AWS Elastic Compute Cloud - EC2

AWS Region and AZs
What is EC2?
EC2 - Instance Types
EC2 - Instance Lifecycle
Elastic IP Address
Amazon Machine Images
Amazon Elastic Block Store (EBS)
EC2 - Pricing

AWS Region and AZs

Region:

- AWS has the concept of a Region, which is a physical location around the world where we cluster data centers.
- A AWS Region is represented by a code like **us-east-1** for **North Virginia** Location.

Availibility Zones:

- Each Region has multiple, isolated locations known as Availability Zones.
- An Availability Zone is represented by a Region code followed by a letter identifier, **us-east-1a , us-east-1b**

To know more on Region and AZs

What is EC2

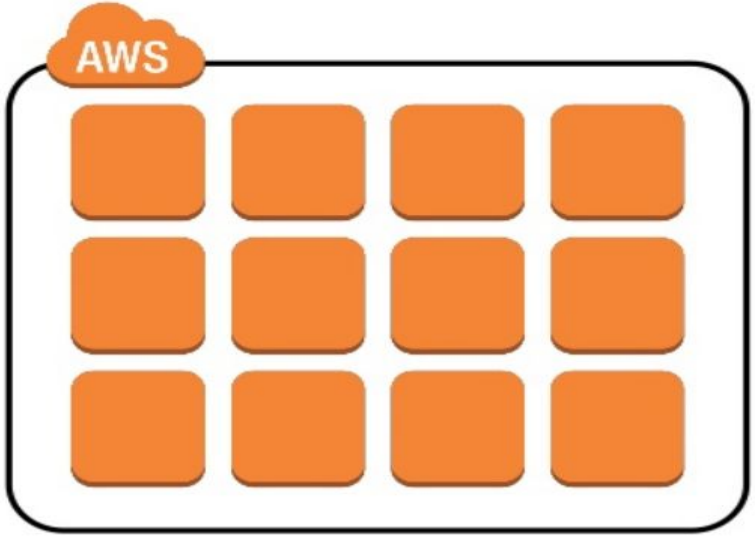
- Provides resizable compute capacity in Cloud.
- Designed to make web-scale cloud computing easier.
- A true virtual computing environment.
- Launch instances with a variety of operatings systems.
- Run as many or few systems as you desire.
- **AWS EC2 Service is Region Specific.**

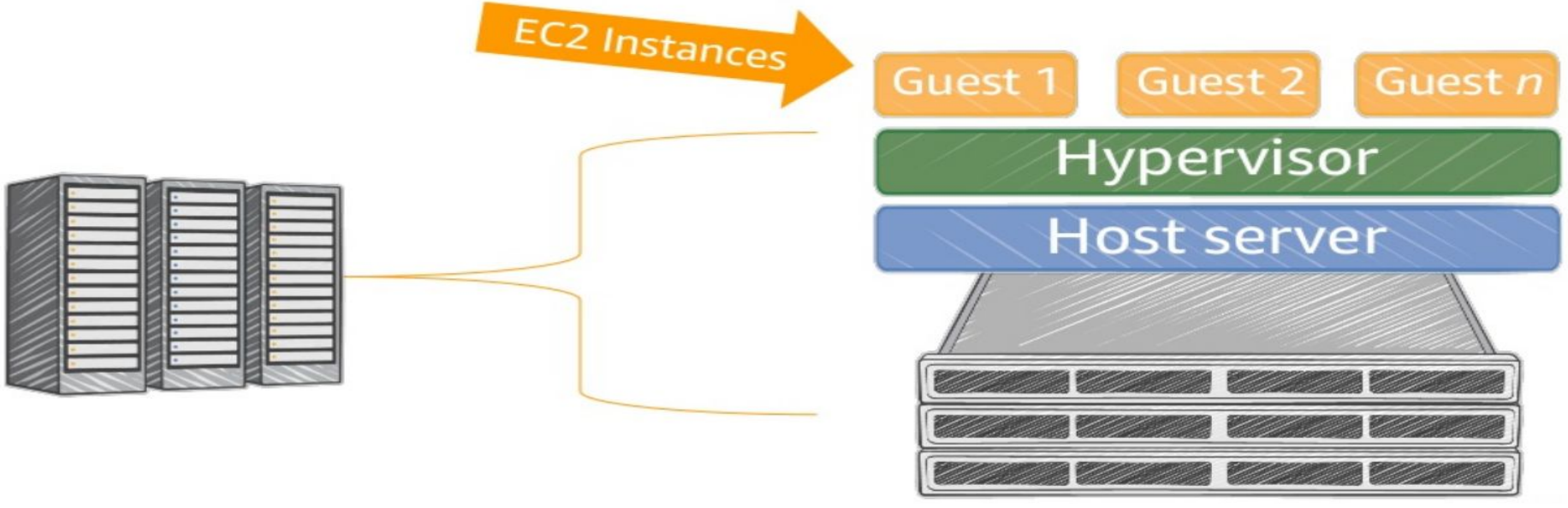
Elastic Compute Cloud

- ✓ Application Server
- ✓ Web Server
- ✓ Database Server
- ✓ Game Server
- ✓ Mail Server
- ✓ Media Server
- ✓ Catalog Server
- ✓ File Server
- ✓ Computing Server
- ✓ Proxy Server

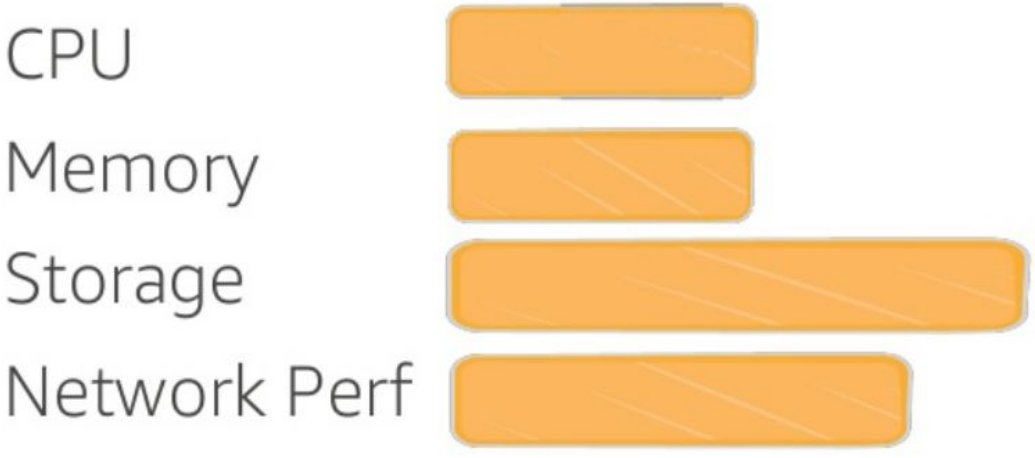
Amazon EC2 Instances

- Pay-as-you-go
- Broad selection of HW/SW
- Global hosting
- *Much more* (aws.amazon.com/ec2)




EC2 Instances
Guest 1
Guest 2
Guest n
Hypervisor
Host server

AWS EC2 : Instance Characteristics



EC2 - Instance Types

Families	Description	Example Use Cases
t2, m4, m3	**General Purpose** Balanced Performance	Websites, web applications, Dev, code repos, micro services, business apps
c3, c4, cc2	**Compute Optimized** High CPU Performance	Front-end fleets, web-servers, batch processing, distributed analytics, science and engineering apps, ad serving, MMO gaming, video-encoding
g2, p2	**GPU Optimized** High-end GPU	Amazon AppStream 2.0, video encoding, machine learning, high perf databases, science
r3, r4, x1, cr1	**Memory Optimized** Large RAM footprint	In-memory databases, data mining
d2, i2, i3, hi1, hs1	**Storage Optimized** High I/O, High density	NAS, data warehousing, NoSQL

General Purpose Instance Workloads

Web/app servers

Enterprise apps

Gaming servers

Caching fleets

Analytics applications

Dev/test environments

Choosing the Right Amazon Instances?

- EC2 Instance types are optimized for different use cases, workloads, and come in multiple sizes. This allows you to optimally scale resources to your workload requirements.
- AWS utilizes Intel® Xeon® processors for EC2 Instances providing customers with high performance and value.
- Consider the following when choosing your instances: core count, memory size, storage size & type, network performance, I/O requirements, and CPU technologies.
- Hurry Up & Go Idle - A larger compute instance can save you time and money, therefore paying more per hour for a shorter amount of time can be less expensive.

Amazon Machine Images

- AMIs (Amazon Machine Images) are immutable images that are used to launch preconfigured EC2 instances.
- They come in both public and private flavors.
- Access to public AMIs is either freely available (shared/community AMIs) or bought and sold in the AWS Marketplace.
- Many operating system vendors publish ready-to-use base AMIs.
- Modern deployments will usually be with **64-bit EBS-backed HVM.**
- You can create your own custom AMI by snapshotting the state of an EC2 instance that you have modified.

Amazon Machine Images

Amazon maintained

Set of Linux and Windows images

Kept up-to-date by Amazon in each region

Community maintained

Images published by other AWS users

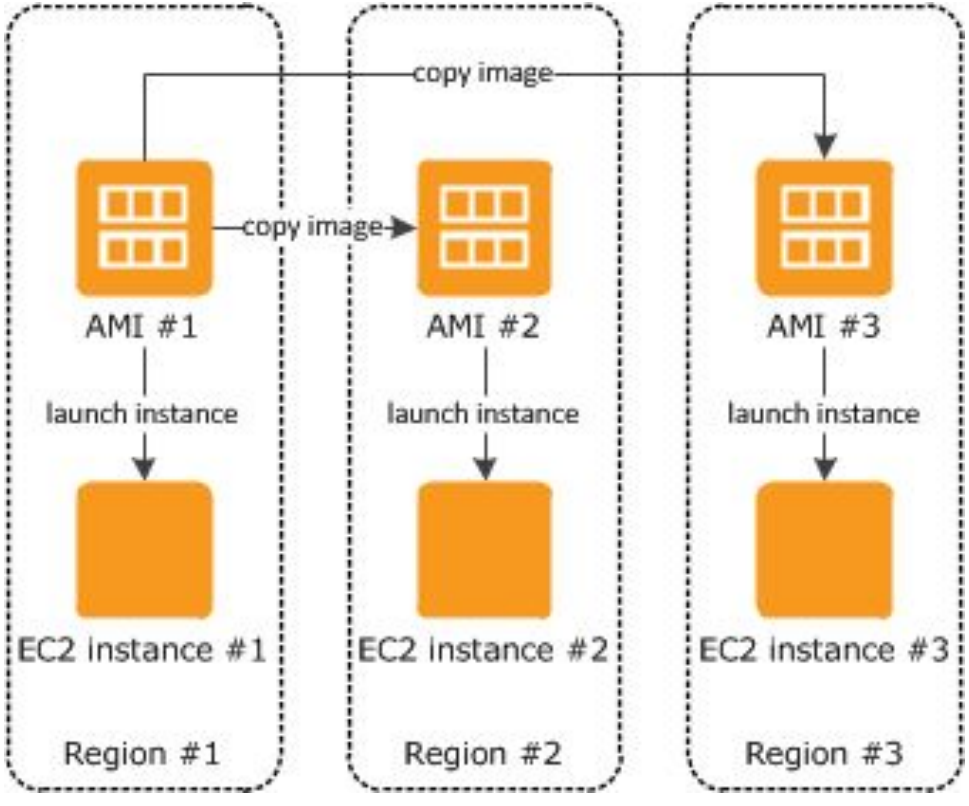
Managed and maintained by Marketplace partners

Your machine images

AMIs you have created from EC2 instances

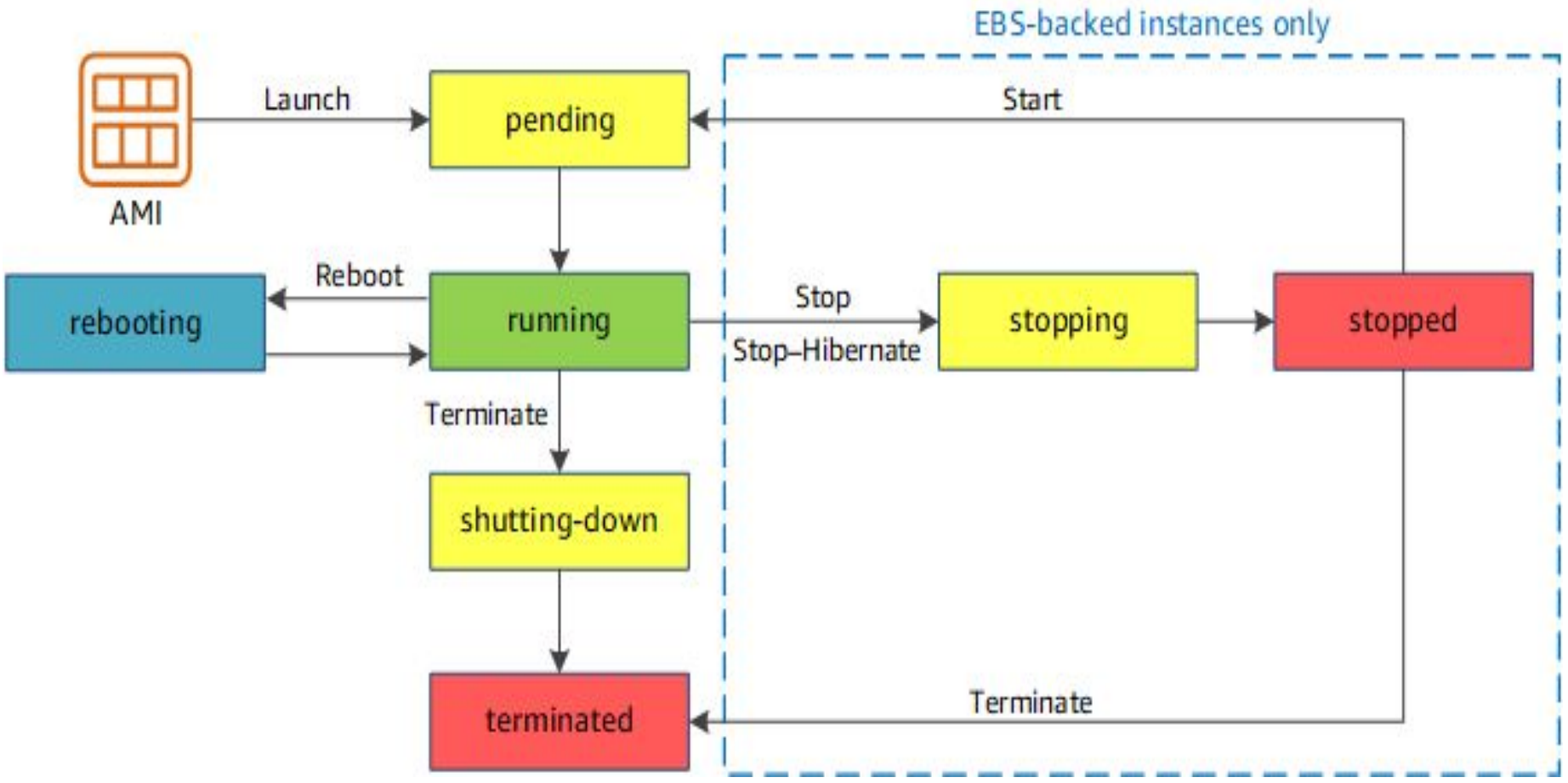
Can be kept private or shared with other accounts

Launching an Instance and AMI Copy



- When you launch an instance, you must select an AMI that's in the same Region.
- If the AMI is in another Region, you can copy the AMI to the Region you're using.
- AMI ID will be different in every Region.

EC2 - Instance Lifecycle



Elastic IP Address

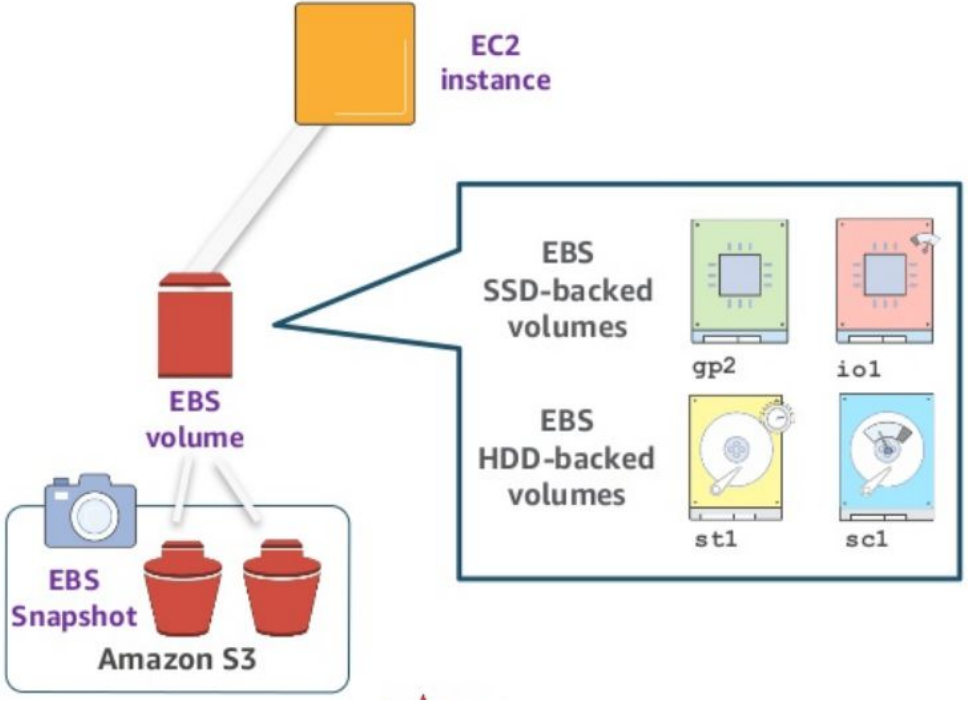
- A Normal Public IP attached to a EC2 instance is changed if an EC2 instance is stopped and started.
- **Elastic IPs** are **static IP** addresses you can rent from AWS to assign to EC2 instances.
- By default, all AWS accounts are limited to five (5) Elastic IP addresses per Region, because public (IPv4) internet addresses are a scarce public resource.
- Its possible to increase via Service Limit Increase Request.
- If an Elastic IP is not attached to an active resource there is a **minimal hourly fee.**
- Elastic IPs have no extra charge as long as it is attached to a running EC2 instance.
- They have a (small) cost when not in use, which is a mechanism to prevent people from squatting on excessive numbers of IP addresses.

- Click here to view EIP Pricing.

Elastic Block Store

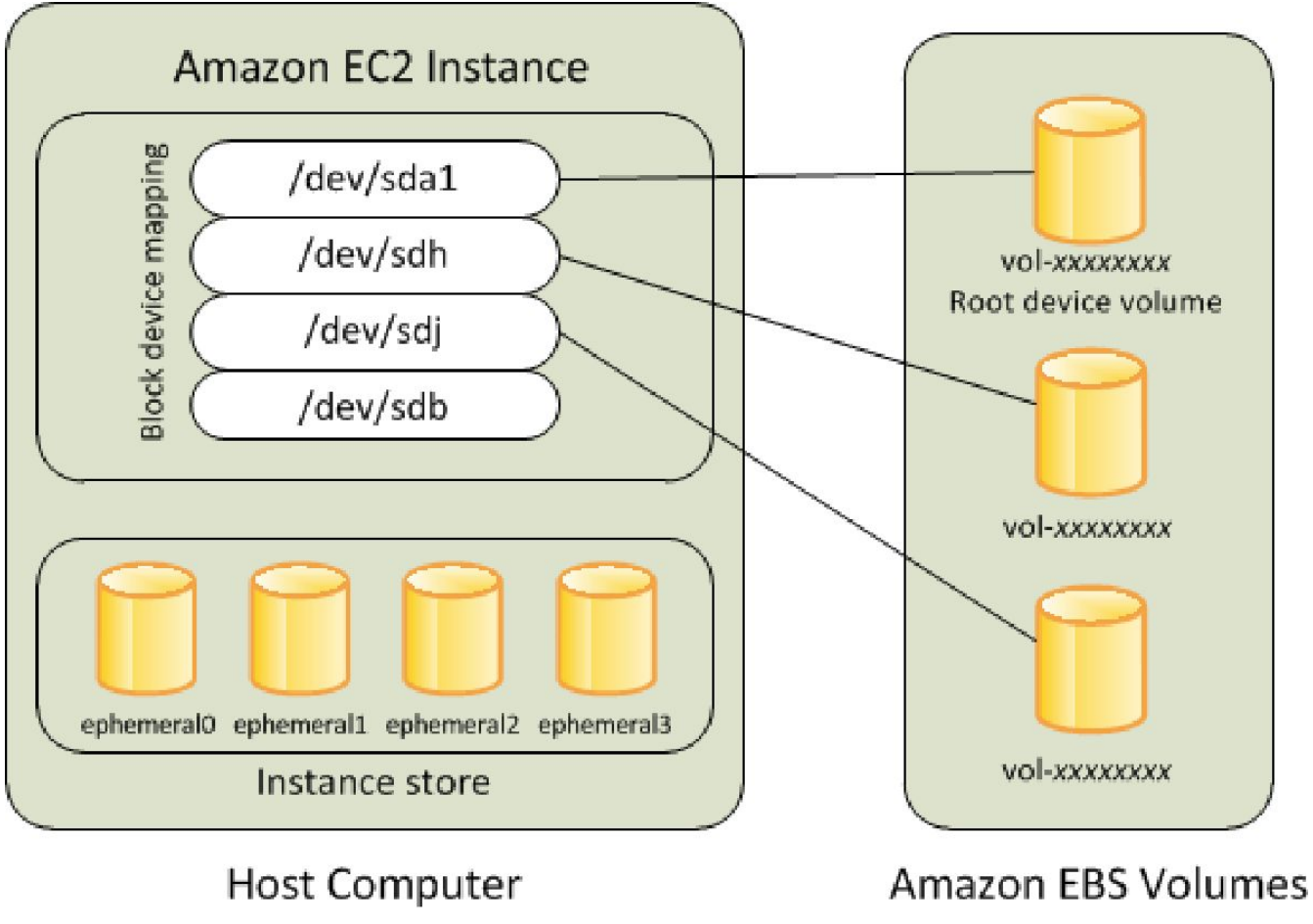
- **EBS (Elastic Block Store)** provides block level storage. That is, it offers storage volumes that can be attached as filesystems, like traditional network drives.
- EBS volumes can only be attached to one EC2 instance at a time.
- Elastic Volumes capability allows you to increase storage, tune performance up and down, and change volume types without any disruption to your workloads.
- Snapshots can be used to quickly restore new volumes across a region's Availability Zones, enabling rapid scale.
- EBS Snapshots allow you to easily take backups of your volumes for geographic protection of your data.
- **Data Lifecycle Manager (DLM)** is an easy-to-use tool for automating snapshot management without any additional overhead or cost.

Amazon Elastic Block Store (EBS)



- Block storage as a service
- Create, attach volumes through an API
- Service accessed over the network
- Select storage and compute based on your workload
- Volumes persist independent of EC2
- Detach and attach between instances
- Choice of magnetic and SSD-based volume types
- Supports Snapshots: Point-in-time backup of modified volume blocks

EBS Volumes


Amazon EC2 Instance
Block device mapping
/dev/sda1
/dev/sdh
/dev/sdj
/dev/sdb
ephemeral0
ephemeral1
ephemeral2
ephemeral3
Instance store
Host Computer
vol-xxxxxxxx
Root device volume
vol-xxxxxxxx
vol-xxxxxxxx
Amazon EBS Volumes

EBS Volumes

Characteristics

- Persistent and customizable block storage for EC2 instances
- HDD and SSD types
- Use Snapshots for backups
- Easy and transparent encryption
- Elastic

Availability

- Durable and automatically replicated

Drive Types

- Storage that best fits your needs
- Magnetic or SSD
- Performance and price requirements

EBS Volumes

Snapshots

- Point-in-time snapshots
- Recreate a new volume at any time

Encryption

- Encrypted EBS volumes
- No additional cost

Elasticity

- Increase capacity
- Change to different types

EBS Volumes

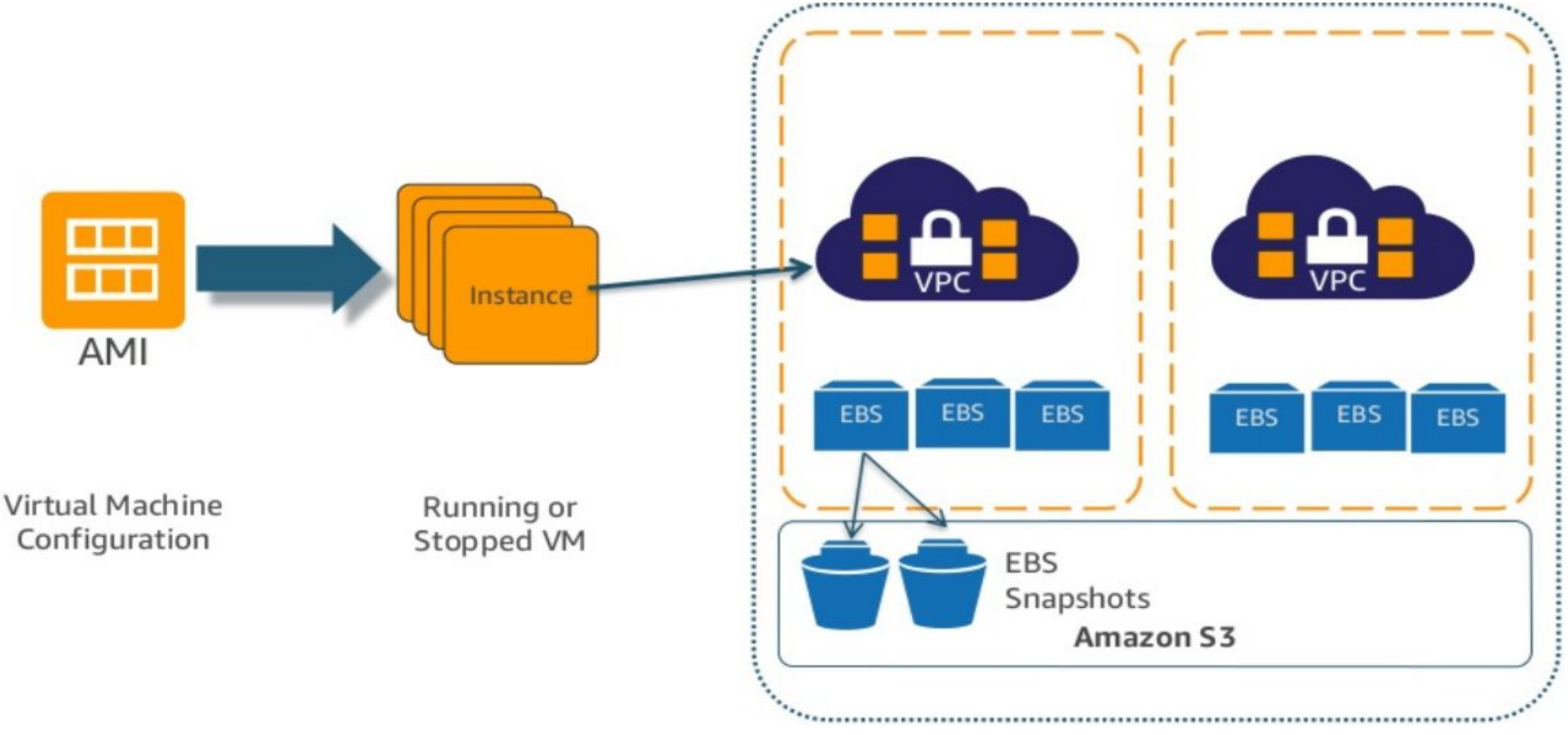
- Features
  - Persistent and customizable block storage for EC2 instances
  - HDD and SSD types
  - Replicated in the same Availability Zones
  - Easy and transparent encryption
  - Elastic volumes
  - Back up using snapshots

Click here to view EBS Pricing.

EBS Snapshots

- An EBS snapshot is a point-in-time copy of your Amazon EBS volume, which is lazily copied to Amazon Simple Storage Service (**Amazon S3**).
- Snapshots are stored in S3 but are not visible under S3 Service View.
- You cannot directly attach a snapshot to an EC2 instance.
- To make Snapshot available in another Region:
  - Copy Snapshot to another region, create Volume and attach to EC2 instance.

EC2 Resources Recap



EC2 Purchasing Options

On-Demand

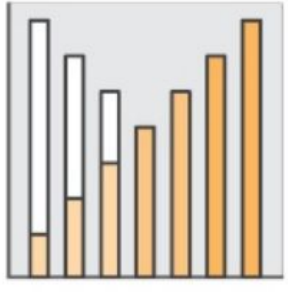
Pay for compute capacity **by the second** with no long-term commitments

Spiky workloads, to define needs

Reserved

Make a 1- or 3-year commitment and receive a **significant discount** off of On-Demand prices

Committed, steady-state usage



Spot

Spare EC2 capacity at a **savings of up to 90%** off of On-Demand prices

Fault-tolerant, dev/test, time-flexible, stateless workloads

NEW! *Per Second Billing for EC2 Linux instances & EBS volumes* NEW!

EC2 - Pricing

On demand:

- Pay for compute capacity by per hour or per second depending on which instances you run.
- No longer-term commitments or upfront payments are needed.
- You can increase or decrease your compute capacity depending on the demands of your application and only pay the specified per hourly rates for the instance you use.

Spot instance

- Amazon EC2 Spot instances allow you to request spare Amazon EC2 computing capacity for up to 90% off the On-Demand price

Reserved Instance

- Reserved Instances provide you with a significant discount (up to 75%) compared to On-Demand instance pricing

Dedicated host

- A Dedicated Host is also a physical server that's dedicated for your use.
- With a Dedicated Host, you have visibility and control over how instances are placed on the server

To view more on EC2 Pricing.

EC2 Spot Pricing

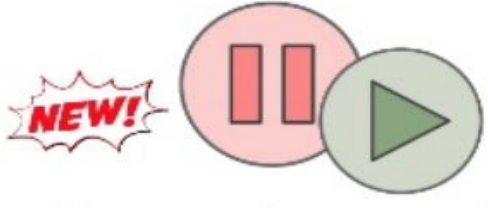
Spare EC2 Capacity that AWS can reclaim with 2-minutes notice

Savings up to 90% off of the On-Demand price

Turbo Boost your results with Spot Fleet

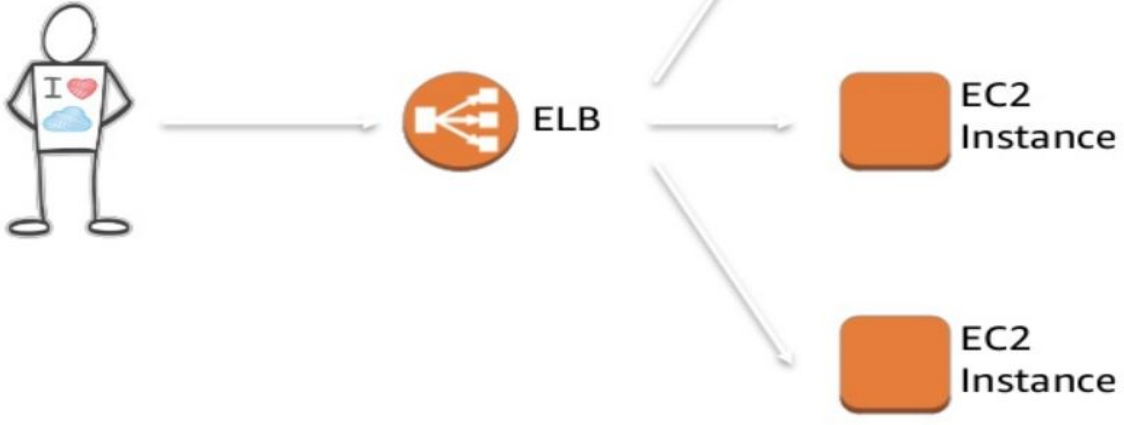
Eliminate the bid!

No need to learn new APIs



Pause and resume with Stop/Start and Hibernate

Elastic Load Balancing



Load balancer used to route incoming requests to multiple EC2 instances, Containers, or IP addresses in your VPC

Elastic Load Balancing provides **high-availability** by utilizing multiple Availability Zones

ANY Questions?