

10.2 AWS Elastic Compute Cloud AWS EC2

In Amazon Web Service (AWS) cloud environment, **Elastic Compute Cloud (EC2)** is a service. Using EC2 you can create a virtual machine which is known as EC2 Instance. AWS provides a virtual environment where you create an EC2 Instance which replaces your hardware need. Using EC2 you can save a huge investment cost to run your physical environment. You can develop and deploy web applications in less time. EC2 created instances are scalable to handle changes and provide multiple layers of security.

As we know that AWS Cloud Computing provides delivery of IT resources on-demand, and it is based on pay-as-you-go pricing model. If you compare running cost of your physical environment with AWS cloud services, it is very economical and provides many features as well.

1. Preconfigured Machine Template – AMI:

In AWS Cloud Computing, you create instances by using templates of instances, which is known as Amazon Machine Image (AMI). By using a preconfigured template of a machine, you can create numbers of Instances of similar instances in very less time. AMI contains image of a running an Operating System (OS) such as Linux or Window OS. In AMIs you can also include some specific application and configurations as per requirement.

You can use available AMIs from AWS which are available in four different resources:

- i. **Quick Start:** In this resource, most frequently used AMIs of various OS are available.

For Example:

- a. Amazon Linux AMI 2017.09.1 (HVM), SSD Volume
 - b. Ubuntu Server 16.04 LTS (HVM), SSD Volume
 - c. Amazon RDS
 - d. Microsoft Windows Server 2016 Base with Containers
 - e. Microsoft Windows Server 2016 with SQL Server 2017 Web etc.
- ii. **My AMIs:** You can create your own AMIs to create Instances and you can also share those AMIs to other AWS accounts.
 - iii. **AWS Marketplace:** Some vendors provide their customized or product AMIs, you can search and purchase them. You can search number of developer tools, business software and software infrastructure AMIs. The list of AMIs at AWS Marketplace is being increased regularly.
 - iv. **Community AMIs:** In this resource, you can find number of popular and OpenSource communities. They are providing AMIs of various Operating Systems for example Amazon Linux, CentOS, Fedora, RedHat, Windows, Ubuntu etc.

2. Instance Type:

In order to launch an instance (virtual machine), you do require two things. First is an Operating System which is available for selection in terms of AMI. Second is hardware resource on which this Operating System will be installed, for example a combination of CPU, memory, storage and other resources.

You can select an instance type which suits your requirement to run the instance. After launching an Instance, it exactly looks like a traditional machine. You can take remote of the Instance by using remote protocol such as RDP in Windows or SSH in Linux.

In Instance type you can find various types of combinations of CPU, memory, storage and other resources.

- a. General Purpose (t2)
- b. Memory Optimized (r4 or x1e)
- c. GPU Compute (p2)
- d. Storage Optimized (d2, i2 or i3)
- e. Compute Optimized (c5)

Some Examples of Instance Types:

Model	vCPU	CPU Credits / hour	Mem (GiB)	Storage
t2.nano	1	3	0.5	EBS-Only
t2.micro	1	6	1	EBS-Only
t2.small	1	12	2	EBS-Only
t2.medium	2	24	4	EBS-Only
t2.large	2	36	8	EBS-Only
t2.xlarge	4	54	16	EBS-Only
t2.2xlarge	8	81	32	EBS-Only

Model	vCPU	Mem (GiB)	SSD Storage (GB)	Dedicated EBS Bandwidth (Mbps)
m5.large	2	8	EBS-only	Up to 2,120
m5.xlarge	4	16	EBS-only	Up to 2,120
m5.2xlarge	8	32	EBS-only	Up to 2,120
m5.4xlarge	16	64	EBS-only	2,120
m5.12xlarge	48	192	EBS-only	5,000

m5.24xlarge	96	384	EBS-only	10,000
-------------	----	-----	----------	--------

Model	vCPU	Mem (GiB)	SSD Storage (GB)	Dedicated EBS Bandwidth (Mbps)
x1e.32xlarge	128	3,904	2 x 1,920	14,000
x1e.16xlarge	64	1,952	1 x 1,920	7,000
x1e.8xlarge	32	976	1 x 960	3,500
x1e.4xlarge	16	488	1 x 480	1,750
x1e.2xlarge	8	244	1 x 240	1,000
x1e.xlarge	4	122	1 x 120	500

Model	vCPU	Mem (GiB)	Networking Perf.	SSD Storage (GB)
r4.large	2	15.25	Up to 10 Gigabit	EBS-Only
r4.xlarge	4	30.5	Up to 10 Gigabit	EBS-Only
r4.2xlarge	8	61	Up to 10 Gigabit	EBS-Only
r4.4xlarge	16	122	Up to 10 Gigabit	EBS-Only
r4.8xlarge	32	244	10 Gigabit	EBS-Only
r4.16xlarge	64	488	25 Gigabit	EBS-Only

Model	GPUs	vCPU	Mem (GiB)	GPU Memory (GiB)
p2.xlarge	1	4	61	12
p2.8xlarge	8	32	488	96
p2.16xlarge	16	64	732	192

3. Instance Security Groups

In a physical environment, you often install a firewall to filter unwanted traffic. In a similar way Security Group in AWS is linked with EC2 Instance as a virtual firewall, which allows traffic filtration as per configuration. A Security Group is a bunch of firewall rules that decide what to allow or deny. By default for a Security Group all traffic is deny. You would write rules to allow incoming traffic for the instance. Rules in Security Group contain TCP/IP protocols, port numbers, source IPs or IP range/s. One Security Group can be associated with multiple EC2 Instances. And in Security Groups, you can write both Inbound and Outbound rules.

For example, if you want to write a rule to allow remote for a Linux Instance using SSH protocol from any IP address of IPv4 and IPv6, you can see the rule below:

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group: ☒ Create a new security group
☐ Select an existing security group

Security group name:

Description:

Type	Protocol	Port Range	Source	Description
SSH	TCP	22	Anywhere 0.0.0.0/0, ::/0	SSH from anywhere

Warning

Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

(Fig. Security Group)

You find rules written in Security Group are permissive in nature, its mean you cannot create rules that deny access.

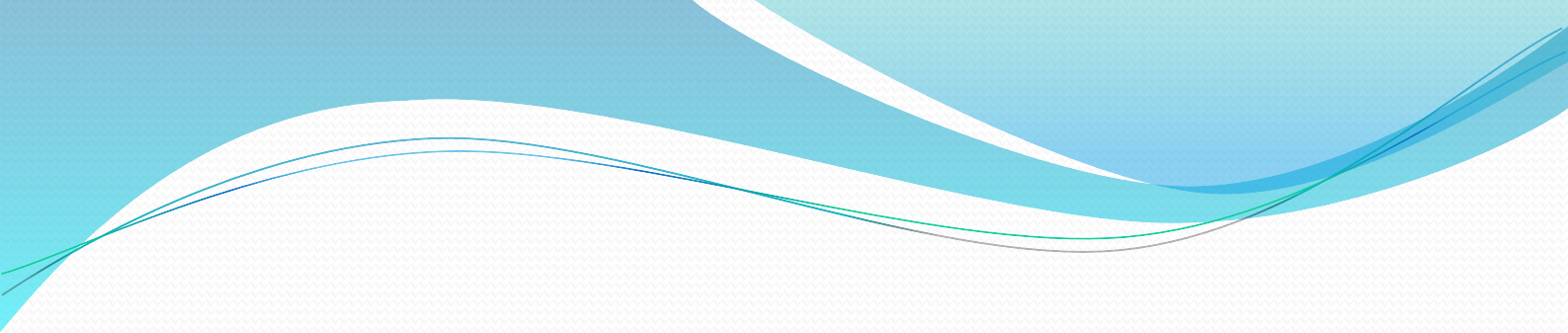
Security group are stateful in nature. In stateful characteristics, when you send a request from your instance, acknowledgement traffic for that request is allowed.

4. EC2 Key Pairs

AWS uses public key cryptography to login into EC2 Instance. Public key cryptography contains two keys, one is public key and another one is private key. Public key is used to encrypt the information and Private Key is used to decrypt the encrypted information. Both keys (Key Pair) are a part of a set.

AWS uses a pair of keys, in which Public key used to encrypt the password information and Private Key is used to decrypt the password.

You create key pair for login into EC2 Instance. One key pair can be used with multiple Instances for login purpose. Or you can create a separate



key pairs for each Instance. If needed, you can also import key pair from your external resources (third party tool).

AWS generated key uses 2048-bit and SSH-2 RSA algorithms.

Your AWS account can have up to 5,000 key pairs per region.

By default, you get a public key in .pem file format.

You can change the key pair that is used to access the default system account of your instance. For example, if a user in your organization requires access to the system user account using a separate key pair, you can add that key pair to your instance. Or, if someone has a copy of the .pem file and you want to prevent them from connecting to your instance (for example, if they've left your organization), you can replace the key pair with a new one.

5. Network Interfaces

On AWS, an EC2 Instance is a virtual machine. And every machine which is running on TCP/IP network requires an IP address to communicate other Instances. For that AWS provides a logical networking interface in a Virtual Private Cloud (VPC). Each EC2 Instance in a VPC has one default network interface **eth0**.

AWS allows you to create network interfaces in your account and attach the new network interface to EC2 instance in your network.

6. Public and Private IP Address

AWS does support both IPv4 and IPv6. By default, your instance will get one public IP and One Private IP address. Default public IP address is dynamic in nature, it is not static. EC2 Instance can have different IP address when you reboot the Instance. Therefore to fix one Public IP address with the Instance, you do use Elastic IP Address.

If you do not want the public IP on your Instance, you can disallow public IP for your Instance from VPC configuration dashboard. AWS support CIDR block in IPv4.

For internal communication amongst instances in the same network, AWS provides internal DNS hostname.

For example, ip-172-230-100-10.ec2.internal.

7. Elastic IP Address (EIP)

If you want to provide a fix public IP address (Static IP) to your EC2 instance or a network interface, Elastic IP address is a best solution. An Elastic IP address is linked with your account. AWS charges for EIPs. You can associate or disassociate EIP with your Instance or network resource any time.

Important about EIP:

- An EIP is a public static IP address, which is reachable from any Instance.
- Static IP address based instance can be reachable from the anywhere on internet.
- In order to provide an entry for reverse DNS, an Elastic IP Address is needed in EC2 Instance.
- Each AWS accounts have limit of 5 Elastic IP addresses per region.
- In case of failure on an instance which is running an application on AWS, a new healthy instance can have same static IP address to provide application available on the Internet.
- If there is a need of more static IP address for your architecture, you can raise your request for additional EIPs to AWS support.

8. Domain Name System (DNS) Name:

Every EC2 Instance on AWS has a public IP and Private IP. And to reach that machine over the Internet AWS provides a public DNS Name which

looks like this **ec2-13-127-228-5.ap-south-1.compute.amazonaws.com**. DNS Name of the instance consists the Public IP address, the region and the service. DNS Name of the Instance can be mapped with public registered DNS to reach the machine. To verify the reachability from the Internet, you can run ping command using public DNS. The Instance will give response only when its ICMP traffic is enabled.

```
C:\Users\SanjayPC>ping ec2-13-127-228-5.ap-south-1.compute.amazonaws.com

Pinging ec2-13-127-228-5.ap-south-1.compute.amazonaws.com [13.127.228.5] with 32 bytes of data:
Reply from 13.127.228.5: bytes=32 time=80ms TTL=241
Reply from 13.127.228.5: bytes=32 time=65ms TTL=241
Reply from 13.127.228.5: bytes=32 time=73ms TTL=241
Reply from 13.127.228.5: bytes=32 time=72ms TTL=241

Ping statistics for 13.127.228.5:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
    Approximate round trip times in milli-seconds:
        Minimum = 65ms, Maximum = 80ms, Average = 72ms
```

In the Description tab of the Instance, you can see all details related to machine including Public DNS Name and IP addresses.

Instance: **i-0f984698e23b7ea08** Public DNS: **ec2-13-127-228-5.ap-south-1.compute.amazonaws.com**

Description

Status Checks

Monitoring

Tags

Instance ID	i-0f984698e23b7ea08	Public DNS (IPv4)	ec2-13-127-228-5.ap-south-1.compute.amazonaws.com
Instance state	running	IPv4 Public IP	13.127.228.5
Instance type	t2.micro	IPv6 IPs	-
Elastic IPs		Private DNS	ip-172-31-5-87.ap-south-1.compute.internal
Availability zone	ap-south-1b	Private IPs	172.31.5.87
Security groups	AdvitiyaLinuxSG01 view inbound rules	Secondary private IPs	
Scheduled events	No scheduled events	VPC ID	vpc-5e4cfa36
AMI ID	amzn-ami-hvm-2017.09.1.20180115-x86_64-gp2 (ami-531a4c3c)	Subnet ID	subnet-d88d6e94
Platform	-	Network interfaces	eth0
IAM role	-	Source/dest. check	True
Key pair name	Linux 16 Dec	T2 Unlimited	Disabled
EBS-optimized	False 🔗	Owner	172850876842
		Launch time	March 7, 2018 at 12:45:17 PM UTC+5:30



For more details on DNS, you will study about AWS service Route53.

9. Regions and Availability Zones (AZ):

AWS Cloud Computing resources available world-wide. They are segregated in various Regions and Regions further are classified in Availability Zones. AWS Regions are separated geographically. By default your resources do not replicate across AWS Regions. If you want to replicate them across resources, you need to specify.

10. Instance Metadata:

It is data about EC2 instance that you can use to manage and configure the running instance.

Instance metadata divided into four categories:

1. The instance ID
2. The Instance type
3. The Instance linked security group
4. Information about AMI used to launch the instance

You can use curl command in Linux Instance to see Instance Metadata.

```
[ec2-user@ip<~>]$ curl http://169.254.169.254/latest/meta-data/
```

```
[ec2-user@ip-172-31-5-87 ~]$ curl http://169.254.169.254/latest/meta-data/  
ami-id  
ami-launch-index  
ami-manifest-path  
block-device-mapping/  
hostname  
instance-action  
instance-id  
instance-type  
local-hostname  
local-ipv4  
mac  
metrics/  
network/  
placement/  
profile  
public-hostname  
public-ipv4  
public-keys/  
reservation-id  
security-groups  
services/[ec2-user@ip-172-31-5-87 ~]$
```

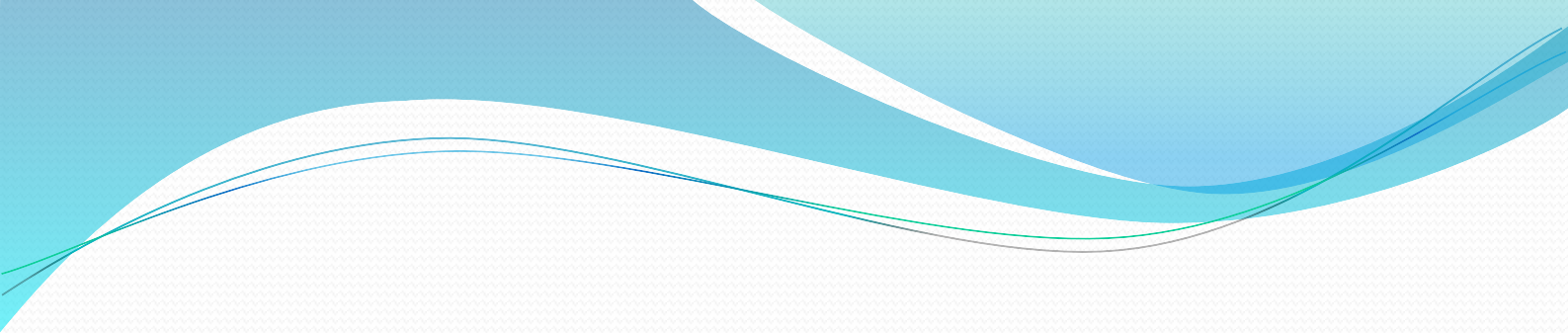
11. VM Import/Export

This feature of AWS is useful to migrate virtual machine from your existing environment to AWS EC2 Instance. In VM Import/Export you import your virtual machine as an AMI. Later on if it is needed, you can export them back to your on-premises environment. You cannot export EC2 Instances created within AWS using AWS provided AMIs.

This technology to import or export machines support with VMWare ESXi, VMWare Workstation, Microsoft Hyper-V and Citrix Xen virtualization formats.

Challenge Note: *Here we leave one excellent opportunity to complete a demanding LAB, in which you can create a virtual machine using VMWare Workstation in your environment and try to import that virtual machine on AWS EC2 Instance.*

12. Bootstrapping:



In some cases, after installation of an operating system in your instance, you want to install, upgrade or configure software. AWS cloud computing provides an amazing benefit to pass a script to manage these tasks automatically for the Instance. You can pass this script or code into the launch wizard as a file or simple text.

It's easy to understand, the process of passing code to be executed on an Instance at launch time is known as bootstrapping.

Challenge Note: *Why don't you write a list of command in a sequence to install and configure an Apache Server in Linux Instance, and pass it on to the instance at the time of launch.*

13. EC2 Instance Storage – Elastic Block Storage (EBS):

EBS volumes in virtual machines (Instances) behave like hard disk drives in physical volumes. EBS is a block level storage. EBS volume to be attached with EC2 instances. And multiple volumes can be attached to an instance. EBS volumes provide high availability and durability.

Types of EBS in AWS:

1. General Purpose SSD
2. Provisioned IOPS SSD
3. Magnetic Volumes

	SSD	SSD	
Purpose	System Boot	Critical Business	Cold Workloads
	Used for small to medium size business needs	Application require better performance and higher IOPS rate	Used where infrequent of data access required
		Raid 0 can be implemented	
IOPS	Performance of 3 IOPS/GiB	High performance maximum up to 20000 IOPS	100 IOPS as an average and burst up to a few hundred
Throughput	160 MB	320MB	40-90 MB
Volume Size	1 GiB to 16 TiB	4 GiB to 16 TiB	1 GiB to 1 TiB

Snapshots:

This is a backup and recovery method of data for your EBS volumes. The snapshots are point-in-time backup of an EBS volume. The EBS Snapshots are incremental and cost effective solution. If multiple backups are taken of a volume, they are incremental. An incremental backup is a type of backup method which copies files that have changed since the last backup. The EBS data restore process creates a volume from snapshots that can be attached with EC2 instance. EBS volumes can be encrypted.