THE BCS PROFESSIONAL EXAMINATION
Professional Graduate Diploma

April 2001

EXAMINERS' REPORT

Distributed and Parallel Systems

This paper had a 100% pass rate although the numbers attempting were low. The questions attempted were evenly distributed over the five questions. The examiners are satisfied that whilst the questions covered some different topics to those of previous years, candidates were obviously well prepared to discuss topics from all parts of the syllabus. The answer points for each question are as follows:

QUESTION ONE

a) Describe the alternative methods by which security violations may be perpetrated in distributed systems (8 marks)

b) Explain how public-key encryption can provide secure communication (8 marks)

c) A public examination requires that examiners annually prepare a set of papers together with specimen answers and marking scheme. Discuss safeguards that would ensure Internet transmission of examination materials between examiners and a central office (7 marks)

Answer Pointers

a) Methods of attack include eavesdropping, masquerading, message tampering, replaying and infiltration (via virus, worm or Trojan horse). All depend upon obtaining access or establishing a pseudo connection to a principal with authority.

b) Public-key encryption is based upon the use of the product of two large primes, relying on their computational intenseness for security. The RSA algorithm may be described or current practice through PEM or PGP.

c) Safeguards include password protection (using either a 'Zipped' RTF file or a Word file), the password having to be communicated verbally. The candidate may describe the steps involved to achieve the safeguard and/or general guidelines about choice of passwords, file names, or email subject headers.

QUESTION TWO

a) How is a 'cluster' of computers distinct from:

1. **A distributed system (7 marks)**
2. **A parallel system (6 marks)**

b) Given that a single system image is the illusion that a collection of computing elements is a single resource, discuss how well the concept of a

'cluster' fits the notion of a single system image. Make reference to application and sub-system levels as appropriate (12 marks)

Answer Pointers

a) In summary, a cluster consists of a collection of interconnected 'whole' computers and is used as a single, unified computing resource distinct from a distributed system by its internal anonymity, and peer relationships and distinct from a parallel system by its high availability, scaling and system management. Clusters are not so much different but more akin to a subparadigm of distributed or parallel systems.

b) Given that every SSL can be viewed as having boundaries with support at different levels, discussion of some of these boundaries (possibly using application, subsystem, OS kernal and hardware boundaries) and corresponding levels of abstraction would be appropriate.

QUESTION THREE

Sorting is one of the most common activities performed on serial computers.

a) Specify either formally or pragmatically a sorting algorithm that would be appropriate for a parallel or distributed architecture. Provide a detailed explanation of the specification technique used. In addition outline the target architecture (13 marks)

b) Compare the theoretical performance of the specified algorithm on the target architecture and on a conventional architecture. Comment on the validity of the comparison (12 marks)

Answer Pointers

a) There are many possible parallel sorting algorithms. Quinn, 1994, describes 10 different variations. There are also many specification techniques that can be used. Pseudo code is acceptable, so are formal techniques such as CSP, CCS etc.

b) Performance may be presented as big O notation or by counting operations and their times or any comparable method. The parallel performance will be the cost of the operations for the sort on the individual parts and the communications necessary and the parts that have to be done sequentially. Expect a formula something like:

P(N) + C(N) + S for N processors

Where P is the time things can be done in parallel. C is the communication time and S is the time that has to be sequential.

The sequential architecture won't have any communications so it will be something like:

P(1) + S

The validity of the comparison is questionable since an algorithm that is well suited to a parallel machine may not be best for a sequential machine. Maybe we should compare a good parallel algorithm with a good sequential one

Marks Breakdown

a)	Suitable algorithm for described target architecture	(5 marks)
	Specification of algorithm	(5 marks)
	Detailed description of specification technique	(5 marks)
b)	The two formulas and comparison	(10 marks)
	Commentary on the validity	(5 marks)

QUESTION FOUR

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

a) Identify a programming language that allows a user to define and use threads. Explain with examples, how threads are defined and co-ordinated in this language (13 marks)

b) Compare the execution of a threaded program on a single processor and on a multi-processor. Illustrate your answer with examples with which you are familiar (12 marks)

Answer Pointers

a) There are many programming languages that support threads. Any language is acceptable. For example Delphi allows the user to use threads. The programmer can define threads based on the TThread object. A thread is an object that runs code. Threads can be given priorities but in general which one is run is non-deterministic.

Lock and Unlock methods can be used to protect critical parts.

Acquire and Release can be used to update globally.

BeginRead, BeginWrite and matching ends give a different degree of protection.

b) In a sequential environment threads share the same global data space, but maintain private copies of their local data. In a parallel version the global data has to be replicated across all processors.

Performance issues may include:

1. More threads can be handled – so a program that could not run on a single processor may be able to run on a multi-processor.

2. There should be less need for swapping, so time may be saved in storing local information.

3. However, it is important that global data is kept consistent across processors. This may be by keeping copies of global data on all processors and ensuring they are all updated as necessary.

4. Alternatively global data can be kept in a single location and fetched from that one place when required.

The first two have a positive effect on performance. The latter two are negative. In addition, the division of work may mean that not all processors are busy all of the time, this idle time will also reduce performance again.

It cannot be assumed that sharing an 8 thread task over 4 processors will make a job 4 times as fast. Different issues may be raised by different examples (for example a client/server example would not consider global data).

Marks Breakdown

a) Example in an appropriate language	(5 marks)
Illustrating thread definition	(5 marks)
Co-ordination	(5 marks)

QUESTION FIVE

You have agreed to talk for 30 minutes at the next meeting of your local BCS branch. The title of your talk is "Client-Server Architectures – an overview of their use".

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

25 marks

Answer Pointers

The question defines the audience so the candidate should bear this in mind when answering. The numbers of slides indicate to the candidate that they should spend approximately 5 minutes on each slide and the candidate should remember only notes are required.

The uses to be discussed are diverse. A good answer may favour a particular use but should reflect the potential in a diverse range of uses.

Marks Breakdown

Each slide	(3 marks up to 8)
Bonus	(6 marks)