The Link Layer and LANs

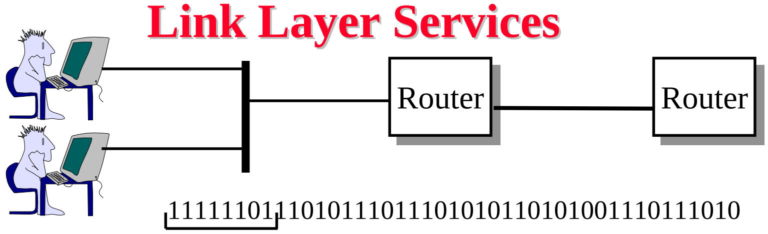
Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@wustl.edu

Audio/Video recordings of this lecture are available on-line at:
http://www.cse.wustl.edu/~jain/cse473-09/

1. Datalink Services
2. Error Detection
3. Multiple Access
4. Bridging
5. Point-to-Point Protocol and MPLS

Note: This class lecture is based on Chapter 5 of the textbook (Kurose and Ross) and the figures provided by the authors.

Link Layer Services

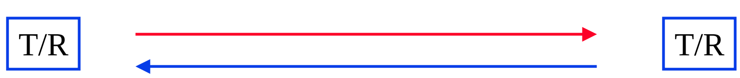


- Link = One hop
- Framing: Bit patterns at begin/end of a frame
- Multiple Access: Multiple users sharing a wire
- Flow Control
- Error Detection/Correction
- Reliable Delivery:
- Duplex Operation

Line Duplexity

- Simplex: Transmit or receive,
  e.g., Television

- Full Duplex: Transmit and receive simultaneously,
  e.g., Telephone



- Half-Duplex: Transmit and receive alternately,
  e.g., Police Radio

Error Detection

- Parity Checks
- Check Digit Method
- Modulo 2 Arithmetic
- Cyclic Redundancy Check (CRC)
- Popular CRC Polynomials

Parity Checks

1	0	1	1	1	0	1	0	
1	2	3	4	5	6	7	8	9

- Odd Parity

1	0	1	1	1	0	1	0	0
1	2	3	4	5	6	7	8	9

0	0	1	1	1	0	1	0	0
1	2	3	4	5	6	7	8	9

1-bit error

0	0	0	1	0	0	1	0	0
1	2	3	4	5	6	7	8	9

3-bit error

0	0	0	1	1	0	1	0	0
1	2	3	4	5	6	7	8	9

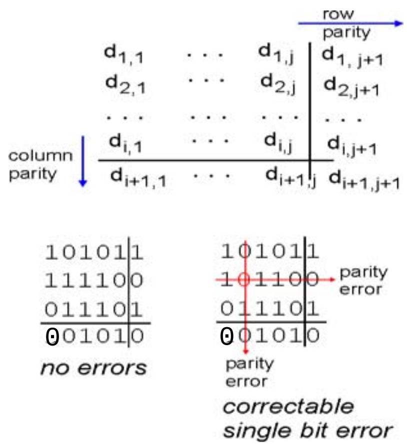
2-bit error

- Even Parity

1	0	1	1	1	0	1	1	0
1	2	3	4	5	6	7	8	9

Two Dimensional Parity

- Detect and correct single bit errors



Check Digit Method

- Make number divisible by 9

Example: 823 is to be sent

1. Left-shift: 8230
2. Divide by 9, find remainder: 4
3. Subtract remainder from 9: 9-4=5
4. Add the result of step 3 to step 1: 8235
5. Check that the result is divisible by 9.

Detects all single-digit errors: 7235, 8335, 8255, 8237

Detects several multiple-digit errors: 8765, 7346

Does not detect some errors: 7335, 8775, ...

Modulo 2 Arithmetic

```
  1111      11001          110
 +1010     ×    11    11  | 1010
 --------  --------      /  11
  0101      11001          --------    010   2
           11001            x11        011   3
           ---------         11        ----  --
           101011           -------    001   1  Mod 2
                              x00      101   5  Binary
                               00
                             -----
                               x0
```


Cyclic Redundancy Check (CRC)

- **Binary Check Digit Method**
- Make number divisible by P=110101 (n+1=6 bits)

Example: M=1010001101 is to be sent

1. Left-shift M by n bits 2^nM= 101000110100000
2. Divide 2^nM by P, find remainder: R=01110
3. ~~Subtract remainder from P~~ ← Not required in Mod 2
4. Add the result of step 2 to step 1 : T=101000110101110
5. Check that the result T is divisible by P.

Modulo 2 Division

```
        Q=1101010110
P=110101)101000110100000=2^nM
          110101
           111011            010110
           110101            000000
            011101            101100
            000000            110101
             111010            110010
             110101            110101
              011111            001110
              000000            000000
               111110            01110 = R
               110101
```


Checking At The Receiver

```
       1101010110
110101)101000110101110
       110101
        111011            010111
        110101            000000
         011101            101111
         000000            110101
          111010            110101
          110101            110101
           011111            00000
           000000
            111110
            110101
```

Polynomial Representation

- Number the bits 0, 1, ..., from right
  $b_n b_{n-1} b_{n-2} .... b_3 b_2 b_1 b_0$
  $b_n x^n + b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + ... + b_3 x^3 + b_2 x^2 + b_1 x + b_0$
- Example:
  543210
  ↓↓↓↓↓
  110101 = $x^5+x^4+x^2+1$
  1101 1001 0011 = $x^{11}+x^{10}+x^8+x^7+x^4+x+1$
  11 10 9 8 1 0

Cyclic Redundancy Check (CRC)

Polynomial Division Method

Make $T(x)$ divisible by $P(x) = x^5+x^4+x^2+1$
(Note: n=5)

Example: M=1010001101 is to be sent
$M(x) = x^9+x^7+x^3+x^2+1$

1. Multiply M(x) by x^n, $x^nM(x) = x^{14}+x^{12}+x^8+x^7+x^5+$ ....
2. Divide $x^nM(x)$ by $P(x)$, find remainder:
   $R(x)$=01110=x^3+x^2+x

CRC (Cont)

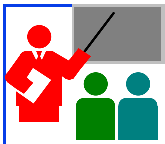
3. Add the remainder $R(x)$ to $x^n M(x)$:
   $T(x)= x^{14}+x^{12}+x^8+x^7+x^5+x^3+x^2+x$
4. Check that the result $T(x)$ is divisible by $P(x)$.

Transmit the bit pattern corresponding to $T(x)$:
101000110101110

Popular CRC Polynomials

- CRC-12: $x^{12}+x^{11}+x^3+x^2+x+1$
- CRC-16: $x^{16}+x^{15}+x^2+1$
- CRC-CCITT: $x^{16}+x^{12}+x^5+1$
- CRC-32: Ethernet, FDDI, ...
  $x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}$
  $+x^{10}+x^8+x^7+x^5+x^4+x^2+x+1$

Even number of terms in the polynomial
$\Rightarrow$ Polynomial is divisible by 1+x
$\Rightarrow$ Will detect all odd number of bit errors



Error Detection: Review

1. Parity bits can help detect/correct errors
2. Remainder obtained by diving by a prime number provides good error detection
3. CRC uses binary division
4. Binary numbers can be represented as polynomials
5. CRC-32 is a degree 31-polynomial used in most LANs.

Homework 5A

- Find the CRC of 1001100 using a generator polynomial of x^3+x+1. *Use polynomial or mod 2 division. Show all steps.*

Review Exercises

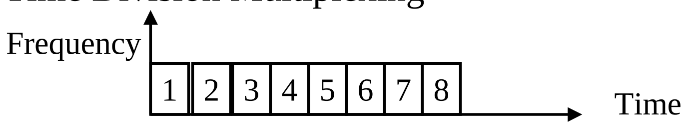
- Do not submit
- P1, P2, P5, P6, p7

Ethernet and ARP

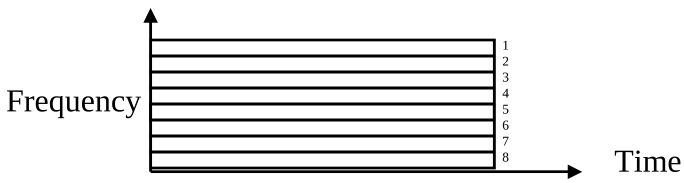
1. Multiple Access
2. CSMA/CD
3. IEEE 802.3 CSMA/CD
4. Ethernet Standards
5. CSMA/CD Performance
6. Distance-B/W Principle
7. Ethernet vs. Fast Ethernet
8. IEEE 802 Address Format
9. Address Resolution Protocol

Multiple Access

- How multiple users can share a link?
- Time Division Multiplexing



- Frequency Division Multiplexing



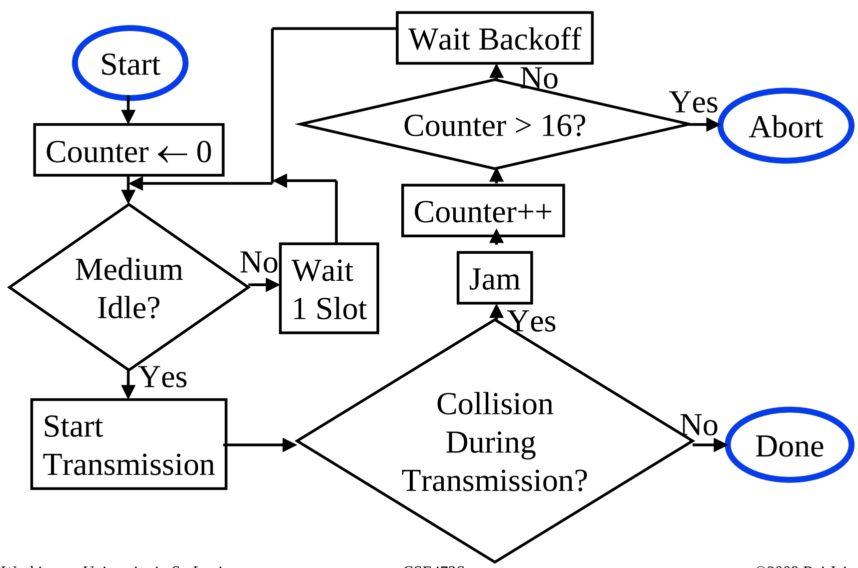
CSMA/CD

- 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

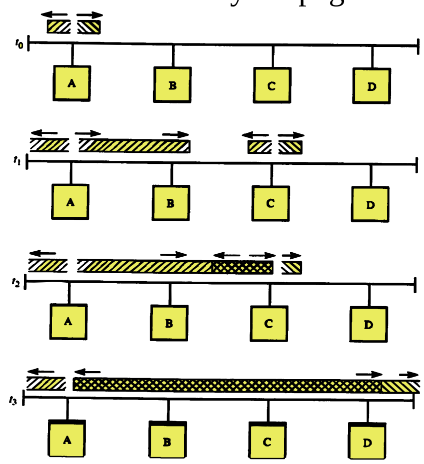
- If the medium is idle, transmit (1-persistent).
- If the medium is busy, wait until idle and then transmit immediately.
- If a collision is detected while transmitting,
  - Transmit a jam signal for one slot (= 51.2 μs = 64 byte times)
  - Wait for a random time and reattempt (up to 16 times)
  - Random time = Uniform[0,2min(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

IEEE 802.3 CSMA/CD Flow Chart



CSMA/CD Operation

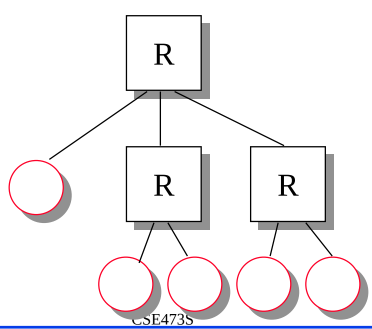
- Collision window = 2 × One-way Propagation delay = 51.2 μs



One way delay = 25.6 μs
Max Distance <2.5 km

10BASE-T

- Collision detected by the hub.
- Activity on two or more channels ⇒ Collision
  Collision presence (CP) transmitted by hub to all stations
  Collision window = 2× One-way delay between farthest stations

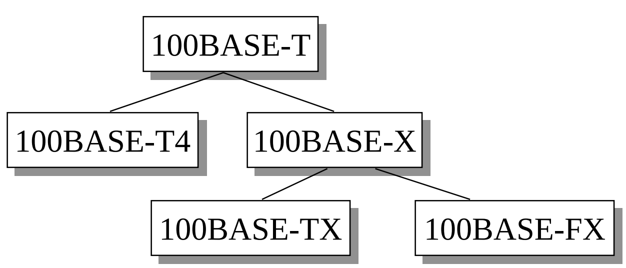


Ethernet Standards

- 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-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

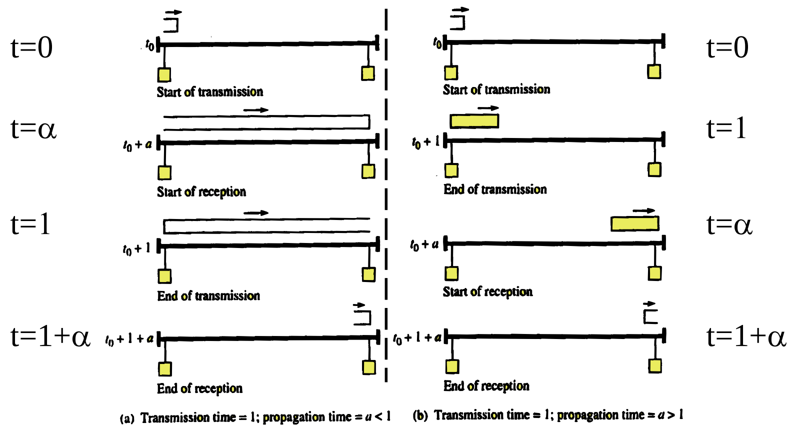
Ethernet Standards (Cont)

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



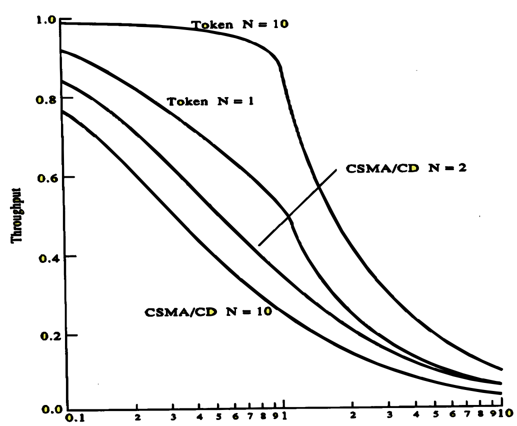
CSMA/CD Performance

- α = Propagation delay/Frame time
- U = Frame Time/(Propagation delay+Frame Time) = $1/(1+\alpha)$

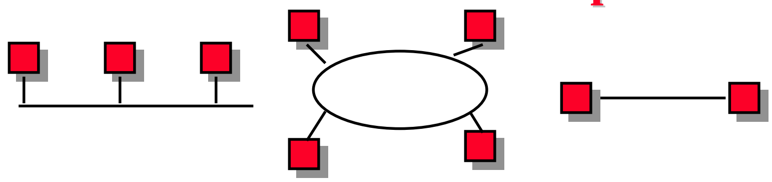


CSMA/CD Performance (Cont)

- $U=1/[1+2\alpha(1-A)/A]$, where $A = (1-1/N)^{N-1} \rightarrow e^{-1}$
- Worst case $U = 1/(1+3.44\alpha)$ with N= ∞



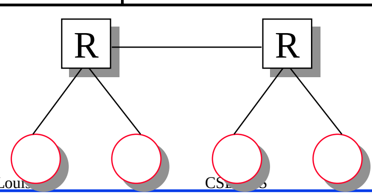
Distance-B/W Principle



- Efficiency = Max throughput/Media bandwidth
- Efficiency is a decreasing 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

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



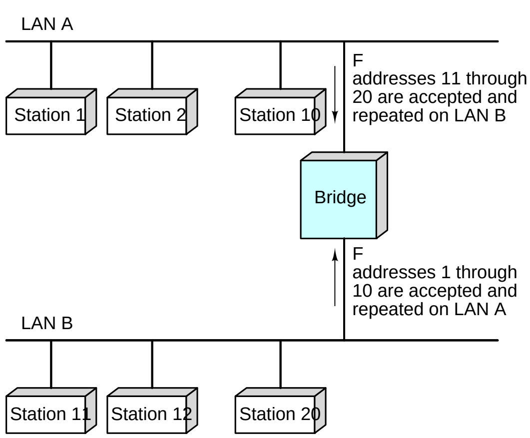
IEEE 802 Address Format

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

Organizationally Unique Identifier (OUI)			24 bits assigned by OUI Owner
Individual/ Group	Universal/ Local		
1	1	22	24

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

Bridges

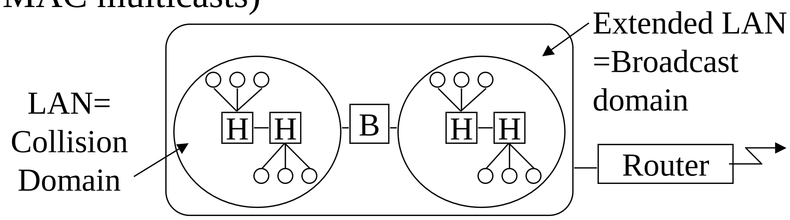


Bridge: Functions

- Monitor all frames on LAN A
- Pickup frames that are for stations on the other side
- Retransmit the frames on the other side
- Knows or learns about stations are on various sides
  Learns by looking at source addresses⇒ **Self-learning**
- Makes no modification to content of the frames.
  May change headers.
- Provides storage for frames to be forwarded
- Improves reliability (less nodes per LAN)
- Improves performance (more bandwidth per node)
- Security (Keeps different traffic from entering a LAN)
- May provide flow and congestion control

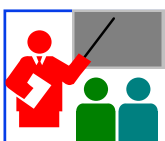
Interconnection Devices

- **Repeater**: PHY device that restores data and collision signals
- **Hub:** Multiport repeater + fault detection, notification and signal broadcast
- **Bridge:** Datalink layer device connecting two or more collision domains
- **Router:** Network layer device (does not propagate MAC multicasts)



Address Resolution Protocol

- Problem: Given an IP address find the MAC address
- Solution: Address resolution protocol
- The host broadcasts a request:
  "What is the MAC address of 127.123.115.08?"
- The host whose IP address is 127.123.115.08 replies back:
  "The MAC address for 127.123.115.08 is 8A-5F-3C-23-45-5616"
- A router may act as a proxy for many IP addresses



Ethernet and ARP: Review

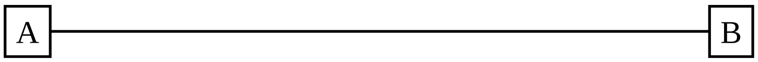
1. CSMA/CD = Listen while transmitting and stop on collision
2. IEEE 802.3 uses a *truncated binary exponential backoff*.
3. Protocol efficiency is a non-increasing function of α (propagation delay to transmission time ratio)
4. 100Base-T is $1/10^{th}$ the distance of 10Base-T in CSMA/CD mode
5. Ethernet uses 48-bit addresses of which the first bit is the universal/local, 2^{nd} bit is multicast/unicast, 22-bits are OUI (Organizationally unique identifier).
6. Address Resolution Protocol (ARP) is used to find the MAC address for a given IP address and vice versa.

Review Exercises

- Do not submit
- Review questions R1-R4, R8-R12
- Problems: P14-P23, P27, P28, P32

Homework 5B

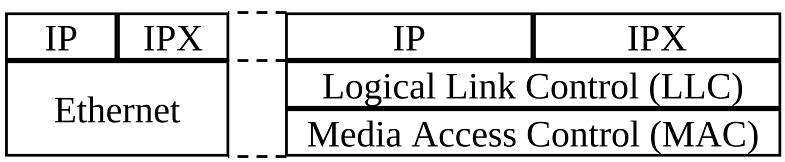
- Submit answer to the Problem 18:
- Suppose nodes A and B are on the same 10 Mbps Ethernet bus, and the propagation delay between the two nodes is 325 bit times. Suppose node A begins transmitting a frame and, before it finishes, node B begins transmitting a frame. Can A finish transmitting before it detects that B has transmitted? Why or why not? In the worst case when does B's signal reach A? (Minimum frame size is 512+64 bits).



LLC, VLANs, PPP, and MPLS

1. Ethernet vs. IEEE 802.3
2. Logical Link Control (LLC) Header
3. Full-Duplex Ethernet
4. Virtual LAN
5. PPP
6. Multiprotocol Label Switching (MPLS)

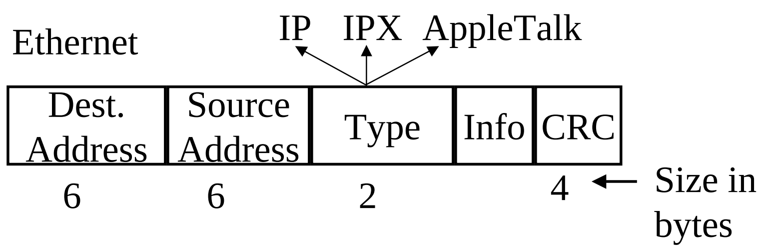
Ethernet vs. IEEE 802.3



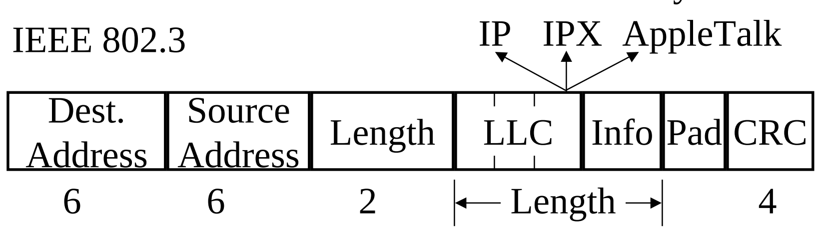
- In 802.3, datalink was divided into two sublayers: LLC and MAC
- LLC provides protocol multiplexing. MAC does not.
- MAC does not need a protocol type field.

Ethernet and 802.3 Frame Formats

- Ethernet



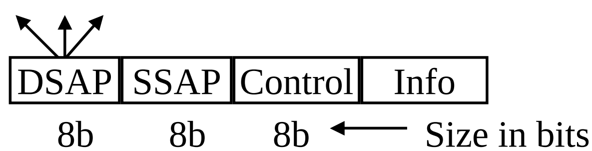
- IEEE 802.3



- Length > 1518 ⇒ It is a protocol type ⇒ Ethernet

LLC Header

- Multiple network layer protocols can share a datalink
- Each protocol is identified by a "service access point (SAP)"



- First bit of DSAP indicates Individual/Group
- First bit of SSAP indicates command/response
- Eight-bit SAP ⇒ Only 256 standard values possible
- Even IP couldn't get a standard SAP.
  Use Subnetwork Access Protocol SAP (SNAP SAP)

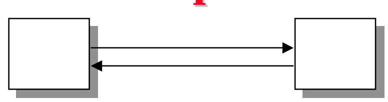
SNAP SAP

- SubNetwork Access Protocol Service Access Point
- When DSAP=AA, SSAP=AA, Control=UI,
  protocol ID field is used for multiplexing

DSAP	SSAP	Control		
AA	AA	03	Protocol ID	Info
			40 bits	

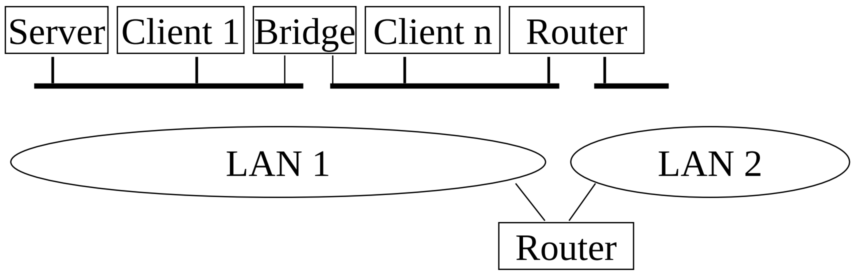
- Protocol ID is 40 bit long. The first 24 bits are Organizationally Unique Identifiers (OUI). OUI of 0 is used. The Ethernet type values are used in the last 16 bits.
  Protocol ID = 00-00-00-xx-xx

Full-Duplex Ethernet



- Uses point-to-point links between TWO nodes
- Full-duplex bi-directional transmission ⇒Transmit any time
- Not yet standardized in IEEE 802
- Many vendors are shipping switch/bridge/NICs with full duplex
- No collisions ⇒ 50+ Km on fiber.
- Between servers and switches or between switches
- CSMA/CD is no longer used (except in old 10/100 hubs)
- 1G Ethernet standard allows CSMA/CD but not implemented.
- 10G and higher speed Ethernet standards do not allow CSMA/CD

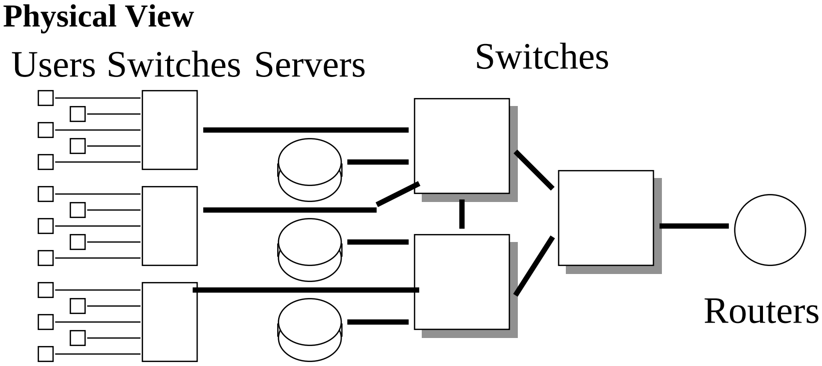
What is a LAN?



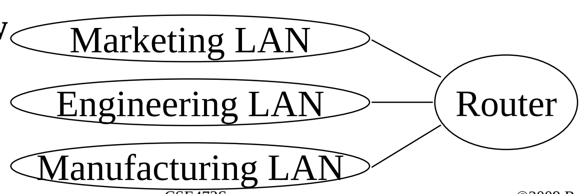
- LAN = Single broadcast domain = Subnet
- No routing between members of a LAN
- Routing required between LANs

What is a Virtual LAN

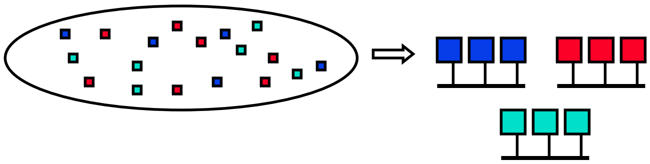
- **Physical View**



- **Logical View**



Virtual LAN



- Virtual LAN = Broadcasts and multicast goes only to the nodes in the virtual LAN
- LAN membership defined by the network manager $\Rightarrow$ Virtual

VLAN: Why?

- Virtual is Better than Real
  - Location-independent $\Rightarrow$ Marketing LAN can be all over the building
  - Users can move but not change LAN
  - Traffic between LANs is routed $\Rightarrow$ Better to keep all traffic on one LAN
  - Switch when you can, route when you must $\Rightarrow$ Do not VLAN over expensive WAN links
  - Better security

Types of Virtual LANs

- Layer-1 VLAN = Group of Physical ports
- Layer-2 VLAN = Group of MAC addresses
- Layer-3 VLAN = IP subnet

Switch Port	VLAN 1	VLAN 2
A1	√	
A2		√
A3	√	
B1		√
B1	√	

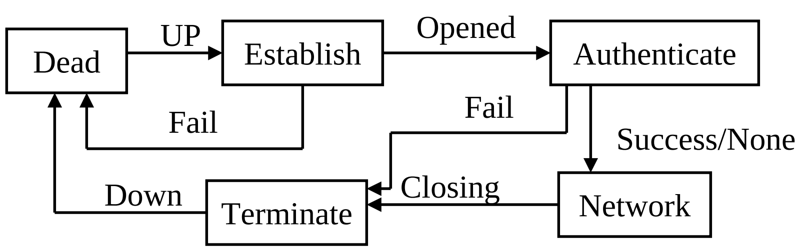
VLAN1	VLAN2
A1B234565600	21B234565600
D34578923434	634578923434
1345678903333	8345678903333
3438473450555	9438473450555
4387434304343	5387434304343
4780357056135	6780357056135
4153953470641	9153953470641
3473436374133	0473436374133
3403847333412	8403847333412
3483434343143	8483434343143
4343134134234	0343134134234

VLAN1: 23.45.6

VLAN2: IPX

PPP: Introduction

- Point-to-point Protocol
- Originally for User-network connection
- Now being used for router-router connection
- Three Components: Data encaptulation, Link Control Protocol (LCP), Network Control Protocols (NCP)



PPP (Cont)

- Typical connection setup:
  - Home PC Modem calls Internet Provider's router: sets up physical link
  - PC sends Link Control Protocol (LCP) packets
    - Select PPP (data link) parameters. Authenticate.
  - PC sends Network Control Protocol (NCP) packets
    - Select network parameters, E.g., Get IP address
- Transfer IP packets

PPP Design Requirements [RFC 1557]

- **Packet Framing**: Bit stream to frames
- **Protocol Multiplexing**: carry any network layer protocol (not just IP) at same time
- **Bit Transparency**: must carry any bit pattern in data
- **Error Detection**: (no correction)
- **Connection Liveness**: Signal link failures
- **Network Layer Address Negotiation**: Endpoints can learn/configure each other's network address
- **Non-Goals**:
  - No error correction/recovery
  - No flow control
  - Out of order delivery OK
  - No need to support multipoint links (e.g., polling)

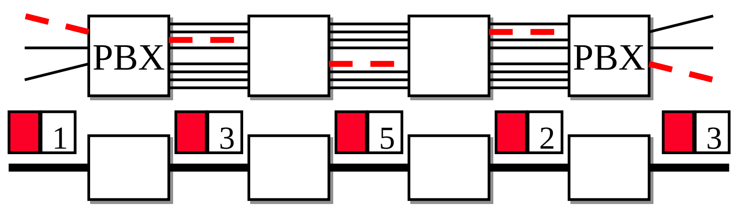
PPP in HDLC-Like Framing

Flag	Address	Control	Protocol
01111110	11111111 (Broadcast)	00000011	

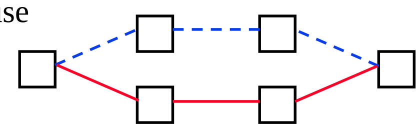
Info	Padding	CRC	Flag
			01111110

- Flag = 0111 1110 = 7E
- Byte Stuffing: 7E ⇒ 7D 5E
  7D ⇒ 7D 5D
- *Byte stuffing method indicated in the textbook is incorrect.*
- Address=FF ⇒ All stations. Control=03 ⇒ Unnumbered
- 16-bit CRC default. 32-bit CRC can be negotiated using LCP

Multiprotocol Label Switching (MPLS)



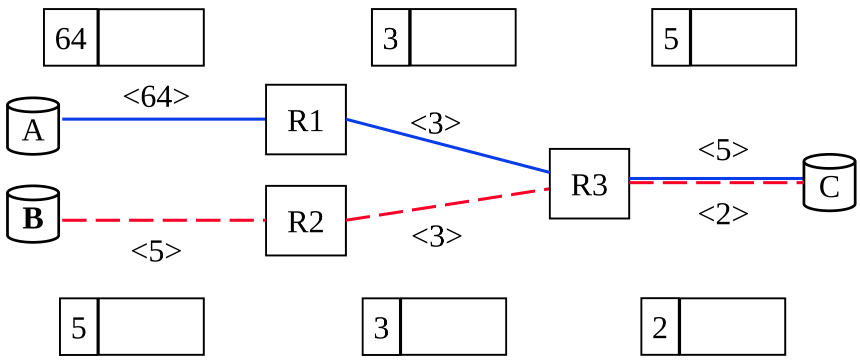
- Allows virtual circuits in IP Networks (May 1996)
- Each packet has a virtual circuit number called 'label'
- Label determines the packet's queuing and forwarding
- Circuits are called Label Switched Paths (LSPs)
- LSP's have to be set up before use
- Allows traffic engineering



Label Switching Example

Ethernet Header	IP Header	Payload

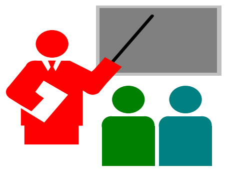
Ethernet Header	Label	IP Header	Payload



LLC, VLANs, PPP, MPLS:Review

1. Ethernet and IEEE 802.3 differ in the protocol type/length field. Length > 1518 ⇒ Protocol Type
2. IEEE 802.3 uses protocol type in the LLC header. 40-bit protocol type is specified using a SNAP SAP
3. Full-Duplex mode allows continuous data transmission on a point-to-point Ethernet. Most of the new equipment uses full-duplex mode. CSMA/CD is not used.
4. Virtual LANs allow hosts to be moved to different broadcast domains (subnets).
5. Point-to-Point protocol (PPP) is used for link and network layer configuration and framing
6. Multiprotocol Label Switching (MPLS) allows label-switched paths (LSPs) in IP networks.

Summary



1. CRC uses binary division using polynomial representation for binary numbers
2. IEEE 802.3 uses a *truncated binary exponential backoff*.
3. Address Resolution Protocol (ARP) is used to find the MAC address for a given IP address and vice versa.
4. PPP is used for configuration and framing on point-to-point links
5. MPLS allows virtual circuits (LSPs) on IP networks.

Review Exercises

- Do not submit.
- Try the following textbook problems: P36

Homework 5C

- Submit answer to Problem P35:
- Consider the MPLS network shown in Figure 5.36 and the labels described on page 503. Suppose that routers R5 and R6 are now MPLS enabled. Suppose that we want to perform traffic engineering so that packets from R6 destined for A are switched to A via R6-R4-R3-R1 and packets from R5 destined for A are switched via R5-R4-R2-R1. Show the MPLS tables in R5 and R6 as well s the modified table in R4 that would make this possible.