One and a half hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Fundamentals of Distributed Systems

Date:     Friday 18th May 2012
Time:     09:45 - 11:15

Please answer any TWO questions.

Use a SEPARATE answerbook for EACH question.

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.
1.  a) Give a brief answer to each of the following:

i) Why was the sampling rate 44.1 kHz adopted as standard for compact disc (CD) recordings of music? (1 mark)

ii) What is the bit-rate for standard CD recordings in stereo? (1 mark)

iii) What sampling rate is most commonly used for ‘narrow-band’ telephone-quality speech and VoIP? (1 mark)

iv) Why was this sampling rate decided upon? (1 mark)

v) How is the ‘ITU-G7111’ standardised bit-rate of 64kb/s achieved? (1 mark)

b) Explain how the VoIP application of distributed systems demonstrates that the latency of network links may not be negligible. What is meant by ‘one-way’ delay and ‘round trip’ delay and how do such delays affect interactive VoIP telephony? (5 marks)

c) What is meant by ‘timing jitter’ as observed at a VoIP receiver and what causes it? How does timing jitter affect VoIP voice quality and how can any adverse effects be reduced? (5 marks)

d) Explain the difference between Transmission Control Protocol (TCP) and UDP (User Datagram Protocol). Why are both these protocols unsuitable for interactive VoIP? (6 marks)

e) Why is RTP (Real Time Transfer Protocol) a preferred alternative to TCP and UDP for interactive VoIP? Is it considered a reliable protocol and does this matter for VoIP? (4 marks)
2. a) ‘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 systems. (4 marks)

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

c) What is meant by an ‘atomic’ transaction? Give an example of where an atomic transaction is crucial for the operation of a distributed system (you may choose any application area for your example). Describe the steps involved in the transaction, and comment upon the potential consequences of the overall transaction not being atomic. (3 marks)

d) What is the purpose of Two Phase Commit? Describe the protocol. (5 marks)

e) Describe how the Bully Algorithm can be used to select a master or co-ordinating node within a distributed system. State any assumptions about the nature of the distributed system. (4 marks)

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

g) 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. (3 marks)
3. **This question contains five sections, (a) to (e), and continues onto the next page.**

   a) Recall that one motivation for distributed systems given in the course was to make continuously-evolving, remote resources accessible for sharing. Give an example of a class of websites (or web businesses) you are familiar with that achieves this purpose. Your answer should briefly explain the overall idea behind the chosen class of websites but need not mention any concrete, specific, existing website, though your answer may be clearer if you do. (4 marks)

   b) The motivation referred to in the previous item can be said to be of a functional nature. Briefly explain what is meant by this observation using your answer to the previous item. (2 marks)

   c) One form of cloud computing is exemplified by servers (e.g., Google Calendar) that deliver functionality (e.g., keeping/managing diaries and task lists) also available in clients (e.g., Microsoft Outlook, or Apple iCal). It could be argued that one motivation for using cloud applications is the fact that most people now use many different devices. Briefly explain what problems arise from using many different devices in terms of the transparencies discussed in the course, then briefly suggest how a cloud application can better provide the desired transparencies than multiple clients could. (6 marks)

   d) Briefly explain:

   i) The notion of a barrier as a synchronization primitive for massively-distributed computing; (2 marks)

   ii) Why the use of a barrier requires attention to workload balancing; (2 marks)

   iii) What effect would result from using a barrier with poor workload balancing. (4 marks)

   **This question continues on the next page...**
e) Suppose you belong to a team of developers working for a web-based company that is an information portal on restaurants and places to eat. The website collects information about restaurants, and also stores user reviews. The company is considering a move towards adding the capability to book tables at a restaurant through the website. The company’s goal is to charge the restaurants a commission on each booking made through its website.

i) At a technical meeting, Sarah (one of your colleagues) stated that this new capability would force the company to change its approach to distributed systems architecture. Sarah argued that the current direct-message exchange style would not work anymore because of the need to enforce transactional semantics when booking tables and receiving the commission from the restaurant. Briefly illustrate what might go wrong in this scenario if transactional semantics were not enforced.

(2 marks)

ii) Sunita, the technical team leader, disagreed with Sarah. She said the need for transactional semantics did not necessarily require a change from a direct-message exchange. She then asked you: “We just need to use a transaction manager, but this doesn’t mean that the message exchange needs to be mediated, does it?” Briefly provide an answer to Sunita’s question.

(3 marks)