

# Problem Set #4

Due Wed, December 11<sup>th</sup> by 3:00 PM

## Problem 1 Hashing (20pts)

The hash table has 13 slots, and integer keys are hashed into the table with the following hash function H

```
int H (int key)
{
    x = ( key + 5 ) * ( key - 3 );
    x = int( x / 7 ) + key;
    x = x % 13;
    return x;
}
```

- (a) Fill in the final hash table with the following keys: 17, 22, 73, 56, 310, 100, 230, 12, 42, 18, 19, 24, 49.

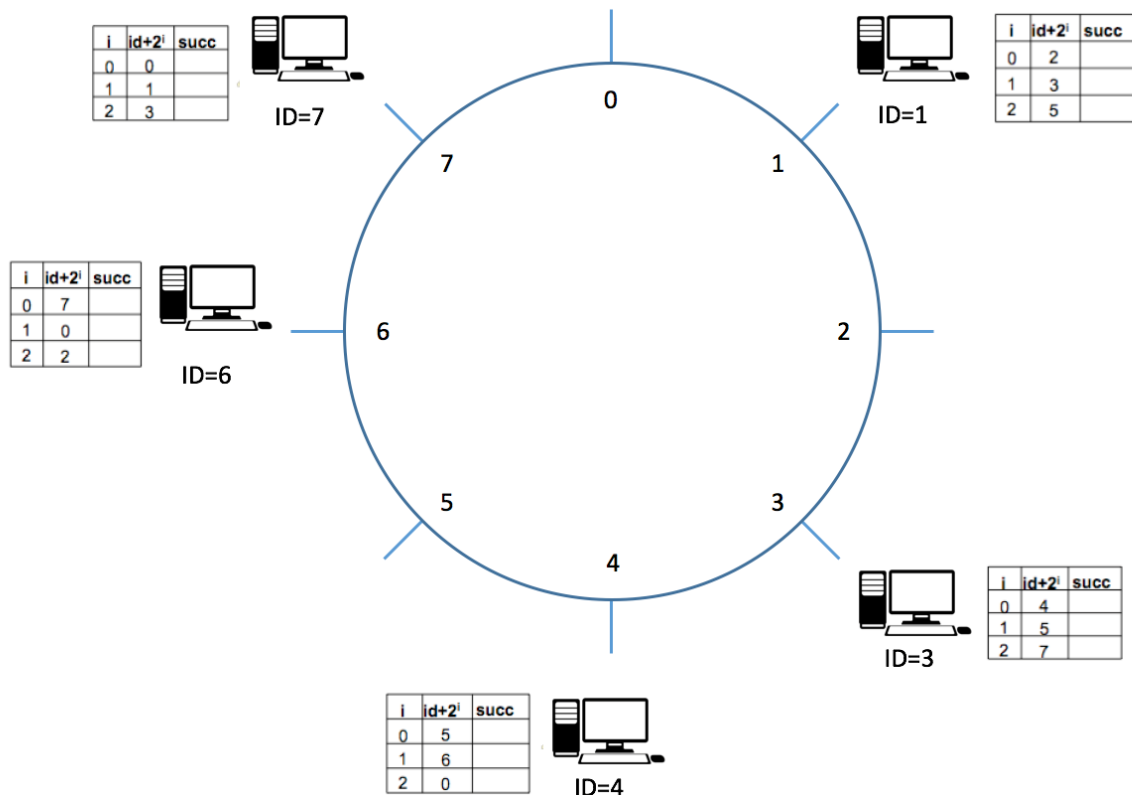
Slot	0	1	2	3	4	5	6	7	8	9	10	11	12
Contents													

- (b) List 2 methods that can handle collision in hashing.

## Problem 2 Distributed Hash Tables (20pts)

There is a Chord DHT in Figure 1. with 5 nodes. The finger tables are listed beside the nodes. Each node may be storing some items according to the Chord rules (Chord assigns keys to nodes in the same way as consistent hashing)

Figure 1. Chord DHT for Problem 2



(a) Fill in the table for node id=1 and 7

	$ID + 2^i$	successor
0	2	
1	3	
2	5	

Table for ID=1

	$ID + 2^i$	successor
0	0	
1	1	
2	3	

Table for ID=7

- (b) List the node(s) that will receive a query from node 1 for item 5 (item named by key 5)

### Problem 3 Bloom Filters (10pt)

**Derive** the probability of false positive rate after 10 keys (or elements) are inserted into a table of size 100. Assume that 5 hash functions are used to setup bit positions in the table for the keys (elements).

5 Hashing functions means: For each key (element) there will be 5 bits to be set to 1 in the table (it can also be less than 5 if the hash functions generate same outputs).

Hint:

Assume **m** is the number of bits in the filter, **n** is the number of elements and **k** is the number of hash functions used.

After inserting one key, the probability of a particular bit being 0 is  $(1 - \frac{1}{m})^k$ . This is because **k** hash functions are independent, and each hash function will have  $(1 - \frac{1}{m})$  probability for a particular bit remain 0. Then after inserting **n** keys, the probability for a particular bit remain 0 is  $(1 - \frac{1}{m})^{kn}$

Now you have to apply this to the case of false positive.

#### Problem 4 P2P system (10pt)

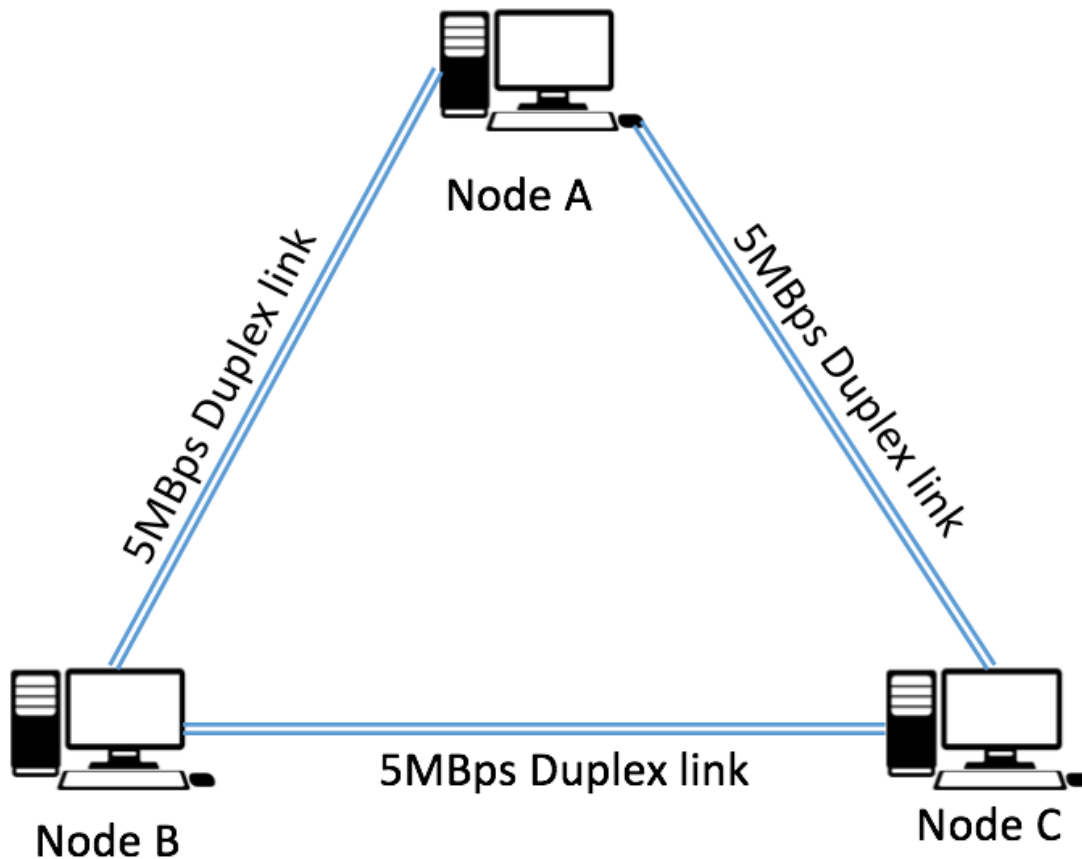


Figure 2. Network topology for Problem 4

3 nodes A, B and C are connected with each other via 5MBps duplex link as is now in Figure 2. Node A want to share a 2500MB file to Node B and C. During the actual transmission 2.5MB piece of the file can be send on the links each time. In the problem we ignore the RTT delay.

- (a) What is the time of the sharing process using centralized approach? (The process ends when B and C all received the file, and the centralized approach means B and C are communicating with A independently and there is no communication between B and C)
- (b) What is the ideal minimum time of the sharing process using P2P approach? (P2P approach means after B and C received a piece from A, they immediately share the their piece to other party)

### Problem 5 File Distribution (20pt)

Consider distributing a file of  $F = 20$  Gbits to  $N$  peers. The server has an upload rate of  $u_s = 30$  Mbps, and each peer has a download rate of  $d_i = 2$  Mbps and an upload rate of  $u$ . For  $N = 10$  and  $100$  and  $u = 300$  Kbps and  $2$  Mbps, prepare a chart giving the minimum distribution time for each of the combination of  $N$  and  $u$  for both client-server distribution and P2P distribution.

#### Client Server

U	N= 10	N= 100
300 Kbps		
2 Mbps		

#### Peer to Peer

U	N= 10	N= 100
300 Kbps		
2 Mbps		

### Problem 6 BGP (20pt)

1. Give the types of business relationships in BGP peering and mention who pays whom. What conditions make the Internet stable? Explain each condition in a line or two.

2. Please identify which of the following paths are valid, which of them are invalid based on the network topology of Figure 3.

Path 1 3 d  
Path 1 4 d  
Path 8 d  
Path 6 d  
Path 4 d  
Path 7 5 d  
Path 7 5 3 d  
Path 2 1 3 d  
Path 1 4 6 d

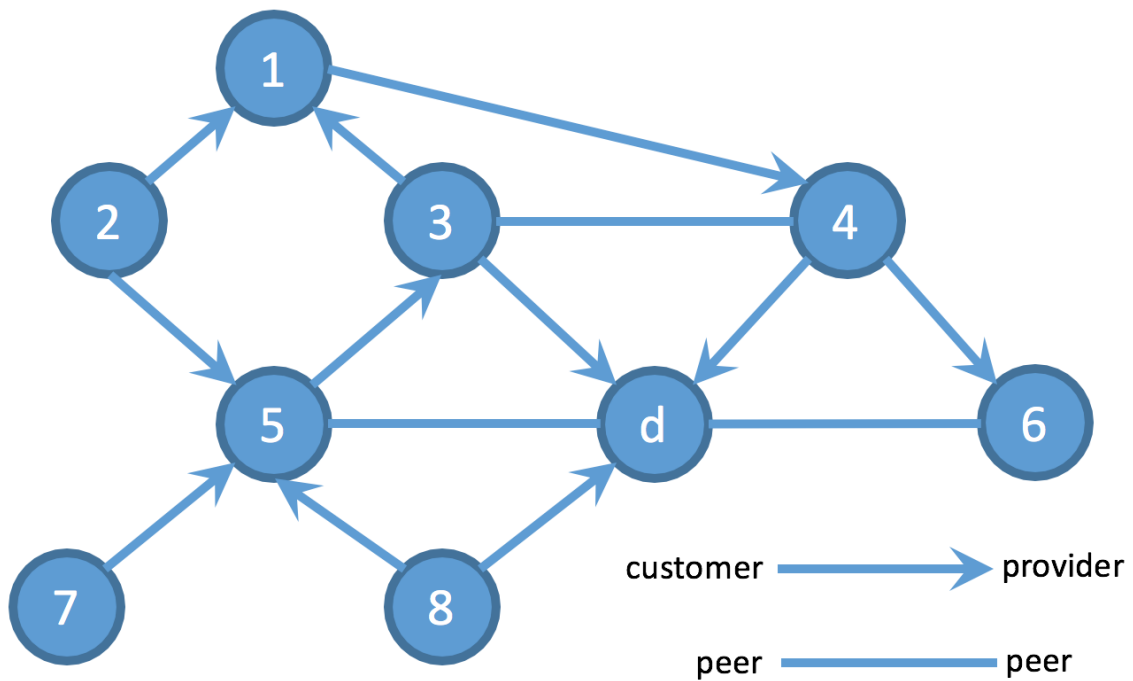


Figure 3. Network topology for Problem 6