**Certified Solutions Architect - Associate**

AWS- stands for Amazon Web services, it’s a cloud computing platform started in 2006 as a startup.

It’s an Infrastructure as a service (IAAS).AWS certification launched in 2013.

History: - Andy Jessy (SVP->AWS)

AWS Exams:-

Professional Tier:-

1. Certified Solutions Architect Professional
2. DevOps Professional

Associate Tier:-

1. Certified Developer Associate
2. Certified Sysops Administrator Associate
3. Sysops Administrator Associate

**Concepts and Components:-**

1. Region: - A Region is a geographical area, each region consist of 2 or more Availability zones.

2. Availability zones: - AZ is simply a Data Centers.

3. Edge Location: - is a CDN (Content Distribution Network) End Point for Cloud front. Edge Location > Regions

**Networking:-**

1. VPC (Virtual private Cloud) it’s a completely separate and isolated set of resources from a logical perspective.

2. Direct Connect is use to connect AWS environment with internet.

3. Route 53 is an amazon's DNS server 53 is a port DNS sites on.

**Compute:-**

1. EC2 (Elastic Compute Cloud) is a virtual server.

2. EC2 Container Service some time known as Amazon ECS, it’s a highly scalable container service. It is use to Run, Scale and terminates Docker. Basically it’s a Docker Container. (Doesn’t come in Exam).

3. Elastic Beanstalk is use to deploy web apps.

4. Lambda is allows you to run code without managing services. You have only need to pay when your code is running.

**Storage:-**

1. S3 is an object based storage (Secure, scalable and easy to use).(Big part of exam)

2. Cloud front is a CDN (Content Delivery web service).

3. Glacier is a secure, durable and low cost storage area.it is for offline storage. Its Block based storage.

4. EFS (Elastic File System) are file storage for EC2 isntances.it can connect to multiple Instances and act as centerline storage in the cloud. It’s a Block level storage. (Doesn’t come in Exam).

5. Snowball is an AWS Import and Export service.

6. Storage gateway is a cloud base storage for on premises organization. It’s a Virtual Machine.

**Database:-**

1. RDS (Relational Database Service) it consist of SQL servers and it’s a relational database.

2. Dynamo dB it’s a NOSQL Server and non-relational database.

3. Elastic ache is way of caching data.

4. Redshift is an Amazon's Data ware Housing service.

5. DMS (Data Migration Service) (Doesn’t come in Exam).

**Analytics:-**

1. EMR (Elastic MapReduce) is used to process data.

2. Data Pipeline is use to pass data from one service to another.

3. Elastic search is use to deploy elastic services, It is an open source service. (Doesn’t come in exam)

4. Kinesis is a platform for streaming data for AWS.

5. Machine learning (Doesn’t come in exam)

6. Quick Sight it’s a very fast Business Intelligent service to perform Ad-hoc analysis. (Doesn’t come in exam)

**Security & Identity:-**

1. IAM (Identity Access Management) is where you can control your Users, groups, Password rotation policies etc.

2. Directory Service

3. Inspector it allows you to store agents on EC2 (Doesn’t come in exam)

4. WAF (Web Application File) (Doesn’t come in exam)

5. Cloud HSM (Hardware Security Module) is a way of securing cloud based infrastructure (Doesn’t come in exam)

6. KMS (Key management service)

**Management Tools:-**

1. Cloud Watch

2. Cloud Formation

3. Cloud trail is a way of providing Audit access. If someone make any changes in any services that is recorded by Cloud Trail.

4. Opsworks is a configuration management service is helping you to configure and operate apps.

5. Config it provides inventory, configuration history (Doesn’t come in exam)

6. Service catalog is allows you to manage catalogs (Doesn’t come in exam)

7. Trusted Advisor is an Automated Service that scans your environment where you can save your money.

**Application Services:-**

1. API Gateway (Doesn’t come in Exams)

2. AppStream is allows you to stream your windows apps to cloud without any kind of code modification. (Doesn’t come in Exams)

3. Cloud Search it manage service that is simple and cost effective to setup it supports 34 languages. (Doesn’t come in Exams)

4. Elastic trance coder is a media transcoder service in cloud it can we use developer to Convert their media file into source code. (It’s a way of trance coding media files).

5. SES (Simple Ebound Service) it allows you to send transactional email and high quality contain.

6. SQS (Simple Query Service) it’s a way of decoupling your infrastructure. It’s a first service ever launched by AWS.

7. SWF (Simple Workflow service) in that developer run and builds background process in parallel and sequential steps.

**Developer Tools:-**

1. Code Commit is a fully managed source control service makes company to host, secure, manage private repository. (Doesn’t come in Exams)

2. Code Deploy is a service that automates any code to deploy on any instance. (Doesn’t come in Exams).

3. Code Pipeline is a continuous delivery service fast and reliable apps update.

**Mobile Services:-**

1. Mobile Hub is a way of monitoring uses of your mobile.

2. Cognito allows you to save your mobile data.

3. Device Farm is use to test the quality of OS to use of web apps

4. Mobile Analytics is a way of monitoring act uses and act returning.

5. SNS (Simple Notification Service) it is use to send notification over the email.

Enterprise Applications:-

1. Workspaces is a virtual desktop in the cloud.

2. WorkDocs is a fully managed secure enterprise storage/Sharing service.

It’s a Dropbox for enterprise. (Doesn’t come in Exams).

3. Work Mail is an Email service (Doesn’t come in exam).

**Internet of Things:-**

1. Internet of Things (Doesn’t come in exam)

**IAM (Identity Access Management):-**

1. it’s about setting your users and granting them permissions.

2. Centralized control of your AWS account.

3. Share Access to your AWS account.

4. Identity Federation (Including Active Directory, Facebook, LinkedIn etc.)

5. Multifactor Authentication

6. Provide temporary access for users/devices and services where necessary.

7. Allows you to set up your own password rotation policy.

8. Integrates with many different AWS services.

9. Support PCI DSS compliance.

**Terms:-**

1. Users-End Users

2. Groups-A collection of users under one set of permissions.

3. Roles-You create roles and can then assign them to AWS resources

4. Policies-A document that defines one/more permissions

**S3 (Simple Storage Service)**

It provides highly secure, scalable, durable object based storage. It is a place to store data objects not OS or Databases.

**S3 is a Key value store.**

* Key is a name of object.
* Value is a data that made up of sequence of bytes.
* Version id is important for versioning. Versioning can't be disable it can only be suspended.(Delete the Bucket)
* Metadata is data about data.
* Sub resources
* Access control list

2. Files can be from 0 byte to 5 TB, it’s an unlimited storage.

3. Files are stored in a Bucket (Directory).

4. S3 is a universal name space that means bucket names are Unique.

5. S3 Read after write consistency for PUTS of new Objects. That means you can read object as you put them in bucket, but as you made some changes in object it will take some time to read.

6. Amazon guarantees 99.99% availability and 99.99999999999% durability.

7. Tiered storage available.

8. Lifecycle management

* Can be used in conjunction of versioning.
* Can be applied to current and previous versioning.
* Following action can be done:-
* Transition to the standard. Infrequent access to storage class (130kb and 30 days after creation date).
* Archive to Glacier storage class after 30 days.
* Permanently Delete.

9. Versioning (Backing up of data)

10. Encryption

11. Secure your data using Access control list and Bucket policy

12. Reduce Redundancy storage means 99.99% availability and 99.99999999999% durability.

13. S3 Charges for Storage, Data transfer pricing, Requests

**Glacier:-**

1. Glacier is a low cost data storage service for data archival.

2. It stores data for low cost $0.01/GB/month.

3. It optimizes data that is infrequently access and for which retrieval time is 3 to 5 hours are suitable. You have to wait 3 to 5 hours to access your data in Glacier.

**Cloud Front:-**

1. A CDN (Content Delivery Network) is a system of distributed delivery servers (Network) that deliver webpages and other web content to user based on geographical location of the user, the origin of the web page and a content delivery server.

**Terminology:-**

* Edge Location- is a location where content will be cached. This is separate to the AWS Region/AZ. Edge location is not just read you can write also. (I.e. put some objects on to them).
* Objects are cached for the life of the TTL (Time to Live) (TTL is always in seconds).you can clear cached objects but you will be charged for that.
* Origin-This is the origin of the entire file that the CDN will distribute. This can be S3 Bucket, EC2 Instance, Route 53, and Elastic Load Balancer.
* Distribution- It is a name given to CDN it is a collection of Edge locations.

There are two type of distribution Web Distribution and RTMP.

1. Cloud Front can be used to deliver your entire website including static, dynamic, streaming and interactive content using a global network of edge location. Requests for your content are automatically routed to the nearest edge location, so the content is delivered with the best possible performance.

2.CloudFront is optimized to work with amazon services like S3,EC2,ELB,Route 53 and amazon also work with non-AWS origin servers which stores the original ,definitive version of files.

**Terminology:-**

* Web Distribution –Typically used for websites.
* RTMP –used for media streaming.

**Elastic Compute Cloud (EC2)**

1. EC2 is a web service that provides resizable compute capacity in the cloud.
2. It Allows you to Scale capacity both up and down.as per the requirements.

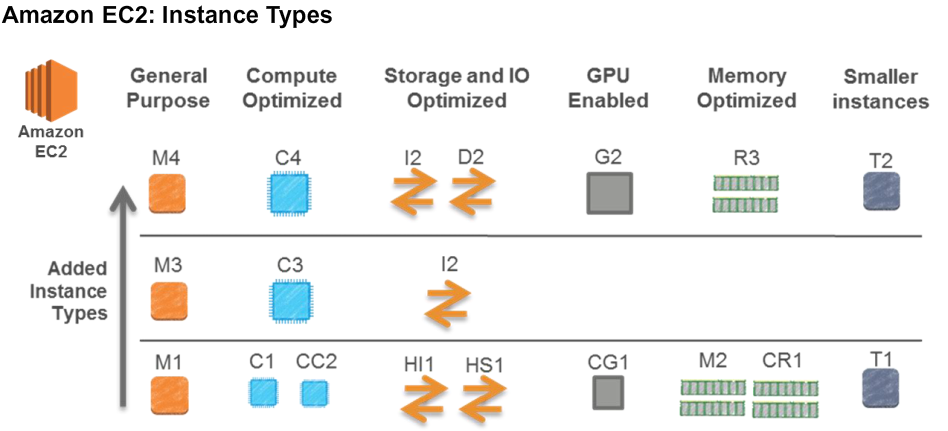
**Revenue Model of EC2:-**

1. **On Demand: -** Allows you to pay fixed rate hourly basis with no commitments.
2. **Reserved: -** Provide you with the capacity reservation, and offers significant discount on hourly charge of instance.1 year or 3 year terms.
3. **Spot: -** Enable you to bid whatever price you want for instance capacity, providing for even greater saving if your application have flexibility start and end times. Once you bid price is less than instance price your instance will be terminated with one hour of notice period. It is just like a stock market.

**Tip: -** If your spot instance is terminated by amazon when bid price increases you will not be charged for partial hour. If you will terminate it you will be charged for an hour.

1. **Scheduled:-**Scheduled Reserved Instances (Scheduled Instances) enable you to purchase capacity reservations that recur on a daily, weekly, or monthly basis, with a specified start time and duration, for a one-year term. You reserve the capacity in advance, so that you know it is available when you need it.
2. **Dedicated Host:-**An Amazon EC2 Dedicated Host is a physical server with EC2 instance capacity fully dedicated to your use. Dedicated Hosts allow you to use your existing per-socket, per-core, or per-VM software licenses, including Windows Server, Microsoft SQL Server, SUSE, Linux Enterprise Server, and so on.

**EC2 Instance Type:-**

****

|  |  |  |
| --- | --- | --- |
| **Family** | **Specialty** | **Use Cases** |
| T2 | Lower Cost, General Purpose | Web Servers/Small DBs |
| M4 | General Purpose | Application Servers |
| M3 | General Purpose | Application Servers |
| C4 | Compute Optimize | CPU Intensive Apps/ DBs |
| C3 | Compute Optimize | CPU Intensive Apps/ DBs |
| R3 | Memory Optimize | Memory Intensive Apps/ DBs |
| G2 | Graphics/General Purpose GPus | Videos Encoding/Machine Learning/3D App Streaming |
| I2 | High Speed Storage | NoSQL DB, Data Warehousing etc. |
| D2 | Dense Storage | File Servers/Data Warehousing/Hadoop |

**Acronym:-** D-Density

I-input/output

R-Ram

T-General Purpose

M-general purpose

C-Compute Optimize

G-Graphics/General Purpose

**Elastic Block Storage (EBS):-**

1. It's disk storage in a cloud.
2. EBS allows you to create storage volume and attach it to your instance.
3. Once you attach volume you can attach file system, run a database or you can use then the way you want.
4. Amazon EBS volume is placed in specific Availability zones.
5. You can’t attach single EBS with multiple ec2.

**EBS Volume Types:-**

1. General Purpose SSD (GP2):- It provides 99.999% availability. Ration of 3 IOPS/GB up to 10,000 IOPS and ability to burst up to 3000 IOPS for short periods for volume under 1 Gib.
2. Provisioned IOPS SSD (IO 1):-Designed for I/O Apps such as large relational or NoSQL Dbs. Use if you need more than 10,000 IOPS.
3. Magnetic (Standard):- Lowest cost for per GBs of all EBS volumes. Magnetic Volumes are idea for workloads where data is accessed infrequently, and apps where the lowest storage cost is important.

**LAMBDA**

AWS Lambda is a compute service that runs your code in response to event and automatically manages the underlying compute resource for you.

It is launched in re-invent of 2014 by AWS.

**What Event can trigger Lambda:-**

1. Modification of objects in Amazon S3 buckets.
2. Messages arriving in kinesis stream
3. Table update in Dynamo DB.
4. AWS API Calls logs created by Cloud Trails.
5. Custom Events from Mobile Apps, Web Apps and other web Services.

**Supported Programming Languages:-**

1. JavaScript
2. Java
3. Python
4. C#

AWS lambda is designed to available 99.99% for service itself and for the functions to operate.

**Pricing:-**

1. Request:-
   1. First 1 Million Requests are free. $0.20 per 1 Million request after that.
2. Duration:-
   1. Duration is calculated from a time your code begins execution until it returns or terminates, rounded up to the nearest 100ms.
   2. The price depends upon the memory you allocated to your function. You are charged $0.00001667 for every GB -Second Used.

**Free Tier:-**

You will get 1 Million free Request per month and 400, 000 GB Computer time per month, the lambda size you choose determines how long they can run in the free tier.

The Lambda free tier does not expire at the end of 12 months at Lambda free tier term, but it is available both existing and new customer indefinitely.

**Security Groups**

You cannot deny any rule in Security groups you can always allow rules in SGs.

1. All inbound traffics are blocked by default.
2. You can assign maximum 5 SG to instance.
3. It applies instance level not Subnet level.
4. All outbound traffics are allowed by default.
5. Changes in SGs can take effect immediately.
6. You can use one SG in multiple instances.
7. SGs are STATEFUl. Network Access control lists, RDS, Dynamo DB and Elastic Caches are STATELESS.
8. Security Groups exists within the vpc, doesn't spam outside of VPC’s.

If you create an inbound rule allowing traffic in, that traffic is automatically allowed back out again.

**Volumes, Snapshots And RAID**

1. Volumes (virtual Hard Disk) exist on EBS.
2. Snapshot exists on S3
3. You can take snapshot of volume and this will store that volume in S3
4. Snapshots are point in time copies of volume.
5. Snapshots are incremental, means that only the blocks that have changed since your last snapshots are moved to S3.
6. If this is your first Snapshot it may take time to create.
7. Snapshot of encrypted volume are automatically encrypted.
8. Volume restored from encrypted Snapshots is encrypted automatically.
9. You can share snapshots but only when they are unencrypted. There snapshots can share with other AWS account if you make them public.

**Root Device Volume:-**

1. You should stop the instance before taking root volume snapshot.

**RAID: -** Redundant Array of Independent Disks it puts whole bunch of disks together and act as one disk.

1. RAID 0:-Striped, No Redundancy, Good Performance, used in Gaming
2. RAID 1:-Mirrored ,Redundancy
3. RAID 5:-Good for Read, Bad for write, AWS does not recommend to put RAID 5 on EBS.
4. RAID 10:- Combination of RAID 1 & 0, Striped, Mirrored, Good redundancy, Good Performance

How to take a Snapshot of RAID Array:-

1. Freeze the file system.
2. Unmount the RAID Array.
3. Shutting down the associated ec2 instance.

Problem with Snapshot:-The snapshot excludes data held in the cache by apps and the OS .this tends not to a matter on a single volume, however using multiple volumes in a RAID array, this can be a problem due to interdependencies of the array.

**Solution:-**we need to take application consistent snapshots.

**EBS vs Instance store:-**

1. Instance store volume is sometime called ephemeral storage.
2. Instance store volume can't be stopped if the underlying host fails, you will lose your data.
3. EBS backed instance can be stop you will not lose your data.
4. You can reboot both and you will not lose your data.
5. By default, both root volume will be deleted on termination, however with EBS volume, you can tell AWS to keep the root volume device.

**AMI (Amazon Machine Image):-**

1. A Template for the root volume for the instance (ex OS,an application server and applications).
2. Launch permission that contains which AWS account can use the AMI to launch the instance.
3. A block device mapping that specifies the volume to attach to the instance when it’s launched.

Tips: - AMI’s are region specific. You can launch AMI in a region in which they are stored. However you can copy AMI’s to other region through console or cli.

Ami stores in S3 like snapshots.

**Elastic Load Balancer:-**

1. In service or out of service
2. Health check
3. ELB have their own DNS name .you never given a Static IP address.

**Cloud Watch:-**

1. Standard Monitoring - 5 Mint
2. Detailed Monitoring - 1 Mint

**Cloud Trail:-**

1. Cloud Trail is used for user activity and API call logs.
2. Cloud Trail is not used for performance monitoring.
3. Used for Auditing.

**IAM Role:-**

1. You can only assign Role to Ec2 while creation any modify and change at any point of time.
2. Roles are universal; you can use them in any of that and launch it into a placement group.

**Instance Meta Data:-**

* <Url:-> <http://169.254.169.254/latest/meta-data/>
* You cannot fetch user data using upper Url.

**Placement Group:-**

Placement groups are logical grouping in instance within a single AZ group. Using placement Group enables apps to participate in a lower latency,10gbps n/w placement Group are recommended for apps that benefits from low n/w latency, high n/w throughput or both.

* Placement groups cannot span in Multiple AZ’s.
* The name you specify for a placement group must be unique within a AWS account.
* Only certain type of instance can be launched in placement group.(Compute Optimize, GPU, Memory Optimize, Storage Optimize)
* AWS recommends homogenous (same size and family) instance with a placement group.
* You cannot merge Placement groups.
* You can launch existing instance into a placement group. You can create an AMI of that and launch it into a placement group.

**Elastic File System:-**EFS is a file storage service for EC2, it provides a simple user interface that allows you to create and configure file system quickly and easily. With EFS, storage capacity is elastic, growing and shrinking automatically as you add and remove files, so your apps have the storage they need, when they need it.

1. EFS are block based storage.
2. Supports Network file system version4 (NFSv4) protocol.
3. You pay only for storage what you use.
4. Can scale up to petabyte.
5. Can support thousands of concurrent NFS Connections.
6. Data is stored across multiple AZ with a same region.
7. Read after write consistency.
8. Data may lose in EFS.

**DNS and Route 53**

ELB Don't have pre-defined IPV4 address, you resolve them using DNS Name. Given the choice; always choose alias record over CNAME

[A Format 192.0.0.0](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#AFormat)

[AAAA Format 2001:0db8:85a3:0:0:8a2e:0370:7334](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#AAAAFormat)

[CNAME Format hostname.example.com](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#CNAMEFormat)

[MX Format 10 mail.example.com](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#MXFormat)

[NAPTR Format](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#NAPTRFormat)

100 50 "u" "E2U+sip""!^(\\+441632960083)$!sip:\\1@example.com!" .

100 51 "u" "E2U+h323" "!^\\+441632960083$!h323:operator@example.com!" .

100 52 "u" "E2U+email:mailto" "!^.\*$!mailto:info@example.com!" .

[NS Format ns-1.example.com](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#NSFormat)

[PTR Format hostname.example.com](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#PTRFormat)

[SOA Format ns-2048.awsdns-64.net hostmaster.awsdns.com 1 1 1 1 60](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#SOAFormat)

[SPF Format "v=spf1 ip4:192.168.0.1/16 -all"](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#SPFFormat)

[SRV Format 10 5 80 hostname.example.com](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#SRVFormat)

[TXT Format "This string includes \"quotation marks\"](http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/ResourceRecordTypes.html#TXTFormat)

**Route 53 Routing Policies:-**

1. Simple

User -----> Route 53 -----> Region

There is no built in intelligence this is a default one.

1. Failover

Active↗ Region 1 (went off) User -----> Route 53

Passive↘ Region 2(Traffic will send here)

Used when you want to use Active/Passive set up.

1. Weighted

10%↗ Region 1 User -----> Route 53

90%↘ Region 2

1. Geolocation

European Customer ↘ ↗ US Customer Route 53

US Customer ↗ ↘ European Customer

Traffic will send according to the geolocation of user

1. Latency

300ms↗ Region 1 User -----> Route 53

54ms↘ Region 2

* Used to provide lowest latency for user and for fastest response.

**Databases**

|  |  |
| --- | --- |
| Relational Database | Non-Relational Database |
| SQL Server | Dynamo DB |
| Oracle | MongoDB |
| MySQL Server | Cache Db |
| PostgreSQL |  |
| Maria DB |  |
| Aurora |  |

**Relational Database service:-**

* **Backup**

1. Automated

* Retention Period 1 to 35 days
* You can recover data at any point of time

1. Snapshots(User Initiated)

Encryption - Can be done Using KMS Key

* **Multi AZ**

It is only for Disaster Recovery ant for Performance

It supports below databases

1. SQL Server
2. Oracle
3. MySQL Server
4. PostgreSQL
5. Maria DB

* **Read Replica**

1. It is a Read only Copy of your database it is for performance.
2. Used for Scaling not for Disaster Recovery
3. Must have automatic backup turned on in order to deploy Read replica.
4. You can have up to 5 Read Replica copies of any DB.
5. You can create Read Replica of Read Replica (but watch out for latency) and it Asynchronous.
6. Each Read Replica will have its own DNS endpoints.
7. You can't have Read Replica that have Multi AZ.
8. You can create Read Replicas of Multi AZ sources databases however.
9. Read Replica can be promoted to be their own databases. This breaks the replication.
10. Read Replica in a second region for MySQL and Maria DB not for PostgreSQL.

|  |
| --- |
| EC2 ↔ RR →RR  ↑  EC2 ↔ DB  ↓  EC2 ↔ RR → RR |

**Dynamo DB Vs RDS**

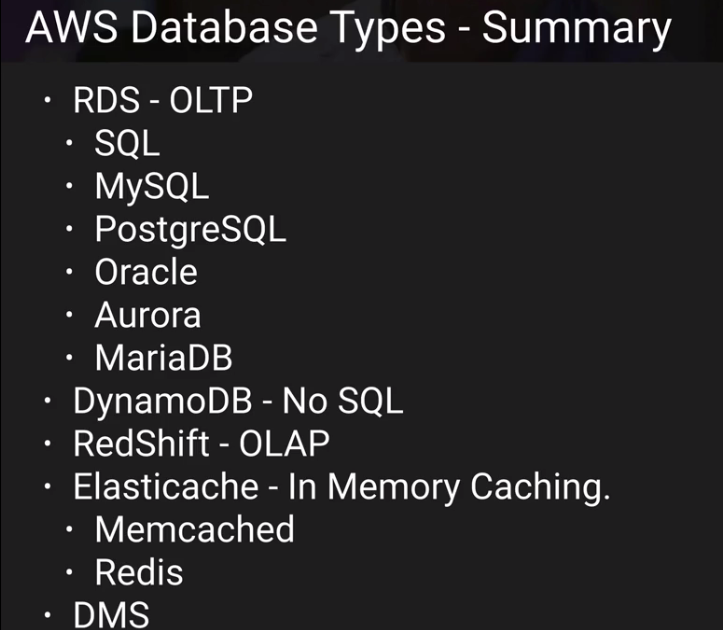
* Dynamo DB offers push button Auto Scaling, means you can scale your database on the fly without any downtime.
* RDS is not so easy you have to use a bigger size instance or to add a Read Replica
* **Database :-**
  + Collection --------------> Table
  + Document --------------> Row
  + Key value pair --------------> Fields

**Data Warehousing:-**

Used for business intelligence, used to pull very large and complex data sets. Usually used by management to do queries on data (such as current performance vs target etc.)

**Elasticache:-**

Elasticache is a web service that makes it easy to deploy, and scale an in-memory cache in cloud, this service improve the performance of web applications by allowing you to retrieve information from fast, managed ,in memory cache ,instead of relying entirely on slower disk-based databases. Elasticache supports two open source in-memory caching engines.

* + Memchached
  + Redis
* **DMS (Data Migration service) :-** You can migrate your on premise database to AWS.

**Dynamo DB:-**

* Is a Fast and flexible NoSQL database service for all apps that need consistent, single digit millisecond latency at any scale.
* It's a fully managed service.
* It supports document and key value data model.
* Stored on SSB storage.
* Spread across 3 geographically distinct data centers.
* Eventual Consistent Reads(Default)
  + Consistency across all copies of data is usually reached within a second. Repeating a read after a short time should return the updated data.(Best read performance)
* Strongly Consistent Reads
  + Strongly consistent read return a result that reflects all writes that received a successful response prior to read.

**Pricing:-**

* Provisioned throughput capacity.
  + Write throughput $0.0065 per hour for every 10 units.
  + Read throughput $0.0065 per hour for every 50 units.
* Storage Cost $0.25GB per month.

**Redshift:-**

RedShift is a fast and powerful, fully managed, petabyte-Scale data warehouse service.

**Configuration:-**

* Single Node(160GB)
* Multi Node
  + Leader Node (Manages client connections and receives queries)
  + Compute Node(Store data and perform queries and computation).Up to 128 compute Nodes
* Redshift store data in the form of columns instead of row
* Columns Compression
* It supports MPP massive parallel processing.(10 time faster)
* It automatically compresses the data you have no need to worry about that.

**Pricing:-**

* You will be charge for compute node only not on leader node(1 unit per node per hour )
* Backup
* Data Transfer(only within VPC, not outside it)

**Security:-**

* Encrypted in transit using SSL
* Encrypted at rest using AES-256 encryption
* By default redshift takes care of Key Management
  + Manage your own key through HSM(hardware security modules)
  + AWS key management service

**Availability:-**

* Currently available only in 1 AZ(No Multi AZ)
* Can restore snapshot to new AZ’s in the event of an outage.

**Elasticache:-**

Elasticache is a web service that makes it easy to deploy, operate and scale in memory cache in cloud. That service improves the performance of web apps by allowing you to retrieve information from fast, managed, in-memory caches, instead of relying entirely on slower disk-based databases.

**Types:-**

* Memcached
  + A widely adopted memory object caching system. Elasticache is a memory compliant with Memcached, So popular tool that you use today with existing Memcached environment will work seamlessly with the service.
* Redis
  + It's a popular open-Source in-memory key value store that supports data structure such a sorted sets and lists. Elasticache supports master/Slave and Multi AZ which can be used to achieve cross AZ redundancy.

**Aurora:-**

Aurora is a MySQL Compatible, relational database engine that combined the speed and availability of high end commercial database with simplicity and cost effective of open source databases.

**Aurora Scaling:-**

* Start with 10Gb,Scales in 10Gb increments to 64Tb(Storage Auto Scale)
* Compute resource can scale up to 32 vCPUs and 244 Gb of Memory.
* It contains two copies of your database in each availability zones, with minimum of 3 availability zones.6 copies of your data.
* Aurora is designed to transparently handle the loss of up to two copies of data without affecting database write availability and up to three copies without affecting read availability.
* Aurora is also self-healing; data blocks and disks are continuously scanned for errors and repaired automatically.

**Aurora Replicas:-**

* 2 Types of replicas are available.
* Aurora replica currently 15.
* MySQL read replicas currently 5

**Virtual Private Cloud(VPC)**

VPC is a virtual data center in the cloud. VPC lets you provision logically isolated section of the AWS cloud where you can launch aws resources in a virtual network that you define. You have complete control over the Virtual networking environment including selection of own IP address range, creation of subnets and configuring route table and network gateway.

**What can you do with VPC?**

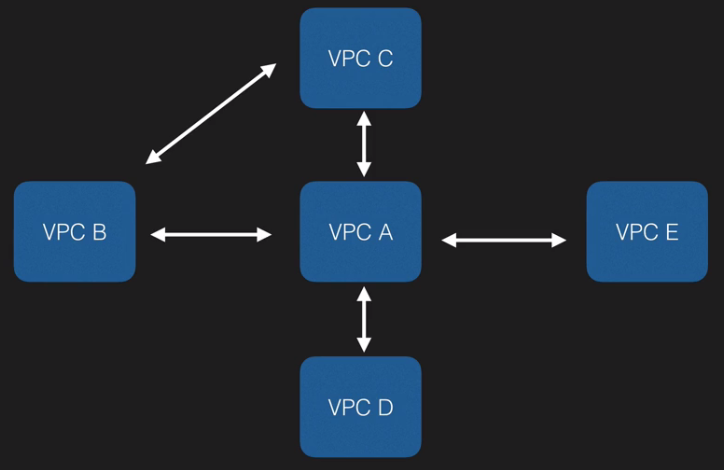
* Launch instances into the subnets
* Assign custom IP address ranges in each subnets
* Configure route tables between subnets
* Create internet gateway and attach them to subnets (or not)
* Much better security control over your AWS resources
* Instance security group
* Subnet network access control lists(ACLS)

**Default VPC and Custom VPC:-**

* Default vpc is a user friendly, allowing you to immediately deploy your instances
* All subnets in default vpc have an internet gateway attached
* Each ec2 have both public and private IP address
* If you delete your default VPC the only way to get it back is contact AWS

**VPC Peering:-**

* Allows you to connect one VPC to another via a direct network route using private IP addresses.
* Instances behave as if they were on the same private network.
* You can peer VPC’s with other AWS accounts as well as with other VPC’s of same account.
* Peering is in star configuration, i.e. 1 central VPC peers with 4 other, No Transitive peering!



**Points to remember:-**

* When you create a VPC by default route table, Network ACL’s and Security Groups used to create with them
* Subnets are always mapped to one AZ. Subnets can’t be across multi AZ’s
* One Subnets equals to one AZ
* When you create internet gateway by default it is detached.
* You can have only one Internet gateway per VPC
* Security groups can stretch across different AZ and different subnets. But exists within VPC
* Security groups are STATEFUL and Network ACL’s are STATELESS
* Subnets can’t stretch across different AZ and different subnets.
* You will always lose 5 IP addresses every time you create subnet 1st four and last one reserved by AWS.

|  |
| --- |
| **VPC** - logical datacenters in AWS  Can span multiple AZ, but can’t span multiple regions, PEER VPC, but no Transitive Peering  Custom VPC has to be /16 can’t go higher than that /8 is not allowed  When you create Custom VPC it creates default security group, default network ACL and default route table., it doesn’t create default Subnet  One Subnet == one AZ, you can have security group spanning multiple AZ, ACL’s span across AZ (assign sg and ACL to two different subnets)  any CIDR block 5 reserved IPs (.0, .1, .2, .3, .255)  so for CIRD block /24: 2^8 - 5 = 256 - 5 = 251 available IP address space  when you create internet gateway, by default its detached, attach it to VPC then, only 1 IGW per VPC  When you create a VPC Default Route table(Main Routable) is created where the default Routes are,  10.0.0.0/16 Local <— all subnets inside VPC will be able to talk to each other  Don’t touch Main route table  Create another route table for route out to internet (0.0.0.0/0 IGW) <— route out to the internet  Last thing you associate this new route table to one of the subnet which will make it public. (you can enable auto assign public IP for the public subnet)  1 subnet can have 1 route table  ICMP is for ping / monitor  NAT instance and NAT gateway  NAT Instance - disable source / destination check., always behind security group, must be in public subnet, must have an EIP, ,must be a route out of the private subnet to NAT  Increase the instance size if bottleneck  Change the main route table - add a route (0.0.0.0/0 NAT Instance target)  NAT Instance is a single point of failover (put it behind a ASG),  NAT gateway - released in 2016 - amazon handled  Amazon maintains it for you, no need to handle yourself. (security patches applied by AWS)  You can just create the gateway and assign EIP (put it in public subnet) (automatically assigned)  Change the main route table - add a route (0.0.0.0/0 NAT gateway target)  No need for disable source/destination check or no need to put it behind a security group - it handles it for you.  Highly available / redundancy no need for ASG.NAT gateways are little bit costly - always use it in production scale automatically up to 10Gbps  **ACL vs SG**  Security groups are state full - any inbound rule , applies to outbound as well (Only Allow rules)  by default all inbound deny, all outbound allow  can span across AZ  ACL are stateless -  For default ACL, all inbound and outbound rules are allowed by default - associated with all subnets in VPC by default  for Custom ACL, all inbound and outbound traffic is denied by default - not associated with any subnet  1 subnet is only associated with ACL. granular rules for ACLs, numbered rules (recommended steps of 100)  Rule no. 99 takes precedence over rule no. 100 (if 99 is blocked and 100 is allowed) 99 will be executed.  Can SPAN across AZ  Ephemeral port - 1024 - 65535 should be allowed to take traffic.  if you want to BLOCK IP address then must use ACL, because security group doesn’t have deny  Bastion - keep it in public subnet to allow SSH / RDP into instances into private subnets (High availability - Bastion in two public subnets and also ASG - Route 53 running Health checks on those Bastion)  VPC Flow logs: to capture all the traffic information into logs - logs everything (create IAM role and create cloud watch log group - and log stream)  VPC Cleanup: can’t delete VPC if you have active running instance or ELB is running |

**Network Address Translation service (NAT):-**

You got a net instance in public subnet and your instances are still not able to communicate then you have to disable Source and destination check from the NAT Instances.

**NAT Instance:-**

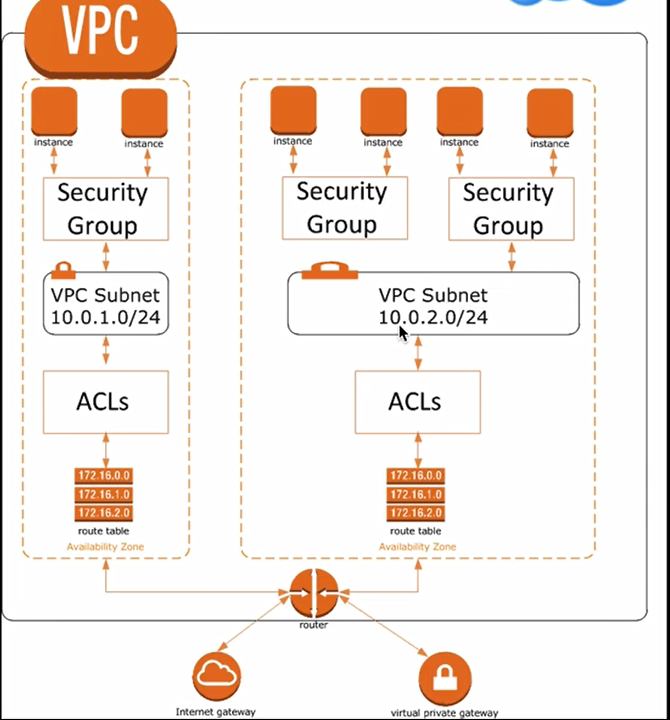
* When creating a NAT instance, disable the source and destination security checks on the instance.
* NAT instance must be in public subnet.
* There must be route out of the private subnet to the NAT instance in order for this to work.
* The amount of traffic that NAT instance can supports that depends on the instance size, if you are bottlenecking increase the instance size.
* You can create high availability using auto scaling groups, multiple subnets in different AZ’s, and script to automate this.

**NAT Gateway:-**

* Preferred by enterprise.
* Scale automatically up to 10GBps
* No need to patch
* Not Associated with security groups
* Automatically assigned public ip address.
* Remember to update your route table
* No need to disable source destination check.

**Network Access Control List (NACL):-**

* VPC comes with default NACL and by default it allows all the inbound and outbound traffic.
* You can create custom NACL and by default it denies all the inbound and outbound traffic until you go and add this in rules.
* In the rule always starts with 100 and we can increment the rule to 100 + something but it will always be evaluated with 100 means 100 will always be greater.
* Every VPC has default NACL attached to all subnets.
* We can’t have multiple NACL for subnets. We can assign 1 subnet to 1 NACL. If you associate subnet with new NACL it will be removed from old one.
* We can use same NACL for multiple subnets. We can assign 1 NACL to multiple Subnets.
* If you disassociate subnet from any NACL by default they will associate back to the default NACL.
* You can associate one subnet to one NACL only.

****

**VPC Flow Log:-**

* You cannot associate Flow logs for VPC that are peered with your VPC unless the peer VPC is in your account.
* You cannot tag a Flow log.
* After you created Flow logs you cannot change its configuration, example you cannot associate different IAM role to flow log.

**Not All IP Address is monitored.**

* Traffic generated by instance when they contact the Amazon DNS Server. If you use your own DNS Server, then all the traffic to the DNS Server is logged.
* Traffic generated by the windows instance for Amazon windows license activation.
* Traffic to end from 169.254.169.254 for instance metadata.
* DHCP Traffic.
* Traffic to reserved IP address for the default VPC Router.

**NAT Vs Bastion:-**

* NAT used to provide internet traffic to EC2 instance in the private Subnet.
* A Bastion used to provide securely Administer EC2 instances (using SSH/RDP) in private subnet.in Australia we call the Jump Boxes.

**Application Services**

**Simple Queue Service (SQS):-**

AWS SQS is a web service that gives you access to a message queue that can be used to store message while waiting for a computer to process them.

* SQS is a temporary repository that is waiting for processing.
* Message can contain 256KB of text in any format.
* SQS ensures delivery on each message at least once, and supports multiple reader and writer interacting with the same queue.
* SQS does not guarantee First in First out delivery on Message.

SQS Example:-

* Asynchronous pulls the task message s from the queue.
* Retrieves the named file
* Processed the conversions
* Writes the image back to S3
* Write the task complete message to another queue
* Deletes the original task message
* Checks the more messages in the worker queue.

Types:-

1. Standard Queue(default)(nearly unlimited transactions per second)
2. FIFO Queue (Limited to 300 transactions per second)

Exam Tips:-

* SQS is a Pull based not Push based
* Maximum Visibility timeout window is 12 hour
* Message will deliver at least once
* 256KB is a max size of message
* Billed at 64KB “Chunks”
* A 256KB message will be 4\*64KB “Chunks”
* First 1 Million Amazon SQS Requests per month are free, its $0.50 per 1 million requests thereafter
* A single request can have from 1 to 10 messages, up to a maximum total payloads of 256KB
* Each 64Kb chunk of payload is billed as 1 request.

**Simple Workflow Service (SWF):-**

SWF is a web service that makes it easy to coordinate work across distributed application components.

**SQS Vs SWF:-**

* SQS has Message retention period 14 days where SWF has up to 1 Year for Workflow execution.
* SWF presents a task-oriented API. Whereas SQS offers a message-oriented API.
* SWF ensures that a task is assigned only once and is never duplicated. With SQS you need to handle duplicate messages and may also need to ensure that a message is proceed only once.
* SWF keeps track of all the tasks and event in an application. With SQS, you need to implement your own application-level tracking, especially if your application uses multiple queues.

**SWF Actors:-**

* **Workflow Starter:-**An application that can initiate (start) a workflow. Could be your ecommerce website when placing an order or a mobile app searching for bus timings.
* **Decider:-**Controls the flow of activity tasks in workflow execution. If something has finished or fails in workflow a decider will decide what to do next.
* **Activity Worker:-**Carry out the activity tasks.

**Simple Notification Service (SNS):-**

SNS is a web service that makes it easy to set up, operate and send notification from the cloud.

Benefits:-

* Instantaneous, push based delivery(No Polling)
* Simple API and easy integration with applications
* Flexible message delivery over multiple transport protocols.
* Pay as you go with no upfront cost.
* Web based AWS management console offers the simplicity of a point and click interface

Protocols:-

* HTTP
* HTTPS
* Email-JSON
* Email
* SQS
* Lambda
* Application

**SNS Vs SQS:-**

* Both are Messaging services
* SNS -PUSH
* SQS-Polls(PULLS)

**Pricing:-**

* $0.50 per million SNS request
* $0.06 per 100,000 over HTTP
* $0.75 per 100 over SMS
* $2.00 over Email

**Elastic Transcoder:-**

* Media Transcoder in the cloud.
* Converts media files from their original source code format to the other.
* Pay based on the minutes that you transcode and the resolution at which you transcode.

**API Gateway:-**

API Gateway is a fully managed service that makes it easy for developers to publish, maintain, monitor, and secure API’s at any scale.

* Low cost and efficient.
* API Gateway has caching capability to increase performance
* Scale effortlessly
* You can throttle request to prevent attacks Connect to cloud watch to log all requests

**Cross Origin resource sharing:-**

CROS is a mechanism that allows restricted resources (e.g. fonts) on a web page to be requested from another domain outside the domain from which the first resource was served.

**Error:-**

Origin policy cannot be read at the remote resource? You need to enable CORS on API gateway.

**Amazon Kinesis:-**

It's a platform to send streaming data. Its makes easy to load and analysis streamed data.

**Kinesis Services:-**

* Kinesis Stream

1. Kinesis Stream consists of Shards
2. 5 transection per second for read and up to maximum total data rate of 2 MB/sec and up to 1000 records per records for writes and up to maximum total data write rate of 1MB/Sec.
3. You can have multiple Shards for Stream.
4. You can retain data by 24 hours to 7 days
5. You can send data from shards to AWS services like S3,DynamoDB,EMR and redshift

**Data processing for Kinesis Stream:-**

Producers (ec2, mobile, data center) →Shards (Kinesis Stream) →EC2 (Consumers) →Aws Service (S3, Dynamo DB, EMR and redshift)

* Kinesis Firehose
  1. There is no data retention concept.

**Data processing for Kinesis Firehose:-**

Producers (ec2, mobile, data center) →Kinesis Firehose (Optional)→S3

* Kinesis Analytics

**Data processing for Kinesis Analytics:-**

Producers (ec2, mobile, data center) →Kinesis Firehose and Kinesis Stream →S3, Redshift and Elastic search Cluster