Tutorial 1: Introduction to Cloud Computing

Question Set:

Q1: What is the NIST (National Institute of Standards and Technology) definition of Cloud Computing?

ANSWER:
Cloud Computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Q2: Give one example of virtualization application.

ANSWER:
Consider a company that needs servers for three functions:

- Store business email securely
- Run a customer-facing application.
- Run internal business applications.

Each of these functions has different configuration requirements:

- The email application requires more storage capacity and a Windows operating system.
- The customer-facing application requires a Linux operating system and high processing power to handle large volumes of website traffic.
- The internal business application requires iOS and more internal memory (RAM).

To meet these requirements, the company sets up three different dedicated physical servers for each application. The company must make a high initial investment and perform ongoing maintenance and upgrades for one machine at a time. The company also cannot optimize its computing capacity. It pays 100% of the servers' maintenance costs but uses only a fraction of their storage and processing capacities.
Efficient hardware use: With virtualization, the company creates three digital servers, or virtual machines, on a single physical server. It specifies the operating system requirements for the virtual machines and can use them like the physical servers. However, the company now has less hardware and fewer related expenses.
Infrastructure as a service: The company can go one step further and use a cloud instance or virtual machine from a cloud computing provider such as AWS. AWS manages all the underlying hardware, and the company can request server resources with varying configurations. All the applications run on these virtual servers without the users noticing any difference. Server management also becomes easier for the company's IT team.

Q3: Define cloud and IT resource and provide some illustrative examples.

ANSWER:

- **Cloud** is a distinct IT environment that is designed for the purpose of remotely provisioning scalable and measured IT resources. VMs on Google Cloud Platform, AWS, Oracle are all well-developed business cloud products.

- **IT resource** is a physical or virtual IT-related artifact that can be either software-based or hardware-based. E.g., physical server, virtual server, storage devices, etc.

Q4: Explain what scaling is and differentiate between Horizontal Scaling and Vertical Scaling. Support your answer with examples.

ANSWER:

- **Scaling:** the ability of the IT resource to handle increased or decreased usage demands.
- **Horizontal scaling:** allocating or releasing of IT resources (same type)
- **Vertical scaling:** higher or lower capacity of the current IT resources (less common due to downtime)
- **Comparison:**

Horizontal Scaling	Vertical Scaling
Less expensive	More expensive
IT resources instantly available	IT resources normally instantly available
Resource replication and automated scaling	Additional setup is normally needed
Additional IT resources needed	No additional IT resources needed
Not limited by hardware capacity	Limited by maximum hardware capacity

Q5: Identify the six essential characteristics of Cloud Computing and provide a brief introduction to each characteristic.

ANSWER:

1. **On-Demand Usage:** A cloud consumer can unilaterally access cloud-based IT resources giving the cloud consumer the freedom to self-provision these IT resources. Once configured, usage of the self-provisioned IT resources can be automated, requiring no further human involvement by the cloud consumer or cloud provider. This results in an on-demand usage environment. Also known as "on-demand self-service usage," this characteristic enables the service-based and usage-driven features found in mainstream clouds.
2. **Ubiquitous Access:** Ubiquitous access represents the ability of a cloud service to be widely accessible. Establishing ubiquitous access for a cloud service can require support for a range of devices, transport protocols, interfaces, and security technologies. To enable this level of access generally requires that the cloud service architecture be tailored to the particular needs of different cloud service consumers.
3. **Multitenancy**: The characteristic of a software program that enables an instance of the program to serve different consumers (tenants) whereby each is isolated from the other, is referred to as multitenancy. A cloud provider pools its IT resources to serve multiple cloud service consumers by using multitenancy models that frequently rely on the use of virtualization technologies. Through the use of multitenancy technology, IT resources can be dynamically assigned and reassigned, according to cloud service consumer demands.
4. **Elasticity**: Elasticity is the automated ability of a cloud to transparently scale IT resources, as required in response to runtime conditions or as pre-determined by the cloud consumer or cloud provider. Elasticity is often considered a core justification for the adoption of cloud

computing, primarily because it is closely associated with the Reduced Investment and Proportional Costs benefit. Cloud providers with vast IT resources can offer the greatest range of elasticity.

5. **Measured Usage**: The measured usage characteristic represents the ability of a cloud platform to keep track of the usage of its IT resources, primarily by cloud consumers. Based on what is measured, the cloud provider can charge a cloud consumer only for the IT resources used and/or for the timeframe during which access to the IT resources was granted. In this context, measured usage is closely related to the on-demand characteristic.
6. **Resiliency**: Resilient computing is a form of failover that distributes redundant implementations of IT resources across physical locations. IT resources can be pre-configured so that if one becomes deficient, the processing is automatically handed over to another redundant implementation.

Q6: What are the four deployment models in cloud computing? Please read these reading materials and answer this question in detail.

ANSWER:

1. **Private cloud:** The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off-premises.
2. **Community cloud:** The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off-premises.
3. **Public cloud:** The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.
4. **Hybrid cloud:** The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

Q7: Discuss the differences between cloud delivery models and make examples of delivery models.

ANSWER:

Deliver Model	Control Level	Functionality	Consumer Activities	Provider Activities
SaaS	Usage and usage-related configuration	Access to front-end user-interface	Uses and configures cloud services	Implements, manages, and maintains cloud service Monitors usage by cloud consumers
PaaS	Limited administrative	Moderate level of administrative control over IT resources relevant to cloud consumer's usage of the platform	Develops, tests, deploys, and manages cloud services and cloud-based solutions	Pre-configures platform and provisions underlying infrastructure, middleware, and other needed IT resources, as necessary

				Monitors usage by cloud consumers
IaaS	Full administrative	Full access to virtualized infrastructure-related IT resources and possibly to underlying physical IT resources	Sets up and configures bare infrastructure, and installs, manages, and monitors any needed software	Provisions and manages the physical processing, storage, networking, and hosting required Monitors usage by cloud consumers

Examples:
IaaS: virtual machines (EC2 etc)
PaaS: Google App Engine
SaaS: Dropbox, Office 365

Q8: What do Cloud-Enabling Technologies (CETs) include?

ANSWER:
1) Broadband Networks and Internet Architecture
2) Virtualization Technology
3) Data Centre Technology
4) Web Technology
5) Multitenant Technology

Q9: Please summarize the key points of Virtualization Technology for servers.

ANSWER:
1) Server virtualization is the process of abstracting IT hardware into virtual servers using virtualization software.
2) Virtualization provides hardware independence, server consolidation, and resource replication, and further supports resource pooling and elastic scalability.
3) Virtual servers are realized through either operating system-based or hardware-based virtualization.

Q10: What is Google Cloud Compute Engine? What are the benefits of Google Cloud Compute Engine?

ANSWER:
Compute Engine refers to customizable virtual machines in Google Cloud. There are three different machine type families, i.e., general-purpose, compute-optimized and memory-optimized.
- General-purpose machines are best suited for general servers, websites and databases.
- It is recommended to use a compute-optimized machine for computing intense applications, e.g., high-performance computing, gaming, and electronic design automation.
- Memory-optimized machine is for memory-intensive workloads, such as in-memory databases, in-memory analytics, and machine learning.

Benefits: live migration, right size recommendations, container support, sustained use saving.

References

[1]. Peter Mell, Timothy Grance, The NIST Definition of Cloud Computing, National Institute of Standards and Technology, September 2011.
[2]. Erl, Thomas, Ricardo Puttini, and Zaigham Mahmood. Cloud computing: concepts, technology & architecture. Pearson Education, 2013.
[3]. Virtualization definition from IBM: https://www.ibm.com/topics/virtualization
[4]. Virtualization example from: https://aws.amazon.com/what-is/virtualization/
[5]. Erl, Thomas, Ricardo Puttini, and Zaigham Mahmood. Cloud computing: concepts, technology & architecture. Pearson Education, 2013.
[6]. Google Cloud Tech (https://www.youtube.com/watch?v=KBeyQHoAcrQ)