Uva Wellassa University, Sri Lanka

Faculty of Science and Technology

Computer Sciences and Technology Degree Program

2nd Semester Examination September/October 2012

CST 326-2 Distributed and Parallel Computing

Time Duration : Two Hours (02)
Number of Pages : Three (03)
Index Number :

Instructions to Candidates
Answer **all** questions.
This paper will contain four **Structured Essay questions**.
**Please return question paper with your answer script.

Structured Essay Questions

1. (25 Marks)
   a. Give an example of each transparency type. Use as examples
      i. a distributed file system
      ii. printing service (a group of printers), which are intended for
      iii. mobile-user applications.
   b. Why is it not always a good idea to implement transparency?
   c.
      i. What kinds of failure types can you expect in

a Local Area Network	
UDP transport service	
on the air link in GPRS?	

      ii. How can you detect such failures and how can you mask them? Is it possible that masking process leads to new failures?
   d. Create a small application example where you utilize RPC calls (in a Java environment). Explain the principles of RPC.

2. (25 Marks)

a. Three different parallel architecture models have been presented during lectures which model real-life parallel architectures and applications. Explain how they differ, and what are their strengths and weaknesses. The models are:

	how they differ	strengths	weaknesses
PRAM			
BSP			
PPM			

b. For two N-dimensional vectors A and B, write an algorithm that calculates their inner product on the architecture models of the previous question (2 a). Compare their complexity.

Choose a PRAM machine with n processors. Count the sum S = A(1) + A(2) + ... A(n) when A contains $n = 2^k$ numbers. Make an algorithm for the sum function that can be executed in each processor. Do notice global and local variables!

c. How would you emulate a message passing architecture using a shared memory machine? Let us presume that both machines have p processors. How about vice versa?

d. Compare the following parallel architectures. Describe two unique features of each of them that cannot be found in any of the other architectures.

i. Parallel vector processor (PVP)
ii. Symmetric multiprocessor (SMP)
iii. Massively parallel processor (MPP)
iv. Cluster of workstations (COW)
v. Distributed shared machine (DSM)

3. (25 Marks)

a. Let's say you are modeling the air flow in the room, and you decide to break the room into a 1000 x 1000 x 1000 grid. Thus, you are storing variables (temperature, pressure, etc.) at each grid point. All of your physical quantities are represented at these grid locations If 80% of the total operations occur in parallel sections (the operations on grid values, of course) what is the maximum theoretical speedup on:

   i. 2 processors?
   ii. 4 processors?
   iii. 8 processors?
   iv. What is the maximum theoretical speedup?

b. Consider a sender *S* and a receiver *R* that are connected on a network where the one-way latency is *250 ms* (milliseconds), that the data rate is *32,000 b/s* (bits per second), and that each frame is *1,000 b* (bits). Assume that *S* and *R* are using sliding windows with selective repeat. How large must the window for *S* be in order to maximize utilization? Based on that window size, how many bits must be used in each frame as a frame identifier (that is, a frame number)?

4. (25 Marks)

a. Describe how the distributed name service (DNS) works.

b. You can create a distributed application by, for example, partitioning the data (and storing each partition on a different server) or replicating the data (and storing each replicate on a different server). Both solutions have their advantages and disadvantages. What kinds? The aspects to be considered are availability, reliability (what is "reliability"?), performance, and scalability. Use the internet directory system DNS as an example. The baseline in the evaluation is a traditional centralized system.

c. The performance of a distributed system can be improved using caches and proxies.

   i. How do these techniques affect the performance and scalability of the distributed system?
   ii. Under what circumstances are these techniques useful?
   iii. What problems can you encounter when using these techniques?