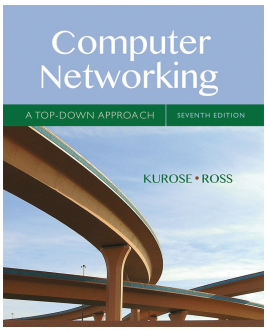
Chapter 6: The Link Layer and LANs

Our goals:

- understand principles behind data link layer services:
  - error detection, correction
  - sharing a broadcast channel: multiple access
  - link layer addressing
  - local area networks: Ethernet, VLANs
- instantiation and implementation of various link layer technologies



Computer Networking: A Top Down Approach
7th edition
Jim Kurose, Keith Ross
Pearson/Addison-Wesley
April 2016

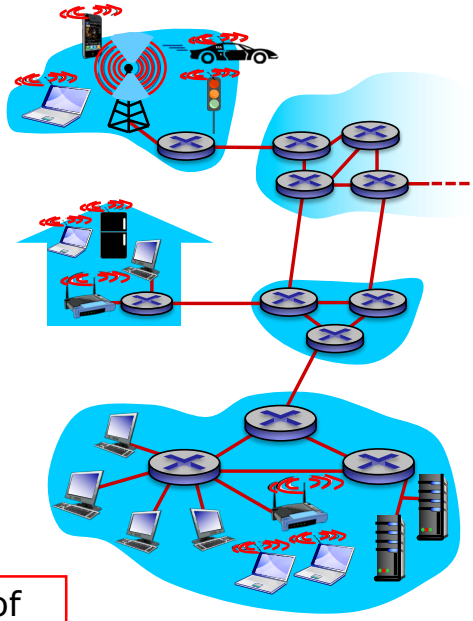
Link Layer, LANs: Outline

- 6.1 Introduction and services
- 6.2 Error detection and correction
- 6.3 Multiple access protocols
- 6.4 LANs
  - addressing, ARP
  - Ethernet
  - switches
  - VLANs
- 6.5 Link virtualization: MPLS
- 6.6 Data center networking
- 6.7 A day in the life of a web request

Link Layer: Introduction

Terminology:

- hosts and routers are **nodes**
- communication channels that connect adjacent nodes along communication path are **links**
  - wired links
  - wireless links
  - LANs
- layer-2 packet is a **frame**, encapsulates datagram



data-link layer has responsibility of transferring datagram from one node to physically adjacent node over a link

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 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 = datagram
- transport segment = communication link
- transportation mode = link layer protocol
- travel agent = routing algorithm

Link Layer Services

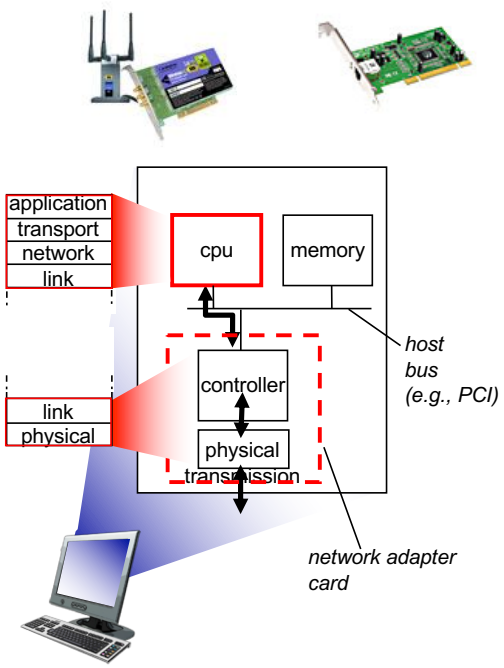
- Framing, link access
  - encapsulate datagram into frame, adding header, trailer
  - multiple access if shared medium
  - "MAC" addresses used in frame headers to identify source, destination
    - ✧ different from IP address
- Reliable delivery between adjacent nodes
  - we learned how to do this already (chapter 3)!
  - seldom used on low bit error link (fiber, some twisted pair)
  - wireless links: high error rates
    - ✧ Q: why both link-level and end-end reliability?

Link Layer Services (more)

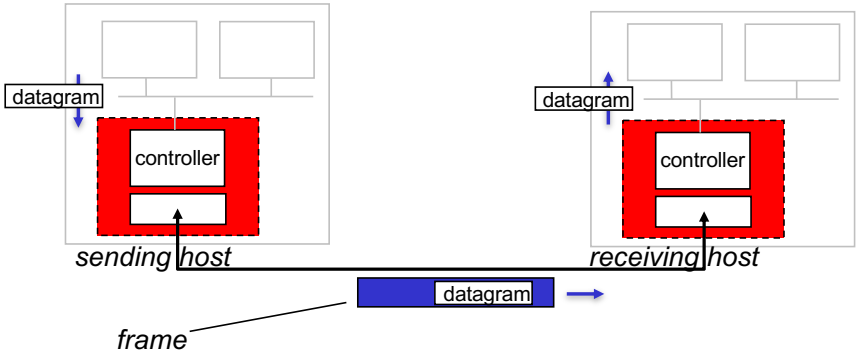
- Flow Control
  - pacing between adjacent sending and receiving nodes
- Error Detection
  - errors caused by signal attenuation, noise.
  - receiver detects presence of errors:
    - ✧ signals sender for retransmission or drops frame
- Error Correction
  - receiver identifies *and corrects* bit error(s) 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

Where is the link layer implemented?

- In each and every host
- Link layer implemented in "adaptor" (aka network interface card NIC) or on a chip
  - Ethernet card, 802.11 card, Ethernet chipset
  - implements link, physical layer
- Attaches into host's system buses
- Combination of hardware, software, firmware



Adaptors Communicating

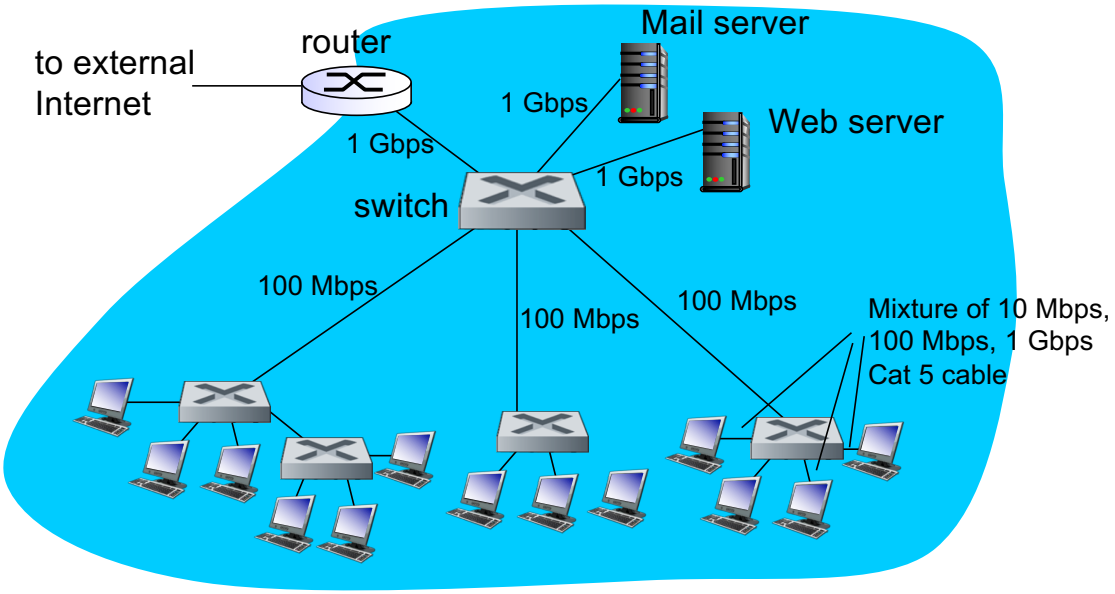


- Sending Side
  - encapsulates datagram in a frame
  - adds error checking bits, rdt, flow control, etc.
- Receiving Side
  - looks for errors, rdt, flow control
  - extracts datagram, passes to receiving node

Link Layer, LANs: Outline

- 6.1 Introduction and services
- 6.2 Error detection and correction
- 6.3 Multiple access protocols
- 6.4 LANs
  - addressing, ARP
  - Ethernet
  - switches
  - VLANs
- 6.5 Link virtualization: MPLS
- 6.6 Data center networking
- 6.7 A day in the life of a web request

Institutional network



IP and MAC Addresses

32-bit IP (Internet Protocol) address:

- *network-layer* address, assign to each network interface
- 32-bit (i.e., 4 bytes) for IPv4
  - e.g., 137.82.52.194
- Function: get datagram to the destination IP network

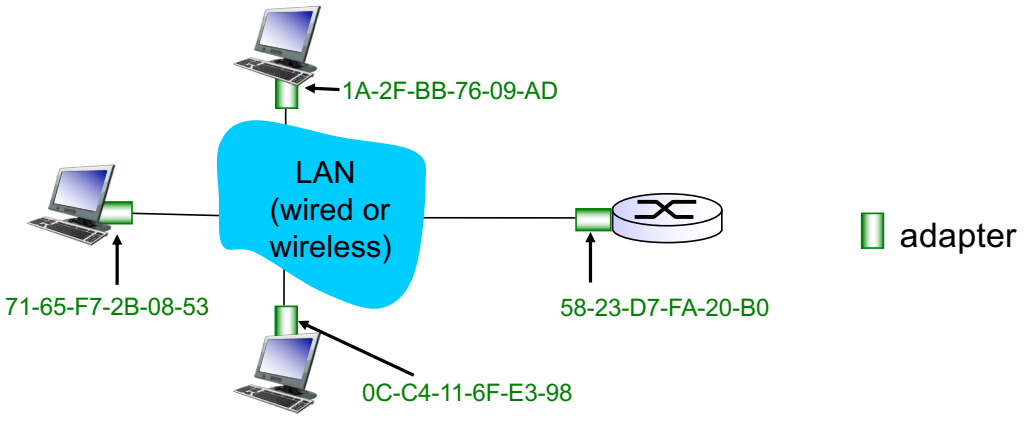
LAN (or MAC or physical or Ethernet) address:

- Each adapter (i.e., NIC) has a link-layer address
- Function: get datagram from one interface to another physically-connected interface (same network)
- 48 bit (i.e., 6 bytes) MAC 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)

MAC addresses

each adapter on LAN has a unique MAC address



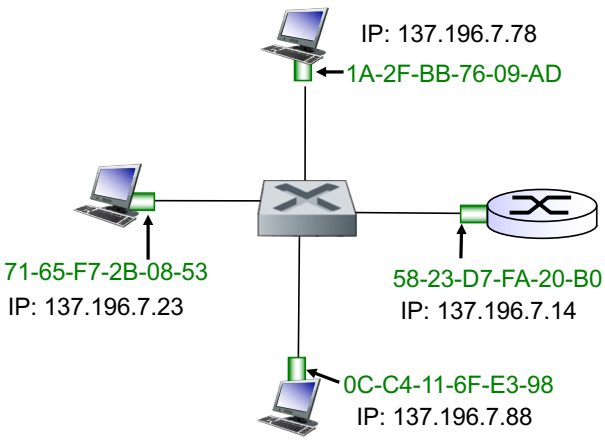
Broadcast address = FF-FF-FF-FF-FF-FF (i.e., 48 consecutive 1s)

MAC Address (more)

- MAC address allocation administered by IEEE
- manufacturer buys portion of MAC address space (to assure uniqueness)
  - IEEE allocates the chunk of 2^{24} addresses by fixing the first 24 bits of a MAC address and let the company create unique combinations of the last 24 bits for each adapter.
- analogy:
  - MAC address: Social Security Number (flat addr structure)
  - IP address: like postal address (hierarchical addr structure)
- MAC flat address ➜ portability
  - can move LAN card from one LAN to another
- IP hierarchical address NOT portable
  - depends on IP subnet to which node is attached

ARP: Address Resolution Protocol

Question: how to determine interface's MAC address, knowing its IP address?



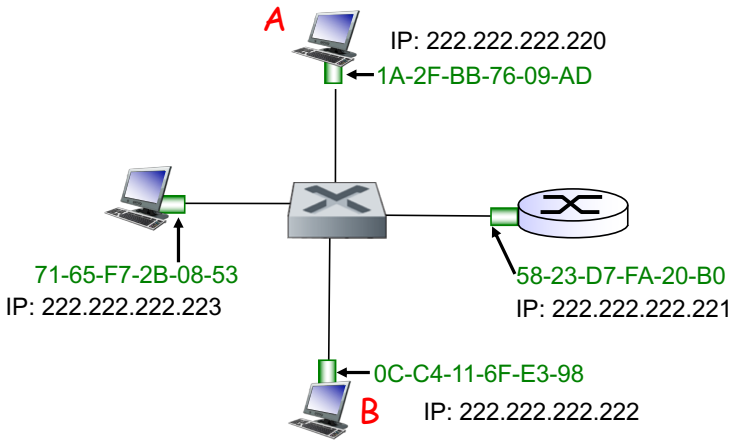
ARP table: each adapter in an IP node (host, router) on LAN has a table

- 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)

ARP: Address Resolution Protocol



- A possible ARP table in node *A*

IP Address	MAC Address	TTL
222.222.222.221	58-23-D7-FA-20-B0	13:45:00
222.222.222.223	71-65-F7-2B-08-53	13:52:00

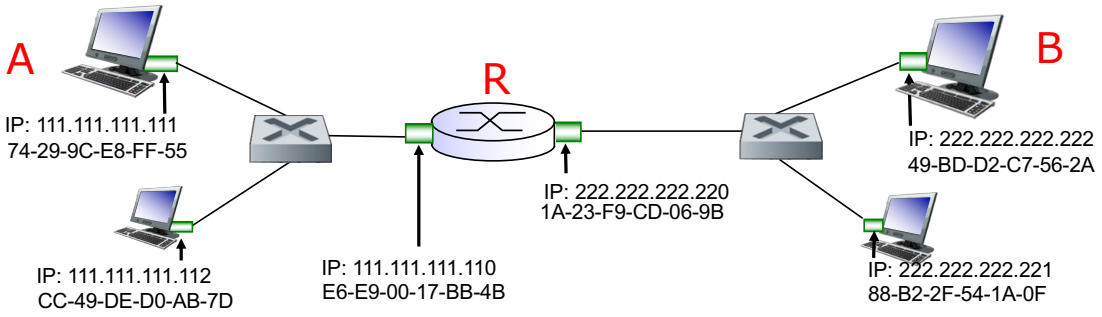
ARP (Address Resolution Protocol) [RFC 826]

- A wants to send datagram to B, and A knows B's IP address.
- Suppose B's MAC address *is not* in A's ARP table.
- A broadcasts ARP query packet, containing B's IP address and A's MAC address
  - all nodes on LAN receive ARP query
- B receives ARP packet, replies to A with its (B's) MAC addr
  - frame sent to A's MAC address (unicast)
- 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 net administrator.

Addressing: routing to another LAN

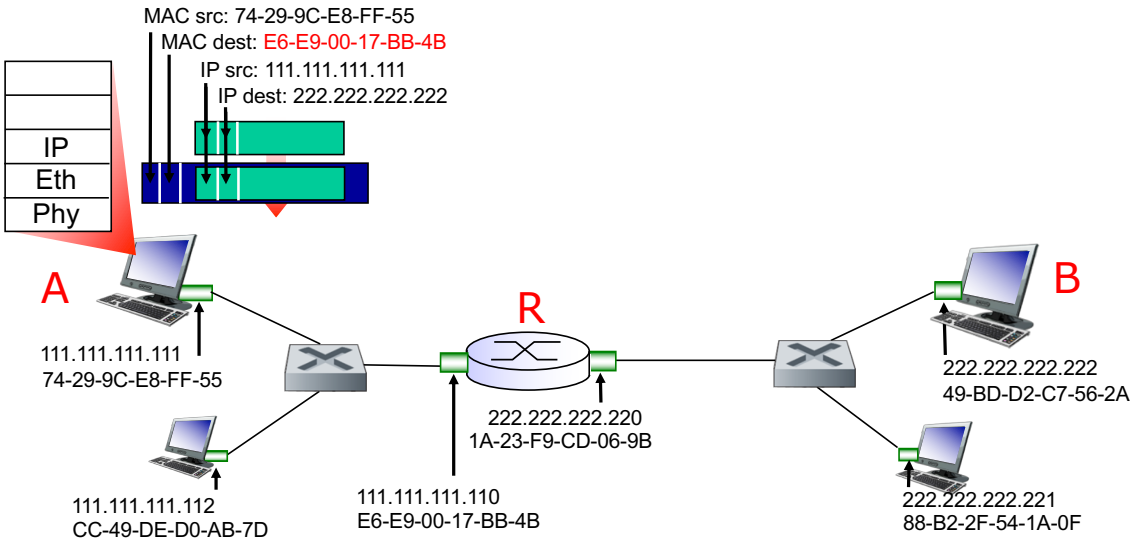
walkthrough: send datagram from A to B via router R

- focus on addressing – at IP (datagram) and MAC layer (frame)
- assume A knows B's IP address
- assume A knows IP address of first hop router, R (how?)
- assume A knows R's MAC address (how?)



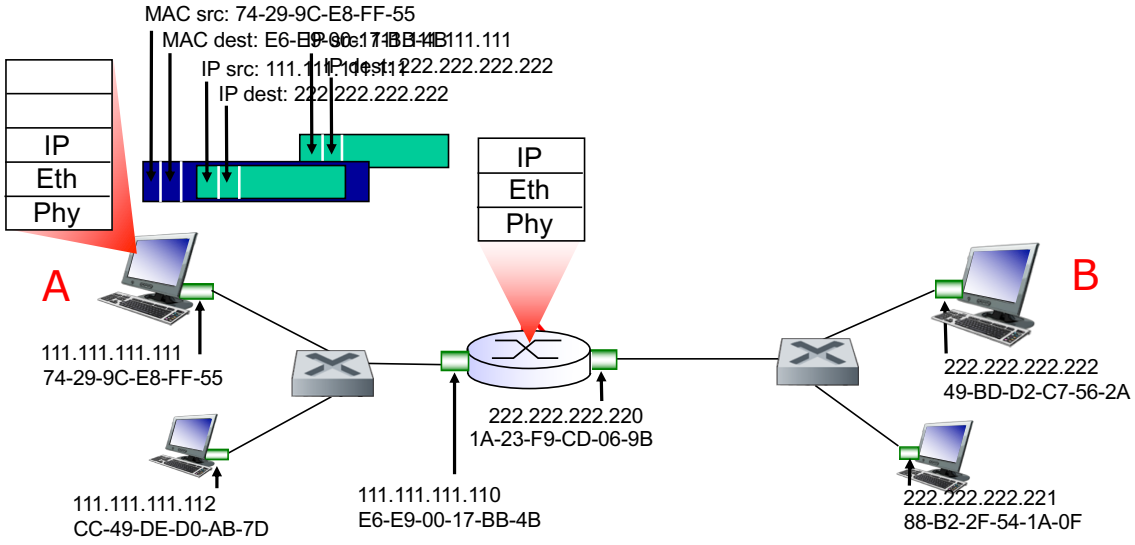
Addressing: routing to another LAN

- A creates IP datagram with IP source A, destination B
- A creates link-layer frame with R's MAC address as destination, frame contains A-to-B IP datagram



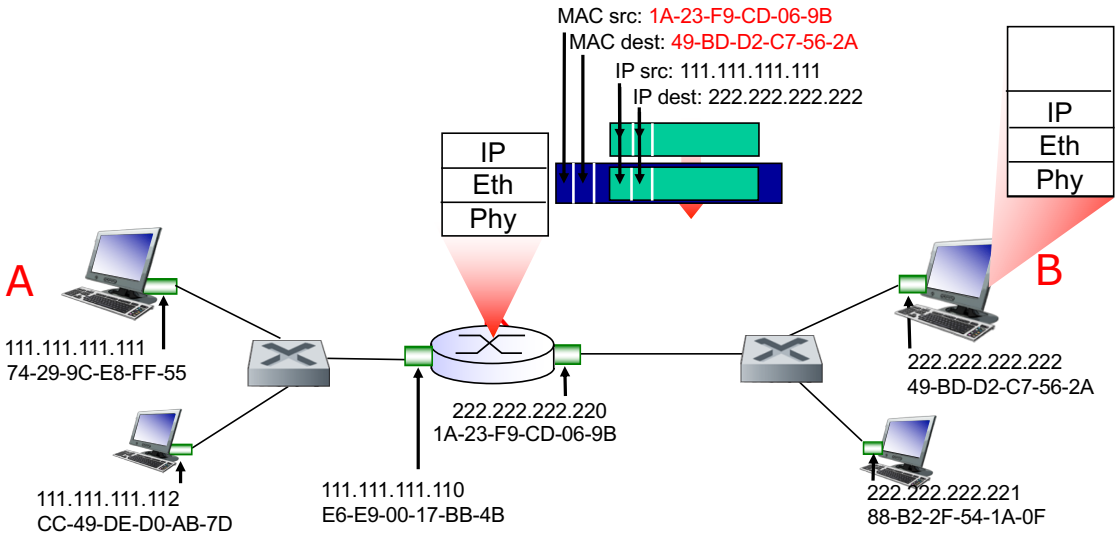
Addressing: routing to another LAN

- frame sent from A to R
- frame received at R, datagram removed, passed up to IP



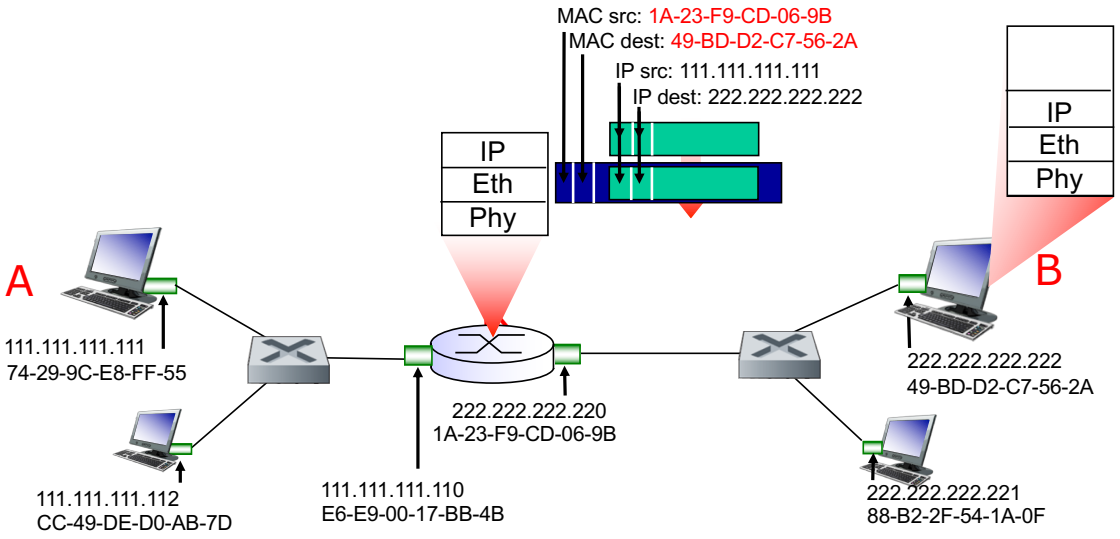
Addressing: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as dest, frame contains A-to-B IP datagram



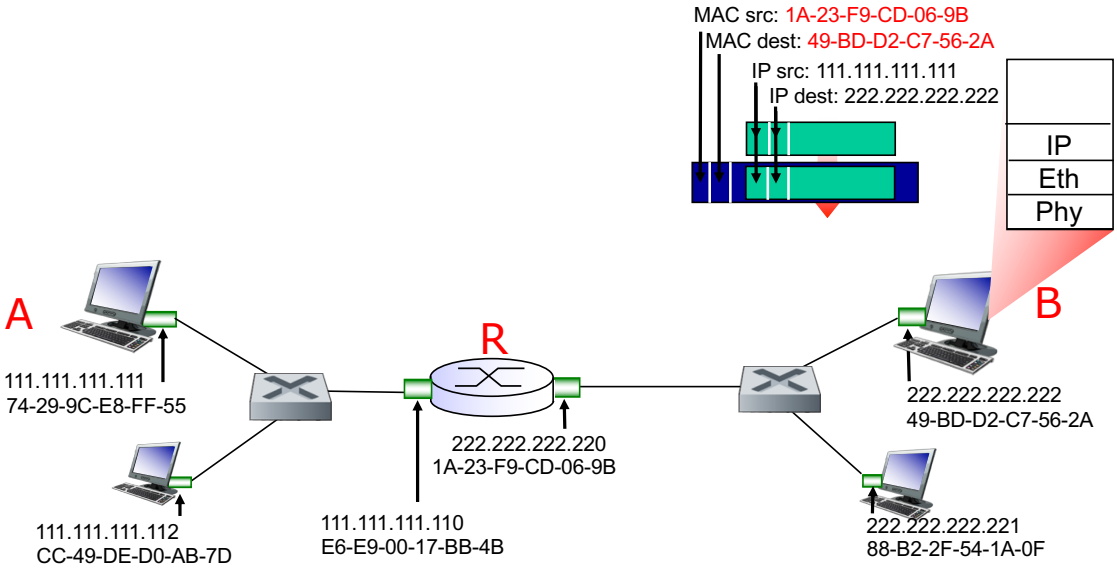
Addressing: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as dest, frame contains A-to-B IP datagram



Addressing: routing to another LAN

- R forwards datagram with IP source A, destination B
- R creates link-layer frame with B's MAC address as dest, frame contains A-to-B IP datagram



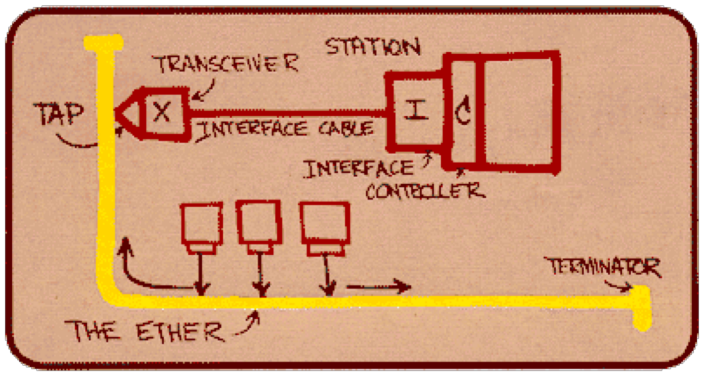
Link Layer, LANs: Outline

- 6.1 Introduction and services
- 6.2 Error detection and correction
- 6.3 Multiple access protocols
- 6.4 LANs
  - addressing, ARP
  - Ethernet
  - switches
  - VLANs
- 6.5 Link virtualization: MPLS
- 6.6 Data center networking
- 6.7 A day in the life of a web request

Ethernet

"dominant" wired LAN technology:

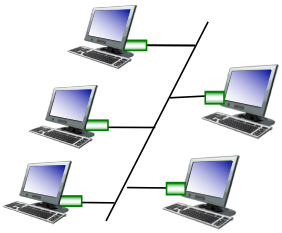
- first widely used LAN technology
- Simpler, cheaper than token LANs and ATM
- Kept up with speed race: 10 Mbps – 10 Gbps



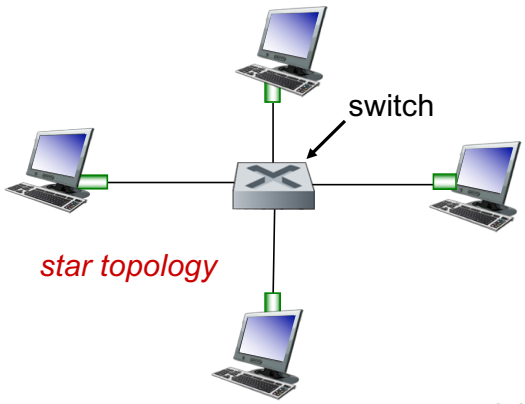
Bob Metcalfe's Ethernet sketch

Ethernet: physical topology

- Bus topology: popular through mid 90s
  - all nodes in same collision domain (can collide with each other)
- Star topology: prevails today
  - active switch in center
  - each "spoke" runs a (separate) Ethernet protocol (frames do not collide with each other)

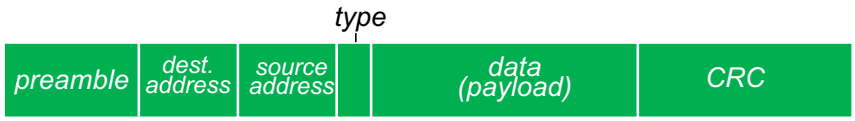


bus: coaxial cable



Ethernet Frame Structure

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



Preamble (8 bytes):

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

Destination address and source address (6 bytes, 6 bytes):

- If adapter receives frame with matching destination MAC address, or with broadcast address (e.g., ARP packet), it passes payload in frame to network layer.
- Otherwise, adapter discards frame

Ethernet Frame Structure (cont.)

Type (2 bytes):

- indicates higher layer protocol (e.g., IP, ARP, Novell IPX, AppleTalk)
- IP: 0800 (hex); ARP: 0806 (hex)

Data field (46 to 1500 bytes):

- Maximum transmission unit (MTU) = 1500 bytes
- If IP datagram < 46 bytes, stuffing is used.

Cyclic Redundancy Check (CRC) (4 bytes):

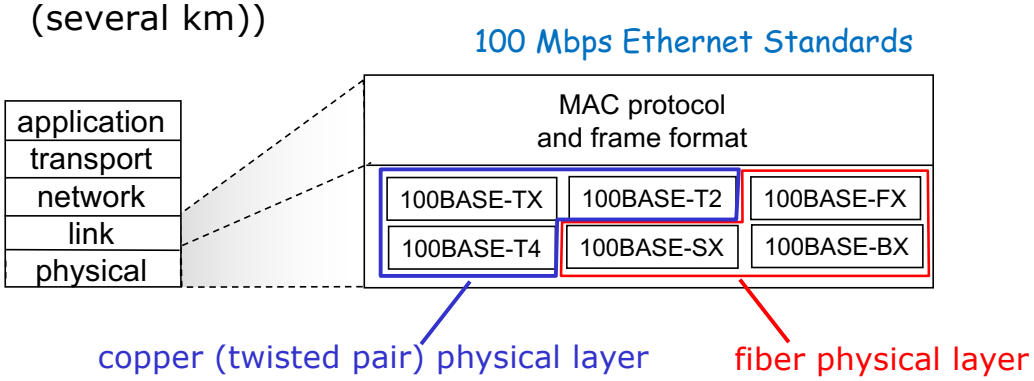
- checked at receiver, if bit errors are detected, the frame is simply dropped

Ethernet: Unreliable, Connectionless

- Connectionless: no handshaking between sending and receiving adapter
- Unreliable: receiving adapter doesn't send an ACK or NACK to sending adapter
  - 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 exponential backoff

IEEE 802.3 Ethernet Standards: Link & Physical Layers

- *many* different Ethernet standards
- common MAC protocol and frame format
- different speeds: 2 Mbps, 10 Mbps, 100 Mbps, 1 Gbps, 10 Gbps, 40 Gbps
- different physical layer (fiber, cable) and maximum distance between nodes (e.g., 100 Mbps: twisted pair (< 100 m), fiber (several km))



Link Layer, LANs: Outline

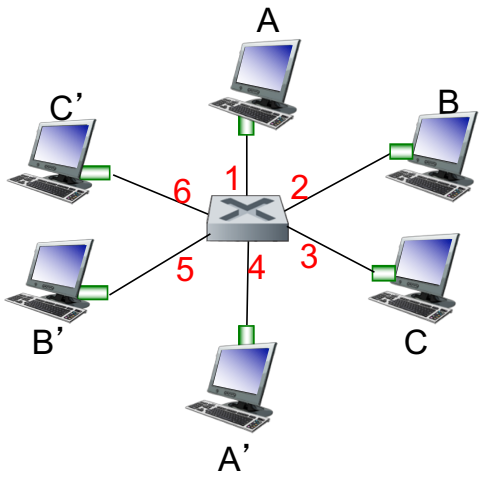
- 6.1 Introduction and services
- 6.2 Error detection and correction
- 6.3 Multiple access protocols
- 6.4 LANs
  - addressing, ARP
  - Ethernet
  - switches
  - VLANs
- 6.5 Link virtualization: MPLS
- 6.6 Data center networking
- 6.7 A day in the life of a web request

Ethernet Switch

- link-layer device: takes active role
  - store and forward Ethernet frames
  - examine incoming frame's MAC destination 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

Switch: allows multiple simultaneous transmissions

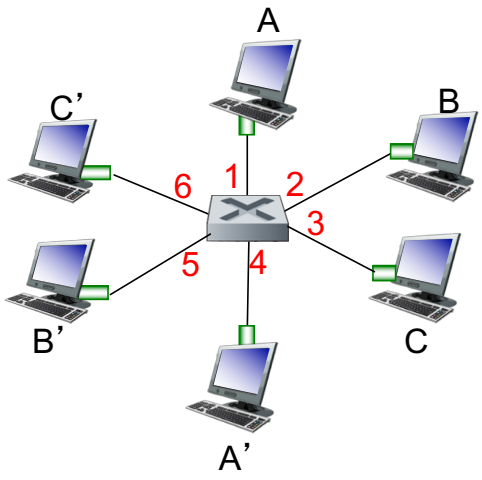
- hosts have dedicated, direct connection to switch
- switches buffer packets
- Ethernet protocol used on each incoming link, but no collisions; full duplex
  - each link is its own collision domain
- switching: A-to-A' and B-to-B' simultaneously, without collisions



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

Switch Table

- Q: how does switch know that A' reachable via interface 4, B' reachable via interface 5?
- A: each switch has a switch table, each entry:
  - (MAC address, interface leads toward that MAC address, time stamp)
  - looks like a routing table!
- Q: how are entries created, maintained in switch table?
- something like a routing protocol?

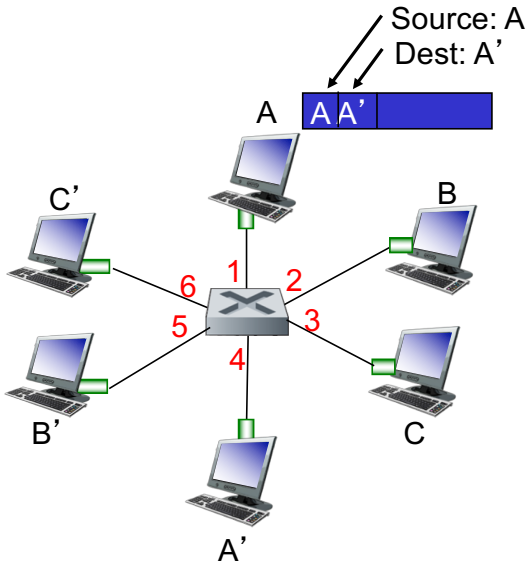


Switch Table

MAC Address	Interface	Time
62-FE-F7-11-89-B3	1	5:12
7D-34-15-C8-3F-12	4	5:45
...	...	...

Switch: 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	Time
A	*1*	*9:39*

Switch table (initially empty)

Switch: Frame Filtering/Forwarding

When a frame is received at switch:

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 the frame

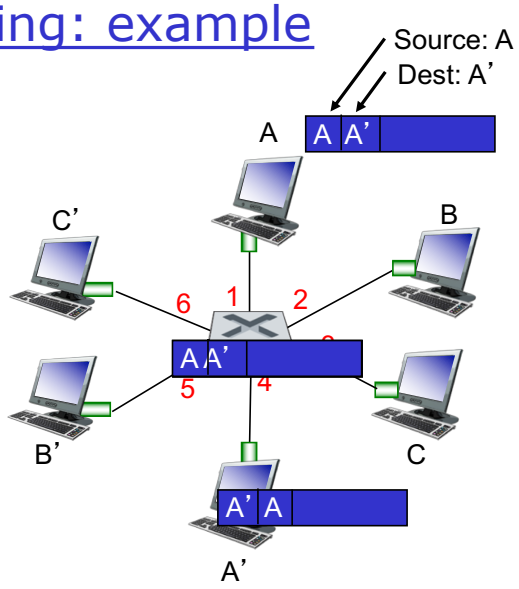
else forward the frame on interface indicated

}

else flood ← *forward to all interfaces except arriving interface*

Self-learning, forwarding: example

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

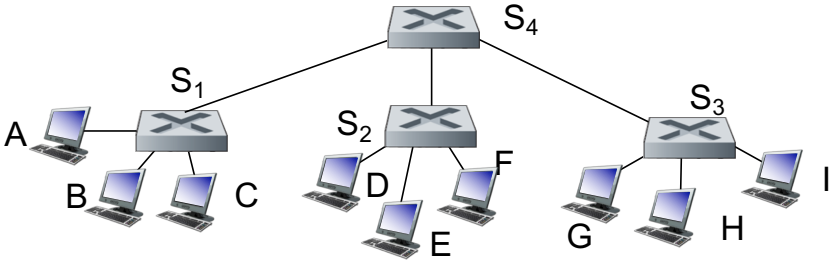


MAC addr	Interface	Time
A	*1*	*9:39*
A'	*4*	*9:41*

switch table (initially empty)

Interconnecting switches

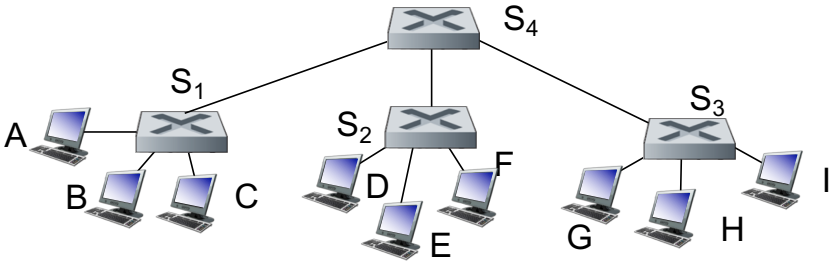
- switches can be connected together



- Q: sending from A to G - how does S_1 know how to forward frame destined to G via S_4 and S_3?
- A: self learning! (works exactly the same as in single-switch case!)

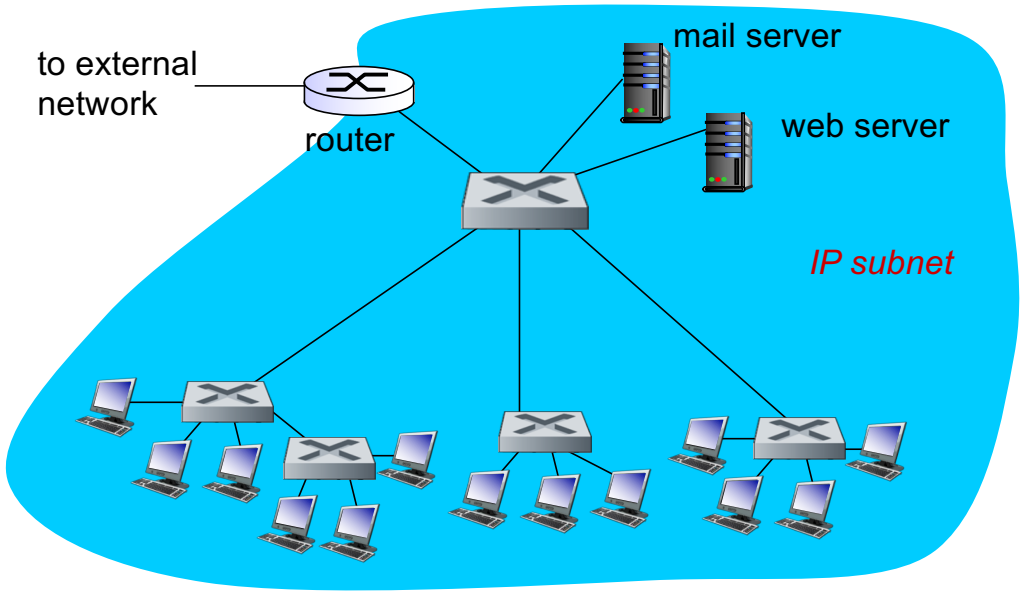
Self-learning multi-switch example

- Suppose C sends frame to I, I responds to C



- Q: show switch tables and packet forwarding in S_1, S_2, S_3, S_4

Institutional network



- Elimination of collisions
- Heterogeneous links
- Management

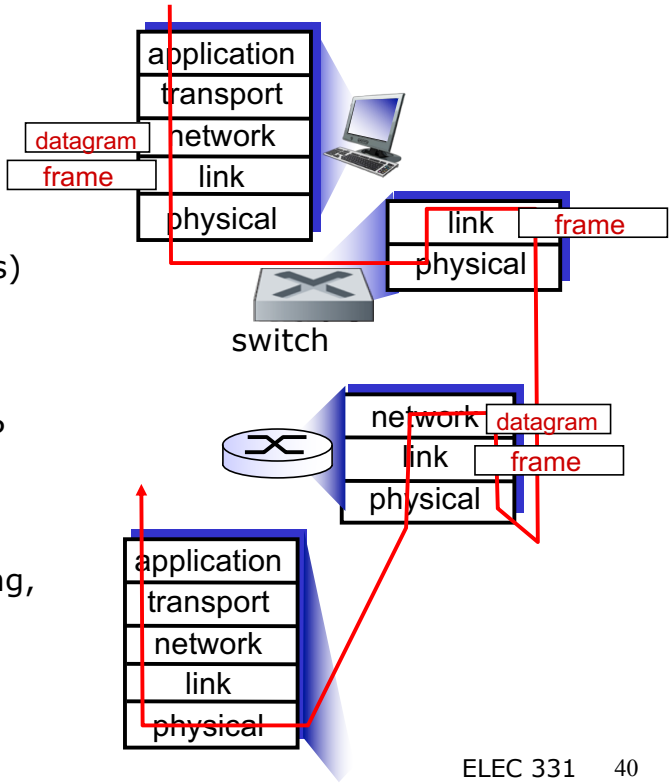
Switches vs. Routers

both store-and-forward

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

Both have forwarding tables

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



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
- synthesis: a day in the life of a web request

Chapter 6: let's take a breath

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