AWS Exam Points

IAM

- Provides centralized control of you AWS account.
- Shared access to your AWS account.
- Granular permissions.
- Multifactor authentication.
- Allows you to set up your own password rotation policy.
- Integrates with different AWS services.
- Users : End users(people)
- Groups: Group of users under one set of permissions.
- Roles: You create roles and can then assign them to AWS resources.
- Policies: A document that contains one or more permissions. We can assign policy to user, group and role.

S3

- S3 is object based storage.
- The data is spread across multiple devices and facilities.
- Files can be from 0 bytes to 5TB
- Unlimited storage.
- S3 is universal namespace that is bucket names must be unique globally.
- Bucket link: https://s3-ap-south-1.amazonaws.com/vision2020bucket
- Success code for upload : HTTP 200
- Data consistency model: Read after write consistency for PUTS of new objects.
Eventual consistency for overwrite PUTS and DELETES.
- S3 does support bit torrent protocol.
- Availability: 99.99%, Durability: 99.999999999%
- Encryption possible in S3.

S3 storage tiers

- S3 - availability: 99.99%, durability: 99.999999999%
- S3-IA: 99.9%, durability: 99.999999999%
- RRS: availability: 99.99%, durability: 99.99%
- Glacier: Very cheap, it takes 3-5 hours to restore from glacier.

S3 is simple key-value store

Key: name of the object.

Value: Data of the object. Sequence of bytes.

Version ID:

Metadata: Data about data. Ex: created date.

S3 Versioning

- Once versioning enabled we can't remove, we can only be disable.
- Deleting particular version can't be retrieved.
- Deleting entire object can be retrieved.
- We can enable MFA delete capability for objects in S3.

Cross region replication

To enable cross region replication we must enable versioning on source and target buckets.

- Only new objects will be replicated after enabling the replication.
- All versions of the object will be replicated while uploading new version.
- Permissions will be shared replicated from 1 bucket to another.
- Multiple cross region replication will not supported.
- Delete markers will be replicated, but deleting delete markers will not be replicated.
- Deleting specific versions will not be replicated.

S3 life cycle management

- Can be used in conjunction with versioning.
- Can be applied to current and previous versions.
- Can be permanently delete objects.
- Transition to the standard-infrequent Access Storage class objects must be 128KB in size.
- We can load files to S3 much faster by enabling multipart upload.

Glacier

- Glacier is tightly coupled with S3. It is for data archival and extremely low cost.

Cloud Front CDN

- CDN: content delivery network.
- There are 50 Edge location currently.
- Edge location: This is the location where the content will be cached, this is separate to an AWS region/AZ
- Origin: Origin is the source of files that the CDN will distribute. This can be an S3 bucket, EC2 instance, ELB or Route53.
- Distribution: This is the name given the CDN which consists of a collection of edge locations.
- We can have our own origin.
- Edge locations are not just for read only, we can write to them too.
- Objects are cached in edge locations for the life of the TTL (Time to live in seconds).
- We can clear cache objects, but you will be charged.
- We can have multiple origins as part of single distribution.
- Cloud Front supports GET, HEAD, OPTIONS, PUT, DELETE, PATCH and POST http methods.
- Using signed URLs and signed cookies we can restrict the viewers to Cloud Front distribution.

Types of distributions

- Web distribution
- RTMP(User for media streaming)

S3 security and Encryption

- By default all newly created buckets are private.
- We can setup access control using:
  - Bucket policies:
  - Access control lists:
- S3 buckets can be configured to create access logs into another bucket within account or another account.

Encryption

- **In Transit**
  - SSL/TLS
- **At Rest**
  - Server side encryption:
    - S3 manager keys **SSE-S3**
    - AWS key management service, managed keys **SSE-KMS**
    - Customer provided keys **SSE-C**
  - Client Side Encryption:

Storage Gateway

Four types of Storage Gateways

- **File Gateways (NFS)** Flat files stored in s3.
- **Volume Gateways(iSCSI):**
  - **Stored volumes:** Entire dataset stored at site and is asynchronously backed up to S3.
  - **Cached volumes:** most recent dataset will be stored at site and entire dataset will be stored in S3.
- **Tape Gateways (VTL):** used for backup and uses popular backup applications like netback up, Veam etc.

Snowball

- We were moving data to AWS using Import/Export method earlier.
- **Types of snowballs:**
  - **Snowball:** It is a PB scale data transport solution. 80TB max capacity.
  - **Snowball Edge:** On board storage along with compute capabilities. 100TB max capacity. We can run lambda function from snowball Edge.
  - **Snowmobile:** Exa Bytes data transfer. 100PB per snowmobile.
- Snowball can import to S3 and export from S3.

S3 transfer acceleration

- This will utilize the local Edge locations and cache the data for certain time period (TTL).

S3 Static websites

- You can use S3 to host static websites.
- Server less
- Very cheap, scales automatically.
- STATIC only, cannot host dynamic sites.

EC2

- EC2 is a web service that provides resizable compute capacity in the cloud.
- **EC2 Options:**
  - On Demand: Fixed rate per hour with no commitment.
  - Reserved: Provide you with a capacity reservation. Kind of contract for 1 or 3 years.
  - Spot: Enable you to bid whatever price you want.
  - Dedicated Hosts: Physical EC2 server dedicated for your use. Per hour rate.
- **EC2 instance types:**
  - D Dense Storage
  - R Memory
  - M General Purpose
  - C Compute
  - G Graphics
  - I through put optimized
  - F FPGA
  - T Cheap general purpose
  - P Graphics
  - X Extreme Memory
- **EBS volume types:**
  - **General Purpose SSD**:
    - 3 IOPS per GB with up to 10,000 IOPS
  - **Provisioned IOPS SSD**:

        - Designed for I/O intensive applications such as large relational or NoSQL databases.
        - Use if you need more than 10000 IOPS.
        - Can provision up to 20000 IOPS.
    - Throughput optimized HDD(ST1)
        - Big Data
        - Data warehouses.
        - Log processing
        - Can't be a boot volumes.
    - Cold HDD (SC1)
        - Lowest cost storage for infrequently accessed workloads.
        - File server.
        - Cannot be a boot volume.
    - Magnetic (Standard)

      Lowest cost per GB of all EBS volume types that is bootable.

- You can't mount single EBS volume to multiple EC2 instance, instead use EFS.
- Two types of virtualizations:
    - HVM – Hardware Virtual Machine.
    - PV – ParaVirtual.
- One subnet is always equal to one AZ.
- We can't encrypt the root volume of AMI provided by amazon. To encrypt the root volume, first create an EC2 instance and create your own AMI and then launch an EC2 instance from the newly created AMI. We can also use third party tools such as bit locker to encrypt root volume.
- Additional volumes can be encrypted.
- Termination protection will be turned off by default.
- On an EBS-backed instance, the default action is root EBS volume to be deleted on instance termination.
- EBS volumes can be changed on the fly.
- Best practice is to stop EC2 instance always and then change the volume.
- You can change volume type by taking snapshot and then using the snapshot to create new volume.
- You can scale EBS volumes up only.
- Volumes must be in the same AZ as the EC2 instance.

Security Groups

- All inbound traffic is blocked by default.
- All outbound traffic is allowed by default.
- One instance can have multiple security groups.
- We can assign single security group to multiple EC2 instances.
- Security groups are region dependent.
- Security groups are STATEFULL. As soon as inbound rule added, outbound rules are allowed automatically.
- We cannot block specific IP addresses using security groups, instead use ACL.
- We can specify allow rules, but not deny rules.
- Any rule you make to security group will applies immediately.

Encrypting the root volume

- Stop the EC2 instance -> Create snapshot of root volume -> Copy the snapshot with encrypt option checked -> Create AMI from the copied snapshot -> Launch new EC2 instance from custom AMI.
- You can share snapshots, but only if they are unencrypted.
- Volumes restored from encrypted snapshots are encrypted automatically.
- Snapshots of encrypted volumes are encrypted automatically.
-

Instance stores vs EBS

- We can select AMI based on:
  - Region
  - Operating System
  - Architecture(32/64 bit)
  - Launch Permissions
  - Storage for root device
    - Instance store
    - EBS backed volumes
- Instance store instances can’t be stopped whereas instances backed by EBS volumes can be stopped.
- Both instances can be rebooted and no data loss.
- We will loss the data in case instances backed by instance store are failed.
- Instance store volumes will not appear in volumes page.

- By default both root volumes will be deleted on termination, however with EBS volumes, toy can tell AWS to keep the root device volume.

Elastic load balancers

- Two types of load balancers.
    - Application load balancer
        - Layer 7 load balancer.
        - This makes routing decisions to only application layer.
    - Classic load balancer
        - Layer 4 load balancer.
        - This makes routing decisions to transport layer as well as application layer.
- Have their own DNS name, you never given IP address.
- Health checks check the instance health by talking to it via HTTP/S.

CloudWatch

- Basic monitoring monitors every 5 minutes, whereas detailed monitoring monitors every 1 minutes.
- Default metrics:
    - CPU related
    - Disk related
    - Network related
    - Status check
- We can create custom metrics in CloudWatch.
- CloudWatch contains:
    - Dashboard: We can make our own dashboards with different metrics for different EC2 instances, ELB's etc.
    - Alarms: We can create alarms which will trigger an email to subscribed mail id at certain threshold reached of a chosen metric.
    - Events: These helps you to respond to state changes of AWS resources.
    - Logs: Helps you to aggregate, monitor and store logs.
    - Metrics
- CloudWatch is for monitoring.
- CloudTrail is for auditing what AWS users doing inside the account.

AWS CLI

- In order to access AWS from command line we need to first configure the AWS in our EC2 instance using below command.
    - Aws configure -> provide access key id, secret access key, default region name and default output format (empty).
    - After configuration keys will be stored under **~/.aws/credentials** file.

IAM Roles

- Storing access keys in EC2 instances is security risk.
- IAM roles provide access EC2 instances to have access to other AWS services through CLI.
- Role types:
    - AWS service role
    - AWS service-linked role
    - Role for cross-account access
    - Role for identity provider access
- IAM roles are global.
- IAM role can be attached to running instance now.
- It's always better to provide --region command line option in our aws s3 command.

Bash Scripting

```
#!/bin/bash
yum update -y
yum install httpd -y
service httpd start
chkconfig httpd on
cd /var/www/html
echo "<html><h1>Hello Venky</h1></html>" > index.html
```

Key Pairs

- Key pairs are region dependent.
- For login from putty we need to convert .pem file into .ppk file using PutyKeyGen.
- Key pair has public and private keys. Public key is pad lock. Private Key is the key which unlocks the pad lock.

Placement Groups

- Placement group is used to achieve very low network latency and high throughput or both.
- Placement group can’t span to multiple AZ.
- The name of the placement group must be unique with in AWS account.
- Only certain types of instances can be launched in a placement group (Compute optimized, storage optimized, memory optimized, GPU).
- AWS recommend homogeneous instances within placement group
- You can't merge placement group.
- You can’t move existing instance into placement group. Create AMI from existing instance and launch new instance from AMI into placement group.

EFS Elastic File System

- Supports the NFS version 4 protocol.
- You only pay for the storage you use.
- You can scale up to PB.
- Can support thousands of concurrent NFS connections.
- Data is stored across multiple AZ within a region.
- Read after write consistency.
- We no need to have multiple copies of data.

Route53

- SOA Records
- NS Records Name Server records are used by top level domain servers to direct traffic to the content DNS server.
- A Records Address records. User to translate the name of a domain to IP address.
- TTL Records Time to live records.

- **CNAMES Records** Canonical Records. Can be used to resolve one domain name to another.
- **Alias records** are very similar to CNAMES.
- ELB's do not have pre-defined IPv4 addresses. You resolve them to using DNS name.

Databases Section

- RDS does not support increasing storage on a SqlServer DB instance.
- Automatic backups are turned on by default.
- Encryption supported by all RDBMS databases except Aurora.
- We can't encrypt the existing database. In order to encrypt the existing database we need to create the new database with encryption enabled and need to import the data.
- Multi AZ supported by all the databases except Aurora.
- Multi AZ are for DR. Not for performance and scaling.
- Read replicas supported by MySQL Server, PostgreSQL and MariaDB.
- Read replicas are for scaling.
- Must have automatic backups turned on in order to deploy a read replica.
- You can have up to 5 read replica copies of any database.
- You can have read replicas of read replicas.
- Each replica will have its own DNS end point.
- You can't have read replicas that have Multi-AZ.
- Read replicas can be promoted to be their own database. This breaks the replication.
- Read replica is read only. We can't write.
- We can have read replica in second region also but only for MySQL and MariaDB.

DynamoDB

- Always stored on SSD
- Supports OLTP model.
- Spread across 3 geographically distinct data centers.
- Built-In redundancy.
- **Eventual consistency reads default**. Repeating read after a short time should return updated data.

- **Strongly consistency reads** a strongly consistent read returns result that reflects all writes that received a successful response prior to the read.

Redshift

- It is fast and powerful, fully managed, petabyte-scale data warehouse service in cloud.
- Supports OLAP model.
- Single Node 160GB
- Multi node Leader node, Compute node. This can scale up to 128 compute nodes.
- Columnar data storage.
- Support advanced compression. The data can be easily compressed with columnar storage.
- Does not require indexes or materialized views, so uses less space compares to RDBMS.
- **MPP:** Massively Parallel Processing Redshift automatically splits the load across all nodes.
- No charge for leader node. Only charge for compute nodes.
- Charge applies for Backup and data transfer within VPC, not outside it.
- Encrypted in transit using SSL.
- Encrypted at rest using AES-256.
- By default Redshift takes care of key management.
- Redshift is available in single AZ.
- Can restore snapshots to new AZ's in the event of outage.

Elasticache

- Web service that makes it easy to deploy, operate, and scale an in-memory cache in the cloud.
- This improves the performance of application by allowing us to get the data from fast, managed, in-memory caches, instead of relaying entirely on disk.
- Two types engines.
    - **Memcached** Elasticache does not supports Multi-AZ.
    - **Redis** open source in-memory key-value store that supports data structures such as sets and lists. Elasticache support master-slave replication and Multi-AZ which can be used to achieve cross AZ redundancy.

Aurora

- Aurora will run only on AWS infrastructure.
- It's a MySQL compatible, relational database engine that combines the speed and availability.
- We will get all the features of oracle.
- Aurora scaling start with 10GB, scales in 10GB increments to 64TB (Storage auto scaling).
- Compute resources can also scale up to 32vCPUs and 244GB of memory.
- 2 copies of data in each of AZ, with minimum of 3 AZ. 6 copies of our data.
- It's really highly redundant.
- Aurora storage if self-healing.
- Types of replicas
  - **Aurora replicas** up to 15 replicas. Auto failover occur to Aurora replica.
  - **MySQL read replicas** up to 5 replicas. Auto failover not possible.

VPC

- VPC can't span multiple regions but they can span multiple AZ within a region.
- Using subnet we can categorize our resources into public and private.
- We can leverage multiple security using security groups and network access control lists.
- Additionally we can create hardware VPN connection between the corporate datacenter and your VPC.
- Security groups and Network ACL's and route tables can span to multiple subnets and multiple AZ within a region.
- One subnet is equal to 1 AZ.
- **What we can do with a VPC**
  - Launch instances into a subnet of you choosing.
  - Creating subnets.
  - Configure route tables between subnets.
  - Create internet gateway and assign it to VPC. We can have one internet gateway per VPC.
- **Default VPC vs Custom VPC**
  - Default VPC is user friendly, allowing you to immediately deploy instances once created AWS account.
  - All subnets in default VPC have a route out to internet.
  - In default VPC each instance has both public and private IP addresses.
  - If you delete default VPC, the only way to contact Amazon to get back that.

- **VPC peering**
    - We can have multiple VPC's within region.
    - VPC peering allows us to connect one VPC to another using private IP address.
    - Instances behave as if they were in same private n/w.
    - You can peer VPC's with another AWS account as well as with other VPC's in the same account.
    - Peering is in a star configuration. No transitive peering allowed.
- When we create VPC, by default route table, security group and network ACLs will be created.

SQS Simple Queue Service

- SQS is pull based.
- Messages may be processed more than once.
- Visibility timeout the time gap between message pull and timeout. Once timeout reached a message will be visible again in SQS.
- Maximum visibility timeout is **12 hours**.
- Messages contain hold up to **256KB** of text in any format. Any component can access the messages programmatically using SQS API.
- We can implement auto scaling based queue size.
- Two types of queue
    - **Standard queues** default. No order. Unlimited no of transactions per second. Generally tries to deliver messages in the order of FIFO but not guarantees. Occasionally more than 1 copy of message will be delivered.
    - **FIFO queues** Ordered No duplicates. 300 transactions per second.
- Messages can be kept in queue from **1 minute to 14 days**. Default is 4 days.
- Long polling Long polling is the way of pulling messages from queue periodically, while short polling returns immediately and tries to pull messages from SQS even though queue is empty. This will be increase the cost.
- Ec2 instances required to pull messages.

SWF Simple Workflow Service

- SWF can have retention of up to **1 year**, whereas SQS can hold messages up to **14 days**.
- SWF offers task-oriented API, whereas SQS provides message-oriented API.
- Task never duplicated.
- SWF keeps track of all the tasks and events in the application, whereas with SQS, you need to implement your own application level tracking.
- SWF Actors
  - **Workflow starters** is an application that initiate a workflow.
  - **Deciders** Control the flow of activity tasks in a workflow execution.
  - **Activity workers** carry out the activity tasks.
- No need of EC2 instances.

SNS Simple Notification Service

- SNS is push messaging system. No polling.
- SNS can deliver notifications by SMS text to mobiles, to HTTP endpoints, can email, can push message to SQS, and can trigger lambda function.
- To prevent loss of messages, all messages published to SNS are stored redundantly across multiple AZ.
- SNS messages can hold up to **256KB** data.

Elastic Transcoder

- Media/video transcoder in the cloud.

API Gateway

- It is fully managed service that makes it easy for developers to publish, maintain, monitor and secure APIs at any scale.
- Scales automatically.
- API gateway has caching capabilities to increase performance.
- This can throttle to prevent attacks.
- We can log results to cloud watch.

Kinesis

Kinesis Streams

- Stores data with minimum retention of **24 hours** and maximum of **7 days**.
- The data will be stored in shards.
- Consumers will be fetching data from shards.
- Shards gives **5 transactions per second** for reads with **2MB** per second read rate.
- Shards gives **1000 transactions per second for writes** with **1MB** per second rate.

Kinesis Firehose

- We don't have to worry about shards, streams, and manually adding shards.
- This is completely automated when compare with Kinesis Streams.
- No retention.
- We can write data from Firehose to Redshift, ElasticSearch etc.

Kinesis Analytics

- Allows you to run SQL queries of the data which is present either in Streams/Firehose.
- Final data can be stored in Redshift, S3, and ELK.

Scenario Questions

- A particular database is under lot of stress/load. Which service you should use to solve this.
    - **ANS:** We need to choose any of Elasticache, read replicas, Redshift or DynamoDB based on the full scenario given.