

Sample Question: Web Caching (please discuss)

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 the average request rate from the institution's browsers to the origin server is 1.5 requests per second. 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 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.

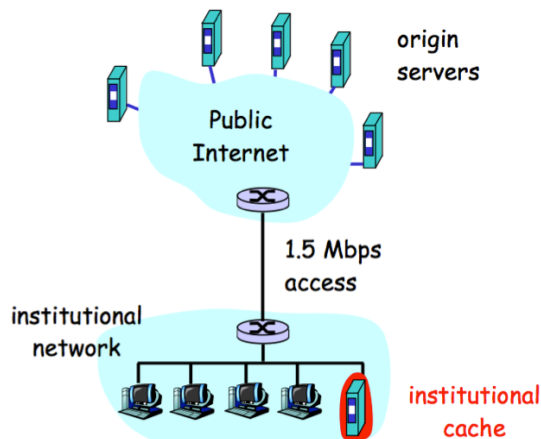


Figure 1: Figure for web cache problem

Question 1. Find the total average response time?

Question 2. Now suppose a cache is installed in the institutional LAN. Suppose the cache hit rate is 0.4. Find the total response time.

Resource created about a month ago (Monday 16 July 2018, 02:50:37 PM), last modified 26 days ago (Friday 27 July 2018, 11:19:46 AM).

Comments

🔍 (/COMP3331/18s2/forums/search?forum_choice=resource/17223) 🗨 (/COMP3331/18s2/forums/resource/17223)

💬 Add a comment



Daniel Hocking (/users/z5184128) 13 days ago (Thu Aug 09 2018 14:14:57 GMT+1000 (Australian Eastern Standard Time))

Assuming that the institutional network is much faster than the access link, there will be basically no delay added by that part of the journey, also assuming that cache hit is basically an instant response.

1 object = 900,000 bits

access link = 1,500,000 bits / sec

time to send object (A) = 0.6 sec

request rate (B) = 1.5 requests / sec

internet response time = 2 sec

average access delay = $A / (1 - AB) = 0.6 / (1 - 0.9) = 6$ sec

Total response time = 8 sec (or would it be +0.6 sec for time to send object as well?)

Q2) Suppose there is a cache hit rate of 0.4

now request rate (B) = 0.9 requests / sec (requests not found in cache)

average access delay = $A / (1 - AB) = 0.6 / (1 - 0.54) = 1.3$ sec

Total response time = 3.3 sec

(Unsure if I interpreted the equation correctly)

Reply



Nadeem Ahmed (/users/z3003139) 13 days ago (Thu Aug 09 2018 21:30:43 GMT+1000 (Australian Eastern Standard Time))

Q1. 8 sec is correct.

Q2. Almost there. 3.3 sec is the delay involved for the cache miss event. The average response time would be $0.6 \cdot 3.3 + 0.4(0) = 1.98$ sec (neglecting the delay for the cache hit case, 40% of the time object transferred from the institutional cache).

Reply