VLAN, MPLS, and Data center

CE 352, Computer Networks

Salem Al-Agtash

Lecture 22

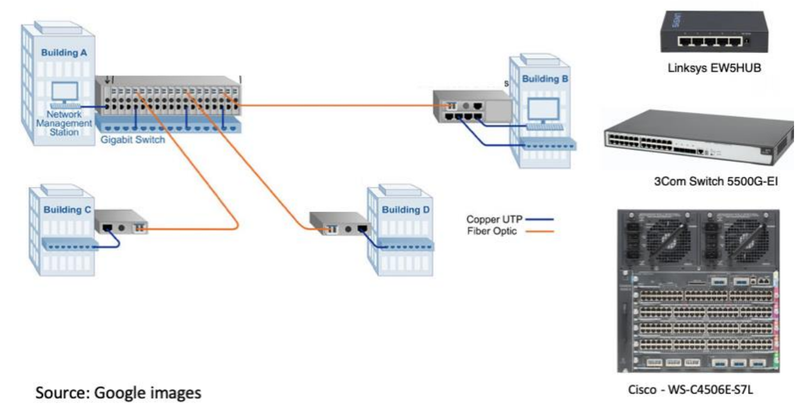
Recap (Ethernet switch)

link-layer device with *active* role

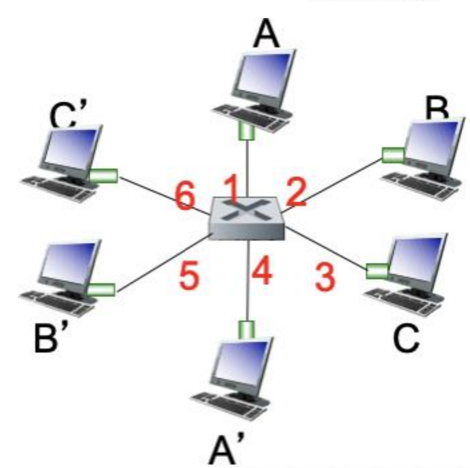
- store, forward Ethernet frames
- examine incoming frame's MAC address, selectively forward frame to one-or-more outgoing links

plug-and-play, self-learning

- switches do not need to be configured



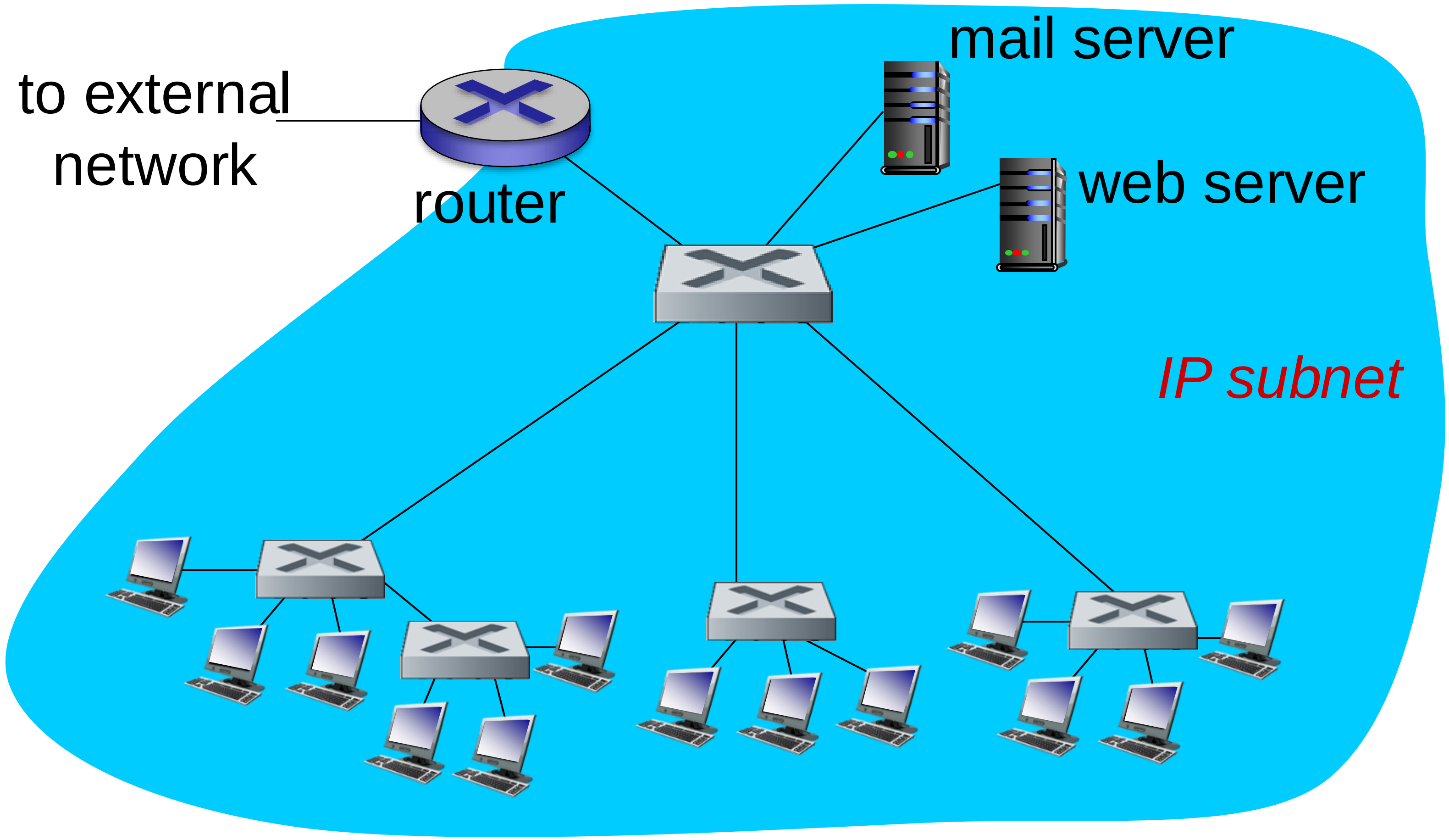
Source: Google images



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

switch table

Recap (institutional network)



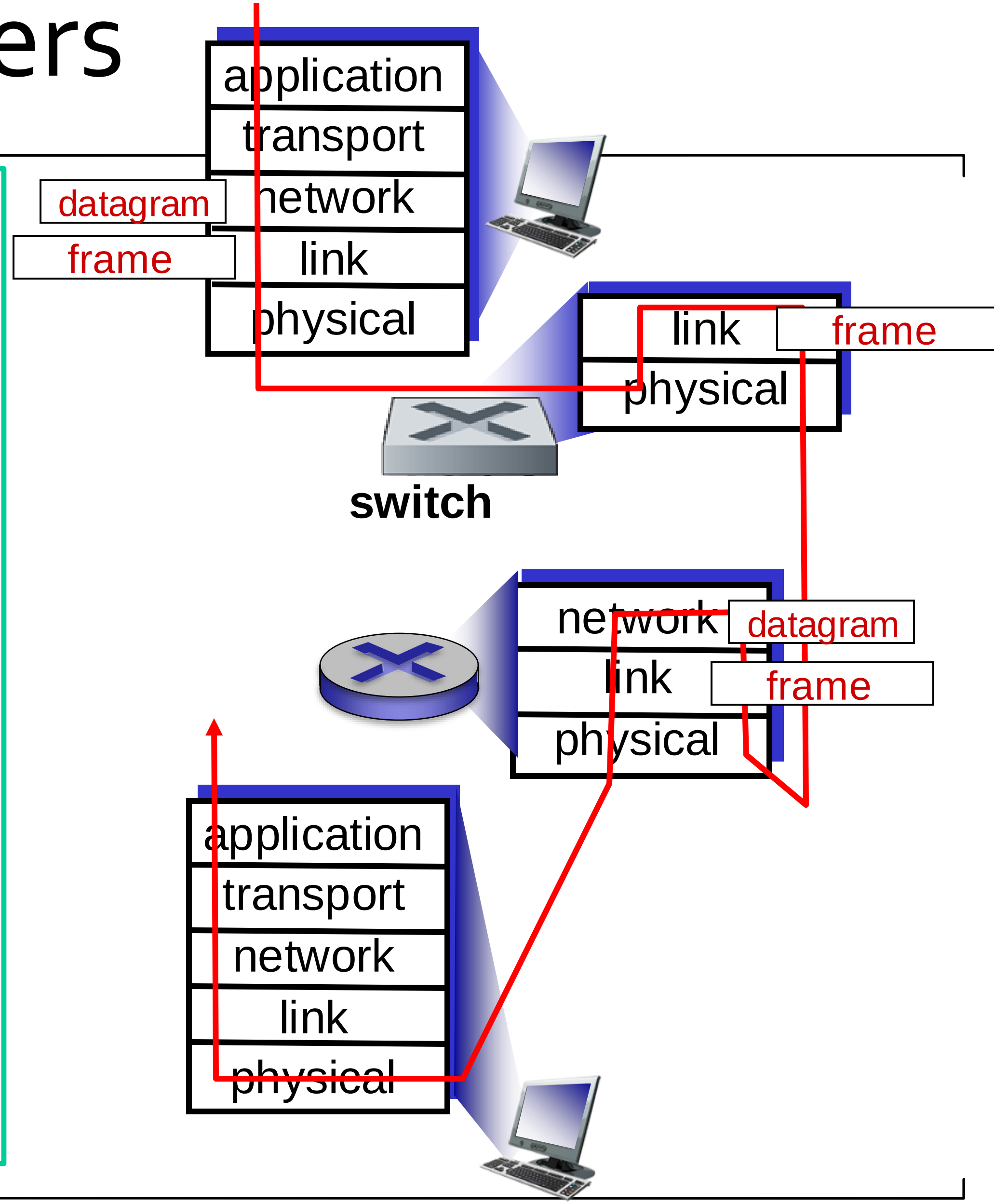
Switches vs. routers

both are store-and-forward:

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

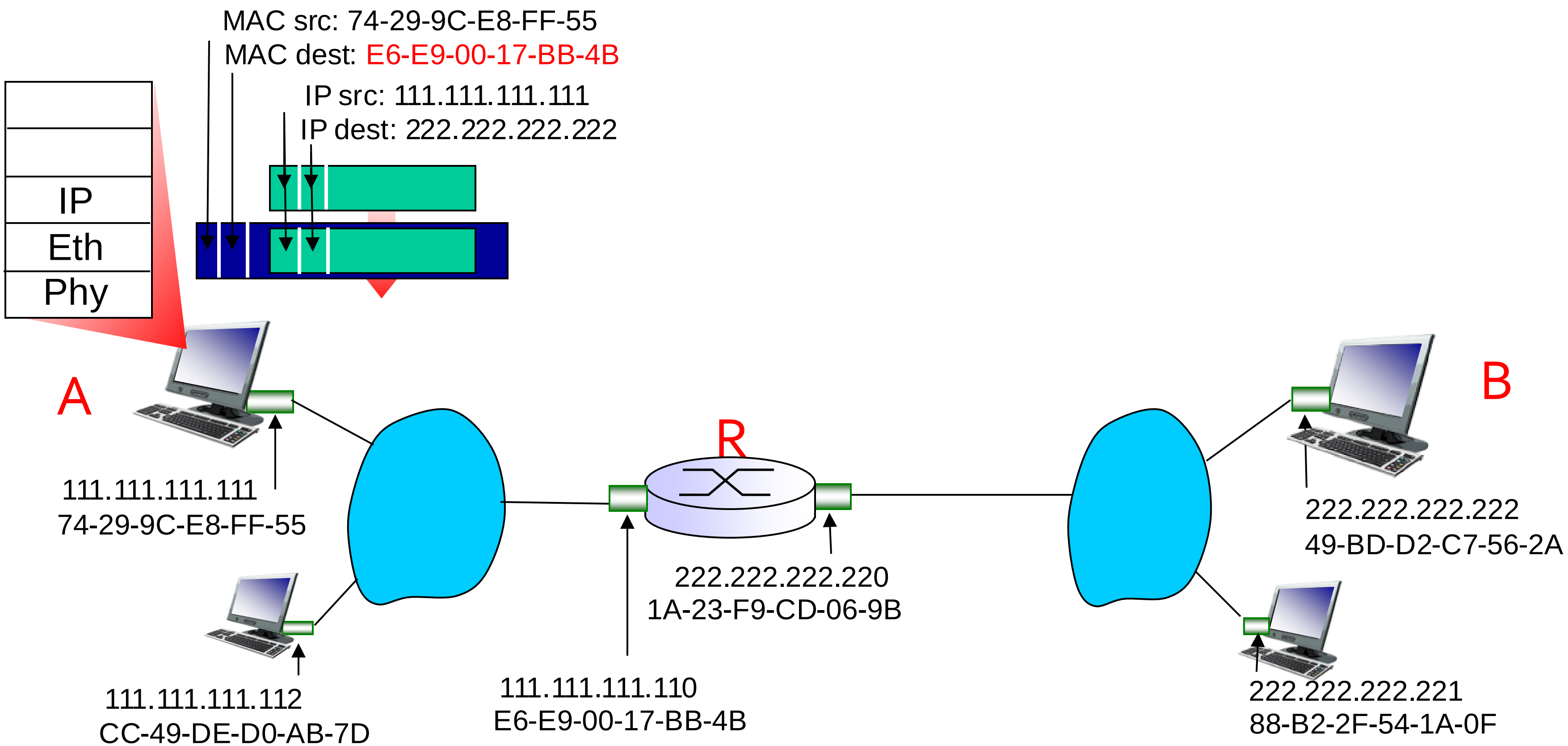
both have forwarding tables:

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



Recap (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 address, frame contains A-to-B IP datagram



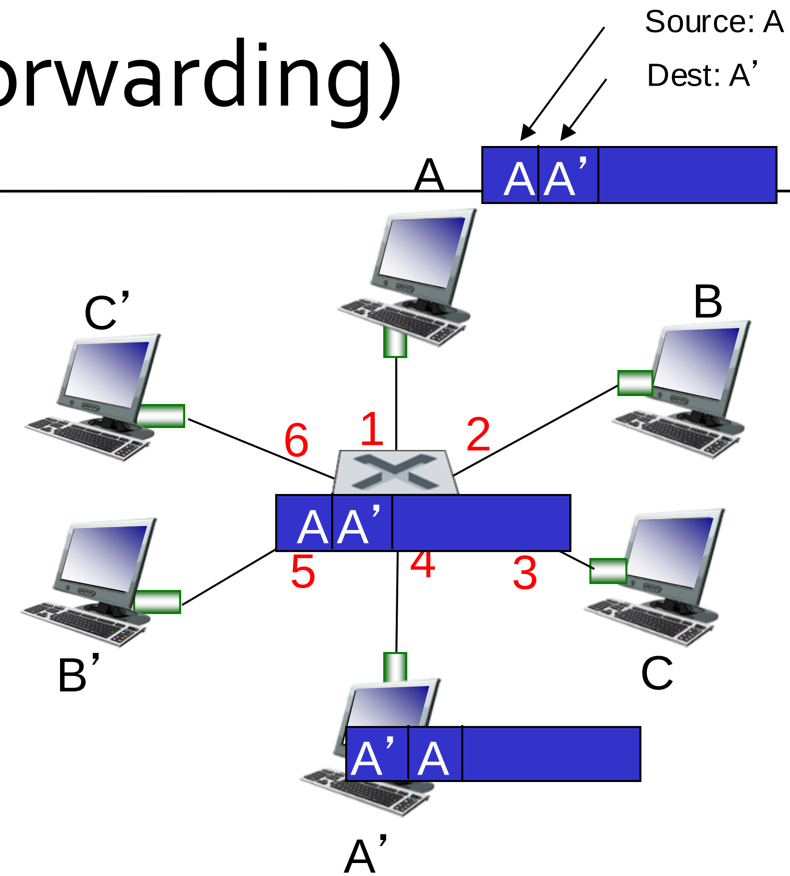
Recap (Self-learning, forwarding)

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*

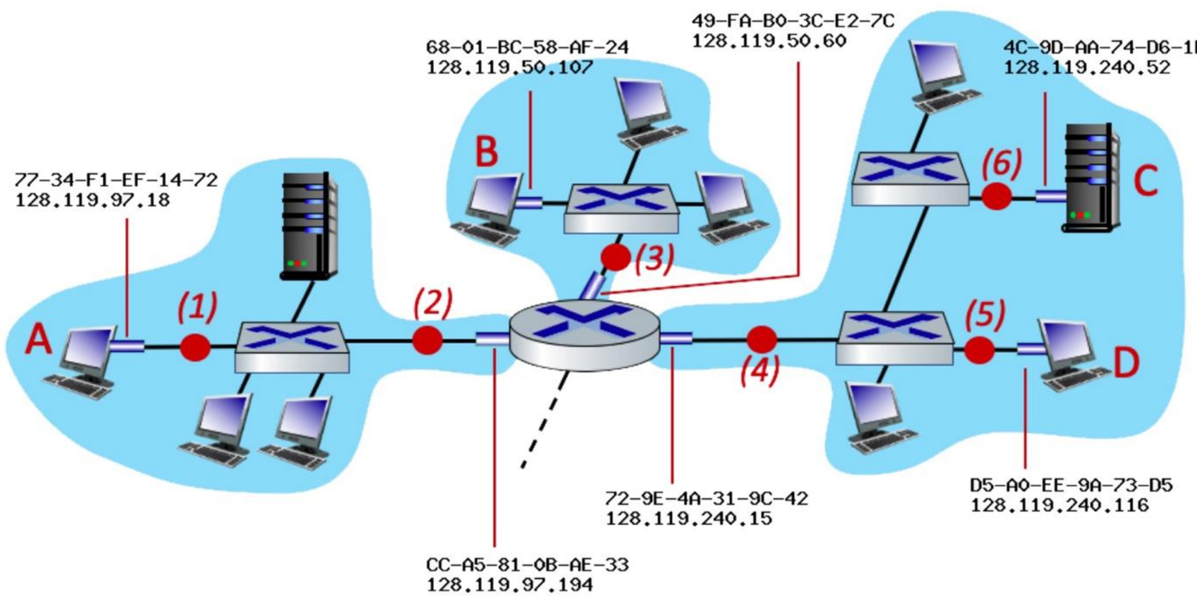
switch table (initially empty)



Question

- Consider the network shown below. The IP and MAC addresses are shown for hosts A, B, C and D, as well as for the router's interfaces. Consider an IP datagram being sent from node B to node D.
- What is the source and destination MAC addresses on the frame at point (4)?

- What is the source and destination IP addresses of the datagram at point (4)?



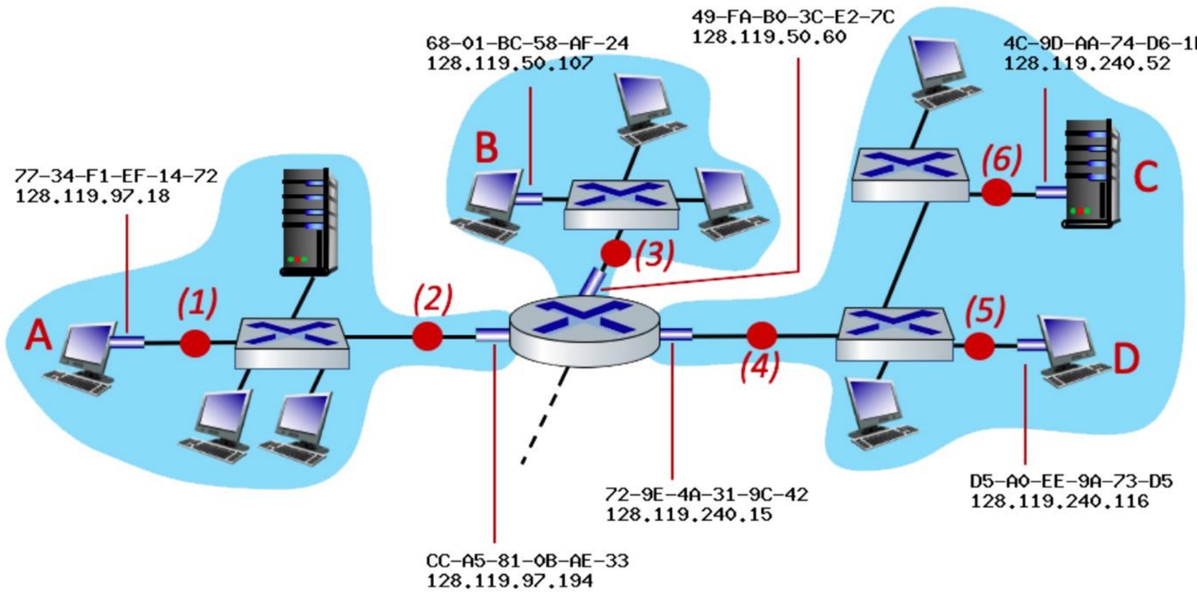
Question

- Consider the network shown below. The IP and MAC addresses are shown for hosts A, B, C and D, as well as for the router's interfaces. Consider an IP datagram being sent from node B to node D.
- What is the source and destination MAC addresses on the frame at point (4)?

Source (MAC): 72-9E-4A-31-9C-42 Destination (MAC): D5-A0-EE-9A-73-D5

- What is the source and destination IP addresses of the datagram at point (4)?

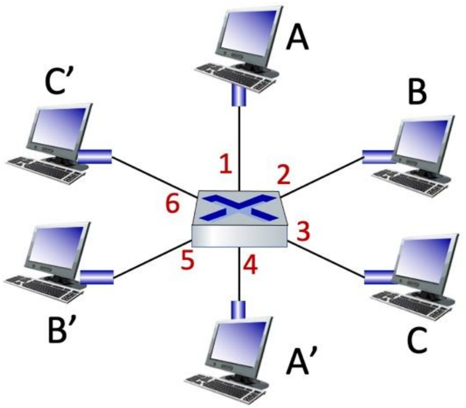
Source (IP): 128.119.50.107 Destination (IP): 128.119.240.116



Question

- Consider the network below with six nodes, star-connected into an Ethernet switch. Suppose that A sends a frame to A, A' replies to A, then B sends a message to B' and B' replies to B. Enter the values that are present in the switch's forwarding table after B'-to-B frame is sent and received. Assumed that the table is initially empty and that entries are added to the table sequentially.

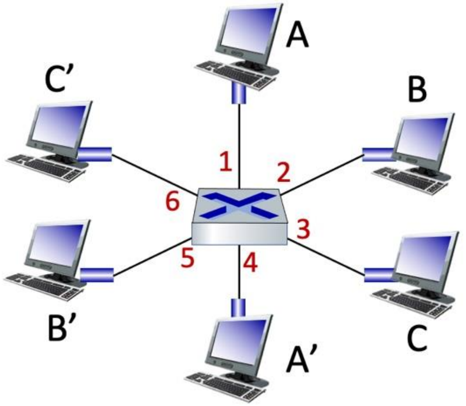
MAC Address	Interface	TTL
		60
		60
		60
		60



Question

- Consider the network below with six nodes, star-connected into an Ethernet switch. Suppose that A sends a frame to A, A' replies to A, then B sends a message to B' and B' replies to B. Enter the values that are present in the switch's forwarding table after B'-to-B frame is sent and received. Assumed that the table is initially empty and that entries are added to the table sequentially.

MAC Address	Interface	TTL
A	1	60
A'	4	60
B	2	60
B'	5	60



Question

- We've seen caching used in many different places in our study of networking. List the different forms of caching that occur with the protocol or network service.

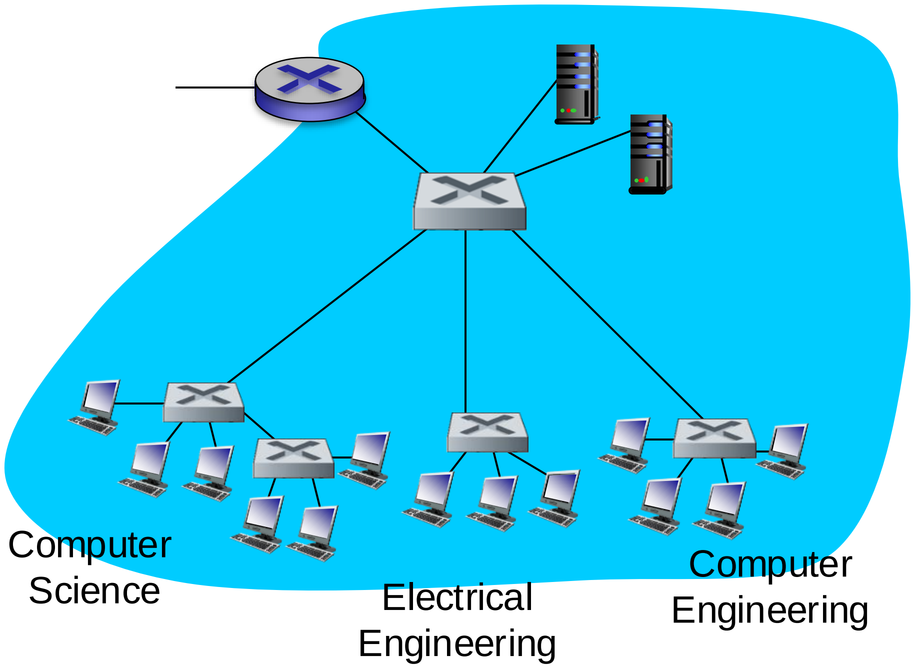
Protocol/ Network service	Form of Caching
Web pages for client	
Web pages for a set of clients	
Translation pairs between a domain name and IP address	
Translation pairs between an IP address and a MAC address	
Translation pairs between a MAC address and a layer-2 switch port	

Question

- We've seen caching used in many different places in our study of networking. List the different forms of caching that occur with the protocol or network service.

Protocol/ Network service	Form of Caching
Web pages for client	Web browser caching
Web pages for a set of clients	Web proxy
Translation pairs between a domain name and IP address	Local DNS caching
Translation pairs between an IP address and a MAC address	ARP cache
Translation pairs between a MAC address and a layer-2 switch port	Learning switch catching

VLANs: motivation



consider:

CS user moves office to EE, but wants connect to CS switch?

single broadcast domain:

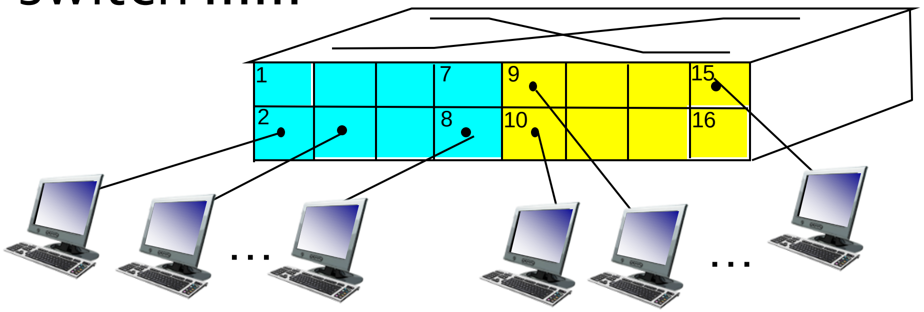
- all layer-2 broadcast traffic
- security/privacy, efficiency issues
- Use ports to configure VLANs
- So physical location does not matter

https://www.youtube.com/embed/IRPRX2fVJYY?feature=oembed

VLANs: motivation

VLANs

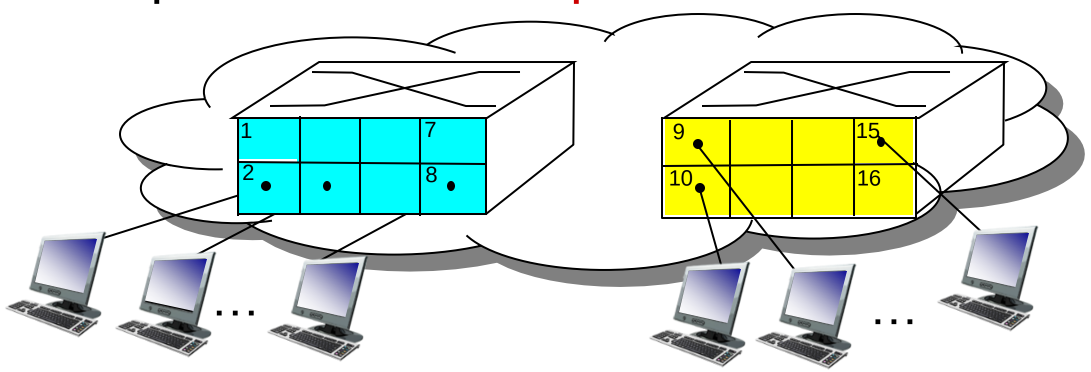
port-based VLAN: switch ports grouped (by switch management software) so that *single* physical switch ......



Virtual Local Area Network

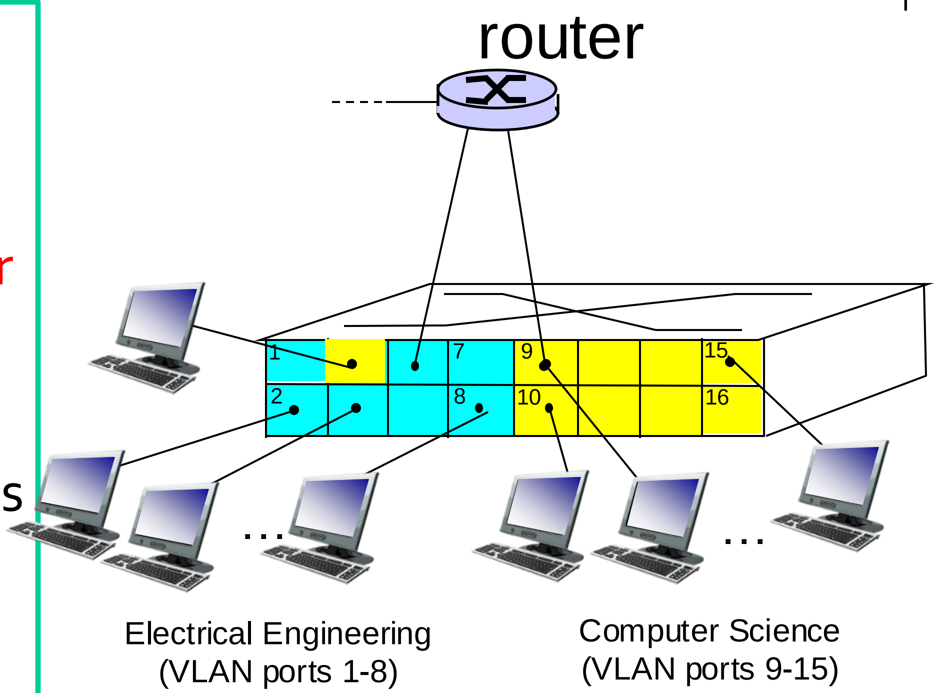
switch(es) supporting VLAN capabilities can be configured to define multiple ***virtual*** LANS over single physical LAN infrastructure.

... operates as multiple virtual switches

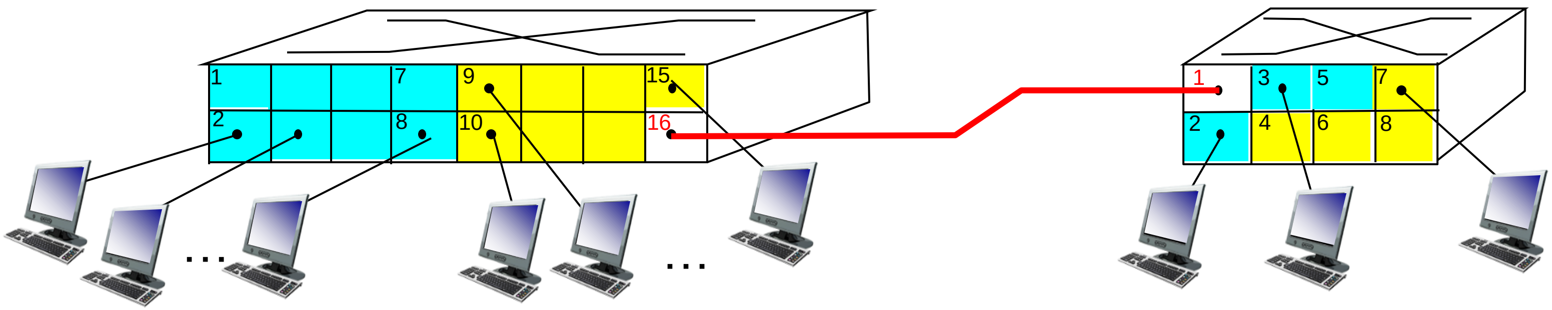


Port-based VLAN

- *traffic isolation:* frames to/from ports 1-8 can *only* reach ports 1-8
- VLAN cab also be defined based on MAC addresses of endpoints, rather than switch port
- dynamic membership: ports can be dynamically assigned among VLANs
- forwarding between VLANS: done via routing/ switching (just as with separate switches)
- Scale to NVLANs by interconnecting multiple switches



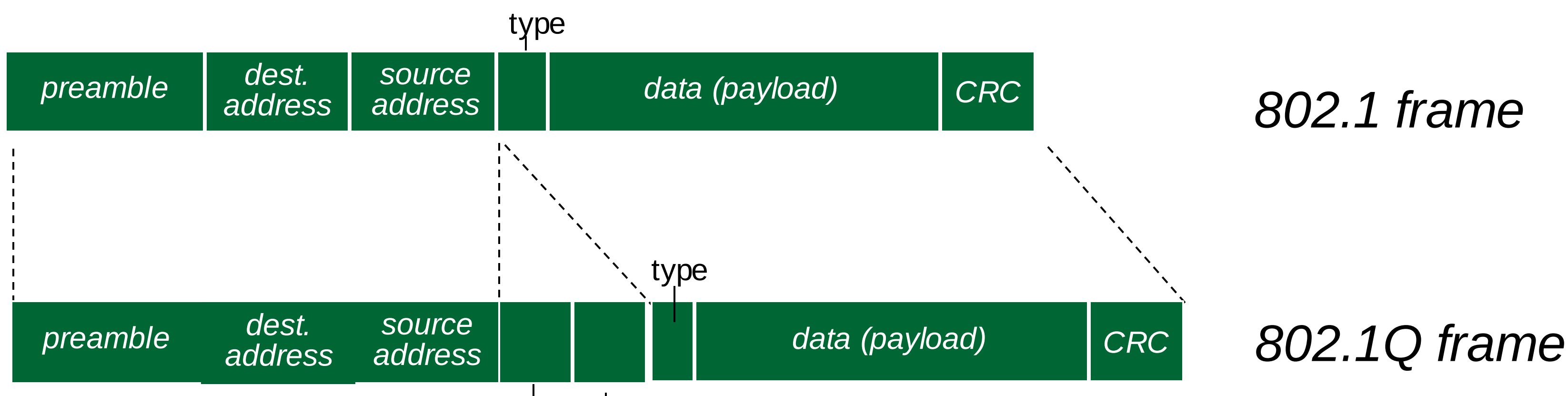
VLANS spanning multiple switches



trunk port: carries frames between VLANS defined over multiple physical switches

- frames forwarded within VLAN between switches can't be 802.1 frames (must carry VLAN ID info)
- 802.1q protocol adds/removes additional header fields for frames forwarded between trunk ports

802.1Q VLAN frame format



2-byte Tag Protocol Identifier (value: 81-00) to identify the frame as an IEEE 802.1Q-tagged frame

Tag Control Information (12 bit VLAN ID field (→ 2^{12} (4096) VLANs),

3 bit priority field like IP TOS)

Multiprotocol label switching (MPLS)

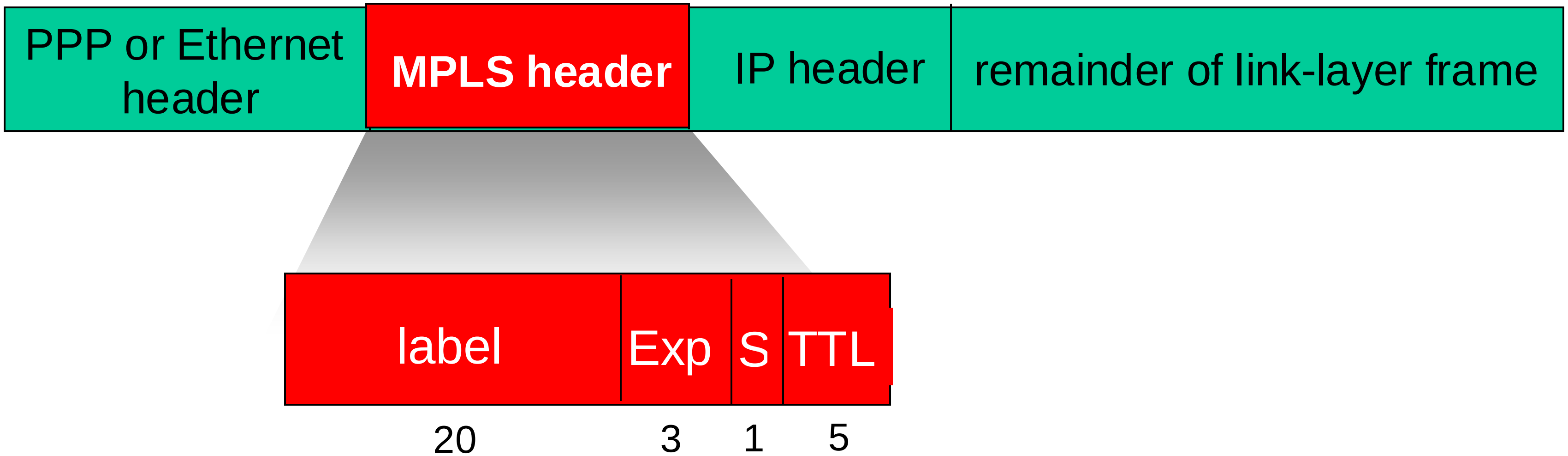
- Link virtualization: A network as a link layer
  - MPLS is a link layer technology that serves to interconnect IP devices
  - MPLS: a packet-switched virtual-circuit network with its own packet format and forwarding behavior
- Initial goal: high-speed IP forwarding using fixed length label (instead of IP address)
  - fast lookup using fixed length identifier (rather than shortest prefix matching)
  - borrowing ideas from Virtual Circuit (VC) approach (widely used in setting up VPN)
  - but IP datagram still keeps IP address!

https://www.youtube.com/watch?v=E5Ud1m9h0yc

Multiprotocol label switching (MPLS)

Multiprotocol label switching (MPLS)

- Runs on MPLS-capable routers
- Blending VC technology into a routed datagram network
- Link-layer frame transmitted between MPLS-capable routers
  - 20 bits Label, 3 bits for experimental use, 1 bit to indicated end of series of stacked MPLS headers, and 5 bits for “time-to-live”



MPLS capable routers

a.k.a. label-switched router

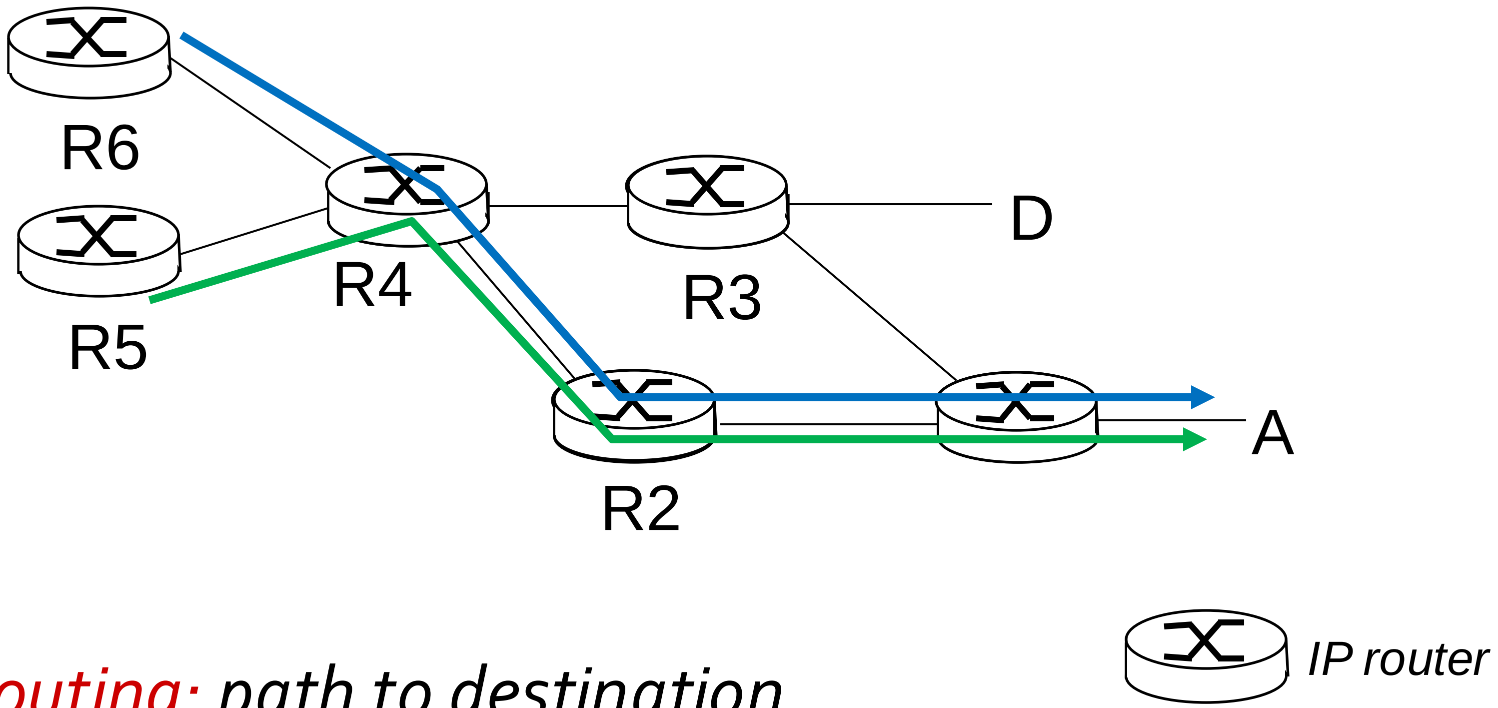
forward packets to outgoing interface based only on label value (*don't inspect IP address*)

- MPLS forwarding table distinct from IP forwarding tables

flexibility: MPLS forwarding decisions can *differ* from those of IP

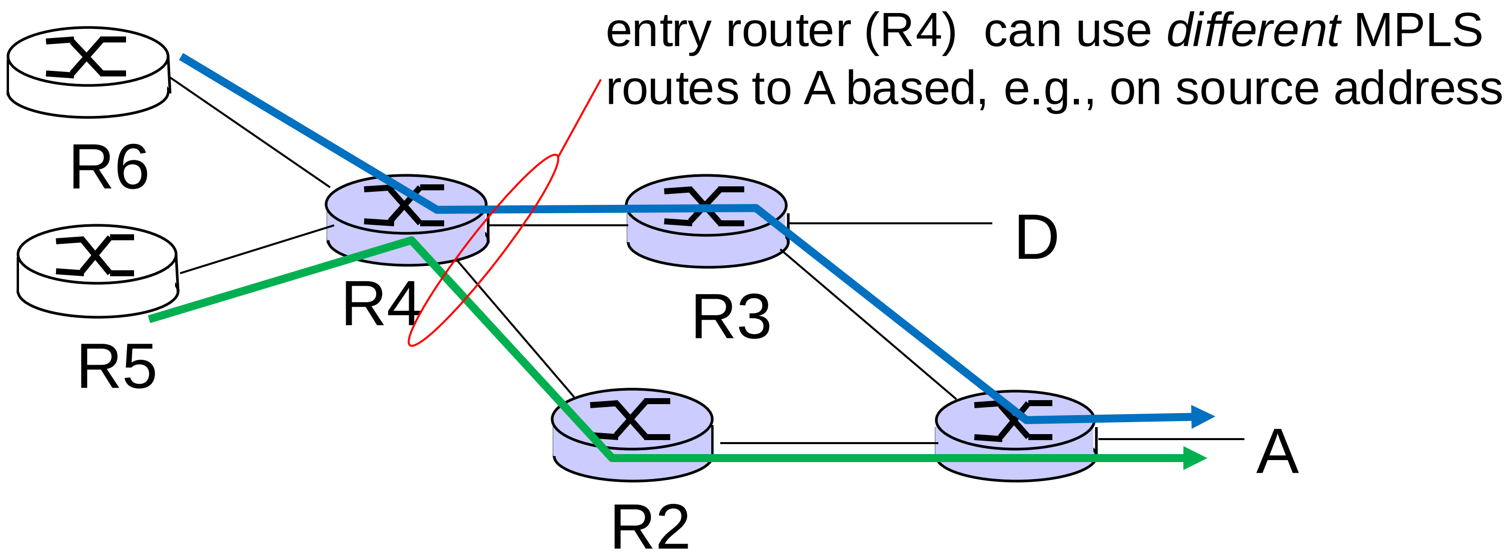
- use destination *and* source addresses to route flows to same destination differently (traffic engineering)
- re-route flows quickly if link fails: pre-computed backup paths (useful for VoIP)

MPLS versus IP paths

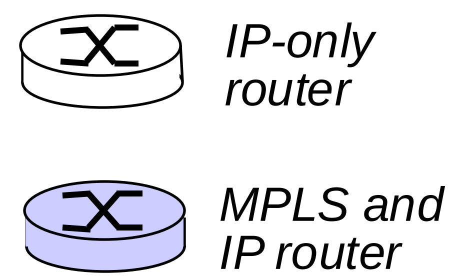


- *IP routing: path to destination determined by destination address alone*

MPLS versus IP paths

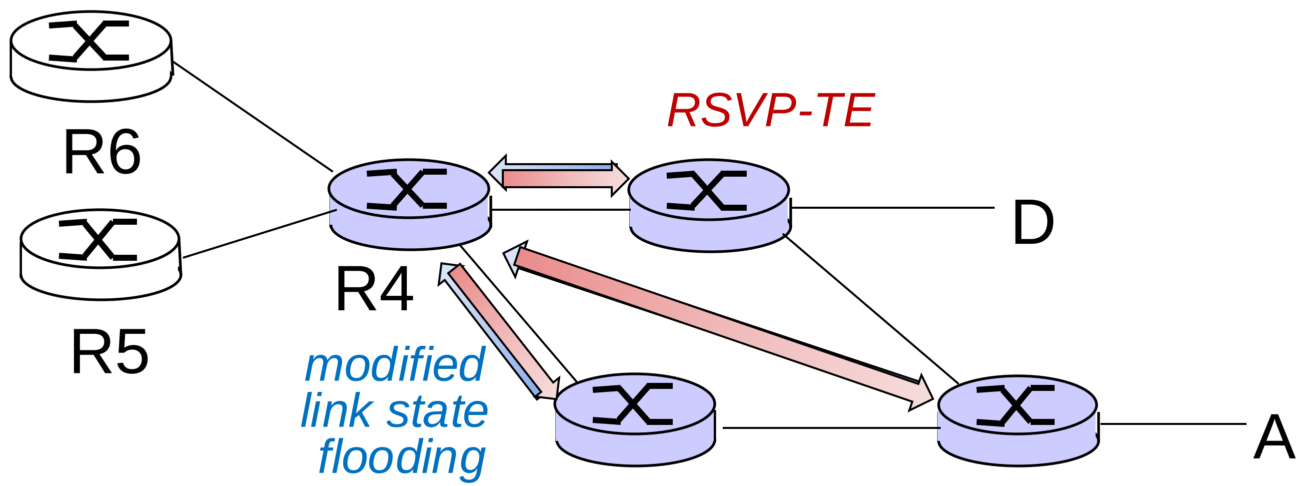


- *IP routing:* *path to destination determined by destination address alone*
- *MPLS routing:* path to destination can be based on source *and* destination address
  - *fast reroute:* precompute backup routes in case of link failure



MPLS signaling

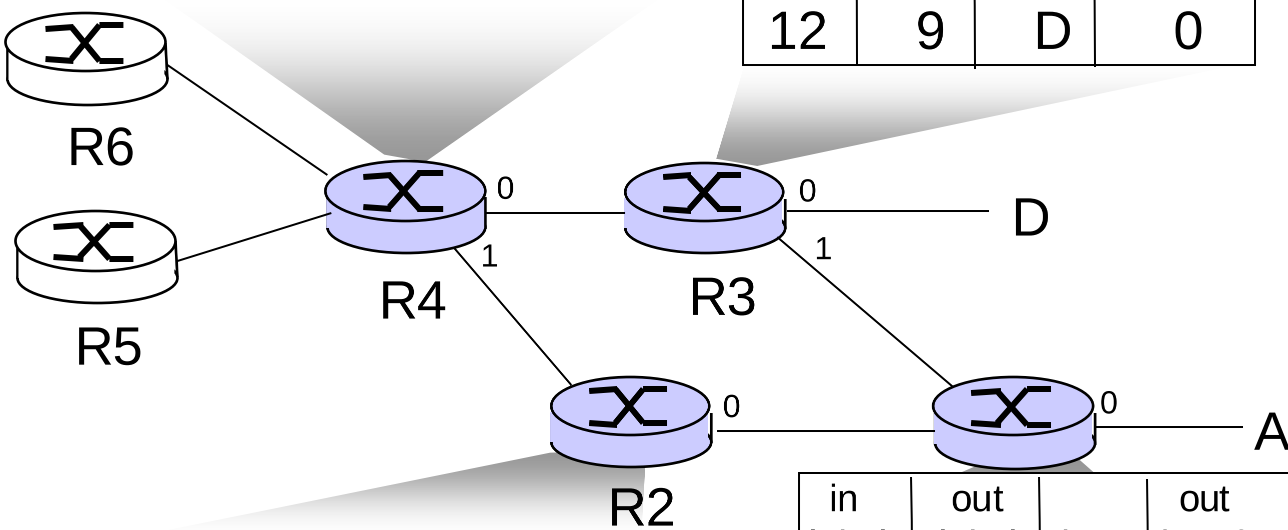
- modify OSPF link-state flooding protocols to carry info used by MPLS routing,
  - e.g., link bandwidth, amount of "reserved" link bandwidth
- entry MPLS router uses RSVP-TE signaling protocol to set up MPLS forwarding at downstream routers



MPLS forwarding tables

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

Data center networks

10's to 100's of thousands of hosts, often closely coupled, in close proximity:

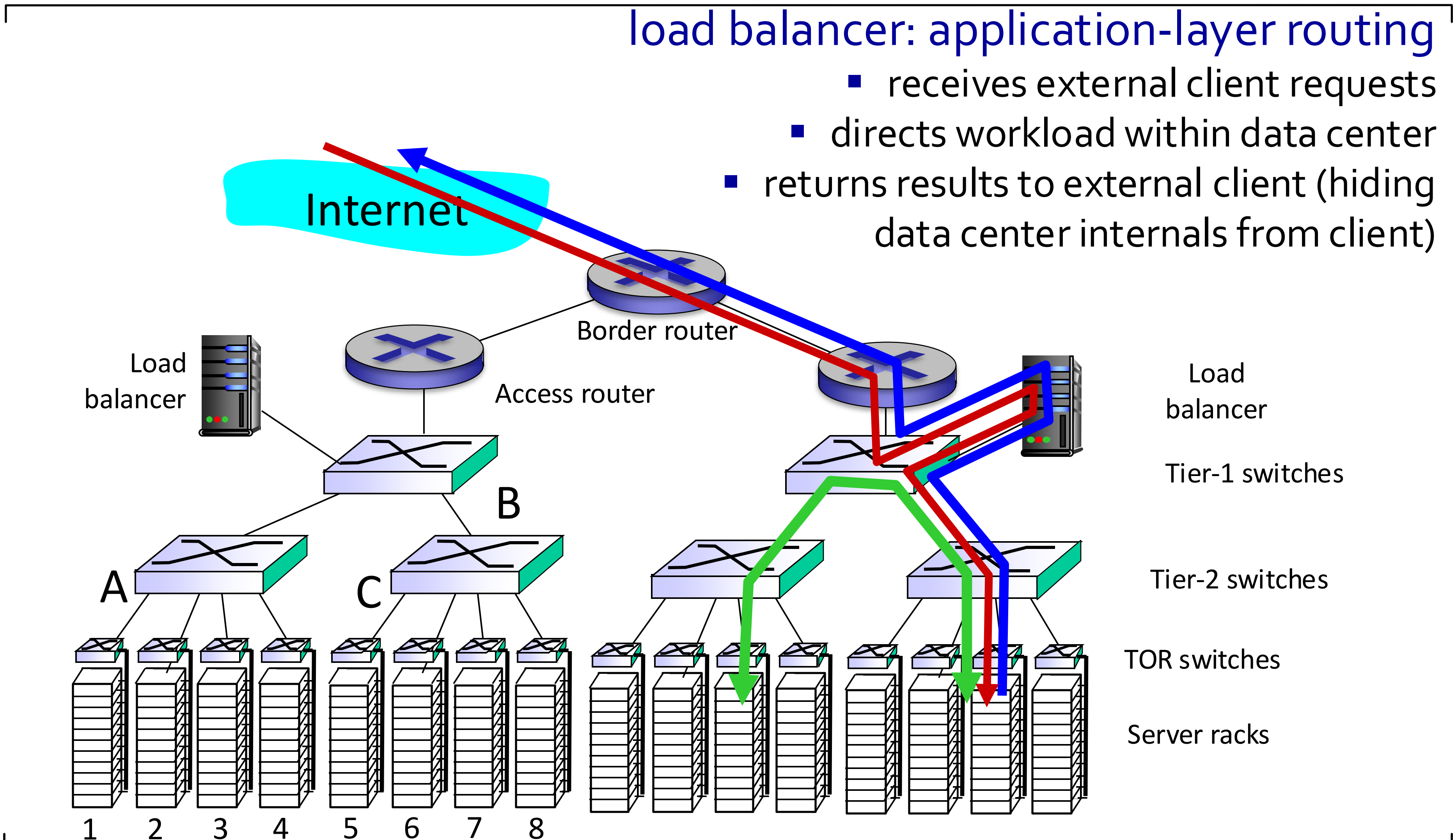
- e-business (e.g. Amazon)
- content-servers (e.g., YouTube, Akamai, Apple, Microsoft)
- search engines, data mining (e.g., Google)

challenges:

- multiple applications, each serving massive numbers of clients
- managing/balancing load, avoiding processing, networking, data bottlenecks

Inside a 40-ft Microsoft container, Chicago data center

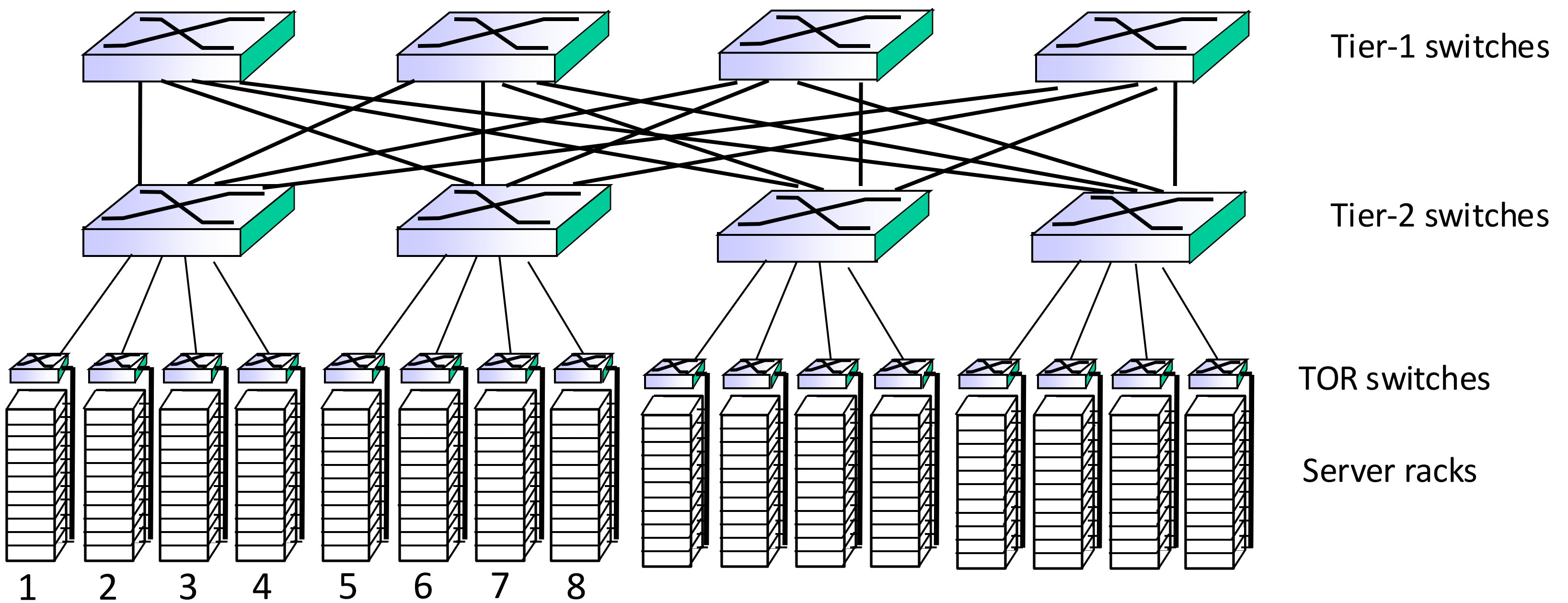
Data center networks



Data center networks

Rich interconnection among switches, racks:

- increased throughput between racks (multiple routing paths possible)
- increased reliability via redundancy



Synthesis: a day in the life of a web request

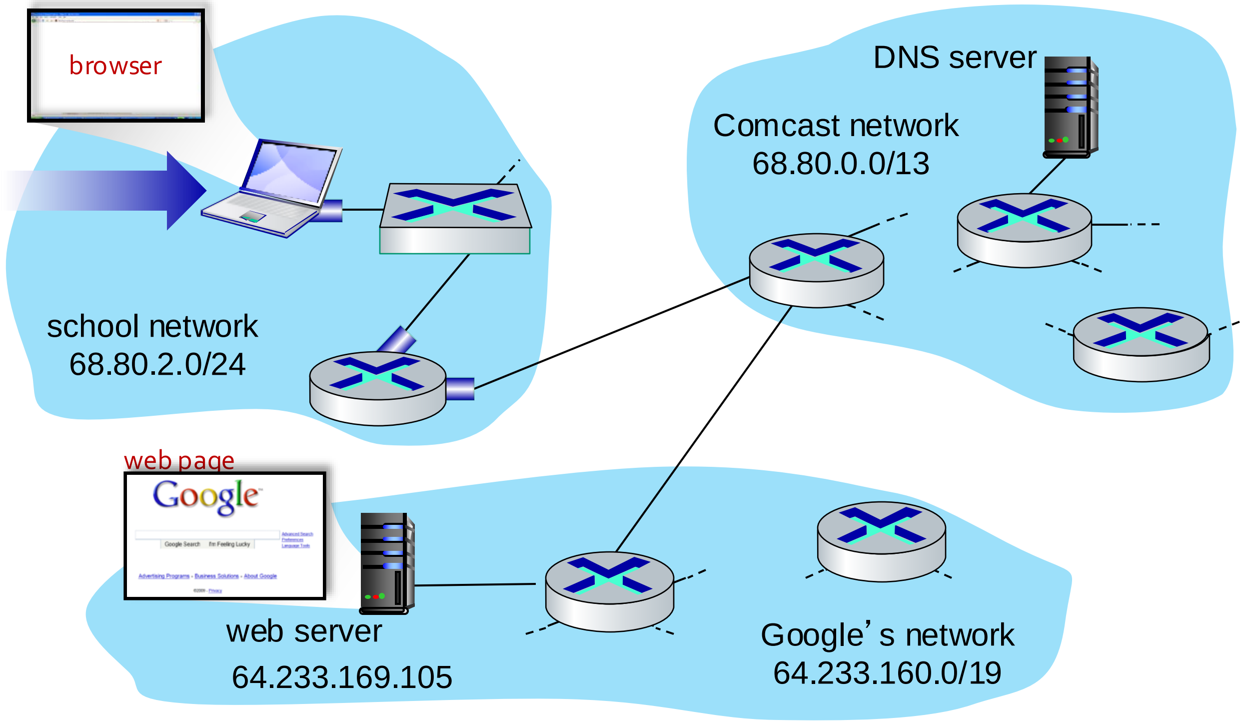
journey down protocol stack complete!

- application, transport, network, link

putting-it-all-together: synthesis!

- *goal:* identify, review, understand protocols (at all layers) involved in seemingly simple scenario: requesting www page
- *scenario:* student attaches laptop to campus network, requests/receives www.google.com

A day in the life: scenario

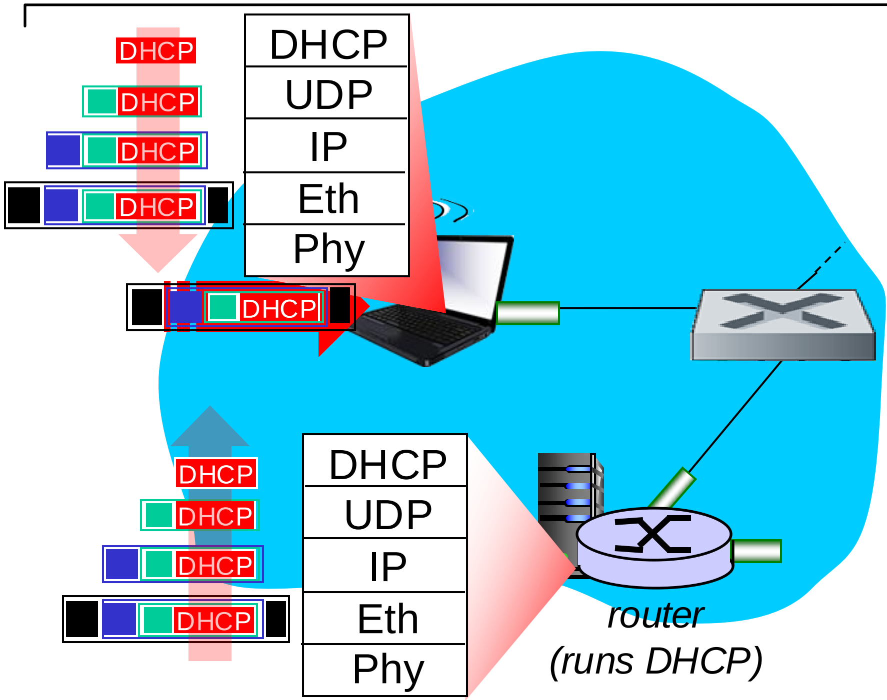


scenario:

- arriving mobile client attaches to network …
- requests web page: www.google.com

Sounds simple!

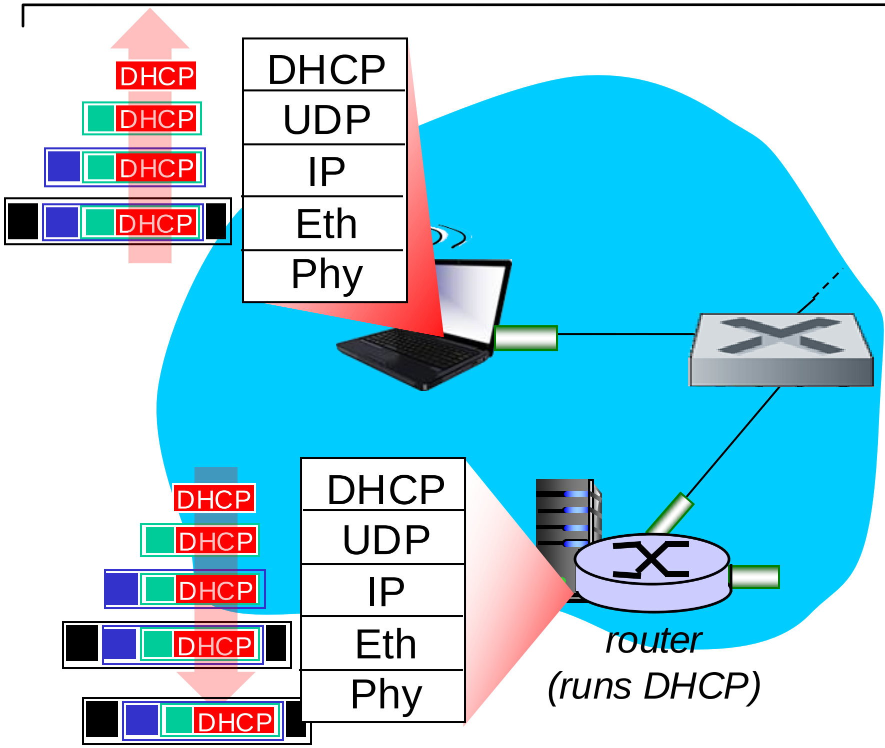
A day in the life… connecting to the Internet



connecting laptop needs to get its own IP address, addr of first-hop router, addr of DNS server: use *DHCP*

- DHCP request encapsulated in UDP, encapsulated in IP, encapsulated in 802.3 Ethernet
- Ethernet frame broadcast (dest: FFFFFFFFFFFF) on LAN, received at router running DHCP server
- Ethernet demuxed to IP demuxed, UDP demuxed to DHCP

A day in the life... connecting to the Internet

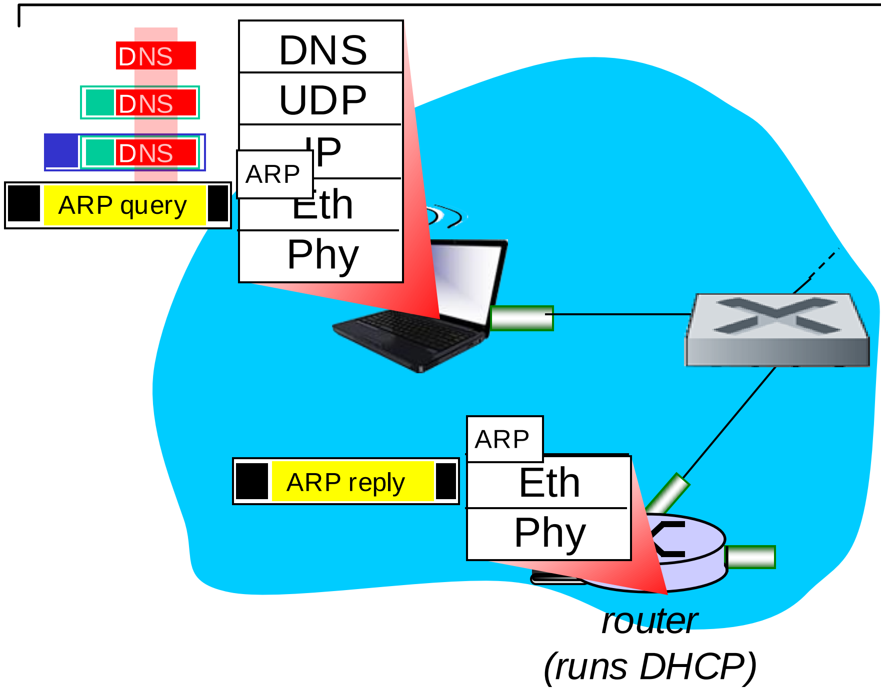


DHCP server formulates *DHCP ACK* containing client's IP address, IP address of first-hop router for client, name & IP address of DNS server

- encapsulation at DHCP server, frame forwarded (switch learning) through LAN, demultiplexing at client
- DHCP client receives DHCP ACK reply

Client now has IP address, knows name & addr of DNS server, IP address of its first-hop router

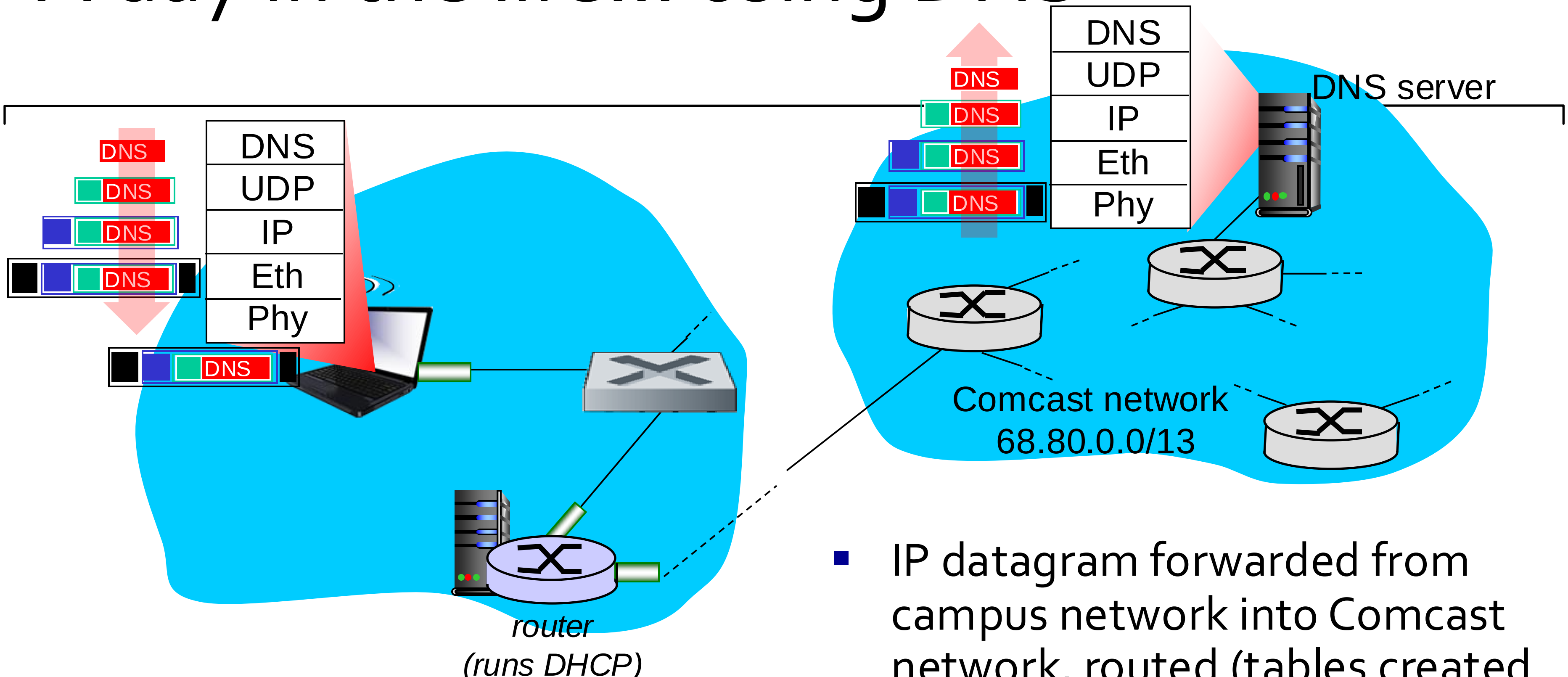
A day in the life… ARP (before DNS, before HTTP)



before sending *HTTP* request, need IP address of www.google.com: *DNS*

- DNS query created, encapsulated in UDP, encapsulated in IP, encapsulated in Eth. To send frame to router, need MAC address of router interface: ARP
- ARP query broadcast, received by router, which replies with ARP reply giving MAC address of router interface
- client now knows MAC address of first hop router, so can now send frame containing DNS query

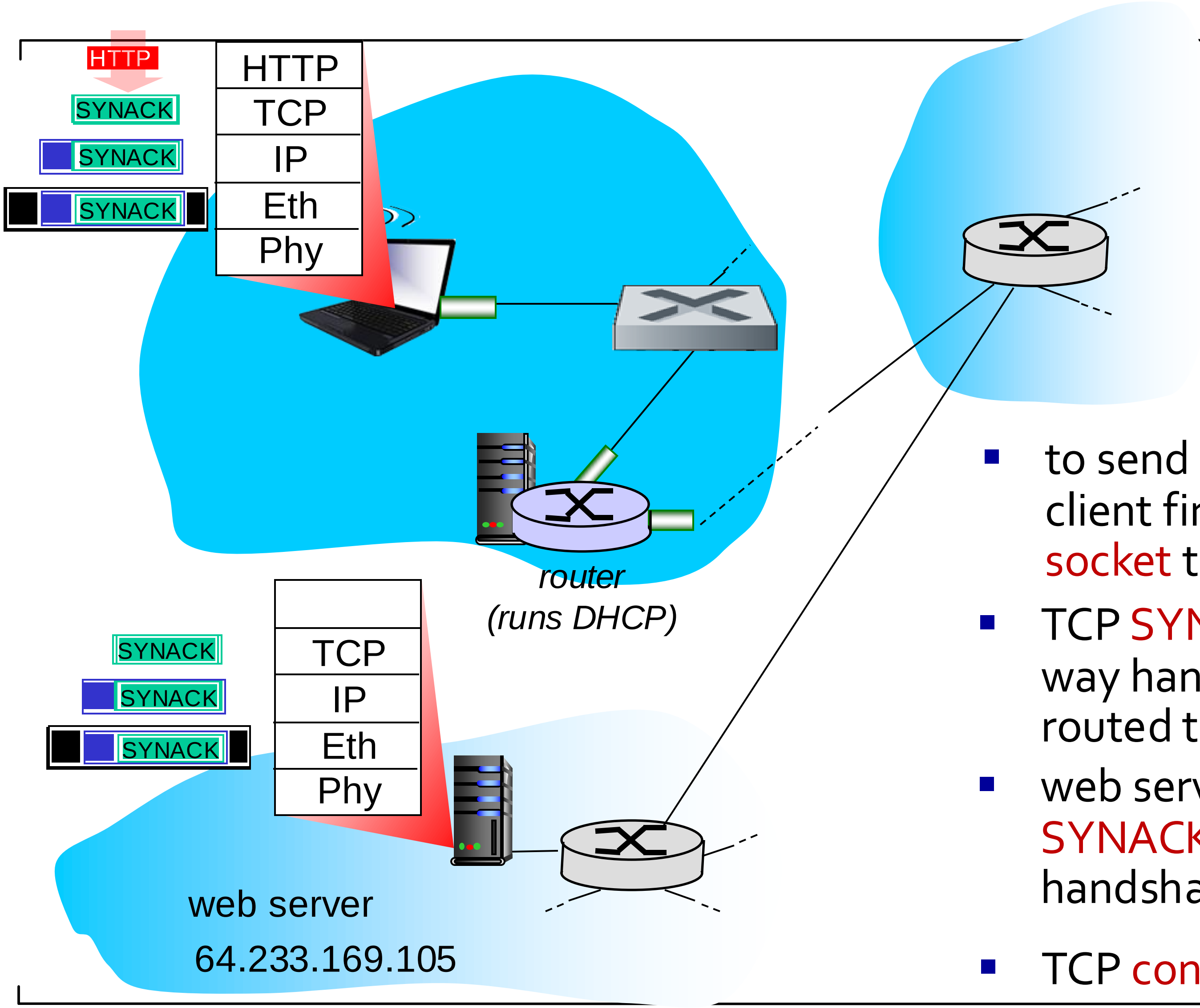
A day in the life… using DNS



- IP datagram containing DNS query forwarded via LAN switch from client to 1st hop router

- IP datagram forwarded from campus network into Comcast network, routed (tables created by RIP, OSPF, IS-IS and/or BGP routing protocols) to DNS server
- demuxed to DNS server
- DNS server replies to client with IP address of www.google.com

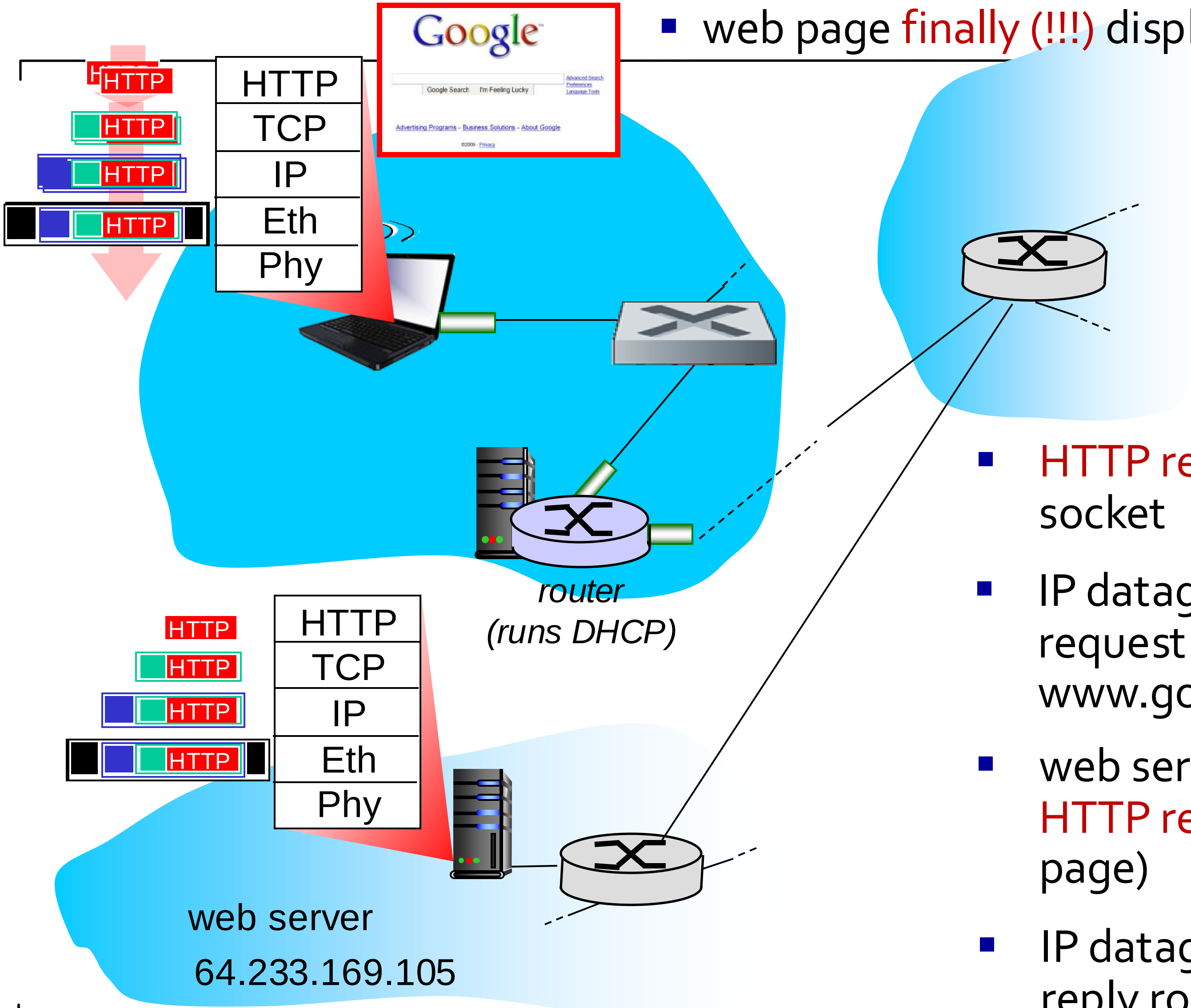
A day in the life…TCP connection carrying HTTP



- to send HTTP request, client first opens TCP socket to web server
- TCP SYN segment (step 1 in 3-way handshake) inter-domain routed to web server
- web server responds with TCP SYNACK (step 2 in 3-way handshake)
- TCP connection established!

A day in the life… HTTP request/reply

- web page finally (!!!) displayed



- HTTP request sent into TCP socket
- IP datagram containing HTTP request routed to www.google.com
- web server responds with HTTP reply (containing web page)
- IP datagram containing HTTP reply routed back to client

Summary

Today:

- VLANs
- Virtualized networks as a link layer: MPLS
- Data centers

Canvas discussion:

- Reflection
- Exit ticket

Next time:

- Read 7.2, 7.3, 7.4 of KR (Wireless Links, WiFi, Cellular Internet)
  follow on Canvas! material and announcements

Any questions?