Computer Networks - CS132/EECS148 - Spring 2013 Instructor: Karim El Defrawy Assignment 2

Deadline : April 25th – 9:30pm (hard and soft copies required)

Problem 1 (1 points) - What is the difference between pull and push network protocols? Explain the difference by using two example protocols.

Pull : HTTP, POP, IMAP : client requests (pulls) data from server Push : SMTP , server sends (pushes) its data to client whenever new data is available

Problem 2 - (Chapter 1 problem 31 , 5 points) In modern packet-switched networks, including the Internet, the source host segments long, application layer messages (for example an image or a music file) into smaller packets and sends the packets into the network. The receiver then resembles the packets back into the original message. We refer to this process as message segmentation. Consider a message that is 8*10^6 bits long that is to be sent from a source to a destination which are separated by two routers in between. Suppose each of the three link in this path is 2 Mbps. Ignore propagation, queueing and processing delays.

a. Consider sending the message without message segmentation. How long does it take to move the message from the source host to the first router ? Keeping in mind that each router uses a store-and-forward packet switching, what is the total time to move the message from source to destination?

b. Now suppose the packet is segmented into 800 packets with each packet being 10000 bits long. How long does it take to move the first packet to the first router ? When the first packet is being sent from the first router to the second one, the second packet is on its way to the first router. At what time will the second packet be fully received by the second router ?

c. What is the total to move the message to destination host ? Compare this result with your answer in part a.

d. In addition to reducing delay, what are the reasons to use message segmentation?

e. Discuss the drawbacks of message segmentation.

a) Time to send message from source host to first packet switch = $8*10^{6} / 2*10^{6} = 4$ sec. With store-and-forward switching, the total time to move message from source host to destination host = $4 \sec 3$ hops = $12 \sec 3$

b) Time to send 1st packet from source host to first packet switch = $1*10^{4}/2*10^{6} = 5$ m sec. Time at which second packet is received at the first switch = time at which 1 st packet is received at the second switch = 2×5 m sec = 10 m sec

c) Time at which 1st packet is received at the destination host = 5 m secx 3 hops = 15 m sec . After this, every 5msec one packet will be received; thus time at which last (800 th) packet is received = 15 m sec + 799 * 5m sec = 4.01 sec . It can be seen that delay in using message segmentation is significantly less (almost 1/3rd).

d)

i. Without message segmentation, if bit errors are not tolerated, if there is a single bit error, the whole message has to be retransmitted (rather than a single packet).ii. Without message segmentation, huge packets (containing HD videos, for example) are sent into the network. Routers have to accommodate these huge packets. Smaller packets have to queue behind enormous packets and suffer unfair delays.

e)

i. Packets have to be put in sequence at the destination.

ii. Message segmentation results in many smaller packets. Since header size is usually the same for all packets regardless of their size, with message segmentation the total amount of header bytes is more.

Problem 3 - (Chapter 2 problem 1, 5 points) True or False ?

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.

b. Two distinct web pages (for example, <u>www.mit.edu/research.html</u> and <u>www.mit.edu/students.html</u>) can be sent over the same persistent 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.

d. The *Data:* header in HTTP response message indicates when the object in the response was last modified.

e. HTTP response messages never have an empty message body.

- a) False
- b) True
- c) False
- d) False
- e) False

Problem 4 - (Chapter 2 problem 7, 4 points) Suppose within your web browser you click on a link to obtain a Web page. The IP address of the associated URL is not cached in your local host, so a DNS lookup is necessary to obtain the IP address. Suppose N DNS servers are visited before your host receives the IP address from DNS: the successive visits incur an RTT of RTT1, ..., RTTn. Further suppose that the webpage 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?

The total amount of time to get the IP address is RTT1 + RTT2 + \dots + RTTn .

Once the IP address is known, RTTO elapses to set up the TCP connection and another RTTO elapses to request and receive the small object. The total response time is

2 RTTo + RTT1 + RTT2 + ... + RTTn

Problem 5 (2 points)- Briefly sketch and explain the process of manipulation of packets in different layers of network on server and client side for retrieving CS132 website first page. Please include the name of layers and protocols which are used in each layer.

The sketch must show the stack of layers (Application, Transport, ...) in TCP/IP on both sides and explain how HTTP packet is initiated in application layer and is passed to TCP and TCP header is added and etc (similar to slides on general networking example)

Problem 6 (2 points) – In what layer of the network does UDP live in? By using an example, explain why it is necessary to have UDP. Name one application protocol which uses UDP as its supporting lower protocol.

Transport layer. The example can be either DNS or a simple replicated database architecture that the master uses UDP to find available servers. DNS uses UDP as its transport protocol.

Problem 7 - (Chapter 2 problem 8, 6 points) Referring to problem 4, suppose the HTML file references 8 very small objects on the same server. Neglecting transmission times, how time elapses with

a. Non-persistent HTTP with no parallel TCP connections?

b. Non-persistent HTTP with the browser configures for 5 parallel connections?

c. Persistent HTTP?

a) RTT1 + * + RTTn + 2 RTTo + 8 \cdot 2 RTTo

= 18RTTo + RTT1 + * + RTTn

b) RTT1 + * + RTTn + 2 RTTo + 2 · 2 RTTo

= 6 RTTo + RTT1 + * + RTTn

c) RTT1 + * + RTTn + 2 RTTo + RTTo

= 3RTTo + RTT1 + * + RTTn

Problem 8 - (Chapter 2 problem 9, 10 points) Consider Figure 2.12 in your book, 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 institution 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. Model the total average response time as the sum of the average access delay (that is the delay from the Internet router to institution router) and the average Internet delay. For the average access delay use Delta/(1 - Delta * Beta). Where Delta is the average time required to send an object over the access link and Beta is the arrival rate of objects to the access link.

a. Find the total average response time.

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

a) The time to transmit an object of size L over a link or rate R is L/R. The average time is the average size of the object divided by R:

 $\Delta = (850,000 \text{ bits})/(15,000,000 \text{ bits/sec}) = .0567 \text{ sec}$

The traffic intensity on the link is given by $\beta \Delta = (16 \text{ requests/sec})(.0567 \text{ sec/request}) = 0.907$. Thus, the average access delay is $(.0567 \text{ sec})/(1 - .907) \approx .6$ seconds. The total average response time is therefore .6 sec + 3 sec = 3.6 sec.

b) The traffic intensity on the access link is reduced by 60% since the 60% of the requests are satisfied within the institutional network. Thus the average access delay is (.0567 sec) /[1 - (.4)(.907)] = .089 seconds. The response time is approximately zero if the request is satisfied by the cache (which happens with probability .6); the average response time is .089 sec + 3 sec = 3.089 sec for cache misses (which happens 40% of the time). So the average response time is (.6)(0 sec) + (.4)(3.089 sec) = 1.24 seconds. Thus the average response time is reduced from 3.6 sec to 1.24 sec

Problem 9 - (Chapter 2 problem 15, 4 points) Read RFC 5321 for SMTP. What does MTA stand for? Consider the spam email included in your book below this problem. Assuming the originator of this spam email is malicious and all other hosts are honest, identify the malicious host that has generated this email.

MTA stands for Mail Transfer Agent. A host sends the message to an MTA. The message then follows a sequence of MTAs to reach the receiver's mail reader. We see that this spam message follows a chain of MTAs. An honest MTA should report where it receives the message. Notice that in this message, "asusus-4b96 ([58.88.21.177])" does not report from where it received the email. Since we assume only the originator is dishonest, so "asusus-4b96 ([58.88.21.177])" must be the originator.

Problem 10 (8 points) - Consider DNS protocol.

a. Sketch the main architecture of domain name servers currently used for the internet. You should include a client, a server and different types of DNSs.

Sketch on page 138 of book is good.

b. Explain four different types of record which are kept in a DNS list. According to your sketch in part a, in what point is each of these types used ?

A, CNAME, NS and MX. CNAME is used in authoritative or TLD servers. NS is used in authoritative servers. A is kept in local DNSs. MX is used for email applications.

c. Name two different types of a DNS query. Complete your sketch with some arrows to explain how these two approaches work.

Iterative and Recursive. In iterative approach, client only talks to local DNS, local talks to Root DNS, Root talks to TLD, TLD talks to authoritative and In recursive approach, client talks to local DNS, then talks to Root, then talks to TLD, then to authoritative and d. What is the necessity of having Root Domain Name Servers ?

Adding a new TLD requires updating only 13 entities in the network.

Problem 11 (Modified version of Chapter 2 problem 22, 6 points) - Consider distributing a file of F = 15 Gbits to N peers. The server has an upload rate of Us = 30 Mbps, and each peer has a download rate of Di = 2 Mbps and an upload rate of U. We have N peers in total.

a. Assume U = 700Kbps. Do the math and compare the total amount of time needed to distribute the file for N = 10 and N = 100. How does increasing N affect this comparison ?

Direct communication : max {10 (or 100) * 15000 / 30 , 15000 / 2}

P2P communication : max $\{15000 / 30, 15000 / 2, 10 \text{ (or } 100) * 15000 / (30 + 10(or 100)*0.7)\}$ If number of peers is increased enough, P2P becomes greater finally.

b. Assume N = 50. Compare this time for U = 300 Kbps and U = 700 Kbps. How does increasing peer's upload rate (U) affects this comparison?

Direct communication : max {50 * 15000 / 30 , 15000 / 2}

P2P communication : max {15000 / 30 , 15000 / 2 , 50 * 15000 / (30 + 50*0.3(or 0.7))}

If upload rate of peers is increased enough, P2P becomes greater finally.

Problem 12 (2 points) - Consider a P2P overlay network.

a. One example of this network is Circular Architecture. Depict this architecture and briefly explain how a file is distributed among peers in this approach.

Page 153 in book

b. Add shortcuts to your answer for part a. How is total upload time affected by adding shortcuts? What is the drawback of adding more edges to the overlay network ?

Total time is decreased because of shortcuts but managing the network becomes more difficult due to peer churn.

Problem 13 - (Chapter 2 problem 29, 5 points) Because an integer in $[0, 2^n - 1]$ can be expressed as an n-bit binary number in a DHT, each key can be expressed as k = (k0,k1,..kn-1), and each peer identifier can be expressed as p = (p0, p1, ..., pn-1). Let's now define the XOR distance between a key k and a peer p as

 $d(k, p) = |k0 - p0|^{*}2^{0} + ... + |kn-1 - pn-1|^{*}2^{(n-1)}$

Describe how this metric can be used to assign (key, value) pairs to peers.

For each key, we first calculate the distances (using d(k,p)) between itself and all peers, and then store the key in the peer that is closest to the key (that is, with smallest distance value).

Extra Credit

You can submit the answers to Problem 19, or Wireshark Lab problems of chapter 2 for 10 points extra credit.