

Chapter 2 Application Layer

Monday, March 1, 2021 9:01 PM

2.1 Principles of Network Applications

- Software run on multiple end systems
- no need to write software that runs on network core devices (routers, switches, work at network layer and below)
- confined application software to end systems

2.1.1 Network Application Architectures

- Client-server architecture or Peer to peer P2P architecture
- **Client server**
 - o always on host - server, services requests from many other hosts - clients
 - send requested object to client host
 - clients do not directly communicate with each other
 - o server has fixed IP address, client can contact server by sending packet to IP address
 - o ex. Web, FTP, Telnet, e-mail
 - o data center has large number of hosts - powerful virtual server
 - o service providers pay recurring interconnection and bandwidth costs
- **P2P**
 - o minimum/no reliance on dedicated servers in data centers
 - o direct comm bw hosts - peers, not owned by service provider
 - o hosts communicate without passing through server
 - o ex. BitTorrent
 - o self-scalability: each peer adds service capacity to the system by distributing files to other peers
 - o cost effective: don't require significant server infrastructure and bandwidth
 - o challenges security, performance, reliability - decentralized

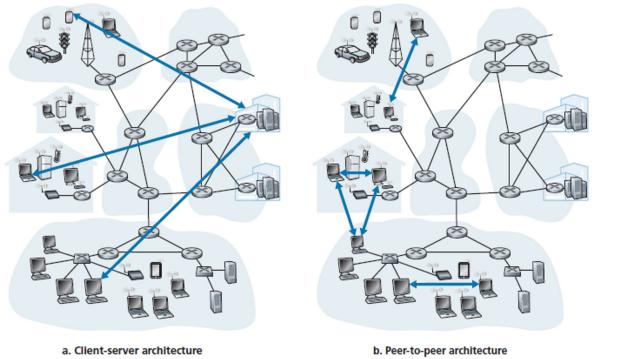


Figure 2.2 • (a) Client-server architecture; (b) P2P architecture

2.1.2 Processes Communicating

- process: program running within end system, can communicate with each other with interprocess comm on same end system
- different end system processes exchange messages across computer network
- **Client Server Processes**
 - o A process can be both client and server - P2P
 - o Web: browser process (client) contacts web server process (server)
 - o P2P: peer A (client) asks peer B(server) to send file
- **Interface Between Process and Computer Network**
 - o process sends messages into and receives messages from software socket
 - o transportation infrastructure on other side of socket to transport message to socket of destination process
 - o message passes through process door

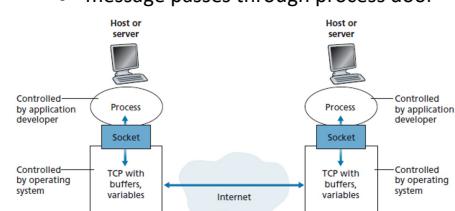


Figure 2.3 • Application processes, sockets, and underlying transport protocol

- Addressing Processes

- o to identify receiving process, need to specify:

- address of host: IP Address
- identifier that specifies receiving process in destination host: receiving socket running in host - destination port number

2.1.3 Transport Services Available to Applications

- application sending pushes messages through socket
- transport layer protocol gets messages to socket of receiving process
- **Services offered by transport layer protocol:**
 - **Reliable Data Transfer**
 - guarantee that data sent is delivered to other end of app
 - sending process pass data into socket, data will arrive without errors
 - if not reliable transfer, some data may never arrive - acceptable for multimedia apps (glitch in audio/video)
 - **Throughput**
 - communication session bw 2 processes along network path, rate at which sending process can deliver bits to receiving process
 - guaranteed available throughput at specified rate
 - app can request guaranteed throughput of r bits/sec
 - if transport protocol cannot provide throughput, app need to encode at lower rate or give up
 - bandwidth sensitive applications ex. mail, file transfer, elastic apps
 - **Timing**
 - timing guarantee
 - every bit sender pumps into socket arrives at receiver socket no more than # msec later
 - ex. multiplayer games, virtual env, conference, tight timing constraints
 - **Security**
 - sending host transport protocol encrypt all data transmitted
 - receiving host, transport layer protocol decrypt data before delivering data to receiving process

2.1.4 Transport Services Provided by Internet

Application	Data Loss	Throughput	Time-Sensitive
File transfer/download	No loss	Elastic	No
E-mail	No loss	Elastic	No
Web documents	No loss	Elastic (few kbps)	No
Internet telephony/ Video conferencing	Loss-tolerant	Audio: few kbps–1 Mbps Video: 10 kbps–5 Mbps	Yes: 100s of msec
Streaming stored audio/video	Loss-tolerant	Same as above	Yes: few seconds
Interactive games	Loss-tolerant	Few kbps–10 kbps	Yes: 100s of msec
Smartphone messaging	No loss	Elastic	Yes and no

- **TCP**
 - reliable data transfer
 - Connection oriented
 - Handshake procedure: client and server exchange transport layer control info before application level messages flow
 - TCP connection now exists between sockets of two processes
 - processes can send messages to eachother at same time
 - connection is torn down when done
 - Reliable data transfer service
 - data is transferred with no missing or duplicate bytes
 - congestion control
 - limit each TCP connection to share of network bandwidth
 - TLS provide security services for encryption
- **UDP**
 - unreliable data transfer service, no guarantee that message reaches receiving process
 - messages may arrive out of order
 - no congestion control mechanism
 - can pump data into the layer at any rate
- **Services Not Provided by Internet Transport Protocols**
 - Internet can provide service to time sensitive applications but no timing and throughput guarantee

Application	Application-Layer Protocol	Underlying Transport Protocol
Electronic mail	SMTP [RFC 5321]	TCP
Remote terminal access	Telnet [RFC 854]	TCP
Web	HTTP 1.1 [RFC 7230]	TCP
File transfer	FTP [RFC 959]	TCP
Streaming multimedia	HTTP (e.g., YouTube), DASH	TCP
Internet telephony	SIP [RFC 3261], RTP [RFC 3550], or proprietary (e.g., Skype)	UDP or TCP

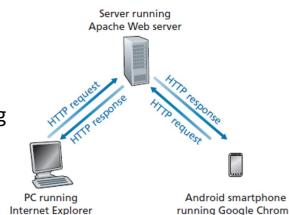
2.1.5 Application Layer Protocols

- Application layer protocol defines how app processes pass messages to each other
 - o type of messages exchanged ex. request messages and response messages
 - o syntax of various message types, fields
 - o semantics of fields, meaning of info
 - o rules for determining when and how process sends/responds msg
- some specified in RFCs public domain ex. HTTP

2.2 Web and HTTP

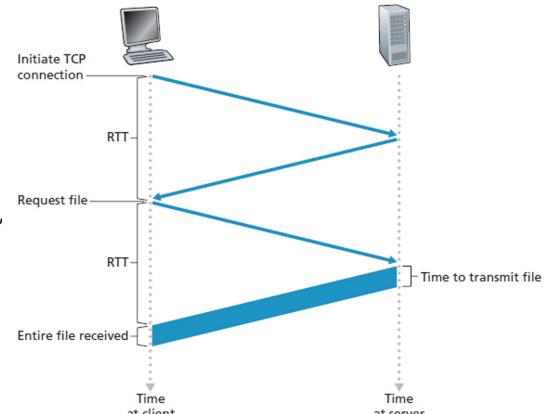
2.1.1 HTTP

- HyperText Transfer Protocol, defined in RFC
- client and server program executing on different end systems, exchange HTTP msg
- Web page consists of objects, ex. images, HTML file
- URL has hostname and houses object and object path name
- Uses TCP as transport protocol
 - o TCP connection with server, socket interface
 - o client sends HTTP request message into socket, receives HTTP response
- Stateless Protocol: does not store state info about client
 - o if client asks for same object twice, server resends object



2.2.2 Non-Persistent and Persistent Connections

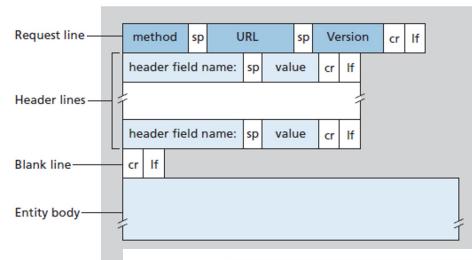
- back to back requests, sent over separate TCP connections or same
- **HTTP and Non-persistent Connections**
 - o ex. <http://www.someSchool.edu/someDepartment/home.index>
 - o HTTP client TCP connection to server with socket
 - o client sends request message to server with path name
/someDepartment/home
.index
 - o server receives request message, retrieves object, encapsulates in HTTP response, sends in socket
 - o HTTP server tells TCP to close
 - o HTTP client receives response, TCP connection closes, client extracts file from response, examines HTML file, finds references to 10 JPEG objects
 - o repeat for each JPEG objects
 - o Total response time: 2 RTTs + transmission time at server of HTML file
 - o Cons:
 - new connection established and maintained for each requested object
 - TCP buffers allocated and TCP vars kept in client and server, burden on Web server
 - each object delivery delay of 2 RTTs - establish TCP and request/receive object
- **HTTP with Persistent Connections**
 - o HTTP/1.1 server leaves TCP connection open after sending response
 - o entire web page sent over single TCP connection
 - o requests back to back without waiting for replies to pending requests (pipelining)
 - o server sends objects back to back



2.2.3 HTTP Message Format

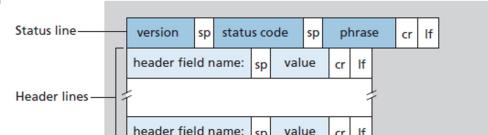
- HTTP Request Message

```
GET /somedir/page.html HTTP/1.1
Host: www.someschool.edu
Connection: close
User-agent: Mozilla/5.0
Accept-language: fr
```



- HTTP Response Message

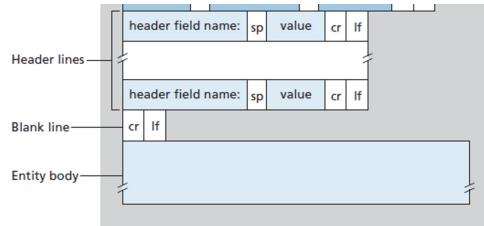
```
HTTP/1.1 200 OK
Connection: close
Date: Tue, 18 Aug 2015 15:44:04 GMT
Server: Apache/2.2.3 (CentOS)
Last-Modified: Tue, 18 Aug 2015 15:11:03 GMT
Content-Length: 6821
```



```

Date: Tue, 18 Aug 2015 15:44:04 GMT
Server: Apache/2.2.3 (CentOS)
Last-Modified: Tue, 18 Aug 2015 15:11:03 GMT
Content-Length: 6821
Content-Type: text/html
(data data data data data ...)

```

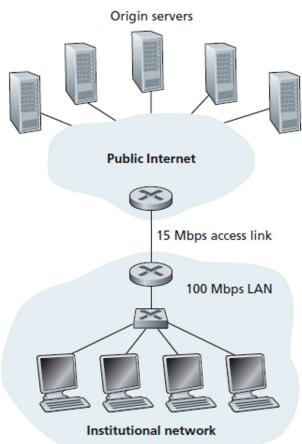
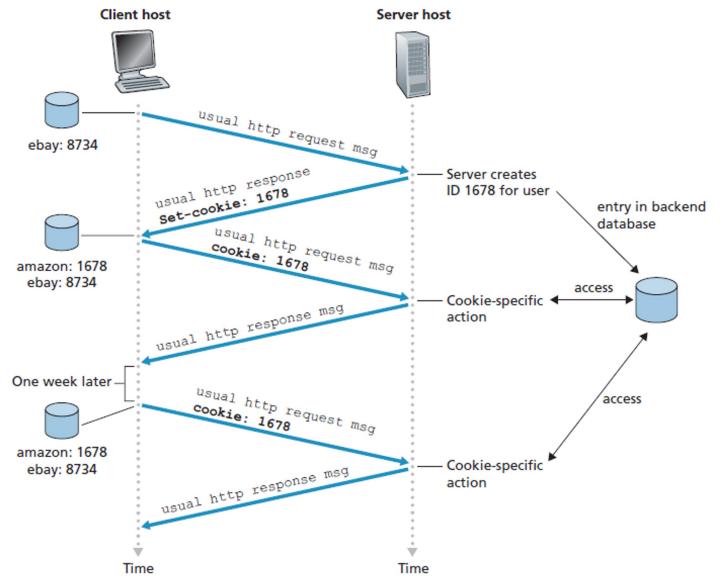


2.2.4 User-Server Interaction: Cookies

- HTTP server stateless
- Website wants to identify users, user cookies, allow sites to keep track of users
 - o cookie header line in HTTP response msg
 - o cookie header line in HTTP request msg
 - o cookie file kept on users end system managed by user browser
 - o back end database at web site
- Set-Cookie: # header contains id number HTTP response and request

2.2.5 Web Caching

- proxy server, satisfies HTTP requests on behalf of origin Web server
- own disk storage, keeps copies of recently requested obj
- When browser request object
 - o browser establish TCP connection, sends HTTP request for obj in web cache
 - o if copy of obj stored, cache returns obj to client browser
 - o else cache opens TCP connection to server, sends HTTP request for obj into cache to server TCP, server sends obj to cache
 - o cache receives obj, stores copy, sends copy to client browser
- cache is server and client at same time
- purchased and installed by ISP
- reduce response time for client request
- reduce traffic on institution access link to internet, no need to upgrade bandwidth as quickly
- reduce web traffic as a whole, improve performance for apps



a) The average time required to send an object over the access link Δ can be determined by:

$$\Delta = \frac{750,000 \text{ bits}}{15,000,000 \text{ bits/sec}} = 0.05 \text{ sec}$$

The traffic intensity on the link is given by $\Delta\beta = 0.05 NK$. Therefore, the average access delay is $\frac{0.05}{1-0.05NK}$. The total average response time is $(\frac{0.05}{1-0.05NK} + 3)$ sec.

b) When a cache is installed, the traffic intensity on the link is given by $\Delta\beta = 0.05(1-p)NK$. Therefore, the average access delay is $\frac{0.05}{1-0.05(1-p)NK}$.

Note that, the probability that a request can be satisfied by the cache is p , which is the hit rate. Hence, the average response time when a cache is installed can be determined by:

$$p \cdot 0 + (1-p) \cdot \left(3 + \frac{0.05}{1-0.05(1-p)NK} \right) = (1-p) \cdot \left(3 + \frac{0.05}{1-0.05(1-p)NK} \right).$$

- Conditional GET

- o copy of obj in cache may be stale, obj in server may have been modified
- o HTTP allows cache to verify that obj are up to date
- o If-Modified-Since: header line

```

GET /fruit/kiwi.gif HTTP/1.1
Host: www.exotiquecuisine.com
If-modified-since: Wed, 9 Sep 2015 09:23:24
  o telling server to send obj only if obj modified since specified date
  o sends response msg without requested obj, Not modified

```

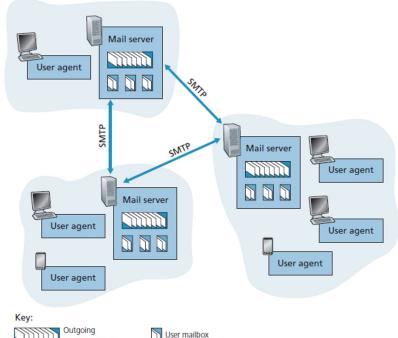
2.2.6 HTTP/2

- goal to reduce latency by enabling request and response multiplexing over single TCP connection
- request prioritization and server push
- efficient compression of HTTP header fields
- reduce number of parallel TCP connections for single Web page, reduce number of sockets
- allow TCP congestion control to operate as intended
- need carefully designed mechanism to avoid HOL blocking
 - o Head of Line HOL

- o ex. large video clip, many small obj below, low-medium speed bottleneck link
 - o using single TCP connection, video clip takes long time to pass through, small obj delayed
 - o HTTP/1.1 work around by opening multiple parallel TCP
- **HTTP/2 Framing**
- o HOL solution: break msg into small frames, req and res on same TCP
 - o ex. 1 large video, 8 small obj
 - 9 requests, video 1000 frames same length, small obj 2 frames
 - frame interleaving, one frame from video, one frame from each small obj
 - o ability to break down HTTP msg into independent frames, interleave them, then reassemble on other end
 - o framing sublayer binary encode frames, more efficient to parse
- **Response Msg Prioritization and Server Pushing**
- o customize rel priority of req to optimize app perf
 - o client prioritize res it is req, assign weight 1-256 to each msg, 256 high
 - o server send first frames for high priority res, dependency stated
 - o send multiple res for single client req, server can push additional obj to client in addition to original res
 - HTML base page indicates obj needed to fully render, can send to client before receiving explicit req
- **HTTP/3**
- o QUIC new transport protocol over UDP
 - o msg multiplexing/interleaving, per stream flow control, low latency connection

2.3 Electronic Mail in the Internet

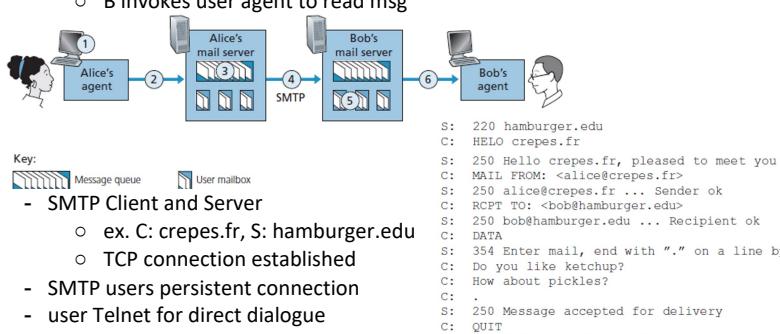
- User agents, mail servers, Simple Mail Transfer Protocol SMTP



- use reliable data transfer service TCP to transfer sender mail to receiver mail server

2.3.1 SMTP

- restricts body of all msg to 7 bit ASCII
- encoded and decoded back to binary
- reliable data transfer TCP
- **Sending Email**
 - o A provide B email address, tells user agent to send msg
 - o A user agent sends msg to mail server, msg queue
 - o client of SMTP on A mail server opens TCP connection to SMTP server running on B mail server
 - o SMTP handshaking, SMTP client sends A msg in TCP connection
 - o B mail server SMTP receives msg, B mail server places msg in B mailbox
 - o B invokes user agent to read msg



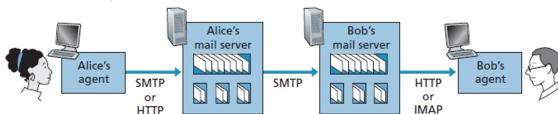
2.3.2 Mail Message Formats

- From, To, Subject header line

From: alice@crepes.fr
To: bob@hamburger.edu
Subject: Searching for the meaning of life.

2.3.3 Mail Access Protocols

- B host would have to always remain on and connected to internet to arrive new mail that could arrive, impractical
- user runs user agent on local host, but accesses mailbox stored on always on shared mail server, shared with other users



- Retrieve msg from mail server
 - o using web based email, using HTTP to retrieve, and SMTP to communicate with A server
 - o using IMAP Internet Mail Access Protocol
 - manage folders maintained in B mail server

2.4 DNS Internet's Directory Service

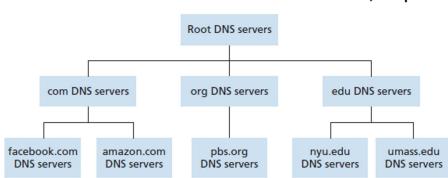
- hosts identified by hostname - mnemonic
- host also identified by IP addresses

2.4.1 Services Provided by DNS

- domain name system DNS
 - o distributed db implemented in hierarchy of DNS servers
 - o application layer protocol allowing hosts to query distributed db
- Obtain IP address of hostname
 - o user machine runs client side of DNS app
 - o browser extracts hostname from URL and pass to client side of DNS
 - o DNS client sends query containing hostname to DNS server
 - o DNS client receives reply, includes IP address for hostname
 - o browser initiates TCP connection to HTTP server process at port 80 of IP
- IP address often cached in DNS server, reduce DNS traffic and delay
- **Host aliasing**
 - o host with complex hostname can have 1+ alias names - canonical hostname
- **Mail server aliasing**
 - o email address simplified even if yahoo mail server has a canonical hostname of relay1.west-coast.yahoo.com
- **Load distribution**
 - o DNS perform load distribution among replicated servers/webservers
 - o busy sites are replicated over multiple servers with each running on diff end system and IP address
 - o set of IP associated with one alias hostname
 - o DNS db contains set of IP addresses, server responds with entire set of IP to query for a name
 - o DNS rotation distributes traffic among replicated servers
 - o used for email so that many mail servers can have same alias name

2.4.2 DNS Overview

- hostname to IP address translation
- Centralized DNS server design problems
 - o single point of failure
 - o traffic volume
 - o distant centralized database
 - not close to all querying clients
 - o Maintenance
 - o doesn't scale
- **Distributed Hierarchical Database**
 - o DNS has many servers around the world
 - o mapping for all hosts in internet, distributed across DNS servers
 - o 3 classes of DNS servers: root DNS, Top level domain TLD, authoritative DNS



- o ex. finding IP for www.amazon.com

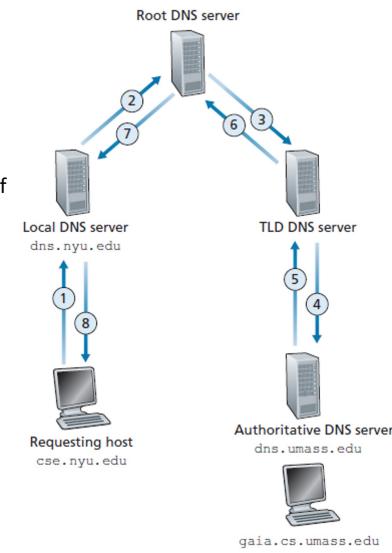
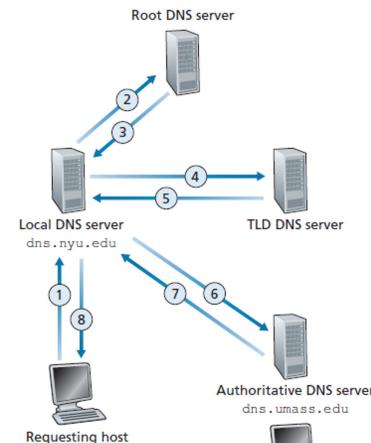
- client contacts root servers, return IP address for TLD server for .com domain
- client contacts TLD server, returns IP of authoritative server for amazon.com
- client contacts authoritative server for amazon.com, returns IP for hostname www.amazon.com
- **Root DNS servers**
 - more than 1000 instances over world
 - copies of 13 diff root servers managed by 12 organizations
 - provide IP addresses of TLD servers
- **Top-level domain TLD servers**
 - com, org, net, edu, gov, country top lvl domains uk, fr, ca, jp
 - provide IP addresses for authoritative DNS servers
- **Authoritative DNS servers**
 - DNS records that map hostnames to IP addresses
 - organization's authoritative DNS server houses authoritative DNS records
 - can also pay to store records in another DNS server of service provider
- **Local DNS Server**
 - not part of hierarchy
 - each ISP has local DNS server
 - when host connects to ISP, provides host with IP address of 1+ local DNS servers
 - close to host, may be on same LAN
 - forwards host DNS queries into DNS server hierarchy
- TLD server does not always know authoritative DNS server for hostname, may only know of an intermediate DNS server leading to desired mapping
- **recursive query:** asks to obtain mapping on behalf of other server
- **iterative query:** all of replies are directly returned to client server
- **DNS Caching**
 - reduce delays and number of DNS messages
 - when DNS server receives DNS reply, can cache mapping in local memory
 - if other query asks for same hostname, DNS server can provide IP address even if not authoritative for hostname
 - can also cache IP of TLD servers, bypass root DNS servers on query chain

2.4.3 DNS Records and Messages

- The DNS servers store resource records RRs that provide hostname to IP mapping

(Name, Value, Type, TTL)

- **TTL**
 - time to live, when resource should be removed from cache
- **Type**
 - **A**
 - name = hostname, value = IP address for hostname
`(relay1.bar.foo.com, 145.37.93.126, A)`
 - **NS**
 - name = domain, value = hostname of authoritative DNS
`(foo.com, dns.foo.com, NS)`
 - **CNAME**
 - value = canonical hostname for alias hostname Name
`(foo.com, relay1.bar.foo.com, CNAME)`
 - **MX**
 - Value = canonical name of mail server with alias hostname Name
`(foo.com, mail.bar.foo.com, MX)`
- **DNS Messages**
 - 12 bytes header section
 - 16 bit number identifying query, copied into reply msg to query
 - 1 bit flag indicating query/flag
 - 1 bit authoritative flag in reply msg when DNS server is authoritative
 - 1 bit recursion-desired flag set when DNS server perform recursion if no record
 - 1 bit recursion-available set in reply if DNS supports recursion
 - 4 number field in header for num of occurrences in 4 types of data sections under header
 - question section: info about query, includes name field, type field
 - answer section: reply from DNS server return RRs
 - authority section contains records of other authoritative servers
 - additional section contains other helpful records
 - use nslookup program to send DNS query from host to DNS
- **Inserting Records into DNS Database**
 - register domain name at registrar
 - provide names and IP addresses of primary and secondary authoritative DNS servers
 - Type A resource record for web server



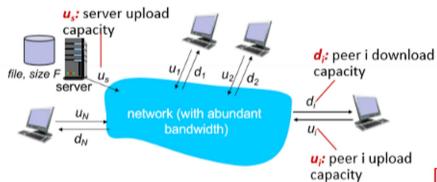
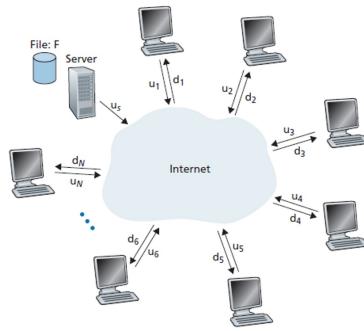
Identification	Flags	12 bytes
Number of questions	Number of answer RRs	
Number of authority RRs	Number of additional RRs	Name, type fields for a query
Questions (variable number of questions)		
Answers (variable number of resource records)		RRs in response to query
Authority (variable number of resource records)		Records for authoritative servers
Additional information (variable number of resource records)		Additional "helpful" info that may be used



- register domain name at registrar
- provide names and IP addresses of primary and secondary authoritative DNS servers
- Type A resource record for web server

2.5 Peer-to-Peer File Distribution

- little to no reliance on always on infrastructure servers
- pairs of connected hosts, peers communicate directly
- peers exchange IP addresses, request service from others
- peers may come and go
- ex. BitTorrent
- **Scalability of P2P Architecture**
 - self scalability: new peers bring new service capacity and service demands
- **File Distribution: Client-Server vs P2P**
 - Distributing file size F from one server to N peers

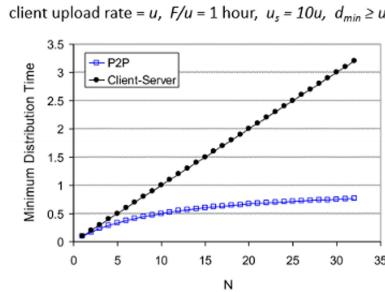


time to distribute F to N clients using client-server approach

$$D_{c-s} \geq \max\{NF/u_s, F/d_{min}\}$$

increases linearly in N

- **Client-Server:**
 - must send N file copies $N*F/u_s$
 - min client download time F/d_{min}
 - d_{min} client download rate
- **P2P:**
 - time to send one copy F/u_s
 - client must download file copy F/d_{min}
 - clients aggregate must download $N*F$ bits



tracker: tracks peers participating in torrent

torrent: group of peers exchanging chunks of a file

... but so does this, as each peer brings service capacity

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 ...



- Requesting, Sending File Chunks

- peers have different subsets of chunks
- host asks each peer for list of chunks they have and requests rarest first
- **Rarest first**
 - send and request chunks
 - determine among the missing chunks, request from neighbours the rarest chunks
 - redistributed quicker to equalize number of copies
- host gives priority to peers sending her chunks at highest rate
- other peers are choked, do not receive chunks from her
- randomly select another peer, unchoke and send chunks

2.6 Video Streaming and Content Distribution Networks

- Distributed application level infrastructure

2.6.1 Internet Video

- sequence of images 24-30 images/second, uncompressed digitally encoded image consists of array of pixels, each encoded into number of bits to represent luminance and color
- important performance measure for streaming is avg end-to-end throughput
- network must provide an avg throughput to streaming application that is as large as bit rate of compressed video

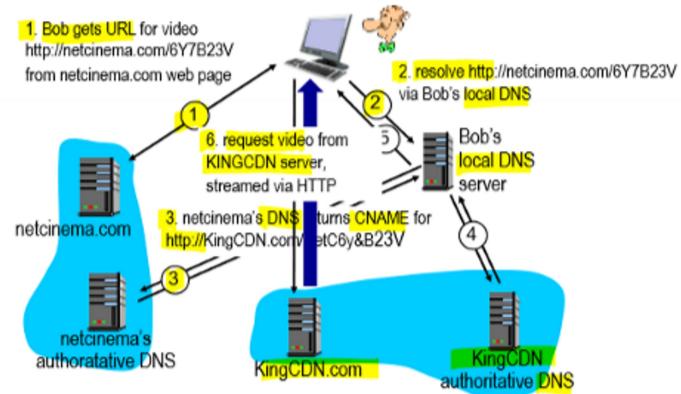
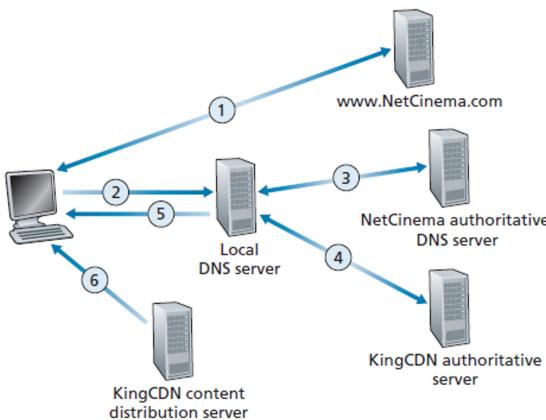
2.6.2 HTTP Streaming and DASH

- video stored at HTTP server as file with URL
- server sends video file within HTTP response message as quickly as underlying network protocols and traffic conditions allow
- bytes collected in client app buffer, once number of bytes in buffer exceeds threshold, client app begins playback, grabbing video frames from buffer

- Con: all clients receive same encoding of video despite variations in bandwidth
- **Dynamic Adaptive Streaming over HTTP (DASH)**
 - o video encoded into several different versions w diff bit rate, diff quality level
 - o client requests chunks of video of a few sec in length, when bandwidth is high, client selects chunks from high rate version
 - o allows clients with different internet access rates to stream at different encoding rates
 - o lots of buffered video, receive high bitrate chunks

2.6.3 Content Distribution Networks

- How to stream content to simultaneous users
- store/serve multiple copies of videos at multiple geographically distributed sites CDN
- **enter deep:** push CDN servers deep into many access networks close to users
- **bring home:** bring home smaller number of larger clusters
- stores copies of content at CDN nodes
- subscriber requests content from CDN, retrieves content or choose different copy if congested
- DNS redirects user request to CDN server



2.7 Socket Programming: Creating Network Applications

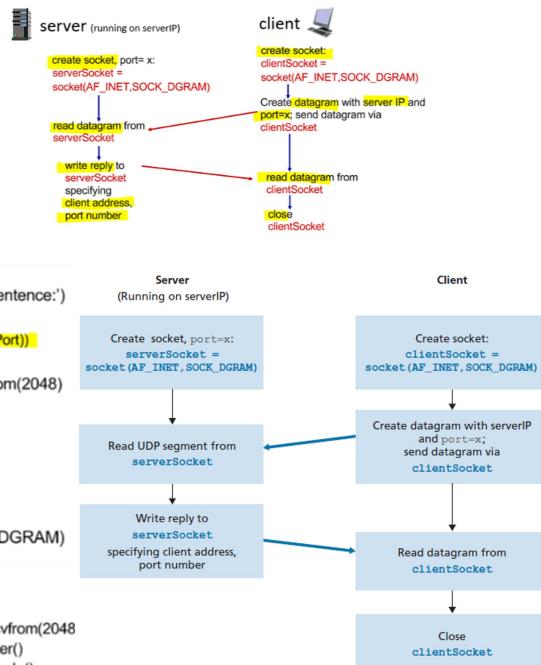
- socket: door bw application process and end-to-end transport protocol
- UDP: unreliable datagram
- TCP: reliable, byte stream oriented

2.7.1 Socket Programming UDP

- no handshaking or connection bw client and server
- sender attaches IP address and port # to each packet
- receiver extracts sender IP address and port # of packet
- transmitted data may be lost or out of order

```
Python UDPCClient
include Python's socket library → from socket import *
serverName = 'hostname'
serverPort = 12000
create UDP socket for server → clientSocket = socket(AF_INET,
                                                       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()

Python UDPServer
from socket import *
serverPort = 12000
create UDP socket → serverSocket = socket(AF_INET, SOCK_DGRAM)
bind socket to local port number 12000 → serverSocket.bind(('', serverPort))
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)
```



2.7.2 Socket Programming TCP

- client must contact server
- server process is running and created socket
- client creates TCP socket, specifying IP and port number of server
- client creates TCP connection



- Client must contact server
- server process is running and created socket
- client creates TCP socket, specifying IP and port number of server
- client creates TCP connection
- server TCP creates new socket to communicate with client
- can have multiple sockets and clients

Python TCPClient

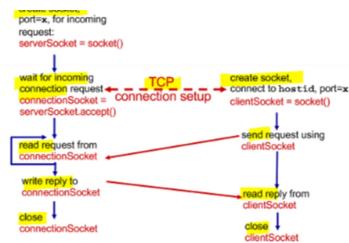
```

from socket import *
serverName = 'servername'
serverPort = 12000
create TCP socket for server,
remote port 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()
No need to attach server name, port
  
```

Python TCPServer

```

create TCP welcoming socket
from socket import *
serverPort = 12000
serverSocket = socket(AF_INET,SOCK_STREAM)
serverSocket.bind(('',serverPort))
serverSocket.listen(1)
print 'The server is ready to receive'
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)
connectionSocket, addr = serverSocket.accept()
sentence = connectionSocket.recv(1024).decode()
capitalizedSentence = sentence.upper()
connectionSocket.send(capitalizedSentence.
encode())
  
```



Server (Running on serverIP)

