CIS 457 - Data Communications

Nathan Bowman

Images taken from Kurose and Ross book

HTTP

HyperText Transfer Protocol (HTTP) is the protocol for the web

Web clients and servers exchange messages using this protocol

How is a web page represented in the computer?

Web page -- collection of objects

Object -- file with a URL. Could be HTML file, image, video, etc.

Web page generally consists of base HTML file specifying layout and giving URLs of referenced objects

Page containing HTML text and five images, for example, consists of six objects

URL consists of

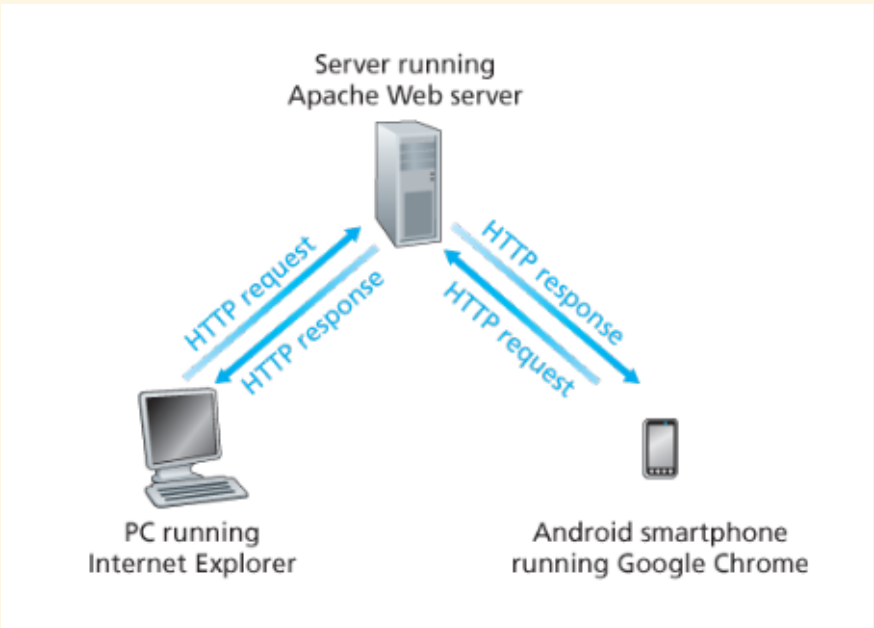
- hostname of server housing object
- objects pathname on that server

Example URL:

```
http://www.someSchool.edu/someDepartment/picture.gif
```

hostname: `www.someSchool.edu`

path: `/someDepartment/picture.gif`


Server running
Apache Web server
HTTP request
HTTP response
HTTP response
HTTP request
PC running
Internet Explorer
Android smartphone
running Google Chrome

One thing to note about the HTTP protocol is that it is *stateless*

For a given HTTP request, an HTTP server can provide the requested object without knowing what other requests this particular client has made in the past

Non-Persistent and Persistent Connections

HTTP runs on top of TCP, which establishes a connection before transferring application-layer data

If HTTP begins new TCP connection every time request is made, connections are **non-persistent**

On the other hand, it is possible to make more than one HTTP request during the same TCP connection. In that case, connections are **persistent**

HTTP allows for either persistent or non-persistent connections

Non-Persistent Connections

Imagine a client requests a web page containing an HTML file and 10 images, all stored on the same web server

URL is `http://www.someSchool.edu/someDepartment/`

We will study messages sent assuming connection is non-persistent

1. HTTP client initiates TCP conection to server (`www.someSchool.edu`) on port 80
2. HTTP client sends HTTP request for path `/someDepartment/home.index`
3. HTTP server receives request, retrieves requested object, and sends HTTP response message containing object to client
4. HTTP server process tells TCP to terminate connection

...continued...

5. HTTP client receives response and TCP connection terminates
6. Client sees references to 10 images in HTML file, repeats steps 1-4 for each image

Note that HTTP does not determine how document will be displayed in browser

HTTP simply specifies how to transfer information

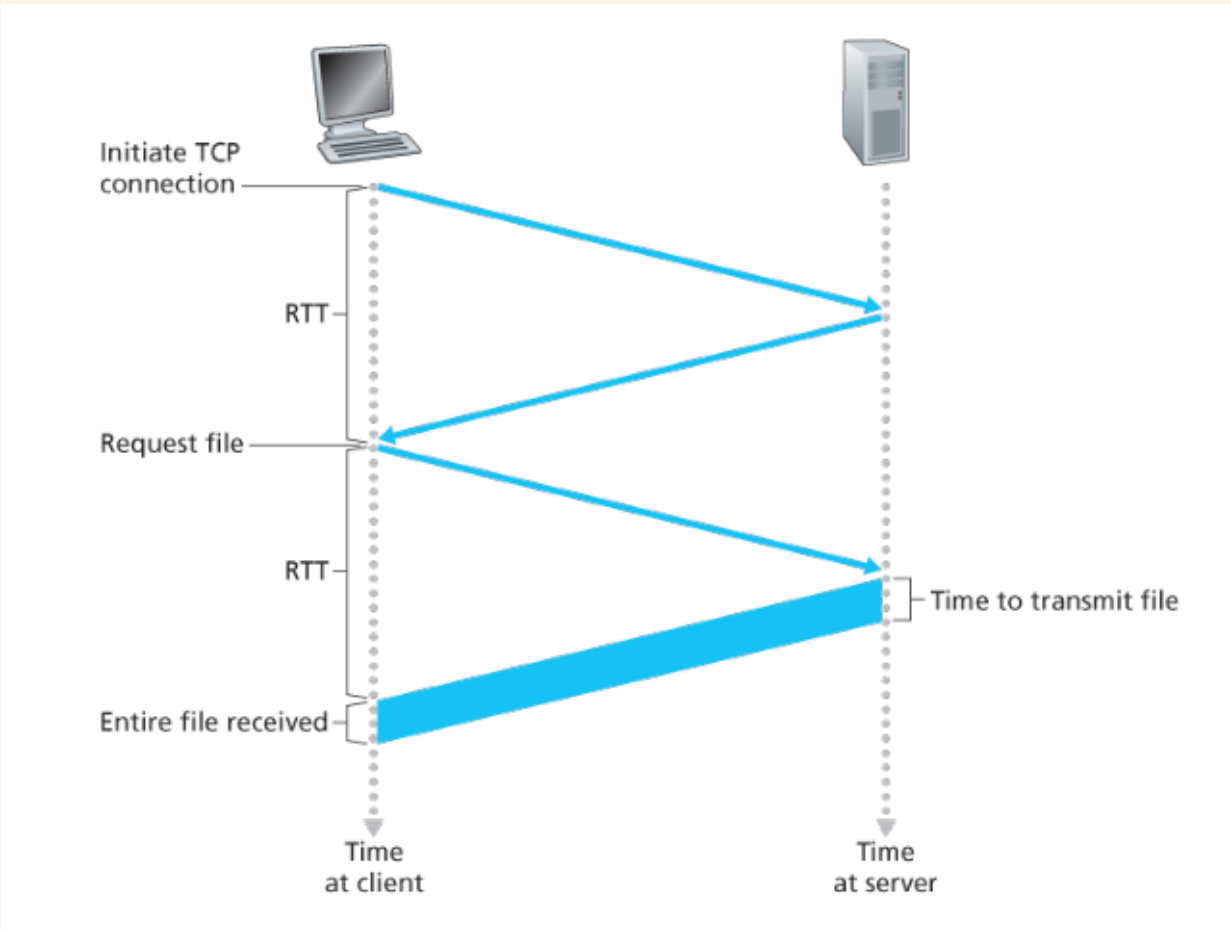
Due to non-persistent connections, new TCP connection must be established for all 11 requests

Before examining persistent connections, good to know more about timing of messages

Goal: determine how much time elapses from client requesting HTML file to client receiving HTML file

Define **round-trip time** (RTT) as time it takes for small packet to travel client -> server -> client

Note that TCP connections begin with so-called "three-way handshake" -- connection not established until three messages have been sent


Initiate TCP
connection
RTT
Request file
RTT
Time to transmit file
Entire file received
Time
at client
Time
at server

Total response time roughly:

2*RTT + time to transmit file

With non-persistent connections, incur cost of (2*RTT + transmit time) or *each* object in web page

With persistent connections (available in HTTP 1.1), TCP connection stays open for more than one HTTP request so additional TCP handshakes not necessary

All objects making up one web page, and even more than one web page stored on same server, can be transferred through single TCP connection