CS 494/594 Homework 2: Part A (Spring 2019)

Instructor: Dr. Nirupama Bulusu

Due Date: 5/2/2019

Jiacheng Zhao - 494

594 Students:

Besides the homework problems below, please select one of the following three papers, and submit your review for your selected paper.

- John Dilley, Bruce M. Maggs, Jay Parikh, Harold Prokop, Ramesh Sitaraman, and Bill Weihl. **Globally Distributed Content Delivery.** IEEE Internet Computing, Vol 6. Issue 5. September 2002.
  URL: https://dl.acm.org/citation.cfm?id=613741

- John S. Otto, Mario A. Sánchez, John P. Rula, and Fabián E. Bustamante. **Content delivery and the natural evolution of DNS - Remote DNS Trends, Performance Issues and Alternative Solutions**, In *Proceedings of the ACM Internet Measurement Conference*, November 2012.
  URL: http://typica.cs.northwestern.edu/wp-content/uploads/2019/02/OttoIMC2012.pdf

- Affiliate Crookies: Characterizing Affiliate Marketing Abuse, Neha Chachra, Stefan Savage, and Geoffrey M. Voelker, Proceedings of the ACM Internet Measurement Conference, Tokyo, Japan, October 2015.
  URL: http://www.sysnet.ucsd.edu/~voelker/pubs/crookies-imc15.pdf

1. (10 points) HTTP

You click on a link to a Web page referencing 15 very small objects. The objects are located on the same server. Ignoring transmissions times, how much time elapses (in Round Trip Time (RTT)) for:

(a)Non-persistent HTTP?

RTT1+RTT2+…+RTTn + 2RTT0 + 15*2RTT0 = 32RTT0 + RTT1+ …+RTTn

(b) Persistent HTTP with at most 4 pipelined requests?

RTT1+RTT2+…+RTTn + 2RTT0 + RTT0 = 3RTT0 + RTT1+ …+RTTn

2. (10 points) DHT

Consider a circular Distributed Hash Table (DHT) with node identifiers in the range [0, 255]. Suppose there are six peers with identifiers 15, 47, 63, 127, 191 and 223. For each of the (key, value) pairs listed below that should be stored in the DHT, identify the peer responsible for storing that pair:

(key, value)	Peer Responsible for Storing the Pair
(30, 'thor')	47
(90, 'captain america')	127
(110, 'ironman')	127
(128, 'hulk')	191
(240, 'frozen')	15

3. (10 points) DNS

A network administrator reads an article claiming that a good way to improve performance of DNS is to make a slight modification to the client resolver. For a given name lookup, like looking up the A record of www.facebook.com, the suggestion is to have the client: Issue a recursive query to its local DNS server for the NS record of www.facebook.com. Send an iterative query to the authority listed in the NS record to get the A record.

(i) Does this modification influence the caching performance of NS records at the local DNS server? If so, how?

For NS records, the iterative query is listed. So it has influence for caching perfermance of NS records.

(ii) Does this modification influence the caching performance of A records at the local DNS server? If so, how?

For getting A records, the caching performances of A records is not influenced.

4. (10 points) BitTorrent

Consider the following simplified BitTorrent scenario. There is a swarm of 2^n peers and, during a considered time period, no peers join or leave the swarm. It takes a peer 1 unit of time to upload or download a piece, during which time it can only do one or the other. Initially one peer has the whole file and the others have nothing.

(i) If the swarms target file consists of only 1 piece, what is the minimum time necessary for all the peers to obtain the file? Consider only the upload/download time and ignore everything else.

All peer could have the file after n time units. Each peer with the piece can transmit it to one peer without the piece. The number of peers with the piece doubles with each time unit: 1 peer at time 0, 2 peers at time 1, 4 peers at time 2.

In n unit of time, 2^n peers at time n.

(ii) Let x be your answer to the preceding question. If the swarms target file instead consisted of 2 pieces, what is the minimum time necessary for all peers to obtain the target file?

All peers could have the file after less than 2n time units.
The minimum time necessary for all peers: n+2.

5. (10 points) Web Caching

The Soccer World Cup 2022 is going to take place in Qatar and you have been hired to help get the country's network ready for the event. Suppose that the country is interconnected through a 1 Gbps network. However, the link that connects Qatar to the Internet is only 100 Mbps.

Assume that the delay to retrieve an object from the Internet is on average 3 seconds, while the delay to get an object of similar size residing in the country's network is 0.01 seconds. You then decide to install a Web cache in the country's network whose hit rate is 40%.

(i) What's the average delay to access an object? Compare it to the initial "cache-less" configuration.

Reduction from 1Gbps inside to 100Mbps outside represent a bottleneck cause large cases of delay.

Delay when hit the cache = 0.01sec = 10ms

Delay when retrieve object = 3s * 1000 = 3000ms

Hit rate = 40 = 0.40

Average delay with cache = 0.60 (3000ms) + 0.40 (10ms) = 1800ms + 4 ms = 1804 ms average delay per access

Cache-less: 3000ms average delay per access

(ii) Is there a way to further decrease the average delay? Explain.

Yes, we can increase the hit rate of the cache by increasing the cache size. We can installing additional web caches and distributing requests to increase the hit rate.