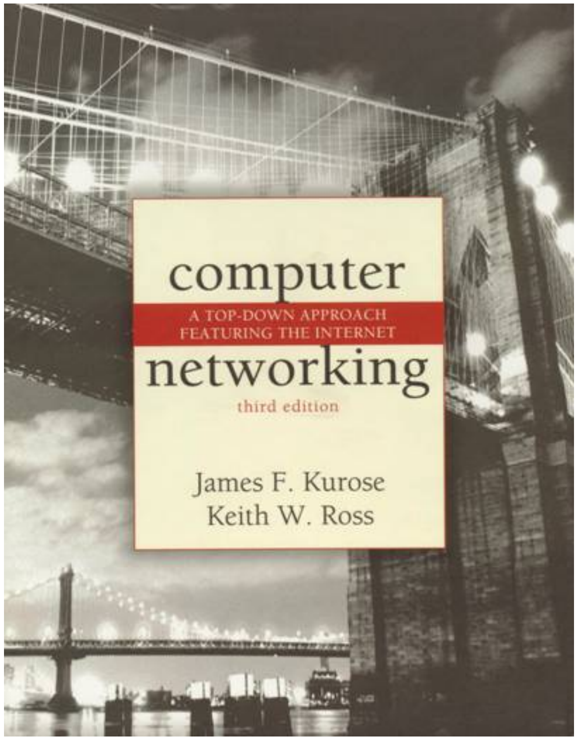
Switching and Routing

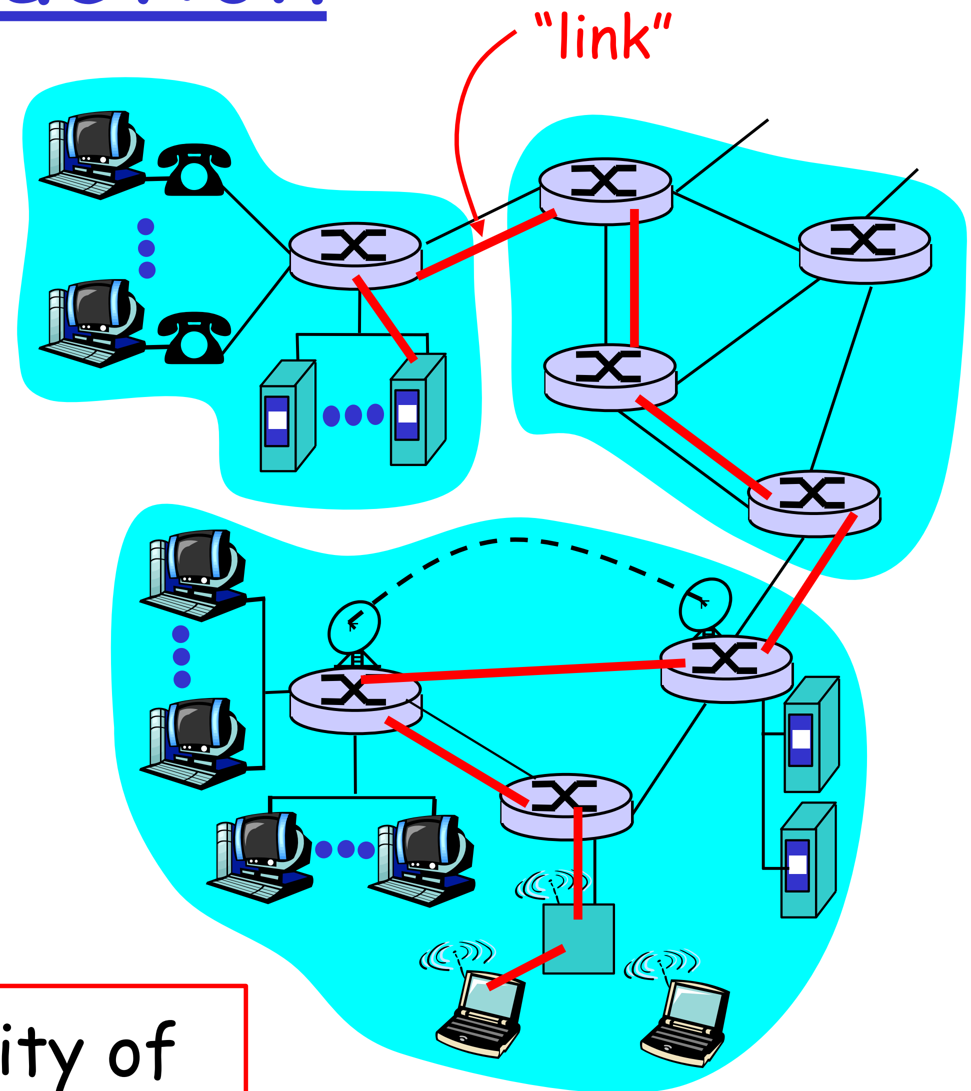


Computer Networking:
A Top Down Approach
Featuring the Internet,
Jim Kurose, Keith Ross
Addison-Wesley

Link Layer: Introduction

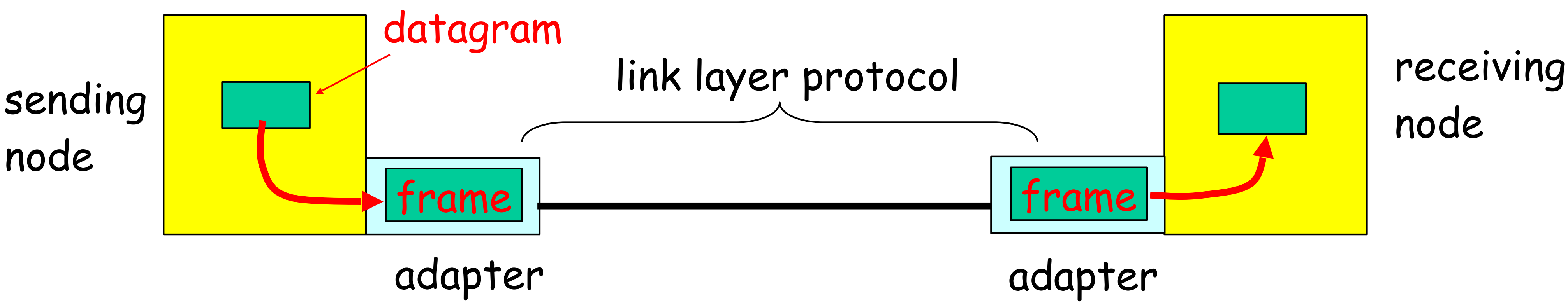
Some terminology:

- 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 adjacent node over a link

Adaptors Communicating



- link layer implemented in "adaptor" (aka NIC)
  - Ethernet card, PCMCI card, 802.11 card
- sending side:
  - encapsulates datagram in a frame
  - adds error checking bits, rdt, flow control, etc.
- receiving side
  - looks for errors, rdt, flow control, etc
  - extracts datagram, passes to rcving node
- adapter is semi-autonomous
- link & physical layers

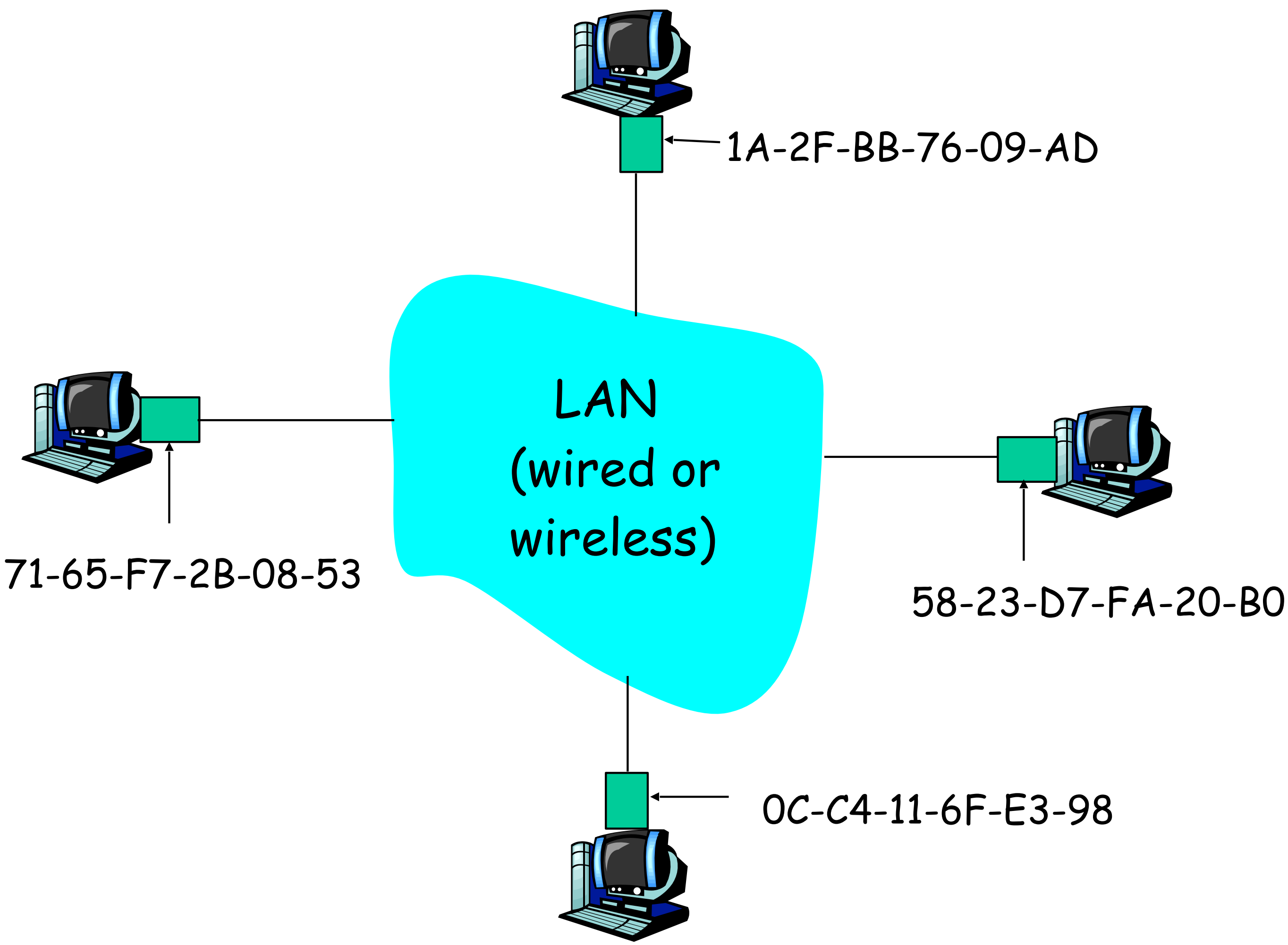
MAC Addresses

- IP address:
  - network-layer address
  - used to get datagram to destination IP subnet

- MAC (or LAN or physical or Ethernet) address:
  - MAC = Media Access Control
  - used to get datagram from one interface to another physically-connected interface (same network)
  - 48 bit MAC address (for most LANs) burned in the adapter ROM

MAC Addresses(more)

Each adapter on LAN has unique 6-byte MAC address



Broadcast address =
FF-FF-FF-FF-FF-FF

MAC Address (more)

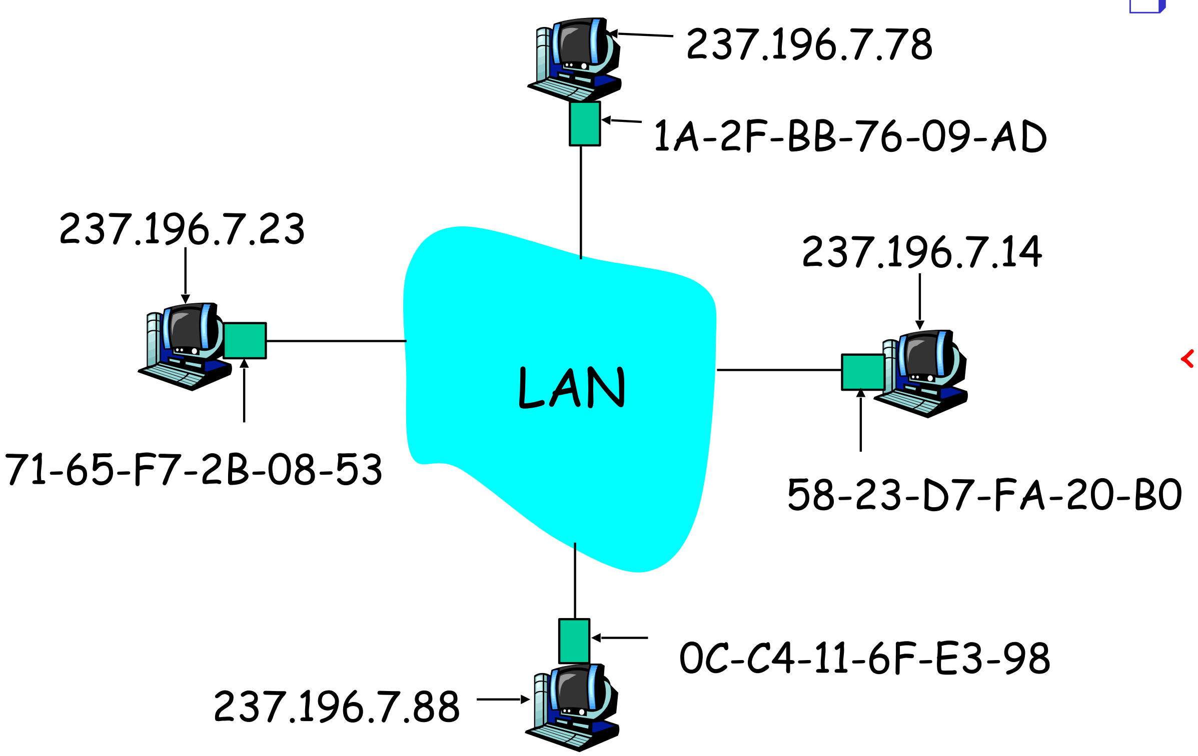
- MAC address allocation administered by IEEE
- manufacturer buys portion of MAC address space (to assure uniqueness)
- Analogy:

  (a) MAC address: like Social Security Number

  (b) IP address: like postal address
- 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 MAC address of a node knowing its IP address?



- Each IP node (Host, Router) on LAN has ARP table
- 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)

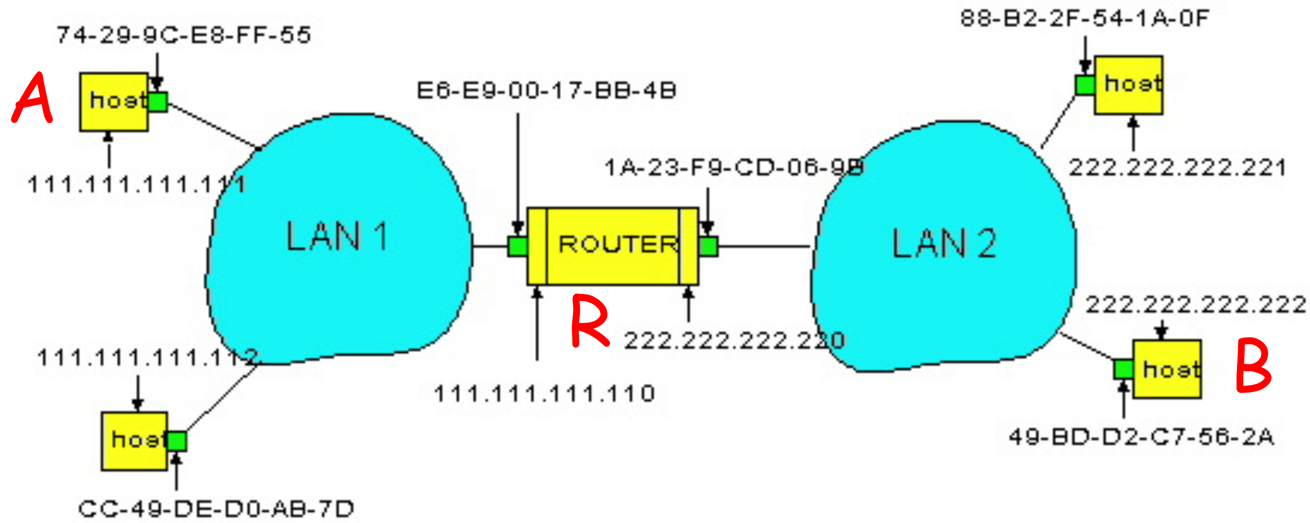
ARP protocol in the Same LAN

- A wants to send datagram to B, and B's MAC address not in A's ARP table.
- A broadcasts ARP query packet, containing B's IP address
  - "Who has B's MAC address?"
  - Dest MAC address = FF-FF-FF-FF-FF-FF
  - all machines 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 (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
- ARP Spoofing: Malicious nodes can lie about another node's MAC address to receive their traffic.

Routing to another LAN

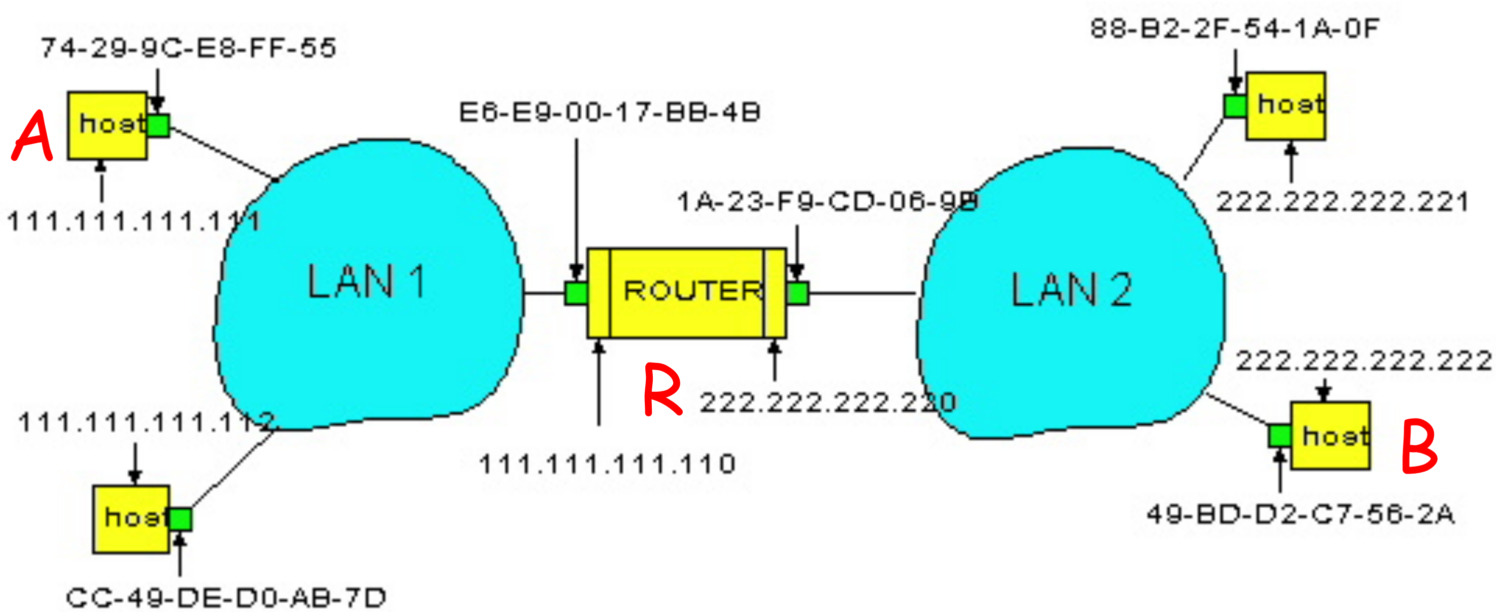
walkthrough: send datagram from A to B via R

assume A know's B IP address



- Two ARP tables in router R, one for each IP network (LAN)
- In routing table at source Host, find router 111.111.111.110
- In ARP table at source, find MAC address E6-E9-00-17-BB-4B, etc

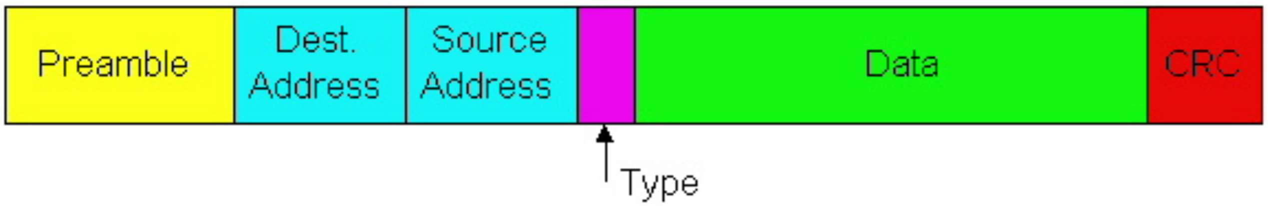
- A creates datagram with source A, destination B
- A uses ARP to get R's MAC address for 111.111.111.110
- A creates link-layer frame with R's MAC address as dest, frame contains A-to-B IP datagram
- A's adapter sends frame
- R's adapter receives frame
- R removes IP datagram from Ethernet frame, sees its destined to B
- R uses ARP to get B's MAC address
- R creates frame containing A-to-B IP datagram sends to B



Ethernet

Ethernet Frame Structure

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



- Preamble: 8 bytes used for synchronizing sender and receiver.
- Addresses: 6 bytes
  - if adapter receives frame with matching destination address, or with broadcast address (eg ARP packet), it passes data in frame to net-layer protocol
  - otherwise, adapter discards frame
- Type: indicates the higher layer protocol (mostly IP but others may be supported such as Novell IPX and AppleTalk)
- CRC: (cyclic redundancy checksum)
  - checked at receiver, if error is detected, the frame is simply dropped

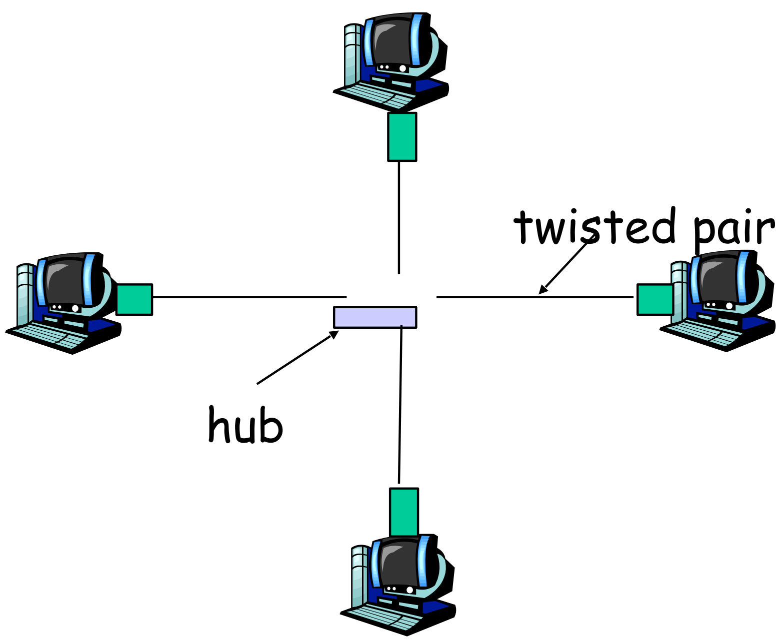
Ethernet uses CSMA/CD

- Adapter transmits whenever data is ready, but listens first.
- Adapter doesn't transmit if it senses that some other adapter is transmitting, that is, carrier sense
- Transmitting adapter aborts when it senses that another adapter is transmitting, that is, collision detection
- Before attempting a retransmission, adapter waits a random time, that is, random access

Hubs

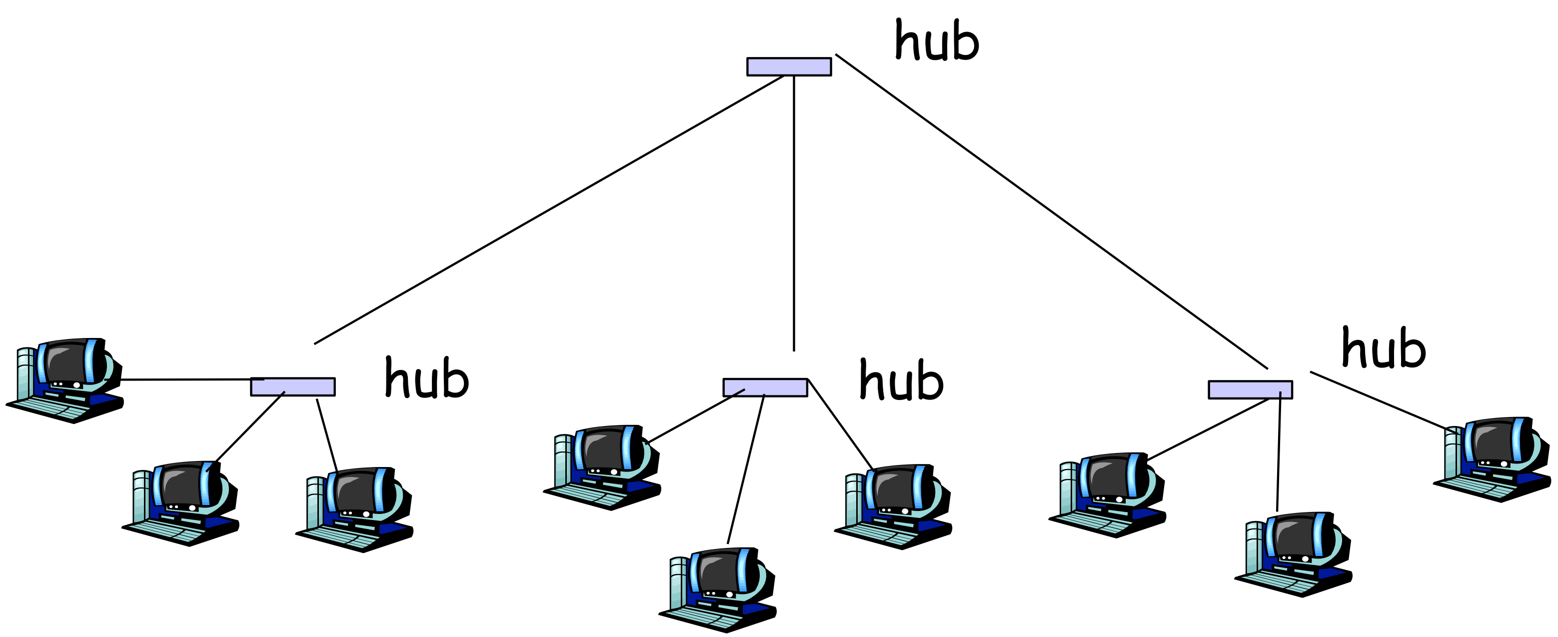
Hubs are essentially physical-layer repeaters:

- bits coming from one link go out all other links
- at the same rate
- no frame buffering
- no CSMA/CD at hub: adapters detect collisions
- provides net management functionality



Interconnecting with hubs

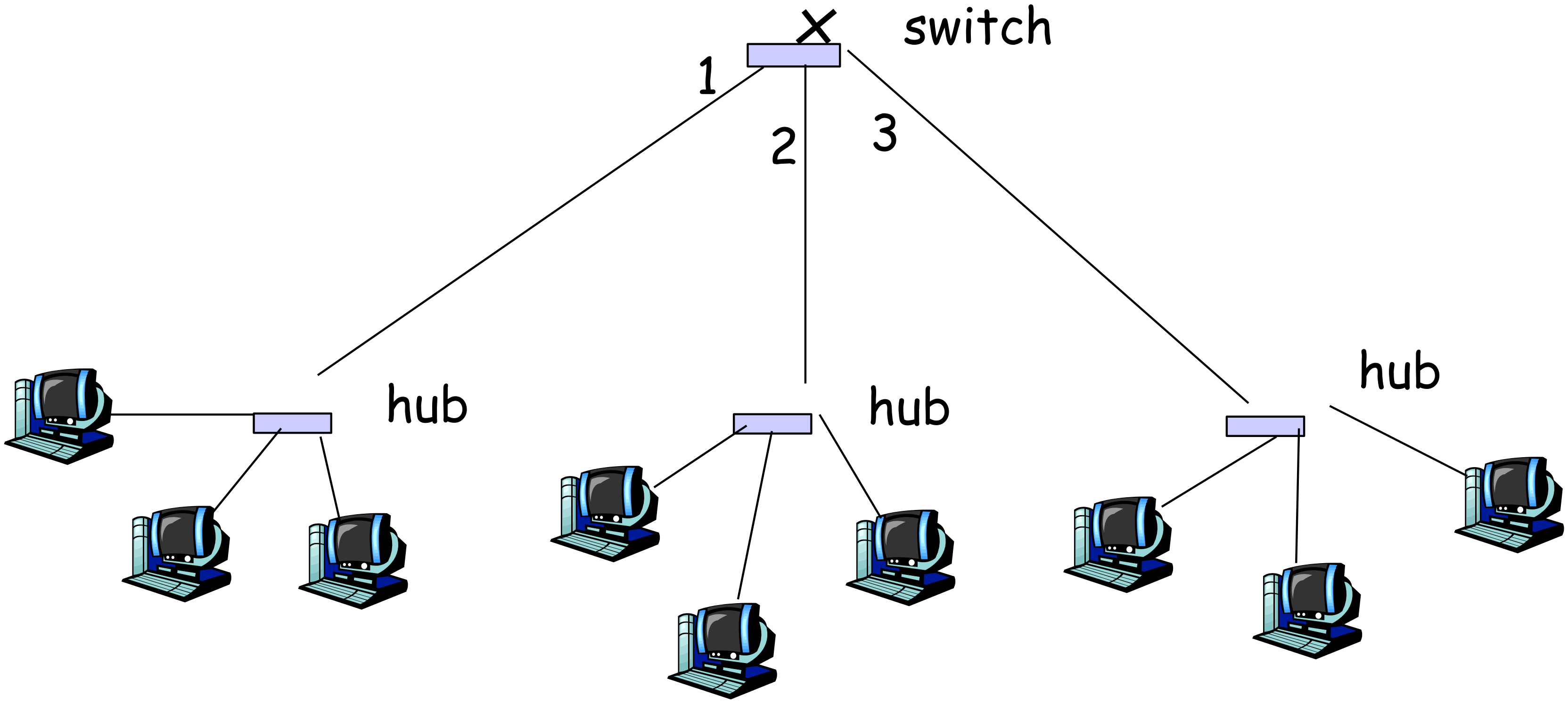
- Backbone hub interconnects LAN segments
- Extends max distance between nodes
- But individual segment collision domains become one large collision domain



Switch

- Link layer device
  - stores and forwards Ethernet frames
  - examines frame header and selectively forwards frame based on MAC dest address
  - 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

Forwarding



- How do determine onto which LAN segment to forward frame?

Self learning

- Also called “backward learning” or “transparent bridging”
- A switch has a switch table
- entry in switch table:
  - [MAC Address, Interface, Time Stamp]
  - stale entries in table dropped (TTL can be 60 min)
- 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

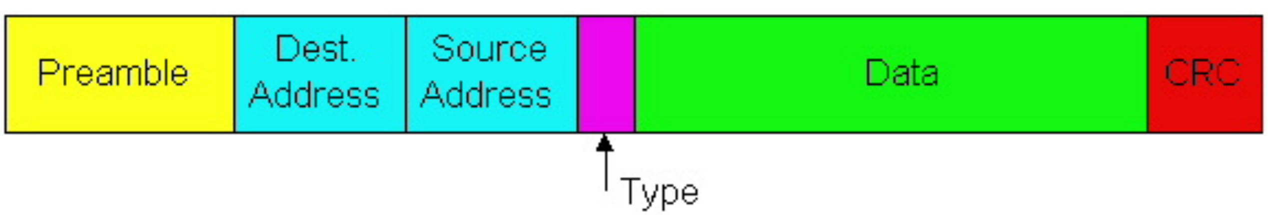
Filtering/Forwarding

When switch receives a frame:

index switch table using MAC dest address

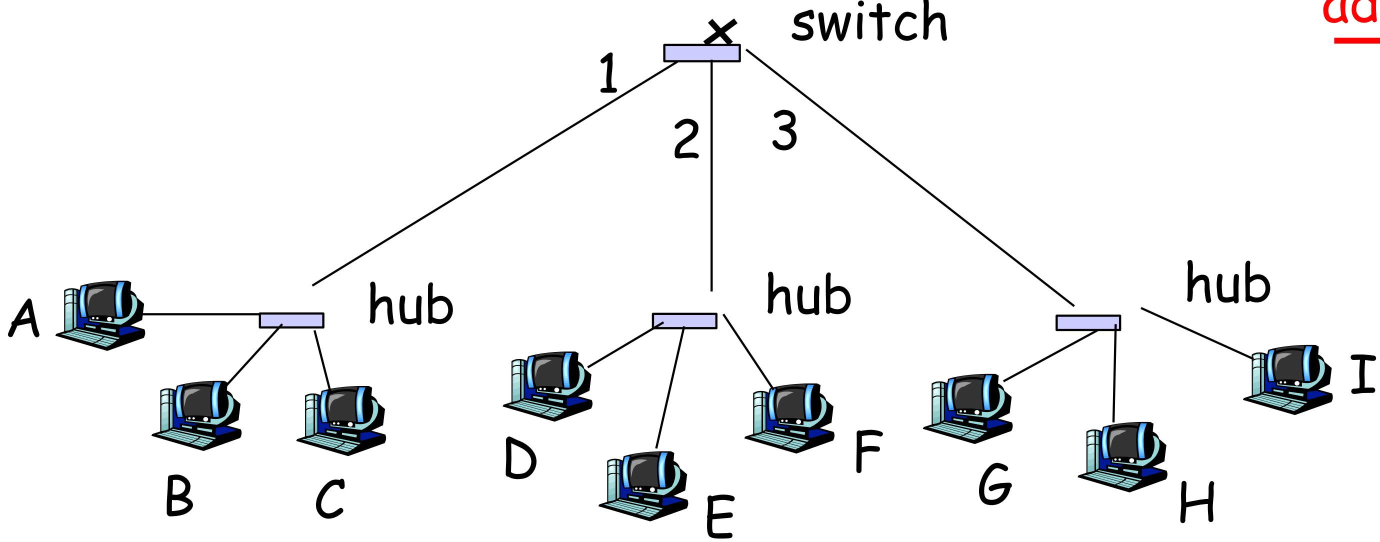
if entry found for destination
    then{
      if dest on segment from which frame arrived
        then drop the frame
        else forward the frame on interface indicated
    }
else flood

forward on all but the interface on which the frame arrived



Switch example

Suppose C sends frame to D

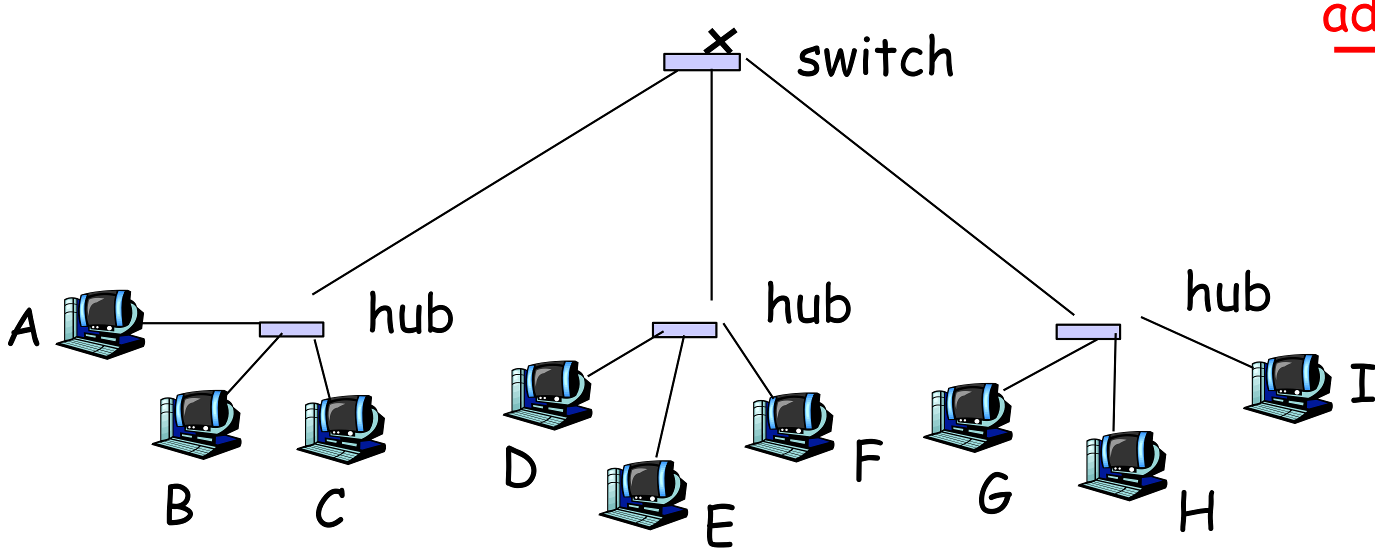


address	interface
A	1
B	1
E	2
G	3

- Switch receives frame from from C
  - notes in bridge table that C is on interface 1
  - because D is not in table, switch forwards frame into interfaces 2 and 3
- frame received by D

Switch example

Suppose D replies back with frame to C.

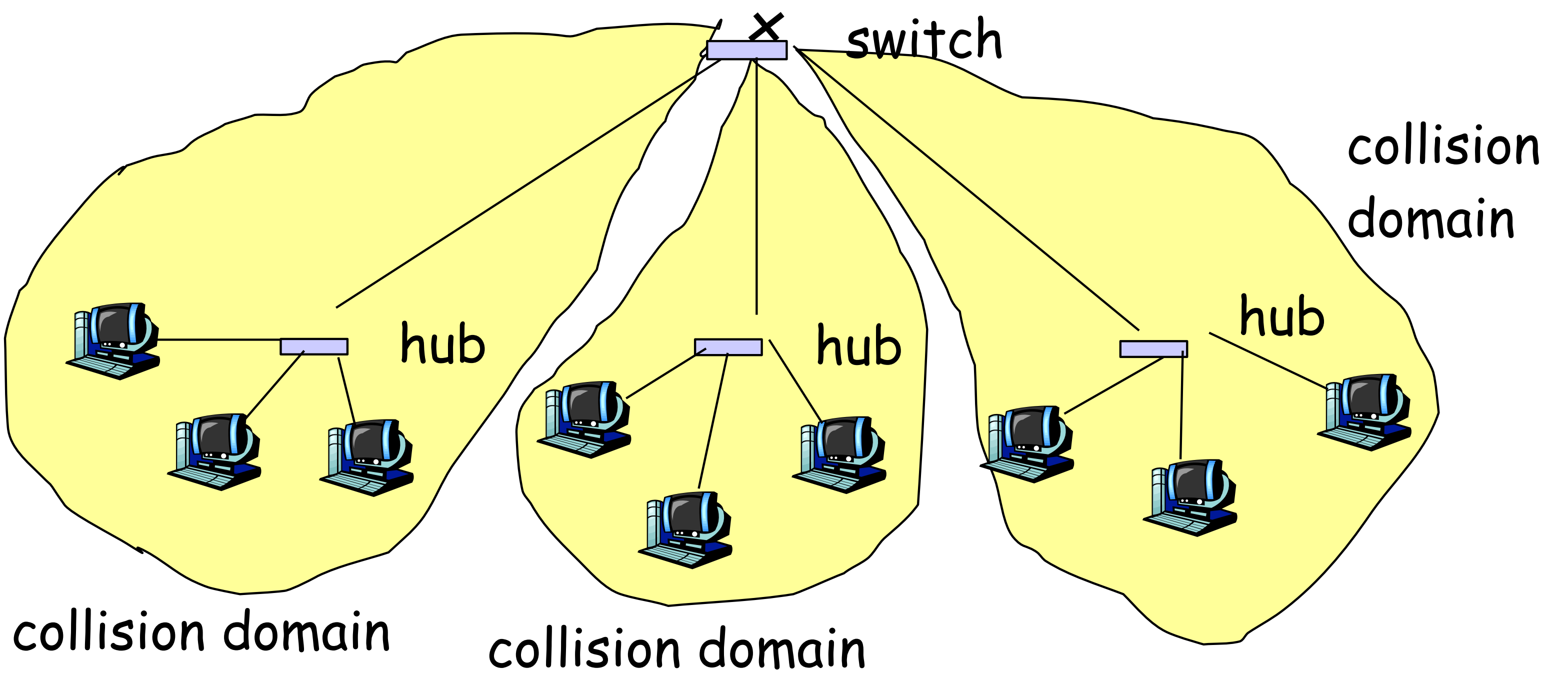


address	interface
A	1
B	1
E	2
G	3
C	1

- Switch receives frame from from D
  - notes in bridge table that D is on interface 2
  - because C is in table, switch forwards frame only to interface 1
- frame received by C

Switch: traffic isolation

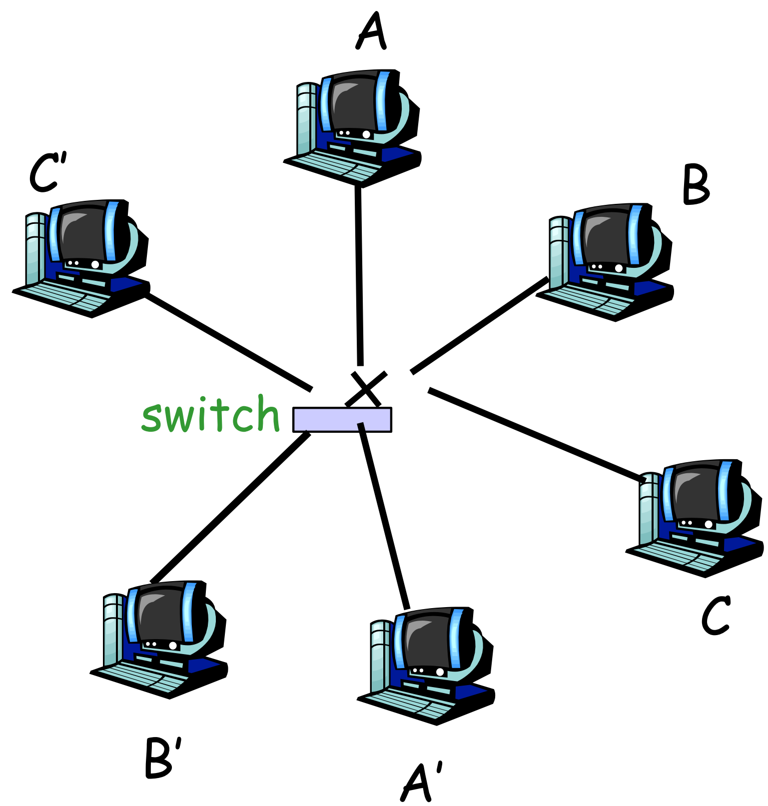
- switch installation breaks subnet into LAN segments
- switch filters packets:
  - same-LAN-segment frames not usually forwarded onto other LAN segments
  - segments become separate collision domains



Switches: dedicated access

- Switch with many interfaces
- Hosts have direct connection to switch
- No collisions; full duplex

Switching: A-to-A' and B-to-B' simultaneously, no collisions

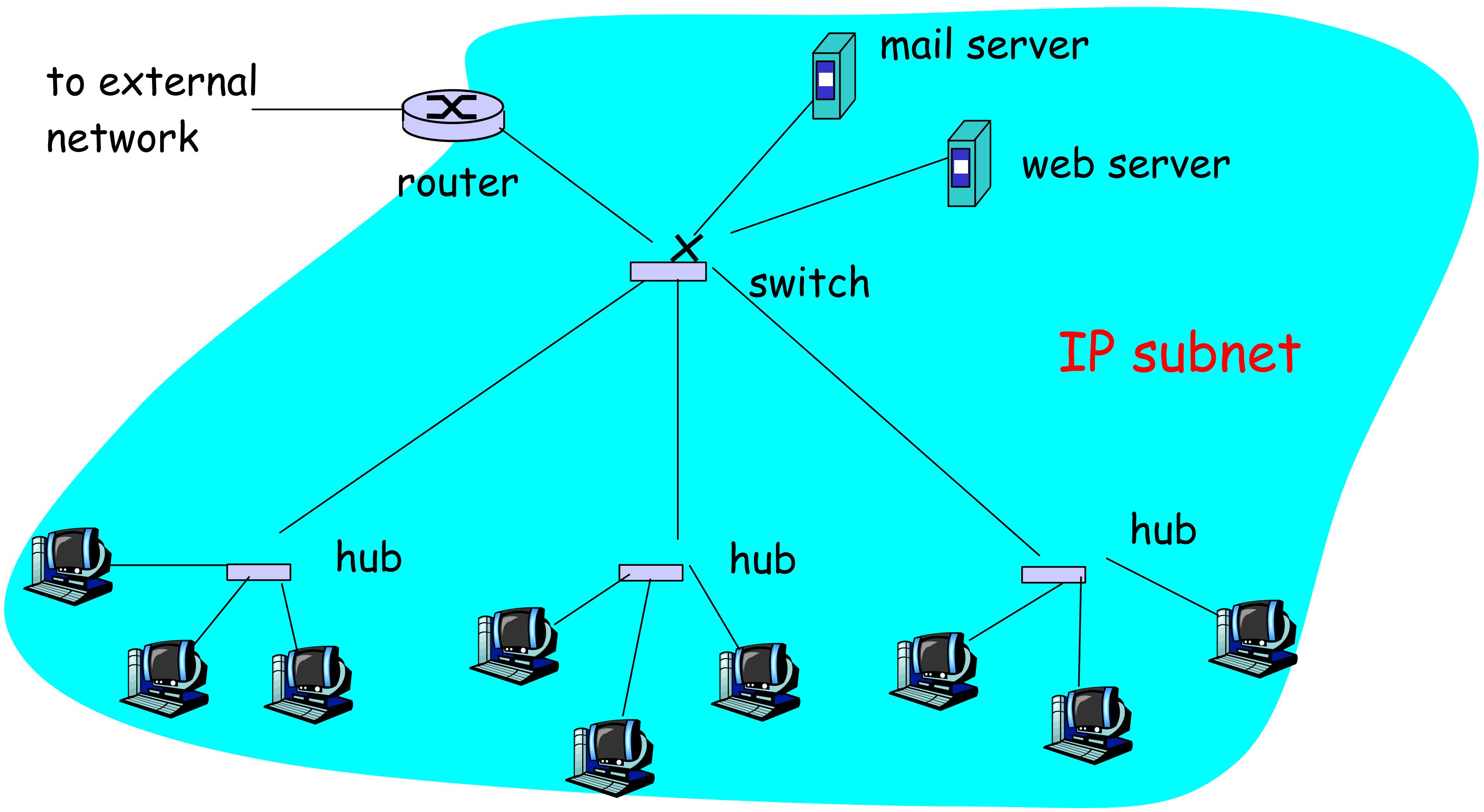


More on Switches

- cut-through switching: frame forwarded from input to output port without first collecting entire frame
  - slight reduction in latency

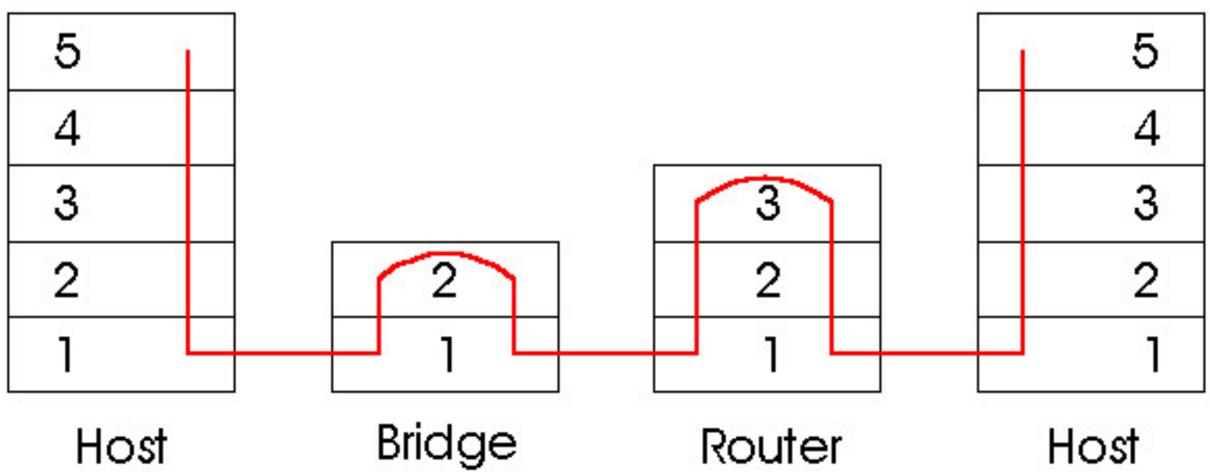
- combinations of shared/dedicated, 10/100/1000 Mbps interfaces

Institutional network



Switches vs. Routers

- both store-and-forward devices
  - routers: network layer devices (examine network layer headers)
  - switches are link layer devices
- routers maintain routing tables, implement routing algorithms
- switches maintain switch tables, implement filtering, learning algorithms



Summary comparison

	hubs	routers	switches
traffic isolation	no	yes	yes
plug & play	yes	no	yes
optimal routing	no	yes	no
cut through	yes	no	yes