Application Layer Sample Questions

Question 2.

Consider the figure below, for which there is an institutional network connected to the Internet. Suppose that the average object size is 900,000 bits and that that the average request rate from the institution's browsers to the origin server is 1.5 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 a HTTP request until it receives the response in two seconds on average. Model the total average response time as the sum of the average access delay and the average Internet delay. For the average access delay, use A/(1-AB) where A is the average time required to send an object over the access link and
B is the arrival rate of objects to the access link. You can assume that the HTTP request messages are negligibly small and thus create no traffic on the network or the access link.

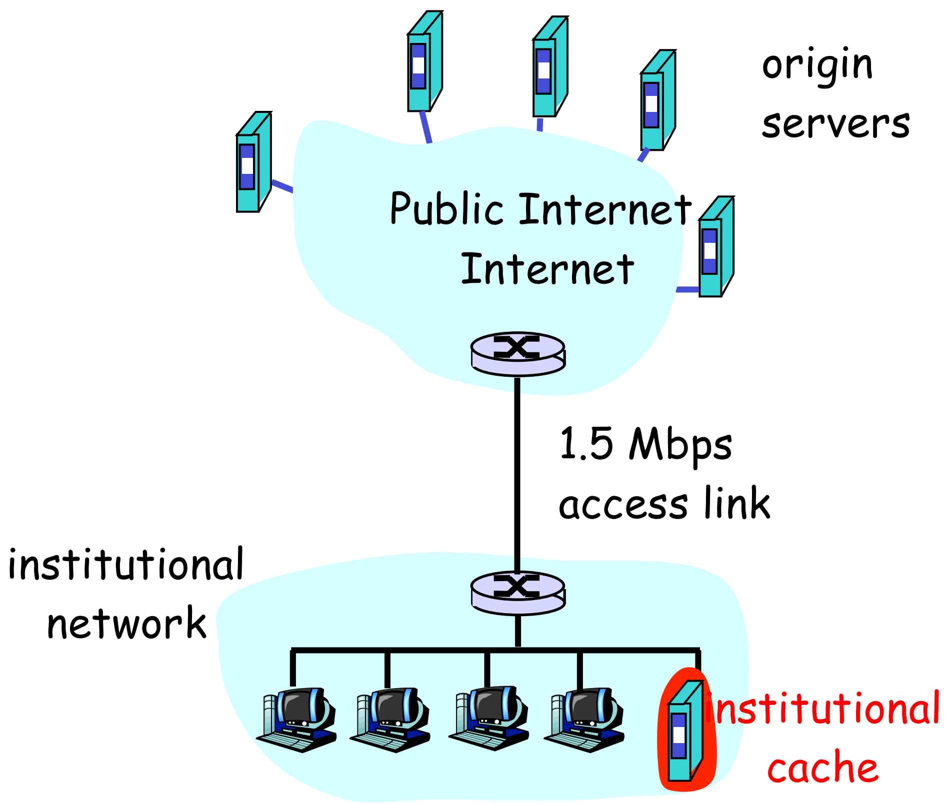

Figure 1: Figure for web cache problem

(a) Find the total average response time.

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

A= (900,000 bits)/(1,500,000 bits/sec) = .6 sec

The traffic intensity on the link is = AB= (.6 msec/request) (1.5 requests/sec) = .9. Thus, the average access delay is (.6 sec)/(1 - .9) = 6 seconds. The total average response time is therefore 6 sec + 2 sec = 8 sec.

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

Answer: The traffic intensity on the access link is reduced by 40% since the 40% of the requests are satisfied within the institutional network. Thus, the arrival rate of the objects to the link also changes since only 60% of the objects need to be fetched from the origin servers (the rest are obtained from the cache). As a result, B=1.5 x 0.6 = 0.9 requests/sec.

Thus the average access delay is (.6 sec)/[1 – (.6)(.9)] = 1.2 seconds. The response time is approximately zero if the request is satisfied by the cache (which happens with probability .4); the average response time is 1.2 sec + 2 sec = 3.2 sec for cache misses (which happens 60% of the time). So the average response time is (.4)(0 sec) + (.6)(3.2 sec) = 1.92 seconds. Thus the average response time is reduced from 8 sec to 1.92 sec.