THE UNIVERSITY OF MELBOURNE

Student Number_____

Faculty/Dept. Computing and Information Systems

Subject Number COMP90024

Subject Name Cluster and Cloud Computing

Writing Time **2** **hours**

Reading Time **15** **minutes**

Open Book Status Closed Book

Number of pages (including this page) 3

Authorised Materials: None

Instructions to Students: This examination is worth 50% of your final mark

Answer 5 out of any 7 questions. Please note that only the first 5 questions will be marked.

Each question carries 10 marks.

The number in square brackets after each sub-question represents the marks allocated to it.

Instructions to Invigilators: Please provide students with standard script books

No calculators are allowed

This paper is NOT to be made available in the library after the examination

Paper to be held by Baillieu Library: Indicate whether the paper is to be held with the Baillieu Library.

Yes ☐ No X

Extra Materials required (please tick & supply)

Graph Paper ☐ Multiple Choice form ☐

Question 1:

A) Cloud computing does not solve the fundamental challenges associated with large-scale distributed systems. Discuss. [8]
B) How has the evolution of service-oriented architectures supported Cloud computing? [2]

Question 2:

A) Define Gustafson-Barsis' law for scaled speed-up of parallel programs. [2]
B) A parallel program takes 120 seconds to run on 8 processors. The total time spent in the sequential part of the program is 12 seconds. What is the scaled speedup? [2]
C) According to Gustafson-Barsis' law, how much faster could the application *theoretically* run if it ran across all 8 processors compared to running on a single processor? [2]
D) Why is theoretically italicized in the above? [4]

Question 3:

A) *Big data* is often associated with data having a range of properties including high volume, high velocity and high variety (heterogeneity). Discuss the advantages, disadvantages and suitability more generally of the following data solutions with regards to these big data properties:
    a. CouchDB [3]
    b. Apache Hadoop Distributed File System (HDFS) [3]
    c. Apache Spark [3]

    Your answer should include the way in which these solutions implement MapReduce.

    What other data properties can be associated with big data challenges? [1]

Question 4:

A) Representational State Transfer (ReST) based web services are often used for creating *Resource-oriented Architectures* (ROA) whilst Simple Object Access Protocol (SOAP)-based web services are often used to implement *Service-oriented Architectures* (SOA). Discuss the similarities and differences between a ROA and a SOA. [3]
B) Discuss the advantages and disadvantages of ReST vs SOAP for web services more generally. [5]
C) HTTP methods can be *safe* or *idempotent*. What is meant by the italicized terms, and give an example of each? [2]

Question 5:

A) Popek and Goldberg laid down the foundations for computer virtualization in their 1974 paper, *Formal Requirements for Third Generation Architectures*.
    a. Identify and explain the different types of classification of instruction sets for virtualization to occur according to the theorem of Popek and Goldberg. You should include the relationships between the instruction sets. [3]
    b. Describe how these principles are realized by modern virtual machine monitors/hypervisors. [2]
    c. Explain the differences between full virtualization and para-virtualisation. Give an example of a hypervisor that uses full virtualization and an example of a hypervisor that uses para-virtualisation. [2]
    d. Describe the role of a virtual machine manager/hypervisor with regards to memory management and shadow page tables. [3]

Please Turn Over

Question 6:

A) Explain what is meant by the following security terms:
- single sign-on [1]
- federated authentication [1]
- authorization [1]
- certification authority [1]
- identity provider [1]

B) Discuss the challenges in supporting finer-grained security in *hybrid* Cloud environments. You may refer to the importance and/or role of (some of) the terms in part A) of this question. [5]

Question 7:

A) The NeCTAR Research Cloud is based on the openStack technology.
   a. Describe the role and features of the following openStack components:
      i. Nova [1]
      ii. Swift [1]
      iii. Glance [1]
      iv. Keystone [1]
   b. Describe the interplay between these components that allows a researcher to create an instance of a virtual machine through a pre-existing snapshot. [3]

B) The NeCTAR Research Cloud has multiple *availability zones*.
   a. What is meant by the term: availability zone? [1]
   b. What are the implications of availability zones with regards to virtual machine instance creation and data volumes offered by NeCTAR? [2]

--- END OF EXAMINATION ---