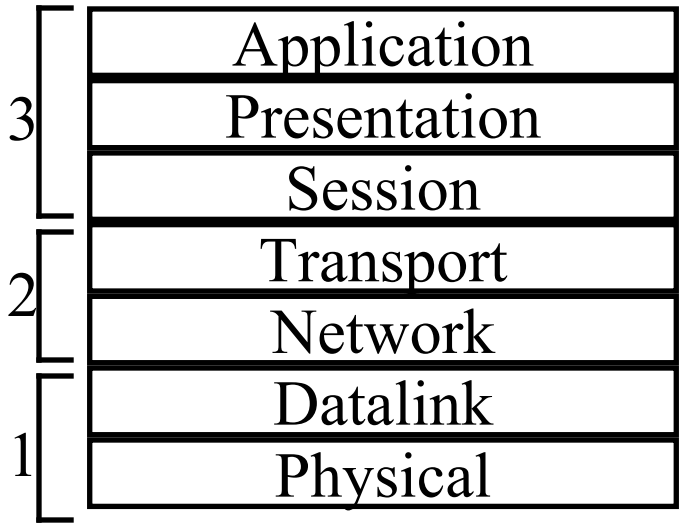
A Review of Key Networking Concepts

Raj Jain

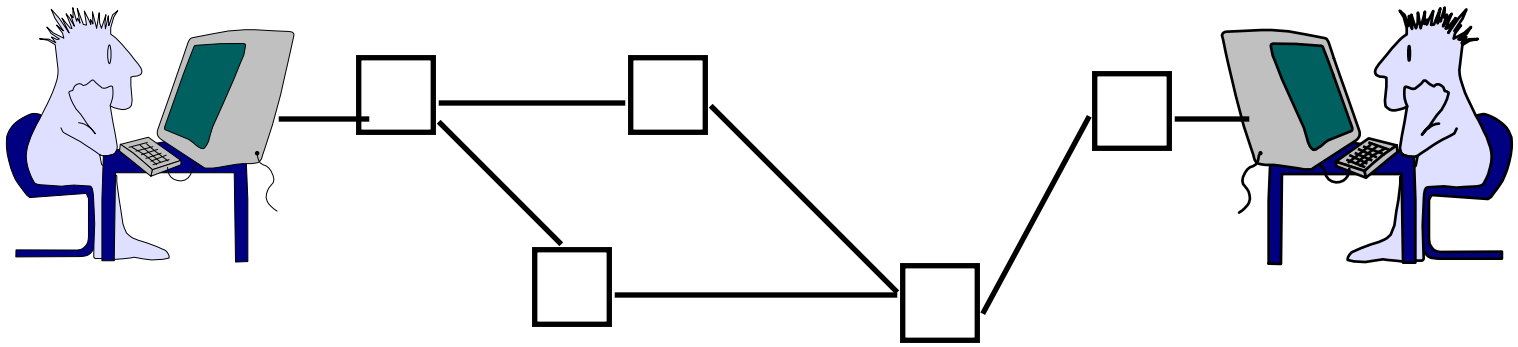
Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

- ISO/OSI and TCP/IP Reference Model
- Ethernet, Fast Ethernet, Gigabit Ethernet
- Interconnecting Devices: Hubs, bridges, routers

ISO/OSI Reference Model

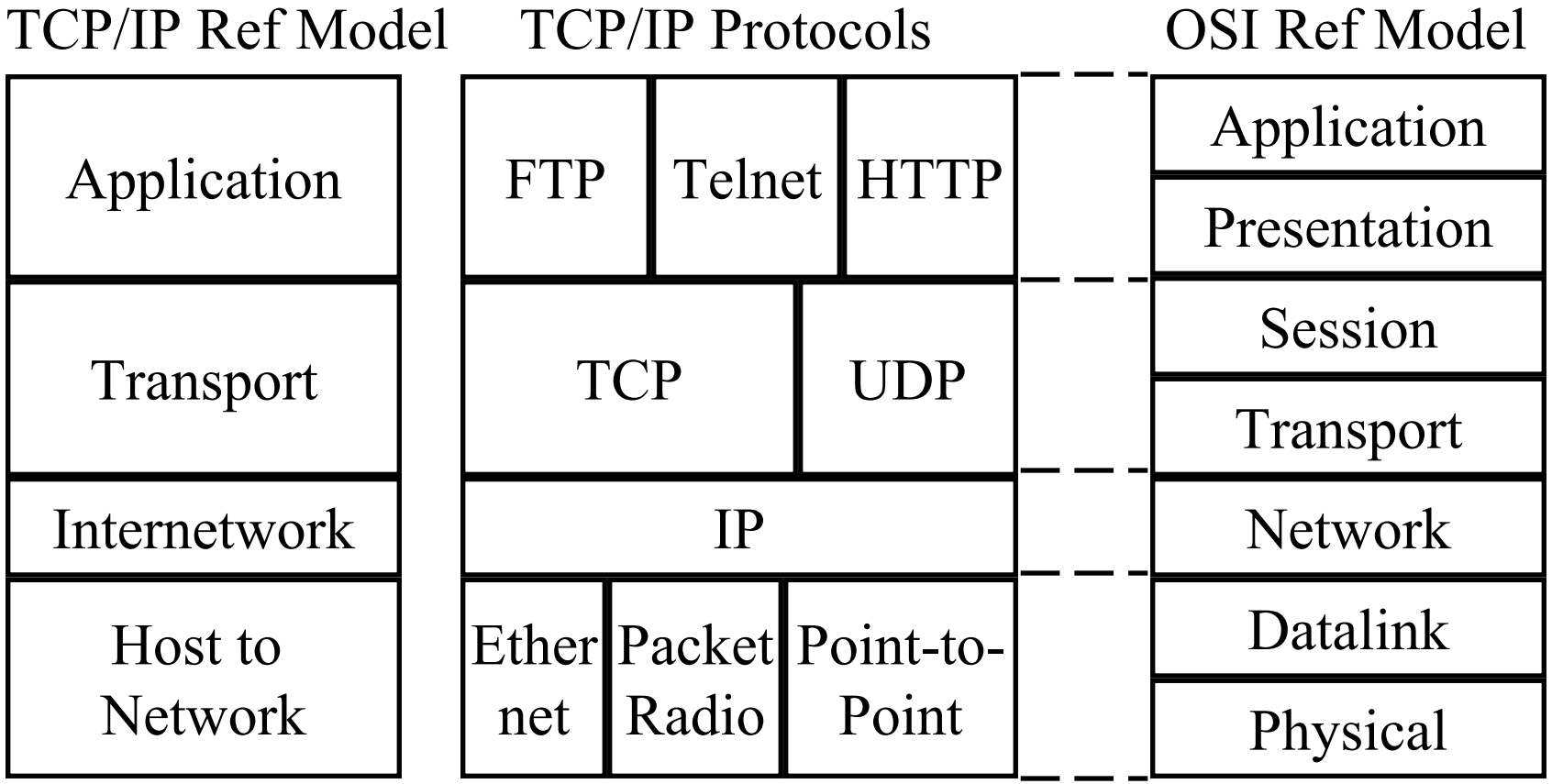


File transfer, Email, Remote Login
ASCII Text, Sound
Establish/manage connection
End-to-end communication: TCP
Routing, Addressing: IP
Two party communication: Ethernet
How to transmit signal: Coding



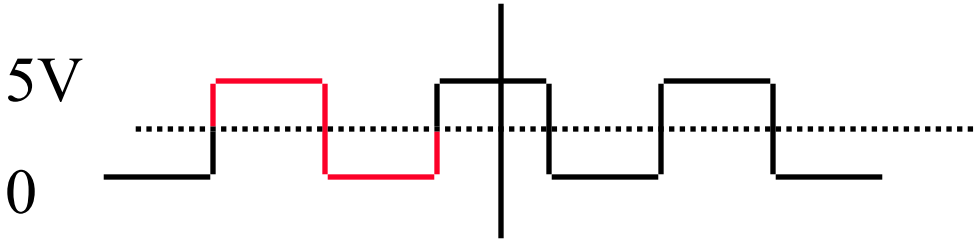
TCP/IP Reference Model

- TCP = Transport Control Protocol
- IP = Internet Protocol (Routing)



Channel Capacity

- Capacity = Maximum data rate for a channel
- **Nyquist Theorem**: Bandwidth = W
  Data rate $\leq$ 2 W
- Bilevel Encoding: Data rate = 2 × Bandwidth



- Multilevel Encoding: Data rate = 2 × Bandwidth × $\log_2$ M

Example: M=4, Capacity = 4 × Bandwidth

Shannon's Theorem

- Bandwidth = H Hz
  Signal-to-noise ratio = S/N
- Maximum number of bits/sec = $H \log_2 (1+S/N)$
- Example: Phone wire bandwidth = 3100 Hz

$$S/N = 30 \text{ dB}$$

$$10 \text{ Log}_{10} S/N = 30$$

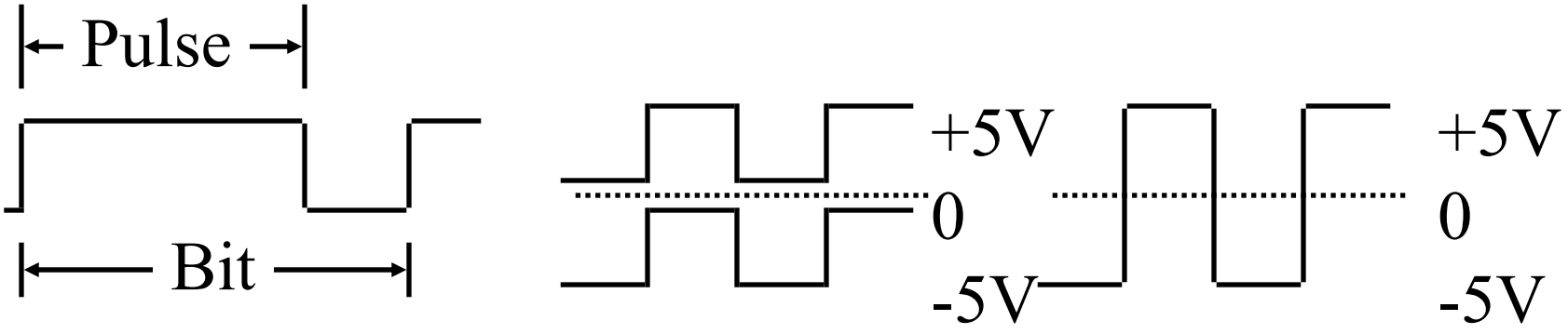
$$\text{Log}_{10} S/N = 3$$

$$S/N = 10^3 = 1000$$

$$\text{Capacity} = 3100 \log_2 (1+1000)$$

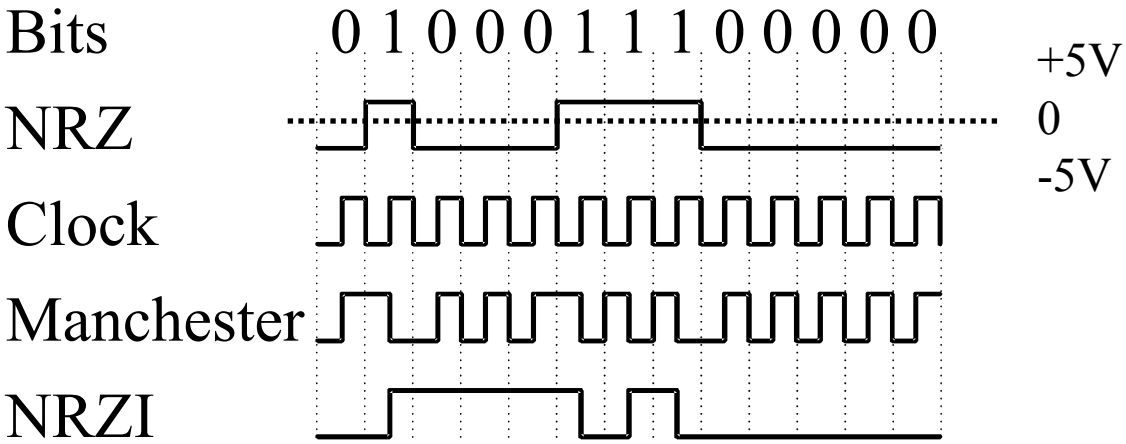
$$= 30{,}894 \text{ bps}$$

Coding Terminology



- Signal element: Pulse
- Modulation Rate: 1/Duration of the smallest element
  =Baud rate
- Data Rate: Bits per second
- Data Rate = Fn(Bandwidth, signal/noise ratio, encoding)

Coding Design



- Pulse width indeterminate: Clocking
- DC, Baseline wander
- No line state information
- No error detection/protection
- No control signals
- High bandwidth
- Polarity mix-up ⇒ Differential (compare polarity)

(a) Multiple Access

(b) Carrier-Sense Multiple Access with Collision Detection

CSMA/CD

- Aloha at Univ of Hawaii:
  Transmit whenever you like
  Worst case utilization = 1/(2e) =18%
- Slotted Aloha: Fixed size transmission slots
  Worst case utilization = 1/e = 37%
- CSMA: Carrier Sense Multiple Access
  Listen before you transmit
- p-Persistent CSMA: If idle, transmit with probability p. Delay by one time unit with probability 1-p
- CSMA/CD: CSMA with Collision Detection
  Listen while transmitting. Stop if you hear someone else

IEEE 802.3 CSMA/CD

- If the medium is idle, transmit (1-persistent).
- If the medium is busy, wait until idle and then transmit immediately.
- If a collision is detected while transmitting,
  - Transmit a jam signal for one slot
    (= 51.2 □s = 64 byte times)
  - Wait for a random time and reattempt (up to 16 times)
  - Random time = Uniform[0,$2^{\min(k,10)}$-1] slots
- Collision detected by monitoring the voltage
  High voltage □ two or more transmitters □ Collision
  □ Length of the cable is limited to 2 km

Ethernet Standards

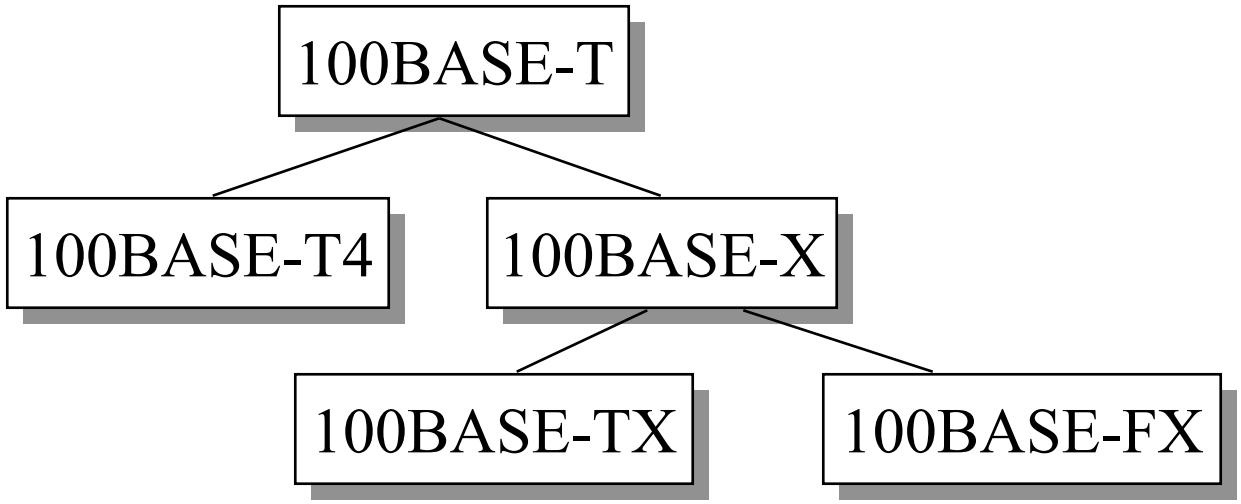
- 10BASE5: 10 Mb/s over coaxial cable (ThickWire)
- 10BROAD36: 10 Mb/s over broadband cable, 3600 m max segments
- 1BASE5: 1 Mb/s over 2 pairs of UTP
- 10BASE2: 10 Mb/s over thin RG58 coaxial cable (ThinWire), 185 m max segments
- 10BASE-T: 10 Mb/s over 2 pairs of UTP
- 10BASE-FL: 10 Mb/s fiber optic point-to-point link
- 10BASE-FB: 10 Mb/s fiber optic backbone (between repeaters). Also, known as synchronous Ethernet.

Ethernet Standards (Cont)

- 10BASE-FP: 10 Mb/s fiber optic passive star + segments
- 10BASE-F: 10BASE-FL, 10BASE-FB, or 10BASE-FP
- 100BASE-T4: 100 Mb/s over 4 pairs of CAT-3, 4, 5 UTP
- 100BASE-TX: 100 Mb/s over 2 pairs of CAT-5 UTP or STP
- 100BASE-FX: 100 Mbps CSMA/CD over 2 optical fiber

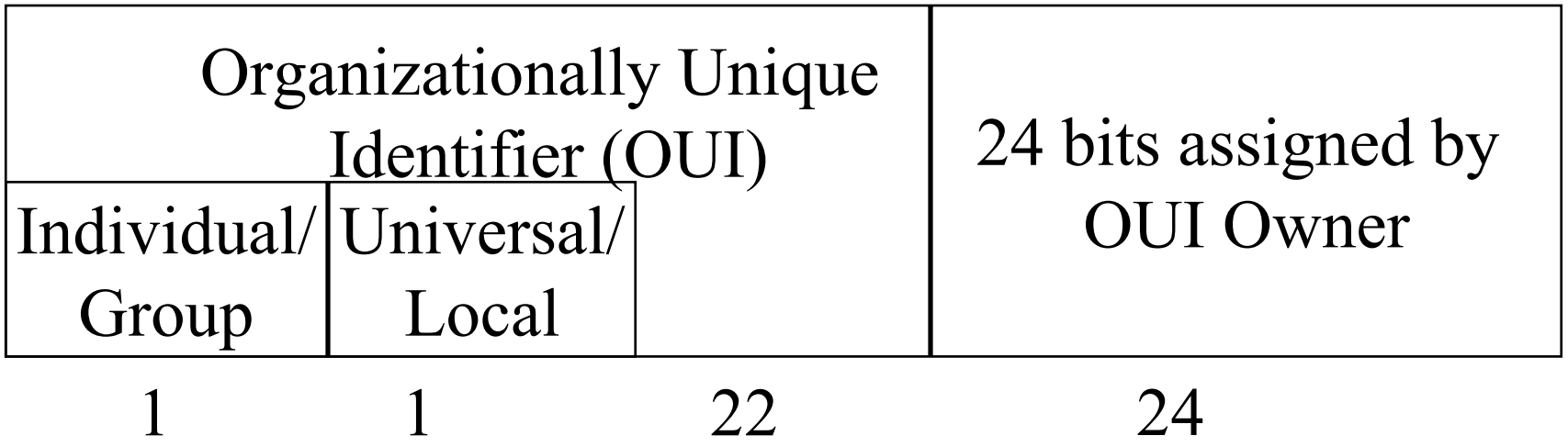
Ethernet Standards (Cont)

- 100BASE-X: 100BASE-TX or 100BASE-FX
- 100BASE-T: 100BASE-T4, 100BASE-TX, or 100BASE-FX
- 1000BASE-T: 1 Gbps (Gigabit Ethernet)



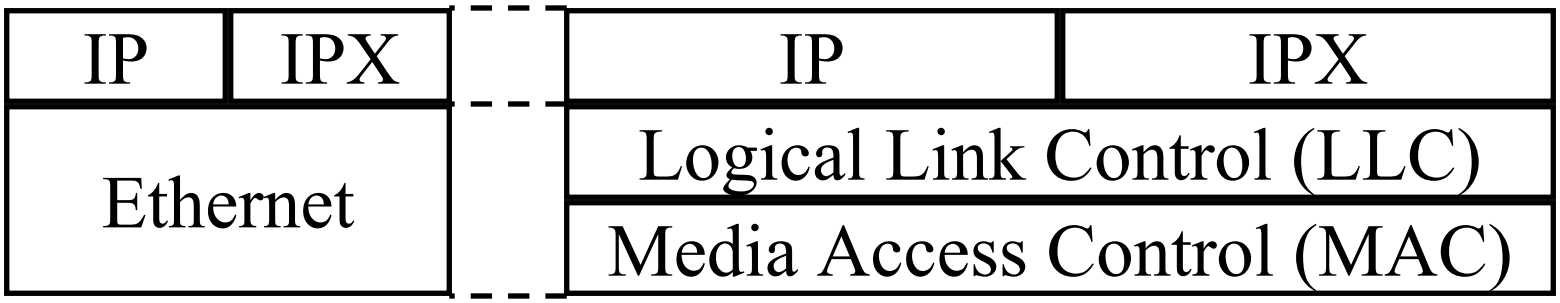
IEEE 802 Address Format

- 48-bit:1000 0000 : 0000 0001 : 0100 0011 : 0000 0000 : 1000 0000 : 0000 1100
  = 80:01:43:00:80:0C



- Multicast = "To all bridges on this LAN"
- Broadcast = "To all stations"
  = 111111....111 = FF:FF:FF:FF:FF:FF

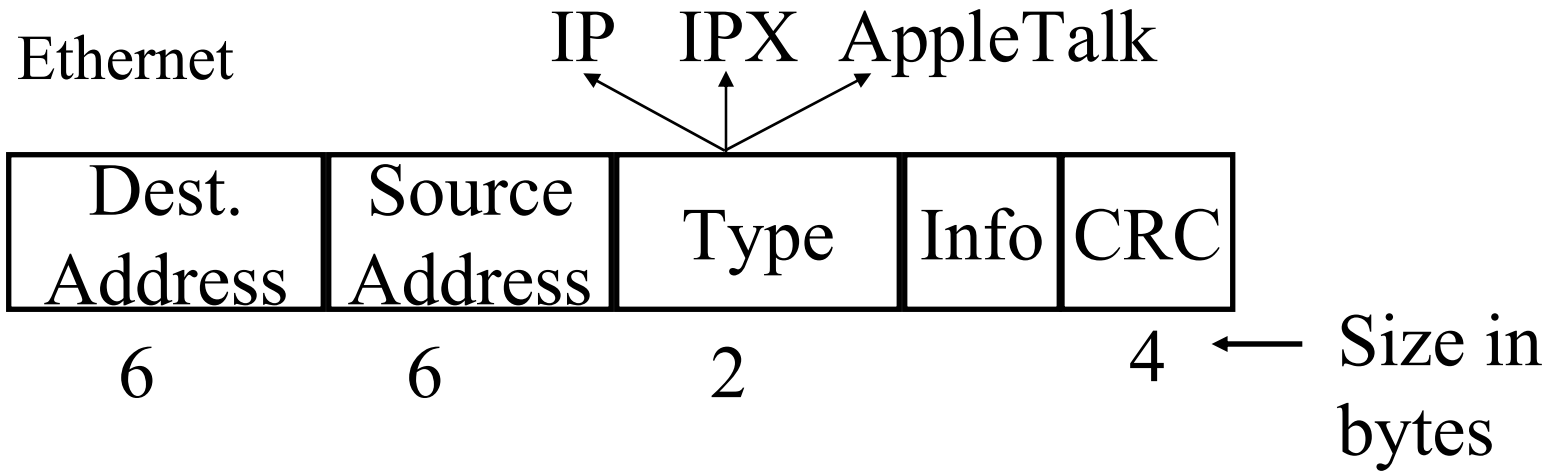
Ethernet vs IEEE 802.3



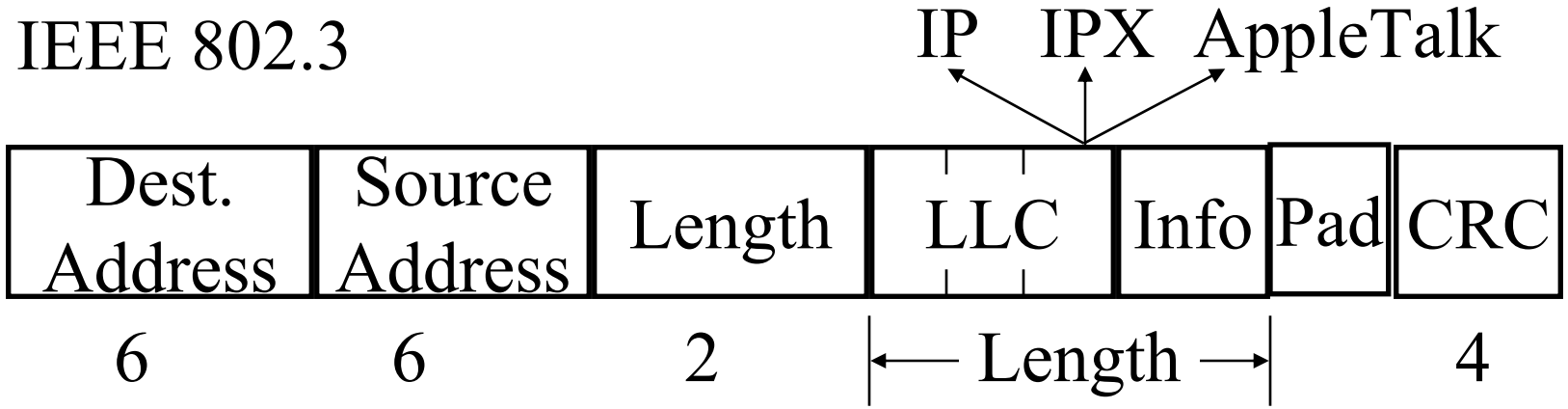
- In 802.3, datalink was divided into two sublayers: LLC and MAC
- LLC provides protocol multiplexing. MAC does not.
- MAC does not need a protocol type field.

Frame Format

- Ethernet



- IEEE 802.3



LLC Type 1

- Unacknowledged connectionless (on 802.3)
  No flow or error control.
  Provides protocol multiplexing.
  Uses 3 types of protocol data units (PDUs):
  UI = Unnumbered informaton
  XID = Exchange ID
  = Types of operation supported, window
  Test = Loop back test

LLC Type 2, 3

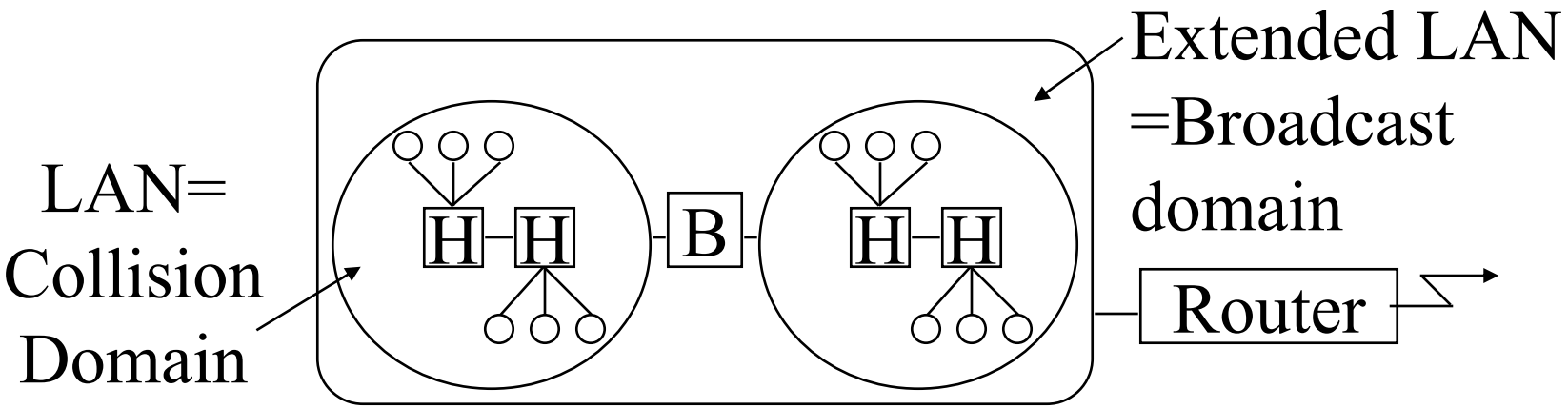
- Type 2: Acknowledged connection oriented (on 802.5)
  Provides flow control, error control. Uses
  SABME (Set asynchronous balanced mode), UA (unnumbered ack), DM (disconneced mode), DISC (disconnect)
- Type 3: Acknowledged connectionless
  Uses one-bit sequence number
  AC command PDUs acked by AC response PDUs

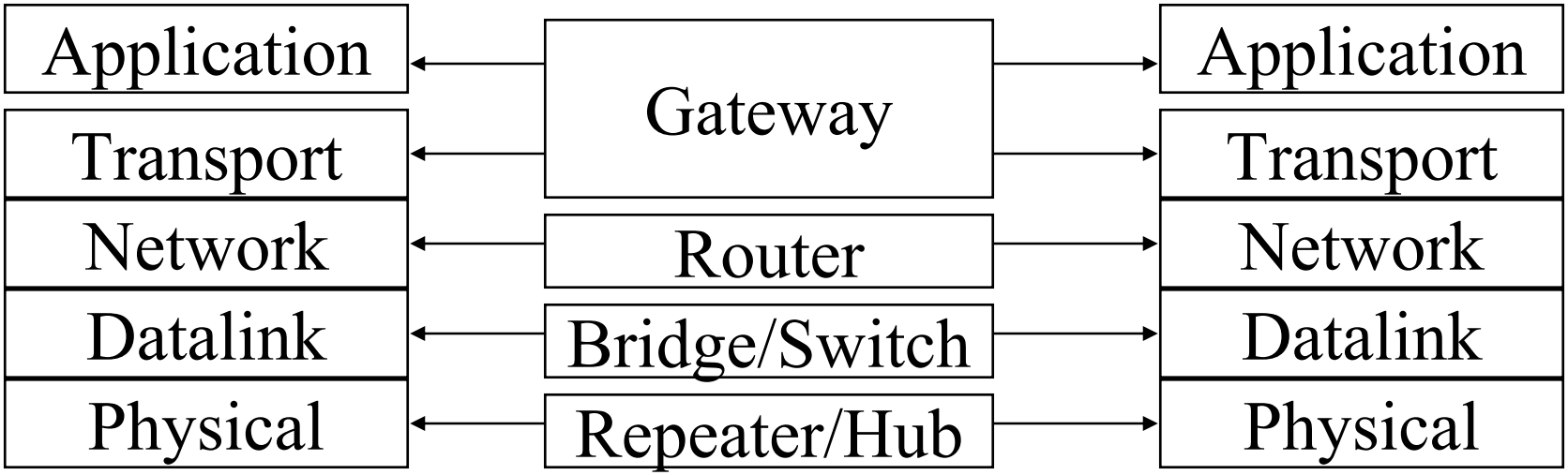
Interconnection Devices

- **Repeater**: PHY device that restores data and collision signals
- **Hub:** Multiport repeater + fault detection and recovery
- **Bridge:** Datalink layer device connecting two or more collision domains. MAC multicasts are propagated throughout "extended LAN."
- **Router:** Network layer device. IP, IPX, AppleTalk. Does not propagate MAC multicasts.
- **Switch**: Multiport bridge with parallel paths

These are functions. Packaging varies.

Interconnection Devices





2. Spanning Tree

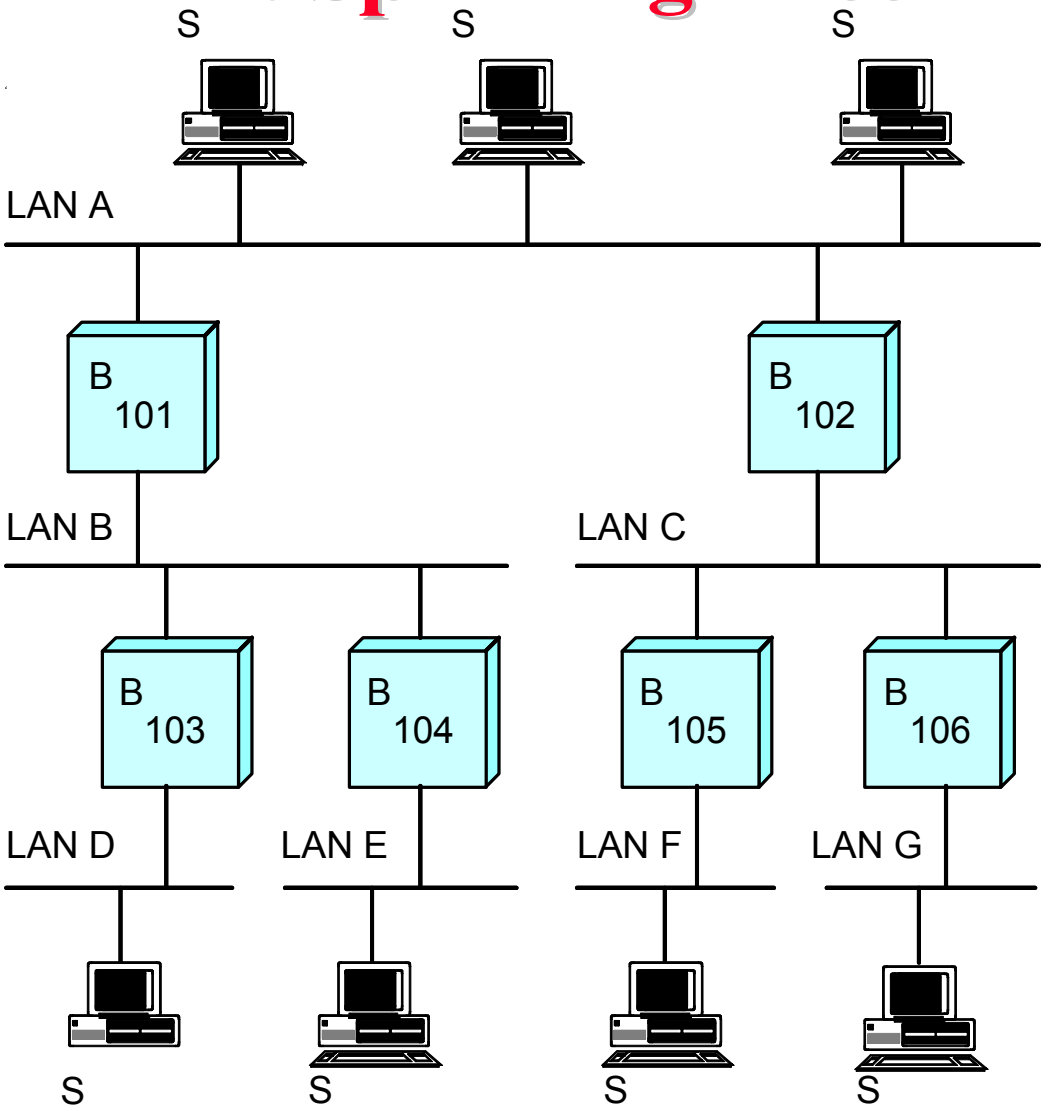


Fig 14.5

Spanning Tree (Cont)

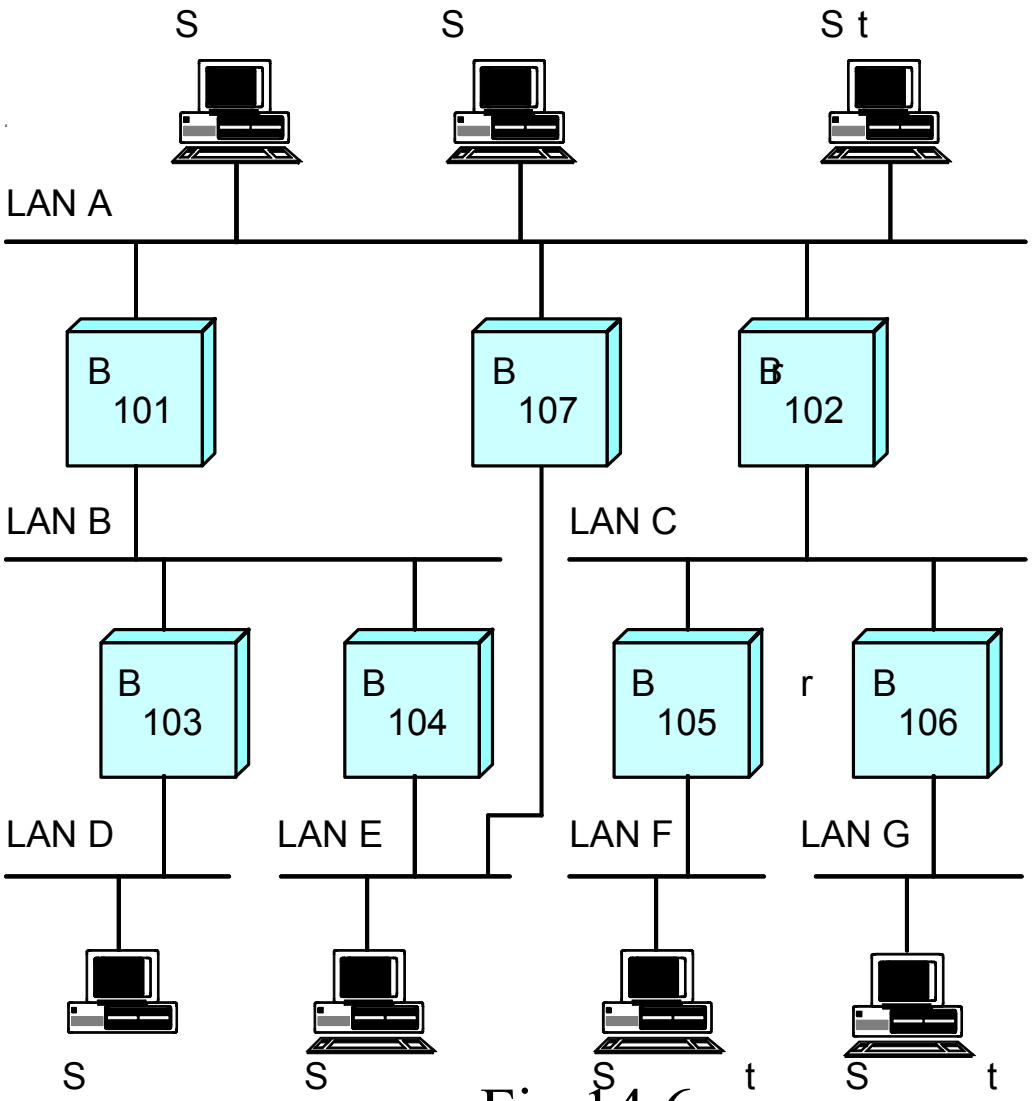
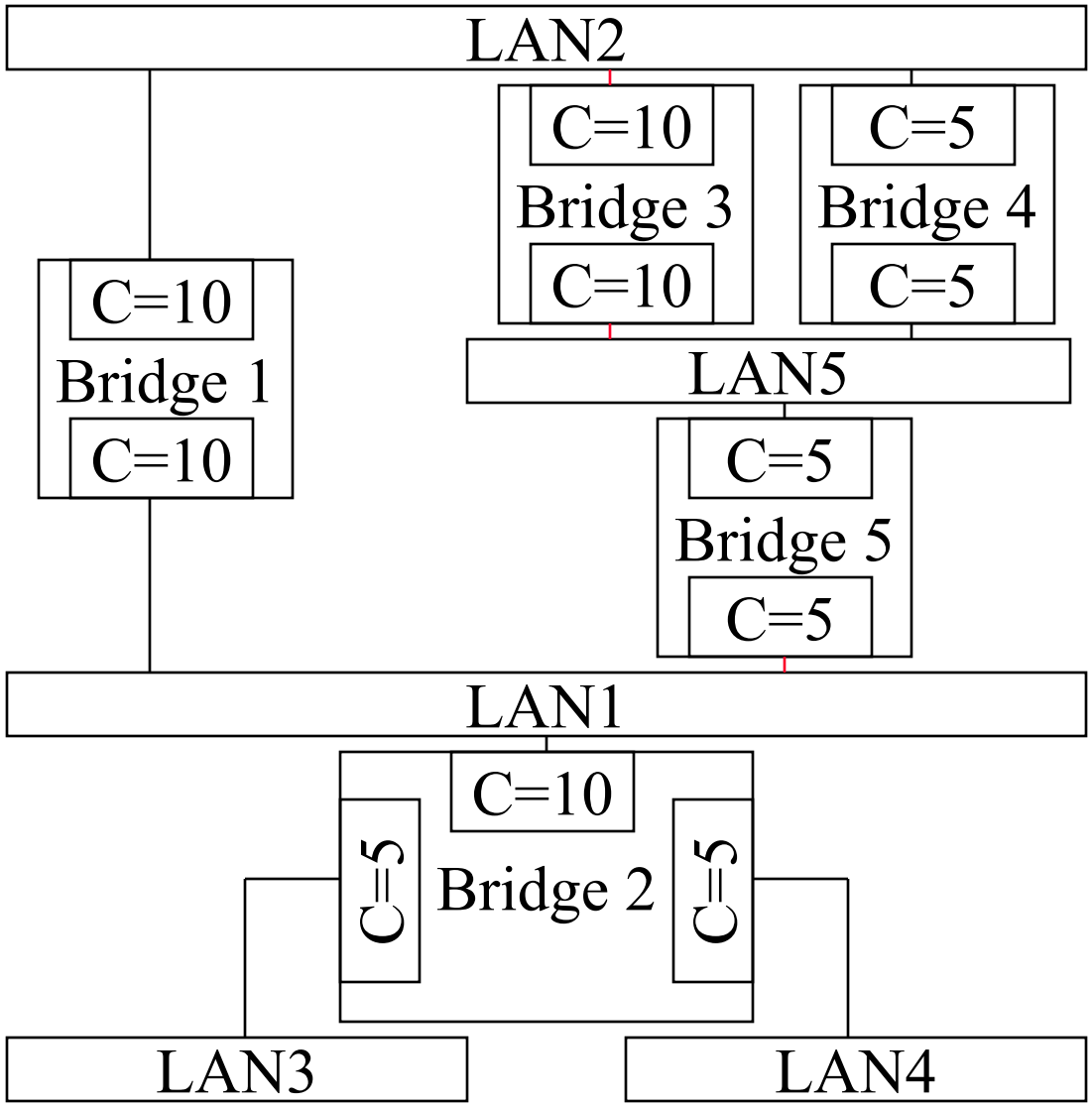


Fig 14.6

Spanning Tree Algorithm

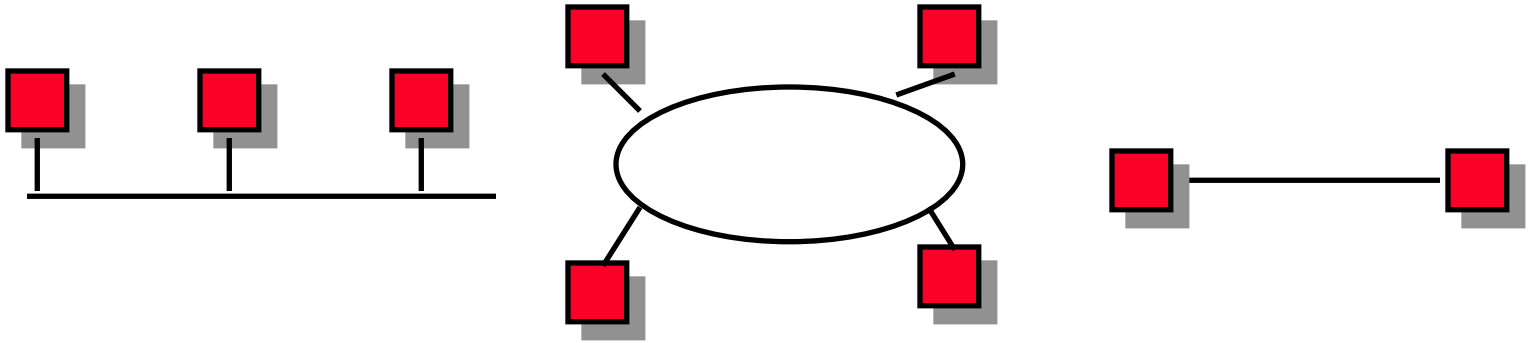
- All bridges multicast to "All bridges"
  - My ID
  - Root ID
  - My cost to root
- The bridges update their info using Dijkstra's algorithm and rebroadcast
- Initially all bridges are roots but eventually converge to one root as they find out the lowest Bridge ID.
- On each LAN, the bridge with minimum cost to the root becomes the Designated bridge
- All ports of all non-designated bridges are blocked.

Spanning Tree Example



The Magic Word α

Distance-B/W Principle



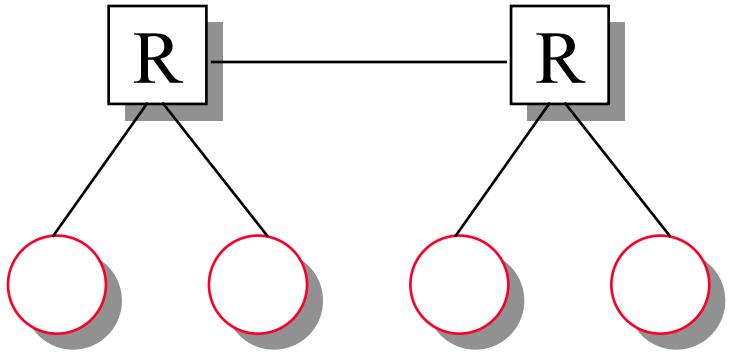
- Efficiency = Max throughput/Media bandwidth
- Efficiency is a nonincreasing function of α
  α = Propagation delay /Transmission time
  = (Distance/Speed of light)/(Transmission size/Bits/sec)
  = Distance×Bits/sec/(Speed of light)(Transmission size)
- Bit rate-distance-transmission size tradeoff.
- 100 Mb/s $\Rightarrow$ Change distance or frame size

(a) Multiple Access

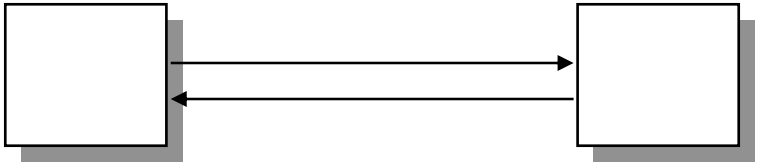
(b) Carrier-Sense Multiple Access with Collision Detection

Ethernet vs Fast Ethernet

	Ethernet	Fast Ethernet
Speed	10 Mbps	100 Mbps
MAC	CSMA/CD	CSMA/CD
Network diameter	2.5 km	205 m
Topology	Bus, star	Star
Cable	Coax, UTP, Fiber	UTP, Fiber
Standard	802.3	802.3u
Cost	X	2X



Full-Duplex Ethernet

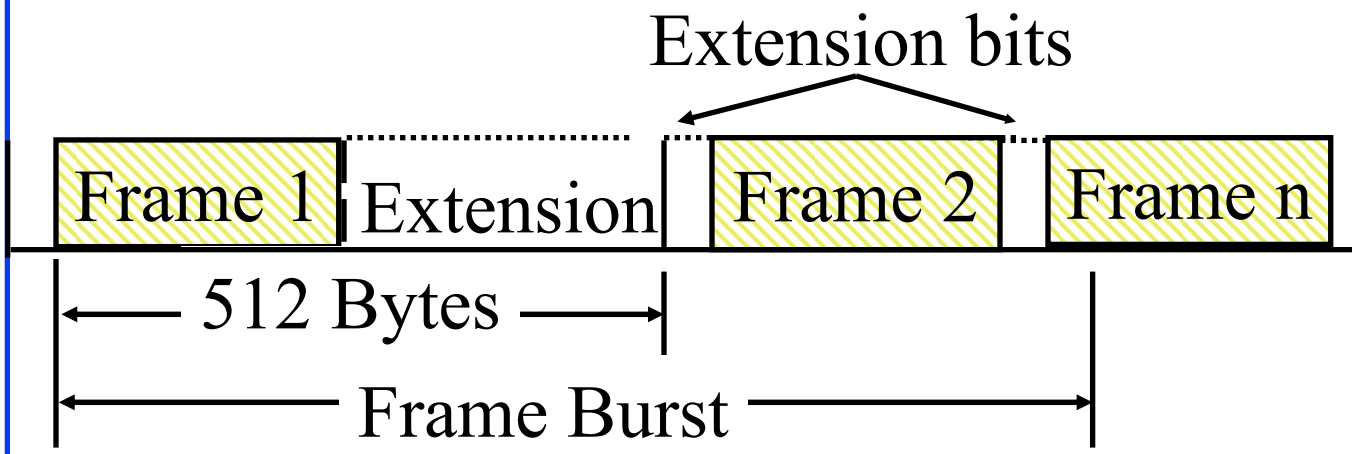


- Uses point-to-point links between TWO nodes
- Full-duplex bi-directional transmission
- Transmit any time
- Many vendors are shipping switch/bridge/NICs with full duplex
- No collisions $\Rightarrow$ 50+ Km on fiber.
- Between servers and switches or between switches

Carrier Extension

- 10 Mbps at 2.5 km ⇒ Slot time = 64 bytes
- 1 Gbps at 200 m ⇒ Slot time = 512 bytes
- Continue transmitting control symbols.
  Collision window includes the control symbols
- Control symbols are discarded at the destination
- Net throughput for small frames is only marginally better than 100 Mbps

Frame Bursting



- Don't give up the channel after every frame
- After the slot time, continue transmitting additional frames (with minimum inter-frame gap)
- Interframe gaps are filled with extension bits
- No no new frame transmissions after 8192 bytes
- Three times more throughput for small frames

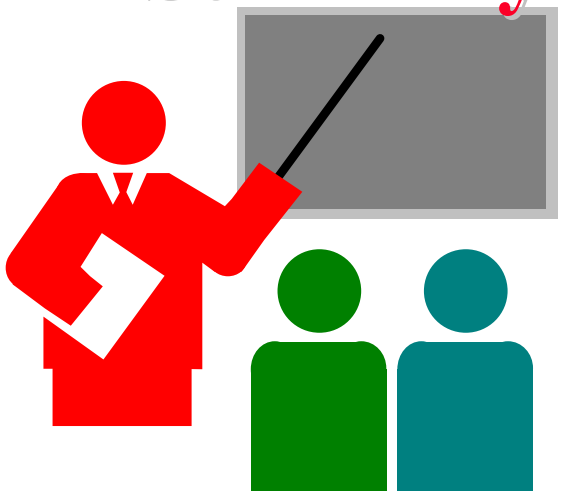
1000Base-X

- 1000Base-LX: 1300-nm laser transceivers
  - 2 to 550 m on 62.5-µm or 50-µm multimode, 2 to 3000 m on 10-µm single-mode
- 1000Base-SX: 850-nm laser transceivers
  - 2 to 300 m on 62.5-µm, 2 to 550 m on 50-µm. Both multimode.
- 1000Base-CX: Short-haul copper jumpers
  - 25 m 2-pair shielded twinax cable in a single room or rack.
    Uses 8b/10b coding ⇒ 1.25 Gbps line rate

1000Base-T

- 100 m on 4-pair Cat-5 UTP
  ⇒ Network diameter of 200 m
- 250 Mbps/pair full duplex DSP based PHY
  ⇒ Requires new 5-level (PAM-5) signaling
  with 4-D 8-state Trellis code FEC
- Automatically detects and corrects pair-swapping, incorrect polarity, differential delay variations across pairs
- Autonegotiation ⇒ Compatibility with 100Base-T
- 802.3ab task force began March'97, ballot July'98, Final standard by March'99.

Summary



- ISO/OSI reference model has seven layers.
  TCP/IP Protocol suite has four layers.
- Ethernet/IEEE 802.3 uses CSMA/CD.
- Configuration rules depend upon physical medium
  10Base5, 10Base2, 10Base-T, 100Base-TX, etc.
- Addresses: Local vs Global, Unicast vs Broadcast.