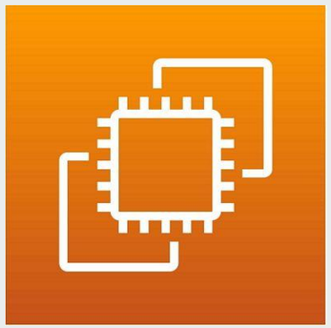


EC2 Overview

Hieu Dinh

Objectives

At the end of this module, you should be able to:

- Identify what is Amazon EC2
- EC2 Terminology
- EC2 Design
- Identify how and when to use EC2
- Basic setup an EC2 instance

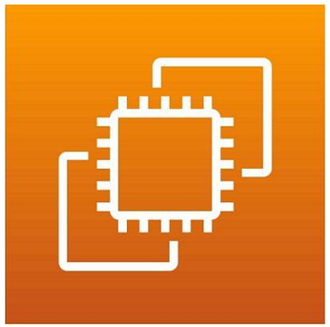
Agenda

1. EC2 Overview

2. EC2 Design

Labs

Choices for Compute



Amazon EC2

Virtual server instances
in the cloud

Amazon ECS, EKS, and Fargate

Container management
service for running
Docker on a managed
cluster of EC2

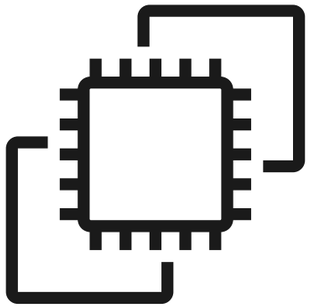
AWS Lambda

Serverless compute
for stateless code execution
in response to triggers

Section 1:

EC2 Overview

Amazon EC2



Amazon EC2

Linux | Windows (OS)

Arm and x86 architectures

General purpose and workload optimized

Bare metal, disk, networking capabilities

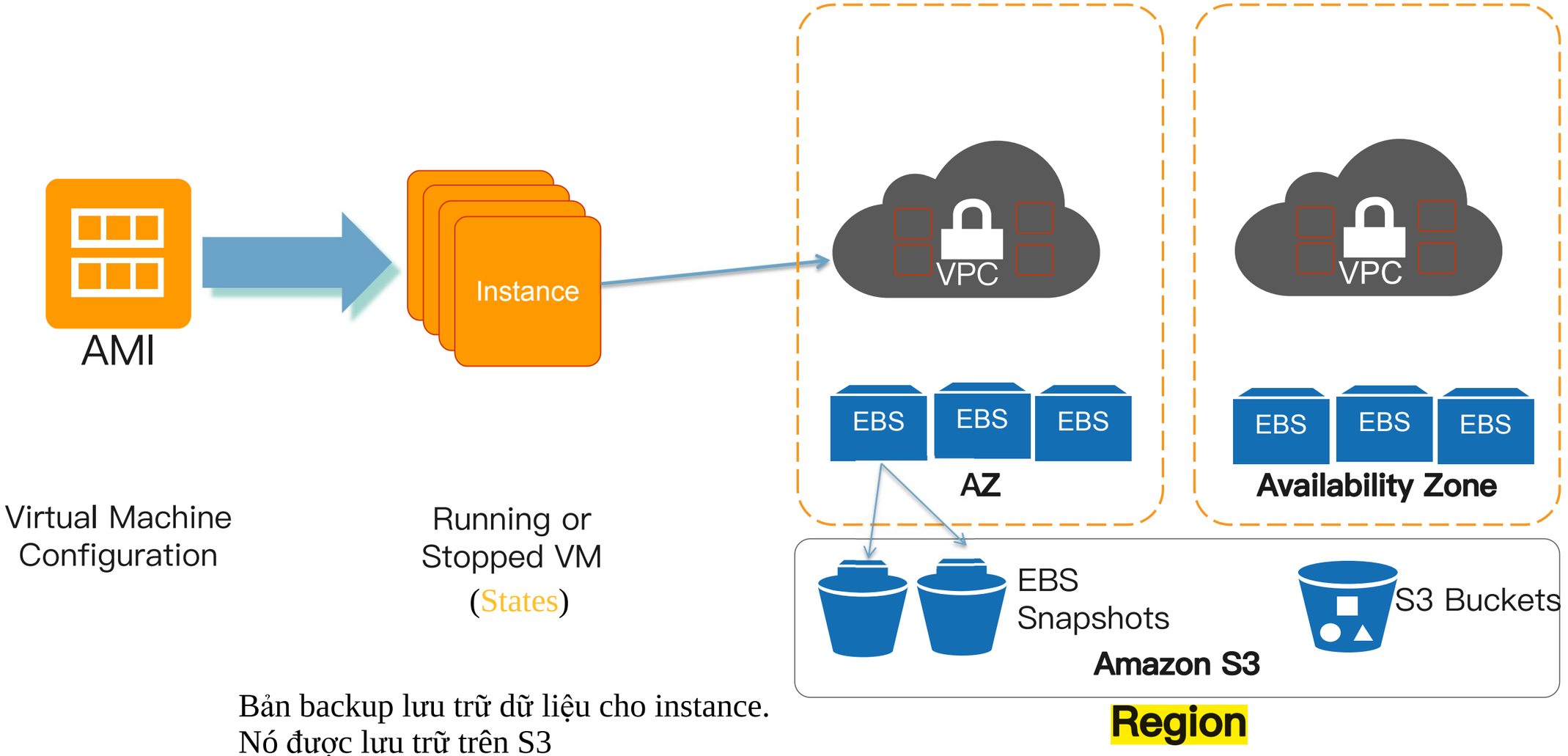
Packaged | Custom | Community AMIs

Multiple purchase options: On-demand, RI, Spot

Amazon Machine Image

A template that contains a software configuration

EC2 Terminology



Virtual Machine Configuration

Running or Stopped VM (States)

Bản backup lưu trữ dữ liệu cho instance.
Nó được lưu trữ trên S3

What’s a virtual CPU? (vCPU)

- A vCPU is typically a hyper-threaded physical core*
- Divide vCPU count by 2 to get core count
- On Linux, “A” threads enumerated before “B” threads
- On Windows, threads are interleaved xen kẽ

- Cores by Amazon EC2 & RDS DB Instance type: https://aws.amazon.com/ec2/virtualcores/

** CPU Optimizing options allow disabling hyperthreading and reduce number of cores*

Memory and Storage

What's a GiB?

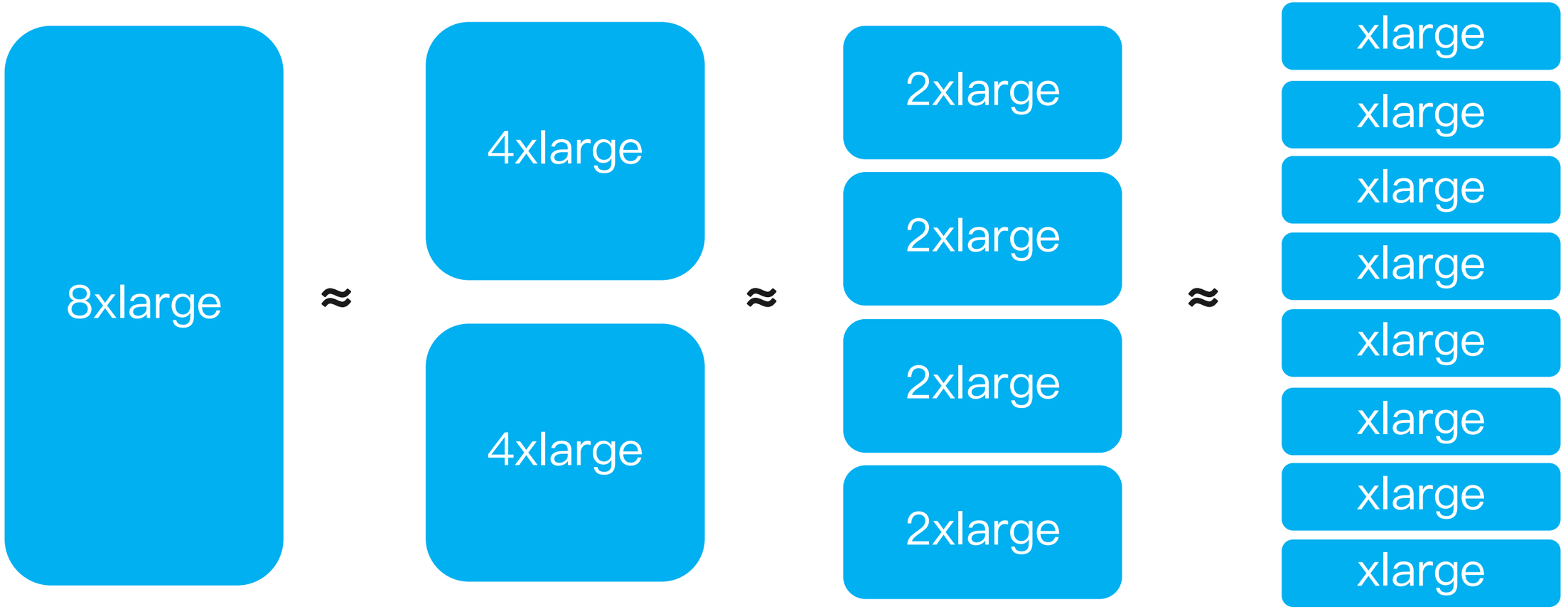
- Memory is presented as GibiBytes (GiB) and not Gigabytes (GB)
- 256 GiB = 275 GB

What about storage?

- Storage is independent of compute
- You allocate drives known as EBS volumes
- Max 16 TiB per volume
- Some instance types provide physically attached (ephemeral) storage

1 EC2 : nhiều EBS
1 EBS - 1 EC2

Instance sizing



c4.8xlarge

2 – c4.4xlarge

4 – c4.2xlarge

8 – c4.xlarge

Resource allocation

- All resources assigned to you are dedicated to your instance with no over commitment*
  - All vCPUs are dedicated to you
  - Memory allocated is assigned only to your instance
  - Network resources are partitioned to avoid “noisy neighbors”
- Curious about the number of instances per host?
  - See “Dedicated Hosts Configuration Table” for a guide.

+ Cơ chế cho phép chạy nhiều instances trên 1 cơ sở vật lý

*Again, the “T” family is special

EC2 Naming Explain

Instance generation

Tip: nên sử dụng các thế hệ mới nhất để tối ưu chi phí

+ (additional) phần được thêm mới nhằm cải thiện 1 attribute nào đó
vd: thể hiện attribute networking được cải thiện

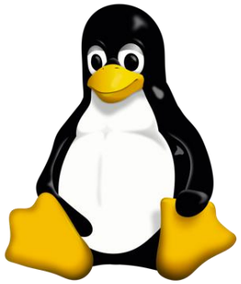
Refer: https://aws.amazon.com/ec2/instance-types/?nc1=h_ls

EC2 Instance Types

	General Purpose		Compute Optimized		Memory Optimized				Accelerated Computing			Storage Optimized		
	Burstable performance	General Purpose	Compute Intensive	Compute +memory up to 100 Gbps	Memory Optimized	In-memory	Memory Intensive	Compute and Memory Intensive	Graphics Intensive	General Purpose GPU	FPGA	High I/O	Dense Storage	Big Data Optimized
intel	T3	M5	C5	C5n	R5	X1	X1e		G3	P3	F1		D2	
Local storage (NVMe SSD)		M5d	C5d		R5d			Z1d				I3		
AMD	T3a	M5a			R5a									
metal		M5m	c5m		R5m		u-12tb1	Z1dm				I3m		
others	A1	M6g	C6g		R6g					P3dn		I3en		

EC2 Operating Systems Supported

- Windows 2003R2*/2008*/2008R2*/2012/2012R2/2016/2019
- Amazon Linux
- Debian
- SUSE
- CentOS
- Red Hat Enterprise Linux
- Ubuntu



for more OSes see: https://aws.amazon.com/marketplace/b/2649367011

Unix: CLI & Window: GUI

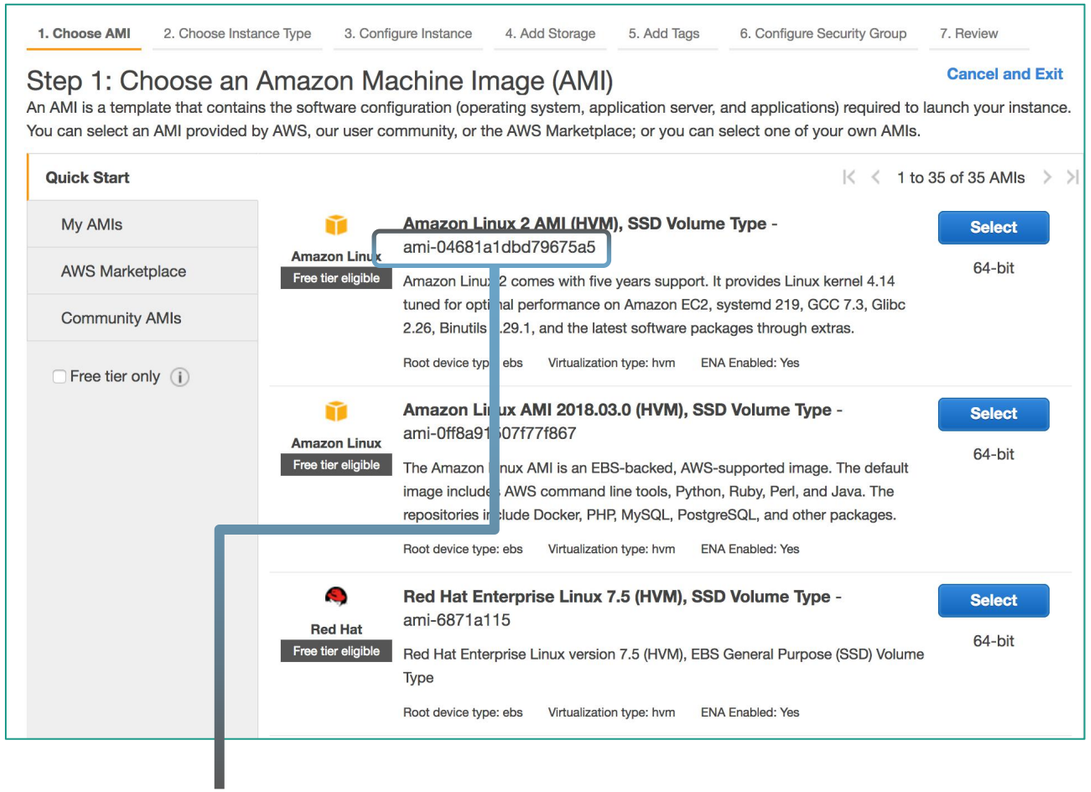
What is an Amazon Machine Image (AMI)

- Provides the information required to launch an instance
- Launch multiple instances from a single AMI
- An AMI includes the following
  - A template for the root volume (for example, operating system, applications)
  - Launch permissions that control which AWS accounts can use the AMI
  - Block device mapping that specifies volumes to attach to the instance

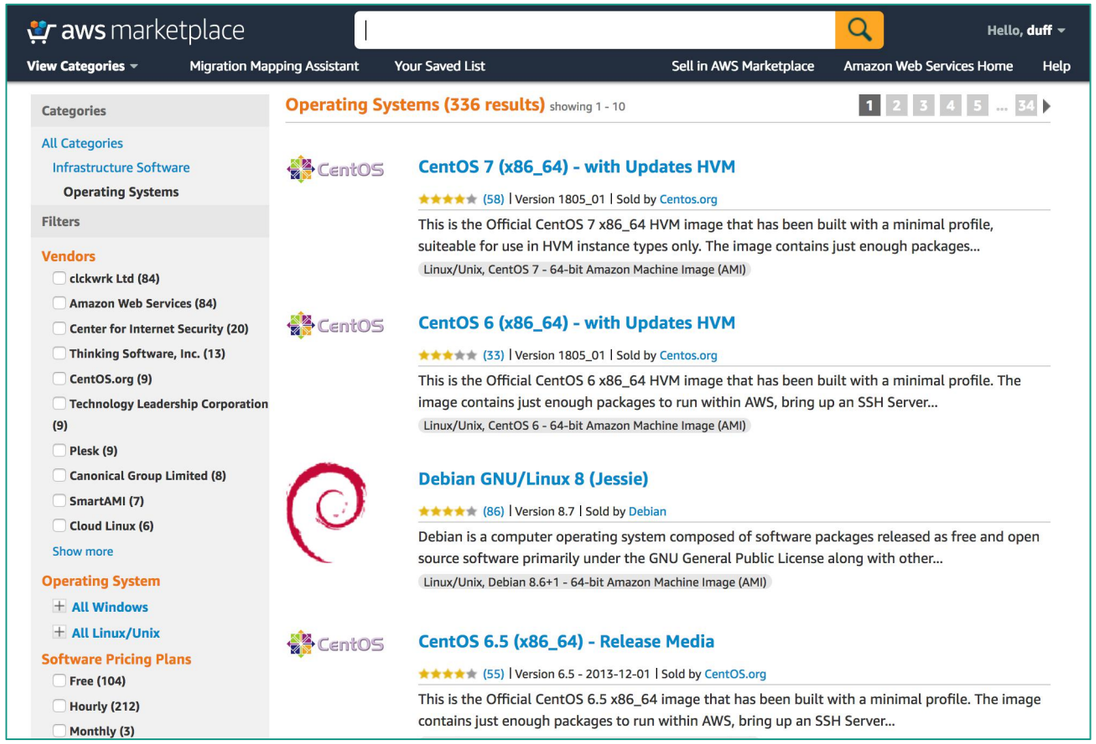
Chỉ định một vùng nhớ cố định

Choosing an AMI

AWS Console



AWS Marketplace



Use the AMI ID to launch through the API or AWS Command Line Interface (AWS CLI)

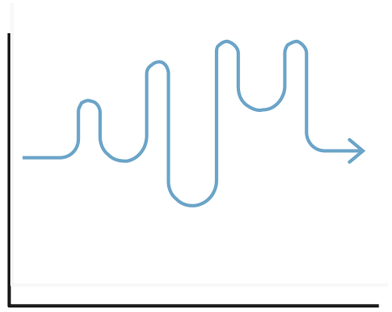
```
aws ec2 run-instances --image-id ami-04681a1dbd79675a5 --instance-type c4.8xlarge --count 10 --key-name MyKey
```

Amazon EC2 Purchase Options

Có thể mix các instances với nhau

On–Demand

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

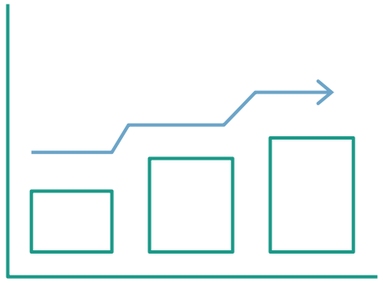


Spiky workloads, to define needs

tư bản chủ nghĩa. Sử dụng khi:
+ Thử nghiệm nhu cầu

Reserved Instances

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



Committed and steady–state usage

Ăn buffer.
+ Ưu: chi phí < On-demand (~ 70%)
+ Nhớ ắn hết
Sư dụng:
+ ứng dụng ổn định, k scale

Savings Plan

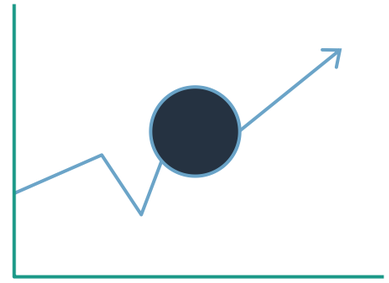
Same great discounts as Amazon EC2 RIs with **more flexibility**

Committed flexible access to compute

Commit theo giờ. Types:
+ Saving plan for compute
  - cover: ec2 + compute (lambda ...)
+ Saving plan for ec2
  - cover: ec2

Spot Instances

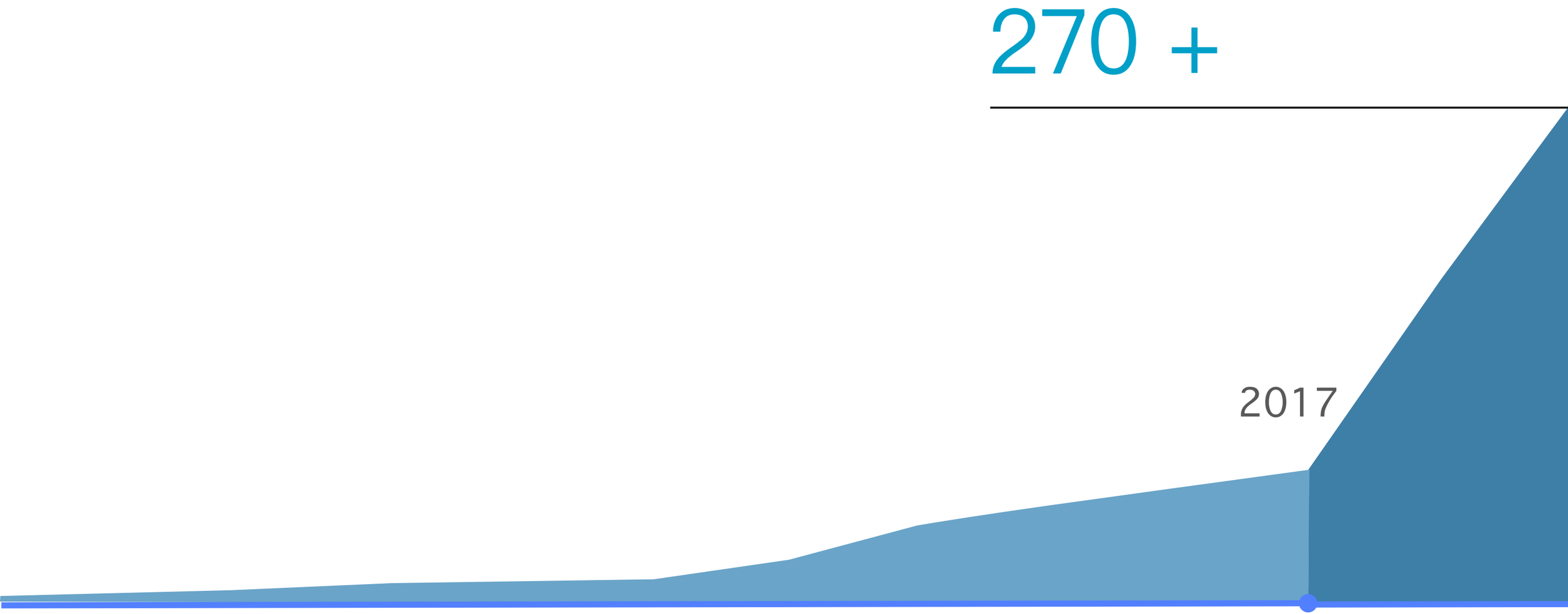
Spare Amazon EC2 capacity at **savings of up to 90%** off On–Demand prices



Fault–tolerant, flexible, stateless workloads

tận dụng các instance rảnh nhưng hạn chế là dễ bị thu hồi bất kì lúc nào.
Sử dụng:
+ testing

270+ instances across 42 instance Families



Section 2:

EC2 Design

Which hypervisor do AWS use?

+ quản lý máy ảo

- **Original host architecture: Xen-based**
  - Hypervisor consumed resources from the underlying host
  - Limited optimization

- **AWS Nitro Hypervisor: Custom KVM based hypervisor**
  - AWS Nitro System (launched on Nov 2017)
  - Less server resources used, more resources for the customer
  - AWS optimized

Rẻ

CPU

- **Bare metal**: **Direct access to processor and memory resources**
  - Built on the AWS Nitro system
  - Enables custom hypervisors and micro-VM runtimes

cho phép tự ảo hóa theo nhu cầu

AWS Nitro System

Nitro Card

Local NVMe storage

Elastic Block Storage

Networking, monitoring, and security

Nitro Security Chip

Integrated into motherboard

Protects hardware resources

Nitro Hypervisor

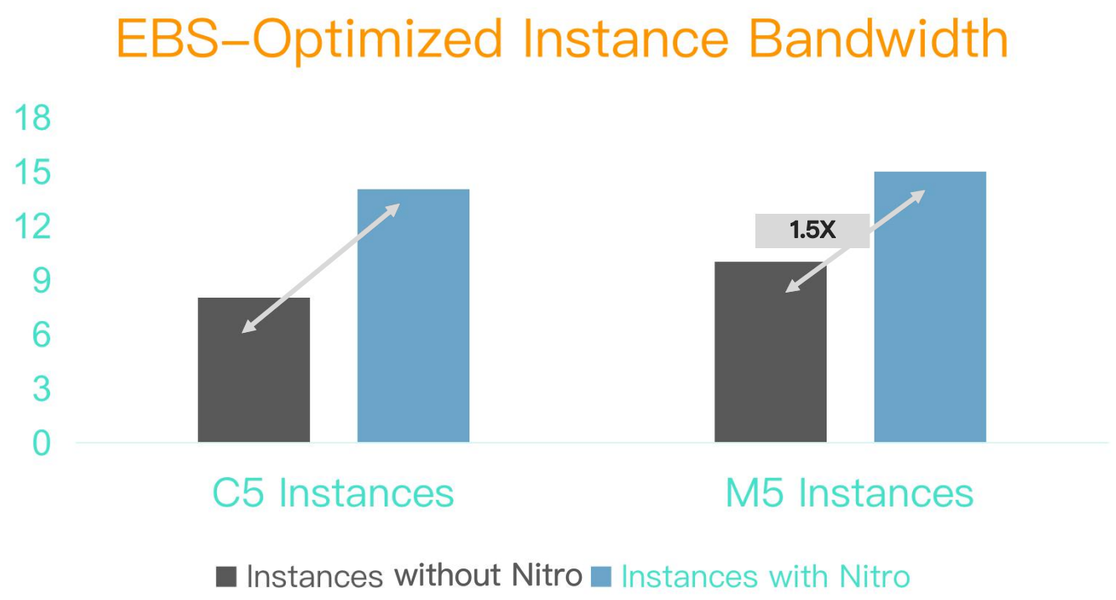
Lightweight hypervisor

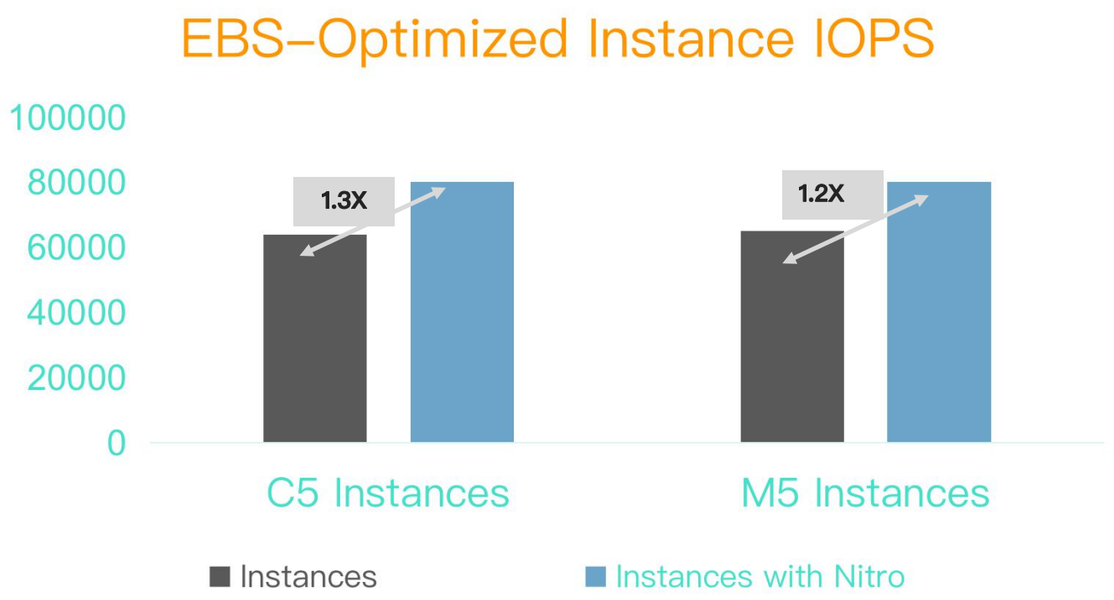
Memory and CPU allocation

Bare metal-like performance

Modular building blocks for rapid design and delivery of EC2 instances

AWS Nitro System



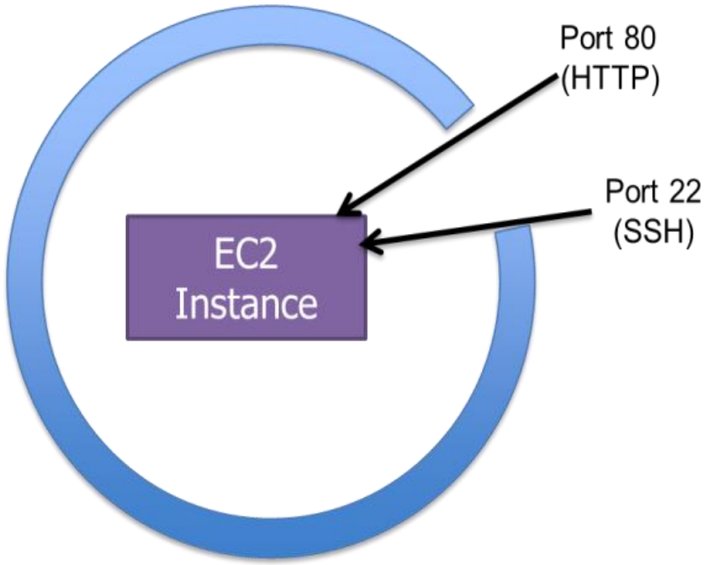


Nitro instances provide **bandwidth, performance, and price improvements** over previous instance generations

EC2 Security Groups

Security Group Rules

- Name
- Description
- Protocol
- Port range
- IP address, IP range, Security Group name

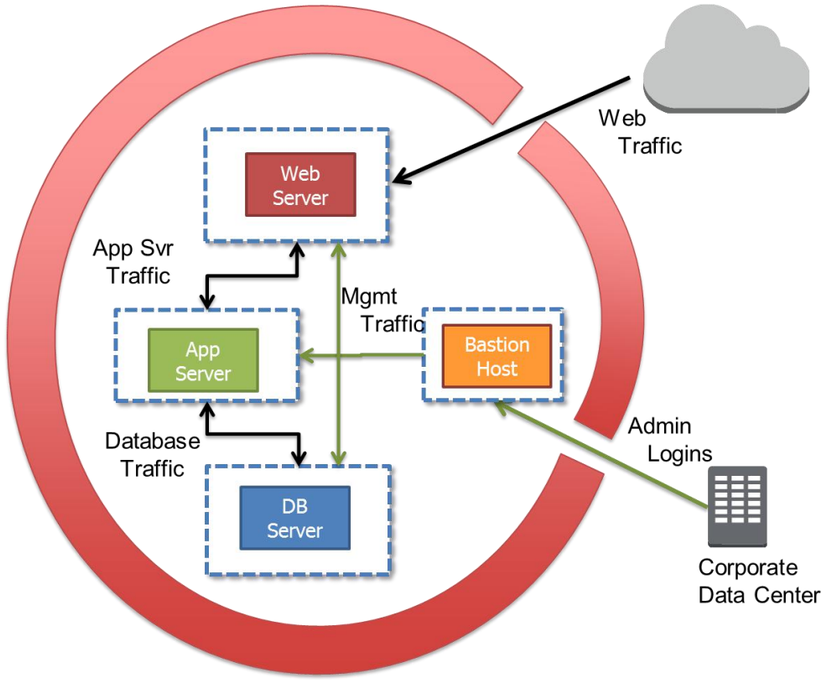


Best practice: hạn chế để resource trên public subnet

Tiered EC2 Security Groups

Hierarchical Security Group Rules

- Dynamically created rules
- Based on Security Group membership
- Create tiered network architectures

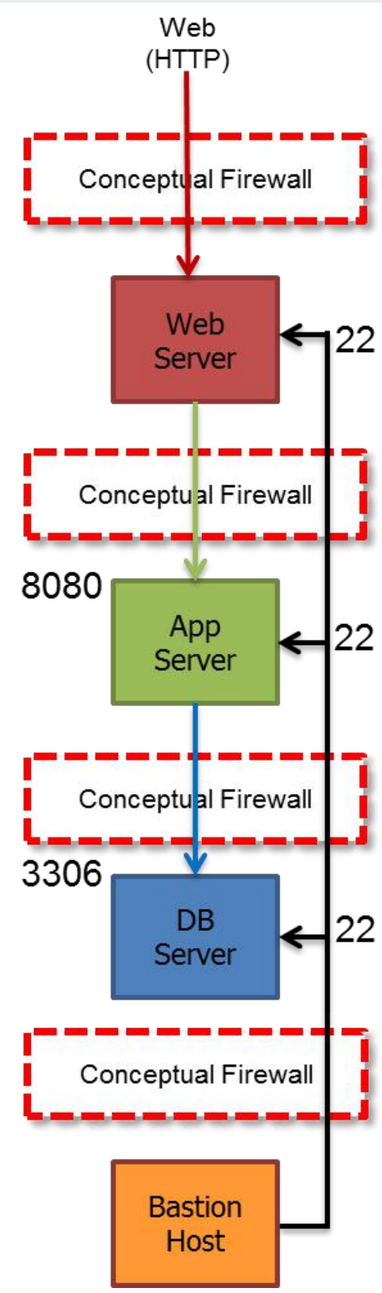


"Web" Security Group:
TCP 80 0.0.0.0/0
TCP 22 "Mgmt"

"App" Security Group:
TCP 8080 "Web"
TCP 22 "Mgmt"

"DB" Security Group:
TCP 3306 "App"
TCP 22 "Mgmt"

"Mgmt" Security Group:
TCP 22 163.128.25.32/32



EC2 IP Addressing

Default VPC	Virtual Private Cloud
Dynamic Private IP	Dynamic or Static Private IP Address
Dynamic Public IP	None by default (can be created with publicIP=true)
Optional Static Public IP (EIP)	Optional Static Public IP (EIP), BYOIP
AWS-provided DNS names • Private DNS name • Public DNS name	AWS-provided public DNS lookup AWS-provided private DNS names Customer-controlled DNS options

EC2-Specific Credentials

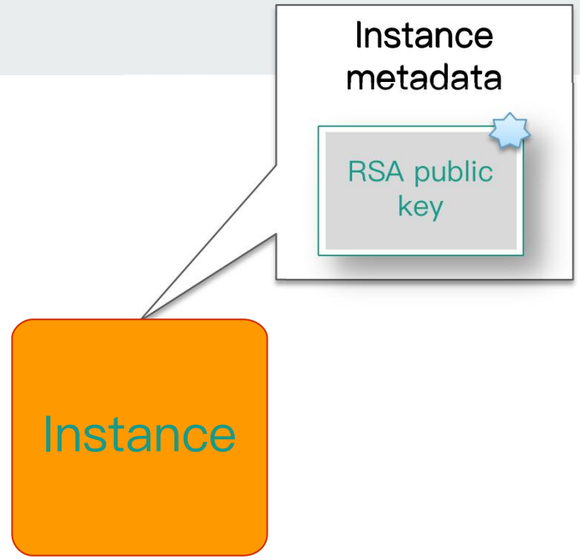
- EC2 key pairs
  - Linux – SSH key pair for first-time host login
  - Windows – Retrieve Administrator password
- Standard SSH RSA key pair
  - Public/Private Keys
  - Private keys are not stored by AWS
- AWS approach for providing **initial** access to a generic OS
  - Secure
  - Personalized
  - Non-generic (NIST, PCI DSS)

Nếu như mất key pairs, ta cần thay mới

EC2 Instance access and Key Pairs

Linux launch (first boot)

- **Public key** made available through metadata
- Public key inserted into ~/.ssh/authorized_keys
- User connects with SSH using their **private key**



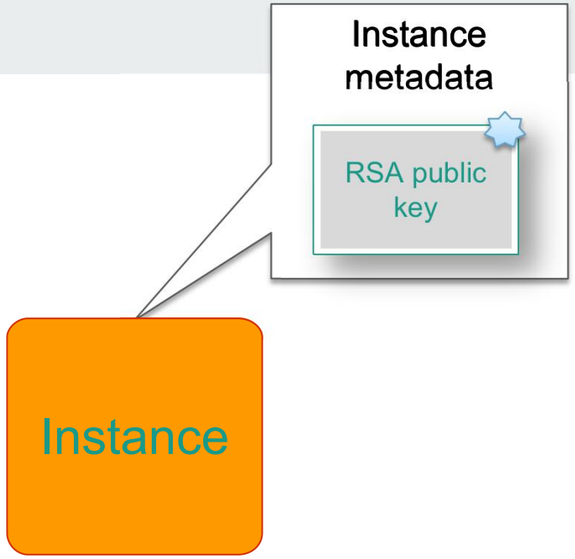
EC2 Instance access and Key Pairs

Linux launch (first boot)

- **Public key** made available through metadata
- Public key inserted into ~/.ssh/authorized_keys
- User connects with SSH using their **private key**

Windows launch (first boot sequence)

- **Public key** made available through metadata
- Sysprep
- Random Administrator password
- Password encrypted with public key
- User decrypts password with their **private key**



EC2 Instance Metadata

Cung cấp dữ liệu mô tả về EC2 instance

http://169.254.169.254/latest/meta-data/ contains a wealth of info

- ami-id
- ami-launch-index
- ami-manifest-path
- block-device-mapping/
- hostname
- instance-action
- **instance-id**
- instance-type
- kernel-id
- local-hostname
- local-ipv4
- mac
- network/
- placement/availability-zone
- profile
- public-hostname
- public-ipv4
- public-keys/

EC2 Userdata

- It is possible to bootstrap our instance using an EC2 User data script
- Bootstrapping means launching commands when a machine starts
- That script is only run once at the instance first start
- EC2 user data is used to automate boot tasks such as:
  - Installing updates
  - Installing software
  - Downloading common files from the internet
  - Anything you can think of

Any Question?

Section 3:

Summary

Summary

At the end of this module, you should be able to:

- Identify what is Amazon EC2
- EC2 Terminology
- EC2 Design
- Identify how and when to use EC2
- Basic setup an EC2 instance

References Document

- AWS EC2 – Resource page
- What is Amazon EC2?
- AWS EC2 FAQs

Thank you