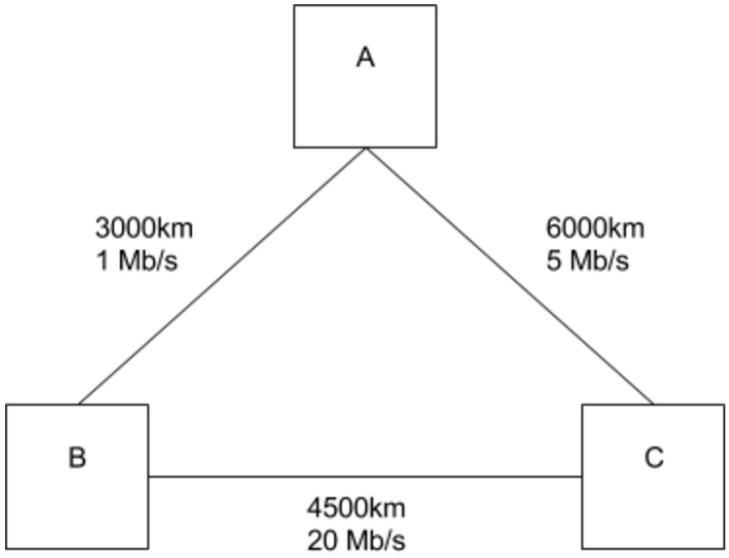
15-440/15-640: Homework 1

Due: September 25, 2017 11:59pm

Name:
Andrew ID:

1 Network Communication

For this section, assume the speed of a signal in a wire is $2 * 10^8$m/s. Use the following diagram for questions 1, 2, and 3.



1. A, B, and C, are nodes attempting to accomplish some task in real time (e.g. playing a video game). The nodes elect one to be the 'host,' and all other nodes connect to the host to send data of an insignificant size. Which of the three nodes should be the host, and explain what factors you're considering. Show all work. **5 points.**

2. Now let's say A, B, and C are nodes involved in some very computationally expensive distributed computing work. They each computed some partial computation, but now need to combine their results in one node. Each node must send all their work to a single node, and you need all the data on hand in order to begin processing it all. A has 1 gigabyte of information, B has 5 gigabytes, and C has 3 gigabytes. Assume the network has no faults, which node would you pick to be the server in order to minimize the total time of combining all this information? Show all work. **10 points.**

3. Assume you can pick which node generates the 1, 3, and 5 gigabytes of information, and you can pick which node becomes the server. What combination will you pick and why? Show all work. **10 points.**

4. (For the following question, ignore the previous diagram.) Consider two nodes A and B in a network, where the bandwidth of network path from A to B is 5KB/s and propagation time is 120ms. On the reverse path i.e. network path from B to A, the bandwidth is 10Kb/s and propagation time is 80ms. Data packets are of size 500 bytes and acknowledgment packets are

of size 100 bytes. Based on this information answer the following questions. Please note that each of the following questions are independent of each other.

(a) What is the throughput that can be achieved when A is transmitting to B using Stop-and-Wait protocol? **5 points.**

(b) Suppose now A is using a sliding window for transmitting packets. Determine the size of the sliding window in terms of packets that A should use to transfer data as fast as possible. **5 points.**

2 Consistency + Classic Synchronization

1. For reference:

   `https://en.wikipedia.org/wiki/Dining_philosophers_problem`

   The dining philosophers have found a potential solution to avoid deadlock. Before eating, each philosopher would flip a coin to decide whether he should pick up the left fork or the right fork first. If the second fork is taken, the philosopher would put the first fork down, and flip the coin again. Do you think that this solution is deadlock-free? Does it guarantee that the philosopher's won't starve? Explain your answers. **5 points.**

2. What are the differences between Mutexes and Semaphores? Explain scenarios where you would prefer using one over the other. One scenario for each situation should suffice. **5 points.**

3 Remote Procedure Calls

1. What would be some of the benefits of having RPC calls where both the server and client stubs are running on the same hardware? **5 points.**

2. There are some drawbacks to using RPC on the same hardware - describe one scenario where it would not be preferable. **5 points.**

3. In an RPC system, programmers may need to pass complex objects or structures to the remote procedures. These complex objects may contain pointer based structures such as reference to a tree node or a linked list. Discuss the implications and approaches of passing this data by value and by reference. **5 points.**

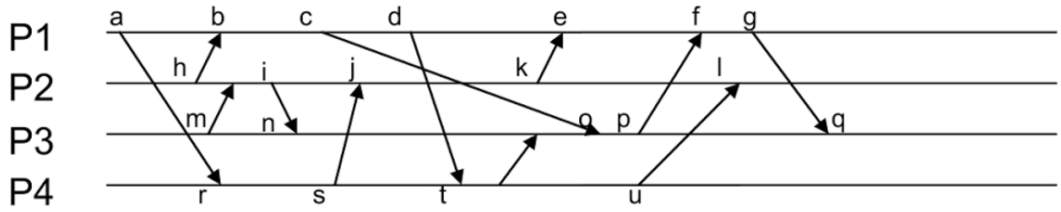
4 Distributed File Systems

1. In a few sentences or less, describe the CAP theorem and what its implications are for distributed file system design. **5 points.**

2. In this spectrum of three qualities with trade-offs, where does AFS lie? **5 points.**

3. If you were to design a DFS for a global bank with many branches across the world to store critical financial records, which DFS qualities would you try to optimize for and why? **5 points.**

5 Time & Synchronization



1. There are 4 processes running in a distributed system, as shown in the figure. Labels a, b, c... indicate the events running on these processes while receiving and sending images. Using this diagram as reference answer the following questions:

   (a) For each labelled event list the Lamport timestamps associated with each of them. Assume that the initial logical clock is set to 0 in each process, and the timestamp is maintained as a single integer value. **10 points.**

   (b) Repeat part (a) with Vector clock timestamps. A table such as the one below would be fine for parts (a) and (b). **10 points.**

Event	Lamport TS	Vector TS

2. A clock at the client side reads 8:20:00. The server's clock reads 8:10:00. What would be the time at the client, if they use the following techniques for synchronization:

   Christian's Algorithm.
   Berkeley's Algorithm.

   Assume message delays are negligible. Show all your work. **5 points.**