UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2001/2002

September 2001

CCS521 – Advanced Distributed Systems Concepts and Design

Duration : 3 hours

INSTRUCTION TO CANDIDATE:

- Please ensure that this examination paper contains **FOUR** questions in **FIVE** printed pages before you start the examination.
- Answer **ALL** questions.
- You can choose to answer either in Bahasa Malaysia or English.

ENGLISH VERSION OF THE QUESTION PAPER

1. (a) Assuming that you have just been appointed as a consultant at a distributed computing company called DCorp. Your boss has requested you to give some advice to the clients. Based on your knowledge gained from CCS521, what kind of comments and advice would you provide to your clients? Below are their questions/concerns:

(i) Client A

"I am working as a Senior Technical Officer at one of the telecommunication companies. I have few basic questions pertaining to distributed computing softwares.

- *There are two popular middlewares that one can use in distributed computing, CORBA and MPI. I only know the acronyms. What do they actually stand for?*
- *Please explain to me the programming paradigm for each and tell me when I should use which.*
- *What are the advantages and disadvantages of each middleware?"*

(ii) Client B

"We are from the Meteorology Department. Among our duties are to forecast the daily weather. Can you guide us on the following:

- *The Meteorology Department in developed countries use parallel computing technologies in simulating the weather (the atmosphere, ocean etc). What do you think their reason for adopting this technology?*
- *We have also come across the term "embarrassingly parallel application". Can you tell us what does that mean and give one example application?*
- *There exists a standard message passing library, MPI, to support us in developing our programs. Can you list the services provided by this library?"*

(iii) Client C

"Our company is the largest warehouse chain in this country. Our aim is to double up the profit gained in the next five years. We can only do this if we know the consumer purchase pattern of every item sold each day. This means that we have to keep track of a database of every item, the information about the consumer and the date and time of purchase. Just imagine how huge the database would be in the next 5 years. By that time any query to the database would take hours to complete. Don't forget that these databases are geographically distributed, with each warehouse owning their own database. We want to be able to analyse the purchase pattern of all these warehouses. Our questions are:

- *We heard about a technology called grid computing. What is grid computing?*
- *Is the technology applicable in our situation? Explain your answer.*
- *What are the challenges that we will face if we were to adopt this technology and what do you think would be the future of the technology (e.g. will it last)?"*

(15/25)

(b) Explain the following terms and state a situation when it will be useful:

(i) Virtual topologies

(ii) Derived data type

(6/25)

(c) Explain the following layers and give **one (1)** protocol example:

(i) Application

(ii) Transport

(4/25)

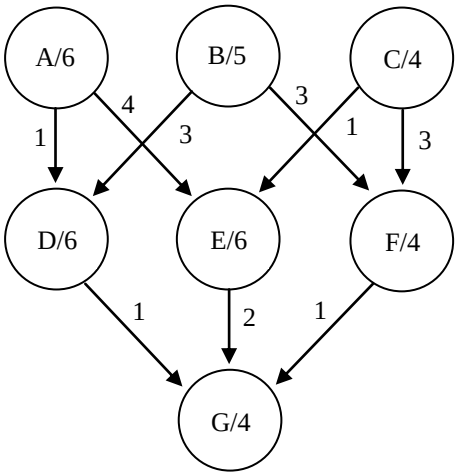
2. (a) Answer the following questions:

(i) Based on your knowledge in doing assignment 1 write a pseudocode which reverse the direction of the ring message transmission. Instead of having process 0 initiates the communication, the last process, process *n*, starts by sending an integer value 2020 to process *n* - 2.

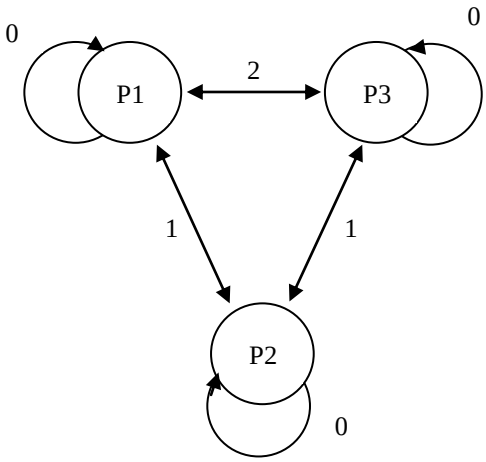
(ii) Discuss a situation in which non-blocking receive is preferred over blocking receive and briefly explain how the performance can be improved.

(10/25)

(b) Given the following Precedence Process Model and Communication System Model:



Precedence process model



Communication system model

Calculate the completion time or the makespan using:

(i) List Scheduling (LS) strategy.

(ii) Extended List Scheduling (ELS) strategy.

(10/25)

(b) Compare and contrast two dynamic load balancing algorithms, namely Diffusion and Dimension Exchange Method.

(5/25)

3. (a) Would you consider Ricart and Agrawala's multicast based mutual exclusion algorithm in a system where each process uses a critical section many times before another process requires it? If so, why and state the benefits expected. If not, why and suggest an improvement.

[6/25]

(b) Compare the Ricart and Agrawala's multicast based mutual exclusion algorithm with Maekawa's algorithm in terms of performance and fault tolerance.

[6/25]

(c) A sensor process writes the current temperature into a variable t stored in a distributed shared memory that implements release consistency. Periodically, a monitor process reads t. Explain the need for synchronization to propagate the updates to t, even though none is otherwise needed at the application level. Which of these processes (sensor or monitor) need to perform synchronization operations?

[7/25]

(d) Referring to 3(c) above, how will the propagation change when eager release is applied? Compare this with lazy release.

[6/25]

4. (a) Given the following situations, which of the services (name, directory or discovery) would be appropriate? Justify your choice.

(i) A user arrives at an airport that she has never visited before carrying a PDA that is capable of wireless networking. She needs information on the local services at that airport.

(ii) A virtual enterprise consists of a collection of companies who are cooperating to carry out a particular project. Each company wishes to provide the others with access to only those objects relevant to the project.

(iii) A user would like to locate a long lost friend whom he knows is studying at a reputable university in United States. Unfortunately, he only remembers his friend's first name.

[6/25]

(b) Briefly describe two transparency features that is hoped to achieve by middleware software like CORBA.

[4/25]

(c) You have been hired to design a directory service for the Internet. Your directory service would advertise Internet services such as information archives, Web servers, and other information providers to potential clients. One of the main goals when designing such a system is scalability. You would base your design in the client-server model, where clients, on behalf of users, would query directory servers.

(i) What kind of information would your directory service provide? How would it be different from DNS?

(ii) To address scalability, how would you structure the system's name space?

(iii) Suggest and explain briefly two techniques that you would use to improve the system's performance.

(15/25)