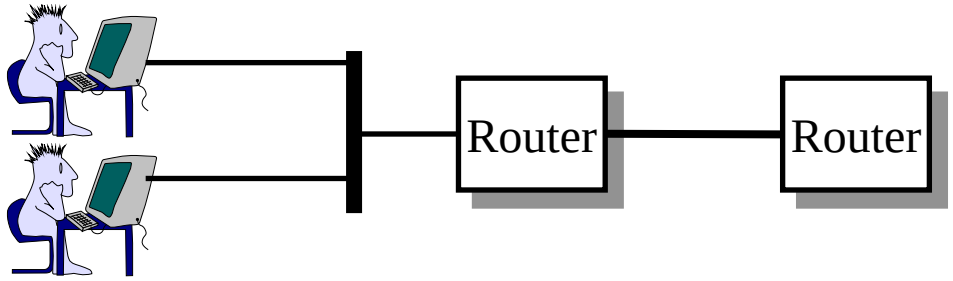
The Link Layer and LANs



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

Audio/Video recordings of this lecture are available online at:

http://www.cse.wustl.edu/~jain/cse473-24/

Student Questions

©2024 Raj Jain
Next Q 6.8

Overview

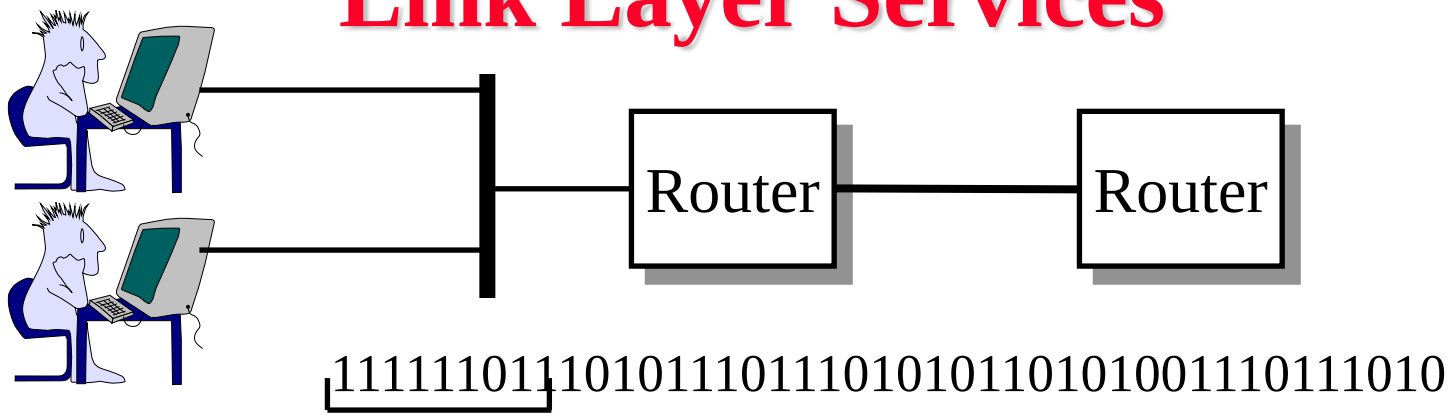
1. Datalink Services
2. Error Detection
3. Multiple Access
4. Bridging
5. MPLS

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

Student Questions

©2024 Raj Jain

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

- What's the frame beginning pattern? Will the pattern change, or is the pattern ensured?

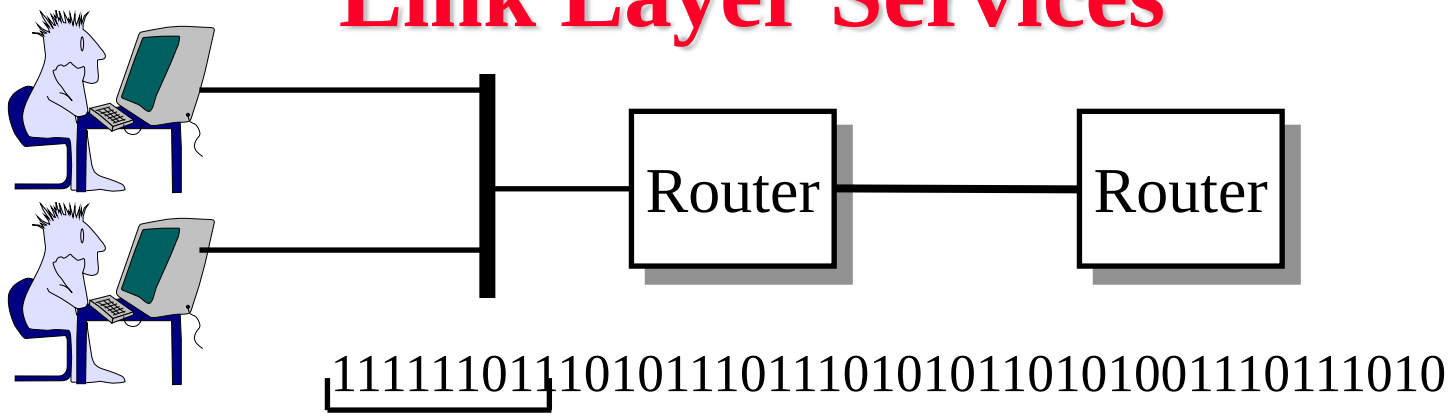
Patterns are fixed but may be different for different speeds.

- Where does the MAC protocol fit under the link layer's services? Is this used to provide multiple access to users on one wire?

It is the critical part of the data link layer. Yes, it provides multiple access.

©2024 Raj Jain

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

- What is a frame?

A physical layer packet

- When pointing at each link, it is not shown on the slides. Could you point them out?

Pointers were not recorded in the live class. I have fixed it on very complex slides.

- How does the implementation of links differ when they connect routers and computers? E.g., router to router vs. router to the computer.

Router-to-router used to be different. But now, all of them use Ethernet.

- Is there a maximum number of users that can share on the wire?

It depends on the protocol. The original Ethernet had a limit of 1024.

©2024 Raj Jain

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

- Is a bus network half-duplex or full duplex?

Most buses have separate lines for each direction.

- How does the half-duplex decide which side to transmit data to? Is it a kind of multiplexing?

Like a tin-can phone. Ten-Four.

- Does Simplex / Half Duplex imply single and Full Duplex double wires?

The number of wires is like lanes on the highway. Duplexity is related to the direction of traffic, not lanes.

©2024 Raj Jain

Error Detection

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

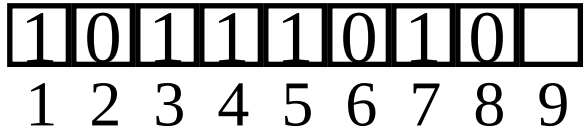
Student Questions

- How can we know which error check to use in various situations?

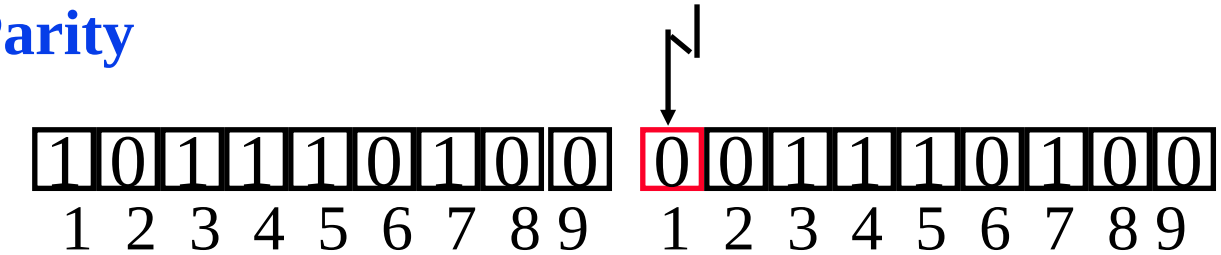
Each protocol designer has already selected one.

©2024 Raj Jain

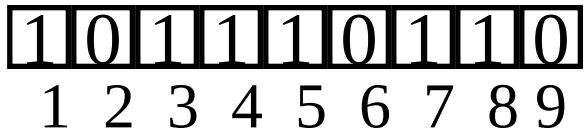
Parity Checks



- **Odd Parity**



- **Even Parity**



Student Questions

- Is there an advantage to using odd parity rather than even parity checks? *No.*
- Why do we use parity check if half of the errors cannot be detected? Why do not just use checksum instead?

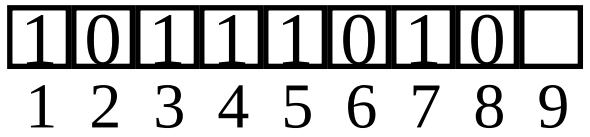
Each byte of memory uses parity. I can't do checksum for each byte.

- Both odd and even parity can detect odd-bit errors but not even-bit ones. Yes
- Is the only difference in how they add their 9th bit? *Yes*
- So, do we use odd and even at the same time? *No*
- Can you explain even parity vs. odd? When do we use each one? *Both do the same thing.*
- Do we have two nine-bit sequences here? According to the number of 1s in one sequence, 1 or 0 is filled in another. Is this true? *No.*
- Can you explain odd parity for 2-bit errors vs. 1-bit parity for detection for all odd numbers of bit errors?

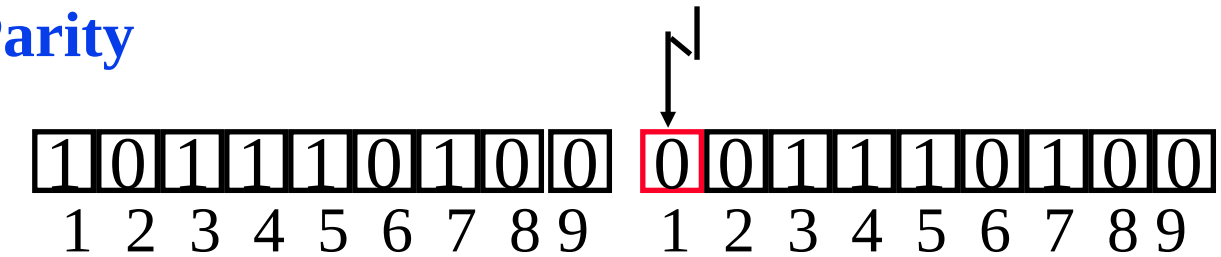
Odd parity=odd # of 1's
Even parity=Even # of 1's

©2024 Raj Jain

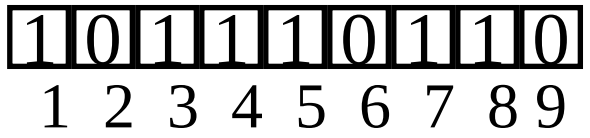
Parity Checks



- **Odd Parity**



- **Even Parity**



Student Questions

- The slide doesn't show it, but for the 1-bit and 3-bit error for odd parity, would the 9th bit be 1?

In the example shown, for 1-bit errors, the 9th bit would be 0. The same applies to 3-bit errors.

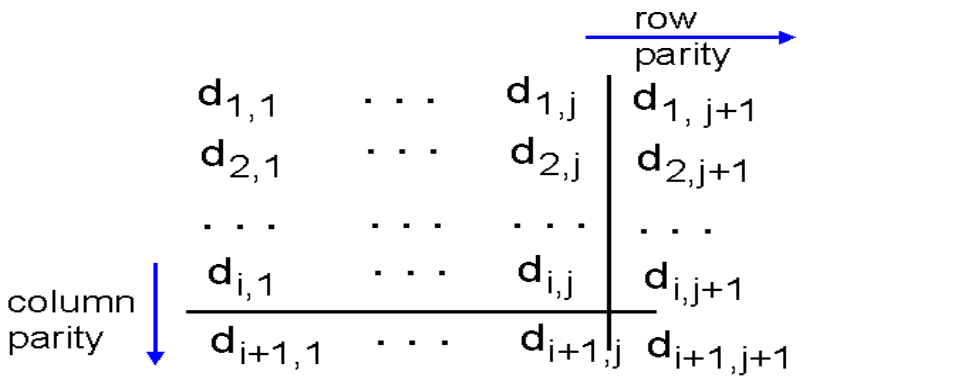
- Is there a specific time when we should use odd parity vs even, or is it just a personal choice?

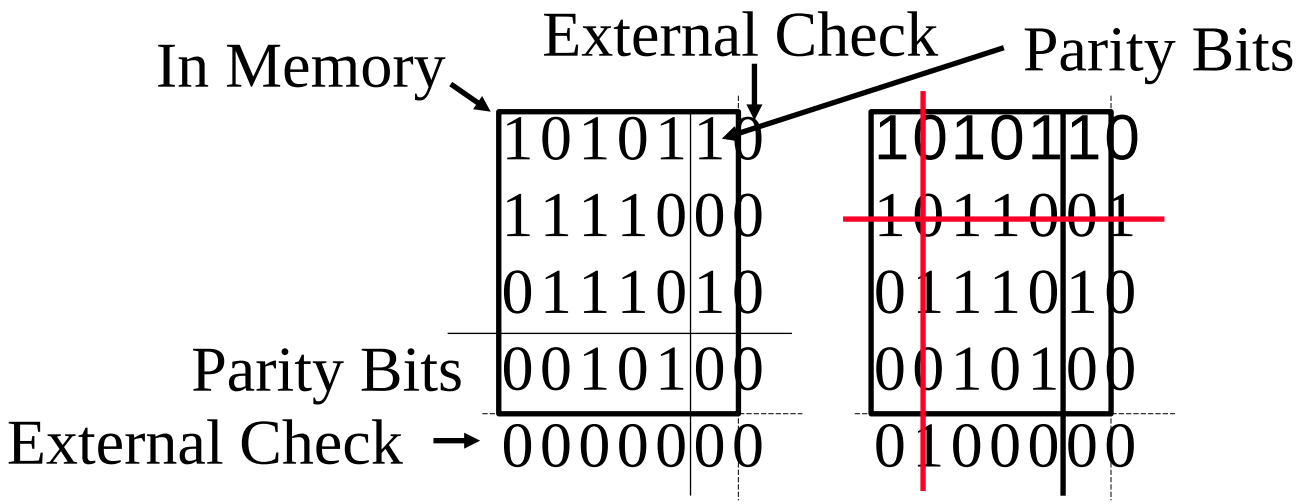
No difference. Personal choice.

©2024 Raj Jain

Two-Dimensional Parity

- Detect and correct single bit errors





Student Questions

- Why are there no errors on the left chart and an error on the right? Does it have to do with even/odd numbers of 1s?

Yes. See the revised picture.

- Can we send and process data in bits instead of bytes? Do 9 bits take up 2 bytes in computer systems?

Each 8-bit byte is stored as 9 bits in the memory.

- What do you mean by "this can only detect single-bit errors"? Can't it also detect any odd number of bit errors?

Some odd number bits may not be detected.

- Would the extra bits also suffer from errors and thus introduce more error frames?

Of course, the error probability in longer frames is slightly higher. But any single bit error in the data or protection bits is detected.

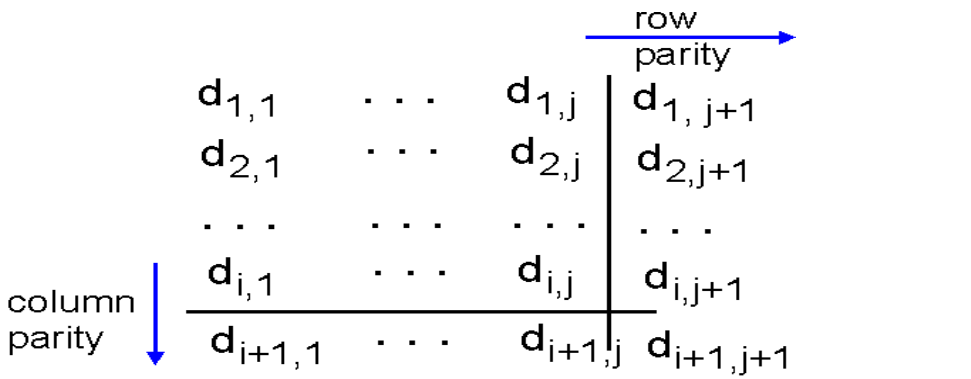
- In the book, it says it can detect any combination of two-bit errors, while we said it could only detect single-bit errors. Which one is correct?

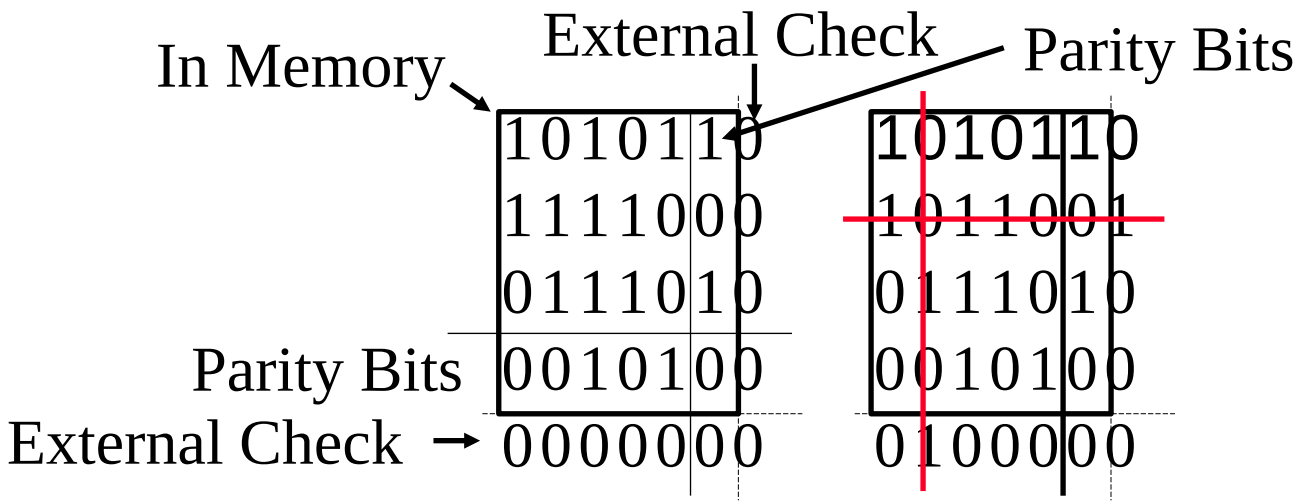
*It can detect and **correct** single-bit errors. It can detect two-bit errors.*

©2024 Raj Jain

Two-Dimensional Parity

- Detect and correct single bit errors





Student Questions

- When would you use one-dimensional parity, and when would you use two-dimensional parity?

Two-dimensional is more costly and is used only when the application deserves it. This is just the tip of the Information Theory. There are a million other Error detecting and correcting methods.

©2024 Raj Jain

Check Digit Method

- Make a number divisible by 9

Example: 823 is to be sent

1. Left-shift: 8230
2. Divide by 9, and find the remainder: 4
3. Subtract the 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 except 0-9 substitutions: 7235, 8335, 8255, 8237

Detects several multiple-digit errors: 8765, 7346

It does not detect some errors: 7335, 8775, etc.

Does not detect transpositions: 2835

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

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

Student Questions

- Does the check digit method provide any error correction at all?

Not with single check digit

- If the message we want to send is 80, send 801. If the message is altered to 891, it is still divisible by 9.

Yes, 0-9 substitutions are not detected.

- Does single check digit have any error correction?

No.

❖ Will we be tested on checking errors with the check digit method or just CRC?

Both

❖ Could you go over the check digit method again?

Sure

©2024 Raj Jain

Next Q 6.10c

Modulo 2 Arithmetic

```
  1111        11001          110
 +1010      ×     11     11  | 1010
 --------    --------        / 11
  0101        11001          ---------
              11001            x11
              ---------          11
              101011          --------
                                x00
                                 00
                               -----
                                 x0
```

010	2	
011	3	
---	--	
001	1	Mod 2
101	5	Binary

See Slides 6.70 and 6.71 for more Mod-2 examples.

Student Questions

- Can you redo the example? This slide was difficult for me to follow. *Sure.*
- Is the modulo-2 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 the divisor is one can I get 1 for the quotient. Otherwise, the quotient should be 0, right? *Yes.*

©2024 Raj Jain

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 the 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)

- So, will the P be given to us in the exam? do we need to find it ourselves

P will be given

- What is the relationship between n and the length of M?

N is the size of CRC.
M is the message.

- Will the P value be sent along with the original data to ensure everyone has the same P?

No, it is specified in the protocol.

- Can we determine the location of the bit error with CRC?

No

©2024 Raj Jain

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 the 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

- The textbook says that we use checksum at the transport layer and CRC more often at the link layer because the software on the transport layer needs to be able to calculate more rapidly. At some point, when computers become fast enough, will we begin using CRC on the transport layer?

No. With a faster computer, links will also be faster, and more bits will need to be sent.

- Are steps 4 and 5 the checks performed at the receiver?

Only Step 5 is performed at the receiver.

- We've talked about CRC needing specifically shift registers, but can CRC be used in software anyway?

Sure. But it would be slow.

©2024 Raj Jain

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 the 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

- What is the difference between CRC and checksum?

Checksum is based on sum and is more straightforward than CRC, which is based on division.

❖ What are popular CRC polynomials?

Polynomials are not a part of this course. They are covered in CSE567. Standards bodies have already selected one polynomial (CRC-32).

❖ Is there ever an advantage to using party checks for more advanced techniques like CRC?

Yes. Quantity, Cost, and Quality determine the optimal.

©2024 Raj Jain

Next Q 6.19c

Modulo 2 Division

Q=1101010110

P=110101)1010001101 00000=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.

- In the first subtraction, isn't 101000<110101? Then why put 1 for the most significant digit?

Exclusive-or is not a subtraction. In the Mod-2 division, we look at the first bit only.

- Could you do modulo 2 division in detail?

Sure.

- For CRC, is it required to show all digit place calculations? Is it ok to skip some unnecessary steps?

Yes, show all the steps.

©2024 Raj Jain

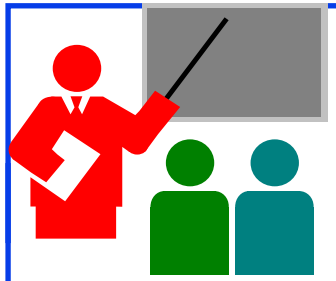
Checking At The Receiver

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

```
010111
000000
 101111
 110101
  110101
  110101
   00000
```

Student Questions

©2024 Raj Jain



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

Student Questions

- Why can't CRC be done with software? If we can do it by hand, shouldn't software also be able to do it?

CRC uses big divisors and can't be done by hand. It will take a long time with the software. It is trivial with hardware.

- In the summary, it says it is good to use prime numbers for error detection, so why did we choose to use 9 in this case?

It is better than some other numbers, e.g., 11.

Ref: Section 6.2, Review questions R2, R3

©2024 Raj Jain

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

- What is meant by "generator"? Is it P? *Yes*

- When will the homework be due, given that spring break is next week?

As specified in the Canvas.
Monday after the break.

©2024 Raj Jain

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

- For Multiple access, if two hosts send bits to a router simultaneously, will one host need to wait for another host, or can they send simultaneously?

They wait.

©2024 Raj Jain

Multiple Access

(a) Aloha Multiple Access

(b) Carrier-Sense Multiple Access with Collision Detection

Student Questions

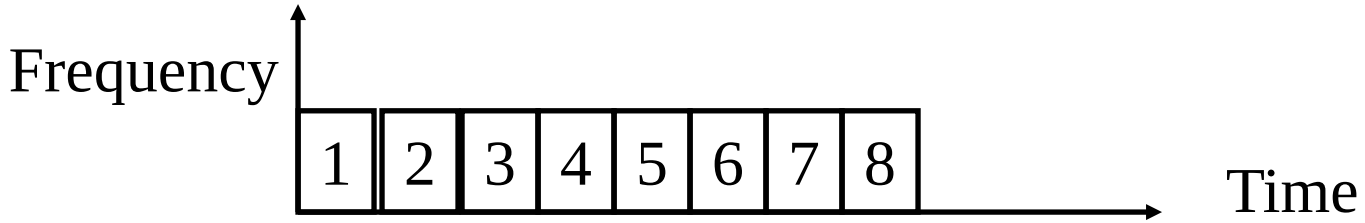
- Could you go over the difference between TDM and time-slotted Aloha again?

In TDM, all slots are pre-reserved. In time-slotted Aloha, there is contention for each slot.

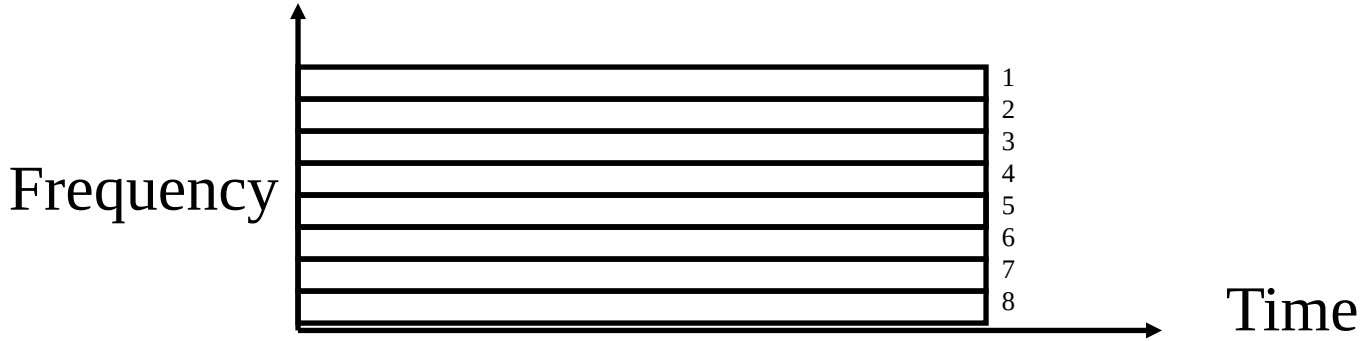
©2024 Raj Jain

Multiple Access

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



- **Frequency Division Multiple Access**



Student Questions

- Is there any difference between this TDM/FDM and the multiplexing that we saw in the past at the network or application layer?

Multiplexing=combining
TDM=Combining into slots
TDMA=Accessing using slots

©2024 Raj Jain

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 leading to 18% and 37% and how those are a "worst case"?

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

- How do Aloha nodes know that their message has been received correctly? Would this method be very error-prone?

They would need an ack at this or a higher layer if required.

- What is the one-time unit? Is it seconds?

It can be any time, e.g., 64 microseconds.

©2024 Raj Jain

IEEE 802.3 CSMA/CD

- If the medium is idle, transmit (1-persistent).
- If the medium is busy, wait until idle and 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
  ⇒ The length of the cable is limited to 2.5 km.

Student Questions

- Why is k in the random time formula?

k^{th} retransmission

- Do original CSMA transmitters listen while transmitting?

*No. They listen **before** transmitting.*

- Why do we wait for random time instead of a fixed time like timeout?

A fixed time will result in colliding again.

- Why is the distance limit 2.5 km?

See slide 6.21

- In the book, it says that when the media is idle, we transmit immediately. Why is the slide different from the book? Which one should we follow?

We say the same. See the first line of this slide.

- Are collisions the most common reason for errors in transmissions?

No. With CSMA/CD, collisions are rare.

- What happens when a collision is detected 16 random times in a row?

The transmission fails. The packet is dropped. The connection will eventually break due to overload.

.

©2024 Raj Jain

IEEE 802.3 CSMA/CD

- If the medium is idle, transmit (1-persistent).
- If the medium is busy, wait until idle and 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
  ⇒ The length of the cable is limited to 2.5 km.

Student Questions

- Does 1-persistent here mean the persistent level is 1?

Yes, p=1. See p-persistence in the previous slide.

- What is the meaning of k here?

k^{th} reattempt

- When it says reattempt does it mean reattempt normal transmission or jam signal?

Normal transmission.

- What determines the length of the jam signal?

Specified in the standard. It was computed so that it could reach everyone from one end of the wire to the other.

©2024 Raj Jain

IEEE 802.3 CSMA/CD

- If the medium is idle, transmit (1-persistent).
- If the medium is busy, wait until idle and 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
  ⇒ The length of the cable is limited to 2.5 km.

Student Questions

- Is there priority on which packet or router goes first once the medium becomes idle?

Everyone starts immediately for the first time. Other times, they go by the random number they have drawn.

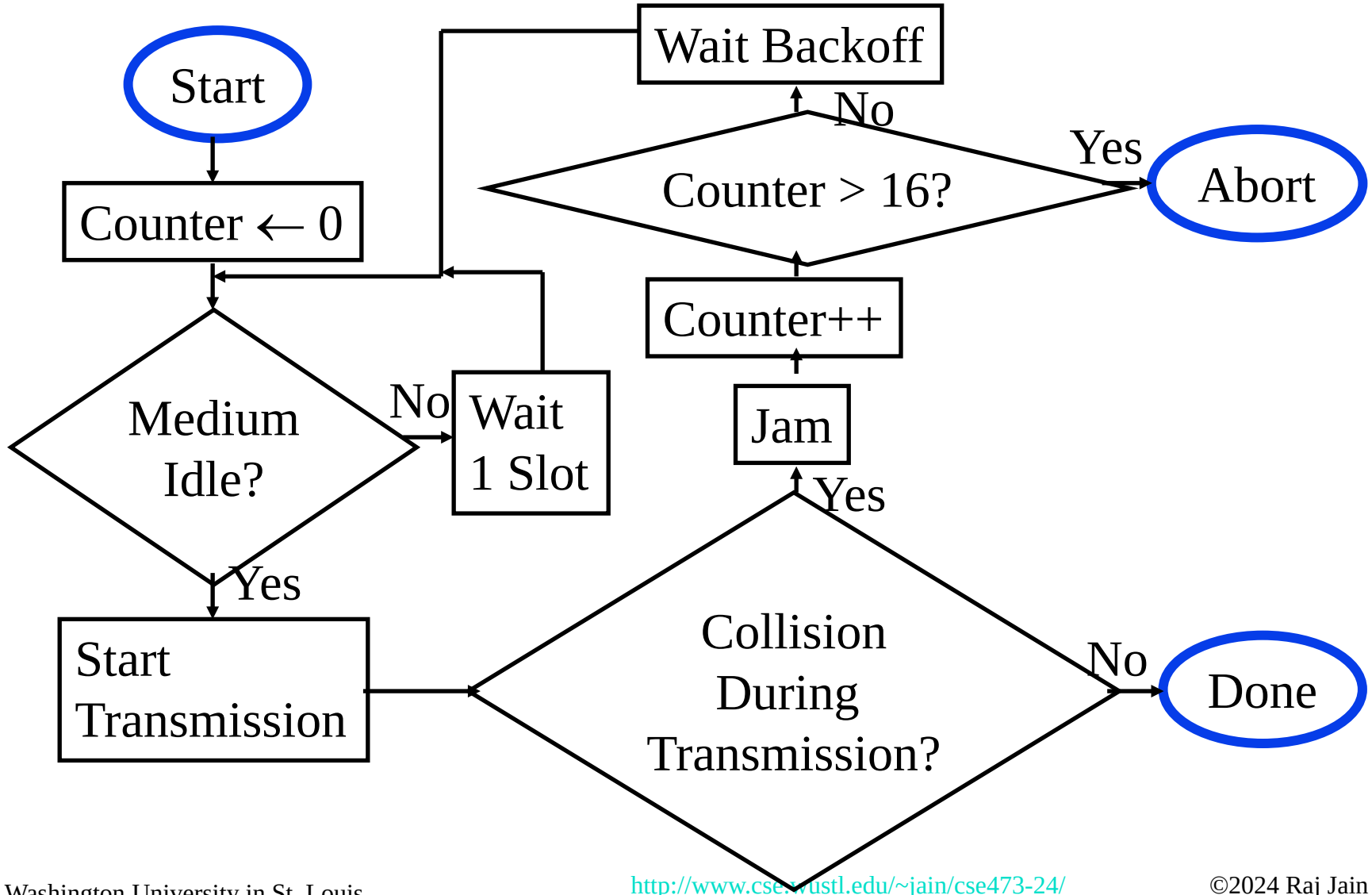
❖ How does a jam signal work?
High voltage on the wire. Any station can't miss it.

❖ What is truncated in truncated binary backoff?
k

❖ What uses CSMA?
Wi-Fi

©2024 Raj Jain

IEEE 802.3 CSMA/CD Flow Chart



Student Questions

- ❑ What does this counter count?

Number of attempts

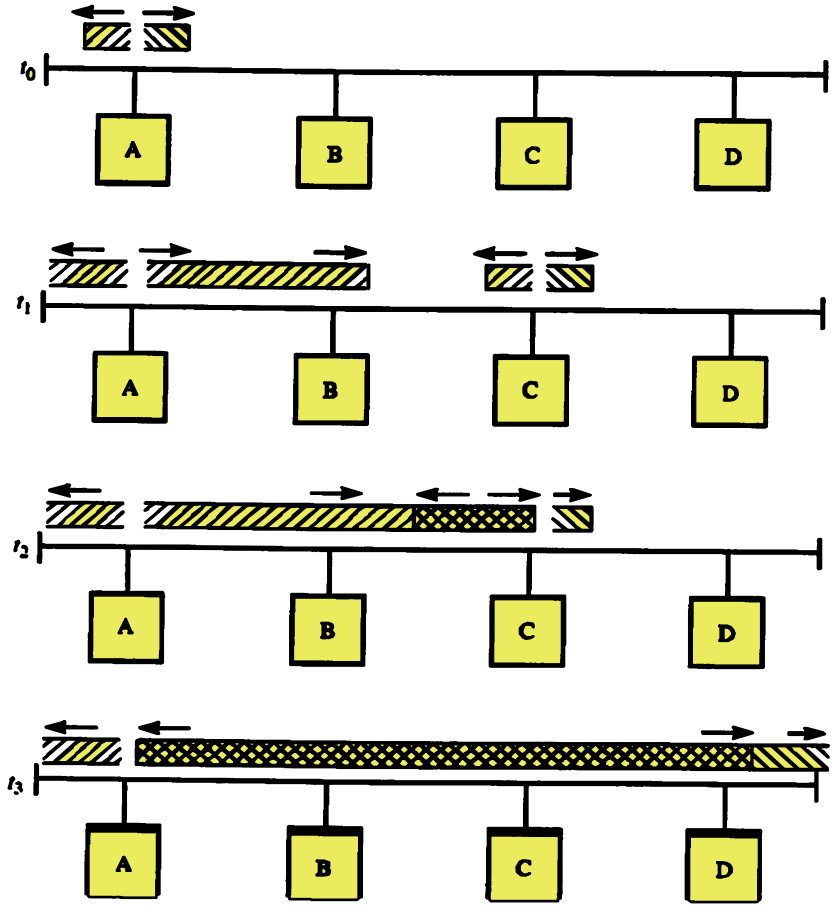
- ❑ What happens to the data when transmission is aborted?

The packet is dropped.

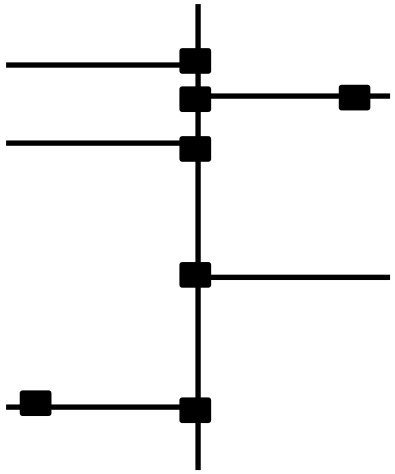
©2024 Raj Jain

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 the 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.

- Why don't A and C experience collision at the same time?

Since the signals start at different times.

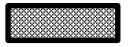
- Is the propagation delay the same for every distance value?

No. It is linearly proportional to distance.

- What does collision window mean?

Time until when the collision can happen.

- Where is the jam signal in the diagram?

 *is the jam signal*

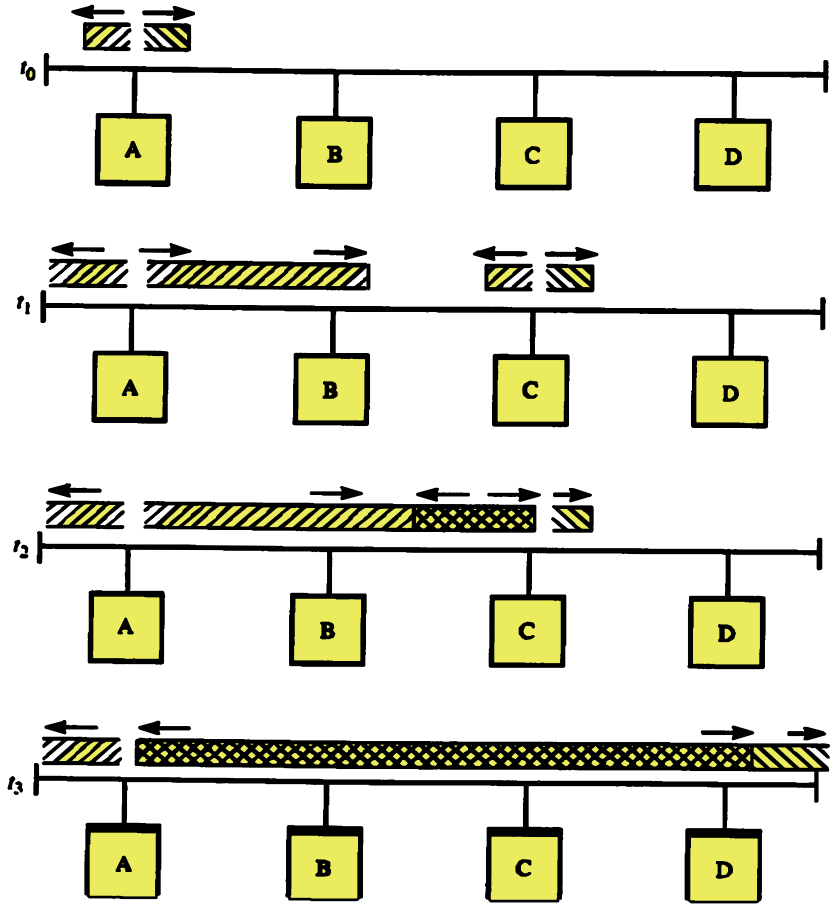
- Why is the time to notice a jam two times the propagation delay? Isn't it that a host knows it is in a jam when it receives a signal from another host while transferring?

They know it is a collision in one round-trip delay. Jam is higher voltage.

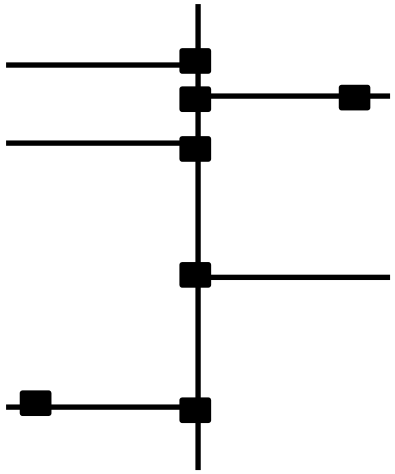
©2024 Raj Jain

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

- Is this a multi-access protocol per link? Would this only apply to broadcast wires?

Broadcast is a type of packet, not wires. The protocol applies to all packets. In CSMA/CD, all nodes hear everything.

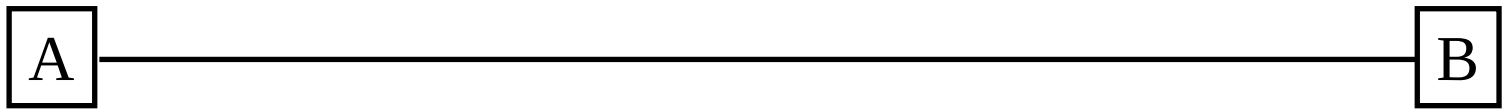
- Can you go over this example again? Why Collision window = 2 × One-way Propagation delay?

A and C are located on extreme ends. C starts transmitting just 1 bit before the first bit of A reaches it. A will sense the collision only after 1 round-trip.

©2024 Raj Jain

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. Node A begins transmitting a frame, and node B begins transmitting a frame before it finishes. 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 the preamble).

- Does downstream use TDM or FDM, or both?

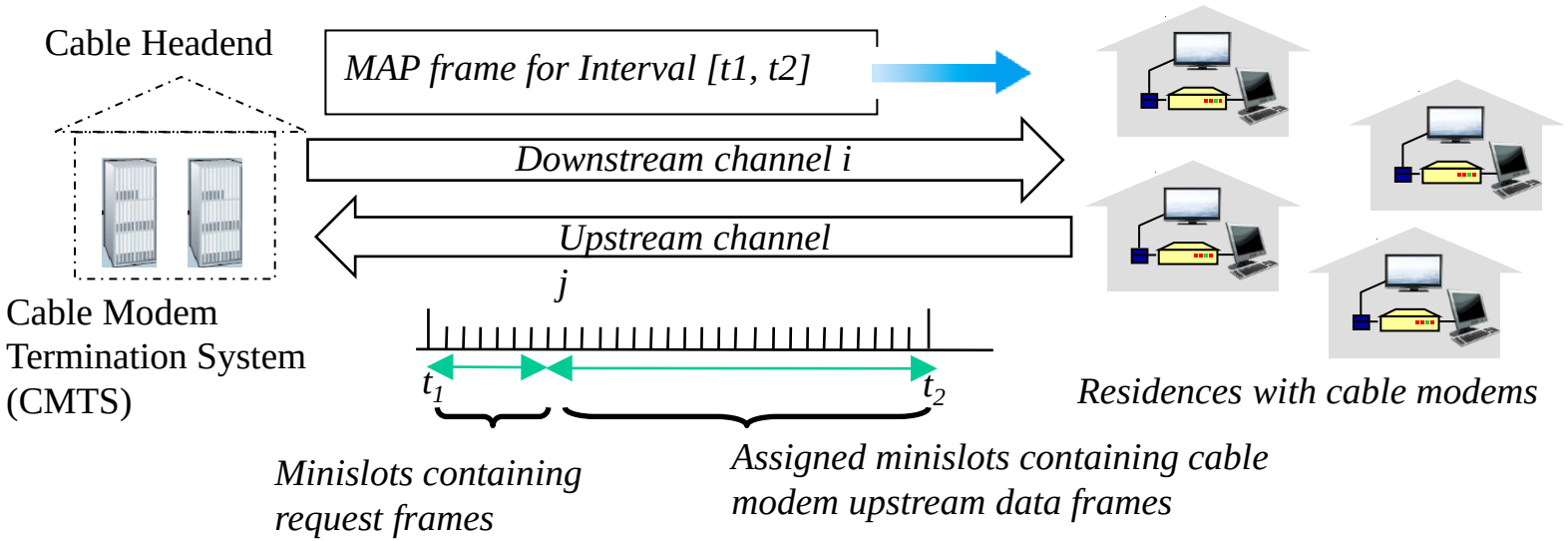
The problem says "Ethernet." It does not use FDM.

- When B starts transmitting before 251-bit times, A will not finish. Does that mean that for the first question, both yes and no are correct?

In the worst case, B's signal will reach A at 650 bits. The minimum frame size is 576 bits. So yes, A may finish transmitting and not know that its frame collided. The bus is longer than allowed.

©2024 Raj Jain

Cable Access Network

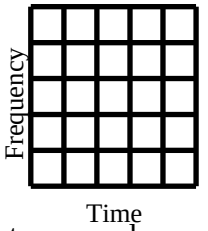


- **DOCSIS**: Data Over Cable Service Interface Specification
- Frequency Division Multiplexed (**FDM**) channels over upstream and downstream
- Time Division Multiplexed (**TDM**) slots in each upstream channel:
  - Some slots are assigned, and some have contention
  - Downstream **MAP** frame: Assigns 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 does the MAP frame arrive over downstream FDM?

FDM is both ways.



- Can you explain the upstream and downstream of the cable access network again? *Sure.*
- What happens if the number of users trying to send upstream traffic is so large that a collision always occurs? Can this shut down the network?

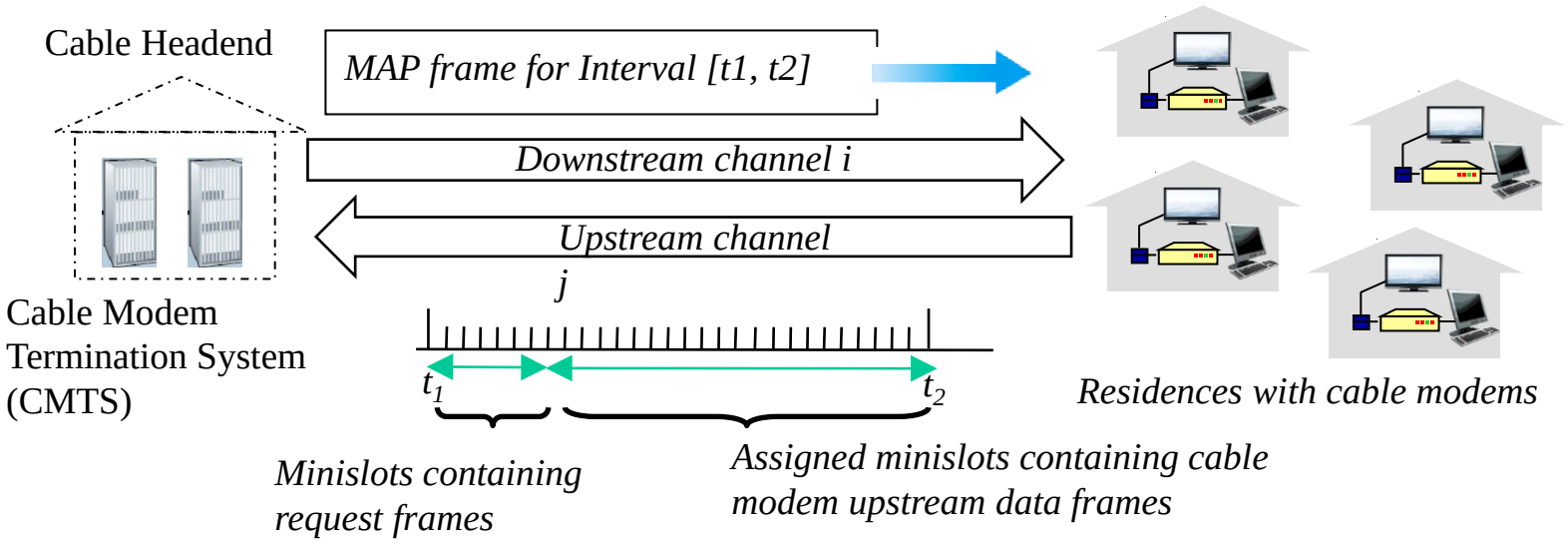
If all upstream transmission slots result in collisions, the stations will try again until they succeed.

- For Cable Access Networks, a modem requests for an upstream timeslot during the contention time slots, and then hopefully, they get sent a MAP frame with an assigned collision-free time slot.

Yes.

©2024 Raj Jain

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 are assigned, and 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

- Does this combination of FDM and TDM have a new name? *DOCSIS*
- How do cable access networks compare to other broadband technologies, such as DSL and fiber optic, regarding speed, reliability, cost, and scalability?

Cable access networks were faster initially.

- The cable access network upstream has a slower rate than downstream because fewer slots have been assigned upstream than downstream. *Yes.*
- What does "some slots have contention" mean?

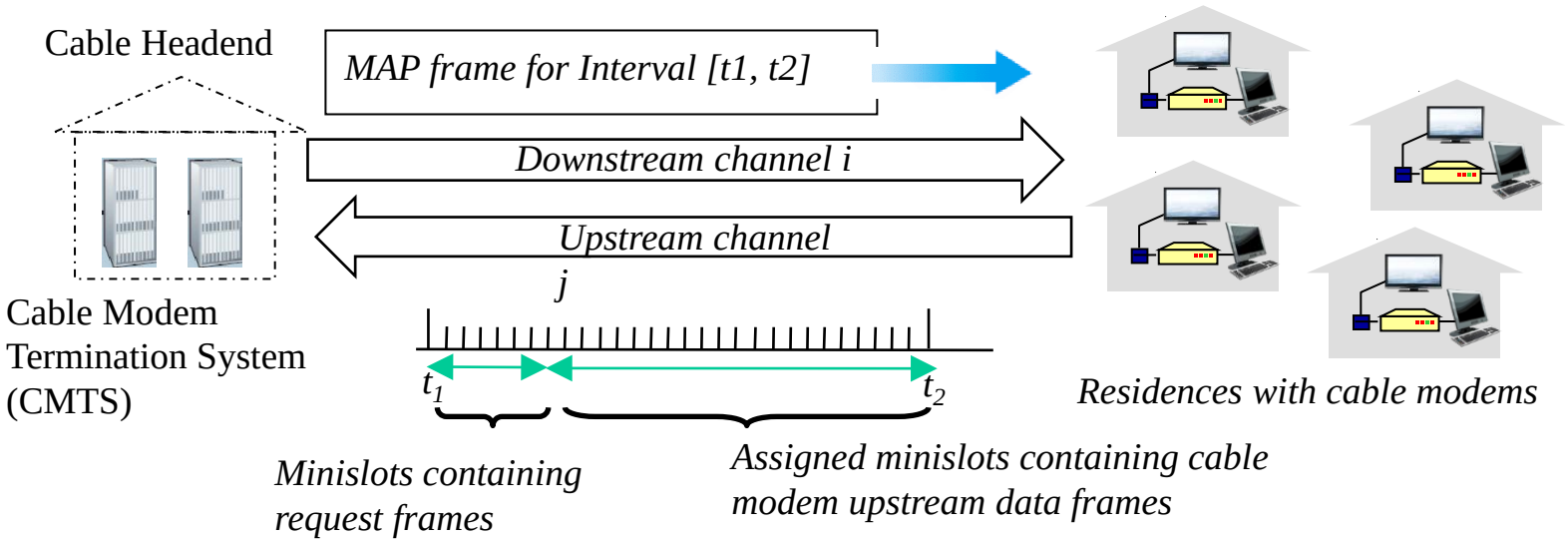
They are available to use by contention. They are not reserved.

- Why do by-assignment and by-contention exist together upstream? What are they for?

The data is transmitted only on assigned slots. The contention is to request slots.

©2024 Raj Jain

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 are assigned, and 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

- Can you clarify the request frame vs. the upstream data frame?

Data frames are used to transmit data in pre-assigned slots. Request frames are used to request these slots.

❖Is MAP literally a map?

Yes, it gives directions on who can transmit and when.

©2024 Raj Jain

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 improve 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. Each channel time division is multiplexed, with some slots reserved for random access.

Student Questions

- Can you go over R4 and R6:

R4: What is the role of the forwarding table within a router?

They are used to determine which port to forward a packet to.

- R6: A router typically consists of input ports, output ports, a switching fabric, and a routing processor. Which of these are implemented in hardware?

Switching fabric may be in software.

- Data plane and control plane. Which is implemented in hardware?

The data plane is high-speed and needs special hardware.

Ref: Section 6.3, Review question R4-R8

©2024 Raj Jain

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

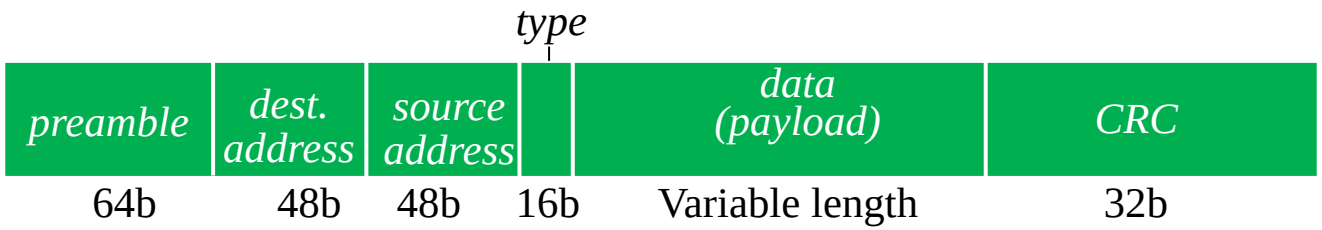
Student Questions

- Textbook, page 478: “It’s important to note, however, that link-layer switches do not have link-layer addresses associated with their interfaces that connect to hosts and routers." Can you explain why they don’t?

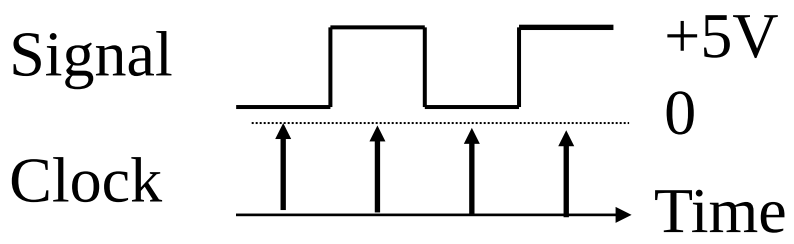
All interfaces have MAC addresses. But switch interfaces pick up frames with other MAC addresses also. The hosts do not have to specify the switch MAC address as the intermediate point.

©2024 Raj Jain

Ethernet Frame Structure



- ***Preamble:*** 7 bytes with pattern 10101010 followed by one byte with pattern 10101011. To synchronize the receiver, the sender clocks
- ***Addresses:*** 6-byte source, destination MAC addresses
- ***Type:*** indicates higher layer protocol
  - IP: 0x0800
  - ARP: 0x0806



- ***CRC:*** Cyclic Redundancy Check
  - If an error is detected, the 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 need more time to turn around and miss the frame.

- For error detection, is only the CRC field used in computation, or is 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, it should come out to zero.

- What is the MTU in Ethernet frame structure, and how does it affect network performance?

MTU is for IP. In Ethernet, the maximum frame size was 1518 bytes initially.

- Why do we use a MAC address in this case?

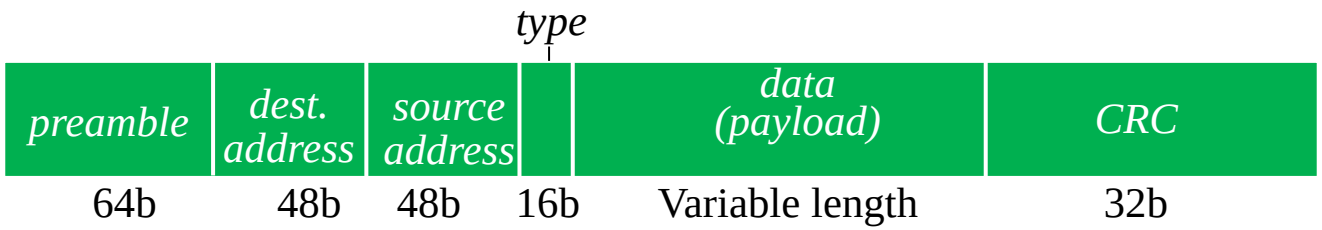
Every layer uses its own addresses.

- Does preamble length have implications in clock synchronizing?

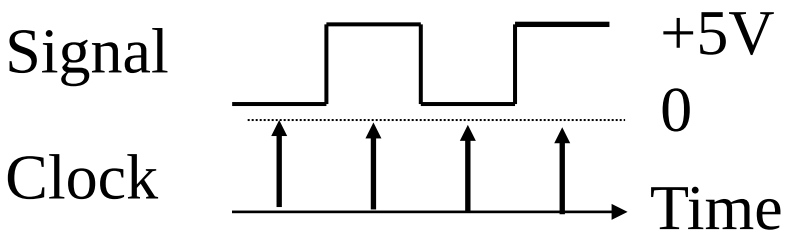
Yes. Higher speeds require more precision.

©2024 Raj Jain

Ethernet Frame Structure



- ***Preamble:*** 7 bytes with pattern 10101010 followed by one byte with pattern 10101011. To synchronize the receiver, the sender clocks
- ***Addresses:*** 6-byte source, destination MAC addresses
- ***Type:*** indicates higher layer protocol
  - IP: 0x0800
  - ARP: 0x0806



- ***CRC:*** Cyclic Redundancy Check
  - If an error is detected, the frame is silently dropped at the receiver
- ***Connectionless***: No need to ask the receiver
- ***Unreliable***: No ack, nack, or retransmissions

Student Questions

- What is the purpose of the preamble?

It helps synchronize the clocks.

- The book says that IEEE will assign blocks of MAC addresses to a company. What if these addresses don't end up being used and are wasted? Will we ever run out of address space?

2^{24} addresses for $1000. We are already using 64-bit addresses for IoT instead of 48-bit IEEE-802 addresses.

- What is meant by silently dropped? How is this different than when a packet is dropped regularly?

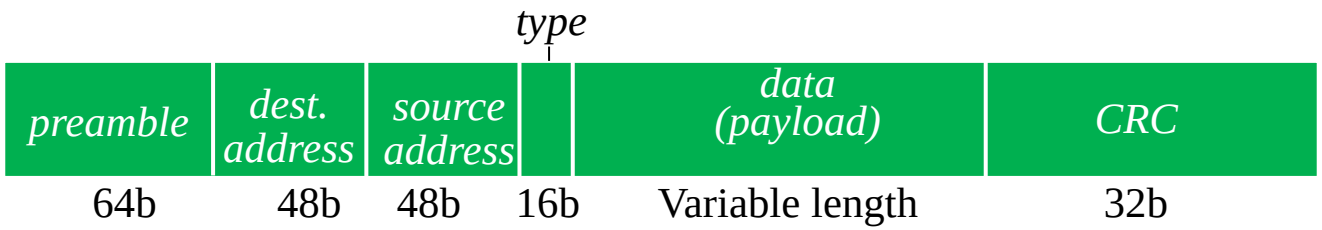
The source needs to be informed, as is done in the upper layers. E.G., ICMP is used to inform the source at layer 3.

- What makes CRC "cyclic"?

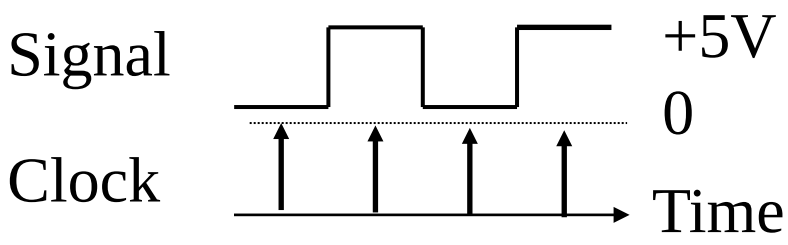
Hardware implementation is cyclic.

©2024 Raj Jain

Ethernet Frame Structure



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



Student Questions

- Why is the problem 7 bytes long? Wouldn't it be better for it to be shorter so you don't have to send as much data?

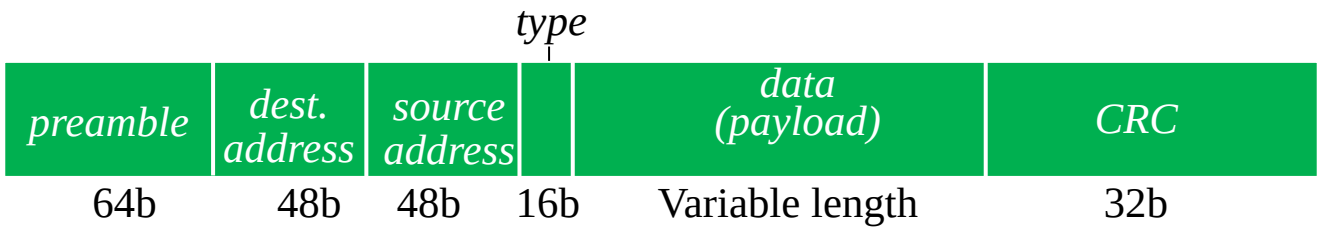
Less bits result in inaccurate clock synchronization and hence more bit errors.

- What does it mean by the sender clocks? What are they clocking and how is synchronization occurring?

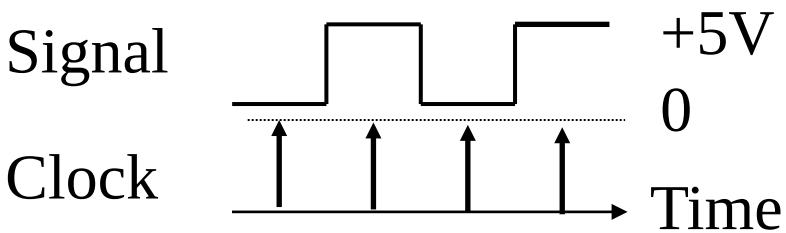
Each bit is sent as a sequence of pulses or as a transition. A clock is used to sample the received signal. If the receiver clock is twice as fast as the sender, the receiver will receive twice as many bits as sent.

©2024 Raj Jain

Ethernet Frame Structure



- ***Preamble:*** 7 bytes with pattern 10101010 followed by one byte with pattern 10101011. To synchronize the receiver, the sender clocks
- ***Addresses:*** 6-byte source, destination MAC addresses
- ***Type:*** indicates higher layer protocol
  - IP: 0x0800
  - ARP: 0x0806



- ***CRC:*** Cyclic Redundancy Check
  - If an error is detected, the frame is silently dropped at the receiver
- ***Connectionless***: No need to ask the receiver
- ***Unreliable***: No ack, nack, or retransmissions

Student Questions

- This seems to imply that the whole packet is received in ethernet, checked if it needs to be dropped, and then given to the host if not dropped. How are buffers managed?

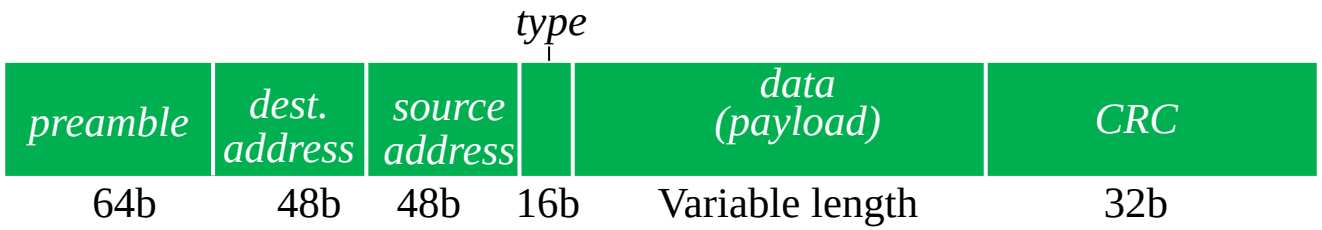
Ethernet chip can store one full Ethernet frame.

- How does a manufacturer guarantee that each device has a unique MAC address?

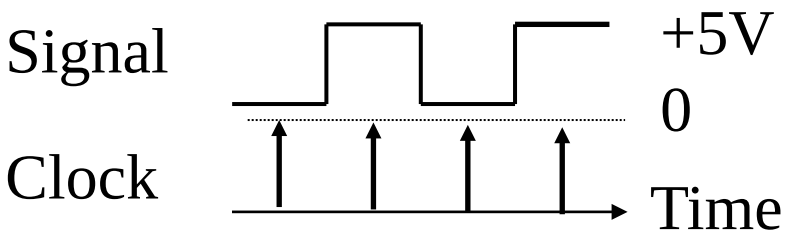
They get their address allocation from IEEE.

©2024 Raj Jain

Ethernet Frame Structure



- ***Preamble:*** 7 bytes with pattern 10101010 followed by one byte with pattern 10101011. To synchronize the receiver, the sender clocks
- ***Addresses:*** 6-byte source, destination MAC addresses
- ***Type:*** indicates higher layer protocol
  - IP: 0x0800
  - ARP: 0x0806



- ***CRC:*** Cyclic Redundancy Check
  - If an error is detected, the frame is silently dropped at the receiver
- ***Connectionless***: No need to ask the receiver
- ***Unreliable***: No ack, nack, or retransmissions

Student Questions

- Why is it Connectionless? Is there no need to ask the receiver? Does it ensure that the message is sent to the correct destination?

Yes. No prior permission/connection. Everything is broadcast.

©2024 Raj Jain

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 two pairs of UTP
- **10BASE2**: 10 Mb/s over thin RG58 coaxial cable (ThinWire), 185 m max segments
- **10BASE-T**: 10 Mb/s over two pairs of UTP
- **100BASE-T4**: 100 Mb/s over four pairs of CAT-3, 4, 5 UTP
- **100BASE-TX**: 100 Mb/s over two 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 the base part, 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.

- When/why did the CAT designations emerge for Ethernet cables? Why do CAT 3, 4, 5, etc., not follow the Ethernet naming standard shown here?

CAT is a physical layer specification. Ethernet is the data link layer.

- What is the most common standard?

10BASE-T, 100BASE-TX, 1000BASE-T

- Do we need to remember the details of each standard, such as what medium they use?

You need to know how to read the names in blue and how many wires that imply.

©2024 Raj Jain

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 two pairs of UTP
- **10BASE2**: 10 Mb/s over thin RG58 coaxial cable (ThinWire), 185 m max segments
- **10BASE-T**: 10 Mb/s over two pairs of UTP
- **100BASE-T4**: 100 Mb/s over four pairs of CAT-3, 4, 5 UTP
- **100BASE-TX**: 100 Mb/s over two pairs of CAT-5 UTP or STP
- **1000BASE-T**: 1 Gbps (Gigabit Ethernet)
- **10GBASE-T**: 10 Gbps
- **40GBASE-T**: 40 Gbps

Student Questions

- What exactly does the third part of the naming mean?

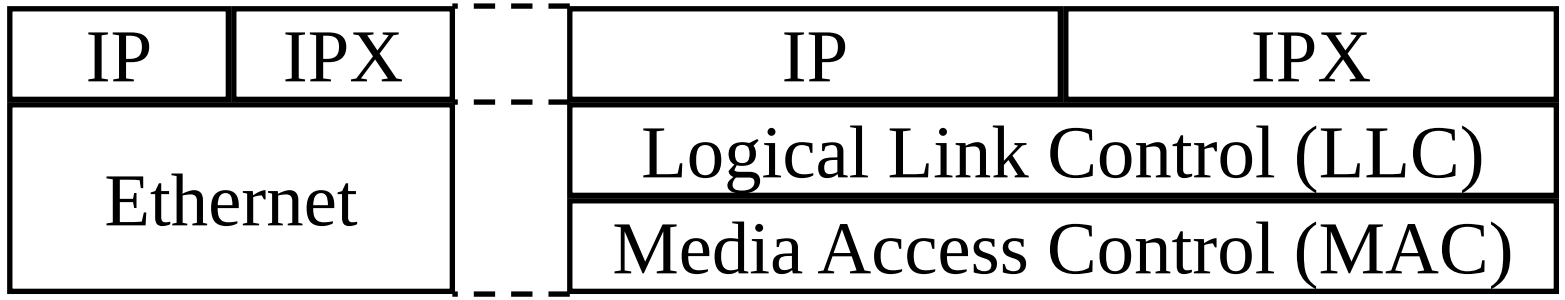
Physical media or distance

- What are you referring to as twisted wire when comparing to 1base5?

Unshielded Twisted pairs were explained in Slides 1.11.

©2024 Raj Jain

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. They are not optional. All 802.3 frames have MAC and LLC. Original Ethernet frames do not have LLC.

- Can you reiterate what MAC is responsible for in the IEEE 802.3 version of Ethernet?

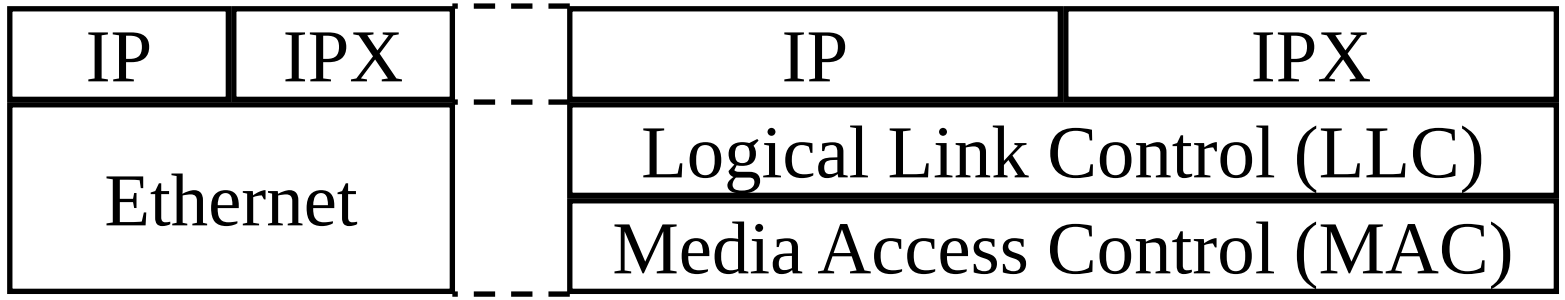
Layer 2 does media sharing receiving and sending frames.

- Why did IEEE deem it important to make these two separate layers?

There were two proposals for LANs: Ethernet and Token Ring. Token ring included flow control. So LLC sublayer was designed for that.

©2024 Raj Jain

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

- The title of one of the slides is 'Ethernet vs. IEEE 802.3' but 802.3 is both a standard OF Ethernet and a predecessor to the original 'Ethernet' and 'Ethernet II' standards correct?

Some people call the original Ethernet as Ethernet II. The original is a predecessor to 802.3 Ethernet.

- And when we work with Lab 6, it is using the older Ethernet standard because of the presence of a 'Type' field?

Yes. Most implementations, even today, use the original Ethernet format.

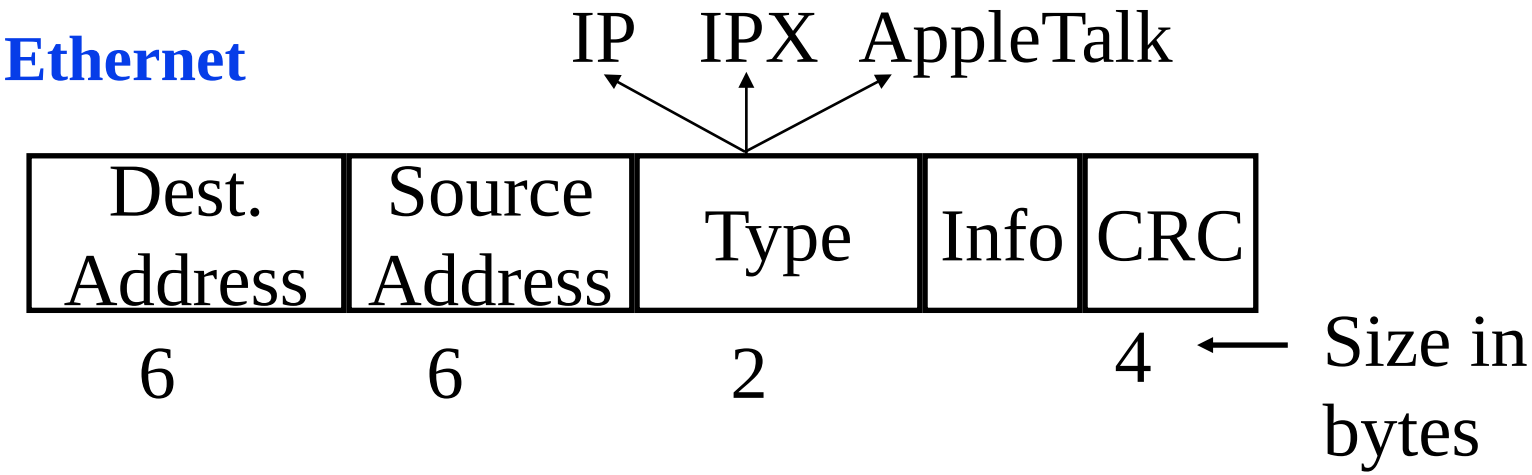
- What does LLC do besides protocol multiplexing?

Flow control, connection setup, …

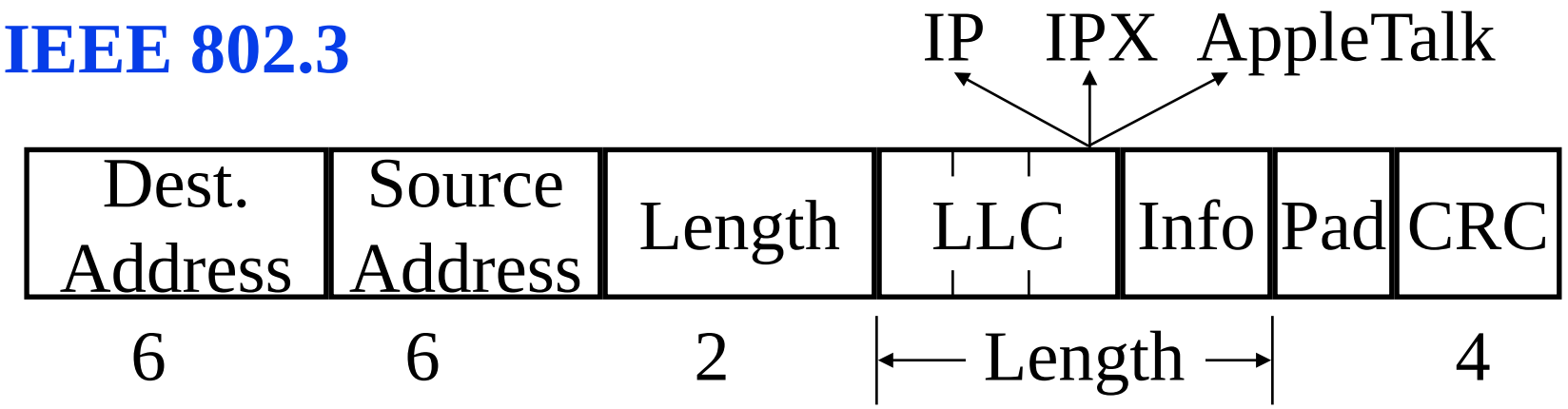
©2024 Raj Jain

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 bytes is the maximum length of Ethernet frames. Eighteen bytes of the 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, do we need to use a layer-2 protocol?

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

- What are IPX and AppleTalk?

Layer 3 protocols that are not IP.

- What is contained in the info field?

Info = payload = Layer 3 PDU

- Why do we need padding for IEEE 802.3

CRC works only if the whole frame is a multiple of 4 bytes.

- Which format is currently in use?

All devices implement both.

- Where does 1518 come from?

1518 was the maximum Ethernet frame size.

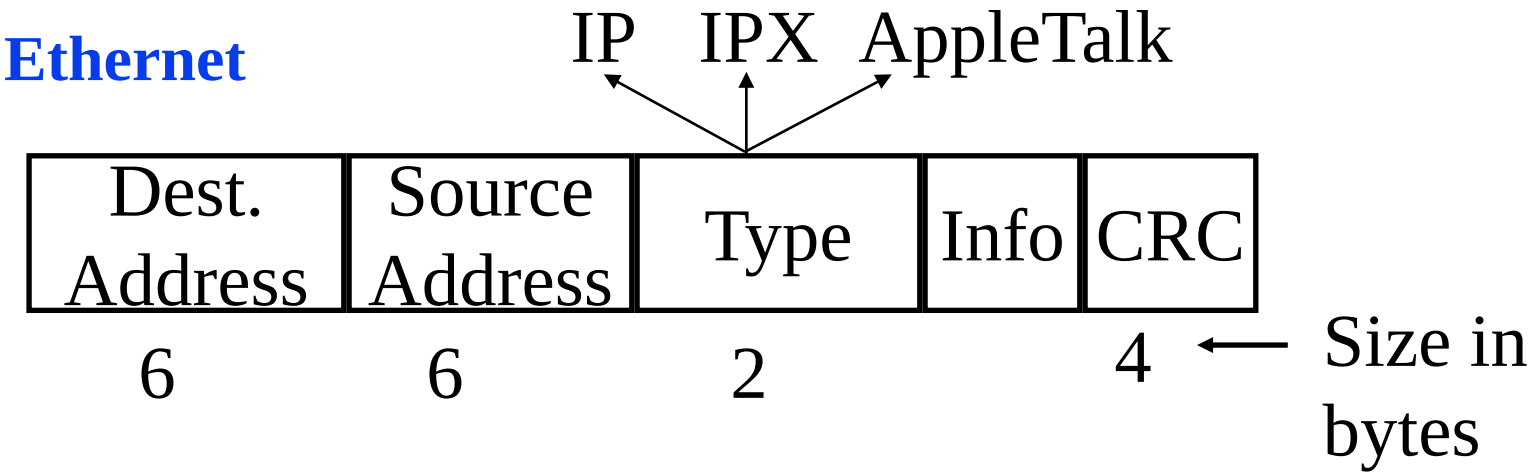
- What is meant by protocol multiplexing?

Allows multiple higher-layer protocols to use this protocol.

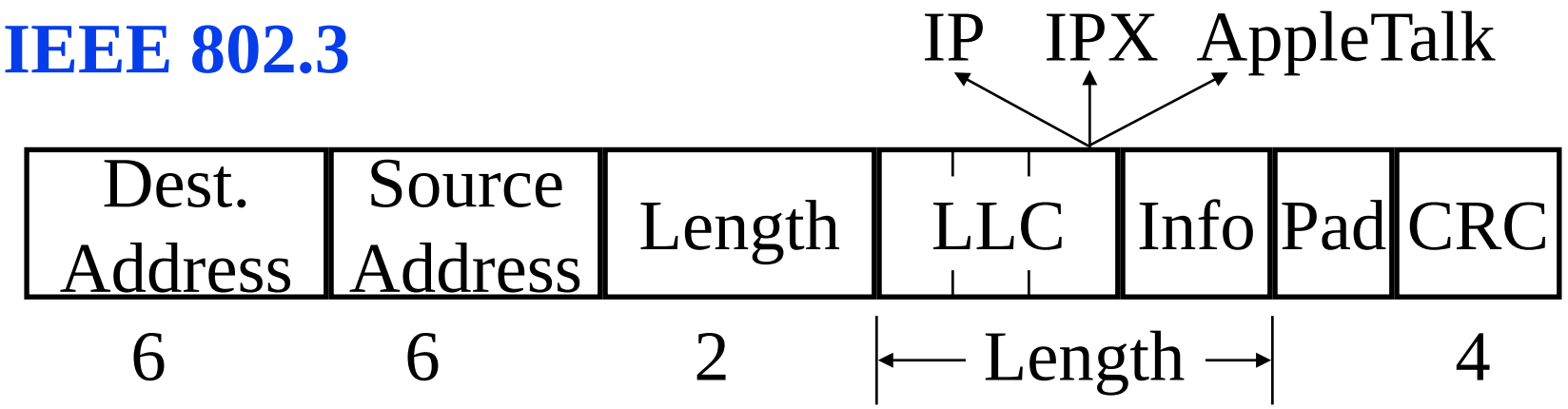
©2024 Raj Jain

Ethernet and 802.3 Frame Formats

- **Ethernet**



- **IEEE 802.3**



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

Student Questions

- What if you have a payload requiring a length larger than 1518?

Ethernet will reject all payloads that will result in frames larger than 1518 bytes. Now there is a version of Ethernet that does not use CSMA/CD. So in those environments, chips allow larger frames. There is no fragmentation or reassembly.

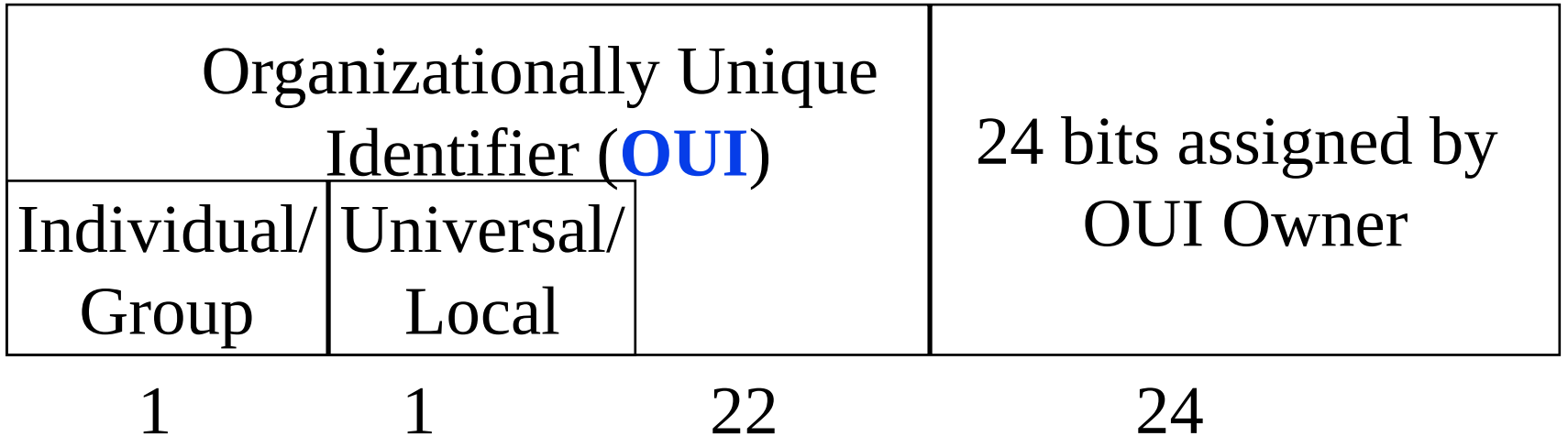
❖ What do you mean by Length > 1518 ⇒ It is a protocol type ⇒ Ethernet?

If that field is more than 1518, it is an original Ethernet format frame.

©2024 Raj Jain

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

Student Questions

- So, each host has its own MAC address, and it does not change wherever the host is, but its IP address reflects the host's local position, right?

Universal MAC address is universal. All IP addresses are locally assigned. They may not indicate "position."

- What do the first two bits of the address stand for? How are they used?

First bit: 1=Group
Second bit: 1=Local

- What's the difference between multicast and broadcast?

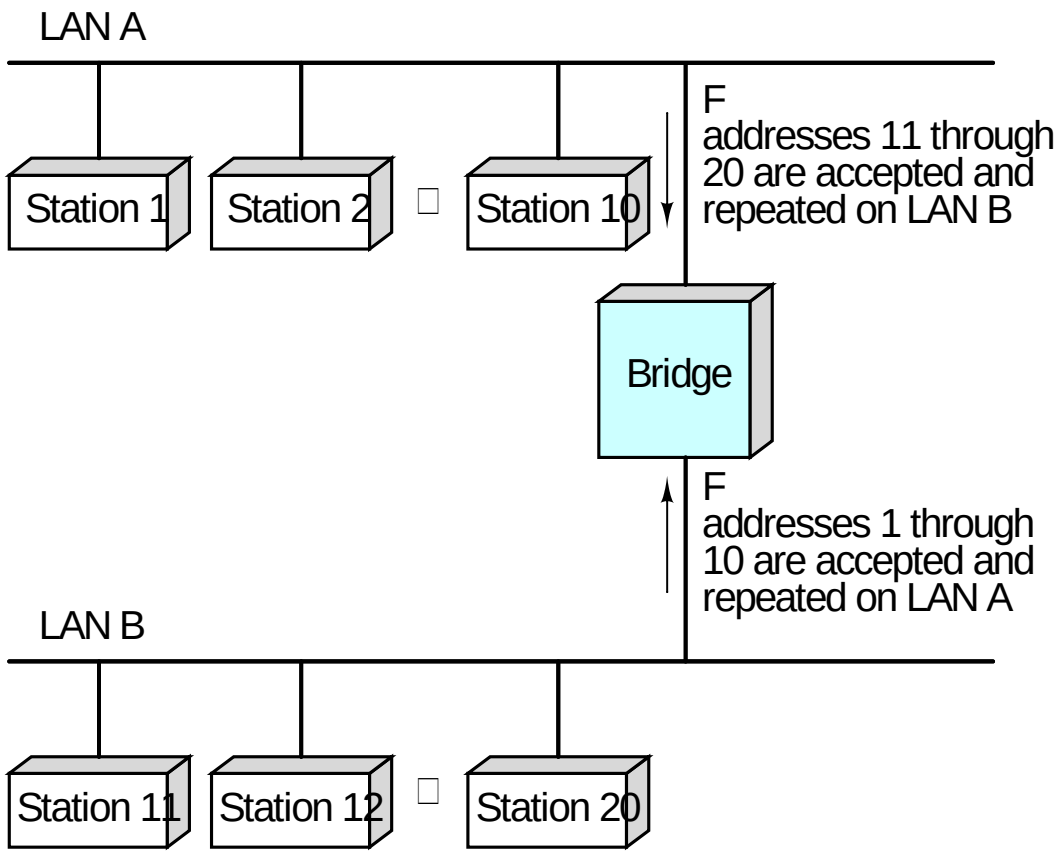
Broadcast = To all
Multicast = To a select group, e.g., all Seniors.

- *What do the "24 bits assigned by OUI owner do"?*

This the address space allocated by IEEE to the chip manufacturer.

©2024 Raj Jain

Bridges



Student Questions

- Should we consider a bridge as 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.*
- What is the difference between a bridge and a switch?

They are synonymous. IEEE exclusive calls the Bridge. Industry exclusively calls the Switch.

- Bridges are the same as link-layer switches? *Yes.*

©2024 Raj Jain

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 on various sides
  Learns by looking at source addresses ⇒ **Self-learning**
- Does not modify the content of the frames
  ⇒ **Transparent**
  May change headers.
- Provides storage for frames to be forwarded
- Improves reliability (fewer 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

- How do switches do flow control?

By giving the source a window or pause. Token rings, now dead, used sophisticated flow control. Ethernet does not. KISS principle.

- Do we need to remember the process by which switches become self-learning?

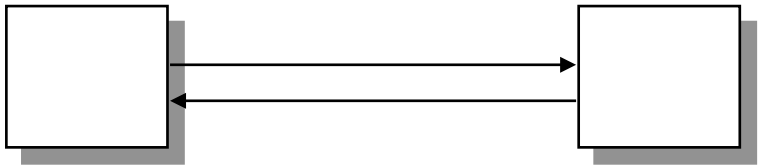
Yes.

- Do bridges (switches) use Content Addressable Memory like routers do to provide fast service between two wires?

Yes.

©2024 Raj Jain

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 is not implemented.
- 10G and higher speed Ethernet standards do not allow CSMA/CD

Student Questions

- What are the roles of a NIC?

The network interface card is either an Ethernet, WiFi, or BlueTooth interface. Some (e.g., Ethernet are visible. Others are invisible)

- Do IoT devices utilize NICs? If so, are the hardware implementations different from those of a traditional computer?

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

- Is it possible to link three or more user nodes mutually?

You could clamp up to 1024 hosts on a single bus. But we no longer use that.

- What do you mean by no collisions?

No CSMA/CD.

- Do you think Ethernet could be more efficient, albeit more trouble to manage, if it used CSMA/CD? Why or why not?

The full duplex is a bit costlier but more reliable.

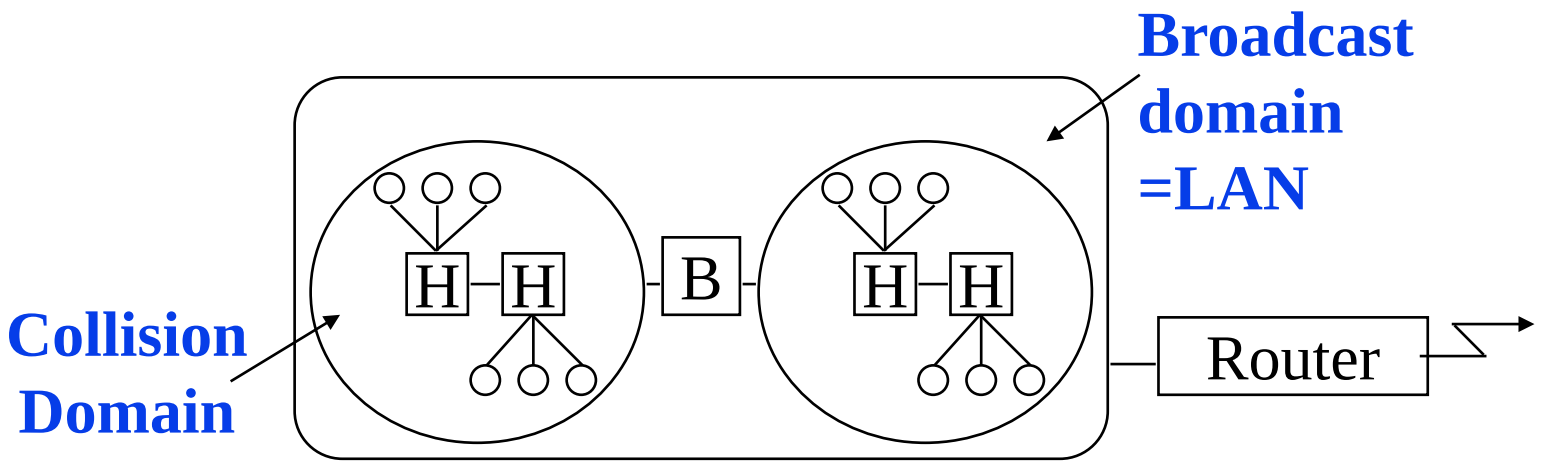
- What does "All vendors are shipping switch/bridge/NICs with full duplex" mean?

Nowadays, all interfaces are full-duplex.

©2024 Raj Jain

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 multiple bridges connect more than 2 Ethernets in a system, how can a bridge know to forward it to the next bridge to reach further away Ethernets?

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

- Are hubs and bridges on the same layer?

Hub is a physical layer device. The bridge is a data link layer device.

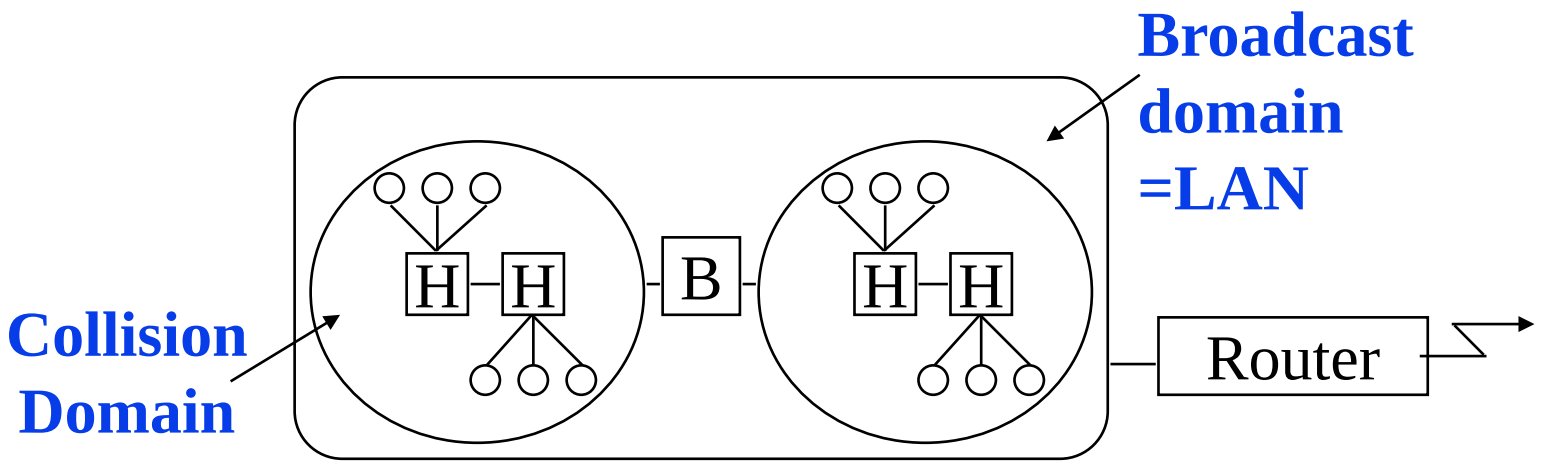
- Is it correct to say, "collision domain may not always collide"?

Two nodes in a collision domain may collide. "May" includes "May not."

©2024 Raj Jain

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

- We have discussed devices that work both as a bridge and a router. Does that mean its collision domain would be the same as the broadcast domain?

A box may have two functions inside. Each function will have its own domain.

- Are hubs not used at all anymore or are there any niche uses for them?

Old hubs can still exist. They still work.

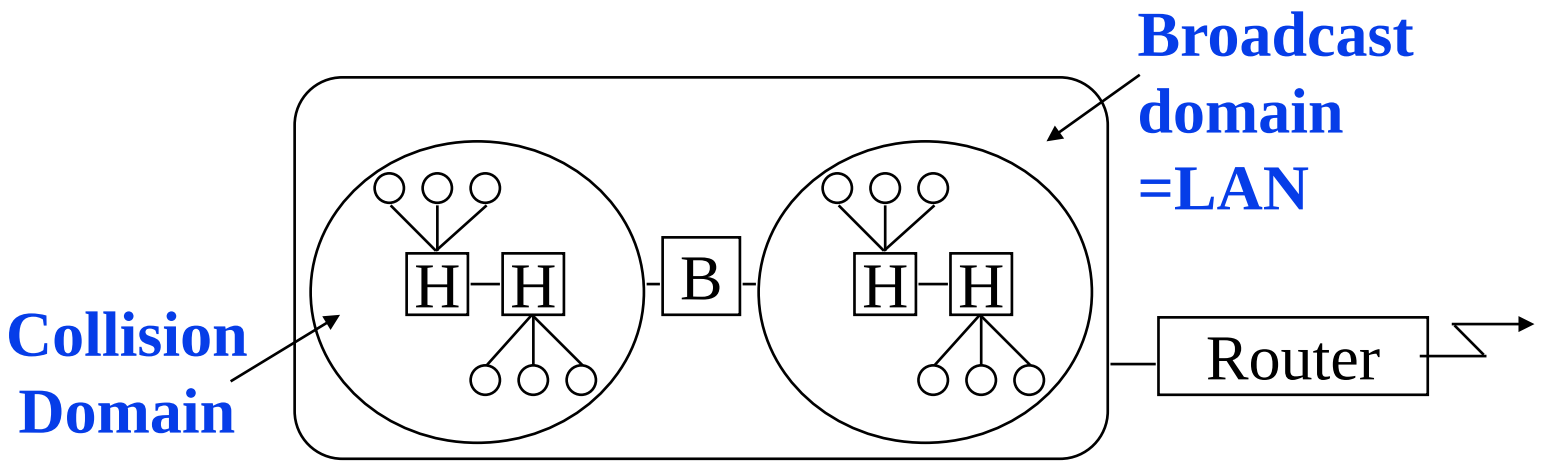
- Are hubs just repeaters with extra features, or are there other differences?

No differences, but multi-port.

©2024 Raj Jain

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

- We have discussed devices that work both as a bridge and a router. Does that mean its collision domain would be the same as the broadcast domain?

A box may have two functions inside. Each function will have its own domain.

- What is a collision domain? And how is it different from broadcast domain?

Two stations in one collision domain will collide if they speak simultaneously. Two stations in one broadcast domain will hear all broadcasts.

- When it does fault detection, how does the hub distinguish between a short circuit and just no activity on the wire.

Electronics circuits can distinguish shorts, open, and idle.

©2024 Raj Jain

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 an "ARP table." You can list the ARP table using the "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?

IP, IPX, and other L3 protocols can use ARP. 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.

- Do we have ARP before every TCP packet in Wireshark?

No. You ARP once and remember the answer until you get errors.

- The video starts to be very unstable from this slide. Is it only on my computer?

Maybe.

©2024 Raj Jain

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 an "ARP table." You can list the ARP table using the "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

- Can there ever be a case where an ARP table is outdated, and the message is sent to the wrong device?

Yes. So, ARP entries are timed out after a while.

- What does the protocol address represent?

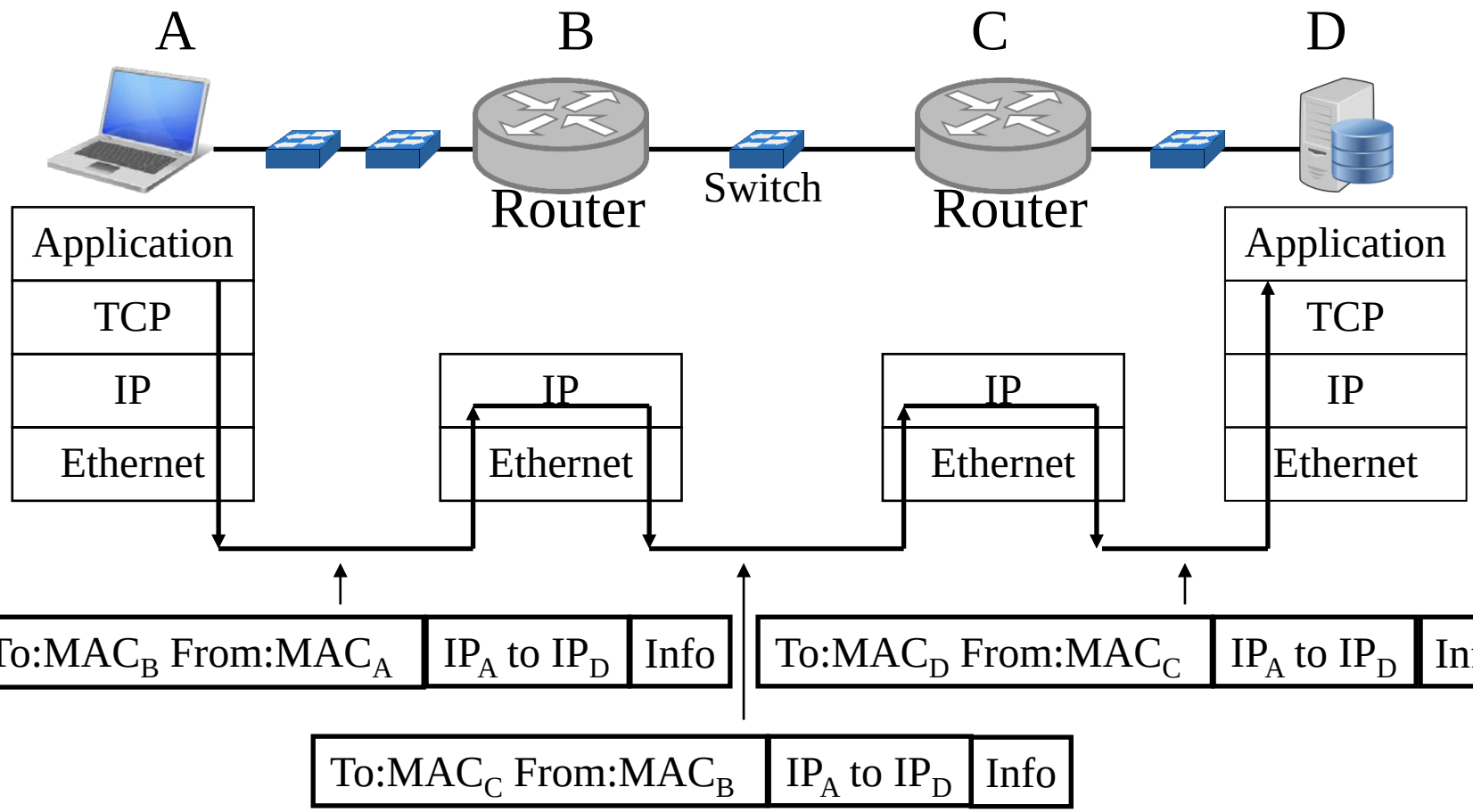
ARP can be used for all L2 protocols. For Ethernet, protocol address = Ethernet Address

- Since a broadcast request is sent to everyone in the same subnet, how does ARP find the MAC of an IP coming from somewhere?

All traffic on the Ethernet already has the source MAC address.

©2024 Raj Jain

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 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. The 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.

- Why is there a quote around the subnet?

For emphasis. Like blue/bold.

- Do switches change Ethernet layer information?

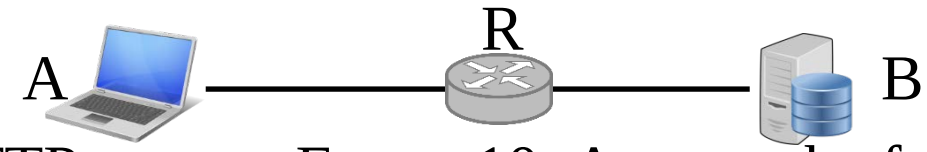
No. Switches are generally transparent. But, VLAN switches do add/remove VLAN fields.

©2024 Raj Jain

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 an 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? To whom does it belong: A, B, or R?
   B. What is the 48-bit Ethernet destination address? To whom does it belong: 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

©2024 Raj Jain

Lab 6 (Cont)

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

2. Examine the HTTP OK response. (Frame **12 … 16**).

A. What is the Ethernet source address? To whom does it belong: A, B, or R?

B. What is the destination address in the Ethernet frame? To whom does it belong: 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 start of the Ethernet frame does the ASCII "O" in "OK" appear in the Ethernet frame? How many bytes are used up in the Ethernet header, IP header, and TCP header before the first byte of the HTTP message?

Student Questions

©2024 Raj Jain

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 are the target MAC and IP addresses in the ARP "question"?

Student Questions

©2024 Raj Jain

Lab 6 (Cont)

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

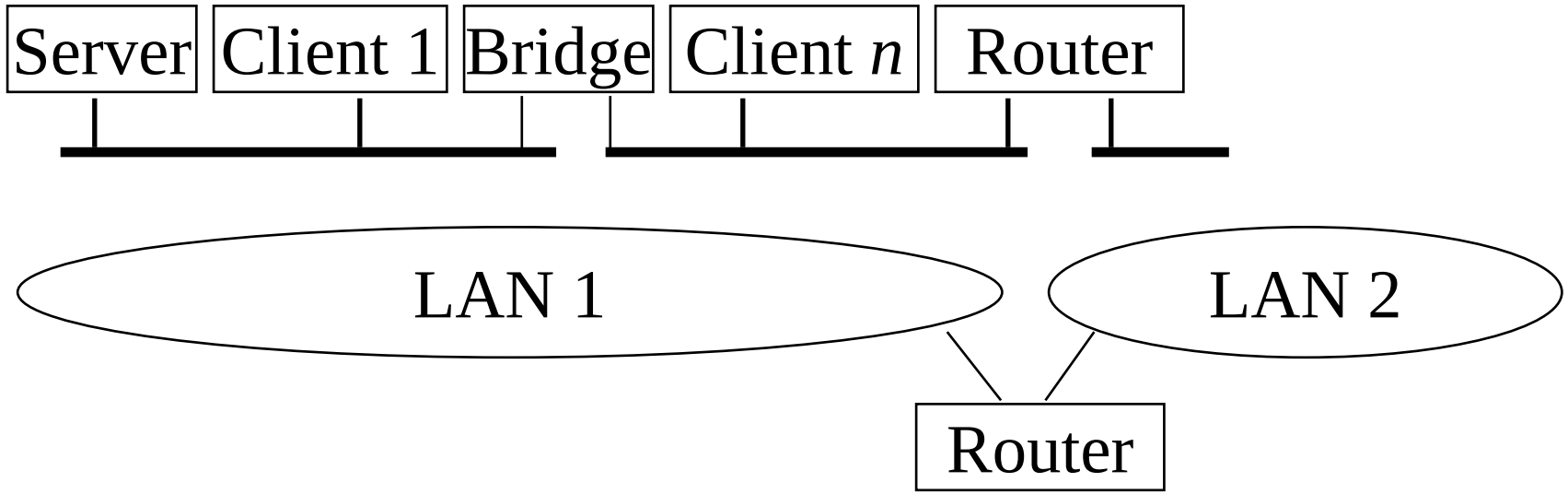
A. What are the hexadecimal values for the source and destination addresses in the Ethernet frame containing the ARP response message?
B. What is the hexadecimal value for the two-byte Frame type field? What upper layer protocol does this correspond to?
C. How many bytes from the very beginning of the 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. There is no need to add screen captures.

Student Questions

©2024 Raj Jain

What is a LAN?



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

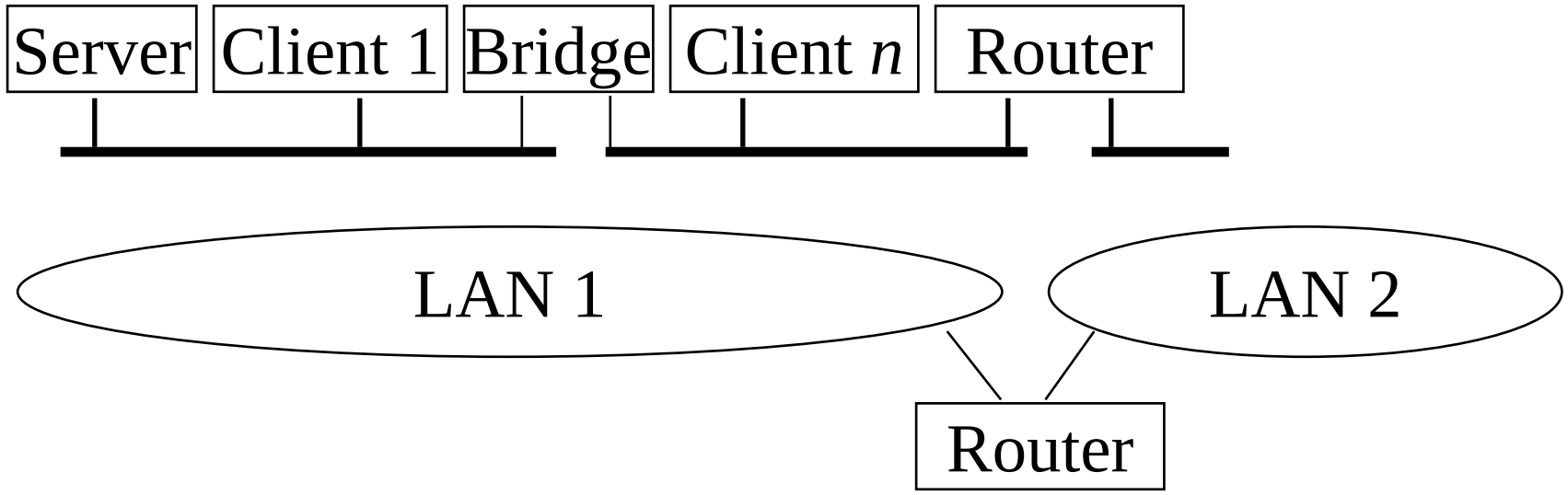
Student Questions

- Can you talk about WAN and its relation to LAN

LAN is one trusted environment.
WAN consists of multiple LANs
=> Multiple trusted environments.
It would be best to have a router connecting 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.

©2024 Raj Jain

What is a LAN?



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

Student Questions

- What does saying "no routing is required between members in a LAN" mean? Is it because the addresses of LAN members are known to each other?

LAN members communicate using the LAN layer. They don't need a router.

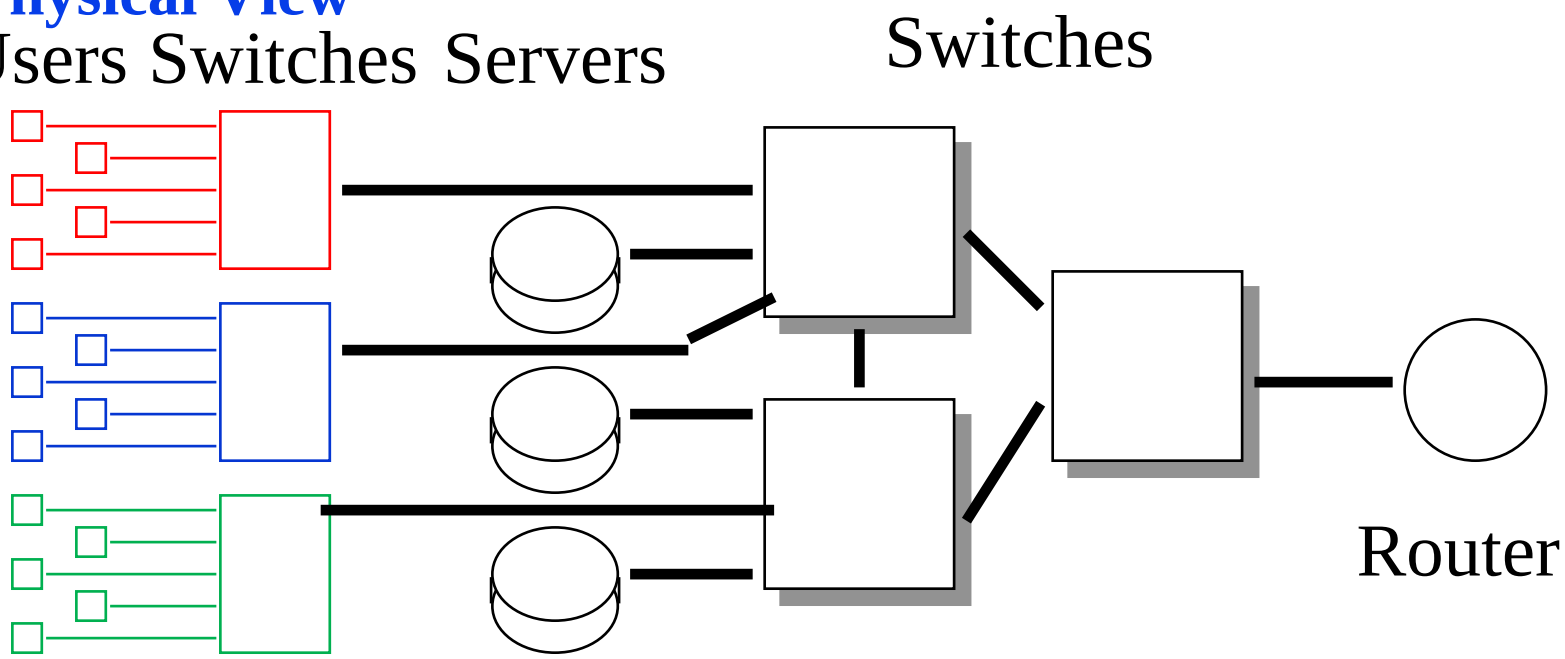
- Does this imply that multicast won't work across LANs?

L2 multicast won't work across two LANs that are connected via a router. You will need a L2 bridge to connect them.

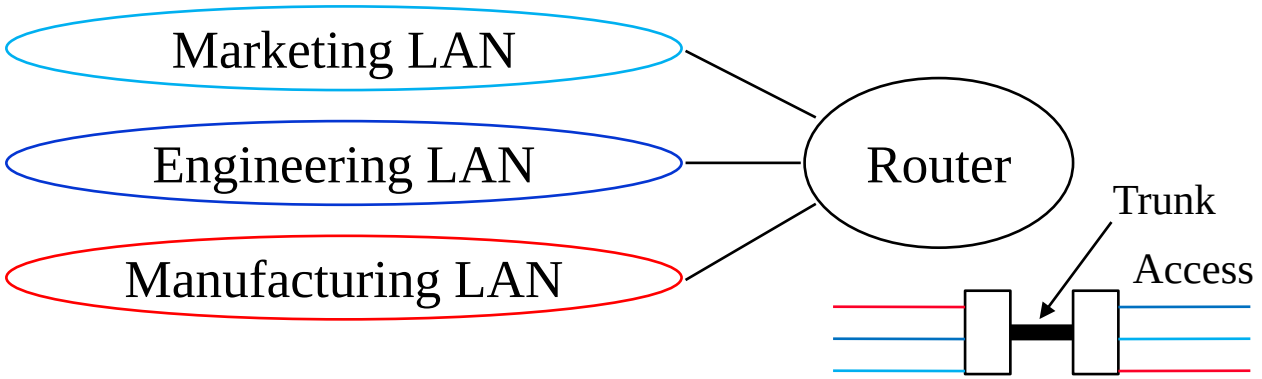
©2024 Raj Jain

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 a trunk port/access port. Could you discuss how these are used for inter-VLAN communication?

The advanced topic for CSE 574S:
Ref: https://www.educba.com/trunk-port-vs-access-port/

- Why do we want to isolate VLANs?

The organization using a VLAN wants its privacy.

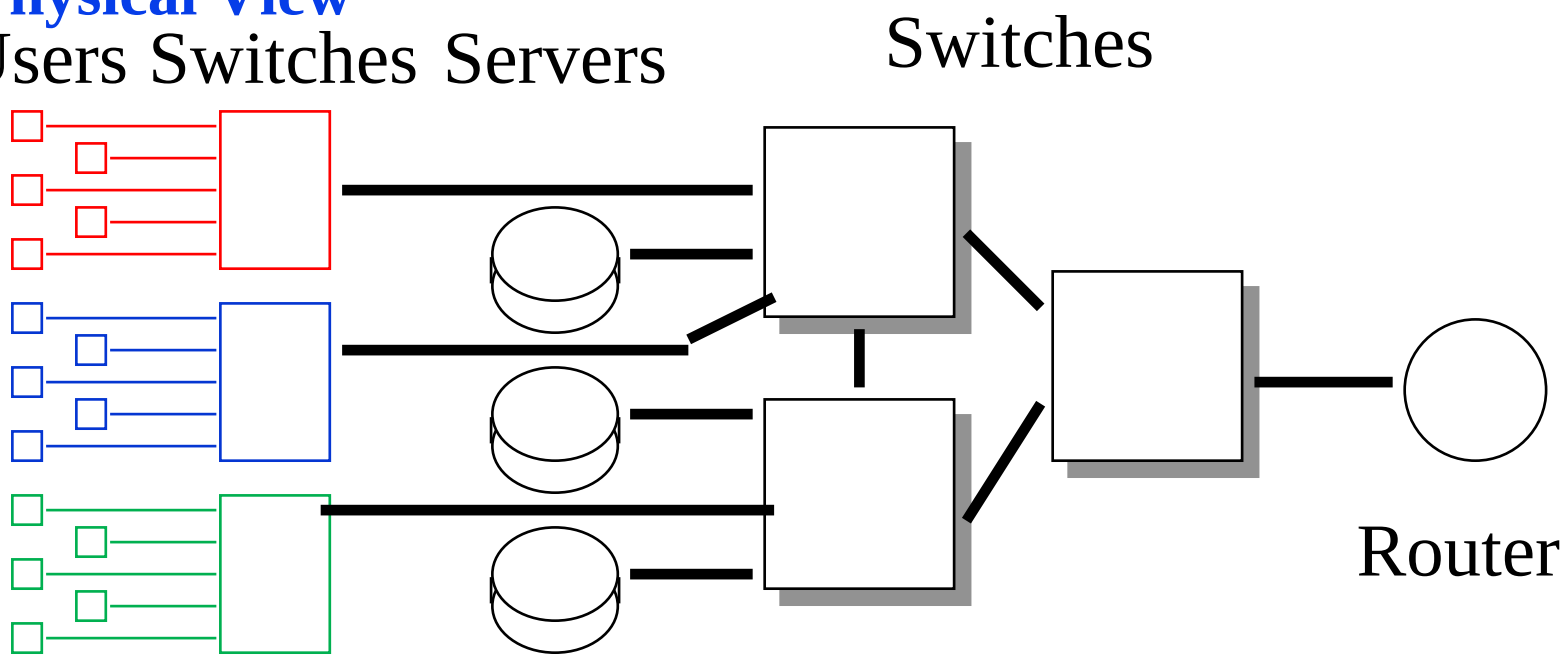
- What is the difference between a physical and a virtual LAN, and is there any difference in their structure? If so, what advantages and disadvantages do both have with each other?

VLANs allow a single network infrastructure to be used by many departments.

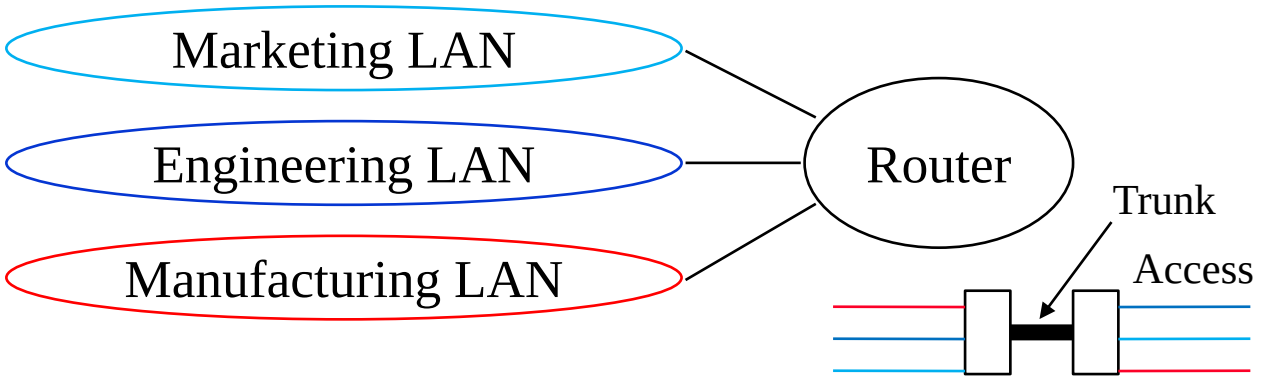
©2024 Raj Jain

What is a Virtual LAN

- **Physical View**



- **Logical View**



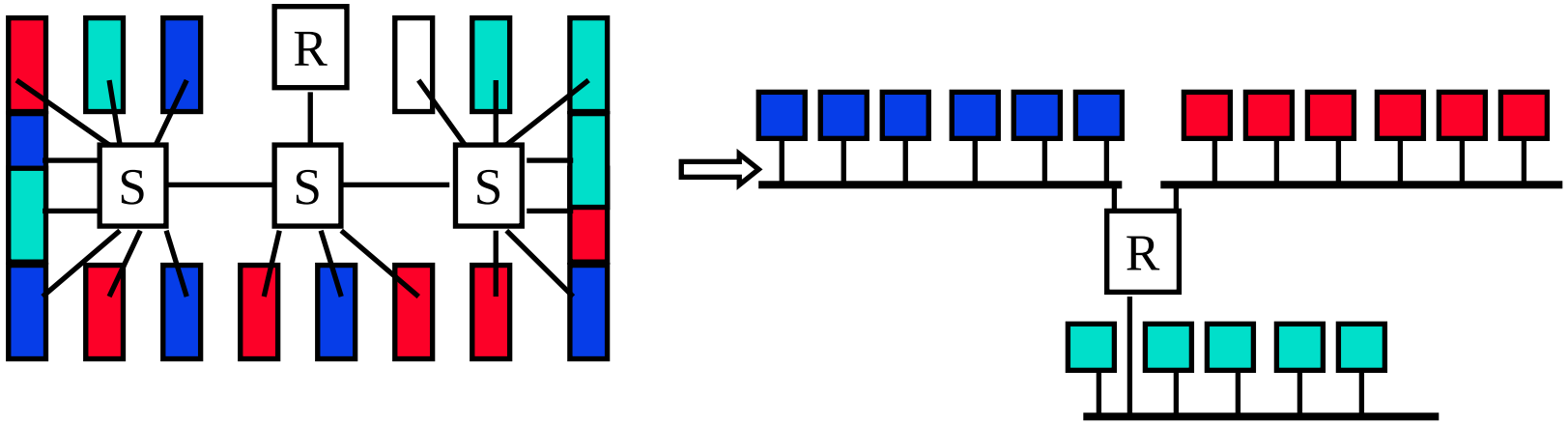
Student Questions

- What is the purpose of virtual LAN?

See above.

©2024 Raj Jain

Virtual LAN



- Virtual LAN = Broadcasts and multicasts go only to the nodes in the virtual LAN
- The network manager defines LAN membership
  ⇒ Virtual

Student Questions

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

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

- When broadcasting to a VLAN, do all messages get broadcast to the entire network, and is the packet selectively sent only at the last link?

All switches may hear all traffic. Intermediate switches know if there are any nodes for a VLAN on one of its ports.

©2024 Raj Jain

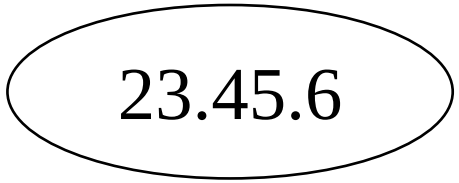
Types of Virtual LANs

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

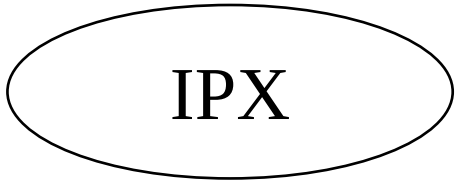
Switch Port	VLAN 1	VLAN 2
A1	√	
A2		√
A3	√	
B1		√
B2	√	

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

VLAN1



VLAN2



Student Questions

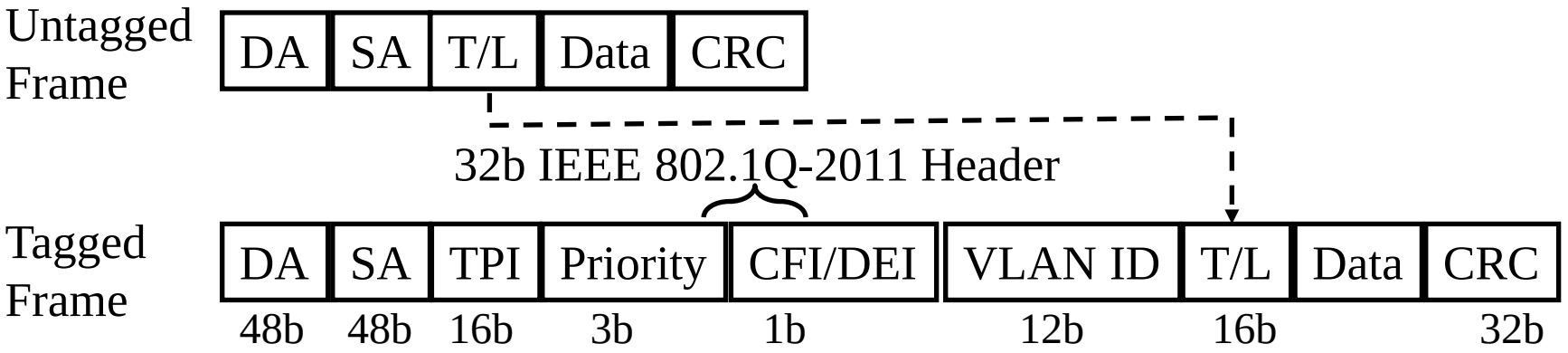
- How are the three VLAN layers related?

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

©2024 Raj Jain

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 is 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

©2024 Raj Jain

Student Questions

- Why are VLAN IDs needed in the link layer if the subnet mask in the network layer takes care of that? *Ethernet switches do not speak IP. The 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.*
- Do switches add these tags?

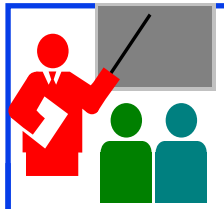
Yes.

- There is a lot of detail here. Are we required to know everything here?

Yes.

- Why the CFI is only one bit? How does it know which non-canonical format it is?

There are only two formats. Canonical (Ethernet) and non-standard (Token ring).



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, and 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. The IEEE 802.1Q tag in Ethernet frames allows a LAN to be divided into multiple VLANs. Broadcasts are limited to each VLAN; you need a router to go from one VLAN to another.

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 this layer in the future?

I am not aware of ML in the Link layer.

- What exactly does self learning imply for Ethernet bridges?

You don't have to manually program which node is on which port. They need to exchange tables like in routing algorithms.

Ref: Section 6.4, Review Questions R9-R16

©2024 Raj Jain

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

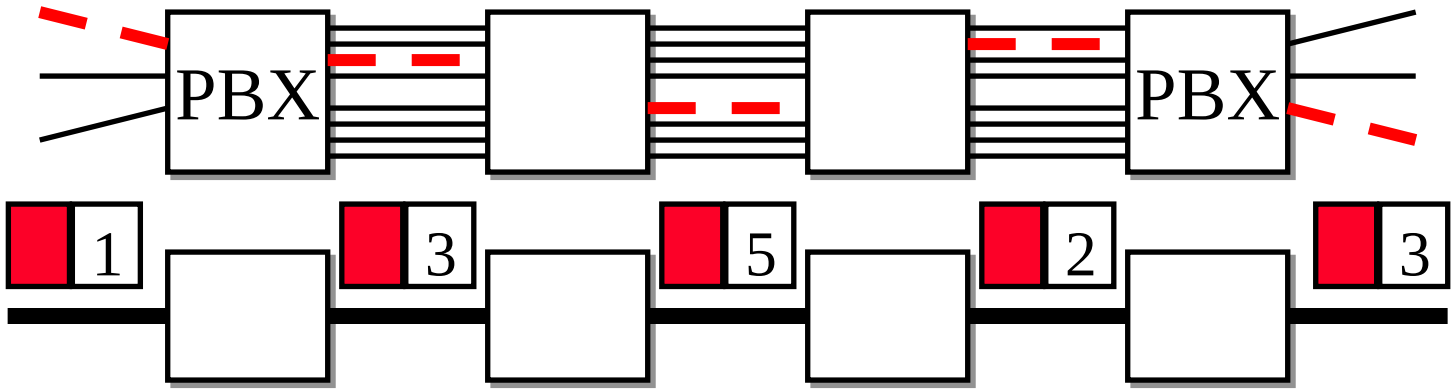
Student Questions

- Which protocols are most commonly used for MPLS?

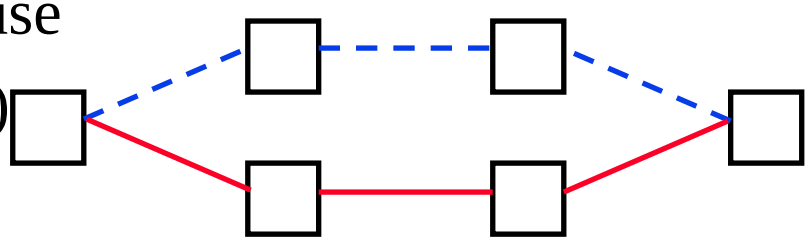
All telecoms use MPLS

©2024 Raj Jain

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)
- LSPs have to be set up before use
- Label-switching routers (LSRs) allows traffic engineering



Student Questions

- How are labels and SDN used in conjunction? As for load-balancing, would an LSR automatically start using some 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.

- Does MLPS reroute possibly cause troubles in the IP routing path calculation?

No. LSRs are designed for this.

- What is the main reason for having MPLS?

Telecoms need to guarantee the quality of service. They can not charge for the best-effort service. MPLS allows them to do that on the Internet.

- Do most routers support it?

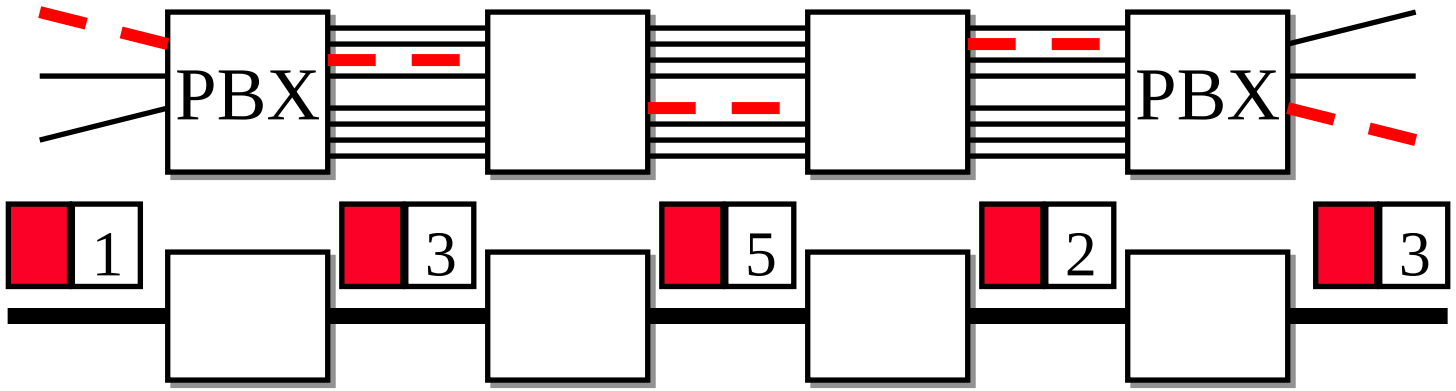
It would be best if you had special routers – LSRs.

- Why do we need the label? Why don't we use the table in layer 3?

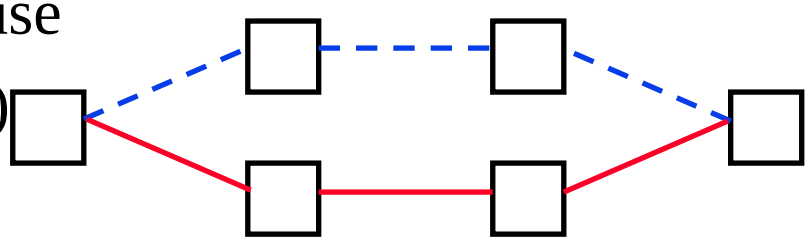
See the 3rd question above.

©2024 Raj Jain

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)
- LSPs have to be set up before use
- Label-switching routers (LSRs) allows traffic engineering



Student Questions

- Apart from the traffic labels, what else makes MPLS quality-assured?

Prior reservation when setting up the labeled paths.

- How does label-switched IP suddenly become reliable?

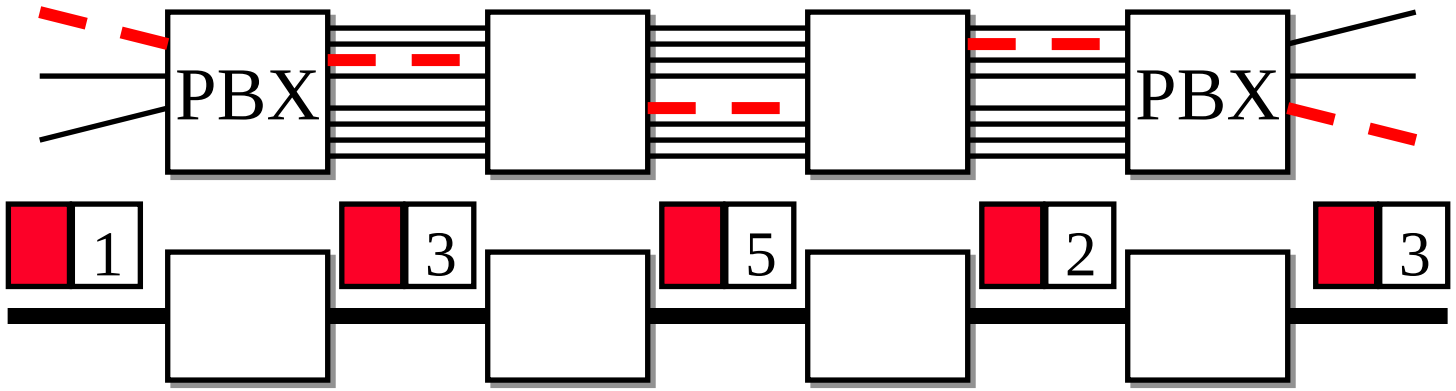
Prior reservation results in lower/negligible loss. UDP-TCP difference becomes smaller.

- So, are LSPs in hardware and LSRs in software, meaning that MPLS is both SDN and hardware?

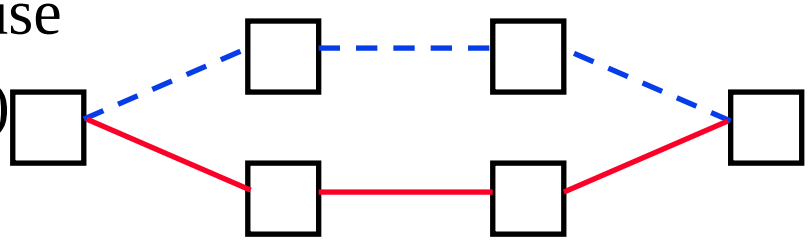
SDN ≠ Software.
SDN = Centralized control.
LSRs can use software/hardware.
LSPs are paths. They are neither software nor hardware.

©2024 Raj Jain

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)
- LSPs have to be set up before use
- Label-switching routers (LSRs) allows traffic engineering



Student Questions

- Can LSPs switch on other header fields or only on the label?

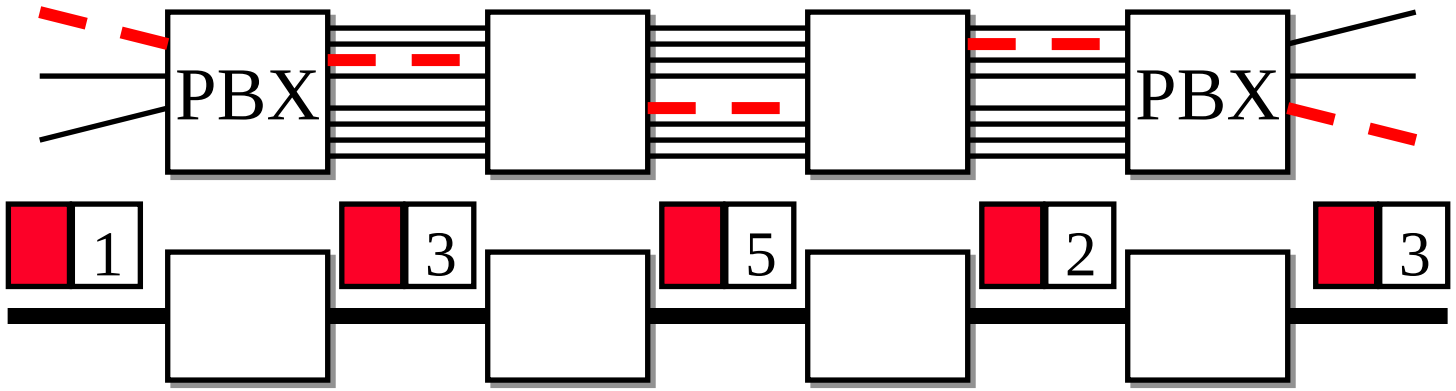
You mean LSRs, not LSPs. Also, you mean route, not switch. Yes, MPLS routers can route non-labeled packets via the IP header, but non-MPLS routers cannot handle labeled packets.

- Can you please explicitly define what a virtual circuit is?

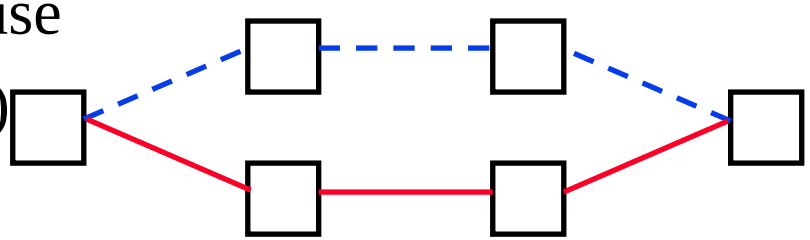
Circuit = A path set up in advance. Physical circuit = hardware switches (or operators) connecting links to make the path.
Virtual = Not physical.
You can create many virtual circuits on one physical circuit on demand.

©2024 Raj Jain

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)
- LSPs have to be set up before use
- Label-switching routers (LSRs) allows traffic engineering

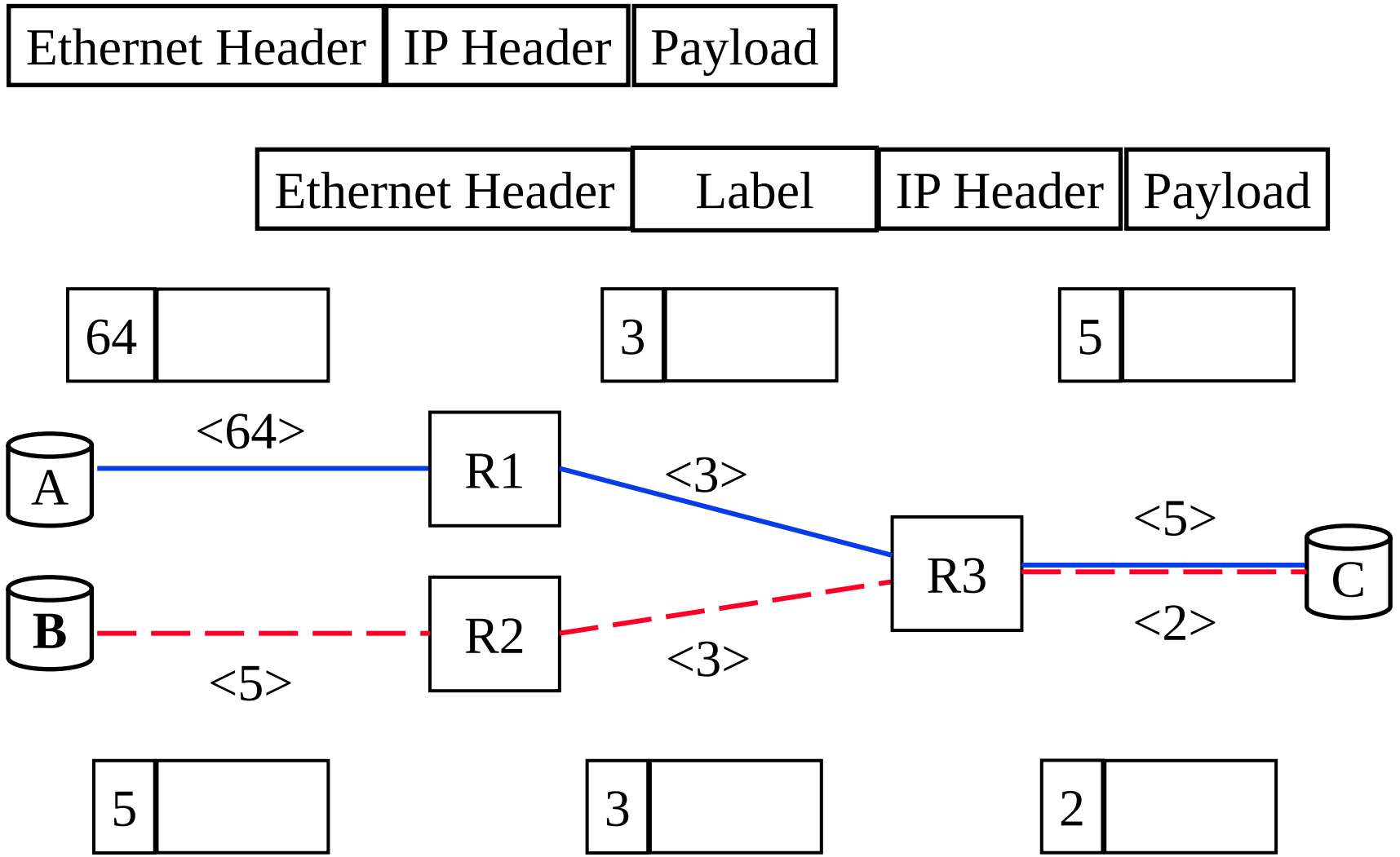


Student Questions

- What hardware is necessary to support virtualization? Is any virtualization done in software? *Hardware is needed to look up label numbers in tables. It can be done in software but is usually done in hardware.*

©2024 Raj Jain

Label Switching Example



Student Questions

- Do the Ethernet Q or MPLS tags use up some of the 1500B used for the L3 payload? *Yes, they do.*
- What is meant by virtual circuits?

VC was defined in the previous slide.

- The ports on the slide you referred to are all physical ports on the router, correct? *Yes*
- How big are the tables for each port on a router typically?

Varies with the size of the routers. You may have several thousand VCs on a port.

- Is layer 2.5 between routers also?

Between special routers - LSRs.

- Are labels attached at the source and removed when switching?

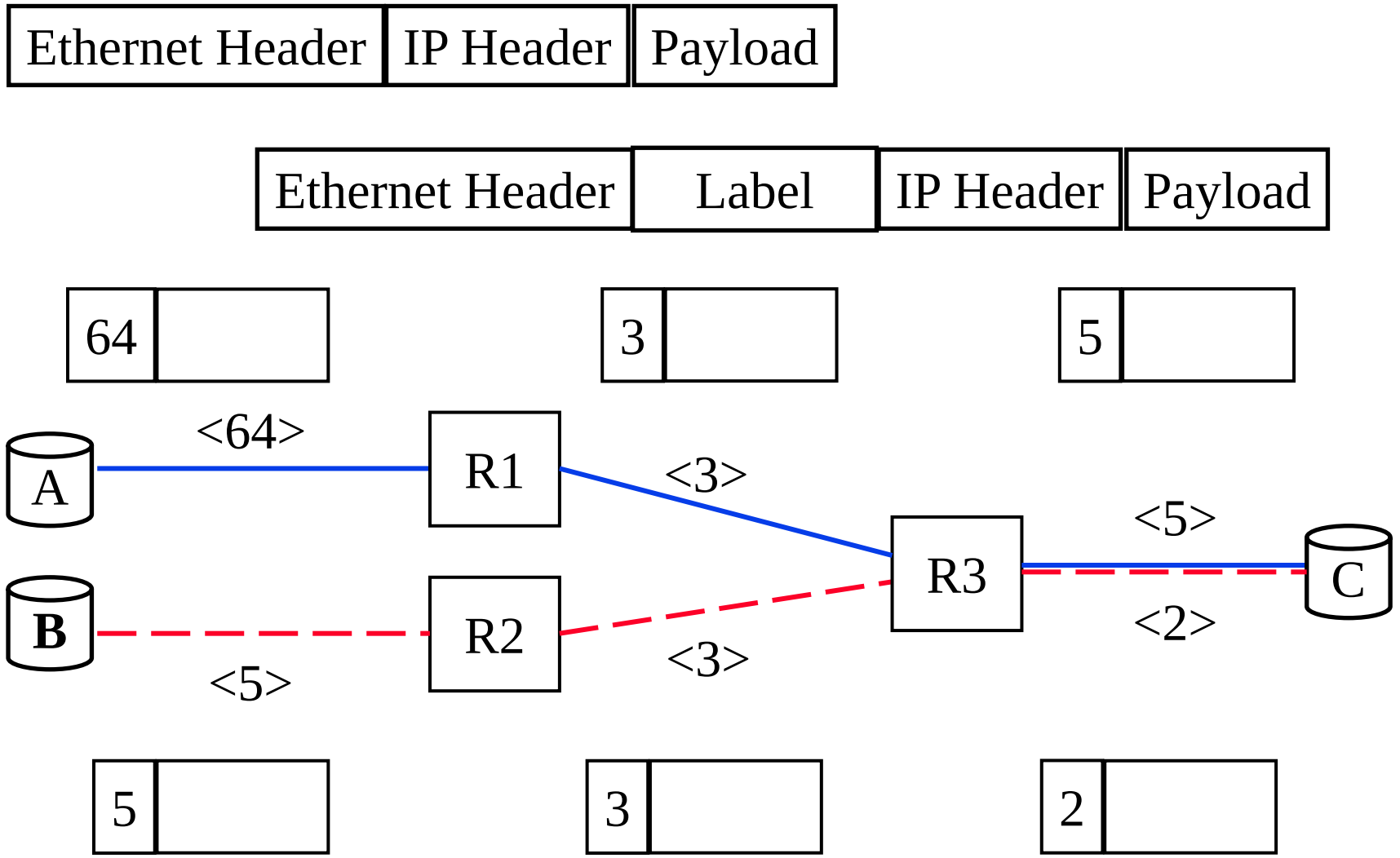
Attached is the first node that understands MPLS. Replaced when switching. Removed at the last MPLS node.

- Can we attach multiple labels to one packet? *Yes. That's an extension.*
- What is the relationship between label and address?

The first LSR looks up the address and attaches a label.

©2024 Raj Jain

Label Switching Example



Student Questions

- How does MPLS simplify the routing compared to traditional IP routing?

Traditional IP routing is simpler than MPLS. However, traditional IP is "best effort" only and unsuitable for telecom carriers that guarantee throughput, etc.

©2024 Raj Jain

MPLS Forwarding Tables

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

R6

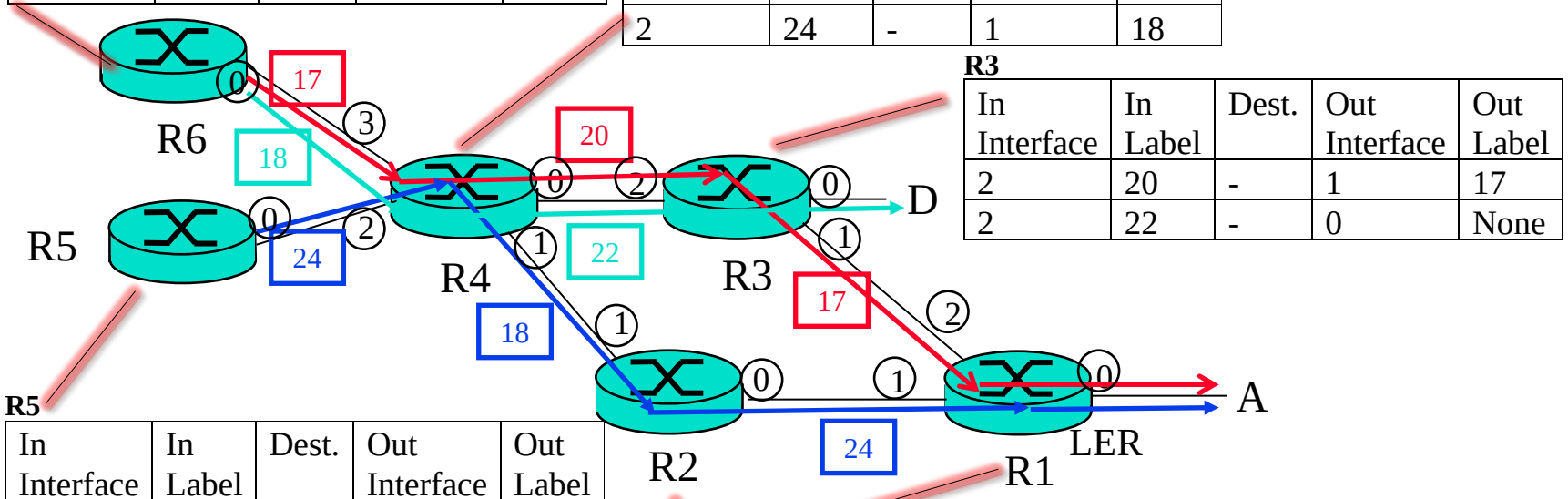
In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	17
-	None	D	0	18

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. A Label indicates a path. An address indicates a destination. Green and red are different paths, although they share several intermediate points.

- Why don't we 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.

- How is the label assigned?

By row numbers in the MPLS table.

- Does the destination have a table?

Not for incoming packets. Maybe for packets going back.

- Is the router LER that lists destinations in the table? For example, R6.

All routers at the edge are LERs. Edge is defined where the next/previous node does not understand MPLS.

©2024 Raj Jain

MPLS Forwarding Tables

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

R6

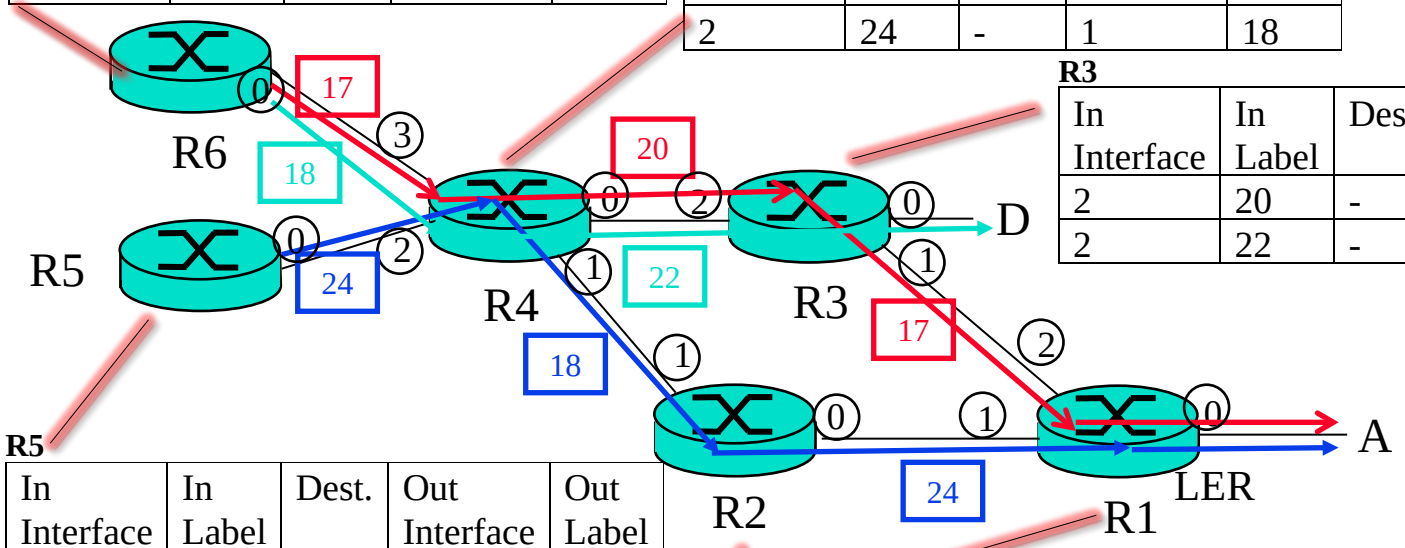
In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	17
-	None	D	0	18

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

- What does the interface number represent?

Interface = Router port.

- How are the terms "interfaces" and "ports" related in MPLS?

Interface = Router port.

- Does the number of ports equal the number of interfaces? *Yes.*
- Are MPLS forwarding tables updated with new routes whenever a host sends traffic to a new IP address?

A new VC is set up every time a host sends traffic to a new host, just like phone calls.

- What would happen if there is congestion or failure on the red path to A? Would it be considered the blue path at R4?

No. Congestion cannot happen to MPLS traffic. All resources are reserved in advance.

- Can you explain the MPLS forwarding table one more time?

Sure.

- What would happen if the interface and label do not match any table entry?

The packet would be dropped. All labels are set up in advance.

©2024 Raj Jain

MPLS Forwarding Tables

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

R6

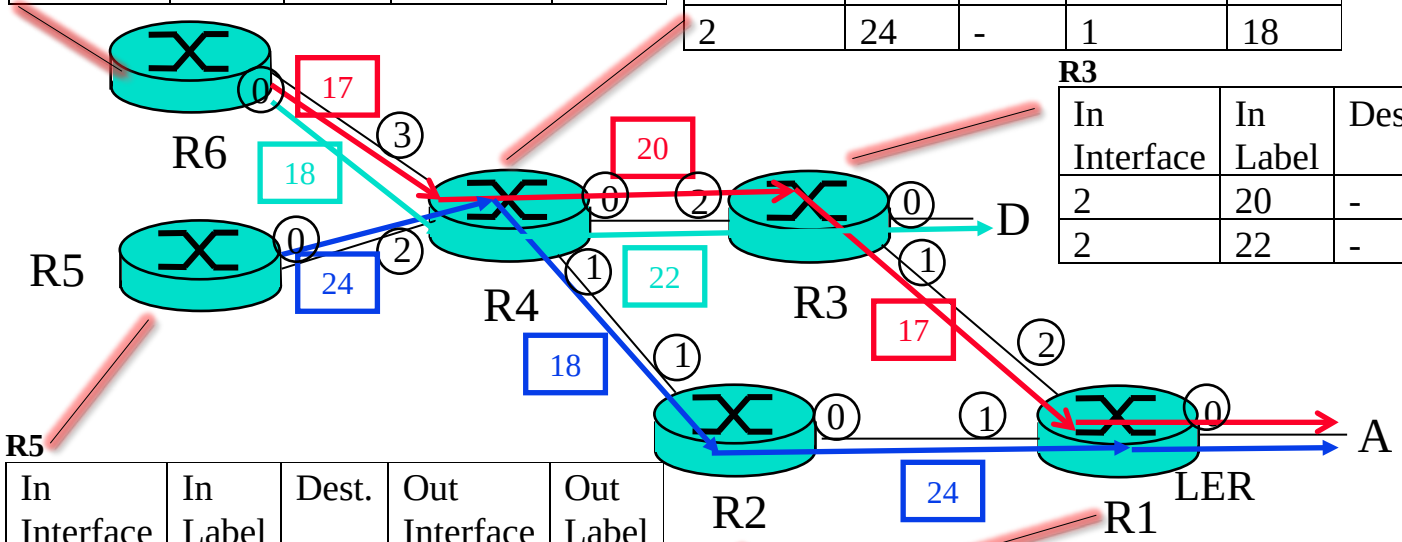
In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	17
-	None	D	0	18

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

- Can you go over LSR and LER again?

LERs are MPLS routers that are on edge. They connect non-MPLS nodes to the MPLS network.

- How are labels determined? On the slide, the numbers go from 17-24 but there's no label 19, for example.

Each label is assigned by the receiving node and communicated to the sending node during circuit setup.

- Are packets that have out label not dropped?

Edge routers put/remove labels on all packets entering/leaving the MPLS domain.

©2024 Raj Jain

MPLS Forwarding Tables

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

R6

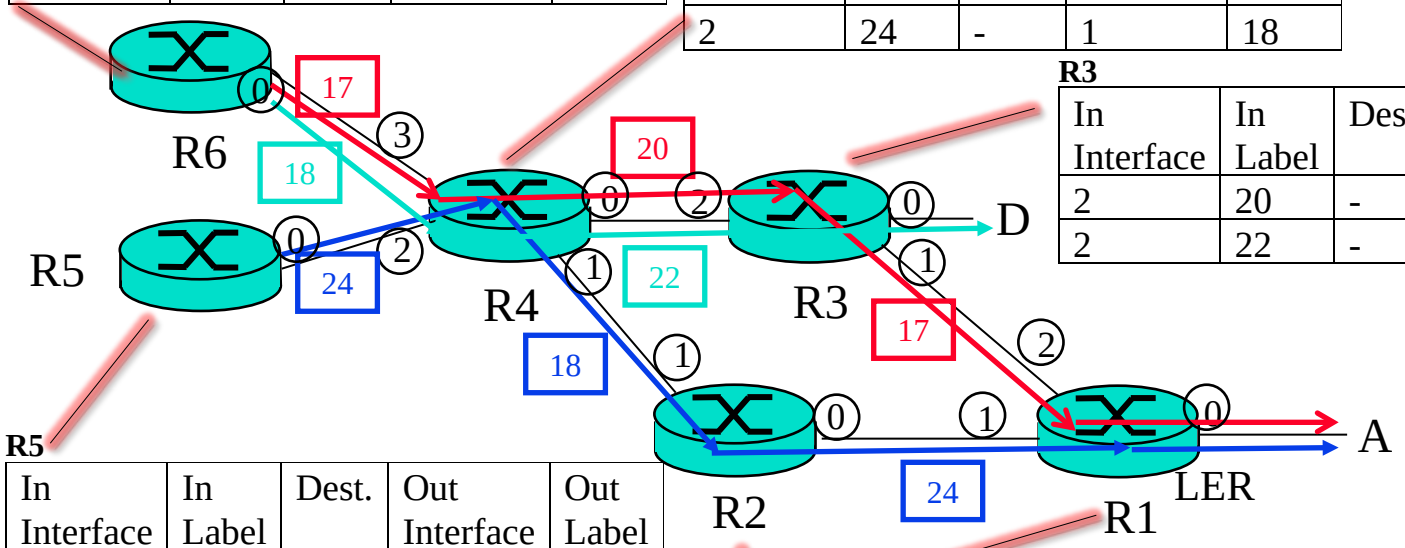
In Interface	In Label	Dest.	Out Interface	Out Label
-	None	A	0	17
-	None	D	0	18

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

- What is happening before and after making the MPLS forwarding tables?

All MPLS traffic needs to follow a fixed path (circuit) set up before the first packet of the flow. The labels are exchanged during setup, and the path may be disconnected when not needed.

- What happens when packets go between MPLS-capable routers and non-capable?

Labels are added or removed.

- How does MPLS handle old circuits and delete outdated entries?

Paths are disconnected when not needed.

Notes

- **Error in the Book:**
  - The tables are per interface, not per router.
  - For compatibility, we have kept the table per router but added the input interface column.
  - The book lists no input interface in the table.
  - The same label #s can 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 is an equal number of reverse circuits that have their own labels unrelated to forward labels.
- Out Label=None ⇒ MPLS Tag is removed.
  In Label=None ⇒ Packet arrives with no MPLS tag.

Student Questions

- Why are the tables of MPLS 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 the next router can 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 added the input interface column to show the rows used. If per-interface tables were kept, that column would not be required, and all rows would refer to the same interface.*

©2024 Raj Jain

MPLS Label Switched Paths (LSPs)

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

Student Questions

- What is meant by bandwidth being reserved? *The sum of the traffic cannot exceed the capacity of airlines.*
- Can the paths be altered later? Or are they permanently set? *The paths can be deleted and reset from the beginning.*
- Would change one path later potentially conflict with other paths?

No. All new paths have to honor previous reservations.

- Why are labels 0-15 reserved?

For internal use.

- Would it be correct to say that each LSP has its own independent label ID set?

No. Labels of two LSPs coming on the same wire cannot be identical, so they are not independent.

©2024 Raj Jain

MPLS Label Switched Paths (LSPs)

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

Student Questions

- Why is indexing an MPLS table faster than searching an IP table?

Think of a similar example. Is looking up "Page 530" faster than looking up "MPLS" in the textbook?

- What are labels 0-15 reserved for?

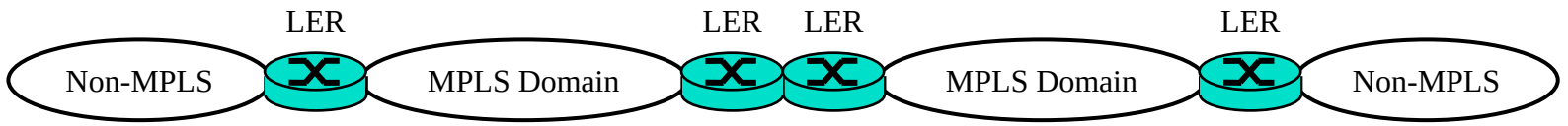
Management and control.

- Do labels being local mean that each time a packet visits a router, it may arrive and leave with a new label?

Each router assigns labels for all packets that it expects to receive. Another router cannot use the info.

©2024 Raj Jain

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 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 and other considerations, such as source address, QoS, etc.
- Other LSRs forward based solely on the label and the frame's incoming interface. They **do not** look at the destination address field.

Student Questions

- Is MPLS used instead of other routing protocols we discussed within MPLS domains? Or does it use a mix?

Other routing protocols may be used to determine the MPLS path.

- Does domain manager mean the MPLS domain?

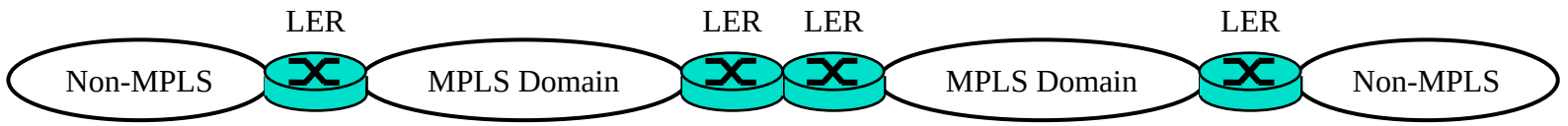
Yes. MPLS domains may represent different phone companies.

- So because of this, could a T-Mobile outage directly affect AT&T customers?

Only those customers that were using part of the T-Mobile network will be affected. AT&T will quickly setup alternate paths for those customers that are affected. So the disruption will be minimal.

©2024 Raj Jain

Label Edge Routers (LERs)



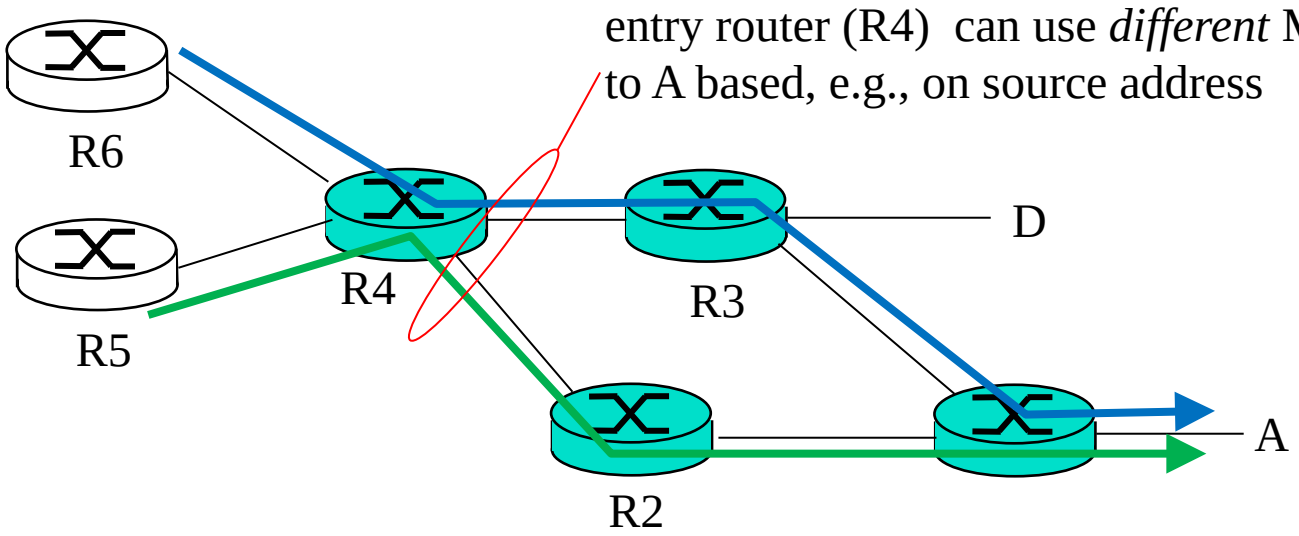
- Routers connected to non-MPLS routers or nodes or routers of other MPLS domains are called Label Edge Routers (**LERs**)
- LERs add labels to frames from non-MPLS nodes or remove their labels if forwarding to non-MPLS nodes or other domains.
- The labels added by LERs **may be** based on destination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

Student Questions

- What are the Non-MPLS and MPLS domains?

Non-MPLS domains do not use MPLS. They use regular IP datagrams.

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

IP-only router

MPLS and IP router

Student Questions

- In the 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.

- Why is it beneficial to base routing on more than the destination address?

Guaranteed quality of service

- What advantages do interior routing protocols have on MPLS, then?

MPLS provides guaranteed QoS.

- Why isn't this used instead of other routing protocols if it has these additional features?

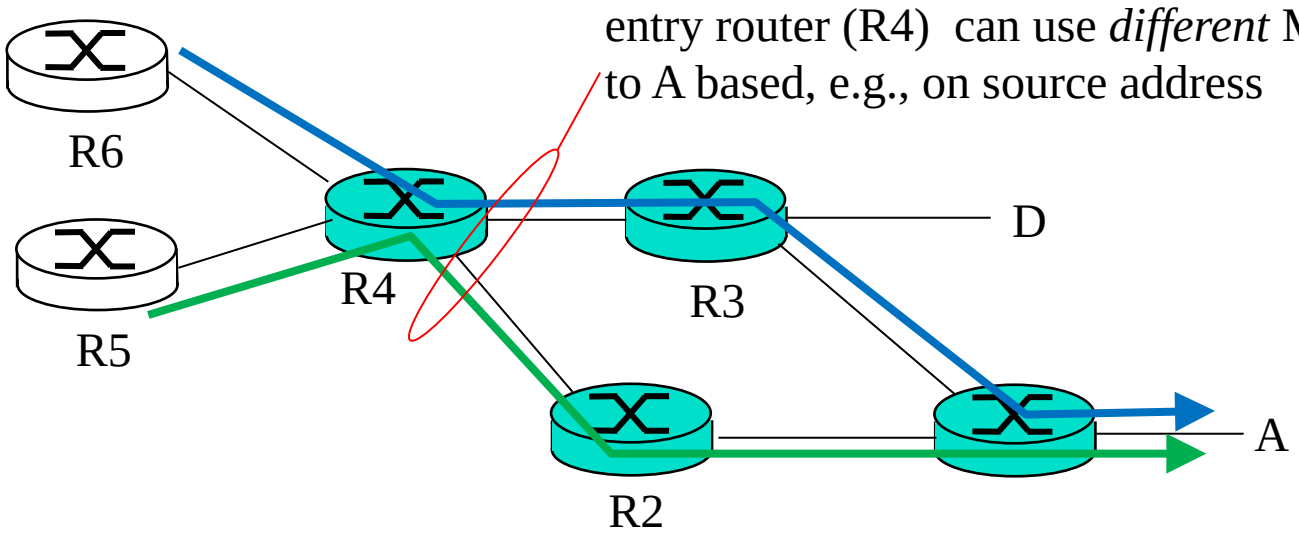
It is extensively used in telecom networks.

- If LSPs have to be set up beforehand, do the backup routes for fast rerouting also need to be preordained?

Yes, if required.

©2024 Raj Jain

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

IP-only router

MPLS and IP router

Student Questions

- Why would multipath IP make things arrive out of order?

Different delays for each packet.

- Does MPLS routing work as an implementation or a replacement for IP routing?

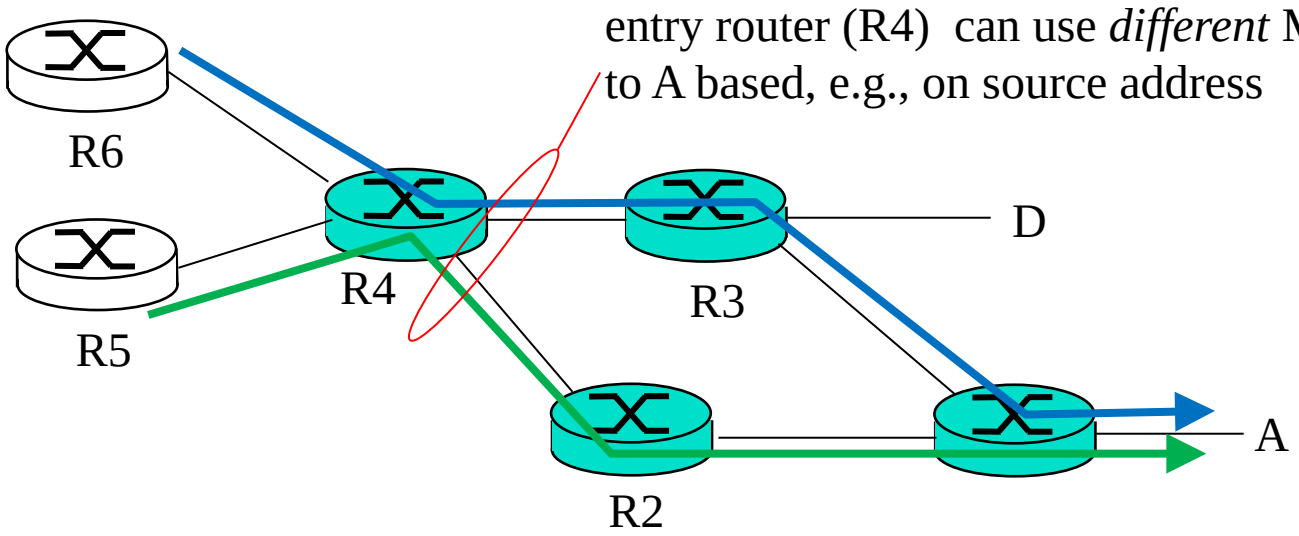
Both. You can use it as an "enhanced implementation" or replacement.

- MPLS appears to be more versatile than IP. Is this true, and if so, why would IP be used?

MPLS is used in carrier networks where guarantees justify the cost. IP is used in enterprise and home networks where cost is not justified.

©2024 Raj Jain

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

IP-only router

MPLS and IP router

Student Questions

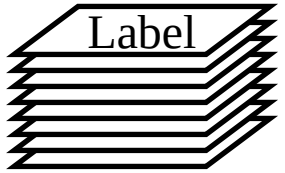
- Is MPLS Routing similar to the Generalized Forwarding because their paths depend on multiple criteria, not only the IP addresses.

What do you mean by generalized forwarding? Does it require IP addresses? MPLS does not.

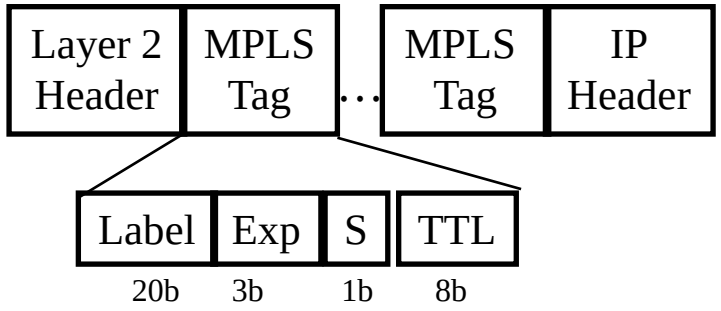
©2024 Raj Jain

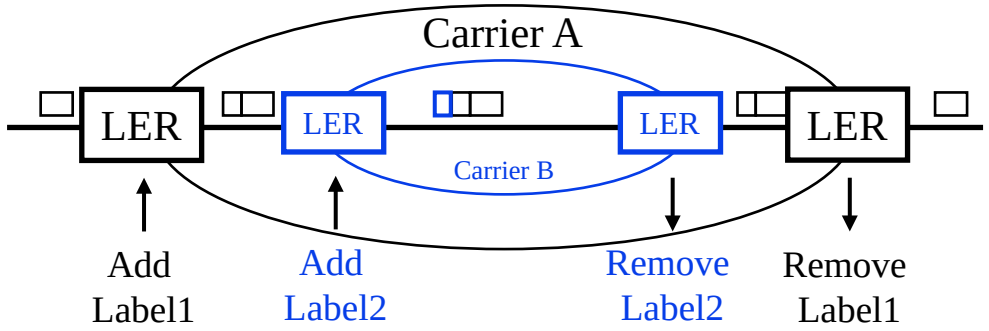
MPLS Label Format

- MPLS label is inserted after the layer-2 header but before the 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 a label, the TTL field from the IP header is copied to the MPLS tag.
  When removing a label, the TTL field from the MPLS tag is copied to the IP Header.
- MPLS Signaling:
  - OSPF has been extended to help prepare label tables
  - There are several other "*Label Distribution Protocols*"





Student Questions

- So, VPNs are a form of overlay network?

In a sense, yes. However, most people do not call it that.

We generally use overlay in L3 rather than L2.

- What is a class of service?

Service quality – lower delay, higher throughput, etc.

- Is it correct that LSRs only change the label but do not add or remove them?

All routers that understand MPLS are LSRs. The LSRs at the edge add/remove labels. Those not at the edge only replace them.

- Why must we keep the TTL field in the IP header after removing the label?

To keep track of how many hops the packet has traveled.

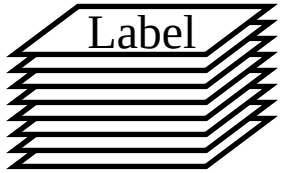
- How is TTL decremented when the packet is forwarded between LSRs within an MPLS domain (since only the LERs decrement when they forward)?

All LSRs decrement the TTL field. LERs copy the IP TTL field to the MPLS TTL field.

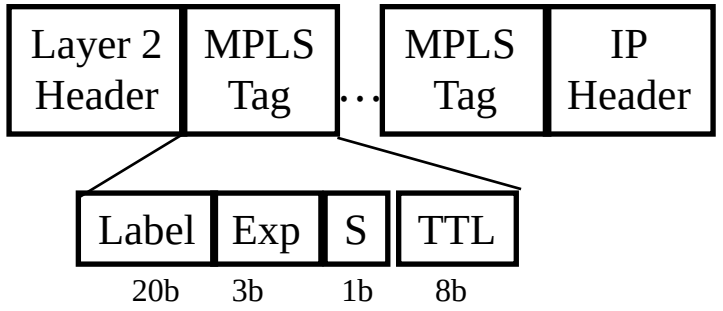
©2024 Raj Jain

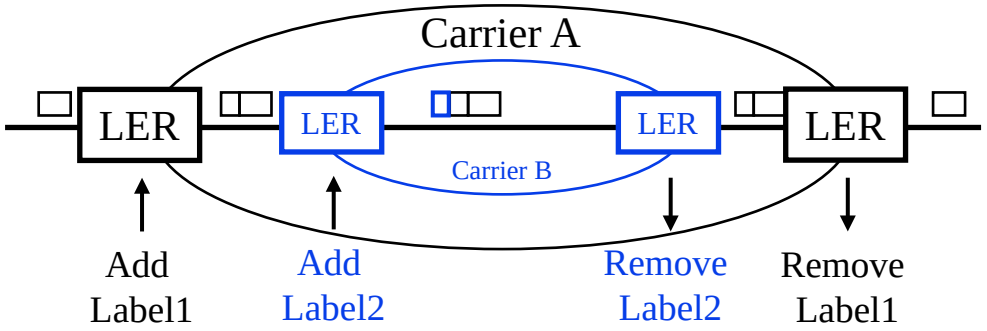
MPLS Label Format

- MPLS label is inserted after the layer-2 header but before the 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 a label, the TTL field from the IP header is copied to the MPLS tag.
  When removing a label, the TTL field from the MPLS tag is copied to the IP Header.
- MPLS Signaling:
  - OSPF has been extended to help prepare label tables
  - There are several other "*Label Distribution Protocols*"





Student Questions

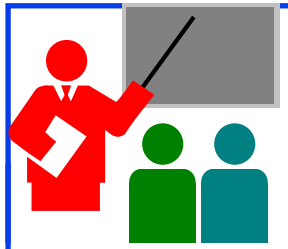
- What exactly are the experimental bits used for?

No standard. One example is the "Class of service."

- How does carrier nesting work? Does the router uses a upper-layer protocol field somewhere in the header to identify that the packet has MPLS header instead of standard IP header?

Different payload type fields in the L2 (Ethernet) header distinguish MPLS vs IP. A bit in the MPLS header distinguishes MPLS vs IP inside.

©2024 Raj Jain



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 the input interface and the label
4. The label table is prepared using 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 on one link.

- Is there a limit on how many MPLS tags can be stacked?

No.

©2024 Raj Jain

Homework 6C: MPLS

- [6 points] Consider the MPLS network in the "MPLS Forwarding Tables" slide. Suppose 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

©2024 Raj Jain

Homework 6C (Cont)

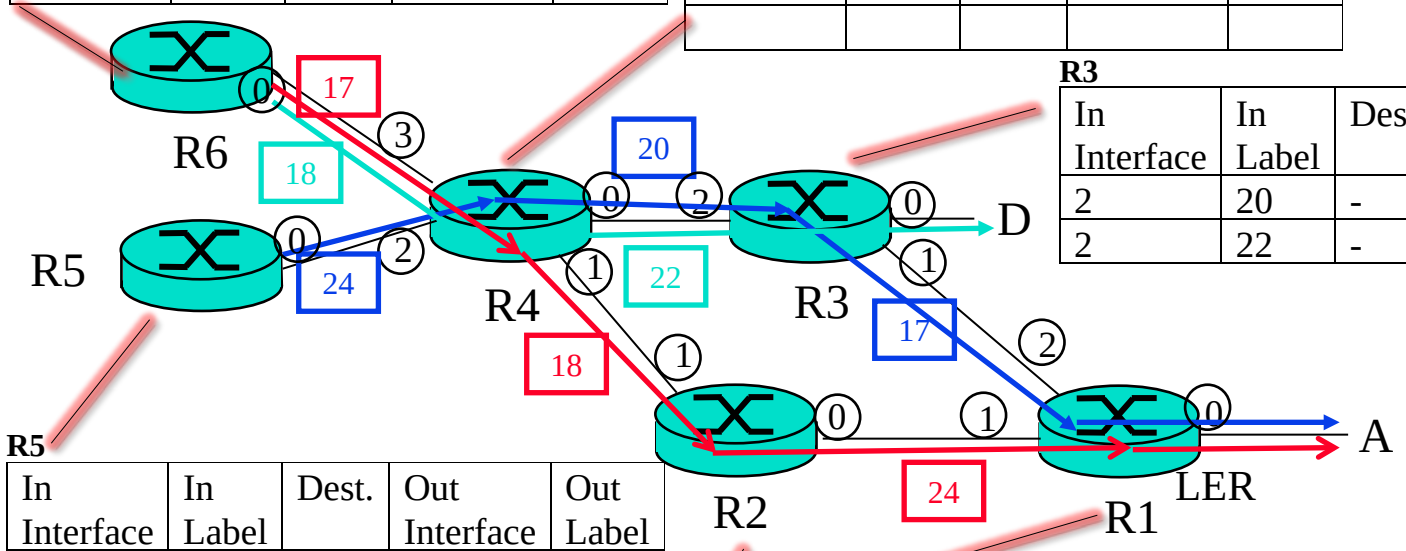
- Interface numbers are in circles. Label IDs are in rectangles.

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

©2024 Raj Jain

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

©2024 Raj Jain

Homework 6D (Cont)

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

R6

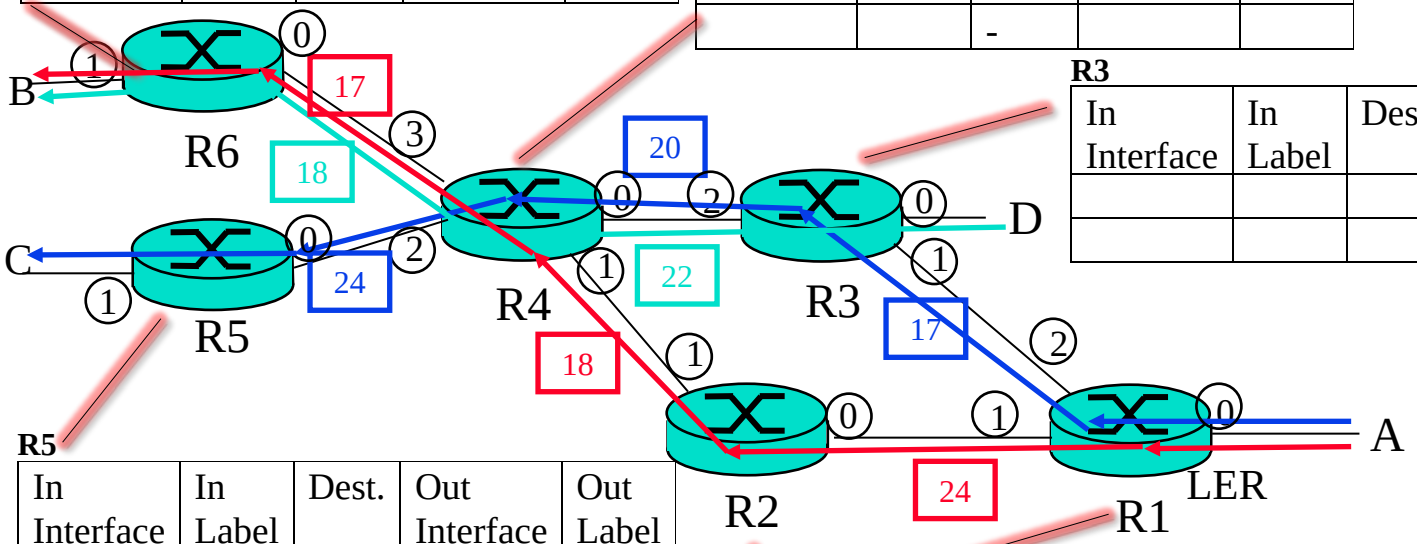
In Interface	In Label	Dest.	Out Interface	Out Label

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

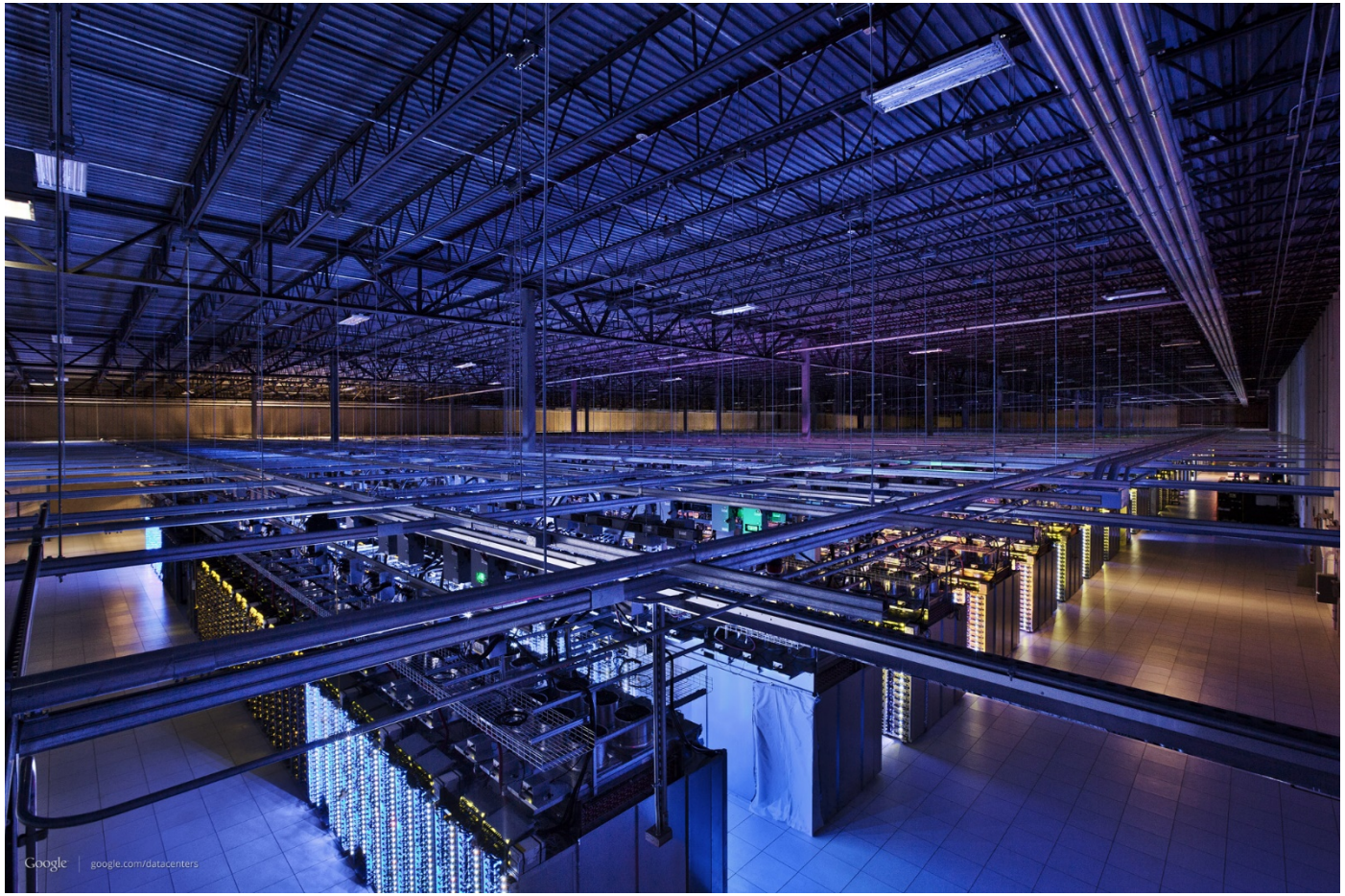
R1

In Interface	In Label	Dest.	Out Interface	Out Label

Student Questions

©2024 Raj Jain

Google's Data Center



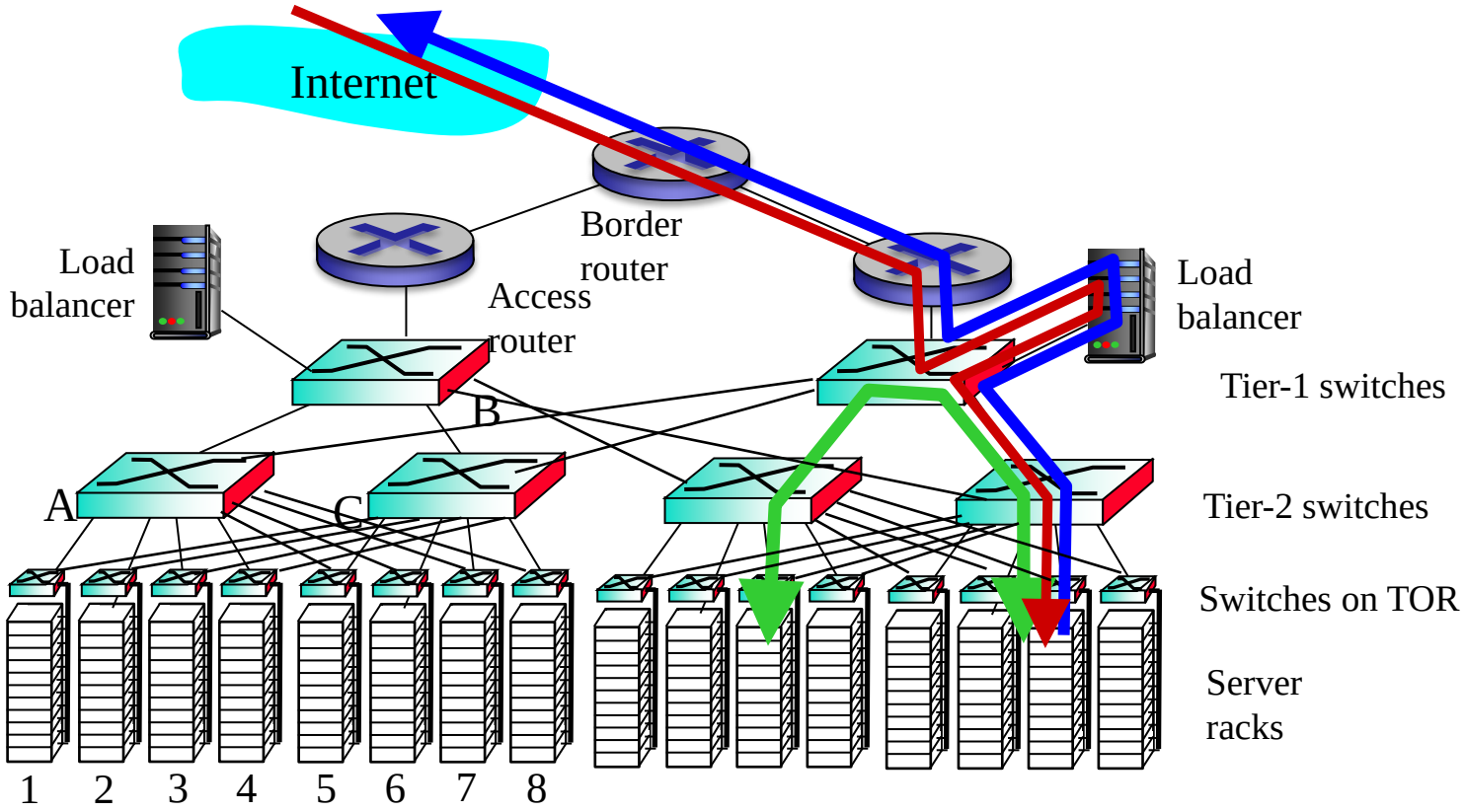
Source: http://webodysseum.com/technologyscience/visit-the-googles-data-centers/

Student Questions

©2024 Raj Jain

Data Center Networks Topology

- **3-Tier Architecture**: Server switches, Aggregation, Core
- **Middleboxes**: Load balancer, Firewall, Intrusion detection, …
- Rich Interconnection between switches
- Server switches on "top of rack" (TOR) or "end of rack" (EOR)



Student Questions

- Is TOR the bottom tier, aggregation the middle tier, and core the upper tier?

No. TOR is the position, not the tier.

- Do load balancers only interface with Tier-1 switches, and if so, why?

Load balancers can be used anywhere. Here, they are shown to separate the traffic at entry.

- Where should EOR appear on this picture?

It's the last rack in a row.

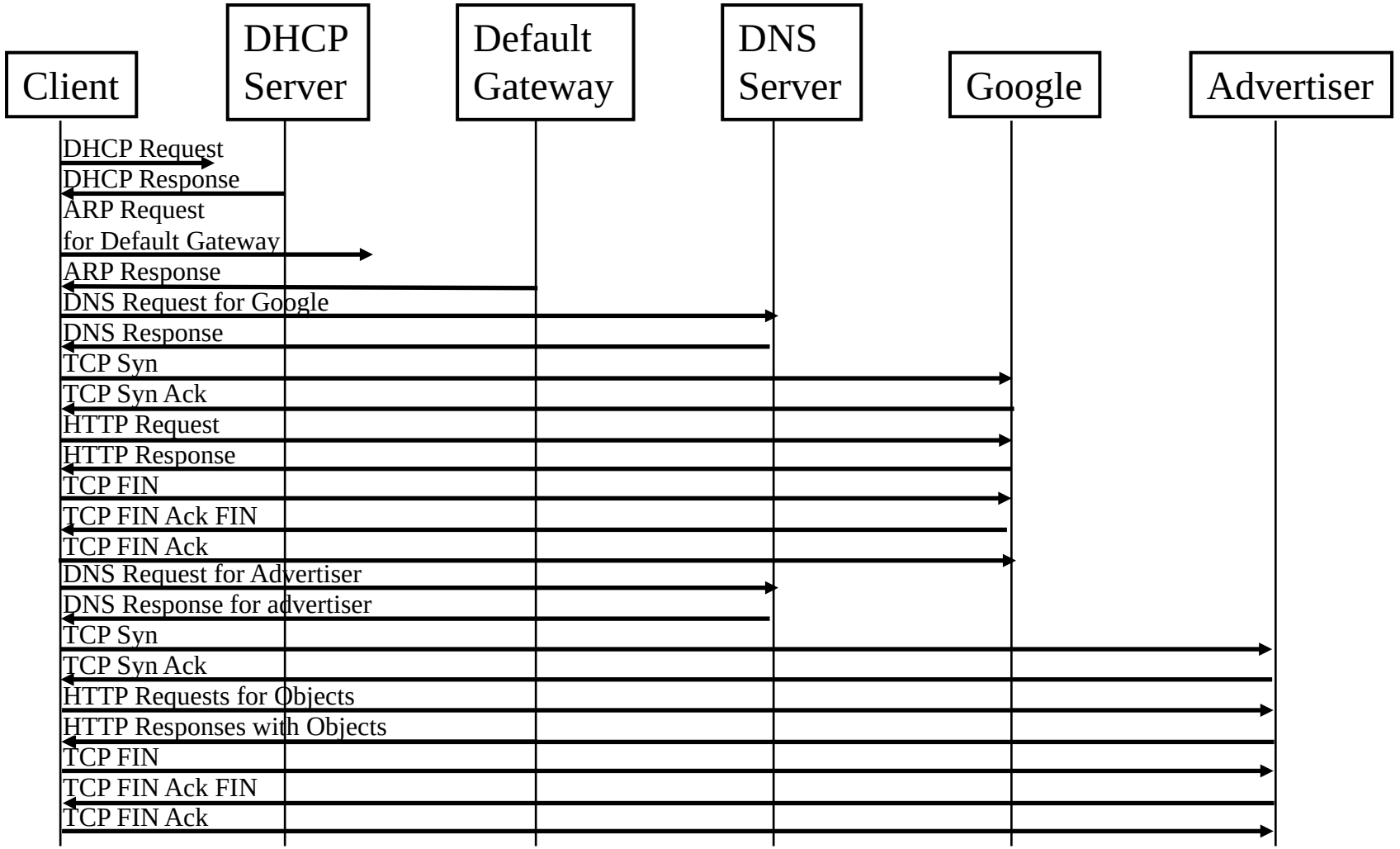
- Where are the middleboxes in the picture?

Load balancers optimally divide incoming requests among multiple servers. Firewalls and IDS are for security.

©2024 Raj Jain

Protocols: Complete Picture

Task: Connect and search in www.google.com



Student Questions

- Are these protocols in order?

Yes.

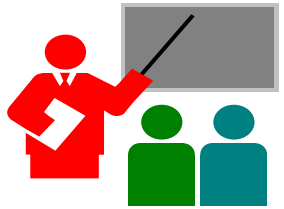
- Do we need to remember this picture in our exam?

No. But you need to understand what needs to happen during a single search.

Ref: Section 6.7

©2024 Raj Jain

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.

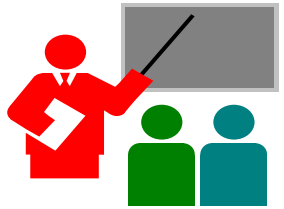
The physical layer codes bits into waveform on the media and detect malformed bits.

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

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

©2024 Raj Jain

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

❖ Chapter 6.4, page 514, paragraph 3: What if a host's IP changes (because of DHCP use) and Layer 2 forwards packets to the wrong/nonexistent host? Would the error propagate up to layer 4 (lack of ACK)?
Yes.

❖When will we see the adjusted exam scores?
Today

End of Q

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.

©2024 Raj Jain

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.

©2024 Raj Jain

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

©2024 Raj Jain

Scan This to Download These Slides

Raj Jain
http://rajjain.com

http://www.cse.wustl.edu/~jain/cse473-24/i_6lan.htm

Student Questions

- When should iBGP be used instead of OSPF?

OSPF prepares a routeing table and distributes it. 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.

- Will this exam cover any mobile and wireless concepts?

No

©2024 Raj Jain

Modulo 2 Arithmetic: More Examples

Addition:

1-bit				2-bit				3-bit
1	0	0	1	00	01	10	11	110
+1	+0	+1	+0	+11	+11	+11	+11	+101
0	0	1	?	11	10	01	??	???

Multiplication:

1-bit				2-bit				3-bit
1	0	0	1	00	01	10	11	110
×1	×0	×1	×0	×11	×11	×11	×11	×101
1	0	0	?	00	01	10	??	???
				00	01	10	??	???
				000	011	110	???	????

Student Questions

©2024 Raj Jain

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

©2024 Raj Jain