

AWS EC2

CS516 – Cloud Computing

Computer Science Department

Maharishi International University

Maharishi International University - Fairfield, Iowa



All rights reserved. No part of this slide presentation may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying or recording, or by any information storage and retrieval system, without permission in writing from Maharishi International University.

Content

- Elastic Cloud Compute (EC2)
 - Instance Types
 - IP addressing
 - Placement group
 - Pricing options
- Amazon Machine Images (AMI)
- Storage
 - Block vs File vs Object storages
 - EBS/EFS
 - Snapshots
 - Root & additional volumes
- Security groups



A Web Server

Web server is a computer that hosts a website.

In the early days of the internet, if you want a server, you need to run it yourself. You need to buy one, set it up, power it up, plug it in, connect it to the internet, and you needed to keep it running, 24/7, keep your internet connection on.

There was a lot of work involved, just to manage a server.

A Web Server in the Cloud

A cloud server is a powerful computer that you're able to access via internet.

Typically rack-mounted or put into blade chassis, so they're quite flat. A cloud server does not have a monitor that you can connect to, you have to connect into them remotely.



Lenovo System x3650 M5 Rack Server

AWS Elastic Cloud Compute (EC2)

EC2 is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale cloud computing easier for developers.

Basically, a virtual computer, very similar to the desktop or laptop you use at home, and commonly referred to as an **instance**.

You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure security and networking, and manage storage.

Amazon EC2 autoscaling enables you to scale in or out to handle changes in requirements or spikes in popularity, reducing your need to forecast traffic.

Read more: [Amazon EC2](#)

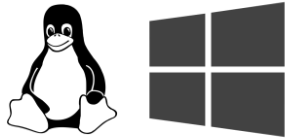
Computer and EC2 Instance



Computer



EC2 Instance



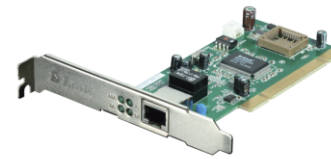
Operating System
AMIs
(Linux or Windows)



CPU
Instance Type



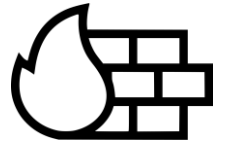
Hard Drive
EBS



Network Adapter
ENI



RAM



Firewall
Security Groups

Amazon Machine Image - AMI

Preconfigured and required to launch an EC2 instance that includes an operating system, software packages and other required settings.



Amazon Machine Image (AMI) **provides the information required to launch an instance.** You specify an AMI when you launch an instance.

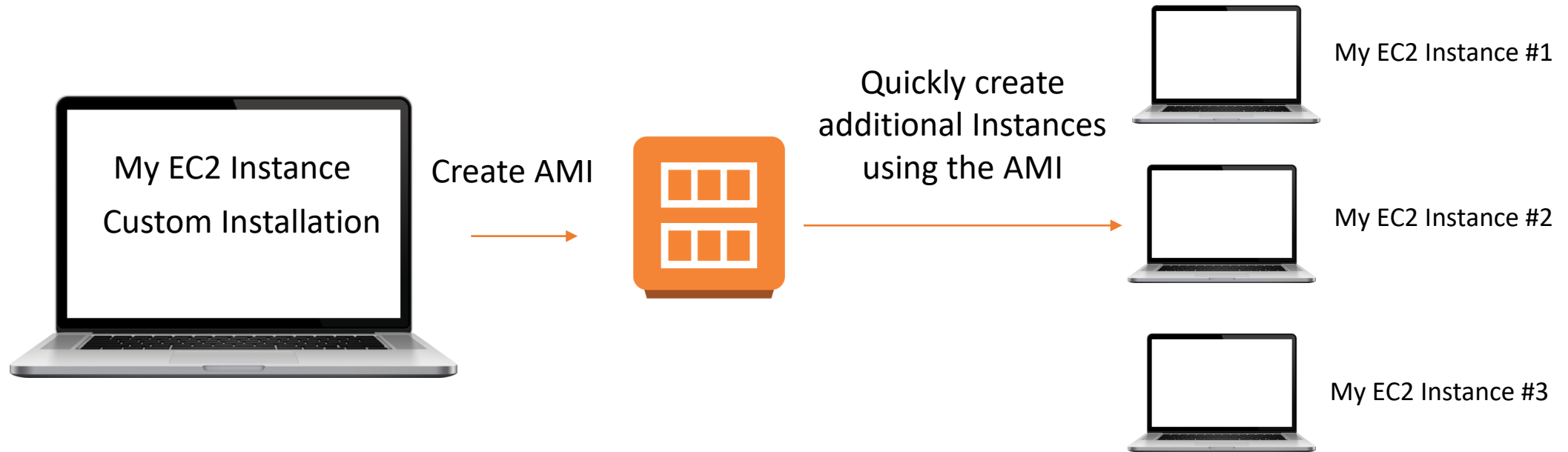
Read more: [AMI](#)

Custom AMI

Assume that you need to run your app in hundreds of servers. It is not possible to configure every single server one by one. Instead, you can create a custom AMI and use it on as many servers as you want. A custom AMIs contain:

- Operating System (OS)
- OS configuration – Installing agents for monitoring, security configs etc.
- Packages or dependencies your application needs
- Your application configuration – For example, your app needs to download a file from the internet before starting.

Understanding AMI



Instance Types

The Instance Type is the **CPU** (Central Processing Unit) of your instance.

When you launch an instance, the instance type that you specify determines the hardware of the host computer used for your instance.

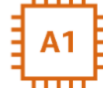





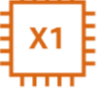



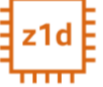

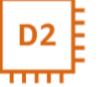
Each instance type offers different **compute**, **memory**, and **storage**, **graphic** capabilities and are grouped in instance families based on these capabilities.

Read more: [Instance Types](#)

Instance Types

- **t**=tiny, **m**=medium, **c**=compute, **r**=RAM, **x**=xtreme, **p**=pictures, **h**=HDD, **i**=IOPS, **d**=dense

| | Family | Type | vCPUs ⓘ | Memory (GiB) | Instance Storage (GB) ⓘ | EBS-Optimized Available ⓘ | Network Performance ⓘ | IPv6 Support ⓘ |
|-------------------------------------|-----------------|--------------------------------|---------|--------------|-------------------------|---------------------------|-----------------------|----------------|
| <input type="checkbox"/> | General purpose | t2.nano | 1 | 0.5 | EBS only | - | Low to Moderate | Yes |
| <input type="checkbox"/> | General purpose | t2.micro Free tier eligible | 1 | 1 | EBS only | - | Low to Moderate | Yes |
| <input checked="" type="checkbox"/> | General purpose | t2.small | 1 | 2 | EBS only | - | Low to Moderate | Yes |
| <input type="checkbox"/> | General purpose | t2.medium | 2 | 4 | EBS | | | |
| <input type="checkbox"/> | General purpose | t2.large | 2 | 8 | EBS | | | |
| <input type="checkbox"/> | General purpose | t2.xlarge | 4 | 16 | EBS | | | |
| <input type="checkbox"/> | General purpose | t2.2xlarge | 8 | 32 | EBS | | | |
| <input type="checkbox"/> | General purpose | m4.large | 2 | 8 | EBS | | | |
| <input type="checkbox"/> | General purpose | m4.xlarge | 4 | 16 | EBS | | | |
| <input type="checkbox"/> | General purpose | m4.2xlarge | 8 | 32 | EBS | | | |

| General Purpose | Compute Optimised | Memory Optimised | Accelerated Computing | Storage Optimised |
|--|--|--|--|--|
|  A1 ARM based core and custom silicon |  C4 Compute - CPU intensive apps and DBs |  R4 RAM - Memory intensive apps and DB's |  P2 Processing optimised - Machine Learning |  H1 High Disk Throughput - Big data clusters |
|  T2 Tiny - Web servers and small DBs | |  X1 Xtreme RAM - For SAP/Spark |  G3 Graphics Intensive - Video and streaming |  I3 IOPS - NoSQL DBs |
|  M4 Main - App servers and general purpose | |  z1d High Compute and High Memory - Gaming |  F1 Field Programmable - Hardware acceleration |  D2 Dense Storage - Data Warehousing |

Elastic Block Storage - EBS

EBS is a storage volume for an EC2 instance. It provides block level storage volumes for use with EC2 instances.

EBS volumes are **highly available and reliable** storage volumes that can be attached to any running instance that is **in the same Availability Zone**.

EBS volumes that are attached to an EC2 instance are exposed as storage volumes that persist **independently** from the life of the instance.

Read more: [Amazon EBS Volumes](#)



=



Amazon Elastic File System (Amazon EFS)

It is an elastic file system that lets you share file data without provisioning or managing storage.

Amazon EFS is designed to provide massively parallel **shared** access to thousands of Amazon EC2 instances.

Amazon EFS is well suited to support a broad spectrum of use cases from home directories to business-critical applications.



Read more: [Amazon EFS](#)

Block vs File vs Object storages

File storage stores data as a single piece of information **in a folder** to help organize it among other data.

Block storage takes a file apart **into singular blocks** of data and then stores these blocks as separate pieces of data. Faster than the file storage.

Object storage **is a flat structure** in which files are spread out among hardware. Unlimited scaling.

Read more: [File storage, block storage, or object storage?](#)

Input/Output Operations Per Second - IOPS

IOPS is the amount of data that can be written to or retrieved from EBS per second.

The operations are measured in KiB, and the underlying drive technology determines the maximum amount of data that a volume type counts as a single IO.

More IOPS means better volume performance (faster R/W speeds).

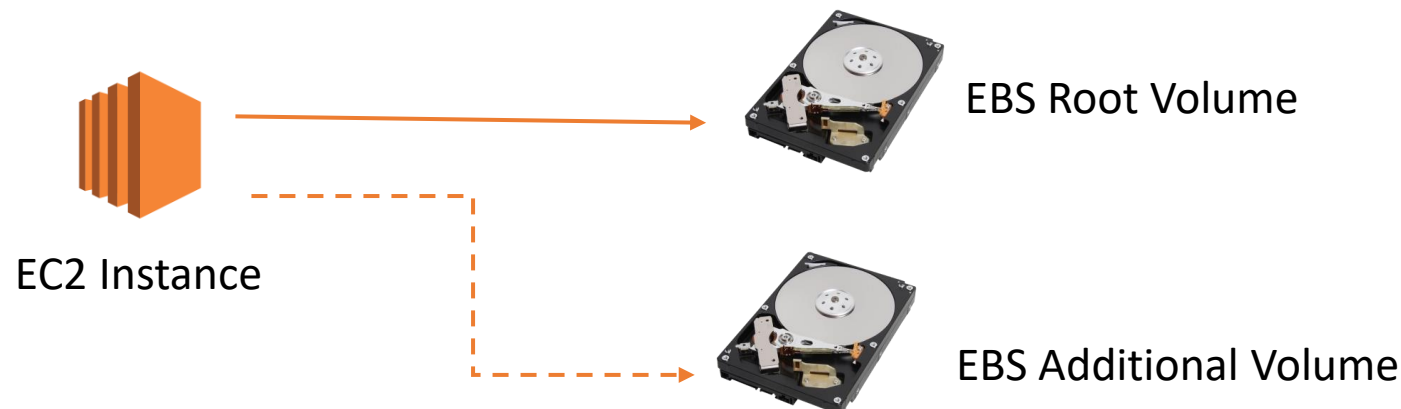
More expensive. Recommended for production environment.

Root vs. Additional EBS Volumes

Every EC2 instance must have a **root volume** that the AMI is restored.

You can add **additional EBS Volumes** to an instance at anytime. Any additional volume can be attached or detached from the instance at any time.

Additional EBS volumes are NOT deleted (the default) and you still pay when the instance is terminated whereas root volume gets deleted when the instance is terminated.



Create Additional Volume

[Volumes](#) > Create Volume

Create Volume

Volume Type General Purpose SSD (gp2) ⓘ

Size (GiB) 100 (Min: 1 GiB, Max: 16384 GiB)

IOPS 300 / 3000 (Baseline of 3 IOPS per GiB with a minimum of 100 IOPS, burstable to 3000 IOPS)

Availability Zone* us-east-1a ⓘ

Throughput (MB/s) Not applicable ⓘ

Snapshot ID Select a snapshot ⓘ ⓘ

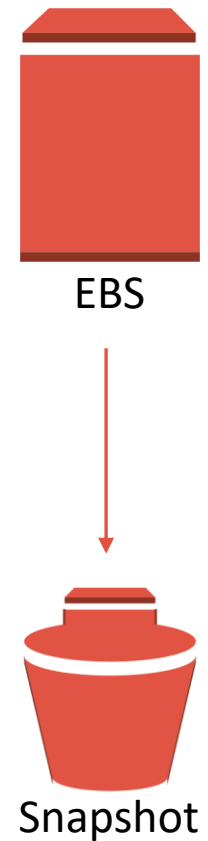
Encryption ☐ Encrypt this volume ⓘ

Snapshots

A snapshot is an **image** of an EBS volume that can be stored as an **incremental backup** of the volume or used to create a duplicate.

A snapshot is not an active EBS volume. You cannot attach or detach a snapshot to an EC2 instance.

To restore a snapshot, you need to create a new EBS volume using the snapshot as its template.



AMI vs Snapshots

- AMI is an entire EC2 instance definition whereas a snapshot is used for a backup of a single volume.
- AMI includes all EBS snapshots plus some metadata like kernel, AMI name, description, block device mappings, and more.
- AMIs could be used to deploy your applications in different regions easily. But you need to recover and mount a snapshot to an EC2 instance.

Think you have 2 volumes. A root volume has your app configurations and gets changed once a year. Your app stores images on an additional volume that you need to backup images every day.

Would you take an AMI or snapshot of the additional volume for backing up images?

Security Groups

Security groups are found on the **instance level** (as opposed to the subnet Level in NACL). They act as a **virtual firewall** that controls the traffic for one or more instances.

When you launch an instance, you associate one or more security groups with the instance. You add rules to each security group that **allow traffic** to or from its associated instances.

All the rules from the security groups that are associated with the instance are evaluated.

Read more: [Security Groups](#)

Inbound/Outbound Rules

Name

Group ID

Group Name

VPC ID

sg-b63da5fa

default

vpc-af9b48d5

Description

Inbound

Outbound

Tags

Edit

Type ⓘ

Protocol ⓘ

Port Range ⓘ

Source ⓘ

Description ⓘ

All traffic

All

All

your IP

Description

Inbound

Outbound

Tags

Edit

Type ⓘ

Protocol ⓘ

Port Range ⓘ

Destination ⓘ

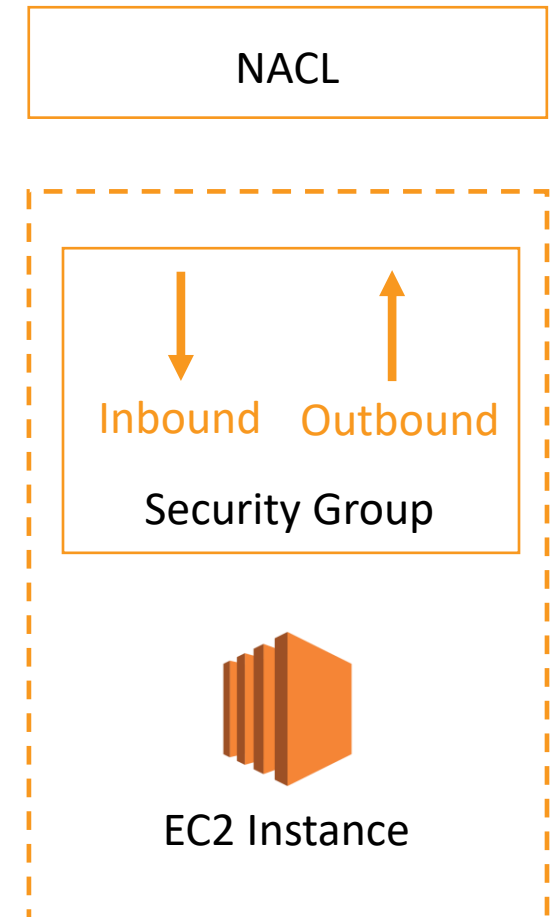
Description ⓘ

All traffic

All

All

0.0.0.0/0



Create Security Group

When you create a new Security Group, by default:

- All inbound traffic is denied
- All outbound traffic is allowed

Create Security Group

Security group name ⓘ

Description ⓘ

VPC ⓘ vpc-af9b48d5 (default)

Security group rules:

Inbound

Outbound

| Type ⓘ | Protocol ⓘ | Port Range ⓘ | Source ⓘ | Description ⓘ |
|----------------------------------|------------|--------------|----------|---------------|
| This security group has no rules | | | | |

Add Rule

Inbound

Outbound

| Type ⓘ | Protocol ⓘ | Port Range ⓘ | Destination ⓘ | Description ⓘ |
|-------------|------------|--------------|------------------|----------------------------|
| All traffic | All | 0 - 65535 | Custom 0.0.0.0/0 | e.g. SSH for Admin Desktop |

Add Rule

Network ACL & Security Groups

| Security group | Network ACL |
|---|--|
| Operates at the instance level | Operates at the subnet level |
| Supports allow rules only | Supports allow rules and deny rules |
| Stateful: Return traffic is automatically allowed, regardless of any rules | Stateless: Return traffic must be explicitly allowed by rules |
| We evaluate all rules before deciding whether to allow traffic | Rules are processed in order, starting with the lowest numbered rule, when deciding whether to allow traffic |
| Applies to an instance only if someone associates the security group with the instance | Automatically applies to all instances in the subnets that it's associated with |
| Rules include IP addresses and other security groups | Rules only include IP addresses |

Read more: [VPC Security](#)

IP Addressing

An IP address is the EC2 instance address on the network.

Private IP Address:

- All EC2 instances have a private IP address.
- Private IP addresses allow instances to communicate with each other.

Public IP address:

- All EC2 Instances can be launched with or without a public IP address.
- Public IP addresses are required for the instance to communicate with the Internet.

Read more: [IP Addressing](#)

Elastic IP

When you stop and then start an EC2 instance, it may change its public IP. If you need to have a fixed public IP for your instance, you need an Elastic IP.

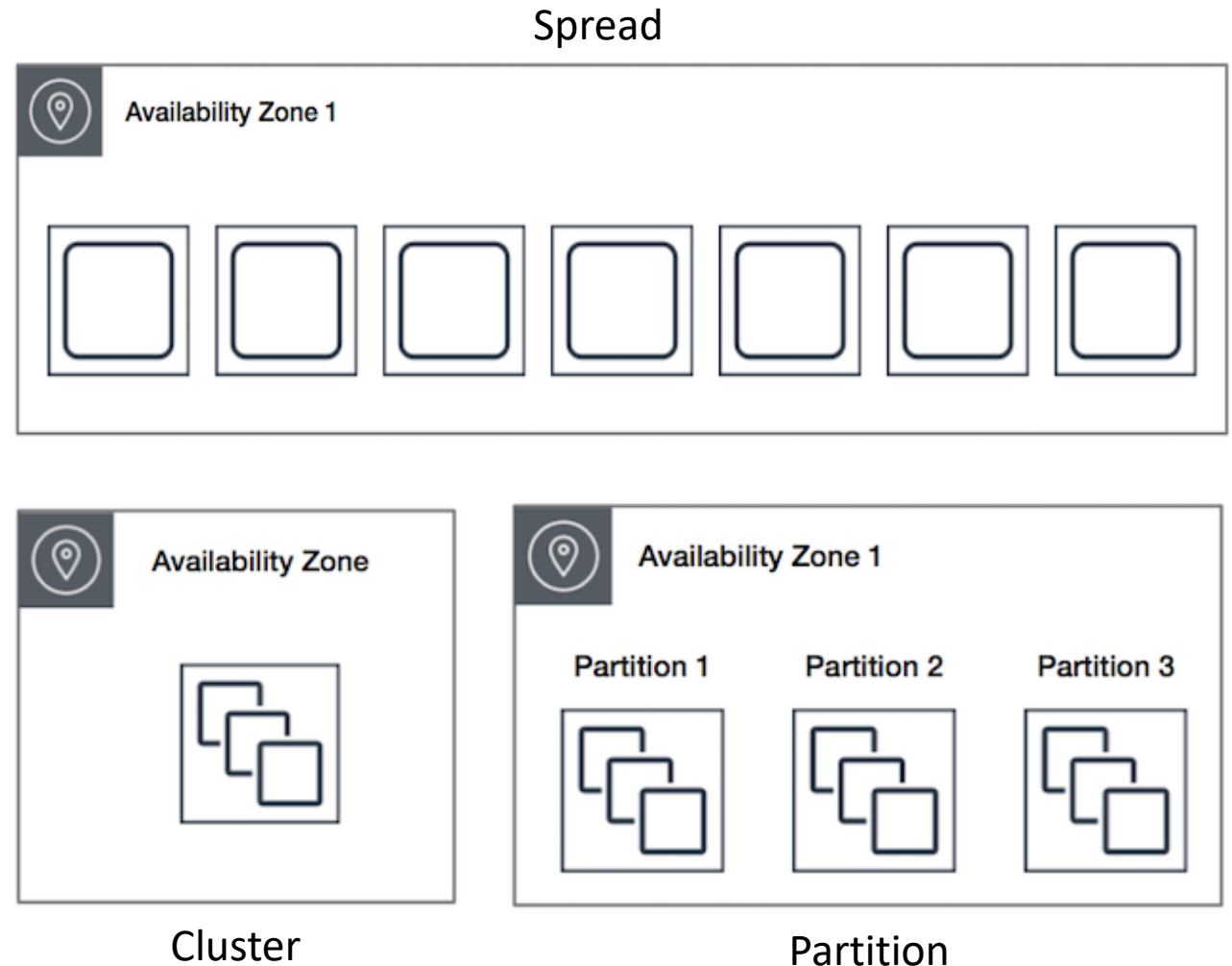
- An Elastic IP is a public IPv4 IP you own.
- You can attach it to one instance at a time.
- By default, you can have 5 elastic IPs in AWS.
- It charges when you are NOT using. When using, it is free.

Placement groups

Cluster – Packs instances close together.
Low latency and high performance.

Partition – Spreads instances across logical partitions. One partition don't share the underlying hardware. Large distributed and replicated workloads such as Hadoop, Cassandra, and Kafka.

Spread – Places a small group of instances across distinct underlying hardware to reduce correlated failure.



EC2 Instance Options - On-Demand

On-demand purchasing allows you to choose any instance type you like and provision/terminate it at any time (on-demand).

- The most expensive purchasing option.
- You are only charged when it is running (billed by the hour).
- You can provision/terminate an on-demand instance at anytime.

EC2 Instance Options - Reserved

Reserved purchasing allows you to purchase an instance for a set time period of 1 or 3 years.

- This allows for a significant price discount over using on-demand.
- You can select to pay upfront, partial upfront, no upfront.
- Once you buy a reserved instance, you own it for the selected time period and are responsible for the entire price - regardless of how often you use it.
- About 20% savings with one-year term and 30% savings with three-year term.

EC2 Instance Options - Spot

Spot pricing is a way for you to bid on an instance type, then only pay for and use that instance when the spot price is equal to or below your bid price.

- Amazon sells unused instances, for short amounts of time, at a substantial discount.
- Spot prices fluctuate based on supply and demand.
- You are charged by the minute.
- A provisioned instance automatically terminates when the spot price is greater than your bid price.