CSE434 Computer Networks (Fall, 2009)
Homework 3
Due: Monday, October 12, 2009

Submission Procedure: No late submissions will be accepted. Submit a hardcopy of your answers before the class.
Note: Please note that not all questions will be graded. Be clear and technically precise arguments for your answers. Answers without proper justification will not get full credit even thought the final result is correct. A student might need to read a section in the book. If you need a copy, contact the TA (su.kim@asu.edu).

1. What aspect of IP addresses makes it necessary to have one address per network interface, rather than just one per host? In light of your answer, why does IP tolerate point-to-point interfaces that have nonunique addresses or no addresses?

2. Suppose an IP packet is fragmented into 10 fragments, each with a 1% (independent) probability of loss. To a reasonable approximation, this means there is a 10% chance of losing the whole packet due to loss of a fragment. What is the probability of net loss of the whole packet if the packet is transmitted twice,

   a) assuming all fragments received must have been part of the same transmission?

   b) assuming any given fragment may have been part of either transmission?

3. Having ARP table entries time out after 10 – 15 minutes is an attempt at a reasonable compromise. Describe the problems that can occur if the timeout value is too small or too large.

4. IP currently uses 32-bit addresses. If we could redesign IP to use the 6-byte MAC address instead of the 32-bit address, would we be able to eliminate the need for ARP? Explain why or why not.

5. Suppose hosts A and B have been assigned the same IP address on the same Ethernet, on which ARP is used. B starts up after A. What will happen to A’s existing connections? Explain how “self-ARP” (querying the network on startup for one’s own IP address) might help with this problem.

6. Suppose an IP implementation adheres literally to the following algorithm on receipt of a packet, P, destined for IP address D:

```
If (<Ethernet address for D is in ARP cache>)
        <send P>
Else
        <send out an ARP query for D>
        <put P into a queue until the response comes back>
```

a) If the IP layer receives a burst of packets destined for D, how might this algorithm waste resources unnecessarily?

b) Sketch an improved version.

c) Suppose we simply drop P, after sending out a query, when cache lookup fails. How would this behave? (Some early ARP implementations allegedly did this.)

7. 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 $RTT_1, \ldots, RTT_n$. Further suppose that the Web page associated with the link contains exactly one object, consisting of a small amount of HTML text. Let RTT_0 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?

8. Referring to Question 7, suppose the HTML file references 8 very small objects on the same server. Neglecting transmission times, how much time elapses with

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

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

   c) Persistent HTTP?

9. (Optional) Consider Figure 1, 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 in the text book). 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.

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

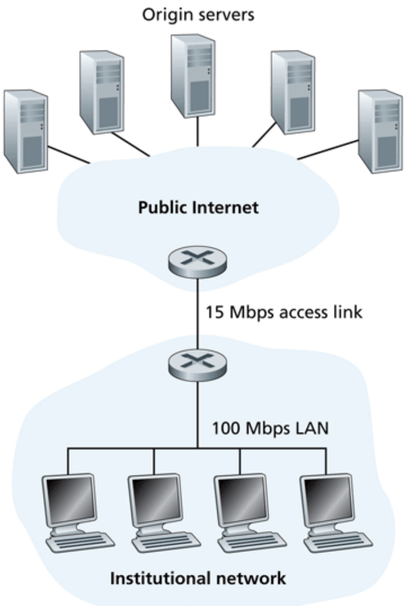

Figure 1: Bottleneck between an institutional network and the Internet

10. Consider distributing a file of F = 15 Gbits to N peers. The server has an upload rate of u_s = 30 Mbps, and each peer has a download rate of d_i = 2 Mbps and an upload rate of u. For N = 10, 100, and 1000 and u = 300 Kbps, 700 Kbps, and 2 Mbps, prepare a chart giving the minimum distribution time for each of the combinations of N and u for both client-server distribution and P2P distribution.

11. (Optional) Consider distributing a file of F bits to N peers using a client-server architecture. Assume a fluid model where the server can simultaneously transmit to multiple peers, transmitting to each peer at different rates, as long as the combined rate does not exceed u_s.

    a) Suppose that $u_s /N \leq d_{min}$. Specify a distribution scheme that has a distribution time of NF/u_s.

    b) Suppose that $u_s /N \geq d_{min}$. Specify a distribution scheme that has a distribution time of F/ d_{min}.

    c) Conclude that the minimum distribution time is in general given by max $\{NF/u_{s,} F/ d_{min}\}$.

12. (Optional) In this problem, we are interested in finding out the efficiency of a BitTorrent-like P2P file sharing system. Consider two peers Bob and Alice. They join a torrent with M peers in total (including Bob and Alice) that are sharing a file consisting of N chucks. Assume that at a particular time t, the chunks that a peer has are uniformly at random chosen from all N chucks, and no peer has all N chucks. Answer the following questions.

a) What is the probability that Bob has all the chunks that Alice has, given that the numbers of chunks that Bob and Alice have are denoted by n_b and n_a?

b) Remove part of the conditioning in part a) to find out the probability that Bob has all the chunks that Alice has, given that Alice has n_a chunks?

c) Suppose that each peer in BitTorrent has 5 neighbors. What is the probability that Bob has data that is of interest to at least one of his five neighbors?

13. In the circular DHT example in Section 2.6.2, suppose that peer 3 learns that peer 5 has left. How does peer 3 updat its successor state information? Which peer is now its first successor? Its second successor?

14. In the circular DHT example in Section 2.6.2, suppose that a new peer 6 wants to join the DHT and peer 6 initially only knows peer 15's IP address. What steps are taken?

15. Consider a circular DHT with node and key identifiers in the range [0, 63]. Suppose there are eight peers with identifiers 0, 8, 16, 24, 32, 40, 48, and 56.

    a) Suppose each peer can have one shortcut peer. For each of the eight peers, determine its shortcut peer so that the number of messages sent for any query (beginning at any peer) is minimized.

    b) Repeat (a) but now allow each peer to have two shortcut peers.

16. As DHTs are overlay networks, they may not necessarily match the underlay physical network well in the sense that two neighboring peers might be physically very far away; for example, one peer could be in Asia and its neighbor could be in North America. If we randomly and uniformly assign identifiers to newly joined peers, would this assignment scheme cause such a mismatch? Explain. And how would such a mismatch affect the DHT's performance?