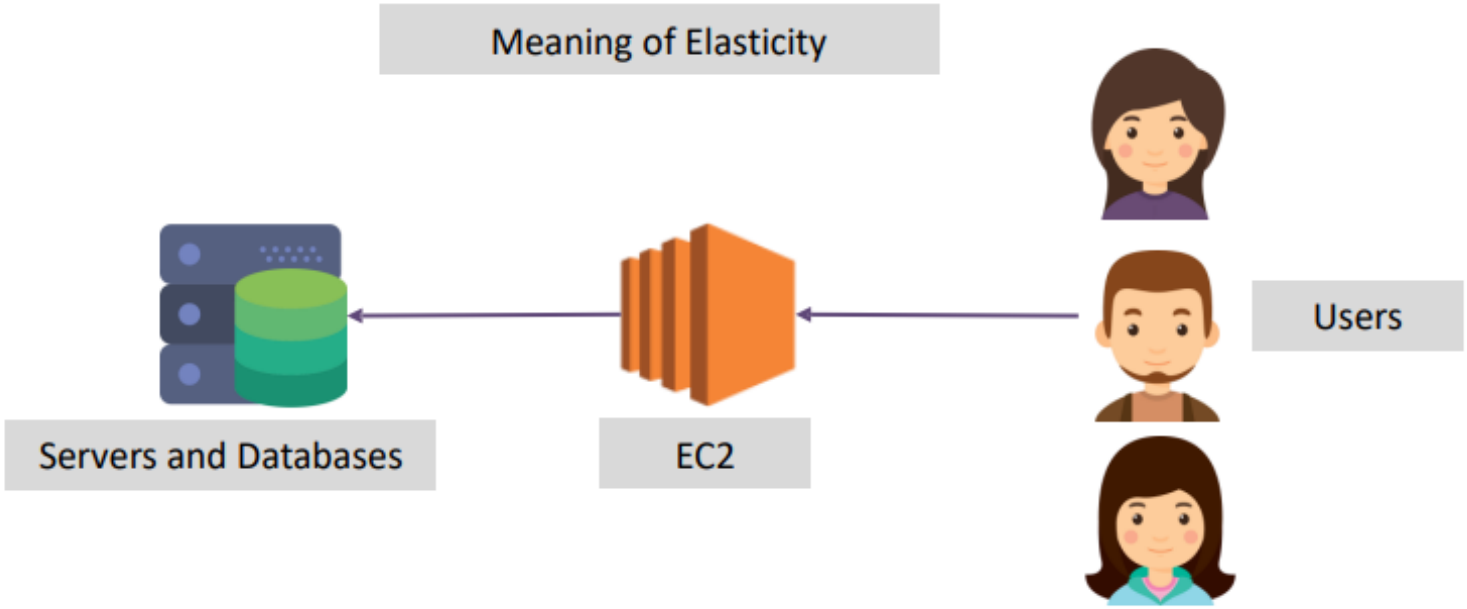
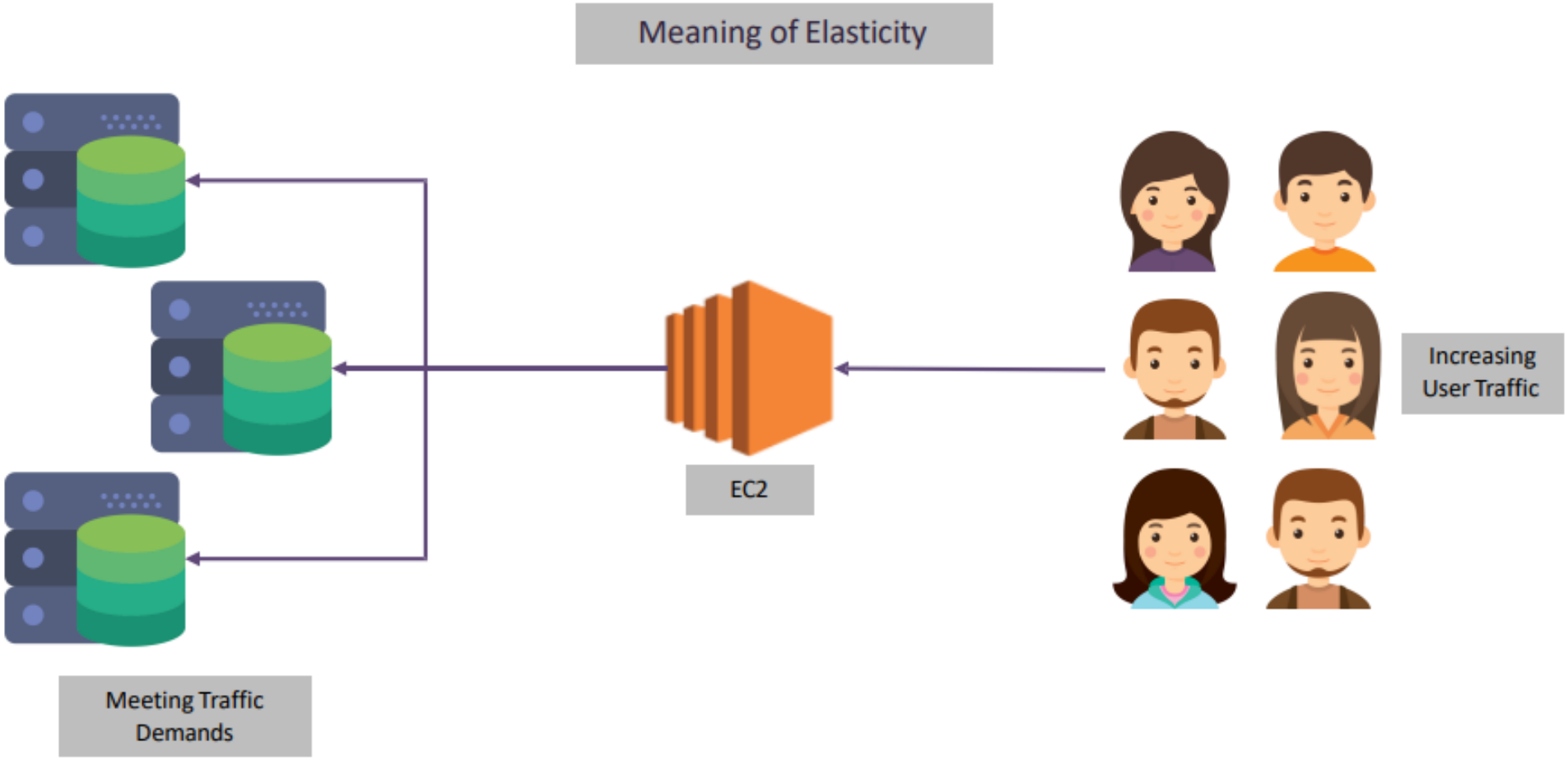
Introduction To EC2

Elastic Compute Cloud

Elastic: It is the level at which a system is able to adapt to workload changes by provisioning and de provisioning resources such that the resources meet the current demand as closely as possible



Introduction To EC2



EC2:Regions and Availability Zones

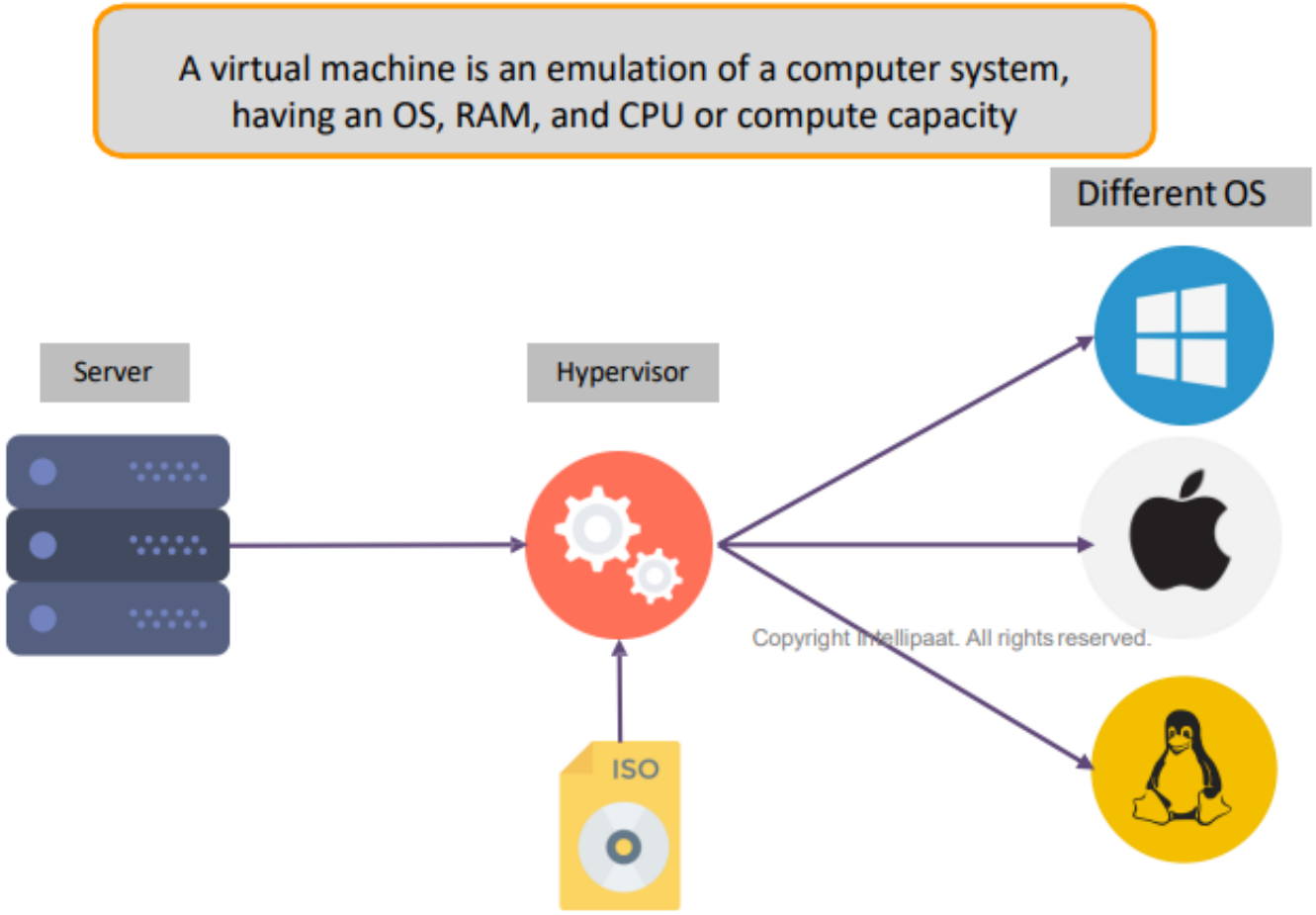
Regions are geographical locations where AWS data centers reside. Following are AWS region names and their subdivisions:

- US East: N. Virginia (us-east-1), Ohio (us-east-2)
- US West: N. California (us-west-1), Oregon (us-west-2)
- Asia Pacific: Mumbai (ap-south-1), Seoul (ap-northeast-2), Singapore (ap-southeast-1)
- EU: Frankfurt (eu-central-1), Ireland (eu-west-1), London (eu-west-2), Paris (eu-west-3)

For instance, 'us-east-1' contains 6 data centers or availability zones:

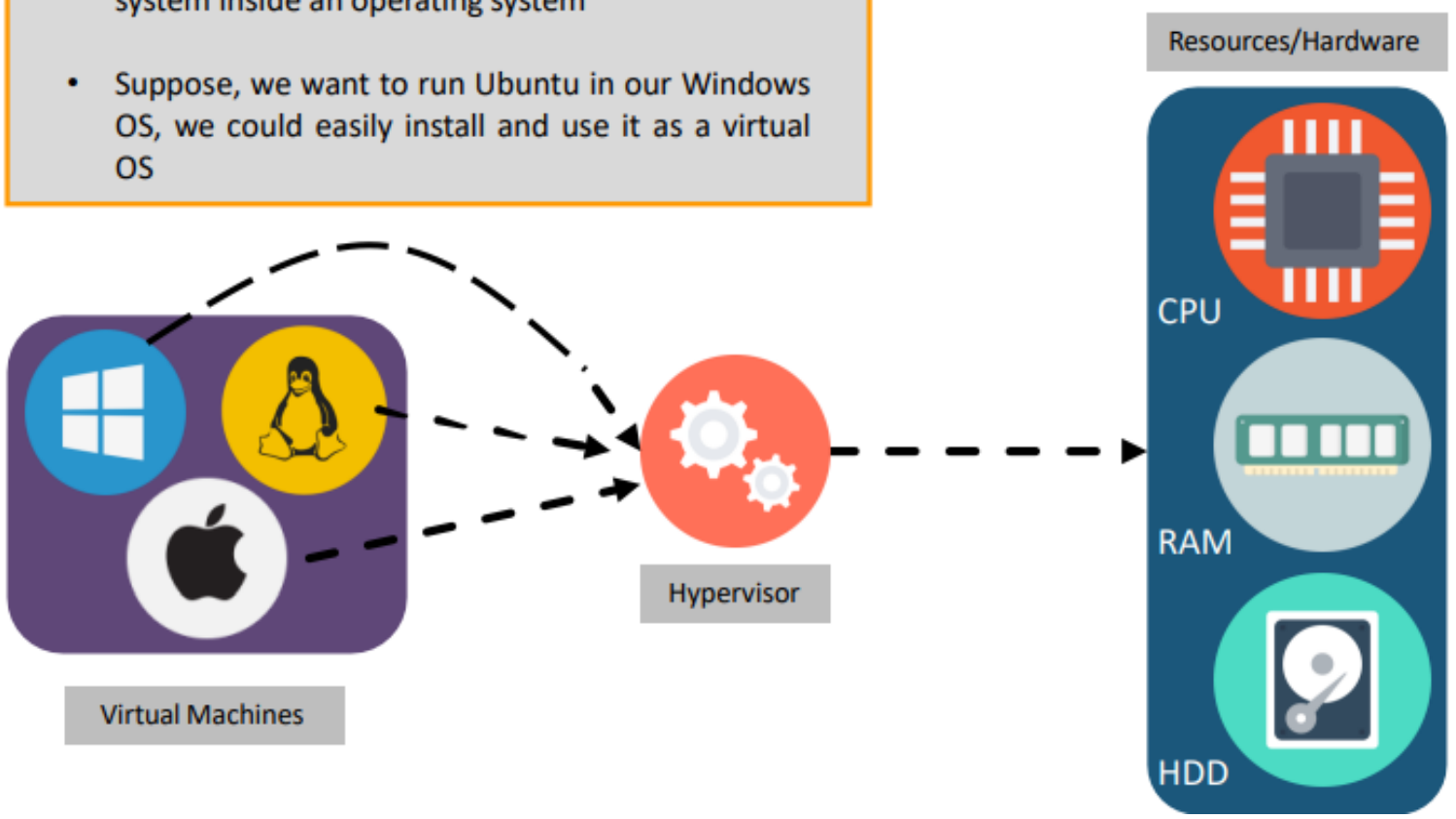
- us-east-1a
- us-east-1b
- us-east-1c
- us-east-1d
- us-east-1e
- us-east-1f

Pre-EC2

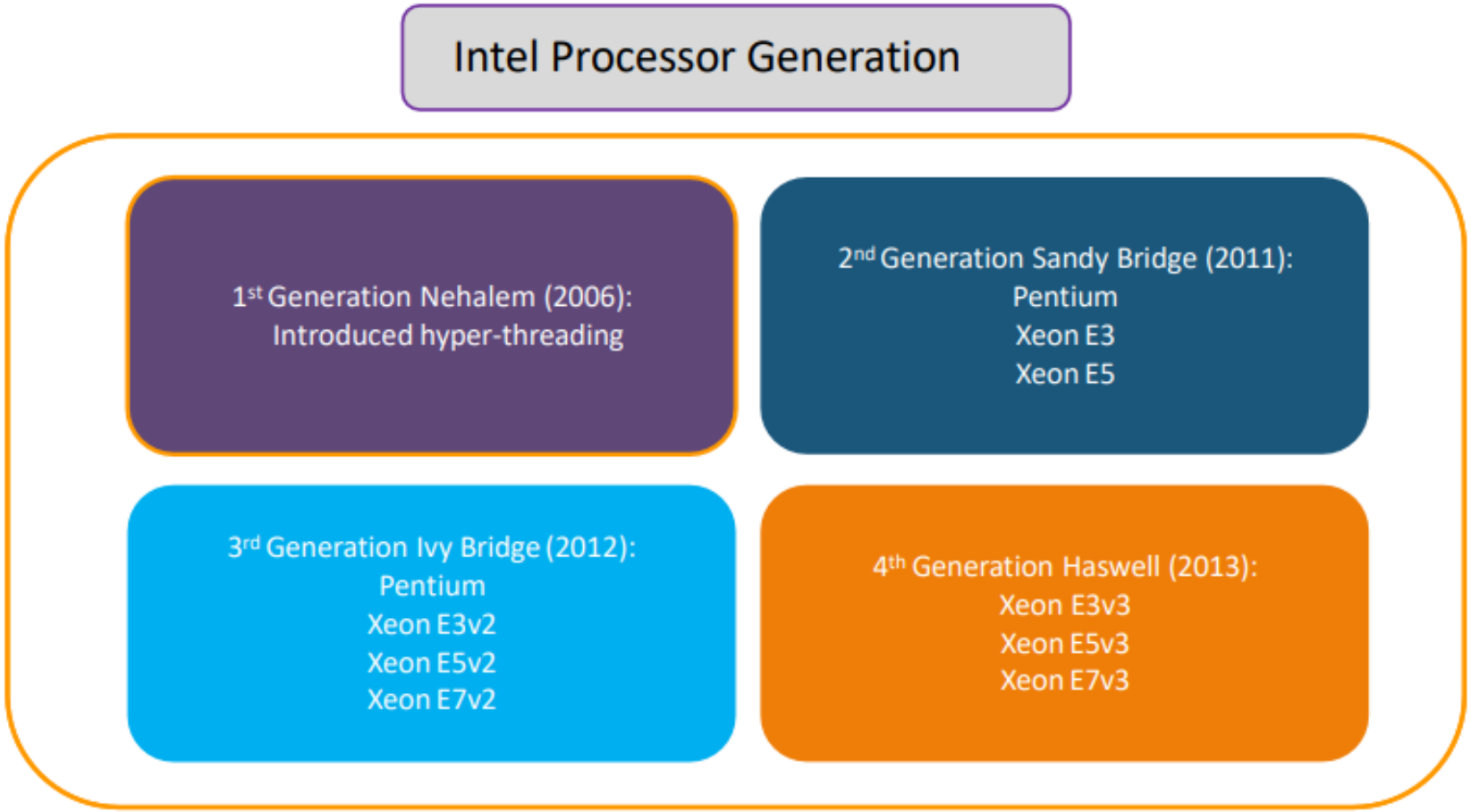

A virtual machine is an emulation of a computer system, having an OS, RAM, and CPU or compute capacity
Different OS
Server
Hypervisor
ISO
Copyright Intellipaat. All rights reserved.

Pre-EC2

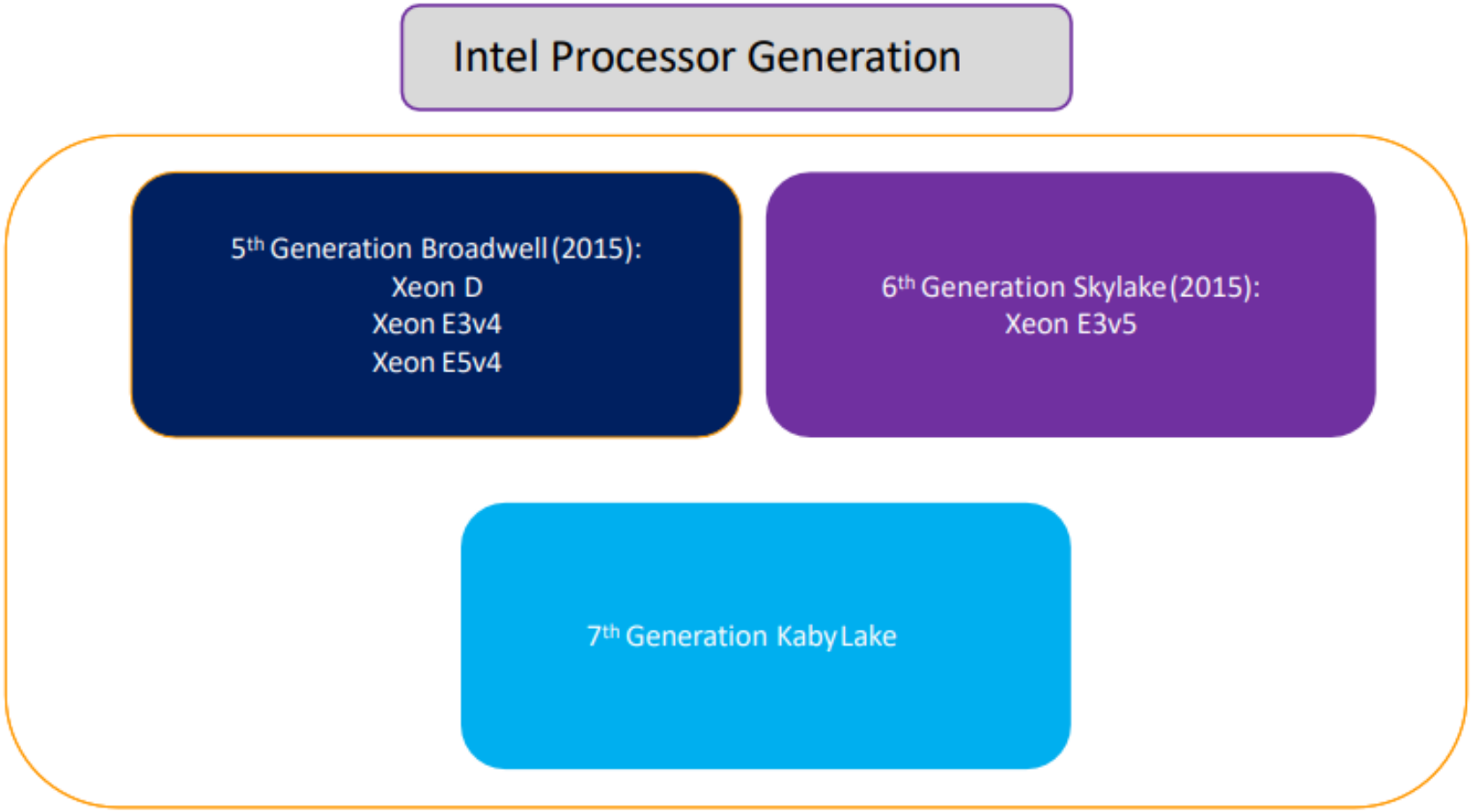
- In simple terms, it is running a virtual operating system inside an operating system
- Suppose, we want to run Ubuntu in our Windows OS, we could easily install and use it as a virtual OS



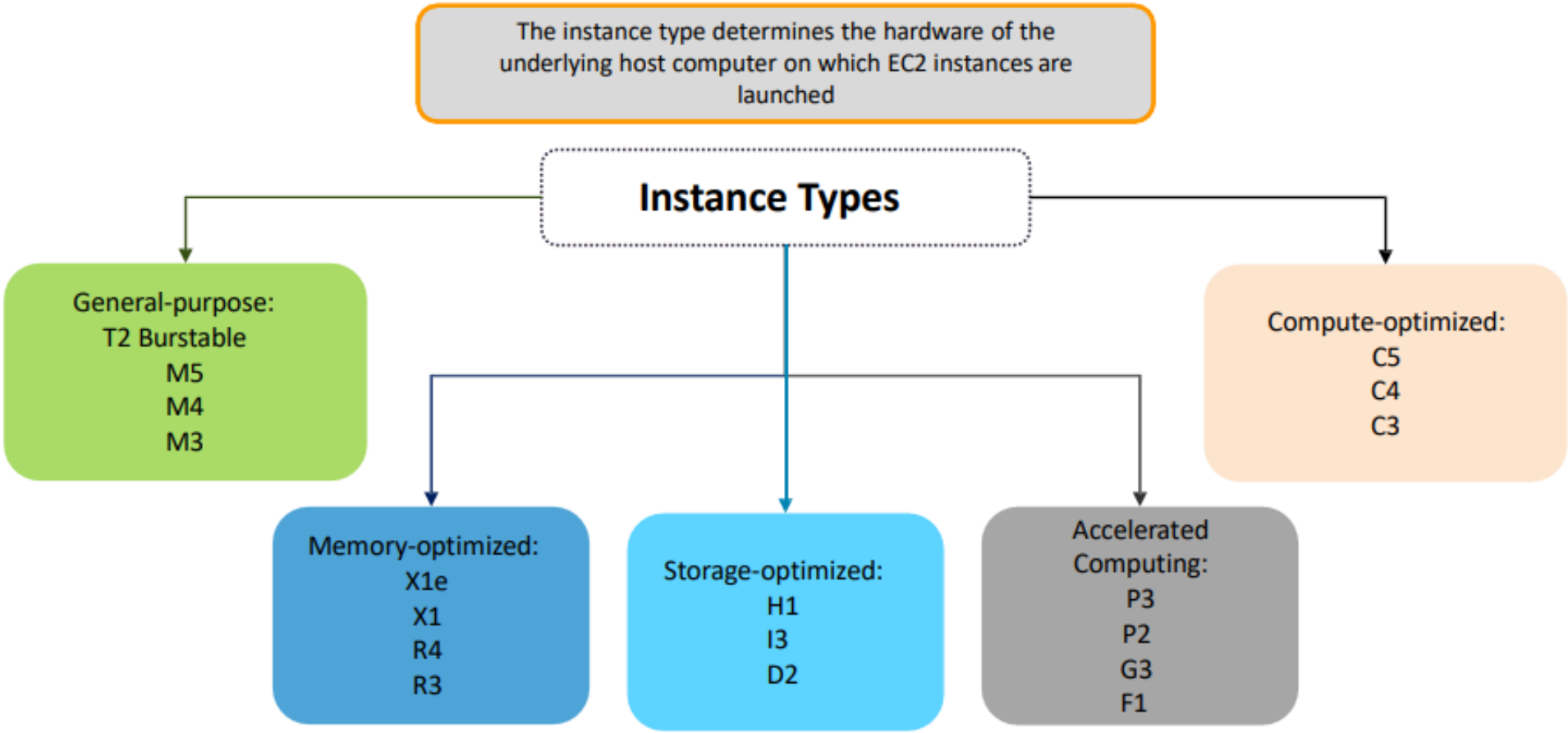
Pre-EC2


Intel Processor Generation
1st Generation Nehalem (2006):
Introduced hyper-threading
2nd Generation Sandy Bridge (2011):
Pentium
Xeon E3
Xeon E5
3rd Generation Ivy Bridge (2012):
Pentium
Xeon E3v2
Xeon E5v2
Xeon E7v2
4th Generation Haswell (2013):
Xeon E3v3
Xeon E5v3
Xeon E7v3

Pre-EC2

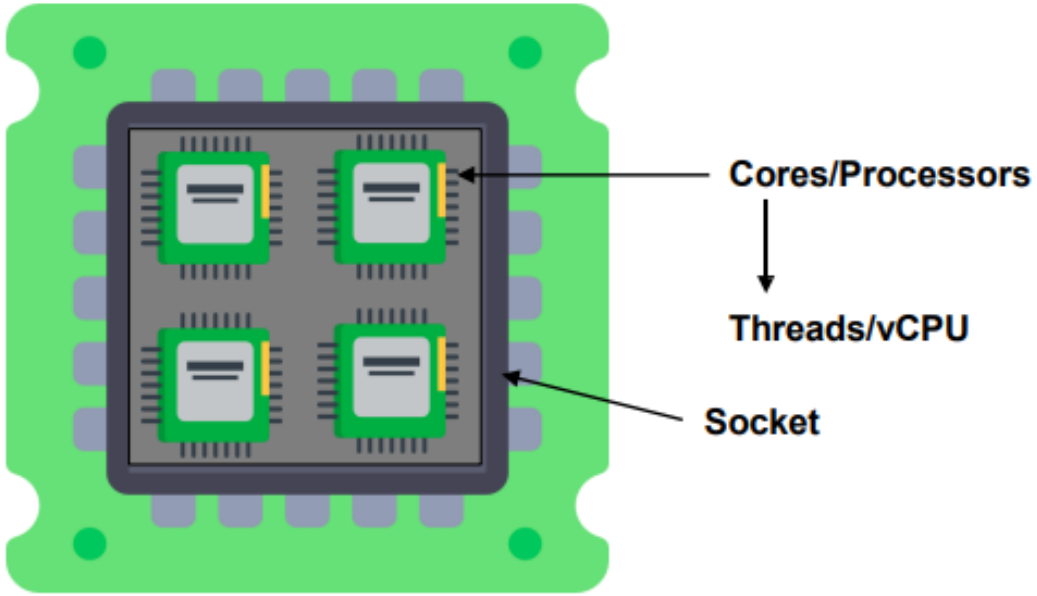

Intel Processor Generation
5th Generation Broadwell (2015):
Xeon D
Xeon E3v4
Xeon E5v4
6th Generation Skylake (2015):
Xeon E3v5
7th Generation Kaby Lake

EC2 Instance Types



The vCPU and the Root Device Volume

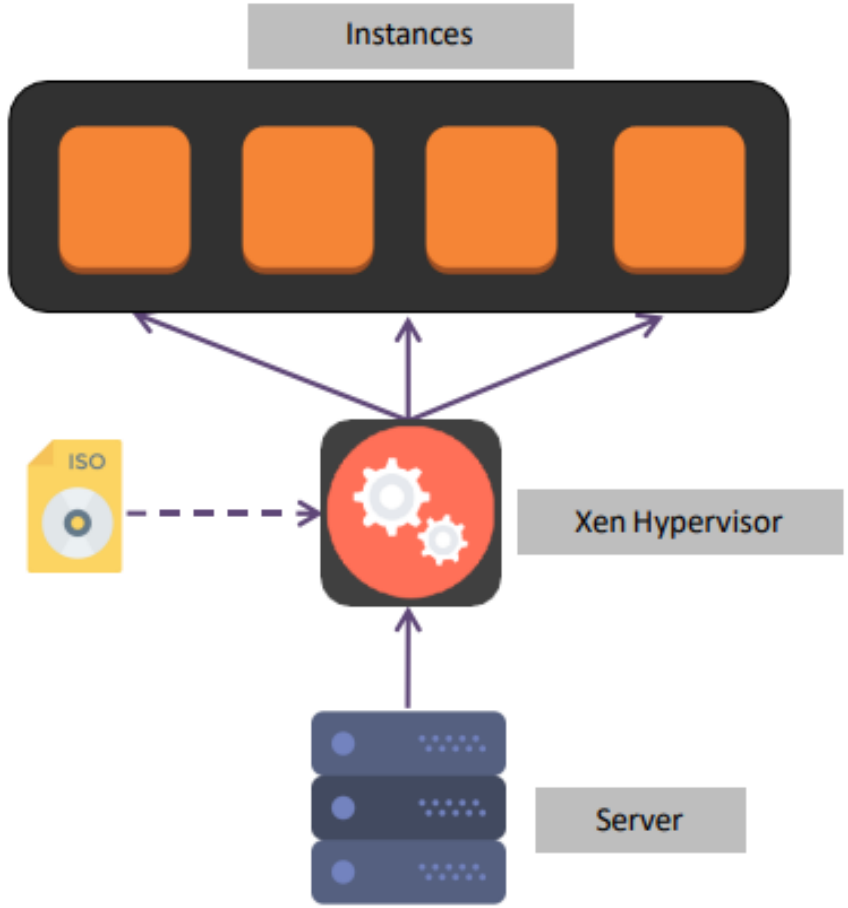
Each vCPU is a hyper-thread of an Intel Xeon core except for t2 and m3.medium instances (AWS Definition) The Root Device Volume contains the image using which the instance is booted



What is an AMI?

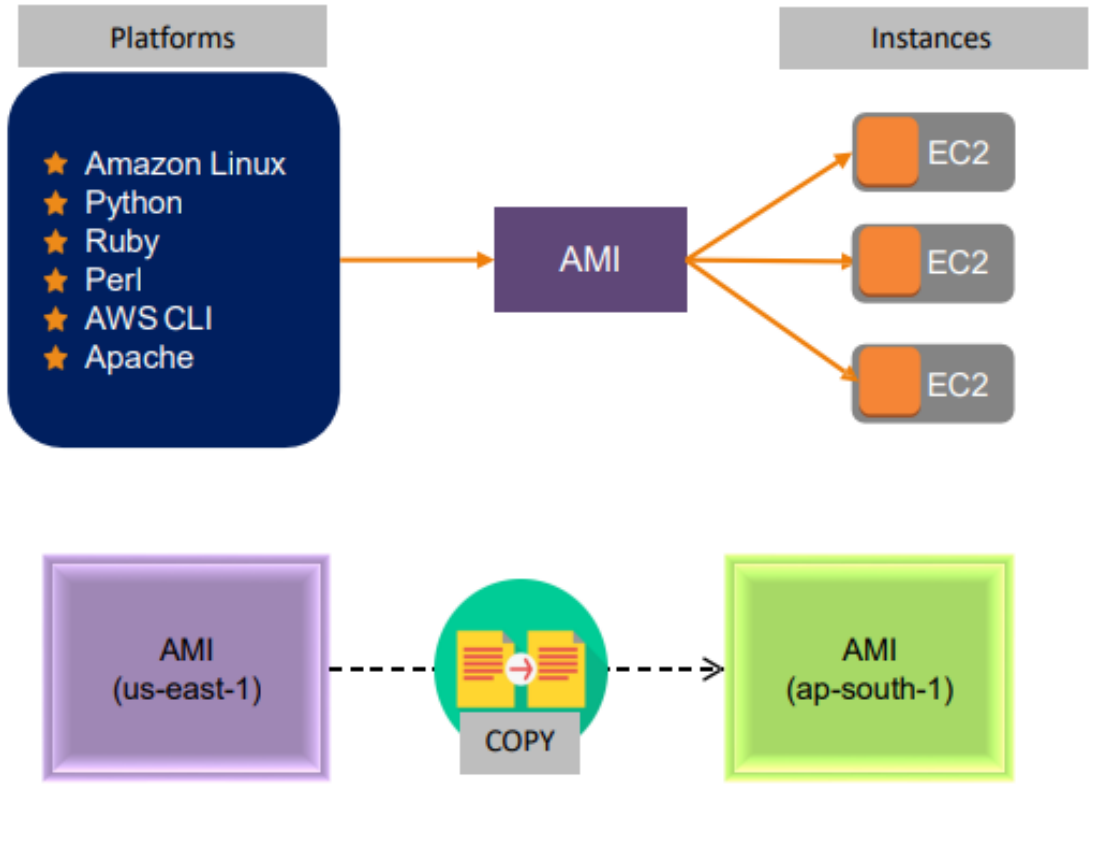
Amazon Machine Image (AMI) contains the information required to launch an instance

- Operating system
- Architecture
- Storage for the root device (Instance store or EBS-backed)
- Virtualization type (HVM or PV)



Creating and Copying an AMI

- Create an AMI from an instance
- Launch multipl instances from it
- Copy the AMI
- AMI permissions



Creating and Copying an AMI

Creating an AMI

1. Select the instance we created in the last demo
2. Click on the **Actions** button, and choose: Image → Create image
3. Provide a name and a small description for the image, and then click on **Create AMI**
4. Now, click on **AMIs** under the Images group in the left-side scroll bar

We will see that the AMI has been created!

Copying an AMI to Another Region

1. Select the created AMI, and click on the Actions button (Actions → Copy AMI)
2. Choose the destination region, and click on **Copy AMI**
3. Go back to the AMIs view, and wait until it is available

We have now successfully created and copied an AMI!

Public vs Elastic IP

Public IP

- It is not associated with an AWS account
- No charges for the public IP, even if it is not being used while the instance is running
- Whenever the instance is re-launched, the public IP changes

Elastic IP

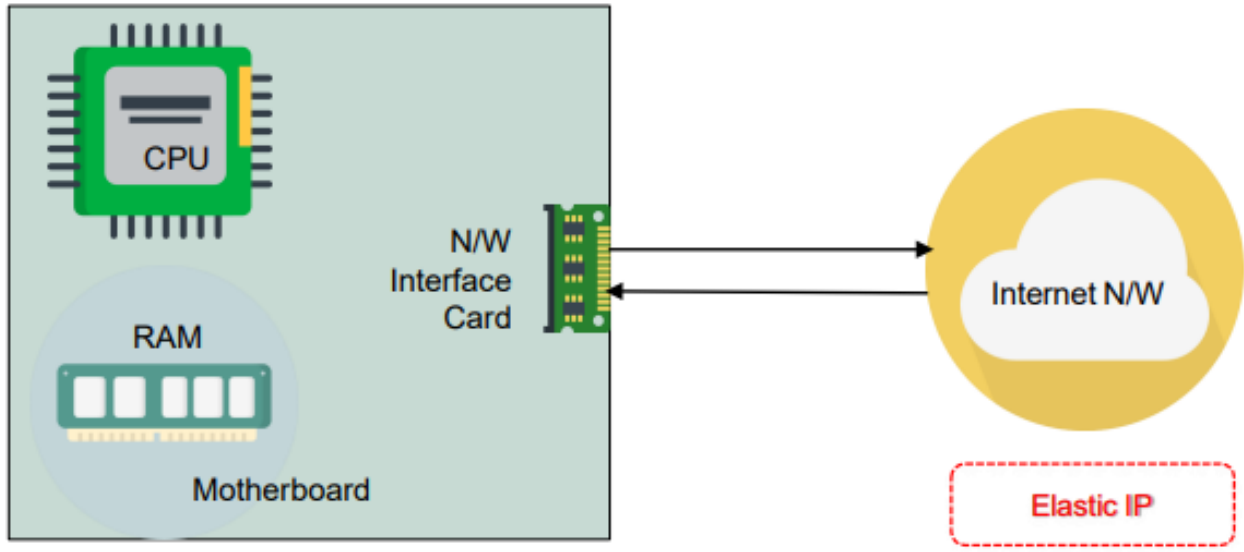
- It is associated with the AWS account
- Charges will be applied if the same is done with the elastic IP
- The elastic IP is the same and static for every launch until we manually release it

Elastic Network Interface

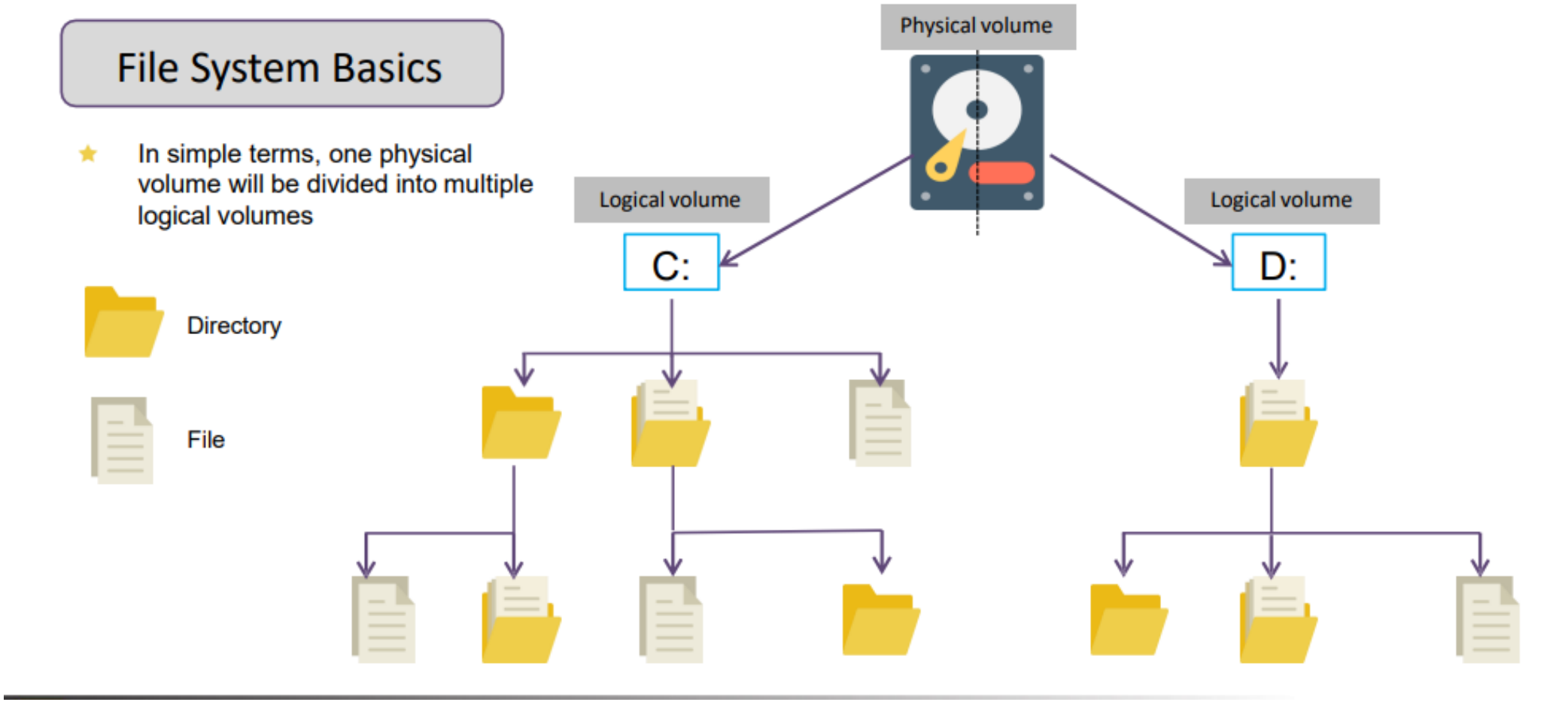
A network interface is the interface between a computer and an Internet network. The network IO happens through n/w interface cards

N/W interfaces contain:

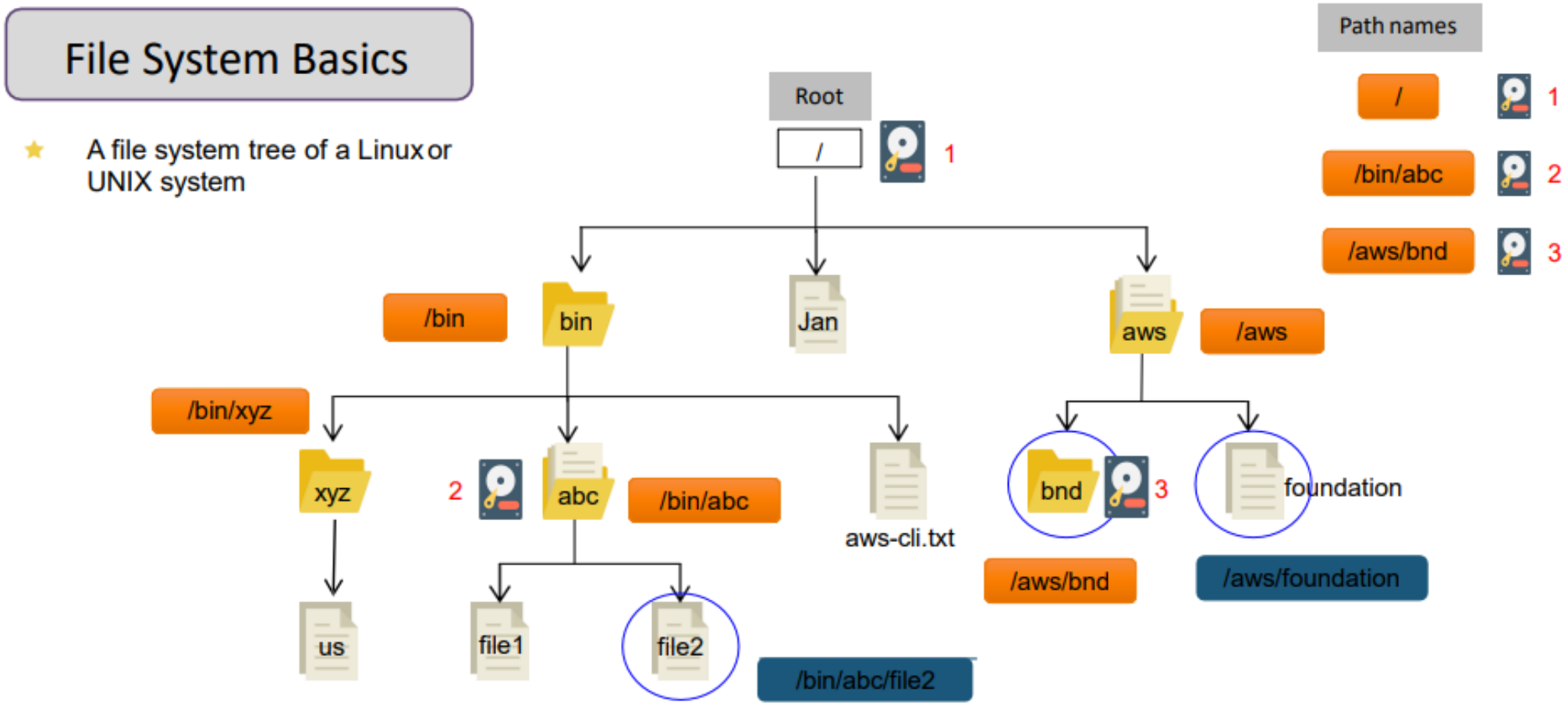
- Elastic IP
- Public IP
- Private IP
- Security Groups



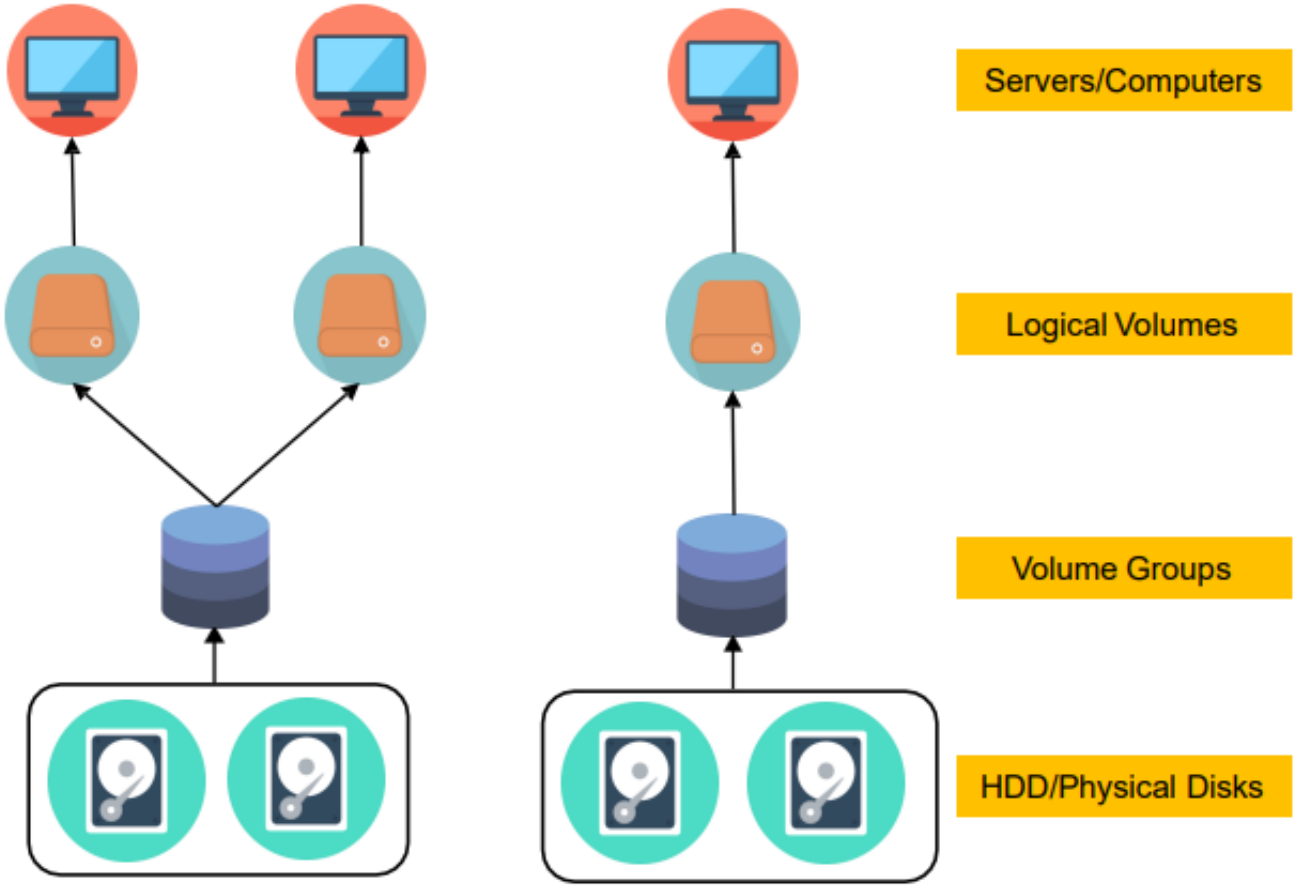
Introduction to EBS



Introduction to EBS

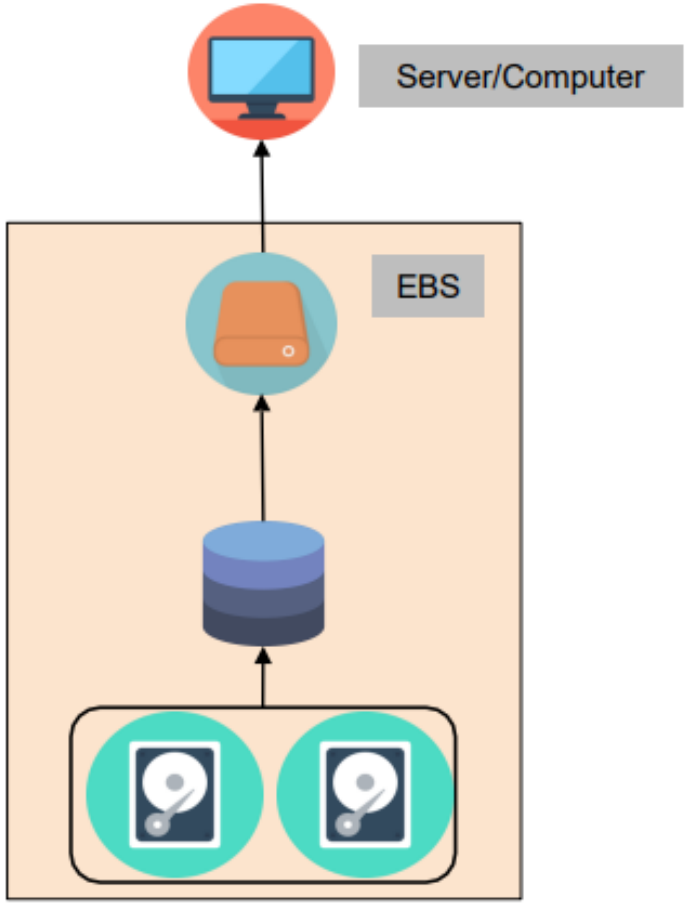


Pre-EBS Storage Layers



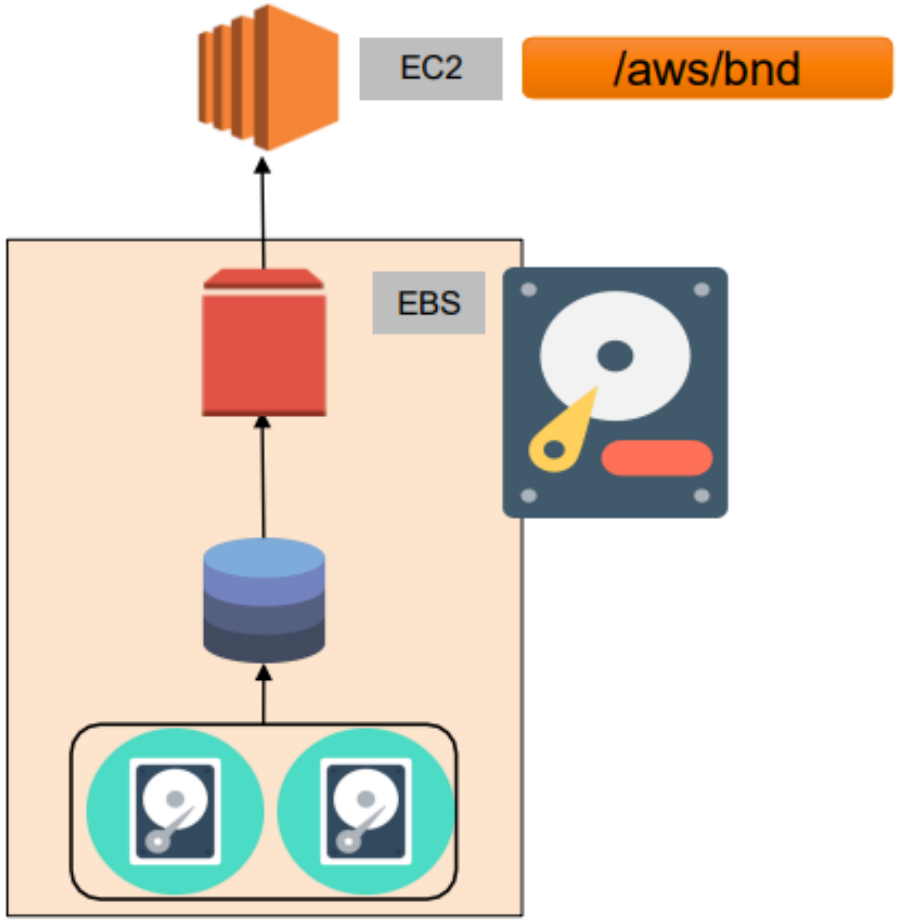
Elastic Block Store

- In an EBS block-level storage, the server-base operating system connects with the raw volumes that are created through a fiber channel
- Then, they are used as individual disks, and if it is very versatile, it can be used as file storage, database storage, and virtual machine volumes



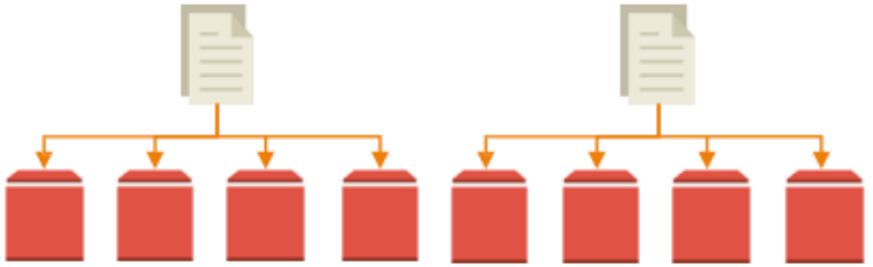
Elastic Block Store

- An EC2 instance is directly connected to EBS
- While the instance is running, a volatile memory called ephemeral storage will be attached to the instance
- If the instance is stopped, the ephemeral memory will be detached

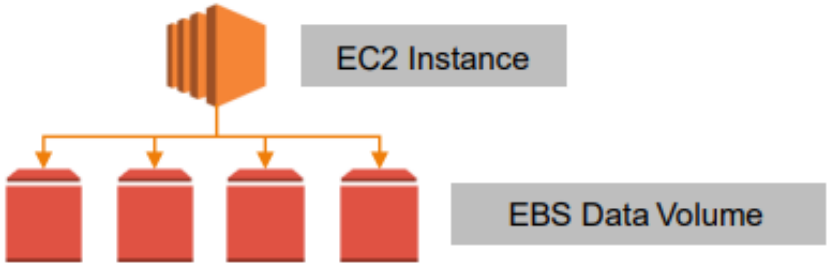


EBS Concepts

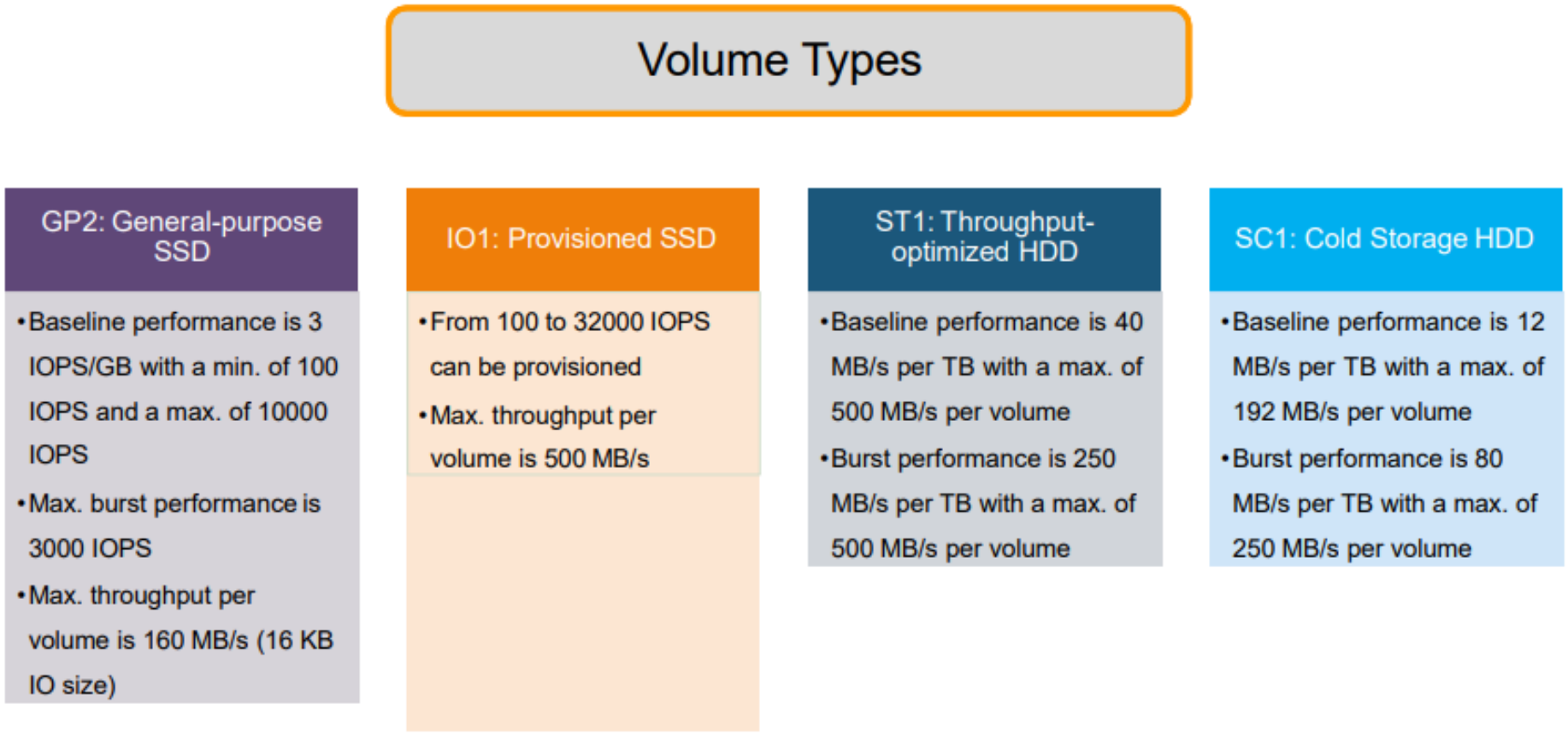
- It is the raw unformatted block-level storage; it is exposed as raw device to the EC2 instance
- EBS volumes persist independently from the life of the EC2 instance
- An EBS volume is automatically replicated within an availability zone
- **Throughput**: It is the sequential transfer rate that an SSD or HDD will maintain continuously



- **IOPS**: It is the measure of the number of I/O operations a drive, SSD, or HDD can handle per second with each block being read from or written to a RANDOM location in the disk



EBS Concepts



EBS Concepts

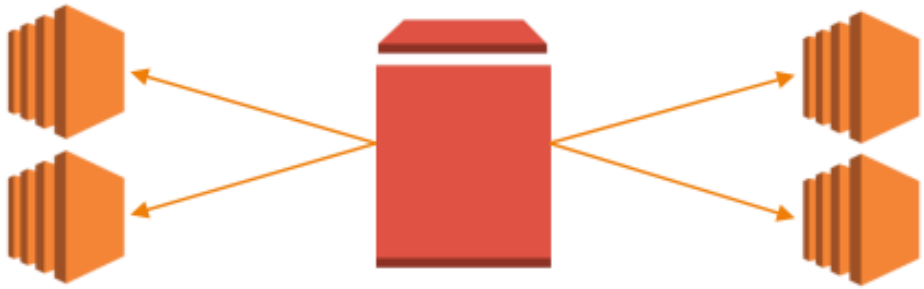
Volume Types

Volume Type	Size Limit	Maximum IOPS	Maximum Throughput	Maximum Burst
GP2	1 GB – 16 TB	10000	160 MB/s	3000
IO1	4 GB – 16 TB	32000	500 MB/s	NA
ST1	500 GB – 16 TB	500	500 MB/s	500 MB/s
SC1	500 GB – 16 TB	250	192 MB/s	250 MB/s

EBS Concepts

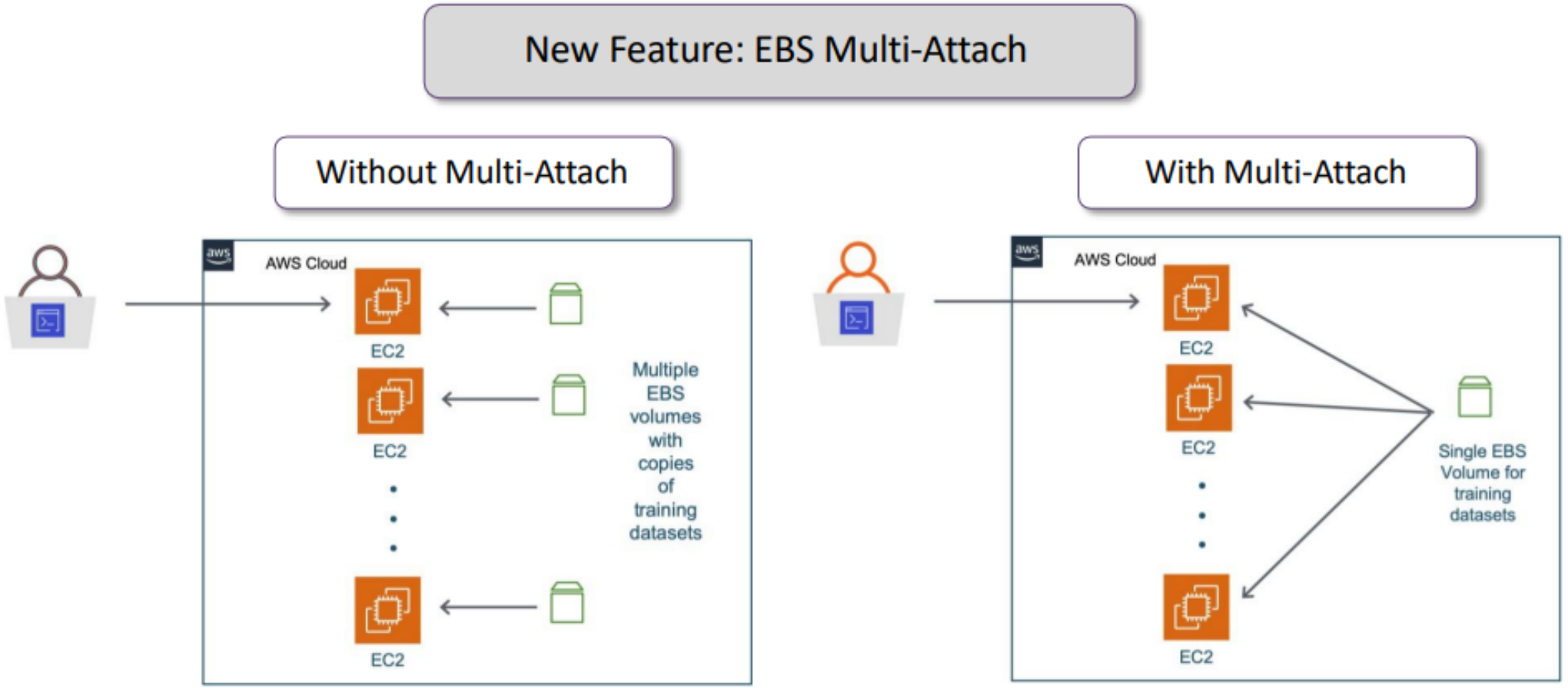
New Feature: EBS Multi-Attach

Amazon EBS Multi-Attach is now available on Provisioned IOPS io1 volumes



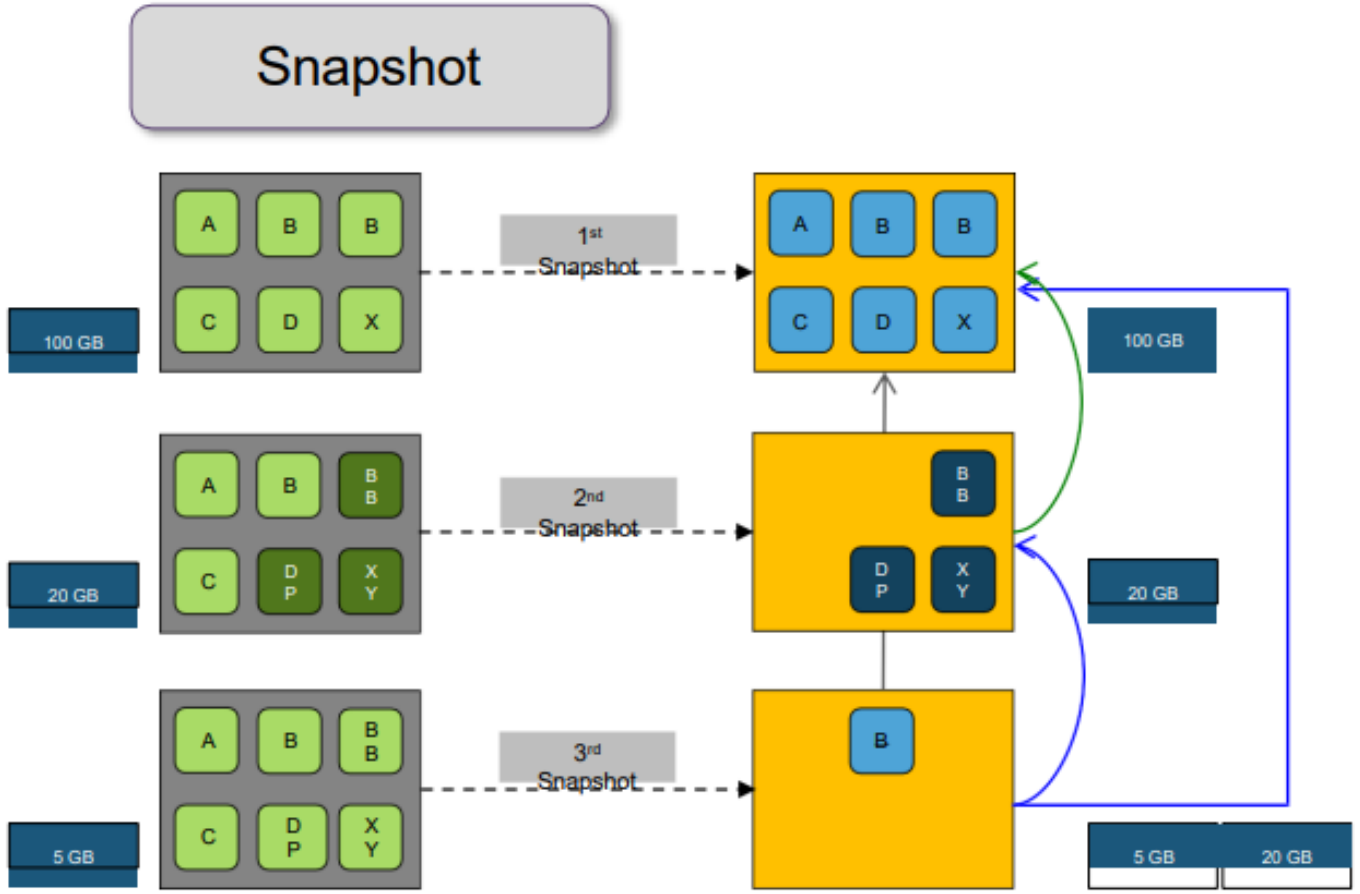
We can now enable Multi-Attach on Amazon EBS Provisioned IOPS io1 volumes to allow a single volume to be concurrently attached to up to 16 AWS Nitro System-based Amazon EC2 instances within the same availability zone

EBS Concepts

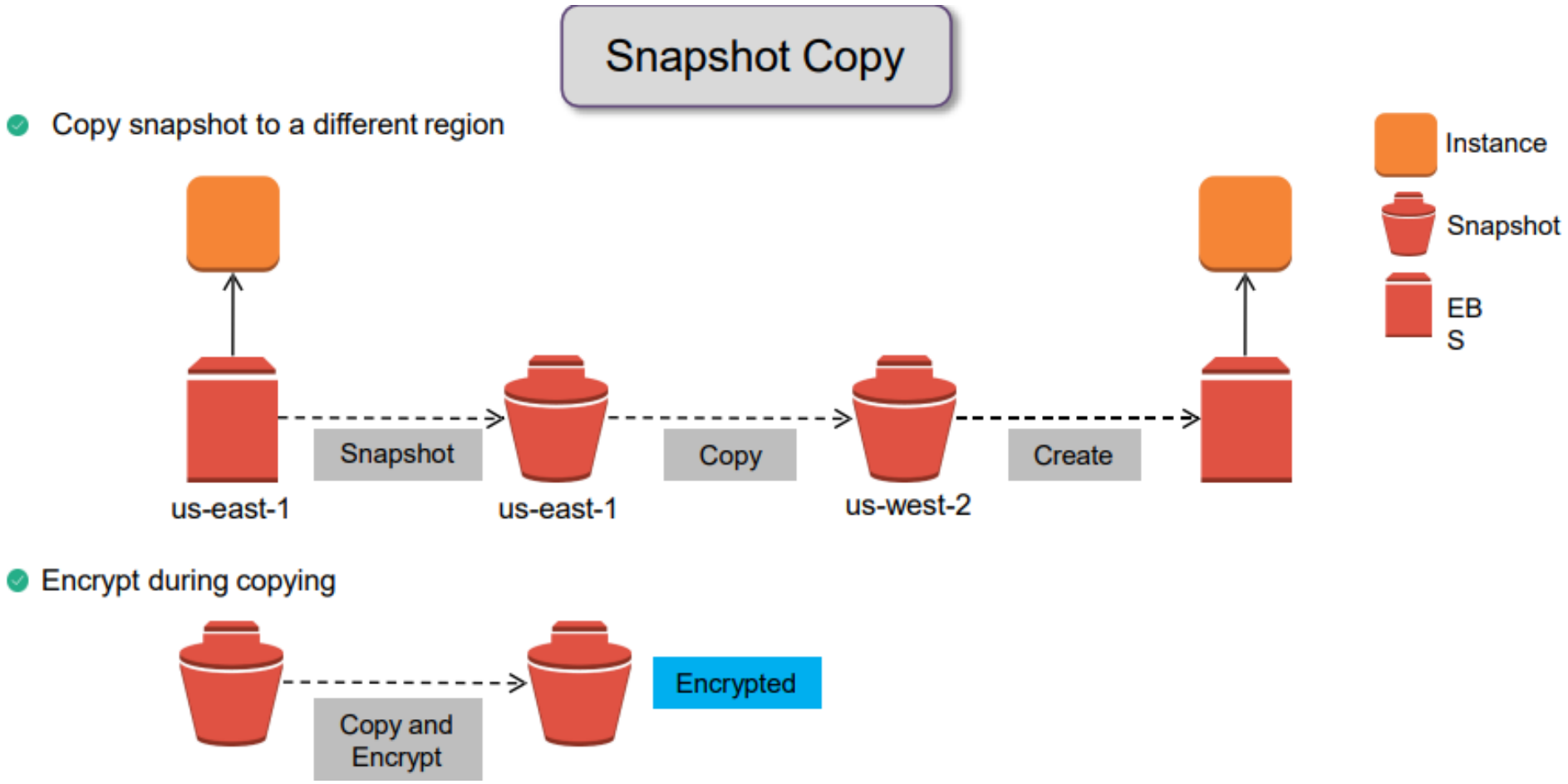


EBS Snapshots

- Snapshots are used to backup data on EBS volumes
- All snapshots are incremental backups except for the first one
- Snapshots are copied to Amazon S3

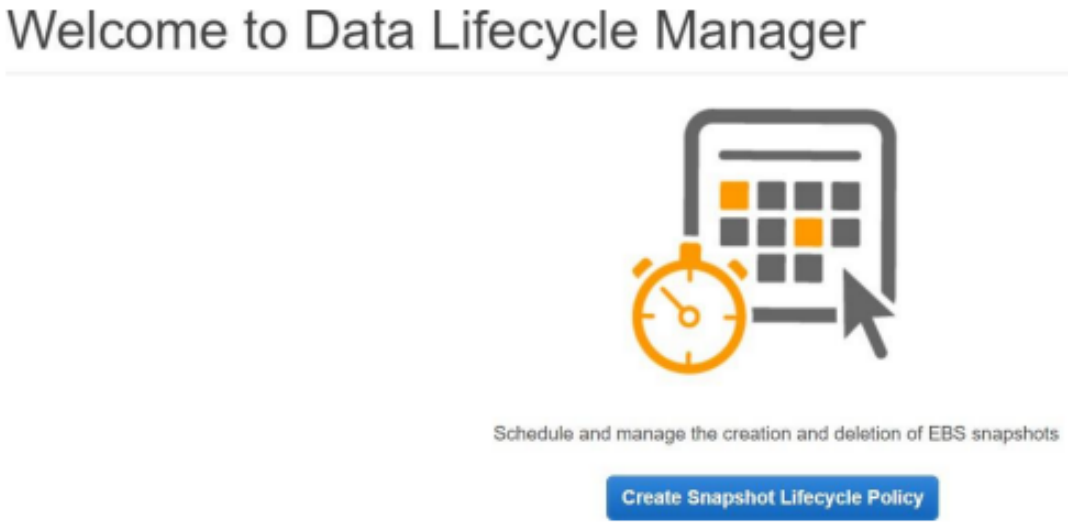


EBS Snapshots



EBS Snapshots

New Feature: Data Lifecycle Manager for Snapshots



- Amazon DLM supports Amazon EBS volumes and snapshots
- We can define backup and retention schedules for EBS snapshots by creating lifecycle policies based on tags
- It is free to use
- We no longer need to create custom scripts for backup and restore

EBS Snapshots

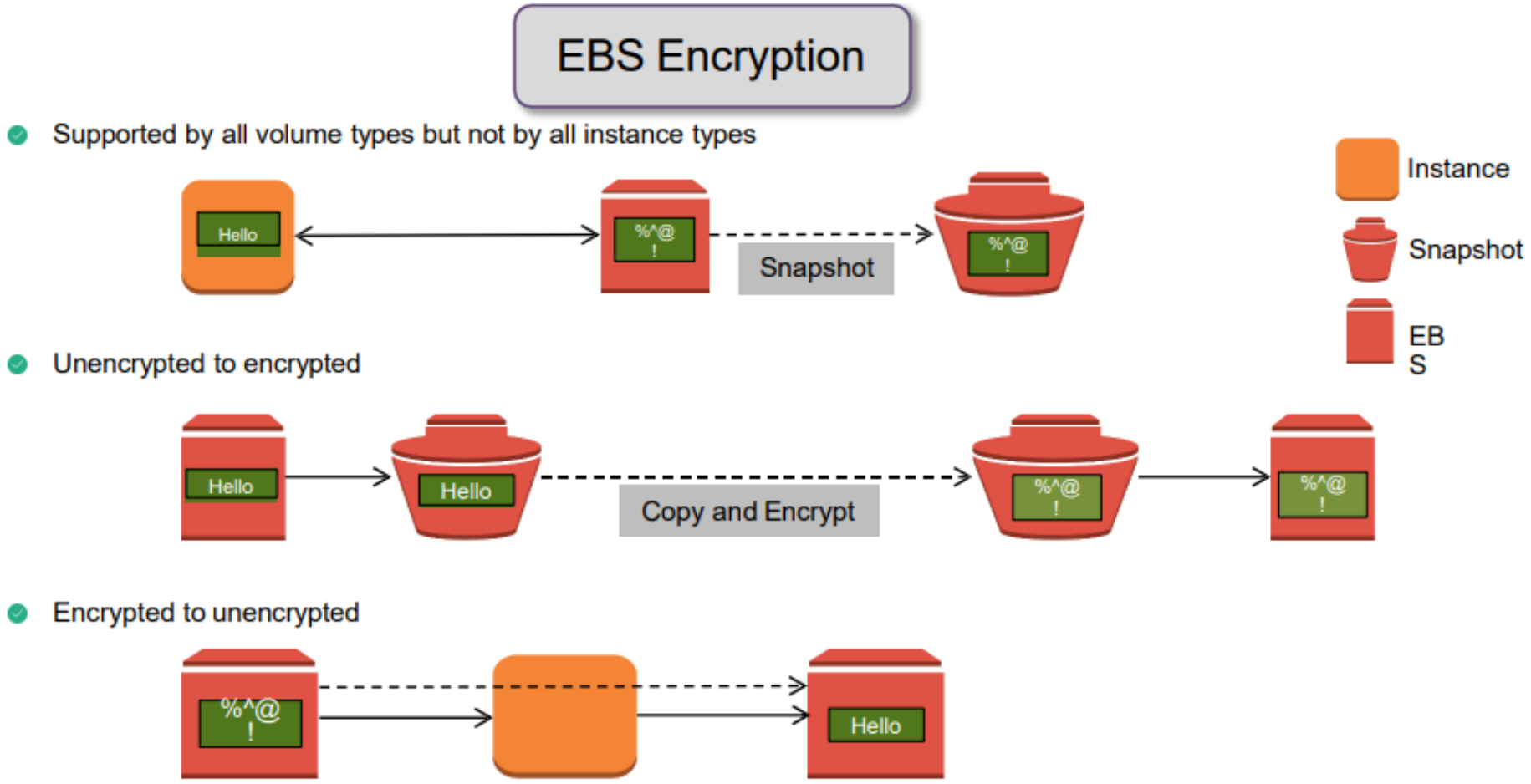
Automating the snapshot cycle helps with:

- Protecting valuable data by enforcing a regular backup schedule
- Retaining backups as required by auditors or internal compliance
- Reducing storage costs by deleting outdated backups

Quotas for AWS DLM:

- We can create up to 100 lifecycle policies per region
- We can add up to 45 tags per resource
- We can create one schedule per lifecycle policy

EBS Snapshots



Introduction to EFS

Amazon Elastic File System

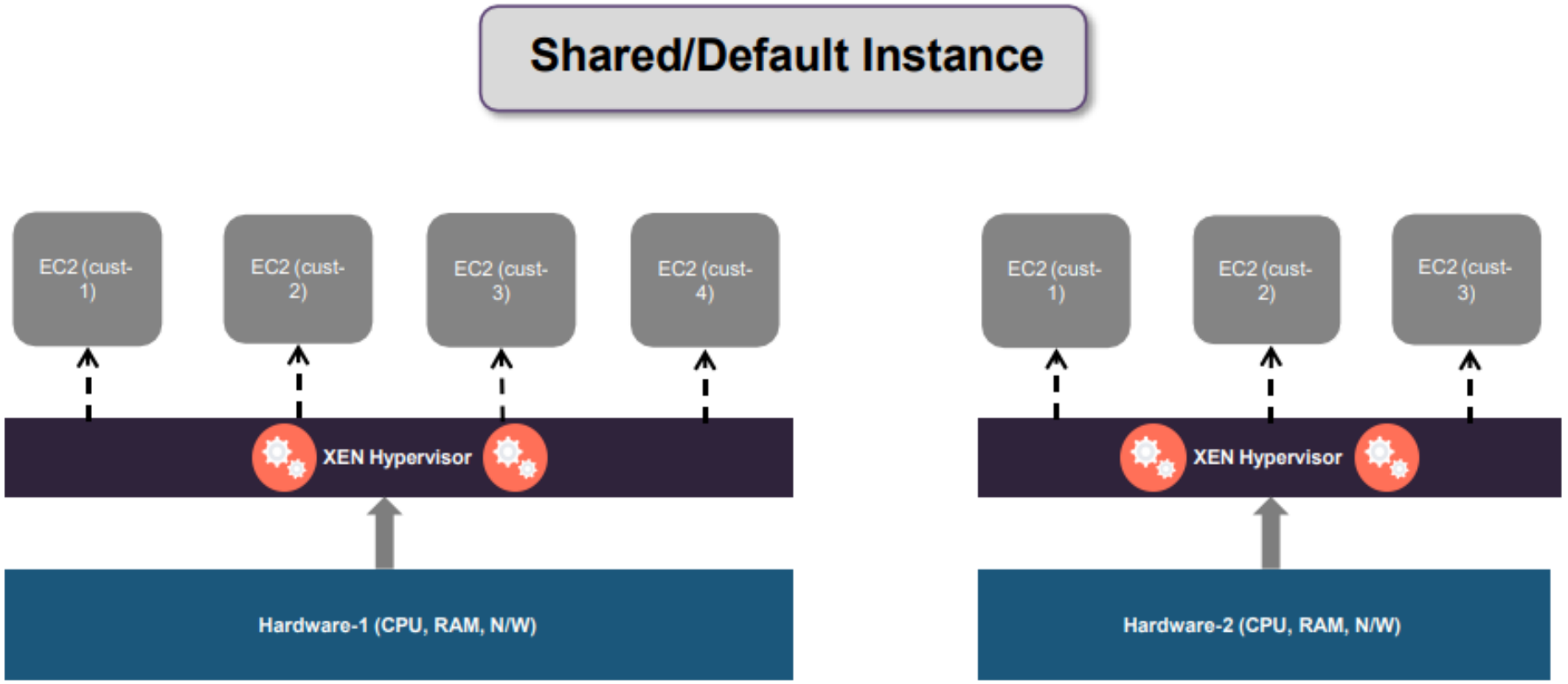
Amazon EFS (**Elastic File System**) is a cloud-based file storage service for applications and workloads that run in the **Amazon Web Services** (**AWS**) public cloud

Why do we need EFS?

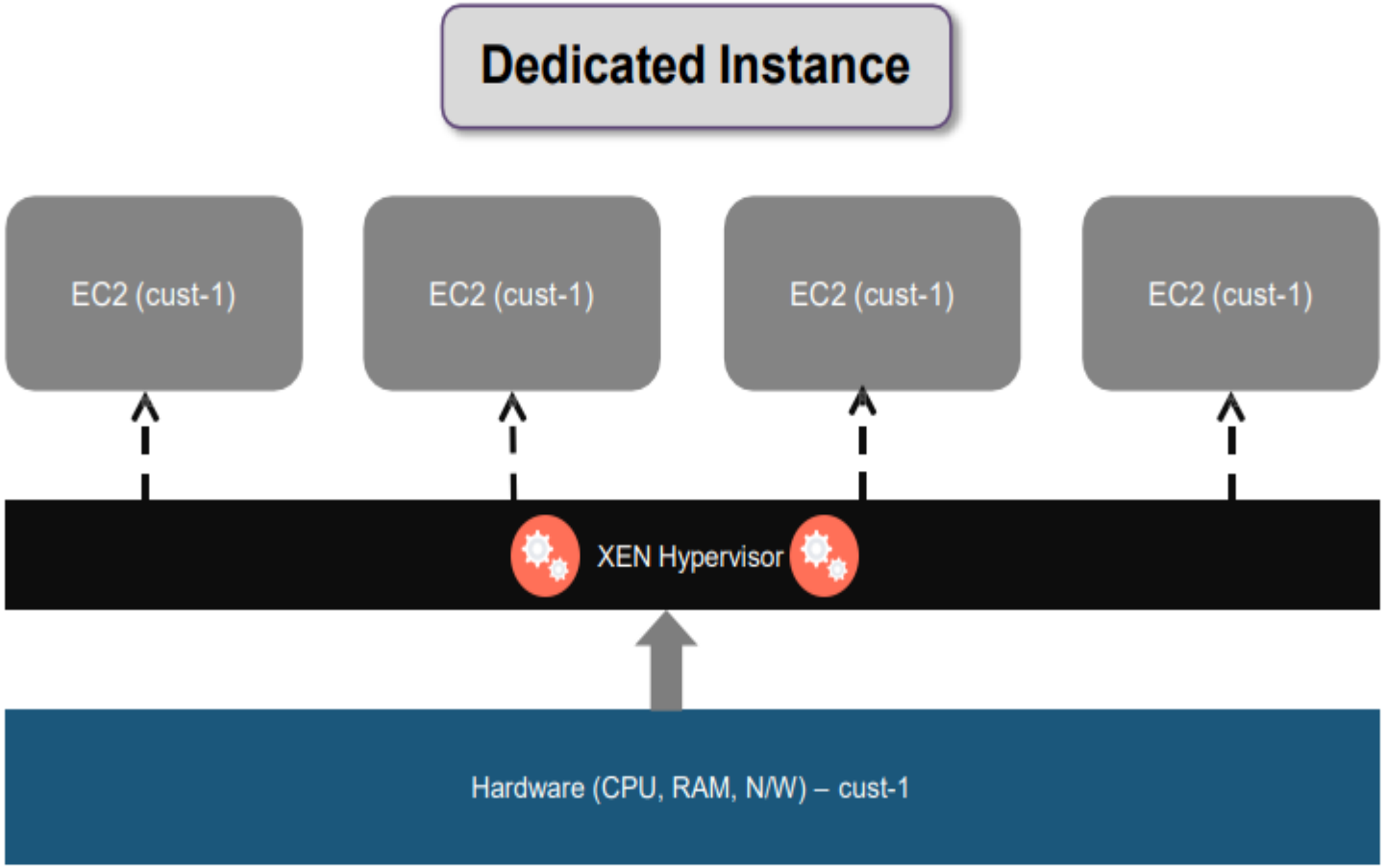
If our application is running on Amazon EC2 and needs a file system or in any use case where a file system is needed

EFS Benefits

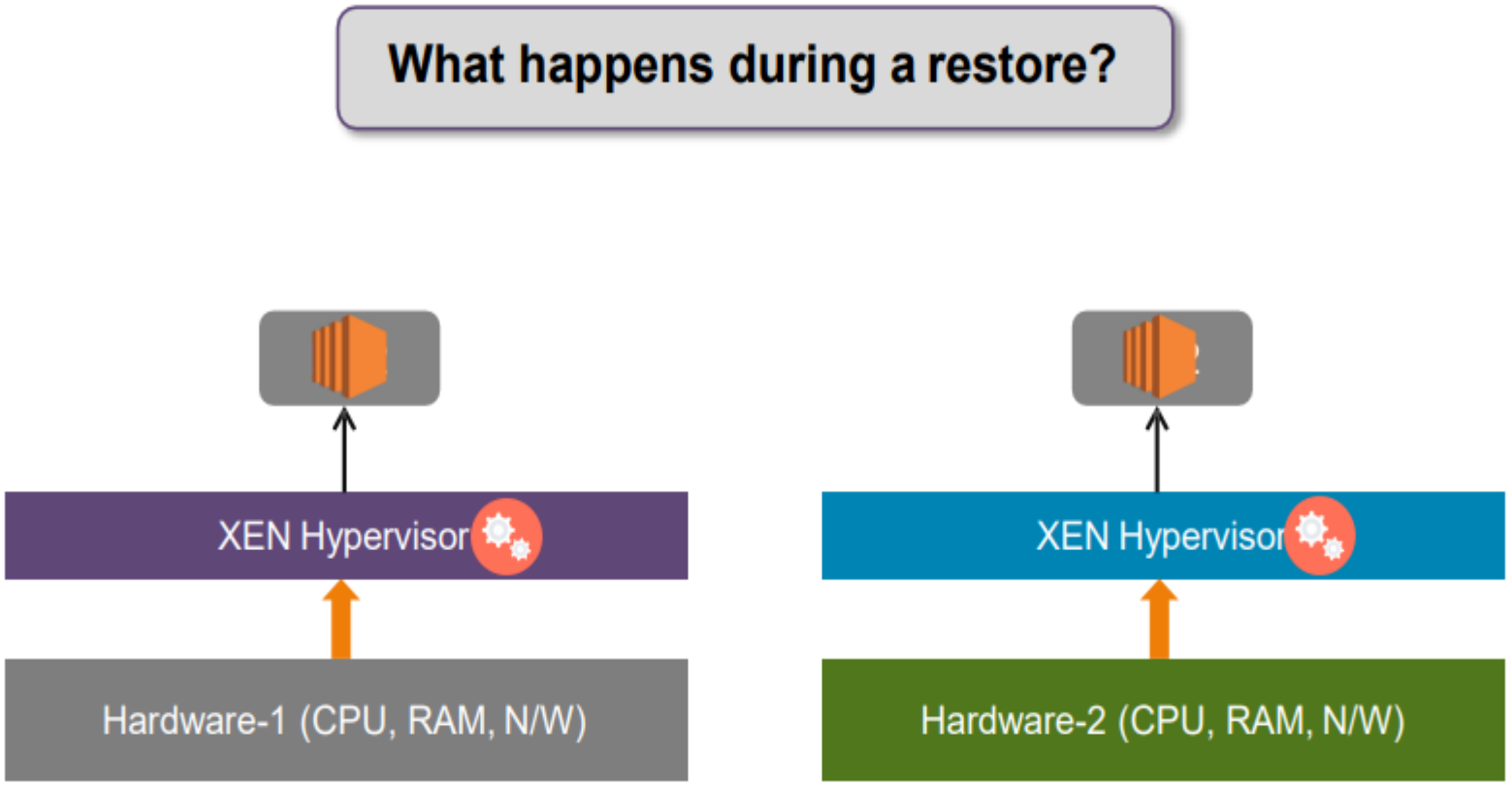
Instance Tenancy



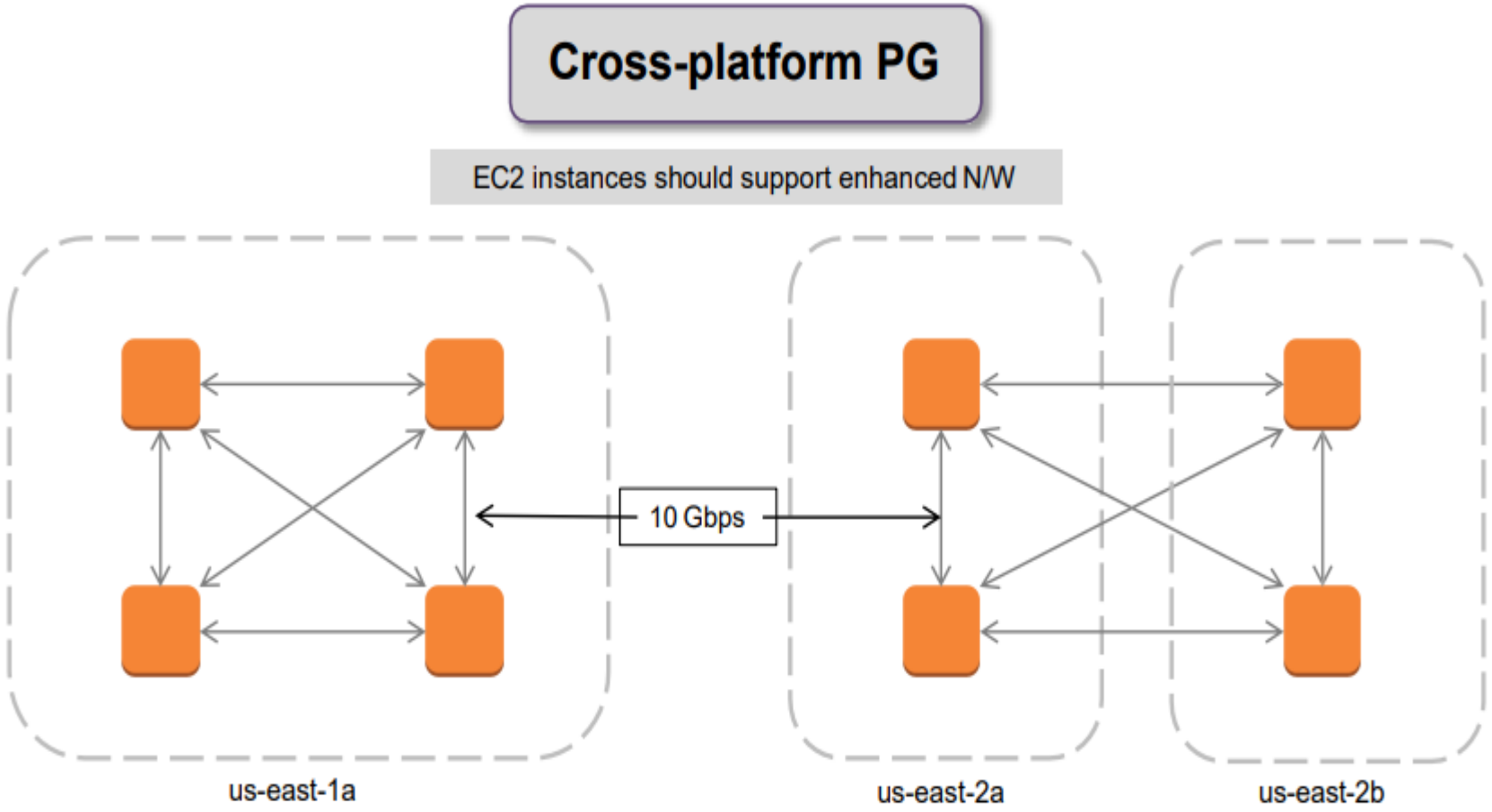
Instance Tenancy



Instance Restart



Placement Group



Reserved and Spot instances

Reserved Instances

- Regional RI – AZ and Instance Size Flexibility (Both default and dedicated tenancy)
- Resources and capacity is reserved until the contract period ends
- Scheduled RI

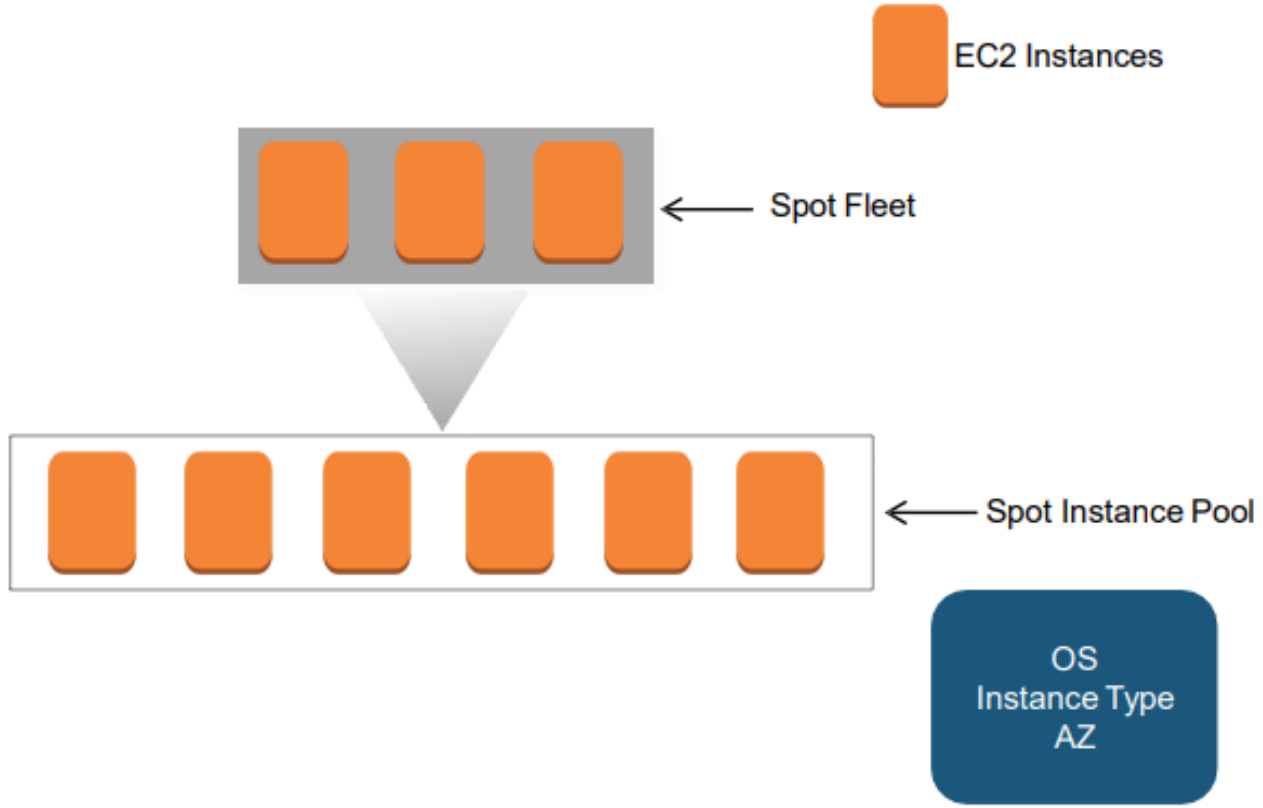
Running Instance	RI bought
4 m3.large Linux, default tenancy in AZ us-east-1a	4 m3.large, Linux, default tenancy, AZ us-east-1a
2 m4.4xlarge Amazon Linux, default tenancy in us-east-1b	4 m4.large, Amazon Linux, default tenancy, region us-east-1
c4.xlarge RHEL dedicated tenancy in AZ us-east-1c	C4.large, RHEL, default tenancy, region us-east-1

Instance size	Normalization factor
nano	0.25
micro	0.5
small	1
medium	2
large	4
xlarge	8
2xlarge	16
4xlarge	32
8xlarge	64
9xlarge	72
10xlarge	80
12xlarge	96
16xlarge	128
18xlarge	144
24xlarge	192
32xlarge	256

Reserved and Spot instances

Spot Instances

- Unused EC2 instances available for lesser price than the on-demand price
- Instances are terminated if the spot price increases than the bid price
- Significant price reduction



Pricing

EC2 Pricing (us-east-1)

- Pay as you use
- Free Tier: 750 hours per month of Amazon Linux, RHEL, SLES, Windows t2.micro single instance usage

On-demand price:

- m5.large = US$0.096/hour
- c5.large = US$0.085/hour
- r4.large = US$0.133/hour

Data Transfer IN:

FREE from anywhere

SLA = 99.99% Uptime

Data Transfer OUT: From EC2 to

- S3, Glacier, DynamoDB, SES, and SQS in same region = FREE
- S3, Glacier, DynamoDB, SES, and
  - SQS in different region =
  - US$0.020/GB
- EC2, RDS, Redshift, Elasticache, ELB, and ENI in same AZ = FREE with private IP and US$0.010/GB with
  - public IP
- EC2, RDS, Redshift, Elasticache, ELB, and ENI in different AZ = US$0.010/GB

EC2 Purchasing Options (RI)

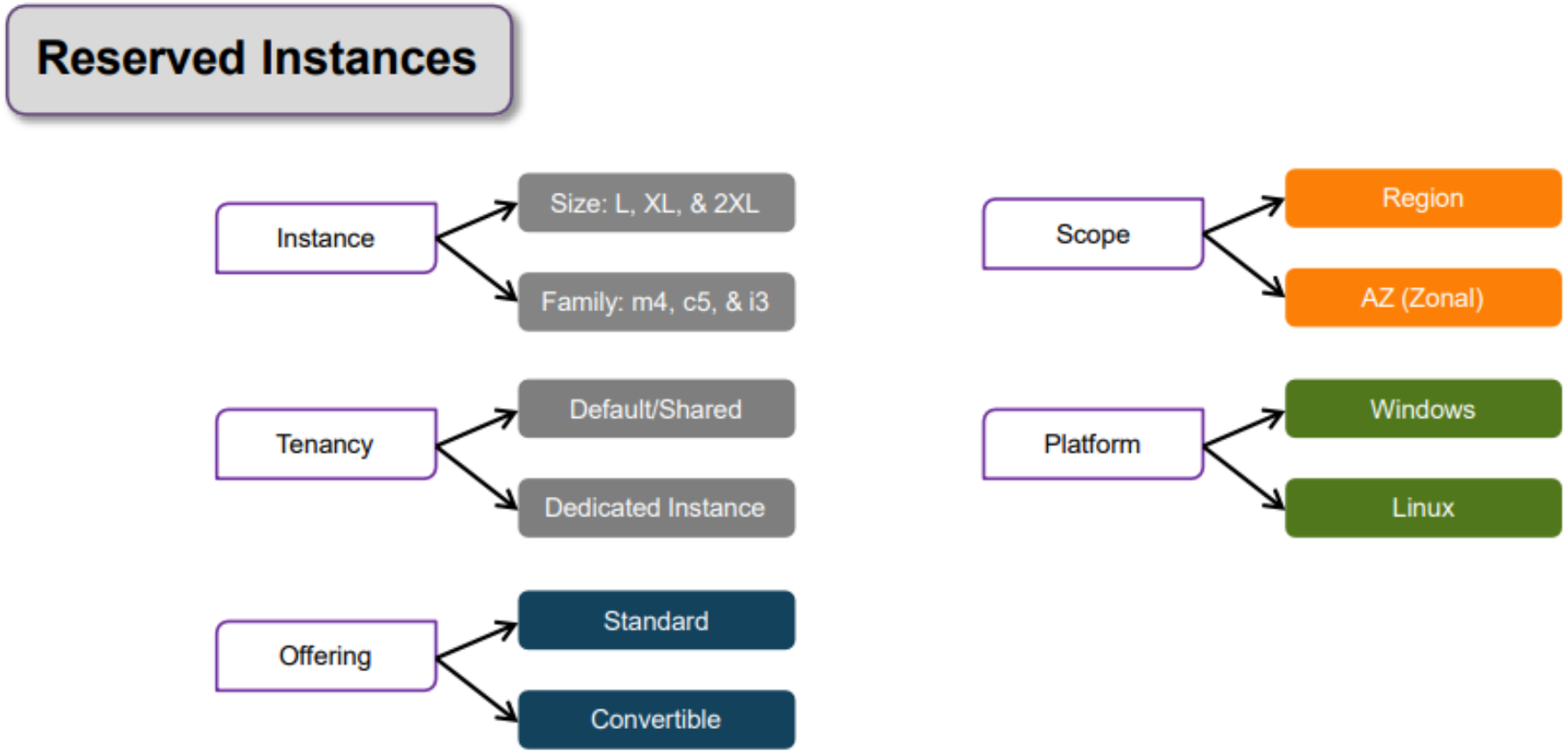
- Reserved Instance: 1 to 3-year terms
- Pricing (on-demand us-east-1 region)

M5.XLARGE = US$0.192/hr

Yearly = US$1681.92

Payment Type	One Time Payment	Total Yearly Cost	Savings
No Upfront	US$0	US$89.79*12 = US$1077.48	36%
Partial Upfront	US$512	US512 + (42.34*12) = US$1020.08	39%
Full Upfront	US$1003	US$1003	40%

EC2 Purchasing Options (RI)



EBS Pricing

- gp2: US$0.1 per GB per month
- io1: US$0.125 per GB per month and US$0.065 per provisioned IOPS per month
- st1: US$0.045 per GB per month
- sc1: US$0.025 per GB per month
- EBS snapshot to Amazon S3: US$0.05 per GB per month
- Free Tier: 30 GB/month, a combination of gp2 and magnetic. 2,000,000 IO with magnetic, 1 GB of snapshot storage
- Visit https://aws.amazon.com/ebs/pricing/ for details

Uptime SLA: 99.99%

Design Patterns