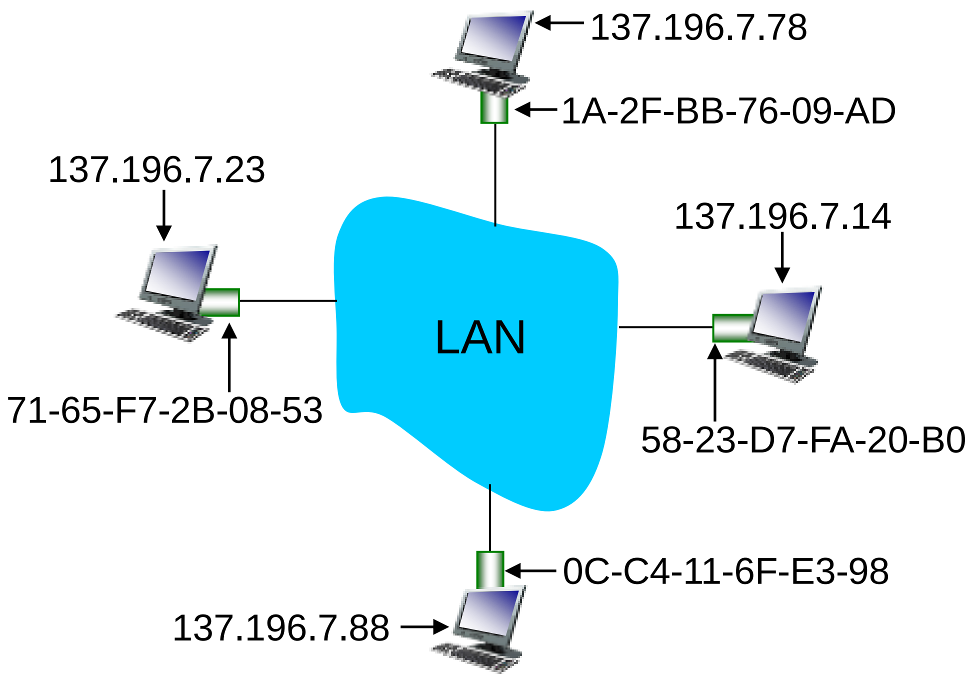
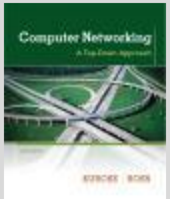
Link-layer addressing, Ethernet, VLANs



Computer Networking: A Top Down Approach
6th edition
Jim Kurose, Keith Ross
Addison-Wesley



Link layer, LANs: outline

5.1 Introduction, services

5.2 Error detection, correction

5.3 Multiple access protocols

5.4 LANs

- Addressing, ARP
- Ethernet
- Switches
- VLANS

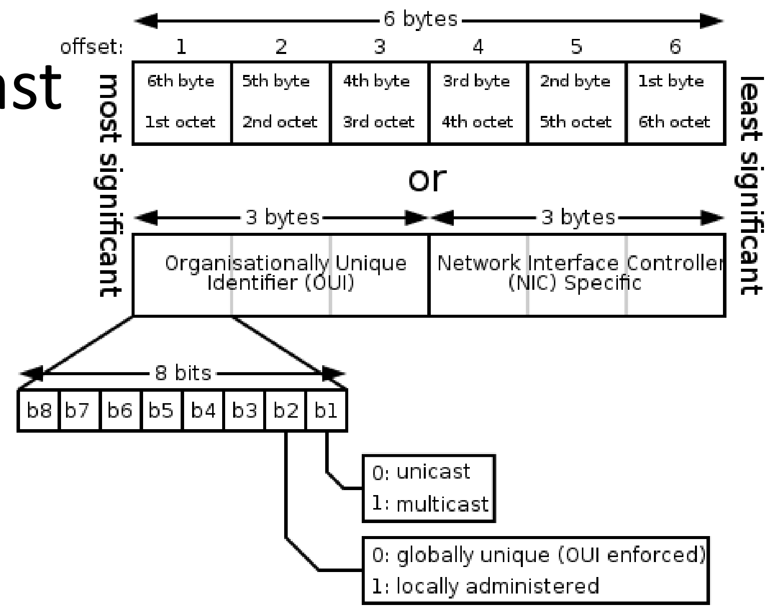
5.5 Link virtualization: MPLS

5.6 Data center networking

5.7 A day in the life of a web request

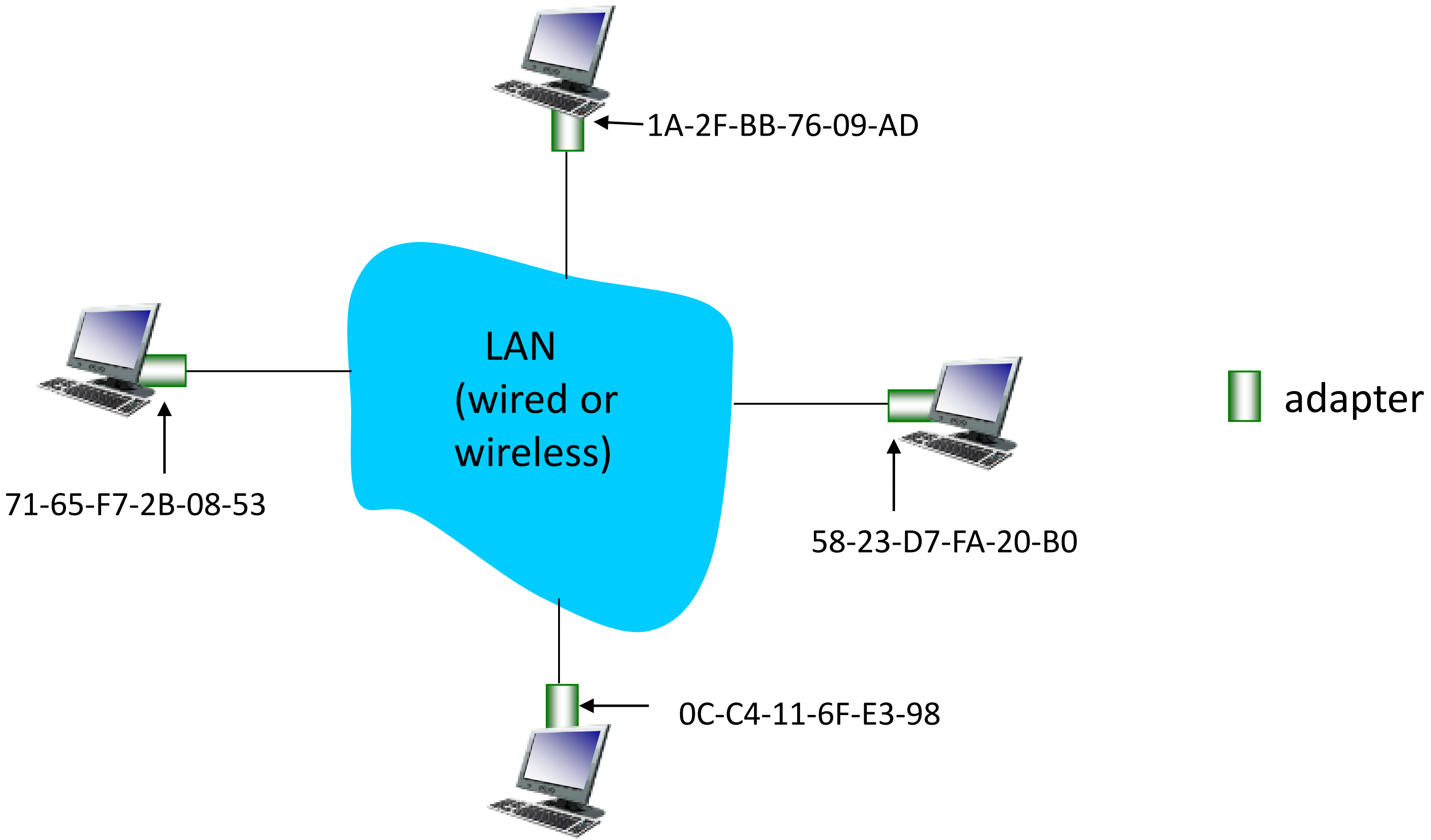
Link-layer addressing

- Media Access Control address (MAC)
  - 48-bit globally unique address
    - 281,474,976,710,656 possible addresses
    - Should last till 2100
    - e.g. 01:23:45:67:89:ab
  - Address of all 1's is broadcast
    - FF:FF:FF:FF:FF:FF



LAN addresses and ARP

Each adapter on LAN has unique *MAC* address
Including routers, NAT boxes, etc.



Address translation

- Problem:
  - How does host send a message to someone on their own network? Their default router?
  - IP address is not the link-level address (e.g. MAC)
- Solution:
  - Host maintains table: IP address -> link address
  - Using the Address Resolution Protocol (ARP)

ARP procedure

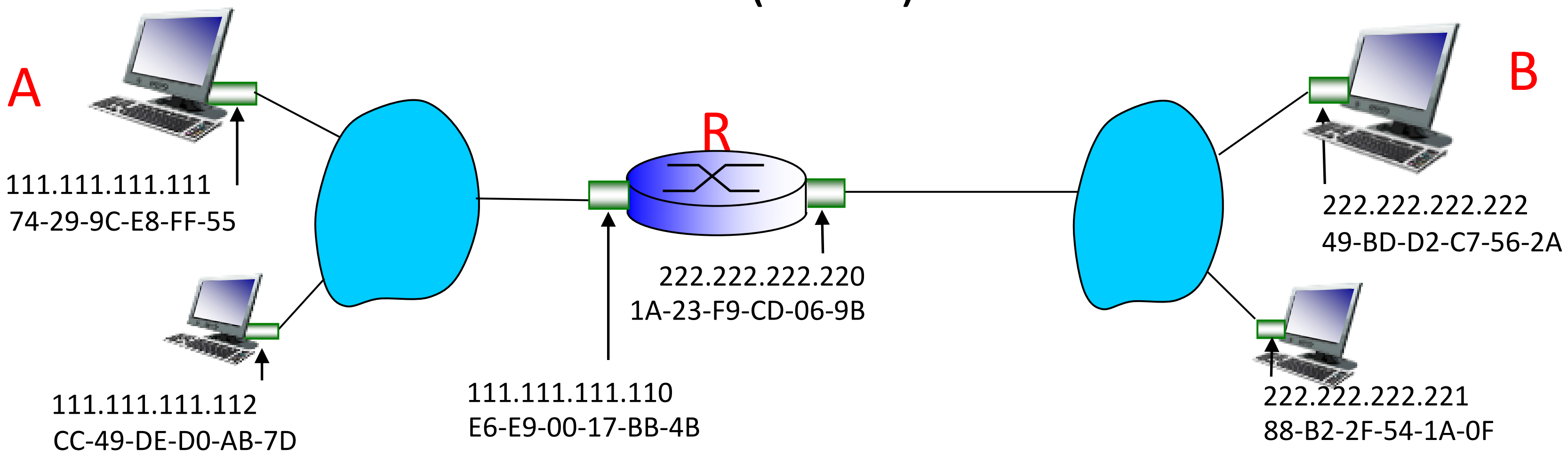
- If destination IP in sender's ARP table:
  - Fire off link-layer packet
  - Otherwise send ARP query using broadcast address
- ARP query:
  - IP address you're looking for
  - Your own IP and hardware address
  - Destination responds with hardware address
  - Other hosts can ignore or refresh their ARP tables
- Plug-and-play, no intervention from admin

```
vertanen@katie:/usr/sbin$ arp
Address                  HWtype  HWaddress           Flags Mask            Iface
10.33.73.39              ether   c8:2a:14:53:da:e5   C                     eth0
lugnut                   ether   00:16:3e:c2:af:f0   C                     eth0
10.33.73.254             ether   00:12:f2:81:ca:74   C                     eth0
10.33.73.120             ether   78:2b:cb:ac:9c:0a   C                     eth0
mtcs32.mtech.edu         ether   b8:ac:6f:45:56:ef   C                     eth0
10.33.73.121             ether   78:2b:cb:ac:9c:1b   C                     eth0
vertanen@katie:/usr/sbin$
```

Addressing: routing to another LAN

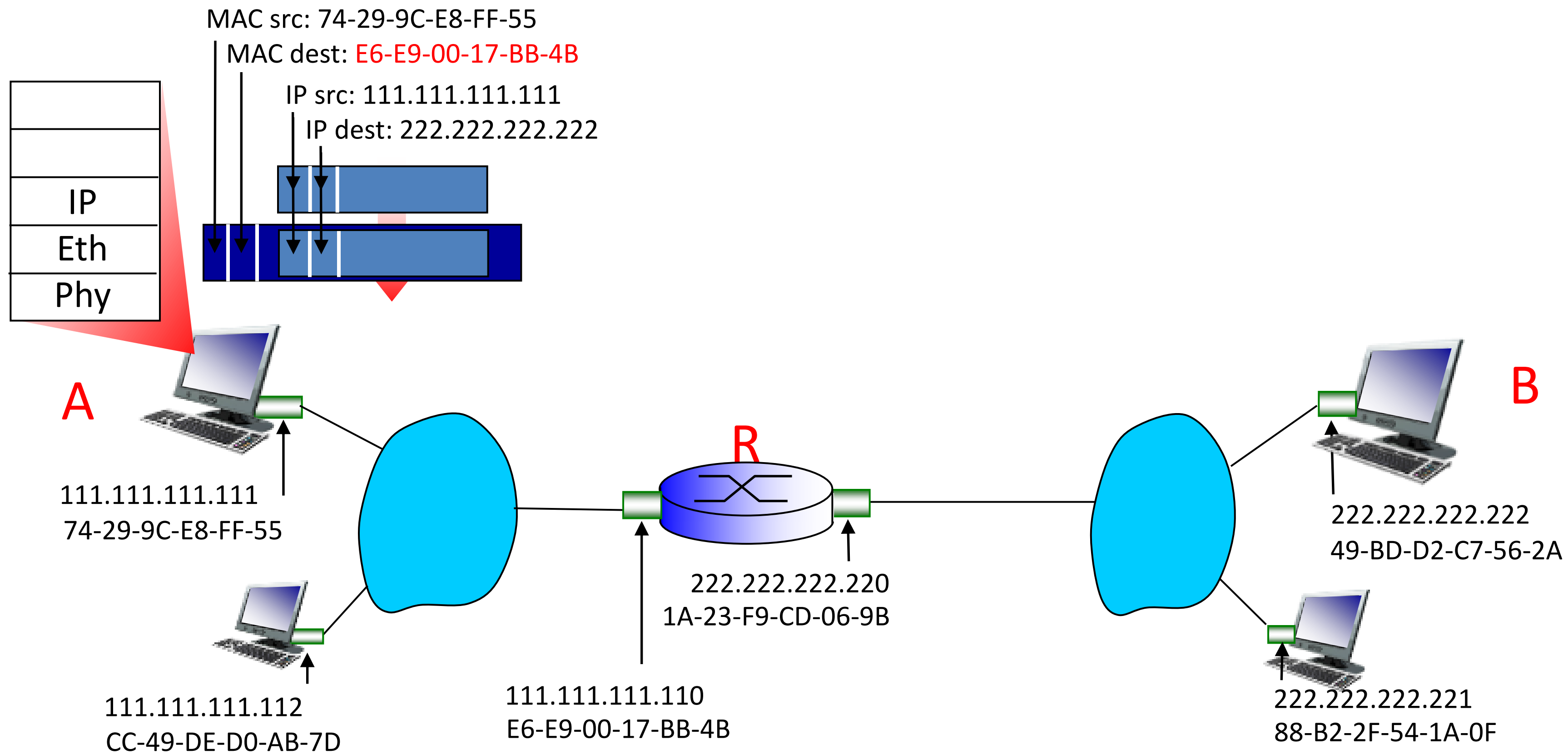
Walkthrough: Send datagram from A to B via R

- Focus on addressing
  - At IP (datagram) and MAC layer (frame)
- Assume:
  - A knows B's IP address
  - A knows IP address of first hop router, R (how?)
  - A knows R's MAC address (how?)



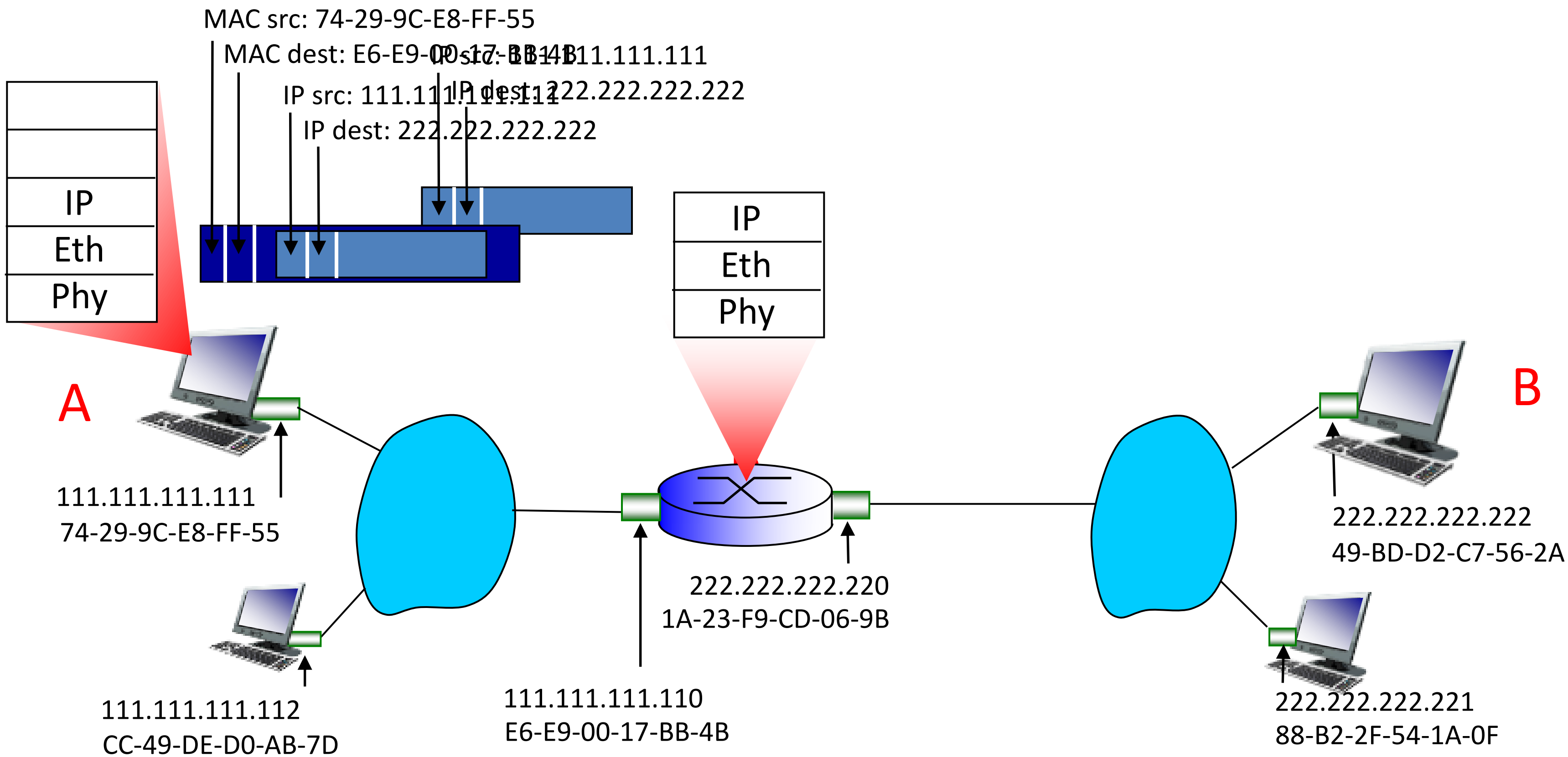
Addressing: routing to another LAN

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



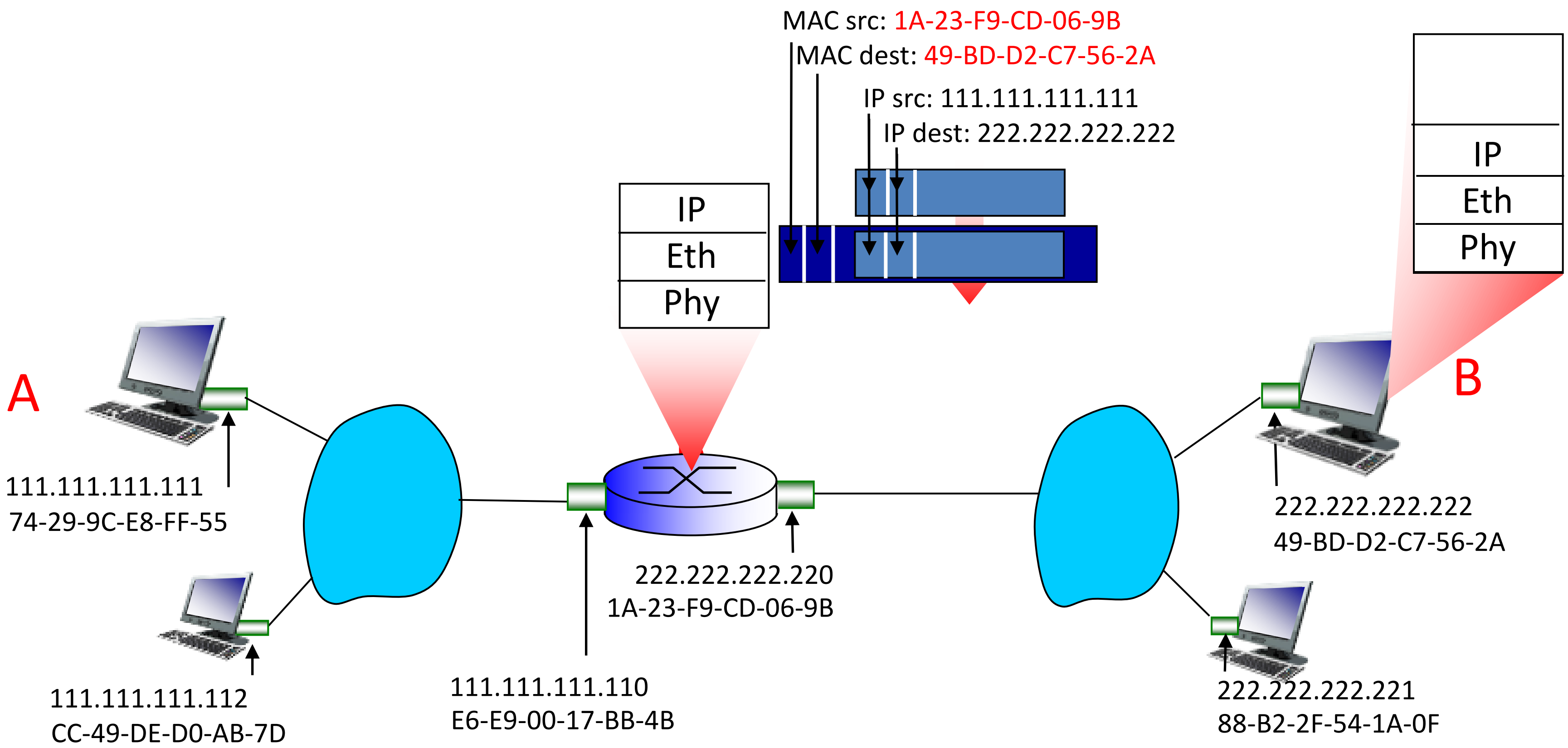
Addressing: routing to another LAN

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



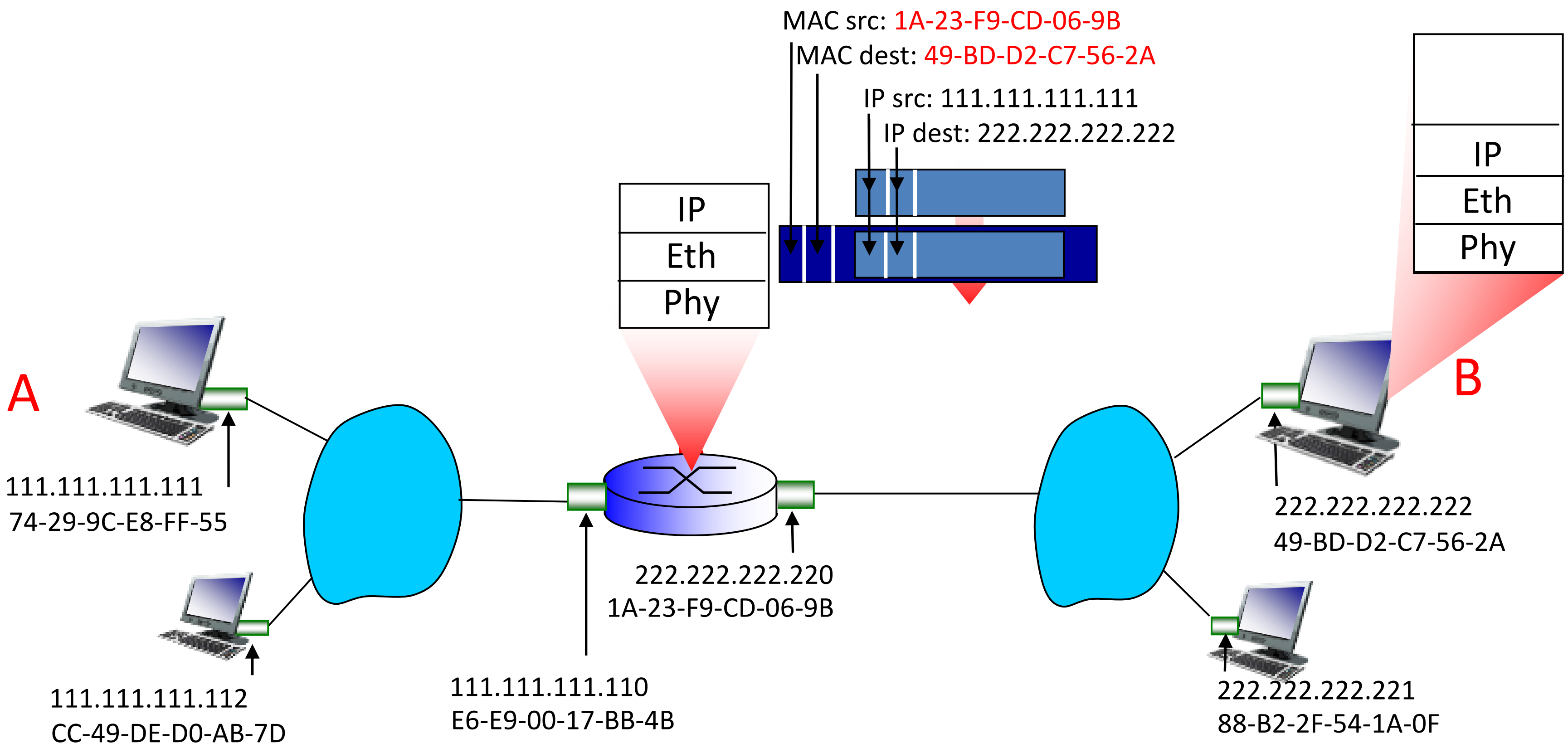
Addressing: routing to another LAN

- 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



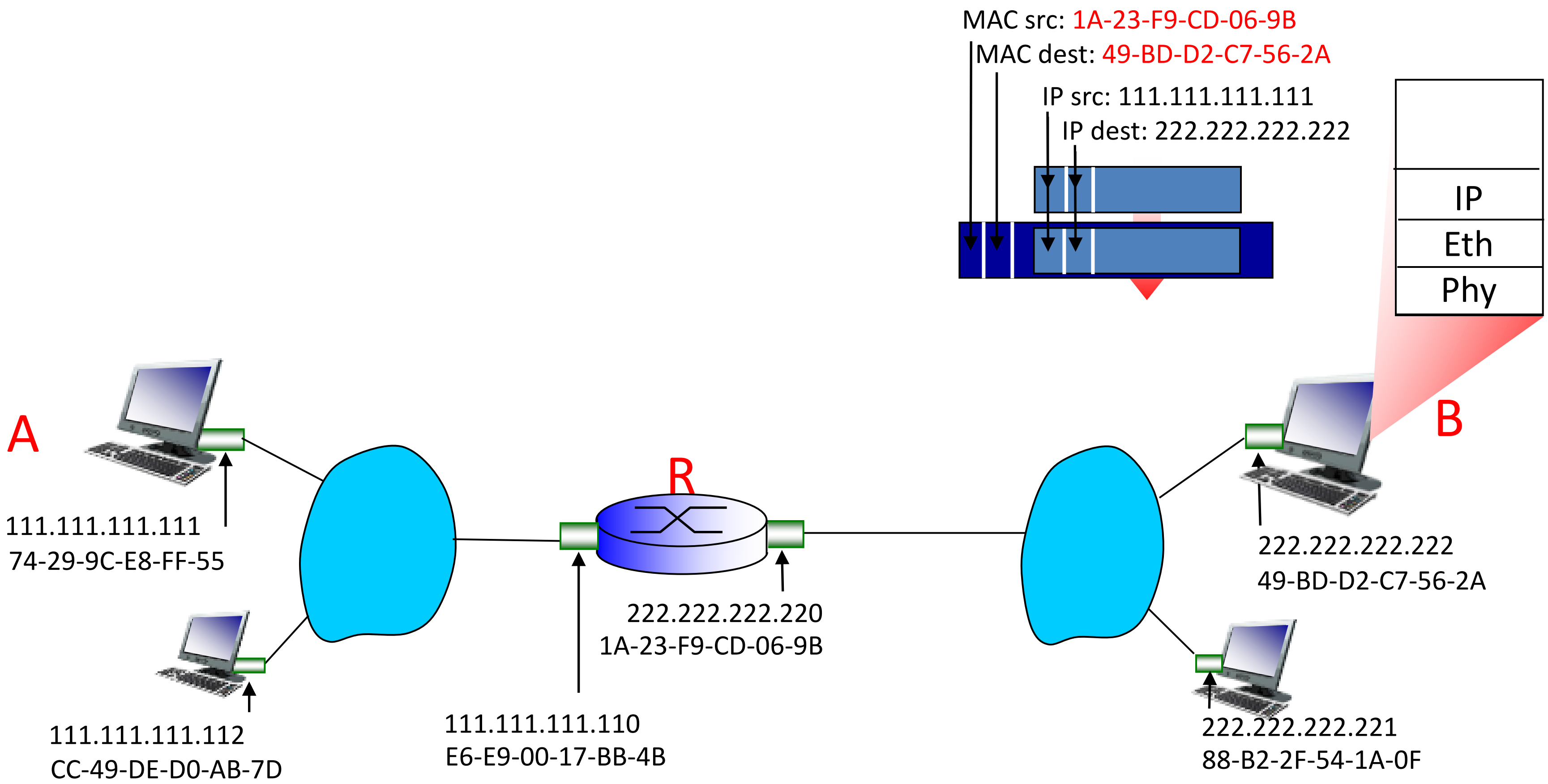
Addressing: routing to another LAN

- 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



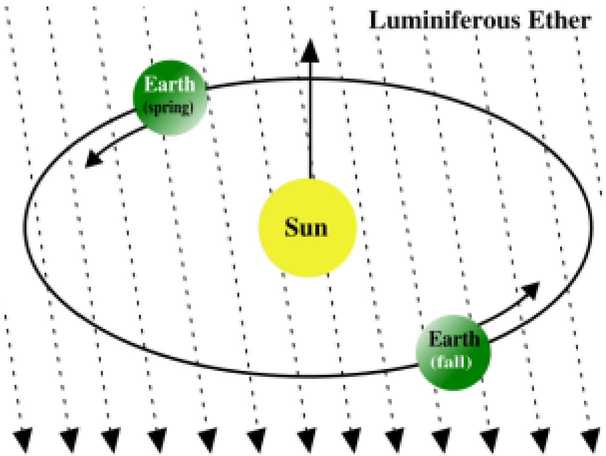
Addressing: routing to another LAN

- 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



Classic Ethernet

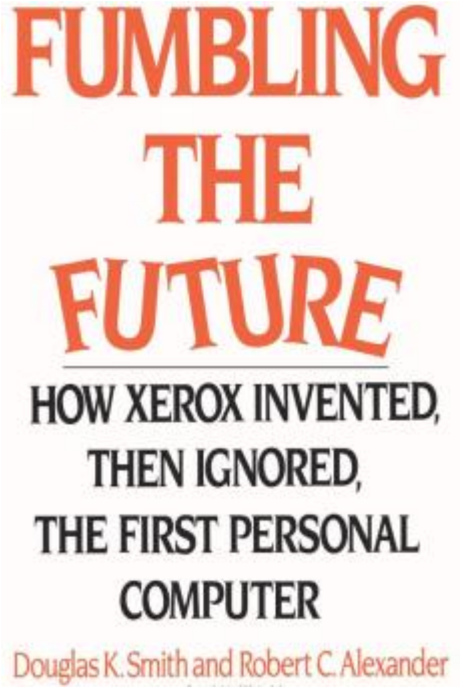
- Ethernet
  - luminferous ether through which electromagnetic radiation once thought to propagate
  - Carrier Sense, Multiple Access with Collision Detection (CSMA/CD)
  - IEEE 802.3



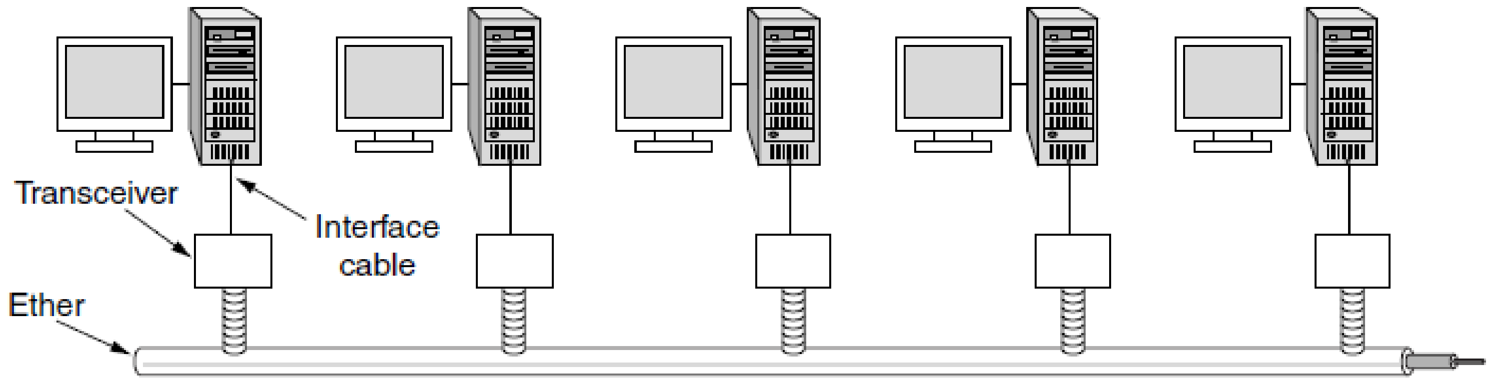
Robert Metcalfe, co-inventor of Ethernet

Classic Ethernet

- Ethernet
  - Xerox Ethernet standardized as IEEE 802.3 in 1983
  - Xerox not interested in commercializing
  - Metcalfe leaves and forms 3Com



Classic Ethernet connectivity



- Shared medium
  - All hosts hear all traffic on cable
  - Hosts tapped the cable
  - 2500m maximum length
  - May include repeaters amplifying signal
  - 10 Mbps bandwidth

Classic Ethernet cabling

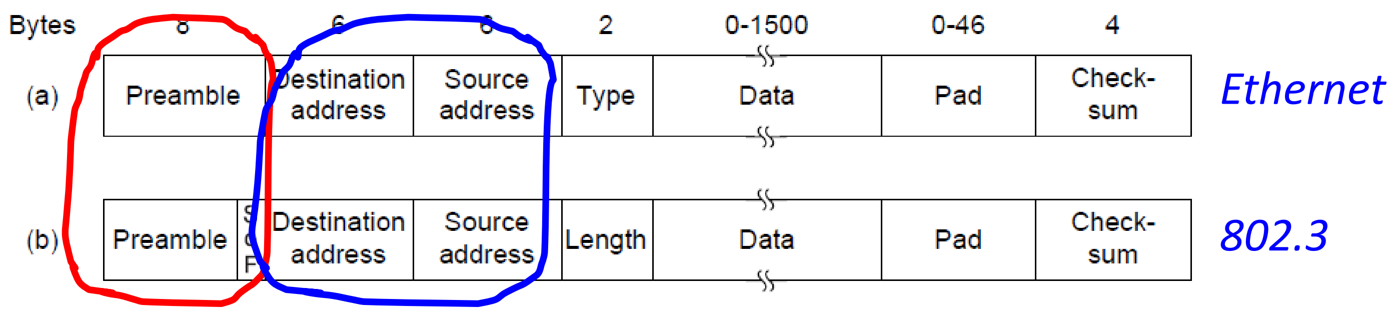
Thick Ethernet cable (yellow), 10BASE-5 transceivers, cable tapping tool (orange), 500m maximum length.

Cable after being "vampire" tapped.

Thin Ethernet cable (10BASE2) with BNC T-connector, 185m maximum length.

Ethernet frame format

- Frame format
  - Manchester encoded
  - Preamble products 10-Mhz square wave
    - Allows clock synch between sender & receiver
  - Pad to at least 64-bytes (collision detection)



Alternating 0's and 1's (except SoF of 11)

48-bit MAC addresses

Ethernet receivers

- Hosts listens to medium
  - Deliver to host:
    - Any frame with host's MAC address
    - All broadcast frames (all 1's)
    - Multicast frames (if subscribed to)
    - Or all frames if in promiscuous mode

MAC sublayer

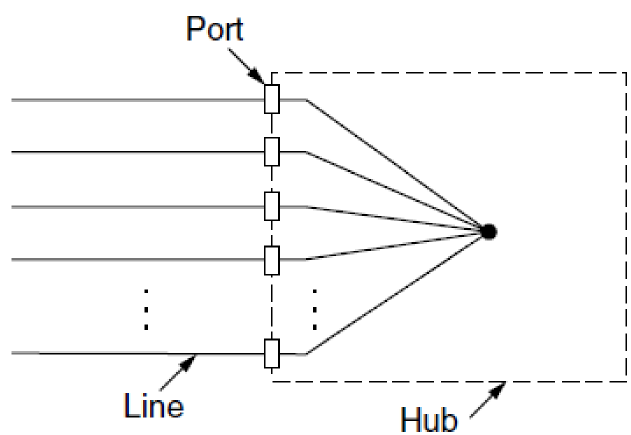
- Media Access Control (MAC) sublayer
  - Who goes next on a shared medium
  - Ethernet hosts can sense if medium in use
  - Algorithm for sending data:
    1. Is medium idle? If not, wait.
    2. Start transmitting data, listen for collision.
    3. If collision detected, transmit 32-bit jamming sequence. Stop transmitting and go to backoff procedure.

Backoff procedure

- Binary exponential backoff
  - First collision
    - Wait 0-1 timeslots (chosen at random)
  - Second collision
    - Wait 0-3 timeslots
  - In general, i^{th} collision
    - Wait a random number of timeslots between 0 and 2^i - 1 (max of 1023 slots)
  - Give up after 16 or so retries
  - Timeslot = 51.2 μs

Ethernet hubs

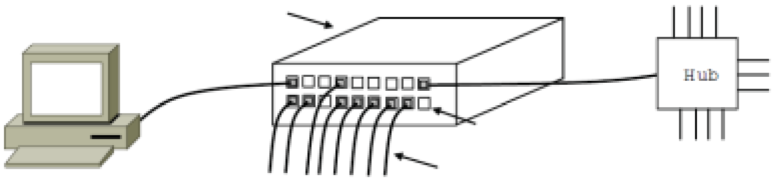
- Long single cable
  - Hard to find breaks or loose connections
- Different wiring pattern
  - Each host wired straight to hub
  - Hub simply connected all wires together
  - Using existing office twisted pair phone lines



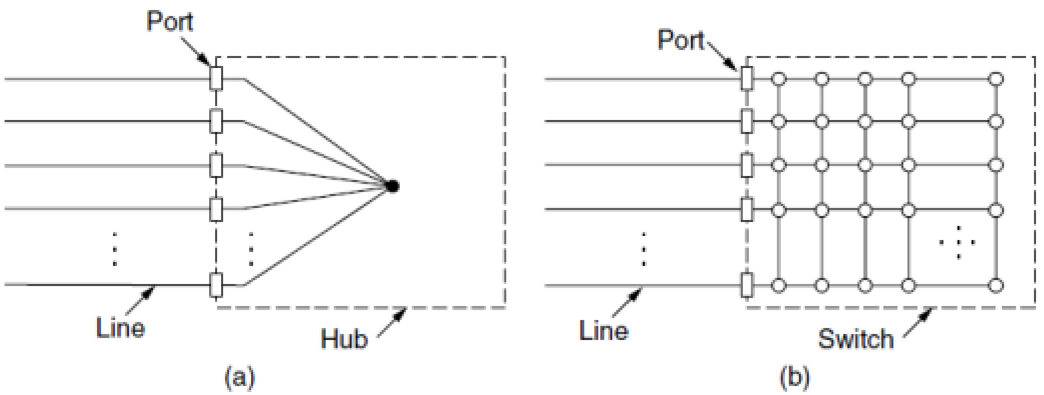
(a)

Switched Ethernet

- Hubs
  - Made network easier to manage
  - But did not address capacity problem



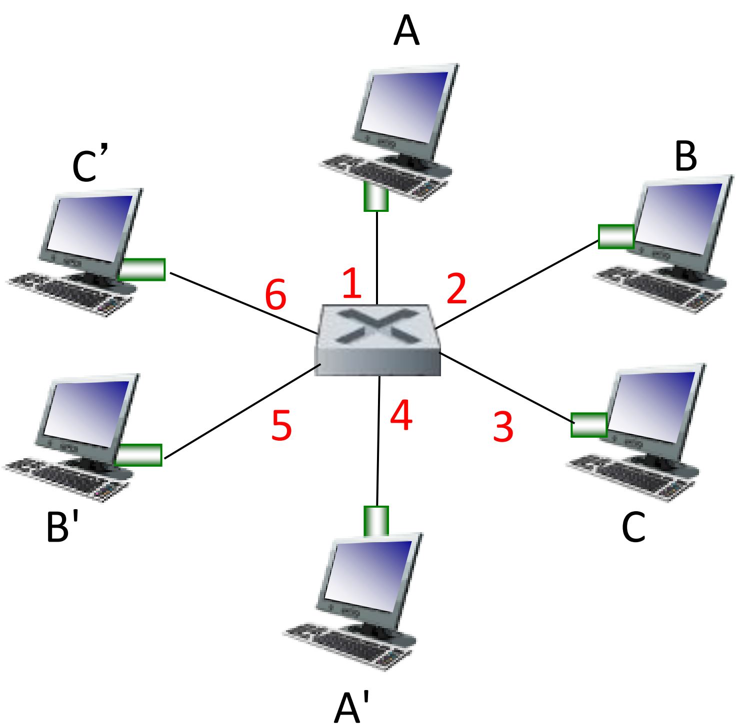
- Switches
  - High-speed backplane connecting all ports
  - Only output frame to destination port
  - Isolates traffic, no collisions, better security



Switch forwarding table

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

- A: Each switch has a switch table, each entry:
  - (MAC address of host, interface to reach host, time stamp)
  - Looks like a routing table!



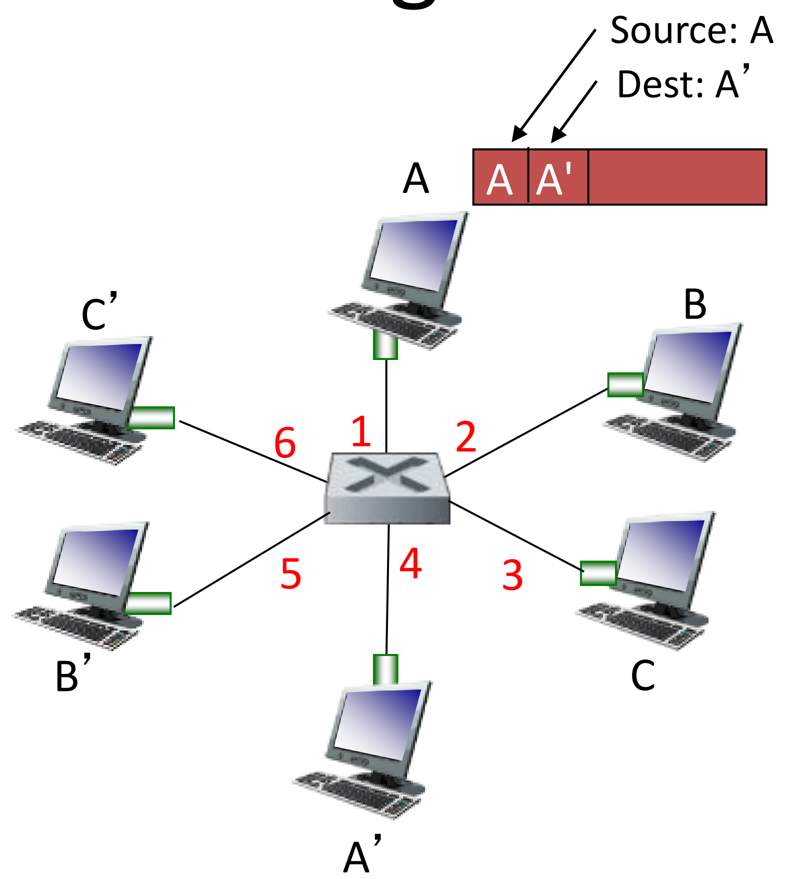
Switch with six interfaces (1,2,3,4,5,6)

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

- Something like a routing protocol?

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	TTL
A	*1*	*60*

Switch table (initially empty)

Switch: frame filtering/forwarding

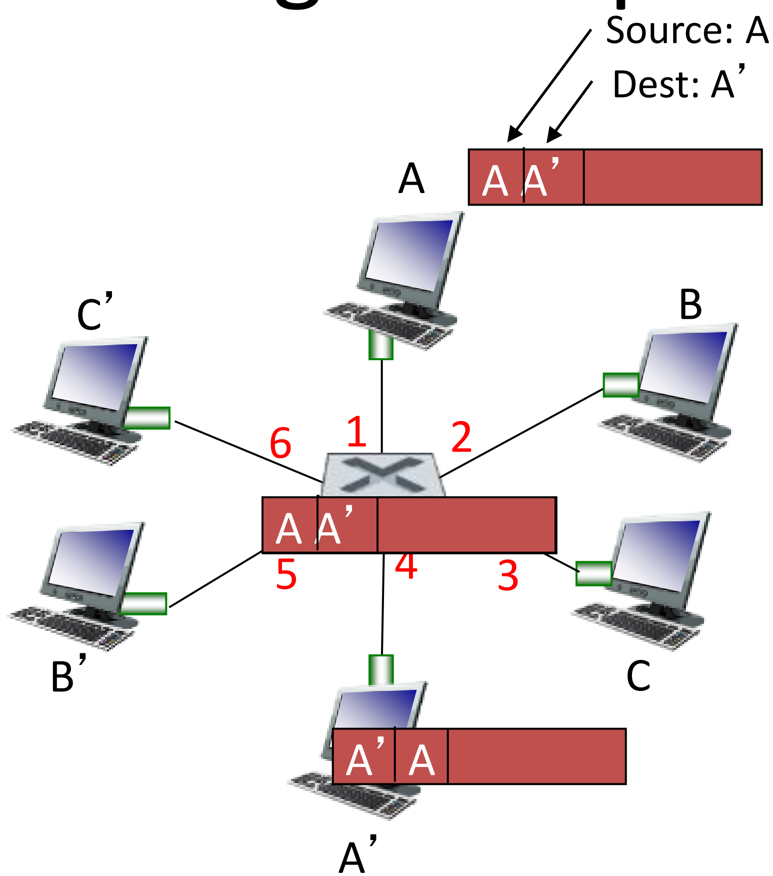
When frame received at switch:

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 frame
      else forward frame on interface indicated by entry
   }
   else flood /* fwd on all ports except arriving interface */

Self-learning, forwarding: example

- Frame destination, A', location unknown: *flood*

❖ Destination A location known: selectively send on just one link

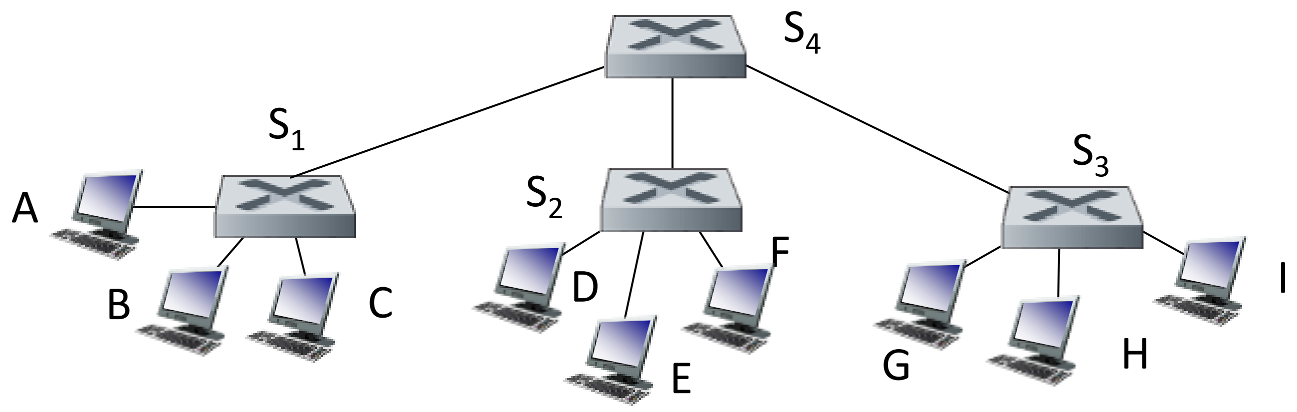


MAC addr	interface	TTL
A	*1*	*60*
A'	*4*	*60*

switch table (initially empty)

Interconnecting switches

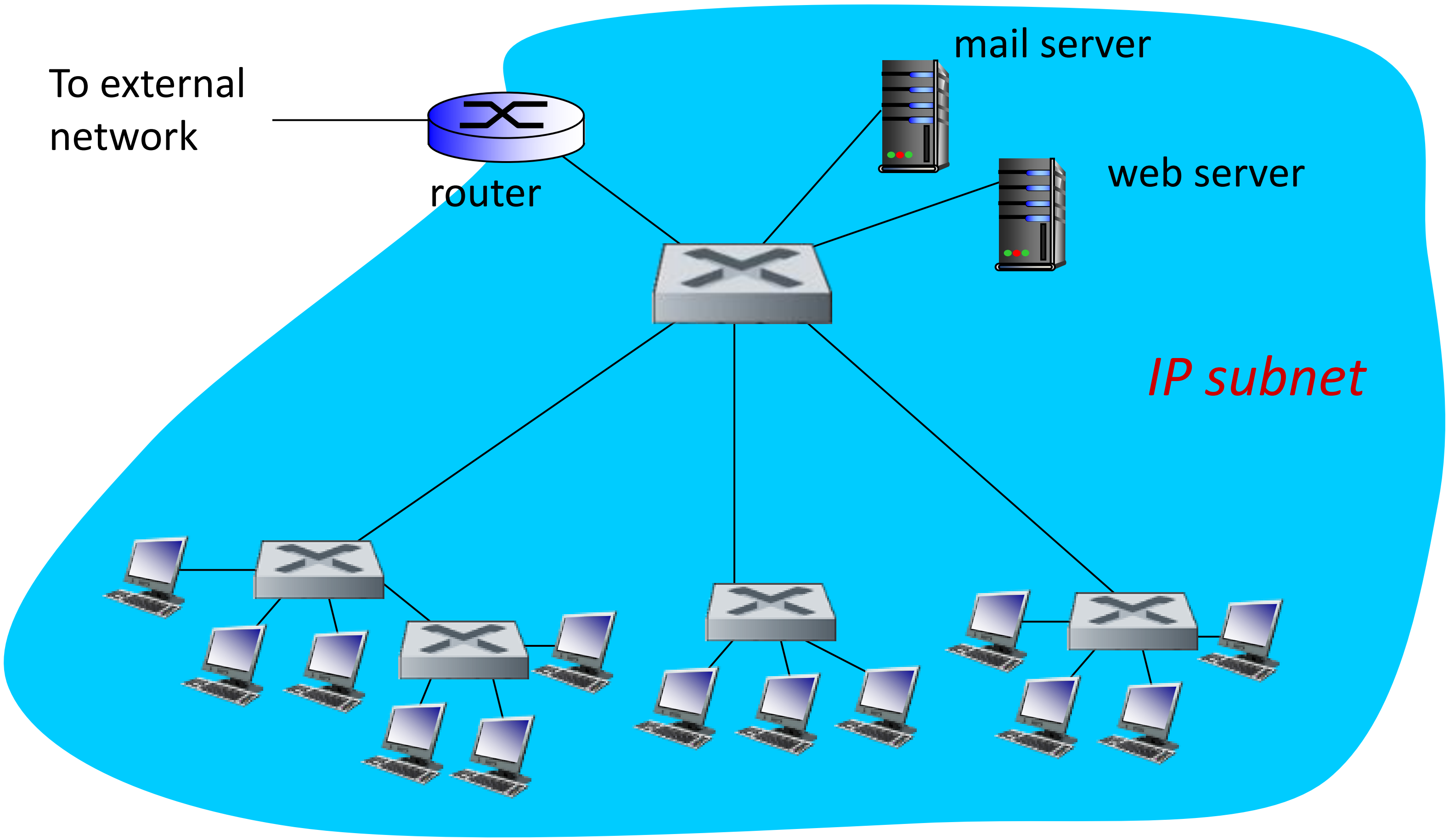
- Switches can be connected together



Q: Sending from A to G - how does S_1 know to forward frame destined to F via S_4 and S_3?

- A: Self-learning! (works exactly the same as in single-switch case!)

Institutional network



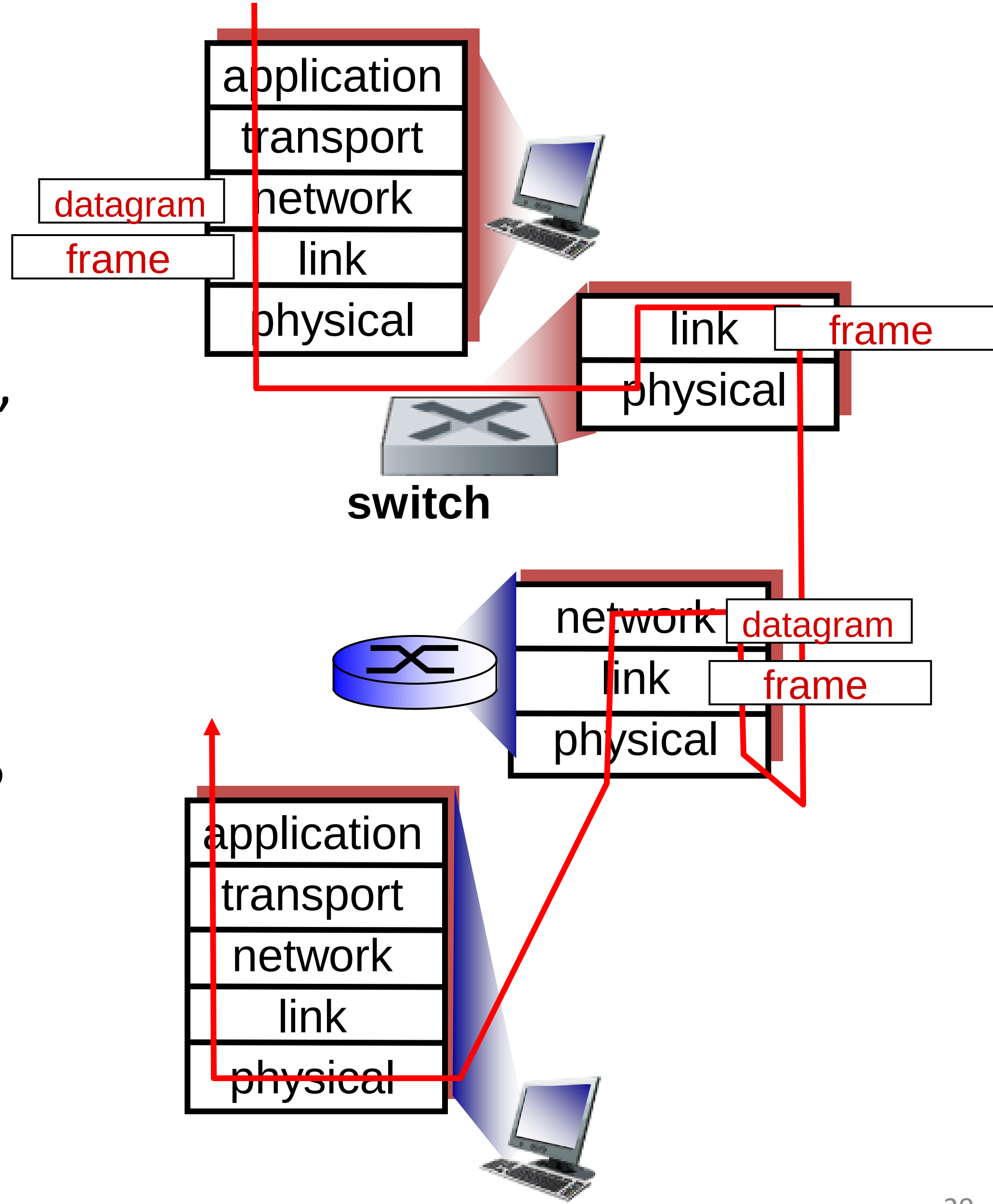
Switches vs. routers

Both are store-and-forward:

- *Routers:* network-layer devices, layer-3 packet switch
- *Switches:* link-layer devices, layer-2 packet switch

Both have forwarding tables:

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



Fast Ethernet

- Fast Ethernet
  - IEEE 802.3u
  - Keep all the classic Ethernet frame formats, etc.
  - Reduce the bit time from 100 nsec to 10nsec
  - 100 Mbps
  - No more multidrop cables or vampire taps

Name	Cable	Max. segment	Advantages
100Base-T4	Twisted pair	100 m	Uses category 3 UTP
100Base-TX	Twisted pair	100 m	Full duplex at 100 Mbps (Cat 5 UTP)
100Base-FX	Fiber optics	2000 m	Full duplex at 100 Mbps; long runs

Gigabit Ethernet

- Gigabit Ethernet
  - IEEE 802.3ab
  - 1000 Mbps
  - Unacknowledged datagram service
  - Addition of flow control
  - Unofficial support for jumbo frames
    - Up to 9KB (instead of limit of 1500 bytes)

Name	Cable	Max. segment	Advantages
1000Base-SX	Fiber optics	550 m	Multimode fiber (50, 62.5 microns)
1000Base-LX	Fiber optics	5000 m	Single (10 µ) or multimode (50, 62.5 µ)
1000Base-CX	2 Pairs of STP	25 m	Shielded twisted pair
1000Base-T	4 Pairs of UTP	100 m	Standard category 5 UTP

Even faster

- 10-Gigabit Ethernet
  - 1000x faster than original Ethernet
  - Inside data centers, long haul trunks

Name	Cable	Max. segment	Advantages
10GBase-SR	Fiber optics	Up to 300 m	Multimode fiber (0.85μ)
10GBase-LR	Fiber optics	10 km	Single-mode fiber (1.3μ)
10GBase-ER	Fiber optics	40 km	Single-mode fiber (1.5μ)
10GBase-CX4	4 Pairs of twinax	15 m	Twinaxial copper
10GBase-T	4 Pairs of UTP	100 m	Category 6a UTP

- 40 and 100-Gigabit Ethernet
  - starting to be deployed

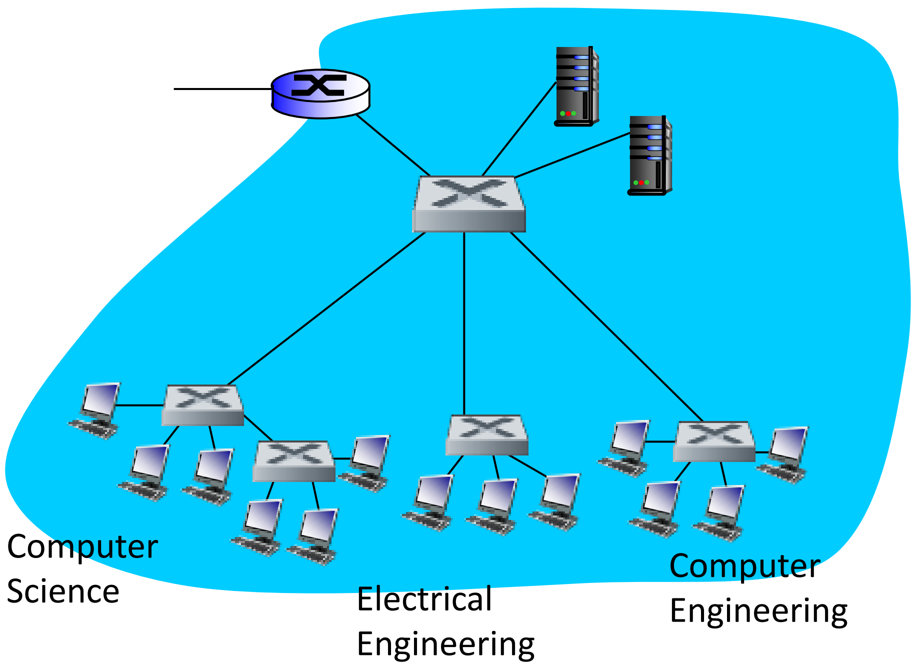
Extreme Networks **BlackDiamond X8** Chassis Switch
187.4 lbs, 18" x 30" x 25"
Designed for large, virtualized data centers and clouds, the Extreme Networks **BlackDiamond X8** switch provides high density 10GbE and 40GbE ...

$20,537
from 3 stores
Compare prices

Ethernet retrospective

- Why so popular?
  - Easy to administer, no routing or config tables
  - Cheap hardware and wiring
  - Plays nice with TCP/IP
    - Ethernet and IP are connectionless protocols
    - Alternates like ATM were not
  - Periodic speed increases
    - Order of magnitude every few years without throwing away existing infrastructure
  - Borrowed good ideas from other (failed) networking technologies (FDDI, Fiber Channel)

VLANs: motivation

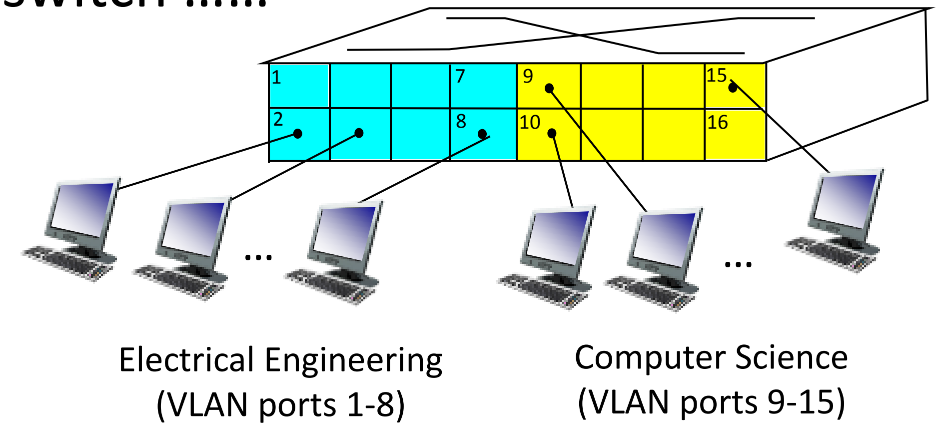


Consider:

- 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

VLANs

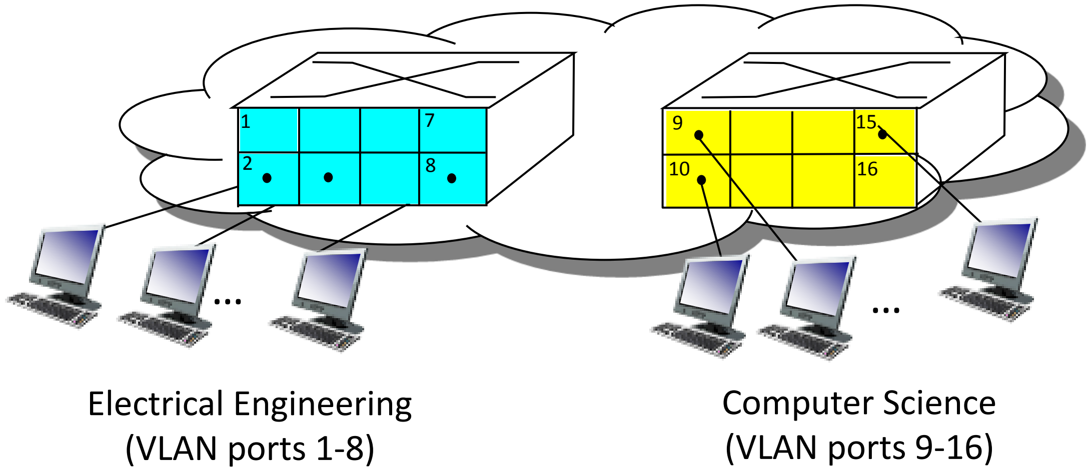
Port-based VLAN: Switch ports grouped (by switch management software) so that *single* physical switch ……



Virtual Local Area Network

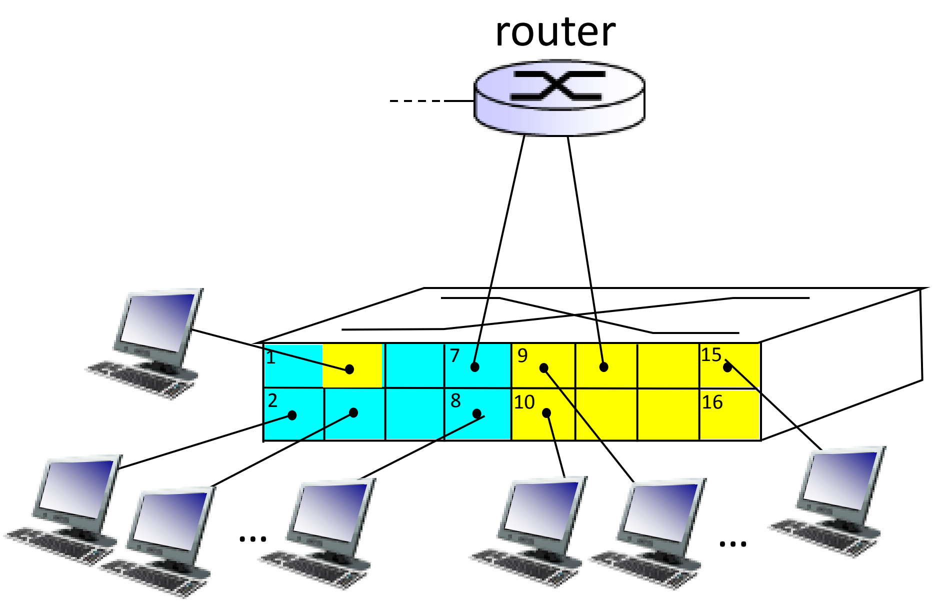
Switches supporting VLAN capabilities can be configured to define multiple ***virtual*** LANS over single physical LAN infrastructure.

… operates as multiple virtual switches



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



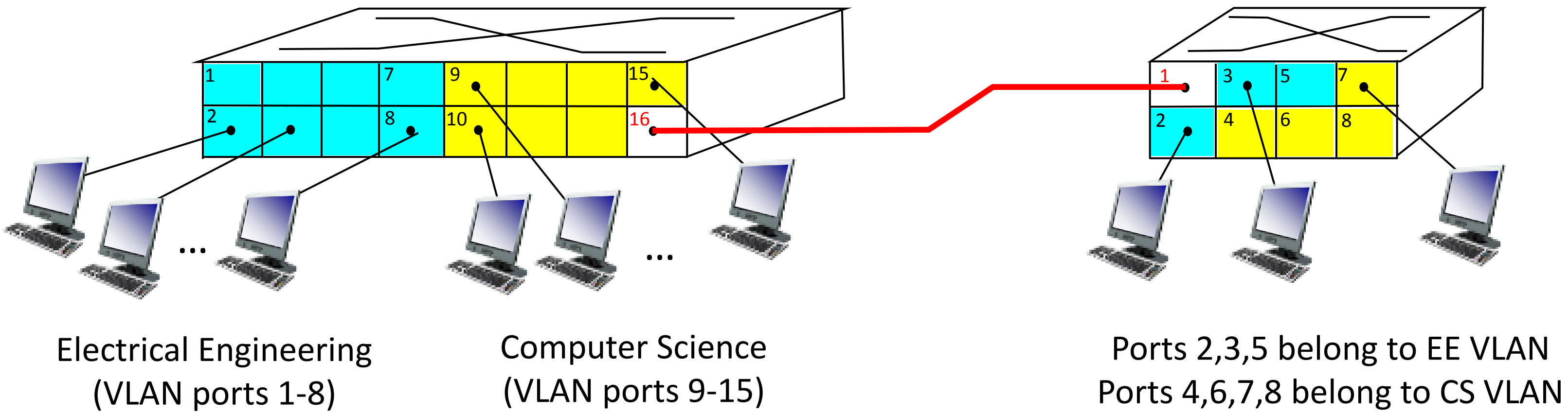
Electrical Engineering
(VLAN ports 1-8)

Computer Science
(VLAN ports 9-15)

- Dynamic membership:
  Ports can be dynamically assigned among VLANs

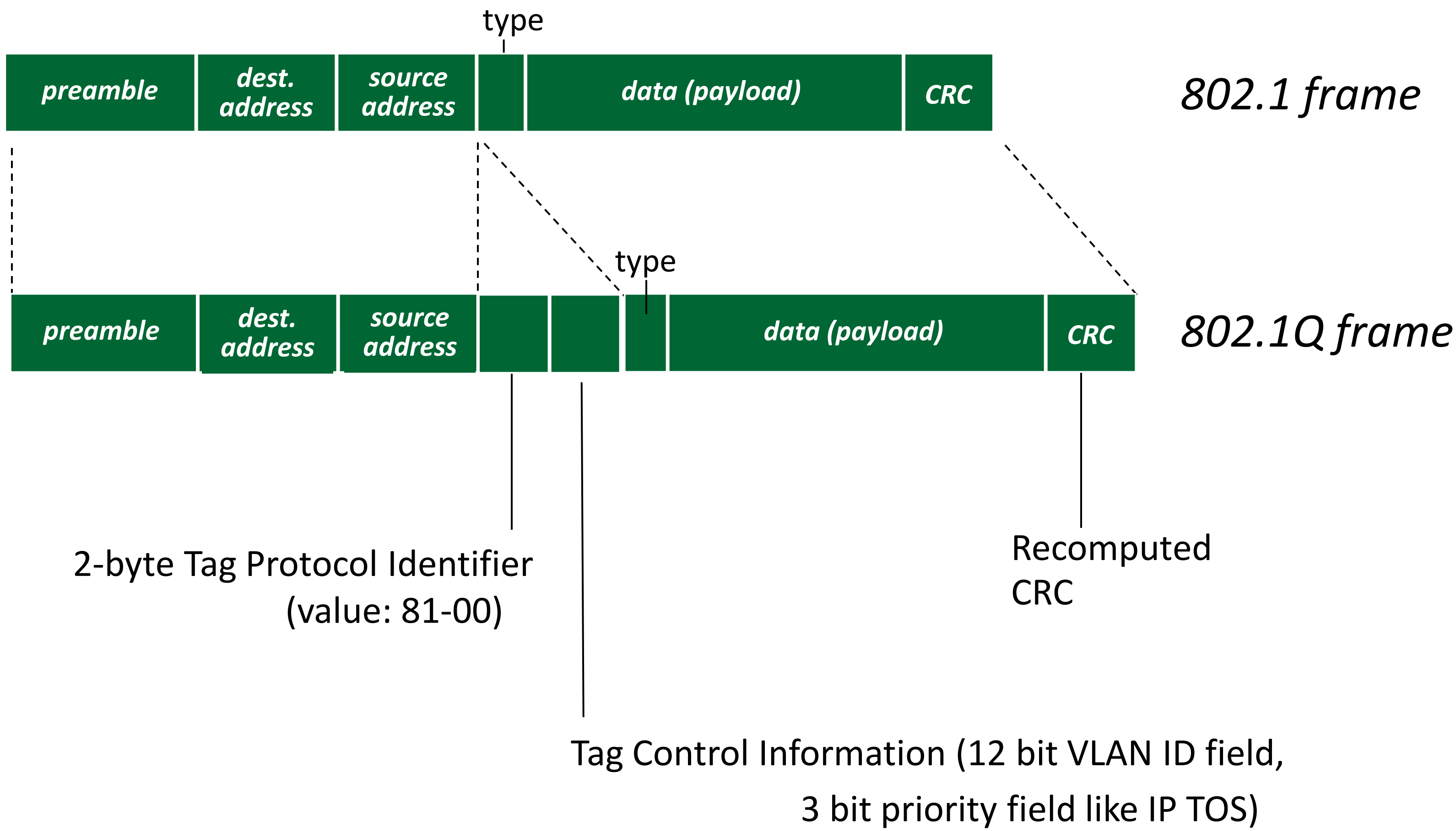
- Forwarding between VLANS: done via routing
  - Just as with separate switches
  - In practice vendors sell combined switches plus routers

VLANS spanning multiple switches



- *Trunk port:* 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

802.1Q VLAN frame format



Summary

- Address Reservation Protocol (ARP)
  - Mapping between link-layer addresses (MAC) and network-layer addresses (IP)
  - Cached table in operating system
  - Broadcast queries for IP destinations with unknown MAC
- Wired Ethernet
  - Long history and widely adoptedHubs vs. switches vs. routers
  - Order of magnitude bit rate increase every few years
  - Careful attention to backwards compatibility
- VLANs
  - Allows virtual segregation of hosts into isolated groups