CARDIFF UNIVERSITY

Class Test

Academic Year: 2021-2022

Module Code: CMT202

Class Test Title: Distributed and Cloud Computing

Duration: Two and a half hours

Please read the following information carefully:

Structure of class test:

There are THREE pages.
There are FOUR questions in total.
There are no appendices.
The maximum mark for the class test is **60 marks** and the mark obtainable for a question or part of a question is shown in brackets alongside the question.

Instructions for completing the class test:

Answer **THREE** questions.

Important note: if you answer more than the number of questions instructed, then answers will be marked in the order they appear only until the above instruction is met. Extra answers will be ignored. Clearly cancel any answers not intended for marking.

1 (a) Distributed systems are increasingly becoming a more important type of computing system. Describe and justify three reasons for this growing importance. **[5]**

(b) Describe two examples of real-world computer systems which use a distributed computing paradigm and explain why these systems would be difficult to implement using a non-distributed computing paradigm. **[5]**

(c) Five types of failures in a distributed system are crash failure, omission failure, timing failure, response failure and Byzantine failure. Describe a real-world example of each of these types of failures. **[5]**

(d) A hash function is a function with a number of properties including the property that if two messages are not equal their hashes are not equal. That is, if $M \neq M'$ then $H(M) \neq H(M')$ where M and M' are messages and H is a hash function. Give an example of a function which does not have this property. **[5]**

2 (a) Within the context of a client-server architecture, describe one potential benefit of having a thin client and one benefit of having a fat client. **[5]**

(b) Describe a real-world application of edge computing. Explain the benefits of using edge computing in this application relative to using a centralized computing paradigm. **[5]**

(c) In a structured peer-to-peer (P2P) system, the chord algorithm may be used to locate data. Describe the network topology used by this algorithm and the motivation for its use. What is the purpose of shortcut links or connections in this algorithm? **[5]**

(d) Distribution transparency is an important concept in distributed systems. Briefly describe a method for achieving access transparency, a method for achieving migration transparency and a method for achieving replication transparency. **[5]**

3 (a) Describe two benefits of using code migration in a distributed system. Describe two challenges to implementing code migration in a distributed system. Describe a real-world example of code migration in a distributed system. **[5]**

(b) Consider a distributed system where a single computer in this system has a Universal Coordinated Time (UTC) receiver. Describe a method for performing clock synchronization in this system. **[5]**

(c) Consider an online banking system where individuals deposit and withdraw money from banking accounts. If this system was implemented as a distributed system, describe a potential unwanted consequence of clock drift in the system. **[5]**

(d) Describe how process replication can support fault tolerance in a distributed system. Describe how data replication can increase availability in a distributed system. **[5]**

4 (a) Interception, interruption, modification and fabrication are four types of security threats. Briefly describe a real-world example of each of these types of security threats. **[5]**

(b) Describe one advantage of using processes instead of threads in a distributed system and one

advantage of using threads instead of processes in a distributed system. Describe a real-world application of threads in a distributed system. What is the benefit of using threads instead of processes in this context? **[5]**

(c) Mary and John have regular face-to-face meetings in a location where their conversations cannot be overheard. Mary and John want to set up a secure channel for communication outside of these meetings. Why, in this context, is using a symmetric (secret-key) cryptosystem, as opposed to an asymmetric (secret-key) cryptosystem, a suitable solution? **[5]**

(d) Provide pseudocode for map and reduce functions in a MapReduce computer program which returns a value 1 if a specified word exists in a collection of documents and returns a value 0 otherwise. **[5]**