Application Layer

Application architectures

possible structure of applications:

- client-server
- peer-to-peer (P2P)

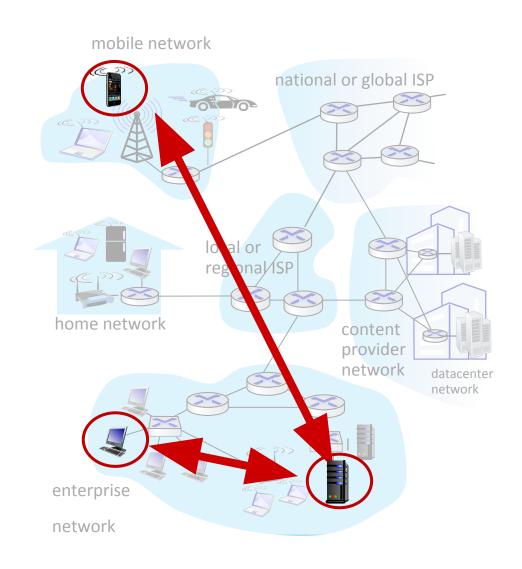
Client-server paradigm

server:

- always-on host
- permanent IP address
- often in data centers, for scaling

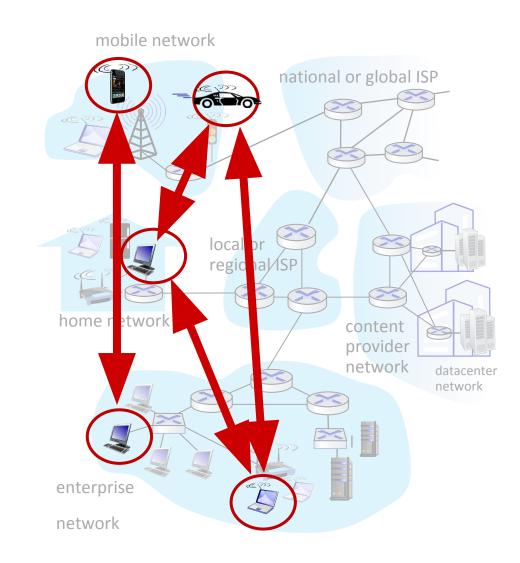
clients:

- contact, communicate with server
- may have dynamic IP addresses
- do not communicate directly with each other
- examples: HTTP, IMAP, FTP



Peer-peer architecture

- no always-on server
- arbitrary end systems directly communicate
- peers request service from other peers, provide service in return to other peers
 - self scalability new peers bring new service capacity, as well as new service demands
- example: P2P file sharing



Processes communicating

process: program running
 within a host

- within same host, two processes communicate using inter-process communication (defined by OS)
- processes in different hosts communicate by exchanging messages

clients, servers

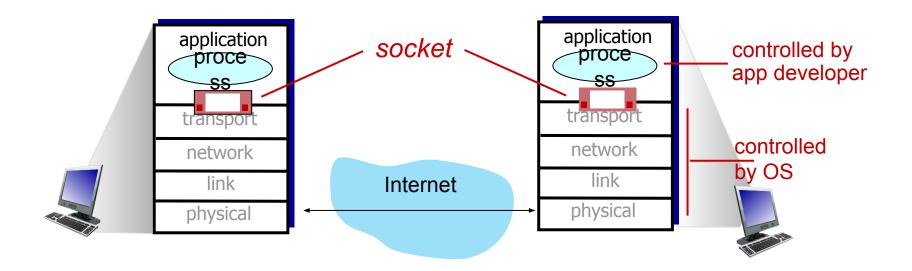
client process: process that initiates communication

server process: process that waits to be contacted

 note: applications with P2P architectures have client processes & server processes

Sockets

- process sends/receives messages to/from its socket
- socket analogous to door
 - sending process shoves message out door
 - sending process relies on transport infrastructure on other side of door to deliver message to socket at receiving process
 - two sockets involved: one on each side



Addressing processes

- to receive messages, process must have *identifier*
- host device has unique 32-bit
 IP address
- •Q: does IP address of host on which process runs suffice for identifying the process?
 - A: no, many processes can be running on same host

- •identifier includes both IP address and port numbers associated with process on host.
- example port numbers:
 - HTTP server: 80
 - mail server: 25
- to send HTTP message to gaia.cs.umass.edu web server:
 - IP address: 128.119.245.12
 - port number: 80
- more shortly...

An application-layer protocol defines:

- types of messages exchanged,
 - e.g., request, response
- message syntax:
 - what fields in messages & how fields are delineated
- message semantics
 - meaning of information in fields
- rules for when and how processes send & respond to messages

What transport service does an app need?

data integrity

- some apps (e.g., file transfer, web transactions) require
 100% reliable data transfer
- other apps (e.g., audio) can tolerate some loss

timing

 some apps (e.g., Internet telephony, interactive games) require low delay to be "effective"

throughput

- some apps (e.g., multimedia) require minimum amount of throughput to be "effective"
- other apps ("elastic apps")
 make use of whatever
 throughput they get

security

encryption, data integrity,

Transport service requirements: common apps

	application	data loss	throughput	time sensitive?
C-1		•		
file transfer	download	no loss	elastic	no
	e-mail	no loss	elastic	no
Web documents		no loss	elastic	no
real-time audio/video		loss-tolerant	audio: 5Kbps-1Mbps	yes, 10's msec
			video:10Kbps-5Mbps	
streaming a	udio/video	loss-tolerant	same as above	yes, few secs
interac	ctive games	loss-tolerant	Kbps+	yes, 10's msec
text	messaging	no loss	elastic	yes and no

Internet transport protocols services

TCP service:

- reliable transport between sending and receiving process
- •flow control: sender won't overwhelm receiver
- •congestion control: throttle sender when network overloaded
- •connection-oriented: setup required between client and server processes
- does not provide: timing, minimum throughput guarantee, security

UDP service:

- unreliable data transfer
 between sending and receiving process
- does not provide: reliability, flow control, congestion control, timing, throughput guarantee, security, or connection setup.

Q: why bother? Why is there a UDP?

Internet applications, and transport protocols

application

application	layer protocol	transport protocol
file transfer/download	FTP [RFC 959]	ТСР
e-mail	SMTP [RFC 5321]	ТСР
Web documents	HTTP 1.1 [RFC 7320]	ТСР
Internet telephony	SIP [RFC 3261], RTP [RFC 3550], or proprietary	TCP or UDP
streaming audio/video	HTTP [RFC 7320], DASH	TCP
interactive games	WOW, FPS (proprietary)	UDP or TCP

Web and HTTP

First, a quick review...

- web page consists of objects, each of which can be stored on different Web servers
- object can be HTML file, JPEG image, audio file,...
- web page consists of base HTML-file which includes several referenced objects, each addressable by a URL, e.g.,

www.someschool.edu/someDept/pic.gif

host name

path name

HTTP overview

HTTP: hypertext transfer protocol

- Web's application-layer protocol
- client/server model:
 - client: browser that requests, receives, (using HTTP protocol) and "displays" Web objects
 - server: Web server sends (using HTTP protocol) objects in response to requests



HTTP overview (continued)

HTTP uses TCP:

- client initiates TCP connection (creates socket) to server, port 80
- server accepts TCP connection from client
- HTTP messages (application-layer protocol messages) exchanged between browser (HTTP client) and Web server (HTTP server)
- TCP connection closed

HTTP is "stateless"

 server maintains no information about past client requests

aside

protocols that maintain "state" are complex!

- past history (state) must be maintained
- if server/client crashes, their views of "state" may be inconsistent, must be reconciled

HTTP connections: two types

Non-persistent HTTP

- 1. TCP connection opened
- 2. at most one object sent over TCP connection
- 3. TCP connection closed

downloading multiple objects required multiple connections

Persistent HTTP

- TCP connection opened to a server
- multiple objects can be sent over single TCP connection between client, and that server
- TCP connection closed

Non-persistent HTTP: example

User enters URL: www.someSchool.edu/someDepartment/home.index (containing text, references to 10 jpeg images)

- 1a. HTTP client initiates TCP connection to HTTP server (process) at www.someSchool.edu on port 80
- 2. HTTP client sends HTTP

 request message (containing
 URL) into TCP connection
 socket. Message indicates
 that client wants object
 someDepartment/home.index

- 1b. HTTP server at host www.someSchool.edu waiting for TCP connection at port 80 "accepts" connection, notifying client
- HTTP server receives request message, forms response message containing requested object, and sends message into its socket

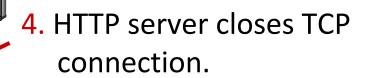


Non-persistent HTTP: example (cont.)

User enters URL: www.someSchool.edu/someDepartment/home.index (containing text, references to 10 jpeg images)



5. HTTP client receives response message containing html file, displays html. Parsing html file, finds 10 referenced jpeg objects



6. Steps 1-5 repeated for each of 10 jpeg objects

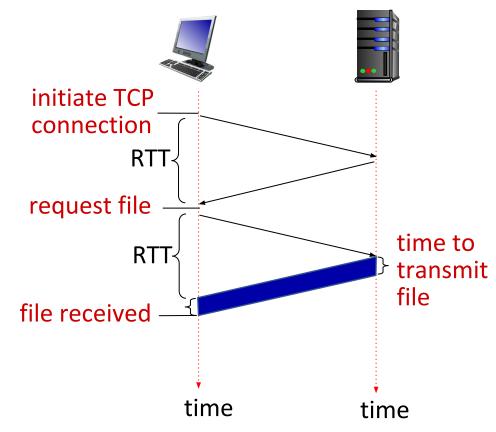


Non-persistent HTTP: response time

RTT (Round Trip Time): time for a small packet to travel from client to server and back

HTTP response time (per object):

- one RTT to initiate TCP connection
- one RTT for HTTP request and first few bytes of HTTP response to return
- obect/file transmission time



Non-persistent HTTP response time = 2RTT+ file transmission time

Persistent HTTP (HTTP 1.1)

Non-persistent HTTP issues:

- requires 2 RTTs per object
- OS overhead for each TCP connection
- browsers often open multiple parallel TCP connections to fetch referenced objects in parallel

Persistent HTTP (HTTP1.1):

- 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

HTTP request messages

POST method (HTTP 1.0):

- web page often includes form input
- user input sent from client to server in entity body of HTTP POST request message

GET method (HTTP 1.0) (for sending data to server):

• include user data in URL field of HTTP GET request message (following a '?'):

www.somesite.com/animalsearch?monkeys&banana

HEAD method (HTTP 1.0):

 requests headers (only) that would be returned if specified URL were requested with an HTTP GET method.

PUT method (HTTP 1.1):

- uploads new file (object) to server
- completely replaces file that exists at specified URL with content in entity body of POST HTTP request message

HTTP response status codes

- status code appears in 1st line in server-to-client response message.
- some sample codes:

200 OK

request succeeded, requested object later in this message

301 Moved Permanently

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

400 Bad Request

request msg not understood by server

404 Not Found

requested document not found on this server

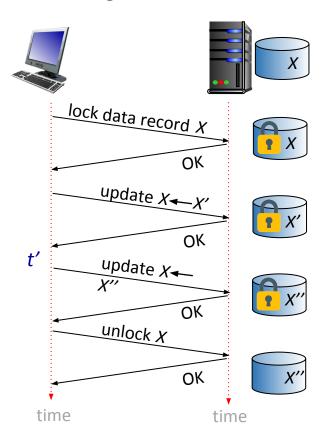
505 HTTP Version Not Supported

Maintaining user/server state: cookies

Recall: HTTP GET/response interaction is *stateless*

- no notion of multi-step exchanges of HTTP messages to complete a Web "transaction"
 - no need for client/server to track "state" of multi-step exchange
 - all HTTP requests are independent of each other
 - no need for client/server to "recover" from a partially-completed-but-never-complete ly-completed transaction

a stateful protocol: client makes two changes to X, or none at all



Q: what happens if network connection or client crashes at t'?

Maintaining user/server state: cookies

Web sites and client browser use cookies to maintain some state between transactions

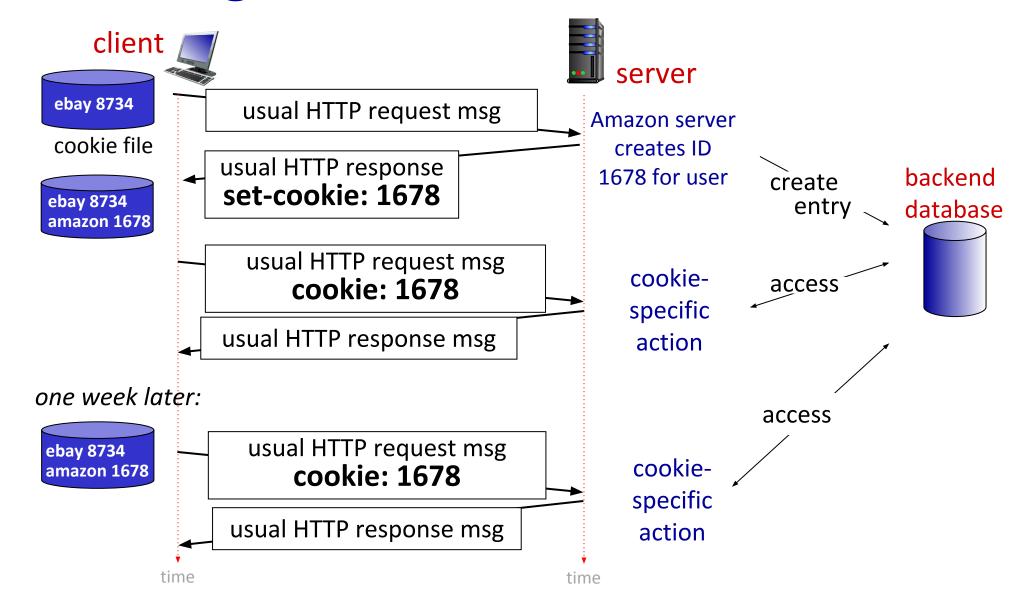
four components:

- 1) cookie header line of HTTP *response* message
- 2) cookie header line in next HTTP request message
- 3) cookie file kept on user's host, managed by user's browser
- 4) back-end database at Web site

Example:

- Susan uses browser on laptop, visits specific e-commerce site for first time
- when initial HTTP requests arrives at site, site creates:
 - unique ID (aka "cookie")
 - entry in backend database for ID
- subsequent HTTP requests from Susan to this site will contain cookie ID value, allowing site to "identify" Susan

Maintaining user/server state: cookies



HTTP cookies: comments

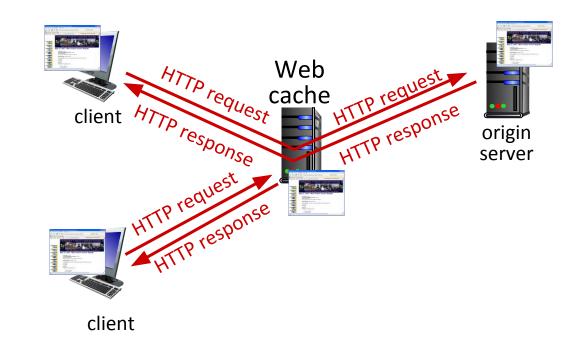
What cookies can be used for:

- authorization
- shopping carts
- recommendations
- user session state (Web e-mail)

Web caches

Goal: satisfy client requests without involving origin server

- user configures browser to point to a (local) Web cache
- browser sends all HTTP requests to cache
 - *if* object in cache: cache returns object to client
 - else cache requests object from origin server, caches received object, then returns object to client



Web caches (aka proxy servers)

- Web cache acts as both client and server
 - server for original requesting client
 - client to origin server
- server tells cache about object's allowable caching in response header:

```
Cache-Control: max-age=<seconds>
```

```
Cache-Control: no-cache
```

Why Web caching?

- reduce response time for client request
 - cache is closer to client
- reduce traffic on an institution's access link

Caching example

Scenario:

- access link rate: 1.54 Mbps
- RTT from institutional router to server: 2 sec
- web object size: 100K bits
- average request rate from browsers to origin servers: 15/sec
 - avg data rate to browsers: 1.50 Mbps

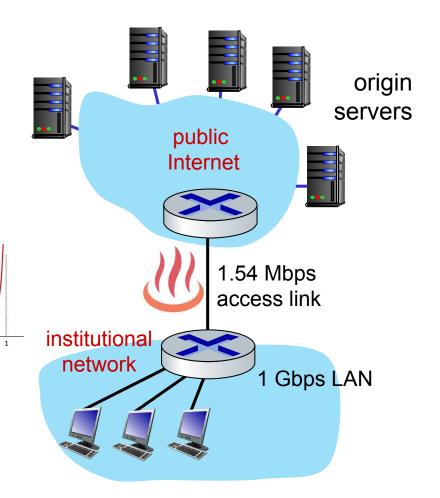
Performance:

- LAN utilization: .0015

problem: large queueing delays at high utilization!

end-end delay = Internet delay +
 access link delay + LAN delay

= 2 sec +(minutes) + usecs



Option 1: buy a faster access link

Scenario:

154 Mbps

- access link rate: 1.54 Mbps
- RTT from institutional router to server: 2 sec
- web object size: 100K bits
- average request rate from browsers to origin servers: 15/sec
 - avg data rate to browsers: 1.50 Mbps

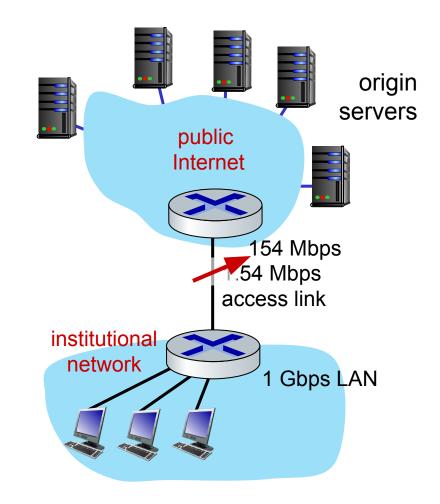
Performance:

- access link utilization = .97 → .0097
- LAN utilization: .0015
- end-end delay = Internet delay +
 access link delay + LAN delay

= 2 sec + minutes + usecs

msecs

Cost: faster access link (expensive!)



Option 2: install a web cache

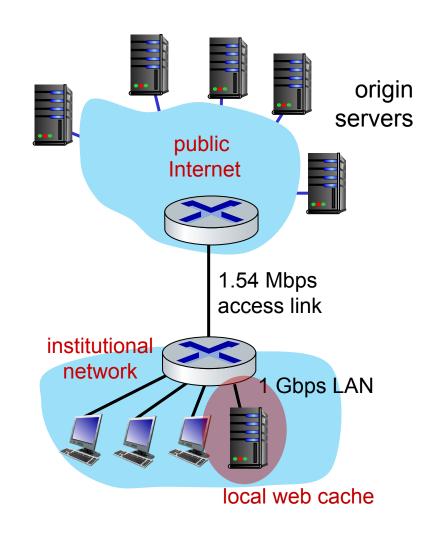
Scenario:

- access link rate: 1.54 Mbps
- RTT from institutional router to server: 2 sec
- web object size: 100K bits
- average request rate from browsers to origin servers: 15/sec
 - avg data rate to browsers: 1.50 Mbps

Cost: web cache (cheap!)

Performance:

- LAN utilization: .? How to compute link
- access link utilization = ? utilization, delay?
- average end-end delay = ?



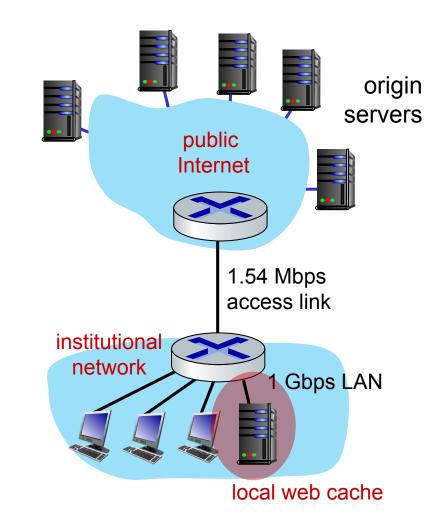
Calculating access link utilization, end-end delay with cache:

suppose cache hit rate is 0.4:

- 40% requests served by cache, with low (msec) delay
- 60% requests satisfied at origin
 - rate to browsers over access link

$$= 0.6 * 1.50 \text{ Mbps} = .9 \text{ Mbps}$$

- access link utilization = 0.9/1.54 = .58 means low (msec) queueing delay at access link
- average end-end delay:
 - = 0.6 * (delay from origin servers) + 0.4 * (delay when satisfied at cache)
 - $= 0.6 (2.01) + 0.4 (^msecs) = ^1.2 secs$



lower average end-end delay than with 154 Mbps link (and cheaper too!)

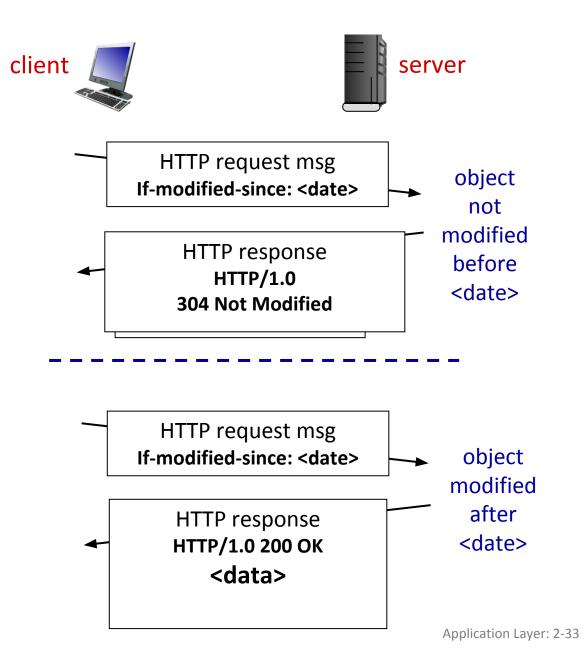
Conditional GET

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

- no object transmission delay (or use of network resources)
- •client: specify date of cached copy in HTTP request

If-modified-since: <date>

server: response contains no object if cached copy is up-to-date:HTTP/1.0 304 Not Modified



HTTP/2

Key goal: decreased delay in multi-object HTTP requests

<u>HTTP1.1:</u> introduced multiple, pipelined GETs over single TCP connection

- server responds in-order (FCFS: first-come-first-served scheduling) to GET requests
- with FCFS, small object may have to wait for transmission (head-of-line (HOL) blocking) behind large object(s)
- loss recovery (retransmitting lost TCP segments) stalls object transmission

HTTP/2

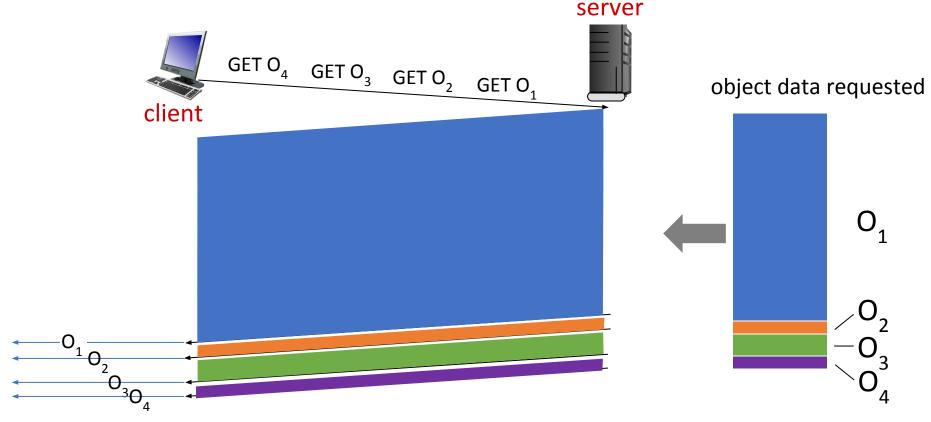
Key goal: decreased delay in multi-object HTTP requests

<u>HTTP/2:</u> [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 object priority (not necessarily FCFS)
- push unrequested objects to client
- divide objects into frames, schedule frames to mitigate HOL blocking

HTTP/2: mitigating HOL blocking

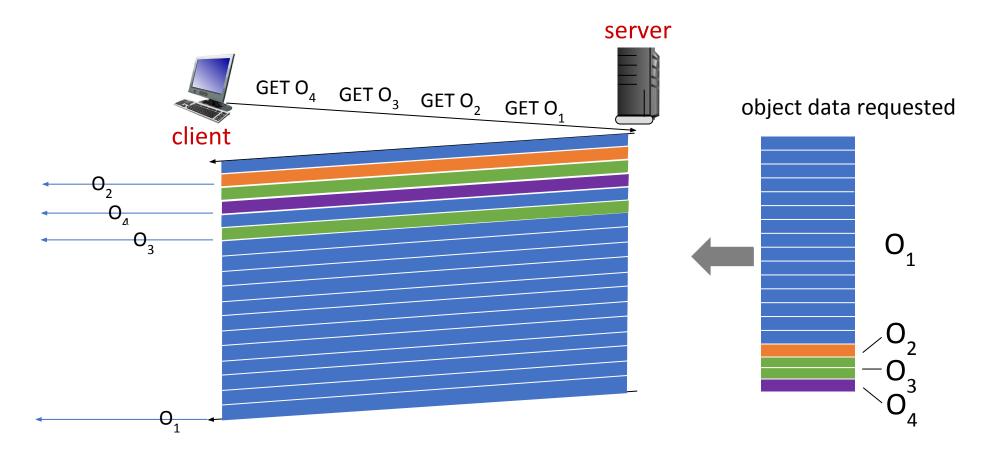
HTTP 1.1: client requests 1 large object (e.g., video file) and 3 smaller objects



objects delivered in order requested: O_2 , O_3 , O_4 wait behind O_1

HTTP/2: mitigating HOL blocking

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



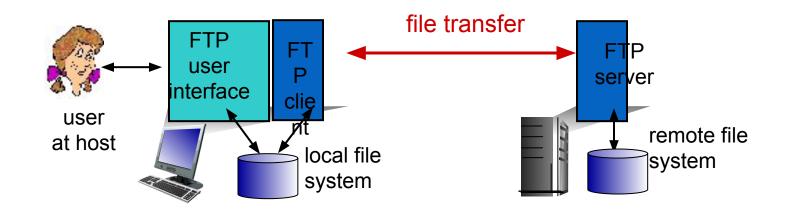
 $O_{2'}$, $O_{3'}$, O_{4} delivered quickly, O_{1} slightly delayed

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- and congestion-control (more pipelining) over UDP
 - more on HTTP/3 in transport layer

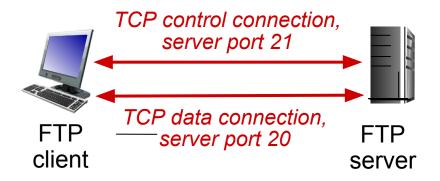
FTP: the file transfer protocol



- transfer file to/from remote host
- client/server model
 - client: side that initiates transfer (either to/from remote)
 - server: remote host
- ftp: RFC 959
- ftp server: port 21

FTP: separate control, data connections

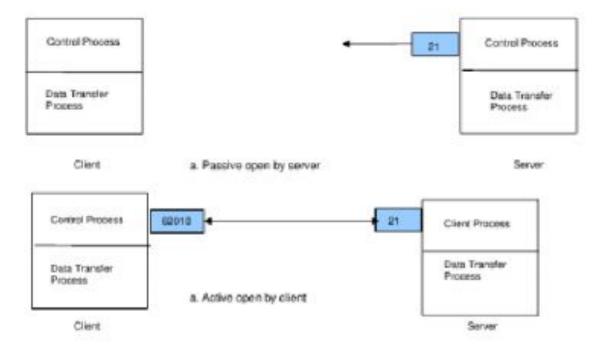
- FTP client contacts FTP server at port 21, using TCP
- client authorized over control connection
- client browses remote directory, sends commands over control connection
- when server receives file transfer command, server opens 2nd TCP data connection (for file) to client
- after transferring one file, server closes data connection



- server opens another TCP data connection to transfer another file
- control connection: "out of band"
- ❖ FTP server maintains "state": current directory, earlier authentication

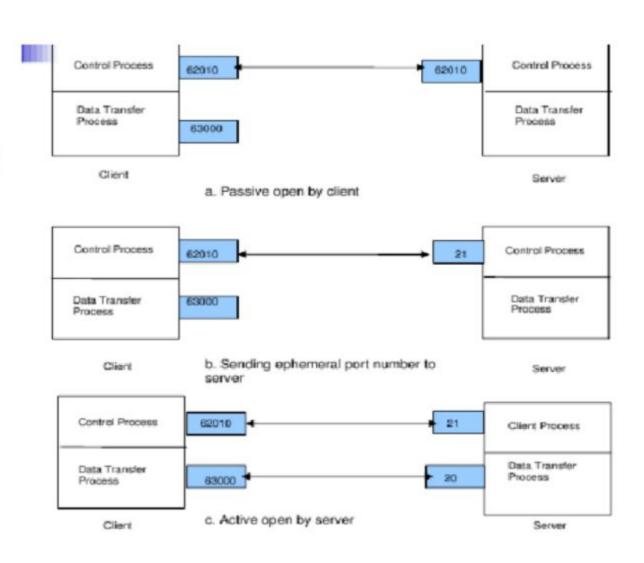
FTP: Control Connections

- Control Connection: Two steps
 - Server issues a passive open on the well-known port 21 and waits for a client
 - Client uses an ephemeral port and issues an active open



FTP: Data Connections

- Data Connection: Three steps
 - Client, not the server, issues a passive open using an ephemeral port.
 - Client sends this port number to the server using the PORT command
 - Server receives the port number and issues an active open using the well-known port 20 and received ephemeral port number.



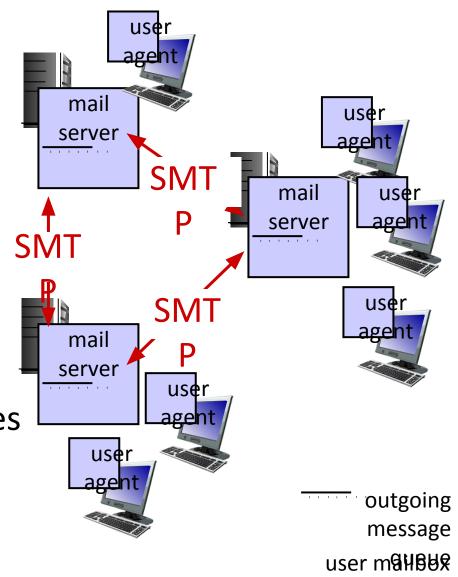
E-mail

Three major components:

- user agents
- mail servers
- simple mail transfer protocol: SMTP

User Agent

- a.k.a. "mail reader"
- composing, editing, reading mail messages
- e.g., Outlook, iPhone mail client
- outgoing, incoming messages stored on server



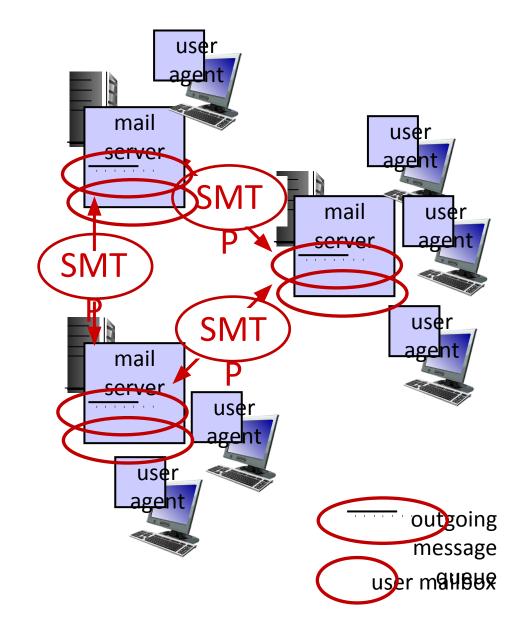
E-mail: mail servers

mail servers:

- mailbox contains incoming messages for user
- message queue of outgoing (to be sent) mail messages

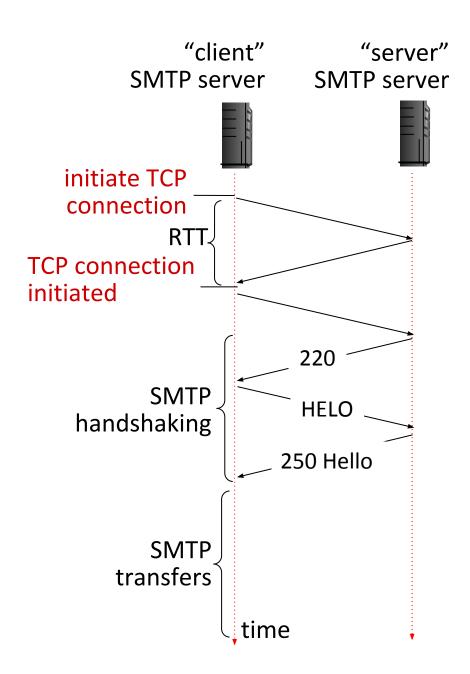
SMTP protocol between mail servers to send email messages

- client: sending mail server
- "server": receiving mail server



SMTP RFC (5321)

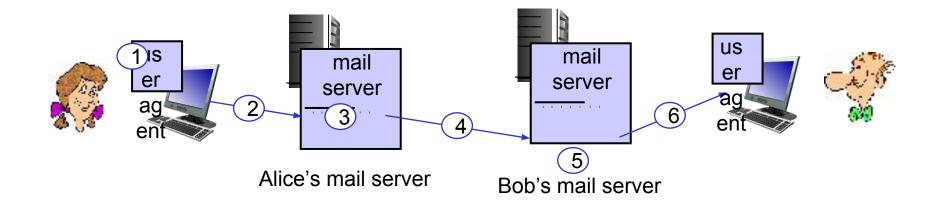
- uses TCP to reliably transfer email message from client (mail server initiating connection) to server, port 25
 - direct transfer: sending server (acting like client) to receiving server
- three phases of transfer
 - SMTP handshaking (greeting)
 - SMTP transfer of messages
 - SMTP closure
- command/response interaction (like HTTP)
 - commands: ASCII text
 - response: status code and phrase



Scenario: Alice sends e-mail to Bob

- 1) Alice uses UA to compose e-mail message "to" bob@someschool.edu
- 2) Alice's UA sends message to her mail server using SMTP; message placed in message queue
- 3) client side of SMTP at mail server 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
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
C: Do you like ketchup?
C: How about pickles?
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 hamburger.edu closing connection
```

SMTP: observations

comparison with HTTP:

- HTTP: client pull
- SMTP: client push
- both have ASCII command/response interaction, status codes
- HTTP: each object encapsulated in its own response message
- SMTP: multiple objects sent in multipart message

- 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

Mail message format

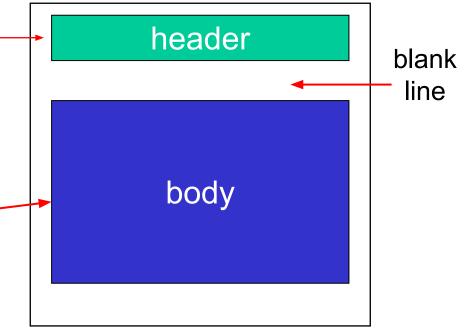
SMTP: protocol for exchanging e-mail messages, defined in RFC 5321 (like RFC 7231 defines HTTP)

RFC 2822 defines *syntax* for e-mail message itself (like HTML defines syntax for web documents)

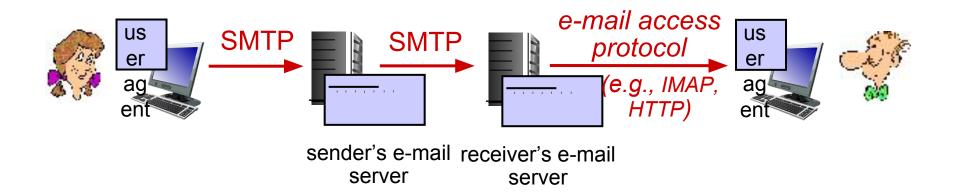
- header lines, e.g.,
 - To:
 - From:
 - Subject:

these lines, within the body of the email message area different from SMTP MAIL FROM:, RCPT TO: commands!

Body: the "message", ASCII characters only



Retrieving email: mail access protocols



- SMTP: delivery/storage of e-mail messages to receiver's server
- mail access protocol: retrieval from server
 - IMAP: Internet Mail Access Protocol [RFC 3501]: messages stored on server, IMAP provides retrieval, deletion, folders of stored messages on server
- HTTP: gmail, Hotmail, Yahoo!Mail, etc. provides web-based interface on top of STMP (to send), IMAP (or POP) to retrieve e-mail messages

POP3 protocol

authorization phase

- client commands:
 - user: declare username
 - pass: password
- server responses
 - +OK
 - -ERR

transaction phase, client:

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

```
S: +OK POP3 server ready
C: user bob
S: +OK
C: pass hungry
S: +OK user successfully logged on
C: list
S: 1 498
S: 2 912
S: .
C: retr 1
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 (more) and IMAP

more about POP3

- previous example uses POP3 "download and delete" mode
 - Bob cannot re-read e-mail if he changes client
- POP3
 "download-and-keep":
 copies of messages on
 different clients
- POP3 is stateless across sessions

IMAP

- keeps all messages in one place: at server
- allows user to organize messages in folders
- keeps user state across sessions:
 - names of folders and mappings between message IDs and folder name

DNS: Domain Name System

people: many identifiers:

SSN, name, passport #

Internet hosts, routers:

- IP address (32 bit) used for addressing datagrams
- "name", e.g., cs.umass.edu used by humans

Q: how to map between IP address and name, and vice versa?

Domain Name System (DNS):

- distributed database implemented in hierarchy of many name servers
- application-layer protocol: hosts, DNS servers communicate to resolve names (address/name translation)
 - note: core Internet function, implemented as application-layer protocol
 - complexity at network's "edge"

DNS: services, structure

DNS services:

- hostname-to-IP-address translation
- host aliasing
 - canonical, alias names
- mail server aliasing
- load distribution
 - replicated Web servers: many IP addresses correspond to one name

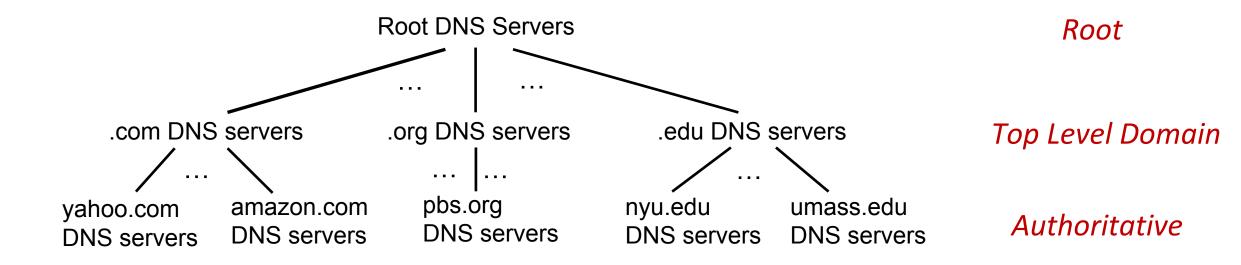
Q: Why not centralize DNS?

- single point of failure
- traffic volume
- distant centralized database
- maintenance

A: doesn't scale!

- Comcast DNS servers alone: 600B DNS queries/day
- Akamai DNS servers alone:2.2T DNS queries/day

DNS: a distributed, hierarchical database

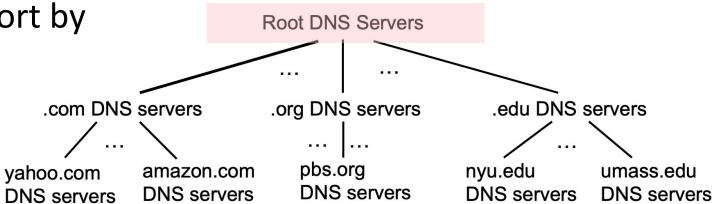


Client wants IP address for www.amazon.com; 1st approximation:

- client queries root server to find .com DNS server
- client queries .com DNS server to get amazon.com DNS server
- client queries amazon.com DNS server to get IP address for www.amazon.com

DNS: root name servers

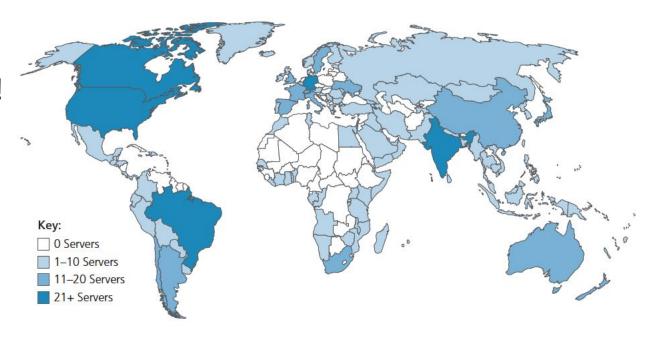
 official, contact-of-last-resort by name servers that can not resolve name



DNS: root name servers

- official, contact-of-last-resort by name servers that can not resolve name
- incredibly important Internet function
 - Internet couldn't function without it!
 - DNSSEC provides security (authentication, message integrity)
- ICANN (Internet Corporation for Assigned Names and Numbers) manages root DNS domain

13 logical root name "servers" worldwide each "server" replicated many times (~200 servers in US)

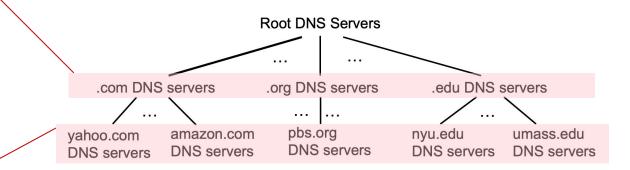


Top-Level Domain, and authoritative servers

Top-Level Domain (TLD) servers:

- responsible for .com, .org, net, .edu, .aero, .jobs, .museums, and all top-level country domains, e.g.: .cn, .uk, .fr, .ca, .jp
- Network Solutions: authoritative registry for .com, .net TLD

Educause: .edu TLD



authoritative DNS servers:

- organization's own DNS server(s), providing authoritative hostname to IP mappings for organization's named hosts
- can be maintained by organization or service provider

Local DNS name servers

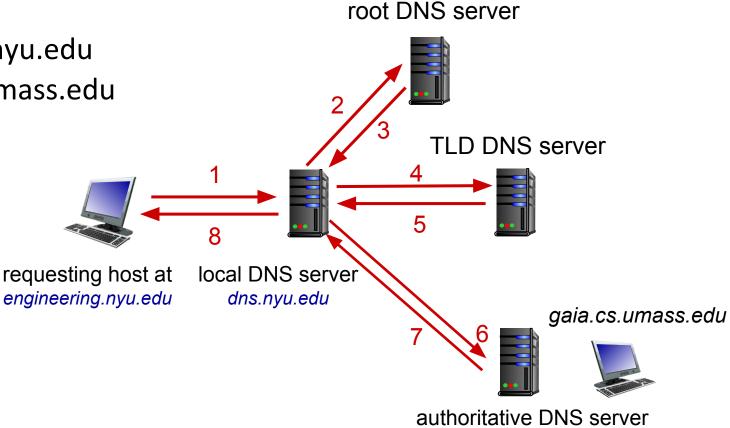
- when host makes DNS query, it is sent to its local DNS server
 - Local DNS server returns reply, answering:
 - from its local cache of recent name-to-address translation pairs (possibly out of date!)
 - forwarding request into DNS hierarchy for resolution
 - each ISP has local DNS name server; to find yours:
 - MacOS: % scutil --dns
 - Windows: >ipconfig /all
- local DNS server doesn't strictly belong to hierarchy

DNS name resolution: iterated query

Example: host at engineering.nyu.edu wants IP address for gaia.cs.umass.edu

Iterated query:

- contacted server replies with name of server to contact
- "I don't know this name, but ask this server"



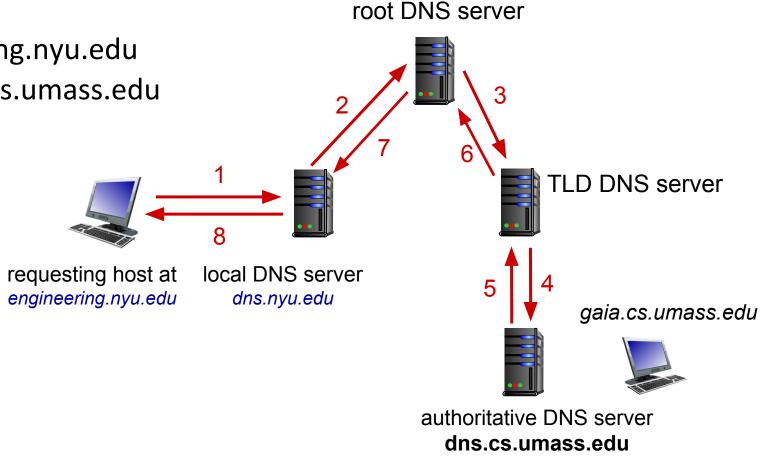
dns.cs.umass.edu

DNS name resolution: recursive query

Example: host at engineering.nyu.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?



Caching DNS Information

- once (any) name server learns mapping, it caches mapping, and immediately returns a cached mapping in response to a query
 - caching improves response time
 - cache entries timeout (disappear) after some time (TTL)
 - TLD servers typically cached in local name servers
- cached entries may be out-of-date
 - if named host changes IP address, may not be known Internet-wide until all TTLs expire!
 - best-effort name-to-address translation!

DNS records

DNS: distributed database storing resource records (RR)

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
- www.ibm.com is really servereast.backup2.ibm.com
- value is canonical name

type=MX

 value is name of SMTP mail server associated with name

Getting your info into the DNS

example: new startup "Network Utopia"

- register name networkuptopia.com at DNS registrar (e.g., Network Solutions)
 - provide names, IP addresses of authoritative name server (primary and secondary)
 - registrar inserts NS, A RRs into .com TLD server:

```
(networkutopia.com, dns1.networkutopia.com, NS)
(dns1.networkutopia.com, 212.212.212.1, A)
```

- create authoritative server locally with IP address 212.212.212.1
 - type A record for www.networkuptopia.com
 - type MX record for networkutopia.com