Introduction to Computer Networks
Homework 2: Solution

Part I

2.1 (20%)

(a) A user requests a Web page that consists of some text and three images. For this page, the client will send one request message and receive four response messages. False. Since there is one request per each object, the client will send 4 requests and receive 4 responses.

(b) Two distinct Web pages (for example, `www.mit.edu/research.html` and `www.mit.edu/students.html`) can be sent over the same persistent connection. True. In a persistent HTTP connection, files requested by a client on the same server can be sent over a single TCP connection.

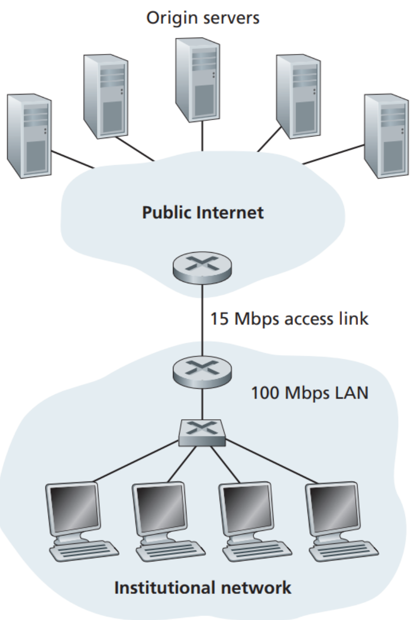
(c) With nonpersistent connections between browser and origin server, it is possible for a single TCP segment to carry two distinct HTTP request messages. False. Every non-persistent connection between the client and the server will create a new TCP segment for every new request.

(d) The `Date:` header in the HTTP response message indicates when the object in the response was last modified. False. The `Date` header indicates the date and time that the message was generated. On the other hand, the `Last-Modified` header is the one that indicates the date and time when the resource was last modified.

(e) HTTP response messages never have an empty message body. False. Certain HTTP status codes, such as 204 (No Content) or 304 (Not Modified), may result in an empty response body.

2.9 (20%)

Consider Figure 2.12, for which there is an institutional network connected to the Internet. Suppose that the average object size is 850,000 bits and that the average request rate from the institution's browsers to the origin servers is 16 requests per second. Also suppose that the amount of time it takes from when the router on the Internet side of the access link forwards an HTTP request until it receives the response is three seconds on average (see Section 2.2.5). Model the total average response time as the sum of the average access delay (that is, the delay from Internet router to institution router) and the average Internet delay. For the average access delay, use $\Delta/(1-\Delta\beta)$, where Δ is the average time required to send an object over the access link and β is the arrival rate of objects to the access link.



(a) Find the total average response time.

$L = 850000\ bits, R = 15\ Mbps$

$$\Delta = \frac{L}{R} = \frac{8.5 \times 10^5}{1.5 \times 10^7} = 0.0567\ (sec)$$

$$\Delta\beta = 0.0567\ (sec/req) \times 16\ (req/sec) = 0.9072$$

$$\text{Access delay} = \frac{\Delta}{(1-\Delta\beta)} = \frac{0.0567}{1-0.9072} \approx 0.6\ (sec)$$

Total average response time = access delay + internet delay = $0.6 + 3 = 3.6\ (sec)$

(b) Now suppose a cache is installed in the institutional LAN. Suppose the miss rate is 0.4. Find the total response time.

$\Delta = 0.0567\ (sec)$

$$\text{Average response time for cache miss} = \frac{0.0567}{1 - 0.4 \times 0.9072} + 3 \approx 0.089 + 3 = 3.089\ (sec)$$

Since the response time is nearly 0 when the request is fulfilled by the cache, with a probability of 0.6, the total response time $= 0.6 \times 0 + 0.4 \times 3.089 \approx 1.24\ (sec)$

Part II. Additional problems.

II.1 (20%)

Consider an HTTP client that wants to retrieve a web document at a given URL (for example, `http://www.cs.nthu.edu.tw`). The IP address of the HTTP server is initially unknown and therefore the HTTP client needs to get the corresponding IP address by using DNS lookup.

(a) What transport-layer protocol does DNS lookup need (in this scenario)?

UDP. DNS uses the User Datagram Protocol (UDP) on port 53 to serve DNS queries. UDP is preferred because it is fast and has lower overhead. It doesn't use a time-consuming three-way hand-shake procedure to start the data transfer like TCP does.

(b) What transport-layer protocol does HTTP service need?

TCP. The HTTP service relies on the Transmission Control Protocol (TCP) as its essential transport-layer protocol. TCP is chosen for its reliability and connection-oriented nature, which is crucial for the consistent and ordered exchange of data in the context of web communication.

II.2 (40%)

Suppose within your web browser you click on a link to obtain a webpage. The IP address for the associated URL is not cached in your local host, so a DNS lookup is necessary to obtain the IP address.

- Suppose that n DNS servers are visited before your host receives the IP address from DNS. The successive visits incur round-trip times of RTT_1, RTT_2, ..., RTT_n, respectively.
- Suppose that the round-trip time between the local host and the HTTP server (associated with the URL) is RTT_0.
- Suppose that the base HTML file of the webpage references 8 objects on the same server.

Neglecting the transmission times of the base HTML file and the objects, how much time elapses with

(a) Non-persistent HTTP with no parallel TCP connections?

1. The total amount of time to get the IP address:
   $RTT_1 + RTT_2 + \cdots + RTT_n$
2. Set up the TCP connection: RTT_0
3. Base html file: RTT_0
4. Object: RTT_0

Therefore, the total response time:

$$\begin{aligned}&RTT_1 + RTT_2 + \cdots + RTT_n + (RTT_0 + RTT_0) + (RTT_0 + RTT_0) \times 8\\ &= RTT_1 + RTT_2 + \cdots + RTT_n + 18 \cdot RTT_0\end{aligned}$$

(b) Non-persistent HTTP with the browser configured for 5 parallel TCP connections?

$8/5 = 1.6 \Rightarrow 2$

$$\begin{aligned}&RTT_1 + RTT_2 + \cdots + RTT_n + (RTT_0 + RTT_0) + (RTT_0 + RTT_0) \times 2\\ &= RTT_1 + RTT_2 + \cdots + RTT_n + 6 \cdot RTT_0\end{aligned}$$

(c) Persistent HTTP with pipelining?

$$\begin{aligned}&RTT_1 + RTT_2 + \cdots + RTT_n + (RTT_0 + RTT_0) + RTT_0 \times 1 \text{ (pipelining)}\\ &= RTT_1 + RTT_2 + \cdots + RTT_n + 3 \cdot RTT_0\end{aligned}$$

(d) Persistent HTTP without pipelining (and without parallel connections)?

$$\begin{aligned}&RTT_1 + RTT_2 + \cdots + RTT_n + (RTT_0 + RTT_0) + RTT_0 \times 8\\ &= RTT_1 + RTT_2 + \cdots + RTT_n + 10 \cdot RTT_0\end{aligned}$$