THE BCS PROFESSIONAL EXAMINATION
Professional Graduate Diploma

April 2004

EXAMINERS' REPORT

Distributed & Parallel Systems

General

The performance was generally good with an average mark of a shade higher than 62%. The pass rate was above 75%, with a number of students producing good answers. The candidates seemed to be more comfortable with distributed systems than with parallel systems.

Although most candidates had a good grasp of the concepts, their answer formats, particularly the construction of sentences in English, and handwriting, need improvement.

Candidates struggled to gain marks in this year's paper although the number of pages submitted per script remained reasonably high. This year it was particularly noticeable that many candidates were submitting scripts that often contained pages of written work for which no mark could be awarded. This was particularly noticeable in questions 1b, 2b and 2c where there were about 25 pages submitted overall that just did not answer the question posed and as a consequence remained unrewarded.

Question 1

1. In a parallel or distributed system the issue of load balancing is usually crucial to the performance of the system.

a) Explain what is meant by load balancing and differentiate between static and dynamic load balancing. **(10 marks)**

b) Using an example from a parallel or distributed system you have studied, evaluate how replication of processes within a system can enhance performance. **(15 marks)**

Answer Pointers

a) Answer must demonstrate the understanding of i) the concept of load balancing- computation work load equally distributed across all processors (4 marks) ii) the static load balancing in which the distribution of computation work load at compile time (3marks) iii) The dynamic load balancing of the computation work load occurs " on the fly" during run time. (3 marks).

b) Examples such as replicated web servers, processor farm, replicated processes in distributed data bases could be chosen. Answer should show good understanding replicated process- its role (3 marks), its purpose in achieving performance improvement (3 marks), method of work allocation to achieve this, static and dynamic (3 marks) and advantages-including scalability and reliability and disadvantages- costs, data consistency etc.. (6 marks).

[Pfister p 77, 276, 480, Coulouris p 553]

Examiners' Comments

The concept of load balancing was reasonably well presented. While most of those who answered this question got the static load balancing part right, only a few could describe the dynamic load balancing part correctly. Most of them did present good application examples.

Popular with almost all candidates, this question was nevertheless not well done. Not many knew about dynamic load balancing 'on the fly' during runtime; not many had much depth to their knowledge of how replication of processes can enhance performance.

Question 2

2. Threaded programs may be executed on both single processor and multi-processor computers.

a) Explain what is meant by a thread and explain how threads are executed on a single processor. **(8 marks)**

b) Discuss how the performance of a multi-threaded program would be expected to alter when ported to a multi-processor. **(8 marks)**

b) Using a programming language of your choice, design a threaded program that could be used to measure the performance of the program on a range of multi-processor computers, with between 2 and 20 processors.

Explain the design of your program and how the performance will be measured. (**9 marks)**

Answer Pointers

a) Answer should clearly discuss that;

- A thread is a lightweight process within the environment defined by a task, threads of a task run concurrently share the same global data space, but maintaining private copies of their local data. Threads can be dynamically created, suspended, resumed or destroyed. (4 marks).

- Threads can be prioritised, queued and can be swapped in a non-deterministic manner. Usually a physical limit to the number of threads that can be active at any time in a single processor. Threads can be treated as objects capable of performing computations with low overhead and minimal state representations. (4 marks).

b) Discussion: When ported to a multi-processor, opportunities for parallel execution of threads, and hence performance improvement exists subject to communication and synchronisation constraints imposed by resource-sharing in tightly-coupled multi-processors and message-passing in loosely-coupled multi-processors and effectiveness of load balancing. Also advantage such as less need for swapping and time saved in storing local data. Issues such as consistency of global data when copies of it are kept in memories of processors. (8 marks).

c) Discussion: Performance measurement approach – putting the code to be measured inside a large loop, executing it many times, and determining the average execution time for a single iteration. Mapping of program threads on processors, distributing their numbers on each processors -on to a maximum of 20 processors. One issue could be running parts of codes on different processors (4 marks)

Appropriate language could be used for the program design and code presented. There should be clear indication of starting and terminating parts of threads (5 marks).
[Pfister p 314, Coulouris p22, Magee p25]

Examiners' Comments

While the part a) above was answered correctly, incorrect answers were produced for part b). Only a very few could answer part b). Some were able to produce pseudo codes which if refined within a language format would have improved the outcome of their effort.

Part c) proved to be a particularly testing and disappointingly few were able to indicate the design of a program to measure performance. Candidates should aim to be able to think through the rudiments of a design scenario under exam conditions

Question 3

3. Middleware is a term used to define the software technology designed to link otherwise incompatible computers, networks and applications together.

a) Describe a set of criteria for selecting middleware, which can be used generically for choosing the most appropriate software for a particular area. **(10 marks)**

c) Describe an application that would need to be distributed across at least three computers. Indicate which middleware products would be suitable for linking the distributed parts. Use the criteria developed above to justify your choice. **(15 marks)**

Answer Pointers

a) Description of the following set of criteria:

i) Suitability: integration of software/hardware aspects of architectures.
ii) Integration of applications- standards and middleware technology considerations
ii) Reliability and robustness
iii) Transparency
iv) Risks and cost aspects
iv) Strength of product support
v) Security characteristics

(10 marks)

b) Description:

Any application that illustrates the point, for example a three tier application that uses thick (personal computers) and thin (mobile) clients, web and other application servers including mail and data base servers. (7 marks)

The discussion could centre around the architectural aspects of the application –support for integration of services like web services, and for distributed object connection, security reliability transactions etc. with references to the criteria listed in a) above. (8 marks)

[Couloris p 207]

Examiners' Comments
Comments: Part b) seemed to be more popular than part a). Some good examples were produced and discussions in some papers were crisp.

Only two candidates gained more than half marks for the bookwork-based part a), (although it maybe that the term 'middleware' is less frequently encountered than hitherto).

Question 4

4. *a)* Explain what is meant by a concurrent architecture. **(6 marks)**

b) Name three different concurrent architectures, indicate how they may be modelled and describe situations in which the architectures are used. **(9 marks)**

c) Discuss the merits of the different types of abstract connector that co-ordinate communication between components of a concurrent architecture. **(10 marks)**

Answer Pointers

a) Explanation should say how in a concurrent architecture support at the architectural level is provided to exploit concurrency/parallelism in applications. This is contrasted with a conventional uniprocessor architecture where the architecture supports sequential processing. (6 marks)

b) Architectures could include tightly and loosely coupled architectures in Flynn's SIMD and MIMD examples;: Vector processors- CRAY machines, DAP s, shared memory architecture (tightly coupled multi-processors)- Silicon Graphics work stations, distributed memory architectures – networks. A number of situations can be described as examples in each. SIMD architecures for data intensive processing situations like weather forecasting, MIMD shared memory architectures for real-time rendering tasks like in medical imaging and MIMD distributed memory architectures for client-server type applications – internet –based applications like on-line banking, e-commerce etc..

Discussion on SIMD and MIMD models with fine-grained and coarse-grained Concurrencies. (9 marks)

c) Communication abstraction examples: semaphores and monitors and mailboxes for MIMD shared memory architectures and ports for MIMD distributed memory architectures provided at the operating system's kernel level. Discussion on semaphores to implement simple protocols for accessing critical resources and mutual exclusion of critical accesses, monitors as data structures for non-deterministic and deadlock-free controlled accesses of shared resources, ports for implementing send and receive communication primitives for message-passing etc..

(10 marks)

[Almasi and Gottlieb and Ben-Ari]

Examiners' Comments
Parts a) and b) were almost correctly answered. None got it right in their discussions on part c). Indeed the students got very confused, and produced a variety of wrong answers.

It appears that not many had read the latest reference text (Magee & Kramer, Ch 11) which provided the inspiration for this question. Knowledge of concurrent programming models remains an important element of the syllabus content.

Question 5

5. You have agreed to talk for 30 minutes at the next meeting of your local BCS branch. The title of your talk is "Engineering Software for Parallel and Distributed Systems".

Sketch out approximately eight overhead slides, with associated notes, that you would use for your talk.

(25 marks)

Up to eight slides can be prepared for demonstration. There can be a number of approaches. For example,the first slide could introduce the concepts of parallel and distribution systems. The slides that follow could b e on parallel programming environment-languages, operating systems, concurrent tasks, coordination-synchronisation and interprocess communication and related issues and performance measurement etc.and conclusion.

(3 marks for each slide with a mark for overall marks for topic selection, conciseness etc..)

Examiners' Comments

Most of the students concentrated on only the hardware architectures and very few were able to present slides on software. Some answers were in a descriptive format with no bullets anywhere. This question seemed to be popular with the students

Candidates anticipate the appearance of this type of question and it is generally the best answered. However some candidates spend too much time answering this question for too little reward.