

Lee-23

EC2 = Elastic Compute Cloud

04/08/20

- EC2 is a WebService from AWS, that provides Resizable compute services in the cloud.
- Resizable = Scale UP / Scale Down, Scalability Scale OUT / Scale IN, Elasticity.
- EC2 is Regional.
- Servers = EC2 instances.
- EC2 families / Instance Types
- Instance Type = CPU + Memory
- Scalability can be achieved in AWS by changing the instance type.
- If you start & stop the EC2 instance, Data is not lost because data is stored in EBS volumes.
- You need to stop the EC2 instance to change the instance type (Downtime).

Lee-24

Volumes

↓ (elastic block storage)

EBS volumes.

- Persistent Storage / Permanent
- If you STOP & START the EC2 instance, Data is Not Lost.
- EBS volume max size = 16 TB
- EBS volumes are Billable.

Instance Store Volumes.

- NOT Persistent Storage
- IF you START & STOP EC2 instance, DATA IS LOST
- Instance Store Volumes are free.
- Ephemeris Storage.

* Types of EBS volumes :-

- General Purpose (gp2, gp3) - SSD = GP
- Provisioned IOPS

-EC-25

Snapshot :- Whatever Data in volume taking copy of it in another volume is known as snapshot.

- Backup of the Volume is known as snapshot.
- EBS snapshots are created from EBS volumes.
- You can attach snapshot to the EC2 instance, you have to create a volume & attach to the EC2.
- It is not possible to login or use the snapshot directly.
- Snapshots are stored in S3 (AWS S3)
- * - Snapshot is feature which is provided by some AWS storage services (like EBS, RDS, Amazon Redshift)
 - Snapshot doesn't have any AZ's.
 - It is Regional.
 - It is private (by default), if required, make it public.
- * - You can copy the snapshot from one region to another region in same account.
- * - You can share snapshot from one account to another AWS account using ID.
- EBS volumes can't be moved directly, Create a snapshot & move on.

* IMAGES

- Copy of the OS is called Image.
- Image = AMI = Amazon Machine Image
- Template of OS is called AMI
- AMI contains OS or OS + APPS.
- Copy of the image includes all configurations that we did on original instance.
- Snapshot → Copy of volume
- AMI → Copy of entire EC2 instance including volume

- 1 AMI, can be used multiple times to launch multiple EC2 instances. AMI's are re-usable, AMI's doesn't have any AZ's.
- You can't directly use AMI to login, instead launch EC2 instance from AMI & then login to the instance.
- By Default, AMI's are private, if required we can make it public. AMI's are Regional.
- AMI can be copied from one region to another region, in same account & AMI can be shared one account to another account using AWS ID.
 - x All public images are located at AWS Market Place.
- Image contains OS → Root Volume (EBS) or Root Volume (ISV).
- No need to stop the EC2 instance to create image (But recommended is to STOP).

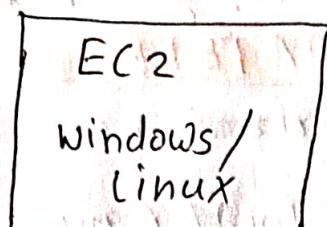
Dec-26]

* Key-Pair

- Key-Pair is used to retire the password of the EC2 instance.
- Key Pair is the combination of Public Key & Private Key.
- The extension of the key-pair is .Pem. Also called as Pem file.
- Whenever we launch the EC2 instance, the console will ask you to create & attach a pem file to the EC2 instance.
- You can create multiple key pairs.
- 1 key pair can be attached to the multiple EC2 instances at the same time.
- EC2 instance can have only 1 key-pair attached at a time.
- Once the key pair attached & launched EC2 instance, we can't change the key pair.
- Once Pem file is lost, it is lost, keep the Pem file very safe & Secure it.

- For every EC2 instance the password is different.
- AWS has public key, customer has private key (Pem file).

| Windows EC2 instance | Linux EC2 instance |
|---|----------------------------|
| Protocol = RDP | Protocol = SSH |
| Port No: - 3389 | Port = 22 |
| Client = Remote desktop connection tool | Client = Putty / MobaXterm |



Windows :-

ip = will be provided by AWS
username = Administrator
password = using key-pair.

Linux :-

ip = will be provided by AWS
username = EC2-user
password = using key-pair.

* Cluster networking instances.

- **Cluster**:- Group of servers / EC2 instances → This group is called placement group.
- When you launch a new EC2 instance, the EC2 service will place the instance such a way that all your instances are spread out across different hardware.
- **Cluster placement group** = Grouping the instances in same / single AZ / same rack, High performance, Risk.
- **Spread placement group** ; EC2 instances are spread across AZ's, High HA, critical application per 1 AZ = 7 Instances.
- **Partial placement group** : Across AZ's. Max partitions = 7.

1 partition can have 100's of EC2 instances.

* IMP

Security Groups

- Security acts as like a firewall to secure the EC2 instance.
 - 1 Security group can be attached to multiple EC2 instances at the same time 1 EC2 can have multiple SG's.
- * If you allow any inbound rule, you must allow an outbound rule also → **STATELESS**.

Firewall - which stops unauthorized access to the network.

firewall - security group Allow/Deny.

Security group - which stops unauthorized access to the EC2 instance, & it is used to secure the EC2 instance.

* Security group has 2 rules:-

Inbound rule - which allows traffic towards EC2 instance.

outbound rule - which sends the traffic outside EC2 instance.

By **DEFAULT**, Inbound rules are deny / outbound rules are allowed.

- In SG, we **ALLOW** protocols **not DENY**.

- AWS EC2 has default security group.

* **NACL** → Network access control list.

NACL - is **STATELESS**.

* If you allow inbound rule, you ^{no} need to allow outbound rule → **STATEFUL**

* Security groups are **STATEFUL**.

- Auto scaling - scale out & scale in EC2 instances based on load
 - Whenever there is a demand on traffic ASG will scale out & scale in EC2 instances automatically.
- * 3 TYPES OF SCALING OPTIONS.
 - + Manual scaling = If you are manually modifying DC value
 - + Scheduled scaling = Based on the time period of a day.
 - + Dynamic scaling = Based on load (Metrics - CPU) matrices (Cloud Watch)
- ELB Does the Health checks for the application, CloudWatch monitors EC2 instances.

* Elastic Load Balancers.

- ELB Distributes the traffic to multiple EC2 instances across AZ's.
- ELB is completely managed by AWS.
- ELB is service not a server.
- ELB can be accessed by DNS Name, you can't login to the ELB.
- ELB has ip address, but there are Dynamic.
- AWS always recommended to use the DNS Name not ip address

Types of Load Balancers.

Application L.B.

ALB

- Latest generation.
- Default choose ALB
- HTTP & HTTPS.
- Best for microservice.
- Routing features.
 - Host based Routing
 - Path - //
 - string parameters routing

Network L.B

NLB

- Latest generation
- TCP, UDP & TLS
- Extreme High Performance.
- Network level
- 1 static ip per AZ

Classic L.B

CLB (HTTP & TCP)

- Previous gen.

Gateway L.B

GSLB

- Latest generation
- 3rd party security appliance

+ TYPES of IP'S

* There are **three types** of IP

- ① Public ip
- ② Private ip
- ③ Elastic IP

① Public IP :-

- which can be accessed from internet
- public ip is optional.
- public ip is dynamic.
- If you start & stop EC2 instance, public ip changes.
- AWS assign public ip to the EC2 instance. By default, public ip is not Enabled.

② Private IP :-

- which can't be accessed from the internet.
- private ip is mandatory.
- private ip's are static.
- Within VPC and through VPN..

③ Elastic IP :-

- same as public IP, but EIP is static.
- If you start & stop the EC2 instance EIP will not change.
- EIP's are FREE
- EIP has to be associated to the EC2 instance.
- If you have not associated EIP to the EC2 instance, AWS will charge for EIP.

* 7 steps for launch EC2 instance

- Select AMI → (linux / window).
- Select instance type → t2.micro.
- Instance configuration. → user data, all configuration.
- Select storage. → EBS volumes
- Select security group. → Default SG.
- Add tags. → (key value pairs)
- Attach PEM file.
(create it or use it)

→ Launch instance.

+ **Global Accelerator :-**

- Leverage the AWS internal network to route to your application 2 Anycast IP's are created for your application.
- The anycast IP send traffic directly to Edge locations
- The Edge locations sends traffic to application,
- GEA is Billable.
- GEA is connected to ELB through EL.

Unicast IP = one server holds one ip address.

Anycast IP = All servers holds the same IP address.

Lec:- 30

Launching template :-

- For edit instah's user data you first need to stop your instance.

Steps:-

- Select created instance ,
- go to Actions
- select image & templates
- Click on Create a template from instance
- Give name to template , & click create template
- You can see created template in launch template section.

* using Launched template Launch instance steps:-

- Select launched template
- choose actions ,
- Click on launch instance from template
- Click on select template.

Create image :- (steps)

- select instance , go to actions , choose image & templates , click create image , give name at 2 times , uncheck reboot
- click on create image .

- (Creating img is taking backup of instance)

- Check that created img in images (AMI) section.

Level 1

Steps :- Attaching Additional Volumes to instance.

- Create instance.
- Connect it to machine.
- Adding storage - go to volumes section
 - Select volume (That should be matched instance volume)
 - Go Action section.
 - Modify Volume
 - Go in size bar & add how you want (GB) & click on modify.

Note:- Volumes can be only increase not decrease.

But the volume will be increase at AWS not in machine, then steps to add volume in machine.

- Steps :-
- ① In machine (virtual machines)
 - ② Right click on windows
 - ③ Choose disk management.
 - ④ Right click on root volume (C-drive)
 - ⑤ Select extend volume
 - ⑥ Click next, next, finish
 - ⑦ Go in file manager & refresh.

Note:- If you want to decrease volume then delete it & recreate it.

* Creating snapshot is backup of volume.

- If we are try to detach volume from instance, we can't first of all we should stop the instance & then we can detach it.

* Creating additional volume (D-drive) :-

- Go in volumes
- Click (Add) volume
- Choose magnetic (standard) (Because it is free)
- Take size in (GB)
- Choose AZ. (Where your instance running choose that)
- Snapshot id (Select fresh) → don't create volume from a snapshot.
- Click on Create Volume.

- Deploying Application (Web) using Load Balancing.

Steps -

- ① Create instance :- Created 2 instances using user data inclusion.
- ② Add protocols - depended upon requirement +
- ③ Create target group
- ④ Create load balance.

Check :- using its IP after TG & using OSS for LB

Now, Steps for How to Create Target group.

- Go to Target group on console. (^{Load balancing} section)
- Click on create target group.
- Do not change anything, scroll down & Add ^{Add group name &} health check path, (ex:- /index.html, /admin)
- Then next.
- Register target :-(which instance want to add in TG)
- Select & click on include as pending below.
- Click on Create target group,

Now, Steps for Create Load Balancing :-

- Go in Load Balancers.
- Click on create Load Balancer
- Choose application Load Balancer.
- Add load balancer name,
- Select Availability zone. (^{Note:-} Select that AZ, where ~~section~~ is your instance)
- In listeners & routing section, choose target group where your instances are included
- And next, click on Create Load Balancer.

* Actions in TARGETED GROUP :-

- Select & Gr., click actions
- ① Delete - we can delete TG.
- ② Register Target : we can add new instance in TG.
- ③ Edit health check : - we can change health check.

Lec - 33]

26/08/2024

* Creating auto-scaling-group.

Steps :-

- Click create ASG.
- Give name to group & select launched temp & Next.
- Select Availability zones (1a, 1b) & Next.
- In load balancing, select Attach to existing load balancer.
- Select Target group from Existing load balancers & Next.
- Turn on ELB health checks from Health Checks & Next.
- Choose desire capacity. Enter no. of instances you launch. (Want to 4)
- Put scaling capacity. (min - 2, max - 10)
- Choose target tracking scaling policy from Automatic scaling.
- Keep Average CPU utilization at Metric type.
- Add target value (30). & next.
- Add notification : click add, create topic & give topic name & email & choose event types (which notification you want update about) & next.
- & Great click create auto-scaling group.

Pre:- create target group.

- Create load balancer.
- Launch template.