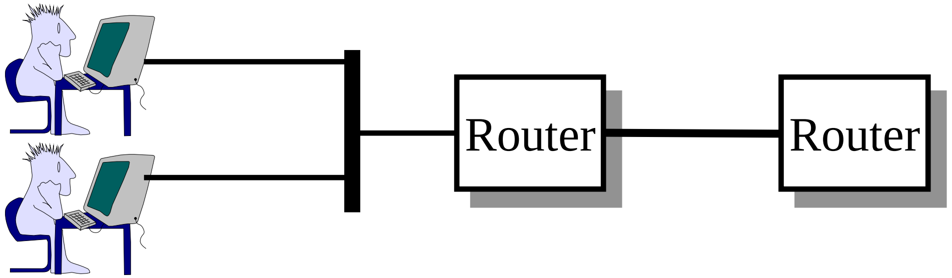
The Link Layer and LANs



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/cse473-22/

Student Questions

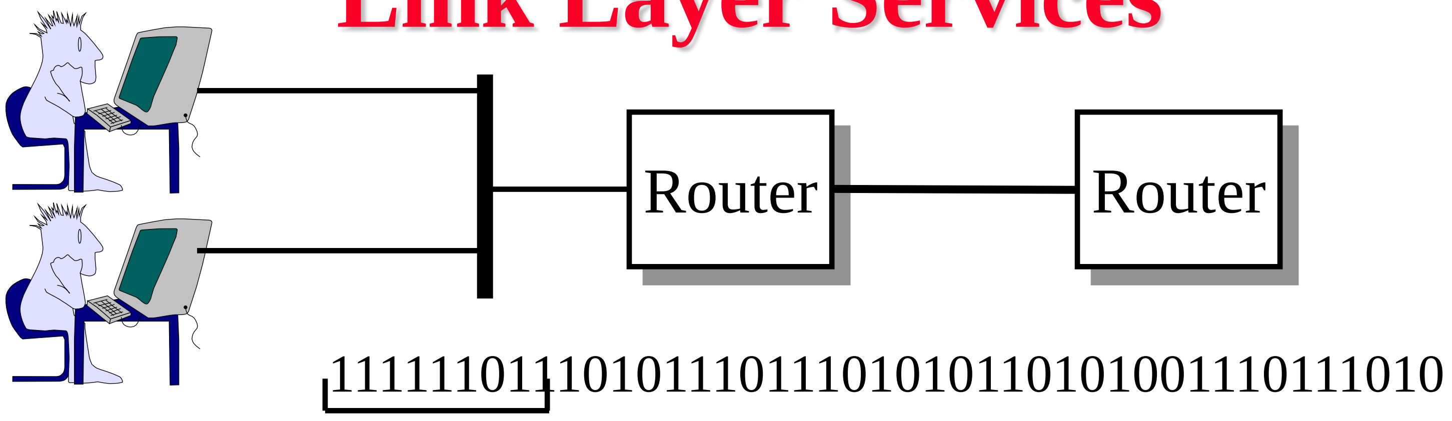
Note: Questions discussed during regular Q&A sessions are marked with ❑. New questions asked during Exam 2 review are marked with ❖

1. Datalink Services
2. Error Detection
3. Multiple Access
4. Bridging
5. MPLS

Note: This class lecture is based on Chapter 6 of the textbook (Kurose and Ross) and the figures provided by the authors.

Student Questions

Link Layer Services



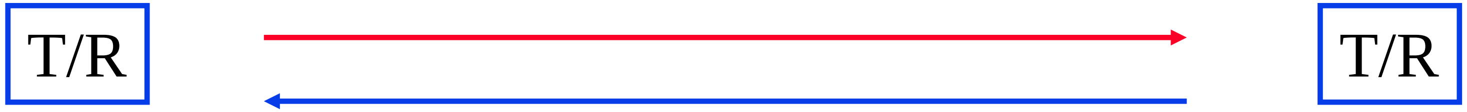
- Link = One hop
- Framing: Bit patterns at begin/end of a frame
- Multiple Access: Multiple users sharing a wire
- Optional (On Lossy wireless links)
  - Flow Control
  - Error Detection/Correction
  - Reliable Delivery
- Duplex Operation

Student Questions

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

Ref: Section 6.1, Review question R1

Student Questions

Error Detection

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

Student Questions

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 → 0 0 1 1 1 0 1 0 0
1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9
1-bit error

0 0 0 1 0 0 1 0 0 0 0 0 1 1 0 1 0 0
1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9
3-bit error 2-bit error

- **Even Parity**

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

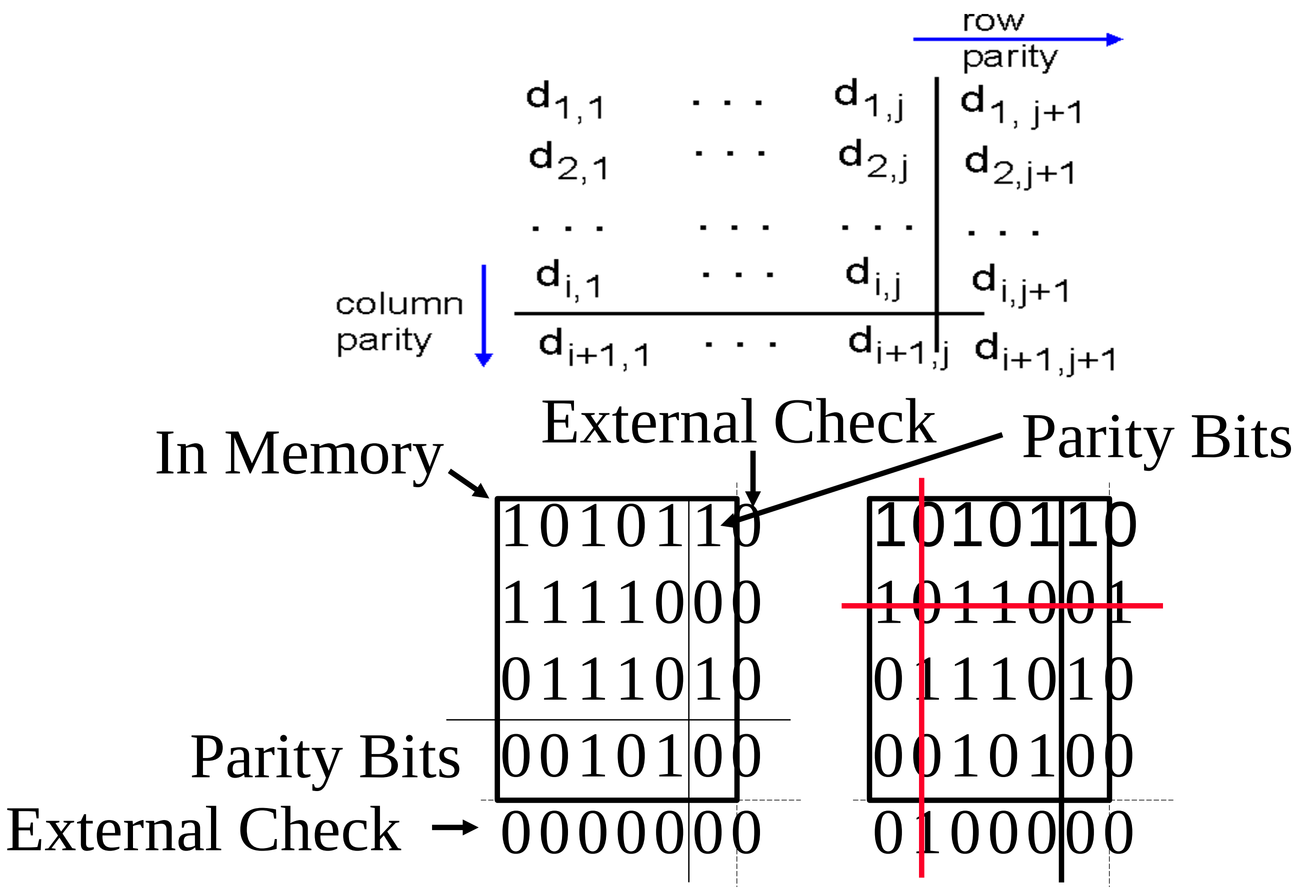
Student Questions

- Is there an advantage to using odd parity rather than even parity checks?

No.

Two Dimensional Parity

- Detect and correct single bit errors



Student Questions

- Why are there no errors in the chart on the left and an error on the right? Does it have to do with even/odd number of 1s?
- *Yes. See revised picture.*

Check Digit Method

- Make a 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, ...

Does not detect transpositions: 2835

Credit card numbers are protected via a similar method called "Luhn Algorithm" which detects most transpositions.

Ref: http://en.wikipedia.org/wiki/Luhn_algorithm

Student Questions

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
```


Student Questions

- Can you redo the example? This slide was difficult for me to follow.

Sure.

- Is the modulo-2 basically a bit-wise XOR? There doesn't seem to be a carry bit between different columns.

Yes. See the illustration on the right.

- For Modulo 2 Division, only if the leftmost bit of divisor is 1 then I can get 1 for quotient, otherwise the quotient should be 0, right?

Yes.

Cyclic Redundancy Check (CRC)

- **Binary Check Digit Method**
- Make a 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.

Student Questions

- Do we perform step 4?

No need to perform step 3. The cross should be on step 3 (not step 4)

Modulo 2 Division

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

Student Questions

- Could you go over modulo 2 division with a laser pointer?

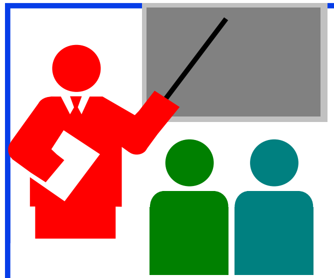
Sure.

Checking At The Receiver

```
              1101010110
110101)101000110101110
       110101
        111011
        110101
         011101
         000000
          111010
          110101
           011111
           000000
            111110
            110101

010111
000000
 101111
 110101
  110101
  110101
   00000
```

Student Questions



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 mod 2 division

Ref: Section 6.2, Review questions R2, R3

Student Questions

Homework 6A: CRC

- [4 points] Find the CRC of 1001100 using a generator *1011*. *Use mod 2 division. Show all steps including the checking at the receiver.*

Student Questions

Multiple Access Links and Protocols

1. Multiple Access
2. CSMA/CD
3. IEEE 802.3 CSMA/CD
4. CSMA/CD Performance
5. Cable Modem Access

Student Questions

Multiple Access

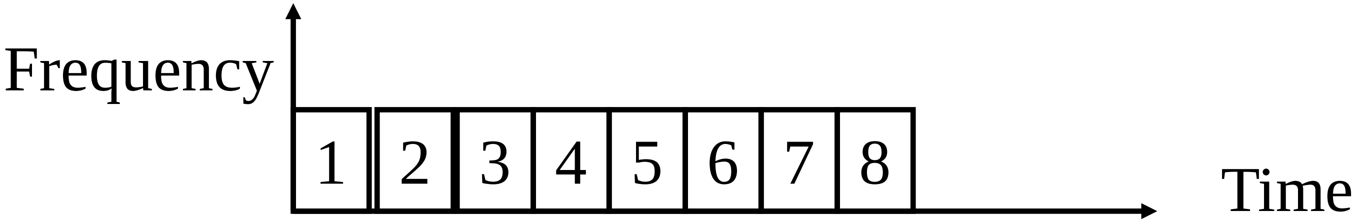
(a) Aloha Multiple Access

(b) Carrier-Sense Multiple Access with Collision Detection

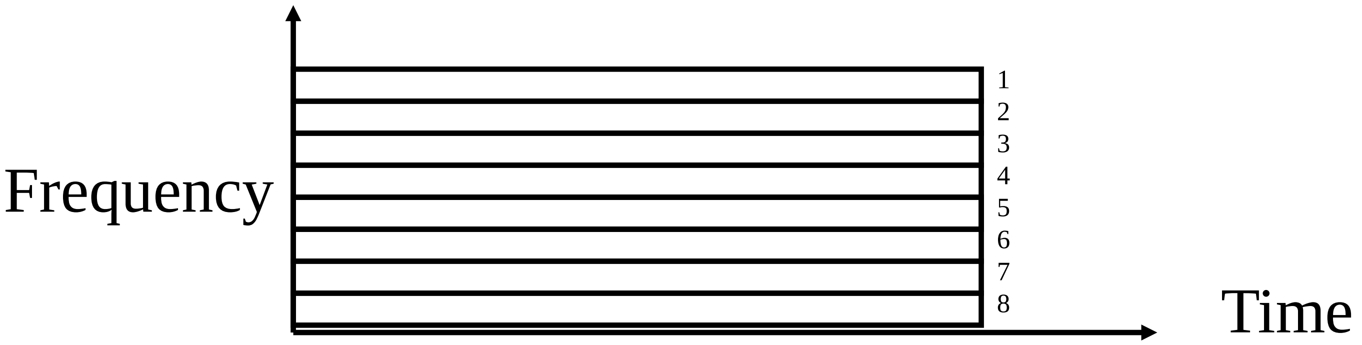
Student Questions

Multiple Access

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



- **Frequency Division Multiplexing**



Student Questions

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

Student Questions

- Can you clarify the situation that would lead to 18% and 37% and how those are a "worst case"?

There are published papers with details. Basically assumes random (Poisson arrivals).

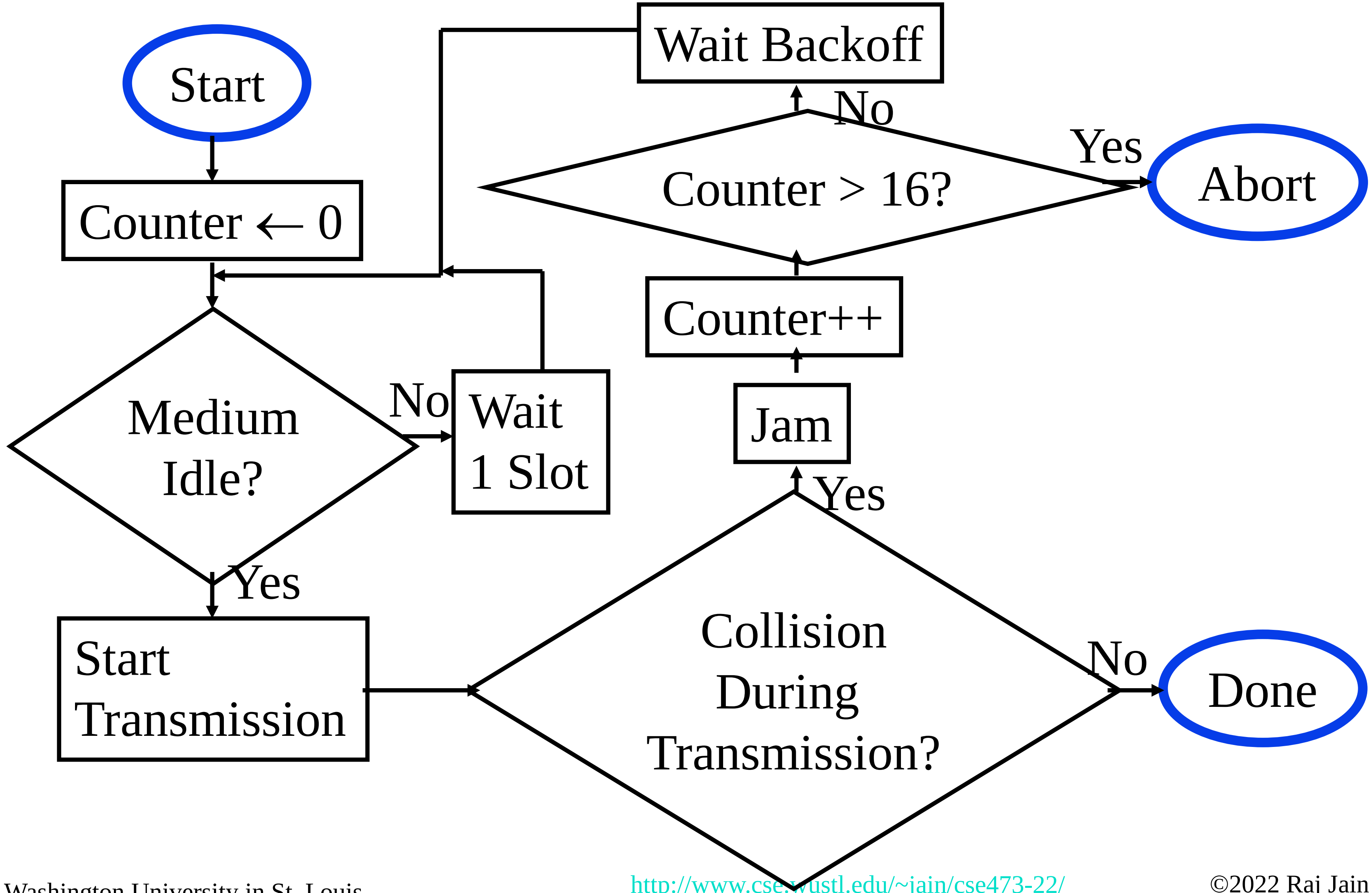
IEEE 802.3 CSMA/CD

- 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,$2^{\min(k,10)}$-1] slots **Truncated Binary Backoff**
- Collision detected by monitoring the voltage
  High voltage ⇒ two or more transmitters ⇒ Collision
  ⇒ Length of the cable is limited to 2.5 km

Student Questions

- why is k increasing to 16 instead of 10 if we are looking for min(k,10)?

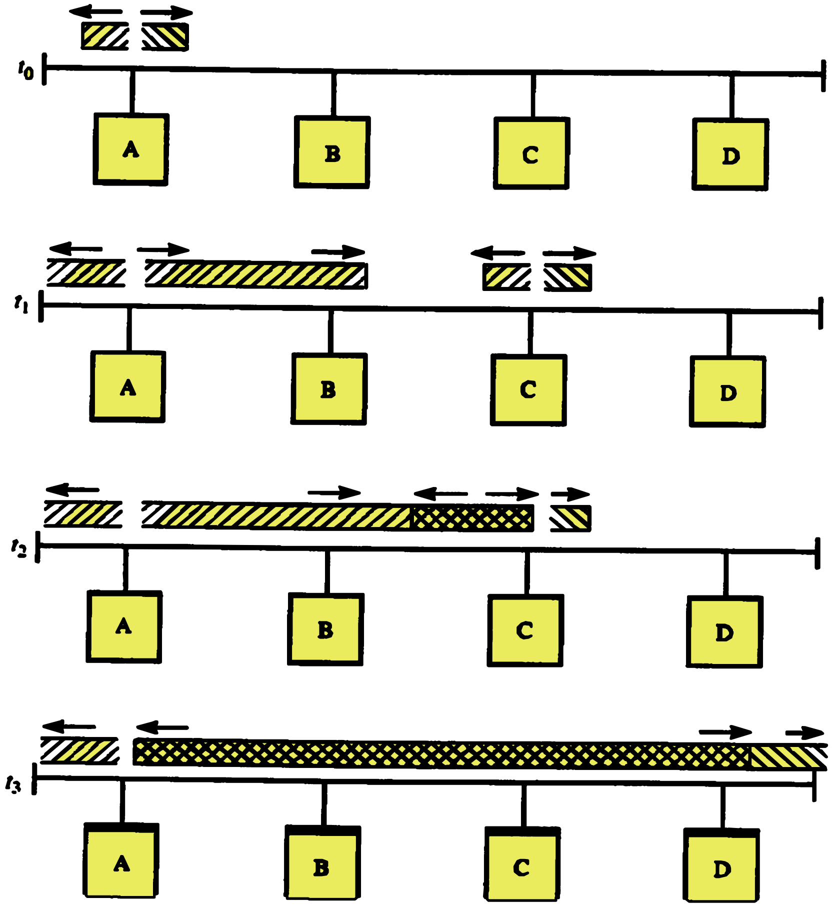
IEEE 802.3 CSMA/CD Flow Chart



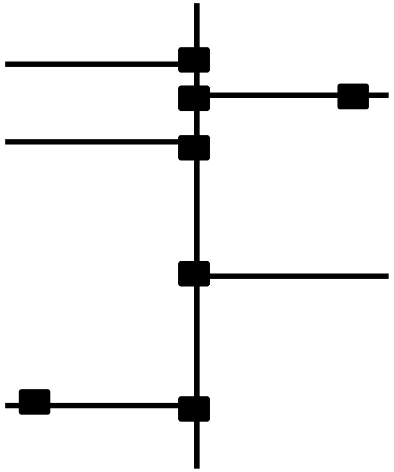
Student Questions

CSMA/CD Operation

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



One way delay
= 25.6 µs
Max Distance
<2.5 km



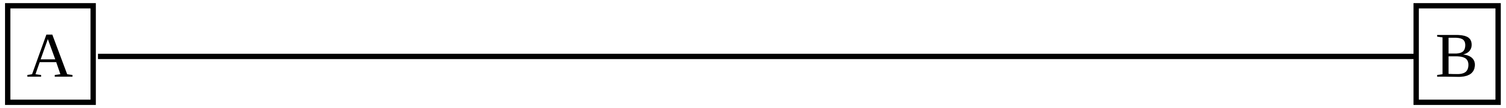
Student Questions

- If electricity in copper travels at 2.5E8m/s and the link is 2500m long, then shouldn't the one-way trip only be 10 microseconds?

Allows for some tolerance in interconnections and hubs.

Homework 6B: Collision Detection

- [6 Points] 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).

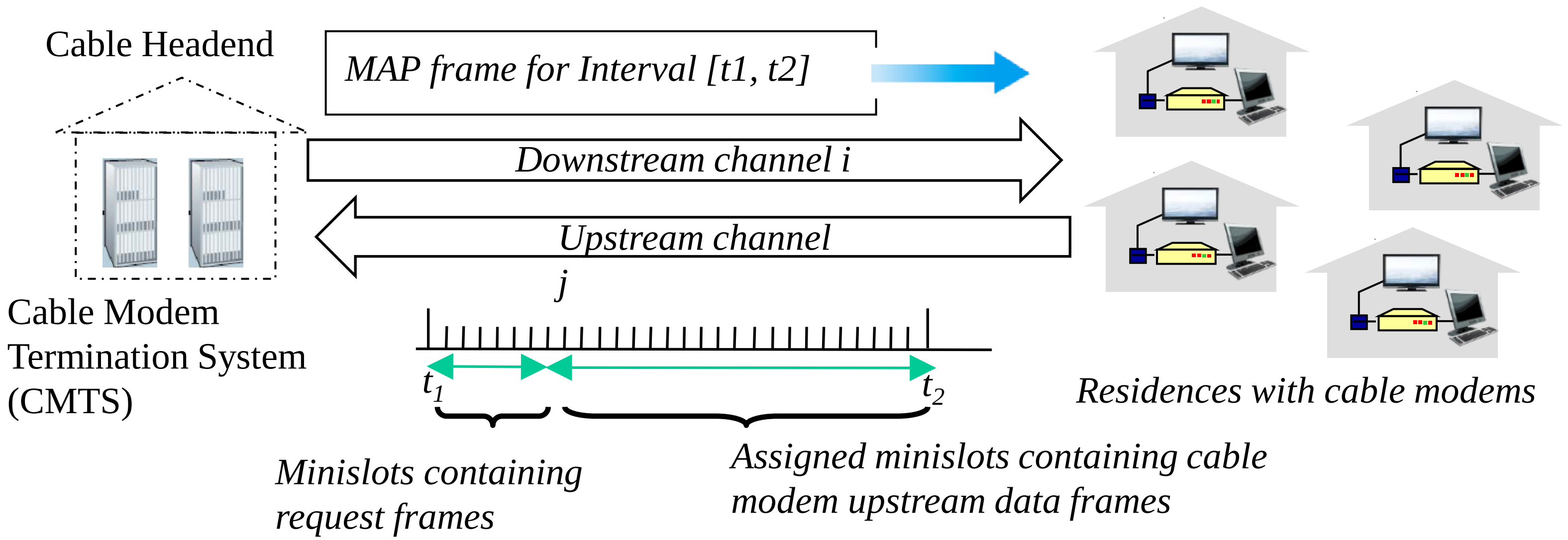


Student Questions

- Can you explain what a 'bit time' is on HW 6B? And is the slide supposed to say '512b=64B'?

With 10 Mbps, each bit is 10^{-7} s or 0.1 microseconds long. The frame size that A may transmit may be as much as 576 bits (64 bits of preamble).

Cable Access Network

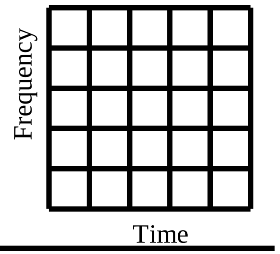


- **DOCSIS**: Data Over Cable Service Interface Specification
- Frequency Division Multiplexed (**FDM**) channels over upstream and downstream
- Time Division Multiplexed (**TDM**) slots in each upstream channel:
  - Some slots assigned, some have contention
  - Downstream **MAP** frame: Assigns upstream slots
  - Request for upstream slots (and data) transmitted random access (binary backoff) in selected slots

Student Questions

- So FDM is not upstream at all? And the MAP frame arrives over downstream FDM?

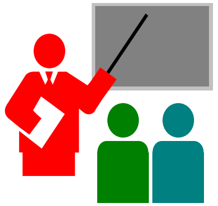
FDM is both ways.



- Can you explain the upstream and downstream of the cable access network again?

Sure.

Multiple Access Links and Protocols: Review



1. Multiple users can share using **TDMA** or **FDMA**
2. Random access is better for data traffic.
3. Aloha has an efficiency of **1/2e**. Slotted Aloha makes it **1/e**.
4. Carrier sense and collision detection improves the efficiency further.
5. IEEE 802.3 uses **CSMA/CD** with **truncated** binary exponential backoff
6. DOCSIS used in cable access networks has **frequency division** multiplexed channels. With each channel **time division** multiplexed with some slots reserved for random access.

Ref: Section 6.3, Review question R4-R8

Student Questions

End of Part 1

Switched Local Area Networks

1. Ethernet Standards
2. IEEE 802 Address Format
3. Address Resolution Protocol
4. Bridging
5. Virtual LANs

Student Questions

Ethernet Frame Structure

preamble	dest. address	source address	type	data (payload)	CRC
64b	48b	48b	16b	Variable length	32b

- ***Preamble:*** 7 bytes with pattern 10101010 followed by one byte with pattern 10101011. To synchronize receiver, sender clocks
- ***Addresses:*** 6 byte source, destination MAC addresses
- ***Type:*** indicates higher layer protocol
  - IP : 0x0800
  - ARP: 0x0806
- ***CRC:*** Cyclic Redundancy Check
  - If error detected: frame is silently dropped at the receiver
- ***Connectionless***: No need to ask the receiver
- ***Unreliable***: No ack, nack, or retransmissions

Student Questions

- What happens when two frames arrive closer together than the minimum gap?

Some receivers may not have enough time to turn around and may miss the frame.

- For error detection, is only the CRC field used in computation or the whole frame used?

The whole frame is used both while sending and receiving. While sending, the CRC field is initially filled with zeros and then replaced with the computed CRC. While receiving the computed CRC should come out to zero.

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
- **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
- **1000BASE-T**: 1 Gbps (Gigabit Ethernet)
- **10GBASE-T**: 10 Gbps
- **40GBASE-T**: 40 Gbps

Student Questions

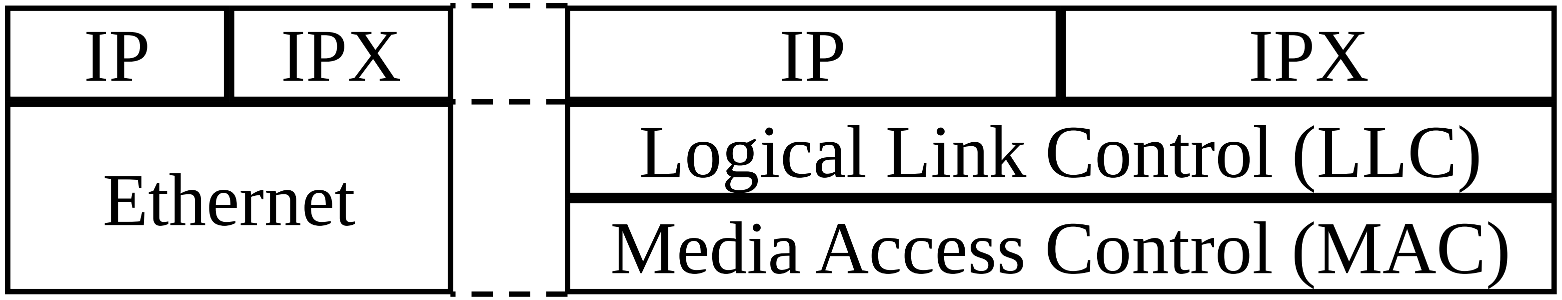
- Which Ethernet is more common today?

1000BASE-T full-duplex.

- Could you explain exactly what the base part is, i.e., base 5?

Base = Baseband ⇒ No Frequency Multiplexing
Broad = Broadband => Frequency multiplexing
The number after Base, if present, indicates the maximum distance. The letters and numbers after the dash indicate the media type.

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.

Student Questions

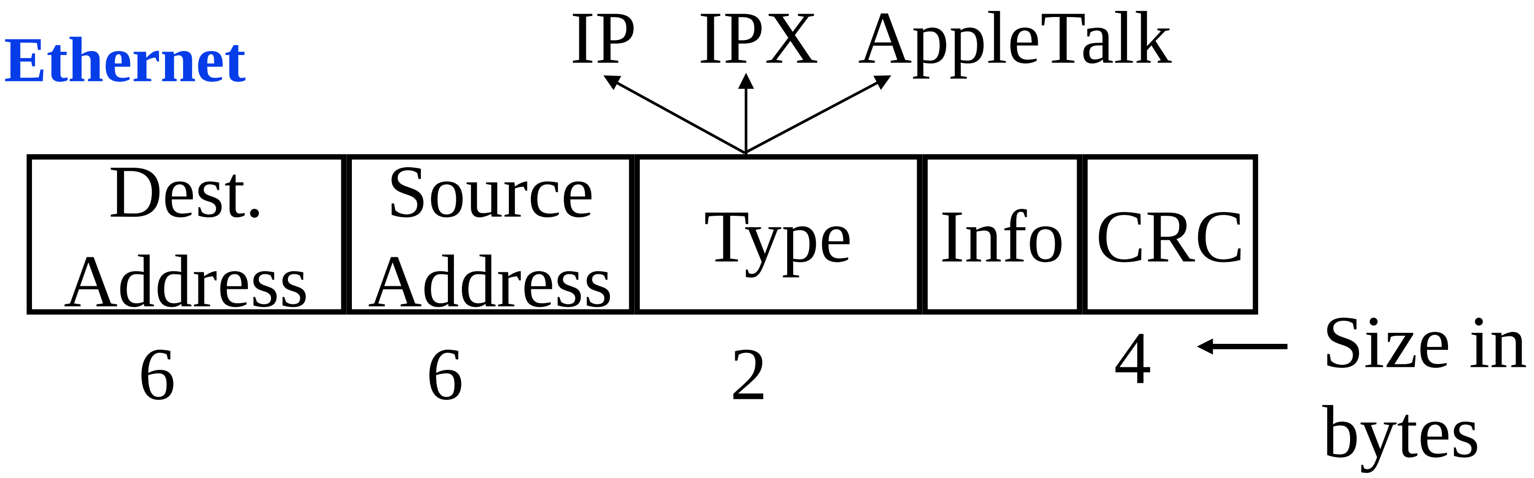
- What is a sublayer? Are these two different forms of an Ethernet frame?

Sublayers are layers inside a layer. There are not optional. All 802.3 frames have MAC and LLC.

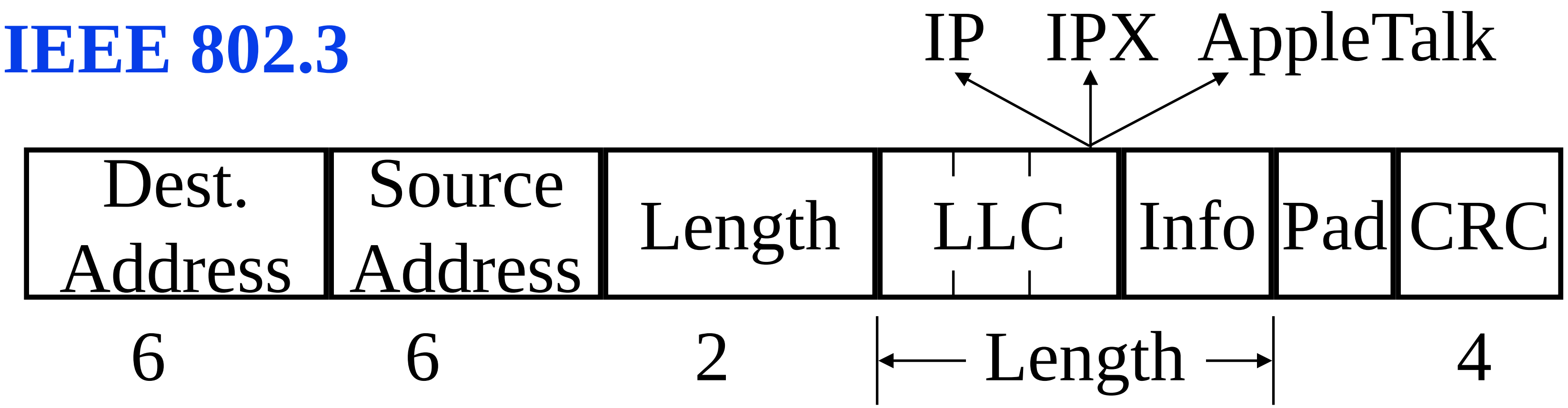
Original Ethernet frames do not have LLC.

Ethernet and 802.3 Frame Formats

- Ethernet



- IEEE 802.3



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

Student Questions

- Can you re-explain the significance of the value 1518

1518 byte is the maximum length of Ethernet frames. 18 bytes of header plus 1500 byte payload.

- Does layer-2 protocol correspond to the movement of frames in wires, and whenever we send a frame into the wire, we need to use a layer-2 protocol?

Yes, Yes. Layer 2 is the "Media Access Control" layer. The wire is the media.

IEEE 802 Address Format

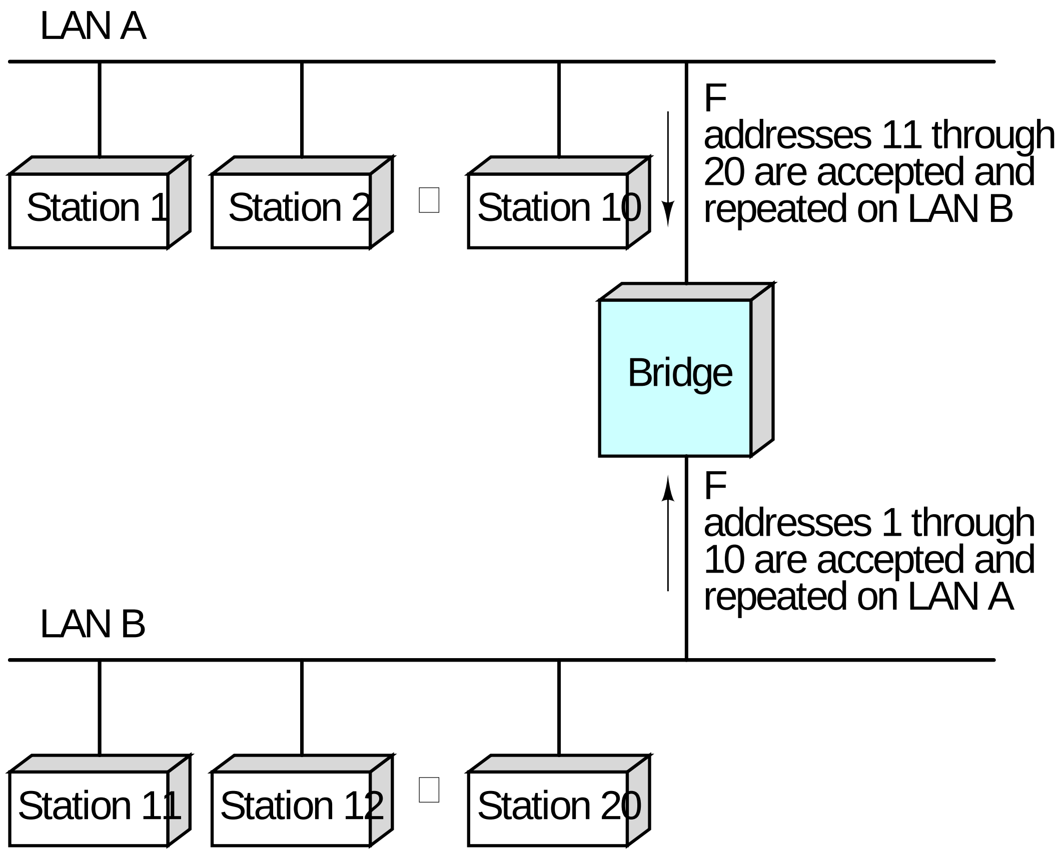
- 48-bit:1000 0000 : 0000 0001 : 0100 0011 : 0000 0000 : 1000 0000 : 0000 1100 = 80:01:43:00:80:0C

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

Student Questions

Bridges



Student Questions

- Should we consider a bridge as being a component internal to one individual LAN at a time?

Sure. As long as you define internal as transparent.

- I have rewatched the video; I still don't know why these two LANS connected with the bridge, which conflicts with slide 41?

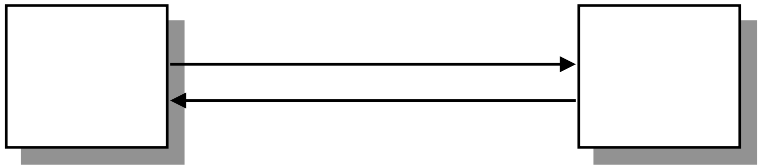
Yes. Thanks for noticing the conflict. In this slide, the combination of LAN A, Bridge, and LAN B results in one "Extended LAN." The word "extended" is often omitted in practice. This picture shows one extended LAN, while Slide 41 shows how to connect two extended LANs using a router. In both figures, the word "extended" was omitted and is inferred from the context.

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
  ⇒ **Transparent**
  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 (in Token Rings)

Student Questions

Full-Duplex Ethernet



- Uses point-to-point links between TWO nodes
- Full-duplex bi-directional transmission ⇒ Transmit any time
- Standardized in IEEE 802.3-2018
- All 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

Student Questions

- What are the roles of an NIC?

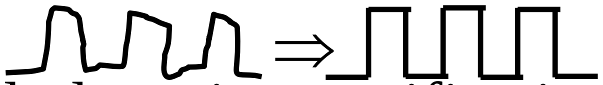
Network Interface Card is either Ethernet/WiFi/BlueTooth interface. Some (e.g., Ethernet are visible. Others are invisible)

- Do IoT devices utilize NICs? if so, are the hardware implementations different than a traditional computer?

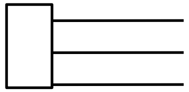
Yes, all networked devices need NICs. They are generally lower speed and less optional features. Most CPU chips nowadays have built in NICs.

Interconnection Devices

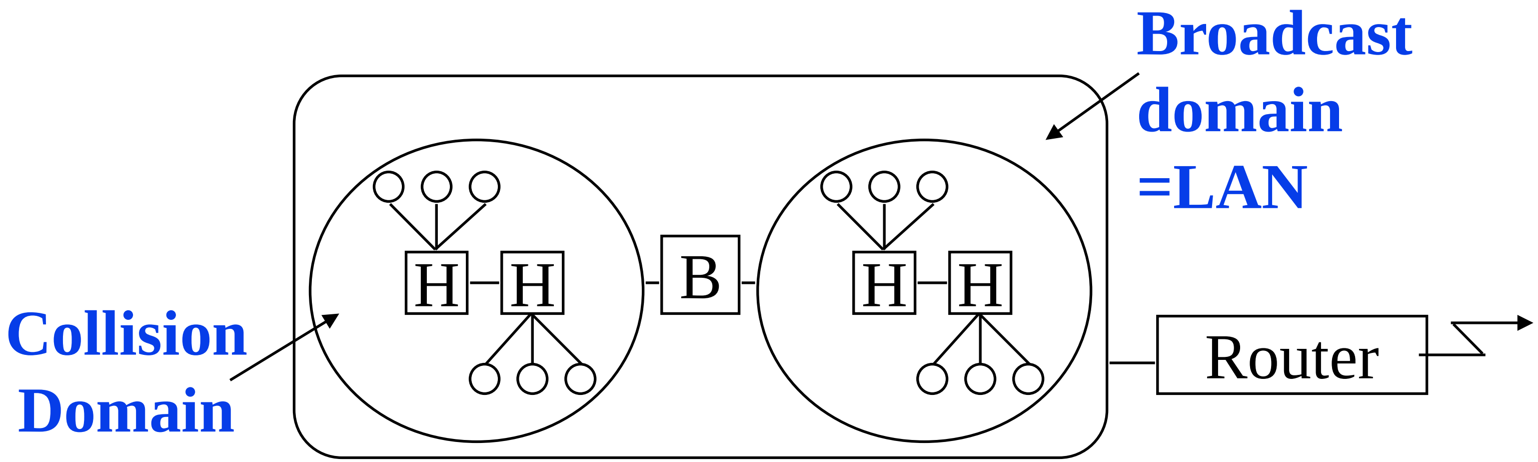
- **Repeater**: PHY device that restores data and collision signals
  Repeater = Digital Amplifier



- **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)



Student Questions

- Could you explain what a hub and repeater do again?

Hub: Multiport Repeater + Fault detection. Every incoming bit is broadcast to all ports.

- Can you again explain what a collision domain means, and in what case is a bridge needed?

Two nodes in one collision domain cannot speak simultaneously. Two nodes in different collision domains can speak simultaneously.

- If there are multiple bridges connecting more than 2 Ethernets in a system, how can a bridge know to forward it to the next bridge to reach further away Ethernets?

When there are multiple bridges, a "Spanning Tree" protocol is used to select one bridge. Other bridges are put into standby mode.

Address Resolution Protocol

- Problem: Given an IP address find the MAC address
- Solution: Address Resolution Protocol (**ARP**)
- The host broadcasts a request (Dest MAC=FFFFFFFF): "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 $8A:5F:3C:23:45:56_{16}$"
- Nodes cache the MAC-IP mapping in a "ARP table" You can list ARP table using "arp –a" command
- Frame Format: Hardware (HW): 0x0001 = Ethernet,
  - Protocol (Prot): 0x0800 = IP,
  - Operation: 1 = Request, 2=Response

HW Type	Prot Type	HW Addr Length	Prot Addr Length	Oper-ation	Sender HW Addr	Sender Prot Addr	Target HW Addr	Target Prot Addr
16b	16b	8b	8b	16b	48b	32b	48b	32b

Student Questions

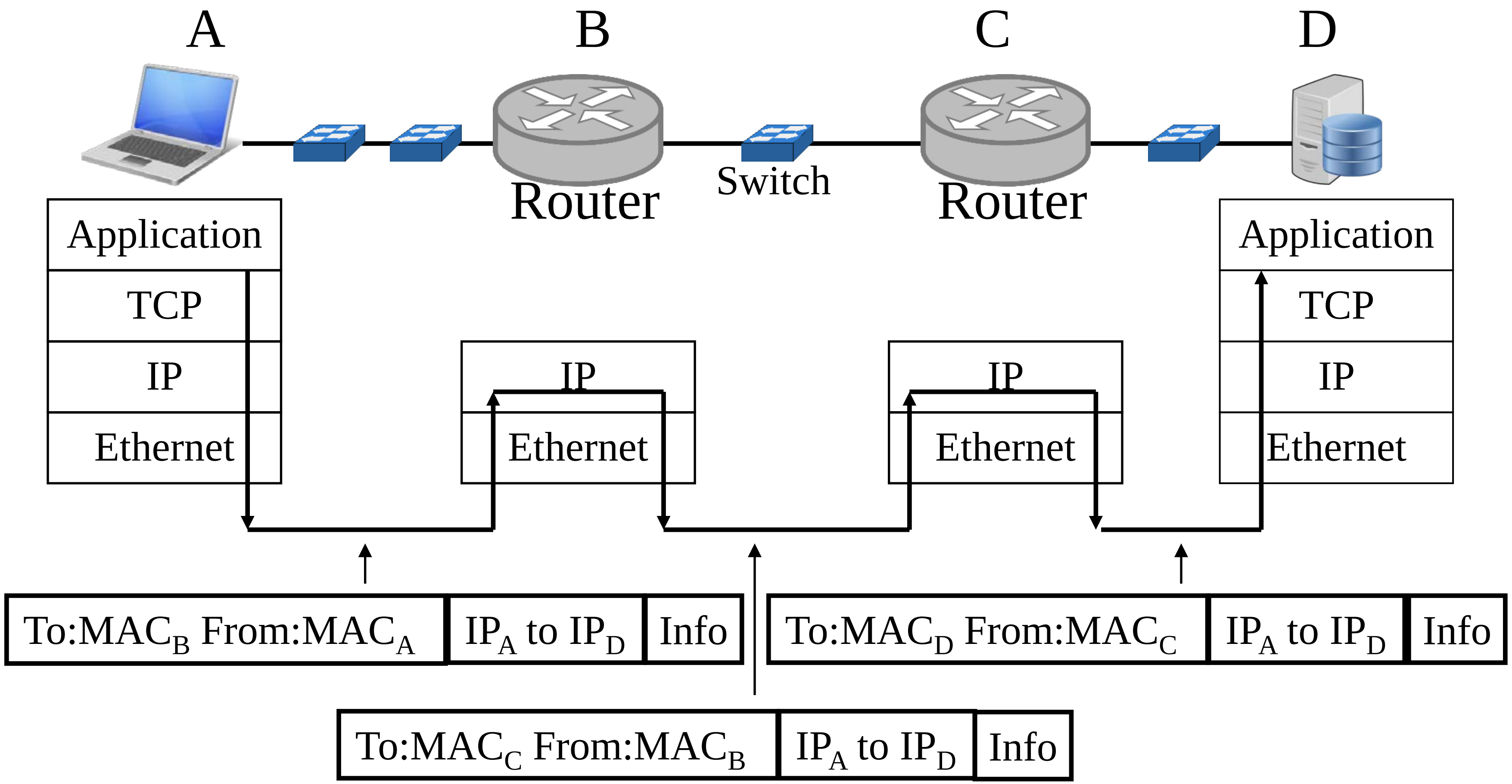
- In the ARP header, what are the protocol addresses?

ARP can be used by IP, IPX, and other L3 protocols. Protocol type and protocol address are L3 type and L3 address.

- Why is the ARP query sent within the broadcast frame?

Since we don't know who can answer.

IP over Multiple Hops



- Switches = Transparent Bridges ⇒ No changes to frames
- ARP required only for nodes on the same “subnet”

Student Questions

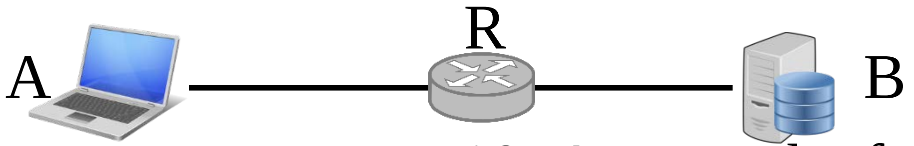
- Can you walk through the ARP messages when a host sends another host in a different subnet with multiple hops?

If A wants to send an IP datagram to C, it uses the routing table and finds that the shortest path is through B.
A knows that B is on the same subnet by looking at the subnet part of B’s IP address. Subnet=LAN. The IP layer in A gives the datagram to the MAC layer. MAC layer needs the MAC address of B to send it using the LAN. It looks up that address in its ARP table. If it is not found, it broadcasts an ARP request. After it receives a response, it stores it in its ARP table and sends the MAC frame to B. ARP is used only for nodes on the same subnet. Subnet=LAN.

Lab 6: Ethernet and ARP

[32 points] Download the Wireshark traces from http://gaia.cs.umass.edu/wireshark-labs/wireshark-traces.zip

Open *ethernet--ethereal-trace-1* in Wireshark. Select **View → Expand All**. This trace shows a HTTP exchange between end host A and Server B via Router R as shown below:



1. Examine HTTP request Frame 10. Answer the following questions.
   A. What is the 48-bit Ethernet source address? Who does it belong to: A, B, or R?
   B. What is the 48-bit Ethernet destination address? Who does it belong to: A, B, or R?
   C. What is the hexadecimal value for the two-byte Frame type field. What upper layer protocol does this correspond to?

Student Questions

Lab 6 (Cont)

D. How many bytes from the very start of the Ethernet frame does the ASCII "G" in "GET" appear in the Ethernet frame? How many bytes are used up in Ethernet header, IP header, and TCP header before this first byte of HTTP message.

2. Examine the HTTP OK response. (Frame **12 … 16**).
   A. What is the Ethernet source address? Who does it belong to: A, B, or R?
   B. What is the destination address in the Ethernet frame? Who does it belong to: A, B, or R?
   C. What is the hexadecimal value for the two-byte Frame type field. What upper layer protocol does this correspond to?
   D. How many bytes from the very start of the Ethernet frame does the ASCII "O" in "OK" appear in the Ethernet frame? How many bytes are used up in Ethernet header, IP header, and TCP header before the first byte of HTTP message.

Student Questions

Lab 6 (Cont)

3. Examine Frame 1. This is an ARP request.
   A. What are the hexadecimal values for the source and destination addresses in the Ethernet frame containing the ARP request message?
   B. What is the hexadecimal value for the two-byte Frame type field. What upper layer protocol does this correspond to?
   C. How many bytes from the very beginning of the Ethernet frame does the ARP opcode field begin?
   D. What is the value of the opcode field within the ARP-payload?
   E. What is the IP address of the sender?
   F. What is the target MAC and IP addresses in the ARP "question"?

Student Questions

Lab 6 (Cont)

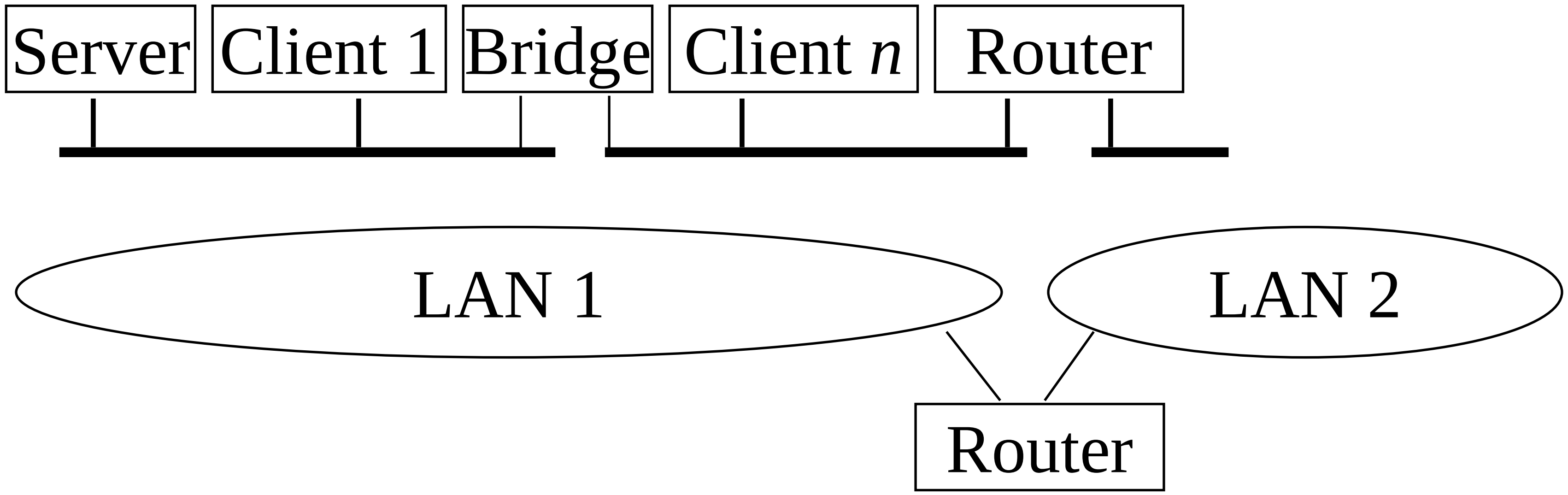
4. Examine Frame 2. This is the ARP response.

A. What are the hexadecimal values for the source and destination addresses in the Ethernet frame containing the ARP response message?
B. What is the hexadecimal value for the two-byte Frame type field. What upper layer protocol does this correspond to?
C. How many bytes from the very beginning of the Ethernet frame does the ARP opcode field begin?
D. What is the value of the opcode field within the ARP-payload?
E. What is the IP address of the sender?
F. What is the target MAC and IP addresses in the ARP "answer"?

For all questions of this lab, please provide numerical answers only. No need to add screen captures.

Student Questions

What is a LAN?



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

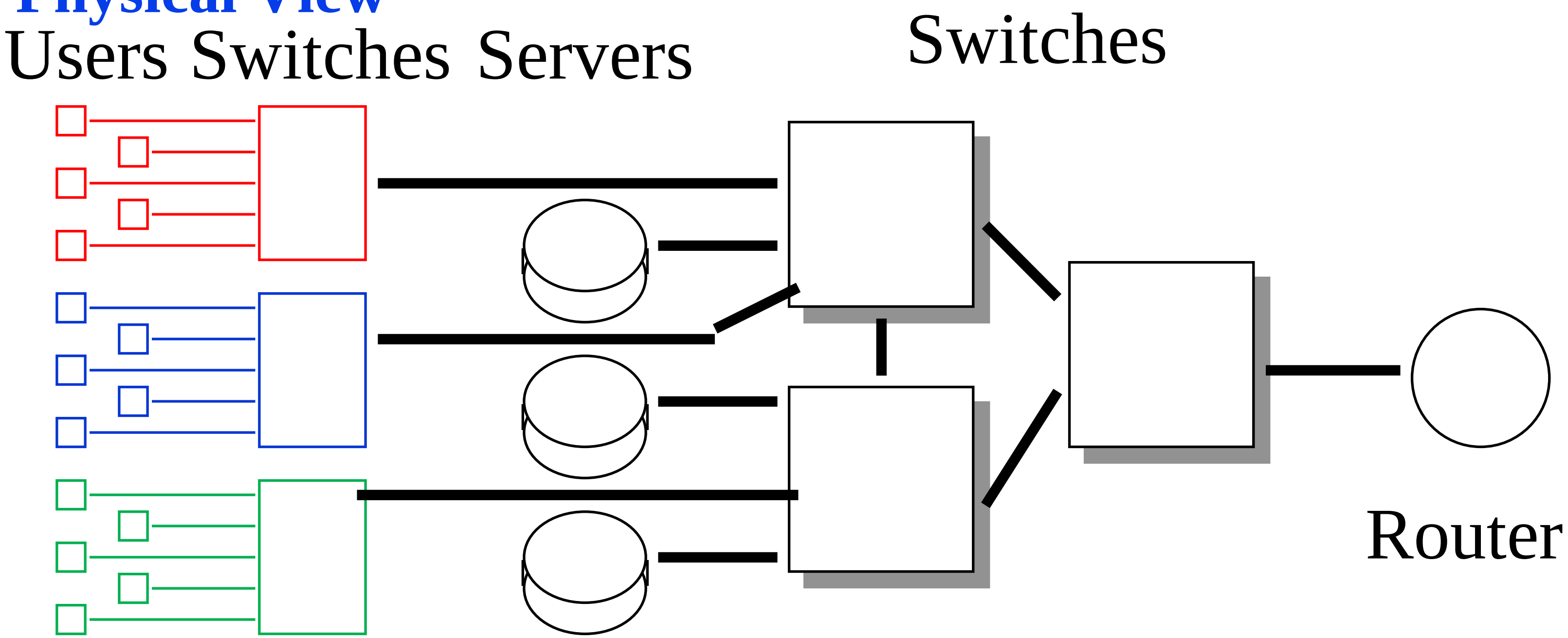
Student Questions

- Can you talk about WAN and its relation to LAN

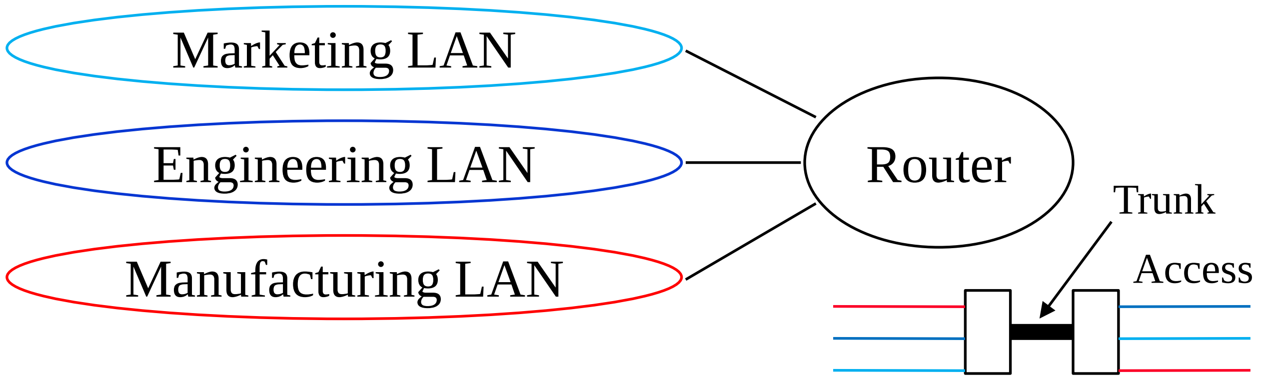
LAN is one trusted environment.
WAN consists of multiple LANs
=> Multiple trusted environments.
You need a router to connect different trust environments. You can use a bridge to connect multiple components of a single trust environment.
Generally, each single-family home = LAN.
You do not need a router to communicate inside a single home; you can use bridges if necessary.
For example, different sections of a large home can be connected via bridges. Other homes are separate LANs. You will need a router to communicate with nodes out of your LAN.

What is a Virtual LAN

- **Physical View**



- **Logical View**



Student Questions

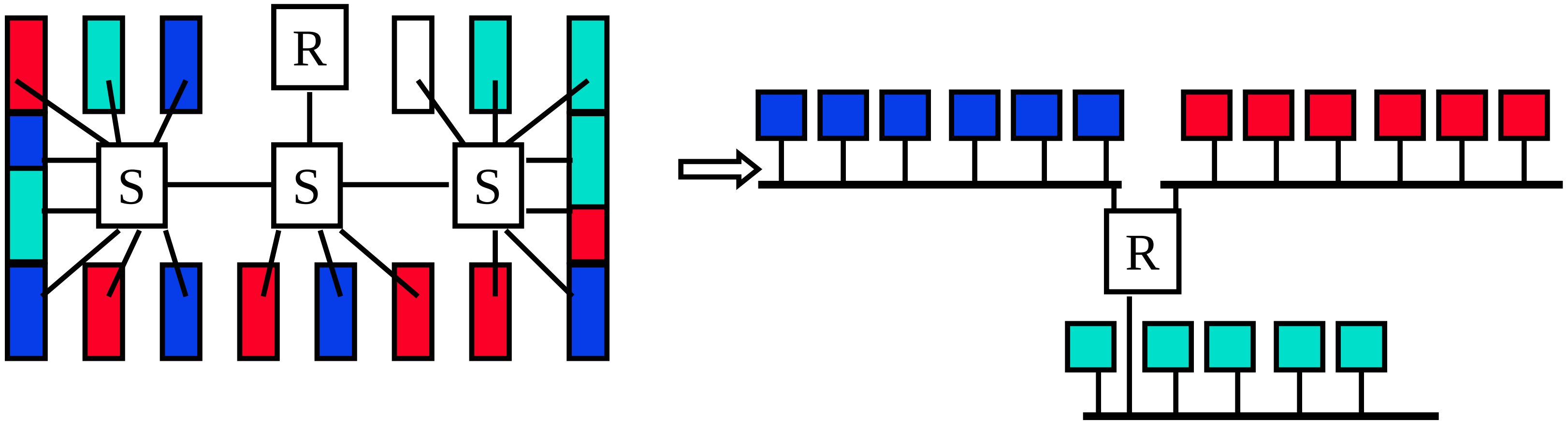
- Given two computers with identical network address *masks* (e.g., 192.168.1.0/24) but different VLANs with an access port link, do we need a router to communicate across the VLANs? If the link is in access mode, I think we can ping across machines.

Each VLAN has a different address mask. Each VLAN is a different subnet.

- While reading more about VLAN, I came across trunk port / access port. Could you discuss how these are used for interVLAN communication?
- Advanced topic for CSE 574S:

Ref: https://www.educba.com/trunk-port-vs-access-port/

Virtual LAN



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

Student Questions

- On the virtual LAN slide you had 3 different departments. Would contact with an onsite mail server become more complicated?

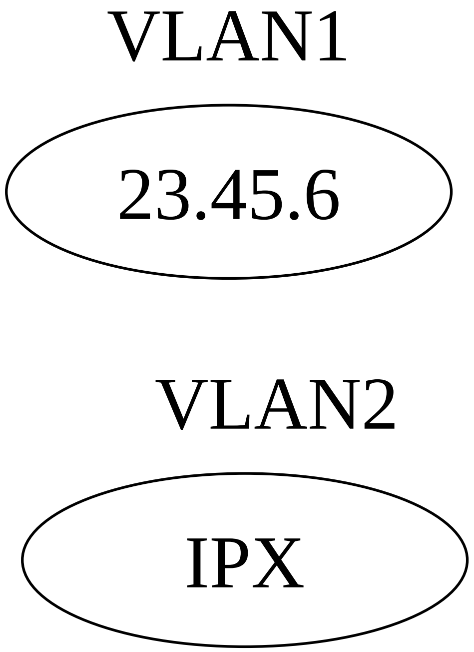
No. IP can go over many LANs (and VLANs.)

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		√
B2	√	

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



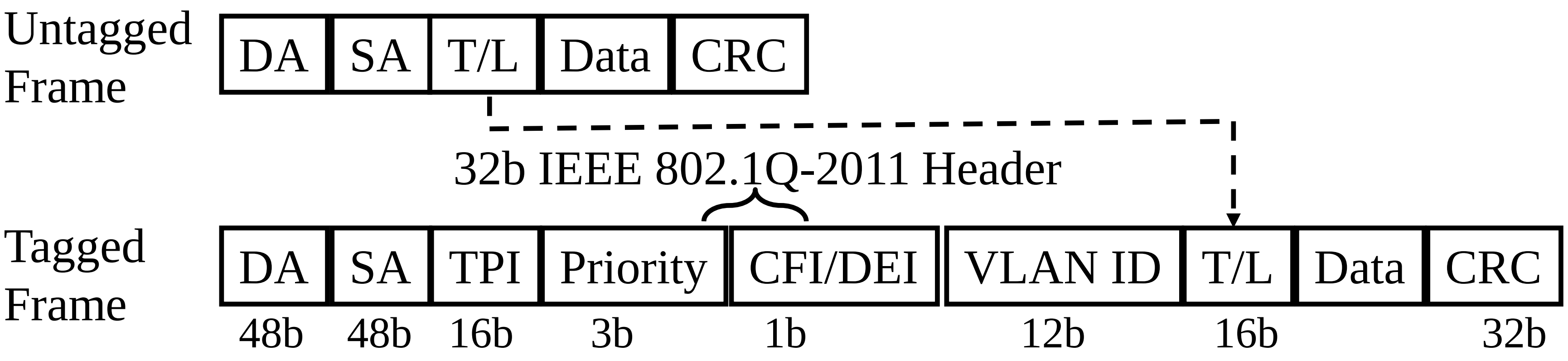
Student Questions

❖ How are the three VLAN layers related?

*These are three **types** of VLANs, not layers of VLANs. You generally use only one type of VLAN.*

IEEE 802.1Q-2011 Tag

- **Tag Protocol Identifier** (**TPI**)
- **Priority Code Point** (**PCP**): 3 bits = 8 priorities 0..7 (High)
- **Canonical Format Indicator** (**CFI**): 0 ⇒ Standard Ethernet,
  1 ⇒ IBM Token Ring format (non-canonical or non-standard)
- CFI now replaced by Drop Eligibility Indicator (DEI)
- VLAN Identifier (12 bits ⇒ 4095 VLANs)
- Switches forward based on MAC address + VLAN ID
  Unknown addresses are flooded.



Ref: Canonical vs. MSB Addresses, http://support.lexmark.com/index?page=content&id=HO1299&locale=en&userlocale=EN_US
Ref: G. Santana, "Data Center Virtualization Fundamentals," Cisco Press, 2014, ISBN:1587143240

Student Questions

- Why are VLAN IDs needed in the link layer if the subnet mask in network layer takes care of that?

Ethernet switches do not speak IP. Subnet mask is an IP/IETF concept. VLAN is an Ethernet/IEEE concept. There are many types of VLANs. L3 VLANs (based on subnet mask) are one of them. There are others.

Switched Local Area Networks : Review

1. IEEE 802.3 uses a *truncated binary exponential backoff*.
2. Ethernet uses 48-bit addresses of which the first bit is the unicast/multicast, 2nd bit is universal/local, 22-bits are OUI (Organizationally unique identifier).
3. Ethernet bridges are transparent and self-learning using source addresses in the frame
4. Bridges are layer 2 devices while routers are layer 3 devices and do not forward layer 2 broadcasts
5. Address Resolution Protocol (ARP) is used to find the MAC address for a given IP address and vice versa.
6. IEEE 802.1Q tag in Ethernet frames allows a LAN to be divided in to multiple VLANs. Broadcasts are limited to each VLAN and you need a router to go from one VLAN to another.

Ref: Section 6.4, Review Questions R9-R16

Student Questions

- Is the OUI 22 bits or 24 bits?

The first two bits are unicast/multicast and Global/Local. So the next 22 bits indicate the "Organization".

- Are there ways that the link layer has been built to incorporate machine learning? If so, Do you foresee machine learning being used in the future in this layer?

Not aware of ML in Link layer.

End of Part 2

Multiprotocol Label Switching

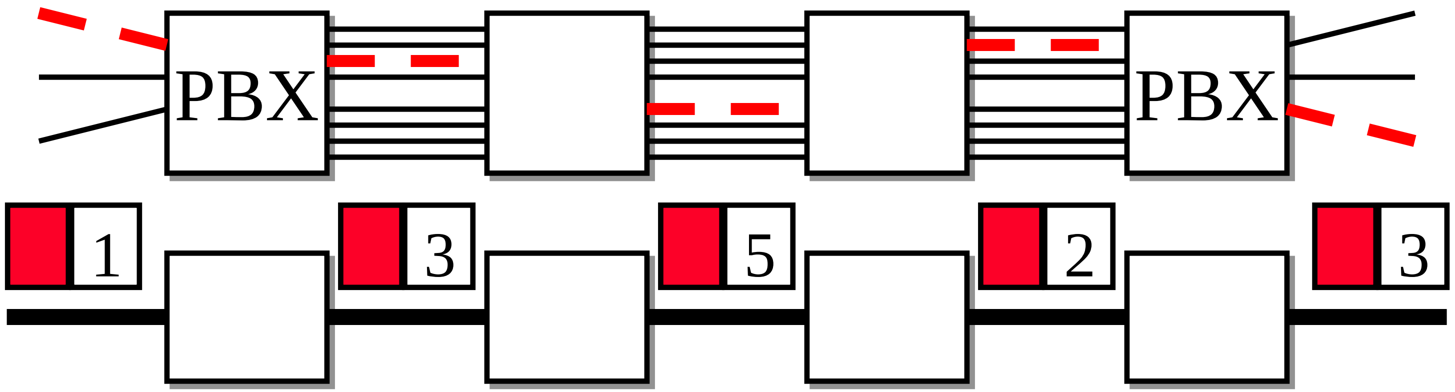
Connection-oriented IP: Paths set up in advance

Borrowed from the Telephone networks

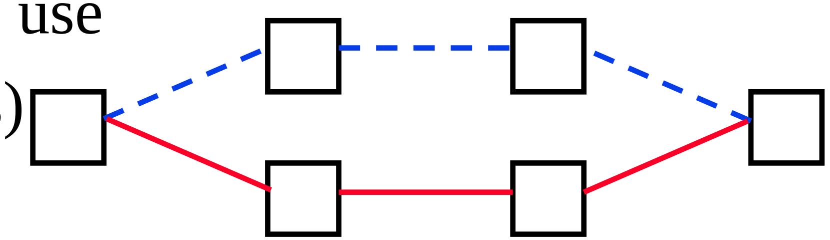
- Multiprotocol Label Switching (MPLS)
- Label Switching Example
- MPLS Forwarding Tables
- MPLS versus IP Paths
- MPLS Label Format

Student Questions

Multiprotocol Label Switching (MPLS)



- Allows virtual circuits in IP Networks (May 1996)
- Each packet has a virtual circuit ID 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
- Label switching routers (LSRs) allows traffic engineering

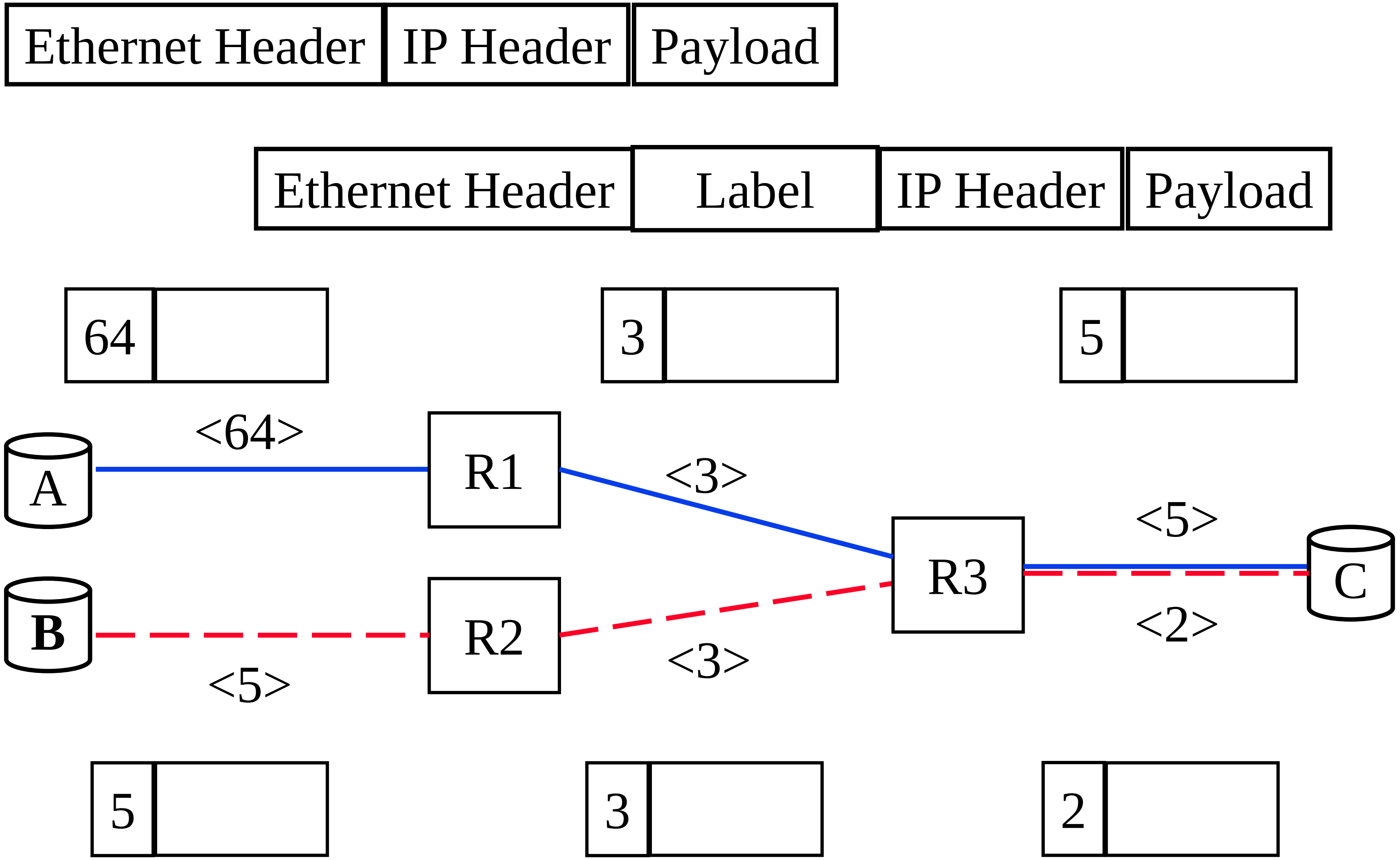


Student Questions

- How are labels and SDN used in conjunction? Like for load-balancing, would an LSR automatically start using some of the backup routes if an ICMP quench message came?

OpenFlow has been extended to include MPLS labels. The tables indicate what to do for each label.

Label Switching Example



Student Questions

- Do the Ethernet Q tag or the MPLS tags use up some of the 1500B used for the L3 payload?

Yes, they do.

MPLS Forwarding Tables

- Interface numbers are in circles. Label IDs are in rectangles.

R6

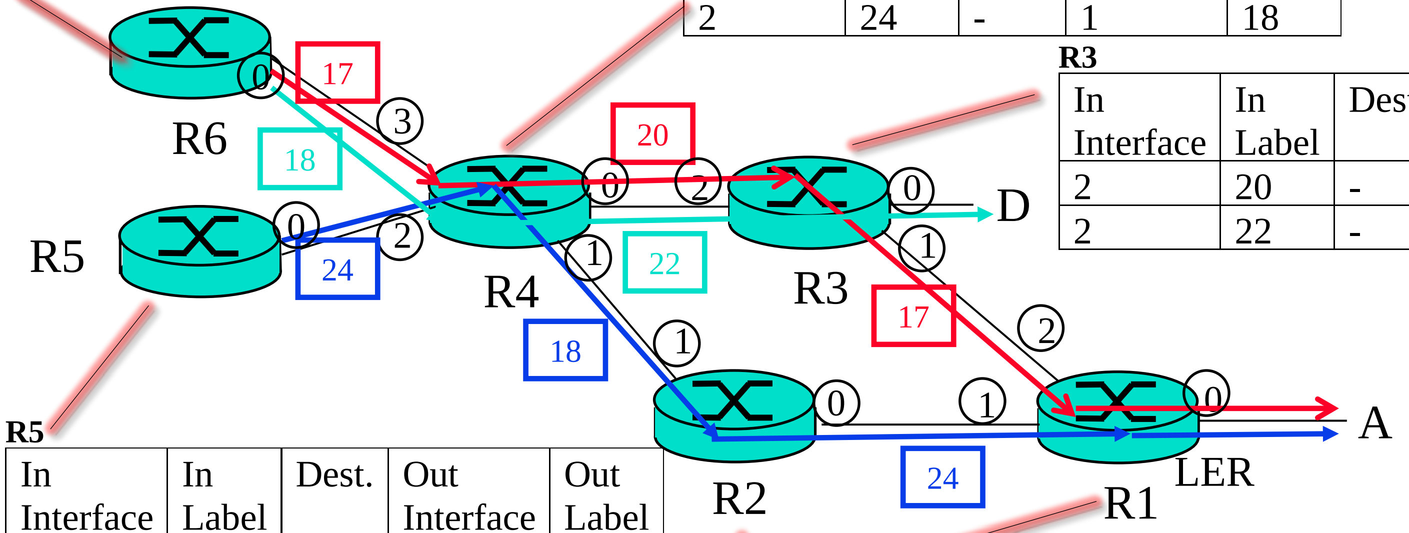
In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	17
-	None	D	0	18

R4

In Interface	In Label	Dest.	Out Interface	Out Label
3	17	-	0	20
3	18	-	0	22
2	24	-	1	18

R3

In Interface	In Label	Dest.	Out Interface	Out Label
2	20	-	1	17
2	22	-	0	None



R5

In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	24

R2

In Interface	In Label	Dest.	Out Interface	Out Label
1	18	-	0	24

R1

In Interface	In Label	Dest.	Out Interface	Out Label
2	17	-	0	None
1	24	-	0	None

Student Questions

- Why aren't both of R6's outgoing labels the same since they are both going from R4 to R3 on the same outgoing interface of R4?

This is the difference between IP address based routing vs. MPLS label based routing. Label indicates a path. Address indicates a destination. Green and red are different paths, even though they share several intermediate points.

- Why don't we need to pass the destination address across different tables?

MPLS frames are forwarded by looking at MPLS labels. The address is not required if the label-switched path (LSP) has been set up. To set up the LSP, you need the address and routing table.

Notes

- **Error in the Book:**
  - The tables are per interface not per router.
  - For compatibility, we have kept table per router but added the input interface column.
  - The book lists no input interface in the table.
  - Same label #s are allowed to be used in different interfaces of the same router. For example, See Router R3 in the "Label Switching Example" slide.
  - The textbook notation will not allow this possibility.
- Only one direction of circuits is shown for clarity.
  - There are equal number of reverse circuits that have their own labels not related to forward labels.
- Out Label=None ⇒ MPLS Tag is removed.
  In Label=None ⇒ Packet arrives with no MPLS tag

Student Questions

- Why the tables of MPLS are per-interface, not per router. *MPLS tables are used only to forward the frame from one interface to another interface inside the same router. And the label is changed so that the next router will be able to do the same.*
- Does each interface on the same router maintain an identical table?

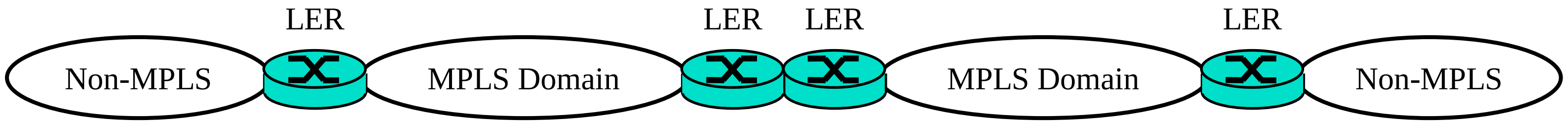
*The tables on each interface are **different.** In the book, only one MPLS table is used per router, so we have added the input interface column to show rows used at that interface. If per-interface tables were kept, that column would not be required, and all rows will refer to the same interface.*

MPLS Label Switched Paths (LSPs)

- Label switched paths (LSPs) are set up before use.
  $\Rightarrow$ Connection oriented
- During set up each router tells the previous router what label it should put on the frames of that LSP.
- The label is actually an index in the MPLS forwarding table.
- Indexing in MPLS table is much faster than searching in IP tables.
- Although speed was one reason for using MPLS but the main reason is that the bandwidth can be reserved along the path.
- Labels are local. The same label ID may be used by different routers for different LSPs.
- The label ID changes along various links of the same LSP.
- Label IDs are 20-bit long $\Rightarrow$ 2^{20}-1 Labels.
  Labels 0-15 are reserved.

Student Questions

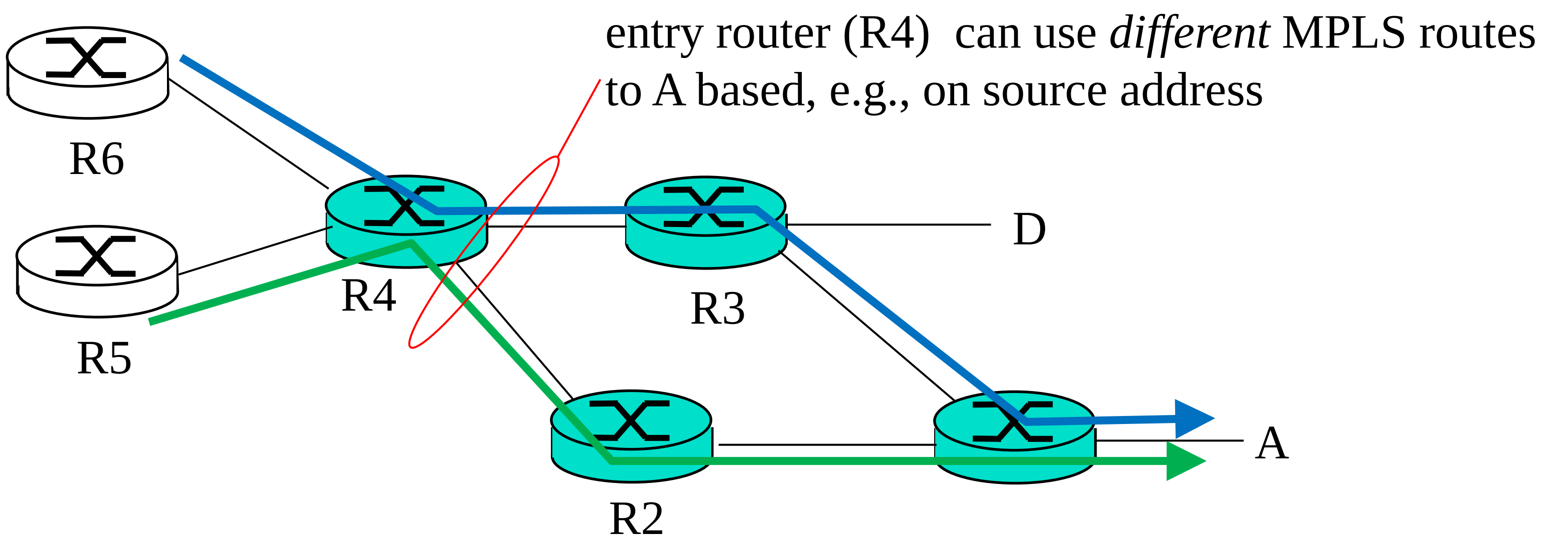
Label Edge Routers (LERs)



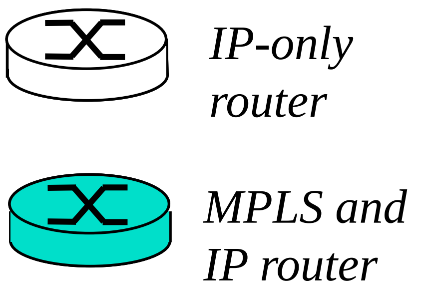
- Routers connected to non-MPLS routers or nodes or routers of other MPLS domains are called Label Edge Routers (**LERs**)
- LERs add labels to frames coming from non-MPLS nodes or remove their labels if forwarding to non-MPLS nodes or other domains.
- The labels added by LERs **may be** based on destination address along with other considerations, such as source address, QoS, etc.
- Other LSRs forward based solely on the label and the interface the frame came in. They **do not** look at the destination address field.

Student Questions

MPLS versus IP Paths



- **IP Routing**: Path determined by destination address alone
- **MPLS Routing**: Path can be based on source and destination address, flow type, …
  - **Fast reroute**: Precompute backup routes in case of link failure



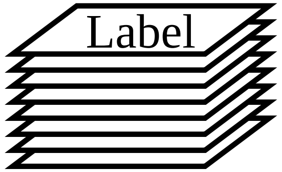
Student Questions

- In exam context, can we assume it's an "MPLS and IP" router if not specified?

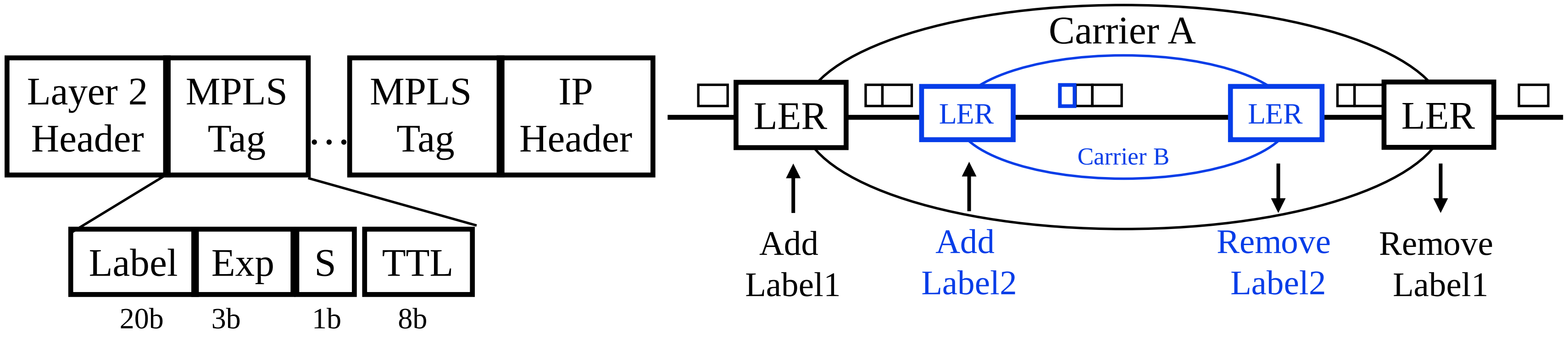
Yes. Every router is always an MPLS and IP router. IP routing is required to set up MPLS paths.

MPLS Label Format

- MPLS label is inserted after layer 2 header but before layer 3 header ⇒ MPLS is **Layer 2.5**
  - 20 bit label
  - 3 bit Experimental: Class of Service
  - 1 bit end-of-stack. A packet may have a stack of labels to allow carrier nesting.



- TTL field is decremented for all forwarded packets.
  When adding label TTL field from IP header is copied to the MPLS tag.
  When removing label TTL field from MPLS tag is copied to IP Header.
- MPLS Signaling:
  - OSPF has been extended to help prepare label tables
  - There are several other "*Label Distribution Protocols*"



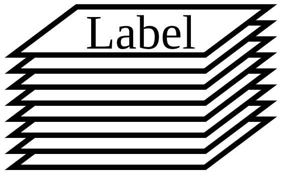
Student Questions

- So VPNs a form of overlay networks?

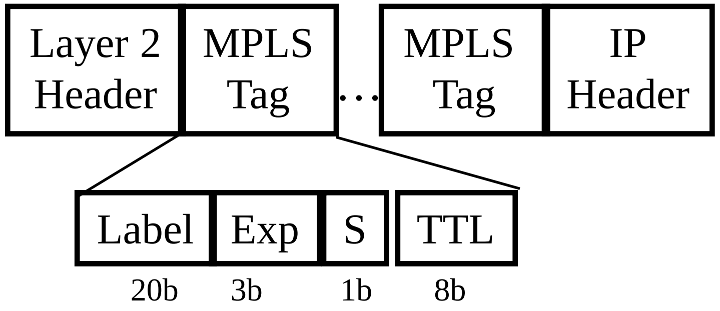
In a sense yes. However, most people do not call it that. We generally use overlay in L3 rather than L2.

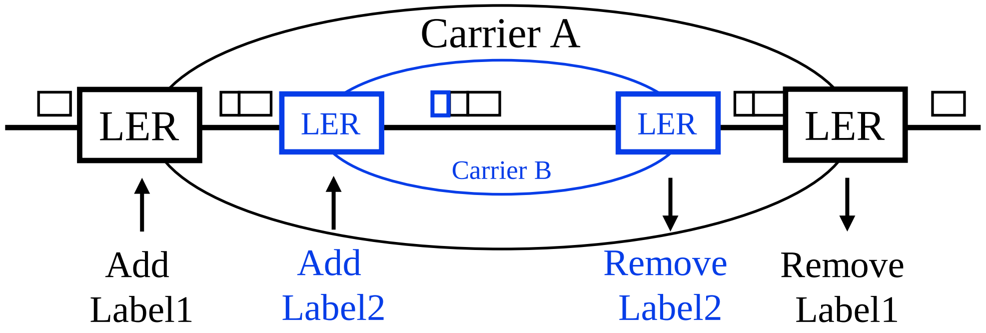
MPLS Label Format

- MPLS label is inserted after layer 2 header but before layer 3 header ⇒ MPLS is **Layer 2.5**
  - 20 bit label
  - 3 bit Experimental: Class of Service
  - 1 bit end-of-stack. A packet may have a stack of labels to allow carrier nesting.



- TTL field is decremented for all forwarded packets.
  When adding label TTL field from IP header is copied to the MPLS tag.
  When removing label TTL field from MPLS tag is copied to IP Header.
- MPLS Signaling:
  - OSPF has been extended to help prepare label tables
  - There are several other "*Label Distribution Protocols*"





Student Questions

MPLS: Review

1. Multiprotocol Label Switching (MPLS) allows virtual circuits called "Label Switched Paths (LSPs)" in IP
2. Each packet has a Layer 2.5 MPLS tag which includes a 20-bit label ID
3. Label switching routers (LSRs) forward based on input interface and the label
4. Label table is prepared by a "Label Distribution Protocol." OSPF is one example of a LDP.
5. MPLS tags can be stacked to allow network nesting

Ref: Section 6.5

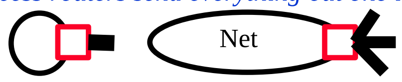
Student Questions

- You mentioned an OSPF extension for distributing labels. Is there such an extension for ICMP?

No.

- Why do you need access routers and border routers?

Border routers have to deal with routing to the world. Access routers send everything out one link.



Homework 6C: MPLS

- [6 points] Consider the MPLS network shown in "MPLS Forwarding Tables" slide. Suppose that we want to perform traffic engineering so that packets from R6 destined for A are switched to A via R6-R4-R2-R1 and packets from R5 destined for A are switched via R5-R4-R3-R1. Show the updated MPLS table in R4 that would make this possible. For simplicity, use the same label values as shown currently. Only LSP paths change and the table at Router R4.

Student Questions

Homework 6C (Cont)

- Interface numbers are in circles. Label IDs are in rectangles.

R6

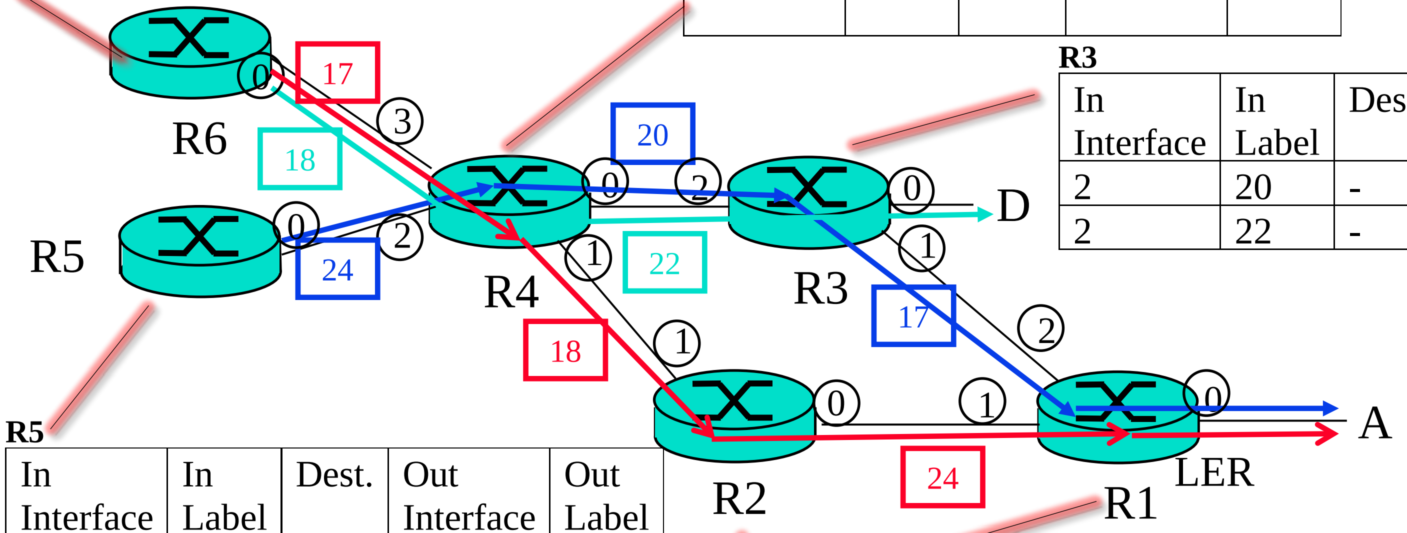
In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	17
-	None	D	0	18

R4

In Interface	In Label	Dest.	Out Interface	Out Label

R3

In Interface	In Label	Dest.	Out Interface	Out Label
2	20	-	1	17
2	22	-	0	None



R5

In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	24

R2

In Interface	In Label	Dest.	Out Interface	Out Label
1	18	-	0	24

R1

In Interface	In Label	Dest.	Out Interface	Out Label
2	17	-	0	None
1	24	-	0	None

Student Questions

Homework 6D: MPLS

- [28 points] The next figure shows the flows on an MPLS network with the reverse direction flows. Using the Labels shown, fill in all the tables.

Student Questions

Homework 6D (Cont)

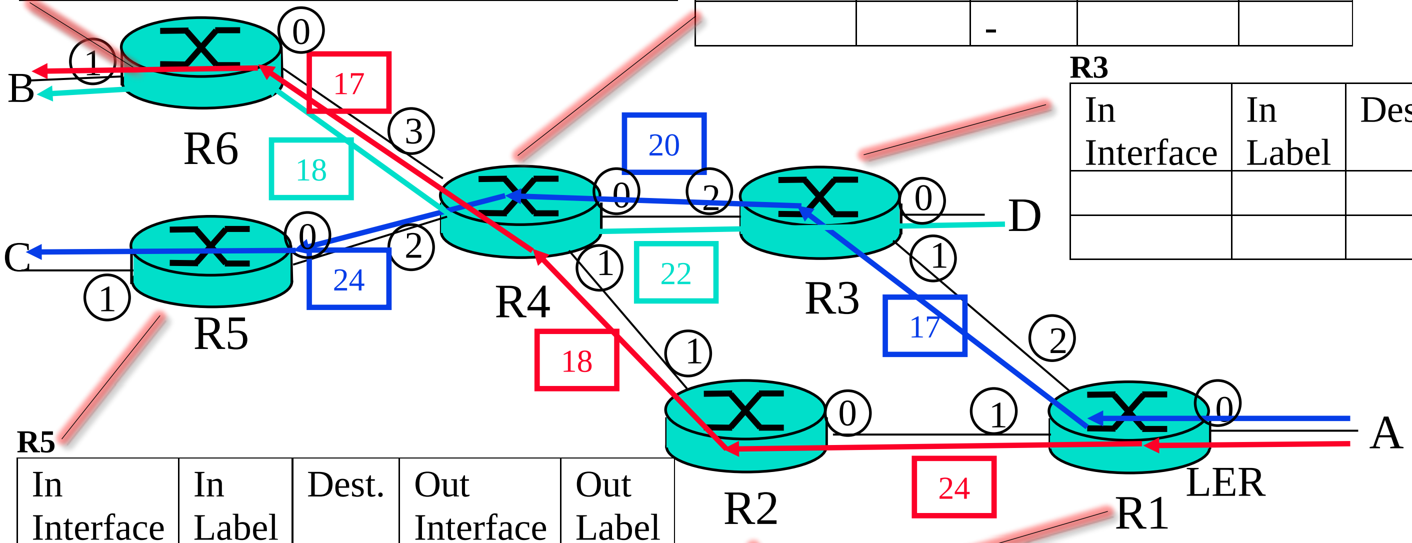
- Interface numbers are in circles. Label IDs are in rectangles.

R6

In Interface	In Label	Dest.	Out Interface	Out Label

R4

In Interface	In Label	Dest.	Out Interface	Out Label
		-		



R3

In Interface	In Label	Dest.	Out Interface	Out Label

R5

In Interface	In Label	Dest.	Out Interface	Out Label

R2

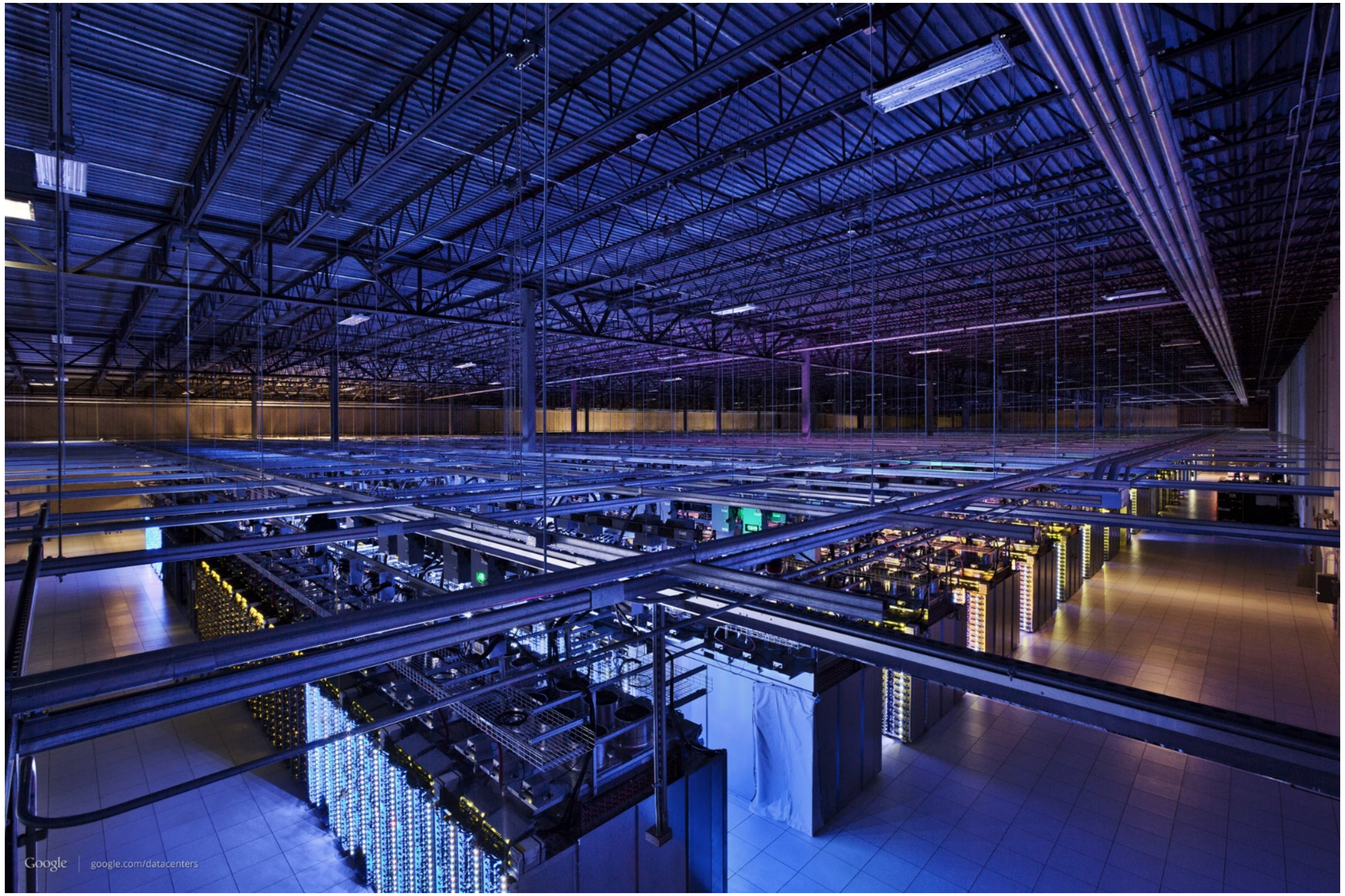
In Interface	In Label	Dest.	Out Interface	Out Label

R1

In Interface	In Label	Dest.	Out Interface	Out Label

Student Questions

Google's Data Center

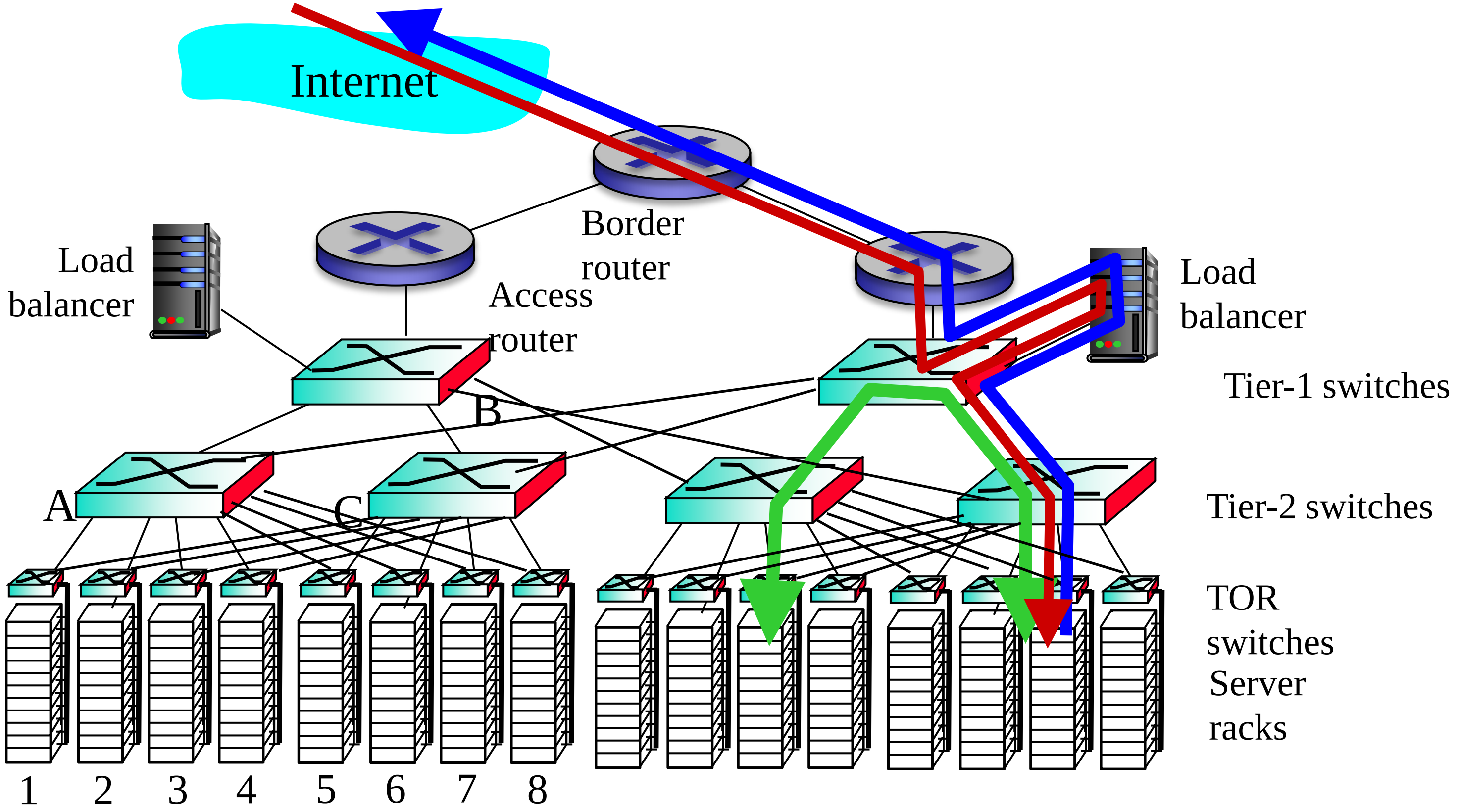


Source: http://webodysseum.com/technologyscience/visit-the-googles-data-centers/

Student Questions

Data Center Networks Topology

- **3-Tier Architecture**: Top-of-Rack, Aggregation, Core
- **Middle boxes**: Load balancer, Firewall, Intrusion detection, …
- Rich Interconnection between switches

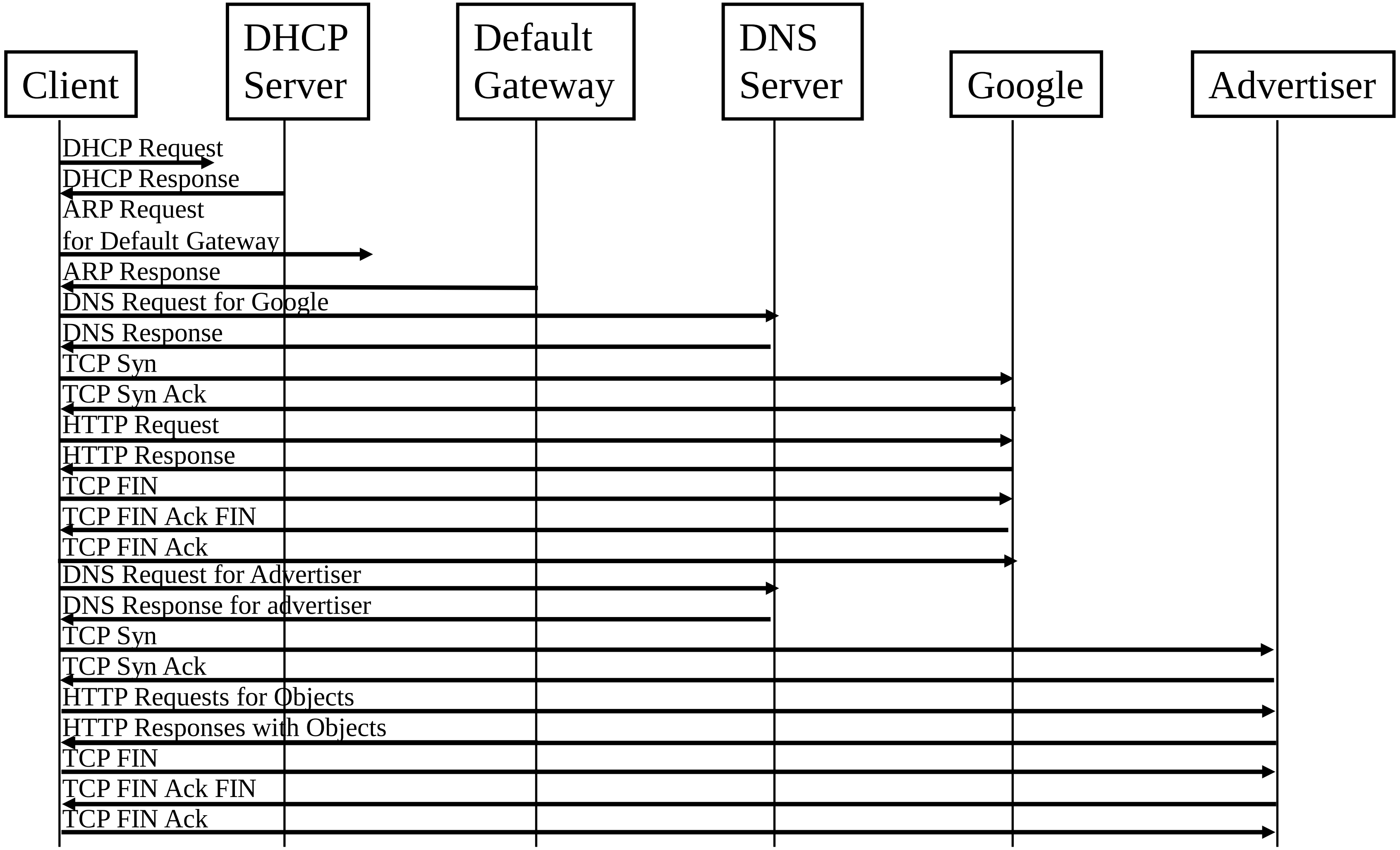


Ref: Section 6.6

Student Questions

Protocols: Complete Picture

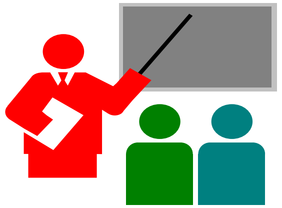
Task: Connect and search in www.google.com



Ref: Section 6.7

Student Questions

Summary



1. CRC uses mod-2 division using specially selected numbers
2. IEEE 802.3 uses a *truncated binary exponential backoff*.
3. Ethernet uses 48-bit universal addresses
4. Ethernet bridges are transparent and self-learning
5. 802.1Q allows several virtual LANs inside a LAN.
6. Address Resolution Protocol (ARP) is used to find the MAC address for a given IP address and vice versa.
7. MPLS allows virtual circuits (LSPs) on IP networks.
8. Data centers use a multi-tier switching architecture with redundancy

Student Questions

- In general, the routing layer builds the table.
  The link-layer follows the routing table, builds switch tables, and sends the bits.
  The data layer monitors the status of bits.
  Is that right?

The routing layer builds the routing tables. The data link layer does not follow the routing table. It builds its switch tables using its observations and sends frames.

The physical layer codes bits into waveform on the media and detects malformed bits.

The data link layer checks the correctness of the frames using CRC.

The routing layer further checks the correctness of datagrams using their checksum.

Acronyms

- ARP Address Resolution Protocol
- ASCII American Standard Code for Information Exchange
- CAT Category
- CD Collision Detection
- CRC Cyclic Redundancy Check
- CSMA Carrier Sense Multiple Access
- DA Destination Address
- DEI Drop Eligibility Indicator
- DHCP Dynamic Host Control Protocol
- DNS Domain Name Server
- DOCSIS Data over Cable Service Interface Specification
- FDMA Frequency Division Multiple Access
- HTTP Hypertext Transfer Protocol
- ID Identifier
- IEEE Institution of Electrical and Electronic Engineers

Student Questions

- How do routers know when an address is CIDR vs classed?

Routers do not need to know. CIDR vs classed is issue at address allocation time. Now all allocations are CIDR. Before CIDR, all allocations were class-based. Now, CIDR.

Acronyms (Cont)

- IP Internet Protocol
- IPX Internetwork Packet Exchange
- LAN Local Area Network
- LDP Label Distribution Protocol
- LLC Logical Link Control
- LSP Label Switched Path
- MAC Media Access Control
- MAP Map
- MPLS Multiprotocol Label Switching
- MSB Most Significant Byte First
- NIC Network Interface Card
- OSPF Open Shortest Path First
- OUI Organizationally Unique Identifier
- PBX Private Branch Exchange
- PCP Priority Code Point
- PHY Physical Layer

Student Questions

- What is the purpose of OpenFlow, and why is it no longer used?

OpenFlow showed the world how to program networks. It turned out to be micro-management and so other protocols are used that are more granular.

Acronyms (Cont)

SA	Source Address
STP	Shielded Twisted Pair
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TOR	Top of the Rack
TPI	Tag Protocol Identifier
TTL	Time to live
TX	Transmit
UTP	Unshielded Twisted Pair
VLAN	Virtual Local Area Network

Student Questions

Scan This to Download These Slides

Raj Jain

http://rajjain.com

http://www.cse.wustl.edu/~jain/cse473-22/i_6lan.htm

Student Questions

- When should iBGP be used instead of OSPF?

OSPF prepares a route table and distributes. BGP is used to distribute externally learned info.

- Can you go over HW 4a and 4b?

- I want to see my former attempts for assignments.

I don't know how to allow that on Canvas.

- Can you clarify whether a one-page cheat sheet can be front and back? I believe I came in with the wrong assumption last time.

You can use both the front and back sides of one 8.5" ×11" sheet.

Modulo 2 Arithmetic: More Examples

Addition:

1-bit

1	0	0	1
+1	+0	+1	+0
0	0	1	?

2-bit

00	01	10	11
+11	+11	+11	+11
11	10	01	??

3-bit

110
+101
???

Multiplication:

1-bit

1	0	0	1
×1	×0	×1	×0
1	0	0	?

2-bit

00	01	10	11
×11	×11	×11	×11
00	01	10	??
00	01	10	??
000	011	110	???

3-bit

110
×101
???
???
????

Student Questions

Modulo 2 Division: More Examples

Long Division:

Decimal Arithmetic

```
13) 1514( 116
    13
    021
     13
     084
      78
      06 ← Remainder
```

Mod-2 Arithmetic

```
10) 1101( 110
    10
    010
     10
     001
      00
      01 ← Remainder
```

```
10) 11011( 101?
    10
    010
     10
     001
      00
      01?
       ??
       ??
```

Student Questions