

15-440/15-640: Homework 2

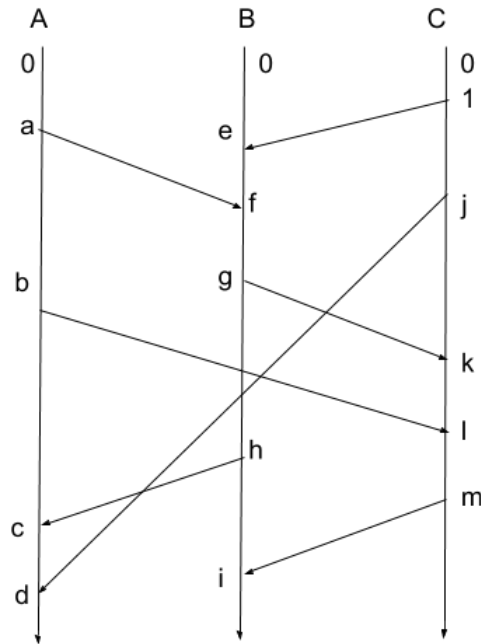
Due: October 15, 2015 10:30am (**NO LATE DAY**)

Name: _____

Andrew ID: _____

1 Time and Synchronization

1. Consider the behavior of two machines in a distributed system. Both have clocks that are supposed to tick 1000 times per millisecond. One of them actually does, but the other ticks only 990 times per millisecond. If UTC updates come in once a minute, what is the maximum clock skew that will occur?
2. Each vertical line represents a process and each arrow represents a message sent from one process to another. Given the start times, fill in the corresponding Lamport times for each letter “a” - “m”.



3. Consider the amazing progress we’ve made in data communication networks over the years. Will it –ever– be possible to sync physical clocks rapidly enough to use as the basis for correct synchronization of a global distributed system? Why or why not?

2 Distributed Mutual Exclusion

1. In class we discussed enforcing mutual exclusion, among other ways, via a central server, majority voting, and token ring.
 - (a) Which of these systems requires the fewest messages under heavy contention? How many messages are required per request?
 - (b) Which of these systems requires the most messages under heavy contention? Why?
 - (c) Which of these systems is most robust to failure? Why?
 - (d) The voting protocol, and the voting district protocol, are based upon participants reaching an agreement as to who can enter the critical section. What happens in the event of a tie that could otherwise risk deadlock?
2. Consider three processes. The system has totally ordered clocks by breaking ties by process ID. It uses the Ricard & Agrawala algorithm. The timestamp for each process of id i is $T(p) = 10 * L(p) + i$, where $L(p)$ is a regular Lamport clock.

Each message takes 2 “real-time” steps to get delivered. Critical section takes 2 real-time steps. Fill in the table with the messages that are being broadcast, sent, or received between the processes until all nodes have executed their critical sections. Write “execute critical section” as the action for a node when it enters its critical section. The first three rows have been filled in for you, and the fourth row has been started. Assume that if a process receives messages from the other two processes at the same time, the message that comes from the lower process ID will be received first.

Action Types: Broadcast (B), Receive (R), Send (S), Execute Critical Section (ExCS) Initial timestamps: $P1 \rightarrow 111$, $P2 \rightarrow 202$ and $P3 \rightarrow 123$

Real Time	Process	Lamport Time	Action(to/from)	Contents	Q at P1	Q at P2	Q at P3
1	2	212	B	(request 212)		212	133
	3	133	B	(request 133)			
2	1	121	B	(request 121)	121	212	133
3	1	221	R from 2	(request 212)	121	133	133
	1	231	R from 3	(request 133)	133	212	212
	2	222	R from 3	(request 133)	212		
	3	223	R from 2	(request 212)			
4	2	232	S to 3	(reply 133)	121	121	121
					133	212	133
					212		212

3 Concurrency Control

1. Is 2-phase commit blocking or non-blocking? What about 3-phase commit? Elaborate on your answers by explaining the situations where a transaction blocks or why it doesn't block.

4 Logging and Crash Recovery

1. Sam developed a database system that uses write-ahead logging and writeback caching to improve performance (Writeback caching allows writes to be reordered). What is the most significant concern that you would expect Sam to have?
2. Describe the pros and cons about Shadow Paging and Write-ahead logging.

5 Fault Tolerance - Detecting and Correcting Local Faults

1. What is the mean time to data loss for a system with 100,000 disk, which are organized into 10,000 10-disk arrays, using data striping and striped parity (i.e., RAID 5)? Assume each disk has an MTBF of 100 years.
2. **(OPTIONAL)** Imagine that you work for a large Internet services company that has 100,000 disks in its data center. If three replicas of any given file block are stored on a random three of the disks, and each disk has an MTBF of 100 years, how many data loss events would you tell your boss to expect in a one-year period (assume no rebuild)?

6 RAID

1. **(OPTIONAL)** Imagine a 5-disk disk array subsystem configured for RAID-5 with a 16 KB stripe unit size. If you were creating a log-structured file system on it, with a file system block size of 4 KB, what would be a good segment size to use? Justify your answer.