Cert Prep: 4 Compute Services

1 Compute Services Design

EC2 Overview

Elastic Compute Cloud (EC2)

Virtual machine in the cloud |Pay as you go model |Integrates with storage, networking, and security
It is FAST!!!! Goes from idea to inception in minutes. While older legacy models take weeks.

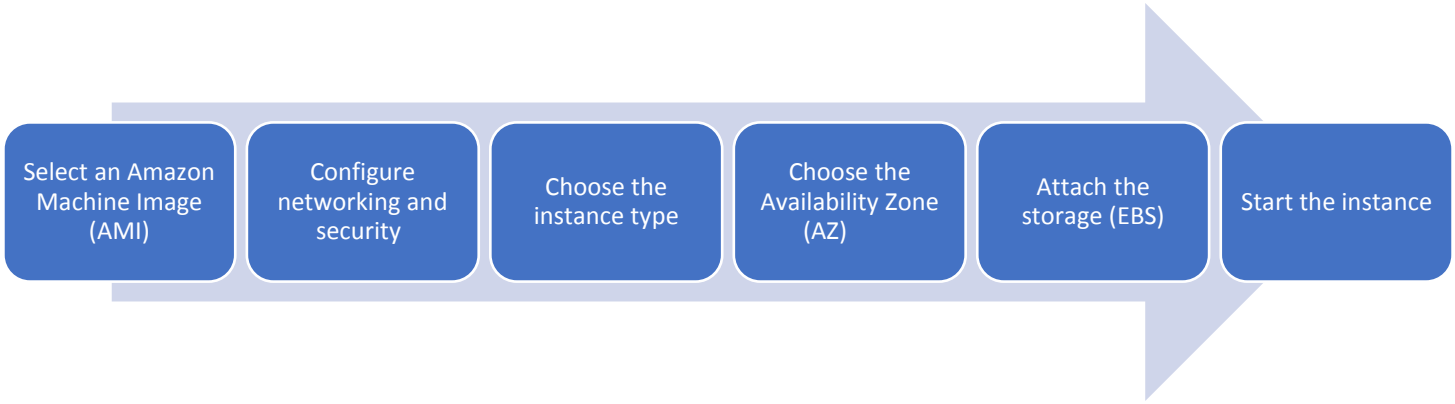
Supported Operating Systems

Windows 2003 R2 through 2016
Amazon Linux |Debian | SUSE | CentOS | Red Hat Enterprise Linux | Ubuntu

EC2 Benefits

Time to market | Scalability | Reliability | Security | Control | Services Integration | Cost Efficiency

EC2 Deployment Six(6) Steps



EC2 Overview Quick Review

Elastic Compute Cloud (EC2) is a service that allows you to run virtual servers (machines) in the cloud.
You can implement EC2 instances in the free tier and instances that cost hundreds or thousands of dollars to run each month.
Common Operating Systems (OS) are supported including Windows and Linux.
Deploy an EC2 instance by selecting an Amazon Machine Image (AMI) and then deploying it using appropriate configuration settings.

EC2 Instance Types

General Purpose

T2 | M5 | M4 | M3 |
Provides a balance of memory and network resources
T2 provides burst performance
    Credits accrue during idle times | Credits used to burst performance
M5 | M4 |M3 | have no burst option
    Good for development, staging, etc.

Compute Optimized

C5 | C4 | C3 | Useful for CPU-intensive applications
Media coding | Intensive batch jobs | Many concurrent users | Gaming Servers | Anything computer-intensive

Memory Optimized

X1e |X1 | R4 | R3
Useful for high memory requirements
Processing large data sets | In-memory databases | Big data processing

Storage Optimized

H1 |I3 | D2
Useful for high sequential read/writes to local storage
Relational databases | Data warehousing | Image storage and processing

Advanced Computing

P3 | P2 | G3 | F1
Useful for specialty hardware compute requirements
(GPU) Graphics Processing Unit |(FPGA) Field-Programmable Gate Array

EC2 Instance Types Quick Review

The T2, M5, M4, and M3 classes provide a balance of memory and network resources
The C5, C4, and C3 classes are useful for CPU-intensive applications
The X1e, X1, R4, and R3 classes are best for high memory demand instances
The Hi, I3, and D2 classes are optimized for storage access
The P3, P2, G3, and F1 classes support specialty hardware such as GPs and FPGAs

EC2 Pricing

Pricing Categories

On-Demand | Charged for usage time at a flat rate | Billed in 60-second increments rounded up |
1 Have the instance running
2 Have the appropriate storage
3 Have the appropriate network throughput
Reserved | Hours | Reserve usage minimum 1 year | Can be less expensive than on-demand
Spot | Bid on unused compute time | Up to a 90% discount over on-demand **frequently at night *must be flexible*

EC2 Pricing Quick Review

On-demand pricing incurs charges based on usage and is billed in 60 second increments.
Reserved pricing is based on at least a 1 year reservation and can be less expensive than on-demand when estimates are correct
Spot pricing is the least expensive because you are using unused compute time

EBS and EC2

EBS | Persistent block storage | Requires an EBS-optimized instance | Magnetic or SSD - General purpose or PIOPS

EBS and EC2 Quick Review

(EBS) Elastic Block Store provides persistent block storage (virtual hard drives) to EC2 instances
To use EBS, an EBS- optimized instance should be deployed
Storage can magnetic or SSD; the former is slower, but less expensive, the latter is faster, but more expensive

Computer Services Design Chapter Quiz

1 What is a simple definition of EC2?

Virtual machines in the cloud |Virtual drive in the cloud | virtual databases in the cloud | virtual interfaces in the cloud

EC2 (Elastic Compute Cloud) provides for virtual machines running in the AWS cloud.

2 While the ___ instance family is optimized for processing large data sets, the ___ family is optimized for intensive batch jobs.

R4:C5 |P3:F1 | C4:H1 | T2:R3 R4:C5 the R family is optimized for memory and the C family is optimized for computer

3 What instance type is used when heavy graphics processing is required?

P3 |T2 | H1| X1e P3 is for advanced computing and can include special hardware for graphics processing.

4 Which pricing category is the most expensive?

On-Demand | Spot | Reserved On-Demand must be available without knowing for how long, they are usually the most expensive.

5 How is EC2 pricing structured?

On-Demand | Flat fee at install | Flat monthly rate | Flat annual rate

On-Demand, rather than using flat rates, AWS EC2 instances are billed for usage time. If you shut down an instance, but do not delete it, you will not be charged for it as it is not active. However, it will be there if/when you later require it.

6 Which statement is true regarding EBS?

It is required to guarantee IOPS | It is available for all instance types | It may be offered for free in some plans | It always requires SSD storage.

It is required to guarantee IOPS. IOPS can only be guaranteed on EBS-optimized instances.

7 If you want to save money and do not require fast storage access, what should you do?

Use magnetic EBS volumes | Use SSD EBS volumes |

Use S3 for the instance drives instead of EBS volumes | Use EFS for the instance drives instead of EBS volumes |

Use magnetic EBS volumes. EC2 instances boot from EBS volumes and not EFS or S3; therefore, an EBS volume must be used. Magnetic volumes are slower, but less expensive than SSD volumes.

2 Compute Services Implementation

Lab: Launching an EC2 Linux Instance

The AWS marketplace has many AMIs of Linux

When selecting a Linux AMI, be sure to select the right distribution and feature set

You will access Linus instances using SSH

Lab: Configuring an EC2 Linux Instance

The Access and Secret key will be required to access a Linux Instance through SSH

After connecting with SSH, you can update the instance and perform other administrative tasks

Always take care to protect the keys used to access your EC2 Instance

Lab: Setting up an EC2 Windows Instance

It is important to remember that Windows Instances require more processing power than Linux Instances in most cases

After launching a Windows Instance, connect to it with Remote Desktop port 3389 and not SSH port 22

You will need an RDP configuration file downloaded from AWS to access the Windows Instance

Shared Tenancy

Multiple customers share the time and space on the physical machine. ***Shared Tenancy is the Default***

Default Instance behavior

Shared Tenancy Considerations

Pros	Cons
Reduced costs	Lower performance
Simpler deployment	Less control

Shared Tenancy Quick Review

The shared tenancy model indicates that multiple instances from multiple customers will be on the same hardware
Shared tenancy is the default behavior of an instance
Using shared tenancy can reduce costs, but it may not comply with internal security policies

Dedicated Hosts

Physical machines | Run the Virtual Machine | Used by one customer | Must be explicitly configured | Not available in free tier

Dedicated Host Considerations

Pros	Cons	
More accurate licensing management	It costs more	***So bring your own licenses****
More detailed reporting		
Compliance management		
Determine host placement during Instance restarts		

Dedicated Hosts Quick Review

Dedicated Hosts are used by one AWS customer
Provides physical machine separation from other Instances
May allow for the installation of some applications that have licenses bound to the hardware

Dedicated Instances

Runs on a physical machine
Only Instance running on that machine | On restart, may be moved
Used by one customer |Must be explicitly configured |Not available in the free tier

Pros	Cons
Runs on hardware dedicated to the customer	Less accurate licensing management
Provides performance advantage of a dedicated host	

Dedicated Instances Quick Review

Dedicated Instances run on a single machine in isolation
They may be moved to another physical machine on restart
Dedicated Instances must be explicitly configured
The free tier services do not allow for Dedicated Instances

AMI Virtualization

Blueprint with server configuration details | Similar to localized imaging solutions

(AMI)Amazon Machine Image

The term “instance” indicates the use of the AMI | All instances are created from an AMI
Sources: Amazon (free) AWS Marketplace (free/$) Community (free)

AMI Launch Permissions

Who can launch an Instance of an AMI? Public | Anyone Explicit | Specified Implicit | Owner
Must be set Defaults to implicit

AMI Creation

Use existing AWS AMIs | Customize existing AMIs | Create from scratch | Use from other public sources **be cautious**

HVM AMIs

Hardware Virtual Machine (HVM)

AMIs fully virtualizes the hardware | Requires hardware assisted virtualization

PV AMIs

Paravirtual (PV)

Run on hosts without specific support for virtualization | Doesn't perform as well as HVM AMIs

Instance Root Volume

Contains the boot sector | Boot sector initiates the boot loader | Boot loader launches the OS

Instance Root Volume

Instance store-backed AMI	EBS-backed AMI
Root volume is stored in S3	Root volume stored in an EBS volume
No support for the stop action	Support for the stop action
On failure, data in the instance store is lost	On failure, data in the EBS volume is not lost

AMI Virtualization Quick Review

An (AMI) Amazon Machine Image is a blueprint including a machine image and details for configuration.
Each Instance (AWS term) is an instance (general term) for an AMI
Free AMIs are available from Amazon and the community
Paid AMIs are available from the AWS marketplace

Compute Services Implementation Chapter Quiz

1 One difference between dedicated instanced and dedicated Hosts is that Dedicated Instances _____.
May move between hosts | can be on the free tier | run on dedicated hardware | can be used by different customers
Restarting a Dedicated Instance may move it to a different physical machine.

2 What is unique about a Dedicated Instance?
It is the only Instance running on a physical machine | It is dedicated to the organization that implemented it
It cannot access the internet | It cannot access other VPCs
A Dedicated Instance runs on a physical machine that is specific to it. No other customers use the same physical machine (host).

3 You create your own AMI and do not specify its launch permission. Who will be able to launch it?
Only you | no one | everyone in your organization | everyone with an AWS account
The default launch permission is implicit if not specified, and only the owner is allowed to launch it.

4 By default, who can launch an Instance of an AMI?
The owner | anyone | all accounts in an enterprise | Amazon only
Be default, only the owner can launch an instance of an (AMI)Amazon Machine Image. However, it can be made public or shard with specific accounts.

Compute Services Implementation Chapter Quiz cont'd

5 You launch a new Instance from a non-encrypted Linux AMI. What volume(s) can you choose to encrypt?
Only addition non-root volumes. The root volume cannot be set to encrypted because it is already non-encrypted, but other volumes can. Only the root volume | any volume | only magnetic volumes

6 How do you install Linux on an AWS instance? ***Launch an AMI that already has Linux installed***
Hundreds of Linux-based AMI's exist and you can start with that. Alternatively, you can create a local VM with Linux installed as you desire and import it as an AMI into AWS. Linux is not supported in AWS | Launch an empty AMI and then install Linux with a Telnet connection

7 What is the default username used when connecting to an Amazon Linux Instance over SSH?
ec2-user. This is the standard username that is used across most EC2 services. ppk | pem | ubuntu

8 What free tool can you use to create an SSH connection to an AWS Linux Instance from Windows?
PuTTY a free terminal emulator that supports SSH and can be used to connect to a Linux instance.
Telnet | FileZilla | CMD.exe

9 When launching a Windows instance with a new security group, what rule will this group have by default?
(RDP) Remote Desktop Protocol will be added as a default rule type. IAM | no rules | SSH

10 What do you use to connect to an AWS Windows Instance from your computer to control the desktop interface?
(RDS) is the Microsoft protocol that works with Windows servers, and you use this to remote into the AWS Windows Instance and control the desktop. Telnet | SSH | Teamviewer

11 What is one reason to avoid using shared tenancy? ***It reduces performance.***
Due to the shared underlying resources, shared instances generally have lesser performance than dedicated instances.
It complicates deployment | it does not offer a region choice | it is more expensive

12 If you do not use a dedicated host nor a dedicated instance, what normal instance hosting do you use?
Shared Tenancy is the default instance hosting method. *More than one customer's instance* may be on the same physical machine at the same time. Single Host | Single Server | Single Instance

13 What is one reason to avoid using dedicated hosts? ***Higher costs*** because you have a dedicated physical server.
Reduced Hardware Control | Lesser Compliance | Difficult license management

14 What specific software constraint may force you to use a dedicated host in AWS?
Per-processor license(s), will ultimately have to go away in the age of the cloud, but for now, they must often be accommodated.
The need for more than 4GB RAM | The need for more than one network interface | The need for more than 1GHz processing power

3 Compute Services Management

Instance Management

Launching Instances

Bootstrapping | Providing code to be run on an Instance at launch
VM Import/Export | Importing existing virtual machines into EC2

Instance Metadata

Security Groups | Instance ID | Instance Type | AMI base of the Instance

Instance Management

Changing Instance type
Stop the Instance | Change the type
Change Security Groups on the fly
Activate termination protection

Instance Management Quick Review

Bootstrapping allows you to launch an instance and have a configuration script run on first launch
You can import virtual machine images from your local environment into EC2 for deployment in the cloud
Instances have metadata such as the Instance ID, Instance type, and AMI base of the instance

Lab: Connecting to Instances

Lab: Connecting to Instances Quick Review

You can initiate a connection to a Windows or Linux Instance from the Management Console
You can save connection configuration information in a terminal emulator package like PuTTY
The RDP configuration file has everything in it required to connect a Windows Instance be sure to protect this file

Working with Security Groups

Security Groups

Limited to five (5) per Instance
Can layer security groups
Instances receive the default security group for the VPC
Default setting | Other Security Group may be attached | Default Security Group may be detached

Security Groups vs ACLs

Security Group	**Network ACL**
Operates at the ***Instance*** level	Operates at the ***Subnet*** level
Supports allow rules only	Supports allow and deny rules
Is stateful: Return traffic is automatically allowed, regardless of any rules	Is stateless: Return traffic must be explicitly allowed by rules
We evaluate all rules before deciding whether to allow traffic	We process rules in number order when deciding whether to allow traffic
Applies to an instance only if someone specifies the security group when launching the instance or associates the Security Group with the Instance later on	Automatically applies to all instances in the subnets it's associated with (therefore, you don't have to rely on users to specify the Security Group

Security Group Constraints

Only “allow” rules are permitted
Separate inbound and outbound rules are used
Stateful
    By default, no inbound traffic is allowed without request
    By default, all outbound traffic is allowed
By default, Security Groups are only bound to the primary network interface
    Can be bound to other network interfaces, including ENIs

Security Group Constraints Quick Review

A maximum of five (5) Security Groups can be associated with an Instance
Instances automatically receive the default Security Group for the VPC in which they are deployed
Security Groups support only allow rules
Security Groups are stateful and the default is to allow no inbound originating traffic

Labs: Working with Security Groups

Labs: Working with Security Groups Quick Review

Remember that Security Groups are associated with Instances
A Security Group can be attached to more than one Instance
It is important to get the order of rules correct or the desired permissions will not be accomplished

Advanced EC2 Management

Advanced EC2 Management Quick Review

Resource optimization in the AWS Cost Management console gives recommendations to help save money
Host recovery restarts EC2 Instances when a problem is detected or when a new host is available
Traffic monitoring copies network traffic from an elastic network interface (ENI) of an EC2 Instance and sends it wherever you want it to go

AWS Batch Quick Review

AWS Batch enables running batch processing of scripts within the AWS environment
Batch jobs can be run at any scale
A job is a unit of work executed by an AWS batch, or basically any script that is run in the AWS environment

(ECS) Elastic Container Service

ECS Features

No Virtual Machine Builds required
Uses Amazon Fargate to automatically build environments
Can use EC2 Instances for more control

Container Usage

Web Server | Application Server | Message Queue Server | Each of the backend worker processes

Container Usage Quick Review

The Elastic Container Service (ECS) can be used to launch apps in AWS without deploying instances directly
A multi-tier application can use separate containers for each tier of the application
The concept of microservices is supported by ECS

Elastic Beanstalk Environment

Elastic Beanstalk Environment Quick Review

Using Elastic Beanstalk, you can create a server instance with the Create New Environment wizard
You cannot change the environment tier after creating an environment
You can use application platforms such as .NET, Java, Node.js, Python, and Ruby.

Compute Service Management Chapter Quiz

1 What do both Security Group and network ACLs offer? ***Allow rules to let traffic in***.
Application to a specific instance | stateful inspection | deny rules

2 Why do you use a Security Group in AWS in relation to an instance? ***To set allow access rules for he instance***
Security Groups are a lot like a client or local firewall. They allow you to control what communications may enter the instance.
To aggregate multiple uses together for access to the instance | to automate security on multiple instances at once
To group the storage in the instance into a Single Secure Container

3 When setting up a Security Group, what will you use for the source to allow incoming traffic from anywhere?
0.0.0.0/0 this IP address range opens the rule to traffic with any source.
443 | TCP | CIDR

4 What is the default Security Group status when creating an Instance? ***Allow outbound, disallow inbound***
The default configuration is like a stateful firewall: outbound traffic and responses to it are allowed and original inbound traffic is not allowed.
Allow outbound to other Instances in the VPC only | disallow outbound allow inbound | allow inbound from within the VPC only

5 What is the name of the EC2 feature that allows you to copy network traffic going into and out of an instance, and then send it somewhere for inspection and monitoring? ***Traffic mirroring*** | port scanning | Interface replication

6 Which statements about Instance tags is incorrect? Tags define ***optional custom code that runs*** when the instance boots.
Tags consist of both a key and a value | One of the tags is the name of the Instance | Any tag can be set to show as a column in the Instance list

Compute Service Management Chapter Quiz cont'd

7 What provisioning model will an AWS Batch Job use? ***spot or on-demand.***
You specify the provisioning model when defining the job's compute environment.
Always on-demand | always dedicated | always spot

8 What term defines the provision of code to be run on an instance at launch time? ***Bootstrapping***
In AWS, providing code that runs at launch time to configure the instance is called bootstrapping.
Snap shooting | scripting | launching

9 What would you use Amazon ECS for? ***To deploy and scale container***. ECS is used to deploy, manage, and scale Docker containers.
To design and create containers | to deploy containers on EC2 Instances | to containerize an EC2 Instance

10 Which AWS service allows you to run docker containers? ***ECS is the Elastic Container Service***, and it supports docker containers, which, in turn, contain everything needed to run an application.
EFS | RDS | EBS

11 When connecting to a Windows Instance using RDP, how can you obtain the necessary password?
It is found by uploading the key pair file. The password can be extracted from the key pair PEM file that you selected at instance launch.
It is included in the RDP shortcut file | It is listed in the Instance metadata | It is the same for all Windows Instance

12 Which AWS service builds out multiple Instances and configurations to form a complete solution?
Elastic Beanstalk is used to build out entire solutions. They can include multiple Instances, databases, VPCs, subnets, and more.
RDS | ECS | AMI

13 What is needed from within AWS, and downloaded to your computer, to create an RDS connection to a Windows Instance? ***RDS configuration file*** is downloaded from AWS and used to connect to the desktop of the Windows Instance.
AMI | Shell script | Bash script