**-----------------------------------2.1.17-----------------------------------------**

1. Cloud Developers

2. SysOps - System Operations Administration

3. Cloud DevOps

4. Cloud Architects

**Cloud offers the following services**

1. Infrastructure as a service (Iaas)

2. Platform as a service (Paas)

3. Software as a service (Saas)

**-----------------------------------3.1.17-----------------------------------------**

**Regions**

Represents a geographical area. In one region, there might be two or more data centers

**Edge Location:** Represents a physical Datacenter in specific region

**VPC: Virtual Private Cloud:**

Services --> Networking --> VPC

* VPC represents Virtual private network in cloud.
* It is logically isolated data center in AWS cloud
* There is no charge for VPC
* VPC is specific to the region.

IQ: Can we create a VPC which covers 2 regions?

Ans: No. VPC is region specific.

**Default VPC**

When we create a cloud account in every region there will be one implicity VPC. We call that VPC as default VPC.

**Note:** If we delete default VPC, we cannot recreate them on our own, however we can write an email to that amazon support team for default VPC.

**Subnet**

* It is a kind of smaller VPC or smaller network.
* Services --> VPC--> Subnet

**-----------------------------------5.1.17-----------------------------------------**

**EC2 - Elastic Compute Cloud**

1) VPC

2) Subnet

3) Create EC2 Instance

- Assign public

4) Create Internet Gateway

- Attach to VPC

5) Modify route table add

**Public IP**

- It is used to access to from outside network.

- If we stop and start, EC2 instance public IP will change.

**EIP-Elastic IP**

- It also Public IP but static.

- This a static IP, IP address will not change.

- We can detach Elastic IP from one EC2 instance and we can attach Elastic IP to another EC2 instance.

**Creating Elastic IP & attaching EIP to EC2 instance**

Go to VPC Dashboard-->Elastic IPs--> Allocate New Address--> Associate Address

**Note:** If EIP is not associated with a running EC2 instance then there is a minimal bill charged for that.

**How to delete/ release Elastic IP?**

Select EIP-->Goto Actions--> Release address

**Route Table**

Routing Table is not associated with EC2 & associated with Subnet.

Any EC2 instance in this VPC can talk to this subnet.

When we create a subnet, the subnet is implicitly associated with main route table.

**What is main Route?**

- It is a default routing table created by amazon in our VPC.

**Public subnet**

- If a subnet is exposed to the internet then we call it as a public subnet.

**Private subnet**

- If a subnet is not exposed to the internet then we call it as a private subnet

**\*\*\***Public subnets are used for Web servers. Database servers must be in private subnets.

\*\*\*EC2 instance in a private subnet requires internet access for security updates. How to give internet access to the private subnet?

Ans: We can use either NAT instance or NAT gateway.

**-----------------------------------6.1.17-----------------------------------------**

1) Create NAT instance

2) Create NAT instance in a public subnet

**Steps to create NAT instance**

Lanch instance--> search for NAT under community AMIs

**Modify Routing table of private subnet and add a route to NAT Instance**

**\*\*\*** While using NAT instance, we need to disable source & destination check.

**NAT Gateways**

Using NAT Gateway, to provide internet access to the private subnet.

Step1: Create a NAT Gateway

Step2: Modify Routing Table

**The Drawbacks of NAT Instance**

1. By default NAT Instance does not provide high availability that is we have only one NAT instance if it fails our subnet will not get internet access.

2. Performance of NAT Instance is in our control, if there is any performance issue we need to take action to improve performance.

3. Even maintenance of NAT will be taken by us

**NAT Gateway**

--NAT Gateway is Amazon managed service

1. It is highly Available(HA). Amazon internally maintenance multiple copies of NAT Gateway, if one fails automatically failovers to the second NAT Gateway.

2. NAT Gateway provides high performance

3. NAT gateway is managed service, the maintenance service will be taken by Amazon.

**Security Groups**

* Security Group is a firewall on top of EC2 instance.
* We can add up to five(5) security groups to the EC2 instance. However we can add more security groups if we have multiple ENI(Elastic Network Interface).
* It controls only incoming traffic or inbound traffic.
* If there is any inbound traffic is allowed, the traffic will go out irrespective of outbound routes
* Changes to security group takes effects on fly.
* Security group is state full. When traffic enters into EC2 it remembers the inbound rule even when a traffic goes out.

**Network Access Control List(NACLs)**

- NACL is a firewall at subnet level.

IQ: We are not able to ping EC2 instance, however it has a public IP & Internet gateway configure probably?

Ans: We need to check Security Group & NACL

**-----------------------------------9.1.17-----------------------------------------**

**NACL - Network Access Control List**

**Characteristics of NACLs**

* Using NACL we can allow or deny the traffic.
* NACL is stateless means Incoming traffic is decided by Inbound rule and Outgoing traffic is decided by Inbound rule
* By default implicitly the NACL is associated with a subnet.
* A default NACL by default allows all Inbound & Outbound traffic.
* A NACL is associated with Rule number
* NACL has Rule numbers. Rules are evaluated in ascending order.
* The evaluation of Rule stops when a matching rule is found.
* AS a best practice Amazon says Rule numbers must be multiples of hundreds because in future we can add rules in between.

**VPC Peering Connections**

* We can combine two or more VPCs using VPC Peering.
* By default all VPCs are isolated i.e. EC2 instances in one VPC can't communicate with EC2 instances in another VPC.
* The use case of VPC Peering is let say we have two companies in cloud with their own isolated VPC and let say one company acquiring another company and we want to club all VPCs into one single network then they can use VPC Peering.

**Steps:**

1) We must have two VPCs and its CIDR block range must be different otherwise we can't peer.

2) Create Peering Connection. The other party should accept the peering Connection request.

VPC->Peering Connections->

3) Modify the Routing table association with all subnets in both the VPCs.

4) Note: Transitive peering is not allowed.

**-----------------------------------10.1.17-----------------------------------------**

**VPC Peering**

Step1: Create VPC1, create subnet & create EC2 instance with public IP.

Create Internet gateway

VPC Dashboard-> select "Start VPC Wizard" -> VPC with a Single Public Subnet ->

Step2: Create VPC2, create a subnet & create EC2 instance with public IP.

Create Internet gateway

**EC2 Termination Protection**

Using this option we can enable or disable Termination Protection. For example: If a Termination Protection is enabled it can't terminate EC2 instance.

Select the Instance-> Actions-> Instance settings-> Change termination protection

**Hosting Static website in the cloud by using EC2 instance**

Installing Apache webserver in EC2 instance

]$ sudo yum install httpd

- Verifying the status of httpd server

~]$ sudo service httpd status

- Starting httpd server

~]$ sudo service httpd start

Create welcome.html file in the apache server

**Note:** create this html under /var/www/html/

**-----------------------------------5.1.17-----------------------------------------**

**Accessing VPC**

Option 1: Have a public subnet with a EC2 instance having public or private Elastic IP.

Option 2: We can have VPN between VPC and ours company network.

- VPN uses internet and it is secure. i.e. data passing through VPN is encrypted.

Option 3: Direct connect. This doesn't use internet. This has a dedicated physical connection to the VPC.

**EC2 instance family:**

**General Purpose:**

T Family:

* This provides base line performance. This can perform above base line if it has CPU credits.
* If CPU is Idle then it gains credit points.

Q: What are the use cases for T family?

1. CI Tools / Build Tools: Where the server is Idle very often.

2. Dev test staging Environments

**Note:** We should not use T family for production.

M family:

* It is same as T family, M family provides consistent performance.
* It is consistence

**Uses:** used for CI Tools, used for Development & testing

**Compute Optimized:**

* Compute Optimized provides highest ratio of CPU
* Compute Optimized is recommended for high traffic frontend web servers, on demand batch processing, web servers

**GPU Instances(Graphical Processing Unit)**

**Memory Optimized**

* If our application demands more RAM we should go for Memory Optimized.

**Storage Optimized:**

* It is good for servers.
* It is optimized for applications with specific disk I/O and storage capacity.

**-----------------------------------18.1.17-----------------------------------------**

**On demand EC2 instances**

* This has little high price billing is done every hour

**Spot EC2 instances**

* We purchase these spot instances by bidding.

**Note:** - Cost is very low when compared with other instances.

- The downside of spot instance is if other tenancy out bids. The instance is immediately taken from our account.

**Reserved Instance**

- We can purchase EC2 instance for a 1year or 3years. We can save up to 70% compared to on demand instance pricing.

**EBS Volumes (Elastic Block Storage)**

* Data store on EBS is permanent.
* EBS Volumes are used as a sub resource of EC2 instance.
* We can remove/detach EBS volume from one EC2 instance & it can be attach to another EC2 instance.
* We can also encrypt data store on EBS.
* **\*\*\*** EBS Volume created in one availability zone cannot be used in another availability zone.

Q: How many ways we can create EBS Volumes?

Ans:

1. At the time of launching EC2 instance

2. Under EC2 Dashboard, we can directly created EBS volume

3. We also can create EBS Volume from EBS snapshot

**EBS Snapshots:**

* Snapshot is the backup of EBS volume.

Q: We have a EBS volume in availability zone 1, can we use same for availability zone 2?

Ans: Directly it is not possible. However we can create snapshot. Using this snapshot we can create a new EBS volume in a different availability zone.

Q: We have EBS volume with general purpose and we want to migrate to provision IPOs?

Ans: If we create a volume from a snapshot we can choose different volume type[Provision IOPs].

**Note:** EBS snapshot uses incremental snapshots. i.e. if we take snapshot of EBS which have 5GB data, snapshot size is 5GB. Let us say after 6 months, our data become 8GB, if we take snapshot of EBS it won't be 8GB. It is 3GB.

**EBS Volumes (Elastic Block Storage) Types**

1. General purpose: This provides a base line performance. We should not use this machine critical applications.

2. Provision IOPS: - This provides a very good consistent performance.

- The size of IO is less. may be MBs/KBs

3. Throughput Optimized HDD: This also provides a very good stable IO performance but the size of IO is huge.

4. Cold HDD: This works similar to Throughput Optimized HDD but it is infrequently access the data.

5. Magnetic: Amazon doesn't accept/suggest Magnetic EBS Volumes

**\*\*\***We can create EC2 instances with Instance Storage

**Instance base storage:**

* This is a temporary storage which resides on virtual machine. If we stop virtual machine, data store on Instance storage is lost.
* **Note:** Instance Storage are very cheap.

**RDS(Relational Database Service)**

* RDS is a Platform as a service(Paas). Infrastructure maintenance is done by AWS.

**-----------------------------------5.1.17-----------------------------------------**

**RDS(Relational Database Service)**

RDS is a Platform as a service(Paas). Infrastructure maintenance is done by AWS

->Services->Database->RDS

Get started now

Step1: Select Engine - Select MySQL

Step2:

**Characteristics of RDS**

1. It is Platform as a Service

2. It provides automatic updates

3. It provides automatic backups up to 35days & default 7days

4. It also provides point in time restore of a database.

5. It supports Multi AZ(Availability Zone) [provides HA(High Availability)] Deployments. Multi AZ is not meant for performance.

6. It supports Read Replica (This is meant for improving performance).

**RDS Multi AZ**

* In Multi AZ one server is Active and other one is Passive.
* Any updates/modifications happening on Active those modifications are synchronously updated in Passive also
* If Active server fails, RDS will automatically failover to the Passive server where Passive becomes Active server.

**Note:** Multi AZ is not meant for performance. It is meant for High Availability.

**Read Replica**

Read Replica is mainly used for improving the Read performance.

---------------------------------------------------------------------------------

**DynamoDB**

* Under AWS->Services->Database, it also supports no SQL database server.
* Dynamo DB is no SQL database server

**Elastic Cache:**

* Elastic Cache internally supports Memcached & Redis framework.

**Red shift**

* Red shift is used for Data warehouse Applications

**ELB(Elastic Load Balancer)**

* It is a Service which keeps track of all the servers associated with this Load balancer.
* It also track of the load of every server
* The idea of Load Balancer is, if there is any request from client route that request to least busy server

**-----------------------------------20.1.17-----------------------------------------**

**ELB(Elastic Load Balancer)**

* It is a Service which keeps track of all the servers associated with this Load balancer.
* It also track of the load of every server
* The idea of Load Balancer is, if there is any request from client route that request to least busy server

-----------------------

Step1: Create two EC2 instances in two different subnets,

Step2: Install Apache server on both EC2 instance

Step3: Create ELB and add EC2 instances to the Load Balancer

1. Creating a VPC with Wizard

2. Select VPC with a Single Public Subnet

3. Creating a Load Balancer

Step1: Services->EC2->Load Balancers

Select Load Balancer-> Continue->

-----------------------------------------------

**Characteristics of ELB**

1. It is Platform as a Service

2. ELB is highly available(Managed by cloud)

3. It supports internal & external Load Balancing

- external Load Balancing meaning it can load balance server expose to the internet.

- internal Load Balancing meaning it can load balancethe servers which are not expose to the internet.

4. If any server under ELB is not healthy it takes the server out of rotation.

5. If server becomes healthy again it brings that server back to the rotation.

6. It can do/perform ssl termination.

ssl termination --> i.e., ELB can take encrypted data & decrypt that and pass that data to the EC2 instance.

Q: What is sticky session in ELB?

Ans: When we enable Sticky Session all the requests coming from a particular client is handled by same server

Q: What is Connection Draining in ELB?

Ans: For example, Connection Draining is 300 secs, if ELB finds any EC2 is a unhelathy state, it give 300secs time for that EC2 to process that current request. after 300secs EC2 is taken out of rotation

**-----------------------------------21.1.17-----------------------------------------**

**User Data in EC2**

Using this we can execute scripts at the time if launching EC2 instance.

Services->EC2->Launch Instance->Step 3: Configure Instance Details->Advanced details->User details

------------------------------

#/bin/bash

yum install httpd -y

service httpd start

chkconfig httpd on

----------------------------

**Creating Amazon Machine Images(AMI)**

->Select EC2 instance->Go to Actions->Image->Create Image

* Using our own AMI we can launch new EC2 instance
* By default custom AMI is private. we also can make it public & we can also share a specific AWS account

**Auto Scaling**

Auto scaling is AWS feature using which we can add more servers when we have more load or throughput and we can remove EC2 instance if we have less load.

- When there is high load we want to add a Ec2instance which has web server and our code.

This can be achieve by creating custom AMI

**Configuring Auto scaling Group**

Step1: we need a load balancer for this

EC2->load balancer-> launch load balancer-> classic load balancer

Step2: Configure Launch configuration.

- Using Launch configuration we can tell auto scaling group to use our custom AMI which has Linux and top of it web server and top of it our code.

- Auto scaling group comes under EC2

Step3:

IQ: While configuring Auto scaling we want tp put maximum instance size is 10, Is it possible?

Ans: Yes