**AWS Elastic Compute Cloud (EC2)**

* provides resizable compute capacity in the Cloud
* designed to make web-scale cloud computing easier
* a true virtual computing environment
* launch instances with a variety of operating systems
* run as many or few systems as you desire

**AWS EC2 Characteristics**

* Completely Controlled
* Flexible
* Web-Scale
* Reliable
* Secure
* Inexpensive
* Easy to Start

**AWS EC2 Use Cases**

* running a high performance game server
* running video rendering servers
* supporting enterprise web applications
* running database instances
* basically, any use case that can be applied to a CPU and an OS

**AWS EC2 Concepts**

**Regions**

* the geographical area where Amazon EC2 will launch the instances that you create
* choose a region to optimize latency, minimize costs, or address regulatory requirements

**Availability Zones (AZs)**

* distinct locations that are engineered to be insulated from failures in other AZs
* provides inexpensive, low latency network connectivity to other AZs in the same Region
* Regions have between 2 and 5 AZs

**Instances**

* wide selection of instance types, each optimized to fit different use cases
  + types comprise varying combinations of CPU, memory, storage, and networking capacity

**AWS EC2 Fundamentals**

* available in different locations ([docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-regions-availability-zones.html](http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-regions-availability-zones.html))

* extensive list of supported operating systems & software ([aws.amazon.com/ec2/faq](http://aws.amazon.com/ec2/faq))
  + RedHat Linux, Windows Server, SuSE Linux, Ubuntu, Fedora, Debian, Cent OS, etc.

* integrated with other AWS Services
  + AWS Elastic Block Store (EBS)
  + AWS CloudWatch
  + AWS Virtual Private Cloud (VPC)
  + AWS Identity and Access Management (IAM)

* other AWS services are dependent upon EC2:
  + Elastic MapReduce (EMR)
  + Relational Database Service (RDS)
  + RedShift

* purchasing options for cost optimization
  + On-Demand Instances
    - pay for compute capacity by the hour with no-long term commitments or upfront payments

* Reserved Instances
  + provides a significant discount (up to 75%) compared to On-Demand Instance pricing
  + instances are reserved for a predetermined amount of time (1-3 years)

* Spot Instances
  + purchase compute capacity with no upfront commitment and at hourly rates usually lower than the On-Demand rate

* import and export virtual machines
  + supported by the AWS EC2 VM Import/Export service
  + allows users to easily import virtual machine images (VMIs) from an existing environment to EC2 instances
  + allows users to easily export VMIs from the EC2 instance back to the on-premises environment
  + great feature for disaster recovery planning

* choice of instance families with differing resource ratios
  + several instance types of varying combinations of computing power, memory allocation, storage capacity, and network capacity:
    - General Purpose: ***M4***
    - Compute Optimized: ***C4***
    - Memory Optimized: ***R3***
    - GPU Optimized: ***G2***
    - Storage Optimized: ***D2***
    - IO Optimized: ***I2***
    - Low cost, burst-able performance: ***T2***

* Example:

**d2.xlarge (Storage-Optimized)**

* accessible via AWS CLI
  + Example command (create an EC2 instance using flags):

**aws ec2 run-instance \**

**--image-id ami-c229c0a2 \**

**--instance-type m3.medium \**

**--count 5 \**

**--security-group-ids sg-9d5c2dfa \**

**--subnet-id subnet-608ebb14 \**

**--key-name ec2-example-demo**

* accessible via SDK
  + Python, C++, Java, Node.js, .NET, Go, and many more..

**AWS EC2 Storage & Networking**

**Data Storage Options**

* ***Instance Storage***
  + storage location is integrated into the machine that is hosting the EC2 instance

* does not survive instance start/stop operations, or instance termination

* type and amount of storage differs by instance type

* ***Elastic Block Store (EBS)***
  + network-attached block devices that can be connected to EC2 instances

* storage types include:
  + Magnetic
  + General Purpose - SSD
  + Provisioned Input/Output Per Second (IOPS) - SSD

* lifecycle is independent of the instance lifecycle
  + this can be configured to have data deleted upon instance termination if desired

* automatically replicated in the AZ which they are created in
  + no need for RAID storage

* can be backed by snapshots (which are stored in S3 buckets)
  + can be used to create replications of instances
  + can be shared across AWS accounts
  + can be copied across AWS Regions

* data stored at rest on the volume, disk I/O, and snapshots are all encrypted
  + encryption occurs on the servers that host the EC2 instance, providing encryption of data-in-transit from the instance to the EBS storage
  + uses the AWS Key Management Service (KMS), but allows for the create of a Customer Master Key (CMK)
    - gives the ability to create, rotate, disable, define access controls, and audit the encryption keys

**Networking**

***Amazon Virtual Private Cloud (VPC)***

* a virtual network in your own ***logically isolated area*** within the AWS cloud, populated by infrastructure, platform, and application services that share common ***security*** and ***interconnection***.

* Components (not required to know for QC)

* Elastic Network Interface (ENI)
  + can be attached to instances

* have public IP addresses

* used to float public IP addresses to different instances

* Subnetting
  + allows users to define network topology
  + allows for more finely-grained access control

* Network Access Control List
  + a stateless filter that will filter traffic flowing in and out user-created subnets

* Route Table
  + allows for routing between subnets

* Internet Gateways
  + optional

* Virtual Private Gateway
  + optional

* Route 53 Private Hosted Zones
  + Optional

***Elastic Load Balancing***

* a highly-scalable, load-balancing service that provides regional traffic management capabilities

* allows users to dynamically and intelligently distribute traffic across a fleet of EC2 instances

* used commonly for web servers, API servers, and variety of other use cases as well

* features:
  + timeout configuration
  + connection draining
  + cross-zone load balancing

* resource:
  + [aws.amazon.com/elasticloadbalancing/](http://aws.amazon.com/elasticloadbalancing/)

**AWS EC2 Monitoring, Metrics, & Logs**

***AWS CloudWatch***

* performs the core of monitoring and metrics tasks over EC2s

* a monitoring service AWS cloud resources and the applications ran on AWS

* used to:
  + collect and track metrics
  + collect and monitor log files
  + set alarms
    - can be delivered using AWS Simple Notification Service (SNS)
    - can be handled automatically using:
      * auto scaling action (scale up or down)
      * EC2 actions:
        + Recover
        + Stop
        + Terminate
* resource:
  + [aws.amazon.com/cloudwatch](http://aws.amazon.com/cloudwatch)

* CloudWatch Logs
  + used to monitor applications and systems using log data
  + stored in a highly durable storage with set retention policies
  + allows users the ability to access files via Web, CLI, or SDK

**AWS EC2 Security & Access Control**

***Access Credentials***

* access key and secret key used to authenticate when accessing AWS APIs
  + access key: identifies the requester
  + secret key: signs requests

***Key Pairs***

* public key and private key used to authenticate when accessing an EC2 instance
  + public key: injected by AWS into the instance via its instance metadata
  + private key: is kept by the user

***IAM Roles***

* strongly recommended
* used to pass access credentials to EC2 instances

**AWS EC2 Deployment**

Unit of deployment: ***Amazon Machine Images (AMIs)***

* Amazon-Maintained AMIs
  + a set of Windows and Linux images
  + kept up to date by Amazon in each region

* Community-Maintained AMIs
  + images published by other AWS users
  + managed and maintained by Marketplace partners

* User Custom
  + AMIs that the user has created from EC2 instances
  + can be kept private or shared with other accounts