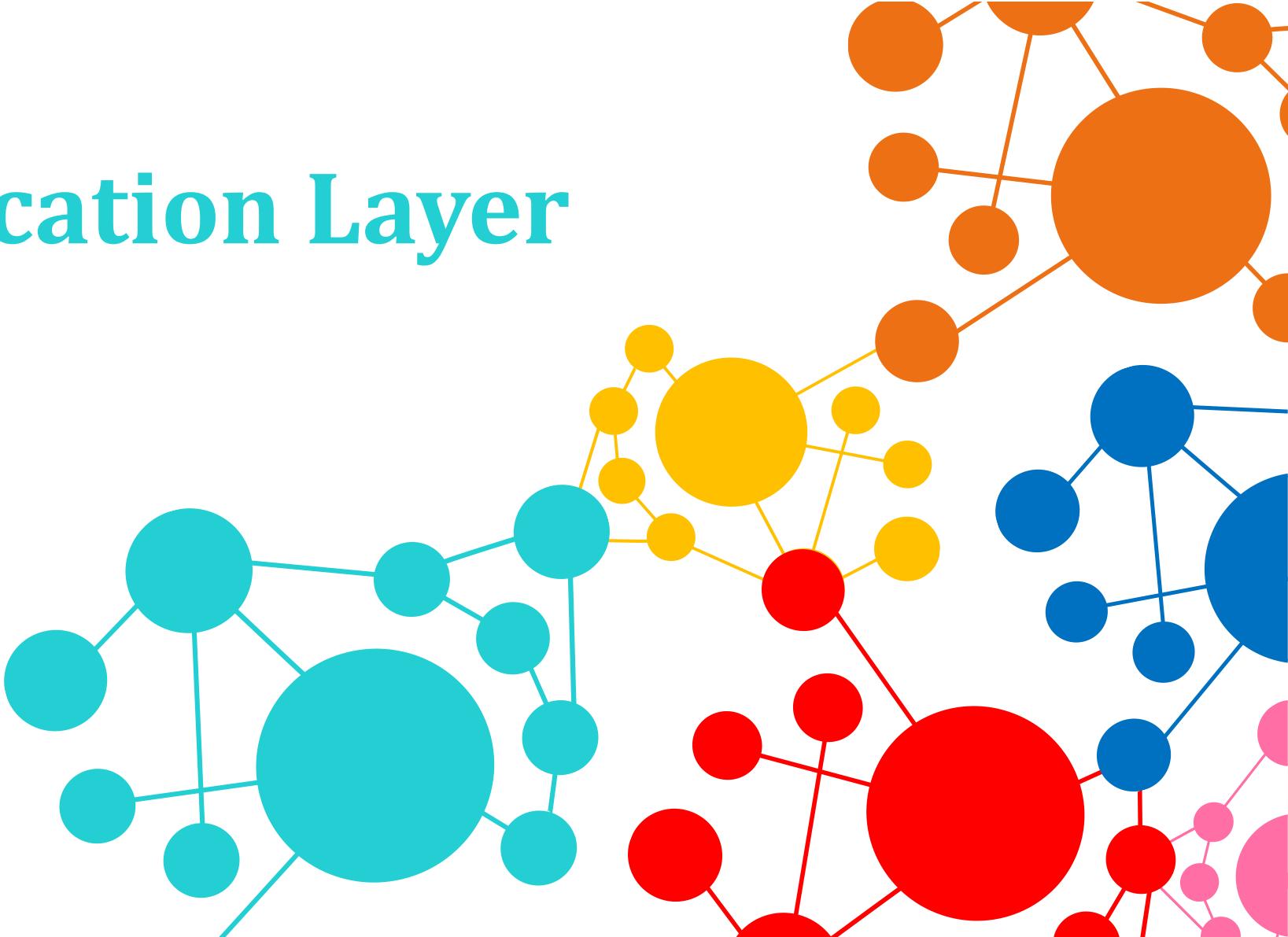
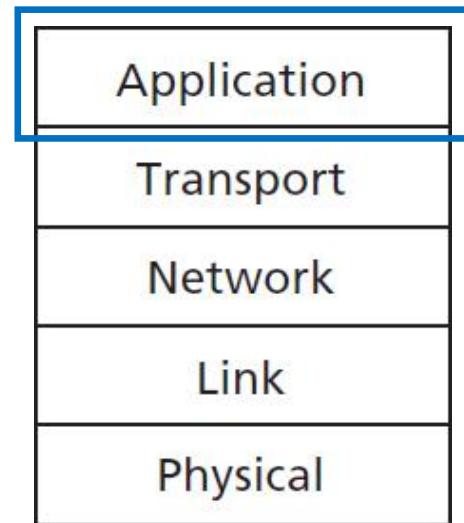


Application Layer



Application Layer

- Network Applications
- Web & HTTP
- E-mail (SMTP, POP3, IMAP)
- Domain Name System (DNS)
- Peer-to-Peer (P2P) Applications
- Video Streaming and Content Distribution Networks (CDN)
- Socket Programming (UDP & TCP)



Network Applications



Network Applications

- Example Network Applications
 - E-mail
 - Web
 - Messaging
 - Remote Login
 - P2P File Sharing (BitTorrent)
 - Multi-user Network Games
 - Video Streaming (YouTube, Hulu, Netflix)
 - Voice Over IP (Skype)
 - Real-Time Video Conferencing
 - Social Networking
 - Search

- Programs that

- Run on (different) end systems
- Communicate over network

Some utilize different protocols

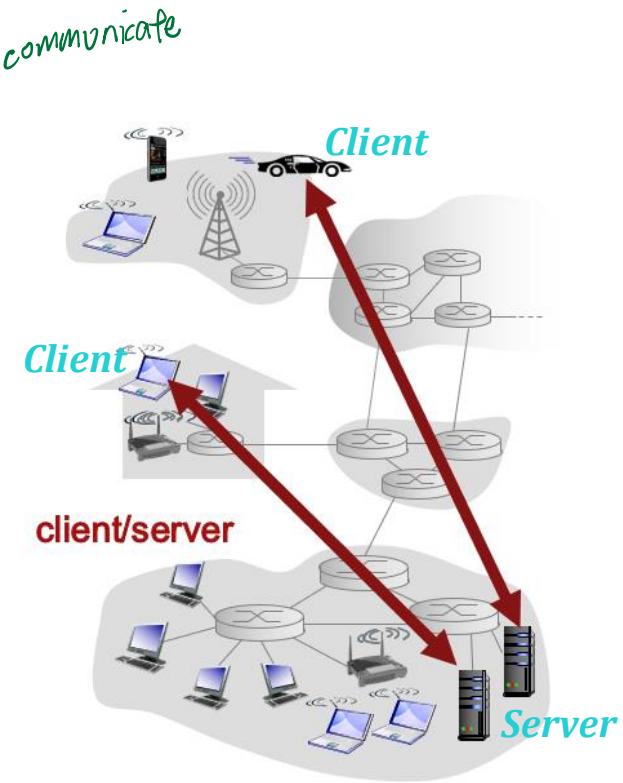
- Application Architectures

- Client-Server
- Peer-to-Peer (P2P)

happens between 2 nodes - P2P Direct while most go to server then receiver
↳ even though might seem

Network Applications: Client-Server

- Server
 - Always-on host ↗
 - Permanent IP address - *Cause if change dunno when to communicate*
 - Data centers for scaling
 - ↳ Handling demand
- Clients
 - Communicate with Server
 - May be intermittently connected
 - May have **dynamic** IP addresses
 - Do not communicate directly with each other



Collective Power

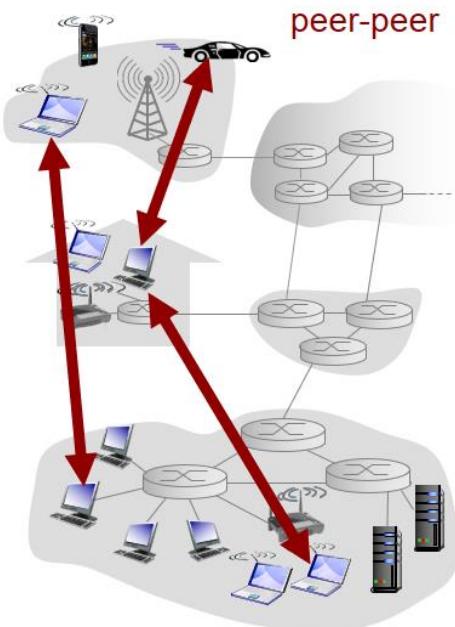
Network Applications: P2P

sets of nodes collectively running something

- No always-on server: Arbitrary end systems directly communicate
 - Peers request service from other peers
 - Peer provide service in return to other peers
- Self Scalability
 - New peers bring new service capacity
 - New peers bring new service demands
- Peers are intermittently connected & change IP addresses
 - Complex Management

Torrenting

Newer users bring capacities



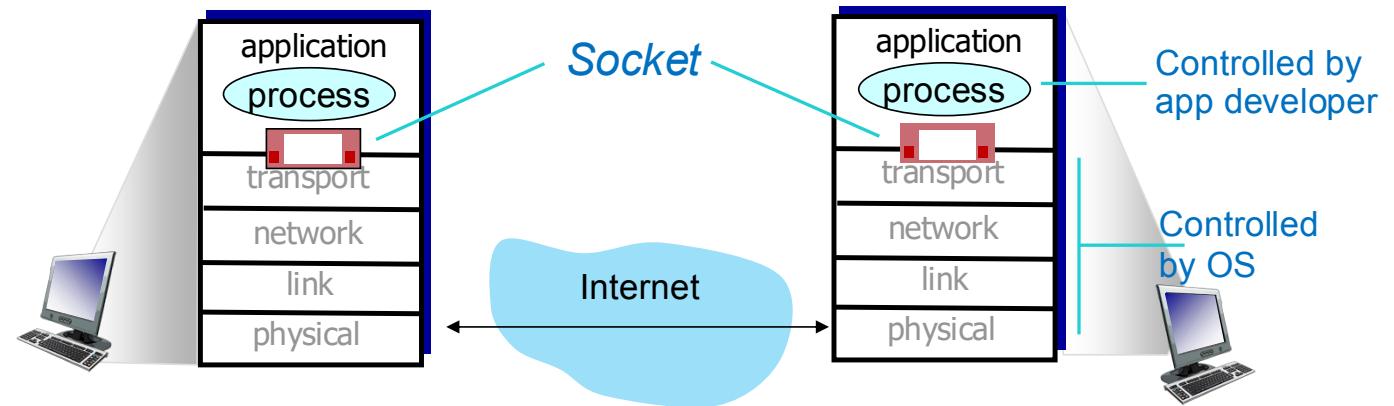
run on end host to client

Network Applications: Processes

- Process: Program running on a host
- Processes Communicating
 - Processes running on the same host
 - Two processes communicate using inter-process communication (Defined by OS)
 - Processes running on different hosts
 - Two processes communicate by message exchange (using **Socket**) over the network

Network Applications: Socket

- Processes send & receive messages using Sockets



- Socket analogous to door *similar to addresses*
 - Sending process sends message out the door
 - Sending process relies on transport infrastructure on the other side of the door to deliver message to socket at the receiving process

Network Applications: Process Addressing

- Address (identifier) needed for processes to receive messages
 - IP Address (Unique 32-bit address)
 - Is IP address enough? **No**
 - Processes running on the same host → IP address + port number
- Example
 - To send HTTP messages to `cs.sfu.ca` web server
 - **IP Address:** 142.58.102.68 – *street address*.
 - **Port:** 80 – *door number*

Network Applications: Protocols

- Application Layer Protocol defines
 - Types of messages exchanged (e.g. request, response,...)
 - Message Syntax (fields, and field delineated)
 - Message Semantics (Meaning of fields)
 - Rules (When & how processes send & respond to messages) *Behavior*
- Protocols
 - Open
 - Examples: HTTP, SMTP
 - Refined in RFCs
 - Allows for interoperability
 - Proprietary
 - Examples: Skype

Network Applications: Needs

- Throughput
 - Minimum throughput guarantees
 - Make use of available throughput
- Data Integrity
 - Total reliability - email
 - Some loss acceptable
- Timing
 - Low delay - facetime
 - Internet telephony
 - interactive games
- Security
 - Encryption - Banking
 - Data Integrity

✓ zoom meetings / online games / live stream
elastic - whatever throughput available works

Application	Data Loss	Throughput	Time-Sensitive
File transfer	no loss	elastic	no
E-mail	no loss	elastic	no
Web documents	no loss	elastic	no
Text messaging	no loss	elastic	yes and no
Real-time audio/video	loss-tolerant (little pixelation)	audio: 5kbps-1Mbps video: 10kbps-25Mbps	yes, 100's msec
Interactive games	loss-tolerant	few kbps up	yes, 100's msec
Stored audio/video	loss-tolerant	same as above	yes, few secs

Which protocol

Network Applications: Transport Protocol Services

- Transport Layer services to Application Layer for delivery of messages

- Transport Layer Protocols

- TCP

- Reliable Data Transfer
- Flow Control - do not want source overload the destination
- Congestion Control - do not want overwhelm the networks.
↳ controls amount of data sent
- Does not provide timing, throughput guarantee, and security
- Connection-oriented
↳ like building pipes & slide packet through
↳ establish connection

- UDP

- Does not provide reliable data transfer, NO GUARANTEES
flow control, congestion control, timing,
throughput guarantee, security, or connection setup
- Connection-less - tries to send packet

↳ like telling delivery to send, n hopefully sends

Networks layer



Covered in detail
in Transport Layer

Network Applications: Transport Protocol Services

<i>Application</i>	<i>Application layer protocol</i>	<i>Underlying transport protocol</i>
E-mail	<i>SMTP [RFC 2821]</i>	<i>TCP</i>
Remote terminal access	<i>Telnet [RFC 854]</i>	<i>TCP</i>
Web	<i>HTTP [RFC 2616]</i>	<i>TCP</i>
File transfer	<i>FTP [RFC 959]</i>	<i>TCP</i>
Streaming multimedia	<i>HTTP (e.g., YouTube), RTP [RFC 1889]</i>	<i>TCP or UDP</i>
Internet telephony	<i>SIP, RTP, proprietary (e.g., Skype)</i>	<i>TCP or UDP</i>

↗

Possible to implement security & reliability info Applications



Web & HTTP

https checks security of server

Web & HTTP

- Web pages consist of a base HTML (Hypertext Markup Language) file, including several references objects including
 - HTML file
 - JPEG image
 - Java Applet
 - Audio file
- Each object **addressed** by a **URL** (Uniform Resource Locator)

- URL example: `www.sfu.ca/computing/research.html`

The URL `www.sfu.ca/computing/research.html` is shown. A curly brace underlines the host name `www.sfu.ca`, and another curly brace underlines the path name `computing/research.html`. The text "Host Name" is written below the first brace, and "Path Name" is written below the second brace.

HTTP

1 request 1 response

- Hypertext Transport Protocol (HTTP)

- Web's application layer protocol

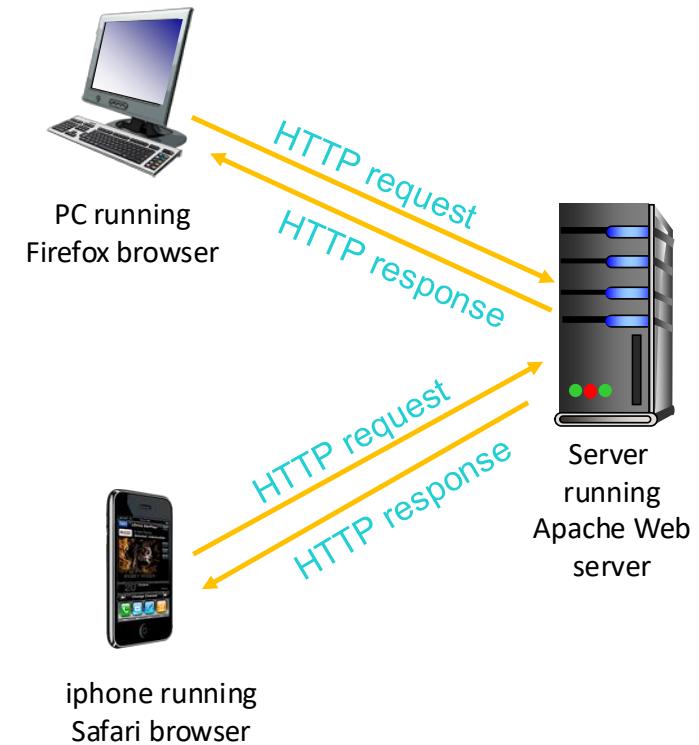
- **Client-Server**

- **Client:** Browser that requests, receives, and displays web objects
 - **Server:** Web server sends object in response to requests

- **Stateless:** No information about past client requests

- Uses **TCP** *connection requires
↳ establishes pipes*

- 1 • Client initiates TCP connection to server port 80
- 2 • Server accepts TCP connection
- 3 • HTTP messages exchange
- 4 • TCP connection close



HTTP Connections

- Non-Persistent HTTP

- At most one object sent over TCP connection
- Downloading multiple objects required multiple connections

→ send request, received, close connection

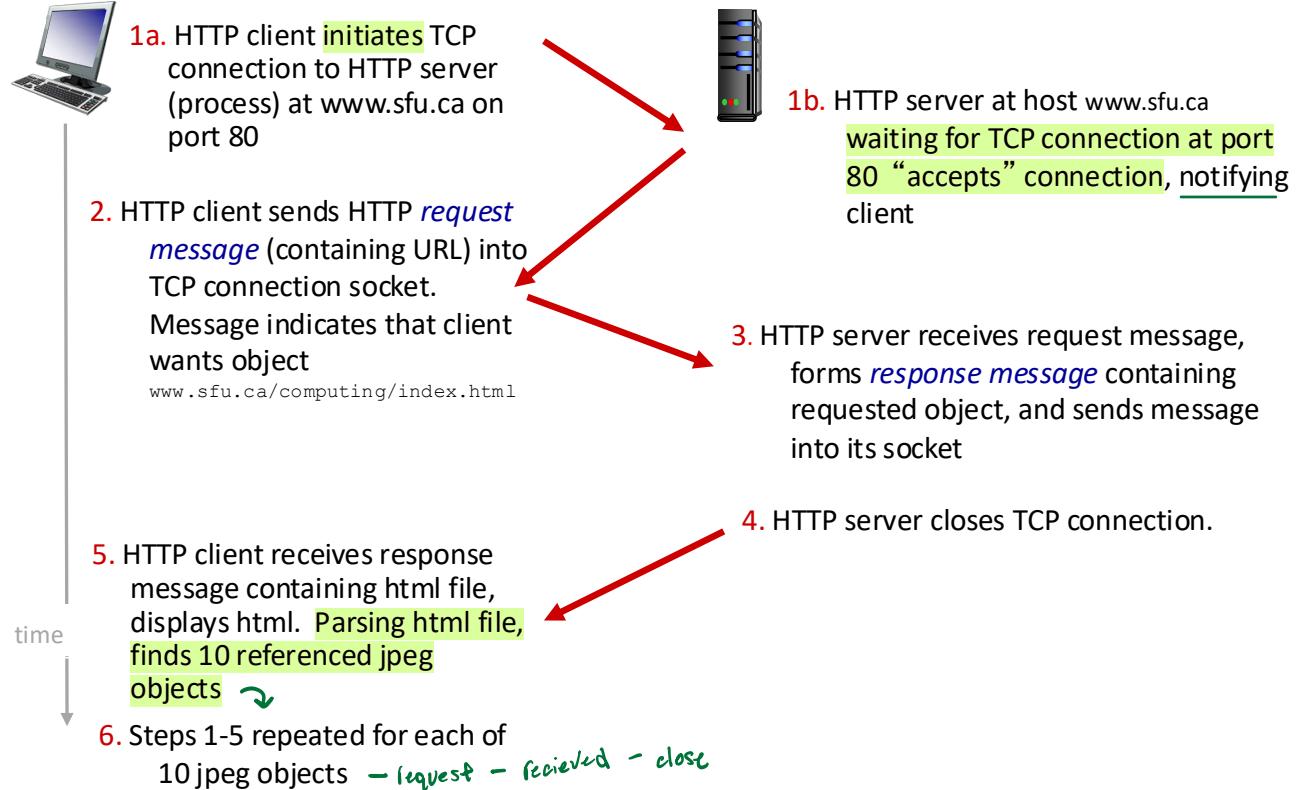
- Persistent HTTP → doesn't close pipes

- Multiple objects can be sent over single TCP connection between client and server

Non-Persistent HTTP

- User enters URL:

www.sfu.ca/computing/index.html



Non-Persistent HTTP

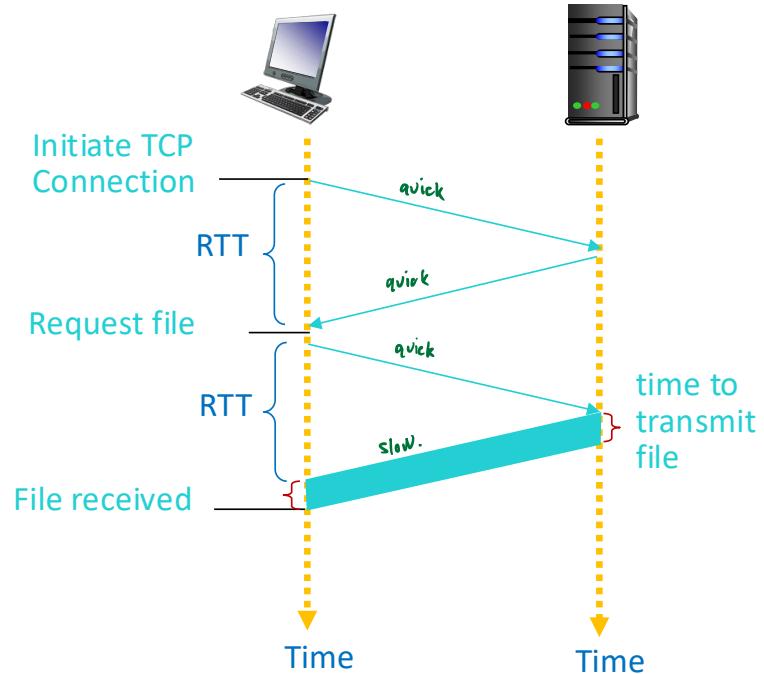
Round Trip Time

- **RTT:** Time for a small packet to travel from client to server & back
- **HTTP response time:**
 - One RTT to initiate TCP connection
 - One RTT for HTTP request and first few bytes of HTTP response to return
 - File transmission time

$$\text{Non-persistent HTTP response time} = 2\text{RTT} + \text{file transmission time}$$

✓
initiate & request

- **Issue**
 - Two RTTs per object
 - OS Overhead for each TCP connection
 - Browsers often open parallel TCP connections to fetch referenced objects



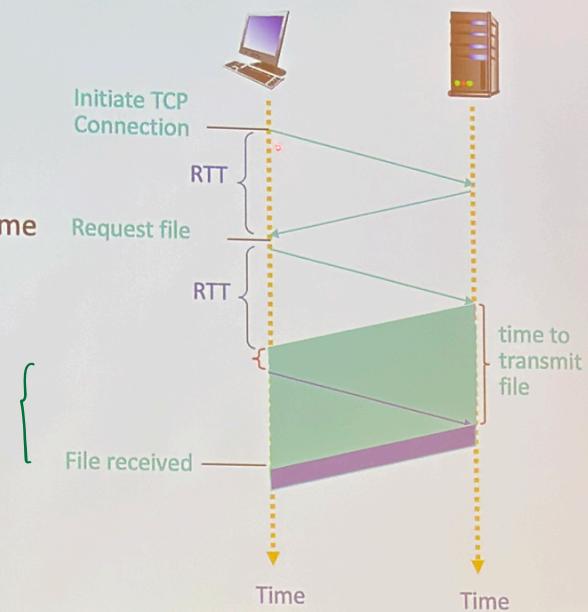
Persistent HTTP

- Server leaves connection open after sending response
- Subsequent HTTP messages between same client server sent over open connection
- Client sends requests as soon as it encounters a referenced object
- As little as one RTT for all the referenced objects

Persistent Connection

- Request for objects
 - As soon as encountered
 - Could overlap with another object download time

transfers everything
Saves time cause no
need a request again

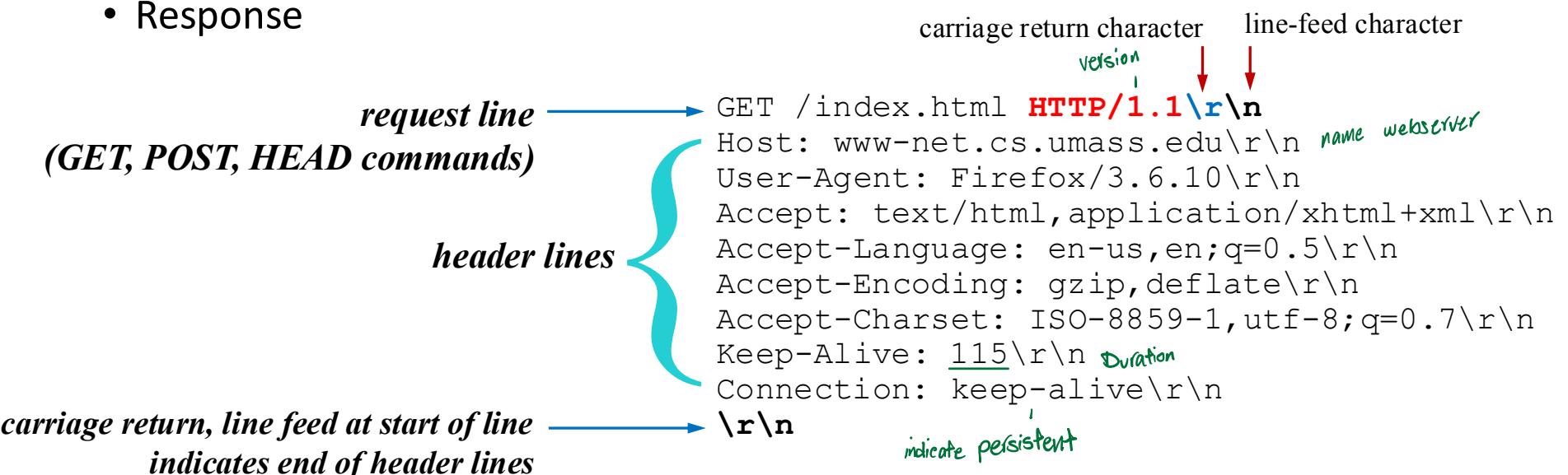


HTTP Messages

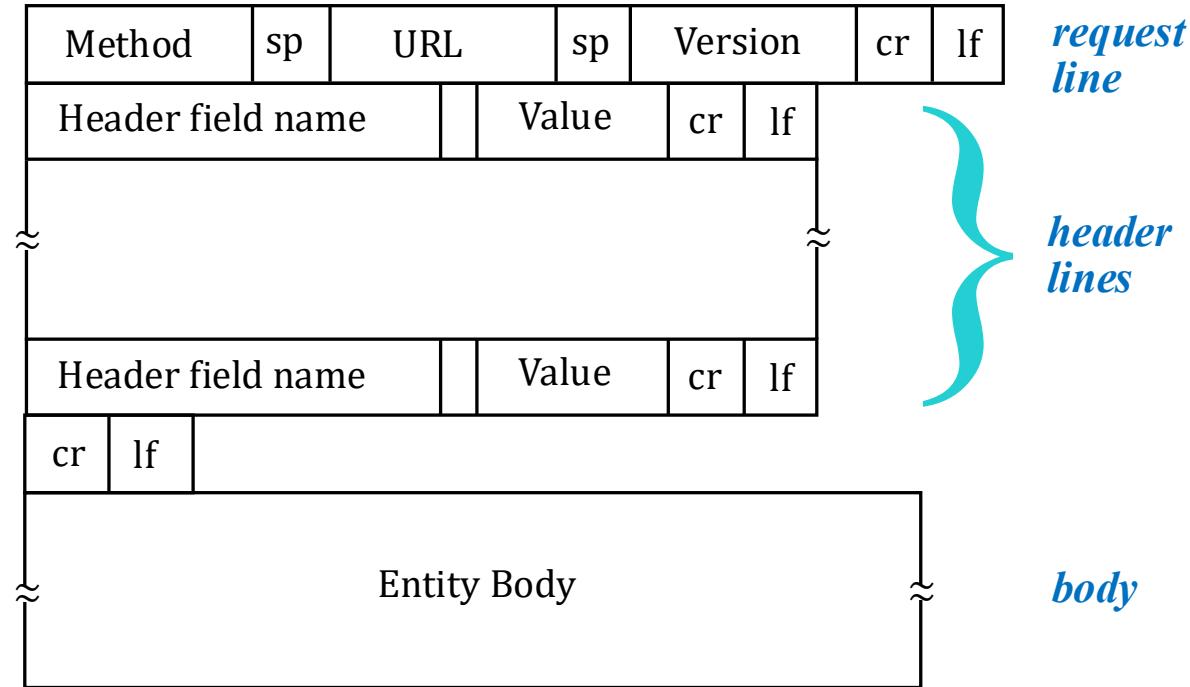
What are the 2 types of HTTP messages?

- Request (ASCII, human readable)
- Response

- Two types of HTTP messages
 - Request (ASCII, human-readable)
 - Response



HTTP Request: General



Method Types

HTTP/1.1:

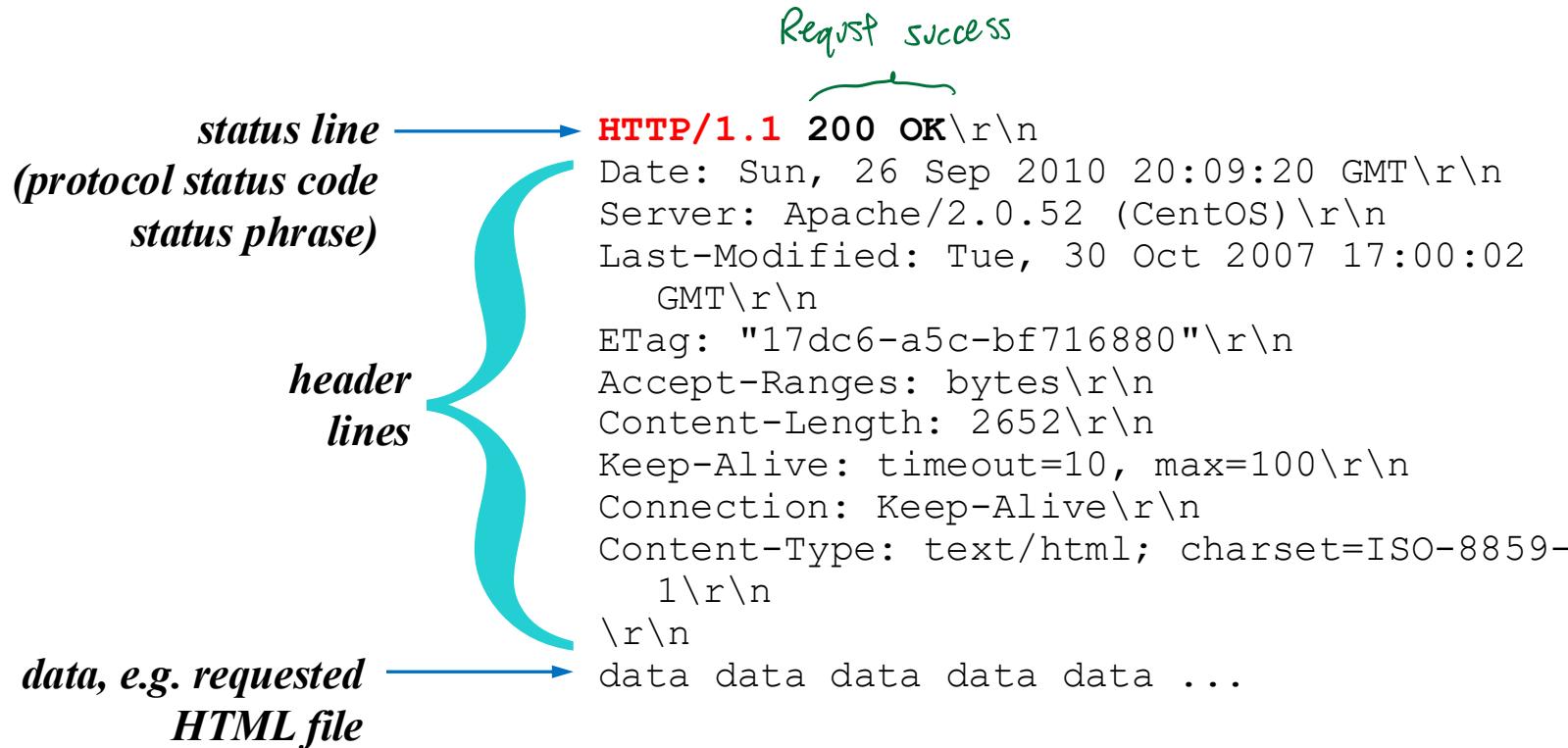
- GET
 - POST
 - HEAD
 - ↳ Only retrieves headers
 - PUT
 - Uploads file in entity body to path specified in URL field
 - DELETE
 - Deletes file specified in the URL field
- frequently used methods
↳ request line commands*

HTTP/1.0:

- GET
- POST
- HEAD
 - Asks server to leave requested object out of response
 - Used to get metadata about an object without transferring it

Only 1 response per request

HTTP Response



HTTP Response

- **Status Code** appear in the first line in the server to client response message
- **Example Codes**

200 OK Request succeeded, requested object later in this message

301 Moved Permanently ex - Redirect

Requested object moved, new location specified later in this message (Location:)

400 Bad Request ex - Client request missing something

Request message not understood by server

404 Not Found Server level issue

Requested document not found on this server

505 HTTP Version Not Supported

403 - client doesn't have permission to access

Trying HTTP (Client-Side)

1. Telnet to your favorite Web server:

```
telnet gaia.cs.umass.edu 80
```

known web App port



Opens TCP connection to port 80 (default HTTP server port) at gaia.cs.umass.edu. Anything typed in will be sent to port 80 at gaia.cs.umass.edu

2. Type in a GET HTTP request:

```
GET /kurose_ross/interactive/index.php HTTP/1.1  
Host: gaia.cs.umass.edu
```



By typing this in (hit carriage return twice), you send this minimal (but complete) GET request to HTTP server

3. Look at response message sent by HTTP server!

(Alternatively, use Wireshark to look at captured HTTP request/response)

Why is HTTP popular? Statelessness - Independent

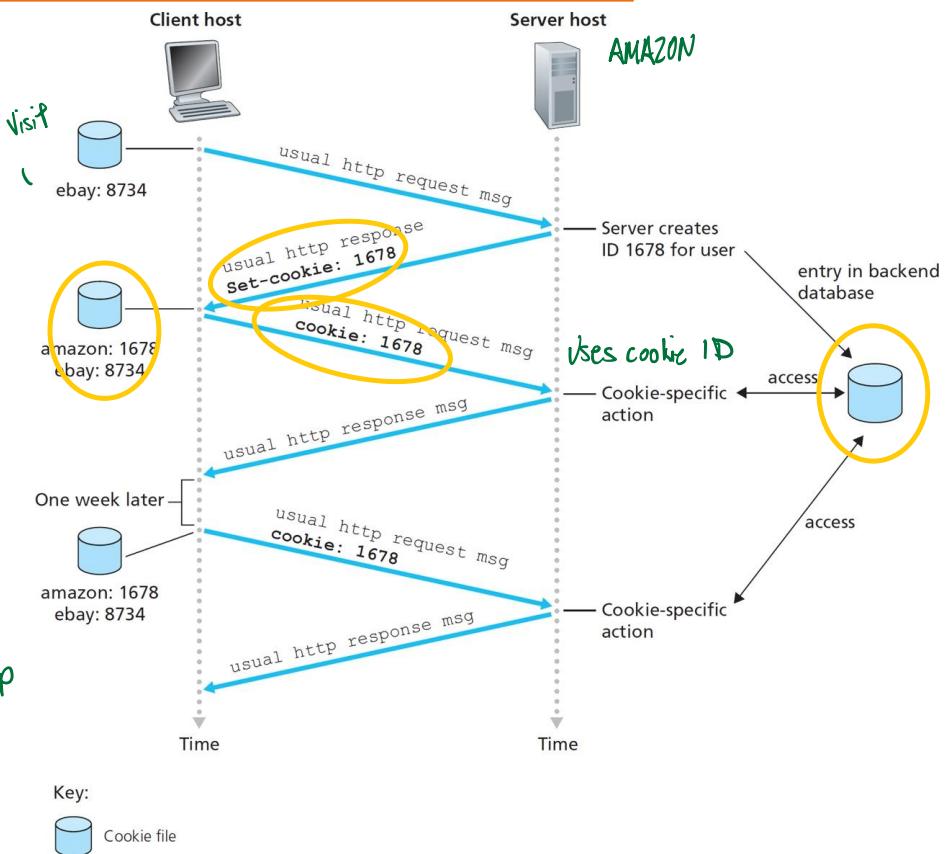
HTTP: User-Server State (Cookies)

- Cookies have four components:

- Cookie header line of HTTP **response** message
- Cookie header line in **next** HTTP **request** message
- Cookie file kept on user's host, managed by user's browser
- Back-end database at Web site

What are cookies? They provide a state for HTTP
(stateless protocol)

User deletes cookies - "new user" next visit



HTTP protocol remains stateless - state kept by
→ Endpoint
→ Cookies.

HTTP State: Cookies

- How to keep **state**?
 - Protocol endpoints: Maintain state at sender/receiver over multiple transactions
 - Cookies: HTTP messages carry state
- Cookies can be used for
 - Authorization
 - Shopping carts
 - Recommendations
 - User sessions state (e.g. web email)
- Cookies permit sites to learn a lot about you – Privacy Concerns?!

GDPR (EU General Data Protection Regulation) and cookies

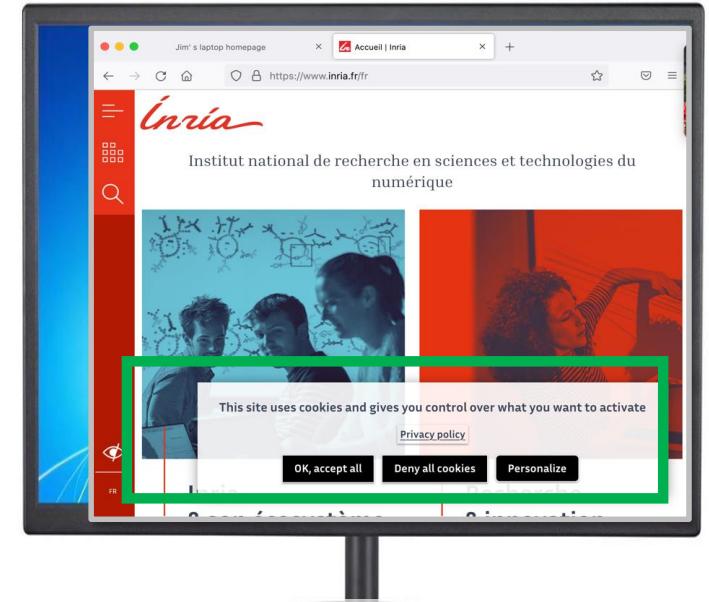
“Natural persons may be associated with online identifiers [...] such as internet protocol addresses, cookie identifiers or other identifiers [...].

This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them.”

GDPR, recital 30 (May 2018)

When cookies can identify an individual, cookies are considered personal data, subject to GDPR personal data regulations.

Basically - when more than basic cookies are needed - must ask



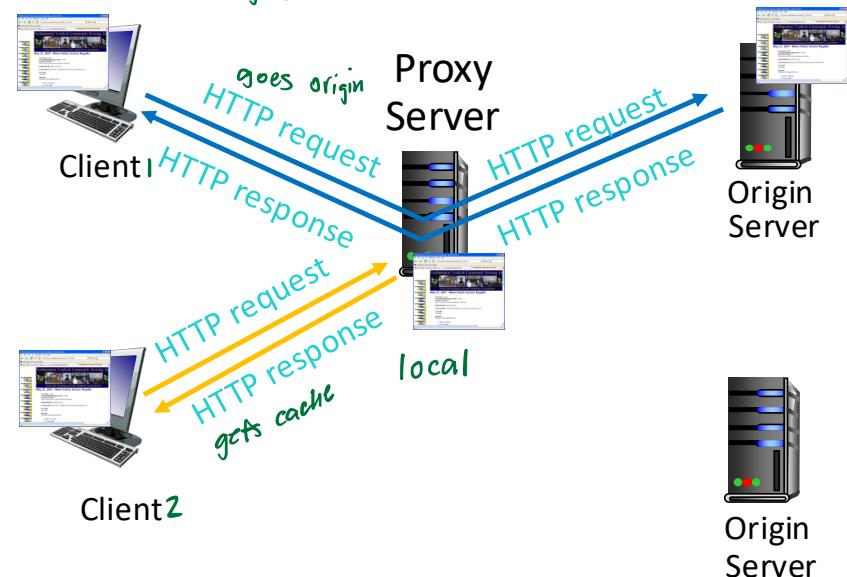
User has explicit control over whether or not cookies are allowed

Web Caches (Proxy Server)

What if many users request same? Cache page locally instead of ↗

Goal: Satisfy client request without involving origin server

- **User sets browser:** Web accesses via cache
- Browser sends all HTTP requests to cache
 - Object in cache: Cache returns object
 - Otherwise cache requests object from origin server, then returns object to client
- Typically cache is installed by ISP
(University, Company, Residential ISP)



Web Caching

- Cache acts as both client and server
 - Server for original requesting client
 - Client to origin server

Why Web caching?

- Reduce response time for client request
- Reduce traffic on an institution's access link
- Enables content providers to cost-effectively deliver content
(same for P2P file sharing)

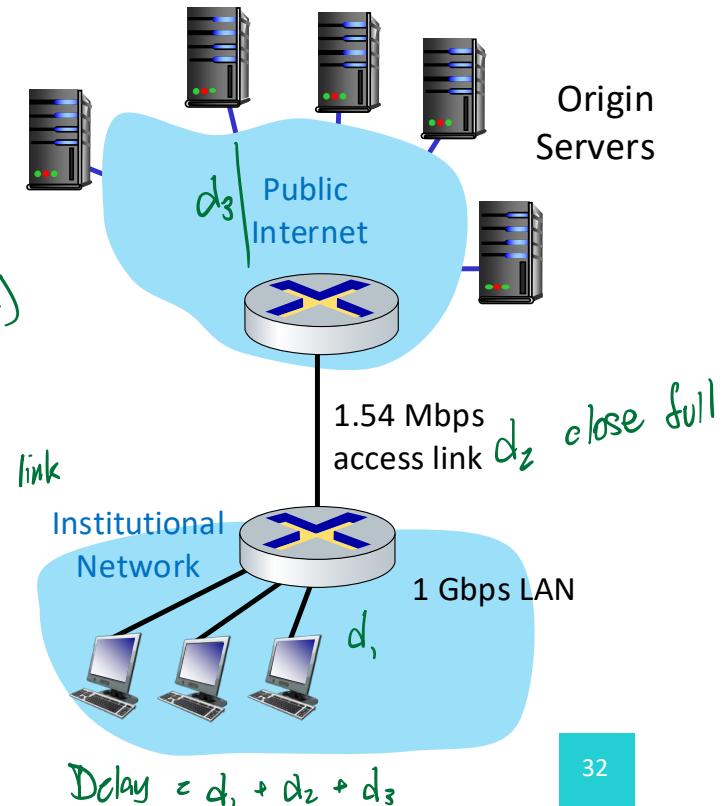
Web Caching Example

Assumptions:

- Access link rate: **1.54 Mbps**
- Average object size: **100K bits**
- Average request rate from browsers to origin servers: **15/sec**
- RTT from institutional router to any origin server: **2 sec**

Consequences:

- Average data rate to browsers: **1.50 Mbps** (object size \times request rate)
 - LAN utilization: 0.0015 ✓ (Rate / LAN) L/R
 - Access link utilization: 0.97 → Problem La/R reaching capacity of link
 \hookrightarrow Data Rate / Access Rate
- Total delay = Internet delay + access delay + LAN delay
= 2 sec + minutes + uSecs



Web Caching Example

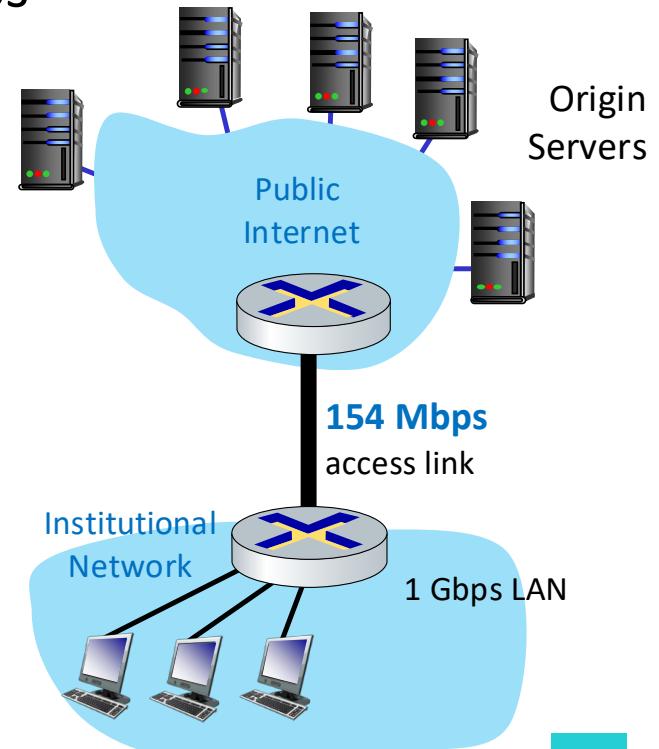
Solution (1): Access link rate change to 154Mbps

Consequences:

- LAN utilization: 0.0015
- Access link utilization: 0.0097
- Total delay = Internet delay + access delay + LAN delay

$$= 2 \text{ sec} + \cancel{\text{minutes}} + \cancel{\text{usecs}}$$

msecs



Cost: Increased access link speed is not cheap!

Web Caching Example

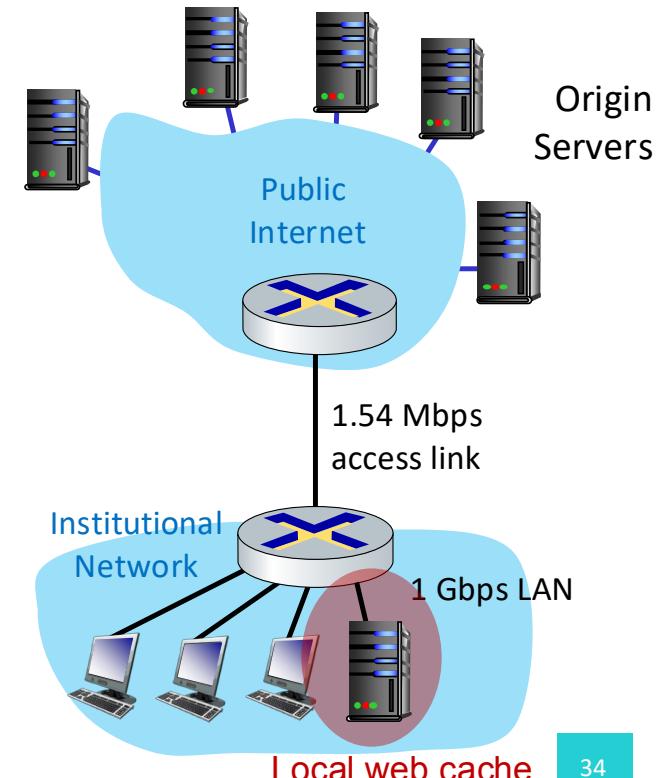
Solution (2) : Install local cache

Consequences

- Suppose cache hit rate is 0.4
 40% requests satisfied at cache
 60% requests satisfied at origin
- Access link utilization: 60% of requests use access link
$$= 0.6 \times 1.5 / 1.54 = 0.58$$

Data Rate to Browsers
- Data rate to browsers over access link
Total delay = 0.6 (2.0) + 0.4 (\sim msecs) = \sim 1.2 secs

Cost: Web cache is not expensive!



What if Server changes page while proxy is giving to client?

↓
client of origin

Conditional Get

Goal: Don't send object if cache has up-to-date **cached** version

Use HEAD command in - retrieve header information

- Cache: Specify date of cached copy in HTTP request

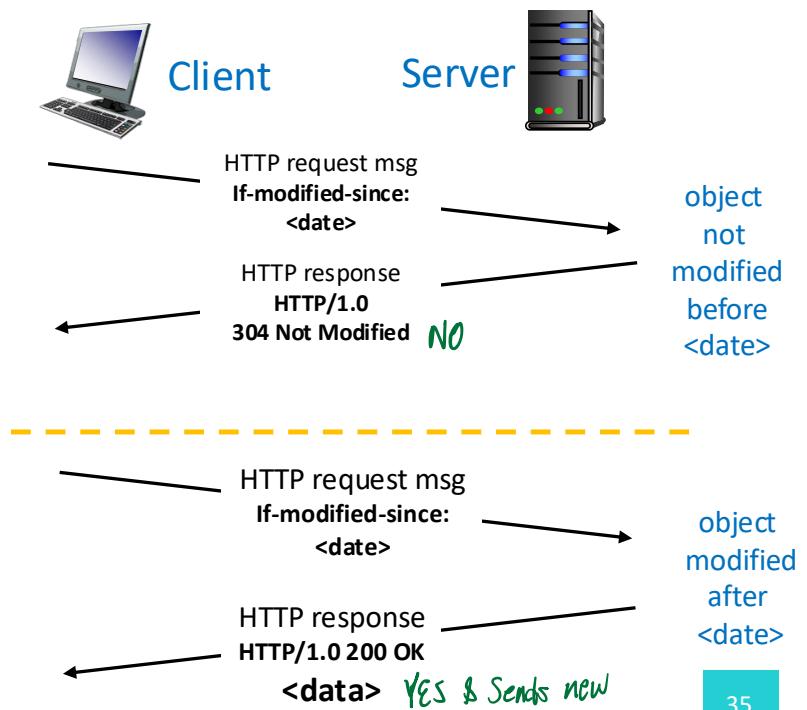
If-modified-since: <date>

↳ only gets if modified since < ?>

- Server: Response contains no object if cached copy is up-to-date:

HTTP/1.0 304 Not Modified

Essentially → (PROXY)
Client checks if modified
↳ Yes → update cache → send client
↳ No → send cache to client



HTTP/2

Goal: Decreased delay in multi-object HTTP requests

HTTP1.0 [RFC 1945, 1996]: Uses non-persistent HTTP

↗ Allow persistent

HTTP1.1 [RFC 2616, 1997]: Introduced **multiple, pipelined GETs over single TCP connection**

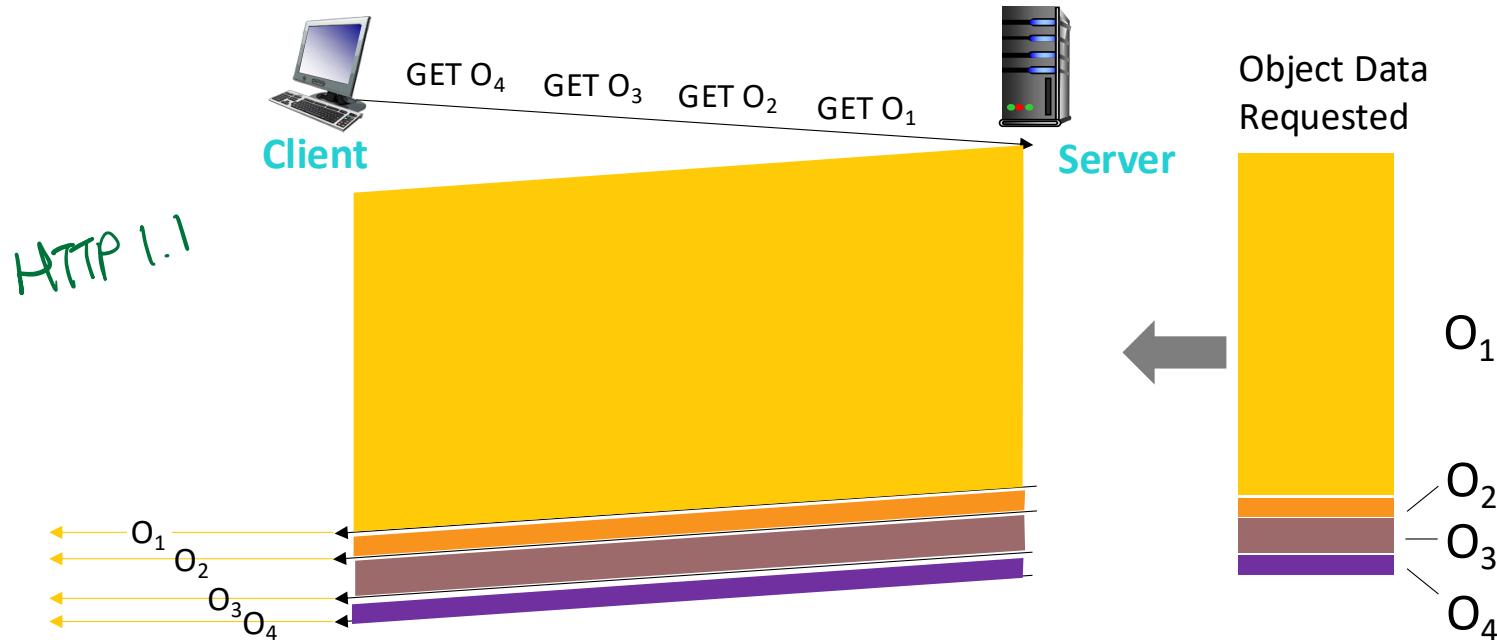
- Server responds First-Come-First-Served (FCFS) to GET requests
- With FCFS, small object may have to wait for transmission behind large object(s)
This is called **Head-Of-Line (HOL) Blocking** Unideal
- Retransmitting lost TCP segments stalls object transmission

HTTP/2 [RFC 7540, 2015]: Increased flexibility at *server* in sending objects to client:

- Methods, status codes, most header fields unchanged from HTTP 1.1
- Transmission order of requested objects based on client-specified priority (not necessarily FCFS)
- Push unrequested objects to client Push potential needs to ↑ speed
- Divide objects into frames, schedule frames to mitigate HOL blocking

Framing → resolves HOL issue

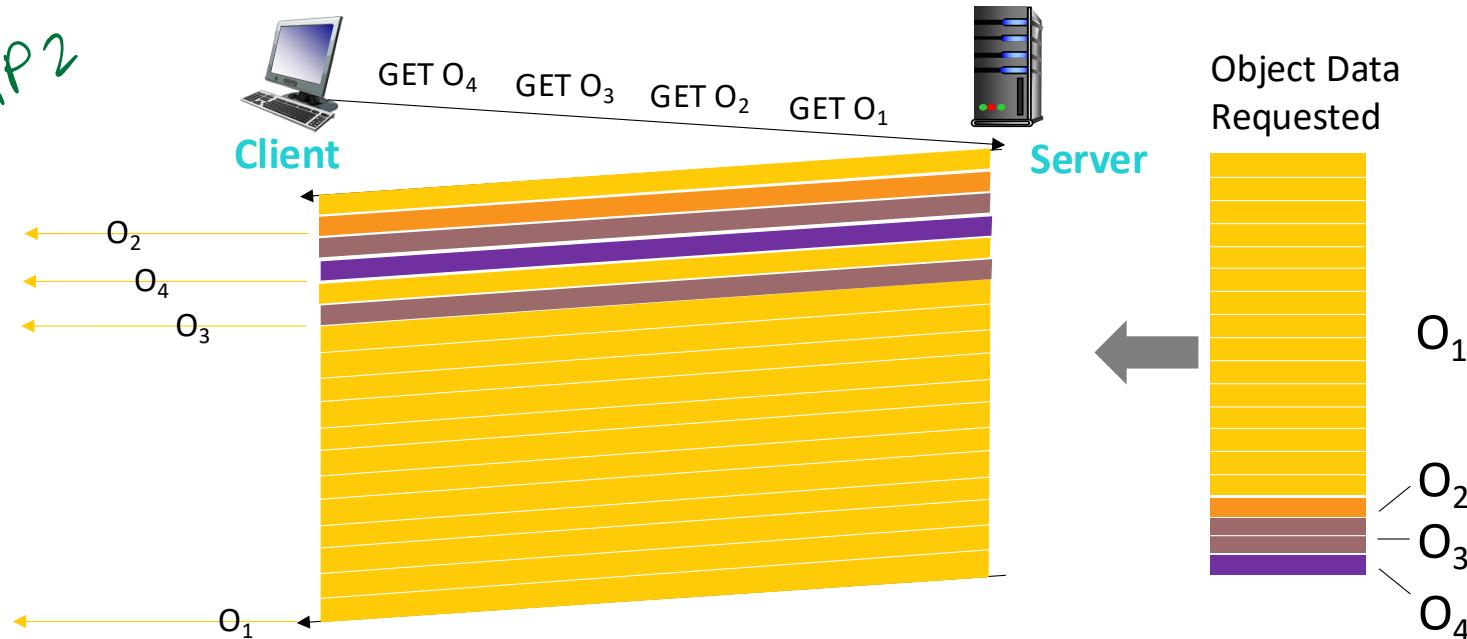
HTTP/2: Mitigating HOL blocking



HTTP 1.1: Client requests 1 large object (e.g., video file, and 3 smaller objects)
Objects delivered in order requested: O_2 , O_3 , O_4 wait behind O_1

HTTP/2: Mitigating HOL blocking

HTTP 2



HTTP/2: Objects divided into frames, frame transmission interleaved

O₂, O₃, O₄ delivered quickly, O₁ slightly delayed

1 frame from every object & repeat
ROUND ROBIN

HTTP/2 to HTTP/3

- HTTP/2 over single TCP connection means:
 - Recovery from packet loss still stalls all object transmissions
 - As in HTTP 1.1 browsers have incentive to open multiple parallel TCP connections to reduce stalling, increase overall throughput
 - No security over vanilla TCP connection
 - HTTP/3 adds security, per object error & congestion-control (more pipelining) over UDP
- HTTP & **QUIC**

Web & HTTP Summary

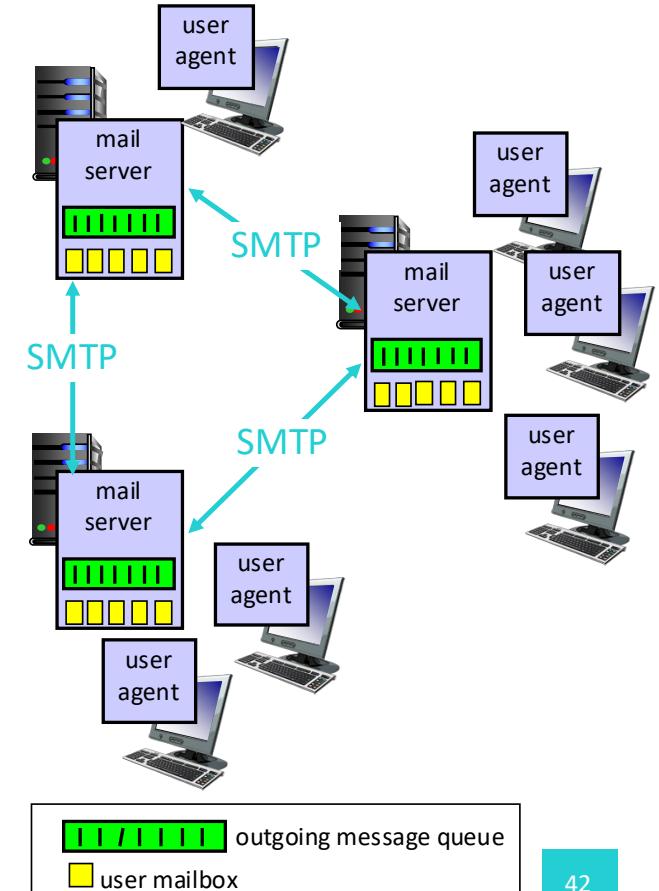
- Messages: Request, Response
 - Methods: GET, POST, HEAD, PUT, DELETE
 - Status Codes and Phrases
- Persistent versus Non-Persistent Connection
- Cookies and State Information
- Caching and Proxy Servers
- Head Of Line Blocking and HTTP/2
- HTTP/2 to HTTP/3



Email

Electronic Mail

- Major components
 - User agents
 - Reading, composing, editing mail messages
 - Mail client examples: Outlook, Thunderbird, iPhone mail client
 - Outgoing and incoming messages stored on server
 - Mail servers
 - Mailbox: Contains incoming mails for user
 - Mail Queue: Contains outgoing (to be sent) mails
 - Simple Mail Transfer Protocol: SMTP
 - Client: Sending mail server
 - Server: Receiving mail server

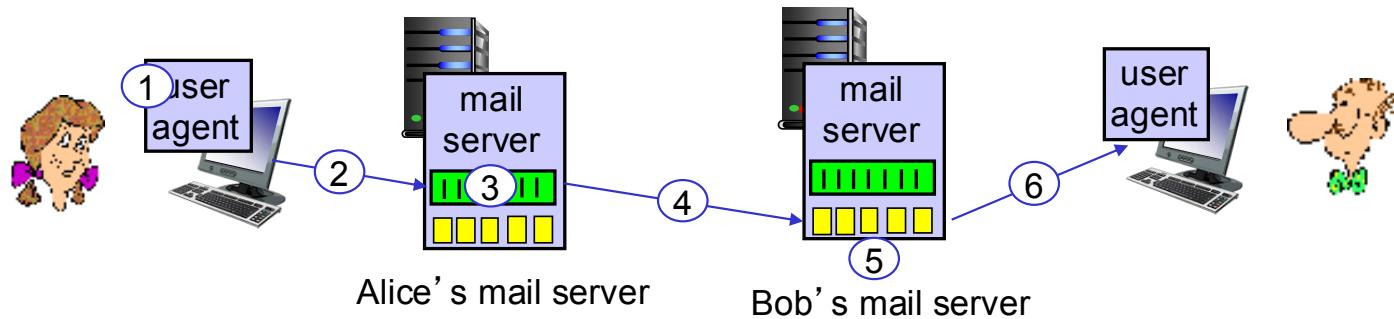


Electronic Mail: SMTP (RFC 5321)

- Uses TCP to reliably transfer email message from client to server (port 25)
- Direct transfer: Sending server to receiving server *servers talk each other*
- Three phases of transfer
 - 1 • Handshaking (greeting)
 - 2 • Transfer of messages
 - 3 • Closure
- Command/response interaction (like HTTP)
 - **Commands:** ASCII text
 - **Response:** Status code and phrase
- Messages must be in 7-bit ASCII

Electronic Mail

1. Alice uses UA to compose message to bob@someschool.edu
2. Alice's UA sends message to her mail server. message placed in message queue.
3. Client side of SMTP opens TCP connection with Bob's mail server
4. SMTP client sends Alice's message over the TCP connection
5. Bob's mail server places the message in Bob's mailbox
6. Bob invokes his user agent to read message



Sample SMTP Interaction

```
S: 220 hamburger.edu      } HANDSHAKE
C: HELO crepes.fr
S: 250 Hello crepes.fr, pleased to meet you
C: MAIL FROM: <alice@crepes.fr>
S: 250 alice@crepes.fr... Sender ok
C: RCPT TO: <bob@hamburger.edu>
S: 250 bob@hamburger.edu ... Recipient ok
C: DATA
S: 354 Enter mail, end with "." on a line by itself SENDS
C: Do you like ketchup?
C: How about pickles?    } RECEIVES
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 hamburger.edu closing connection
```

Try SMTP

- **telnet servername 25**
- See 220 reply from server
- Enter HELO, MAIL FROM, RCPT TO, DATA, QUIT commands
above lets you send email without using email client (reader)

Set up and test that for an exchange server:

<https://learn.microsoft.com/en-us/exchange/mail-flow/test-smtp-telnet>

SMTP

- SMTP uses persistent connections
- SMTP requires message (header & body) to be in 7-bit ASCII
- SMTP server uses CRLF . CRLF to determine end of message

Comparison with HTTP:

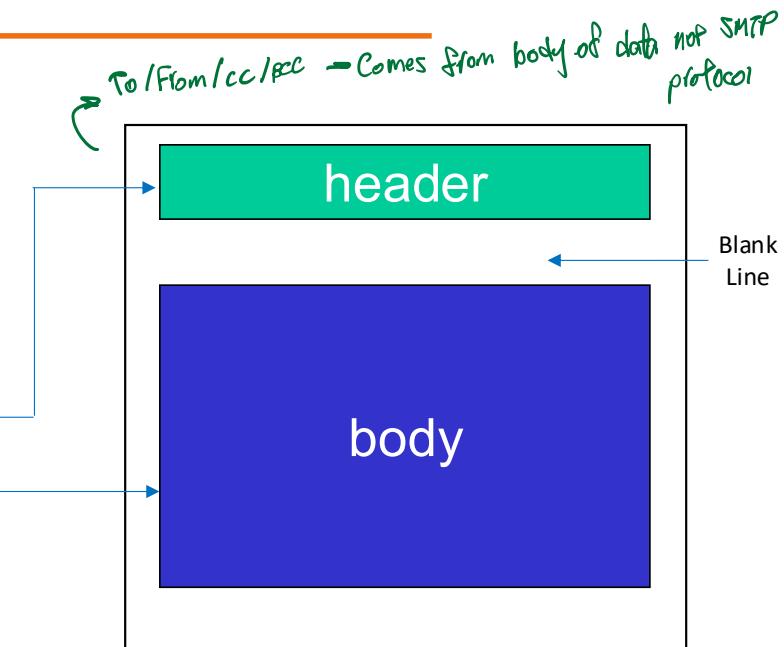
SMTP	HTTP
<i>Push</i>	<i>Pull</i>
<i>Multiple Objects sent in multipart message</i>	<i>Each object encapsulated in its own response message</i>
<i>Both have ASCII command/response interaction & status codes</i>	

Mail Message Format

SMTP [RFC 5321]: Protocol for exchanging e-mail messages

RFC 5322 defines syntax for e-mail message itself

- Header lines, e.g.,
 - To:
 - From:
 - Subject:
 - These lines, within the body of the email message area, are different from SMTP MAIL FROM:, RCPT TO: commands
- Body: The “message” in ASCII characters only



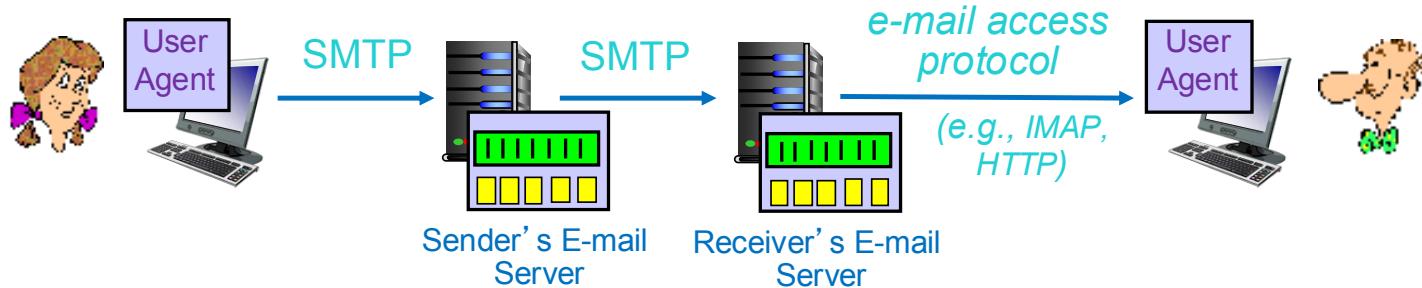
Mail Access Protocols

SMTP: Delivery/storage to receiver's server

Pull protocol
↳ Request comes client

Mail access protocol: Retrieval from server

- **POP:** Post Office Protocol [RFC 1939]: Authorization, download
- **IMAP:** Internet Mail Access Protocol [RFC 3501]: More features, including manipulation of stored messages on server
- **HTTP:** Gmail, Hotmail, Yahoo! Mail, etc.



POP3

Authorization phase

- Client commands
 - User: Declare username
 - Pass: Password
- Server Responses
 - OK
 - ERR

Transaction Phase

- Client
 - list: list message numbers
 - retr: retrieve messages by number
 - dele: delete
 - quit

S: +OK POP3 server ready
C: user bob user defines itself
S: +OK server OR
C: pass hungry
S: +OK user successfully logged on

C: list
S: 1 498
S: 2 912
S: .
C: retr 1 RETRIEVE
S: <message 1 contents>
S: .
C: dele 1
C: retr 2
S: <message 1 contents>
S: .
C: dele 2
C: quit
S: +OK POP3 server signing off

POP3 & IMAP

POP3 (RFC 1939)	IMAP (RFC 3501)
<i>Download & delete mode:</i> <i>User cannot re-read e-mail if he changes client</i>	<i>Keeper all messages in one place: at server</i>
<i>Download & keep mode:</i> <i>Copies of messages on different clients</i>	
<i>No mail folder organization</i>	<i>Allows user to organize messages in folders</i>
<i>POP3 is stateless across sessions</i>	<i>Keeps user state across sessions:</i> <i>Names of folders</i> <i>Mappings between message IDs & folder name</i>
<i>Both Mail Access Protocols</i>	

Mail Security Issues

- Basic electronic mail is not secure
 - No encryption (Plain Text)
 - Eavesdropping and interception is easy
 - Spamming
 - Source of Malware Delivery
 - Viruses, worms, Trojan horses, documents with destructive macros, and malicious code
 - No Source Verification
 - Spoofing sender address

Servers receiving email

- ↳ verify sending server exists
- ↳ verify user exists on sending server

Learn More:

Certificates & Digital Signatures

PGP: Pretty Good Privacy - what to do upon fails

SPF: Sender Policy Framework - verifies sender

DKIM: DomainKeys Identified Mail - public keys / private keys

DMARC: Domain-based Message Authentication Reporting and Conformance

↳ Receiving servers deny messages without certain measures

DNS



One of the most important

DNS -

Internet hosts and routers:

- IP address used for addressing datagrams
- Name: e.g., www.yahoo.com

IP address → What's used send stuff

Analogy: Name and ID number for people / Name & Postal Code

Question: How to map between the IP address and name?

Domain Name System:

- **Distributed database** implemented in hierarchy of many **name servers**
- **Application-layer protocol:** Hosts, name servers communicate to **resolve names** (address/name translation)
- Core Internet function, implemented as application-layer protocol
- Complexity at network's edge

↳ find IP based on name

DNS Services

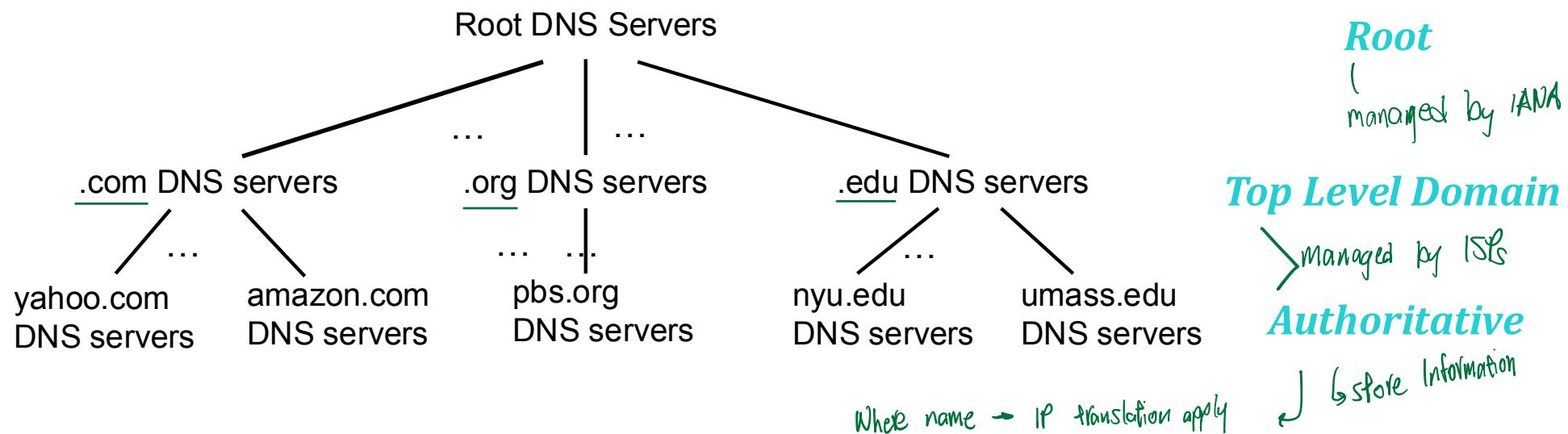
- Hostname to IP address translation - *Google 88.8.8.8.
Cloudflare 1.1.1.1.*
- Host aliasing (canonical, alias names) - *Multiple names same server*
- Mail server aliasing
- Load distribution (Replicated Web Servers: many IP addresses correspond to one name)

DNS: Why Not Centralized?

- Single point of failure - Important cause translates names to IP Addresses
- Traffic volume
- Distant centralized database
- Maintenance

Does not Scale!

DNS: A Distributed Hierarchical Database



DNS

↗ Networks edge (complexity)

Local DNS Name Server *Helps navigate above architecture*

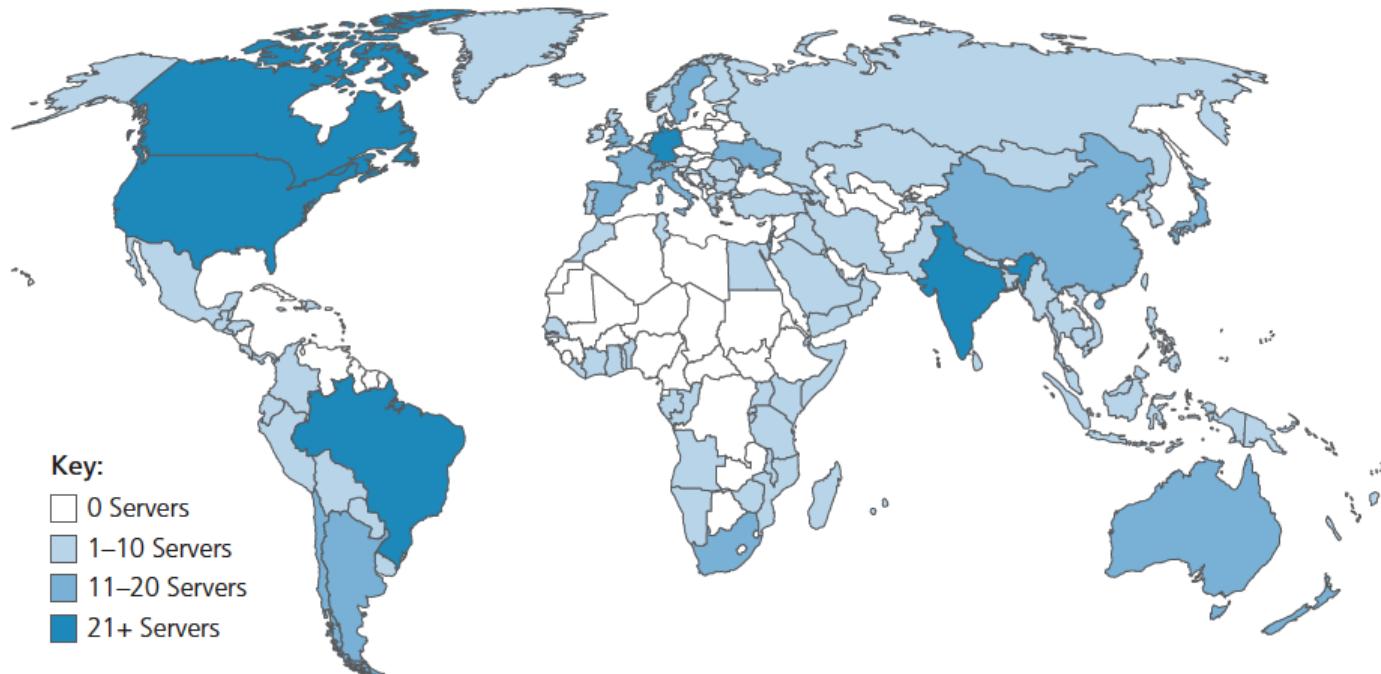
- Does not strictly belong to hierarchy
- Each ISP (residential ISP, company, university) has one
 - also called **default name server / Resolver**
- When host makes DNS query, query is sent to its local DNS server
 - Has local cache of recent name-to-address translation pairs
(but may be out of date!)
 - Acts as proxy & forwards query into hierarchy

DNS

Root Name Servers

- Contacted by local name server that can not resolve name
- Root name server:
 - Contacts authoritative name server if name mapping not known
 - Gets mapping
 - Returns mapping to local name server

DNS



13 logical root name servers worldwide each server replicated many times

iana.org/domains/root/servers

DNS

Refers your query to corresponding AUTHORITATIVE level

Top-level domain (TLD) servers ↑

- Responsible for com, org, net, edu, aero and all top-level country domains like uk, fr, ca

Authoritative DNS servers

- Providing authoritative hostname to IP mappings for named hosts
- Can be maintained by organization or service provider

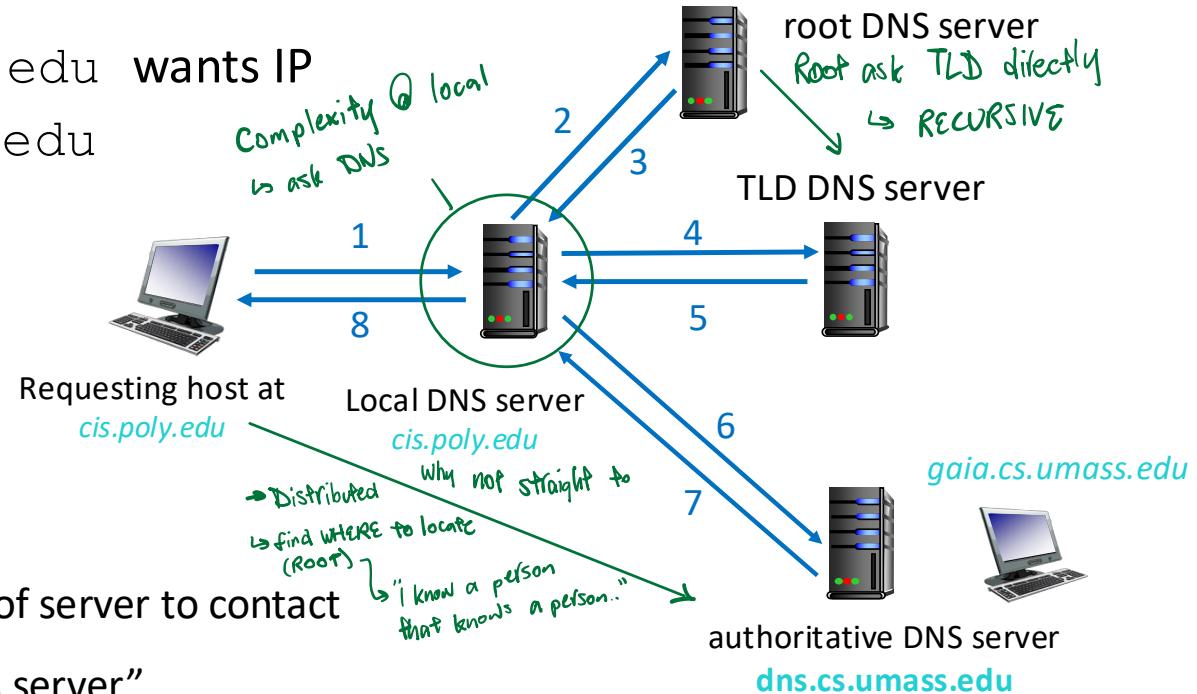
DNS: Name Resolution

Example: Host at `cis.poly.edu` wants IP address for `gaia.cs.umass.edu`

ITERATING

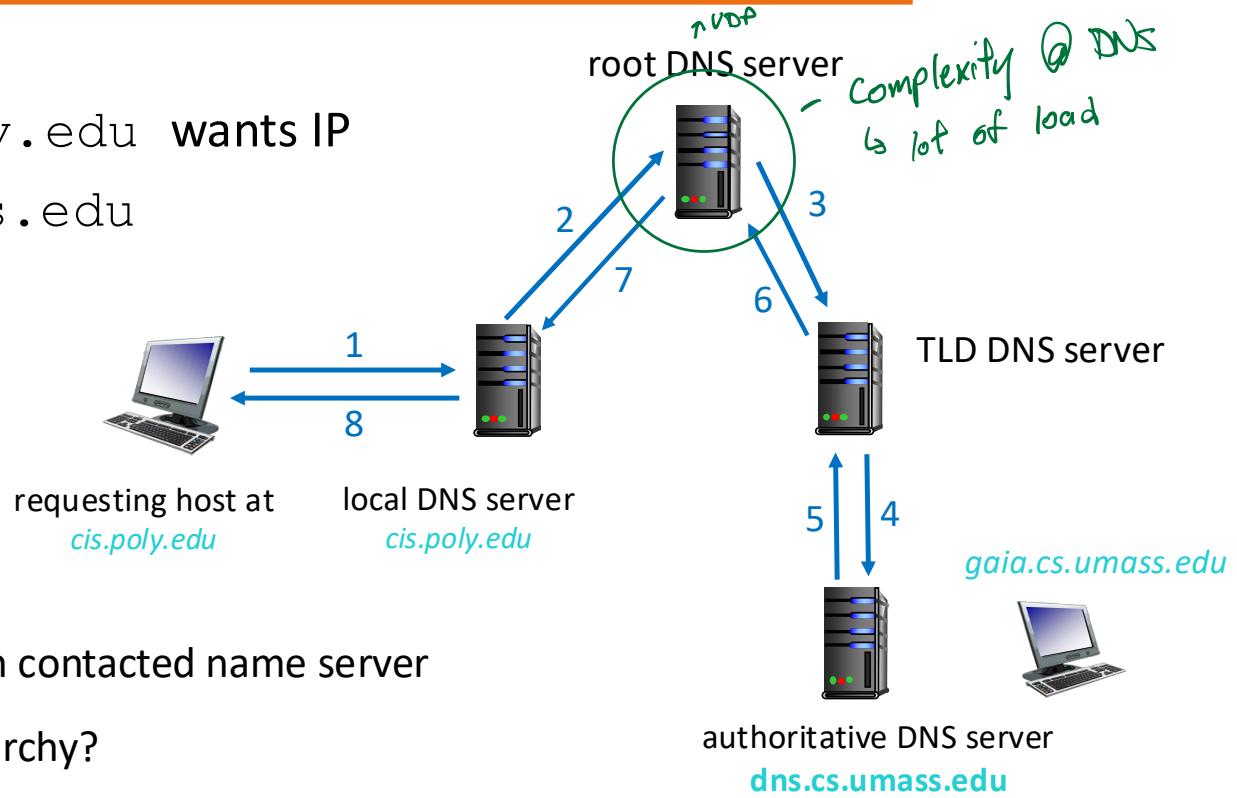
Iterated query

- ① Contact local DNS server
- Contacted server replies with name of server to contact
- OR ② Ask ROOT DNS server
- "I don't know this name, but ask this server"
- OR ③ Ask TLD DNS server
- OR ④ Ask authoritative DNS server



DNS: Name Resolution

Example: Host at `cis.poly.edu` wants IP address for `gaia.cs.umass.edu`



Recursive query

- Puts burden of name resolution on contacted name server
- Heavy load at upper levels of hierarchy?

Some servers disable recursive to avoid excess cache

DNS: Caching & Updating Records

- Once (any) name server learns mapping, it **caches** mapping
 - Cache entries timeout (disappear) after some time (TTL)
 - TLD servers typically cached in local name servers
 - Thus root name servers not often visited
 - ↳ Root DNS usually strong
 - ↳ local DNS often weak → cache poisoning
- Cached entries may be **out-of-date** (best effort name-to-address translation!)
 - If name host changes IP address, may not be known Internet-wide until all TTLs expire
- Update/notify mechanisms proposed IETF standard
 - RFC 2136

DNS Records

- **DNS:** Distributed database storing resource records (**RR**)

domain name

A/NS/CNAME/MX

RR format: (name, value, type, ttl)

type=A : name is hostname

value is IP address

type=NS : name is domain (e.g., [foo.com](#))

value is hostname of authoritative name server for this domain

type=CNAME : name is alias name for some “canonical” (the real) name

E.g. [www.ibm.com](#) is really [servereast.backup2.ibm.com](#)

value is canonical name

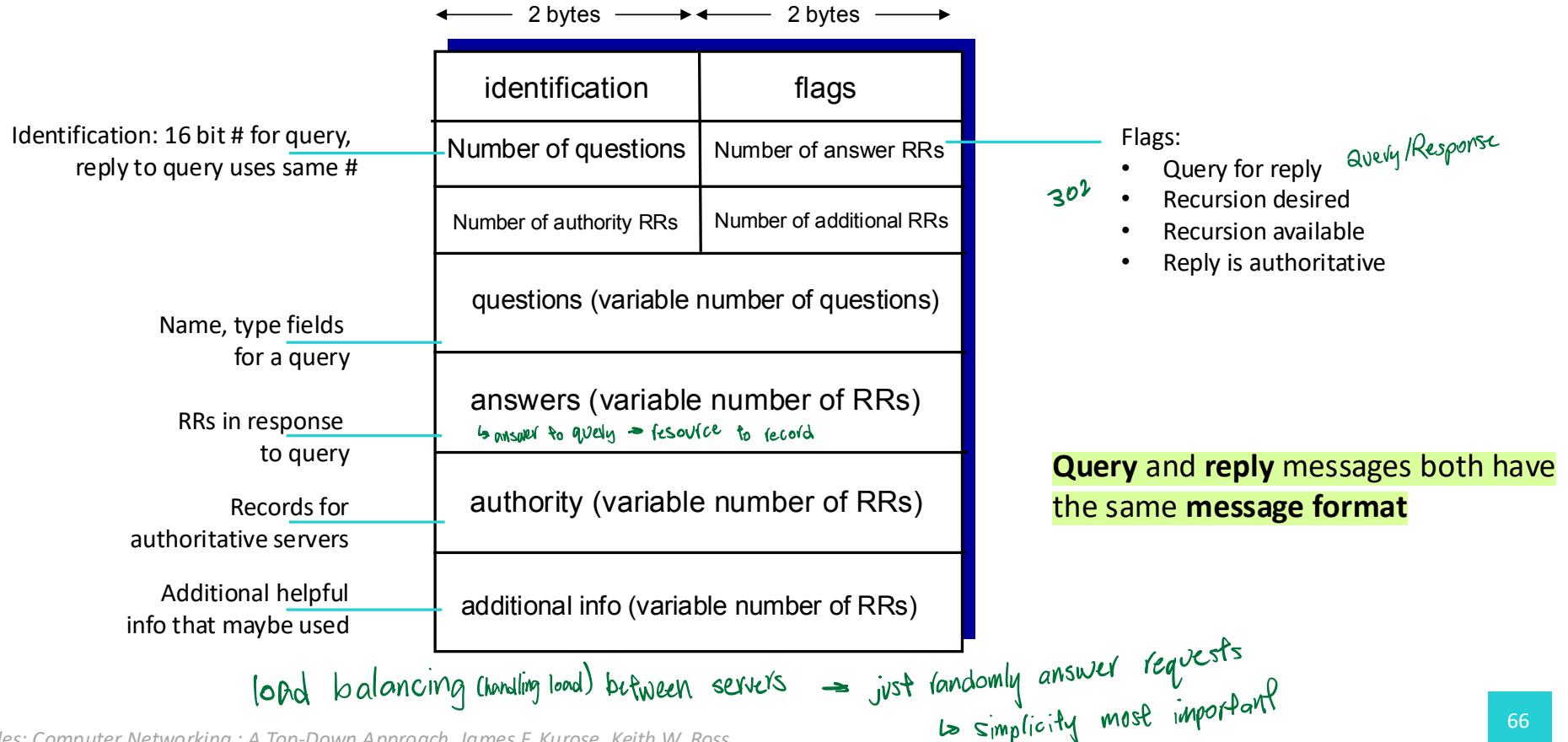
type=MX

value is the name of mailserver associated with name

value

time to live
seconds that record could
be cached before refresh/
discard.

DNS Messages



Inserting Records into DNS

Example: New startup Network Utopia

- Owner
 - Create authoritative server type A record for server and type MX record for mail server
 - Register name networkutopia.com at **DNS registrar** (essentially phone book)
 - Owner provides names and IP addresses of authoritative name servers (primary and secondary)
- **DNS registrar** (list available at internic.net)
 - Inserts NS and A type RRs into .com **TLD server**:

↗ domain has the name ↗
(networkutopia.com, dns1.networkutopia.com, **NS**)
(dns1.networkutopia.com, 212.212.212.1, **A**)
 ↗ name has the IP ↗
(networkutopia.com, dns2.networkutopia.com, **NS**)
(dns2.networkutopia.com, 212.212.212.2, **A**)

2 because reliability & load.

Attacking DNS

DDoS attacks

- Bombard root servers with traffic: Not successful to date
Traffic filtering, Local DNS servers cache IPs of TLD servers, allowing root server bypass
- Bombard TLD servers: Potentially more dangerous

Exploit DNS for DDoS (DNS Amplification)

- Send queries with spoofed source address: target IP
- Requires amplification → Request that returns a bunch of information

Redirect or Hijacking attacks

- Man-in-middle: Intercept queries
- DNS poisoning: Send bogus replies to DNS server, which caches



P2P

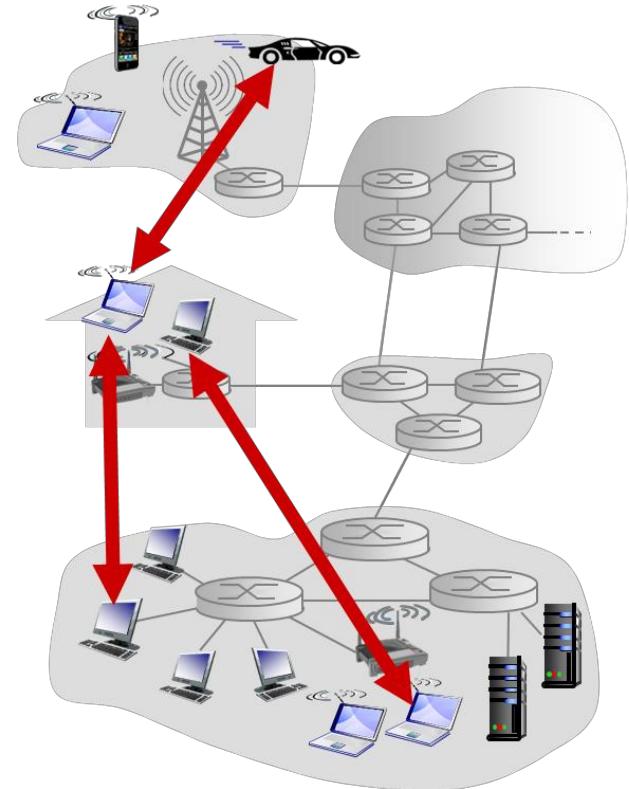


Pure P2P Architecture

- No always-on server
- Arbitrary end systems **directly** communicate
- Peers are **intermittently connected** and change IP addresses

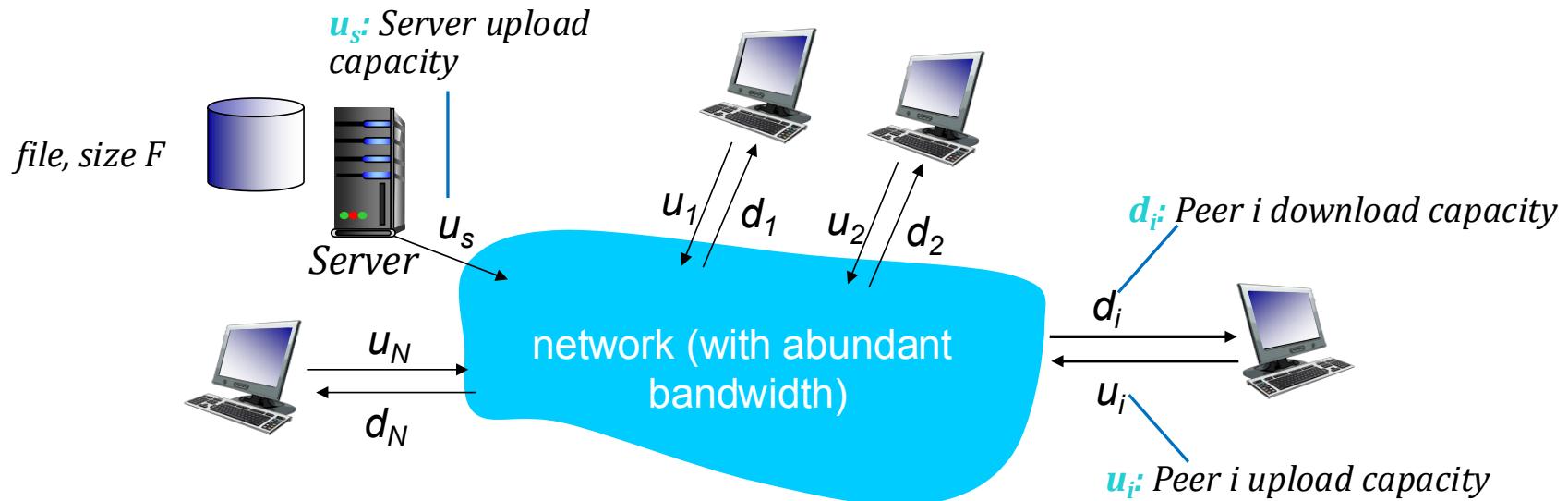
Examples:

- File distribution (BitTorrent)
- Streaming (KanKan)
- VoIP (Skype)



File Distribution

Question: How much time to distribute file (size F) from one server to N receivers?



File Distribution: Client-Server

- **Server transmission:** Must sequentially send (upload) N file copies
 - Time to send one copy: F/u_s
 - Time to send N copies: $N \times F/u_s$
- **Client:** Each client must download file copy
 - Client download time: F/d_i
 - d_{min} = Min client download rate/capacity
 - Min client download time: F/d_{min}

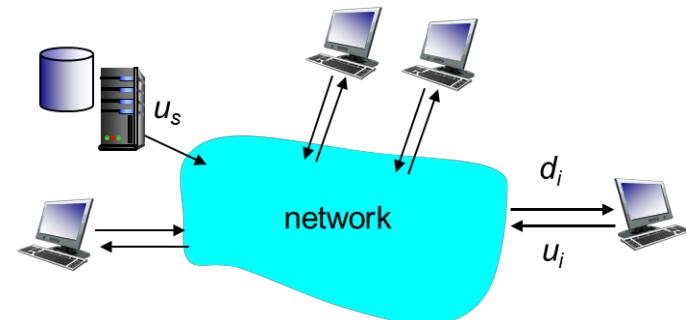
Time to distribute F
to N clients using
client-server approach

$$D_{CS} \geq \max\{NF/u_s, F/d_{min}\}$$

because bottleneck

Time to send copies / time to download slowest

Increases linearly in N



➢ When user downloads, start using as upload

File Distribution: P2P

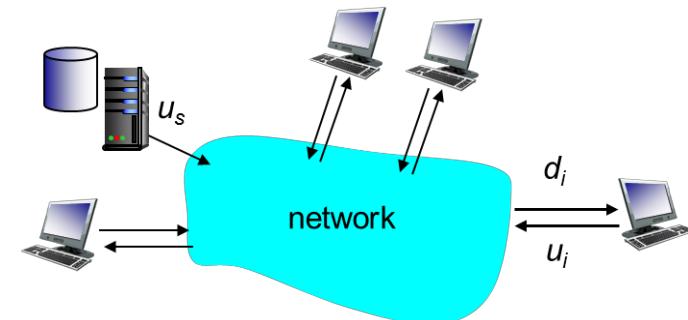
- **Server transmission:** Must upload at least one copy
 - Time to send one copy : F/u_s
- **Client:** Each client must download file copy
 - Min client download time: F/d_{min}
 - Clients as aggregate must download NF bits
- Max upload rate (limiting max download rate) is $u_s + \sum u_i$

N - clients
 F - filesize

Time to distribute F to N clients using P2P approach

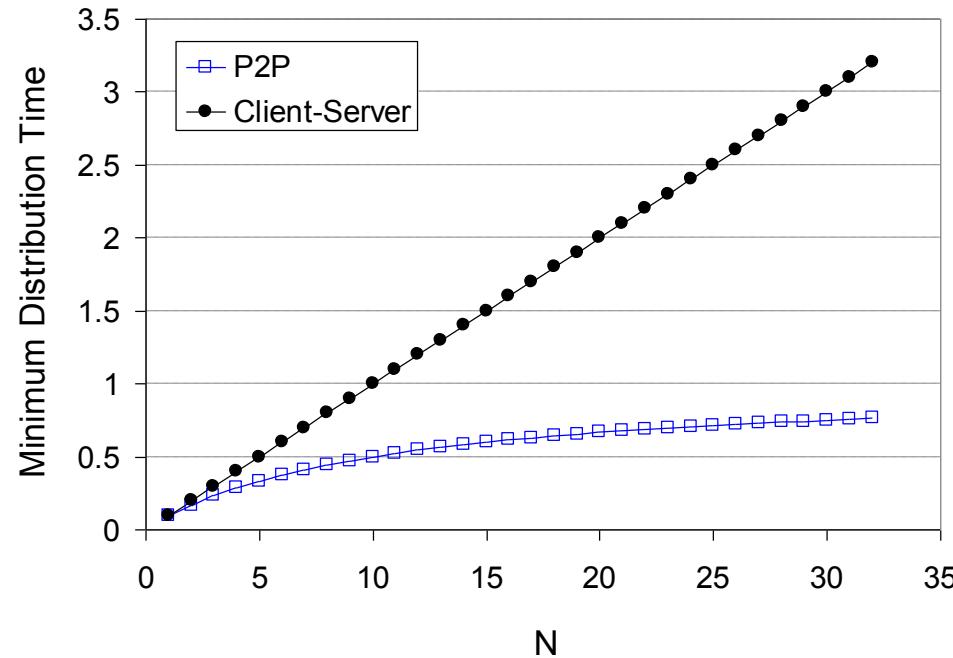
$$D_{P2P} \geq \max\{F/u_s, F/d_{min}, NF/(u_s + \sum u_i)\}$$

Increases linearly in N
but so does this, as each peer brings service capacity



Client-Server vs P2P Example

- Client upload rate = u , $F/u = 1$ hour, $u_s = 10u$, $d_{min} \geq u_s$

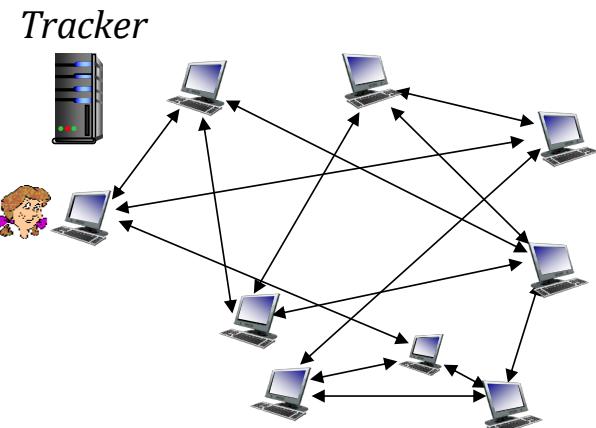


BitTorrent

Request receive chunk

- File divided into 256KB chunks
 - Peers in torrent send & receive file chunks
- Peer joining torrent
 - Has no chunks, but will accumulate them over time from other peers
 - Connects to subset of peers (neighbors)
 - Registers with **tracker** to get list of peers
- While downloading
 - Peer uploads chunks to other peers
 - Peer may change peers with whom it exchanges chunks
- **Churn:** Peers may come and go
 - Once peer has entire file, it may (selfishly) leave or (altruistically) remain in torrent

Peers might not have entire file
AND/OR
capacity to provide to others.



BitTorrent

Requesting File chunks

- At any given time, different peers have different subsets of file chunks
- Periodically, Alice asks each peer for **list of chunks** that they have
- Alice requests **missing chunks** from peers, **rarest first** - available & known when they might leave
 - ↳ Rate = low # peers w/ the chunk
 - ↳ so not reliant if they leave

Sending File chunks

- Alice sends chunks to those four peers currently sending her chunks **at highest rate**
 - Other peers are **choked** by Alice (do not receive chunks from her)
 - Re-evaluate top 4 every 10 secs

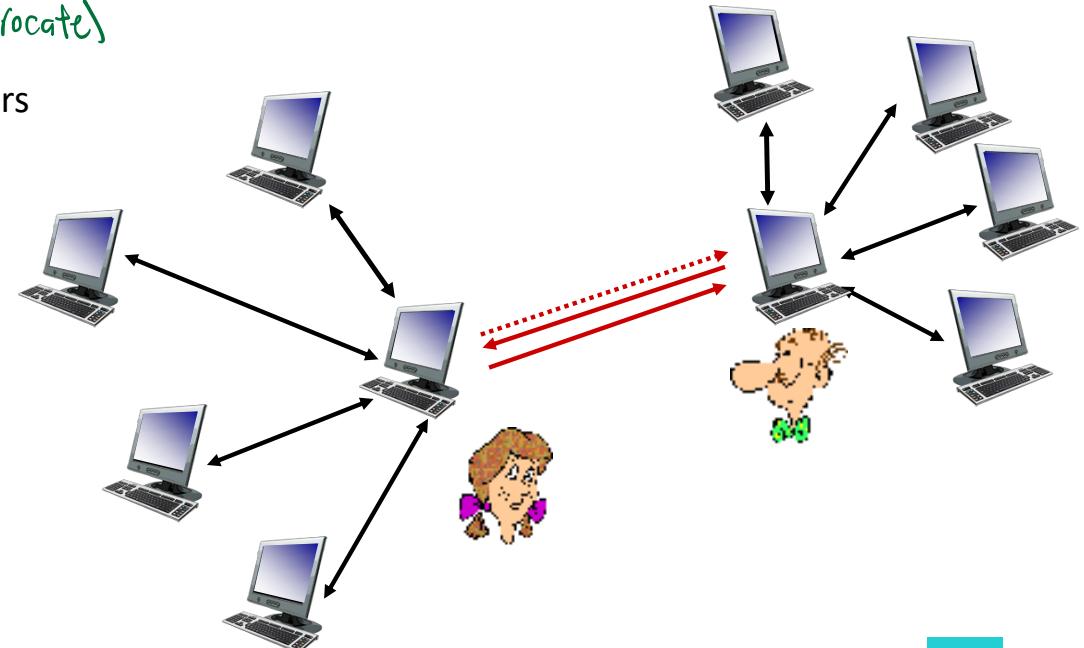
peers provide highest rates

BitTorrent

choke (do not receive chunk)

Tit-for-tat

- Every 30 secs Alice randomly select another peer (Bob)
- Alice **optimistically unchoke** Bob (hopes will reciprocate)
- Alice may become one of Bob's top-four providers
- Bob may reciprocate and therefore become one of Alice's top-four providers





Video Steaming & CDN

Video Streaming & CDN

- Video traffic is a major consumer of Internet bandwidth
large filesize
- Challenges
 - Scale: How to reach ~1B users?
 - Single mega-video server will not work (why?)
 - Heterogeneity
 - Different users have different capabilities (e.g. wired vs mobile; bandwidth rich vs bandwidth poor)

Solution: Distributed Application-Level Infrastructure

30 B devices



Video

- Video: Sequence of images displayed at constant rate
(E.g. 24 images/sec)
- Digital image: Array of pixels (Each pixel represented by bits)
- **Coding:** Using redundancy **within** and **between** images to decrease number of bits used to encode image
 - **Spatial** (within image)
 - **Temporal** (from one image to next)

spatial coding example: instead of sending N values of same color (all purple), send only two values: color value (purple) and number of repeated values (N)



frame i

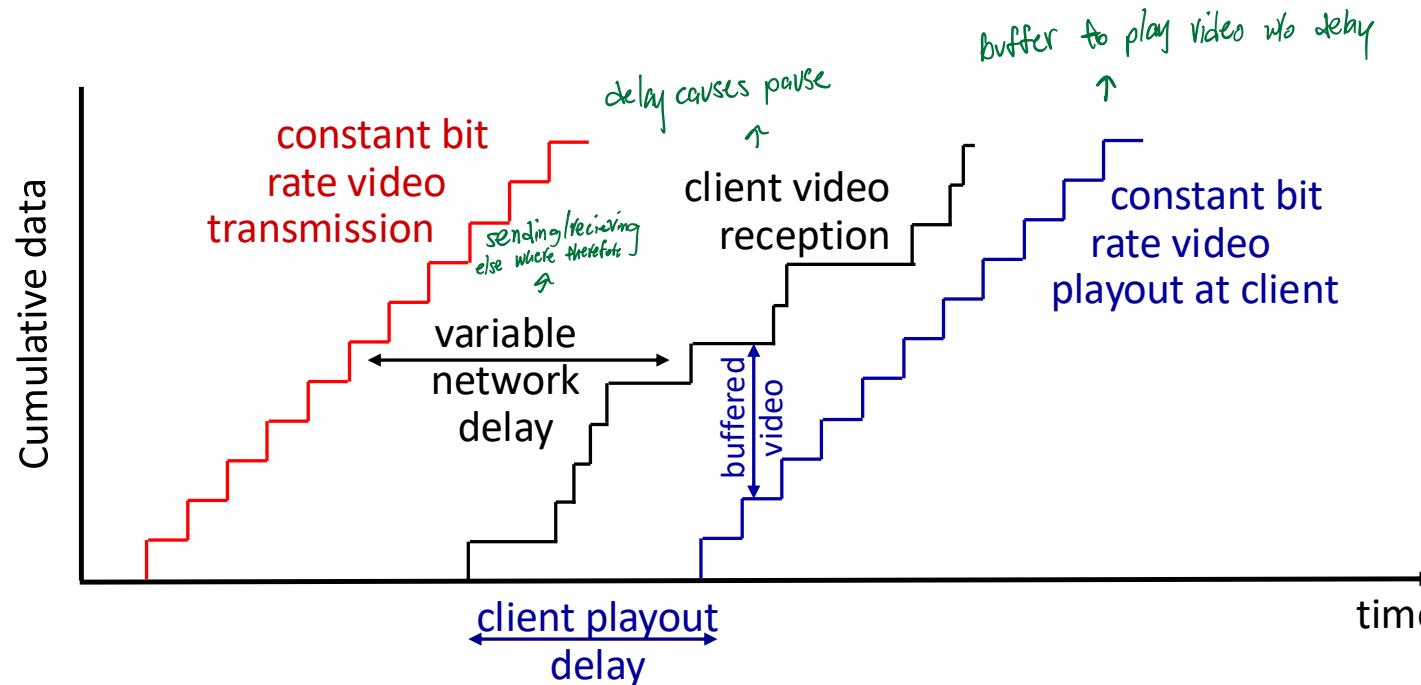


temporal coding example: instead of sending complete frame at $i+1$, send only differences from frame i

Video

- CBR: (**Constant Bit Rate**): Video encoding rate fixed
→ timing & size
- VBR: (**Variable Bit Rate**): Video encoding rate changes as amount of spatial, temporal coding changes
- Examples:
 - MPEG1 (CD-ROM) 1.5 Mbps
 - MPEG2 (DVD) 3-6 Mbps
 - MPEG4 (often used in Internet, < 1 Mbps)

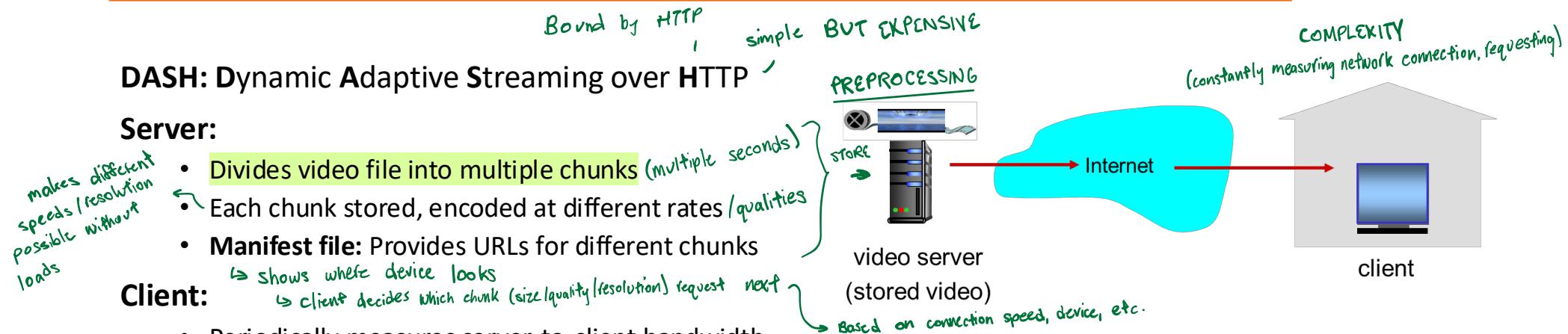
Streaming Video



Client-side buffering and playout delay: Compensate for network-added delay and delay jitter

Streaming Stored Video: DASH

- On TCP uses HTTP



CDN: Content Distribution Networks

→ Do not wanna pay T_1 ISPs

Challenge: How to stream content (from millions of videos) to hundreds of thousands of **simultaneous** users?

Option 1: A single large mega-server

- Single point of failure
- Point of network congestion
- Long path to distant clients
- Multiple copies of video sent over outgoing link

} Downside - Even if scalable, networks clogged
↳ Therefore option 2

This solution **does not scale!**

Option 2: Store/serve multiple copies of videos at multiple geographically distributed sites (**CDN**)

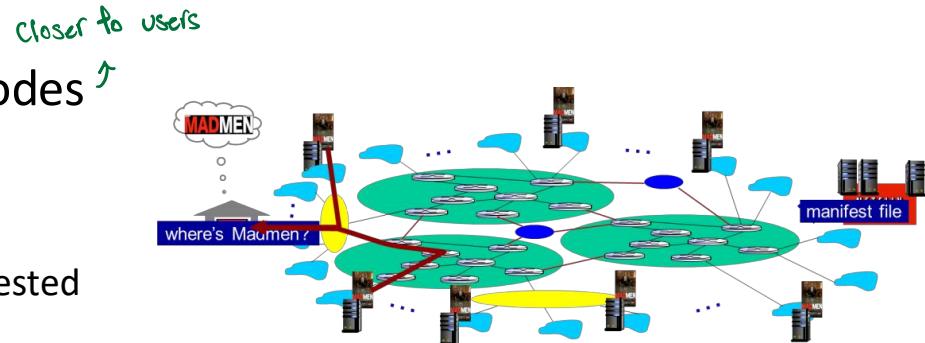
- **Enter deep:** (Access ISPs) Push CDN servers deep into many access networks
 - Close to users
 - Used by Akamai, 1700 locations
- **Bring home:** (IXP) smaller number (10's) of larger **clusters** near (but not within) access networks
 - Used by Limelight

- Closer User / Higher # Servers
- More clusters = lower # servers / further users.

Wouldn't wanna bring too close to users -

CDN

- CDN: Stores copies of content at CDN nodes ↑
- Subscriber requests content from CDN
 - Directed to nearby copy, retrieves content
 - May choose different copy if network path congested



Over the top (OTT): Internet host-host communication as a service

OTT challenges: Coping with a congested Internet

- From which CDN node to retrieve content?
- Viewer behavior in presence of congestion?
- What content to place in which CDN node?

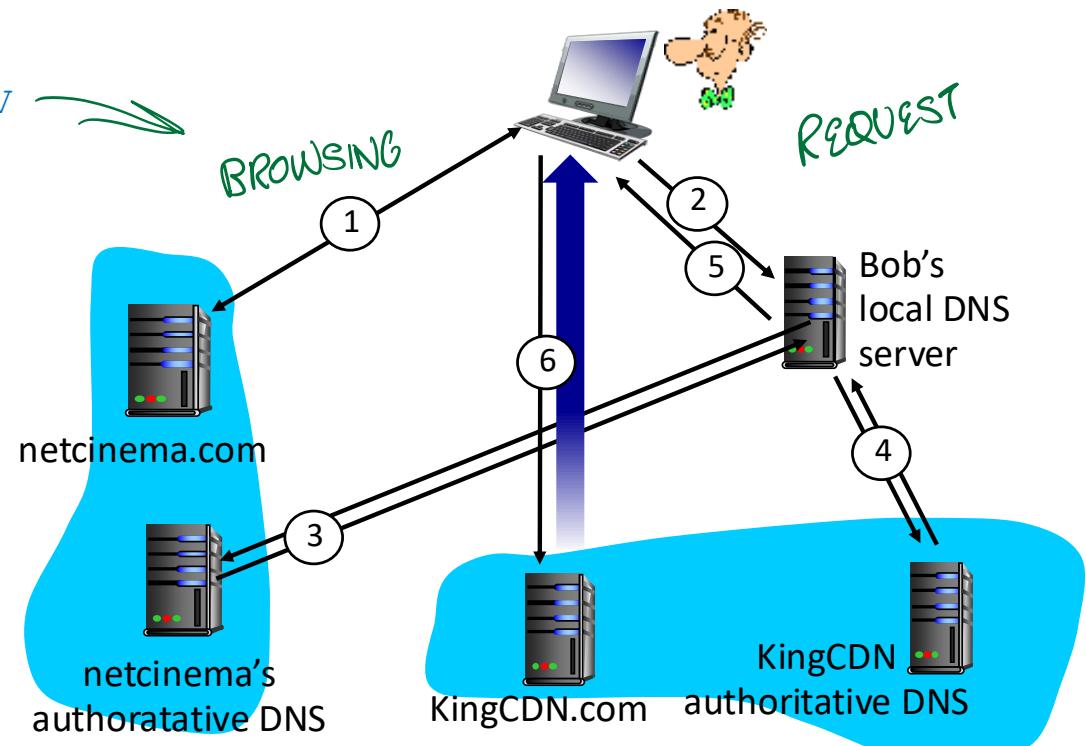
CDN: Example

Bob (client) requests video (Netflix)

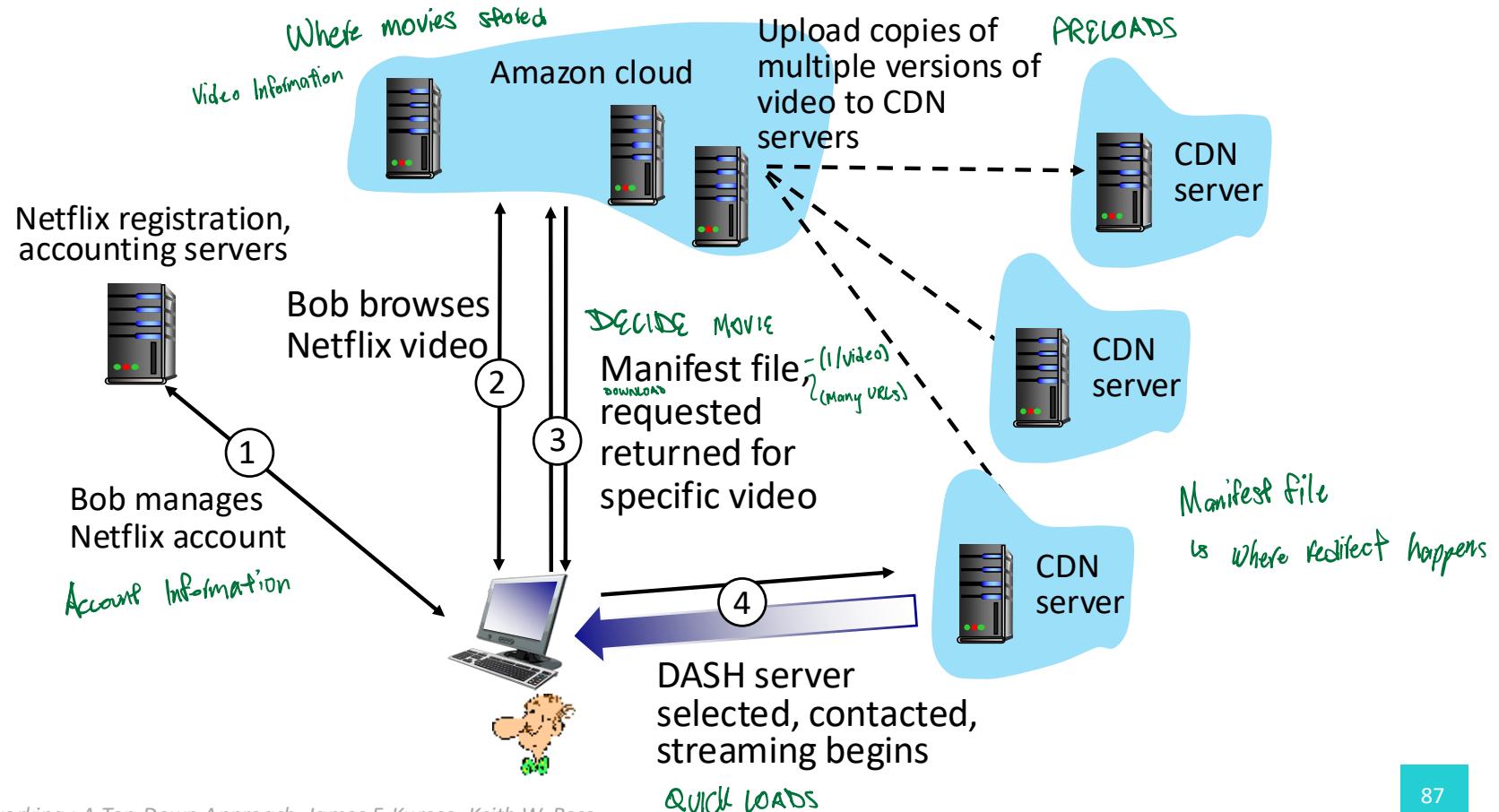
<http://video.netcinema.com/6Y7B23V>

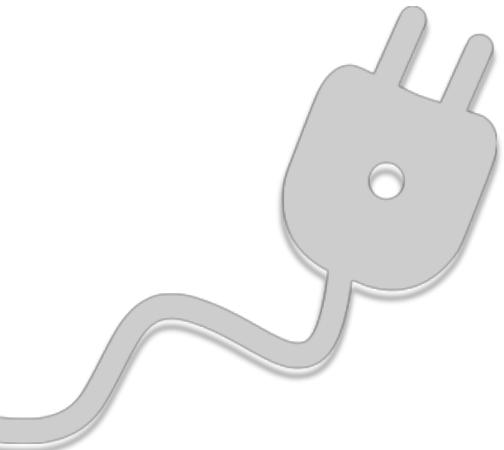
Video stored in CDN at (CDN, Akamai)

<http://KingCDN.com/NetC6y&B23V>



Case Study: Netflix





Socket Programming

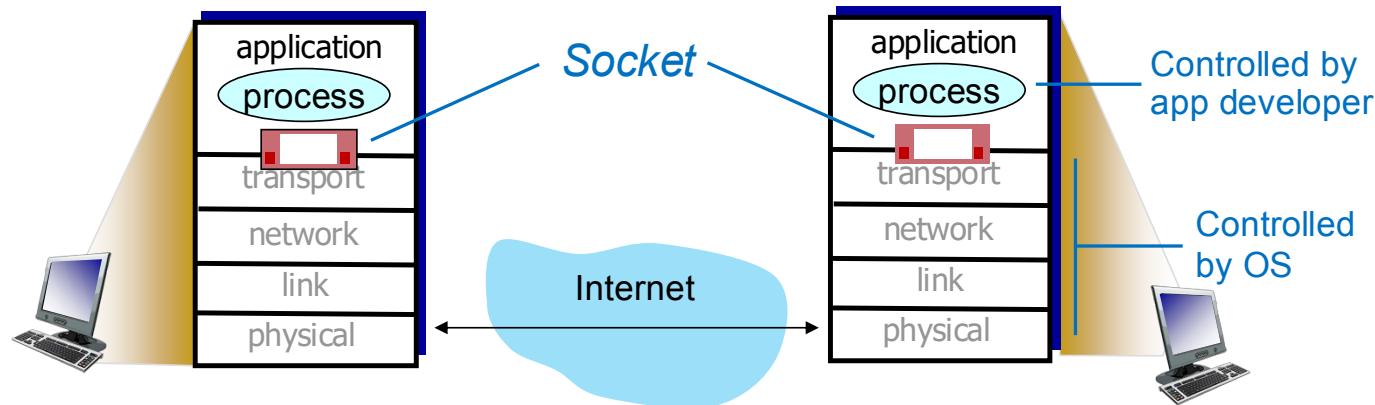


Socket Programming

IP - Address Name
Socket - Address #
Port - Door/Suite #

Goal: Learn how to build client/server applications that communicate using sockets

Socket: Door between application process and end-end-transport protocol



Socket Programming

Two Socket types for two transport services

- **UDP:** Unreliable datagram
- **TCP:** Reliable, byte stream-oriented

Application Example:

1. Client reads a line of characters (data) from its keyboard and sends data to server
2. Server receives the data and converts characters to uppercase
3. Server sends modified data to client
4. Client receives modified data and displays line on its screen

Socket Programming with UDP

UDP: No connection between client & server

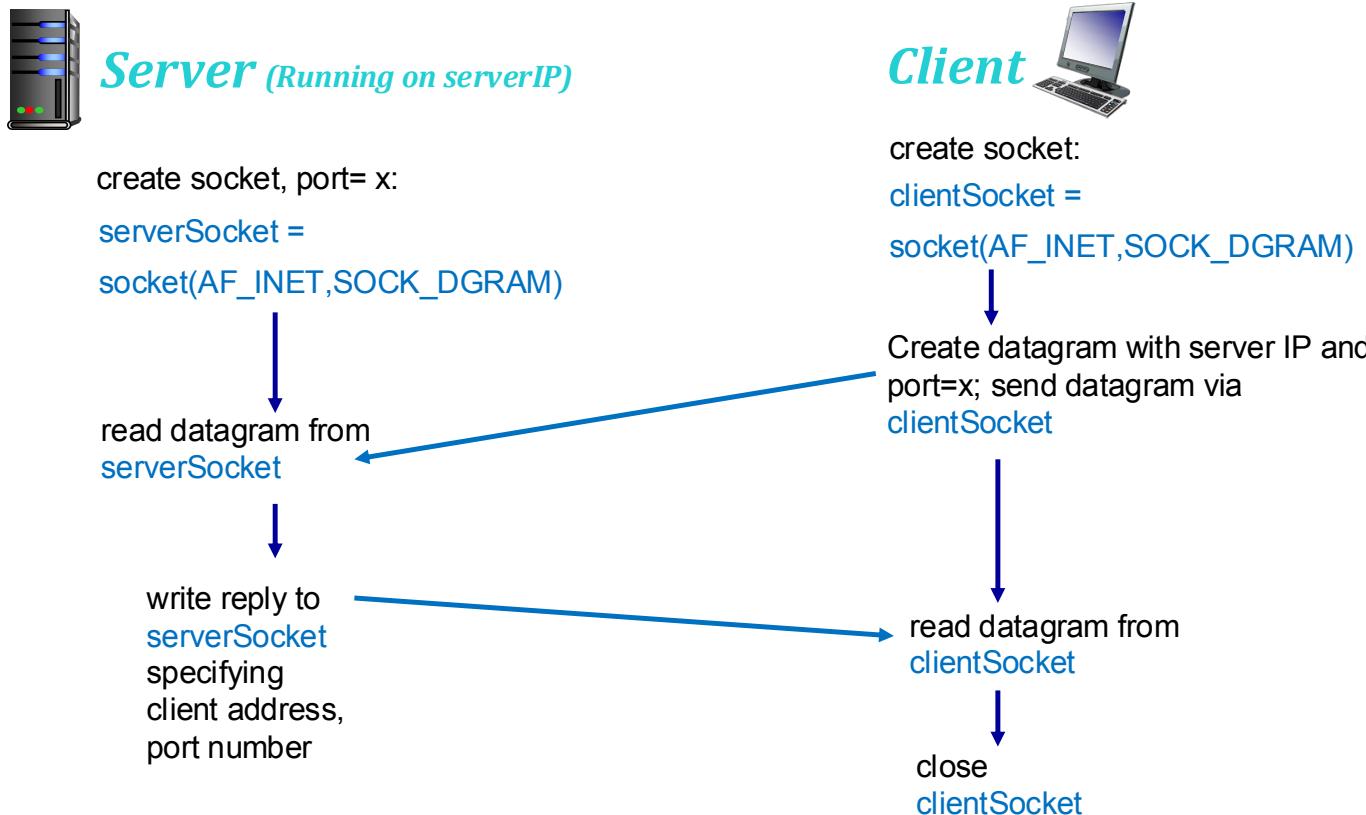
- No handshaking before sending data
- Sender explicitly attaches IP destination address and port number to each packet
- Receiver extracts sender IP address and port number from received packet
- Transmitted data may be lost or received out-of-order

like storing mail to keep sendback address

Application viewpoint

UDP provides **unreliable** transfer of groups of bytes (datagrams) between client and server

Client/Server Socket Interaction: UDP



Example App: UDP Client

Python UDPClient

```
Include Python's socket library → from socket import *
serverName = 'hostname' 127.0.1 localhost / IP Address Device
serverPort = 12000 Certain known ports
Create UDP socket for server → clientSocket = socket(AF_INET, IPv4 _ UDP
                                                SOCK_DGRAM)
Get user keyboard input → message = raw_input('Input lowercase sentence:')
Attach server name, port to message; send into socket → clientSocket.sendto(message.encode(),
                                                               (serverName, serverPort))
Read reply characters from socket into string → modifiedMessage, serverAddress =
                                                clientSocket.recvfrom(2048)
Print out received string and close socket → print modifiedMessage.decode()
                                                clientSocket.close()
```

Example App: UDP Server

Python UDPServer

```
from socket import *
serverPort = 12000 Same as Client

Create UDP socket → serverSocket = socket(AF_INET, SOCK_DGRAM)
Bind socket to local port number 12000 → serverSocket.bind((), serverPort) Bind to socket ⇒ starts listening
                                                print ("The server is ready to receive")

Loop forever → while True:
Read from UDP socket into message, getting →     message, clientAddress = serverSocket.recvfrom(2048)
client's address (client IP and port)           modifiedMessage = message.decode().upper()
                                                serverSocket.sendto(modifiedMessage.encode(),
                                                                    clientAddress)
Send upper case string back to this client →       ↳ Sends using client Address
```

Socket Programming with TCP

Client must contact server

- Server process must first be running
- Server must have created socket (door) that welcomes client's contact

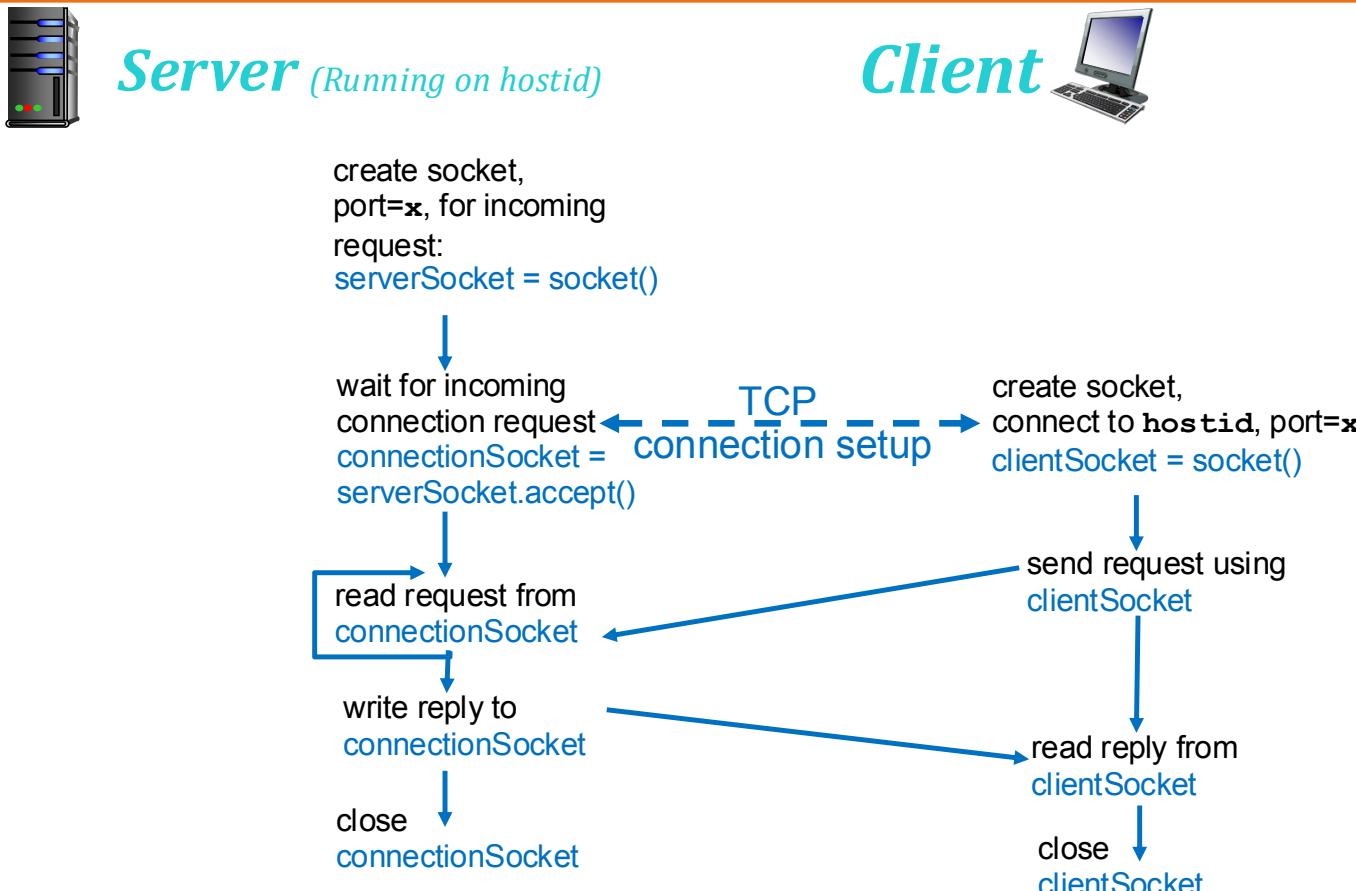
Client contacts server by:

- Creating TCP socket, specifying IP address, port number of server process
- **When client creates socket:** client TCP establishes connection to server TCP
- When contacted by client, **server TCP creates new socket** for server process to communicate with that particular client
 - Allows server to talk with multiple clients
 - Source port numbers used to distinguish clients (more in Transport layer)

Application Viewpoint

TCP provides reliable, in-order byte-stream transfer (pipe) between client and server

Client/Server Socket Interaction: TCP



Different

Example App: TCP Client

Python TCPClient

- ↳ Definition of socket `SOCK_STREAM`
- Create TCP socket for Client →
- Connect to Server TCP Socket →
- ↳ Once connected, send down
 - ↳ no need establish connection every sends
- No need to attach server name, port →

```
from socket import *  
serverName = 'servername'  
serverPort = 12000  
clientSocket = socket(AF_INET, SOCK_STREAM)  
clientSocket.connect((serverName, serverPort))  
sentence = raw_input('Input lowercase sentence: ')  
clientSocket.send(sentence.encode())  
modifiedSentence = clientSocket.recv(1024)  
print('From Server:', modifiedSentence.decode())  
clientSocket.close()
```

} SAME code initialization
ByteStream

Essentially → UDP is "Connection-less"
↳ Receives everything no handshake
→ TCP requires connection

Example App: TCP Server

Python TCP Server

```
from socket import *
serverPort = 12000
serverSocket = socket(AF_INET,SOCK_STREAM)
serverSocket.bind(("",serverPort))
serverSocket.listen(1) | listen @ a time.
print 'The server is ready to receive'
while True:
    connectionSocket, addr = serverSocket.accept()
    | listen to port (door), create connection somewhere else (so no congestion)
    sentence = connectionSocket.recv(1024).decode()
    capitalizedSentence = sentence.upper()
    connectionSocket.send(capitalizedSentence.
                           encode())
    | to UTF8/ASCII
connectionSocket.close()
```

Create TCP welcoming socket →

Server begins listening for incoming TCP requests →

loop forever →

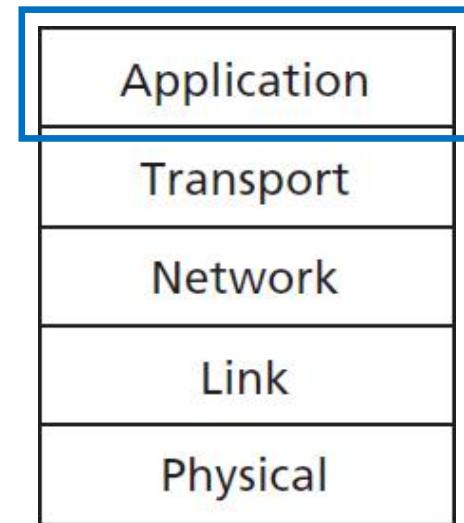
Server waits on accept() for incoming requests, new socket created on return →

Read bytes from socket (but not address as in UDP) →

Close connection to this client (but *not* welcoming socket) →

Application Layer Summary

- Network Applications
- Web & HTTP
- E-mail (SMTP, POP3, IMAP)
- Domain Name System (DNS)
- Peer-to-Peer (P2P) Applications
- Video Streaming and Content Distribution Networks (CDN)
- Socket Programming (UDP & TCP)



Acknowledgements

- The following materials have been used in preparation of this presentation:

[1] Textbook and (edited) Slides: Computer Networking: A Top-Down Approach

James Kurose, Keith Ross

7th and 8th Edition, Pearson

http://gaia.cs.umass.edu/kurose_ross/

[2] Reference: Computer Networks: A Systems Approach

<https://www.systemsapproach.org/book.html>

- Recommended Additional Resources:

[1] Interactive Exercises (Chapter two)

http://gaia.cs.umass.edu/kurose_ross/interactive/