

## Homework 2

*Handed Out: February 12<sup>th</sup>, 2023**Due: 11:59pm, February 19<sup>th</sup>, 2023**TA: Pradhyumna Padmanabha*

- Homework assignments must be submitted online through **GradeScope**. Hard copies are not accepted. Please submit a **pdf file** to GradeScope (<https://www.gradescope.com>). You can either type your solution or scan a **legible** hand-written copy. We will not correct anything we do not understand. Contact the TAs via Campuswire if you face technical difficulties in submitting the assignment. **Note: Please select your page(s) for each question during submission on GradeScope; Otherwise, your submission will not be graded.**
- Homework assignments can be done in **groups of two**, but **only one person** needs to submit on GradeScope. Remember to include your partners name in the solution and on GradeScope by **editing "Group Members"**. It is your responsibility that your partner's name is included. For detailed instruction please refer to ([https://youtu.be/rue7p\\_kATLA](https://youtu.be/rue7p_kATLA)).
- You can use Campuswire to find a partner. We highly recommend working in groups. You will not get extra credit for working alone.
- Please use Campuswire and come to office hours if you have questions about the homework. Failure to understand the solutions will be the student's fault.
- While we encourage discussion within and outside of the class, cheating and copying is strictly prohibited. Copied solutions will result in the entire assignment being discarded from grading at the very least and a report filed in the FAIR system. It is also your responsibility to ensure that your partner obeys the academic integrity rules as well.

## 1 DNS - 14 points

The task requires using the `dig` command to provide answers. To ensure accurate results, it is recommended to perform these steps from a computer located on a campus network. The user can refer to the `dig` documentation to understand how to utilize it.

1. Starting from one of the root servers `a-m.root-servers.net`, perform an *iterative* lookup for the host `www.eecs.mit.edu`. For instance, you can initiate the search by using the following command:

```
dig @h.root-servers.net www.eecs.mit.edu
```

Please provide a list of the following information for each name server you visit during the lookup process:

- (a) Can you specify the domain name of the name server being visited?
- (b) Can you provide the IP address of the name server that is currently being used?
- (c) How long did the query take?

- (d) What is the round-trip time (RTT) to the server, which can be determined by using the "ping" command to connect to the server?
  - (e) For how long can you store the results in cache?
2. Perform a *recursive* query using `resolver.illinois.edu` of the name `www.eecs.mit.edu`. Was this query faster or slower than the sum of the iterative steps? Why do you think that is?
  3. Perform an iterative reverse-mapping query for the address of `www.eecs.mit.edu` you found in the previous steps, using `dig -x`. List again the information asked in part 1.
  4. Can you explain why the DNS protocol tends to utilize UDP rather than TCP, considering what has been previously discussed in this inquiry?

## 2 Client-Server - 14 points

Think about spreading a  $F$ -bit file among  $N$  peers using a client-server structure. Let the server have an maximum upload capacity  $\mu_s$ , and each client  $c$  has a download capacity  $d_c$ . Let  $d_{\min} = \min_c d_c$  be the minimum download rate. Assume that the server can serve multiple clients simultaneously and fluidly set the rate for each client,  $r_c$ , as long as  $\forall_c r_c \leq d_c$  and  $\sum_c r_c \leq \mu_s$ .

1. Suppose that  $\mu_s/N \leq d_{\min}$ . How would you set the rates  $r_c$  so that the file is fully distributed to all clients in a minimum time? (i.e., you are minimizing the time that the slowest client receives the file.) What would the distribution time be?
2. Suppose now that  $\mu_s/N > d_{\min}$ . How would you set the rates  $r_c$  now to fully distribute the file to the clients in a minimum time? And what would this time be?
3. Consider a concrete example with 5 clients with  $d_c = \{06, 12, 18, 24, 30\}$  and a server upload capacity of  $\mu_s = 30$ . How would you set the rates to get the smallest average download time *without* increasing the total distribution time from the previous part?

## 3 BitTorrent - 16 points

BitTorrent employs a choking mechanism for distributing bandwidth to peers. It selects the four peers that have provided the best download performance using a tit-for-tat strategy and one additional peer chosen randomly, referred to as an "optimistic unchoking". The selection of the four best peers and the random choice are updated every second.

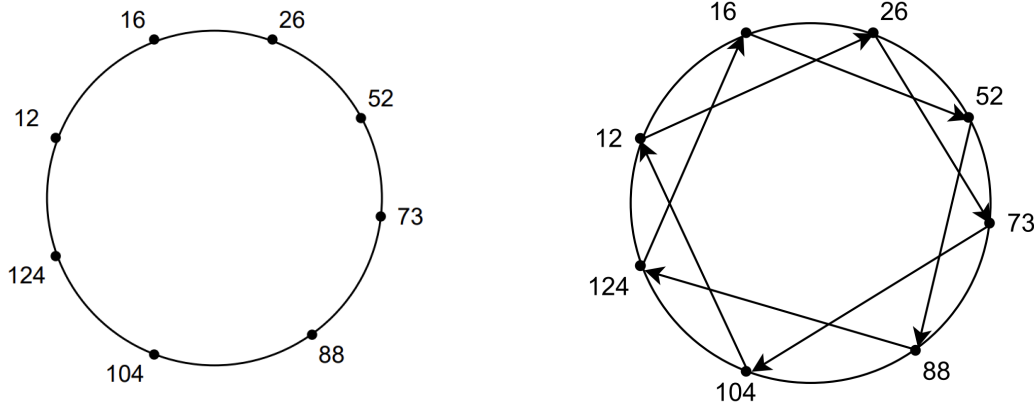
Suppose Bob joins a BitTorrent swarm with 30 other peers. Let each peer, including Bob, have an upload speed of 80 *Mbps* and unlimited download speed. Consider what happens in the phase of the protocol where for each pair of peers A and B, A has some blocks that B wants and vice versa; i.e, any peer can productively download data from any other peer.

1. If Bob intends to not upload any data and become a free rider, what would his average download speed be?

2. What would be the average download rate for the remaining peers?
3. Suppose that Bob runs a second client that pretends to be a separate peer, who also becomes a free rider. How fast can Bob download data now?
4. Suppose Bob switches his two clients to the regular BitTorrent code and they both start uploading as well. Will Bob's performance see a noticeable improvement compared to when he was not contributing?

## 4 DHT - 16 points

Consider a distributed hash table (DHT) with 8 peers and an 8-bit hash ID that operates in a circular fashion. The peers are represented by the values {12, 16, 26, 52, 73, 88, 104, 124}. The key-value pairs are assigned to the peer that immediately follows the key in the circular DHT.



1. What is hashing ID space?
2. In which peer is the key-value pair stored in the following circumstances:
  - (a) The key has a value of 19?
  - (b) The key has a value of 125?
3. If peer 88 initiates a search for the value associated with key 23, as shown in the illustration on the left, how many messages will be necessary to complete the query? This includes the message that initiates the query and the message that returns the value. Provide an explanation for the answer?
4. Assume each peer is given one cord and is now aware of the next two successors as shown in the right figure above. If peer 12 starts a query for the value associated with key 111, then how many messages (including the query message and return value message) are required to resolve this query? Explain why?
5. What is the maximum number of messages to resolve any query if

- (a) each peer is aware of its immediate successor as shown in the left figure above?
  - (b) each peer is given one cord and is now aware of the next two successors?
6. Given that each peer is provided with one cord, can you propose a new arrangement of the cords that would minimize the maximum number of messages needed to resolve any query? Draw a representation of the circular distributed hash table and provide an argument to support why your solution is effective.

## 5 HTTP Connections - 8 points

1. How does a web server handle multiple HTTP connections at the same time, even though all of the data packets received from different sessions are all directed to its TCP port 80? In other words, how does the server ensure that each packet is correctly delivered to its corresponding socket?
2. Is there a restriction on the number of simultaneous connections a single host can make with a web server, such as "a host can only maintain one HTTP connection with the server at any given time"? If such limitations exist, please describe them. If not, please explain the reasons why there are no limitations.

## 6 Reliable Data Transfer – 12 points

1. How can the maximum utilization of the network be calculated, when Alice is sending data to Bob using the stop-and-wait method, given that the bottleneck bandwidth is 40 Mbps and the end-to-end delay is 25 ms, and each packet being sent is 1200 bytes with 32 bits reserved for the checksum and sequence number?
2. How should Alice and Bob set their timeout value?
3. What changes would you make to the protocols if the connection between Alice and Bob is ensured to not cause any packet damage?
4. What would be the impact on performance of this change?