

source: computer-networks-webdesign.com



CSCI 6760 - Computer Networks Spring 2017

Instructor: Prof. Roberto Perdisci
perdisci@cs.uga.edu

These slides are adapted from the textbook slides by J.F. Kurose and K.W. Ross

Chapter 2: Application Layer

Our goals:

- ▶ conceptual, implementation aspects of network application protocols
 - ▶ transport-layer service models
 - ▶ client-server paradigm
 - ▶ peer-to-peer paradigm
- ▶ learn about protocols by examining popular application-level protocols
 - ▶ HTTP
 - ▶ FTP
 - ▶ SMTP / POP3 / IMAP
 - ▶ DNS
- ▶ programming network applications
 - ▶ socket API



Some network apps

- ▶ e-mail
- ▶ web
- ▶ instant messaging
- ▶ remote login
- ▶ P2P file sharing
- ▶ multi-user network games
- ▶ streaming stored video (YouTube)
- ▶ voice over IP
- ▶ real-time video conferencing
- ▶ cloud computing
- ▶ ...
- ▶ ...
- ▶



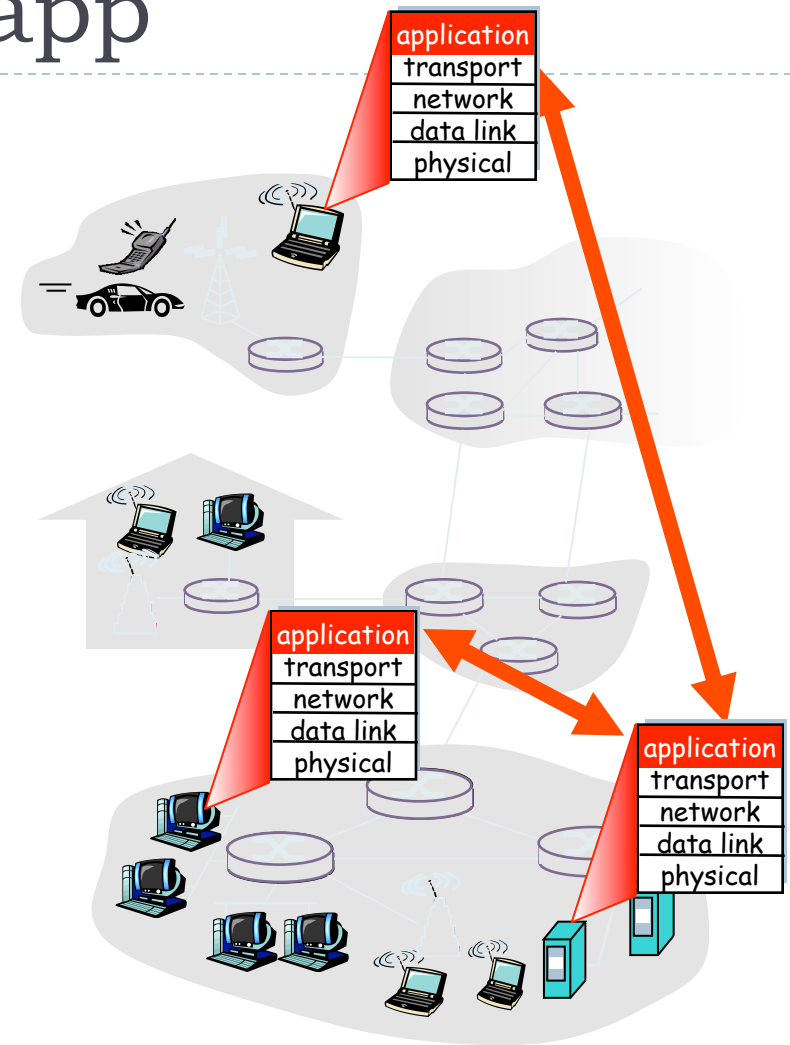
Creating a network app

write programs that

- ▶ run on (different) *end systems*
- ▶ communicate over network
- ▶ e.g., web server software communicates with browser software

No need to write software for network-core devices

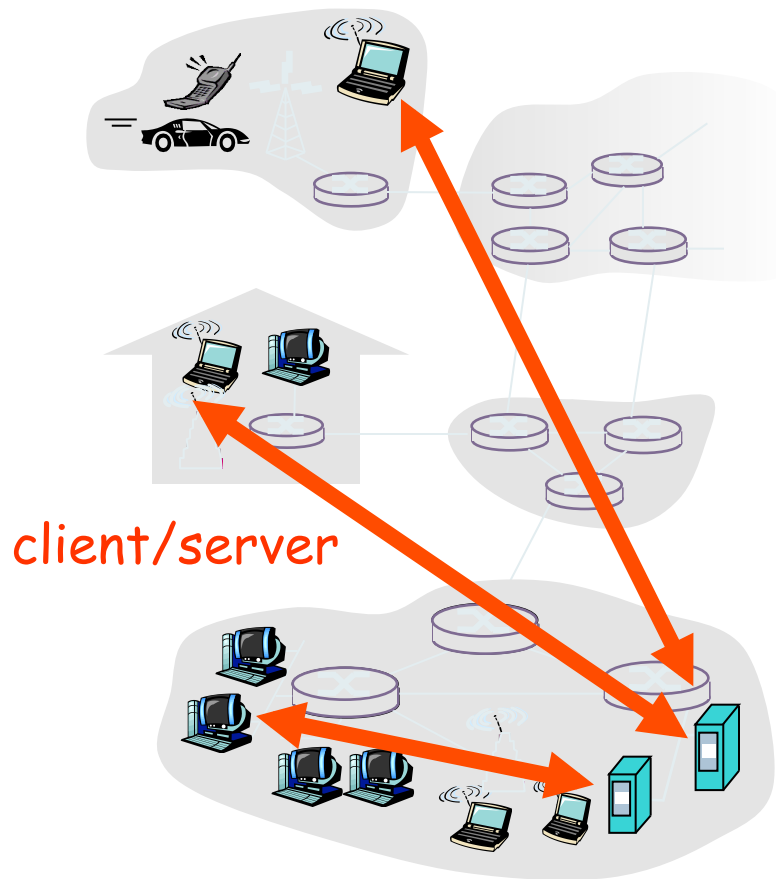
- ▶ network-core devices do not run user applications
- ▶ applications on end systems allows for rapid app development, propagation



Application architectures

- ▶ client-server
- ▶ peer-to-peer (P2P)
- ▶ hybrid of client-server and P2P

Client-server architecture



server:

- ▶ always-on host
- ▶ permanent IP address
- ▶ server farms for scaling

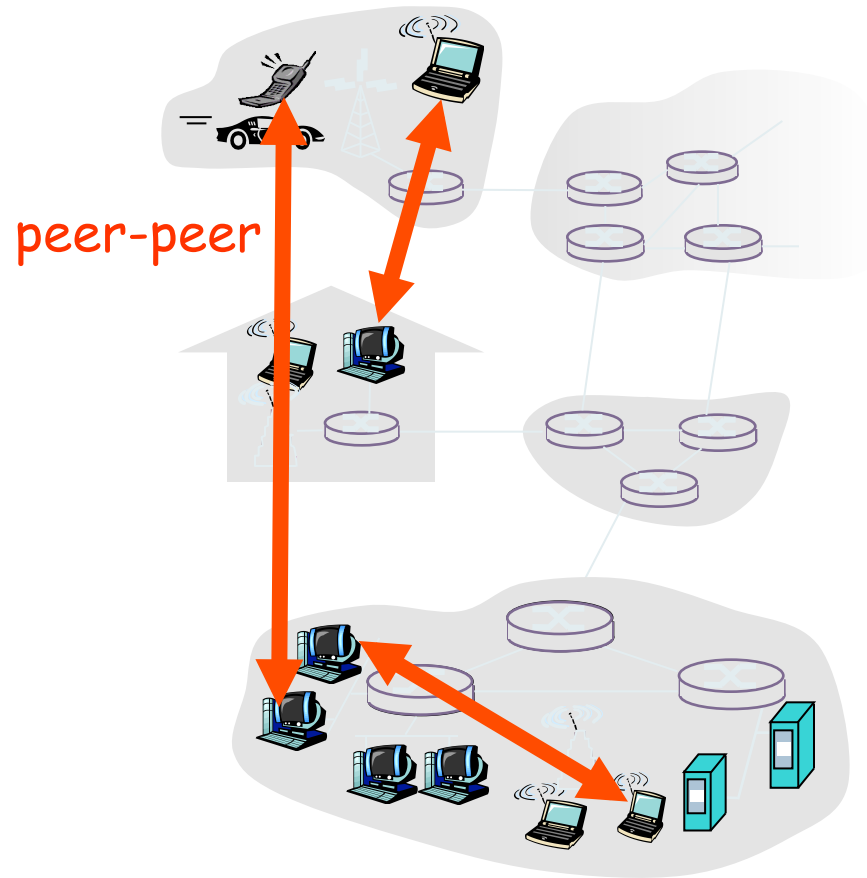
clients:

- ▶ communicate with server
- ▶ may be intermittently connected
- ▶ may have dynamic IP addresses
- ▶ do not communicate directly with each other

Pure P2P architecture

- ▶ *no* always-on server
- ▶ arbitrary end systems directly communicate
- ▶ peers are intermittently connected and change IP addresses

highly scalable but difficult to manage



Hybrid of client-server and P2P

Skype

- ▶ voice-over-IP P2P application
- ▶ centralized server: finding address of remote party:
- ▶ client-client connection: direct (not through server)

Instant messaging

- ▶ chatting between two users is P2P
- ▶ centralized service: client presence detection/location
 - ▶ user registers its IP address with central server when it comes online
 - ▶ user contacts central server to find IP addresses of buddies



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**

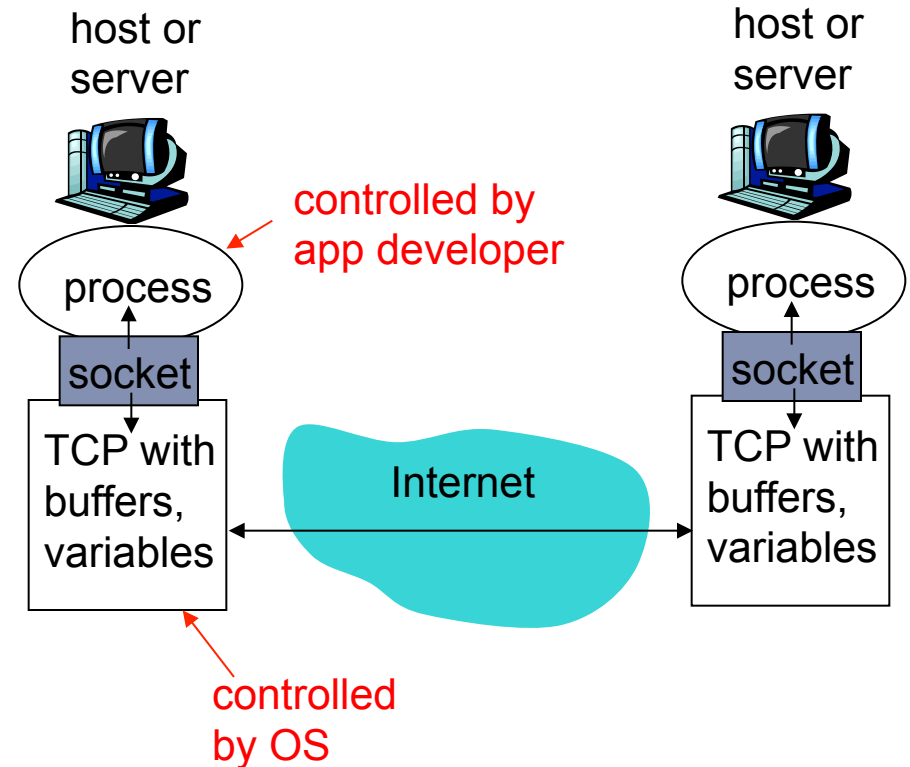
client process: process that initiates communication

server process: process that waits to be contacted

- ❖ **aside:** 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 which brings message to socket at receiving process



- ❖ API: (1) choice of transport protocol; (2) ability to fix a few parameters (lots more on this later)

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?

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...

App-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

public-domain protocols:

- ▶ defined in RFCs
- ▶ allows for interoperability
- ▶ e.g., HTTP, SMTP

proprietary protocols:

- ▶ e.g., Skype



What transport service does an app need?

Reliability

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

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 of common apps

Application	Data loss	Throughput	Time Sensitive
file transfer	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 video: 10kbps-5Mbps	yes, 100's msec
stored audio/video	loss-tolerant	same as above	yes, few secs
interactive games	loss-tolerant	few kbps up	yes, 100's msec
instant messaging	no loss	elastic	yes and no

Internet transport protocols services

TCP service:

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

UDP service:

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

Q: why bother? Why is there a UDP?

Internet apps: application, transport protocols***

Application	Application layer protocol	Underlying transport protocol
e-mail	SMTP [RFC 2821]	TCP
remote terminal access	Telnet [RFC 854]	TCP
Web	HTTP [RFC 2616]	TCP
file transfer	FTP [RFC 959]	TCP
streaming multimedia	HTTP (eg Youtube), RTP [RFC 1889]	TCP or UDP
Internet telephony	SIP, RTP, proprietary (e.g., Skype)	typically UDP



Chapter 2: Application layer

2.1 Principles of network applications

- ▶ app architectures
- ▶ app requirements

2.2 Web and HTTP

2.3 FTP

2.4 Electronic Mail

- ▶ SMTP, POP3, IMAP

2.5 DNS

2.6 P2P applications

2.7 Socket programming with TCP

2.8 Socket programming with UDP



Web and HTTP

First, a review...

- ▶ **web page** consists of **objects**
- ▶ object can be HTML file, JPEG image, Java applet, audio file,...
- ▶ web page consists of **base HTML-file** which includes several referenced objects
- ▶ each object is addressable by a **URL**
- ▶ example URL:

`www.someschool.edu/someDept/pic.gif`

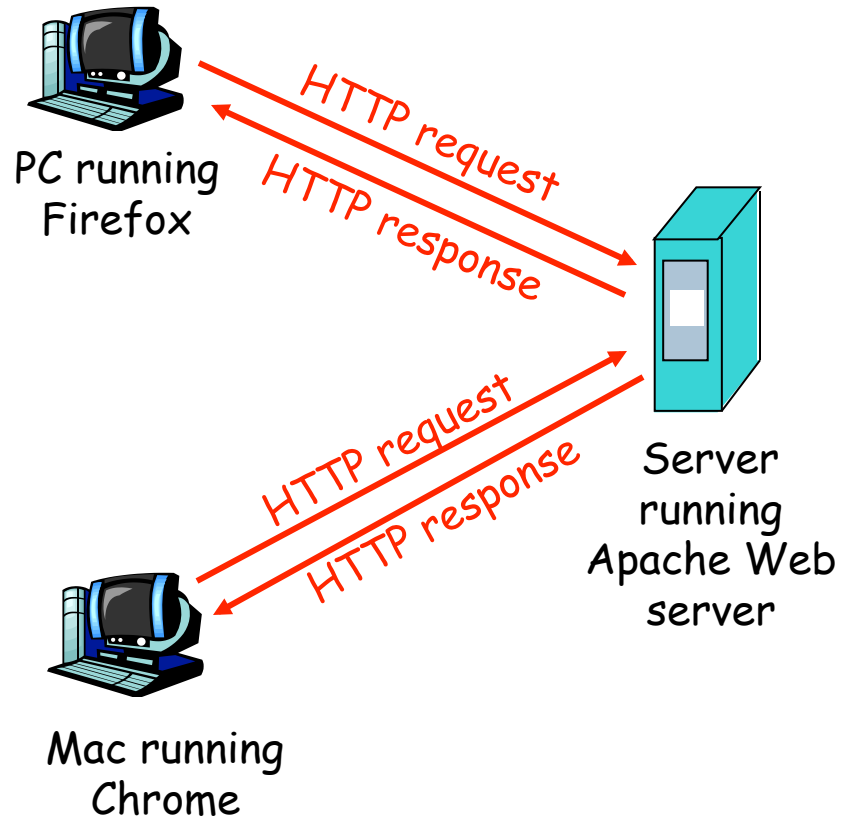
host name

path name

HTTP overview

HTTP: hypertext transfer protocol

- ▶ Web's application layer protocol
- ▶ client/server model
 - ▶ *client*: browser that requests, receives, "displays" Web objects
 - ▶ *server*: Web server sends objects in response to requests



HTTP overview (continued)

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

non-persistent HTTP

- ▶ at most one object sent over TCP connection.

persistent HTTP

- ▶ multiple objects can be sent over single TCP connection between client, server.



Nonpersistent HTTP

suppose user enters URL:

`www.someSchool.edu/someDepartment/home.index`

(contains text,
references to 10
jpeg images)

1a. HTTP client initiates TCP

connection to HTTP server
(process) at `www.someSchool.edu`
on port 80

1b. HTTP server at host

`www.someSchool.edu` waiting
for TCP connection at port 80.
"accepts" connection, notifying
client

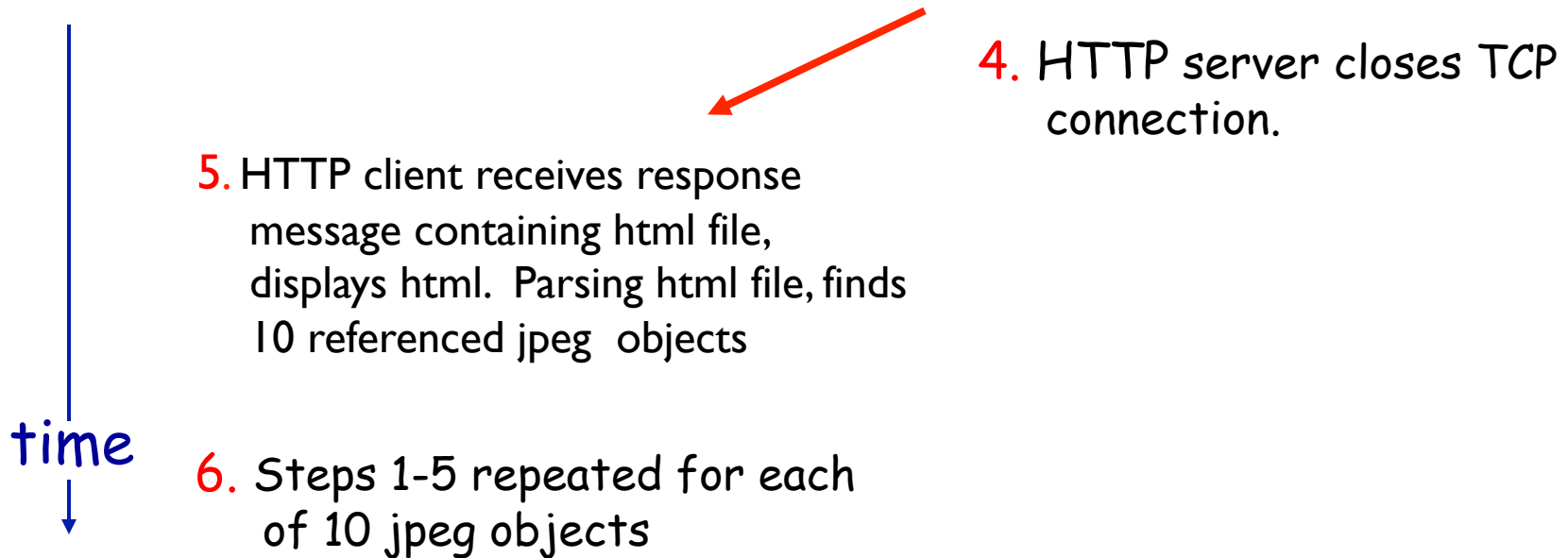
2. HTTP client sends HTTP
request message (containing
URL) into TCP connection
socket. Message indicates
that client wants object
`someDepartment/home.index`

3. HTTP server receives request
message, forms *response
message* containing requested
object, and sends message
into its socket

time



Nonpersistent HTTP (cont.)



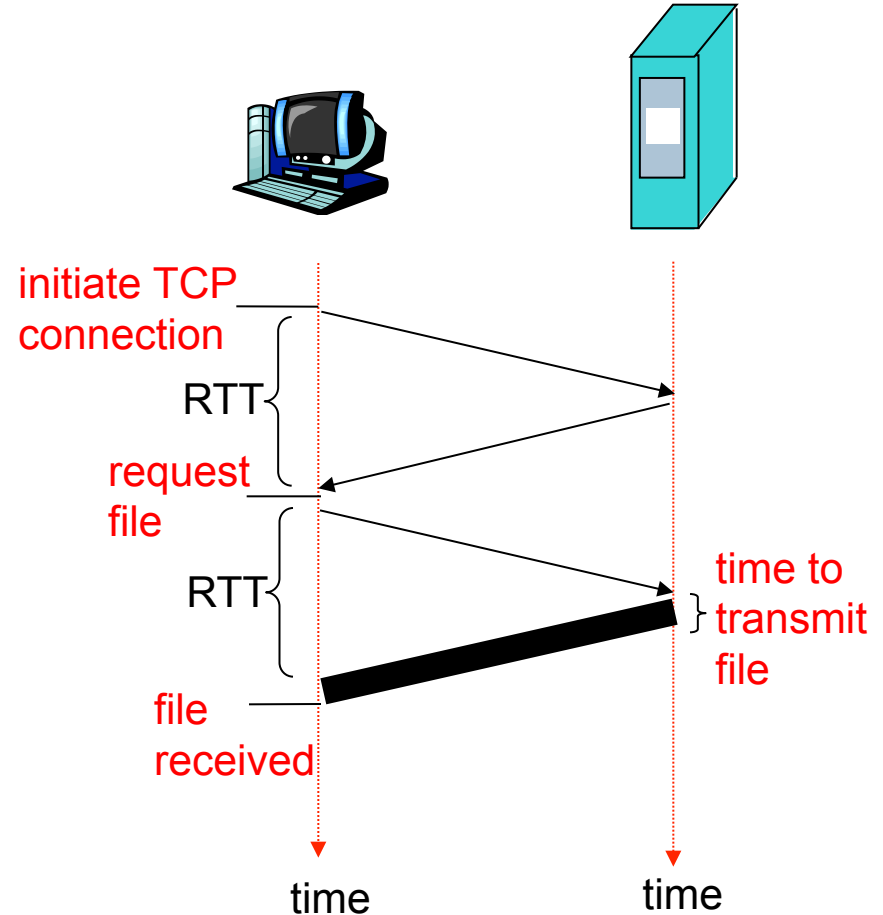
Non-Persistent HTTP: Response time

definition of RTT: time for a small packet to travel from client to server and back.

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

total = 2RTT + transmit time



Persistent HTTP

non-persistent HTTP issues:

- ▶ requires 2 RTTs per object
- ▶ OS overhead for *each* TCP connection

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



Advantage of non-persistent HTTP

non-persistent HTTP:

- ▶ browsers can open parallel TCP connections to fetch referenced objects “at the same time”
 - ▶ Has advantages and disadvantages



HTTP request message

`http://www-net.cs.umass.edu:8080/index.html`

- ▶ two types of HTTP messages: *request, response*
- ▶ **HTTP request message:**
 - ▶ ASCII (human-readable format)

request line
(GET, POST,
HEAD commands)

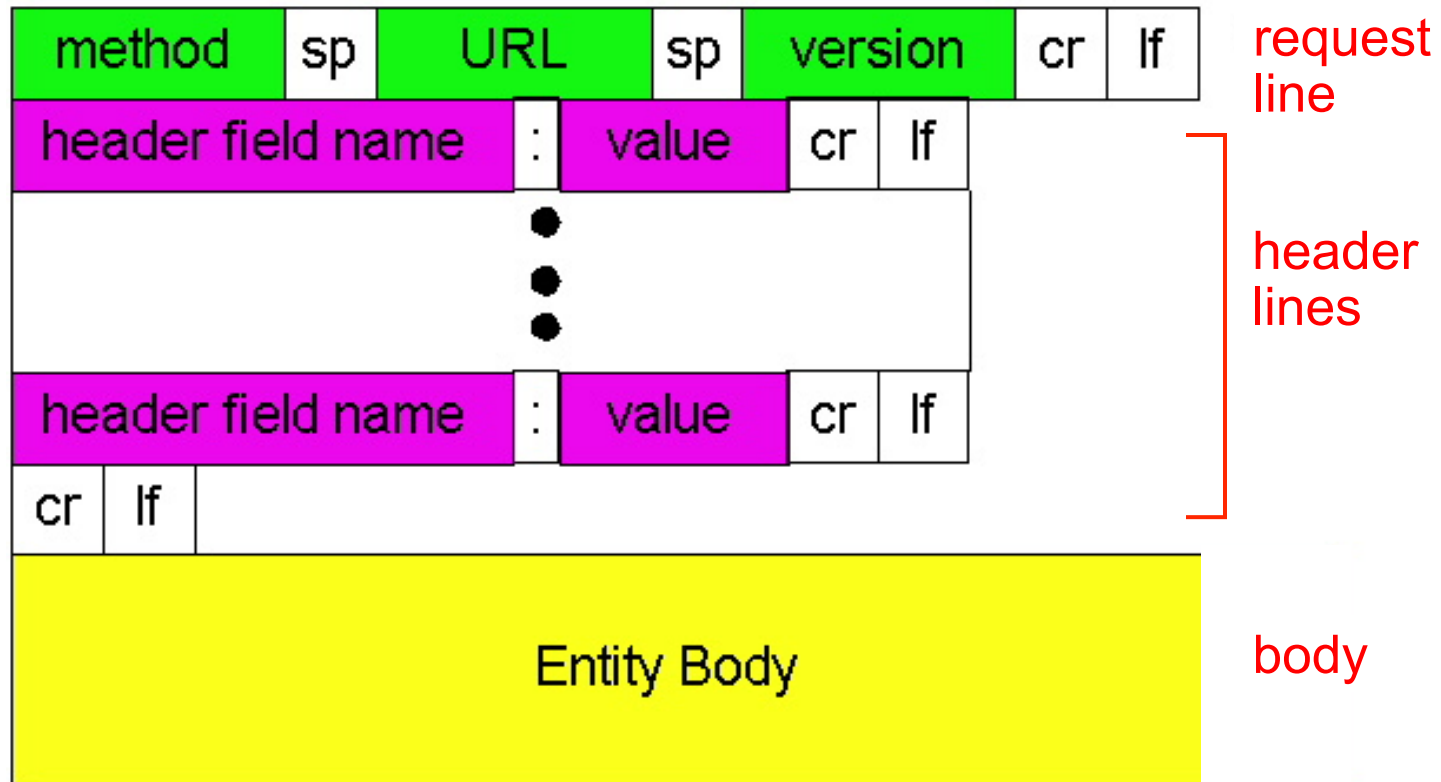
header
lines

carriage return,
line feed at start
of line indicates
end of header lines

```
GET /index.html HTTP/1.1\r\n
Host: www-net.cs.umass.edu\r\n
User-Agent: Firefox/3.6.10\r\n
Accept: text/html,application/xhtml+xml\r\n
Accept-Language: en-us,en;q=0.5\r\n
Accept-Encoding: gzip,deflate\r\n
Accept-Charset: ISO-8859-1,utf-8;q=0.7\r\n
Keep-Alive: 115\r\n
Connection: keep-alive\r\n
\r\n
```

carriage return character
line-feed character

HTTP request message: general format



A simple test... ****

- ▶ \$ nc -l 12345
- ▶ Point your browser to <http://127.0.0.1:12345/testme>
- ▶ If your user-agent looks strange and you curious to know why, read this:
 - ▶ <http://webaim.org/blog/user-agent-string-history/>

Uploading form input

POST method:

- ▶ web page often includes form input
- ▶ input is uploaded to server in entity body

URL method:

- ▶ uses GET method
- ▶ input is uploaded in URL field of request line:

www.somesite.com/animalsearch?monkeys&banana

`www.example.com/animalsearch.php?name=monkeys&age=10`



Method types

HTTP/1.0

- ▶ GET
- ▶ POST
- ▶ HEAD
 - ▶ asks server to leave requested object out of response

HTTP/1.1

- ▶ GET, POST, HEAD
- ▶ PUT
 - ▶ uploads file in entity body to path specified in URL field
- ▶ DELETE
 - ▶ deletes file specified in the URL field



HTTP response message

status line

(protocol

status code

status phrase)

HTTP/1.1 200 OK\r\n

Date: Sun, 26 Sep 2010 20:09:20 GMT\r\n

Server: Apache/2.0.52 (CentOS)\r\n

Last-Modified: Tue, 30 Oct 2007 17:00:02
GMT\r\n

header
lines

ETag: "17dc6-a5c-bf716880"\r\n

Accept-Ranges: bytes\r\n

Content-Length: 2652\r\n

Keep-Alive: timeout=10, max=100\r\n

Connection: Keep-Alive\r\n

Content-Type: text/html;
charset=ISO-8859-1\r\n

\r\n

data, e.g.,

requested

HTML file

data data data data data ...



HTTP response status codes

❖ status code appears in 1st line in server->client response message.

❖ some sample codes:

200 OK

- ▶ request succeeded, requested object later in this msg

301 Moved Permanently

- ▶ requested object moved, new location specified later in this msg
(Location:)

400 Bad Request

- ▶ request msg not understood by server

404 Not Found

- ▶ requested document not found on this server

505 HTTP Version Not Supported



Trying out HTTP (client side) for yourself

1. Telnet to your favorite Web server:

```
telnet www.uga.edu 80
```

opens TCP connection to port 80 (default HTTP server port) at cis.poly.edu. anything typed in sent to port 80 at cis.poly.edu

2. type in a GET HTTP request:

```
GET /profile/mission HTTP/1.1  
Host: www.uga.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!

(or use Wireshark!)

User-server state: cookies

many Web sites use cookies

four components:

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

example:

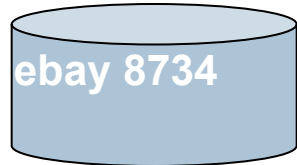
- ▶ Susan always access Internet from PC
- ▶ visits specific e-commerce site for first time
- ▶ when initial HTTP requests arrives at site, site creates:
 - ▶ unique ID
 - ▶ entry in backend database for ID



Cookies: keeping "state" (cont.)

client

server



cookie file



one week later:



usual http request msg

usual http response
Set-cookie: 1678

usual http request msg
cookie: 1678

usual http response msg

usual http request msg
cookie: 1678

usual http response msg

Amazon server
creates ID
1678 for user

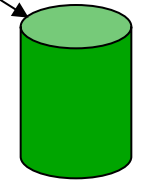
cookie-specific
action

cookie-specific
action

create
entry

access

access



backend
database

Cookies (continued)

what cookies can bring:

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

how to keep "state":

- ❖ protocol endpoints: maintain state at sender/receiver over multiple transactions
- ❖ cookies: http messages carry state

— aside —

cookies and privacy:

- ❖ cookies permit sites to learn a lot about you
- ❖ you may supply name and e-mail to sites



Cookies and Privacy

- ▶ Two types of cookies
 - ▶ Session cookies
 - ▶ Permanent cookies (tracking cookies)
- ▶ Third-party cookies (see <http://tools.ietf.org/html/rfc2965>)
 - ▶ You visit www.example.com, which contains a banner from ads.clicks-for-me.net
 - ▶ in simple terms ads.clicks-for-me.net is third-party because it does not match the domain showed on the URL bar
 - ▶ third-party sites should be denied setting or reading cookies
 - ▶ The browser allows ads.clicks-for-me.net to drop a third-party cookie
 - ▶ Then you visit www.another-example.com , which also loads ads from ads.clicks-for-me.net
 - ▶ ads.clicks-for-me.net can track the fact that you visited both www.example.com and www.another-example.com !!!

Cookies and Security

- ▶ **Authentication Cookies can be stolen**
 - ▶ An attacker may be able to “sniff” your authentication cookies
 - ▶ The attacker will be able to login as you on a website (e.g., Facebook, Twitter, etc...)

- ▶ See FireSheep for a concrete example!
 - ▶ <http://codebutler.com/firesheep>

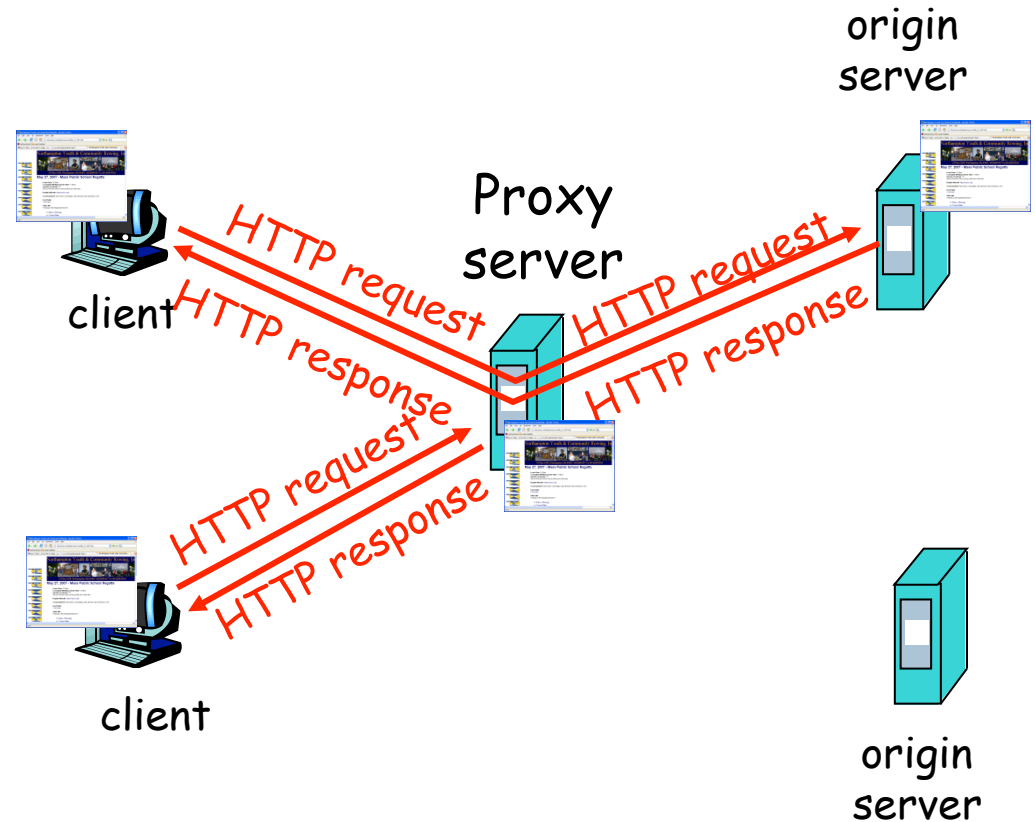
Session IDs

- ▶ Cookies are not the only way you can keep state
 - ▶ Session IDs are commonly used by web applications
 - ▶ http://example.com/index.php?user_id=0F4C26A1&topic=networking
- ▶ What are the main difference between cookies and Session IDs?
 - ▶ Session IDs are typically passed in the URL (added to web app links)
 - ▶ Cookies are passed through HTTP req/resp headers
 - ▶ Cookies are stored in the browser's cache and have an expiration date
 - ▶ Session IDs are volatile: never stored, only used until end of session

Web caches (proxy server)

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
 - ▶ else cache requests object from origin server, then returns object to client



More about Web caching

- ▶ cache acts as both client and server
 - ▶ Splits the TCP connection!
- ▶ typically cache is installed by ISP (university, company, residential ISP)

why Web caching?

- ▶ reduce response time for client request
- ▶ reduce traffic on an institution's access link.
- ▶ Internet dense with caches: enables “poor” content providers to effectively deliver content (but so does P2P file sharing)

Caching in HTTP

<http://www.w3.org/Protocols/rfc2616/rfc2616-sec13.html>

Caching example

assumptions

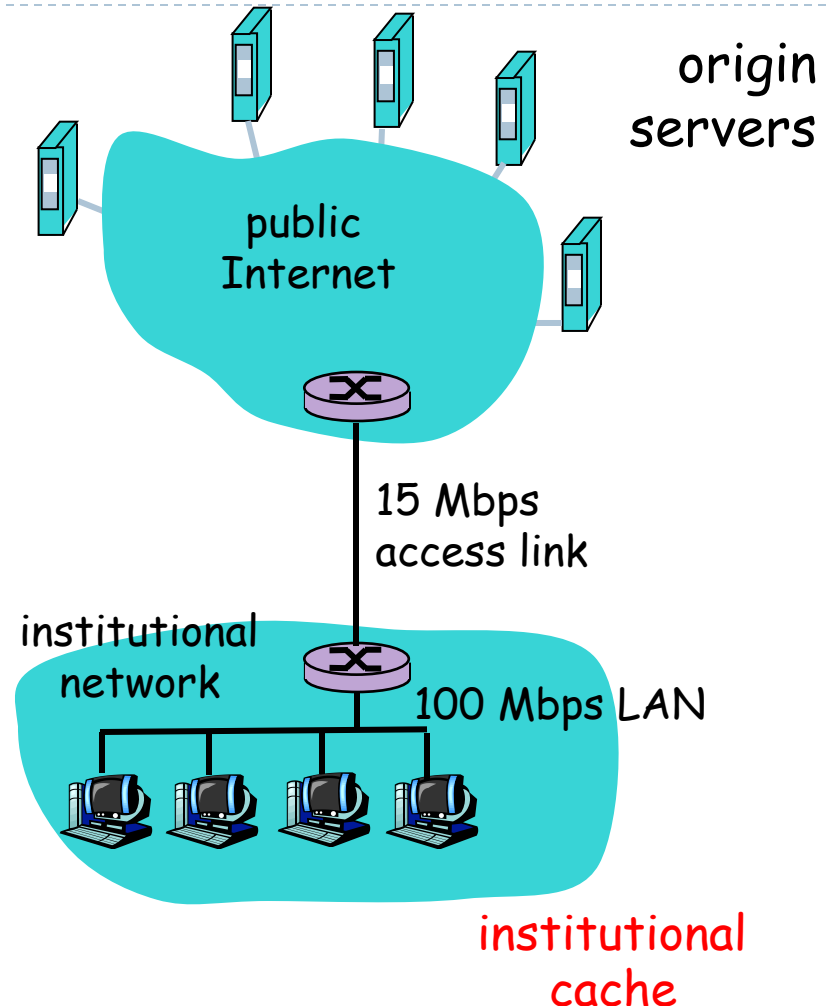
- ▶ average object size = 1M bits
- ▶ avg. request rate from institution's browsers to origin servers = 15/sec
- ▶ delay from "Internet router" to any origin server and back to router = 2 sec

consequences

- ▶ utilization on LAN = 15%
- ▶ utilization on access link = 100%
- ▶ total delay = Internet delay + access delay + LAN delay

$$= 2 \text{ sec} + \text{minutes} + \text{milliseconds}$$

↑
Due to traffic intensity = 1
on the access link



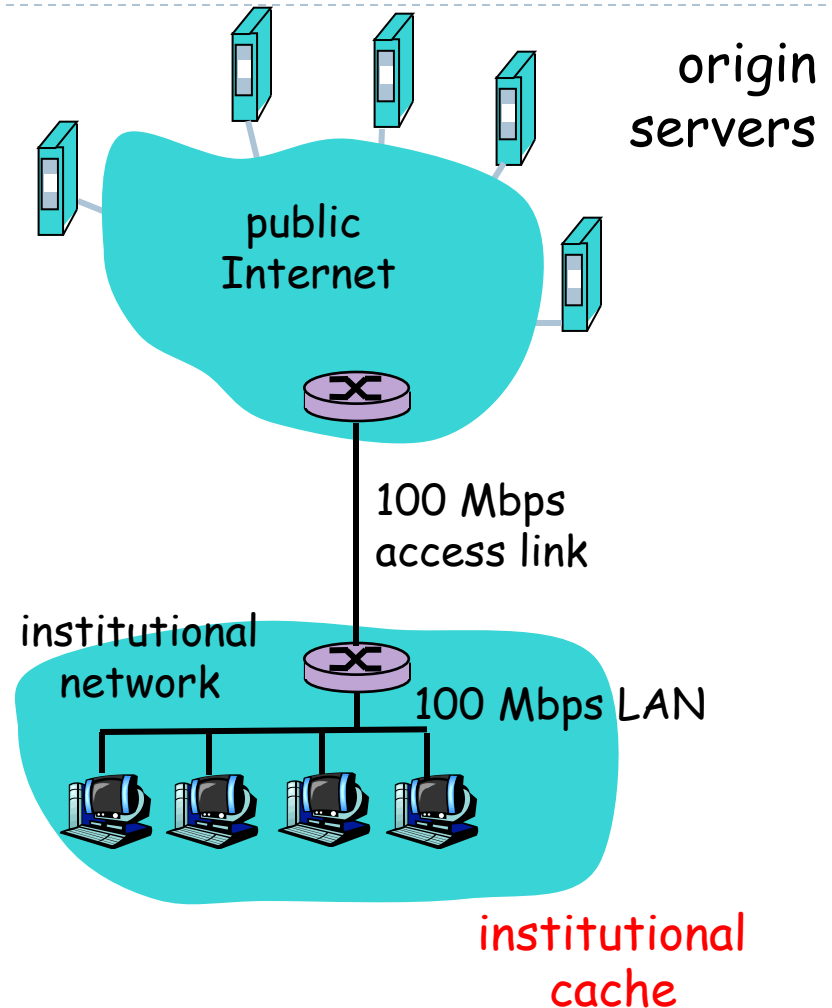
Caching example (cont)

possible solution

- ▶ increase bandwidth of access link to, say, 100 Mbps

consequence

- ▶ utilization on LAN = 15%
- ▶ utilization on access link = 15%
- ▶ Total delay = Internet delay + access delay + LAN delay
= 2 sec + msec + msec
- ▶ often a costly upgrade



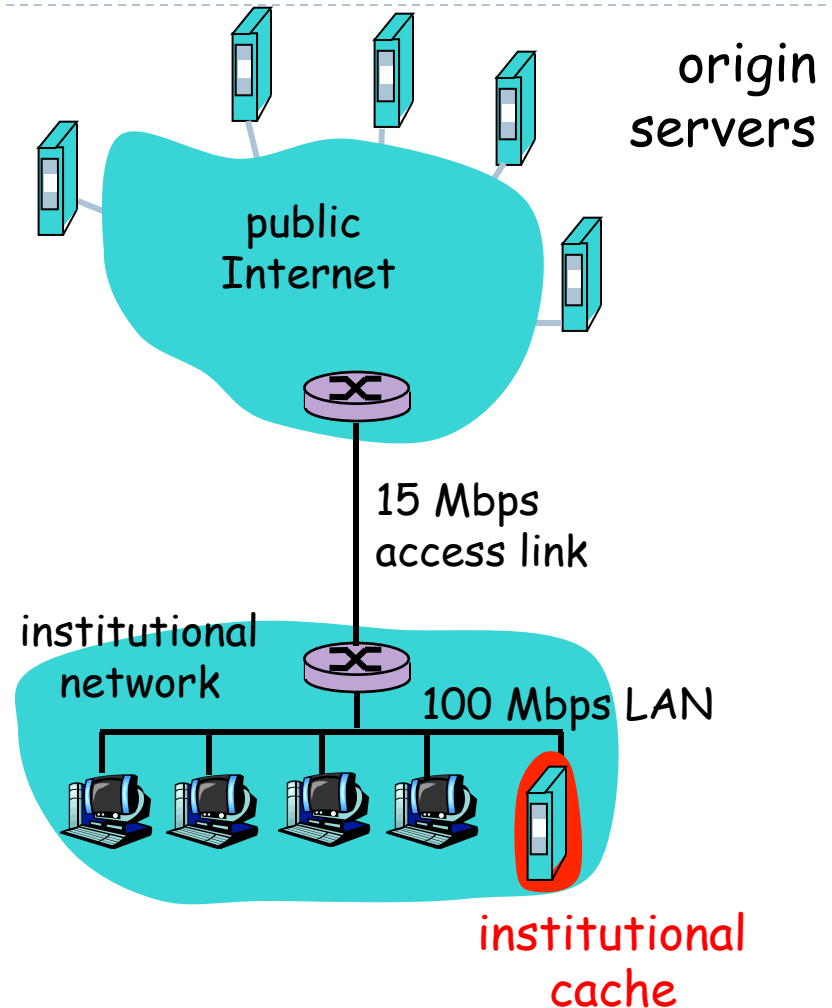
Caching example (cont)

possible solution:

- ▶ install cache

consequence

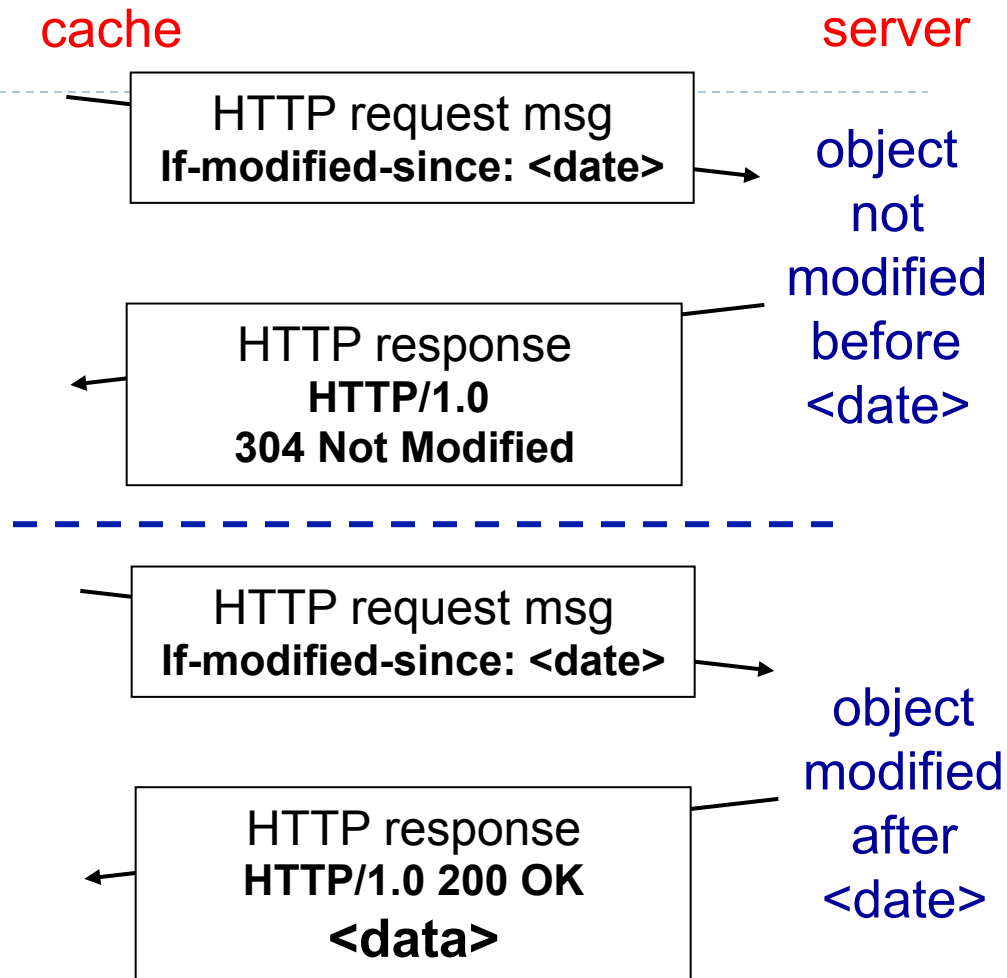
- ▶ suppose hit rate is 0.4
 - ▶ 40% requests will be satisfied almost immediately
 - ▶ 60% requests satisfied by origin server
- ▶ utilization of access link reduced to 60%, resulting in negligible delays (say 10 msec)
- ▶ total avg delay = Internet delay + access delay + LAN delay =
 $0.6 \cdot (2.01) \text{ secs} + 0.4 \cdot \text{milliseconds} < 1.4 \text{ secs}$



Conditional GET

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



In reality, cache entry validation and eviction policies are quite complex

<http://www.w3.org/Protocols/rfc2616/rfc2616-sec13.html#sec13>

HTTP Pipelining and Range

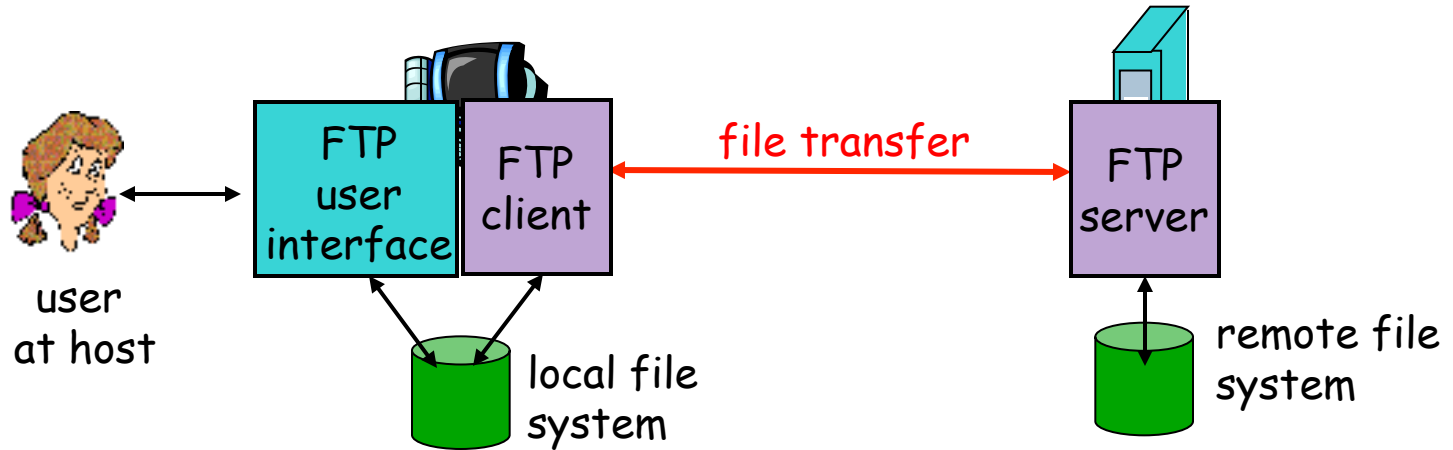
▶ Pipelining

- ▶ The client sends multiple HTTP request without waiting for server response
- ▶ The server sends the response one after the other

▶ Range

- ▶ HTTP allows downloading pieces of objects
- ▶ Example:
 - ▶ 10MB image to be downloaded
 - ▶ We can open 10 different TCP connection and send 10 HTTP requests in parallel
 - ▶ Download 1MB of data from each connection and stitch them back together

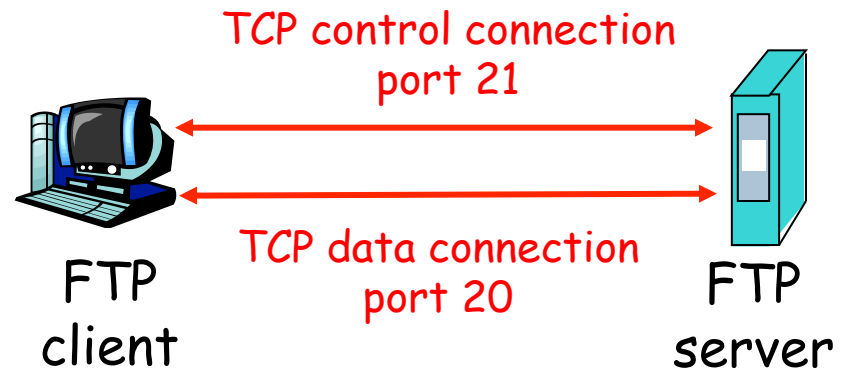
FTP: the file transfer protocol



- ▶ 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, TCP is transport protocol
- ▶ client authorized over control connection
- ▶ client browses remote directory by sending commands over control connection.
- ▶ when server receives file transfer command, server opens 2nd TCP connection (for file) to client
- ▶ after transferring one file, server closes data connection.

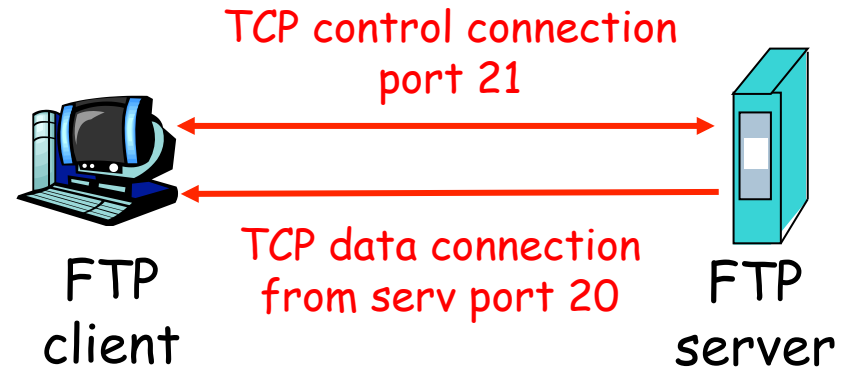


- ❖ server opens another TCP data connection to transfer another file.
- ❖ control connection: "out of band"
- ❖ FTP server maintains "state": current directory, earlier authentication

FTP: separate control, data connections

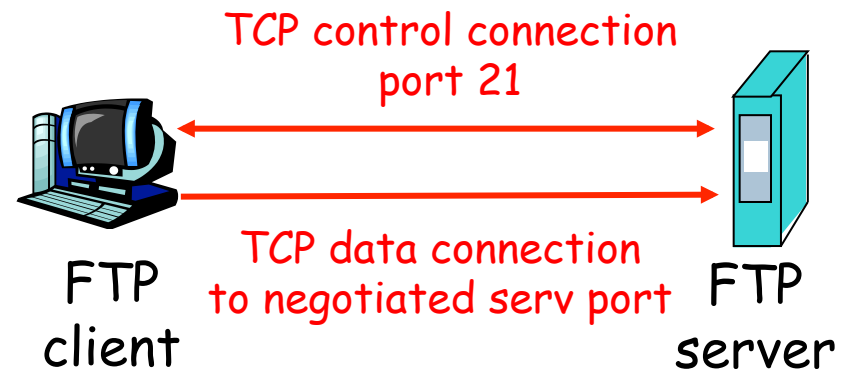
▶ **Active FTP**

FTP server contacts client from TCP src-port 20 to negotiated dst-port



▶ **Passive FTP**

client contacts FTP server at negotiated TCP dst-port



When is Passive FTP useful?

FTP commands, responses

sample commands:

- ▶ sent as ASCII text over control channel
- ▶ **USER *username***
- ▶ **PASS *password***
- ▶ **LIST** return list of file in current directory
- ▶ **RETR *filename*** retrieves (gets) file
- ▶ **STOR *filename*** stores (puts) file onto remote host

sample return codes

- ▶ status code and phrase (as in HTTP)
- ▶ **331 Username OK, password required**
- ▶ **125 data connection already open; transfer starting**
- ▶ **425 Can't open data connection**
- ▶ **452 Error writing file**

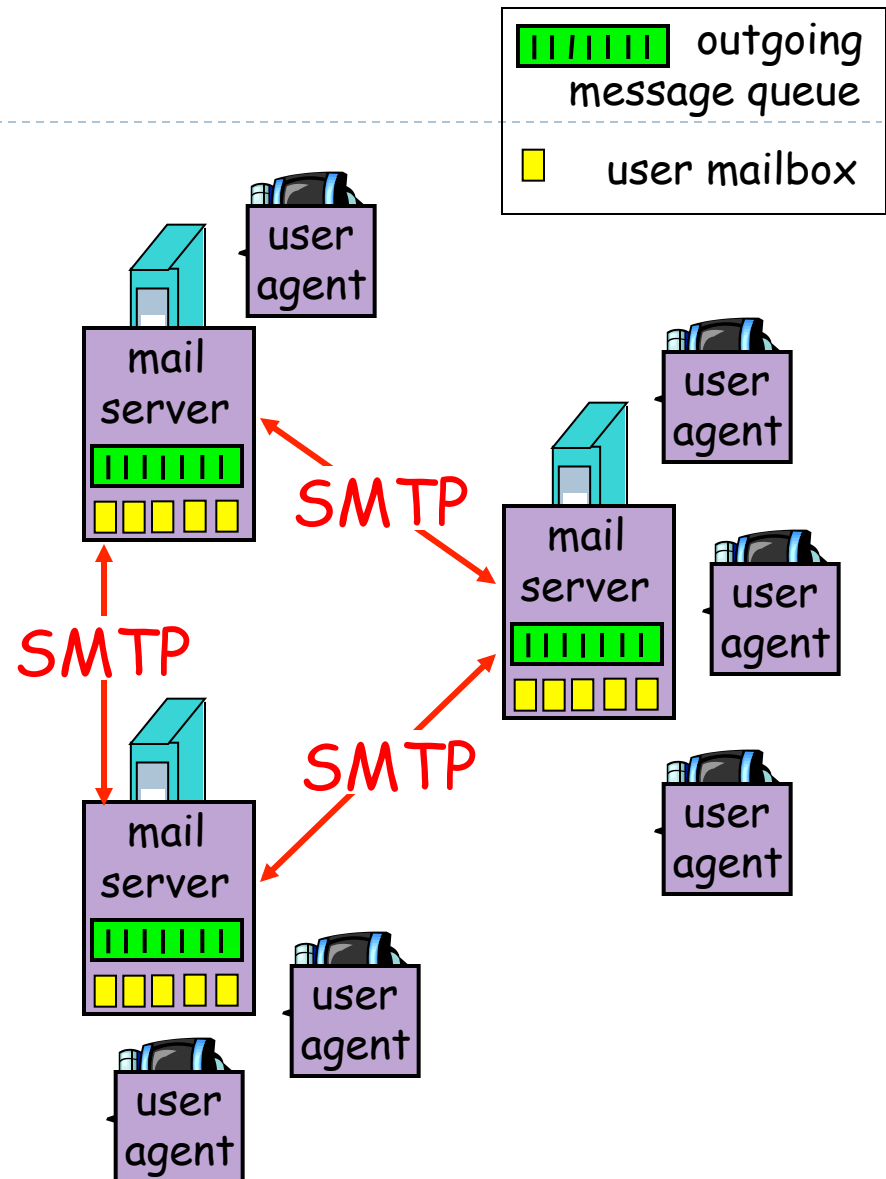
Electronic Mail

Three major components:

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

User Agent

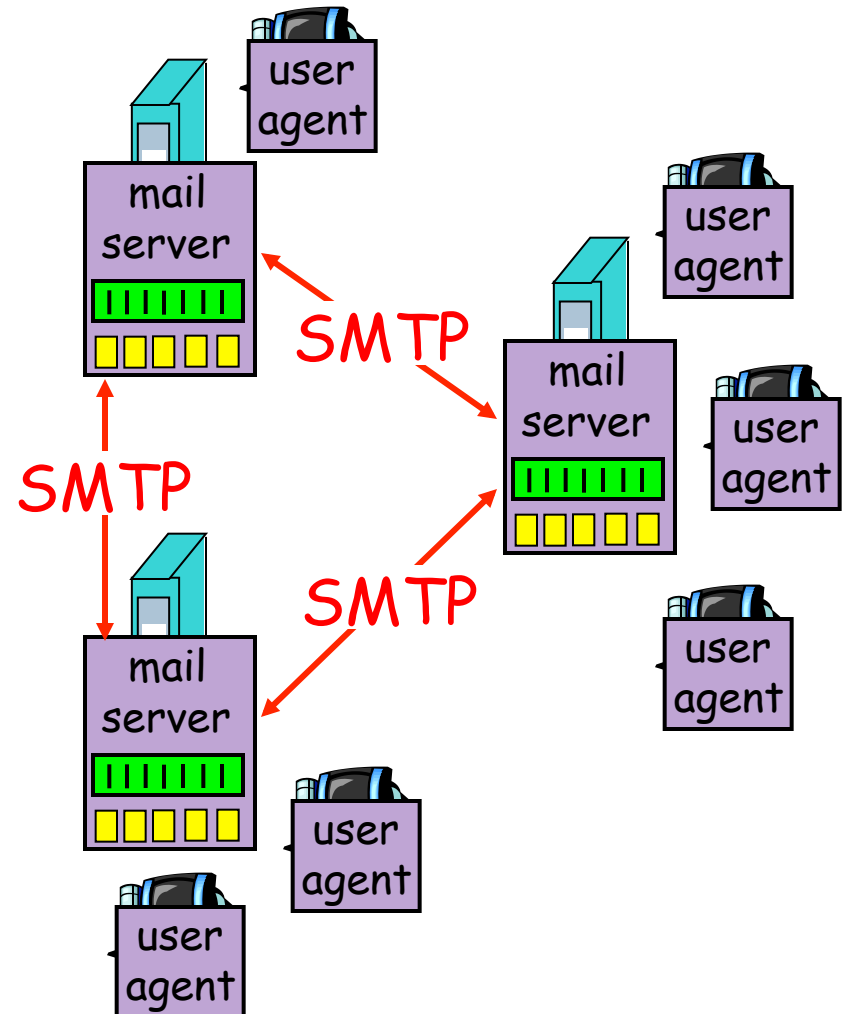
- ▶ a.k.a. “mail reader”
- ▶ composing, editing, reading mail messages
- ▶ e.g., Eudora, Outlook, elm, Mozilla Thunderbird
- ▶ outgoing, incoming messages stored on server



Electronic 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

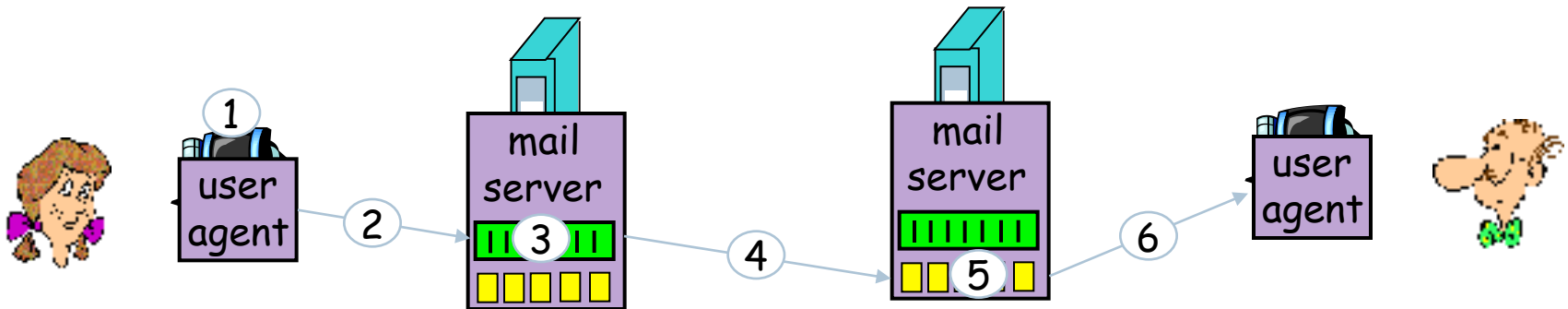


Electronic Mail: SMTP [RFC 2821]

- ▶ uses TCP to reliably transfer email message from client to server, port 25
- ▶ direct transfer: sending server to receiving server
- ▶ three phases of transfer
 - ▶ handshaking (greeting)
 - ▶ transfer of messages
 - ▶ closure
- ▶ command/response interaction
 - ▶ **commands:** ASCII text
 - ▶ **response:** status code and phrase
- ▶ messages must be in 7-bit ASCII

Scenario: Alice sends message to Bob

- 1) Alice uses UA to compose message and “to”
`bob@some school . 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
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: From: Alice
C: To: Bob
C: Subject: Quick question
C: Do you like ketchup?
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 hamburger.edu closing connection
```

Try SMTP interaction for yourself:

- ▶ `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)

Concrete example***

```
$ dig +short -t MX uga.edu
10 1282373658.mail.outlook.com.
$ dig +short -x 198.137.20.113
h198-137-20-113.paws.uga.edu.
$ telnet 1282373658.mail.outlook.com. 25
Trying 216.32.181.178...
Connected to 1282373658.mail.outlook.com.
Escape character is '^]'.
220 CH1EHSMHS014.bigfish.com Microsoft ESMTMP MAIL Service ready at Tue, 29 Jan 2013 15:20:08
HELO h198-137-20-113.paws.uga.edu
250 CH1EHSMHS014.bigfish.com Hello [128.192.4.39]
MAIL FROM: <perdisci@cs.uga.edu>
250 2.1.0 Sender OK
RCPT TO: <perdisci@uga.edu>
250 2.1.5 Recipient OK
DATA
354 Start mail input; end with <CRLF>.<CRLF>
From: Roberto <perdisci@cs.uga.edu>
To: Rob <perdisci@uga.edu>
Subject: Quick question

Do you like ketchup?

.
250 2.6.0 <...ehs.local> [InternalId=21919093] Queued mail for delivery
QUIT
221 2.0.0 Service closing transmission channel
Connection closed by foreign host.
```

Mail message format

SMTP: protocol for exchanging email
msgs

RFC 822: standard for text message
format:

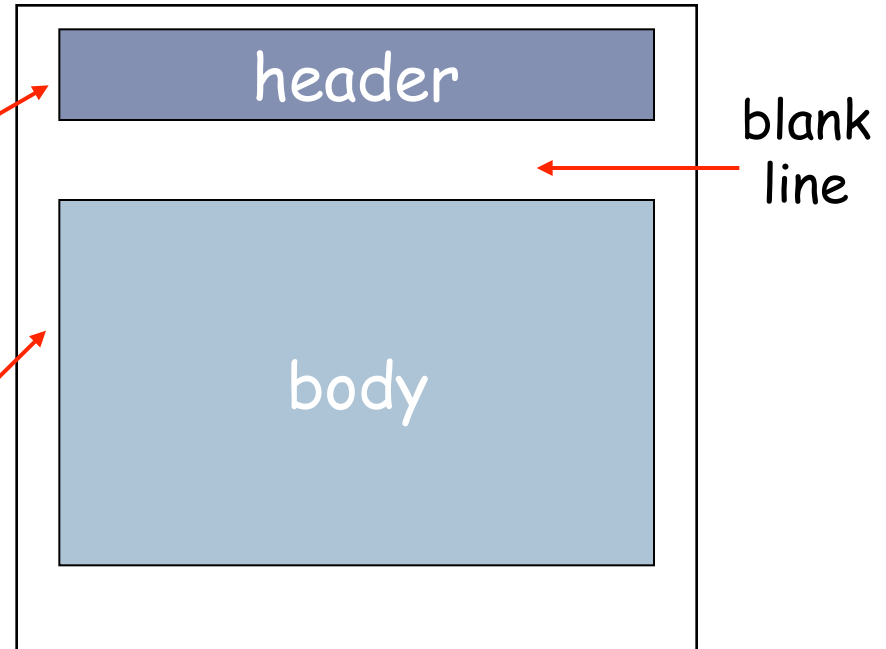
▶ header lines, e.g.,

- ▶ To:
- ▶ From:
- ▶ Subject:

different from SMTP commands!

▶ body

- ▶ the “message”, ASCII characters
only



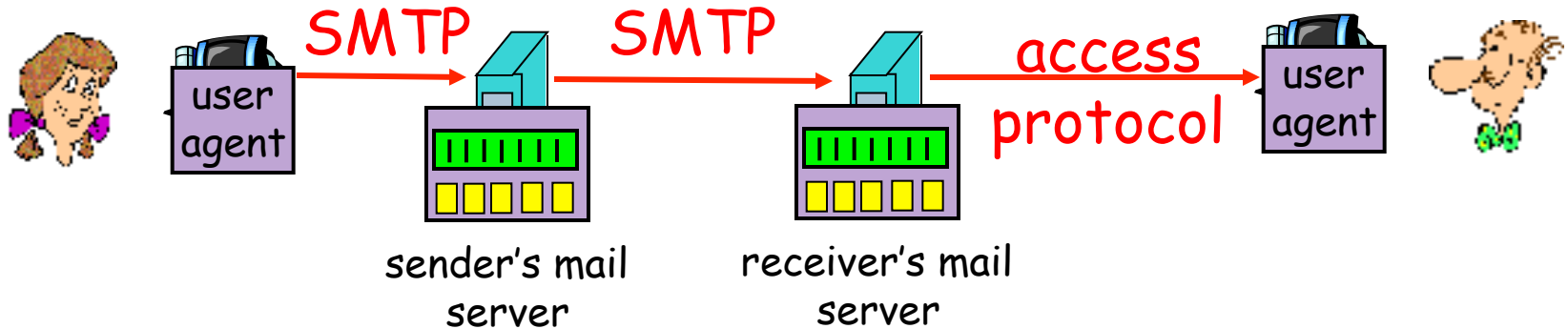
SMTP: final words

- ▶ 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:

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

Mail access protocols



- ▶ SMTP: delivery/storage to receiver's server
- ▶ mail access protocol: retrieval from server
 - ▶ POP: Post Office Protocol [RFC 1939]
 - ▶ authorization (agent <-->server) and download
 - ▶ IMAP: Internet Mail Access Protocol [RFC 1730]
 - ▶ more features (more complex)
 - ▶ manipulation of stored msgs on server
 - ▶ HTTP: gmail, Hotmail, Yahoo! Mail, etc.

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 “download and delete” mode.
- ▶ Bob cannot re-read e-mail if he changes client
- ▶ “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



Chapter 2: Application layer

- ▶ 2.1 Principles of network applications
- ▶ 2.2 Web and HTTP
- ▶ 2.3 FTP
- ▶ 2.4 Electronic Mail
 - ▶ SMTP, POP3, IMAP
- ▶ **2.5 DNS**
- ▶ 2.6 P2P applications
- ▶ 2.7 Socket programming with TCP
- ▶ 2.8 Socket programming with UDP

DNS: Domain Name System

people: many identifiers:

- ▶ SSN, name, passport #

Internet hosts, routers:

- ▶ IP address (32 bit) - used for addressing datagrams
- ▶ “name”, e.g., ww.yahoo.com - used by humans

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

Domain Name System:

- ▶ *distributed database* implemented in hierarchy of many *name servers*
- ▶ *application-layer protocol* host, routers, name servers to communicate to *resolve* names (address/name translation)
 - ▶ note: core Internet function, implemented as application-layer protocol
 - ▶ complexity at network’s “edge”

DNS

DNS services

- ▶ hostname to IP address translation
- ▶ host aliasing
 - ▶ Canonical, alias names
- ▶ mail server aliasing
- ▶ load distribution
 - ▶ replicated Web servers: set of IP addresses for one canonical name

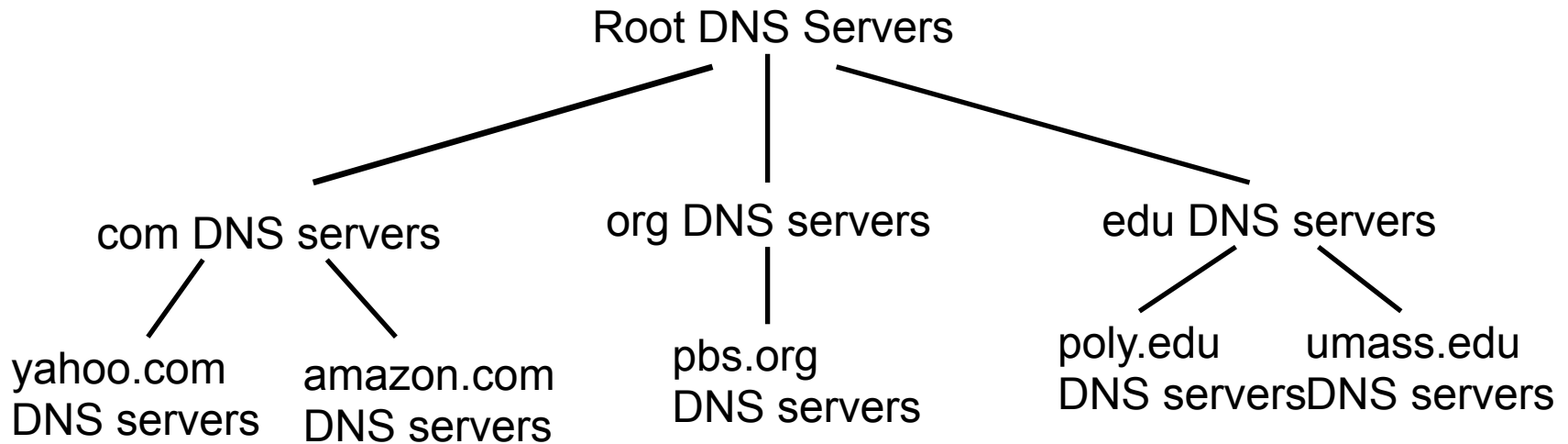
Why not centralize DNS?

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

doesn't scale!



Distributed, Hierarchical Database



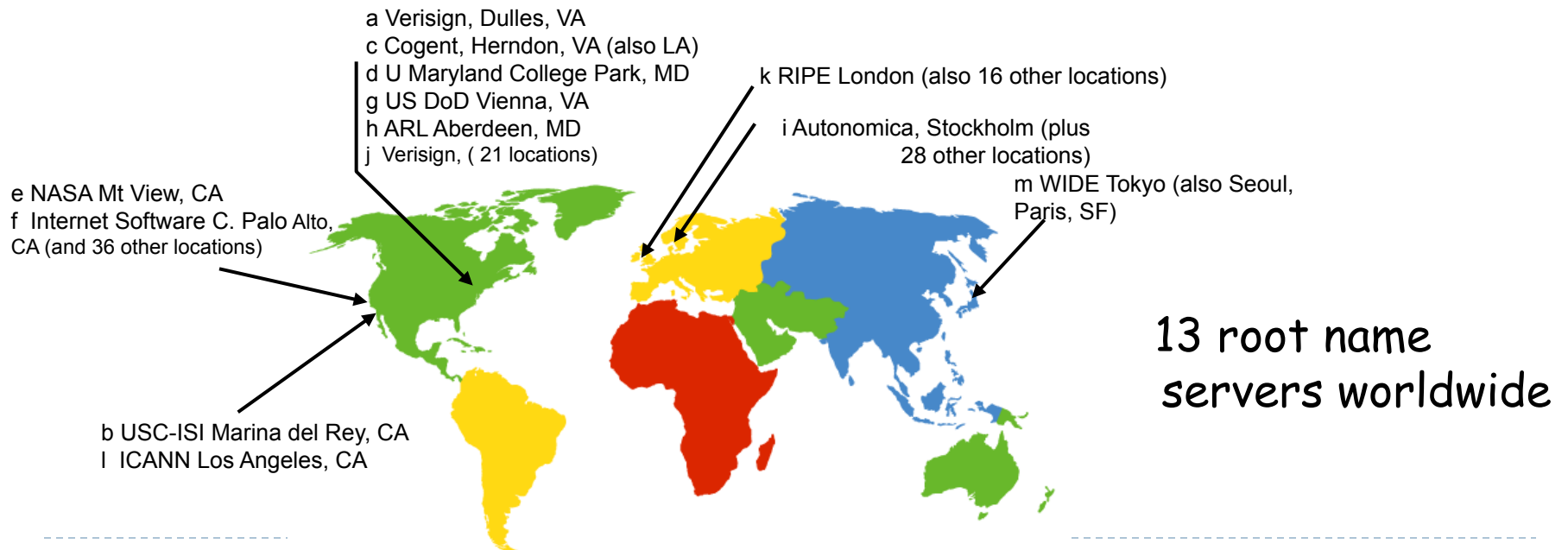
client wants IP for www.amazon.com; 1st approx:

- ▶ client queries a 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

<http://www.internetsociety.org/sites/default/files/DNS%20Root%20Name%20Servers%20Frequently%20Asked%20Questions.doc.pdf>

- ▶ 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



TLD 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.: uk, fr, ca, jp.
- ▶ Network Solutions maintains servers for com TLD
- ▶ Educause for edu TLD

Authoritative DNS servers:

- ▶ organization's DNS servers, providing authoritative hostname to IP mappings for organization's servers (e.g., Web, mail).
- ▶ can be maintained by organization or service provider

Local Name Server

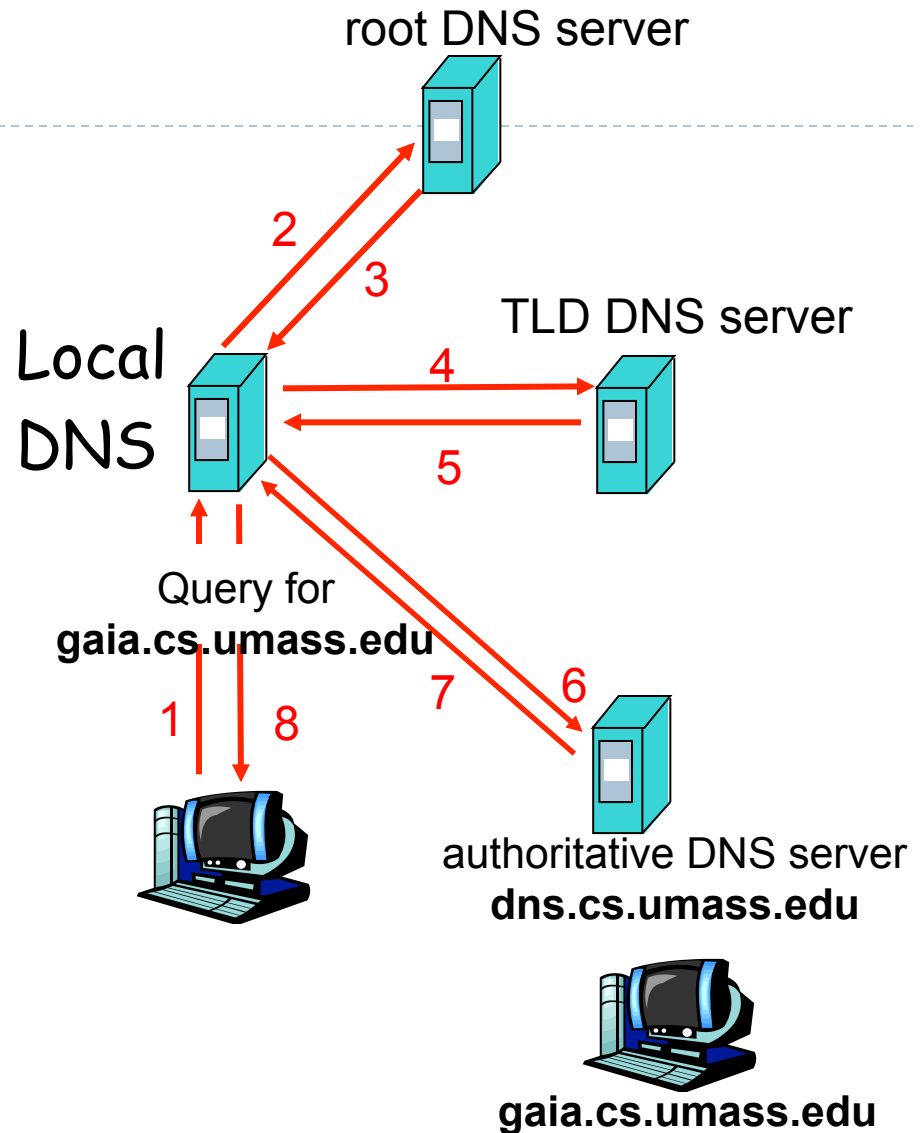
- ▶ does not strictly belong to hierarchy
- ▶ each ISP (residential ISP, company, university) has one.
 - ▶ also called “default name server”
- ▶ when host makes DNS query, query is sent to its local DNS server
 - ▶ acts as proxy, forwards query into hierarchy

DNS name *** resolution example

- ▶ host at cis.poly.edu wants IP address for gaia.cs.umass.edu

iterated query:

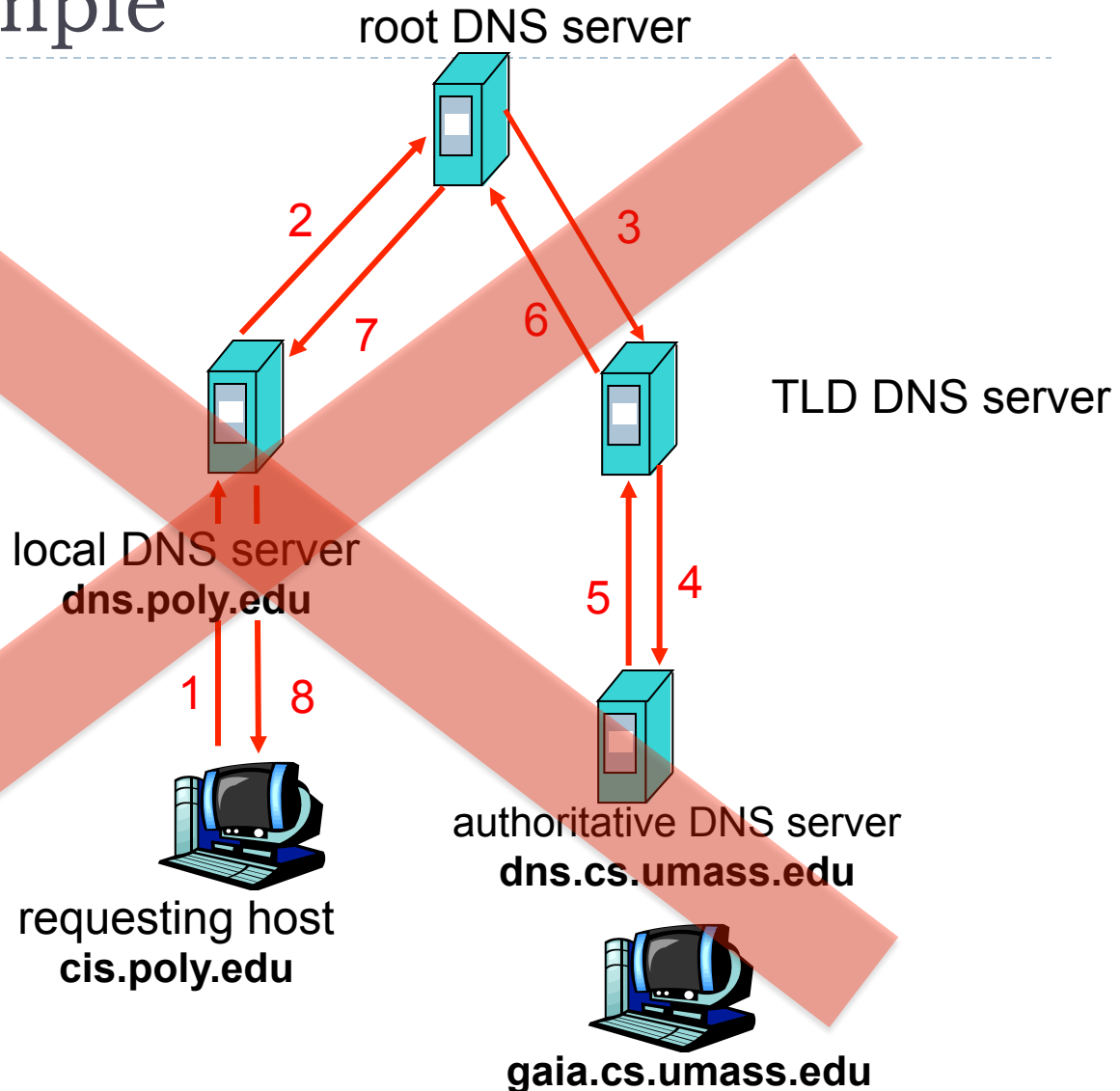
- ❖ contacted server replies with name of server to contact
- ❖ "I don't know this name, but ask this server"



DNS name *** resolution example

recursive query:

- ❖ puts burden of name resolution on contacted name server
- ❖ heavy load?



DNS: caching and updating records

- ▶ once (any) name server learns mapping, it *caches* mapping
 - ▶ cache entries timeout (disappear) after some time
 - ▶ TLD servers typically cached in local name servers
 - ▶ Thus root name servers not often visited

DNS records

DNS: distributed db 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 mailserver associated with name

DNS protocol, messages

DNS protocol : *query* and *reply* messages, both with same *message format*

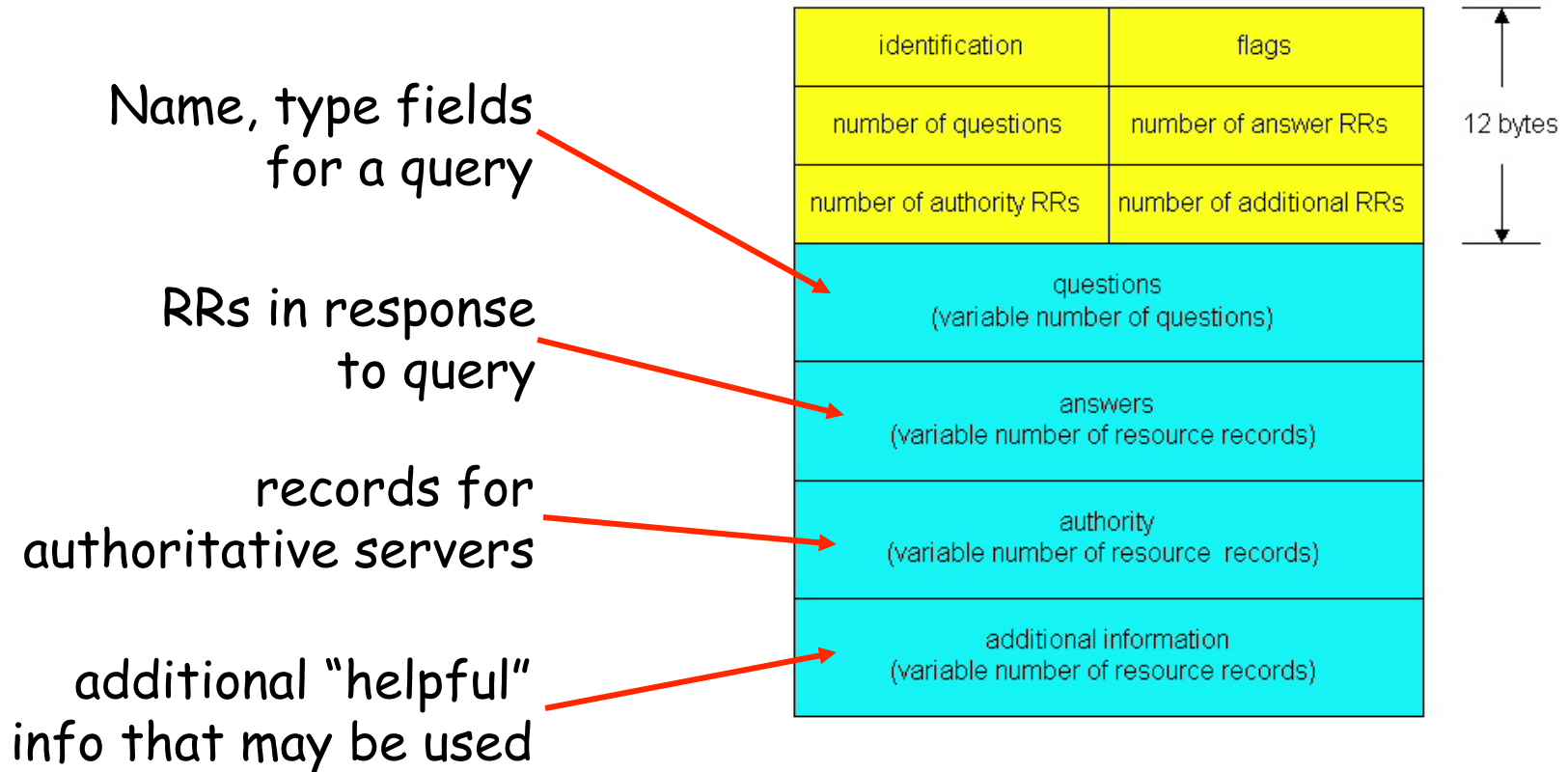
msg header

- ❖ **identification**: 16 bit #
for query, reply to query
uses same #
- ❖ **flags**:
 - query or reply
 - recursion desired
 - recursion available
 - reply is authoritative

identification	flags
number of questions	number of answer RRs
number of authority RRs	number of additional RRs
questions (variable number of questions)	
answers (variable number of resource records)	
authority (variable number of resource records)	
additional information (variable number of resource records)	



DNS protocol, messages



Inserting records into DNS

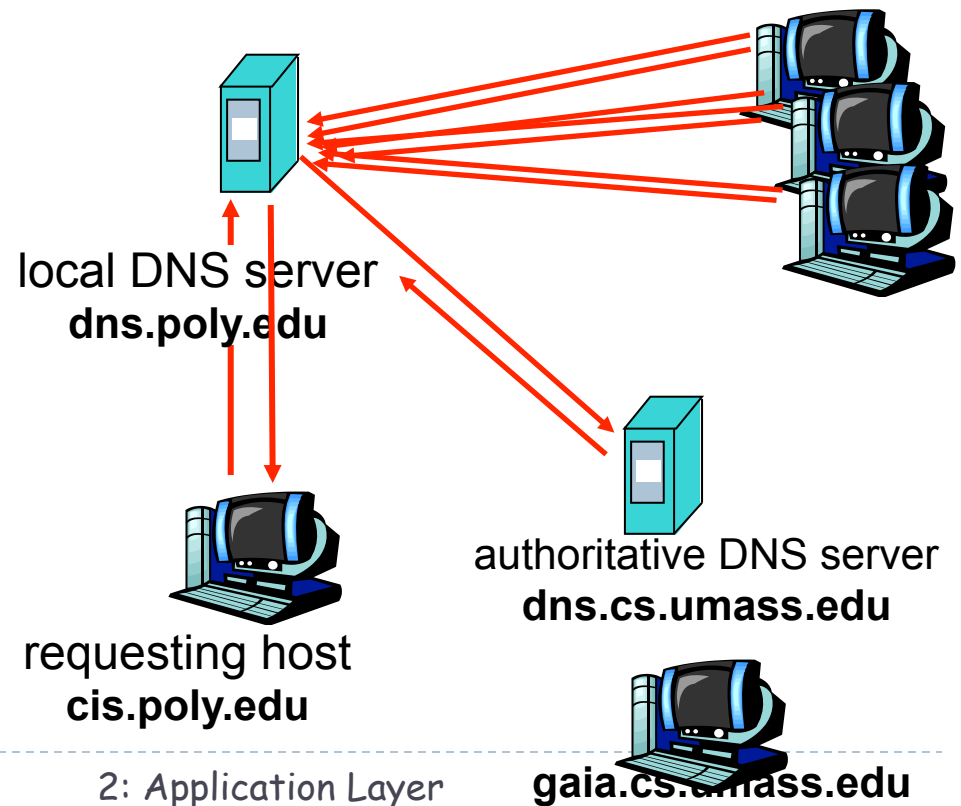
- ▶ example: new startup “Network Utopia”
- ▶ register name networkutopia.com at *DNS registrar* (e.g., Network Solutions)
 - ▶ provide names, IP addresses of authoritative name server (primary and secondary)
 - ▶ registrar inserts two RRs into com TLD server:

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

- ▶ create authoritative server Type A record for `www.networkutopia.com`; Type MX record for `networkutopia.com`
- ▶ **How do people get IP address of your Web site?**

DNS Poisoning

- ▶ DNS uses UDP
- ▶ Source IP address can be spoofed
- ▶ Responses are accepted with a “First Comes First Wins” policy, subsequent
 - ▶ Only check is on TXID
- ▶ What consequences?



DNSSEC

- ▶ DNS “patches”

- ▶ Port randomization
- ▶ 0x20-Bit encoding

- ▶ Better solution: DNSSEC

- ▶ Responses are digitally signed
- ▶ They can be verified by following a *chain of trust* anchored at the roots
- ▶ Not yet fully deployed

Chapter 2: Application layer

2.1 Principles of network applications

2.2 Web and HTTP

2.3 FTP

2.4 Electronic Mail

- ▶ SMTP, POP3, IMAP

2.5 DNS

2.6 P2P applications

2.7 Socket programming with TCP

2.8 Socket programming with UDP

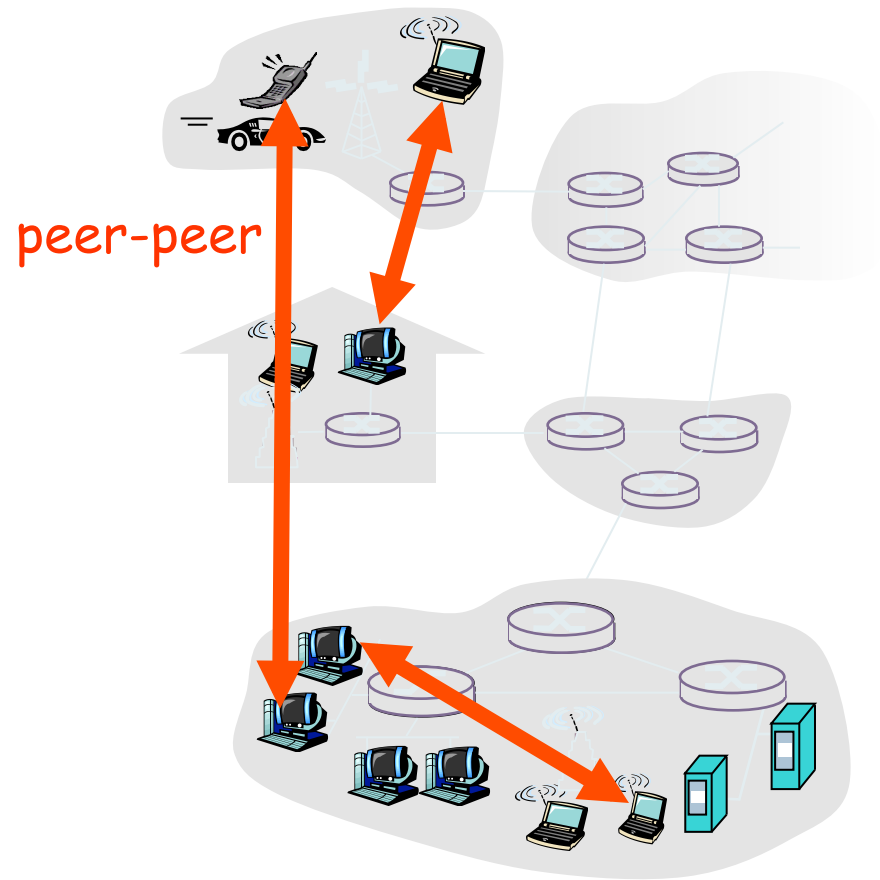


Pure P2P architecture

- ▶ *no* always-on server
- ▶ arbitrary end systems directly communicate
- ▶ peers are intermittently connected and change IP addresses

Applications:

- ▶ file distribution
- ▶ searching for information
- ▶ case Study: Skype



Distributed Hash Table (DHT)

- ▶ **Problem:**
 - ▶ Build a simple DB that can store (key, value) pairs
 - ▶ key: ss number; value: human name
 - ▶ key: file name; value: IP address of peers that have file
 - ▶ Clients can provide a key, and get the value from DB
 - ▶ Centralized solution is trivial (e.g., Napster)
- ▶ **DHT: distributed P2P database**
 - ▶ No central authority
 - ▶ Data distributed across very large number of (unreliable) nodes
- ▶ database has **(key, value)** pairs;
- ▶ peers **query** DB with key
 - ▶ DB returns values that match the key
- ▶ peers can also **insert** (key, value) pairs

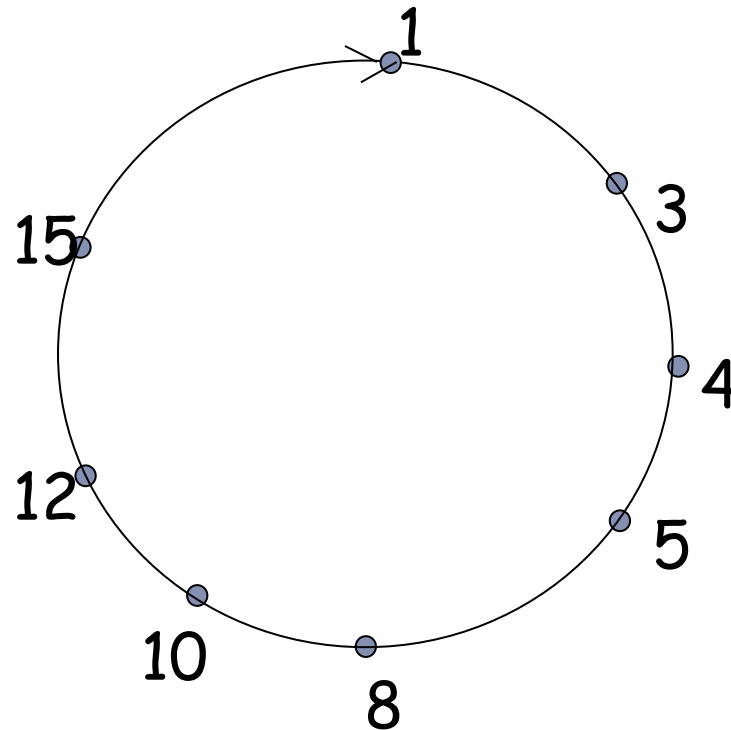
DHT Identifiers

- ▶ assign integer identifier to each peer in range $[0, 2^n - 1]$.
 - ▶ Each identifier can be represented by n bits.
- ▶ require each key to be an integer in **same range**.
- ▶ to get integer keys, hash original key.
 - ▶ e.g., $\text{key} = h(\text{"Led Zeppelin IV"})$
 - ▶ this is why they call it a distributed "hash" table

How to assign keys to peers?

- ▶ central issue:
 - ▶ assigning (key, value) pairs to peers.
- ▶ rule: assign key to the peer that has the **closest** ID.
- ▶ convention in lecture: closest is the **immediate successor** of the key.
- ▶ e.g.,: $n=4$; peers: 1,3,4,5,8,10,12,15;
 - ▶ key = 13, then successor peer = 15
 - ▶ key = 15, then successor peer = 15

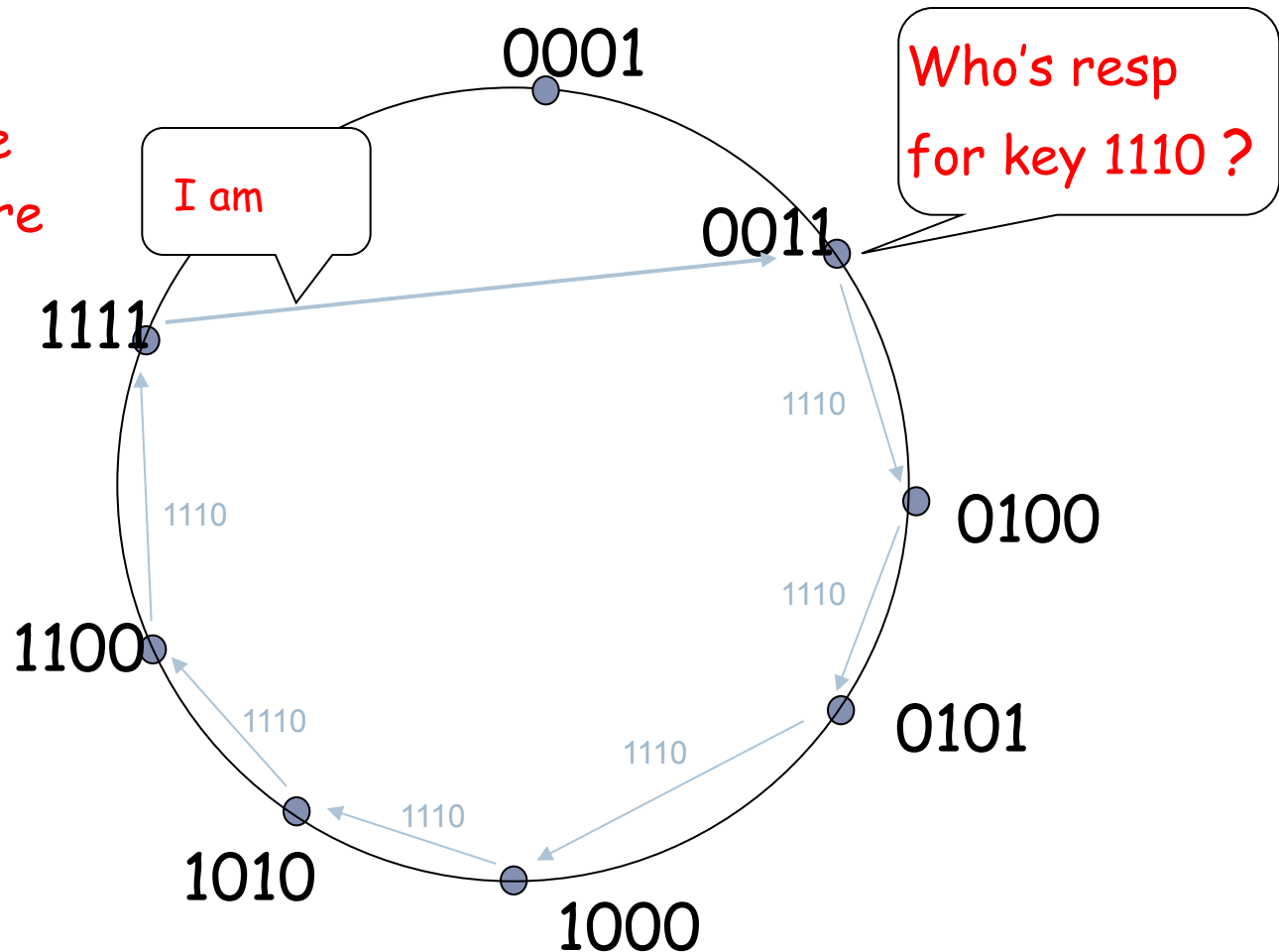
Circular DHT (1)



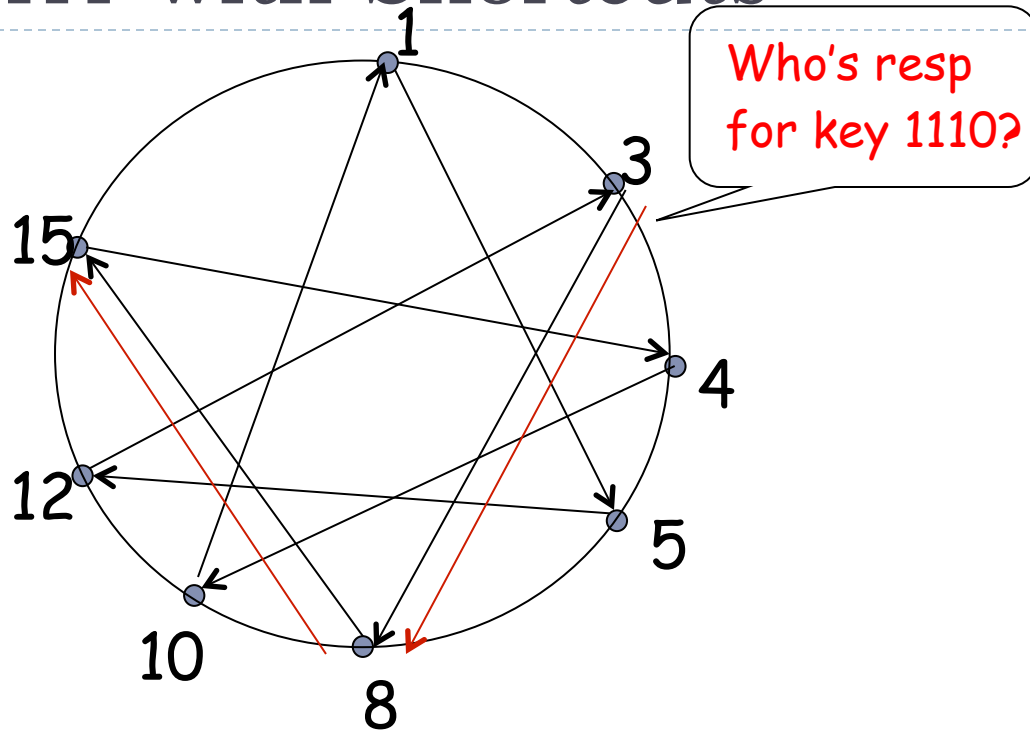
- ▶ each peer *only* aware of immediate successor and predecessor.
- ▶ “overlay network”

Circular DHT (2) ***

$O(N)$ messages on avg to resolve query, when there are N peers

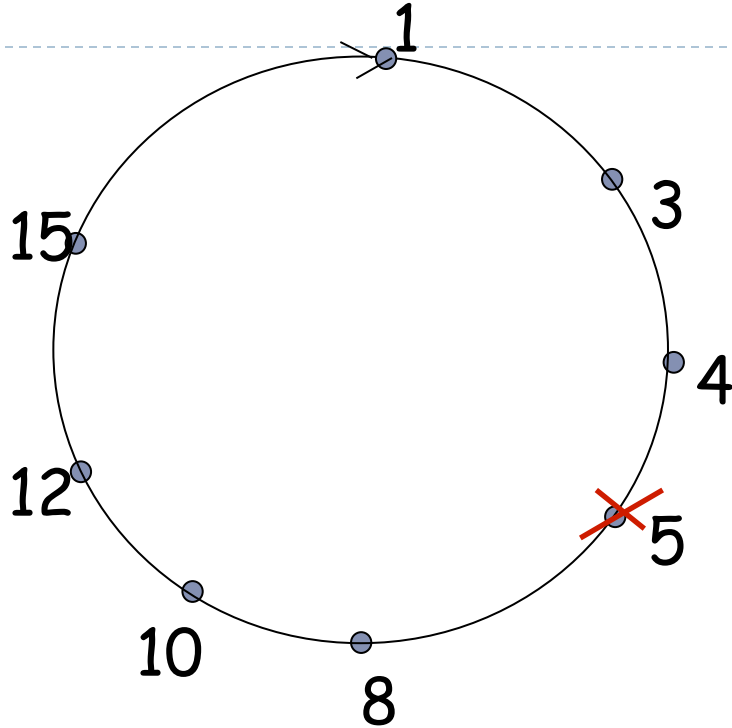


Circular DHT with Shortcuts



- ▶ each peer keeps track of IP addresses of predecessor, successor, short cuts.
- ▶ reduced from 6 to 2 messages.
- ▶ possible to design shortcuts so $O(\log N)$ neighbors, $O(\log N)$ messages in query

Peer Churn

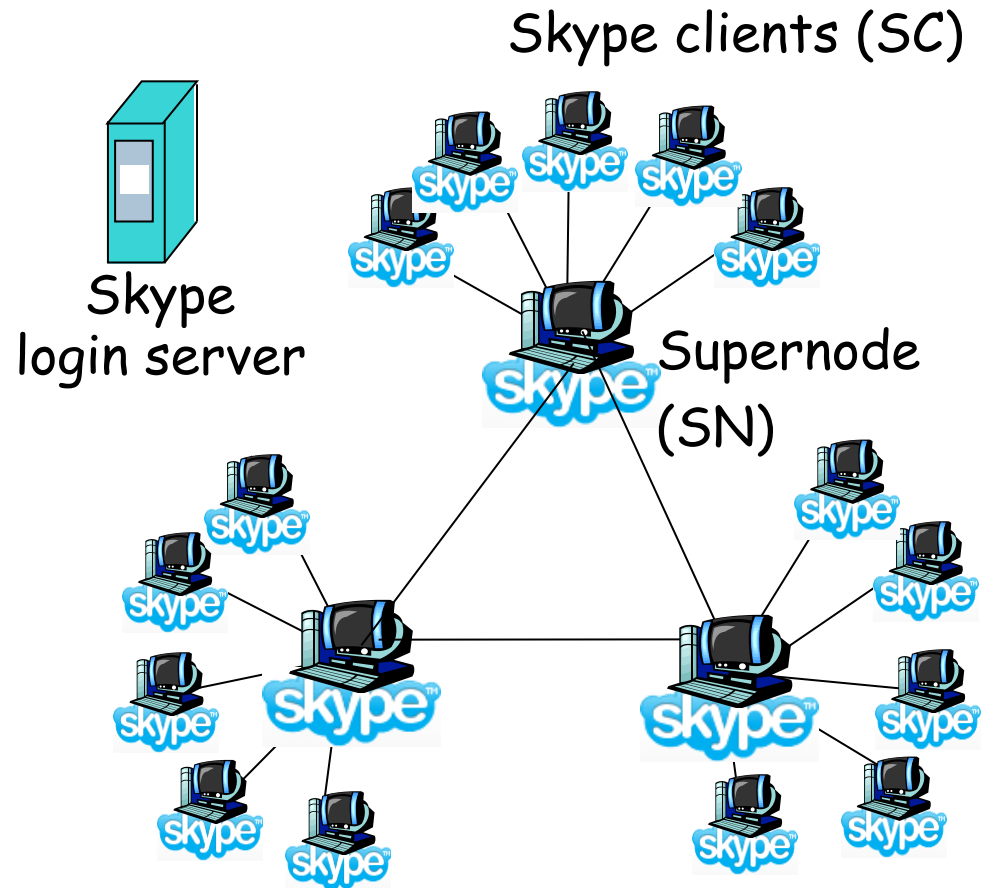


- ❖ To handle peer churn, require each peer to know the IP address of its two successors.
- ❖ Each peer periodically pings its two successors to see if they are still alive.

- ▶ peer 5 abruptly leaves
- ▶ Peer 4 detects; makes 8 its immediate successor; asks 8 who its immediate successor is; makes 8's immediate successor its second successor.
- ▶ What if peer 13 wants to join?

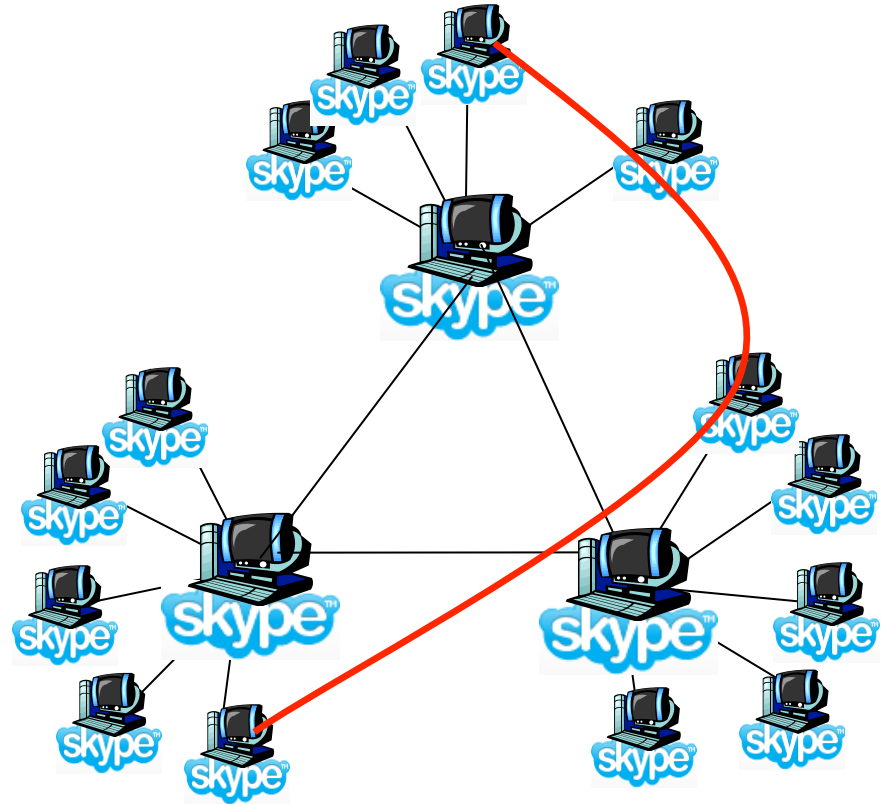
P2P Case study: Skype

- ▶ inherently P2P: pairs of users communicate.
- ▶ proprietary application-layer protocol (inferred via reverse engineering)
- ▶ hierarchical overlay with SNs
- ▶ Index maps usernames to IP addresses; distributed over SNs



Peers as relays

- ▶ problem when both Alice and Bob are behind “NATs”.
 - ▶ NAT prevents an outside peer from initiating a call to insider peer
- ▶ solution:
 - ▶ using Alice’s and Bob’s SNs, *relay* is chosen
 - ▶ each peer initiates session with relay.
 - ▶ peers can now communicate through NATs via relay



Socket programming

Goal: learn how to build client/server application that communicate using sockets

Socket API

- ▶ introduced in BSD4.1 UNIX, 1981
- ▶ explicitly created, used, released by apps
- ▶ client/server paradigm
- ▶ two types of transport service via socket API:
 - ▶ unreliable datagram
 - ▶ reliable, byte stream-oriented

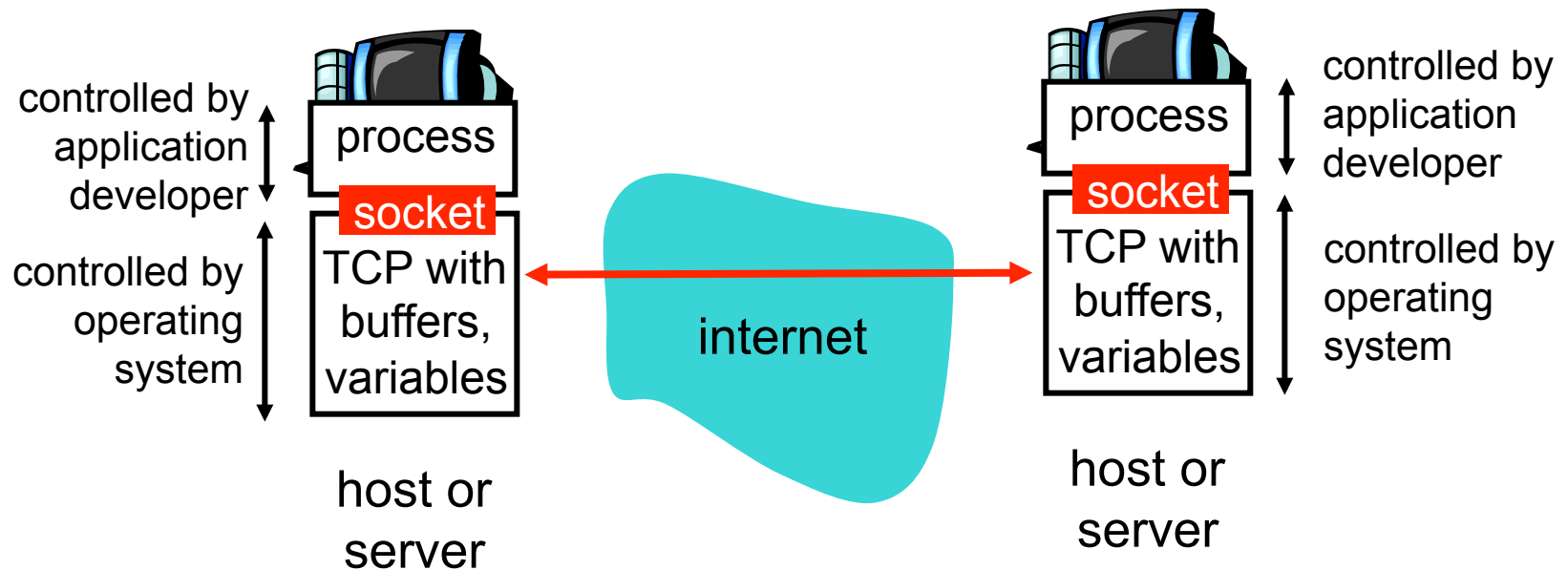
socket

a *host-local*,
application-created,
OS-controlled interface
(a "door") into which
application process can
both send and
receive messages to/from
another application
process

Socket-programming using TCP

Socket: a door between application process and end-end-transport protocol (UCP or TCP)

TCP service: reliable transfer of *bytes* from one process to another



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 client-local 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 client
 - ▶ allows server to talk with multiple clients
 - ▶ source port numbers used to distinguish clients (more in Chap 3)

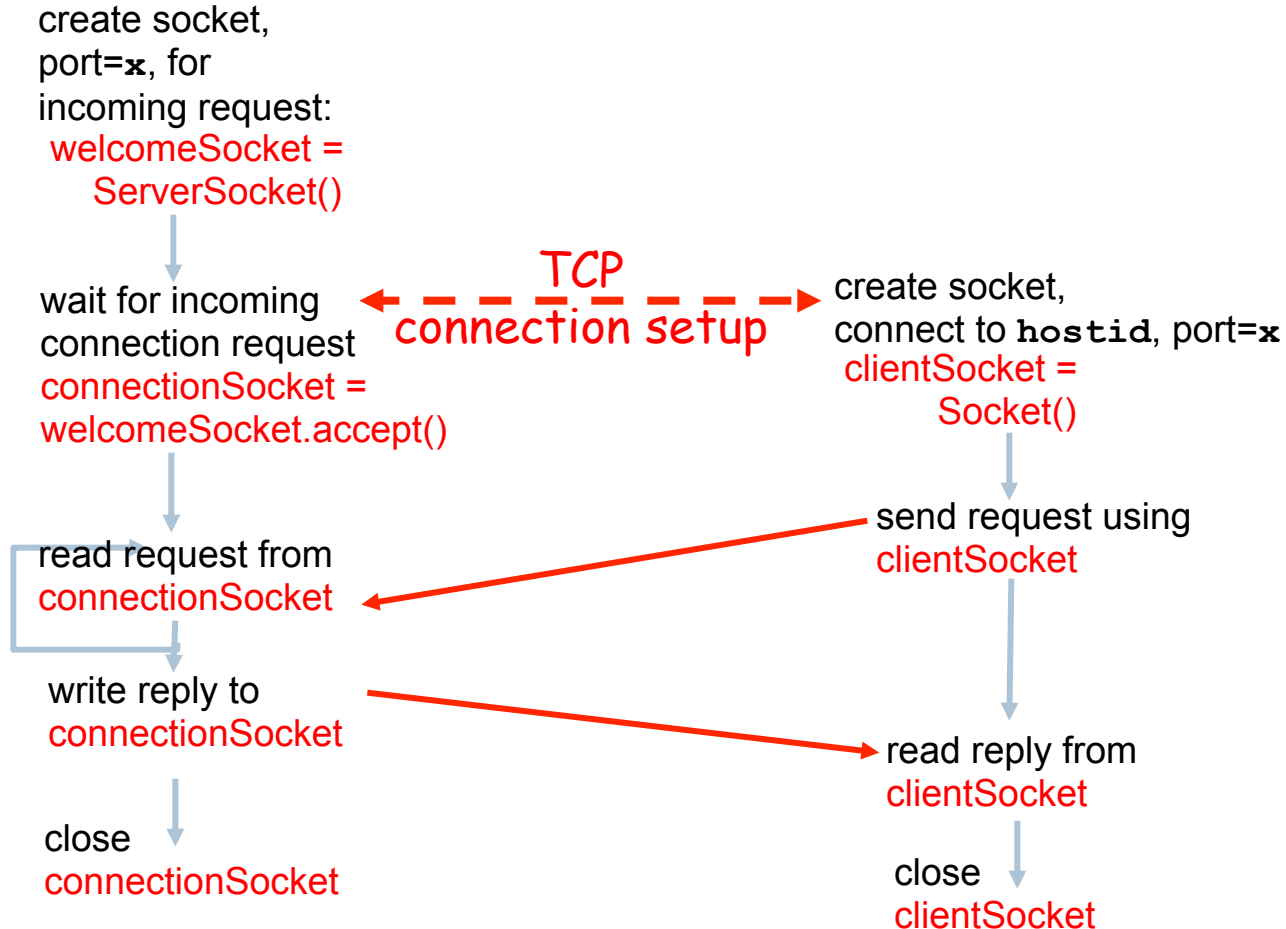
application viewpoint

TCP provides reliable, in-order transfer of bytes ("pipe") between client and server

Client/server socket interaction: TCP

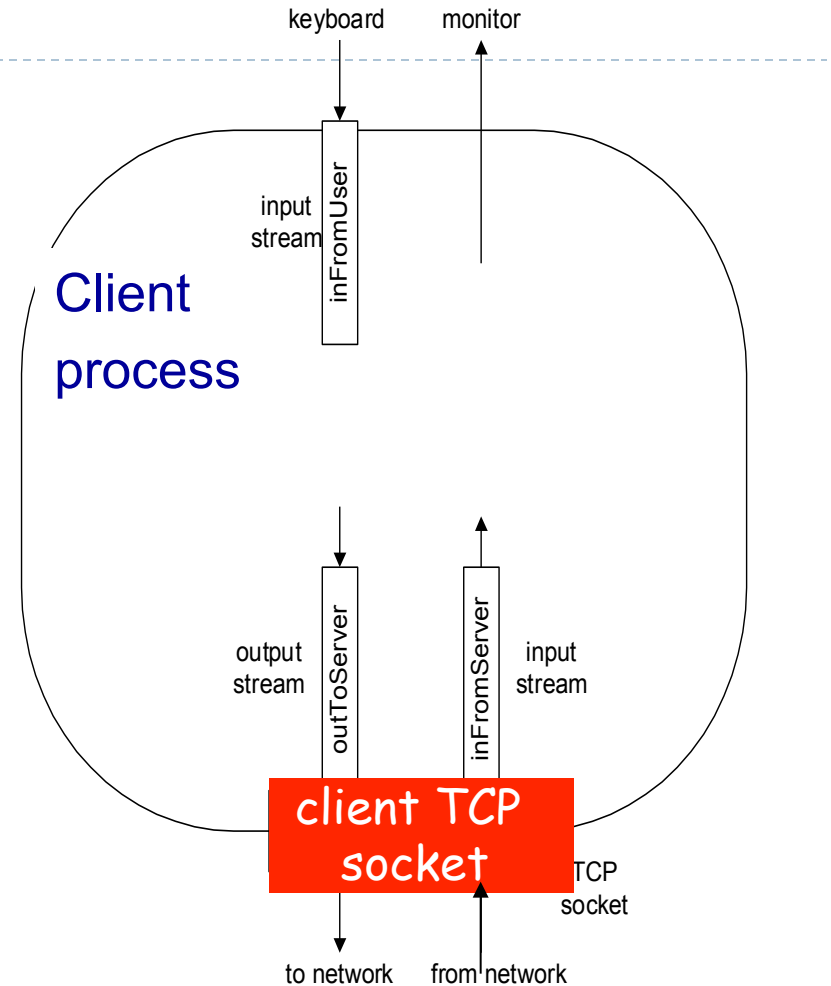
Server (running on `hostid`)

Client



Streams

- ▶ **stream** is a sequence of characters that flow into or out of a process.
- ▶ **input stream** is attached to some input source for the process, e.g., keyboard or socket.
- ▶ **output stream** is attached to an output source, e.g., monitor or socket.



Socket programming with TCP

Example client-server app:

- 1) client reads line from standard input (**inFromUser** stream) , sends to server via socket (**outToServer** stream)
- 2) server reads line from socket
- 3) server converts line to uppercase, sends back to client
- 4) client reads, prints modified line from socket (**inFromServer** stream)



Example: Java client (TCP)

```
import java.io.*;
import java.net.*;
class TCPClient {

    public static void main(String argv[]) throws Exception
    {
        String sentence;
        String modifiedSentence;
```

create
input stream → `BufferedReader inFromUser =
new BufferedReader(new InputStreamReader(System.in));`

create
client socket,
connect to server → `Socket clientSocket = new Socket("hostname", 6789);`

create
output stream
attached to socket → `DataOutputStream outToServer =
new DataOutputStream(clientSocket.getOutputStream());`



Example: Java client (TCP), cont.

create
input stream
attached to socket

```
BufferedReader inFromServer =  
    new BufferedReader(new  
        InputStreamReader(clientSocket.getInputStream()));  
  
sentence = inFromUser.readLine();  
  
send line  
to server
```

```
outToServer.writeBytes(sentence + '\n');
```

read line
from server

```
modifiedSentence = inFromServer.readLine();  
  
System.out.println("FROM SERVER: " + modifiedSentence);  
  
clientSocket.close();  
  
    }  
}
```

Example: Java server (TCP)

```
import java.io.*;  
import java.net.*;
```

```
class TCPServer {
```

```
    public static void main(String argv[]) throws Exception  
    {
```

```
        String clientSentence;  
        String capitalizedSentence;
```

create
welcoming socket
at port 6789

```
        ServerSocket welcomeSocket = new ServerSocket(6789);
```

wait, on welcoming
socket for contact
by client

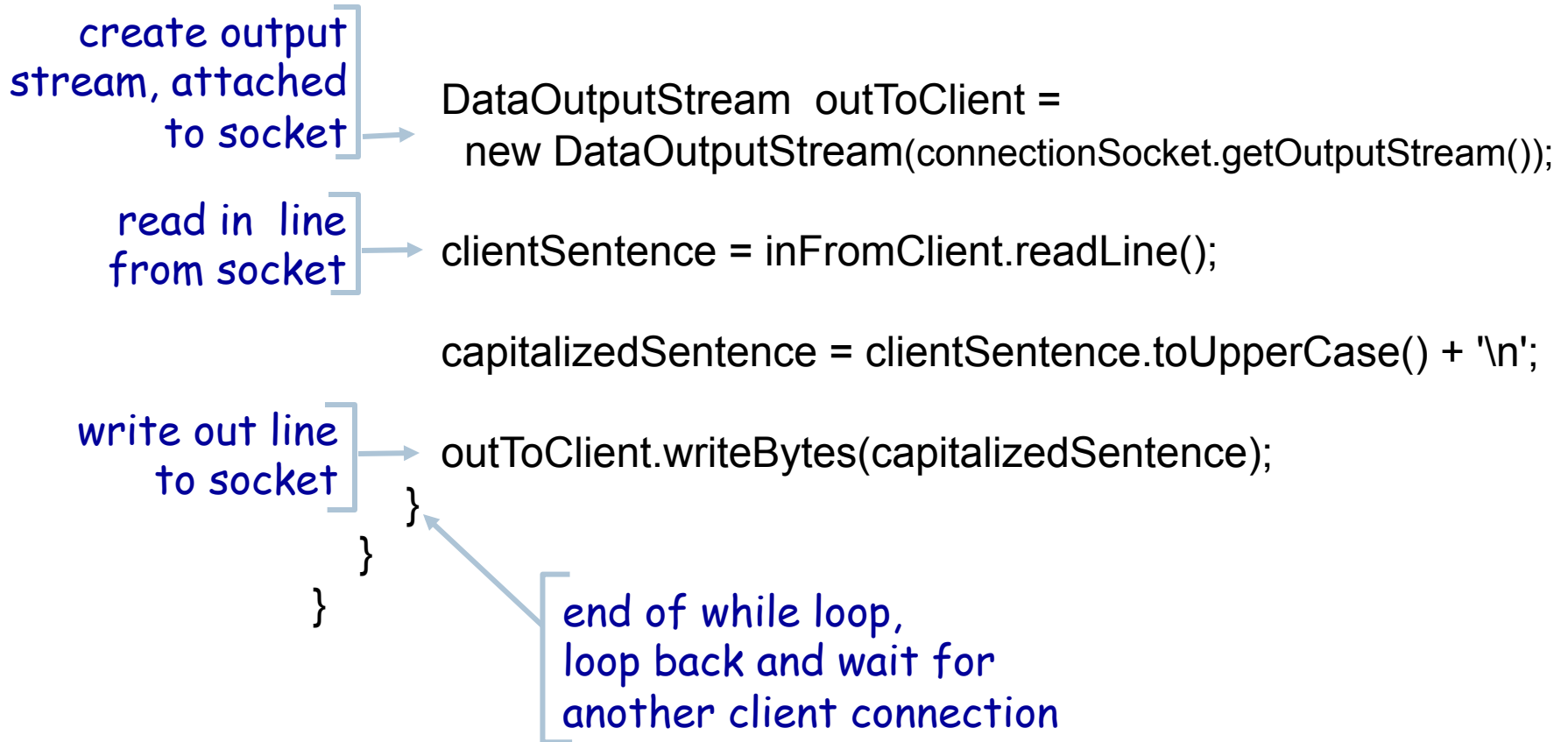
```
        while(true) {
```

```
            Socket connectionSocket = welcomeSocket.accept();
```

create input
stream, attached
to socket

```
            BufferedReader inFromClient =  
                new BufferedReader(new  
                    InputStreamReader(connectionSocket.getInputStream()));
```

Example: Java server (TCP), cont



Socket programming *with UDP*

UDP: no “connection” between client and server

- ▶ no handshaking
- ▶ sender explicitly attaches IP address and port of destination to each packet
- ▶ server must extract IP address, port of sender from received packet

UDP: transmitted data may be received out of order, or lost

application viewpoint:

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

Client/server socket interaction: UDP

Server (running on `hostid`)

Client

create socket,
port= x.

`serverSocket =
DatagramSocket()`

read datagram from
`serverSocket`

write reply to
`serverSocket`
specifying
client address,
port number

create socket,

`clientSocket =
DatagramSocket()`

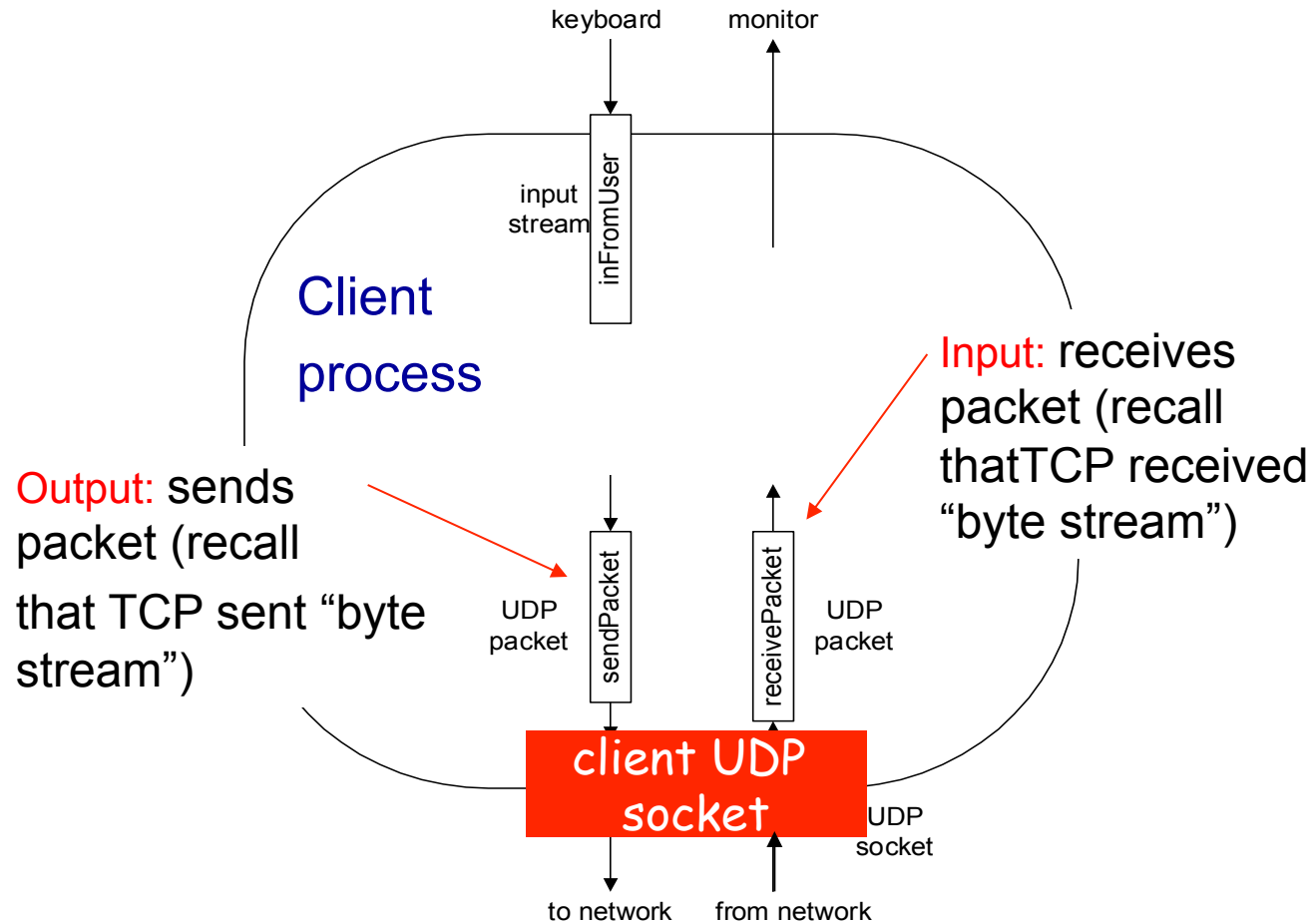
Create datagram with server IP and
port=x; send datagram via
`clientSocket`

read datagram from
`clientSocket`

close
`clientSocket`



Example: Java client (UDP)



Example: Java client (UDP)

```
import java.io.*;
import java.net.*;
```

```
class UDPClient {
    public static void main(String args[]) throws Exception
    {
```

create
input stream

```
        BufferedReader inFromUser =
            new BufferedReader(new InputStreamReader(System.in));
```

create
client socket

```
        DatagramSocket clientSocket = new DatagramSocket();
```

translate
hostname to IP
address using DNS

```
        InetAddress IPAddress = InetAddress.getByName("hostname");
```

```
        byte[] sendData = new byte[1024];
        byte[] receiveData = new byte[1024];
```

```
        String sentence = inFromUser.readLine();
```

```
        sendData = sentence.getBytes();
```

Example: Java client (UDP), cont.

create datagram
with data-to-send,
length, IP addr, port

```
DatagramPacket sendPacket =  
    new DatagramPacket(sendData, sendData.length, IPAddress, 9876);
```

send datagram
to server

```
clientSocket.send(sendPacket);
```

read datagram
from server

```
DatagramPacket receivePacket =  
    new DatagramPacket(receiveData, receiveData.length);
```

```
clientSocket.receive(receivePacket);
```

```
String modifiedSentence =  
    new String(receivePacket.getData());
```

```
System.out.println("FROM SERVER:" + modifiedSentence);  
clientSocket.close();  
}  
}
```

Example: Java server (UDP)

```
import java.io.*;  
import java.net.*;
```

```
class UDPServer {  
    public static void main(String args[]) throws Exception  
    {
```

create
datagram socket
at port 9876

```
        DatagramSocket serverSocket = new DatagramSocket(9876);
```

```
        byte[] receiveData = new byte[1024];  
        byte[] sendData = new byte[1024];
```

```
        while(true)  
        {
```

create space for
received datagram

```
            DatagramPacket receivePacket =  
                new DatagramPacket(receiveData, receiveData.length);
```

receive
datagram

```
            serverSocket.receive(receivePacket);
```

Example: Java server (UDP), cont

```
String sentence = new String(receivePacket.getData());
```

get IP addr
port #, of
sender

```
InetAddress IPAddress = receivePacket.getAddress();
```

```
int port = receivePacket.getPort();
```

```
String capitalizedSentence = sentence.toUpperCase();
```

```
sendData = capitalizedSentence.getBytes();
```

create datagram
to send to client

```
DatagramPacket sendPacket =  
    new DatagramPacket(sendData, sendData.length, IPAddress,  
                        port);
```

write out
datagram
to socket

```
serverSocket.send(sendPacket);
```

```
}
```

```
}
```

```
}
```

end of while loop,
loop back and wait for
another datagram

Useful Debugging Tools

- ▶ telnet
- ▶ nc (netcat)
- ▶ wireshark / tshark
- ▶ tcpdump

Chapter 2: Summary

our study of network apps now complete!

- ▶ application architectures
 - ▶ client-server
 - ▶ P2P
 - ▶ hybrid
- ▶ application service requirements:
 - ▶ reliability, bandwidth, delay
- ▶ Internet transport service model
 - ▶ connection-oriented, reliable: TCP
 - ▶ unreliable, datagrams: UDP
- ❖ specific protocols:
 - HTTP
 - FTP
 - SMTP, POP, IMAP
 - DNS
 - P2P: BitTorrent, Skype
- ❖ socket programming

Chapter 2: Summary

most importantly: learned about protocols

- ▶ typical request/reply message exchange:
 - ▶ client requests info or service
 - ▶ server responds with data, status code
- ▶ message formats:
 - ▶ headers: fields giving info about data
 - ▶ data: info being communicated

Important themes:

- ❖ control vs. data msgs
 - ❖ in-band, out-of-band
- ❖ centralized vs. decentralized
- ❖ stateless vs. stateful
- ❖ reliable vs. unreliable msg transfer
- ❖ "complexity at network edge"