**AWS Cloud Practitioner Essentials**

**MODULE 1: INTRODUCTION TO AMAZON WEB SERVICES**

**Introduction**

Client-server model

Client: something that the person interacts with to make requests

**Cloud Computing**

Cloud computing is the on-demand delivery of IT resources over the internet with pay-as-you-go pricing

Cloud-based Deployment

* Run all parts of the application in the cloud
* Migrate existing applications to the cloud
* Design and build new applications in the cloud
* In a cloud-based deployment model, you can migrate existing applications to the cloud, or you can design and build new applications in the cloud. You can build those applications on low-level infrastructure that requires your IT staff to manage them. Alternatively, you can build them using higher-level services that reduce the management, architecting, and scaling requirements of the core infrastructure.

On-premises Deployment

* Deploy resources by using virtualization and resource management tools
* Increase resource utilization by using application management and virtualization technologies
* On-premises deployment is also known as a *private cloud deployment*. In this model, resources are deployed on premises by using virtualization and resource management tools.

Hybrid Deployment

* Connect cloud-based resources to on-premises infrastructure
* Integrate cloud-based resources with legacy IT applications
* In a hybrid deployment, cloud-based resources are connected to on-premises infrastructure.

Benefits of cloud computing

* Trade upfront expense for variable expense
* Stop spending money to run and maintain data centers
* Stop guessing capacity
* Benefit from massive economies of scale
* Increase speed and agility
* Go global in minutes

**MODUEL 2: COMPUTE IN THE CLOUD**

**Introduction**

EC2 runs on top of physical host machines managed by AWS using virtualization technology. When you spin up an EC2 instance, you aren't necessarily taking an entire host to yourself. Instead, you are sharing the host with multiple other instances, otherwise known as virtual machines. And a hypervisor running on the host machine is responsible for sharing the underlying physical resources between the virtual machines. This idea of sharing underlying hardware is called multitenancy. The hypervisor is responsible for coordinating this multitenancy and it is managed by AWS. The hypervisor is responsible for isolating the virtual machines from each other as they share resources from the host. This means EC2 instances are secure. Even though they may be sharing resources, one EC2 instance is not aware of any other EC2 instances also on that host. They are secure and separate from each other.

**Amazon EC2 Instance Types**

Combinations of CPU, GPU, Memory, Network etc

* General purpose
* Compute optimized
* Memory optimized
* Accelerated computing
* Storage optimized

**Amazon EC2 Pricing**

* On-Demand: ideal for short-term
* Saving Plans: commit to a consistent amount of compute usage for a 1-year or 3-year term
* Reserved Instances: billing discount applied to the use of On-Demand; Standard Reserved and Convertible Reserved Instances for a 1-year or 3-year term, and Scheduled Reserved Instances for a 1-year term
* Spot Instances: ideal for workloads with flexible start and end times; use unised Amazon EC2 computing capacity (if capacity is no longer available for Spot Instance increases, the instance is iterrupted)
* Dedicated Hosts: physical servers with Amazon EC2 instance capacity that is fully dedicated to your use

**Scaling Amazon EC2**

Beginning with only the resources you need and designing your architecture to automatically respond to changing demand by scaling out or in

Amazon EC2 Auto Scaling

**Directing traffic with Elastic Load Balancing**

Where to route the requests?

Elastic Load Balancing

* Regional construct
* A load balancer acts as a single point of contact for all incoming web traffic to your Auto Scaling group
* Stop the flow and terminate all existing process. Then finally terminates extra instances.