One and a half hours

QUESTION PAPER MUST NOT BE REMOVED FROM THE EXAM ROOM

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Fundamentals of Distributed Systems

Date: Tuesday 31st May 2011
Time: 09:45 - 11:15

Please answer ALL questions in Section A
and any TWO questions from Section B

Answer Section A on this Question Paper
Use separate Answerbooks for EACH Question in Section B

For full marks your answers should be concise as well as accurate.
Marks will be awarded for reasoning and method as well as being correct.

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are
not programmable and do not store text.
Section A

Section A is a multiple choice section and is therefore restricted
Section B

Answer two of the three questions, using a separate answer book for each question.

1. a) Explain the terms ‘connection oriented’ and ‘connectionless’ as applied to traditional telephone and computer networks.

(4 marks)

b) By what mechanism do virtual circuits improve the quality of service (QoS) of network layer connections? What improvements are achieved and how do these improvements affect the performance of ‘Voice over IP’ (VOIP) applications.

(6 marks)

c) How does ‘transmission control protocol’ (TCP) achieve reliable and connection-oriented communication at the transport layer over a non-reliable connectionless network layer? Why is this mechanism normally considered unsuitable for voice transmission in VOIP applications?

(4 marks)

d) Explain how ‘unit datagram protocol’ (UDP) and ‘real time transport protocol’ (RTP) differ from TCP and why one of these protocols is considered suitable for voice transmission. What is the normal function of ‘real time transport control protocol’ (RTCP) in VOIP communications?

(6 marks)
2. a) Explain the role and effects of a ‘cache’ in a distributed system. (2 marks)

    b) Give two distinct examples of how the caching of data improves the behaviour of a distributed system. In each case, explain why a cache is needed, and describe the approach taken in that particular system to avoid inconsistent or out-of-date data. (6 marks)

    c) ‘HTTP is a stateless protocol’. Explain what is meant by this statement, and discuss why statelessness is an important factor in the reliability of web-based system. (4 marks)

    d) Describe two distinct mechanisms by which applications built on top of HTTP can achieve stateful behaviour. (4 marks)

    e) Stateful behaviour in a distributed system introduces the potential for deadlock. Explain the concept of deadlock using an example from a web-based application (the application can be a hypothetical one of your own devising, or a real system). (2 marks)

    f) Explain the problems associated with using timeouts to manage the behaviour of a cache, and to break deadlock. Give an example of where the use of a timeout can justifiably be used in a distributed system. (2 marks)
3. a) Consider that famous quote in an 1987 email by Leslie Lamport:

A distributed system is one in which the failure of a computer you didn't even know existed can render your own computer unusable.

Now, using the characteristics of distributed systems introduced in this course unit, answer the following:

i) State which characteristic of distributed systems underlies the fact that it is possible that one computer may fail without all other computers being aware of that and briefly (in one or two sentences) explain how this follows from how distributed systems were defined in this course unit. (3 marks)

ii) Explain (in two to three sentences) why it is possible that there is a computer, other than your own, whose existence your computer is unaware of, even though it may cause your computer to fail? (4 marks)

b) A general principle discussed in this course unit states that we can hope to reduce the total elapsed time taken to perform a computation on a collection of data items if we can apply the computation to partitions of the collection (i.e., smaller chunks of it) at the same time.

i) Assume that, in a setting where the cost of using the interconnect is not negligible, someone in your development team has stated that "it is always better (in terms of reducing the total elapsed time) to work with chunks that are small as possible". State clearly (i.e., a 'yes' or a 'no') whether you agree with your colleague, and explain (in two or three sentences) your reasons for holding this view. (4 marks)

ii) Assume now that there are three machines, of equal processing power, and the collection was broken down into three chunks $C_1$, $C_2$ and $C_3$, where the size of the $C_1$ chunk is double that of the equal-sized $C_2$ and $C_3$ chunks. Further assume that someone in your development team has stated that "this set-up allows us to do the work in a third of the elapsed time that would be needed in a sequential set-up using only one of the machines". State clearly (i.e., a 'yes' or a 'no') whether you agree with your colleague, and explain (in two to three sentences) your reasons for holding this view. (4 marks)

c) This question relates to the types of transparency with which a distributed system aims to deliver its functionality to users.

i) Explain (in two to three sentences) the difference between migration transparency and relocation transparency. (2 marks)

ii) Using the notions and concepts introduced in this course unit, explain (in two to three sentences) why migration transparency is less challenging to enforce than relocation transparency. (3 marks)