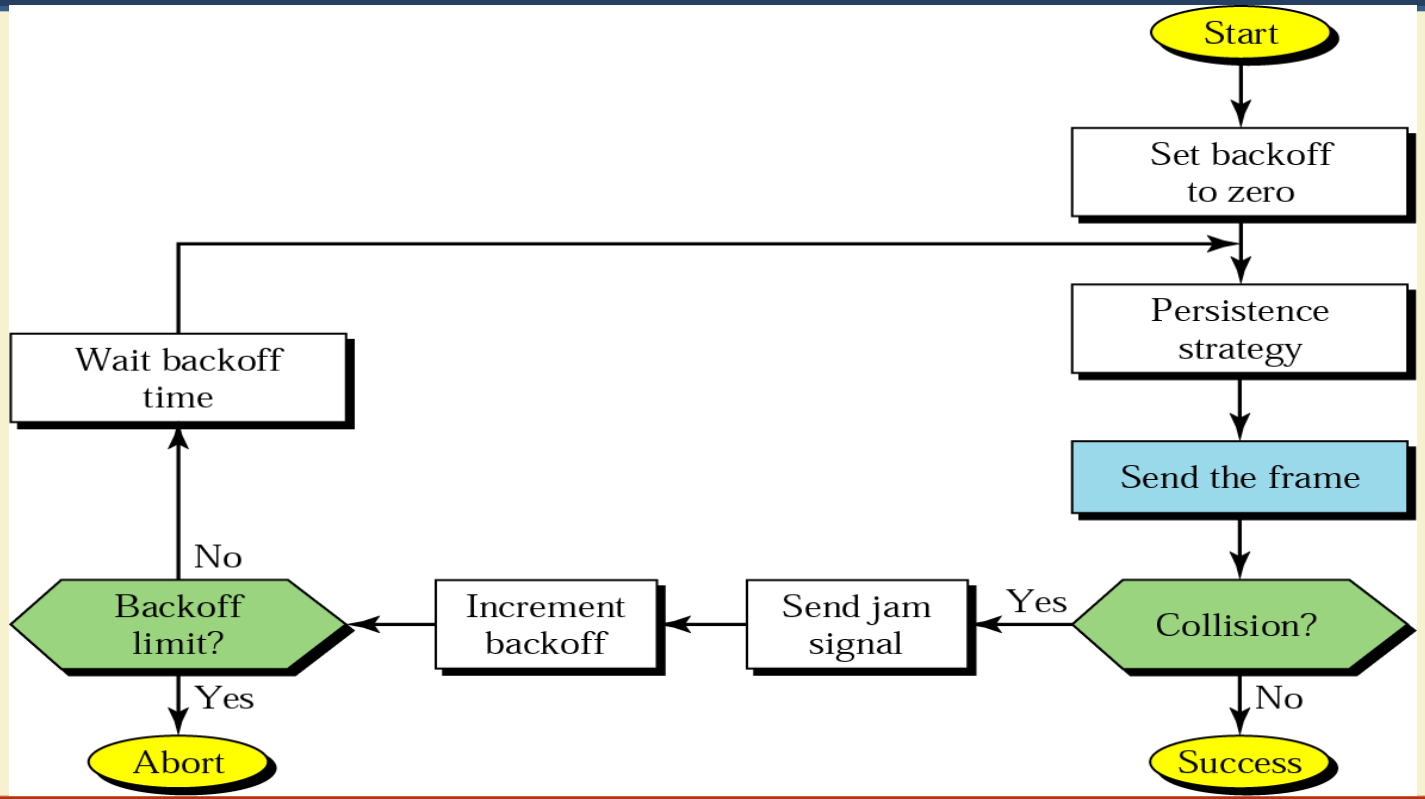
CSMA/CD procedure

Sense the channel

- If idle, transmit immediately
- If busy, wait until the channel becomes idle

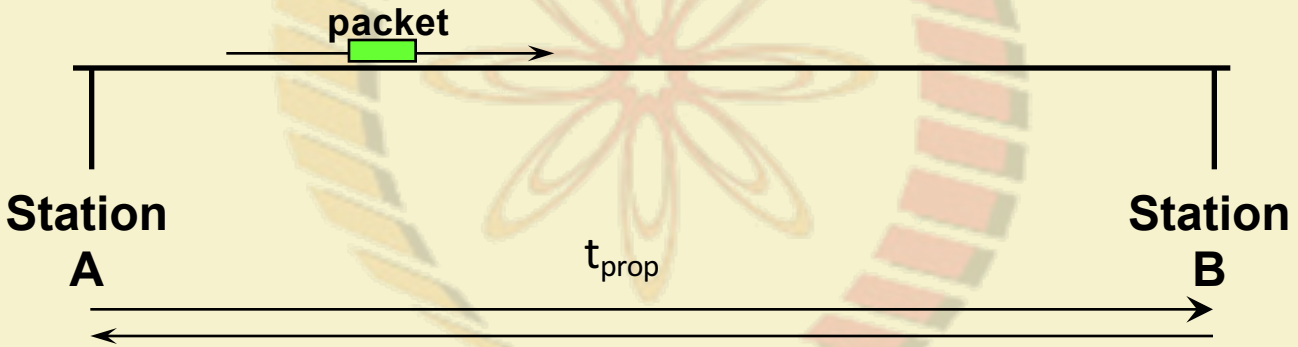
Collision detection

- Abort a transmission immediately if a collision is detected
- Try again later after waiting a random amount of time



Collision detection time

- How long does it take to realize there has been a collision?
- Worst case: 2 x end-to-end propagation delay



Control Token or Token Passing

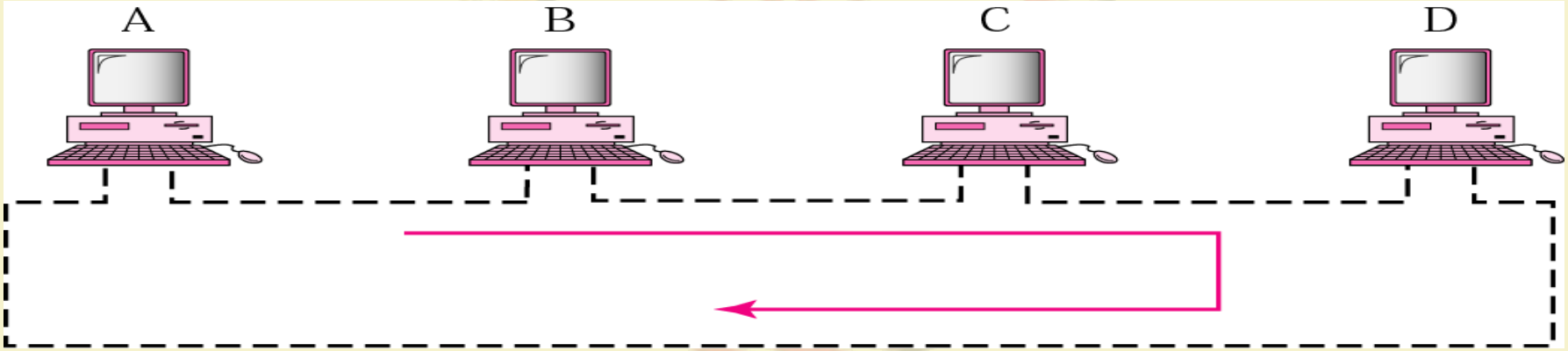
- Another way of controlling access to a shared transmission medium is by a control token (Token Passing)
- The Control Token technique uses a **control** or **permission token** to share the communication resource between a number of nodes. The technique can be applied to both **bus** and **ring** network topologies.
- This token is passed from one station to another according to a defined set of rules
- A station may transmit a frame only when it has possession of the token and after it has transmitted the frame, it passes the token on, to allow another station to access the transmission medium

Control Token Operation

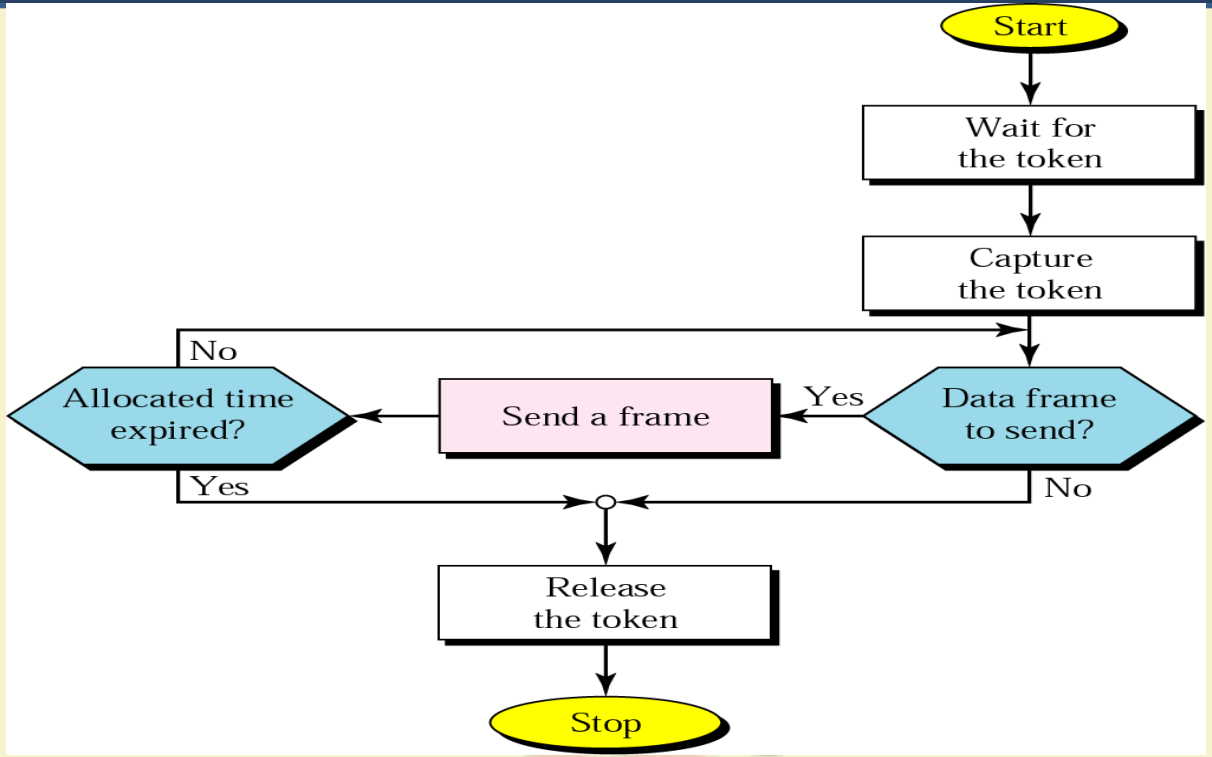
- Whether using a ring or bus topology, a **logical ring** is established which links all the nodes using the physical medium (see next two slides)
- A single control (permission) token is created at one of the nodes
- The token is passed from node to node around the logical ring until it arrives at a node waiting to send a frame
- The node **captures** the token and transmits the frame
- Upon completing transmission, the node releases the token to the next node in the logical ring

Token passing network

- A token always circulates around a ring net.
- A user grabs a token to transmit data



Control Token procedure

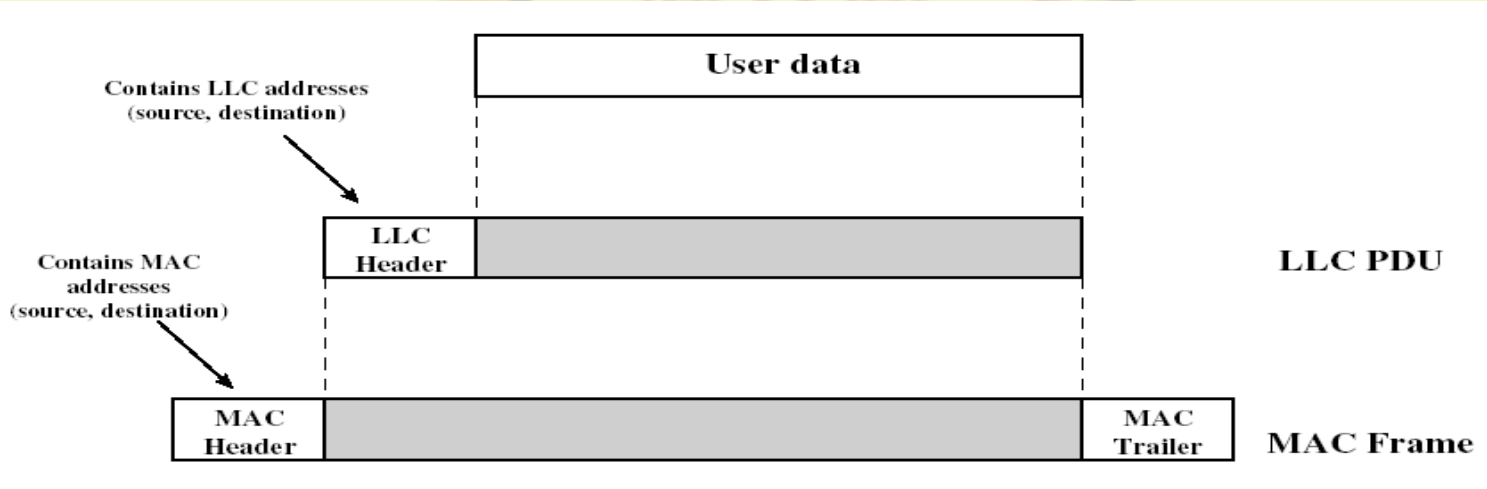


LAN addressing

- Communication involves three agents:
  - Processes
  - Stations
  - Networks
- Transfer of data from one process to another involves
  - getting the data to the station in which the destination process resides, and then
  - getting it to the intended process within the computer
- Two levels of addressing
  - MAC address
  - LLC address

LAN protocol Control info and User data

- LLC user data (IP datagram) are passed down to LLC which appends a header, to make the LLC PDU
- LLC PDU is passed to the MAC entity, which in turn appends a header and a trailer, to make the MAC frame



Levels of addressing

- **MAC address**
  - Identifies a physical interface from the station to the LAN
  - There is one to one relationship between stations and physical addresses

- **LLC address**
  - Identifies an LLC user
  - LLC address (LLC SAP) is associated with a particular user within a station
  - LLC SAP may refer to a process executing on a station or to a hardware port

MAC Protocols: A Taxonomy

Three broad classes:

- **Channel partitioning**
  - Divide channel into smaller "pieces" (time slots, frequency)
  - Allocate piece to node for exclusive use
- **Random access**
  - Allow collisions
  - "Recover" from collisions
- **Shared Access**
  - Tightly coordinate shared access to avoid collisions

Goal: efficient, fair, simple, decentralized


thank you!

COMPUTER NETWORKS AND INTERNET PROTOCOLS

Data Link Layer - Ethernet

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

Shared Access Networks

- Shared Access Networks assume multiple nodes on the same physical link
  - Bus, ring and wireless structures
  - Transmission sent by one node is received by all others
  - No intermediate switches
- Methods for moderating access (MAC protocols)
  - Fairness
  - Performance

Random Access MAC Protocols

- When node has packet to send
  - Transmit at full channel data rate R
  - No *a priori* coordination among nodes
- 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 of random access MAC protocols:
  - ALOHA
  - Slotted ALOHA
  - CSMA and CSMA/CD

Aloha – Basic Approach

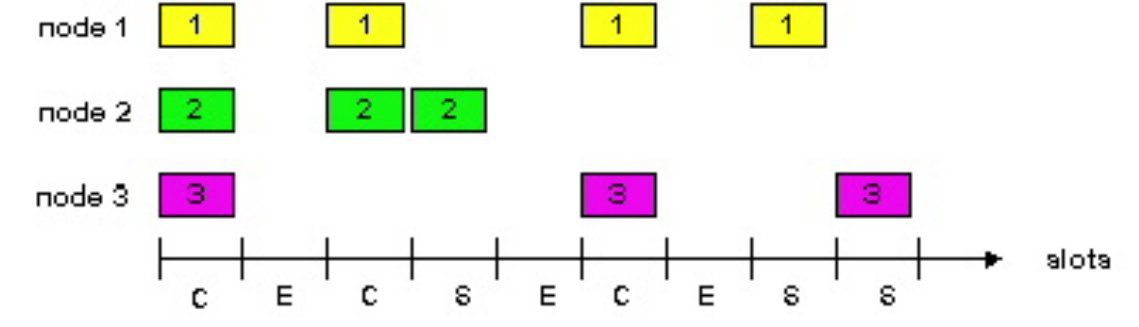
- First random MAC developed
  - For radio-based communication in Hawaii (1970)
- Basic idea:
  - When you're ready, transmit
  - Receiver's send ACK for data
  - Detect collisions by timing out for ACK
  - Recover from collision by trying after random delay
    - Too short → large number of collisions
    - Too long → underutilization

Aloha Networks - Overview

- Developed by Norm Abramson at Univ. of Hawaii for use with packet radio systems
  - Any station can send data at any time
  - Receiver sends an ACK for data
  - Timeout for ACK signals that there was a collision
    - What happens if timeout is poorly timed?
  - If there is a collision, sender will resend data after a random backoff
- Utilization (fraction of transmitted frames avoiding collision for N nodes) was pretty bad
  - Max utilization = 18%
- Slotted Aloha (dividing transmit time into windows) helped
  - Max utilization increased to 36%

Slotted Aloha

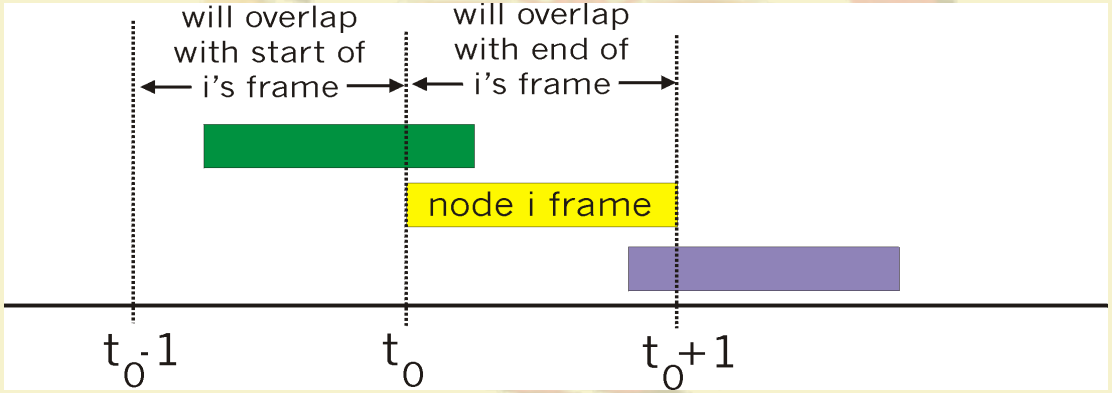
- Time is divided into equal size slots (i.e. packet transmission time)
- Node (w/ packet) transmits at beginning of next slot
- If collision: retransmit packet in future slots with probability p, until successful



Success (S), Collision (C), Empty (E) slots

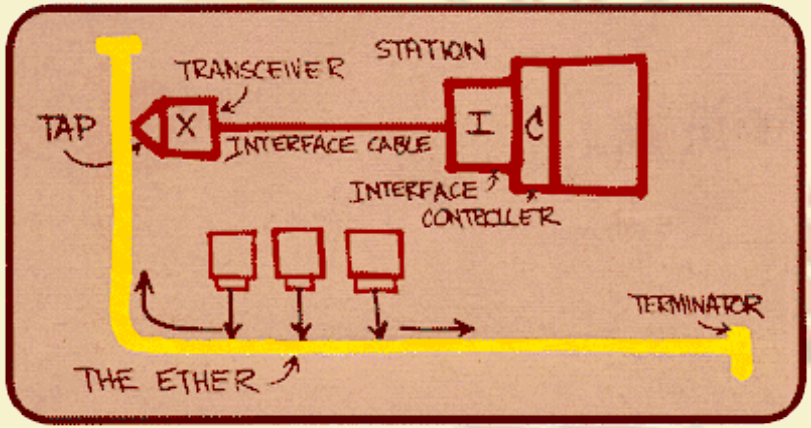
Pure (Unslotted) ALOHA

- Unslotted Aloha: simpler, no synchronization
- Packet needs transmission:
  - Send without awaiting for beginning of slot
- Collision probability increases:
  - Packet sent at t_0 collide with other pkts sent in $[t_0-1, t_0+1]$



Ethernet

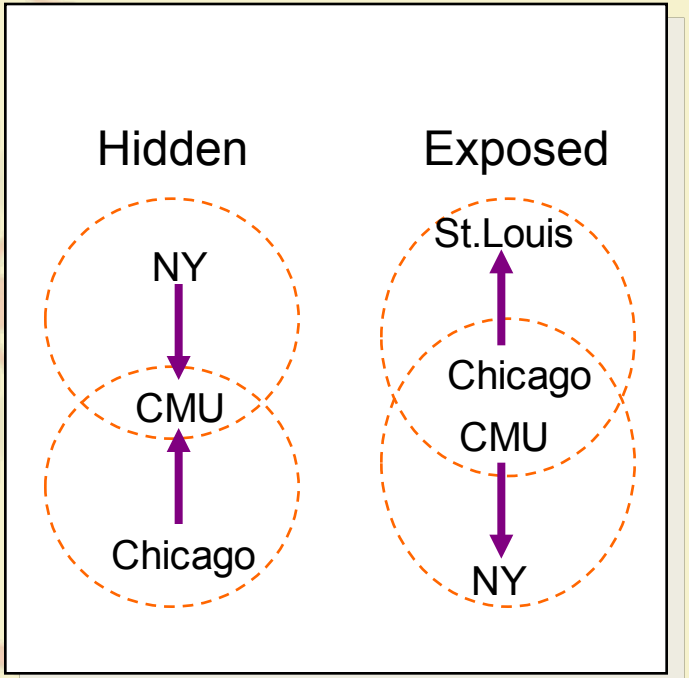
- First practical local area network, built at Xerox PARC in 70's
- "Dominant" LAN technology:
  - Cheap
  - Kept up with speed race: 10, 100, 1000 Mbps



Metcalfe's Ethernet sketch

Ethernet MAC – Carrier Sense

- Basic idea:
  - Listen to wire before transmission
  - Avoid collision with active transmission
- Why didn't ALOHA have this?
  - In wireless, relevant contention at the *receiver*, not sender
    - Hidden terminal
    - Exposed terminal



Multiple Access Methods

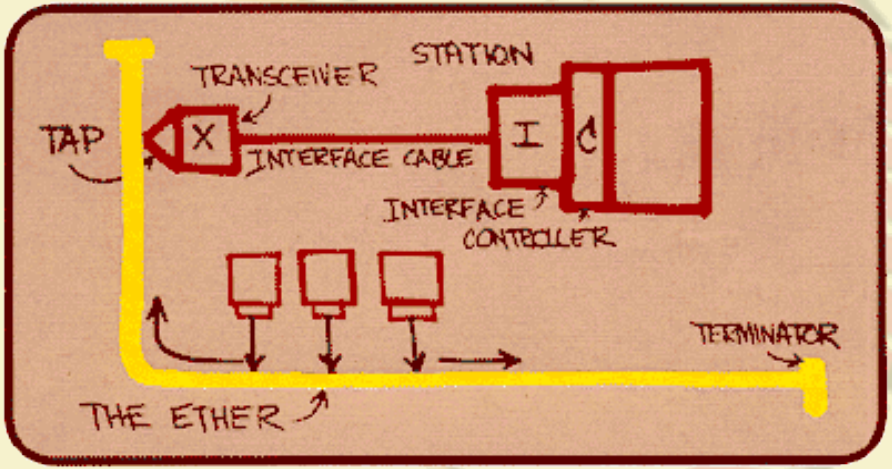
- Fixed assignment
  - Partition channel so each node gets a slice of the bandwidth
  - Essentially circuit switching – thus inefficient
  - Examples: TDMA, FDMA, CDMA (all used in wireless/cellular environments)
- **Contention-based**
  - Nodes contends equally for bandwidth and recover from collisions
  - Examples: Aloha, Ethernet
- Token-based or reservation-based
  - Take turns using the channel
  - Examples: Token ring

Ethernet

- Background
  - Developed by Bob Metcalfe and others at Xerox PARC in mid-1970s
  - Roots in Aloha packet-radio network
  - Standardized by Xerox, DEC, and Intel in 1978
  - LAN standards define MAC and physical layer connectivity
    - IEEE 802.3 (CSMA/CD - Ethernet) standard – originally 2Mbps
    - IEEE 802.3u standard for 100Mbps Ethernet
    - IEEE 802.3z standard for 1,000Mbps Ethernet
- CSMA/CD: Ethernet's Media Access Control (MAC) policy
  - CS = Carrier Sense
    - Send only if medium is idle
  - MA = Multiple Access
  - CD = collision detection
    - Stop sending immediately if collision is detected

Ethernet Standard

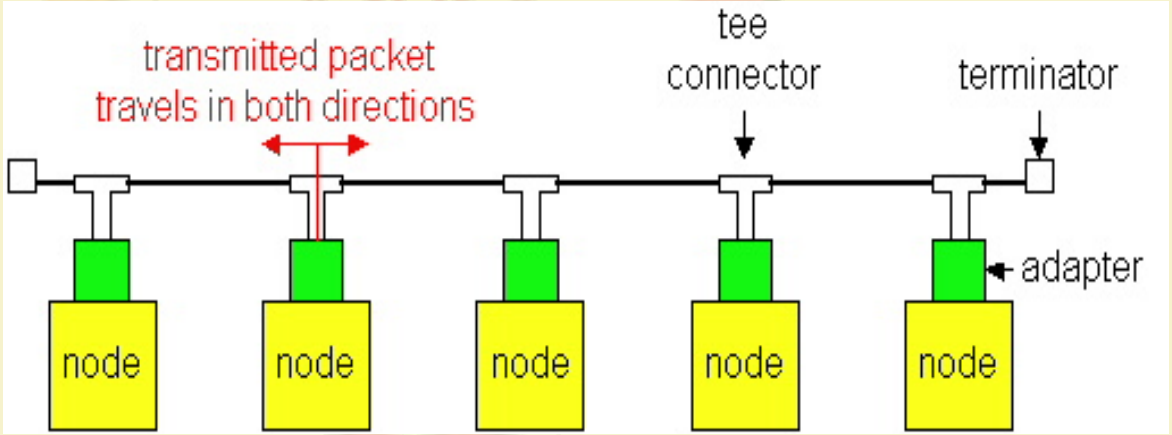
- 802.3 standard defines both MAC *and* physical layer details



Metcalfe's original Ethernet sketch

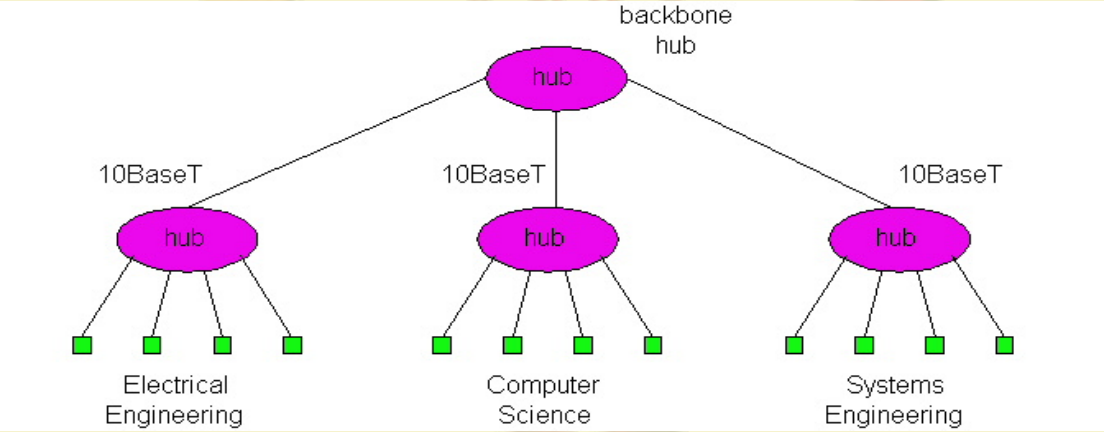
Ethernet Technologies: 10Base2

- ***10:*** 10Mbps; ***2:*** under 185 (~200) meters cable length
- Thin coaxial cable in a bus topology
- Repeaters used to connect multiple segments
  - Repeater repeats bits it hears on one interface to its other interfaces: physical layer device only!



10BaseT and 100BaseT

- 10/100 Mbps rate
- ***T*** stands for Twisted Pair
- Hub(s) connected by twisted pair facilitate "star topology"
  - Distance of any node to hub must be < 100M



Physical Layer configurations for 802.3

- Physical layer configurations are specified in three parts
- Data rate (10, 100, 1000)
  - 10, 100, 1000Mbps
- Signaling method (base, broad)
  - Baseband
    - Digital signaling
  - Broadband
    - Analog signaling
- Cabling (2, 5, T, F, S, L)
  - 5 - Thick coax (original Ethernet cabling)
  - F – Optical fiber
  - S – Short wave laser over multimode fiber
  - L – Long wave laser over single mode fiber

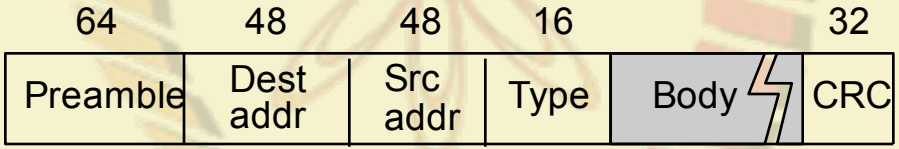
Ethernet Overview

- Most popular packet-switched LAN technology
- Bandwidths: 10Mbps, 100Mbps, 1Gbps
- Max bus length: 2500m
  - 500m segments with 4 repeaters
- Bus and Star topologies are used to connect hosts
  - Hosts attach to network via Ethernet transceiver or hub or switch
    - Detects line state and sends/receives signals
  - Hubs are used to facilitate shared connections
  - All hosts on an Ethernet are competing for access to the medium
    - Switches break this model
- Challenge: Distributed algorithm that provides fair access

Ethernet Overview (contd.)

- Ethernet by definition is a broadcast protocol
  - Any signal can be received by all hosts
  - Switching enables individual hosts to communicate
- Network layer packets are transmitted over an Ethernet by encapsulating

- Frame Format



Switched Ethernet

- Switches forward and filter frames based on LAN addresses
  - It's not a bus or a router (although simple forwarding tables are maintained)
- Very scalable
  - Options for many interfaces
  - Full duplex operation (send/receive frames simultaneously)
- Connect two or more "segments" by copying data frames between them
  - Switches only copy data when needed
    - key difference from repeaters
- Higher link bandwidth
  - Collisions are completely avoided
- Much greater aggregate bandwidth
  - Separate segments can send at once

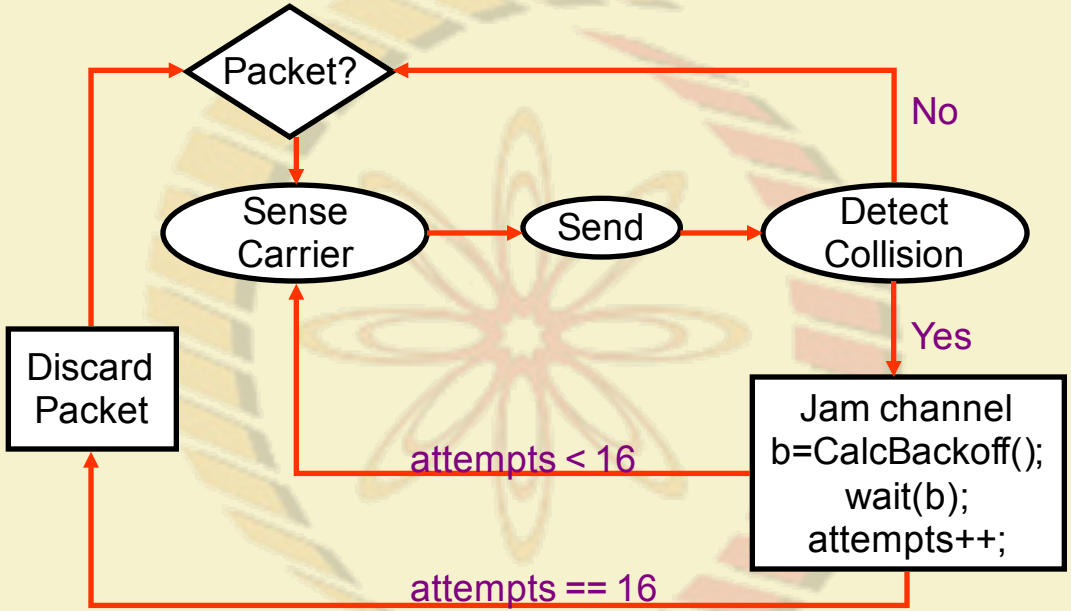
Ethernet Frames

- Preamble is a sequence of 7 bytes, each set to “10101010”
  - Used to synchronize receiver before actual data is sent
- Addresses
  - unique, 48-bit unicast address assigned to each adapter
    - example: 3**8:10:e4:b1:29:07**
    - Each manufacturer gets their own address range
  - broadcast: all **1**s
  - multicast: first bit is **1**
- *Type* field is a demultiplexing key used to determine which higher level protocol the frame should be delivered to
- Body can contain up to 1500 bytes of data

Ethernet's MAC Algorithm

- In Aloha, decisions to transmit are made without paying attention to what other nodes might be doing
- Ethernet uses CSMA/CD – listens to line before/during sending
- If line is idle (no carrier sensed)
  - send packet immediately
  - upper bound message size of 1500 bytes
  - must wait 9.6us between back-to-back frames
- If line is busy (carrier sensed)
  - wait until idle and transmit packet immediately
    - called *1-persistent* sending
- If collision detected
  - Stop sending and jam signal
  - Try again later

State Diagram for CSMA/CD




thank you!

COMPUTER NETWORKS AND INTERNET PROTOCOLS

Data Link Layer – Ethernet (contd.)

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

Ethernet

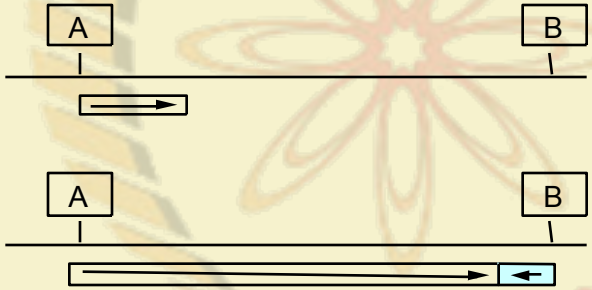
- Background
  - Developed by Bob Metcalfe and others at Xerox PARC in mid-1970s
  - Roots in Aloha packet-radio network
  - Standardized by Xerox, DEC, and Intel in 1978
  - LAN standards define MAC and physical layer connectivity
    - IEEE 802.3 (CSMA/CD - Ethernet) standard – originally 2Mbps
    - IEEE 802.3u standard for 100Mbps Ethernet
    - IEEE 802.3z standard for 1,000Mbps Ethernet
- CSMA/CD: Ethernet's Media Access Control (MAC) policy
  - CS = Carrier Sense
    - Send only if medium is idle
  - MA = Multiple Access
  - CD = collision detection
    - Stop sending immediately if collision is detected

Collisions

Collisions are caused when two adaptors transmit at the same time (adaptors sense collision based on voltage differences)

- Both found line to be idle
- Both had been waiting to for a busy line to become idle

A starts at time 0



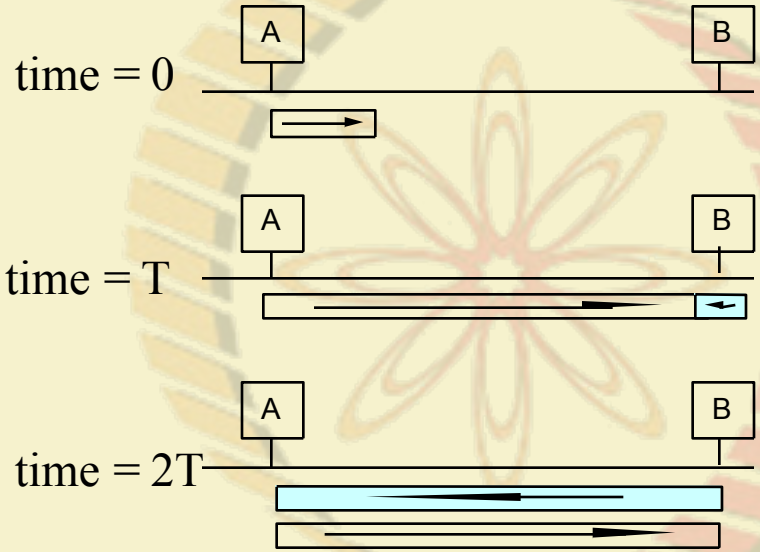
Message almost there at time T when B starts – collision!

How can we be sure A knows about the collision?

Collision Detection

- How can A know that a collision has taken place?
  - There must be a mechanism to insure retransmission on collision
  - A's message reaches B at time T
  - B's message reaches A at time 2T
  - So, A must still be transmitting at 2T
- IEEE 802.3 specifies max value of 2T to be 51.2us
  - This relates to maximum distance of 2500m between hosts
  - At 10Mbps it takes 0.1us to transmit one bit so 512 bits (64B) take 51.2us to send
  - So, Ethernet frames must be at least 64B long
    - 14B header, 46B data, 4B CRC
    - Padding is used if data is less than 46B
- Send jamming signal after collision is detected to insure all hosts see collision
  - 48 bit signal

Collision Detection (contd…)



Exponential Backoff

- If a collision is detected, delay and try again
- Delay time is selected using binary exponential backoff
  - 1st time: choose K from {0,1} then delay = K * 51.2us
  - 2nd time: choose K from {0,1,2,3} then delay = K * 51.2us
  - *nth* time: delay = K x 51.2us, for K=0..2^n – 1
    - Note max value for k = 1023
  - give up after several tries (usually 16)
    - Report transmit error to host
- If delay is not random, then there is a chance that sources would retransmit in lock step
- Why not just choose from small set for K
  - This works fine for a small number of hosts
  - Large number of nodes would result in more collisions

MAC Algorithm from the Receiver Side

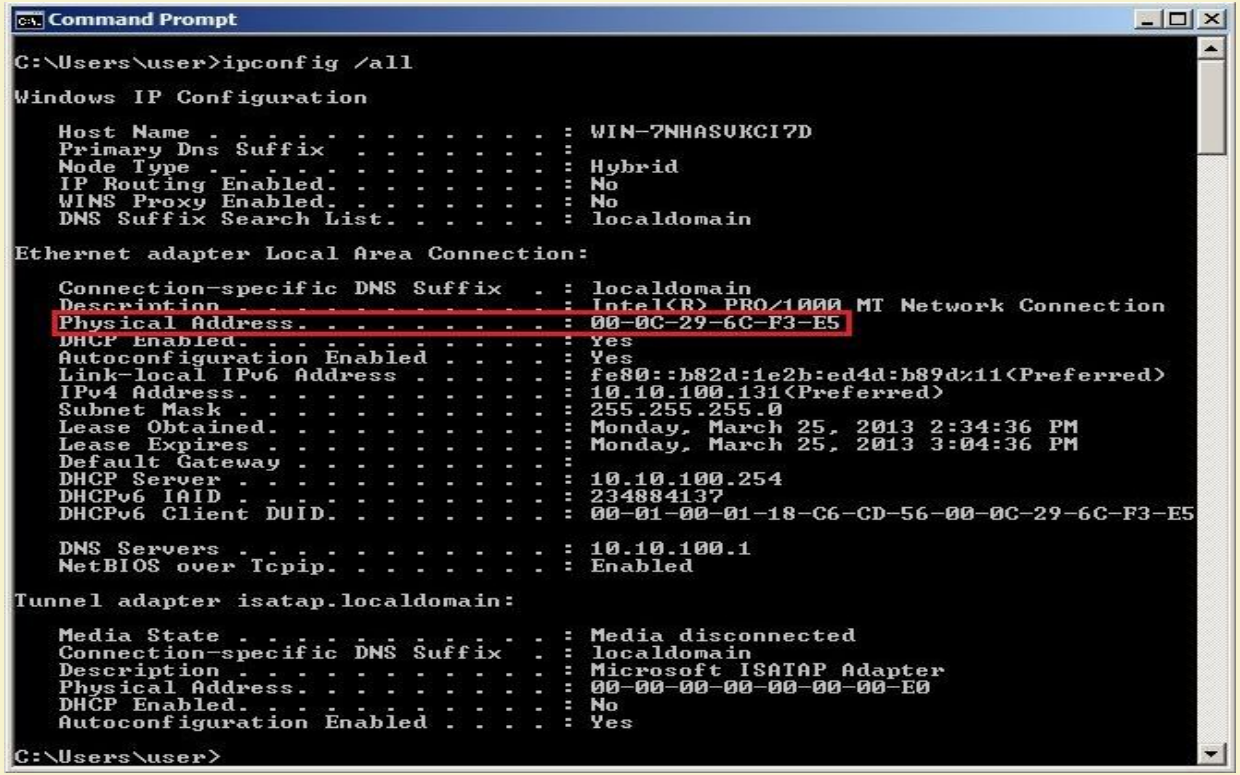
- Senders handle all access control
- Receivers simply read frames with acceptable address
  - Address to host
  - Address to broadcast
  - Address to multicast to which host belongs
  - All frames if host is in promiscuous mode

Ethernet Frame (as per 802.3 standard)

Preamble	SFD	Destination MAC	Source MAC	Type	Data and Pad	FCS
7 Bytes	1 Byte	6 Bytes	6 Bytes	2 Bytes	46-1500 Bytes	4 Bytes

- **Preamble** – informs the receiving system that a frame is starting and enables synchronization.
- **SFD (Start Frame Delimiter)** – signifies that the Destination MAC Address field begins with the next byte.
- **Destination MAC** – identifies the receiving system.
- **Source MAC** – identifies the sending system.
- **Type** – defines the type of protocol inside the frame, for example IPv4 or IPv6.
- **Data** and **Pad** – contains the payload data. Padding data is added to meet the minimum length requirement for this field (46 bytes).
- **FCS (Frame Check Sequence)** – contains a 32-bit **Cyclic Redundancy Check (CRC)** which allows detection of corrupted data.

Ethernet Address?



Fast and Gigabit Ethernet

- Fast Ethernet (100Mbps) has technology very similar to 10Mbps Ethernet
  - Uses different physical layer encoding (4B5B)
  - Many NIC's are 10/100 capable
    - Can be used at either speed
- Gigabit Ethernet (1,000Mbps)
  - Compatible with lower speeds
  - Uses standard framing and CSMA/CD algorithm
  - Distances are severely limited
  - Typically used for backbones and inter-router connectivity
  - Becoming cost competitive
  - How much of this bandwidth is realizable?

Ethernet – Practical facts

- Ethernets work best under light loads
  - Utilization over 30% is considered heavy
    - Network capacity is wasted by collisions
- Most networks are limited to about 200 hosts
  - Specification allows for up to 1024
- Most networks are much shorter
  - 5 to 10 microsecond RTT
- Transport level flow control helps reduce load (number of back to back packets)
- Ethernet is inexpensive, fast and easy to administer!

Ethernet Issues

- Ethernet's peak utilization is pretty low
- Peak throughput worst with
  - More hosts
    - More collisions needed to identify single sender
  - Smaller packet sizes
    - More frequent arbitration
  - Longer links
    - Collisions take longer to observe, more wasted bandwidth
  - *Efficiency cane be improved by avoiding these conditions*


thank you!

COMPUTER NETWORKS AND INTERNET PROTOCOLS

Data Link Layer – Ethernet (contd.)

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

Ethernet

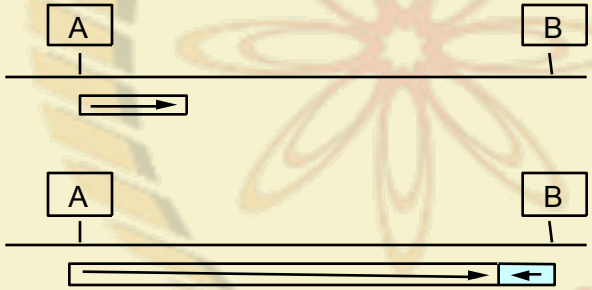
- Background
  - Developed by Bob Metcalfe and others at Xerox PARC in mid-1970s
  - Roots in Aloha packet-radio network
  - Standardized by Xerox, DEC, and Intel in 1978
  - LAN standards define MAC and physical layer connectivity
    - IEEE 802.3 (CSMA/CD - Ethernet) standard – originally 2Mbps
    - IEEE 802.3u standard for 100Mbps Ethernet
    - IEEE 802.3z standard for 1,000Mbps Ethernet
- CSMA/CD: Ethernet's Media Access Control (MAC) policy
  - CS = Carrier Sense
    - Send only if medium is idle
  - MA = Multiple Access
  - CD = collision detection
    - Stop sending immediately if collision is detected

Collisions

Collisions are caused when two adaptors transmit at the same time (adaptors sense collision based on voltage differences)

- Both found line to be idle
- Both had been waiting to for a busy line to become idle

A starts at time 0



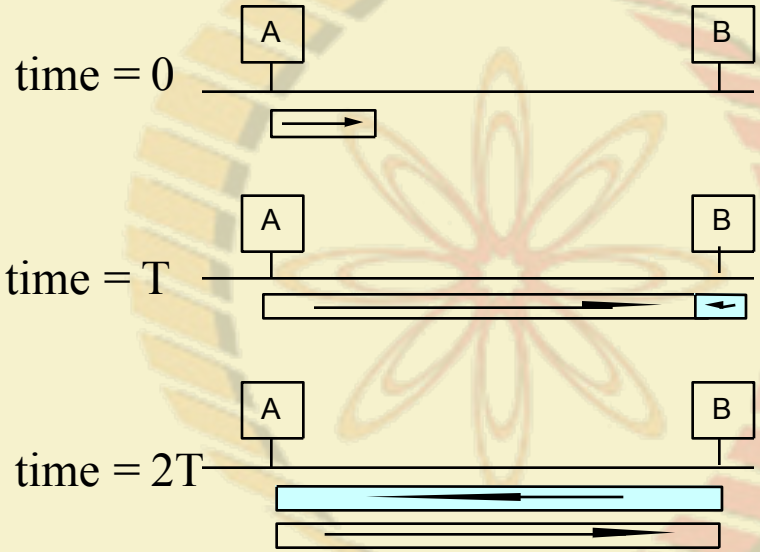
Message almost there at time T when B starts – collision!

How can we be sure A knows about the collision?

Collision Detection

- How can A know that a collision has taken place?
  - There must be a mechanism to insure retransmission on collision
  - A's message reaches B at time T
  - B's message reaches A at time 2T
  - So, A must still be transmitting at 2T
- IEEE 802.3 specifies max value of 2T to be 51.2us
  - This relates to maximum distance of 2500m between hosts
  - At 10Mbps it takes 0.1us to transmit one bit so 512 bits (64B) take 51.2us to send
  - So, Ethernet frames must be at least 64B long
    - 14B header, 46B data, 4B CRC
    - Padding is used if data is less than 46B
- Send jamming signal after collision is detected to insure all hosts see collision
  - 48 bit signal

Collision Detection (contd…)



Exponential Backoff

- If a collision is detected, delay and try again
- Delay time is selected using binary exponential backoff
  - 1st time: choose K from {0,1} then delay = K * 51.2us
  - 2nd time: choose K from {0,1,2,3} then delay = K * 51.2us
  - *nth* time: delay = K x 51.2us, for K=0..2^n – 1
    - Note max value for k = 1023
  - give up after several tries (usually 16)
    - Report transmit error to host
- If delay is not random, then there is a chance that sources would retransmit in lock step
- Why not just choose from small set for K
  - This works fine for a small number of hosts
  - Large number of nodes would result in more collisions

MAC Algorithm from the Receiver Side

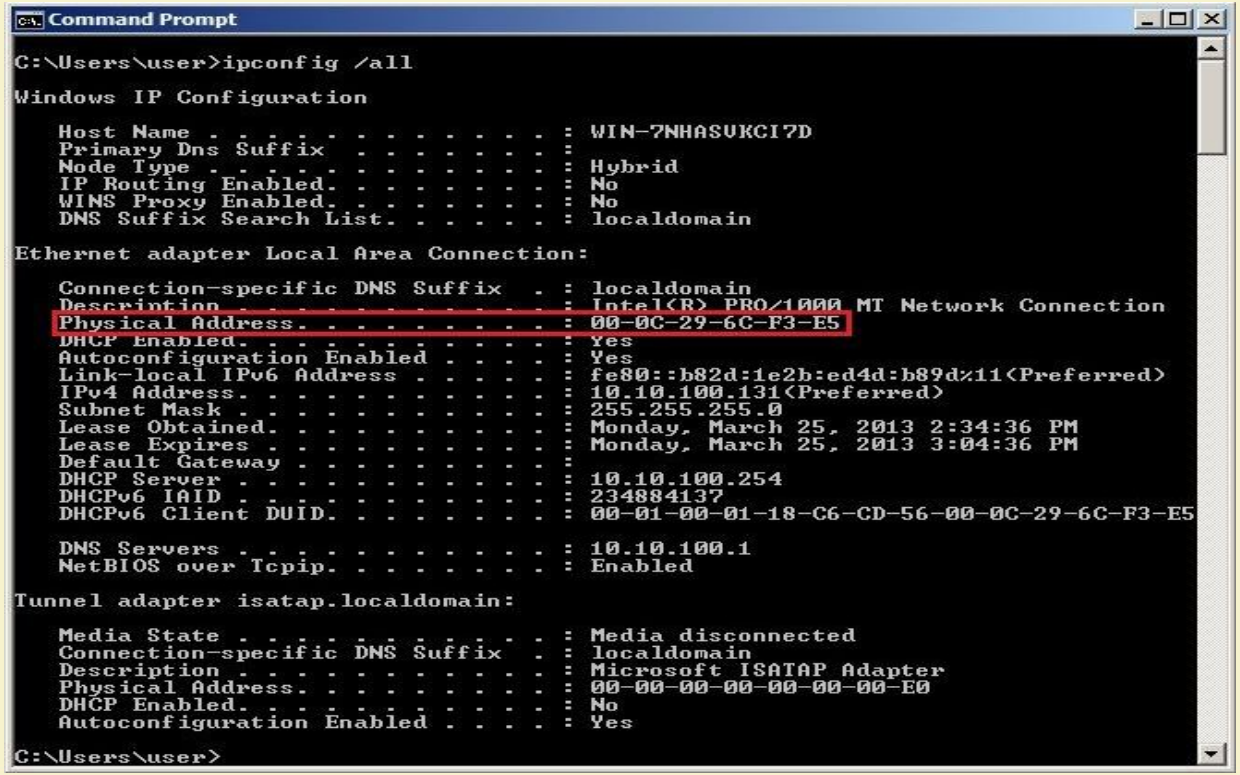
- Senders handle all access control
- Receivers simply read frames with acceptable address
  - Address to host
  - Address to broadcast
  - Address to multicast to which host belongs
  - All frames if host is in promiscuous mode

Ethernet Frame (as per 802.3 standard)

Preamble	SFD	Destination MAC	Source MAC	Type	Data and Pad	FCS
7 Bytes	1 Byte	6 Bytes	6 Bytes	2 Bytes	46-1500 Bytes	4 Bytes

- **Preamble** – informs the receiving system that a frame is starting and enables synchronization.
- **SFD (Start Frame Delimiter)** – signifies that the Destination MAC Address field begins with the next byte.
- **Destination MAC** – identifies the receiving system.
- **Source MAC** – identifies the sending system.
- **Type** – defines the type of protocol inside the frame, for example IPv4 or IPv6.
- **Data** and **Pad** – contains the payload data. Padding data is added to meet the minimum length requirement for this field (46 bytes).
- **FCS (Frame Check Sequence)** – contains a 32-bit **Cyclic Redundancy Check (CRC)** which allows detection of corrupted data.

Ethernet Address?



Fast and Gigabit Ethernet

- Fast Ethernet (100Mbps) has technology very similar to 10Mbps Ethernet
  - Uses different physical layer encoding (4B5B)
  - Many NIC's are 10/100 capable
    - Can be used at either speed
- Gigabit Ethernet (1,000Mbps)
  - Compatible with lower speeds
  - Uses standard framing and CSMA/CD algorithm
  - Distances are severely limited
  - Typically used for backbones and inter-router connectivity
  - Becoming cost competitive
  - How much of this bandwidth is realizable?

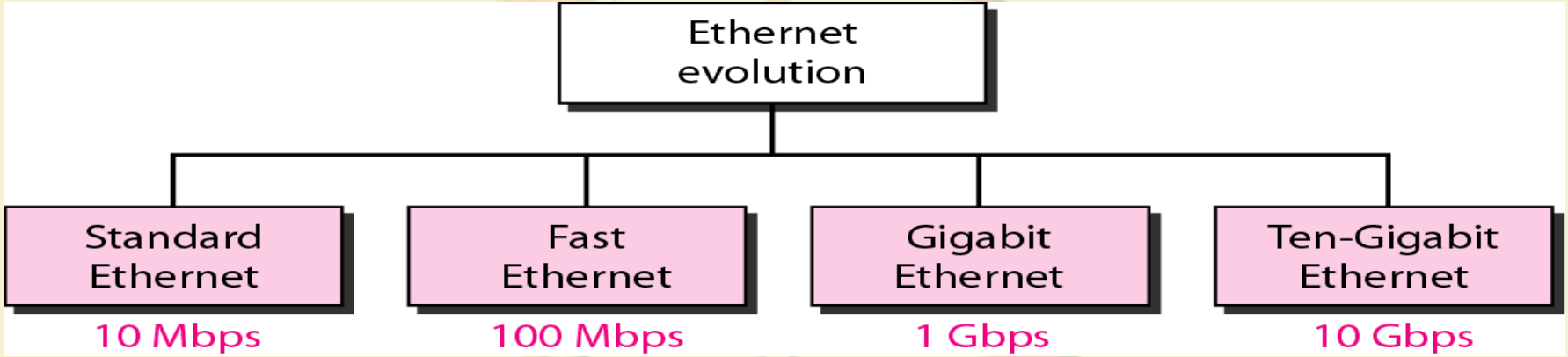
Ethernet – Practical facts

- Ethernets work best under light loads
  - Utilization over 30% is considered heavy
    - Network capacity is wasted by collisions
- Most networks are limited to about 200 hosts
  - Specification allows for up to 1024
- Most networks are much shorter
  - 5 to 10 microsecond RTT
- Transport level flow control helps reduce load (number of back to back packets)
- Ethernet is inexpensive, fast and easy to administer!

Ethernet Issues

- Ethernet's peak utilization is pretty low
- Peak throughput worst with
  - More hosts
    - More collisions needed to identify single sender
  - Smaller packet sizes
    - More frequent arbitration
  - Longer links
    - Collisions take longer to observe, more wasted bandwidth
  - *Efficiency cane be improved by avoiding these conditions*

Ethernet Evolution



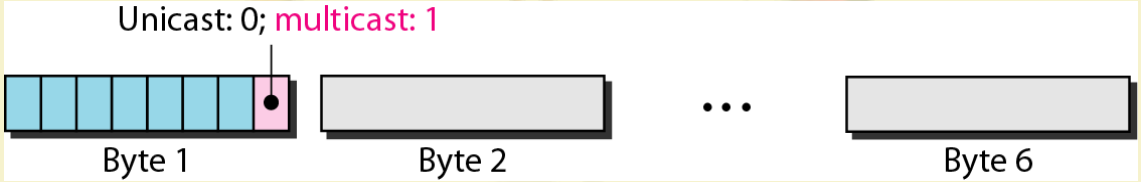
Ethernet Addressing

- Each station on an Ethernet network (such as a PC, workstation, or printer) has its own network interface card (NIC).
- NIC fits inside the station and provides the station with a 6-byte physical address.

17 : 6E : 10 : 06 : 4C : 2A

Unicast, Multicast, and Broadcast Addresses

- A source address is always a unicast address--the frame comes from only one station.
- The destination address, however, can be unicast, multicast, or broadcast. If the least significant bit of the first byte in a destination address is 0, the address is unicast; otherwise, it is multicast.
- The broadcast destination address is a special case of the multicast address in which all bits are 1s.



Ethernet Addressing

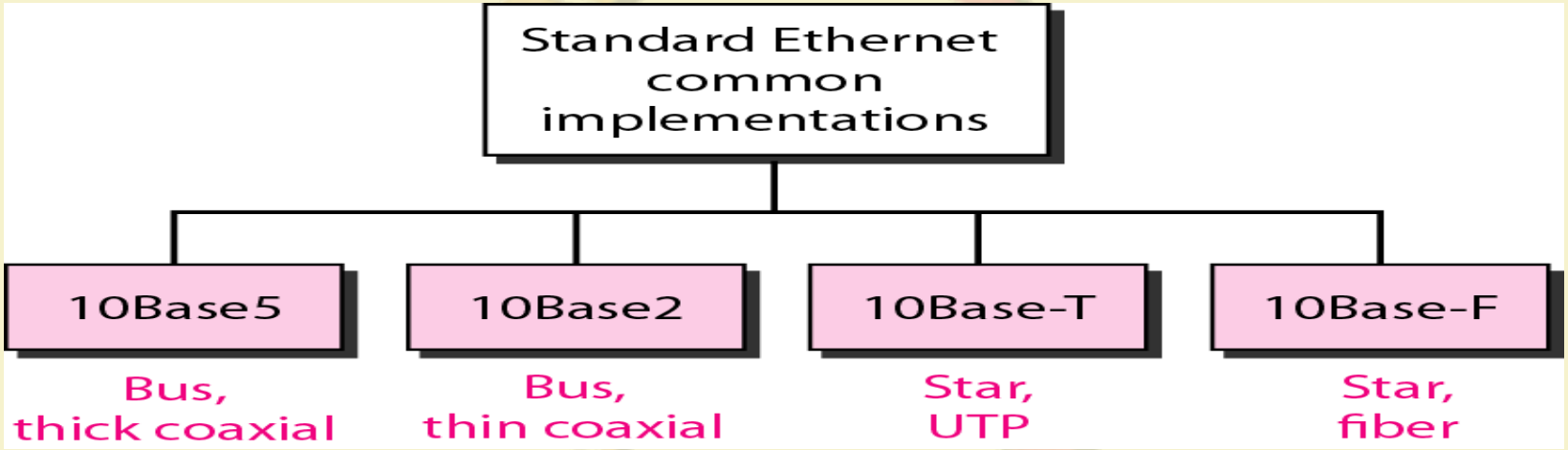
Type of the following destination addresses?

- a. 2A:30:1A:22:11:1A b. 17:2A:1B:2E:08:AE c. FF:FF:FF:FF:FF:FF

a. Unicast b. Multicast c. Broadcast

Representation of the address 47:20:1B:2E:08:EE

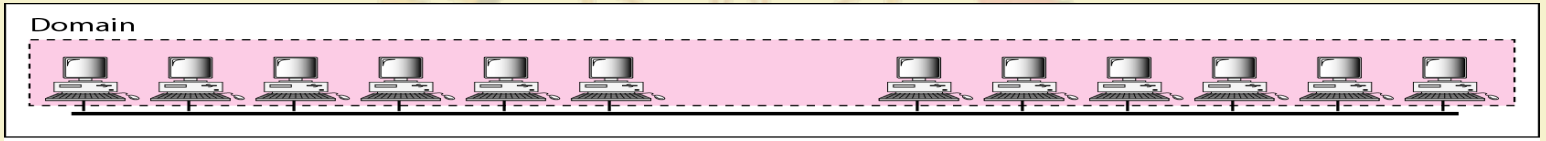
Standard Ethernet



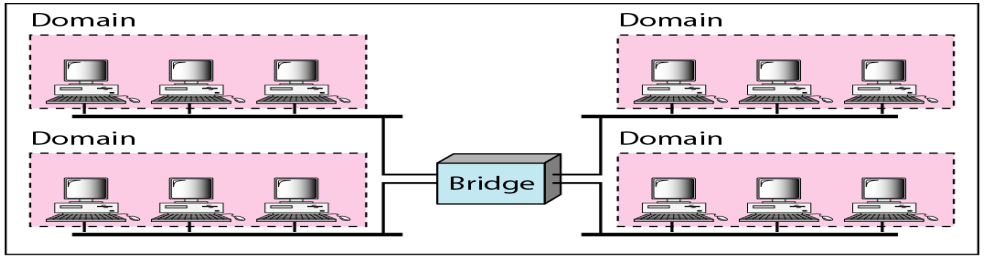
Ethernet Evolution - Bridged Ethernet

Bridged Ethernet

- Dividing LANs using Bridges
- Increased bandwidth and separate collision domains



a. Without bridging

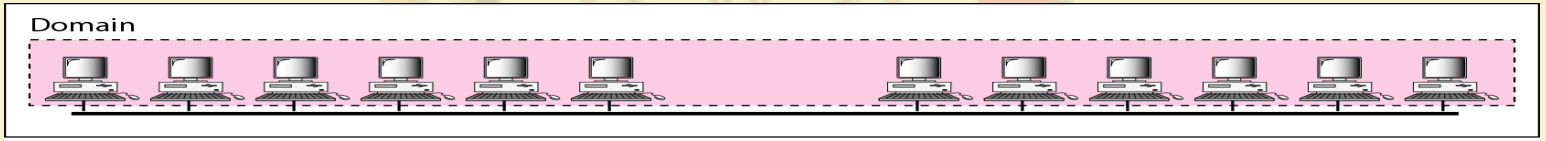


b. With bridging

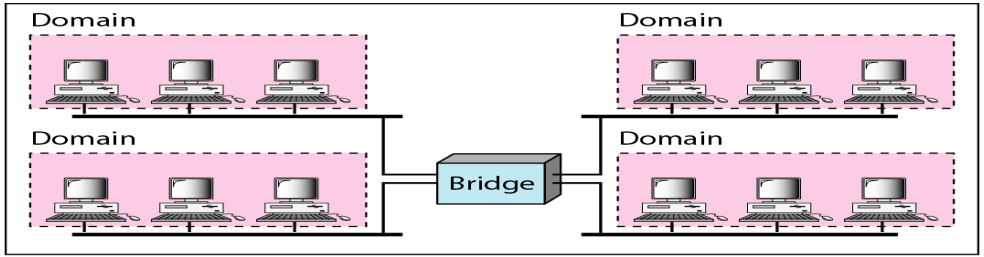
Ethernet Evolution - Bridged Ethernet

Bridged Ethernet

- Dividing LANs using Bridges
- Increased bandwidth and separate collision domains



a. Without bridging

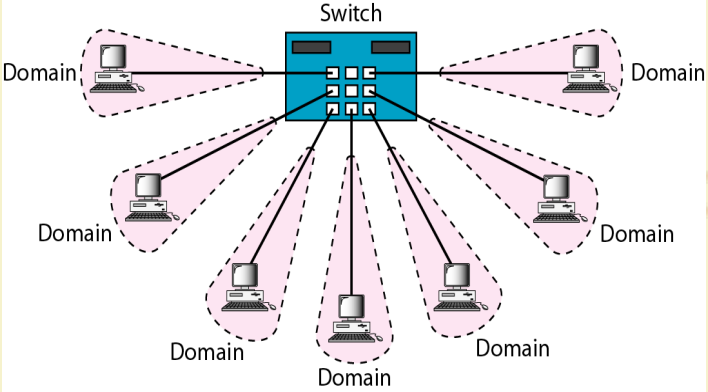


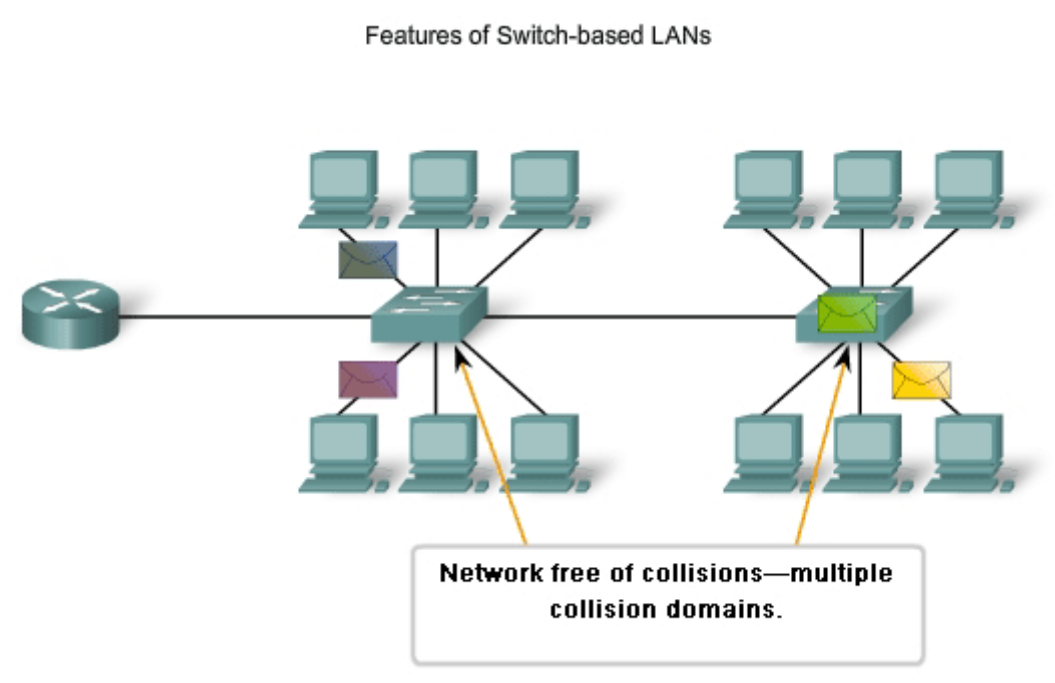
b. With bridging

Ethernet Evolution - Switched Ethernet

Switched Ethernet

- Extension of Bridged Ethernet

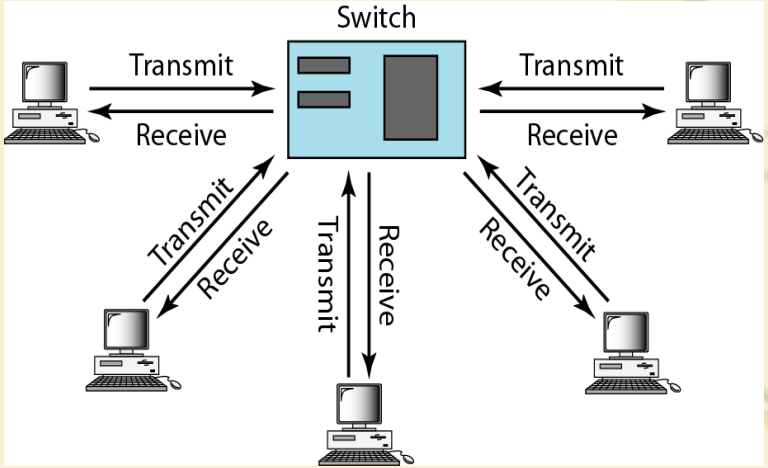


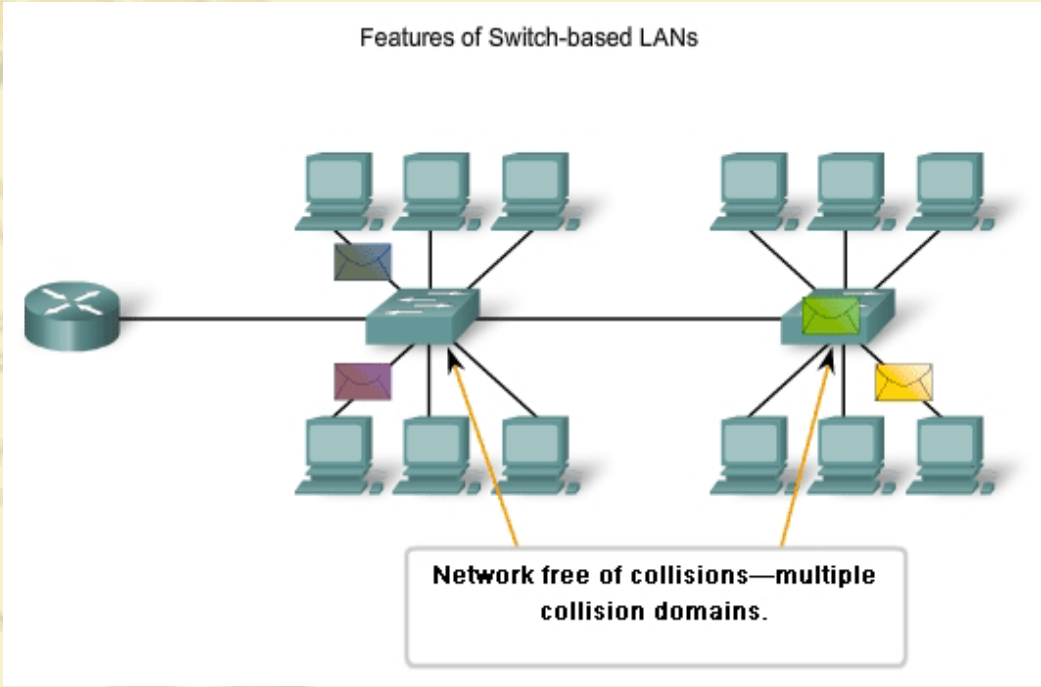


Ethernet Evolution - Full-Duplex Ethernet

Full-Duplex Ethernet:

Full-duplex mode increases the capacity of each domain (10Mbps to 20Mbps)





Ethernet Evolution - Full-Duplex Ethernet

Full-Duplex Ethernet:

No Need for CSMA/CD

- In full-duplex switched Ethernet, there is no need for the CSMA/CD method. In a full-duplex switched Ethernet, each station is connected to the switch via two separate links.
- Each station or switch can send and receive independently without worrying about collision. Each link is a point-to-point dedicated path between the station and the switch. MAC layer becomes much simpler.

MAC Control Layer

- Standard Ethernet was designed as a connectionless protocol at the MAC sublayer.
- To provide for flow and error control in full-duplex switched Ethernet, a new sublayer, called the MAC control, is added between the LLC sublayer and the MAC sublayer.

Ethernet Evolution – Fast Ethernet

- Fast Ethernet was designed to compete with LAN protocols such as FDDI or Fiber Channel. IEEE created Fast Ethernet under the name 802.3u.
- Fast Ethernet is backward-compatible with Standard Ethernet, but it can transmit data 10 times faster at a rate of 100 Mbps.

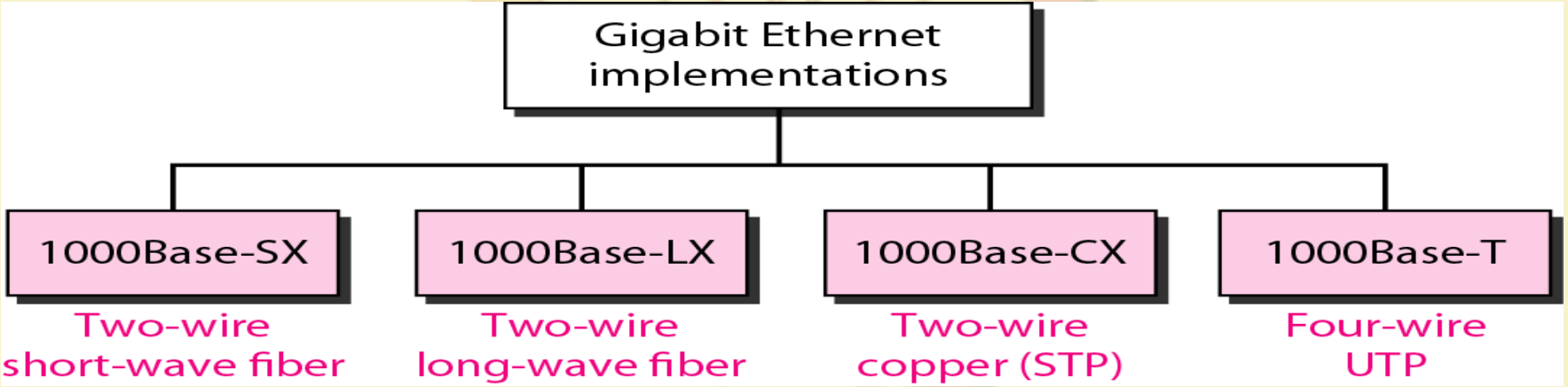
Goals:

1. Upgrade the data rate to 100 Mbps. 2. Make it compatible with Standard Ethernet.
3. Keep the same 48-bit address. 4. Keep the same frame format.
5. Keep the same minimum and maximum frame lengths.

Autonegotiation: A new feature added to Fast Ethernet is called autonegotiation. Autonegotiation allows two devices to negotiate the mode or data rate of operation. Objectives : (i) To allow incompatible devices to connect to one another. (ii) To allow one device to have multiple capabilities. (iii) To allow a station to check a hub's capabilities.

Ethernet Evolution – Gigabit Ethernet

- Higher data rate of 1000 Mbps (IEEE 802.3z standard)
- In the full-duplex mode of Gigabit Ethernet, there is no collision; the maximum length of the cable is determined by the signal attenuation in the cable.




thank you!

COMPUTER NETWORKS AND INTERNET PROTOCOLS

Data Link Layer – Flow and Error Control

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

Flow Control

- Flow control coordinates the amount of data that can be sent before receiving acknowledgement
- Flow control is a set of procedures that tells the sender how much data it can transmit before it must wait for an acknowledgement from the receiver.
- Receiver has a limited speed at which it can process incoming data and a limited amount of memory in which to store incoming data.
- Receiver must inform the sender before the limits are reached and request that the transmitter to send fewer frames or stop temporarily.
- Since the rate of processing is often slower than the rate of transmission, receiver has a block of memory (buffer) for storing incoming data until they are processed.

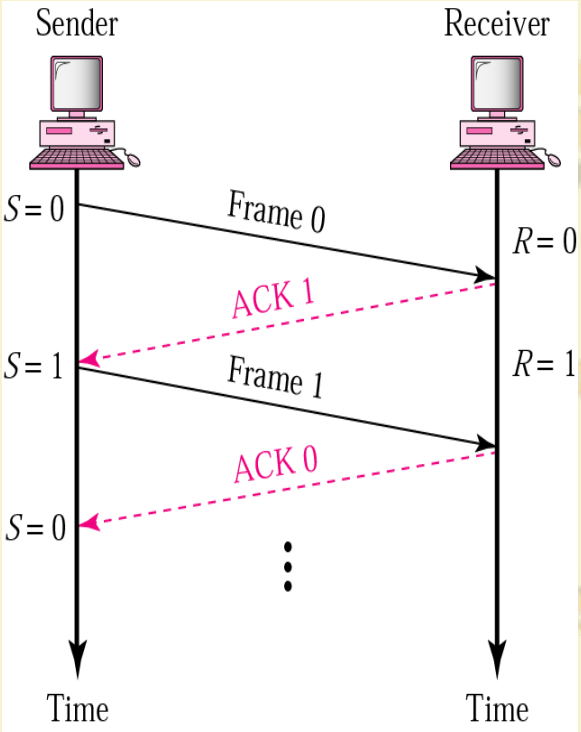
Error Control

- Error control includes both error detection and error correction.
- It allows the receiver to inform the sender if a frame is lost or damaged during transmission and coordinates the retransmission of those frames by the sender.
- Error control in the data link layer is based on automatic repeat request (ARQ). Whenever an error is detected, specified frames are retransmitted.

Error and Flow Control Mechanisms

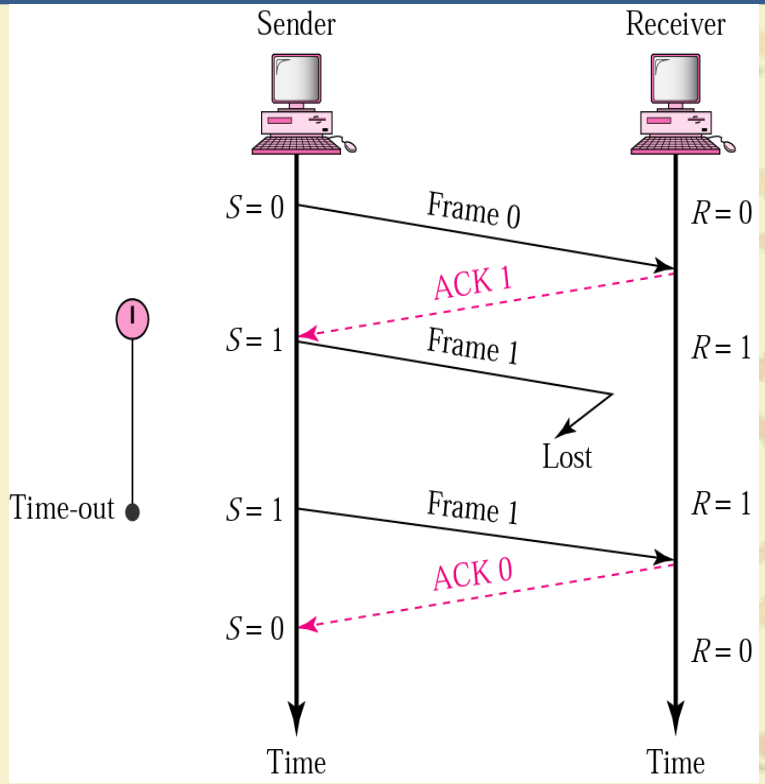
- Stop-and-Wait
- Go-Back-N ARQ
- Selective-Repeat ARQ

Stop-and-Wait



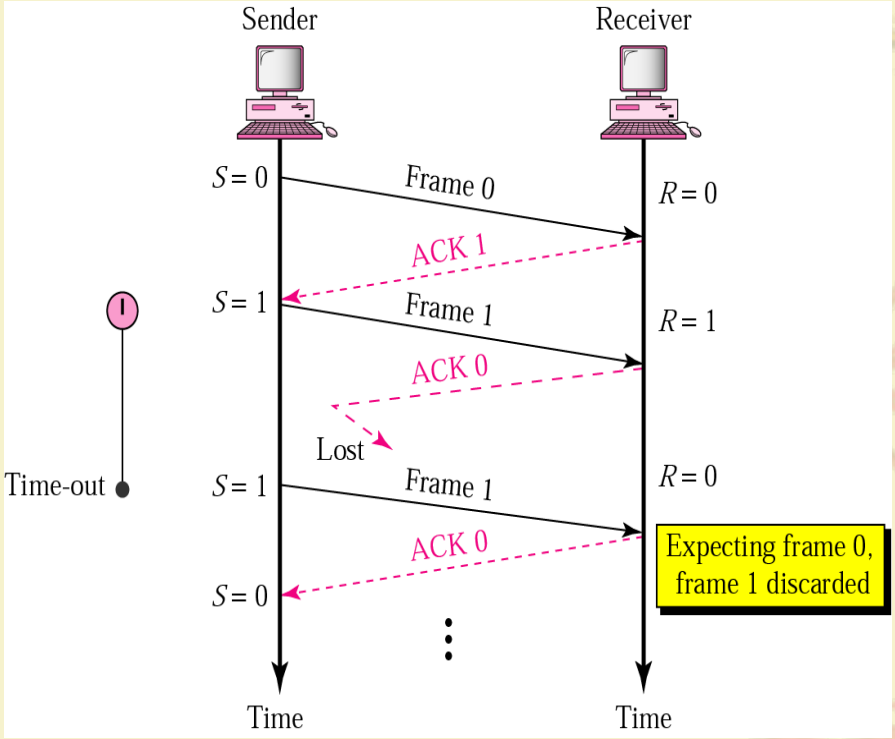
- Sender keeps a copy of the last frame until it receives an acknowledgement.
- For identification, both data frames and acknowledgements (ACK) frames are numbered alternatively 0 and 1.
- Sender has a control variable (S) that holds the number of the recently sent frame. (0 or 1)
- Receiver has a control variable (R) that holds the number of the next frame expected (0 or 1).
- Sender starts a timer when it sends a frame. If an ACK is not received within a allocated time period, the sender assumes that the frame was lost or damaged and resends it
- Receiver send only positive ACK if the frame is intact.
- ACK number always defines the number of the next expected frame

Stop-and-Wait ARQ, lost ACK frame



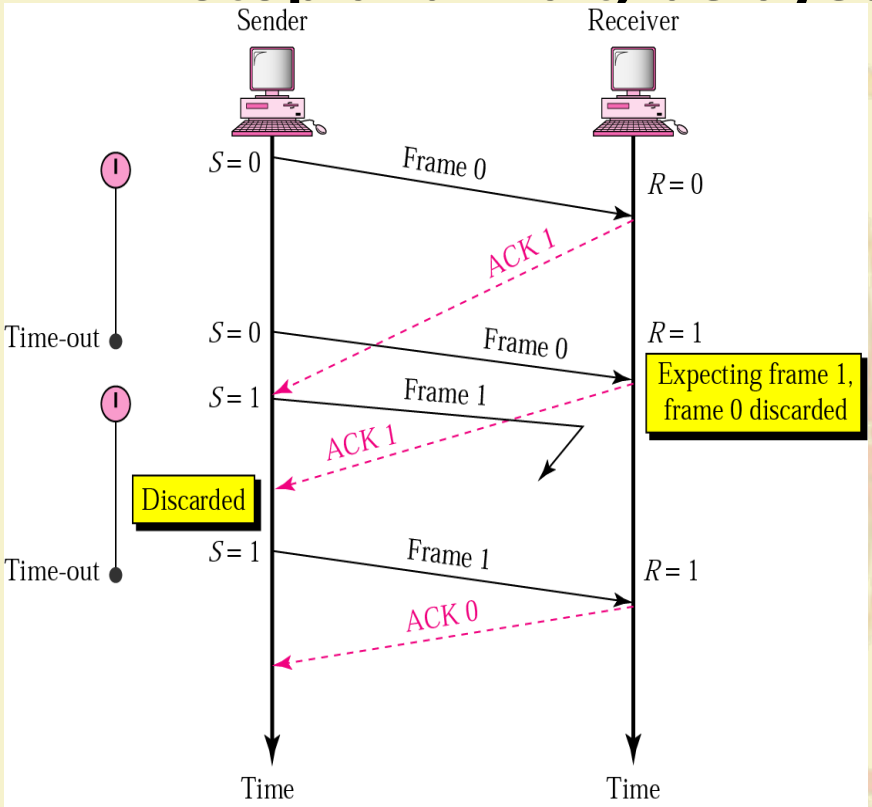
- When a receiver receives a damaged frame, it discards it and keeps its value of R.
- After the timer at the sender expires, another copy of frame 1 is sent.

Stop-and-Wait, lost ACK frame



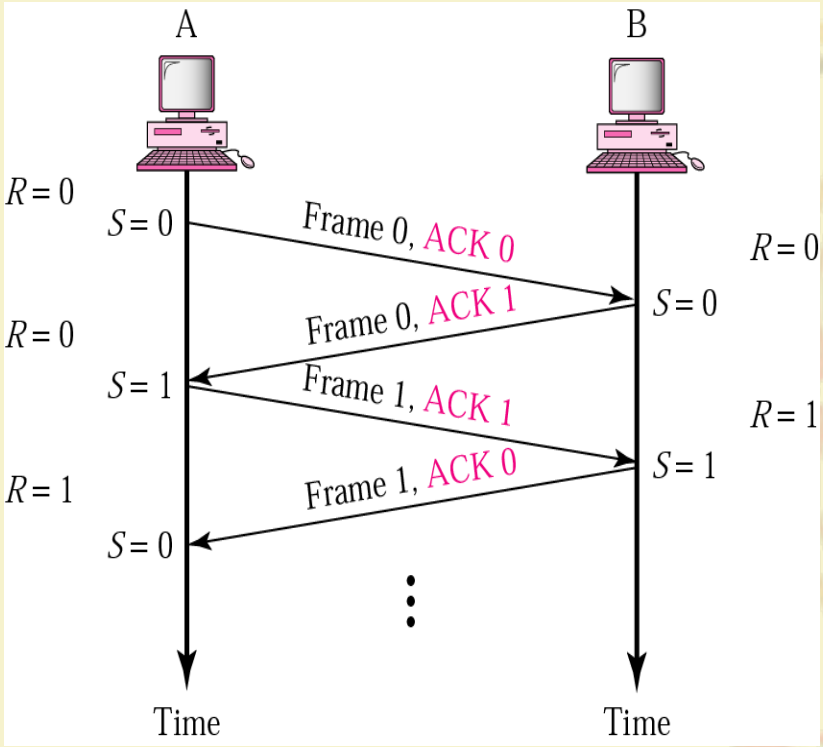
- If the sender receives a damaged ACK, it discards it.
- When the timer of the sender expires, the sender retransmits frame 1.
- Receiver has already received frame 1 and expecting to receive frame 0 (R=0). Therefore it discards the second copy of frame 1.

Stop-and-Wait, delayed ACK frame



- ACK can be delayed at the receiver or due to some problem
- It is received after the timer for frame 0 has expired.
- Sender retransmitted a copy of frame 0. However, R =1 means receiver expects to see frame 1. Receiver discards the duplicate frame 0.
- Sender receives 2 ACKs, it discards the second ACK.

Piggybacking



- A method to combine a data frame with ACK.
- Station A and B both have data to send.
- Instead of sending separately, station A sends a data frame that includes an ACK.
- Station B does the same thing.
- Piggybacking saves bandwidth.

Disadvantage of Stop-and-Wait

- In stop-and-wait, at any point in time, there is only one frame that is sent and waiting to be acknowledged.
- This is not a good use of transmission medium.
- To improve efficiency, multiple frames should be in transition while waiting for ACK.
- Two protocol use the above concept,
  - **Go-Back-N ARQ**
  - **Selective Repeat ARQ**

Go-Back-N ARQ

- We can send up to W frames before worrying about ACKs.
- We keep a copy of these frames until the ACKs arrive.
- This procedure requires additional features to be added to Stop-and-Wait ARQ.

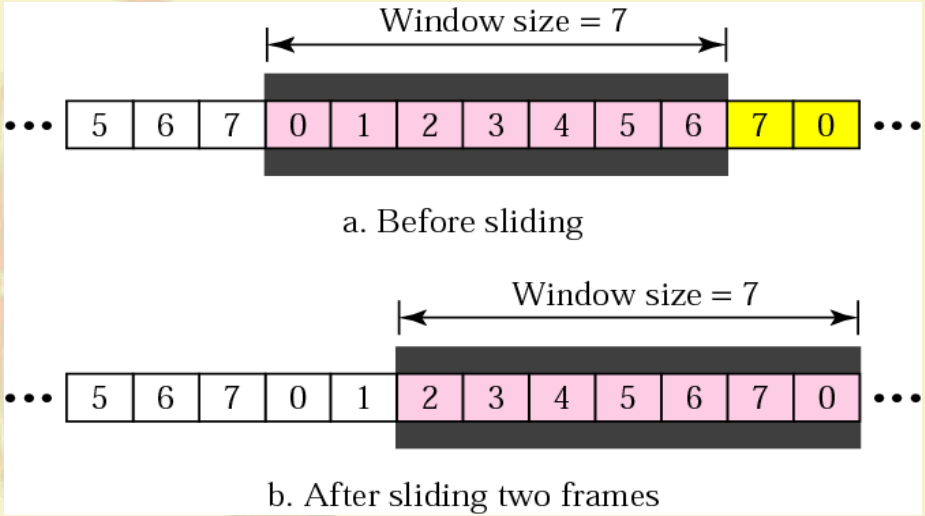
Sequence Numbers

- Frames from a sender are numbered sequentially.
- We need to set a limit since we need to include the sequence number of each frame in the header.
- If the header of the frame allows m bits for sequence number, the sequence numbers range from 0 to $2^m - 1$. for m = 3, sequence numbers are: 1, 2, 3, 4, 5, 6, 7.
- We can repeat the sequence number.
- Sequence numbers are:

  0, 1, 2, 3, 4, 5, 6, 7, 0, 1, 2, 3, 4, 5, 6, 7, 0, 1, …

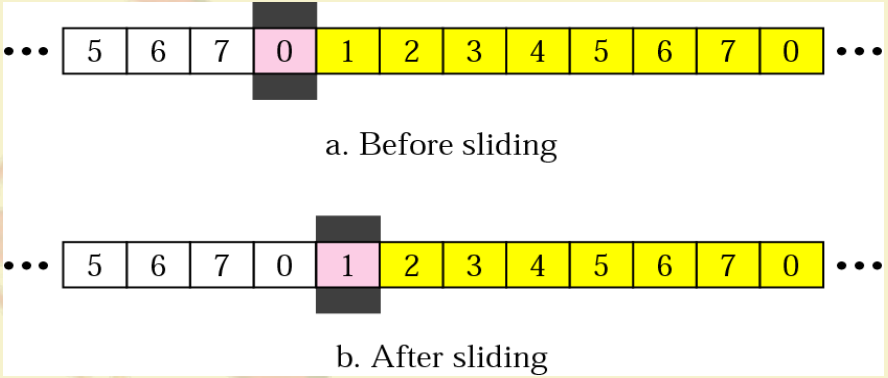
Sender Sliding Window

- At the sending site, to hold the outstanding frames until they are acknowledged, we use the concept of a window.
- Size of the window is at most $2^m - 1$ where m is the number of bits for the sequence number.
- Size of the window can be variable.
- The window slides to include new unsent frames when the correct ACKs are received



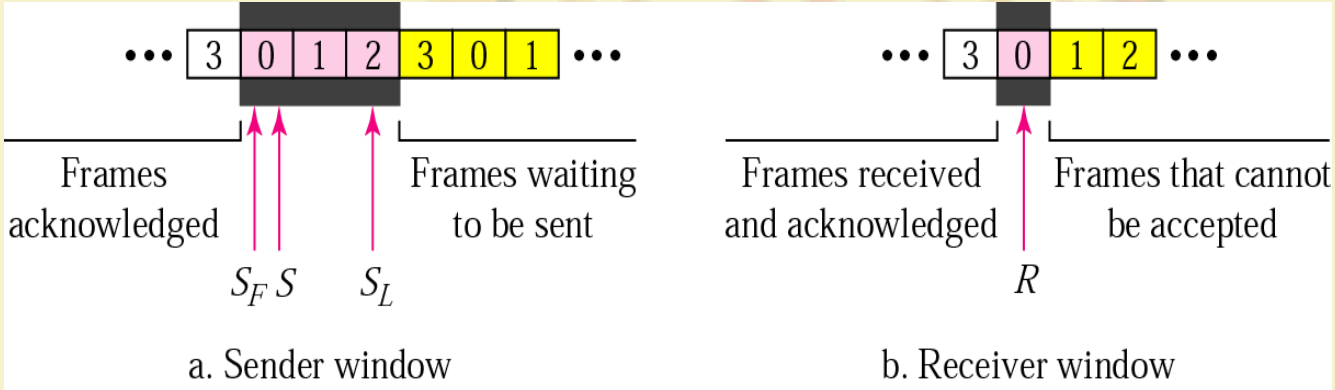
Receiver Sliding Window

- Size of the window at the receiving site is always 1 in this protocol.
- Receiver is always looking for a specific frame to arrive in a specific order.
- Any frame arriving out of order is discarded and needs to be resent.
- Receiver window slides as shown in the figure. Receiver is waiting for frame 0 in part a.



Control Variables

- Sender has 3 variables: S, S_F, and S_L
- S holds the sequence number of recently sent frame
- S_F holds the sequence number of the first frame
- S_L holds the sequence number of the last frame
- Receiver only has the one variable, R, that holds the sequence number of the frame it expects to receive. If the seq. no. is the same as the value of R, the frame is accepted, otherwise rejected.

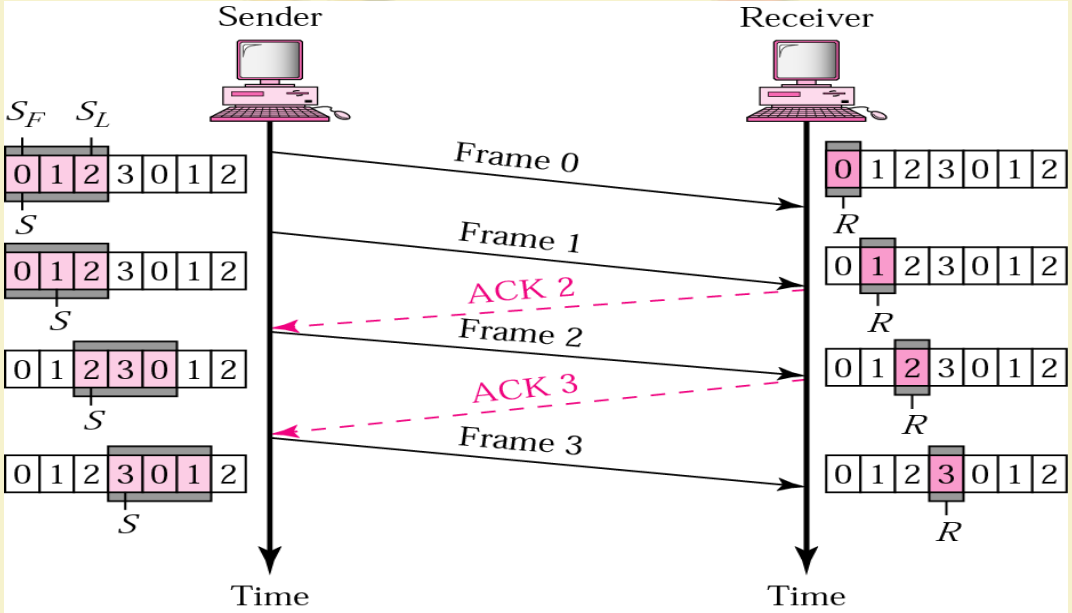


Acknowledgement (ACK)

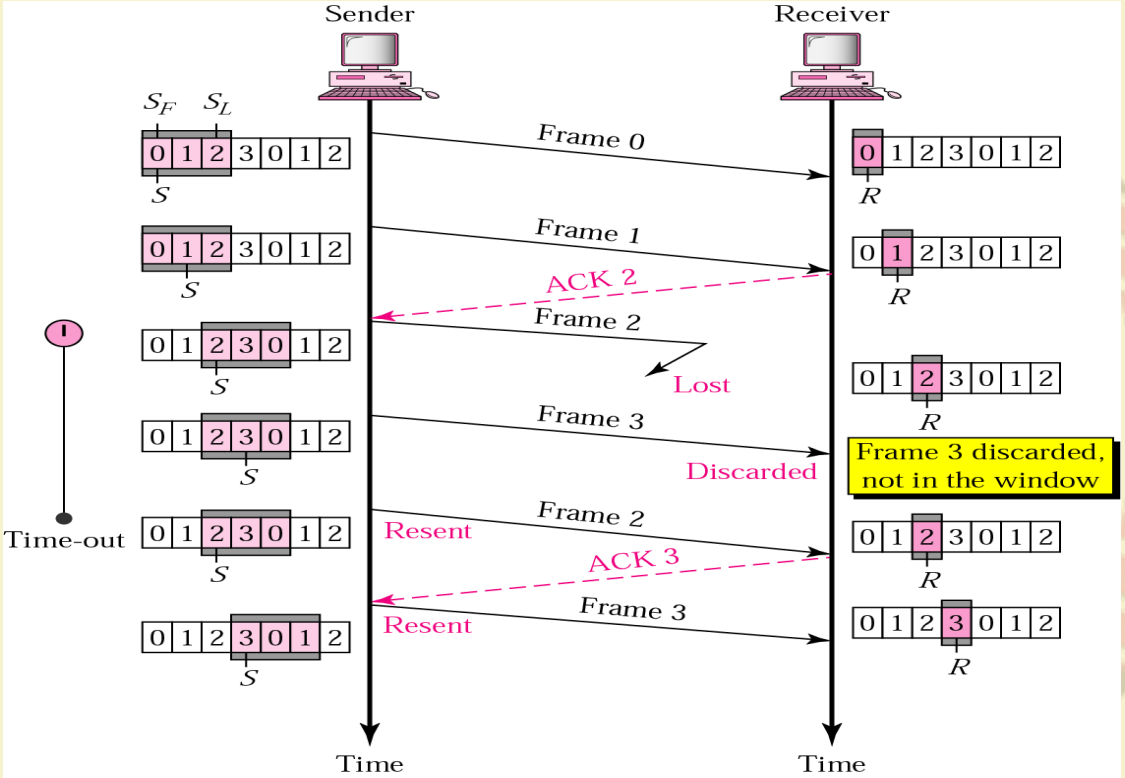
- Receiver sends positive ACK if a frame arrived safe and in order.
- If the frames are damaged/out of order, receiver is silent and discard all subsequent frames until it receives the one it is expecting.
- The silence of the receiver causes the timer of the unacknowledged frame to expire.
- Then the sender resends all frames, beginning with the one with the expired timer.
- For example, suppose the sender has sent frame 6, but the timer for frame 3 expires (i.e. frame 3 has not been acknowledged), then the sender goes back and sends frames 3, 4, 5, 6 again. Thus it is called Go-Back-N-ARQ
- The receiver does not have to acknowledge each frame received, it can send one cumulative ACK for several frames.

Go-Back-N ARQ, normal operation

- Sender keeps track of the outstanding frames and updates the variables and windows as the ACKs arrive.



Go-Back-N ARQ, lost frame



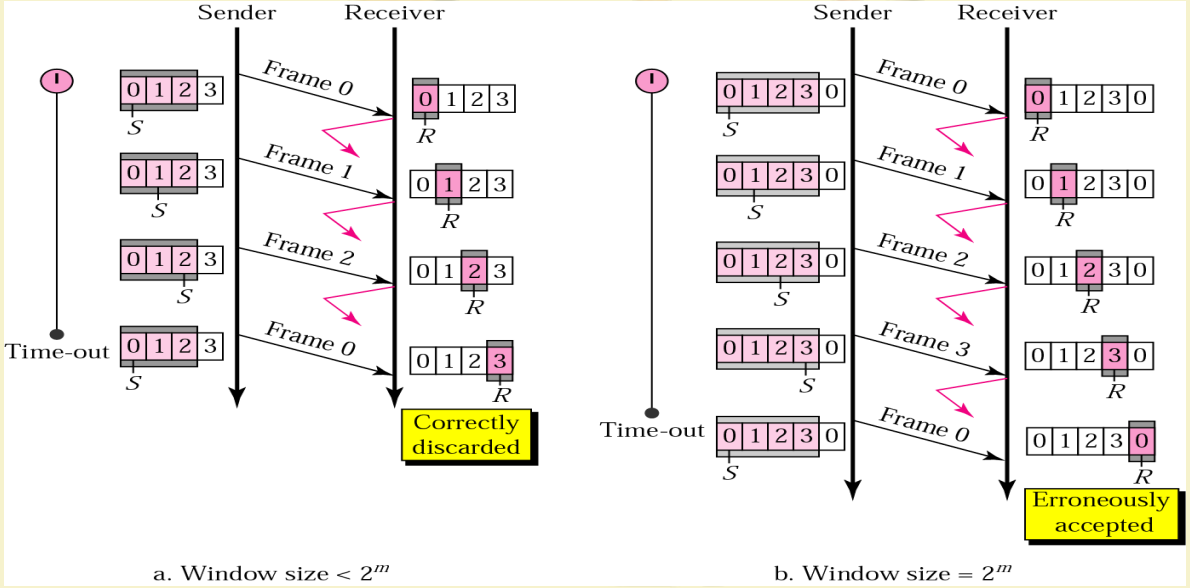
- Frame 2 is lost
- When the receiver receives frame 3, it discards frame 3 as it is expecting frame 2 (according to window).
- After the timer for frame 2 expires at the sender site, the sender sends frame 2 and 3. (go back to 2)

Go-Back-N ARQ, damaged/lost/delayed ACK

- If an ACK is damaged/lost, we can have two situations:
- If the next ACK arrives before the expiration of any timer, there is no need for retransmission of frames because ACKs are cumulative in this protocol.
- If ACK1, ACK2, and ACk3 are lost, ACK4 covers them if it arrives before the timer expires.
- If ACK4 arrives after time-out, the last frame and all the frames after that are resent.
- Receiver never resends an ACK.
- A delayed ACK also triggers the resending of frames

Go-Back-N ARQ, sender window size

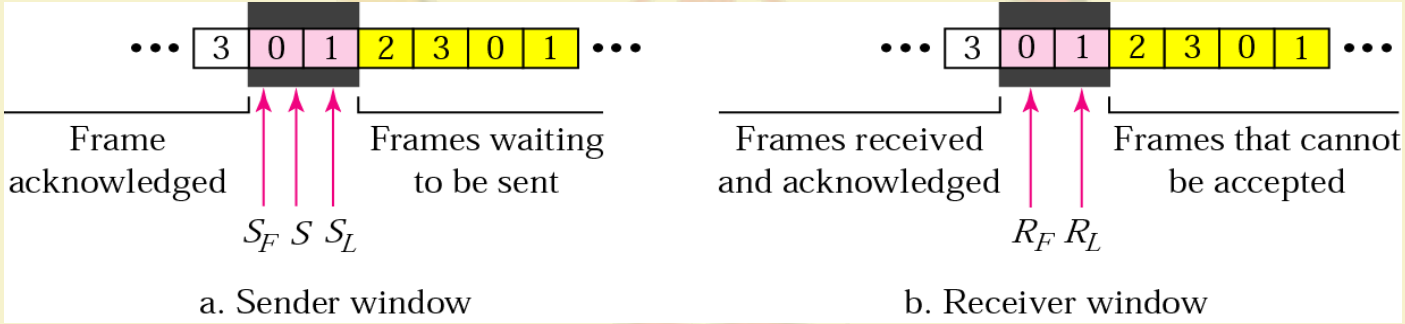
- Size of the sender window must be less than 2^m. Size of the receiver is always 1. If m = 2, window size = $2^m - 1 = 3$.
- Figure compares a window size of 3 and 4.



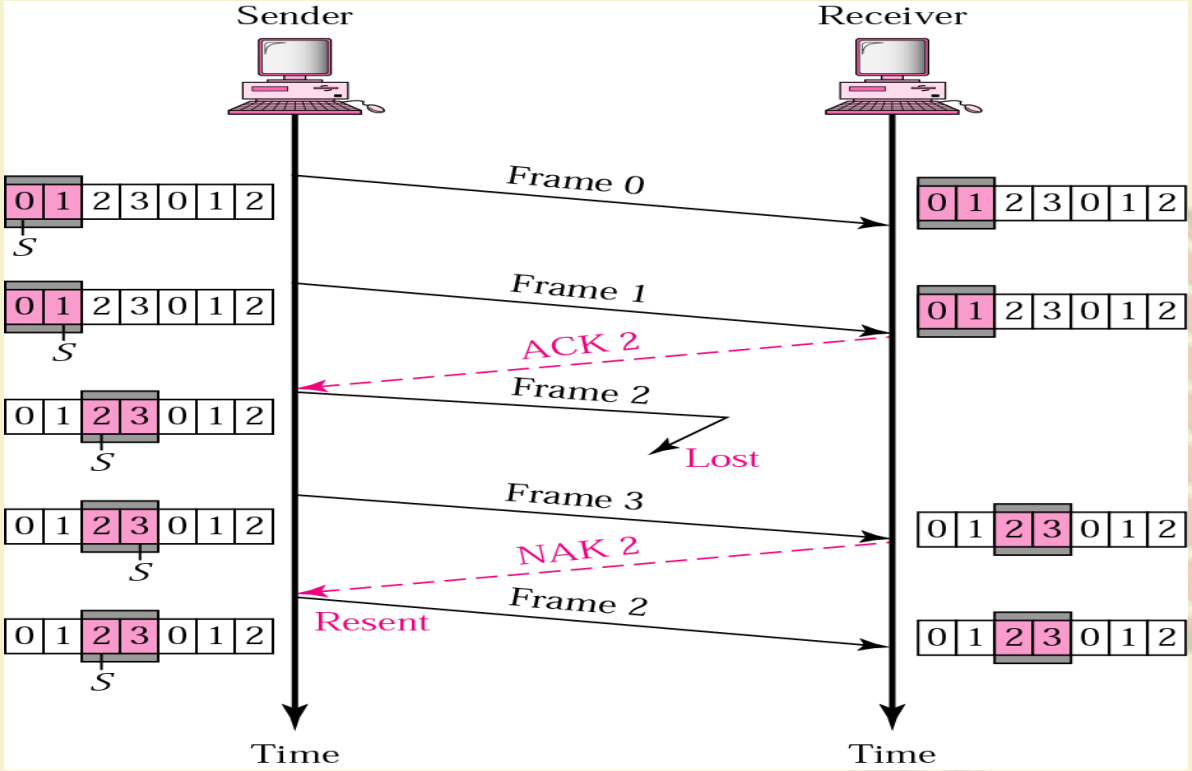
Accepts as the 1st frame in the next cycle-an **error**

Selective Repeat ARQ, sender and receiver windows

- Go-Back-N ARQ simplifies the process at the receiver site. Receiver only keeps track of only one variable, and there is no need to buffer out-of-order frames, they are simply discarded.
- However, Go-Back-N ARQ protocol is inefficient for noisy link. It bandwidth inefficient and slows down the transmission.
- In Selective Repeat ARQ, only the damaged frame is resent. More bandwidth efficient but more complex processing at receiver.
- It defines a negative ACK (NAK) to report the sequence number of a damaged frame before the timer expires.



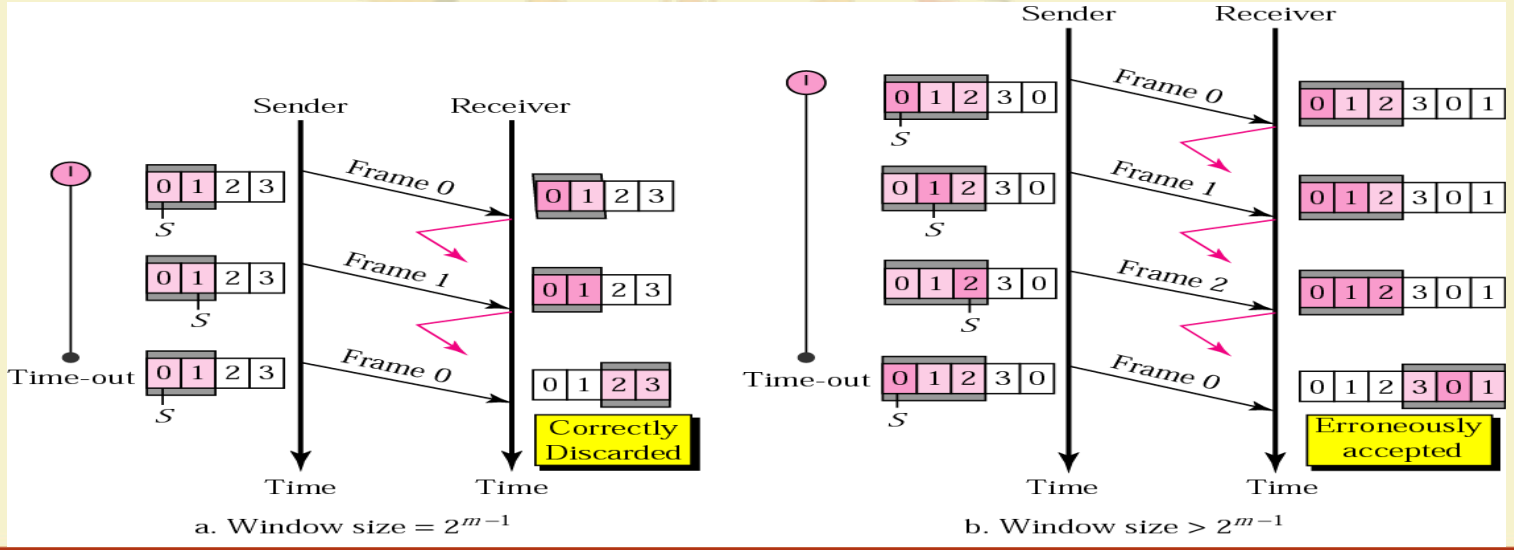
Selective Repeat ARQ, lost frame



- Frames 0 and 1 are accepted when received because they are in the range specified by the receiver window. Same for frame 3.
- Receiver sends a NAK2 to show that frame 2 has not been received and then sender resends only frame 2 and it is accepted as it is in the range of the window.

Selective Repeat ARQ, sender window size

- Size of the sender and receiver windows must be at most one-half of 2^m.
- If m = 2, window size should be $2^m / 2 = 2$.
- Window size is 3 and all ACKs are lost, sender sends duplicate of frame 0, window of the receiver expect to receive frame 0 (part of the window), so accepts frame 0, as the 1st frame of the next cycle – an **error**.




thank you!