EECS 489
Computer Networks

Fall 2020

Mosharaf Chowdhury

Material with thanks to Aditya Akella, Sugih Jamin, Philip Levis, Sylvia Ratnasamy, Peter Steenkiste, and many other colleagues.

Logistics

- Open book/text/notes, but OFFLINE
  - Except for taking the exam over the Internet
- You’re NOT allowed to write/run any programs
- You’re NOT allowed to collaborate with anyone

General guidelines (1)

- Test only assumes material covered in lecture, sections, and assignments after midterm
  - Text: only to clarify details and context for the above
- The test doesn't require you to do complicated calculations
  - Use this as a hint to determine if you're on right track
- You don't need to memorize anything
- You do need to understand how things work

General guidelines (2)

- Be prepared to:
  - Weigh design options outside of the context we studied them in
  - Contemplate new designs we haven't covered in detail but can be put together
    - e.g., I introduce a new IP address format; how does this affect..”
  - Reason from what you know about the pros/cons of solutions we did study

General guidelines (3)

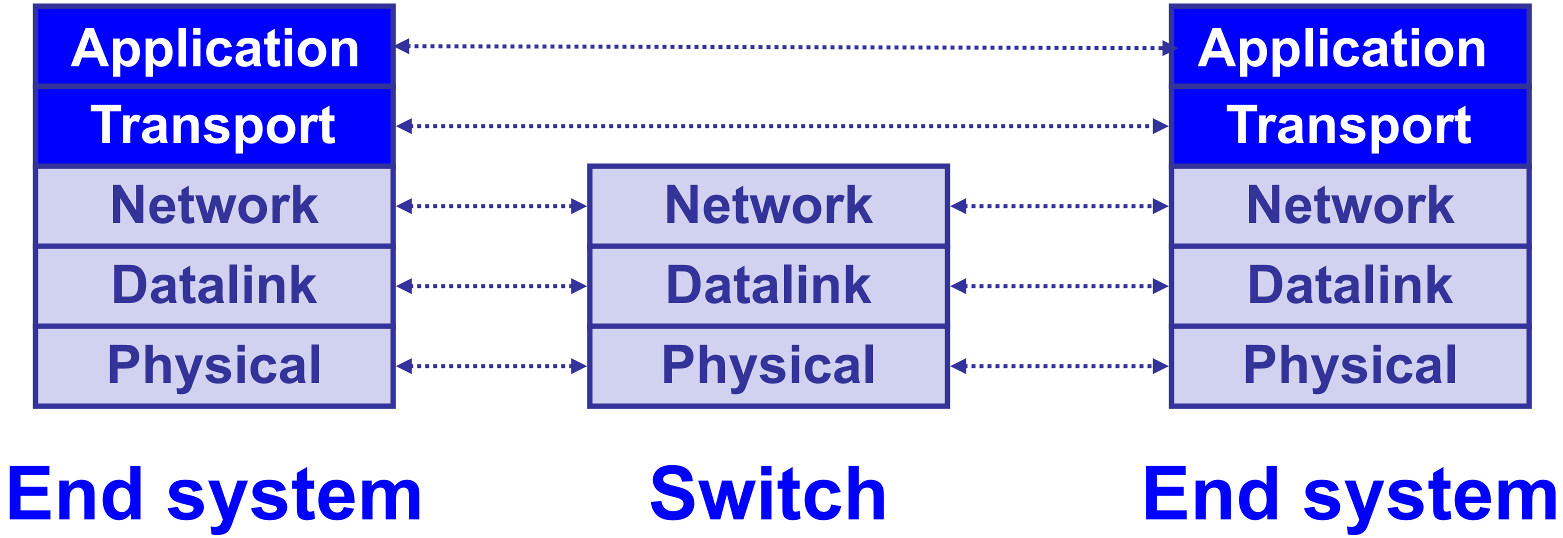
- Exam format
  - Like midterm, but we're working to avoid cascading mistakes
- Questions not ordered in terms of complexity
    - Read all carefully
- Pace yourself accordingly!

This review

- Walk through what you're expected to know since the midterm: key topics, important aspects of each
- Not covered in review does NOT imply you don't need to know it
  - But if it's covered today, you should know it
- Summarize, not explain
  - Stop me when you want to discuss something further!

The networking stack

- Lower three layers implemented everywhere
- Top two layers implemented only at hosts



Topics

- Link layer (lectures 17–19)
  - Ethernet
  - Wireless
- Datacenter networking (lectures 20)

Data link layer

- Provides four primary services
  - Framing
    - » Encapsulates network layer data
  - Link access
    - » Medium access control (MAC) protocol defines when to transmit frames
  - Reliable delivery
    - » Primarily for mediums with high error rates (e.g., wireless)
  - Error detection and correction

Point-to-point vs. broadcast medium

- Point-to-point: dedicated pairwise communication
  - E.g., long-distance fiber link
  - E.g., Point-to-point link b/n Ethernet switch and host
- Broadcast: shared wire or medium
  - Traditional Ethernet (pre ~2000)
  - 802.11 wireless LAN

Random access MAC protocols

- When node has packet to send
  - Transmit at full channel data rate **w/o** coordination
- Two or more transmitting nodes $\Rightarrow$ collision
  - Data lost
- Random access MAC protocol specifies
  - How to detect and recover from collisions
- Examples
  - ALOHA and Slotted ALOHA
  - CSMA, CSMA/CD, CSMA/CA (wireless)

CSMA (Carrier Sense Multiple Access)

- CSMA: listen before transmit
  - If channel sensed idle: transmit entire frame
  - If channel sensed busy, defer transmission
- Human analogy: don't interrupt others!
- Does not eliminate all collisions
  - Why?

CSMA/CD (Collision Detection)

- CSMA/CD: carrier sensing, deferral as in CSMA
  - Collisions detected within short time
  - Colliding transmissions aborted, reducing wastage
- Collision detection easy in wired (broadcast) LANs
  - Compare transmitted, received signals
- Collision detection difficult in wireless LANs

Limits on CSMA/CD network length

- Latency depends on physical length of link
  - Time to propagate a frame from one end to other
- Suppose A sends a frame at time **t**
  - And B sees an idle line at a time just before **t + d**
  - … so B happily starts transmitting a frame
- B detects a collision, and sends jamming signal
  - But A cannot see collision until **t + 2d**

Limits on CSMA/CD network length

- A needs to wait for time **2d** to detect collision
  - So, A should keep transmitting during this period
  - AND keep an eye out for a possible collision
- Imposes restrictions; e.g., for 10 Mbps Ethernet
  - Maximum length of the wire: 2,500 meters
  - Minimum length of a frame: 512 bits (64 bytes)

Efficiency of CSMA/CD

- $d_{prop} \to 0$
  - Efficiency approaches 1
  - Colliding nodes abort immediately
- $d_{trans} \to \infty$
  - Efficiency approaches 1
  - Each frames uses the channel for a long time

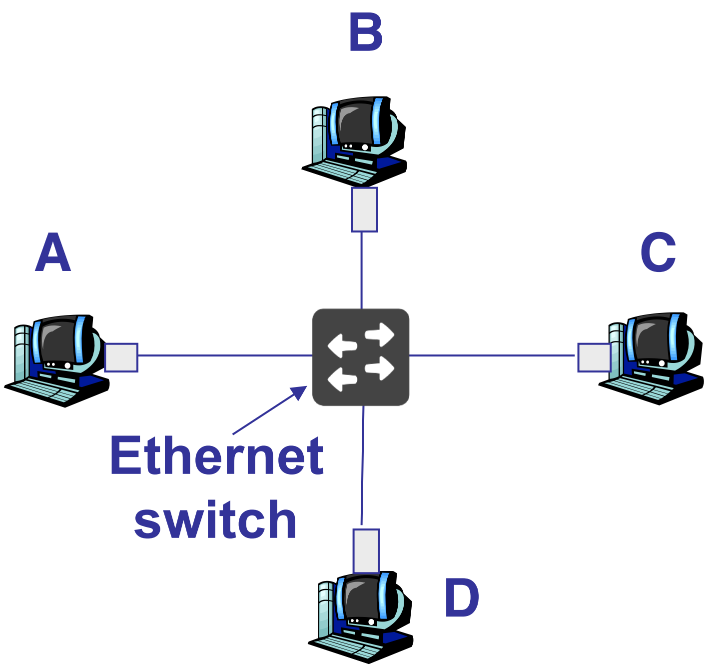
$$\text{Efficiency} \approx \frac{d_{trans}}{d_{trans} + 5\, d_{prop}}$$

Spanning tree protocol (Perlman'85)

- Protocol by which bridges construct a spanning tree
- Nice properties
  - Zero configuration (by operators or users)
  - Self healing
- Still used today

Why switched Ethernet?

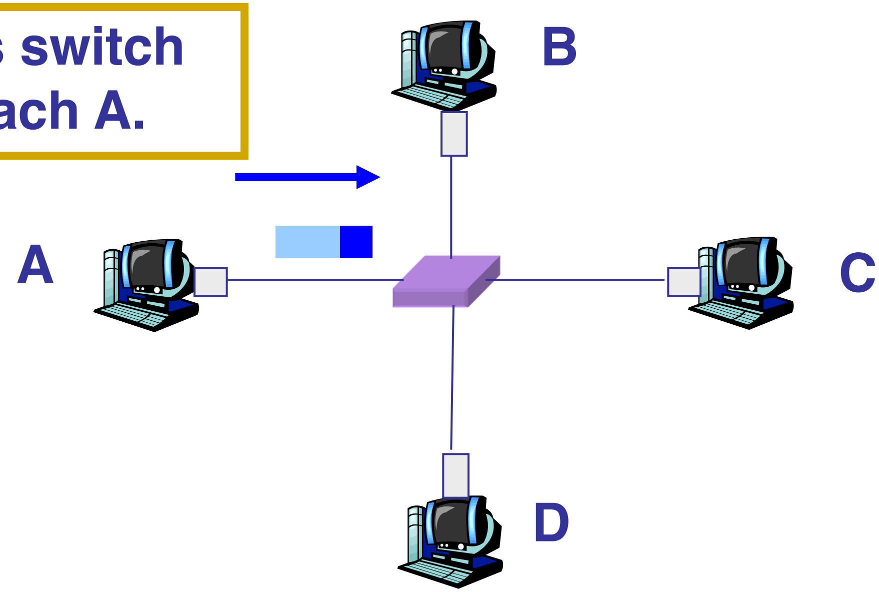
- Enables concurrent communication
  - Host A can talk to C, while B talks to D
  - No collisions and no need for CSMA/CD
  - No constraints on link lengths, etc.



Ethernet switches are "self learning"

- When a packet arrives:
  - Inspect source MAC address, associate with incoming port
  - Store mapping in the switch table
  - Use time-to-live field to eventually forget mapping

Packet tells switch how to reach A.



MAC address vs. IP address

MAC Addresses

- Hard-coded when adapter is built
- Flat name space of 48 bits (e.g., 00-0E-9B-6E-49-76)
- Like a social security number
- Portable, and can stay the same as the host moves
- Used to get packet between interfaces on same network

IP Addresses

- Configured, or learned dynamically
- Hierarchical name space of 32 bits (e.g., 12.178.66.9)
- Like a postal mailing address
- Not portable, and depends on where the host is attached
- Used to get a packet to destination IP subnet

Ethernet vs. IP

IP

- Packets forwarded on all available links
- Addresses can be aggregated
- Routing protocol computes loop-free paths
- Forwarding table computed by routing protocol

Ethernet

- Packets forwarded on subset of links (spanning tree)
- Flat addresses
- “Routing” protocol computes loop-free topology
- Forwarding table derived from data packets(+ spanning tree for floods)

5-MINUTE BREAK!

Announcements

- Teaching evaluations
  - Due by Dec 9
  - 75% or higher completion rate will result in +1 on the final grade for everyone
    - » We are only at **35%**
- Sample final exam on github!
- Final slot sign up by Dec 8
  - https://forms.gle/znKWqTvyjjnnVbws6
  - Defaults to 8AM EST Dec 16

ARP and DHCP

- Link layer discovery protocols
  - ARP → Address Resolution Protocol
  - DHCP → Dynamic Host Configuration Protocol
  - Confined to a single local-area network (LAN)
  - Rely on broadcast capability

Key ideas in both ARP and DHCP

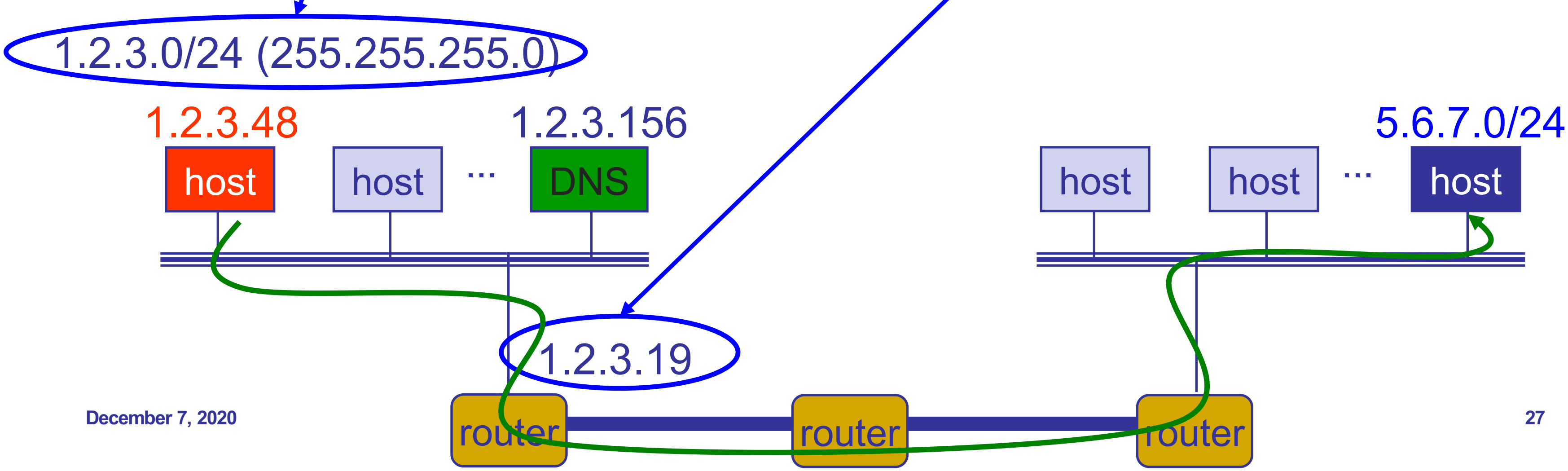
- Broadcasting: Can use broadcast to make contact
  - Scalable because of limited size
- Caching: remember the past for a while
  - Store the information you learn to reduce overhead
- Soft state: eventually forget the past
  - Associate a time-to-live field with the information
  - … and either refresh or discard the information
  - Key for robustness in the face of unpredictable change

ARP: Address Resolution Protocol

- Every host maintains an ARP table
  - List of (IP address → MAC address) pairs
- Consult the table when sending a packet
  - Map dest. IP address to dest. MAC address
  - Encapsulate (IP) data packet with MAC header; xmit
- What if IP address not in the table?
  - Sender broadcasts: Who has IP address 1.2.3.156?
  - Receiver replies: MAC address 58-23-D7-FA-20-B0
  - Sender caches result in its ARP table

What if the destination is remote?

- Look up the MAC address of the first hop router
  - 1.2.3.48 uses ARP to find MAC address for first-hop router **1.2.3.19** rather than ultimate destination IP address
- How does the red host know the destination is not local?
  - Uses netmask (discovered via DHCP)
- How does the red host know about 1.2.3.19?
  - Also DHCP



Wireless link characteristics

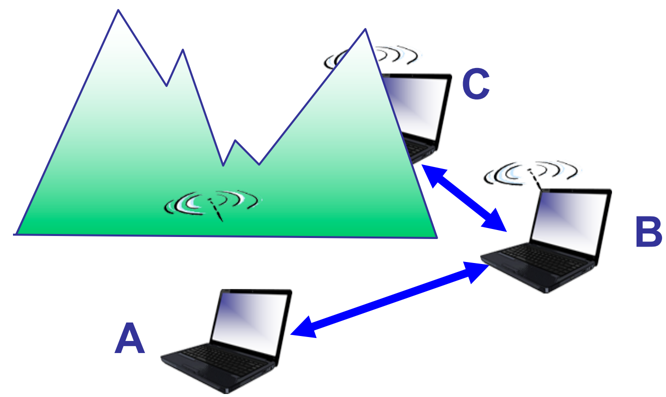
- Three important differences from wired link …
  - Decreased signal strength: Radio signal attenuates as it propagates through matter (path loss)
  - Multipath propagation: Radio signal reflects off objects ground, arriving at destination at slightly different times
  - Interference from other sources: Standardized wireless network frequencies (e.g., 2.4 GHz) shared by other devices (e.g., phone); devices (motors) interfere as well
- … make communication across (even a point-to-point) wireless link much more "difficult"

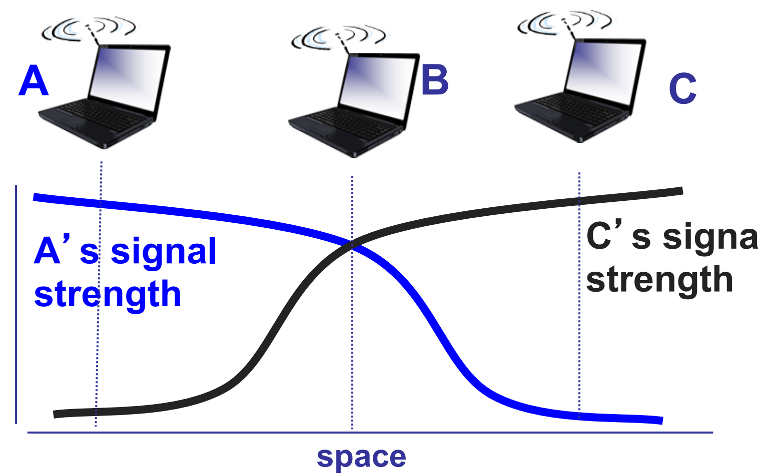
Wireless network characteristics

- Multiple wireless senders and receivers create many problems
  - Multiple access issues (we’ve seen this before)
  - Hidden terminal problem

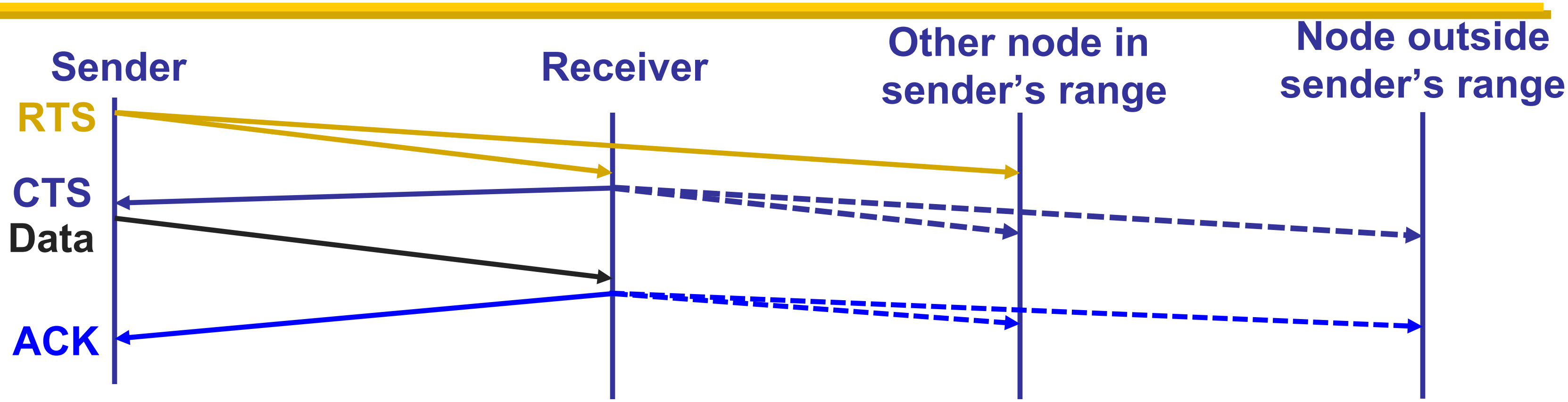
Hidden terminal problem

- B, A hear each other
- B, C hear each other
- A, C can not hear each other
- Hence, A, C are unaware of their interference at B



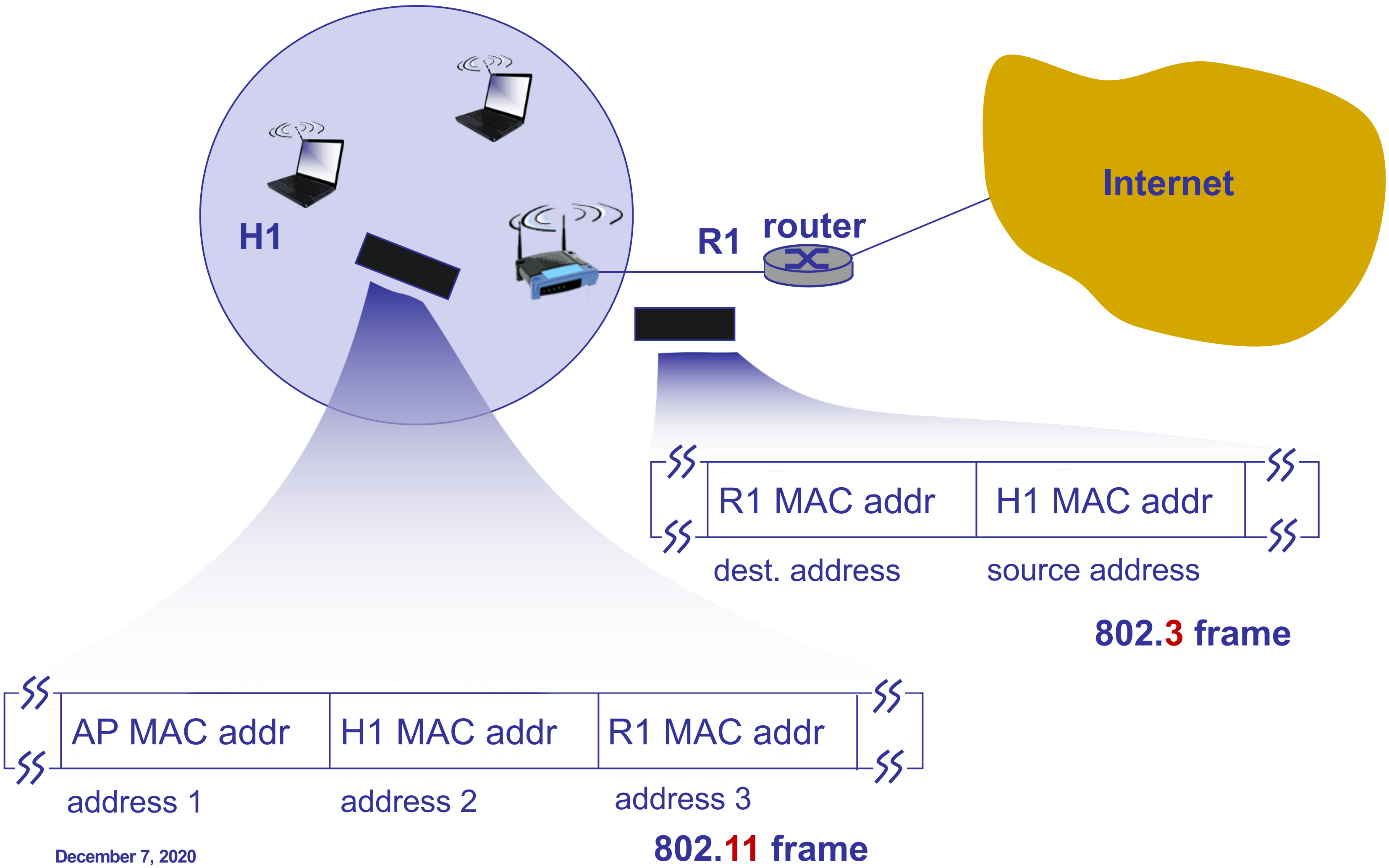


CSMA/CA



- Before every data transmission
  - Sender sends a Request to Send (RTS) frame with the length of transmission and the destination
  - Receiver respond with a Clear to Send (CTS) frame
  - Sender sends data
  - Receiver sends an ACK
- If sender doesn't get a CTS back, it assumes collision

Why do we need Address 3?



Topics

- Link layer (lectures 17–19)
  - Ethernet
  - Wireless
- Datacenter networking (lectures 20)

Datacenter applications

- Common theme: parallelism
  - Applications decomposed into tasks
  - Running in parallel on different machines
- Two common paradigms
  - Partition-Aggregate
  - Map-Reduce

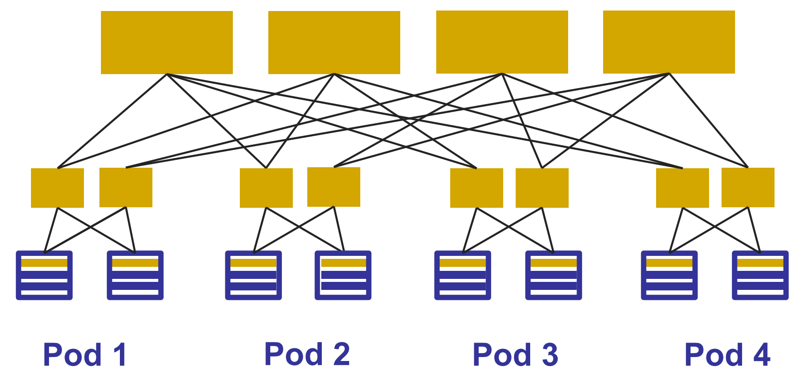
Datacenter traffic characteristics

- Two key characteristics
  - Most flows are small
  - Most bytes come from large flows
- Applications want
  - High bandwidth (large flows)
  - Low latency (small flows)

Clos topology

- Multi-stage network
- k pods, where each pod has two layers of k/2 switches
  - k/2 ports up and k/2 down
- All links have the same b/w
- At most $k^3/4$ machines

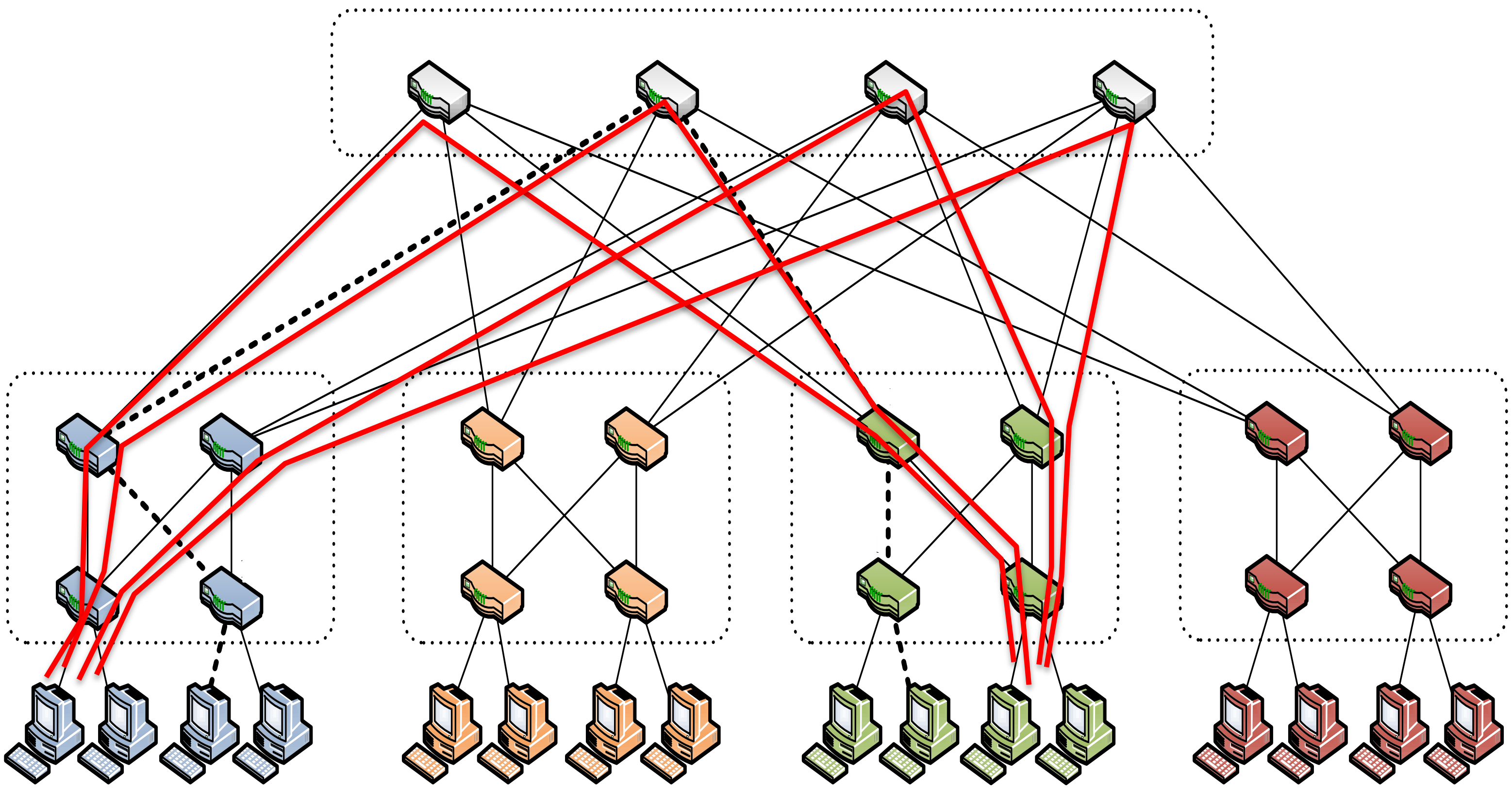
- Example
  - k = 4
  - 16 machines
- For k=48, 27648 machines



Datacenter networking stack

- Networking in modern datacenters
  - L2/L3 design
    - Addressing / routing / forwarding in the Fat-Tree
  - L4 design
    - Transport protocol design (w/ Fat-Tree)
  - L7 design
    - Exploiting application-level information (w/ Fat-Tree)

Using multiple paths well



L2/L3 highlights

- Load balancing while forwarding
  - Per-packet
  - Per-flow
- Hard-coded addressing or via indirection
- Modified LS/DV or source routing

L4 highlights

- Tension between high throughput and low latency requirements
  - Deep queues vs shallow queues
- DCTCP
  - React early, quickly, and with certainty using ECN
  - React in proportion to the extent of congestion, not its presence

L7 highlights

- What do applications care about?
  - Flow completion time (FCT)
  - Coflow completion time (CCT)
    - A coflow is a collection of flows with a shared application-level objective
  - We should strive to optimize as close an objective as possible to the application

Summary

- THANK YOU SO MUCH!!!