Exam in Distributed Systems

Justin Pearson

March., 2012

Cover Sheet

This sheet should be handed in together with the exam.

Each problem must be solved on a separate sheet. Write your name on each sheet. Indicate below which questions you have answered.

Problem no.	Solution provided	Max	Your points
1		12	
2		6	
3		22	
4		8	
5		10	
6		8	
	Total:	58	

Anonymous Exam Code . : .....

Exam Rubric

A mark of 50% is required for a pass, a 4 and 5 are distributed evenly. Answers can be in English or Swedish.

Each full answer should be started on a separate sheet. Please write you name and personal number on each sheet.

I will not be able to come to the exam. If you are unsure about how to answer a question. Make some reasonable assumptions, state your assumptions and answer the exam.

Hjälpmedel:

Pen, pencil, ruler, rubber (eraser for people from the Antipodes or from across the pond) calculator.

1. General Questions on Distributed Systems

   (a) Explain the differences between a distributed system and a multiprocessor system. (**2 points)**

   (b) Fault tolerance is an important property of a distributed system. Over a standard network it is hard to implement a fault tolerant system. Explain, with examples, some of the problems. (**4 points)**

   (c) When implementing a distributed system middleware is often used. Explain, with examples, some of the services that middleware should provide to the programmer. (**6 points)**

2. Remote Procedure Calls (RPCs).

   (a) Explain the client server paradigm. (**2 points)**

   (b) RPCs attempt give the programmer the illusion remote procedures are running locally. Is this really true? (**2 points)**

   (c) When would be be appropriate to move a remote procedure call to a local procedure call?(**2 points)**.

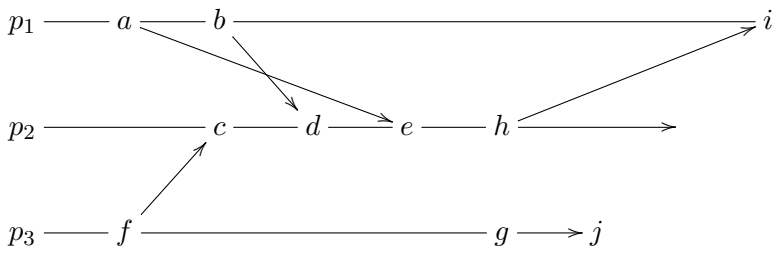
3. Clock Synchronization and Timestamps

   (a) Explain the difference between UTC and TAI. (**2 points)**

   (b) Suppose that you have two clocks both with drift $\rho = 10^{-3} seconds/seconds$. Suppose that you want the clocks synchronized within 0.001 of a second. How often should the clocks be resynchronized to achieve this? (**2 points)** (Obs. Show your workings, no workings no points for a correct or incorrect answer)

   (c) Suppose you have two clocks where the first clock has a drift $\rho_1 = 10^{-6}$ and the second clock has a drift of $\rho_2 = 10^{-4}$. Derive a formula that tells you how often the clocks should be synchronized. (**6 points)**

   (d) Consider the following three processes p_1,p_2 and p_3 with the following pattern of communication:



   Label the each event with a normal Lamport timestamp (**1 points)** and a vector timestamp (**1 points)**

(e) What does it mean for two events to be concurrent and what is the relation of the Lamport timestamps of the two events. (**1 points)**.

(f) Is it possible for two events to have the same Lamport (non vector) timestamp? If it possible give an example, if it is not possible argue why it not possible. (**1 points)**.

(g) A number of processes want to enter a critical section in a distributed fashion, without a central server. Sketch an algorithm that uses Lamport timestamps to decide which processor should enter the critical section first. (**8 points)**. To get full marks you should argue that your algorithm is correct.

4. Cuts

(a) Define a consistent cut in a distributed system (you may use the notation used in the slides or the notation used in the book). If you are not able to define it formally you will still get some points if you define it informally. (**2 points)**

(b) Motivate the above definition: that is, explain why it is useful. (**2 points)**

(c) Explain, with examples, why it hard to record the state of a distributed system. (**4 points)**

5. Distributed Shared Memory.

(a) What is the rôle of memory consistency models in distributed shared memory? (**2 points)**

(b) Define Sequential Consistency. (**2 points)**

(c) Sketch how a system might implement a sequentially consistent memory store. (**2 points)**

(d) Why might a programmer prefer a different memory model from sequential consistency? (**4 points)**

6. Peer to Peer

   (a) Explain the flooding algorithm used in some versions of Gnutella. Explain its advantages and its disadvantages. (**4 points)**

   (b) One of the aims of using a peer-to-peer system is to achieve some sort of load balancing, explain how the Bit-Torrent protocol achieves this. (**4 points)**.