Computer Networking: A Top Down Approach 7th Edition
Jim Kurose, Keith Ross
Pearson
March 2017

A Top-Down Approach

CHAPTER 5

The Link Layer: Links, Access Networks, and LANs

CHAPTER 6

The Link Layer

Our goals:

- understand principles behind link layer services:
  - error detection, correction
  - sharing a broadcast channel: multiple access
  - link layer addressing
  - Local Area Networks: Ethernet, VLANs
- Instantiation (provide tangible example), implementation of various link layer technologies

5-2

CHAPTER 6

The Link Layer

Roadmap:

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs
- Addressing, ARP
- Ethernet
- Switches
- VLANS

6.5 Data center networking

6.6 A day in the life of a web request

5-3

CHAPTER 6

(6.1) Introduction

Terminology:

- hosts and routers: ________
- communication channels that connect adjacent nodes along communication path: ______
  - wired links
  - wireless links
  - LANs
- layer-2 packet: ___________, encapsulates datagram

Data-link layer has responsibility of transferring frame from one node to *physically adjacent* node over a link

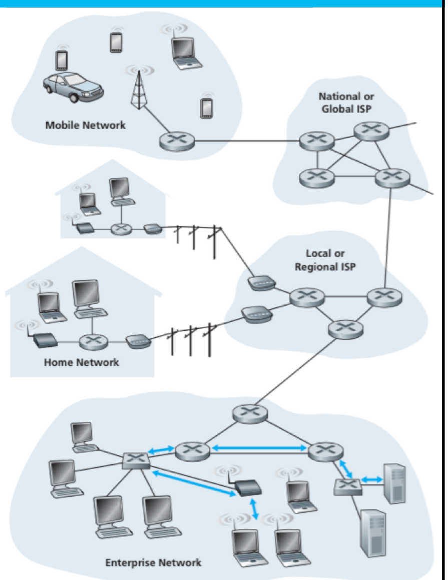


Figure: Six link-layer hops between wireless host and server.

CHAPTER 6

Link Layer: Context

- Datagram transferred by different link protocols over different links:
  - *e.g.*, Ethernet on first link, frame relay on intermediate links, 802.11 (WiFi) on last link
- Each link protocol provides different services
  - *e.g.*, may or may not provide `rdt` over link

Transportation analogy:
- Trip from Princeton to Lausanne
  - Limo : Princeton to JFK
  - Plane : JFK to Geneva
  - Train : Geneva to Lausanne
- Tourist = frame
- Transport segment = communication link
- Transportation mode = link layer protocol
- Travel agent = routing algorithm

JFK (John F. Kennedy, New York)

5-5

CHAPTER 6

Link Layer Services

- *Framing, Link access:*
  - encapsulate _________ into _________, adding header, trailer
  - channel access if shared medium
  - "MAC" addresses (*e.g.*: `74-29-2F-10-54-1A-FF-0F`) used in frame headers to identify source, destination
    - different from IP address (*e.g.*: `161.139.68.204`)!
- *Reliable delivery between adjacent nodes:*
  - we learned how to do this already (chapter 3)!
  - seldom used on low bit-error link (fiber, some twisted pair)
  - wireless links: high error rates

MAC (Medium Access Control)

5-6

CHAPTER 6

Link Layer Services

- *Flow control:*
  - pacing between adjacent sending and receiving nodes
- *Error _______________:*
  - errors caused by signal attenuation, noise.
  - receiver detects presence of errors:
    - signals sender for retransmission or drops frame
- *Error ________________:*
  - receiver identifies and corrects bit error(s) without resorting to retransmission
- *Half-duplex and full-duplex*
  - with half duplex, nodes at both ends of link can transmit, but not at same time

5-7

CHAPTER 6

Where is the Link Layer implemented?

- in each and every host
- link layer implemented in “adaptor” (aka ____________________ (NIC)) or on a chip
  - Ethernet card (wired), 802.11 card (wireless); Ethernet chipset
  - implements link, physical layer
- attaches into host’s system buses
- combination of hardware, software, firmware

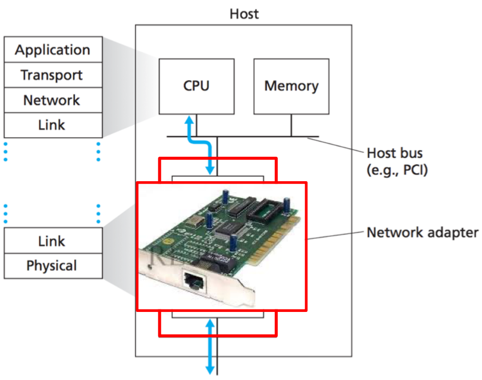


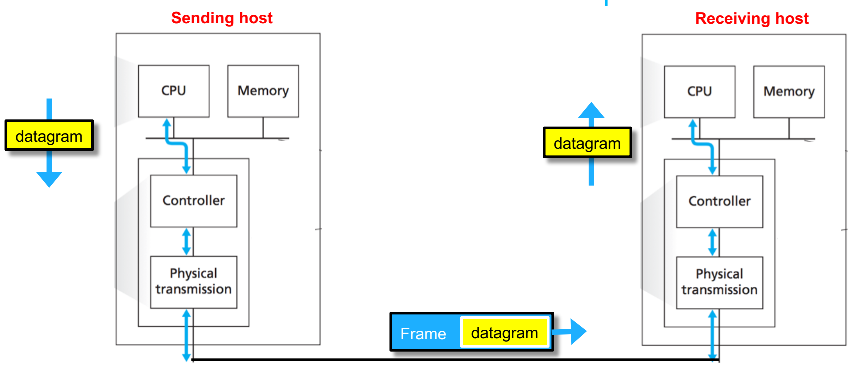
Figure: Network adapter: its relationship to other host components and to protocol stack functionality.

5-8

CHAPTER 6

Where is the Link Layer implemented?

Adaptors communicating



<u>Sending side</u>:

- encapsulates *datagram* in *frame*
- adds error checking bits, `rdt`, flow control, etc.

<u>Receiving side</u>:

- looks for errors, `rdt`, flow control, etc
- extracts *datagram*, passes to upper layer at receiving side

CHAPTER 6

The Link Layer

Roadmap:

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs
- Addressing, ARP
- Ethernet
- Switches
- VLANS

6.5 Data center networking

6.6 A day in the life of a web request

5-11

CHAPTER 6

(6.2) Error-Detection and -Correction Techniques

EDC = *Error Detection and Correction* bits (redundancy)
D = *Data* protected by error checking, may include header fields

Error detection not 100% reliable!
- protocol may miss some errors, but rarely
- larger EDC field yields better detection and correction

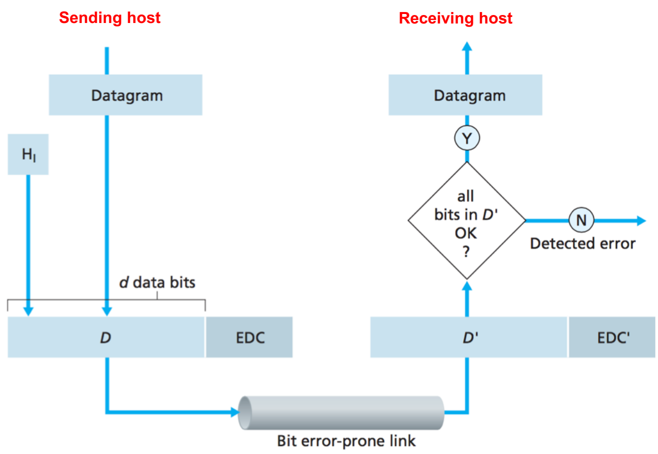


Figure: Error-detection and –correction scenario.

5-12

CHAPTER 6

(6.2) Error-Detection and -Correction Techniques

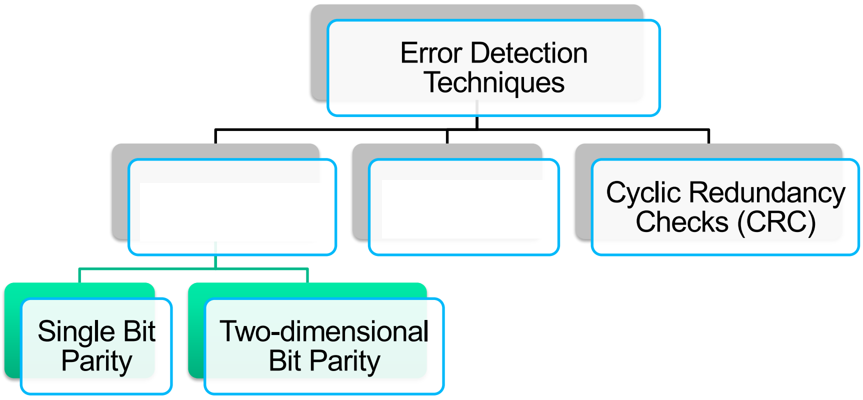


Figure: Techniques for detecting errors in transmitted data.

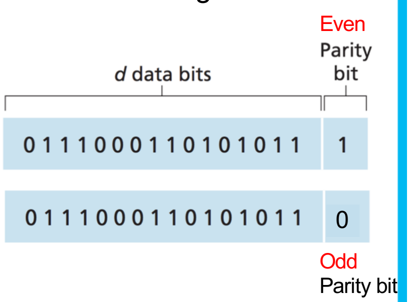
5-13

CHAPTER 6

(a) Parity Checks

Single Bit Parity:

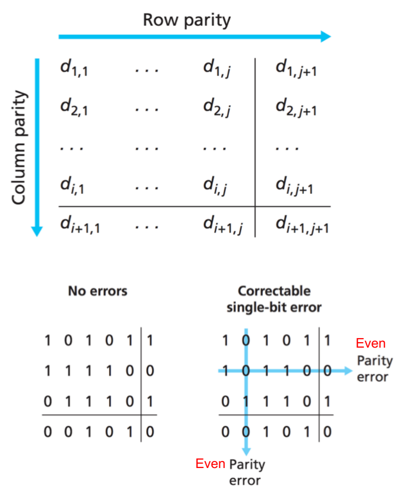
- Detect single bit errors



Example:
Parity bit indicates whether number of bits with "1" in *d* is *even* or *odd*.

Two-dimensional Bit Parity:

- Detect and correct single bit errors



CHAPTER 6

(b) Internet Checksum

Review

Goal:
Detect "errors" (e.g., flipped bits) in transmitted segment

(Note: used at transport layer only)

Sender:

- treat segment contents, including header fields, as sequence of 16-bit integers
- ________________: addition (one's complement sum) of segment contents
- sender puts checksum value into UDP checksum field

Receiver:

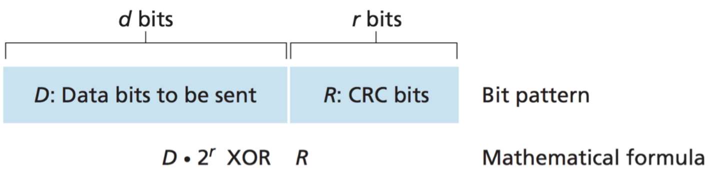
- compute checksum of received segment
- check if computed checksum **equals checksum** field value:
  - NO - error detected
  - YES - no error detected.

5-15

CHAPTER 6

(c) Cyclic Redundancy Check (CRC)

- more powerful error-detection coding
- view data bits, *D*, as a binary number
- choose `r+1` bit pattern (generator), *G*
- Goal: choose `r` CRC bits, *R*, such that
  - `<D,R>` exactly divisible by *G* (modulo 2)
  - receiver knows *G*, divides `<D,R>` by *G*. If non-zero remainder: Error detected!
  - can detect all burst errors less than `r+1` bits
- widely used in practice (Ethernet, 802.11 WiFi, ATM)



CHAPTER 6

(c) Cyclic Redundancy Check (CRC)

1) Agree with receiver that r = 3 bits
2) Append 3 zeros to *D* → 101110000
3) Agree on *G* (must be r+1 = 4 bits). Can choose from {1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111} → Choose 1001
4) Get *R* (r bits)
5) Append *R* to *D*. Send to receiver

What's transmitted ? 101110 011

Example: Sender

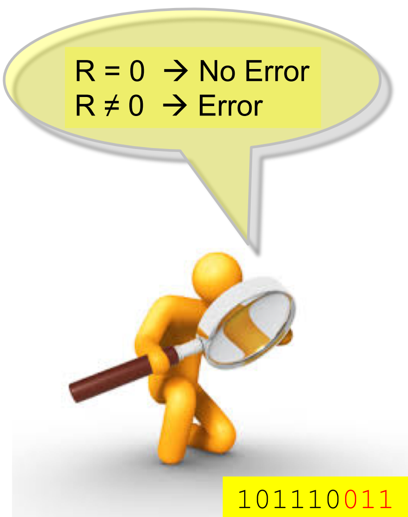
Subtraction of 1011−1001
= 1011 XOR 1001

```
   G             1 0 1 0 1 1
1 0 0 1 ) 1 0 1 1 1 0 0 0 0
          1 0 0 1              D
            0 1 0 1
            0 0 0 0
              1 0 1 0
              1 0 0 1
                0 1 1 0
                0 0 0 0
                  1 1 0 0
                  1 0 0 1
                    1 0 1 0
                    1 0 0 1
                      0 1 1
                        R
```

CHAPTER 6

(c) Cyclic Redundancy Check (CRC)

Example: Receiver




```
               101011
G ← 1001 ) 101110011
           1001
            0101
            0000
             1010
             1001
              0110
              0000
               1101
               1001
                1001
                1001
                 000  → R
```

CHAPTER 6

Exercise 6.1

If the data to be sent is 101110 and the CRC technique is used with $r = 3$ and G = 1010.

a) What is the value of R ?
b) What is the data will be sent?

CHAPTER 6

Exercise 6.2

If the data received is 1011101100 and the CRC technique is used with G = 10100.

a) What is the value of CRC?
b) Is the data error?

5-20

CHAPTER 6

Exercise 6.3a

Consider the generator, *G*, and suppose the value as *D*. What is the value will be transmitted if the *D* and *G* have the values of :

a) 11000111010, *G* = 1001

b) 01101010101, *G* = 1110

c) 11111010101, *G* = 10011

d) 10001100001, *G* = 101

Kurose, J.F. and Ross, K.W. (2017). Computer Networking: A Top-Down Approach (7th Edition). Pearson Education Limited, England. (pp. 536)

CHAPTER 6

Exercise 6.3b

Consider a CRC error detection technique is used. The following table received by a receiver tabulates with *D* and *G* as data and generator with *r* bit pattern, respectively. Complete the table. Show your working steps.

	D	***G***	***r***	***CRC***	***Error*** (Y/N)
a)	101011000011	10011			
b)	11001100011	1110			
c)	1111000100001	11000			
d)	100010001111	101			

(Remarks: r is number of bits, CRC is the binary part from D)

Online CRC Generator: https://asecuritysite.com/comms/crc_div

CHAPTER 6

The Link Layer

Roadmap:

- 6.1 Introduction, services
- 6.2 Error detection, correction
- **6.3 Multiple access protocols**
- 6.4 LANs
  - Addressing, ARP
  - Ethernet
  - Switches
  - VLANS
- 6.5 Data center networking
- 6.6 A day in the life of a web request

5-23

CHAPTER 6

(6.3) Multiple Access Links & Protocols

Types of Network Links

- PPP for dial-up access
- PPP link between Ethernet switch, host

(shared wire or wireless)

- old-fashioned Ethernet
- upstream HFC
- 802.11 wireless LAN

PPP (Point-to-Point Protocol)
HFC (Hybrid Fiber Coax)

5-24

CHAPTER 6

(6.3) Multiple Access Links & Protocols

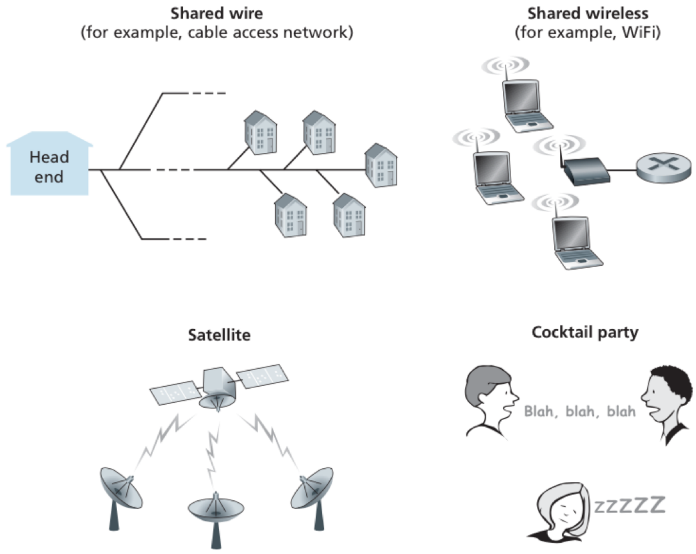


Figure: Broadcast: Various multiple access channels.

5-25

CHAPTER 6

(6.3) Multiple Access Links & Protocols

As humans, we've evolved an elaborate set of protocols for sharing the broadcast channel:

In a single shared broadcast channel :

- *two or more simultaneous transmissions by nodes: ______________*

"Give everyone a chance to speak."
"Don't speak until you are spoken to."
"Don't monopolize the conversation."
"Raise your hand if you have a question."
"Don't interrupt when someone is speaking."
"Don't fall asleep when someone is talking."

- collision if node receives two or more signals at the same time

5-26

CHAPTER 6 (6.3) Multiple Access Links & Protocols

Multiple Access Protocols

Solution

- distributed algorithm that determines how nodes share channel,
  - *i.e.*: determine when node can transmit
- communication about channel sharing must use channel itself!
  - no *out-of-band* channel for coordination

5-27

CHAPTER 6 (6.3) Multiple Access Links & Protocols

Overview conclusion

A **multiple access protocol** for a broadcast channel of rate R bps should have the following *desirable characteristics*:

1) when one node wants to transmit, it can send at rate R *bps*.
2) when M nodes want to transmit, each can send at average rate (R / M) *bps.*
3) Fully decentralized:
   - no special node to coordinate transmissions
   - no synchronization of clocks, slots
4) The protocol is simple.

MAC (Multiple Access Protocol)

5-28

CHAPTER 6

(6.3) Multiple Access Links & Protocols

Multiple access protocols categories

Channel Partitioning Protocols
(Channelization Methods)

- divide channel into smaller "pieces" (time slots, frequency, code)
- allocate piece to node for exclusive use

- channel not divided, allow collisions
- "recover" from collisions

- nodes take turns, but nodes with more to send can take longer turns

TDMA | FDMA | CDMA | Slotted ALOHA | ALOHA | CSMA | CSMA/CD | Polling | Token Passing

5-29

CHAPTER 6

(a) Channel Partitioning Protocols

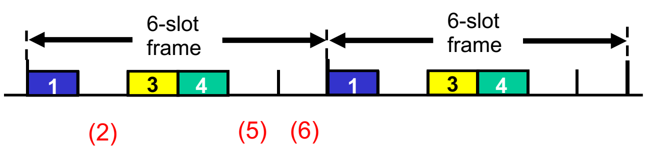
TDMA

TDMA: Time Division Multiple Access

- access to channel in "rounds" ;
- each station gets fixed length slot (length = packet transmission time) in each round;
- unused slots go idle;

Example:

- 6-station LAN, 1,3,4 have packet, slots 2,5,6 idle



5-30

CHAPTER 6

(a) Channel Partitioning Protocols

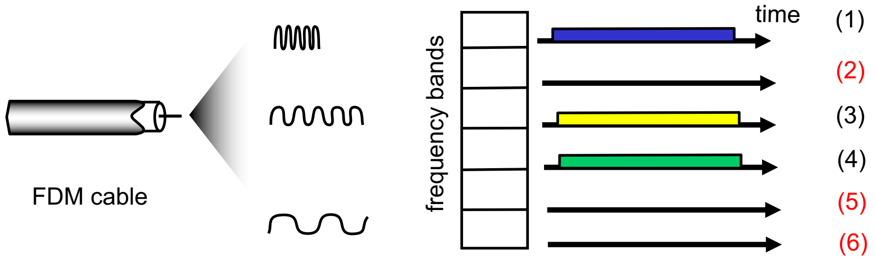
FDMA

FDMA: Frequency Division Multiple Access

- channel spectrum divided into frequency bands;
- each station assigned fixed frequency band;
- unused transmission time in frequency bands go idle;

Example:

- 6-station LAN, 1,3,4 have packet, frequency bands 2,5,6 idle



CHAPTER 6

(a) Channel Partitioning Protocols

CDMA

CDMA: Code Division Multiple Access

- Assigns a different ___________ to each node;
- Each node uses its unique code to encode the data bits it sends;
- If the code chosen carefully, different nodes can transmit simultaneously;

- CDMA has been used in military for some time;
- now has widespread civilian use, particularly in cellular telephony (wireless channel)

5-32

CHAPTER 6

(b) Random Access Protocols

- when node has packet to send:
  - transmit at full channel data rate R;
  - no *a priori* coordination among nodes;
- Two or more transmitting nodes → "______________"
- Random access MAC protocol specifies:
  - How to **detect** collisions?
  - How to **recover** from collisions? (e.g., via delayed retransmissions)
- *Examples of random access MAC protocols:*
  - *slotted ALOHA*
  - *ALOHA*
  - *CSMA, CSMA/CD, CSMA/CA*

5-33

CHAPTER 6

(b) Random Access Protocols

Slotted ALOHA

Assumptions:

- all frames same size
- time divided into equal size slots (time to transmit 1 frame)
- nodes start to transmit only slot beginning
- nodes are synchronized
- if 2 or more nodes transmit in slot, all nodes detect collision

Operation:

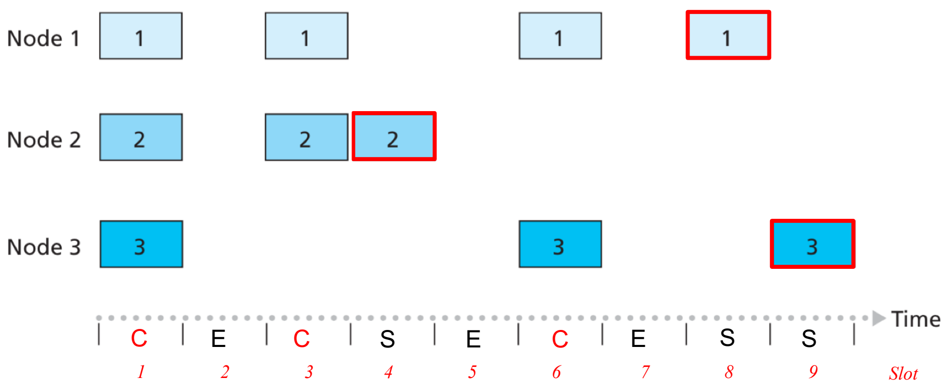
- when node obtains fresh frame, transmits in next slot
  - if *no collision:* node can send new frame in next slot
  - if *collision:* node retransmits frame in each subsequent slot with probability p until success

5-34

CHAPTER 6

(b) Random Access Protocols

Slotted ALOHA



Key:
C = Collision slot
E = Empty slot
S = Successful slot

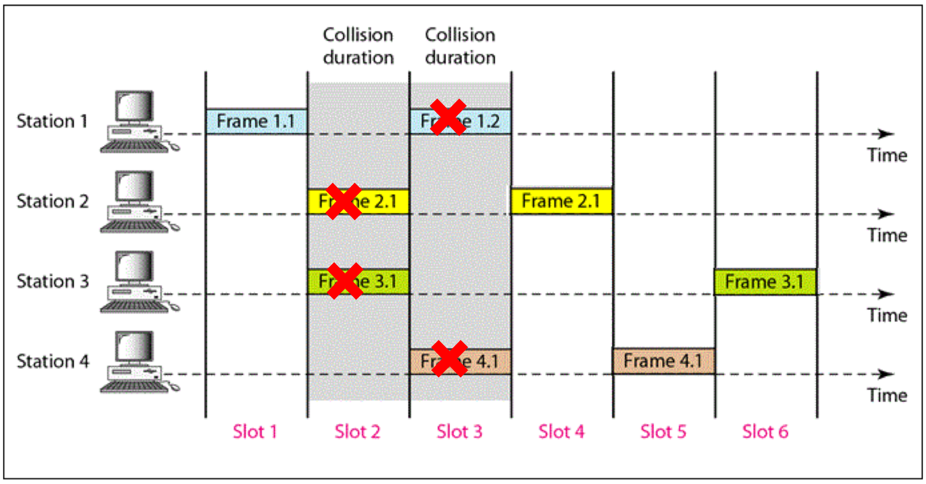
Figure: Nodes 1, 2, and 3 collide in the first slot. Node 2 finally succeeds in the fourth slot, node 1 in the eighth slot, and node 3 in the ninth slot

CHAPTER 6

(b) Random Access Protocols

Slotted ALOHA

Example 2:



http://4.bp.blogspot.com/-Jr1pMHewzbM/U1DF2EM26bI/AAAAAAAABqE/1DPuNeO_5zM/s1600/slotted-aloha.GIF

CHAPTER 6

(b) Random Access Protocols

Slotted ALOHA

Advantages:

- single active node can continuously transmit at full rate of channel
- highly decentralized: each node detects collisions and independently decides when to retransmit
- Simple protocol

Disadvantages:

- Collisions will wasting slots
- idle slots: refrain from transmitting
- require the slots to be synchronized in the nodes

5-37

CHAPTER 6

(b) Random Access Protocols

Pure ALOHA (Unslotted)

- unslotted ALOHA: simpler, no synchronization;
- when frame first arrives → transmit immediately;
- collision probability increases:
  - frame sent at t_0 collides with other frames sent in $[t_0-1, t_0+1]$

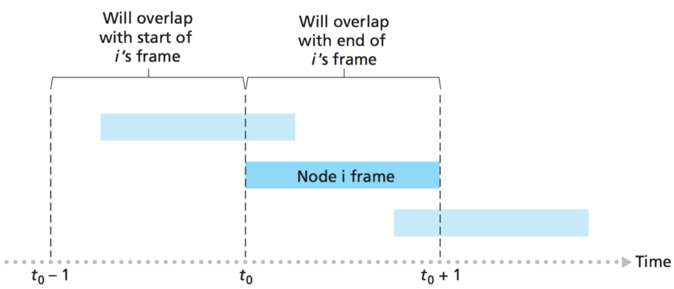


Figure: Interfering transmissions in pure ALOHA.

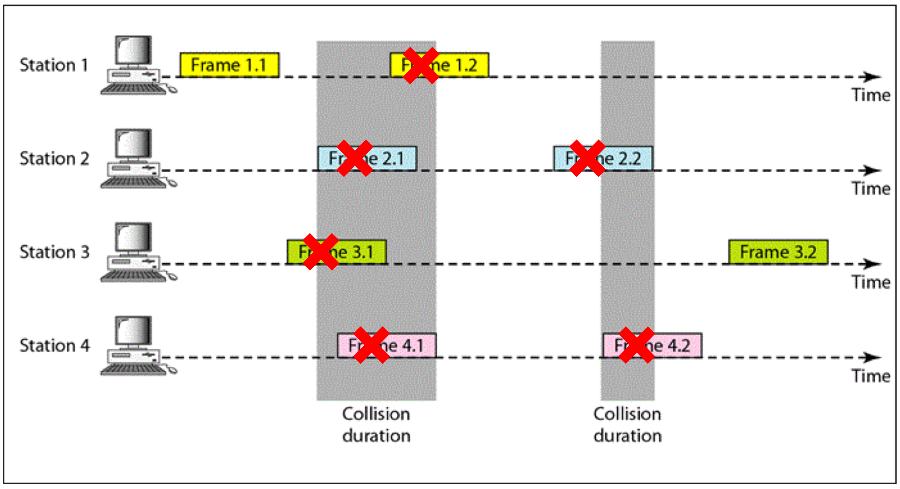
5-38

CHAPTER 6

(b) Random Access Protocols

Pure ALOHA (Unslotted)

Example 2:



http://1.bp.blogspot.com/-nS-6P90-iCc/U1DF2Kei3fI/AAAAAAAABp4/guEz45WrJU4/s1600/aloha.GIF

CHAPTER 6

(b) Random Access Protocols

CSMA

CSMA: Carrier Sense Multiple Access

- Listen before transmit:
  - if channel sensed ________: transmit entire frame
  - if channel sensed ________: defer transmission (abort)
- Human analogy: Don’t interrupt others!

5-40

CHAPTER 6

(b) Random Access Protocols

CSMA

- Collisions *can* still occur: propagation delay means two nodes may not hear each other's transmission.
- Collision: entire packet transmission time wasted
  - *distance* & *propagation delay* play role in in determining collision probability.

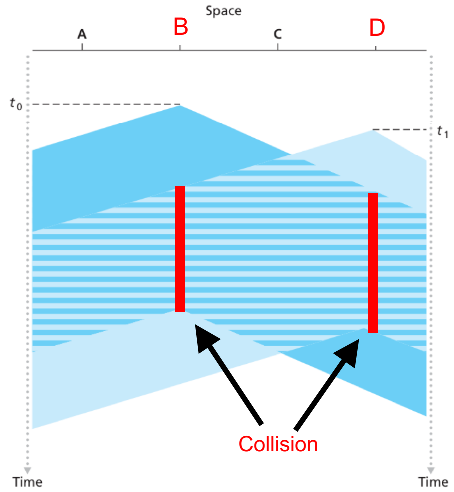


Figure: Space-time diagram of two CSMA nodes with colliding transmission.

CHAPTER 6

(b) Random Access Protocols

CSMA

No collision detection. Both B and D continue to transmit the entire frames even though collision occur.

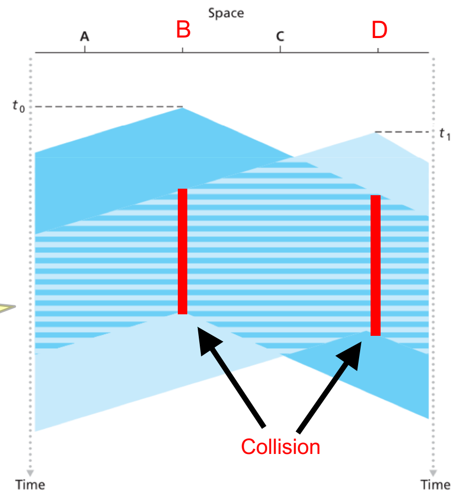


Figure: Space-time diagram of two CSMA nodes with colliding transmission.

CHAPTER 6 (b) Random Access Protocols

CSMA/CD

CSMA/CD:
(Carrier Sense Multiple Access / Collision Detection)

- carrier sensing, deferral as in CSMA
  - collisions *detected* within short time;
  - colliding transmissions aborted, reducing channel wastage;

Collision Detection (CD):

- *easy in wired LANs: measure signal strengths, compare transmitted, received signals;*
- *difficult in wireless LANs: received signal strength overwhelmed by local transmission strength;*

- *human analogy: the polite conversationalist*

CHAPTER 6 (b) Random Access Protocols

CSMA/CD

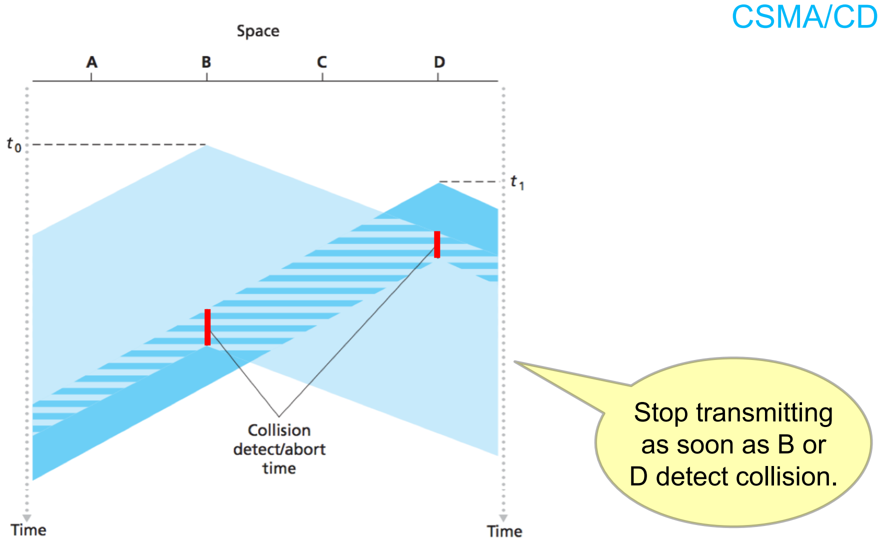


Figure: CSMA with collision detection (CD).

5-44

CHAPTER 6

(b) Random Access Protocols

CSMA/CD

Summary of the operation from the perspective of an adapter attached to a broadcast channel:

1. NIC receives datagram from network layer, creates frame
2. - If NIC senses channel idle, starts frame transmission;
   - If NIC senses channel busy, waits until channel idle, then transmits.
3. If NIC transmits entire frame without detecting another transmission, NIC is done with frame !
4. If NIC detects another transmission while transmitting, aborts and sends jam signal
5. After aborting, NIC enters *binary (exponential) backoff:*
   - after mth collision, NIC chooses K at random from $\{0,1,2, \ldots, 2^m - 1\}$. NIC waits K*512 bit times, returns to Step 2
   - longer backoff interval with more collisions

NIC (Network Interface Card))

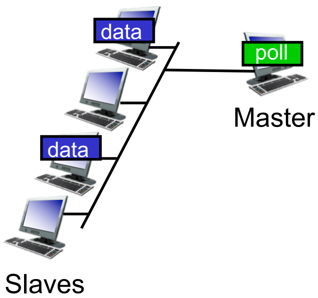
5-45

CHAPTER 6

(c) Taking-Turns Protocols

Polling

- Master node "invites" slave nodes to transmit in turn
- Typically used with "dumb" slave devices
- Concerns:
  - polling overhead
  - latency
  - single point of failure (master)



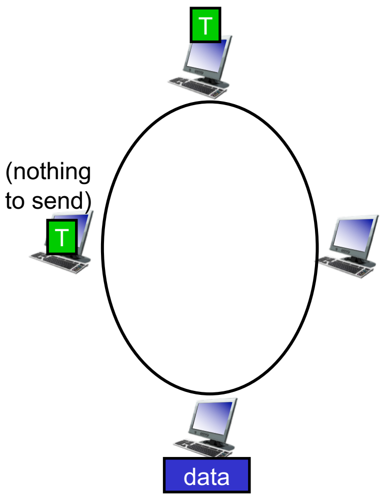
5-46

CHAPTER 6

(c) Taking-Turns Protocols

Token Passing

- Control ________ passed from one node to next sequentially.
- Token message
- Concerns:
  - token overhead
  - latency
  - single point of failure (token)



5-47

CHAPTER 6

Summary of MAC Protocols

- *Channel Partitioning,* by time, frequency or code
  - Time Division, Frequency Division
- *Random Access* (dynamic),
  - ALOHA, S-ALOHA, CSMA, CSMA/CD
  - carrier sensing: easy in some technologies (wire), hard in others (wireless)
    - CSMA/CD used in Ethernet → ____________
    - CSMA/CA used in 802.11 → ____________
- *Taking Turns*
  - polling from central site, token passing
  - bluetooth, FDDI, token ring

Fiber Distributed Data Interface (FDDI)

5-48

CHAPTER 6 The Link Layer

Roadmap:

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs
- **Addressing, ARP**
- **Ethernet**
- **Switches**
- **VLANS**

6.5 Data center networking

6.6 A day in the life of a web request

5-49

CHAPTER 6 (6.4) Switched Local Area Networks

- Having covered broadcast networks and multiple access protocols (MAC) in previous section;
- Switches operate at the link layer :
  - switch link-layer frames;
  - not recognize network-layer addresses
  - not use routing algorithms
  - Use link-layer addresses

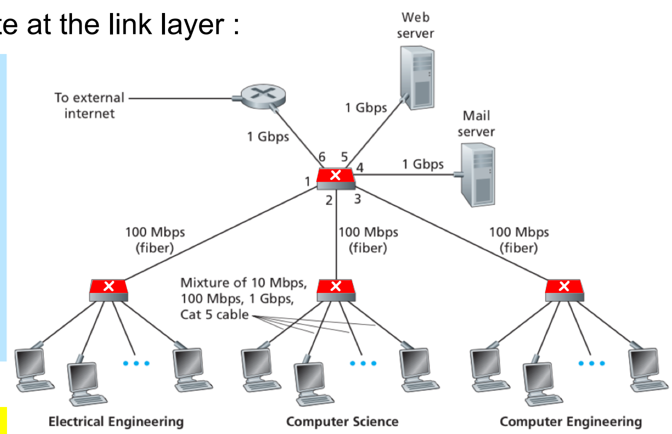


Figure: An institutional network connected together by four switches.

CHAPTER 6
Link Layer Addressing and Address Resolution Protocol (ARP)

- 32-bit IP address:
  - network-layer address for interface
  - used for layer 3 (network layer) forwarding

Hosts and routers have link-layer addresses ?

- MAC (or _________ or __________ or _________) address:
  - function: *used 'locally" to get frame from one interface to another physically-connected interface (same network, in IP-addressing sense)*
  - 48 bit MAC address or 6 bytes address (for most LANs) burned in NIC ROM, also sometimes software settable
  - *e.g.*: `1A-2F-BB-76-09-AD`

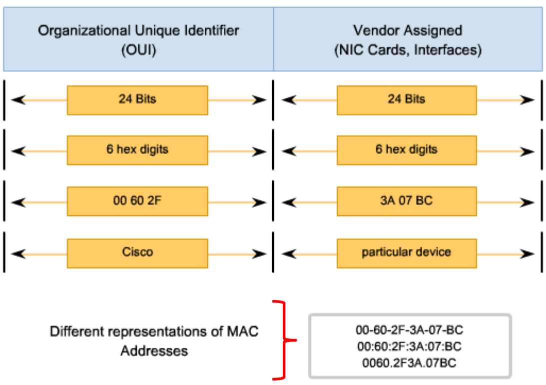
Hexadecimal (base 16) notation (each "number" represents 4 bits)

5-51

CHAPTER 6
Link Layer Addressing and Address Resolution Protocol (ARP)

LAN / MAC addresses

- Every piece of Ethernet hardware has the address "*burned in*" to a chip on the hardware.
- The first 3 bytes of an Ethernet address are the manufacturer's code and the last 3 bytes are a unique sequence number.



5-52

CHAPTER 6

Link Layer Addressing and Address Resolution Protocol (ARP)

LAN / MAC addresses

Each adapter on LAN has unique and unchanged MAC / LAN / physical / ______________ address;

Adapter (interface)

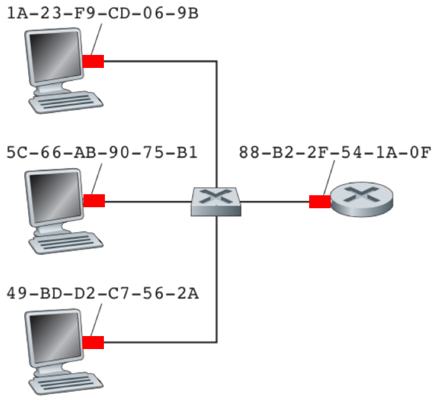


Figure: Each interface connected to a LAN has a unique MAC address.

5-53

CHAPTER 6

Link Layer Addressing and Address Resolution Protocol (ARP)

LAN / MAC addresses

- MAC address allocation administered by IEEE
- manufacturer buys portion of MAC address space (to assure uniqueness)

- Analogy:
  - MAC address: like *Social Security Number*
  - IP address: like *postal address*

- MAC flat address → portability
  - can move LAN card from one LAN to another

- IP hierarchical address → not portable
  - address depends on IP subnet to which node is attached

IEEE (Institute Electrical and Electrical Engineers)

5-54

CHAPTER 6

Link Layer Addressing and Address Resolution Protocol (ARP)

ARP

Q: How to determine interface's MAC address, knowing its IP address?

ARP table: each IP node (host, router) on LAN has an ARP table

- IP/MAC address mappings for some LAN nodes:
  < IP address; MAC address; *TTL*>
- *TTL* (*Time To Live*): time after which address mapping will be forgotten (typically 20 min)

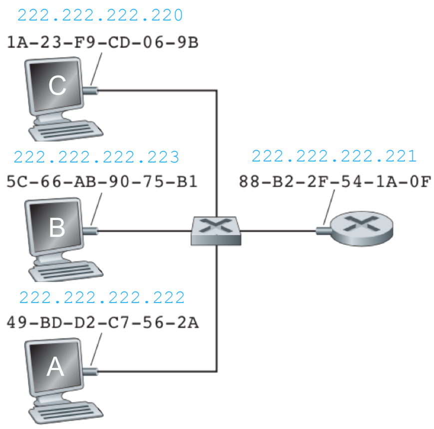


Figure: Each interface on a LAN has an IP address and a MAC address.

CHAPTER 6

Link Layer Addressing and Address Resolution Protocol (ARP)

ARP

IP Address	MAC Address	TTL
222.222.222.221	88-B2-2F-54-1A-0F	13:45:00
222.222.222.223	5C-66-AB-90-75-B1	13:52:00

Figure: A possible ARP table in 222.222.222.220

CHAPTER 6
Link Layer Addressing and Address Resolution Protocol (ARP)
ARP

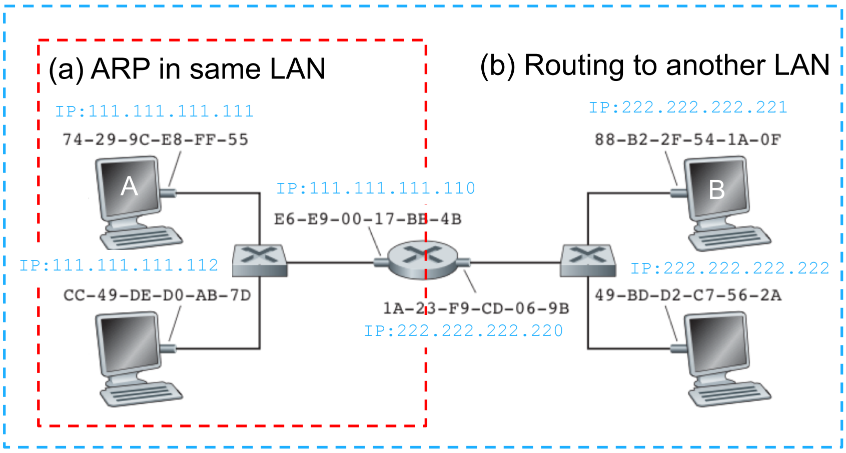
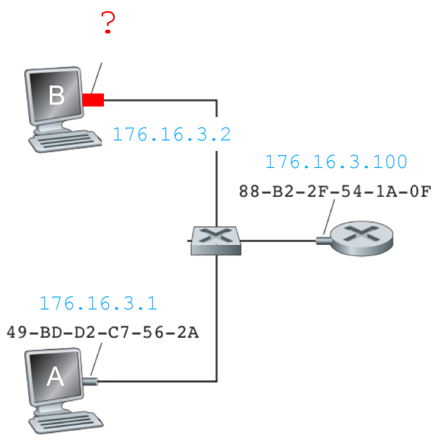


Figure: 2 subnets interconnected by a router.

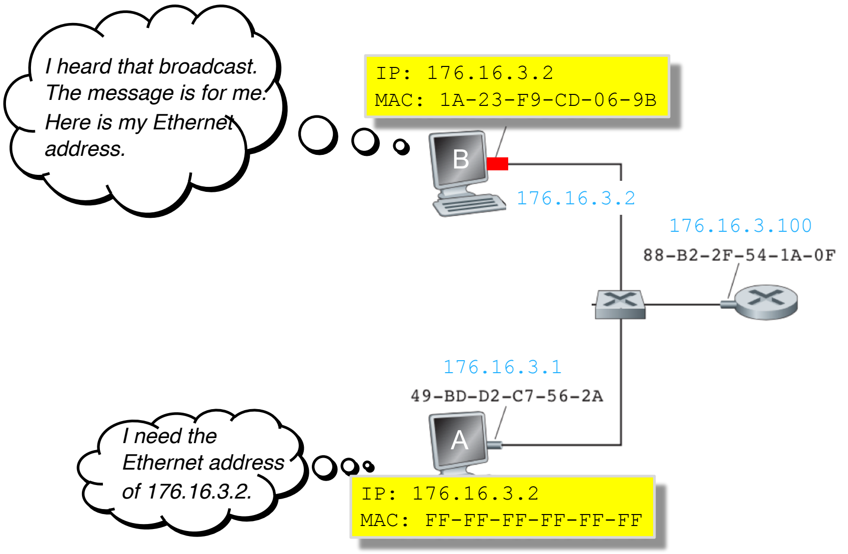
CHAPTER 6
Link Layer Addressing and Address Resolution Protocol (ARP)
(a) ARP protocol: same LAN

- A wants to send frames to B
  - B's MAC address not in A's ARP table.
- A ______________ ARP query packet, containing B's IP address
  - destination MAC address = FF-FF-FF-FF-FF-FF
  - all nodes on LAN receive ARP query
- B receives ARP packet, replies to A with its (B's) MAC address
  - frame sent to A's MAC address (____________)



CHAPTER 6 Link Layer Addressing and Address Resolution Protocol (ARP)

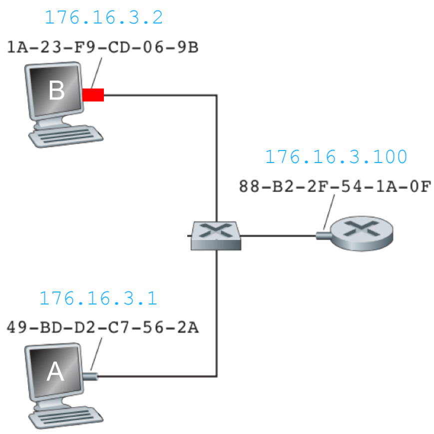
(a) ARP protocol: same LAN



CHAPTER 6 Link Layer Addressing and Address Resolution Protocol (ARP)

(a) ARP protocol: same LAN

- Host A caches (saves) IP-to-MAC address pair in its ARP table until information becomes old (times out)
  - soft state: information that times out (goes away) unless refreshed
- ARP is "plug-and-play":
  - nodes create their ARP tables without intervention from network administrator



CHAPTER 6

Link Layer Addressing and Address Resolution Protocol (ARP)

(b) Addressing: routing to another LAN

Walkthrough: send datagram from A to B via a router

- focus on addressing – at IP (________) and MAC layer (______)
  - assume A knows B's IP address
  - assume A knows IP address of first hop router
  - assume A knows router's MAC address

How ???

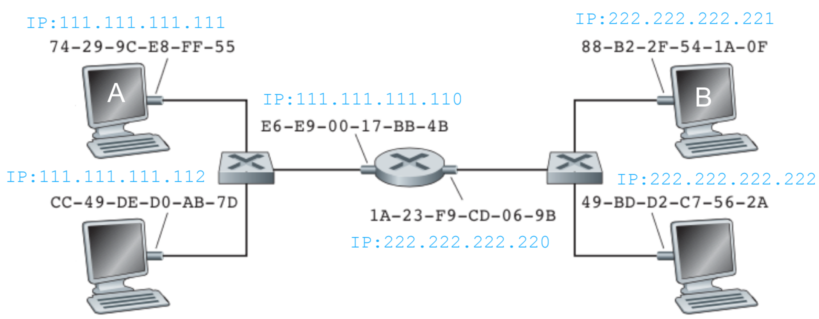


Figure: Two subnets interconnected by a router.

5-61

CHAPTER 6

Link Layer Addressing and Address Resolution Protocol (ARP)

(b) Addressing: routing to another LAN

Link-layer frame
src:74-29-9C-E8-FF-55
dest:E6-E9-00-17-BB-4B

IP datagram
src:111.111.111.111
dest:222.222.222.221

- *frame* received at router, *datagram* removed, passed up to IP
- router forwards *datagram* with IP source A, dest. B
- router creates *link-layer frame* with B's MAC address and contains A-to-B IP datagram

Link-layer frame
src:1A-23-F9-CD-06-9B
dest:88-B2-2F-54-1A-0F

IP datagram
src:111.111.111.111
dest:222.222.222.221

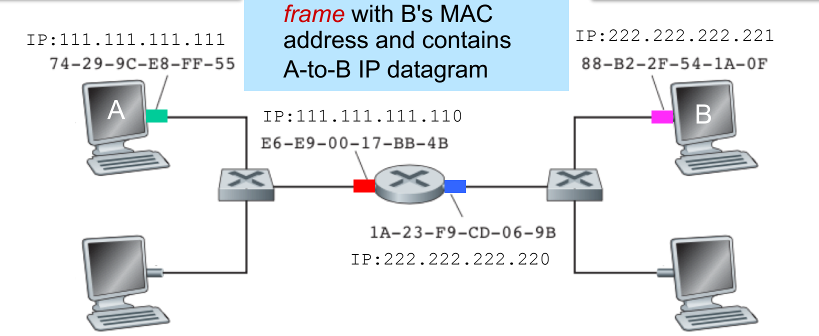


Figure: Two subnets interconnected by a router.

5-62

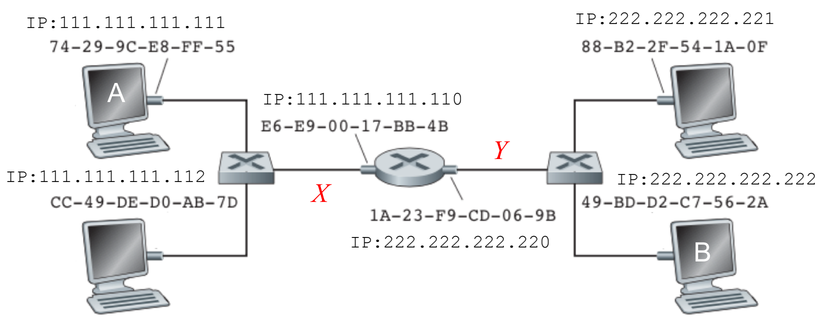
CHAPTER 6

Exercise 6.4

Suppose A sends frame to B.

a) What is the source and destination IP address at X and Y?

b) What is the source and destination Ethernet address at X and Y?



5-63

CHAPTER 6

Solution 6.4

a)

IP address	X	Y
source :		
destination :		

b)

Ethernet address	X	Y
source :		
destination :		

5-64

CHAPTER 6

Ethernet

"dominant" wired LAN technology:
- cheap $20 for NIC
- first widely used LAN technology
- simpler, cheaper than token LANs and ATM
- kept up with speed race: *10 Mbps – 10 Gbps*

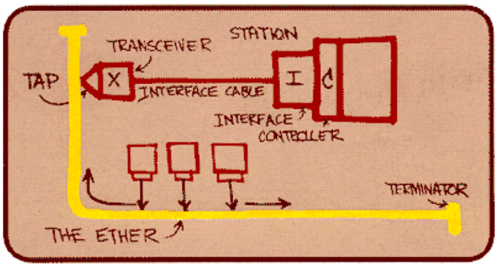


Figure: Melcalfe's Ethernet sketch.

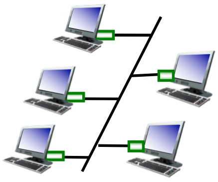
5-65

CHAPTER 6

Ethernet

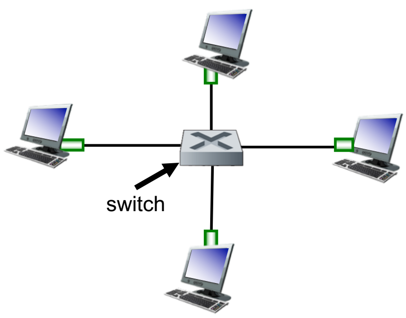
Physical topology

- __________: popular through mid 90s
  - all nodes in same collision domain (can collide with each other)



Coaxial cable

- __________: prevails today
  - active switch in center
  - each "spoke" runs a (separate) Ethernet protocol (nodes do not collide with each other)



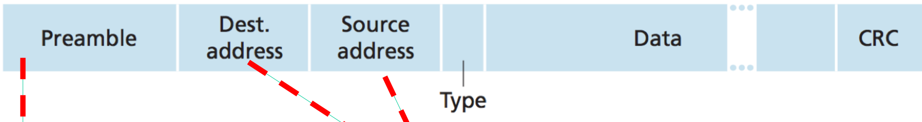
5-66

CHAPTER 6

Ethernet

Frame structure

- sending adapter encapsulates IP datagram (or other network layer protocol packet) in Ethernet frame



Preamble (8 bytes):

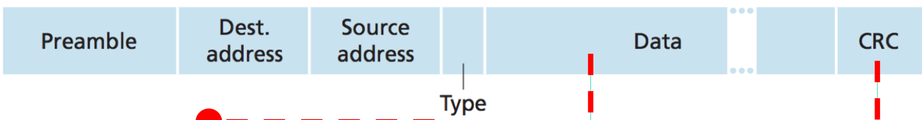
- *7 bytes with pattern* `10101010` *followed by one byte with pattern* `10101011`
- *used to synchronize receiver, sender clock rates*

Addresses: 6 byte source, 6 byte destination MAC addresses

- if adapter receives frame with matching destination address, or with broadcast address (e.g. ARP packet), it passes data in frame to network layer protocol
- otherwise, adapter discards frame

CHAPTER 5

Ethernet



Type (2 bytes):

- indicates higher layer protocol (mostly IP but others possible, e.g., Novell IPX, AppleTalk)

CRC (4 bytes):

- *cyclic redundancy check at receiver error detected: frame is dropped*

Data (payload) (46 – 1500bytes)

- *Min size: 46 bytes (if data less than 46 bytes, use byte stuffing until size become 46 bytes)*
- *Max size: 1500 bytes (MTU size)*

CHAPTER 6

Ethernet

Connectionless, Unreliable

Connectionless:
No handshaking between sending and receiving NICs

Unreliable:
Receiving NIC doesn't send `ACK`s or `NAK`s to sending NIC

- data in dropped frames recovered only if initial sender uses higher layer `rdt` (e.g., TCP), otherwise dropped data lost

- Ethernet's MAC protocol:
  - unslotted *CSMA/CD with binary backoff*

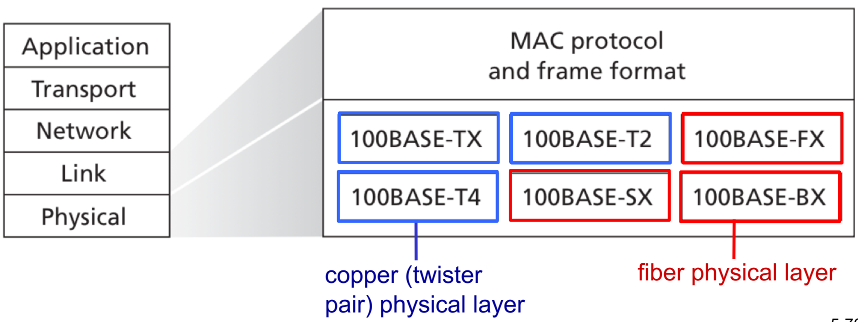
5-69

CHAPTER 6

Ethernet

Ethernet standard: IEEE 802.3 CSMA/CD

- *Many* different Ethernet standards
  - common MAC protocol and frame format
  - different speeds: *2Mbps*, *10Mbps*, *100Mbps*, *1Gbps*, *10Gbps*
  - different physical layer media: fiber, cable



5-70

CHAPTER 6

Link-Layer Switches

- link-layer device: takes an *active* role
  - store, forward Ethernet frames
  - examine incoming frame's MAC address, selectively forward frame to one-or-more outgoing links when frame is to be forwarded on segment, uses CSMA/CD to access segment
- *transparent*
  - hosts are unaware of presence of switches
- *plug-and-play, self-learning*
  - switches do not need to be configured

https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQoUD37sS32DJUtlDM4D7roLy-B8ceHm58lst40JvFeTotqBQeN

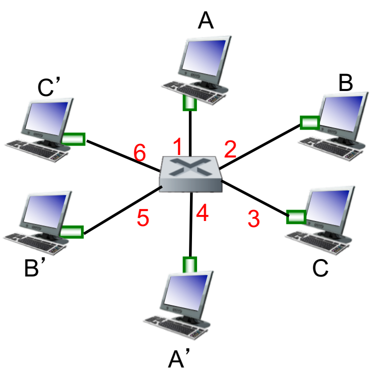
5-71

CHAPTER 6

Link-Layer Switches

Multiple simultaneous transmissions

- hosts have dedicated, direct connection to switch
- switches will buffer the packets
- Ethernet protocol used on *each* incoming link, but no collisions; full duplex
  - each link has its own collision domain
- *switching:* A-to-A' and B-to-B' can transmit simultaneously, without collisions



switch with six interfaces (1,2,3,4,5,6)

5-72

CHAPTER 6

Link-Layer Switches

Forwarding table

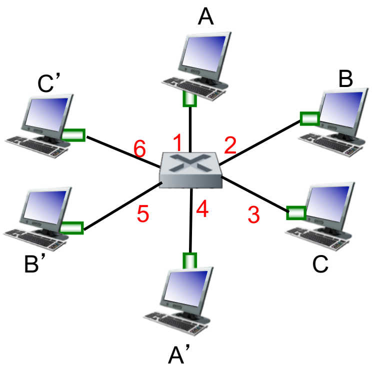
Q: How does switch know A' reachable via interface 4, B' reachable via interface 5?

A: Each switch has a __________,
each entry:

- (MAC address of host, interface to reach host, time stamp)
- looks like a routing table!

Q: How are entries created, maintained in switch table?

- something like a routing protocol?



switch with six interfaces (1,2,3,4,5,6)

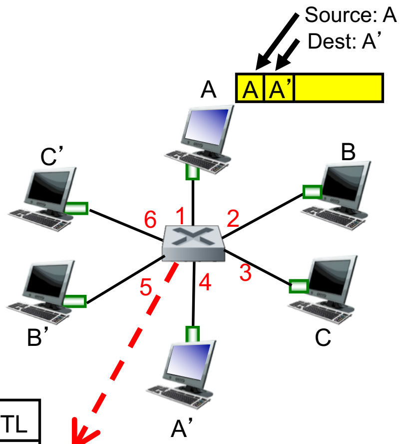
5-73

CHAPTER 6

Link-Layer Switches

Operation: Self-learning

- switch *learns* which hosts can be reached through which interfaces
  - when frame received, switch "learns" location of sender: incoming LAN segment
  - records sender/location pair in switch table



MAC addr.	Interface	TTL
A	*1*	*60*

Switch table (initially empty)

5-74

CHAPTER 6

Link-Layer Switches

Operation: Filtering / Forwarding

<u>When **frame** received at switch</u>:

1- record *incoming link*, *MAC address* of sending host
2- index switch table using MAC destination address

3- *if entry found for destination*
then {
if destination on segment from which frame arrived
then drop frame
else forward frame on interface indicated by entry
}
else flood / forward on all interfaces except arriving interface */*

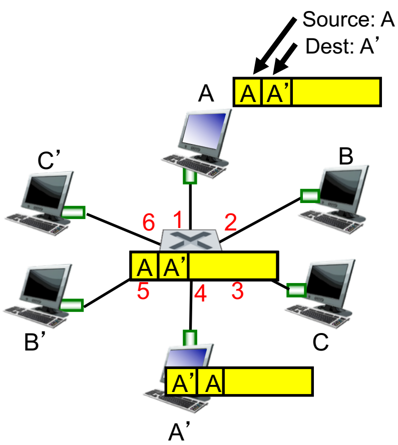
5-75

CHAPTER 6

Link-Layer Switches

Example: Self-learning / Forwarding

- frame destination, A', location unknown: *flood*
- destination A location known: *selectively send on just one link*



MAC addr.	Interface	TTL
A	*1*	*60*
A'	*4*	*60*

switch table (initially empty)

5-76

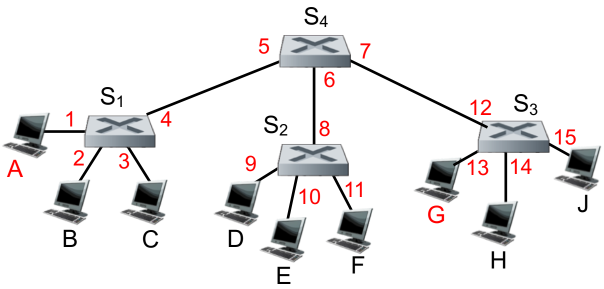
CHAPTER 6

Link-Layer Switches

Interconnecting switches

- switches can be connected together

Q: (Sending from A to G) - How does S_1 know to forward frame destined to G via S_4 and S_3?



A: ____________________

(works *exactly* the same as in single-switch case!)

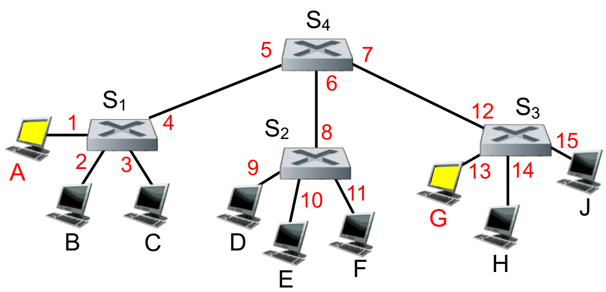
5-77

CHAPTER 6

Exercise 6.5

Suppose A sends frame to G, G responds to A.

Show the switch tables in S_1, S_3, S_4



Switch table S_1:

Mac addr.	Interface

Switch table S_4:

Mac addr.	Interface

Switch table S_3:

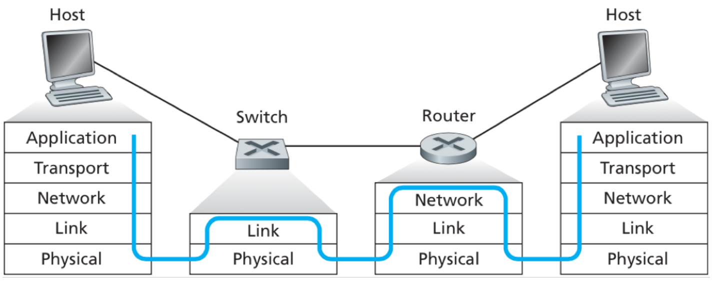
Mac addr.	Interface

5-78

CHAPTER 6

Link-Layer Switches

Switches vs. Routers



Both are

- *routers:* network-layer devices (examine network-layer headers)
- *switches:* link-layer devices (examine link-layer headers)

Both have

- *routers*: compute tables using routing algorithms, IP addresses
- *switches*: learn forwarding table using flooding, learning, MAC addresses

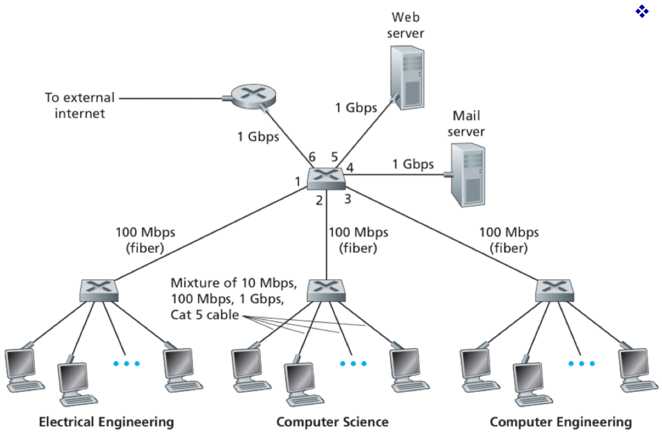
CHAPTER 6

Virtual Local Area Networks (VLAN)

Motivation

Consider **:**

- CS user moves office to EE, but wants connect to CS switch?



- single broadcast domain:
  - all layer-2 broadcast traffic (ARP, DHCP, unknown location of destination MAC address) must cross entire LAN
  - security/privacy, efficiency issues

Figure: An institutional network connected together by four switches.

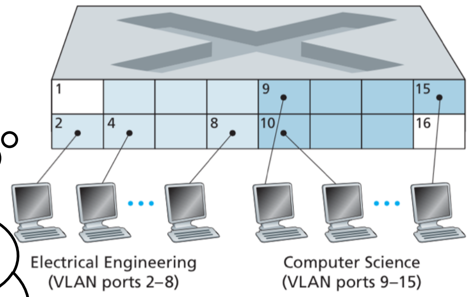
Solution

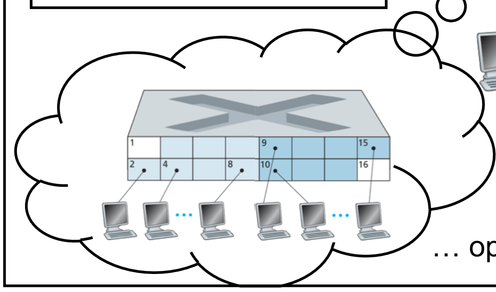
Virtual Local Area Networks (VLAN)

Virtual Local Area Network

Switch(es) supporting VLAN capabilities can be configured to define multiple virtual LANS over single physical LAN infrastructure.

Port-based VLAN: switch ports grouped (by switch management software) so that single physical switch ……





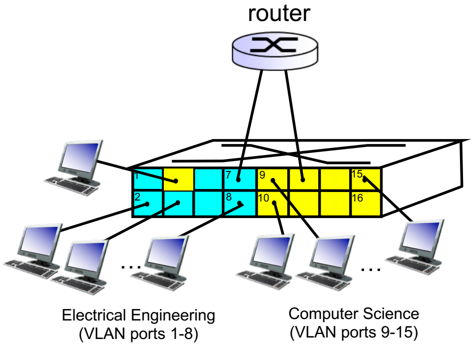
… operates as *multiple* virtual switches

CHAPTER 6

Virtual Local Area Networks (VLAN)

Port-based VLAN

- *traffic isolation:* frames to/from ports 1-8 can *only* reach ports 1-8
  - can also define VLAN based on MAC addresses of endpoints, rather than switch port
- *dynamic membership:* ports can be dynamically assigned among VLANs



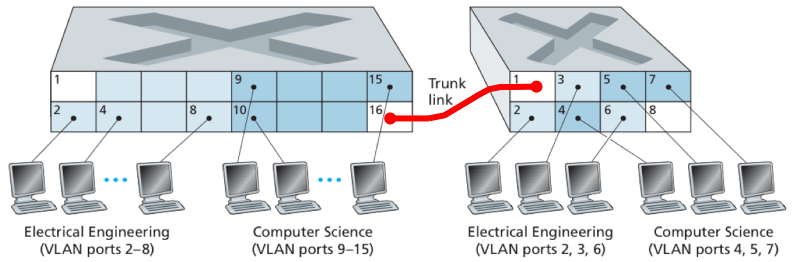
- *forwarding between VLANS:* done via routing (just as with separate switches)
  - in practice vendors sell combined switches plus routers

5-82

CHAPTER 6

Virtual Local Area Networks (VLAN)

VLAN spanning multiple switches



- _______________: carries frames between VLANS defined over multiple physical switches
  - frames forwarded within VLAN between switches can't be vanilla 802.1 frames (must carry VLAN ID info)
  - 802.1q protocol adds/removed additional header fields for frames forwarded between trunk ports

5-83

CHAPTER 6

Virtual Local Area Networks (VLAN)

802.1Q VLAN frame format

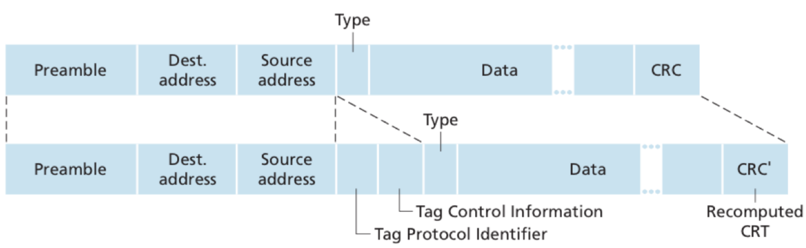


Figure: Original Ethernet frame (top),
802.1Q-tagged Ethernet VLAN frame (below).

5-84

CHAPTER 6

The Link Layer

Roadmap:

6.1 Introduction, services
6.2 Error detection, correction
6.3 Multiple access protocols
6.4 LANs
- Addressing, ARP
- Ethernet
- Switches
- VLANS

6.5 Data center networking
6.6 A day in the life of a web request

5-85

CHAPTER 6

(6.5) Data Center Networking

- 10's to 100's of thousands of hosts, often closely coupled, in close proximity:
  - e-business (e.g amazon)
  - content-servers (e.g. YouTube Akamai Microsoft)
  - search engines, data mining (e.g. Google)
- Challenges:
  - multiple applications, each serving massive numbers of clients
  - managing/balancing load, avoiding processing, networking, data bottlenecks

Inside a 40-ft Microsoft container, Chicago data center

CHAPTER 6

Load balancing

- receives external client requests
- directs workload within data center
- returns results to external client (hiding data center internals from client)

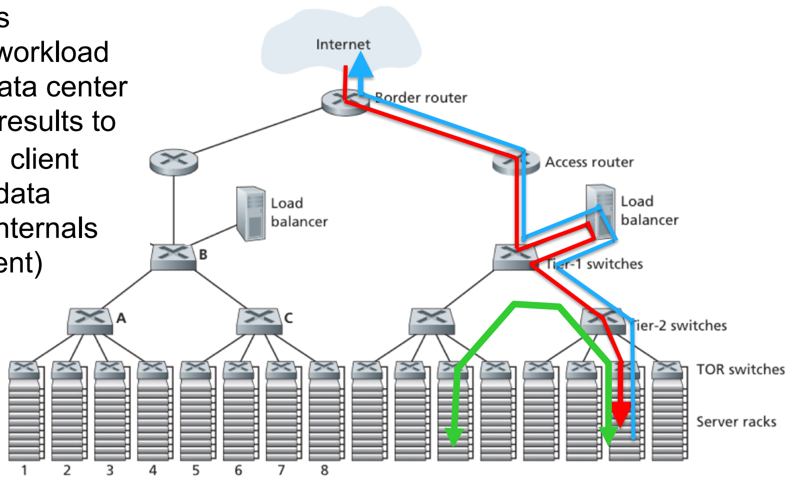


Figure: A data center network with a hierarchical topology.

CHAPTER 6

Trends in data center networking

- Rich interconnection among switches, racks:
  - increased throughput between racks (multiple routing paths possible)
  - increased reliability via redundancy

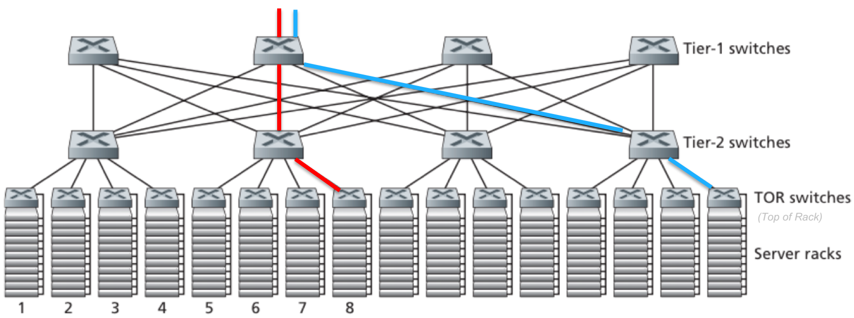


Figure: Highly-interconnected data network topology.

CHAPTER 6

The Link Layer

Roadmap:

6.1 Introduction, services

6.2 Error detection, correction

6.3 Multiple access protocols

6.4 LANs
- Addressing, ARP
- Ethernet
- Switches
- VLANS

6.5 Data center networking

6.6 A day in the life of a web request

5-89

CHAPTER 5

(6.6) Retrospective: A Day in the Life of a Web Page Request

Synthesis

- Journey down protocol stack complete!
  - application, transport, network, link
- putting-it-all-together: **synthesis**!
  - *goal:* identify, review, understand protocols (at all layers) involved in seemingly simple scenario: requesting `www` page
  - *scenario:* student attaches laptop to campus network, requests/receives `www.google.com`

5-90

CHAPTER 6

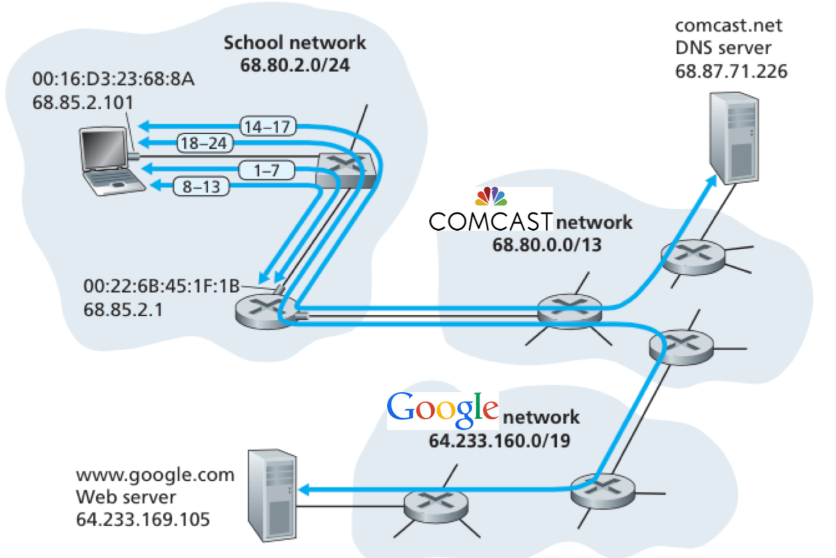
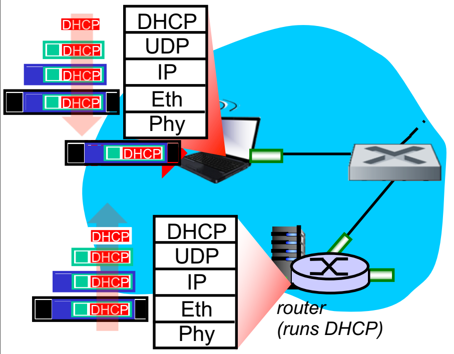


Figure: A day in the life of a Web page request: network setting and actions

CHAPTER 6

Connecting to the Internet

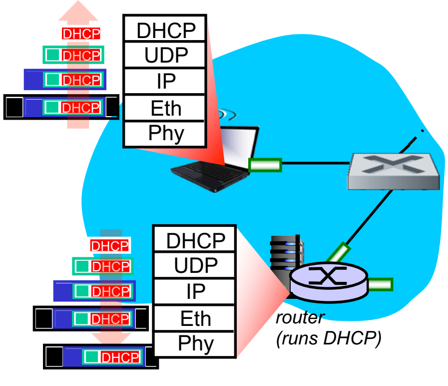


- connecting laptop needs to get its own IP address, addr of first-hop router, addr of DNS server: use *DHCP*
- DHCP request *encapsulated* in *UDP*, encapsulated in *IP*, encapsulated in *802.3* Ethernet
- Ethernet frame *broadcast* (dest: `FF-FF-FF-FF-FF-FF`) on LAN, received at router running *DHCP* server
- Ethernet *demuxed* to IP demuxed, UDP demuxed to DHCP

5-92

CHAPTER 6

Connecting to the Internet



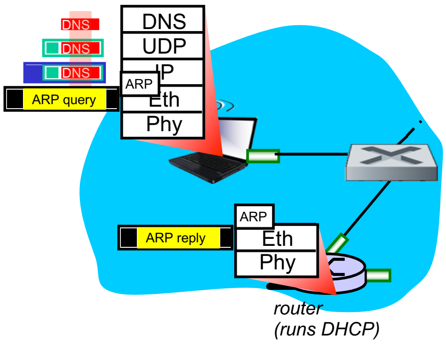
- DHCP server formulates *DHCP ACK* containing client's IP address, IP address of first-hop router for client, name & IP address of DNS server
- encapsulation at DHCP server, frame forwarded (*switch learning*) through LAN, demultiplexing at client
- DHCP client receives DHCP ACK reply

Client now has IP address, knows name & address of DNS server, IP address of its first-hop router

5-93

CHAPTER 6

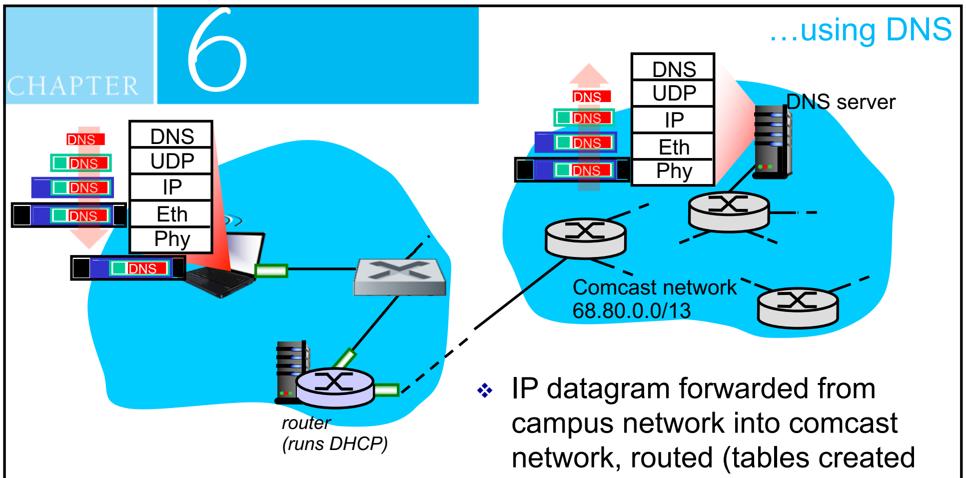
ARP (before DNS, HTTP)



- before sending *HTTP* request, need IP address of `www.google.com`: *DNS*
- DNS query created, encapsulated in UDP, encapsulated in IP, encapsulated in Eth. To send frame to router, need MAC address of router interface: *ARP*
- *ARP query* broadcast, received by router, which replies with *ARP reply* giving MAC address of router interface
- client now knows MAC address of first hop router, so can now send frame containing DNS query

5-94

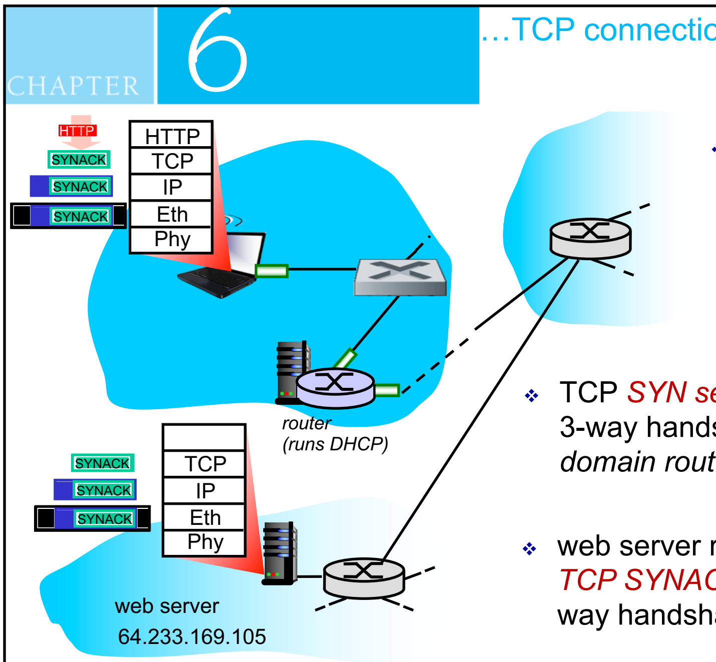
…using DNS



- IP datagram containing DNS query forwarded via LAN switch from client to 1st hop router
- IP datagram forwarded from campus network into comcast network, routed (tables created by *RIP, OSPF, IS-IS* and/or *BGP* routing protocols) to DNS server
- demux'ed to DNS server
- DNS server replies to client with IP address of `www.google.com`

5-95

…TCP connection carrying HTTP

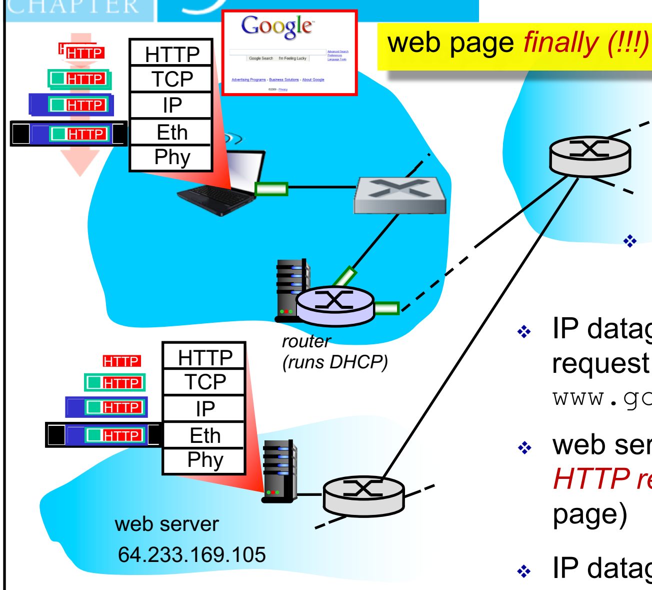


- to send HTTP request, client first opens *TCP socket* to web server
- TCP *SYN segment* (step 1 in 3-way handshake) *inter-domain routed* to web server
- web server responds with *TCP SYNACK* (step 2 in 3-way handshake)
- TCP *connection established!*

CHAPTER 5

…HTTP request / response

web page *finally (!!!)* displayed



- *HTTP request* sent into TCP socket
- IP datagram containing HTTP request routed to `www.google.com`
- web server responds with *HTTP reply* (containing web page)
- IP datagram containing HTTP reply routed back to client

CHAPTER 6

Summary

- principles behind data link layer services:
  - error detection, correction
  - sharing a broadcast channel: multiple access
  - link layer addressing
- instantiation and implementation of various link layer technologies
  - Ethernet
  - switched LANS, VLANs
  - virtualized networks as a link layer: MPLS
- synthesis: a day in the life of a web request

5-98

CHAPTER 6

Summary

- journey down protocol stack *complete* (except PHY)
- solid understanding of networking principles, practice
- ….. could stop here …. but *lots* of interesting topics!
  - wireless
  - multimedia
  - security
  - network management

5-99