CSE 461: Final Review

Autumn 2022

Administrivia

Latest due:

- Dec 12th
  - **Project 3**
- Dec 8th
  - **Everything else**

- Thu, Dec 15th 8:30 – 10:20 AM
  - **Final Exam**

- Please fill out the **course evaluation form**

Final Review Section

- Today: A brief review of lecture materials
  - Concepts, Protocols, Algorithms, …
- What **YOU** should do after this section and before the exam:
  - Go through the lecture slides
  - Think about the **problems** that each protocol/algorithm tries to solve
    - Pros and cons of current approaches?
    - Any other possible solutions?
    - What has not been solved yet?

461

- ~~Physical Layer~~
- ~~Link Layer~~
- Network Layer
- Transport Layer
- Application Layer
- (QoS)

Network Layer

- Network Service Models
- IP Address and Forwarding
- DHCP, ARP, ICMP
- NAT, IPv6
- Routing Algorithms
- BGP

Link Layer

- Multiplexing: Time Division (TDM) / Frequency Division (FDM)
- Multiple Access: ALOHA, CSMA/CD, BEB
- Wireless: Hidden/Exposed Terminal Problem, RTS/CTS
- Switching: Backward Learning
  - Forwarding Loop? -> Spanning Tree Algorithm!
- Software Defined Networking (SDN)
  - Rising of Datacenter Networks
  - Separation of Control Plane and Data Plane

Motivation

- What does the network layer do?
    - Connect different networks (send packets over multiple networks)

- Why do we need the network layer?
    - Switches don’t scale to large networks
    - Switches don’t work across more than one link layer technology
    - Switches don’t give much traffic control

Network Service Models

Datagram Model

- Connectionless service
- Packets contain destination address
- Routers looks up address in its forwarding table to determine next hop
- Example: IP

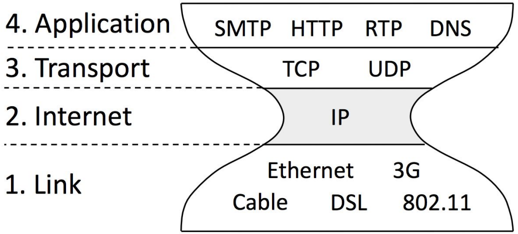
Virtual Circuits

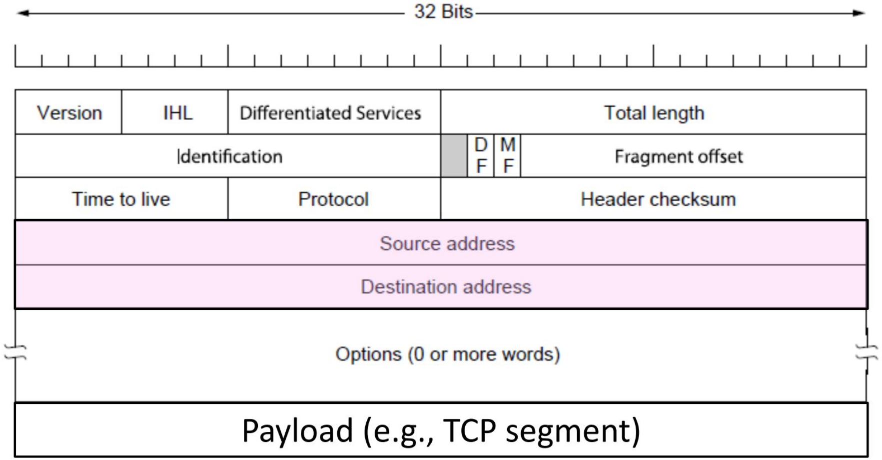
- Connection-oriented service
- Connection establishment → data transfer → connection teardown
- Packets contain label for circuit
- Router looks up circuit in forwarding table to determine next hop
- Example: MPLS

Both of them use **Store-and-Forward packet switching**

Internetworking - IP

- How do we connect different networks together?
- **IP - Internet Protocol**
- Lowest Common Denominator
  - Asks little of lower-layer networks
  - Gives little as a higher layer service



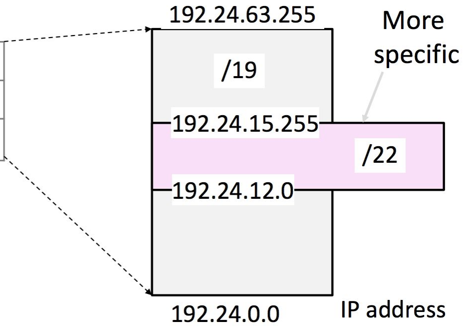


IP Addresses Prefix and Forwarding

- IP prefix a.b.c.d/**L**
  - Represents addresses that have the same first L bits
  - e.g. 128.13.0.0/16 -> all 65536 addresses between 128.13.0.0 to 128.13.255.255
  - e.g. 18.31.0.0/32 -> 18.31.0.0 (only one address)
- **Longest Matching Prefix**
  - find the longest prefix that contains the destination address, i.e., the most specific entry

Prefix	Next Hop
192.24.0.0/19	D
192.24.12.0/22	B

192.24.6.0 → D
192.24.14.32 → B
192.24.54.0 → D



IP Addresses Prefix and Forwarding

Which of the following prefix splits the address space 128.64.24.0/25 in half.

a. 128.64.24.128/26
b. 128.64.24.64/26
c. 128.64.24.128/27
d. 128.64.24.0/26

IP Addresses Prefix and Forwarding

Which of the following prefix splits the address space 128.64.24.0/25 in half.

128.64.24.0/25: 128.64.24.0000 0000 - 128.64.24.0111 1111

Half: 128.64.24.0100 0000

a. 128.64.24.128/26

1000 0000

a. 128.64.24.64/26

0100 0000

a. 128.64.24.128/27

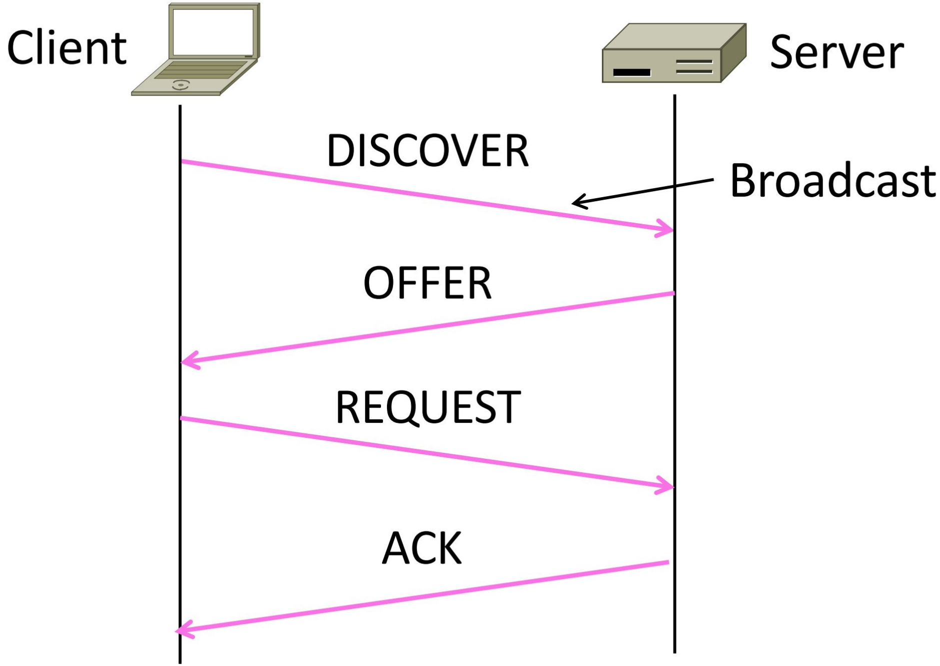
1000 0000

a. 128.64.24.0/26

0000 0000

DHCP - Dynamic Host Configuration Protocol

- Bootstrapping problem
- Leases IP address to nodes
- UDP
- Also setup other parameters:
  - DNS server
  - IP address of local router
  - Network prefix

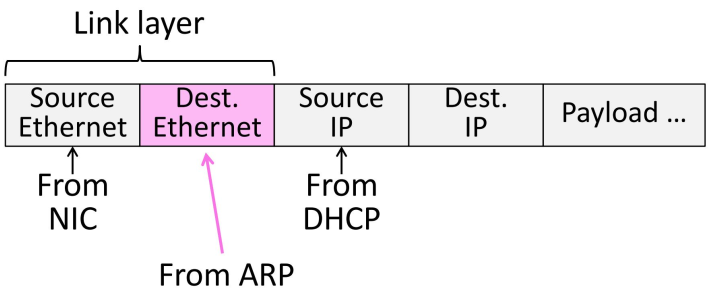


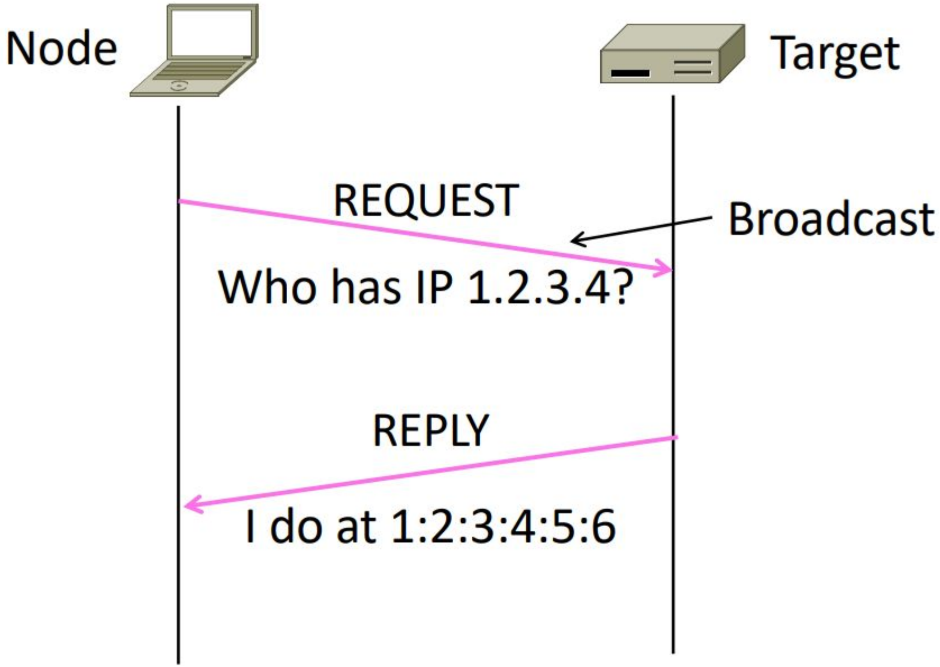
Which of the following is false about DHCP?

- DHCP is a client-server application that uses TCP.
- To get an IP address, a node must first send a broadcast message to the network
- A node can renew an existing lease by sending a REQUEST and receiving an ACK.
- DHCP can also be used to get the local DNS server address.

ARP - Address Resolution Protocol

- MAC is needed to send a frame over the local link
- ARP to map an IP to MAC
- Sits on top of link layer



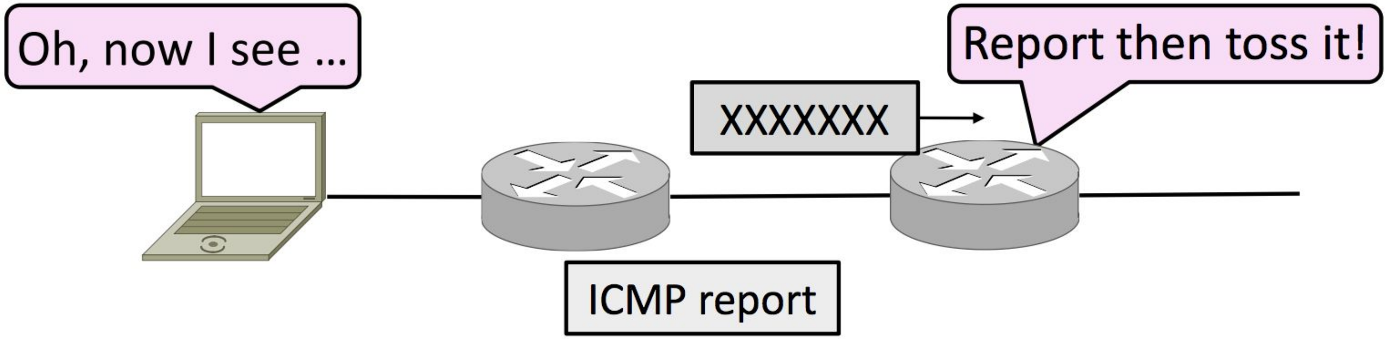


Which of the following is false about ARP?

- ARP is used to fill in the Destination Ethernet address in a packet header.
- ARP involves a central server broadcasting link layer addresses to other nodes.
- ARP is used to map a local IP address to its link layer address.
- ARP is replaced by NDP (Neighbor Discovery Protocol) in IPV6.

ICMP - Internet Control Message Protocol

- Provides error reporting and testing
- Companion protocol to IP
- Traceroute, Ping



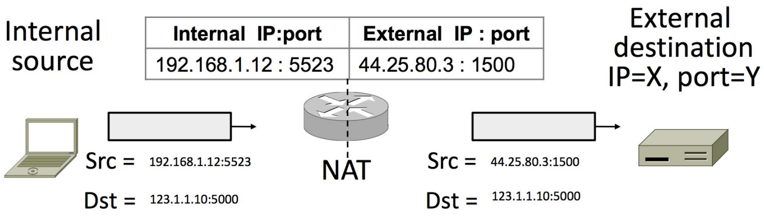
In which of the following cases will an ICMP error report NOT be sent to the source IP?

- Message does not reach destination within the TTL specified in header.
- The destination is unreachable.
- The fragment is bigger than the MTU (Maximum Transmission Unit)
- A packet is lost in the network and needs to be retransmitted.

NAT - Network Address Translation

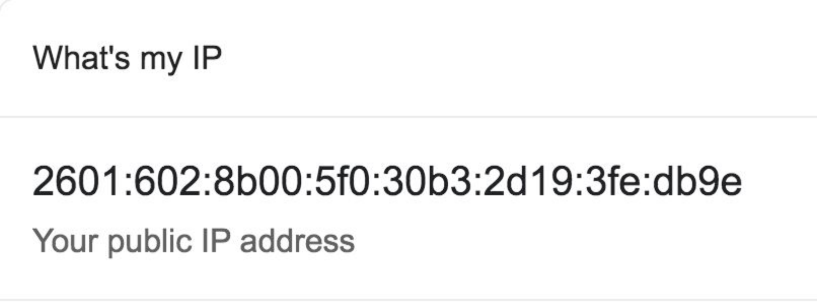
- One solution to **IPv4 address exhaustion**
- Map many private IP to one public IP, with different port number
- True/False?: NAT is many to one. many public IPs are translated to 1 private IP so that hosts can talk to home routers.

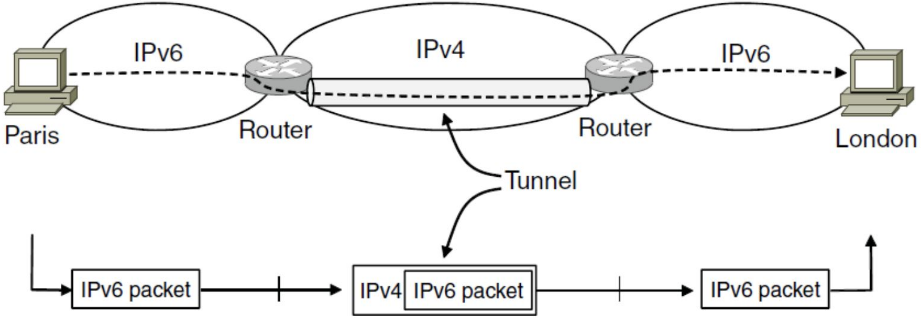
What host thinks	What ISP thinks
Internal IP:port	**External IP : port**
192.168.1.12 : 5523	44.25.80.3 : 1500
192.168.1.13 : 1234	44.25.80.3 : 1501
192.168.2.20 : 1234	44.25.80.3 : 1502



IPv6

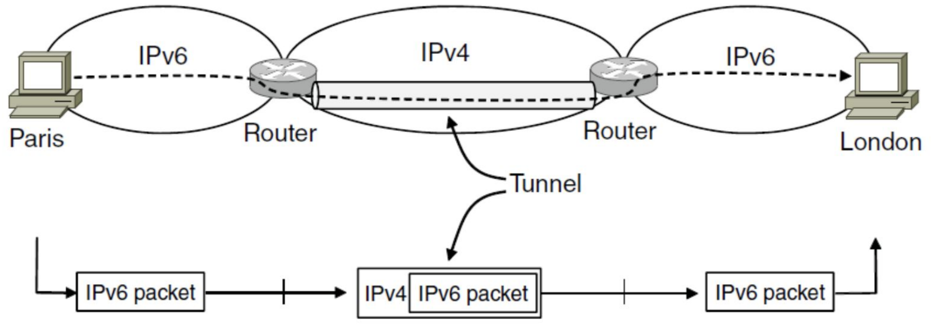
- A much better solution to IPv4 address exhaustion
- Uses 128-bit addresses, with lots of other changes
- IPv6 version protocols: NDP -> ARP, SLAAC -> DHCP
- Problem: being incompatible with IPV4. Solution: Tunnelling





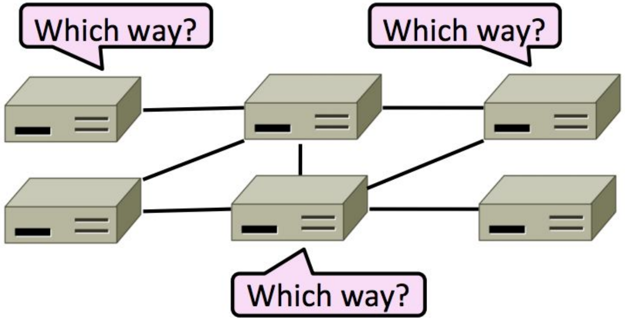
Tunneling
Which of the following is false about tunneling?

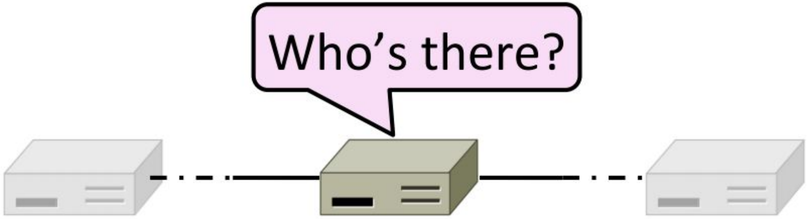
- Setup complexity is high
- packet size increases when using a tunnel
- Adding tunnels as a single link increases routing optimization
- Added layer of security is gained from the use of encrypted tunnels



Routing

- The process of deciding in which direction to send traffic
- Delivery models: unicast, broadcast, multicast, anycast
- Goals: correctness, efficient paths, fair paths, fast convergence, scalability
- Rules: decentralized, distributed setting

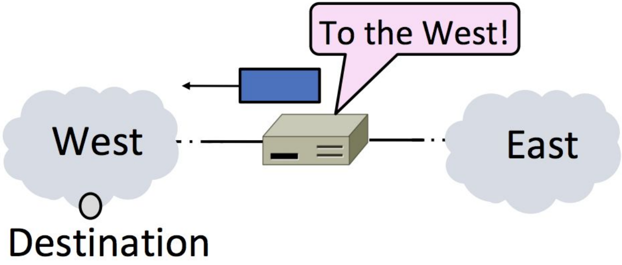




Techniques to Scale Routing

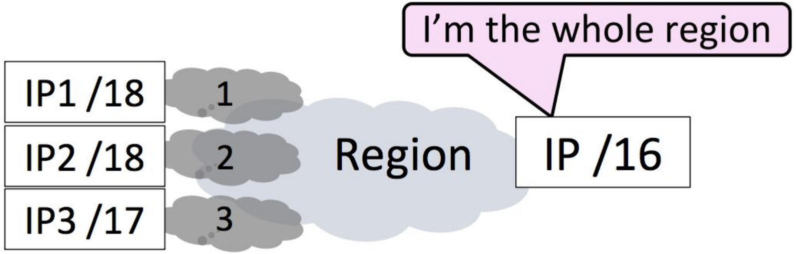
Hierarchical Routing

- Route first to the region, then to the IP prefix within the region



IP Prefix Aggregation and Subnets

- Adjusting the size of IP prefixes
  - Internally split one large prefix
  - Externally join multiple IP prefixes



Best Path Routing

Distance Vector Routing

Each node maintains a vector of distances (and next hops) to all destinations.

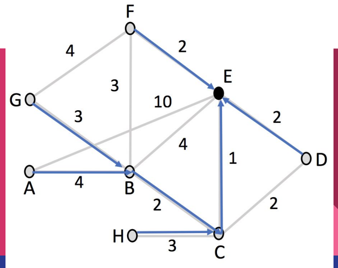
Sometimes doesn't perform very well: count-to-infinity scenario

Link State Routing (widely used)

Phase 1. **Topology Dissemination**: Nodes flood topology

Phase 2. **Route Computation**: running Dijkstra algorithm (or equivalent)

Algorithm details available in lecture slides

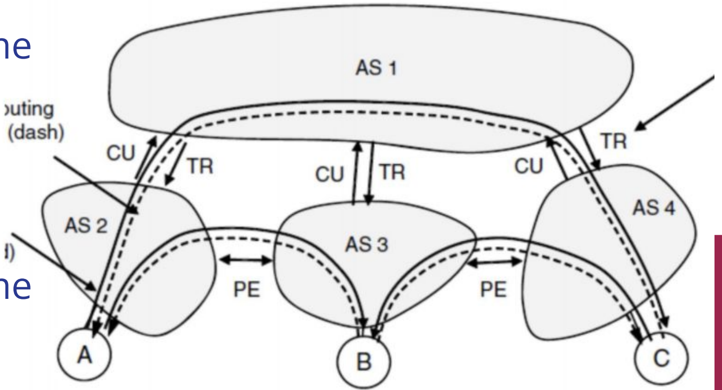


BGP - Border Gateway Protocol

- Internet-wide routing between ISPs (ASes)
  - Each has their own policy decisions
- Peer and Transit (Customer) relationship
- Border routers of ISPs announce BGP routes only to other parties who may use those paths.
- Border routers of ISPs select the best path of the ones they hear in any, non-shortest way

BGP example

- Transit (ISP & Customer)
  - ISP announce everything it can reach to its customer
    - AS1 to AS2: you can send packet to AS4 through me
  - Customer ISP only announce its customers to ISP
    - AS2 to AS1: you can send packet to A through me
- Peer (ISP 1 & ISP 2)
  - ISP 1 only announces its customer to ISP 2
    - AS2 to AS3: you can send packet to A through me



Transport Layer

- Service Models
- TCP vs UDP
- TCP Connections
- Flow Control and Sliding Window
- TCP Congestion Control
- Newer TCP Implementations

Service Models

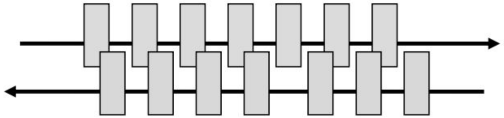
- Transport Layer Services
  - Datagrams (UDP): Unreliable Messages
  - Streams (TCP): Reliable Bytestreams

TCP vs UDP
Which of the following statements about TCP is false?

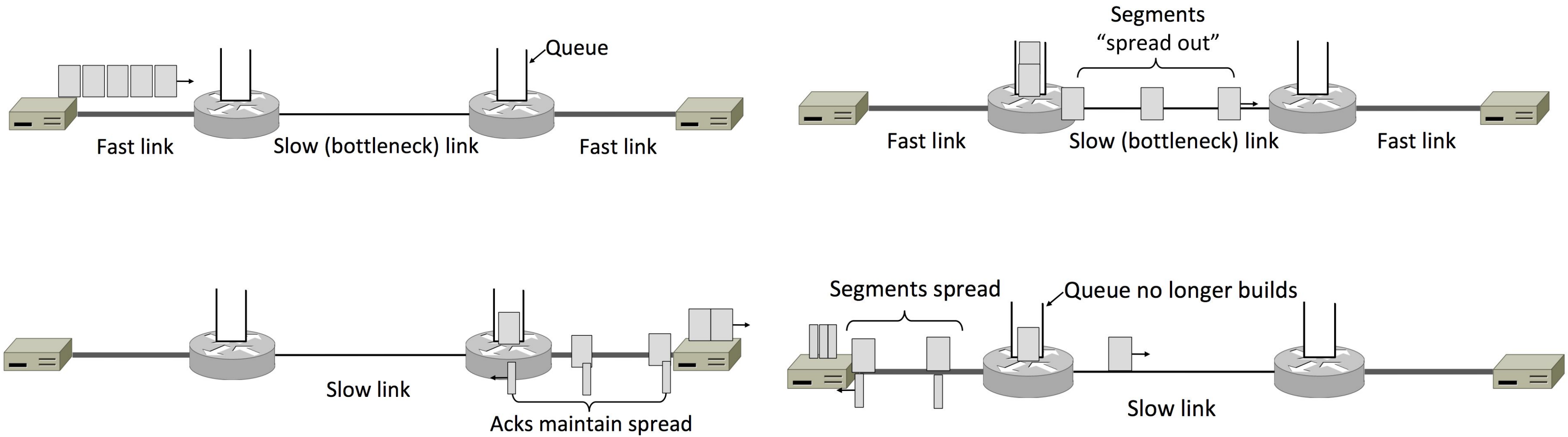
- Unlike UDP, a three-way handshake is used to set up a TCP connection.
- Similar to UDP, TCP can retransmit packets.
- Pipelining can improve performance while maintaining reliability.
- Connection release is symmetric: both sides shutdown independently

Flow Control - Sliding Window Protocol

- Receiver sends ACK upon receiving packets
  - Go-Back-N (p1 part b)
  - **Selective Repeat**
    - Receiver passes data to app in order
    - Buffers out-of-order segments to reduce retransmissions
    - ACK highest in-order segment
- **Selective Retransmission** on sender's side

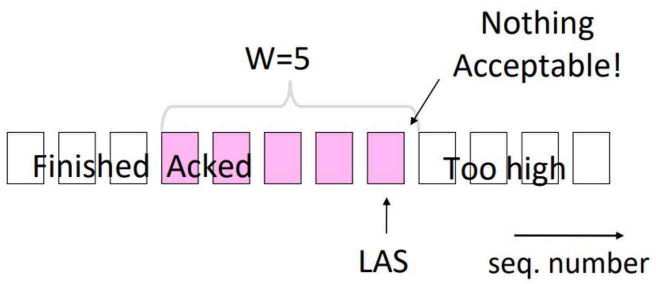


Flow Control - ACK Clock



Flow Control - Sliding Window Protocol (2)

- Avoid loss —-------------------->
  - Let receiver tell sender how much free buffer space receiver has (WIN)
  - Update the value regularly



Flow Control - Sliding Window Protocol (3)

Retransmission?

- Use timer for each segment
- How to set a **timeout**?
  - Adaptively Timeout: determine timeout value based on smoothed estimate of RTT

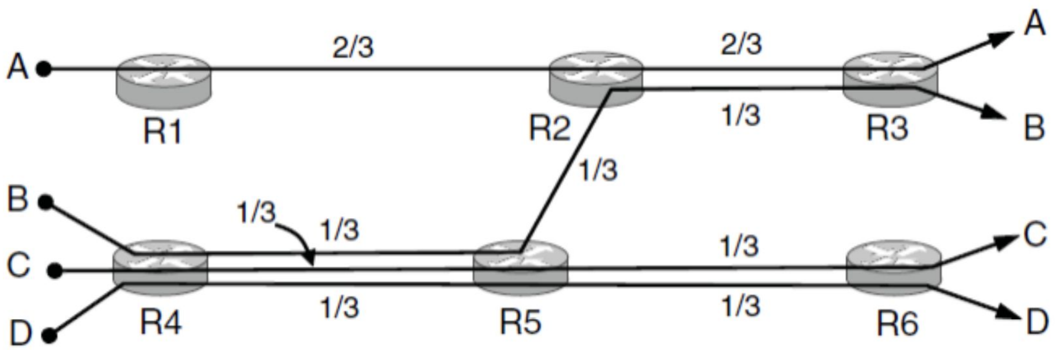
Bandwidth Allocation

Good allocation is both efficient and fair

- Efficient means most capacity is used but there is no congestion
- Fair means every sender gets a reasonable share the network

Max-Min Fair Allocation

- Start with all flows at rate 0
- Increase the flows until there is a new bottleneck in the network
- Hold fixed the rate of the flows that are bottlenecked
- Go to step 2 for any remaining flows

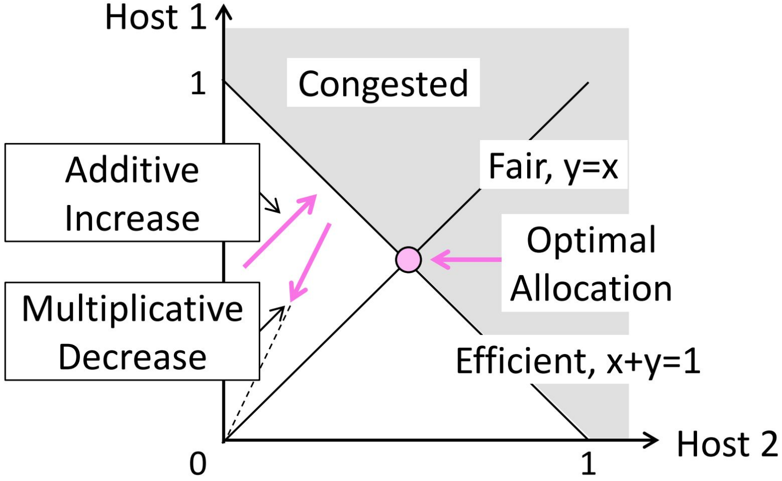


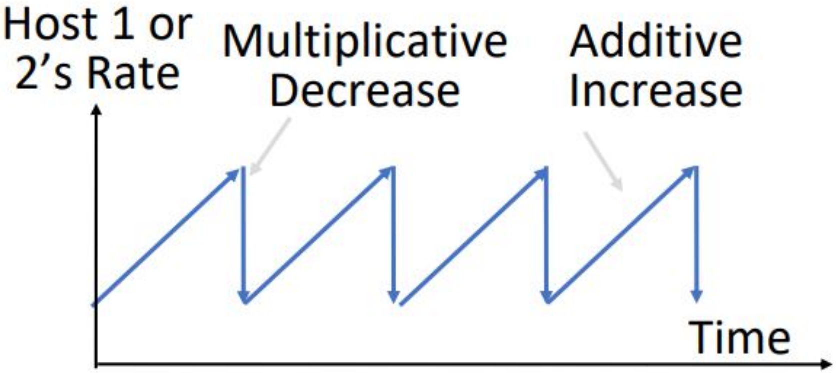
TCP Bandwidth Allocation

- Closed loop: use feedback to adjust rates
  - NOT open loop: reserve bandwidth before use
- Host driven: host sets/enforces allocations
  - NOT network driven
- Window based
  - NOT rate based
- Congestion/Feedback signals
  - Packet loss, Packet delay, Router indication

Additive Increase
Multiplicative Decrease
AIMD

AIMD - Additive Increase Multiplicative Decrease



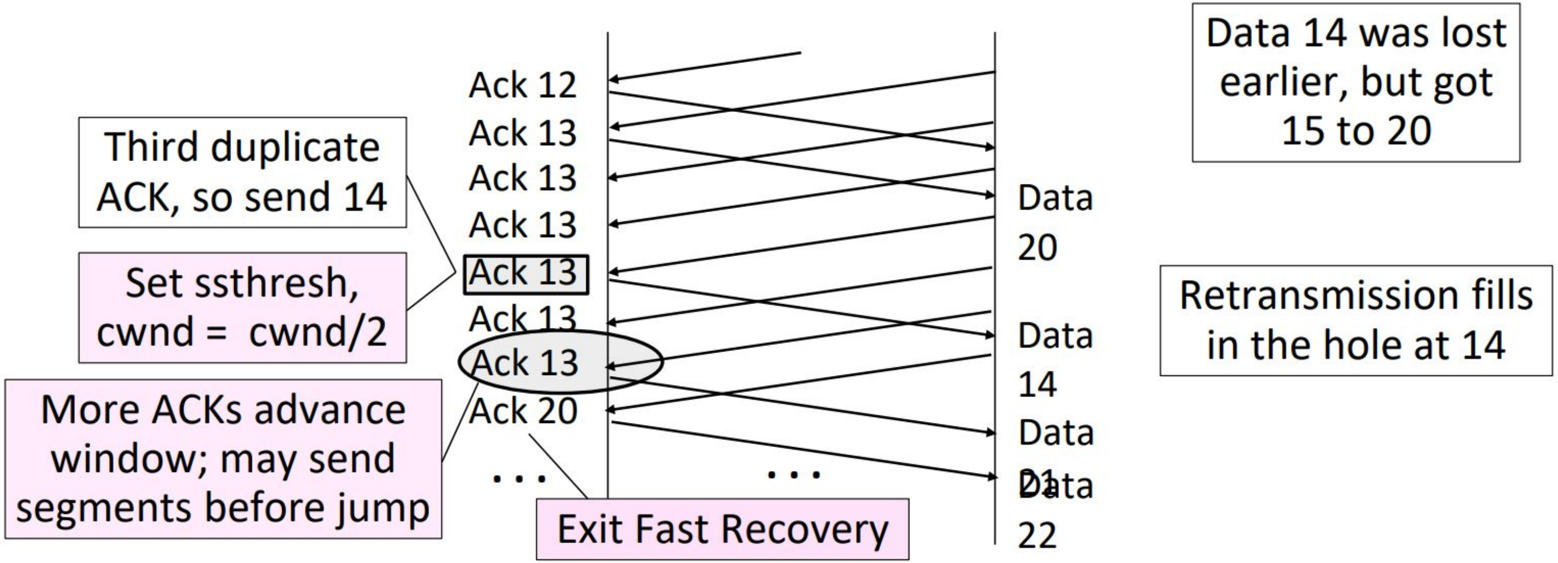


Practical AIMD

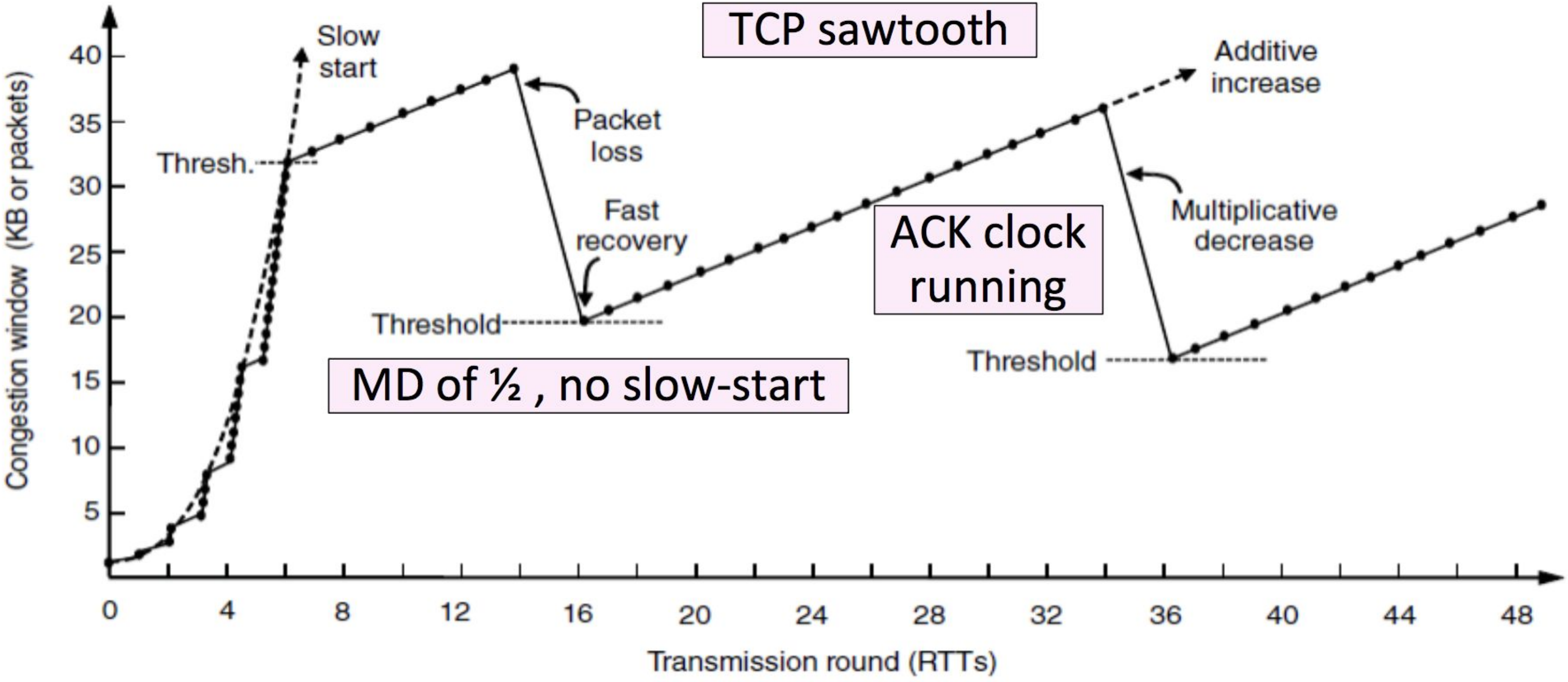
- **Slow-Start** (used in AI)
  - Exponential growth (1, 2, 4, 8, 16, …)
  - Start slow, quickly reach large values
- **Fast-Retransmit** (used in MD)
  - Three duplicate ACKs = packet loss
  - Don’t have to wait for TIMEOUT
- **Fast-Recovery** (used in MD)
  - Instead of timeout or slow-start, pretend duplicate ACKs after fast retransmit are the expected ACKs
  - With fast retransmit, it repairs a single segment loss quickly and keeps the ACK clock running

Practical AIMD -

Fast retransmit + recovery

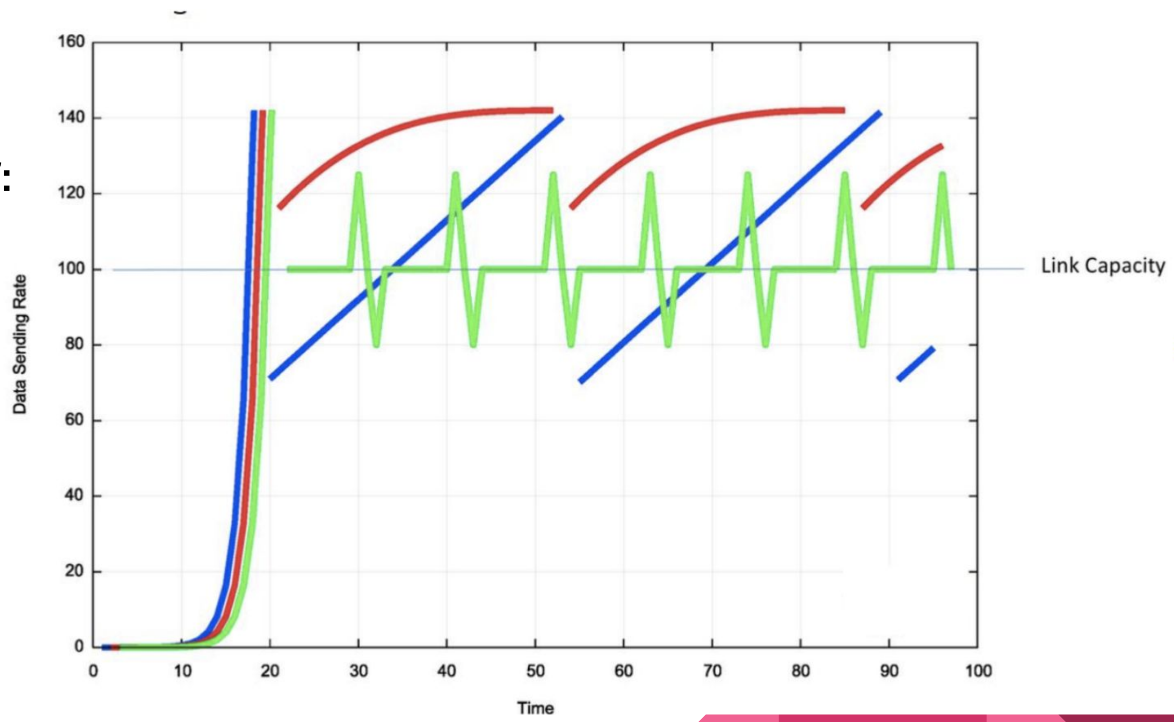


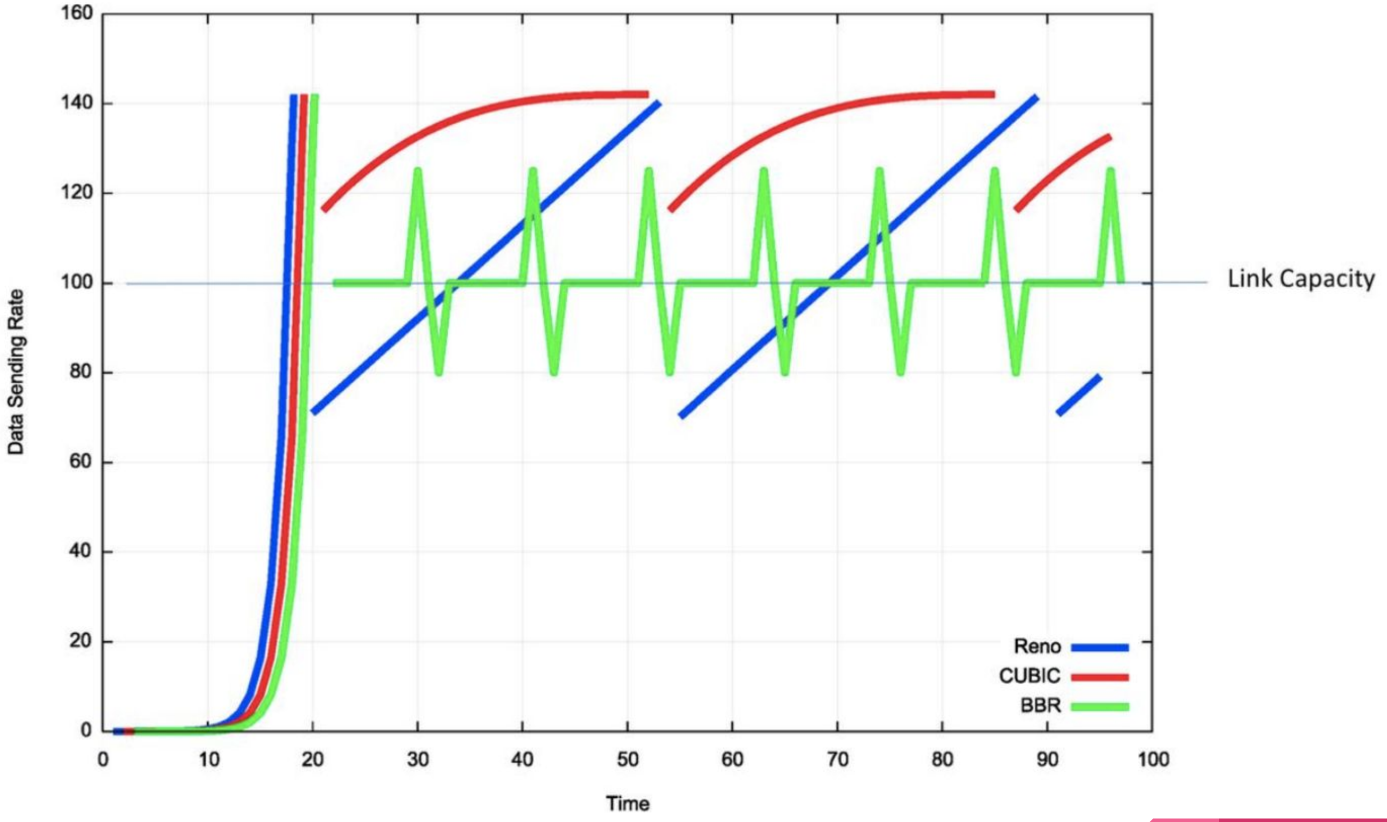
TCP Reno


TCP sawtooth
ACK clock running
MD of ½ , no slow-start
Slow start
Thresh.
Packet loss
Fast recovery
Threshold
Additive increase
Multiplicative decrease
Threshold
Congestion window (KB or packets)
Transmission round (RTTs)
0
4
8
12
16
20
24
28
32
36
40
44
48
5
10
15
20
25
30
35
40

These three lines represent effects of:

- Different queueing algorithms
- Different buffer sizing algorithms
- Different bandwidth allocation strategies
- None of the above





Congestion Avoidance Cubic: Packet Loss
BBR: Packet delay

Network-Side Congestion Control

- Explicit Congestion Notification (**ECN**)
  - Router detects the onset of congestion via its queue.
  - Marked packets treated as loss at receiver.

- Random Early Detection (**RED**)
  - Instead of marking packets (why), drop at random
  - As queue approaches full, increase likelihood of packet drop
  - Example: 1 queue slot left, 10 packets expected, 90% chance of drop

Application Layer

- DNS
- HTTP
- Web Caching / CDN
- Security

DNS

- Terminology
  - Names - higher-level identifiers for resources
  - Addresses - lower-level locators for resources
  - Resolution/lookup - mapping a name to an address
  - Zones - contiguous portions of the namespace
  - Nameserver - server to contact for information about a particular zone
- Maps between host names and address
- Recursive vs iterative query
- Caching
- Built on top of UDP
- Security

HTTP - HyperText Transfer Protocol

Basis for fetching Web pages

Steps to fetch a web HTTP with URL:

- Resolve the server IP
- Setup TCP connection
- Send/Receive HTTP request over TCP
- Fetch embedded resources
- Teardown TCP connection

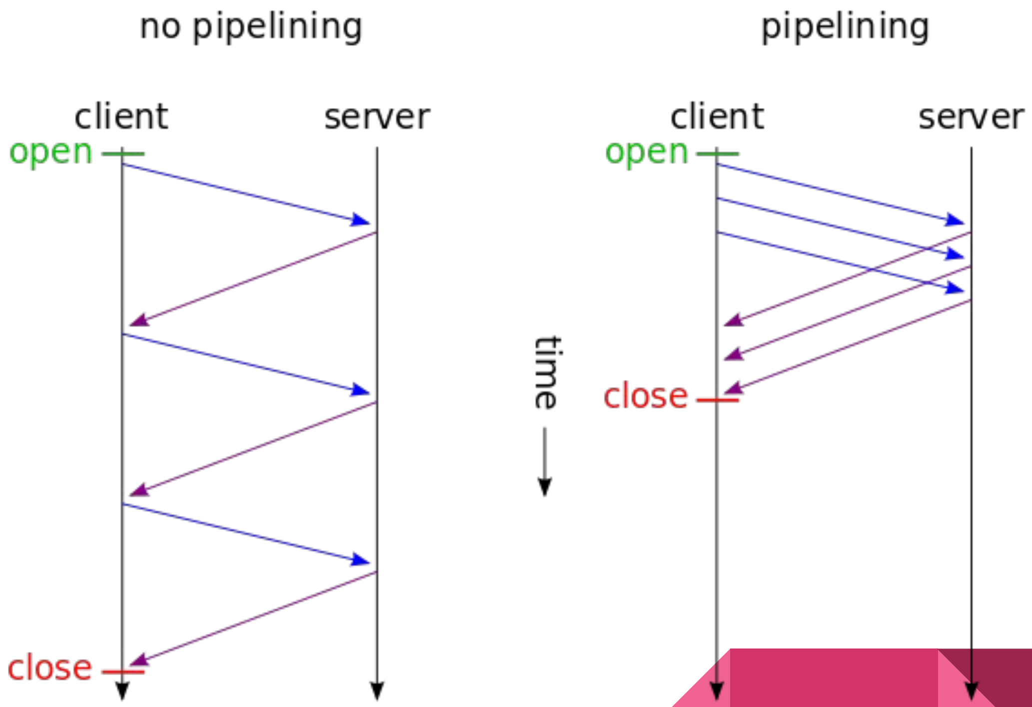
HTTP

- Commands in request- GET/POST/...
- Codes in response - 2xx=Success/4xx=Client Error/...

Page Load Time

How to decrease Page Load Time (PLT)?

- Parallel connections and persistent connections (HTTP1.1)
- HTTP caching and proxies
- Change HTTP protocol
- Move content closer to client (CDNs)



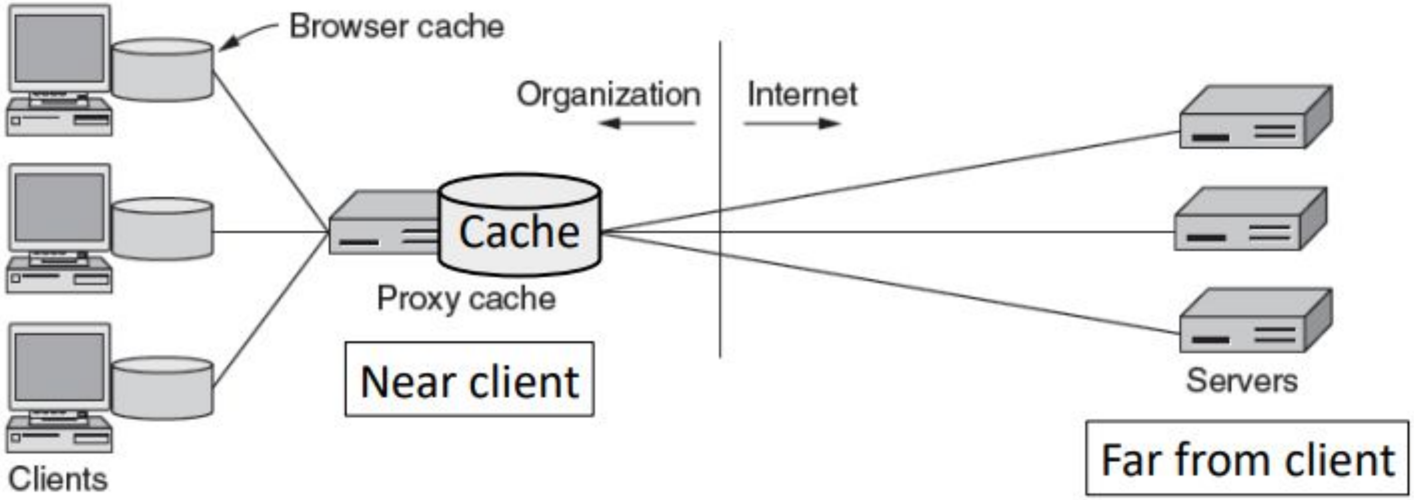
Caching/Proxies

Web Caching

- Local copy on browser
- Revalidate copy with remote server
  - Timestamp
  - Server header

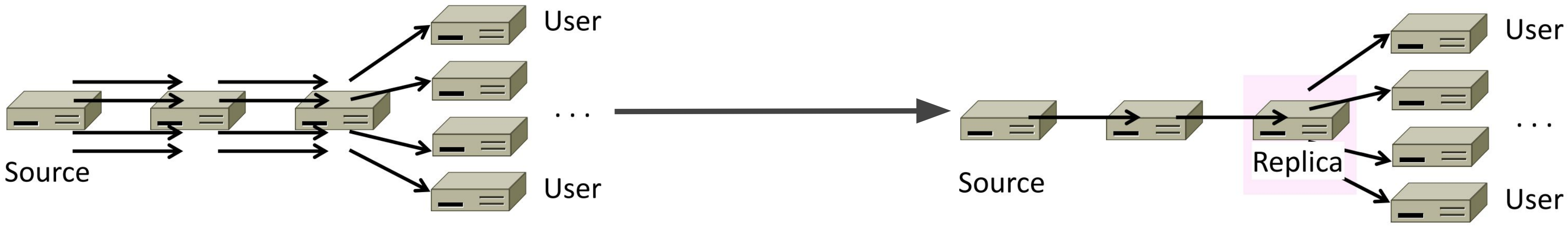
Web Proxies

- Placed near pool of clients
  - Caching
  - Security checking
  - Organization policies



CDNs

- Content Delivery Networks
- Place popular content near clients
    - Use DNS to place replicas across the Internet for use by all nearby clients
    - Reduces server, network load, improves user experience



Which of the following is true about CDNs:

- They manipulate the DNS table to place multiple entries corresponding to each initial entry
- They increase server load (storage magnification), but decrease routing load
- All of the above
- None of the above