

UNIVERSITY OF MALAYA

EXAMINATION FOR THE DEGREE OF MASTER OF DATA SCIENCE

ACADEMIC SESSION 2018/2019 : SEMESTER I

WQD7008 : Parallel and Distributed Computing

Jan 2019

Time : 3 hours

INSTRUCTIONS TO CANDIDATES :

Answer **ALL** the questions (50 marks).

(This question paper consists of 4 questions on 3 printed pages)

1. Cloud service models:

(a) List, explain, and exemplify the service models of cloud computing.

(9 Marks)

(b) Assume a relatively large-scale scenario in which a data processing application is about to deploy on a cloud environment. Many users may access the app at the same time. Can this scenario be deployed on the Microsoft Azure? If yes, provide your justification. If no, discuss all the service models being contributed for such a scenario and justify your answer.

(6 Marks)

2. The main responsibility of the MapReduce framework is to efficiently run a user's program on a distributed computing system.

(a) Define and explain the partitioning function of MapReduce to link the Map and Reduce workers. Draw a simple diagram for MapReduce partitioning function to link the Map and Reduce workers.

(10 Marks)

(b) Draw a simple diagram for the data flow implementation of all data flow steps.

(5 marks)

3. One of the requirements of Distributed File System (DSF) is the transparency.

(a) Explain Five forms of transparency in DSF.

(5 marks)

(b) Name and explain four requirements of DFS except transparency.

(8 marks)

4. Draw the process of different MapReduce's phases such as "Splitting", "Mapping", "Shuffling", "Reducing" and the final result sent to the client.

Input:

Nicole	John	Frank
Sara	Sara	Frank
Frank	Sara	John

(7 marks)

END