CSSE4004/CSSE7014 Distributed Computing
First Semester Examination, June 2006

Question 1 (**8 marks**)

Address the following issues about distributed systems:

(a) Explain why middleware based distributed systems are built above network operating systems, not distributed operating systems. (4 marks)

(b) A process in a distributed system runs on one node and accesses data from another node. After some time, for load balancing purposes, this process relocates to a different node. What kind of transparencies should be provided for this process in a distributed system? (4 marks)

Question 2 (**9 marks**)

Address the following issues about communication in distributed systems:

(a) Explain how dynamic method invocation works. (3 marks)

(b) Do messaging systems such as IBM MQSeries provide transient and synchronous communication? Explain your answer. (2 marks)

(c) Routing tables in current application level messaging systems are configured manually. Describe a protocol which would provide automatic configuration of the routing tables. (4 marks)

Question 3 (**15 marks**)

Answer the following questions related to naming in distributed systems:

(a) Explain why in mobile computing it is better to map names to unique resource identifiers not to addresses. (2 marks)

(b) Is the name "www.uq.edu.au/index.html" location independent? Explain what a "location independent" name is. (2 marks)

(c) In a hierarchical location service with a depth of *n*, how many location records need to be updated at most when a mobile entity changes its location? (3 marks)

(d) An object should be garbage collected if there is no process which has a reference to it. Let's assume that there is a counter in the object which counts the number of processes which have its reference. Processes have to send a message to the object if they acquire its reference from another process – each message increases the counter by 1. They also have to send a message if they are not going to use the reference in the future – this decreases the counter by 1. Is it possible that this object may be garbage collected even if there are processes which have the object's reference and are going to invoke methods on it? Explain. (3 marks)

(e) Briefly describe the concept of intentional naming. (5 marks)

Question 4 (**15 marks**)

Address the following issues related to synchronization in distributed systems:

(a) Give two examples of distributed applications which require total (logical) ordering of events. Explain why. (4 marks)

(b) Diagram 1 illustrates local times in three different nodes and also communication between three processes running on these nodes. Introduce total order to the communication events by applying Lamport's "happened before" relationship. Redraw the diagram to show the changes. (6 marks)

(c) Explain the meaning of concurrency control for transactions. (2 marks)

(d) Several processes access data in a database and most of these accesses are for reading, i.e. writes are quite rare. If you had a choice between optimistic concurrency control and pessimistic concurrency control which one would you choose in this case? Explain why. (3 marks)

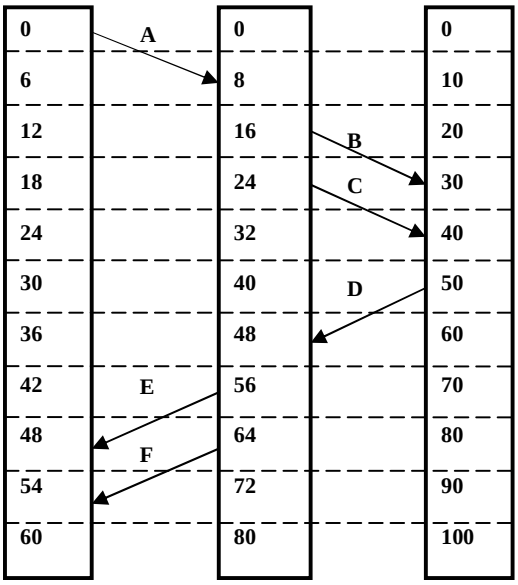

Diagram 1

Question 5 (**14 marks)**

Address the following issues related to the distributed computing platforms:

(a) Explain why Interface Definition Languages are needed in distributed computing platforms. (3 marks)

(b) Explain the sequence of action taken in CORBA if a client uses a persistent CORBA object reference to invoke a method when the server responsible for this object is not running. (5 marks)

(c) Compare .NET and Web Services. (6 marks)

Question 6 (**9 marks**)

The following questions relate to fault tolerance:

(a) Give two examples of *forward recovery*. (2 marks)

(b) There is a group of five replicated processes producing some results. Assume that these processes can only crash silently; what is the maximum number of these processes that may crash while it is still possible to get the correct results? What is the maximum number of Byzantine processes in this group to achieve the correct results? Explain. (4 marks)

(c) Assume that a group of processes has to reach an agreement on electing a coordinator. There are *2* processes in the group which exhibit Byzantine failures. How many processes have to be in this group to reach the agreement? (3 marks)

Question 7 (**8 marks**)

The following questions relate to processes:

(a) Discuss the advantages/disadvantages of a multi-threaded server compared to a server which spawns multiple processes. (4 marks)

(b) We want to compare the performance of a single-threaded and a multi-threaded file server. The following assumptions are made. It takes 15ms to get a request, dispatch it and do the rest of the necessary processing involved in serving the file, assuming the file is cached in main memory. If the file is not cached, a disk operation is needed in which case an additional 60 ms is required, during which the thread sleeps. We assume that for one third of all requests, the file can be served from the cache.

   a. How many requests per second can the single-threaded server handle? (2 marks)

   b. How many requests per second can the multi-threaded server handle? Assume that there is always enough threads. (2 marks)

Question 8 (**12 marks**)

The following questions relate to consistency and replication:

(a) A consistency model uses a small number of locks. Each lock protects updates to a different variable.
   a. What kind of consistency model is this? Explain. (3 marks)
   b. Describe pros and cons for protecting each variable separately. (3 marks)

(b) Which of the following are valid combinations of read and write quorums for Gifford's scheme (quorum-based protocols)? Assume $N = 9$. Explain each answer.

$N_R = 2$, $N_w = 7$ (2 marks)
$N_R = 1$, $N_w = 9$ (2 marks)
$N_R = 3$, $N_w = 7$ (2 marks)

Question 9 (**10 marks**)

The following question relates to distributed file systems and pervasive systems:

(a) Discuss in 3-4 sentences the fault tolerance of Coda. (4 marks)

(b) Context information is data information which is read and evaluated to make adaptation decisions. Why is this data treated differently than any other data applications read, i.e. how does it differ from typical data information. (6 marks)