UNIVERSITY OF EDINBURGH

COLLEGE OF SCIENCE AND ENGINEERING

SCHOOL OF INFORMATICS

INFR11022 DISTRIBUTED SYSTEMS (LEVEL 11)

Monday 12 th May 2014

14:30 to 16:30

INSTRUCTIONS TO CANDIDATES

Answer QUESTION 1 and ONE other question.

Question 1 is COMPULSORY.

All questions carry equal weight.

CALCULATORS MAY NOT BE USED IN THIS EXAMINATION

Year 4 Courses

Convener: I. Stark
External Examiners: A. Cohn, T. Field

THIS EXAMINATION WILL BE MARKED ANONYMOUSLY

1. **You MUST answer this question.**

   (a) Alice is trying to synchronise her clock with Bob's. At exactly 3:01pm, Alice sends Bob a text message asking "what time is it?" Bob replies immediately, and at exactly 3:05pm, Alice receives Bob's reply: "my clock says 3:10".

      i. Assuming that the two messages took the same amount of time, how should Alice set her clock to match Bob's clock? [*3 marks*]

      ii. If we do not have the assumption of the two messages taking equal time, what is the maximum possible difference between Alice and Bob's clocks after Alice sets it to that same value? [*3 marks*]

   (b) Recall the Chandy-Lamport algorithm for computing a global snapshot.

      i. The algorithm is useful for monitoring *stable* predicates. What is the definition of a stable predicate? [*2 marks*]

      ii. The algorithm assumes that messages are not dropped, not duplicated, and sent in first-in-first-out (FIFO) order. Give an example showing what could go wrong if the messages are reordered. [*5 marks*]

   (c) In a connected network of n nodes and m edges how many edges does a spanning tree have? How many edges does a BFS tree have? [*2 marks*]

   (d) Suppose the graph of a network is an unweighted tree.

      i. Describe a distributed algorithm that finds the largest distance of any node from a given root node r. The root must know this value when the execution ends. [*5 marks*]

      ii. Is the distance computed in part (i) above the same as the diameter of the tree? Explain your answer. [*2 marks*]

   (e) Describe any two uses of virtualization in distributed computing. Explain what advantage virtualization provides in each case. [*3 marks*]

2. (a) Consider three processes p_1, p_2, p_3 constituting a small distributed system. Each process maintains an integer state variable n_i whose value changes unpredictably. We wish to detect the property $P(n_1, n_2, n_3) = (n_1 = n_2 \text{ and } n_2 = n_3)$ using the Marzullo-Neiger distributed debugging algorithm. An external process is chosen as a coordinator and receives the following messages from the other processes:

p_1	p_2	p_3
101,6	012,5	001,3
201,7	022,6	002,4
301,8	232,7	023,5
	242,8	024,6

The column headings are the senders of the messages and each message is of the form V_i, s_i where V_i is p_i's vector clock timestamp of the event and s_i is the current value of s_i at p_i. For example, the message $101, 6$ from p_1 indicates that s_i had value 6 at the event timestamped 101.

i. Taking into account the sources and vector clock values of the above messages, reconstruct the event diagram showing each event on a process's timeline and the communications between different events. [*6 marks*]

ii. Does P described above *possibly* hold in the above system? That is, is it ever possible that $s_1 = s_2 = s_3$? If so, describe a *consistent* cut (using the vector clock timestamps) such that P holds. [*4 marks*]

iii. Does P *definitely* hold in the above system? Justify your answer. [*4 marks*]

(b) Suppose you are building a distributed system. Describe at least two criteria that you will use to decide if it should be designed as a peer to peer system. Explain how each of these affect your decision. [*2 marks*]

(c) Recall that Gnutella is a peer to peer file sharing protocol.

i. How does Gnutella solve the problem of finding content? [*1 mark*]

ii. Suppose a Gnutella network has n nodes and each node has at most a fixed constant h number of neighbors. What is the asymptotic communication complexity of a file search? (assume that a message from a node to its neighbor in the Gnutella network has a constant cost c.) [*3 marks*]

iii. What is Gnutella's approach to verifying content? (That is, how does Gnutella check that the data being supplied is what the user has requested?) [*1 mark*]

(d) Deduce the worst case message complexity (in big-oh notation) of the Bully leader election algorithm. Assume that all nodes are active, and none fails during the algorithm operation. (*Hint: the worst case occurs when all nodes choose to initiate election simultaneously.*) [*4 marks*]

3. (a) Sometimes in a distributed system, it is useful to elect a process as a "leader" for the distributed computation. Give two examples of computations or scenarios where having a leader is useful. [*3 marks*]

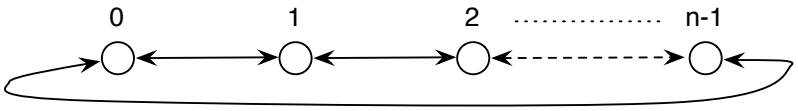

Figure 1: Leader election ring. The numbers next to the processes show their ids.

(b) Consider the ring of n processes shown in Figure 1. Every process has two pointers (or links) called *left* and *right* to its neighbours in the ring. The links at any process naturally point to the processes on its left and right, except at process 0 and $n-1$ where they loop to the other end of the ring as shown in the figure.

Recall that the basic (Chang and Robert's) ring based leader election algorithm sends messages in only one direction and elects the process with the highest id as the leader.

i. Suppose all the processes initiate leader election simultaneously. If every process always sends its message to the process on its *left* link, what will be the asymptotic message complexity of the algorithm (in big-oh notation) on the given ring? (Justify your answer) [*4 marks*]

ii. What will be the message complexity if the processes always send messages to the process on their *right* links? [*3 marks*]

(c) In Bittorrent, a client prefers to download from fast peers, since fast peers usually provide faster downloads. Suppose we modify the protocol so that only peers that are sufficiently fast host the file. What are the advantages and disadvantages of the modified protocol? [*4 marks*]

QUESTION CONTINUES ON NEXT PAGE

(d) Recall the Maekawa voting algorithm for ensuring mutual exclusion. Consider four processes p_1, p_2, p_3, p_4 with the following voting sets:

$$\begin{aligned} V_1 &= \{p_4, p_1, p_2\} \\ V_2 &= \{p_1, p_2, p_3\} \\ V_3 &= \{p_2, p_3, p_4\} \\ V_4 &= \{p_3, p_4, p_1\} \end{aligned}$$

That is, the four processes are arranged in a square, and each process's voting set consists of itself and its two neighbours.

i. Give a possible run of this system resulting in deadlock, i.e., such that each process is blocked waiting for access to the critical section. [*5 marks*]

ii. Give a possible run of the system that violates fairness, i.e., such that some process p_i makes a request before p_j but p_j is granted access before p_i. [*6 marks*]