HW #1 Solutions

ELEN E6761 - Computer Communication Networks
Fall 2000 Prof. Rubenstein

Please turn your homework into Vassilis Stachtos' mailbox, E2 (oustide of the Electrical Engineering office, or fax it by 5pm. On-campus students can fax to Professor Rubenstein at (212) 932-9421 (remember to include your name and that the fax is for Professor Rubenstein). CVN students can fax as above, or fax to the CVN fax number. LATE ASSIGNMENTS WILL NOT BE ACCEPTED!! Note: don't wait until the last minute to fax - only one can go through at a time.

1. Why are the MAC (link layer) addresses fixed on the network card, whereas IP addresses must be assigned by a local administrator?

   Answer: MAC layer addresses are fixed to reduce installation complexity (i.e., there is no need for a sys-admin to configure devices - they are simply plug-n-play). IP addresses are assigned locally to reduce the amount of information that routers must maintain to determine where to forward packets (e.g., they need only identify the longest matching prefix). Were IP addresses fixed on the network card, the state that routers would have to maintain would increase significantly. Were MAC addresses assigned by a local administrator, there would be a significant increase in man-hours whenever the LAN was modified (hosts moved, removed, or added, etc.)

2. When is DHCP beneficial? Why isn't it always used to assign IP addresses to an interface?

   Answer: DHCP is beneficial when a set of hosts can "share" some subset of addresses. For instance, if 10 users have laptops and each user only uses their laptop 1 day during the work-week (2 on Mon, 2 on Tues, etc.), then 2 IP addresses assigned with DHCP can be used to support all 10 laptops, as opposed to needing 10 IP addresses, one for each laptop (i.e., a fixed mapping from the IP address to the MAC address on the laptop's network interface.) DHCP is not always useful, in that sometimes, a fixed mapping is required. For instance, if a host acts as a server, then other hosts need a fixed address in order to contact the host. Assigning the server a different IP address every time it is disconnected from and then reconnected to the network causes needless complications.

3. Typically, a web browser provides a cache that caches web pages. This functionality could be incorporated into the DNS cache as well: rather than simply cache IP address-to-hostname mapping, the DNS cache could store the web-page and return the page upon receiving a request. Give three reasons why this is a bad idea.

   Answer: Reasons are:

   (a) A host that acts as a web-server might perform other operations as well (e.g., telnet, e-mail) and so it should not be assumed that every DNS request for the server pertains to a web-related request.

   (b) The expiry time for a DNS mapping can be different than that for a web page. For instance, the IP address for a stock quotes web page can be fixed for days or months, whereas the web-page info needs to be updated minute-by-minute. Hence, DNS-related processing would increase dramatically, or the web pages returned would be very out of date.

   (c) The memory required to store a web-page is much larger than the 4 bytes required to store an address. Hence, DNS server memory requirements would increase.

4. Assume that the propagation delay between any pair of hosts, A and B (i.e., the time it takes a packet transmitted by A to reach B) is a constant, $C_{A,B}$. Also assume that the time to process a request is the same for a recursive query as it is for an iterated query. Also suppose that every DNS server X that would take a request from A (where A is either a host or another DNS server) and forward that request to B (either recursively or iteratively) knows $C_{A,B}, C_{B,A}, C_{A,X}, C_{X,A}, C_{X,B}$, and $C_{B,X}$. Describe the algorithm that X should use to determine whether to respond to A's request via an iterative or recursive query so that overall DNS query response times are minimized. Assume now that the a query from a host X_0 follows a fixed chain of DNS servers, $X_1, X_2, \cdots, X_n$ (i.e., for $1 \leq i < n$, each host X_i either makes a recursive request to X_{i+1}, or (for the iterative case), tells the host that contacted it [some X_j where $j < i$] where to send a request to X_{i+1}). Does the algorithm you proposed minimize overall DNS query response time. If so, prove the result. If not, give a counterexample.

Answer: A recursive query takes time $C_{A,X} + C_{X,B} + C_{B,X} + CX, A$, and an iterative query takes time $C_{A,X} + CX, A + C_{A,B} + C_{B,A}$. To minimize query time, a recursive query should be used whenever $C_{X,B} + C_{B,X} < C_{A,B} + C_{B,A}$.

Applying this algorithm does not minimize overall DNS query response time. Consider the following scenario with the given propagation delays between DNS servers:

- $C_{X_0,X_1} = 1$
- $C_{X_1,X_2} = 1$
- $C_{X_0,X_2} = .5$
- $C_{X_1,X_i} = .001$ for $i > 2$
- $C_{X_i,X_j} = 10$ for $i = 0$ or $i \geq 2$ and $j > 2$.

Shortest time is $X_0 \rightarrow 1$ and then X_1 contacts all other hosts via iterated queries. But algorithm would have X_0 do iterative query to X_2, hence X_1 would be out of the loop for all remaining transactions.

5. Assume the same model as in problem 4, and also that for all pairs of hosts A, B, $C_{A,B} = C$, i.e., propagation delays between all pairs of hosts are the same. Why would iterative queries be better for this case?

Answer: Iterative queries are better because not only will end-to-end propagation delays be the same as they would were recursive queries being used, but intermediate servers would not need to maintain the state of who contacted them for a DNS lookup. This state is necessary so that the intermediate server knows where to forward its response once received. This could save enormous amounts of memory in servers that handle the processing of numerous requests for a DNS address.

6. Suppose a router (using CIDR) advertises the prefix $w.x.y.z/n$. Is there ever a reason that it would have also advertise $w.x.y.z/m$ where $m > n$? Explain.

Answer: Yes (even if both advertisements are for the same interface). For instance, consider the case where interface A is used to deliver packets with IP addresses 128.150.0.0 through 128.150.127.255 and addresses 128.150.196.0 through 128.150.255.255, and interface B is used to deliver packets with IP addresses 128.150.128.0 through 128.150.195.255. This is most efficiently accomplished by having interface A advertise 128.150.0.0/16 (prefix 10000000 10010110) and 128.150.196.0/18 (prefix 10000000 10010110 11), and having B advertise 128.150.128.0/17 (prefix 10000000 10010110 1).

7. Given the fact that bridges can be used to connect an unlimited number of hosts together, give two motivations for using routers in large-scale networks, even though they cost significantly more than bridges.

Answer: Because MAC addresses are fixed on cards, there is no general efficient means to compress state in bridges (like there is in IP addressing via shortest prefix). Since bridges are layer 2 devices and must operate independent of the network layer (layer 3), they cannot take advantage of the layer 3 addressing structure. Second, bridge routing algorihtms are designed to be simple, not scalable, so a spanning tree would have to be constructed and maintained. Routing protocols, for instance, handle temporary problems routing inconsistencies such as loops in the routing graph.

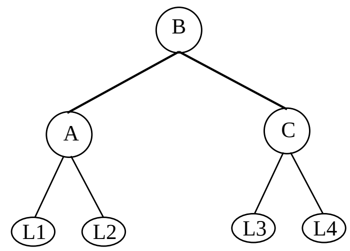

Figure 1:

8. You are designing a network with the topology given in Figure 1, where points A, B, C are ideal locations to place a bridge or a hub. $L1, L2, L3$, and $L4$ are LANs, each of contains 10 hosts. Assume (for simplicity) that transmissions by hosts are done on a round-by-round basis. On any given round, a host may attempt to transmit a frame to another host on any of the four LAN. Any frames that share a wire within a round will collide and cannot be received.

   Assume that in any given round, each host transmits with a probability of p, and that all hosts (on $L1, L2, L3, L4$) have the same likelihood of being the destination of a transmitted frame (i.e., the conditional probability that host y is the destination of a frame transmitted by host x given that host x transmitted a frame is $1/40$).

   Consider a given round and a given host x that lies on $L1$. Let T_x be a random variable that equals 1 if host x transmits a frame, and 0 otherwise. Let S_x be a random variable that equals 1 if host x successfully transmits a frame (i.e., there is), and 0 otherwise. Calculate $\Pr(S_x = 1 \mid T_x = 1)$ when

   (a) hubs are placed at A, B, and C.

   Answer: First, we point out that in general, $\Pr(T_x = 1) = p$. With hubs placed at all locations, any signal transmitted from any host other than x to any location will be broadcast on the entire LAN, and would therefore collide with x's transmission. Thus none of the other 39 hosts can transmit anywhere. Since the other hosts independently transmit with probability p, the probability that none of them transmit is $(1-p)^{39}$, making $\Pr(S_x = 1, T_x = 1) = p(1-p)^{39}$, and so $\Pr(S_x = 1 \mid T_x = 1) = \Pr(S_x = 1, T_x = 1)/\Pr(T_x = 1) = (1-p)^{39}$.

   (b) a bridge is placed at A and hubs are placed at B and C.

   Answer:
   Define D_x to be a random variable that equals 1,2,3, or 4 when x transmits to a host in L_1, L_2, L_3, or L_4 respectively. We also note that because $\Pr(T_x = 1, D_x = i) = p/4$ for $i = 1, 2, 3, 4$. We note that $\Pr(S_x = 1 \mid T_x = 1) = \frac{\sum_{i=1}^{4} \Pr(S_x=1, T_x=1, D_x=i)}{\Pr(T_x=1)}$.
   Let $q_0 = 1-p, q_1 = 1-p/4, q_2 = 1-p/2, q_3 = 1-3p/4$ (as suggested in the hint at the end of the problem).

The following table indicates where hosts, grouped by their location in the network (i.e., on L_1, L_2, L_3, and L_4) can transmit as a function of where host x transmits, and the probabilities that no host on that given portion of the network sends a frame that collides with a frame sent by host x:

x sends to	prob	other L_1 hosts	prob	L_2 hosts	prob	L_3 or L_4 hosts	prob
L_1	$p/4$	nowhere	${q_0}^9$	no L_1	${q_1}^{10}$	no L_1	${q_1}^{20}$
L_2	$p/4$	nowhere	${q_0}^9$	nowhere	${q_0}^{10}$	no L_1, L_2	${q_2}^{20}$
L_3 or L_4	$p/2$	nowhere	${q_0}^9$	no L_1, L_3, L_4	${q_3}^{10}$	nowhere	${q_0}^{20}$

Hence,

$$\begin{aligned}\Pr(S_x = 1 \mid T_x = 1) &= \Pr(S_x = 1, T_x = 1)/\Pr(T_x = 1) \\ &= (\frac{p}{4}{q_0}^9{q_1}^{30} + \frac{p}{4}{q_0}^{19}{q_2}^{20} + \frac{p}{2}{q_0}^{29}{q_3}^{10})/p \\ &= ({q_0}^9{q_1}^{30} + {q_0}^{19}{q_2}^{20} + 2{q_0}^{29}{q_3}^{10})/4.\end{aligned}$$

(c) a bridge is placed at B and hubs are placed at A and C.

Answer:

The table below has the same properties as the previous table, except it lists the properties of the current LAN architecture.

x sends to	prob	other L_1 hosts	prob	L_2 hosts	prob	L_3 or L_4 hosts	prob
L_1 or L_2	$p/2$	nowhere	${q_0}^9$	nowhere	${q_0}^{10}$	no L_1, L_2	${q_2}^{20}$
L_3 or L_4	$p/2$	nowhere	${q_0}^9$	nowhere	${q_0}^{10}$	nowhere	${q_0}^{20}$

Hence,

$$\begin{aligned}\Pr(S_x = 1 \mid T_x = 1) &= \Pr(S_x = 1, T_x = 1)/\Pr(T_x = 1) \\ &= (\frac{p}{2}{q_0}^{19}{q_2}^{20} + \frac{p}{2}{q_0}^{39})/p \\ &= ({q_0}^{19}{q_2}^{20} + {q_0}^{39})/2.\end{aligned}$$

(d) a bridge is placed at C and hubs are placed at A and B.

Answer:

The following is the table for the current architecture:

x sends to	prob	other L_1 hosts or L_2 hosts	prob	L_3 hosts	prob	L_4 hosts	prob
L_1 or L_2	$p/2$	nowhere	${q_0}^{19}$	no L_1 or L_2	${q_2}^{10}$	no L_1 or L_2	${q_2}^{10}$
L_3	$p/4$	nowhere	${q_0}^{19}$	nowhere	${q_0}^{10}$	no L_1, L_2, L_3	${q_3}^{10}$
L_4	$p/4$	nowhere	${q_0}^{19}$	no L_1, L_2, L_4	${q_3}^{10}$	nowhere	${q_0}^{10}$

Hence,

$$\begin{aligned}\Pr(S_x = 1 \mid T_x = 1) &= \Pr(S_x = 1, T_x = 1)/\Pr(T_x = 1) \\ &= (\frac{p}{2}{q_0}^{19}{q_2}^{20} + \frac{p}{2}{q_0}^{29}{q_3}^{10})/p \\ &= ({q_0}^{19}{q_2}^{20} + {q_0}^{29}{q_3}^{10})/2.\end{aligned}$$

(e) bridges are placed at A and C and a hub is placed at B.

(f) bridges are placed at A,B, and C.

Answer: It turns out that (8e) and (8f) have the same solution, i.e., with bridges at A and C, the bridge at B would never filter any packets.

x sends to	prob	other L_1 hosts	prob	L_2 hosts	prob	L_3 hosts	prob	L_4 hosts	prob
L_1	$p/4$	nowhere	${q_0}^9$	no L_1	${q_1}^{10}$	no L_1	${q_1}^{10}$	no L_1	${q_1}^{10}$
L_2	$p/4$	nowhere	${q_0}^9$	nowhere	${q_0}^{10}$	no L_1, L_2	${q_2}^{10}$	no L_1, L_2	${q_2}^{10}$
L_3	$p/4$	nowhere	${q_0}^9$	no L_1, L_3, L_4	${q_3}^{10}$	nowhere	${q_0}^{10}$	no L_1, L_2, L_3	${q_3}^{10}$
L_4	$p/4$	nowhere	${q_0}^9$	no L_1, L_3, L_4	${q_3}^{10}$	no L_1, L_2, L_4	${q_3}^{10}$	nowhere	${q_0}^{10}$

Hence,

$$\begin{aligned}
\Pr(S_x = 1 \mid T_x = 1) &= \Pr(S_x = 1, T_x = 1)/\Pr(T_x = 1) \\
&= (\frac{p}{4}{q_0}^9{q_1}^{30} + \frac{p}{4}{q_0}^{19}{q_2}^{20} + \frac{p}{2}{q_0}^{19}{q_3}^{20})/p \\
&= ({q_0}^9{q_1}^{30} + {q_0}^{19}{q_2}^{20} + 2{q_0}^{19}{q_3}^{10})/4.
\end{aligned}$$

You may assume that the bridges' forwarding tables have stabilized (i.e., the bridge filters frames whenever possible).

Hint: Let $q_0 = 1-p, q_1 = 1-p/4, q_2 = 1-p/2, q_3 = 1-3p/4$. Your solutions should easily be expressed as functions of q_0, q_1, q_2, and q_3.

Assume now that $p = 1/40$, and that bridges cost more than hubs. If you only had a single bridge, would you install it at B? Why or why not? Argue why this would be the case for any $0 < p < 1$.

Answer:

If $p = 1/40$, then having a bridge at just location A means host x's transmission is successful (given it transmits) with probability .484. At B, the probability of success is .427, and at C, the probability of success is .439. Hence, placing the bridge at B yields the lowest probability of success for host x of the three cases. These same values give the probability that a transmission by any host on L_1 or L_2 is successful. Due to symmetry in the model, for a host y on L_3 or L_4, with a bridge placed only at A, the likelihood of y's transmission being successful is the same as the likelihood of x's transmission being successful when a bridge is placed at C. With a bridge placed only at C, the likelihood of y's transmission being successful is the same as the likelihood of x's transmission being successful when a bridge is placed at A. And for a bridge at only B, x and y have the same likelihoods of a transmission being successful. This likelihood is lowest when the bridge is placed at B, so I would never install my only bridge at location B.

For any p, it would always be of more benefit to install a bridge either at A or C instead of at B. Placing a bridge at B only prevents collisions between a host x and y when host x transmits from either L_1 or L_2 to either L_1 or L_2 and host y transmits from either L_3 or L_4 to either L_3 or L_4. But note that a bridge at either A or C would prevent such collisions as well. The set of transmissions that cause collisions if a bridge is only installed at A is a proper subset of the transmissions that can cause collisions if a bridge is installed at B. For instance, installing a bridge at A prevents transmissions from x to other hosts on L_1 from colliding with transmissions on L_2 to other hosts on L_2. These transmissions would collide with a bridge only installed at B. The same can be said for transmissions from a host y on L_4 communicating with hosts on L_3 when a bridge is placed at C. Hence, it is always favorable to put the only bridge at a location other than B.