THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

This exam paper must not be removed from the venue

Venue _____
Seat Number _____
Student Number |__|__|__|__|__|__|__|__|
Family Name _____
First Name _____

School of Information Technology and Electrical Engineering

Semester Two Final Examinations, 2022

INFS3208 Cloud Computing

This paper is for St Lucia Campus students.

Examination Duration: 120 minutes

Planning Time: 10 minutes

Exam Conditions:

•This is a Closed Book examination - no written materials permitted
•Casio FX82 series or UQ approved and labelled calculator only
•During Planning Time - Students are encouraged to review and plan responses to the exam questions
•This examination paper will be released to the Library

Materials Permitted in the Exam Venue:
(No electronic aids are permitted e.g. laptops, phones)

None

Materials to be supplied to Students:
Additional exam materials (e.g. answer booklets, rough paper) will be provided upon request.

1 x Gradescope Bubble Sheet

Instructions to Students:
If you believe there is missing or incorrect information impacting your ability to answer any question, please state this when writing your answer.

Additional exam material rough paper will be provided upon request.
Please answer all the MCQ questions (Part A) on the Gradescope Bubble Sheet and answer all the short-answer questions (Part B) on the examination paper. Total Marks: 50

For Examiner Use Only

Question	Mark

Total _____

Part A – Multiple choice questions (20 marks, 1 mark per question)
Answer all questions on the Gradescope Bubble sheet.

1. Which of the following is NOT the business driver of Cloud Computing?

   a. Ease of use

   b. Capacity planning

   c. Cost reduction

   d. Organisational agility

2. Which of the following are the pre-existing technologies that heavily impacted on Cloud Computing?

   a. Clustering, Big Data, Artificial Intelligence

   b. Database, Grid Computing, Docker

   c. Clustering, Grid Computing, Virtualisation

   d. Artificial Intelligence, Big Data, 5G technology

3. What is the difference between Cloud and Internet?

   a. Cloud has a clear and finite boundary.

   b. Cloud is usually private and offers metered IT resources, while the Internet provides open access to many Web-based IT resources.

   c. Cloud often provides back-end processing capabilities and user-based access to these capabilities, while the Internet is dedicated to the access of content-based IT resources published via the World Wide Web.

   d. All of the above.

4. What are the cloud delivery models for Virtual Machine and Google App Engine, respectively?

   a. SaaS and PaaS

   b. IaaS and PaaS

   c. IaaS and SaaS

   d. None of the above

5. What computing resources do cloud users need to manage in Platform-as-a-service?

   a. Networking and storage

   b. Operation system

   c. Application and data

   d. None of the above

6. Given two servers A and B, which of the following Load Balancing algorithms should be considered if A has more CPU cores than B?

   a. Round Robin

   b. Weighted Round Robin

   c. Least Connections

   d. Random

7. Which of the following statements about load balancer is correct?

   a. To determine the destination, the layer 4 load balancer considers the IP address and ports, as well as the contents in the package.

   b. The layer 7 load balancer cannot direct Internet traffic based on the contents in the packet.

   c. The layer 4 load balancer has faster data transferring performance than the layer 7 load balancer in terms of latency.

   d. None of the above.

8. What are the features of docker?

   a. Easy modelling

   b. Version Control

   c. Application Isolation

   d. All of the above

9. Which of the following tools can help you manage multiple containers on one single machine?

    a. Dockerfile

    b. Docker-compose

    c. Docker machine

    d. None of the above

10. Which of the following tools can help you install Docker Engine on multiple virtual hosts?

    a. Dockerfile

    b. Docker-compose

    c. Docker machine

    d. Docker-swarm and Kubernetes

11. Which of the following statements about dockerfile is correct?

    a. Dockerfile is a text file that defines the environment inside the container.

    b. Dockerfile is a collection of instructions and commands.

    c. Dockerfile is transparent to users as it tells what is contained in the image.

    d. All of the above.

12. Deployments in Kubernetes use _____ to provide self-healing and scaling.

    a. ReplicaSets

    b. Pods

    c. Services

    d. Controller

13. Which of the following statement is WRONG?

    a. NoSQL databases require that schemas be defined before you can add data.

    b. NoSQL databases are built to allow the insertion of data without a predefined schema.

    c. Transaction properties (ACID) in the traditional relational databases are not suitable to be a set of properties to describe NoSQL databases.

    d. All of the above.

14. According to the CAP (Consistency, Availability, and Partition Tolerance) Theorem, which one of the following options is NOT possible to be satisfied for a NoSQL database?

    a. CA

    b. AP

    c. CP

    d. CAP

15. In GFS (Google File System) files are divided in _____ chunks.

    a. Variable size

    b. Fixed size

    c. Both fixed size and variable size

    d. None of the above

16. Which of the following statement is correct about HDFS?

    a. HDFS needs to run on expensive commodity hardware and cannot deal with failures.

    b. HDFS is to handle large files and the block size is 128 MB by default.

    c. HDFS supports multi-users to write one file simultaneously.

    d. In HDFS, low latency is more important than high sustained bandwidth.

17. In Spark programming, the default storage level of cache() is _____

    a. MEMORY_ONLY

    b. MEMORY_AND_DISK

    c. DISK_ONLY

    d. MEMORY_ONLY_SER

18. Which of the following operations is NOT a transformation in Spark?

    a. flatmap

    b. filter

    c. reduce

    d. reduceByKey

19. Which of the following is true for RDD?

    a. We can operate Spark RDDs in parallel with a low-level API.

    b. RDDs are similar to the table in a relational database.

    c. It allows the processing of a large amount of structured data.

    d. It has a built-in optimization engine.

20. Which of the following is the fundamental data structure of Spark?

    a. RDD

    b. DataFrame

    c. Dataset

    d. None of the above

Part B – Short answer questions (30 marks)
There are 7 questions in this part. Please answer all questions in the spaces provided on this examination paper.

Question 1: (**3 marks**) What are horizontal scaling and vertical scaling methods?

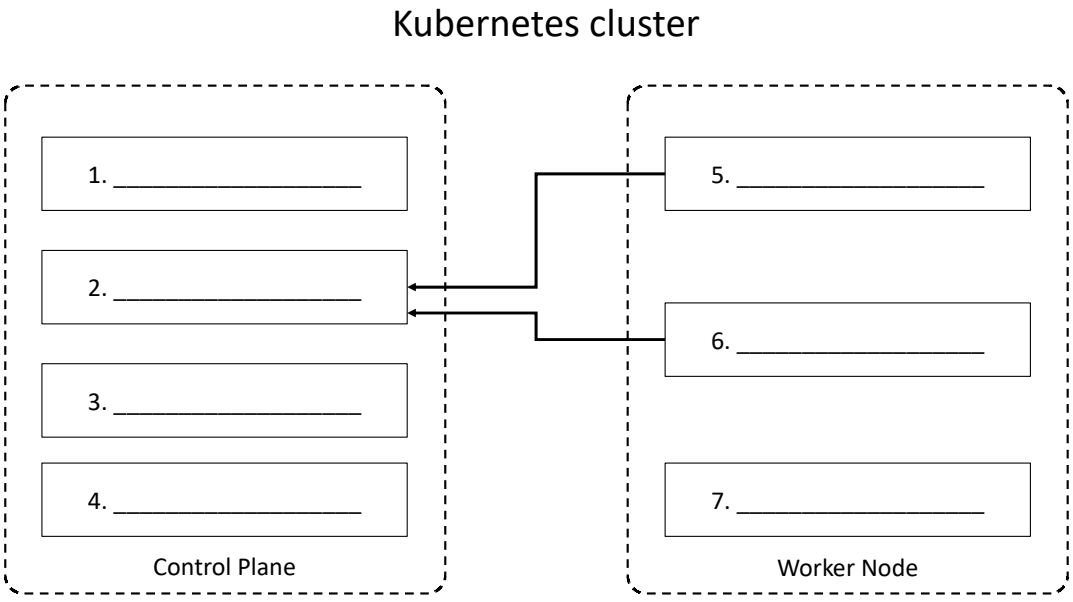
Question 2: (**4 marks**) Compare similarities and dissimilarities between the container and virtual machine technologies.

Question 3: (**5 marks**) Please briefly describe each step in the WRITE operation in Google File System (GFS).

Question 4: (**4 marks**) Please briefly describe what the four types of NoSQL databases are. Please also give at least one example for each type.

Question 5: (**3 marks**) What are the different services within Kubernetes? Please give a brief explanation of each service.

Question 6: (**6 marks**) Given the following conceptual framework of Kubernetes that contains a Control Plane node and a Worker node, please write down each component's name and its function label below. The function labels are listed in the table below.



Label	Function
(A)	makes sure that containers are running in a Pod.
(B)	watches for newly created Pods with no assigned node, and selects a node for them to run on.
(C)	validates and configures data for the API objects which include pods, services, replication controllers, and others.
(D)	is the software that is responsible for running containers.
(E)	is a daemon that embeds the core control loops shipped with Kubernetes.
(F)	provides a consistent and highly available key-value store for all the cluster data.
(G)	maintains network rules on nodes.

Answer Sample:

Component Name	**Function Label**
1. *Controller Manager*	8. *E*

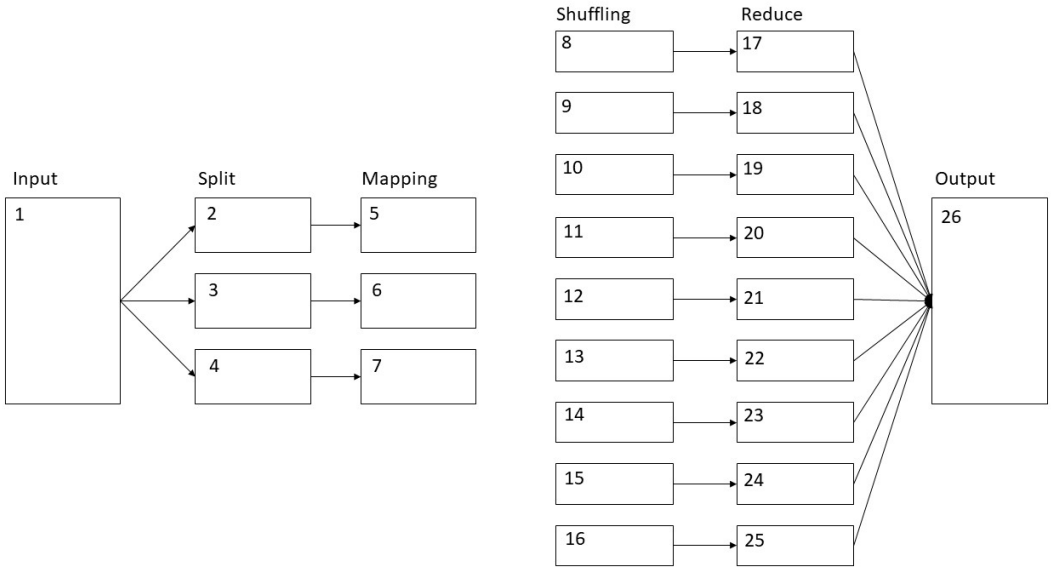
Write down the component name and the respective function label below:

Component Name	**Function Label**
2.	9.
3.	10.
4.	11.
5.	12.
6.	13.
7.	14.

Question 7: (**5 marks**) Given the following text input, please use the below Map-Reduce framework to perform word count (case insensitive). Note that a combiner should be in use if necessary. Please write down the answers according to the box numbers below.

Text Input:

Cloud computing is popular.
Google Cloud Platform is a popular cloud platform.
MapReduce model is a popular computing model.



Answer sample:
1: (Input).
Cloud Computing is popular
Google Cloud Platform is a popular cloud platform
MapReduce model is a popular computing model

Please write the answers in the below tables:

Split		Mapping	
2:		5:	
3:		6:	
4:		7:	

Shuffling		Reduce	
8:		17:	
9:		18:	
10:		19:	
11:		20:	
12:		21:	
13:		22:	
14:		23:	
15:		24:	
16:		25:	

Output	
26:	

END OF EXAMINATION