UNIVERSITY OF LINCOLN
UNITED KINGDOM

Time Constrained Assessment Examination

School	School of Computer Science
Module Title	Cloud Computing
Module Code	CMP2808M
Module Coordinator	Dr. Derek Foster
Duration of Assessment	3 hours (students without a LSP or PASS plan) 4 hours (LSP or PASS plan students)
Date	27th May 2021
Release Time	15:00 BST
Submission Time	18:00 (students without a LSP or PASS plan) 19:00 (students with LSP or PASS plan)

General Instructions to Candidates

1. In sitting this examination you agree to **comply** with the University of Lincoln Code of Conduct in Examinations.

2. You **must** submit your answers as a PDF or MS Word Document to Turnitin on Blackboard **before** the submission time: failure to do so will be classified as misconduct in examinations.
   We strongly recommend you submit 15 minutes prior to the deadline.

3. You **must** also send a copy of your work to the **socssubmissions@lincoln.ac.uk** at the same time. You must place the Module Code and your Student Id in the Subject Field of the Mail.

4. This assessment is an **open resource format**: you may use online resources, lecture and seminar notes, text books and journals.

5. **No collaboration or interaction** with other candidates or individuals using any means of communication or device is permitted during online examinations

6. All work will be **subject to plagiarism and academic integrity checks**. In submitting your assessment you are claiming that it is your own original work; if standard checks suggest otherwise, Academic Misconduct Regulations will be applied.

7. **The duration of the Time Constrained Assessment will vary for those students with Personal Academic Study Support (PASS)**. Extensions do not apply, but Extenuating Circumstances can be applied for in the normal way.

Module Specific Instructions to Candidates

- Provide answers to all **six** questions in the Answers section at the **end** of this document. To strengthen your discussion and responses, support your answers with graphs, figures and references where appropriate to do so. The total marks for each question vary so please thoroughly check as you go through them.

Question 1 – Cloud Computing Models (Total 30 marks)

Define the following main cloud computing models and the benefits they offer. For each model:

i) provide a definition and benefits of the model

(4 marks)

ii) provide two examples of how the model might be used for a real-world cloud application

(3 marks for each example)

a) Infrastructure as a Service (IaaS)

b) Platform as a Service (PaaS)

c) Software as a Service (SaaS)

Question 2 – Virtualisation (Total 15 marks)

For questions 2a) and 2b) below, briefly define the named **hypervisor virtualisation** approaches and give an example where each might be used:

a) Bare-metal virtualisation **(5 marks)**
b) Hosted virtualisation **(5 marks)**

c) In the context of **interoperability**, if a virtual machine is classed as **portable,** what does this mean? You should also provide a short discussion on the associated benefits.

(5 marks)

Question 3 – Networking (Total 10 marks)

Virtual networks (VPNs in this context) are an important component of the cloud and help to provide secure IT infrastructure communication. They provide secure cross-premises connectivity as well as secure communications for employees working remotely. One type of VPN connectivity that businesses use is a **site-to-site** VPN.

i. Name and define the other common VPN type a **remote** employee would use and give a detailed example of how it might be used by a company.

(6 marks)

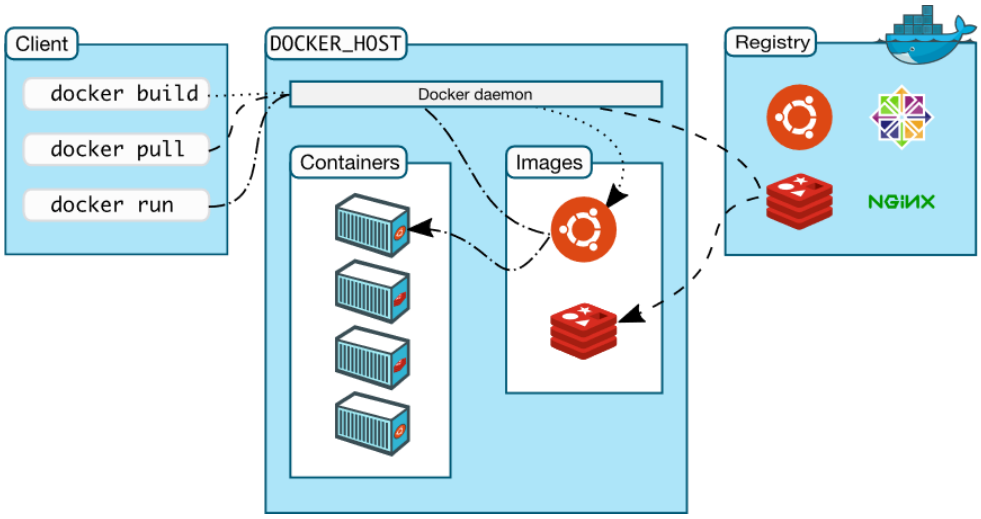
ii. You should illustrate your answer with a diagram. **(4 marks)**

Question 4 – Containers (Total 15 marks)

a) Containerisation is a recognised cloud model commonly named Containers-as-a-Service (CaaS). Give a contextualised example of how CaaS is used in industry and the benefits it offers.

(5 marks)

b) Docker is the most widely used and open source container platform. A basic overview of the Docker architecture is shown in the below diagram:



From the DOCKER_HOST part of the diagram above, define and discuss the **Containers** and **Images** objects. Your discussion should include the technical aspects of the objects, and how they interact with the Docker daemon in the diagram.

(10 marks)

Question 5 – Cloud Ethics and Data Privacy (Total 20 marks)

Data privacy is an important area that a UK business must adhere to by specific Data laws when using cloud services to manage and process customer data. To support businesses who manage their data through cloud services, and to manage expectations around this, Service Level Agreements (SLA) are important contractual documents that businesses agree to when they sign up to cloud storage services. In this example you are the lead developer on a project for an eHealth business that wants to run their core business offering through cloud services.

a) Name and briefly describe within the context of the eHealth business, the **three** main **data privacy** considerations in the SLA when running cloud services that manage sensitive data [5 marks each].

(15 marks)

b) As well as the data privacy considerations from question 5a), another area of concern for the eHealth business are the business continuity considerations with their services hosted in the cloud. Name and briefly describe **one** of the main business continuity considerations that will be part of the SLA.

(5 marks)

Question 6 – Cloud Storage (Total 10 marks)

a) Cloud providers offer a number of **non-relational** data management services for storing data. Name and briefly define **one** such storage service that is offered by the Google Cloud Platform, you should then give an example of how it might be used in scalable software systems.

(5 marks)

b) Briefly define the cloud term **High Availability** in the context of **cloud data storage,** and give an example of how it might be used in a cloud based system.

(5 marks)

Answers

Q1

1) Models

a) IaaS

i) IaaS (Infrastructure as a service) is a cloud computing model where the cloud provider will provide the hardware (Servers, storage and networking) and the virtualisation for a service. The customer would then use this platform to install operating systems onto and run applications as if they were running on a server in a private cloud infrastructure. All configuration, patching and security is controlled by the customer in practice known as "DevOps".

The key benefits to IaaS are that the customer has complete control over the entire application configuration and setup process. This would be useful in instances where customers are currently running on-premise servers and would like to adopt a "lift-and-shift" approach to move their infrastructure to the cloud without redesigning or recreating their applications and configuration. IaaS also allows multiple applications to be run in a single instance which may be more cost effective than a PaaS solution where the costs are for each service.

ii) - An example usage in IaaS is disaster recovery where backed up virtual machines that originally existed on-premise can be run in the cloud as part of a disaster recovery scheme. Due to IaaS not requiring the purchase of hardware, this would allow a customer to immediately start their disaster recovery rather than waiting for purchased hardware.
- IaaS can be suitable for running testing environments. Virtual machines can be created and destroyed quickly allowing development teams to test features or bugs on specific operating systems without having to manage a local server to do this.

b) PaaS

i) PaaS (Platform as a service) is a type of cloud computing model where the hosting provider manages all the hardware, networking and operating systems and where the customer would provide the data and applications to run on the "platform". The benefits to this model is that customers don't have to manage the operating system in terms of security and configuration meaning that deployment of applications on PaaS is both simpler to get started with and faster to setup.

By being simpler to get started with and faster to setup overall, PaaS is available to a wider audience since specialised skills and knowledge are not required to keep the application running smoothly and without any security issues at the networking/operating system level

ii) – Hosting a web application on a platform such as GCPs "App Engine" provides a highly scalable and serverless way to provide a service to users. For example, an API can be written and hosted on App Engine that can handle both small amounts of users and huge amounts of traffic.

- Hosting a database using something like Azure SQL is a PaaS service that gives customers access to a highly performant and scalable database without having to worry about upgrading or backups to the database infrastructure. Since the only thing the customer manages is the data, SQL is easy to start with and doesn't require the company to hire dedicated DevOps team

c) SaaS

i) SaaS (Software as a service) is a fully managed cloud computing model where the only thing the customer interacts with is the final end-product. This product is suited for customers that need a ready-made solution without any development time or resources required. A key benefit of this is cost savings since the customer can start using the product straight away. All aspects of the service including upgrades, security and scaling are all managed by the cloud vendor.

ii) – Office 365 is a subscription SaaS product where companies pay a monthly or annual subscription to use the product from anywhere in the world. Customers can store files, receive emails and edit documents all through the one central platform which is managed by Microsoft.
- Blackboard (As used by the University) is another example of a SaaS product where Blackboard would host all the servers and manage all of the configuration. This can be accessed from all over the world and can be accessed by all University staff and students.

2) Virtualisation

a) Bare-metal virtualisation is a type of virtualisation where the entire operating system is a bare-bones virtualisation engine that only runs the necessary services to provide virtualisation, networking and storage. Because of this, the virtual machines are "closer" to the hardware in that there is less operating system overhead as the virtualisation engine has exclusive access to all the hardware. A key example of this being used is in large datacentres or server farms where less overhead is critical to meet performance and energy usage goals.

An example of bare-metal virtualisation is VMWare's ESXi server which can be deployed in large server farms and be clustered together using vSphere

b) A hosted virtualisation is where a virtualisation engine runs on top of a host operating system such as Windows or Linux. This virtualisation engine runs alongside other applications running on the host operating system. This means there is a higher overhead since the virtualisation engine has to share resources but an advantage of this is that is allows the host computer to be used for more traditional tasks as well as hosting VMs. A key example of where this might be useful is a developer testing application on their computer. They can install a virtual operating system inside the virtual machine and use it like it was a physical computer.

c) If a virtual machine is classed as portable, it means it can be run on any underlying host operating system regardless of the physical hardware or platform that the VMs is being run on. The advantage to this is that the VM can

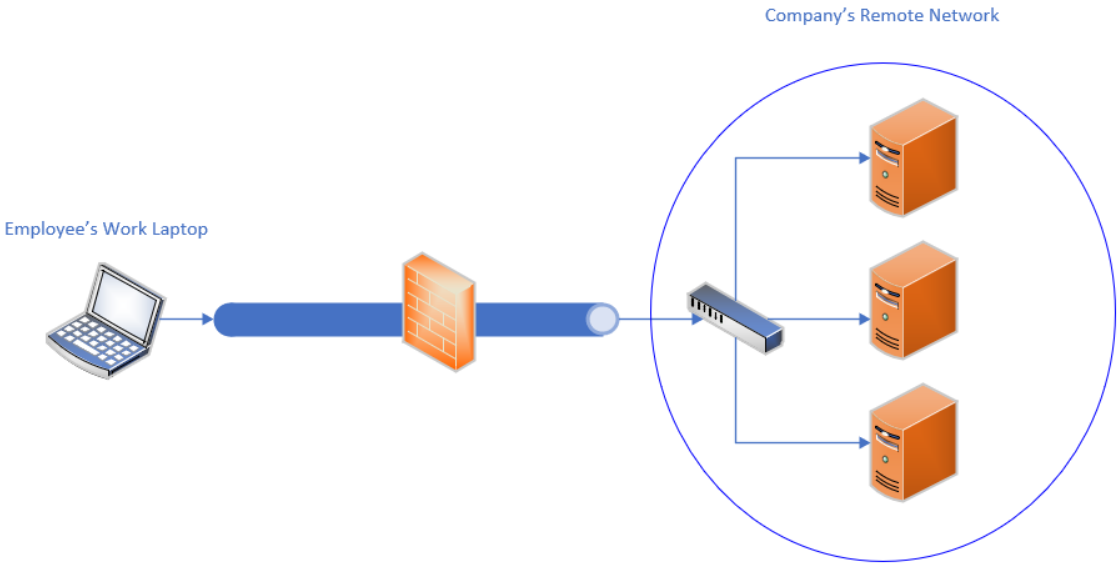
be moved between private and public clouds trivially without recreating or redesigning the application. This is also true for moving between different public clouds such as GCP, AWS and Azure. This is incredibly important to customers since it ensures they don't suffer from vendor lock-out. That is, the ability to be able to move to a different vendor without significant issues.

3) Networking

a) A remote employee would use a type of VPN called a "point-to-site" VPN which is where a "point" i.e. the client's PC or laptop, is connected to a remote network to gain access to company resources.

For example, a finance company does not allow their internal servers to be accessed from the public internet due to security and privacy concerns. To facilitate employees working from home being able to connect to these servers, the company sets up a VPN that employees can connect to using their work laptops to access these server resources. This VPN tunnel ensures that all data being accessed is not going across the public internet unprotected while still allowing convenient access to data and services that the remote employees might need. This VPN is connected to using a client installed on the employee's laptop where they use their work credentials to authenticate to the VPN

b) The diagram below illustrates the employee's laptop using a VPN tunnel to connect to the company remote network to access any of the company's servers and resources.



4) Containers

a) CaaS is commonly used in industry to host applications that have been designed in a micro-service architecture. This architecture is designed to split traditional monolith applications which would be traditionally run using a service such as IaaS or PaaS and split each "service" of the application into a separate container. The key advantage of this is each service of the application is individual to the other services meaning services can be scaled

up and be restarted individually. An example of this is an application that has it's frontend, API backend, statistics collector and database services being run in different containers.

This approach allows the developers to treat each service separately. This means that upgrades and maintenance can be done for each service individually. With load balancing and utilising container replicas (That is, more than one of each service instance running) downtime can be minimized by doing what is called a "rolling upgrade" where each instance of the service is taken offline, upgrades and brought back online one at a time to prevent complete service downtime. This process can usually be done automatically without manual developer intervention saving time and money when the number of services is in the 10s, 100s or 1000s of instances.

Another benefit of containers in this instance is that the container images are typical version controlled in some way. Meaning that if an upgrade is unsuccessful, the whole container replica set could be rolled back to a prior working version without much work.

b) Images and container

i) Containers are an instance of an image. This container using the shared OS kernel provided by the Docker host to run the image. Containers are completely stateless by themselves (When deleted, no files are saved unless volumes are used) and can be created, deleted and moved quickly. As well as the image that can be run inside the container, volumes can be connected to the container to provide persistent storage. Networks can also be connected to the container to provide networking to the image instance.

ii) Images are templates for an operating system to be run inside a container. There are a variety of images provided on the official Docker repository such as Ubuntu or Debian. Using a special type of file called a "Dockerfile", a developer can build upon these base images to install applications or services on the image. Images consist of version-controlled layers meaning that when an image is updated, only the layers that have been modified will have to be redownloaded by the Docker host. This ensures that space is saved on the host while also allowing Docker containers to be updated quickly without updating the entire image unnecessarily.

iii) The Docker daemon controls the life-cycle of both containers and images from creation to deletion. For an image, the daemon will handle the updating and downloading of layers where necessary. This could be downloading from the public Docker repository or a private hosted repository. For containers, the daemon will handle the creation and deletion of containers from their configuration files as well as managing the status of the container (Whether it is running, stopped or restarting)

5) Cloud ethics and data privacy

a) 3 Privacy considerations

i) When dealing with sensitive medial information that the eHealthy industry would be, it is incredibly important to ensure any data store is stored

confidentially. This means that data should be protected from unauthorised access and should only be stored and accessed when necessary. From a cloud platforms perspective, this means all data should be encrypted both at-rest (When stored on the servers) and in-transit (When being accessed over a network). Cloud service providers should have systems in place to prevent both physical access to servers and networking systems to protect network traffic from unauthorised access. Companies that store data should subsequently have access control setup to prevent unauthorised access.

ii) Due to sensitive data being stored, it is imperative that each cloud vendor that handles this data ensures that ownership remains with the company. This ensures that data cannot be kept by a third-party and that the original company has full rights to create, move and delete all data.

Under GDPR, all sensitive data must be stored within the EU or in jurisdictions which provide similar levels of protection. This means that a cloud provider has to be selected carefully according to region to ensure data is handled and stored correctly.

iii) According to GDPR laws, data should only be kept if required. Once the data is no longer required or consent is withdrawn, the data should be immediately deleted and destroyed off the system. This forms a part of the "right to be forgotten" under the GDPR laws. From a physical level, once hard drives have reached the end of their life, they should be securely destroyed to ensure no data could be recovered from them. From a software level, databases should be configured with retention periods so that data is automatically deleted when it hits a certain age.

b) As part of the SLA, a key aspect of business continuity is to have plans in place to ensure that the business will not be interrupted in the case of a disaster. More traditionally, disasters have been seen as natural disasters such as fires or floods but it is more true than ever before that this also applies to cyber attacks such as ransomware. For this reason, there are several key aspects which need to be considered:

i) Data backups – Data should be backed up using the 3-2-1 strategy at minimum (3 copies, 2 different types of media, 1 offsite/offline backup). To prevent against cyber attacks, these backups should be immutable (That is, not able to be deleted/modified once backed up) and all backups should be encrypted to prevent unauthorised access.

ii) System backups – In the event of a disaster, there should be plans in place about what to do when the original system is not available. This may be restoring VM backups to an IaaS service or using a cloud SaaS solution for email for example

6) Cloud storage

a) Google Cloud platform's BigTable service is a NoSQL data storage solution which is primarily used for big data analytics due to it being able to handle petabytes of raw data.

One example of where this might be used is in advertising where there are

masses of realtime data being generated every second that needs to be stored. Once it is stored, the database needs to be able to be queried quickly to gain insights into the data. BigTable is able to be queried using GCPs BigQuery system which allows petabytes of data to be queried quickly. For example, an advertiser might want to find out statistics for a particular demographic in a specific country over a specific time range. BigTable is able to be queried for this data efficiently for such a massive data set.

b) “High availability” is the concept of ensuring a service meets a high level of uptime requirements to ensure there is minimum disruption to the end user. Typically this involves an uptime of 99.999% or higher depending on how critical the workload is.

   In the context of cloud data storage, an example would be in a database solution which contains multiple database server replicas to ensure uptime. An example of this is in Amazon Aurora where there are multiple database replicas that all connect to a single shared cluster volume. There can be as many as 15 replicas in a cluster at once to ensure high availability.