University College Dublin

An Coláiste Ollscoile, Baile Átha Cliath

_____

SEMESTER I EXAMINATION – 2016/2017

_____

COMP 41110 / COMP30520

Cloud Computing

Prof. S. Dobson

Prof. Padraig Cunningham

Prof. M-Tahar Kechadi*

Time allowed: 2 hours

Instructions for candidates

Answer ALL questions of Part A and any ONE question of Part B.

Instructions for invigilators

Closed book exam

Part A

Question 1 *(20 marks)*

1. Explain why cloud computing is not network computing. You may use an example to explain your answer? *(5 marks)*

2. What is the difference between Middleware and Hypervisor in the cloud computing system? *(5 marks)*

3. Give one major challenge of cloud computing system. Explain its effect on the cloud computing's users. *(10 marks)*

Question 2 *(20 marks)*

In the context of Cloud Computing,

1. What is virtualisation? *(5 marks)*

2. Why virtualisation is interesting in a cloud computing environment? *(5 marks)*

3. Briefly describe Server Virtualisation. *(5 marks)*

4. Define paravirtualisation. Give an example of a paravirtualisation system. *(5 marks)*

Question 3 *(35 marks)*

1. What are the platform as a service (PaaS) benefits for developers on a cloud computing system? *(10 marks)*

2. Give five major advantages of a PaaS in cloud computing. *(10 marks)*

3. The use of elastic IP address is very popular in cloud computing environment. Explain what is an elastic IP address. *(5 marks)*

4. The main three components of AWS S3 are: Objects, Bucket, and Keys. Define each of these three components and how they are used in S3. *(10 marks)*

Part B

Question 4a $\qquad$ *(25 marks)*

Consider the following program written in Pig script. What does this program do? Explain your answer.

Algorithm 1 Pig script of a program.

1: A = LOAD 'page_views' AS (user, action, timespent, query_term, ip_addr, timestamp, estimated_revenue, page_info, page_links);
2: B = FOREACH A generate user, estimated_revenue;
3: C = GROUP B by user parallel 40;
4: D = FOREACH C
5:    E = ORDER B by estimated_revenue;
6:    F = E.estimated_revenue;
7:    GENERATE group, SUM(F);
8: ENDFOR
9: STORE D into 'L16out';

Question 4b $\qquad$ *(25 marks)*

Consider a distributed system that consist of S computing nodes. A distributed clustering technique consists of two main steps: 1) Each node executes a local clustering on their local data. 2) Each node sends its results to the server (a node that is elected to be the server) to aggregate the local results of each node to produce global clusters. The main steps of the algorithm are as follows:

1. Step 1:Given S nodes, partition the data objects into S nonempty subsets.

2. Step 2:Distribute the subsets among the S computing nodes.

3. Step 3: Execute on each node a clustering algorithm on its local data.

4. Step 4: Each node sends its results to the server.

5. Step 5: Aggregate the local results to produce global clusters.

Using Map/Reduce model, define the mapper and reducer of this distributed algorithm.

———————————————

oOo