

THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM.
INTERNAL STUDENTS ONLY.

STUDENT'S NAME: _____

STUDENT NUMBER: _____

THE UNIVERSITY OF QUEENSLAND
School of Information Technology and Electrical Engineering
First Semester Degree Examination, June 2005

CSSE4004/CSSE7014

DISTRIBUTED COMPUTING

TWO (2) hours for working
TEN (10) minutes for perusal before examination begins

ANSWER ALL QUESTIONS
THERE ARE **100 MARKS** FOR THE WHOLE PAPER.
QUESTIONS CARRY THE NUMBER OF MARKS INDICATED.

CLOSED BOOK EXAMINATION
NO CALCULATORS PERMITTED

Question 1 (8 marks)

Address the following issues about distributed systems:

- (a) Describe the differences between network operating systems, distributed operating systems and middleware based distributed systems. (4 marks)
- (b) Explain the meaning of access, relocation, concurrency, and failure transparencies in distributed systems. (4 marks)

Question 2 (10 marks)

Address the following issues about communication in distributed systems:

- (a) Explain how dynamic method invocation works. (3 marks)
- (b) Does Remote Procedure Call provide persistent and synchronous communication? Explain your answer. (3 marks)
- (c) Many applications like workflows or operations on collection of databases require asynchronous messaging systems such as IBM MQSeries. Routing tables in IBM MQSeries are configured manually. Describe a simple way to do this automatically. (4 marks)

Question 3 (13 marks)

Answer the following questions related to naming in distributed systems:

- (a) Explain the roles of names, identifiers and addresses. (3 marks)
- (b) What are the similarities and differences between the X.500 naming model and the Internet Domain Name System? (4 marks)
- (c) Is the name “www.itee.uq.edu.au/index.html” location independent? Explain what a “location independent” name is. (2 marks)
- (d) The reference counting method used to check whether an object can be deleted may suffer from race conditions whereas the weighted reference counting method cannot. Explain why. (4 marks)

Question 4 (11 marks)

Address the following issues related to synchronization in distributed systems:

- (a) Explain the meaning of concurrency control for transactions. (3 marks)
- (b) When would you prefer optimistic concurrency control over pessimistic concurrency control? (4 marks)
- (c) Show how totally ordered multicast can be achieved using Lamport timestamps. (4 marks)

Question 5 (14 marks)

Address the following issues related to the distributed computing platforms:

- (a) Describe the role of Interface Definition Languages in distributed computing platforms. (3 marks)
- (b) Outline the roles of the CORBA implementation repository and the POA, in allowing an invocation on a persistent CORBA Object Reference (IOR) to an object for which no running server exists. (5 marks)
- (c) Describe the role of HTTP, SOAP and WSDL in Web Services. (6 marks)

Question 6 (13 marks)

The following questions relate to fault tolerance:

- (a) Explain the multicast related problem called *Virtual Synchrony*. (4 marks)
- (b) Is the recovery of faults in RAID (Redundant Array of Independent Disks) disks considered to be *backward recovery* or *forward recovery*? Explain why. (3 marks)
- (c) In a group of replicated processes there are k processes which exhibit Byzantine failures. How many processes have to be in this group to produce a correct result? Explain. (3 marks)
- (d) Assume that a group of processes has to reach an agreement on electing a coordinator. There are k 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 75ms is required, during which the thread sleeps. We assume that for two thirds of all requests, the file can be served from the cache.
 - i. How many requests per second can the single-threaded server handle? (2 marks)
 - ii. How many requests per second can the multi-threaded server handle? Assume that there is always enough threads. (2 marks)

Question 8 (14 marks)

The following questions relate to consistency and replication:

- (a) Explain what it means that a particular consistency method (e.g. *weak* or *release* or *entry* consistency) uses synchronization variables, i.e. explain main differences between this class of consistency models and the class of models which do not use synchronization variables. (4 marks)
- (b) A consistency model uses a small number of locks. Each lock protects updates to a different variable.
 - i. Is this *release* or *entry* consistency? (2 marks)
 - ii. Describe pros and cons for protecting each variable separately. (2 marks)
- (c) Explain Gifford's scheme (quorum-based protocols). (4 marks)
- (d) Which of the following are valid combinations of read and write quorums for $N = 8$?
 - i. $N_R = 2, N_w = 6$ (1 mark)
 - ii. $N_R = 1, N_w = 8$ (1 mark)

Question 9 (9 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) Why is metadata describing quality and uncertainty of information sensed from logical and physical sensors needed in context-aware applications? (5 marks)