Diane Rai
HW 2
Introduction to Network and their Application

Question 1
Create your own website using the University of Iowa's web servers. Please mention your website address. A simple HTML file uploaded to your homepage is acceptable.

http://homepage.cs.uiowa.edu/~dichharai/hp-dichha/index.html

Question 2
With respect to HTTP:
(I) Explain how Web caching can reduce the delay in receiving a requested object.
A web cache -- also called proxy server -- is a network entity that satisfies HTTP requests on the the behalf of an origin Web server. The Web cache has its own disk storage and keeps copies of recently requested objects in this storage. A user's browser can be configured so that all of the user's HTTP requests are first directed to the Web cache. As an example, suppose a browser is requesting the object *http://www.someschool.edu/campus.gif* , here is a step of what happens:
i) The browser established a TCP connection to the Web cache and sends an HTTP request for the object to the Web cache.

ii) The Web cache checks to see if it has a copy of the object stored locally. If it does, the Web cache returns the object within an HTTP response message to the client browser.

iii) If the Web cache does not have the object, the Web cache opens a TCP connection to the origin server, that is, to *www.someschool.edu*. The Web cache then sends an HTTP request for the object into the cache-to-server TCP connection. After receiving this request, the origin server sends the object within an HTTP response to the Web cache.

iv) When the Web cache receives the object, it stores a copy in its local storage and sends a copy, within an HTTP response message, to the client browser (ever the existing TCP connection between the client browser and the Web cache).

This way Web caching reduces the delay in receiving a requested object by having to get desired content from closer server to a user's host -- eg. LAN.

Numerical example:
Assumptions:
- Average object size = 1Mbits
- Average request rate from browsers to origin servers: 15 req/sec
- Average data rate to browsers: 15Mbits/s
- RTT from institutional router to any origin server: 2s
- Access link (Web cache to origin server link) rate 15Mbits/s
- Web Cache(LAN) link rate: 100Mbits/s

Diane Rai
HW 2
Introduction to Network and their Application

Consequences:
- LAN utilization: 15 req/sec * 1Mbits =15Mbits/s = 15/100 = 0.15 = 15%
- Access link utilization = 15 req/sec * 1Mbits = 15Mbits/s = 15/15 = 1 = 100%
- **Total delay** = Internet delay + access delay + LAN delay
  = 2 sec + minutes + usecs

Increase in access link speed is not cheap if we were to increase its bandwidth.

By web caching we can increase cache hit rate which decreases delay and access link utilization.

Suppose:
- 40% requests satisfied at cache,
- 60% requests satisfied at origin

Access link utilization:
- 60% of requests use access link

Data rate to browsers over access link
= 0.6 * 15Mbits/s = 9Mbits/s
- Utilization = 9/15 = 60%

Total delay
= 0.6 * (delay from origin servers) + 0.4(delay when satisfied at cache
= 0.6 * (2.01) + 0.4(~msec)
= ~1.2 sec

This show Web caching has smaller delay and is cheaper to implement as well.

(II) Consider an HTTP client that wants to retrieve a Web document at a given URL. The IP address of the HTTP server is initially unknown. What transport and application layer protocols besides HTTP are needed in this scenario?

- Before HTTP GET request to be sent for the web document, the HTTP client needs to obtain IP address of the HTTP server hosting the web document.
- For this purpose, a DNS request is sent out to obtain the hostname to IP address mapping. (Resource Record => type = NS, type = A)
- DNS runs over UDP
- Once a mapping is obtained, the HTTP client first establishes a TCP connection with the server. HTTP runs over TCP.
- Following the TCP connection establishment a GET request for the web document is sent over this connection.
- The following are the protocols that are used:

Diane Rai
HW 2
Introduction to Network and their Application

a) Application layer protocols: DNS and HTTP
b) Transport layer protocols: UDP for DNS; TCP for HTTP

(III) Assume a Web browser wants to access a Web page, which in turn includes 8 small objects. Also assume that a TCP connection requires a request and a response (one RTT). With respect to Round Trip Time (RTT), how much it takes to fully load the page with:
(a) Non-persistent HTTP with no parallel TCP connection?

RTT definition: time for a small packet to travel from client to server and back
RTT = prop delays + queuing delays + processing delays
HTTP response time:
- 1 RTT to initiate TCP connection
- 1 RTT for HTTP request and and first few bytes of HTTP response to return
- File transmission time

Therefore, for non-persistent HTTP response time is
= (2RTT + file transmission) * 8

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

In non-persistent HTTP, browsers often open parallel TCP connections to fetch referenced objects. This is an optimization to reduce delay.
Since there are 8 small objects to access, and 5 parallel connections, it can be done in 2 times.
Therefore, non-persistent HTTP response time with 5 parallel connection is = (2RTT + file transmission) * 2

(c) Persistent HTTP?
- Server leaves connection open after sending response
- Subsequent HTTP messages between same client/server are sent over open connection
- Client sends requests as soon as it encounters a referenced object
- In fact, the client does not need to wait for a reply before making another request(pipelining)
- As little as 1RTT is needed for all the referenced objects.

Therefore, persistent HTTP response time is 1RTT.

Diane Rai
HW 2
Introduction to Network and their Application

(IV) Assume that there are u users connected to the University of Iowa's wireless network to connect to the Internet. u = 1000 during the day, and r = 600 during the night. These users want to browse some Web sites, and the average number of HTTP requests generated by all the users is 0.5 × u requests per second. Assume that the size of HTTP request messages is negligible, but each HTTP response message returns an object of size 500 Kbits. If the university is connected to the Internet through a 700 Mbps access link,

(a) What is the percentage of access link's capacity consumed by the HTTP response messages during the day and night?
During day u = 1000
Average number of HTTP requests generated by all the users = 0.5*u requests/s = 500 requests/s
Size of HTTP response message is 500 Kbits = 0.5Mbits

HTTP Response size = 500 requests/s * 0.5Mbits = 250Mbits/s
Utilization of access link = (250/700)*100 = 35.7%

During night u = 600
HTTP Response size = 0.5*600 requests/s * 0.5Mbits = 150Mbits/s
Utilization of access link = (150/700)*100 = 21.4%

(b) What is the percentage of access link's capacity consumed by the HTTP response messages during the day if we use a Web cache which can directly respond to HTTP requests with probability h?
cache -hit = h
Requests through origin-server = 1-h
HTTP Response size = (1-h)*250Mbits/s
% of access link's capacity consumed = (1-h)*(250/700)*100 = (1-h)*(35.7)%

Question 3
With respect to Domain Name System (DNS):
(I) Why DNS is required? (Describe with respect to the Internet protocol stack and Web servers)

- DNS (Domain Name System) is required to translate hostnames to IP addresses. To identify a host, routers prefer fixed-length hierarchically structured IP addresses while people prefer mnemonic hostnames.
- It is an application-layer protocol that allows hosts to query the distributed database
- The DNS protocols run over UDP and uses port 53

Diane Rai
HW 2
Introduction to Network and their Application

- DNS is commonly used by other application-layer -- including HTTP, SMTP and FTP - to translate user-supplied hostnames to IP addresses
- DNS helps in host aliasing and Mail server aliasing. A host with a complicated hostname can have 1 or more alias names for eg. a hostname relay1.west-coast.enterprise.com could have, say, two aliases such as enterprice.com and www.enterprise.com. In this case relay1.west-coast.enterprise.com is said to be a canonical hostname.
- DNS is also used to perform load distribution among replicated servers, such as replicated Web server. Busy sites, such as cnn.com, are replicated over multiple servers, with each server running on a different end system and each having a different IP address. For replicated Web servers, a set of IP addresses is thus associated with one canonical hostname. The DNS database contains this set of IP addresses.

(II) Why DNS is operated in a distributed and hierarchical manner?
DNS is operated in a distributed and hierarchical manner for scaling and security purposes. Without distributed DNS, a single point of failure, a DNS server crash means entire Internet is down. Distributed DNS servers would distribute traffic volume generated by all HTTP requests and e-mail messages. It will decentralize database so that there is no significant delays. It is also very useful in their maintenance.

(III) Run the following command in linux terminal: host -v uiowa.edu. Include the result and discuss about that.

```
======================================================================
Trying "uiowa.edu"
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 49031
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;uiowa.edu.                     IN      A

;; ANSWER SECTION:
uiowa.edu.              70437   IN      A       54.163.225.50

Received 43 bytes from 127.0.1.1#53 in 35 ms
Trying "uiowa.edu"
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 57050
;; flags: qr rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 0
```

Diane Rai
HW 2
Introduction to Network and their Application

;; QUESTION SECTION:
;uiowa.edu. IN AAAA

;; AUTHORITY SECTION:
uiowa.edu. 4268 IN SOA dns0.uiowa.edu. hostmaster.uiowa.edu. 201610295
7200 3600 3600000 21600

Received 79 bytes from 127.0.1.1#53 in 46 ms
Trying "uiowa.edu"
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 43797
;; flags: qr rd ra; QUERY: 1, ANSWER: 4, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;uiowa.edu. IN MX

;; ANSWER SECTION:
uiowa.edu. 32191 IN MX 10 smtp1.its.uiowa.edu.
uiowa.edu. 32191 IN MX 10 smtp4.its.uiowa.edu.
uiowa.edu. 32191 IN MX 10 smtp2.its.uiowa.edu.
uiowa.edu. 32191 IN MX 10 smtp3.its.uiowa.edu.

Received 119 bytes from 127.0.1.1#53 in 26 ms

==

The DNS servers that together implement the DNS distributed store resource records (RRs) including RRs that provide hostname-to IP address mappings. The command host -v uiowa.edu gives information carrying 3 RRs.
The first RR provides hostname-to-IP addressing mappings. The TYPE=A provides the standard hostname-to-IP address mapping.

In the second RR, the the section of TYPE=AAAA, which indicates IPv6 is empty which means that uiowa.edu has still not transferred to IPv6 protocol. In the section of TYPE=SOA which indicates Start of authority record specifies authoritative information about uiowa.edu domain name. It includes primary name server(uiowa.edu), email of the domain administrator (hostmaster.uiowa.edu), domain serial number and several timers relating to refreshing the domain.

The third RR provides Mail Exchange records. It is indicated with TYPE=MX. It maps a domain name to a list of message transfer agents for uiowa.edu domain.

Diane Rai
HW 2
Introduction to Network and their Application

(IV) How can you modify your operating system configurations so that when a DNS query is made, that is sent to a specific DNS server (e.g., Google's)?

In Ubuntu operating system(os),
Following steps are used
i) cd /etc/network
ii) sudo gedit interfaces
iii) in the interfaces file, write dns-nameservers 8.8.4.4
iv) save the file
iv) restart the network using the following commands:

> sudo ifconfig enp0s3 down # for shutting down network, in place where enp0s3 is used put your os's network interface which may be like eth0
> sudo ifconfig enp0s3 up # for restarting network

In windows os, write the following terminal commands
i) Netsh // utility to change system networking settings
ii) interface ip set dns "Local Area Connection" static 192.168.1.1 // specify what network properties to change in this case we changed the network properties for "Local Area Connection" then setup static IP address for a DNS configuration

(V) Suppose within your Web browser you click on a link to obtain a Web page. 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 an RTT of RTT1 +RTT2 +...+RTTn. Further suppose that the Web page associated with the link contains exactly one object, consisting of a small amount of HTML text. Let RTT0 denote the RTT between the local host and the server containing the object. Assuming zero transmission time of the object, how much time elapses from when the client clicks on the link until the client receives the object?

For visits to n DNS servers, total RTTS = RTT1+RTT2+...+RTTn.
Suppose, a web page associated with the link contains exactly one object
Transmission time of the object = 0
RTT between the local host and the server contain the object = RTT0
So RTT0 to set up a TCP connection between local host and server.

Another RTT0 to request and get response for the object.

So, total time = 2RTT0 + RTT1+...+RTTn