COMP 3234B
Computer and Communication Networks

2nd semester 2023-2024

Link Layer (III)

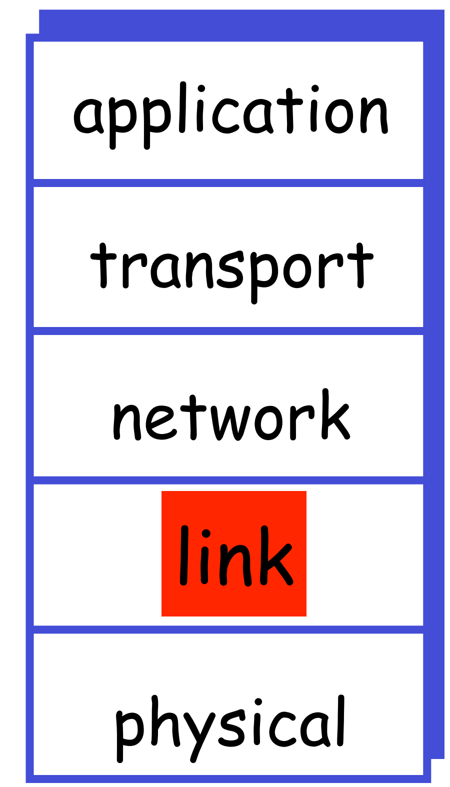
Prof. C Wu

Department of Computer Science
The University of Hong Kong

Roadmap

Link layer

- Principles behind link-layer services (ILO1,2,3)
  - error detection and correction
  - multiple access protocols
  - link-layer addressing
- Implementation of various link-layer technologies (ILO2,3)
  - Ethernet
  - switches
  - 802.11 wireless LAN (WiFi)
  - etc.



Link-layer addressing

- Network-layer addressing
  - IP addresses
  - function: to get datagram to destination IP subnet

- Link-layer addressing
  - MAC (or LAN or physical) address
  - function: get frame from one interface to another physically-connected interface (in the same subnet)

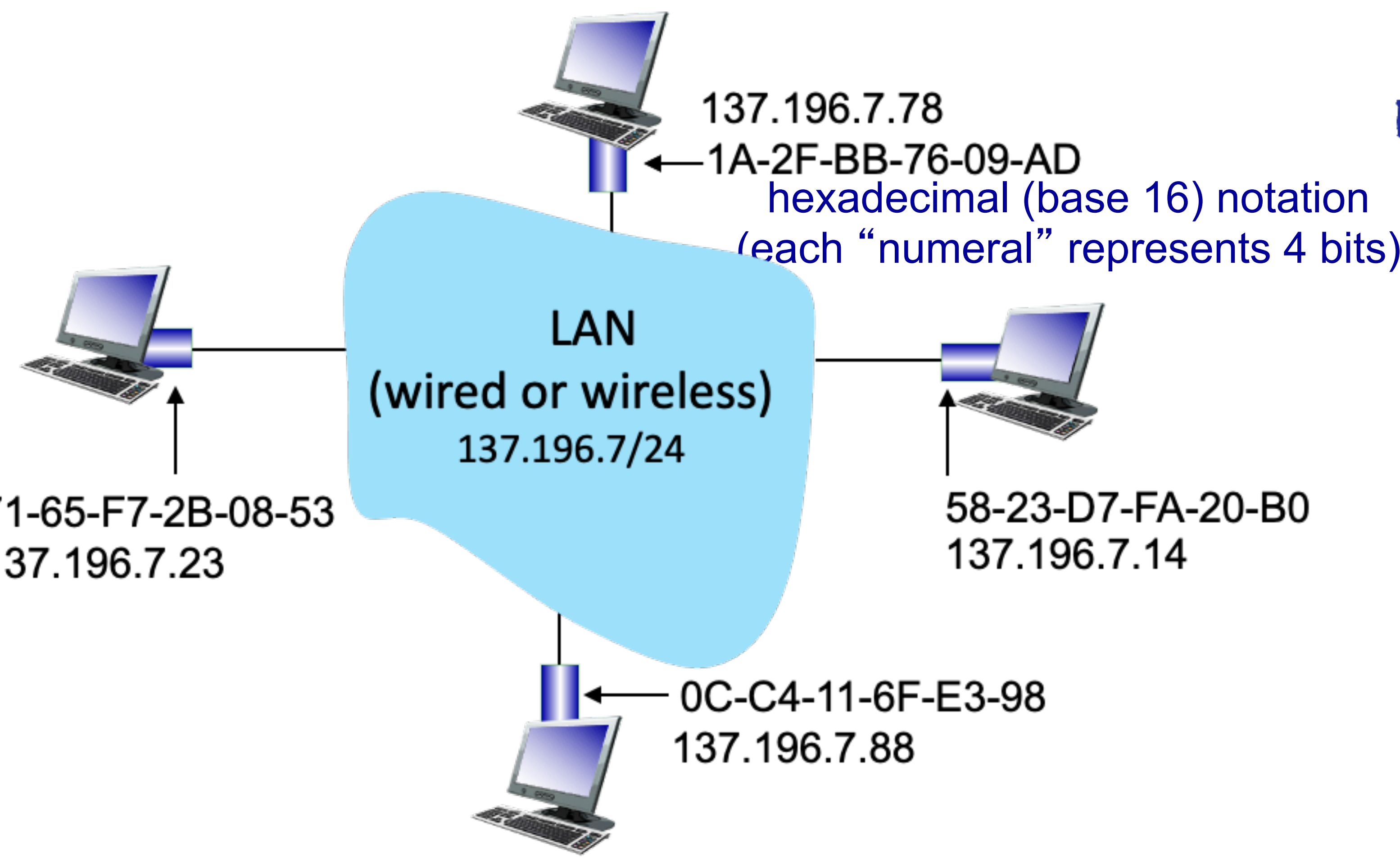
- LAN (Local Area Network)
  - a computer network concentrated in a geographical area

    e.g., in a building, on a university campus
  - route: host → LAN → routers to Internet
  - different LAN technologies (link-layer technologies)

    Ethernet (dominant)

MAC address

- MAC address
  - 48 bit (6 byte) long, for most LANs
    usually expressed in hexadecimal notation
  - each network adapter has a MAC address
    burned in NIC ROM (now possible to change via software)



- How MAC address is used
  - Sender of a frame inserts dest. MAC address into frame
  - frame is received and processed by all adapters on the LAN (a broadcast LAN)
  - Adapter checks whether its own MAC address matches the one in the frame

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

MAC address

- MAC address is unique for every adapter, allocated/administrated by IEEE
  - manufacturer buys portion of MAC address space
- Analogy:
  - MAC address: HKID
  - IP address: postal address
- MAC address is flat (no hierarchy) ➜ portability
  - can move LAN card from one LAN to another

IP hierarchical address not portable:

IP address depends on subnet to which node is attached

ARP: basics

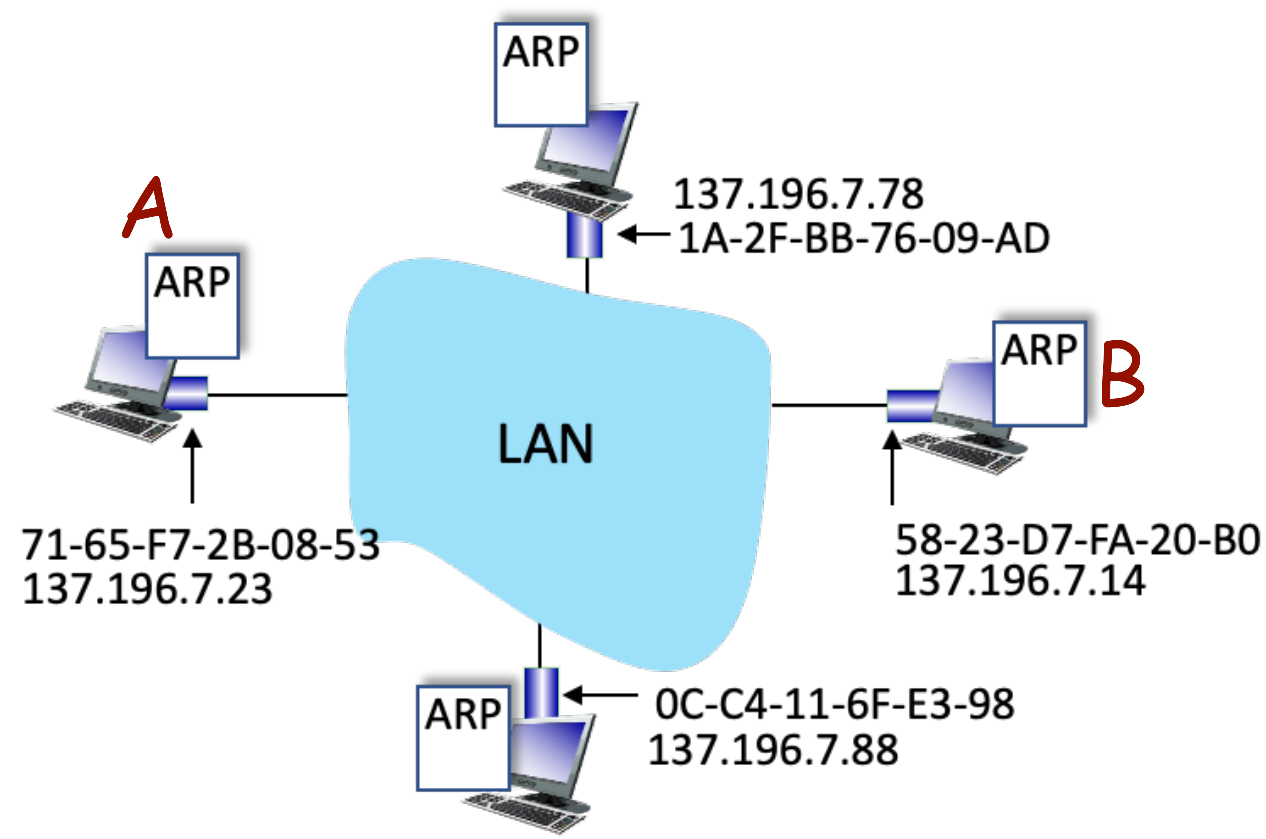
- Address Resolution Protocol
  - resolves IP address to MAC address for nodes in the same LAN

    in order for a node to insert correct MAC address into frame destined to other nodes

- Each node (host, router) on LAN has an ARP table
  - ARP table: IP/MAC address mappings for nodes in the LAN

    < IP address; MAC address; TTL>

    TTL (Time To Live): time after which address mapping will expire (typically minutes to hours)



ARP: protocol

- *A* sends datagram to *B* in the same LAN (*B*'s MAC address not in *A*'s ARP table)
  - *A* broadcasts ARP query packet (a broadcast frame)

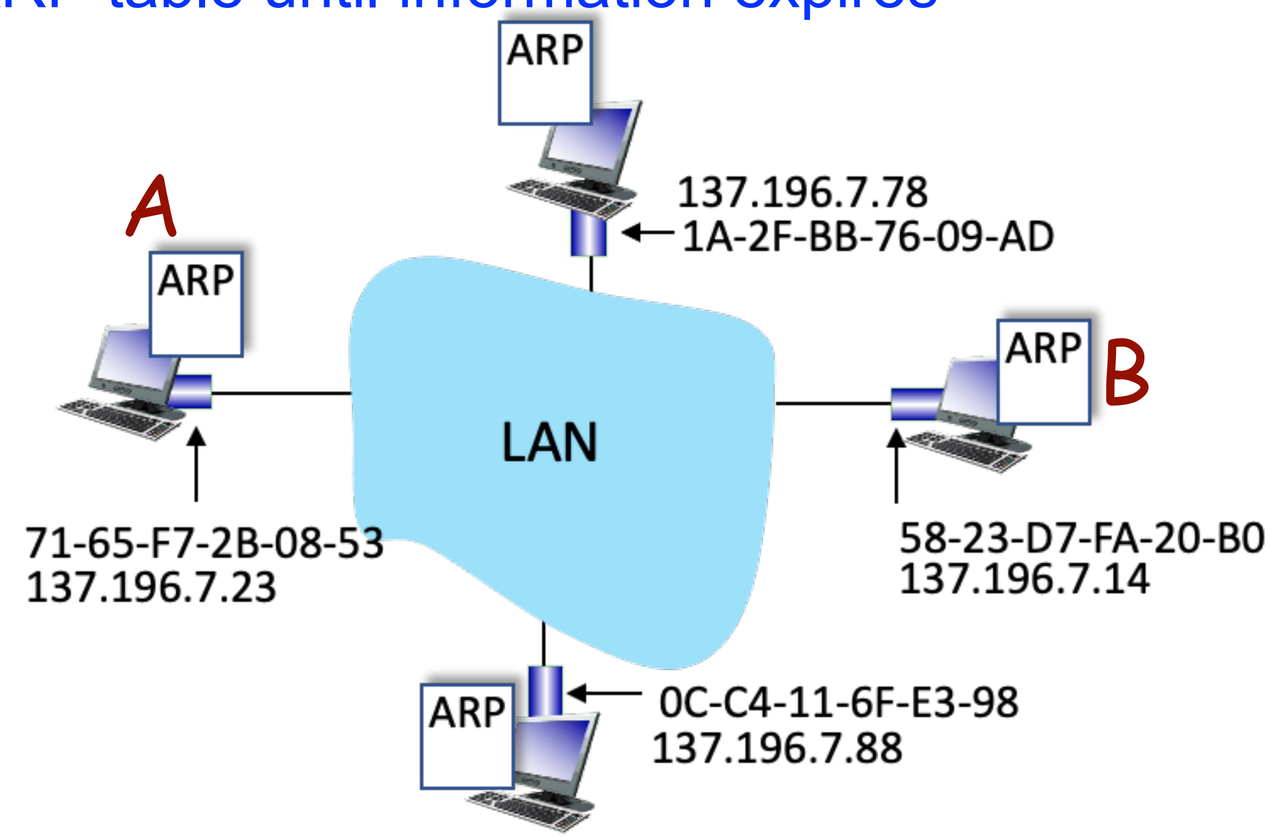
    containing B's IP address

    dest. MAC address = FF-FF-FF-FF-FF-FF
  - all machines on LAN receive the ARP query packet
  - *B* replies an ARP response packet to *A* (a unicast frame)

    containing B's MAC address
  - *A* caches IP-to-MAC address pair in its ARP table until information expires

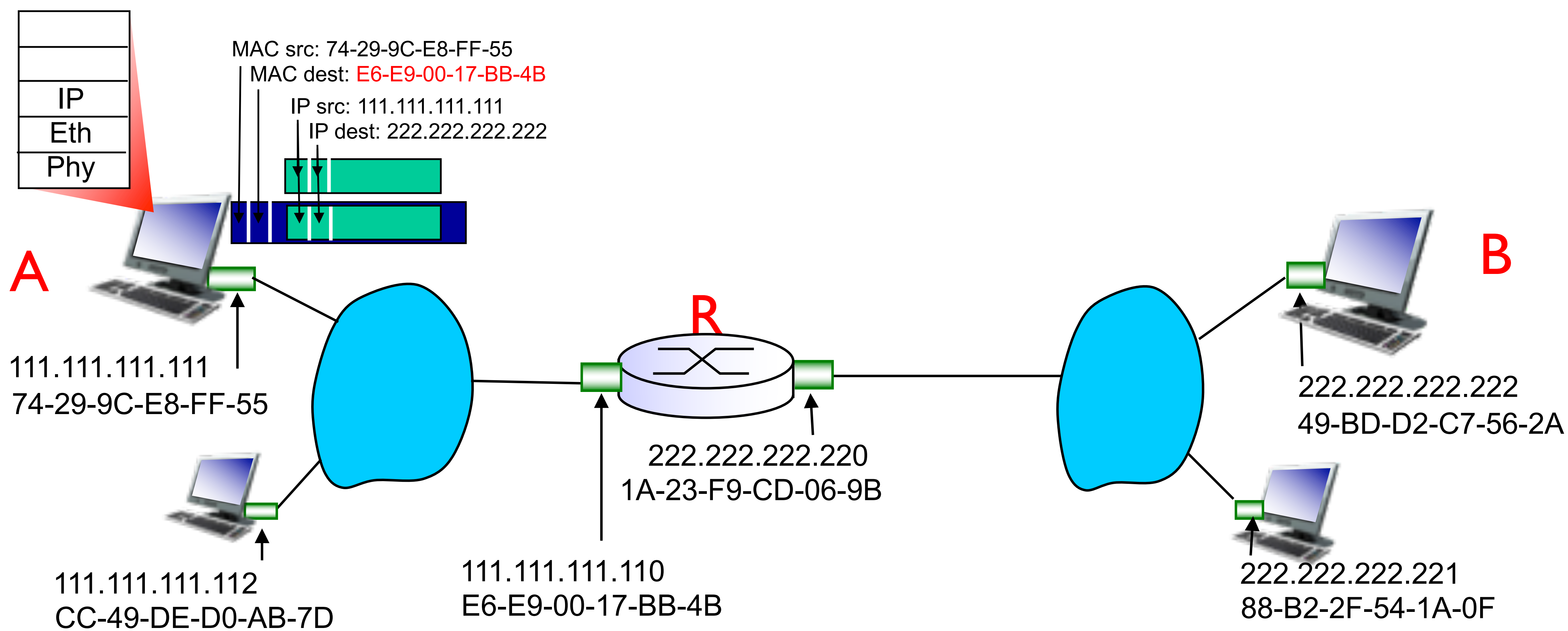
ARP is "plug-and-play":

nodes create ARP tables themselves without intervention from network administrator



Routing to another LAN

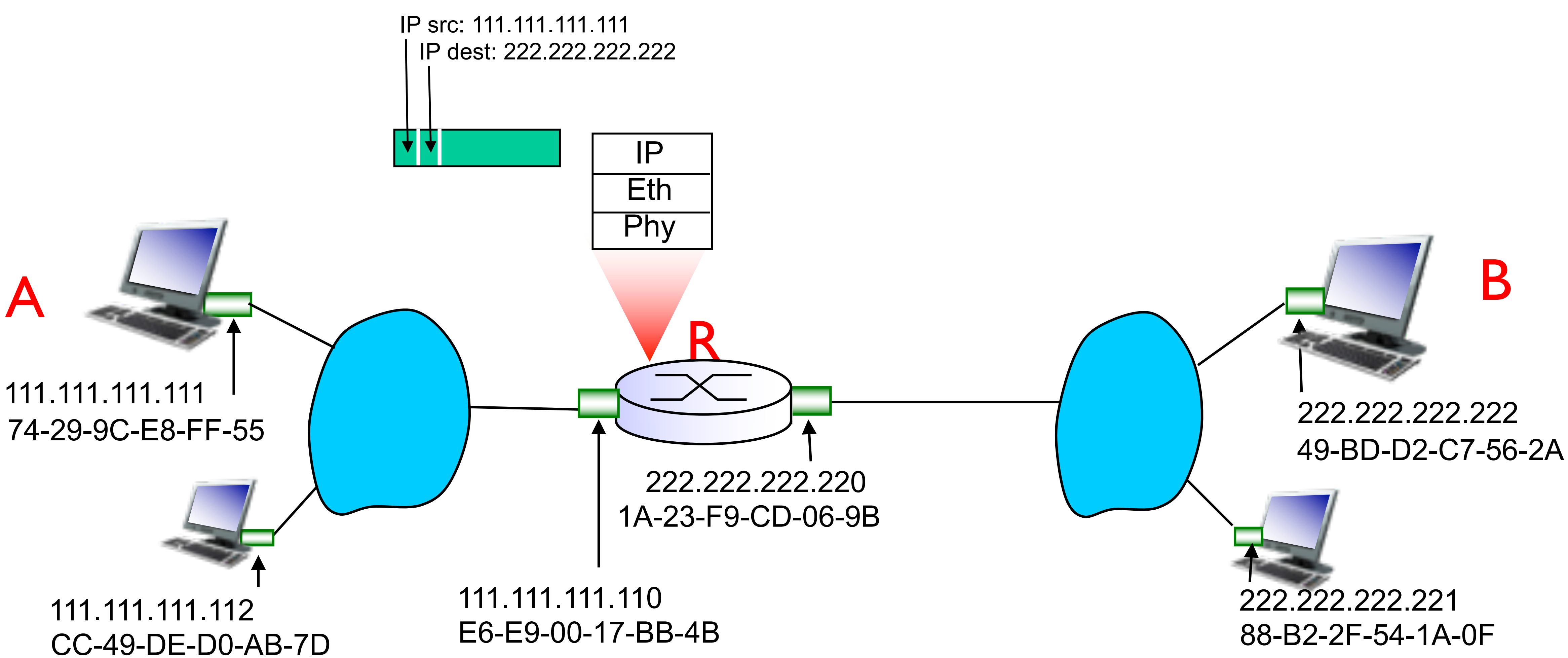
- *A* sends datagram to *B* in a different LAN via router R
  - *A* creates IP datagram with source *A*, destination *B*
  - *A* uses ARP to get R’s MAC address for interface 111.111.111.110
  - *A* creates link-layer frame with R's MAC address as dest, frame contains *A*-to-*B* IP datagram



two ARP tables in router R, one for each LAN

Routing to another LAN (cont'd)

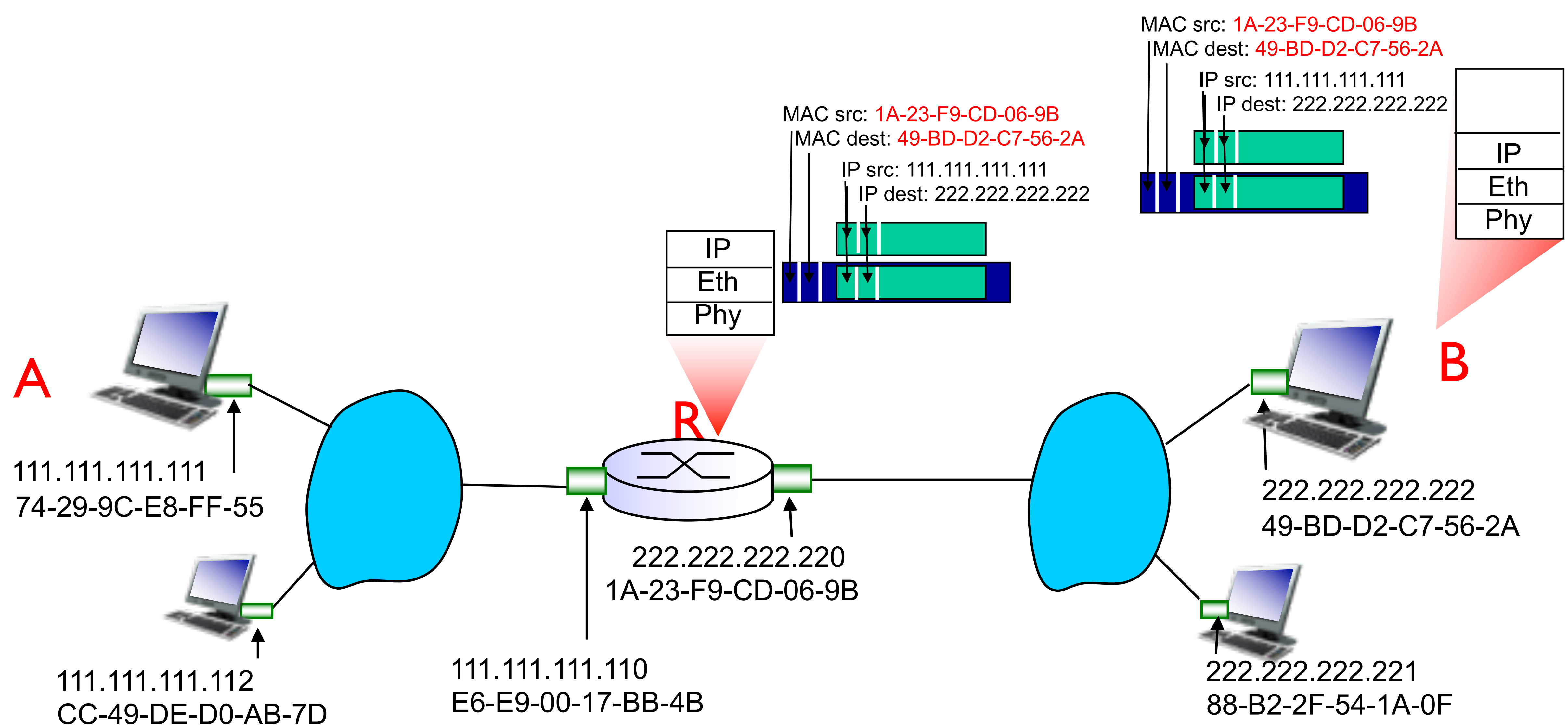
- *A*'s adapter sends frame
- R's adapter (on the left) receives frame
- R removes IP datagram from frame and sees it is destined to *B*
- R uses ARP to get *B*'s MAC address in the right-hand side LAN



two ARP tables in router R, one for each LAN

Routing to another LAN (cont'd)

- R creates frame containing *A*-to-*B* IP datagram and sends it to *B*



two ARP tables in router R, one for each LAN

Ethernet

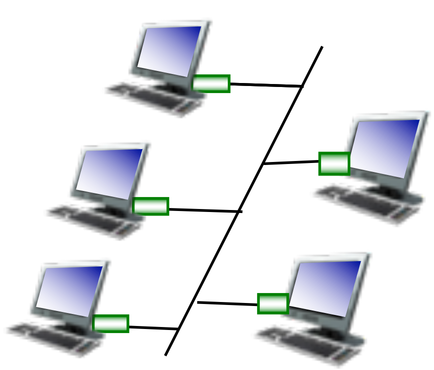
- First widely used LAN technology, invented mid-70's
- "Dominant" wired LAN technology today
- Simpler, cheaper than other LAN technologies
- Kept up with speed race:

1973	1983	1995	1998	2002	2010	2018
3Mbps	10Mbps	100Mbps	1Gbps	10Gbps	40Gbps & 100Gbps	200Gbps & 400Gbps

Ethernet topology

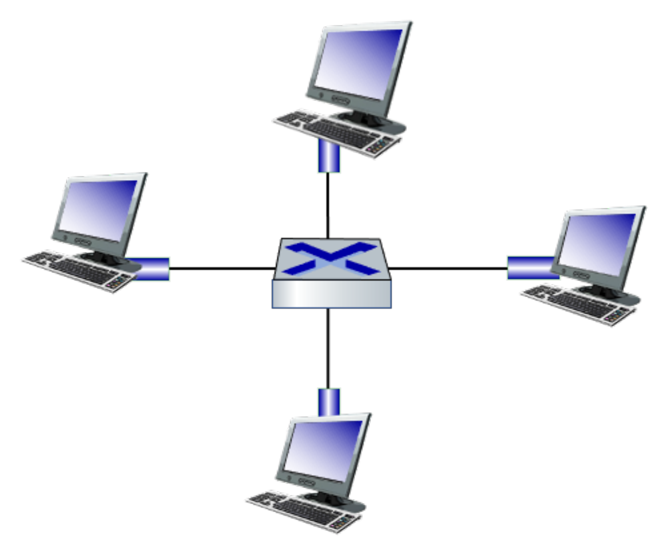
- Bus topology popular before mid 90s
  - all nodes in single collision domain (can collide with each other)
- Star topology – hub-based (late 90s)
  - all nodes are still in a single collision domain
- Star topology – switch-based (prevailing today)
  - each switch port connects to a separate Ethernet link (thus, not confined to a single collision domain)

bus topology



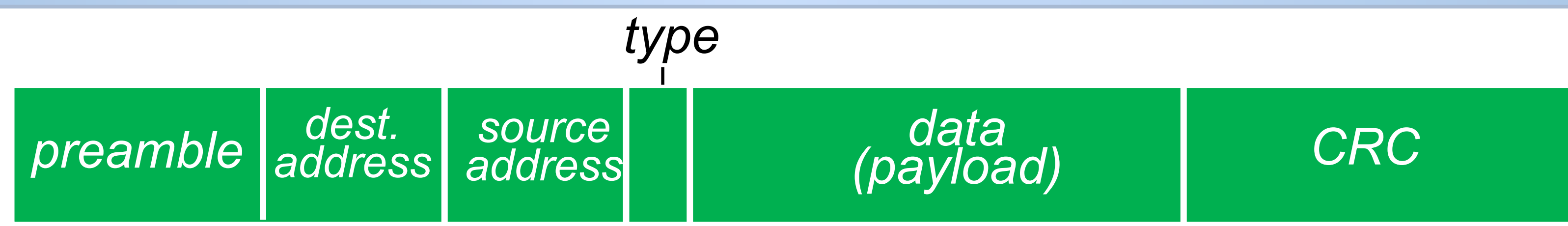
bus: coaxial cable

star topology



switched

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 clocks
  - Addresses: 6 bytes each

    if adapter receives frame with matching dest. address, or with broadcast address, it passes data in frame to network-layer protocol

    otherwise, adapter discards frame
  - Type: 2 bytes

    indicates higher layer protocol (mostly IP but others possible, e.g., ARP)
  - Data: 46 to 1500 bytes (Ethernet MTU)
  - CRC (Cyclic Redundancy Check Codes): 4 bytes

    CRC check: receiver computes CRC using all fields other than CRC and Preamble, and compares with what's carried in CRC field

    drop frame if not match

Ethernet's unreliable connectionless service

- Connectionless
  - no handshaking between sending and receiving adapters
- unreliable (not implementing rdt)
  - receiving adapter does not send ACKs or NAKs to sending adapter
  - datagrams can be dropped at link layer due to error

    depending on upper-layer protocols for rdt, e.g., TCP at transport layer

Ethernet's MAC protocol

- Bus/hub-based Ethernet: it needs a multiple access protocol
  - CSMA/CD with binary exponential backoff
- Switch-based Ethernet
  - CSMA/CD is used on half-duplex link, and is not needed on full-duplex link

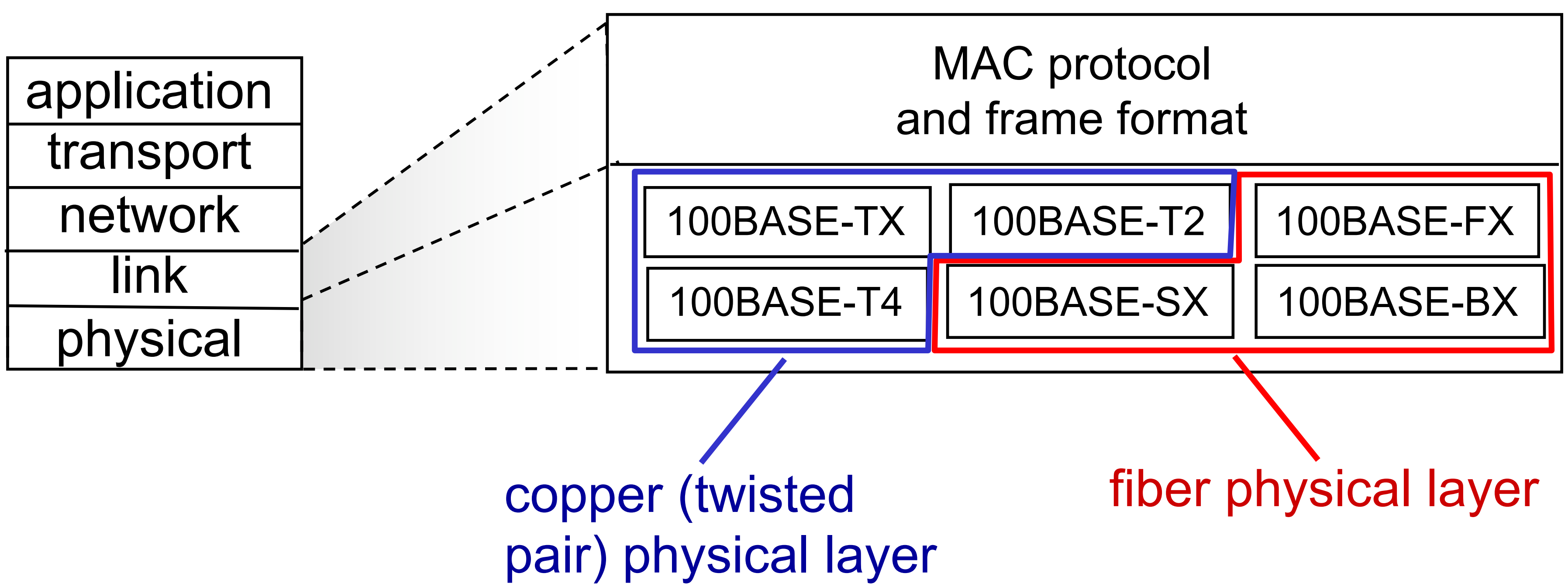
Ethernet standards

- 802.3: specifications for link and physical layers, including many different Ethernet standards
  - common MAC protocol and frame format
  - different speeds: 10Mbps/100Mbps/1Gbps/10Gbps/40Gbps/80Gbps/..
  - different physcial-layer media: 100BASE-TX, 10BASE-2, 100BASE-T, 1000BASE-LX, 10GBASE-T, etc.

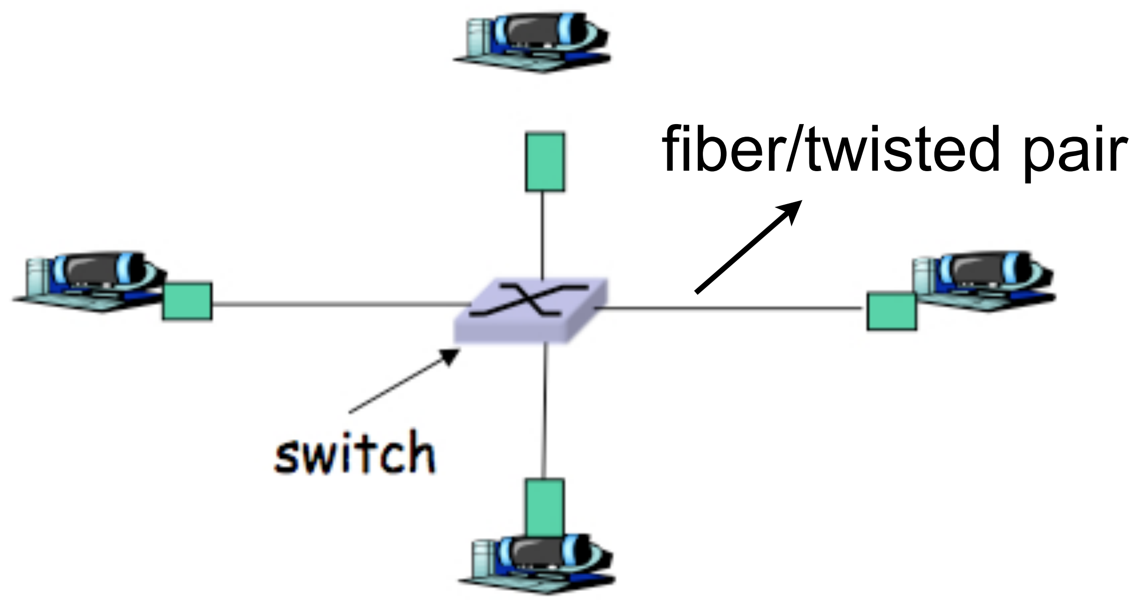
    the first number part: speed

    BASE: baseband Ethernet (physical media only carries Ethernet traffic)

    the last part: physical media, e.g., T for twisted-pair copper wire, X for fiber



Ethernet switch



- Switches
  - link-layer device
  - receive, buffer, and forward Ethernet frames

    examine incoming frame's MAC address, selectively forward frame to one or more outgoing links
  - transparent

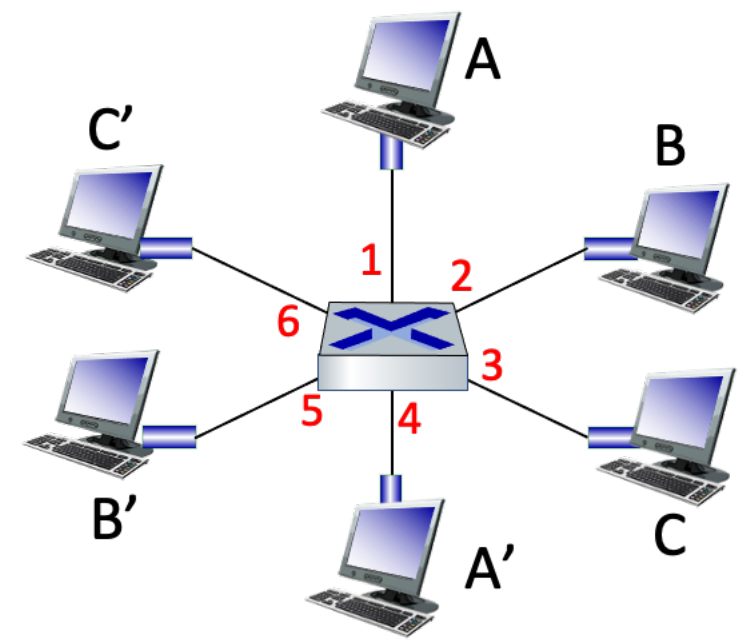
    hosts are unaware of presence of switches => host addresses a frame to another host (or router), rather than switch

Switches

- How does switch know that A' reachable via interface 4, B' reachable via interface 5?
  - each switch has a switch table, with each entry:

(MAC address of host, interface to reach host, time when the entry created)

MAC Address	Interface	Time
62-FE-F7-11-89-A3	4	9:32
7C-BA-B2-B4-91-10	5	9:36
...	...	...



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

looks like a routing table!
difference
- forwarding using MAC address
- constructed differently

Switch table

- Switch table initially empty

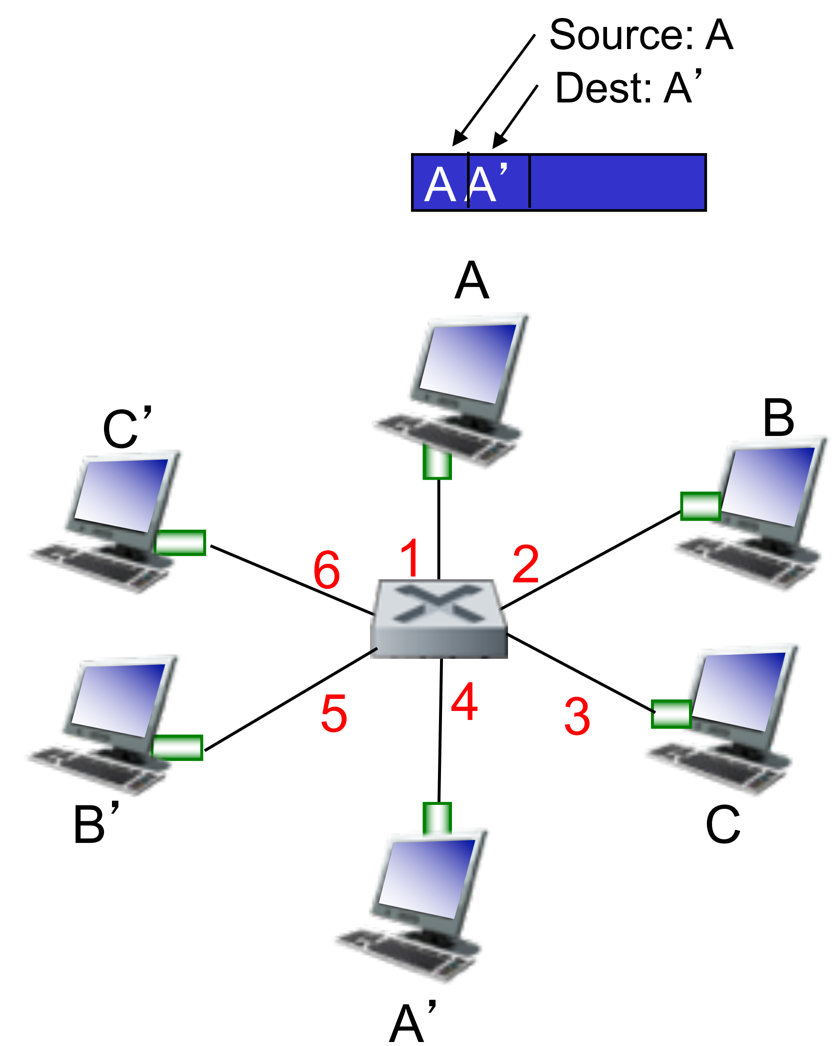
MAC Address	Interface	Time

- Switch learns which hosts can be reached through which interfaces by received frames
  - for each incoming frame, store in the table the MAC address in the source addr. field, interface from which received, and the current time

MAC Address	Interface	Time
54-3E-87-21-91-B6	1	10:36

- Delete entry if no frames received after some aging time

self-learning, plug-and-play



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

Switching

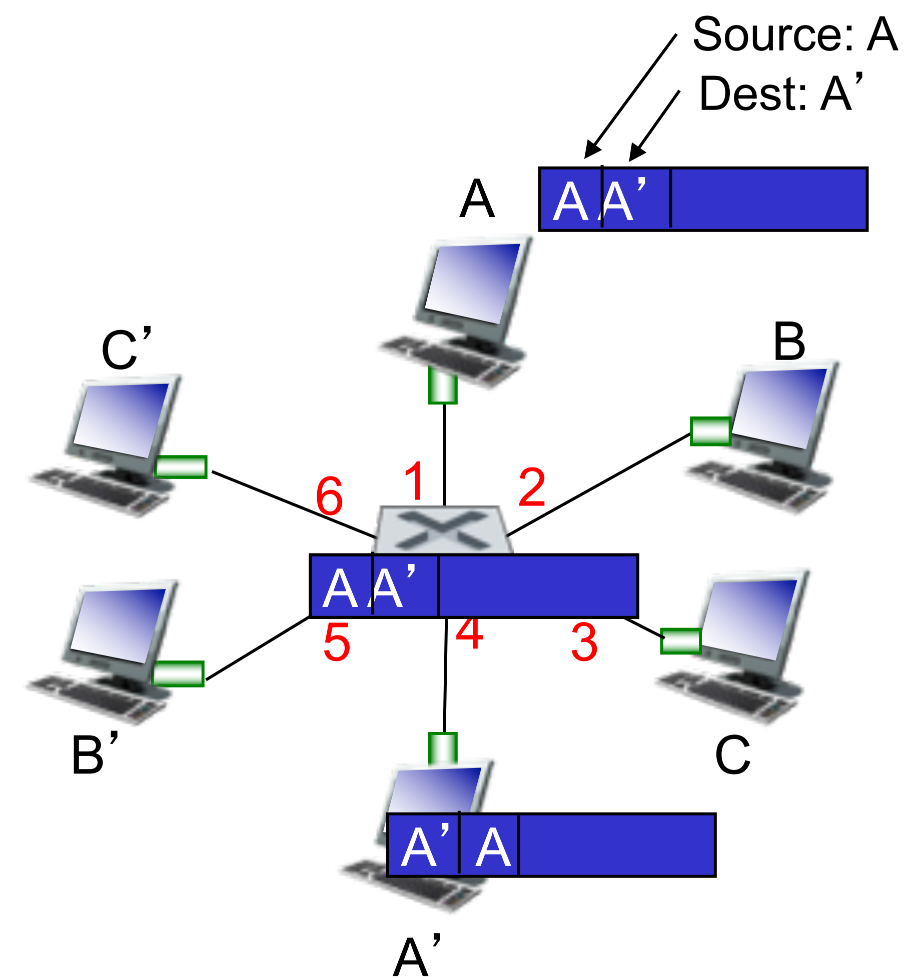
When a frame is received:

1. record switch table entry associated with sending host
2. look up switch table using MAC dest. address
3. **if** entry found for destination
   then
   {
      if dest. on link from which frame arrived
         then drop the frame
         else forward the frame on interface indicated
   }
   else flood

filtering

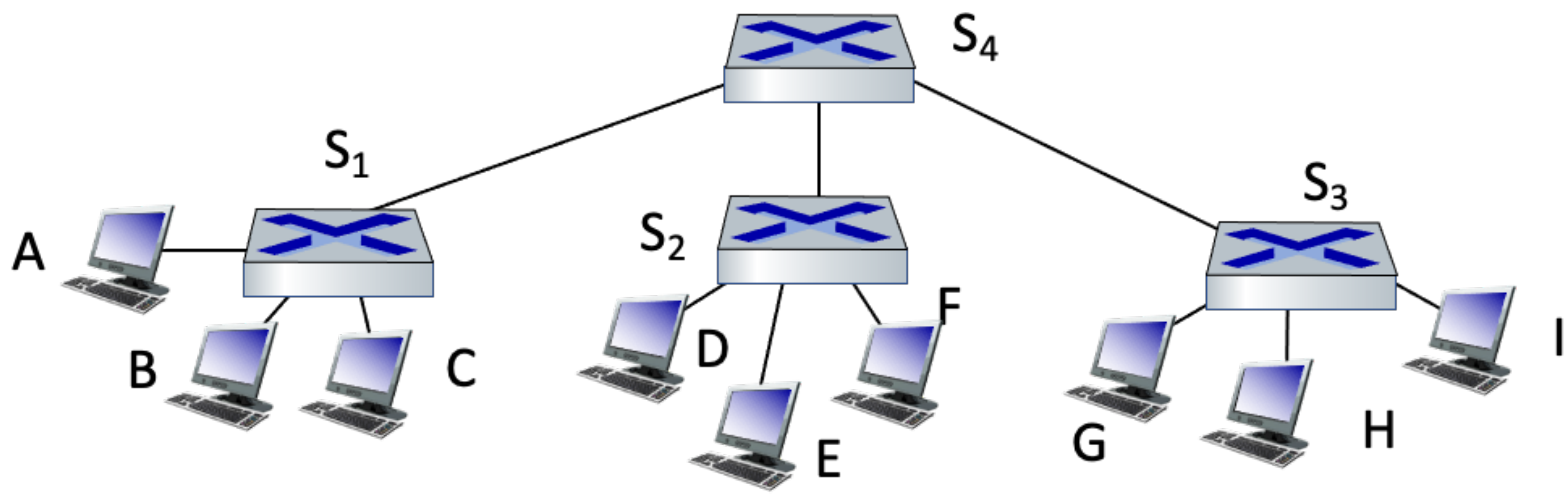
forwarding

forward on all but the interface on which the frame arrived



Interconnecting switches

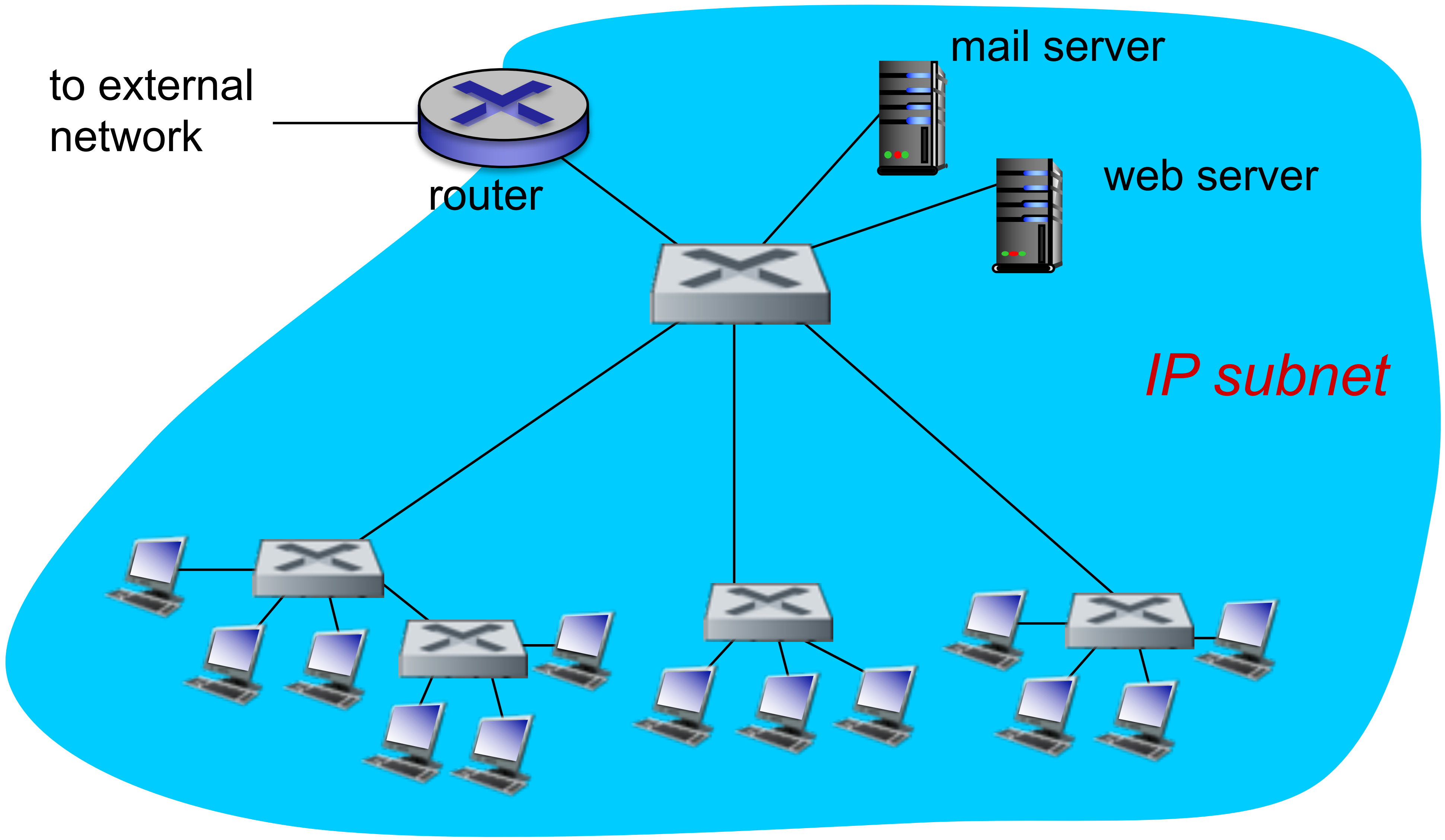
Switches can be connected together



- Example: sending from A to G
  - how does S_1 know to forward frame destined to G via S_4 and S_3?

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

Institutional network (a LAN)



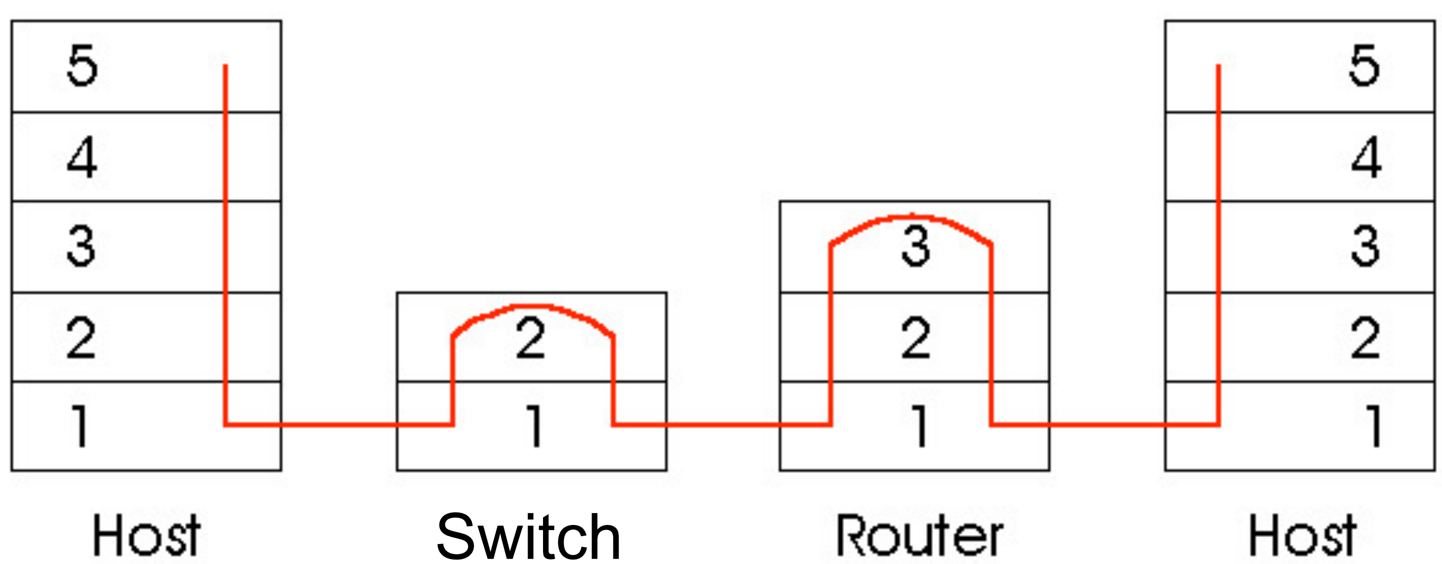
Inside a large LAN, hosts can be connected using routers as well

Switches vs. Routers

- Both store-and-forward devices and can be used for interconnections in LANs

- Routers
  - network-layer devices (examine network-layer headers)
  - maintain routing tables, implement routing algorithms
  - not plug-and-play
  - suitable for larger LAN with thousands of hosts

- Switches
  - link-layer devices (examine link-layer headers)
  - maintain switch tables, implement self-learning
  - plug-and-play
  - suffice for small network with a few hundreds of hosts

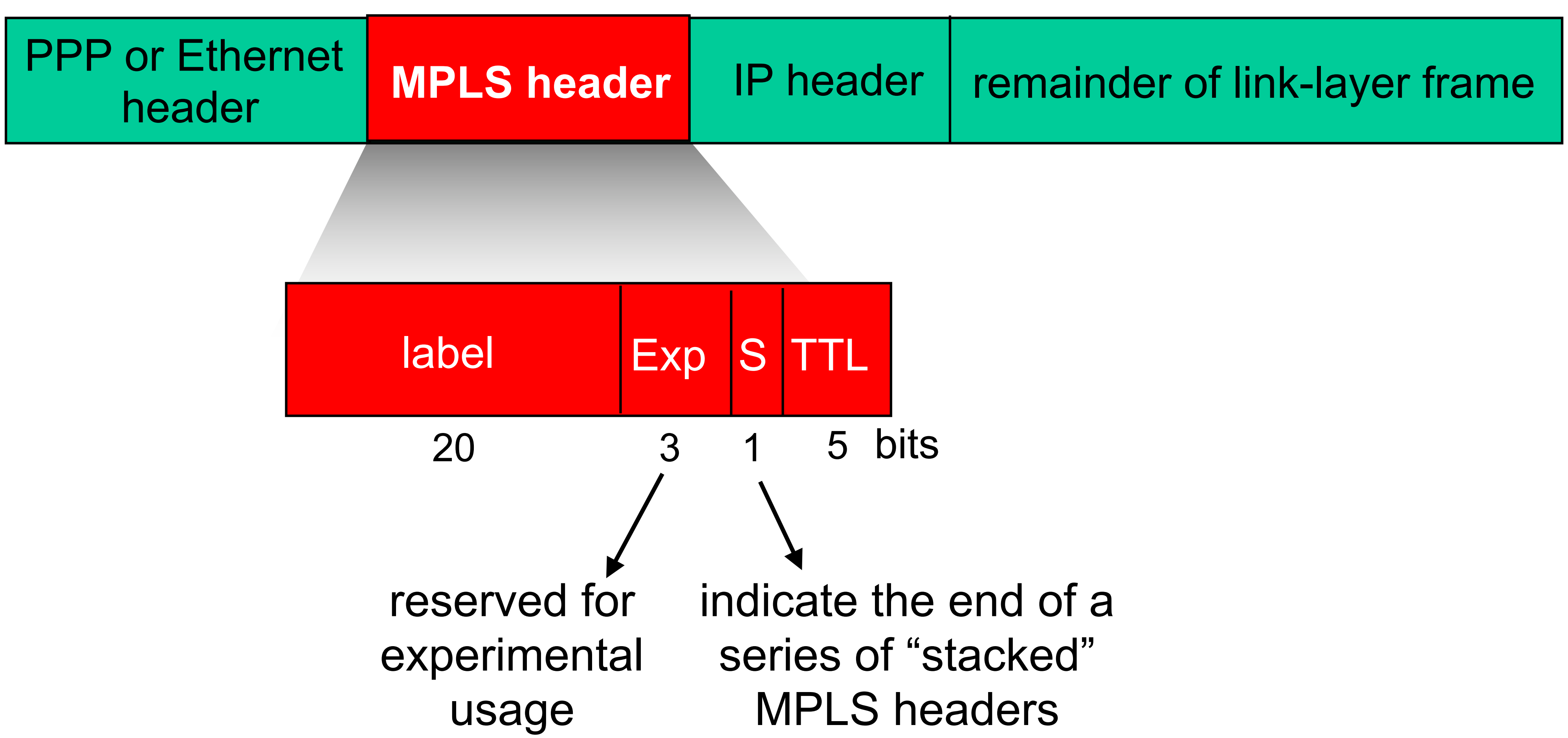


MPLS

- Multiprotocol Label Switching (MPLS)
- Initial goal: high-speed forwarding of selective IP datagrams based on *fixed length label* (instead of destination IP address)
  - fast lookup using fixed length identifier (rather than longest prefix matching)
  - borrowing ideas from Virtual Circuit (VC) approach (network-layer connection service)

MPLS

- Format of link-layer frame transmitted between MPLS-capable devices (e.g., routers)
  - RFC3032 defines format of MPLS header

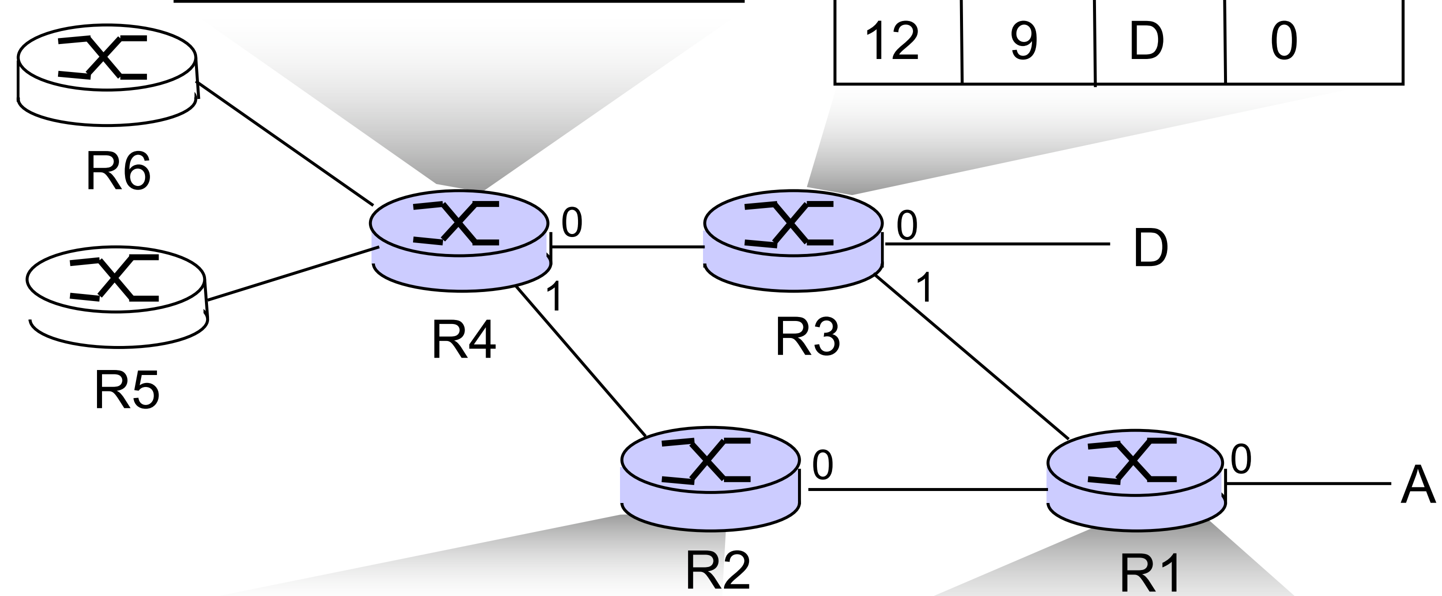


MPLS-capable routers

- a.k.a. label-switched router
- forward packets to outgoing interfaces based only on label value (do not inspect IP address)
  - MPLS forwarding table different from IP forwarding table

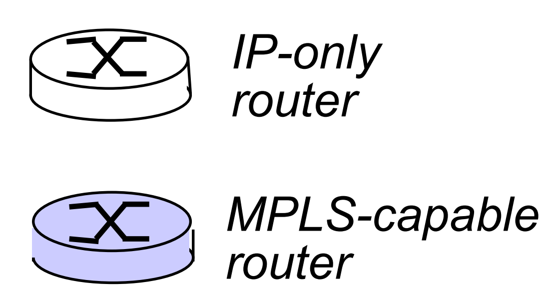
in label	out label	dest	out interface
	10	A	0
	12	D	0
	8	A	1

in label	out label	dest	out interface
10	6	A	1
12	9	D	0



in label	out label	dest	out interface
8	6	A	0

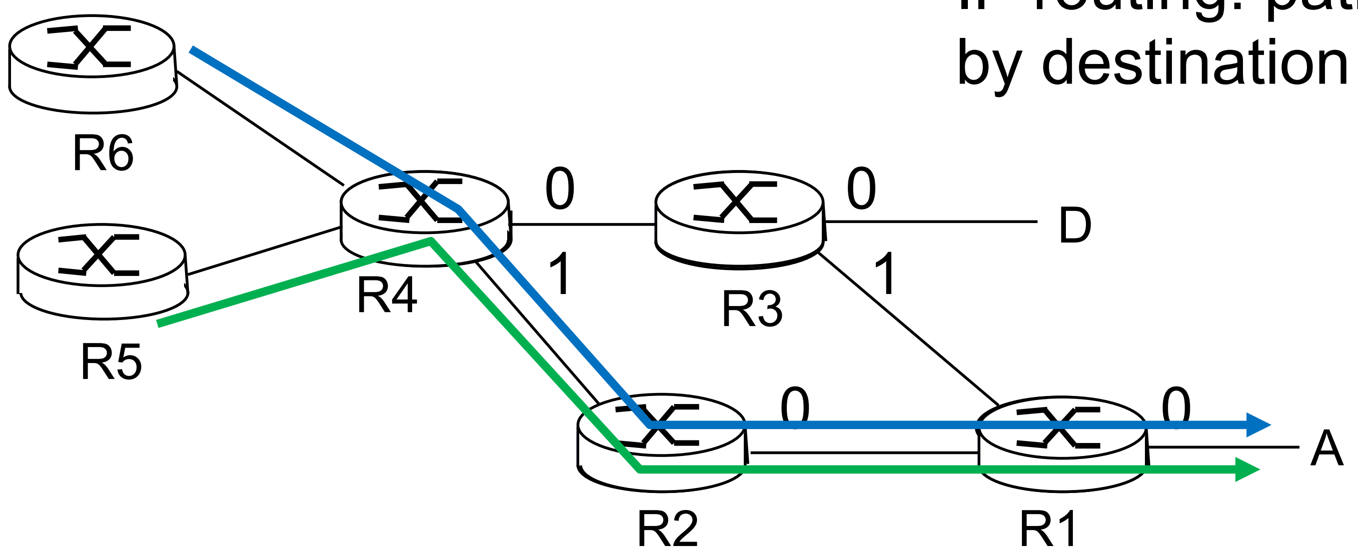
in label	out label	dest	out interface
6	-	A	0



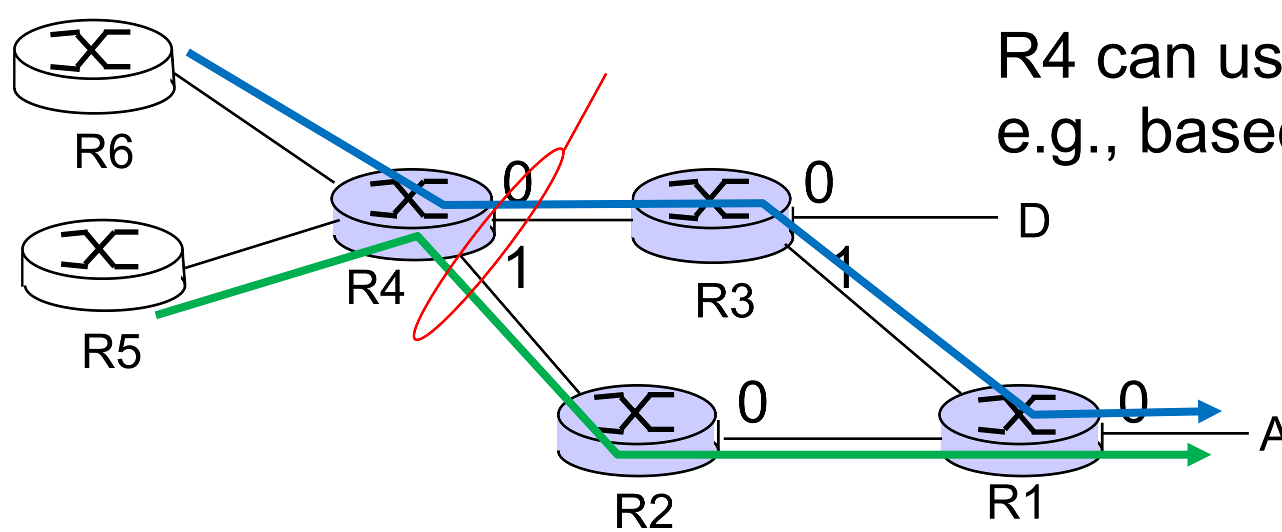
MPLS usefulness

- MPLS enables traffic management and flexibility
  - for example, use both destination and source addresses to route flows to same destination differently (traffic engineering)

IP routing: path to destination determined by destination address alone



R4 can use different MPLS routes to A, e.g., based on source address



MPLS usefulness (cont'd)

- Re-route flows quickly if link fails: pre-computed backup routes (e.g., useful for VoIP)

- MPLS can be used to implement VPN (virtual private network)
  - To implement a VPN for a customer, an ISP uses its MPLS-enabled network to connect together the customer's various networks

Data center network

- Hundreds or thousands of interconnected hosts in a data center

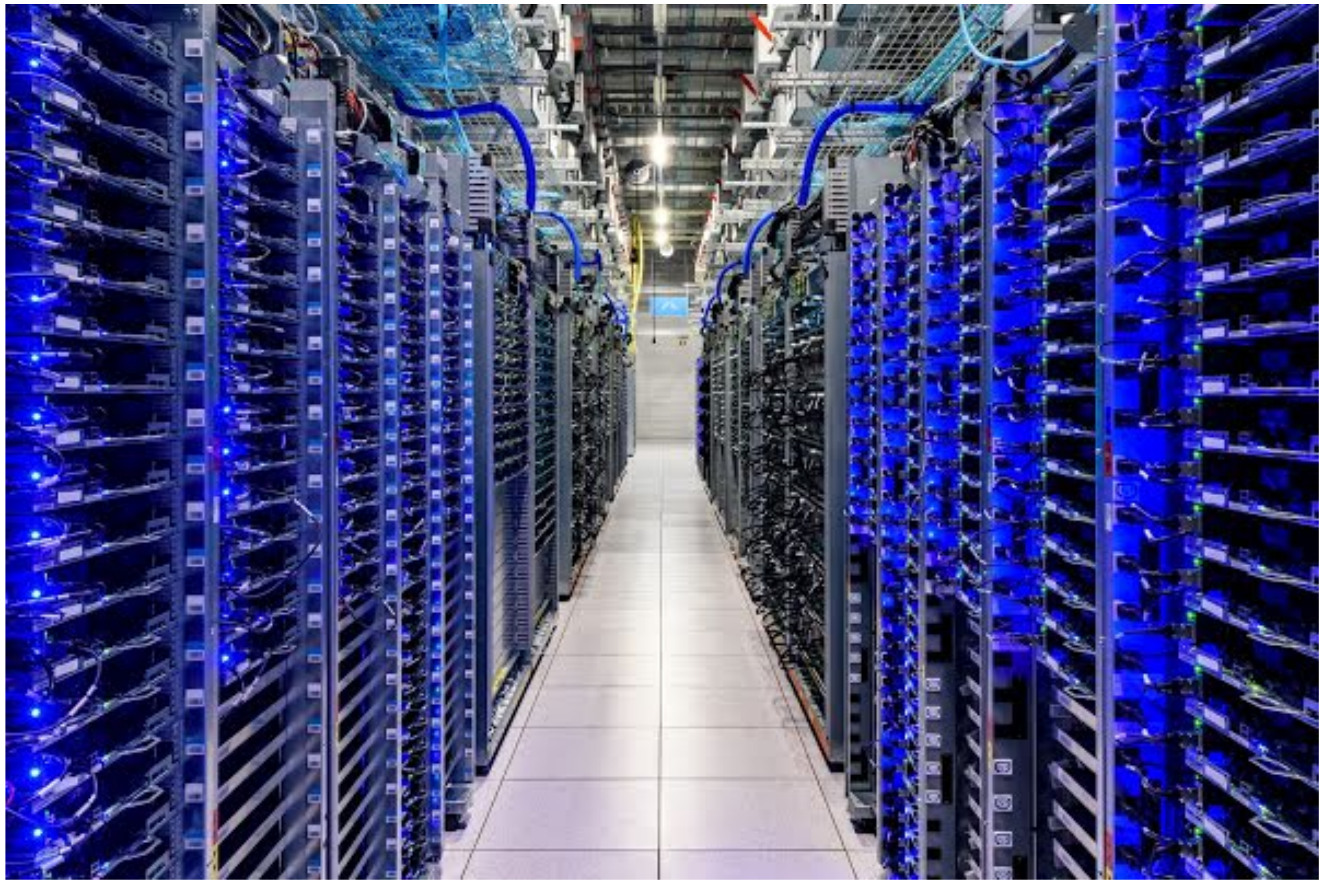
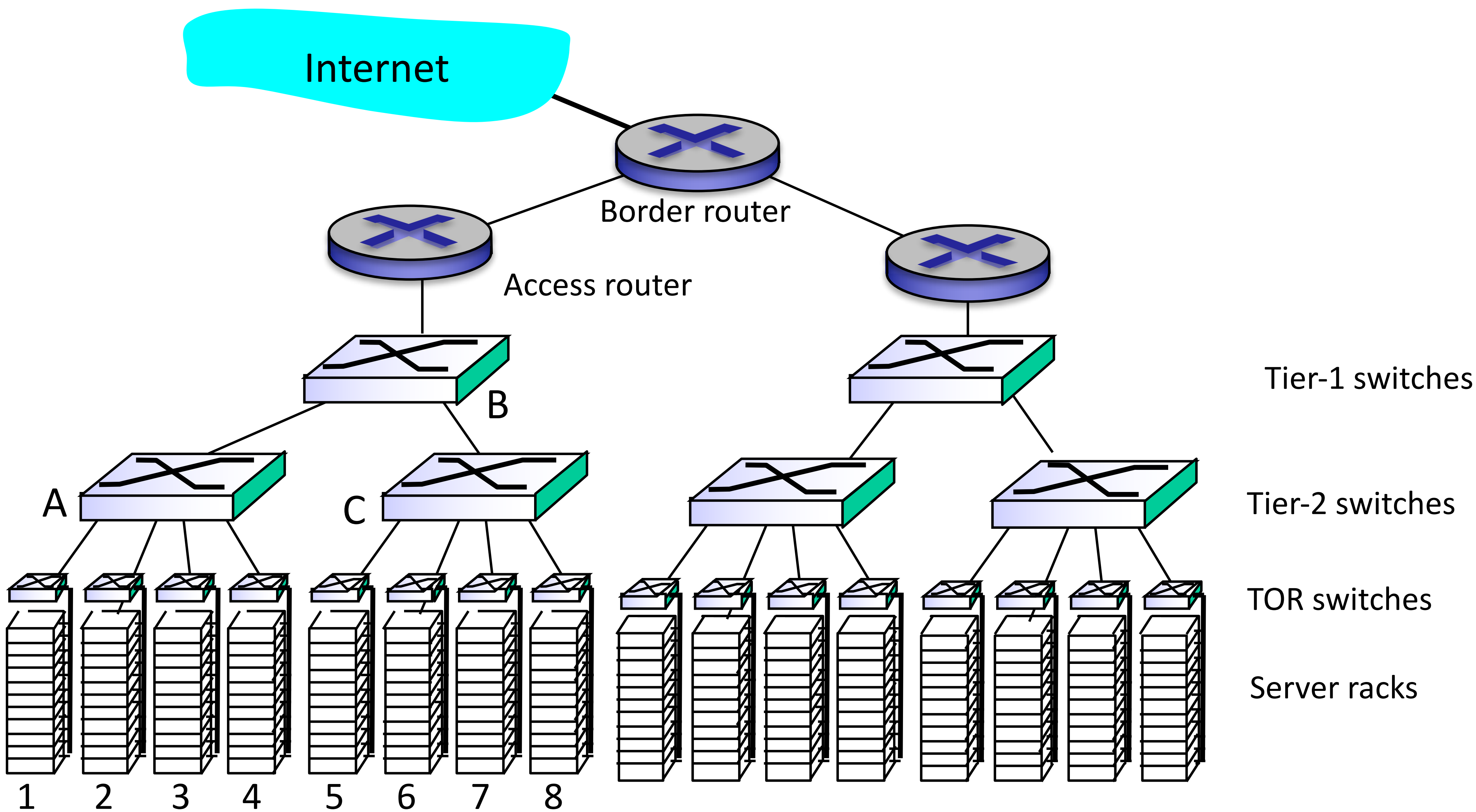


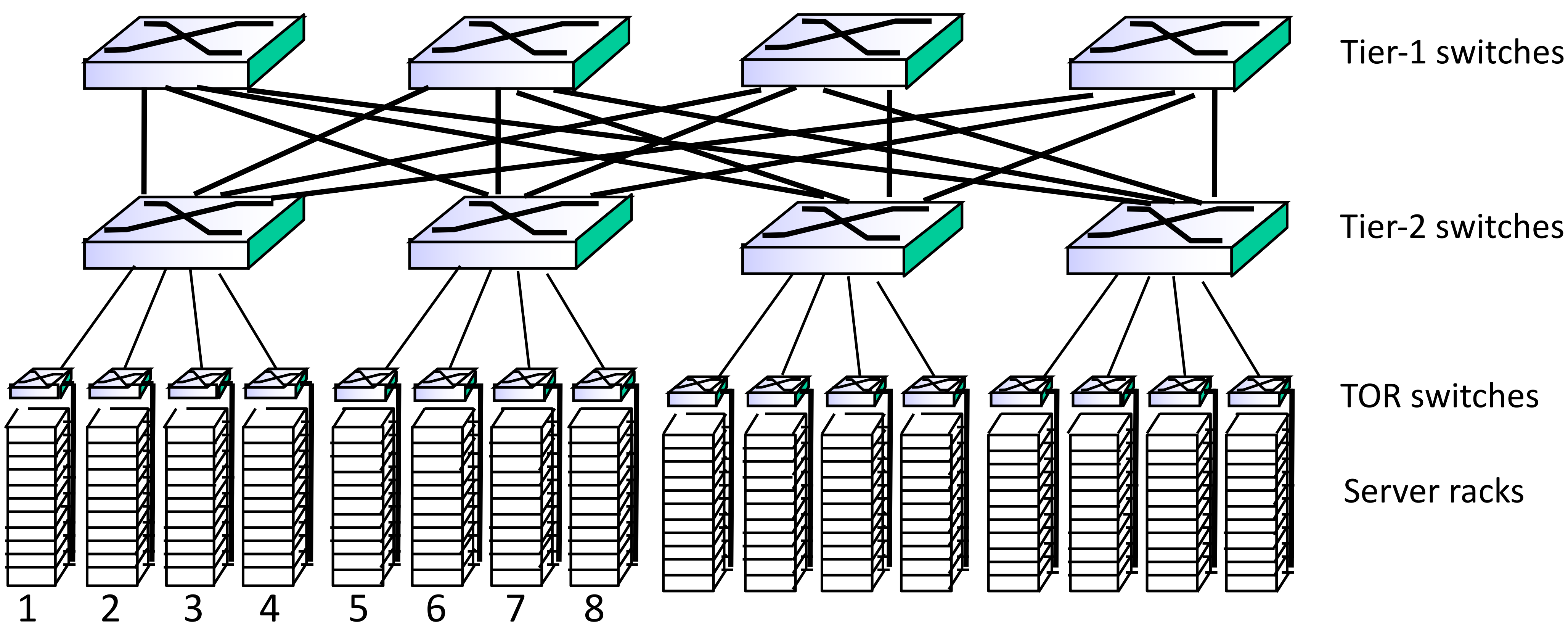
Image from: https://www.google.com/about/datacenters/data-security/

Data center network (cont'd)



Data center network (cont'd)

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



- Required Reading:
  - *Computer Networking: A Top-Down Approach* (8th Edition)
    Ch 6.4.1, 6.4.2, 6.4.3, 6.5, 6.6

- Acknowledgement:
  - Some materials are extracted from the slides created by Prof. Jim F. Kurose and Prof. Keith W. Ross for the textbook.