Three hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Highly Parallel Systems

Date: Monday 2nd June 2008
Time: 09:45 – 12:45

Please answer any THREE Questions from the FIVE questions provided

This is a CLOSED book examination

The use of electronic calculators is NOT permitted
1. a) Recent literature suggests that concurrency is about to emerge as a mainstream programming technique. Yet at the same time, concurrency is viewed as being an order of magnitude harder than sequential programming. Discuss the reasons that are given for this emergence and the potential difficulties that remain. Consider in your answer why concurrency is so difficult. (7 marks)

b) Explain the concept of multi-threading and multi-threaded parallel architectures. Discuss the effects on efficiency on this type of architecture of increasing the following: latency time, context switching time, number of contexts.

In 2008 Cray have launched the Eldorado machine (MTA-3); discuss the motivations for the technology differences between this and the MTA-2. Discuss application areas that are likely to run well on this machine.

Cray in 2006 announced an “adaptive computing” initiative, combining many types of parallel machine, including multi-threaded elements. Speculate on the type of application that might suit such a hybrid architecture. (9 marks)

c) Consider a sequential program, in a particular language, that requires greater performance. Assuming the only practical solution is to utilise parallelism, consider each of the following techniques:

- a compiler to generate parallel code for the language;
- making use of parallel extensions to the existing language to obtain parallel execution;
- rewriting the application in a new (parallel) language.

Consider the possible effects and the advantages and disadvantages of each approach. (4 marks)
2. a) Geoffrey Fox in 2007 commented on the lack of interaction between parallel computing and software engineering. Comment on this statement and consider ways in which software engineering techniques may benefit parallel computing, and the likelihood of their take-up. (7 marks)

b) Consider the detection of termination in distributed processes. Discuss the associated requirements and outline a token-based algorithm for its solution. What properties should be verified for such an algorithm? Consider the effects of using multiple tokens to detect termination. (5 marks)

c) A temporal database supports storage and querying of time-varying relations. Derive and populate a temporal relation, including time-start and time-end, to represent car ownership.

Give a TSQL query to compute the temporal aggregation of the number of car-owners of each type of car and show results.

Suggest how such a query could be executed in parallel in a shared memory environment. Analyse how the query might execute in a distributed memory environment.

Define the term skew-ratio, and analyse the effect that data skew might have on the execution of this query using the parallel schemes you have suggested. (5 marks)

d) Consider a program with which you are familiar. Analyse the program for parallelism considering both control flow and data distribution. Suggest how your application might be represented in a language and how it may execute on a parallel machine. (3 marks)
3. a) Explain ways in which a relation in a distributed database can be fragmented across multiple sites. Discuss what effect the fragmentation schemes have on query execution. (3 marks)

b) Define the concept of deadlock. Explain why deadlock in a distributed system causes significant problems. Discuss the merits of deadlock avoidance and deadlock detection schemes, and consider in what circumstances each might be most appropriately used. Consider reasons why locking schemes are not composable. (5 marks)

c) Consider a supplier-parts distributed relational database as follows, where each of the 3 tables is stored at separate sites X, Y and Z:

- Supplier table: S (S#, CITY)  Cardinality 100K tuples
- Parts table: P (P#, COLOUR)  Cardinality 250K tuples
- Supplier-Part table: SP (S#, P#)  Cardinality 1M tuples;

Given the following query to find supplier number for London suppliers of red parts,

```sql
SELECT S.S# WHERE
EXISTS SP EXISTS P (  
  S.CITY = 'London'  
  AND S.S# = SP.S#  
  AND SP.P# = P.P#  
  AND P.COLOUR = 'Red')
```

discuss a range of strategies for executing this query given the data distribution. Quantify each strategy making clear any assumptions you make. Illustrate your answer as appropriate. (8 marks)

d) The late Jim Gray, formerly of Microsoft, one of the fathers of transaction processing, suggested recently that the transaction concept had always been flawed from a database aspect. Discuss the reasons for and merits of this argument. (4 marks)
4. a) Briefly discuss how you might create a parallel program. (3 marks)

b) Compare and contrast the utility of recursive decomposition and data decomposition as parallel decomposition techniques.

Consider how to apply data decomposition techniques to the problem of computing frequent itemsets, and briefly consider potential performance effects. Illustrate your answer.

For the QUICKSORT algorithm show how recursive decomposition can obtain a restricted form of parallelism. Consider how to parallelise the splitting of the sub-sequences. Why is the choice of pivot element so critical in this regard? (6 marks)

c) Consider the requirements of the Google search engine activity and discuss the system architecture they have developed to meet these needs. The Google system makes significant use of MapReduce: discuss this with particular reference to its potential for parallelism. How is fault-tolerance achieved in this scenario? (6 marks)

d) A shared memory approach that offers a simple programming model, a simple hardware implementation and good performance has long been sought. Transactional memory has recently been proposed as a potential solution. Discuss the concept and its advantages and disadvantages. (5 marks)
5. a) On projections from a few years ago we should now have processors with a clock speed of approximately 10Ghz. Given this is not the case, and estimates suggest now that approximately 8 Ghz clock speed might be a limit, analyse the causes of this situation. Suggest ways in which performance may be increased and speculate on what you might expect to see in a 5 year timeframe. (6 marks)

b) Discuss optimisations within processor architectures that may cause problems for the maintenance of sequential consistency in a parallel system; illustrate the behaviour of the parallel system in each case. Reflect on why optimisations cause such problems. (6 marks)

c) Consider the motivation for the development of less restrictive memory models than sequential consistency. How are such models defined, how might they achieve their motivation and how might they be used? Discuss the type of constraints that are always likely to be enforced in a memory model. (4 marks)

d) Discuss the issues in designing a traffic light system for crossroads, paying particular attention to concerns of concurrency, mutual exclusion, and fail-safety. (4 marks)