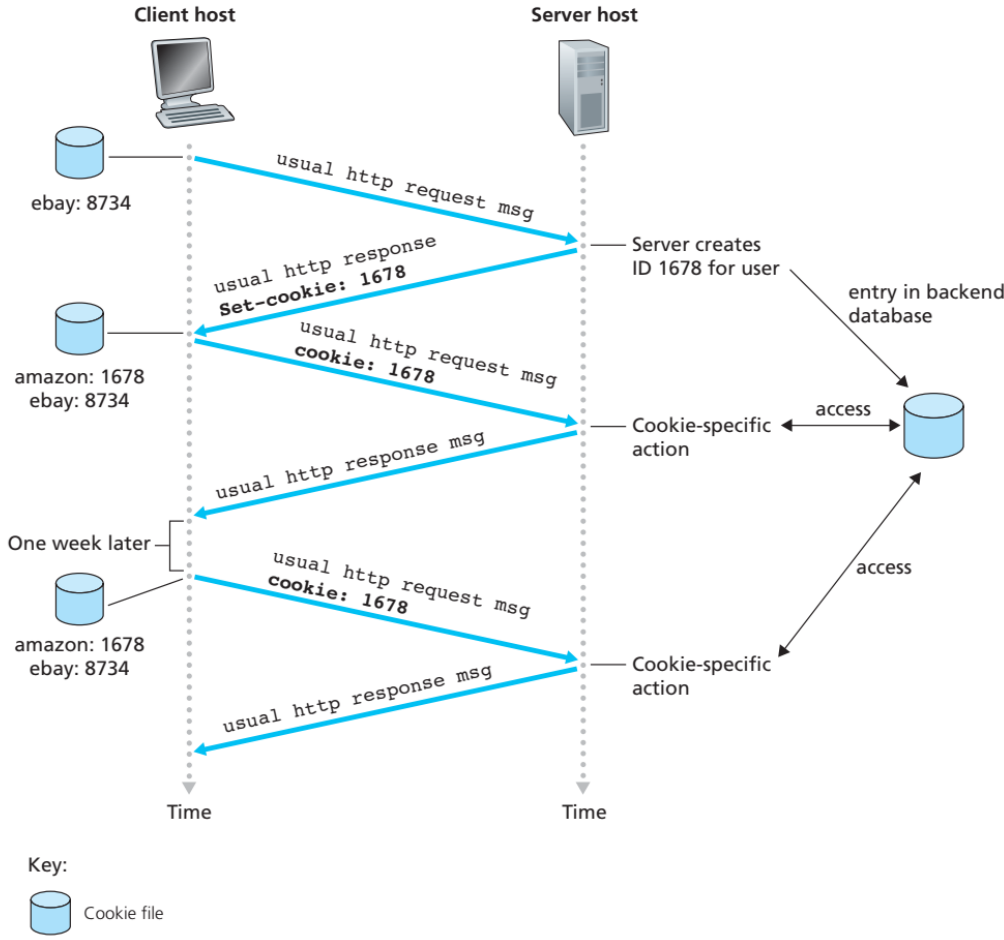
Module 2: Application Layer (Lecture – 2)

Dr. Nirnay Ghosh
Assistant Professor
Department of Computer Science & Technology
IIEST, Shibpur

User-Server Interaction: Cookies

- For future accesses to the Web site, the user's browser consults the cookie file
- Extracts the identification number for this site
- Puts a cookie header line that includes the identification number in the HTTP request
- If the user registers himself/herself with the Web site, then his/her personal information gets attached to the unique identifier
- Cookies: creates a user session layer on top of stateless HTTP

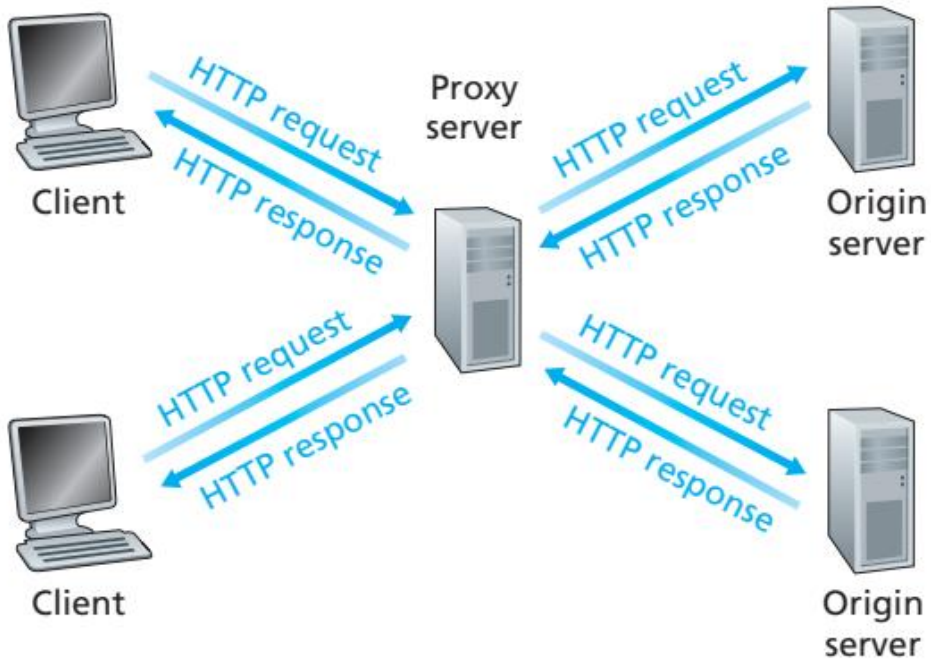


Keeping User State with Cookies

- Drawback:
  - Invasion of privacy
  - Combination of cookies and user supplied information can reveal sensitive personal information and can be potentially sold to a third party

Web Caching

- Web cache/Proxy server: network entity that satisfies HTTP requests on the behalf of an origin Web server
  - Has its own disk storage and keeps copies of recently requested objects in this storage
- A user's browser can be configured to direct all HTTP requests to the Web cache
- The different steps involved in Web cache functioning are as follows:
  - The browser establishes a TCP connection to the Web cache
  - It sends an HTTP request for an object to the Web cache
  - The Web cache checks if it has a local copy of the object
  - If so, it returns the object within an HTTP response message to the client browser
  - If not, it opens a TCP connection to the origin server and sends an HTTP request for the object into the cache-to-server TCP connection

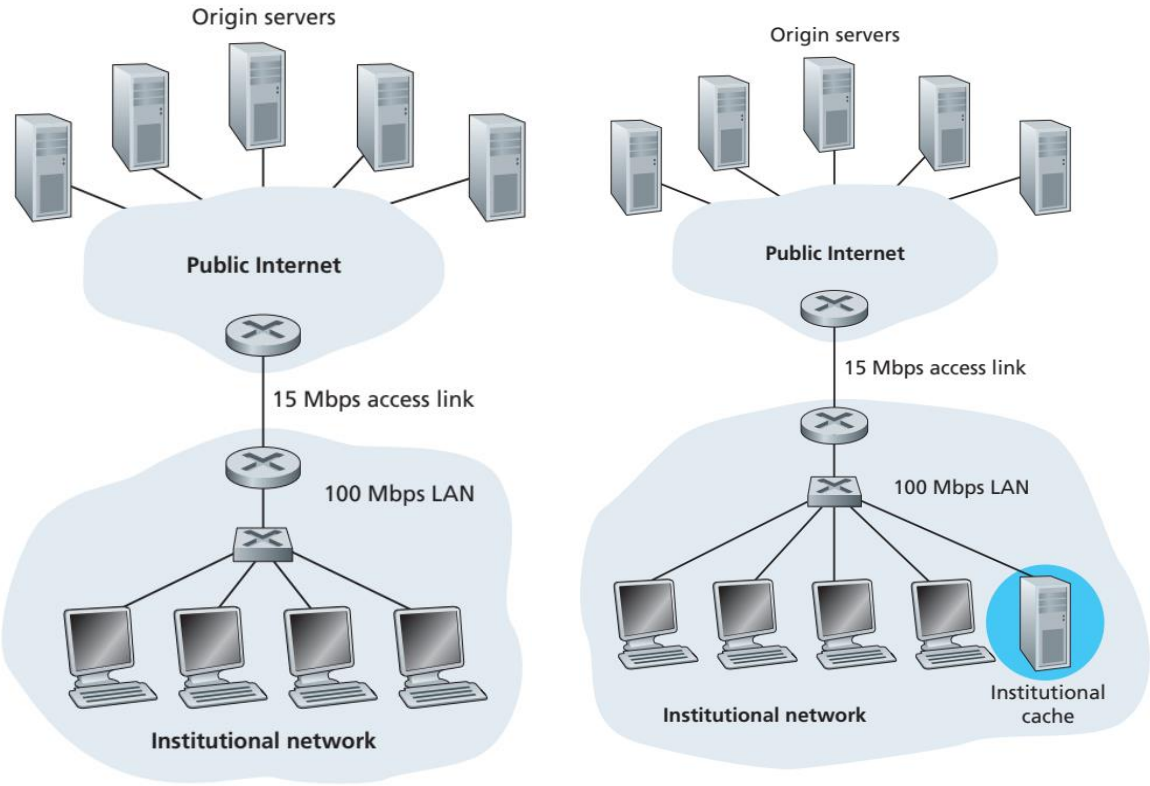


Clients Requesting Objects through a Web Cache

- The origin server sends the object within an HTTP response to the Web cache
- Web cache stores a copy in its local storage and sends the object within an HTTP response message to the client browser

Web Caching

- Web cache: functions both as a server (to the client browser) and as a client (to the origin server)
- Typically a Web cache is purchased and installed by an ISP
  - Cost is low
  - Many caches use public-domain software that runs on inexpensive PCs
  - Example: a university may install a cache on its campus network and configure all campus browsers to point to the cache
- Deployment of Web cache renders the following benefits:
  - Substantially reduces the response time for a client request
    - Generally, the bottleneck bandwidth between client and origin server is much less than that between client and Web cache
  - Significantly reduces the traffic on an institution's access link to the Internet
    - Institution does not have to upgrade bandwidth (an expensive solution than deploying Web cache)
  - Reduces the Web traffic in the Internet as a whole, thereby improving the performance for all applications



Web Cache Reduces the Bottleneck between an Institutional Network and the Internet

- Content Distribution Networks (CDN)
  - Created by installing many geographically distributed caches throughout the Internet
  - Localizes much of the Internet traffic

The Conditional GET

- Drawback of caching:
  - Copy of an object residing in the cache may be stale
  - Object housed in the Web server may have been modified since the copy was cached
- Conditional GET: mechanism that allows a cache to verify that its objects are up-to-date
- Includes an **If-Modified-Since: header** line to perform an up-to-date check
- Suppose the Web cache sends a request message to the Web server:

```
GET /fruit/kiwi.gif HTTP/1.1
Host: www.exotiquecuisine.com
```

- Web server sends a response message with the requested object to the cache:

```
HTTP/1.1 200 OK
Date: Sat, 8 Oct 2011 15:39:29
Server: Apache/1.3.0 (Unix)
Last-Modified: Wed, 7 Sep 2011 09:23:24
Content-Type: image/gif

(data data data data data ...)
```

- Cache forwards the object to the requesting browser and caches it locally
  - Stores the last-modified date along with the object
- The object may or may not be modified in the Web server
- Later if another browser requests for the same object, cache issues a conditional GET to the Web server - performs the up-to-date check

```
GET /fruit/kiwi.gif HTTP/1.1
Host: www.exotiquecuisine.com
If-modified-since: Wed, 7 Sep 2011 09:23:24
```

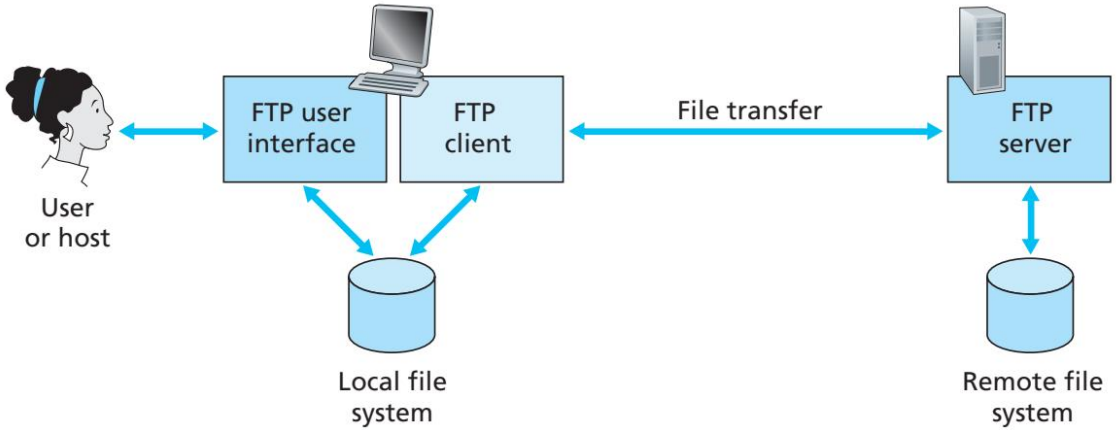
- The value of the **If-modified-since: header** line is exactly equal to the value of the **Last-Modified: header** line that was sent by the server
- Conditional GET: directs the Web server to send the object only if it has been modified since the specified date
- If the object is not modified the server will send the following message

```
HTTP/1.1 304 Not Modified
Date: Sat, 15 Oct 2011 15:39:29
Server: Apache/1.3.0 (Unix)

(empty entity body)
```

File Transfer Protocol (FTP)

- FTP: used to transfer files to or from a remote host
- User interacts with FTP through an FTP user interface
- FTP client process (in the localhost): establishes TCP connection with the FTP server process (in the remote host) on port number 21
- User provides his/her identification and password which are sent over the TCP connection as a part of the FTP command
- Once the server authorizes the user, file transfer to and from it can take place
- FTP uses two parallel TCP connections for file transfer between client and server:
  - Control connection: used for sending control information – user identification, password, commands to change remote directory, and commands to "put" and "get" files
  - Data connection: used to actually send a file



FTP moves files between Local and Remote File Systems

Control and Data Connections

File Transfer Protocol (FTP)

- On receiving the command for file transfer over the control connection, the server side initiates a TCP connection to the client side
- FTP sends exactly one file over the data connection and then closes it
- If the user wants to transfer another file at the same time, a new FTP data connection is opened
- Control connection remains open throughout the user session
  - FTP separates control and data connections
  - The control information it sends is out-of-band (in case of HTTP, it is in-band)
- For every user session, FTP server maintains state information about the user's: control connection; current directory
- Tracking state information of each ongoing user session constraints the total number of simultaneous sessions

- Common FTP commands issued by client:

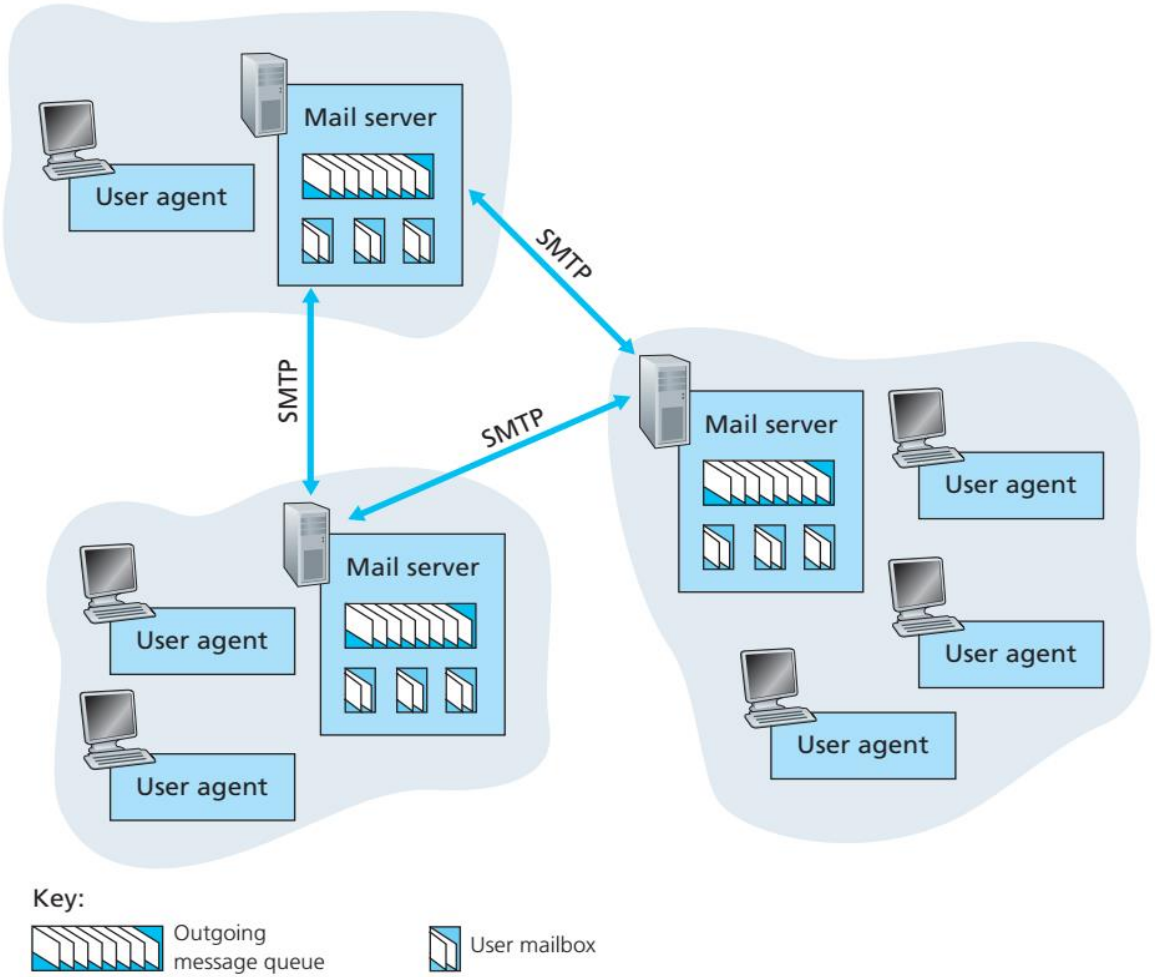
- `USER username:` Used to send the user identification to the server.
- `PASS password:` Used to send the user password to the server.
- `LIST:` Used to ask the server to send back a list of all the files in the current remote directory. The list of files is sent over a (new and non-persistent) data connection rather than the control TCP connection.
- `RETR filename:` Used to retrieve (that is, get) a file from the current directory of the remote host. This command causes the remote host to initiate a data connection and to send the requested file over the data connection.
- `STOR filename:` Used to store (that is, put) a file into the current directory of the remote host.

- Some typical replies from FTP server:

- `331 Username OK, password required`
- `125 Data connection already open; transfer starting`
- `425 Can't open data connection`
- `452 Error writing file`

Electronic Mail in the Internet

- Email: an asynchronous communication medium
- Modern e-mail has many powerful features: messages with attachments, hyperlinks, HTML-formatted text, embedded photos
- Three major components of Internet e-mail system are
  - User agent: allows users to read, reply to, forward, save, and compose messages (e.g., Microsoft Outlook, Apple Mail, etc.)
  - Mail server: core of the e-mail infrastructure – houses the mailbox of each recipient – manages and maintains messages that have been sent to the recipient
  - Simple Mail Transfer Protocol (SMTP): principle application-layer protocol for Internet electronic mail
- If the recipient's mail server is down
  - The message is hold in the message queue hosted by the sender's mail server
  - Transfer attempts are made after every 30 minutes (tentative)
  - Server removes the message from the queue (notifying the sender) if transfer attempts fail for several days



Internet E-mail System

- Typical path followed by a message:
  - *Sender's user agent → Sender's mail server → Recipient's mail server → Recipient's mailbox → Recipient's user agent*