**AWS Module 1**

**What is Cloud Computing?**

Cloud computing is the on-demand delivery of compute power, database storage, applications, and other IT resources through a cloud services platform via the internet with pay-as-you-go pricing.

Six Advantages and Benefits of Cloud computing

* Trade capital expenses for variable expense
* Benefit from massive economies of scale
* Stop guessing capacity
* Increase speed and agility
* Stop spending money on running and maintaining data centres
* Go global in minutes

**Deployment Models: -** There are a range of deployment models, from all on-premises to fully deployed in the cloud. Many users begin with a new project in the cloud, and they might integrate some on-premises applications with these new projects in a hybrid architecture. They might decide to keep some legacy systems on-premises. Over time, they might migrate more and more of their infrastructure to the cloud, and they might eventually reach an all-in-the-cloud deployment.

**Product and Services: -** AWS offers a broad set of global cloud-based products, including compute, storage, databases, analytics, networking, mobile, developer tools, management tools, Internet of Things (IoT), security, and enterprise applications.

**AWS Partner Network (APN): -** APN Partners are focused on your success, and they help customers take full advantage of all the business benefits that AWS has to offer.

**AWS Marketplace: -** The AWS Marketplace is a digital catalog with thousands of software listings from independent software vendors, where you can find, test, buy, and deploy software to run on AWS.

These offering can range from simple web server applications to security, networking business intelligence, databases, DevOps, and media. Many of these applications offer pay-as-you-go or Bring Your Own License (BYOL) models.

**AWS Infrastructure Notes:**

The AWS Cloud infrastructure is built around Regions and Availability Zones. AWS Regions provide multiple, physically separated, and isolated Availability Zones that are connected with low latency, high throughput, and highly redundant networking.

As of the time of publication, the AWS Cloud spans 55 Availability Zones within 18 geographic Regions and 1 Local Region around the world. There are announced plans for 15 more Availability Zones and five more Regions in Bahrain, Hong Kong SAR, Sweden, and South Africa; and a second AWS GovCloud Region in the US.

There is a Local Region in Osaka, Japan (**Osaka-Local**). An AWS Local Region is a single data center that is designed to complement an existing AWS Region. It is available to select AWS customers who request access. Customers who want to use the Asia Pacific (Osaka) Local Region should speak with their sales representative. Like all AWS Regions, AWS Local Regions are completely isolated from other AWS Regions.

Each AWS Region has multiple, isolated locations that are known as Availability Zones. Amazon Relational Database Service (Amazon RDS) provides you with the ability to place resources (such as instances) and data in multiple locations. Resources aren't replicated across AWS Regions unless you do so specifically.

**Compute Services Notes:**

Building and running your application starts with compute, whether you are building enterprise, cloud-native, or mobile applications; or running massive clusters to sequence the human genome.

AWS offers a comprehensive portfolio of compute services that allow you to develop, deploy, run, and scale your applications and workloads in the world’s most powerful, secure, and innovative compute cloud.

Later in this module, you will learn about both Amazon Elastic Compute Cloud (Amazon EC2) and Amazon Lightsail. See those sections for more details.

Both serverless computing and container services are beyond the scope of this class. We will cover these topics in subsequent courses, but the following descriptions provide an introduction to some of the key services for serverless computing and containers.

**AWS Lambda:**

AWS Lambda lets you run code without provisioning or managing servers. You pay only for the compute time you consume--there is no charge when your code isn't running. Additional information about Lambda can be found at: <https://aws.amazon.com/lambda>

**AWS Container Services:**

Amazon Elastic Container Service (Amazon ECS) is a highly scalable, high-performance container orchestration service that supports Docker containers. It allows you to run and scale containerized applications on AWS. You can find more details at: <https://aws.amazon.com/ecs/>

Amazon Elastic Container Service for Kubernetes (Amazon EKS) makes it straightforward to deploy, manage, and scale containerized applications that use Kubernetes on AWS. Details can be found at: <https://aws.amazon.com/eks/>

AWS Fargate is a compute engine for Amazon ECS and Amazon EKS that allows you to run containers without having to manage servers or clusters. You can find more information at: <https://aws.amazon.com/fargate/>

**Amazon EC2 Notes:**

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure and resizable compute capacity in the cloud. It's designed to make web-scale cloud computing easier for developers.

Amazon EC2 presents a true virtual computing environment, and it allows you to use web service interfaces to launch instances with a variety of operating systems, load them with your custom application environment, manage your network’s access permissions, and run your image by using as many or few systems as you want.

Amazon EC2 instance types

\*Required Reading -- Quiz and exam questions might refer to information from the link about instance types below.

Amazon EC2 provides a wide selection of instance types that are optimized to fit different use cases. Instance types comprise varying combinations of CPU, memory, storage, and networking capacity. They give you the flexibility to choose the appropriate mix of resources for your applications. Each instance type includes one or more instance sizes, which allows you to scale your resources to the requirements of your target workload. Current details about available instance types are available at: <https://aws.amazon.com/ec2/instance-types/>

**Amazon Lightsail Notes:**

Amazon Lightsail is the easiest way to get started with AWS for developers, small businesses, students, and other users who need a simple virtual private server (VPS) solution. Lightsail provides developers compute, storage, and networking capacity, and it also provides capabilities to deploy and manage websites and web applications in the cloud. Lightsail includes everything you need to launch your project quickly--a virtual machine, solid state drive (SSD)-based storage, data transfer, Domain Name System (DNS) management, and a static IP--for a low, predictable monthly price.

A more detailed introduction from AWS re:Invent 2017 is available here: <https://www.youtube.com/watch?v=29_LqYnomdg>. Note that pricing has changed (decreased) since this video was created. Specific details are on the [Lightsail web page](https://aws.amazon.com/lightsail" \o "Amazon Lightsail" \t "_blank).

**AWS Module 2**

**CIDR Notation: -** An important concept that’s used in networking on AWS is CIDR or Classless Inter -Domain Routing. CIDR network addresses are allocated in a virtual private cloud (VPC) and in a subnet by using CIDR notation. A /16 block provides 65,536 IPV4 addresses. A /24 block provides 256 address.

**Amazon Virtual Private Cloud: -** Amazon Virtual Private Cloud (Amazon VPC) lets you provision a logically isolated section of the AWS cloud where you can launch AWS resources in a virtual network that you define. You have complete control over your virtual networking environment, including the selection of your own IP address range, the creation of subnets, and the configuration of route tables and network gateways. You can use both IPV4 and IPV6 in your VPC for secure and easy access to resources and applications. You could create up to five non-default VPCs per AWS account per Region.

**Subnets: -** A VPC spans all the Availability Zones in the region. After creating a VPC, you can add one or more subnets in each Availability Zone. When you create a subnet, you specify the CIDR block for the subnet, which is a subset of the VPC CIDR block. Each subnet must reside entirely within one availability zone, and it can’t span Availability Zones.

**Default VPC: -** In each Region, AWS will provision a default VPC. This VPC has a /16 IPV4 CIDR address block of 172.31.0.0/16. This provides 65,536 private IPv4 addresses. In addition, there will be a /20 subnet that is created for each Availability zone in the Region, which provides 4,096 addresses per subnet, with a few addresses reserved for AWS usage. The route table that is associated with the default VPC will have a public route, which in turn is associated with a provisioned internet gateway.

**Amazon Elastic Block Store: -** Amazon EBS provides persistent block storage volumes for use with Amazon EC2 instances in the AWS cloud. Each Amazon EBS volume is automatically replicated inside an Availability zone to protect you from component failure, which offers high availability and durability. Amazon EBS volumes offer the consistent and low-latency performance that you need to run your workloads.

Amazon EBS provides a range of options that allow you to optimize storage performance and cost for your workload. These options are divided into two major categories: SSD-backed storage for transactional workloads, such as databases and boot volumes(performance depends primarily on IOPS), and hard disk drive (HDD)-backed storage for throughput intensive workloads, such as MapReduce and log processing(performance depends primarily on MB/s).

The Elastic Volume features of Amazon EBS allows you to dynamically increase capacity, tune performance and change the type of live volumes with no downtime or performance impact. This allows you to easily right-size your deployment and adapt to performance changes.

**Amazon Simple Storage Service: -** Amazon S3 stores data as objects within resources that are called buckets. You can store as many objects as you want within a bucket, and you can write, read, and delete objects in your bucket. Objects can be up to 5 TB in size.

You can control access to both the bucket and the objects (who can create, delete, and retrieve objects in the bucket for example), and view access logs for the bucket and its objects. You can also choose the AWS Region where a bucket is stored to optimize for latency, minimize costs, or address regulatory requirements. With Amazon S3, you pay only for what you use.

**Amazon Elastic File System: -** Amazon EFS provides simple, scalable, elastic file storage for use with AWS Cloud services and on-premises resources. It is straightforward to use, and it offers a simple interface that allows you to create and configure file systems quickly and easily.

Amazon EFS is designed to provide massively parallel shared access to thousands of Amazon EC2 instances. This enables your applications to achieve high levels of achieve high levels of aggregates throughput and IOPS that scale as a file system grows, with consistent low latencies.

When Amazon EFS file system is mounted on Amazon EC2 instances, it provides a standard file system interface and file system access semantics, which allows you to seamlessly integrate Amazon EFS with your existing applications and tools. Multiple Amazon EC2 instances can access an Amazon EFS file system at the same time, thus allowing Amazon EFS to provide a common data source for workloads and applications that run on more than one Amazon EC2 instance.

**AWS Module 3**

**Amazon Relational Database Service: -** Amazon RDS makes it straightforward to setup, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while automating time-consuming administration tasks such as provisioning hardware, setting up the database, patching, and making backups.

Amazon RDS currently supports six database engines:

* Amazon Aurora.
* PostgreSQL
* MySQL
* MariaDB
* Oracle
* Microsoft SQL server.

You can use the AWS Database Migration Service (AWS DMS) to quickly and securely migrate your databases to AWS.

**Amazon DynamoDB: -** Amazon Dynamo DB is a fast and flexible NoSQL database service for applications that need consistent, single-digit millisecond latency at any scale. It’s a fully managed cloud database, and it supports both document and key-value store models. Its flexible data model, reliable performance, and automatic scaling of throughput capacity make it a great fit for mobile, web, gaming, advertising technology (ad tech), Internet of Things (IOT), and many other applications.

**AWS Module 4**

**Amazon CloudWatch: -** Amazon CloudWatch is a monitoring service for AWS Cloud resources and the applications that you run on AWS. You can use Amazon CloudWatch to collect and track metrics, collect and monitor log files, set alarms, and automatically react to changes in your AWS resources.

**Amazon CloudWatch Events: -** Amazon CloudWatch Events delivers a near real-time stream of events that describe changes in AWS resources. Using simple rules that you can quickly set up, you can match events and route them to one or more target functions or streams. CloudWatch Events becomes aware of operational changes as they occur.

**Amazon CloudWatch Logs Metrics: -** You can use Amazon CloudWatch Logs to monitor, store, and access your log files from Amazon EC2 instances, AWS CloudTrail, Amazon Route 53, and other sources. You can then retrieve the associated log data from CloudWatch Logs.

You can collect metrics from servers by installing the CloudWatch agent on the server. You can install the agent on both Amazon EC2 instances and on-premises servers, and on servers that run either Linux or Windows Server.

**Elastic Load Balancing: -** ELB automatically distributes incoming application traffic across multiple targets, such as Amazon EC2 instances, containers, and IP addresses. It can handle the varying load of your application traffic in a single Availability Zone or across multiple Availability Zones.

ELB offers three types of load balancers that all features the high availability, automatic scaling, and robust security that are necessary to make your applications fault-tolerant.

An application Load Balancer operates at the request level (Layer 7), routing traffic to targets—such as EC2 instances, microservices and containers—within Amazon VPC, based on the content of the request. It’s ideal for the advanced load balancing of Hypertext Transfer Protocol (HTTP) and Secure HTTP (HTTPS) traffic.

A Network Load Balancer operates at the connection level (Layer 4), routing connections to targets—such as Amazon EC2 instances, microservices, and containers—within Amazon VPC, based on IP protocol data. It’s ideal for load-balancing Transmission Control Protocol (TCP) traffic.

This classic Load Balancer provides basic load balancing across multiple Amazon EC2 instances, and it operates at both the request level and the connection level.

**Amazon EC2 Auto Scaling: -** Amazon EC2 Auto Scaling helps you maintain application availability, and it allows you to dynamically scale your Amazon EC2 capacity up or down automatically according to conditions that you define. You can use Amazon EC2 Auto scaling for fleet management of Amazon EC2 instances, which can help maintain the health and availability of your fleet, and ensure that you are running your desired number of Amazon EC2 Instances. You can also use Amazon EC2 Auto Scaling to dynamically scale Amazon EC2 instances. Dynamic scaling automatically increases the number of Amazon EC2 instances during demand spikes to maintain performance decrease capacity during lulls, which can help reduce costs. Amazon EC2 Auto Scaling is well-suited to applications that have stable demand patterns, or applications that experience hourly, daily, or weekly variability in usage.

**AWS Module 5**

**Amazon Shared Responsibility Model: -** Security and compliance are shared responsibilities between us and the customer. This shared model can help relieve a customer’s operational burden because we operate, manage, and control the components from the host operating system and virtualization layer down to the physical security of the facilities where the service operates. The customer is responsible for –and manages—the guest operating system (including updates and security patches) and other associated application software, in addition to the configuration of the AWS-provided security group firewall. Customers should carefully consider the services that they choose because their responsibilities will vary depending on the services that they use, the integration of those services into their IT environment, and applicable laws and regulations. The nature of this shared responsibility also provides the flexibility and customer control that permits the deployment. This differentiation of responsibility is commonly referred to as Security of the Cloud versus Security in the Cloud.

**AWS responsibility**

*Security of the Cloud:* We are responsible for protecting the infrastructure that runs all of the services that are offered in the AWS Cloud. This infrastructure is composed of the Hardware, software, networking, and facilities that run AWS Cloud services.

**Customer responsibility**

*Security in the Cloud:* Customer responsibility will be determined by the AWS cloud services that a customer selects. This determines the amount of configuration work the customer must perform as part of their security responsibilities.

**Cost Management on AWS**

AWS uses a pay-as-you-go pricing model. This allows you to easily adapt to changing business needs without over-committing budgets, and it improves your responsiveness to changes.

As mentioned in the segment on Regions, AWS pricing varies by Region. When you explore costs, make sure that you select the appropriate Region.

Pricing for each service can be found on the AWS website under the specific service listing.

**AWS Pricing Calculator: -** AWS has announced a new pricing tool, The AWS Pricing Calculator. This new tool, which is currently in beta, can be used to calculate Amazon EC2 and EBS pricing.

**AWS Cost Explorer: -** AWS Cost Explorer lets you visualize, understand, and manage your AWS costs and usage over time. You can create custom reports (including charts and tabular data) that analyze costs and usage data, both at a high level (e.g., total costs and usage across all accounts) and for highly specific requests(e.g.,m2.2xlarge costs within account Y that are tagged project:secretProject).

**AWS Trusted Advisor: -** AWS Trusted Advisor is an online resource to help you reduce costs, increase performance, and improve security by optimizing your AWS environment. Trusted Advisor provides real-time guidance to help you provision your resources by following our best practices.