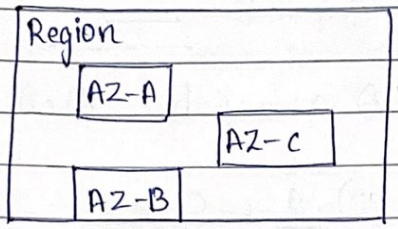
AWS - (SAA-C03)

No. of Regions - 31
No. of Availability Zones - 99 {As of 26 Jan 2023}

Availability Zone → Data Centre → Building filled with servers
Region → two or more Availability Zones



AZ - Availability zone

Edge Location → end points of AWS for caching content
400+ Edge Location (As of 26 Jan 2023)

Key Services to know for the Exam
→ Compute : EC2, Lambda, Elastic Beanstalk
→ Storage : S3, EBS, EFS, FSx, Storage Gateway
→ Databases : RDS, Dynamo DB, Redshift
→ Networking : VPCs, Direct Connect, Route 53, API Gateway, AWS Global Accelerator

Well Architectured Framework - (more info on whitepaper-aws)
Six pillars
- Operational Excellence
- Performance efficiency
- Security
- Cost optimized
- Reliability
- Sustainability

Identity Access Management :

- allows you to manage users and their level of access to the AWS Console.
  - Create users and grant permissions to those users
  - Create groups and roles.
  - Control access to AWS resources.

Root Account:-

The root account is the email address, you used to sign up for AWS. The root account has full administrative access to AWS. It is important to secure this Account.
Use Multi-Factor Authentication to secure your AWS Account.

Four steps to Secure AWS Root Account:-

→ Enable MFA on the root account
→ Create an admin group for your administrators, and assign the appropriate permissions to this group.
→ Create user accounts for your administrators.
→ Add yours users to the admin group.

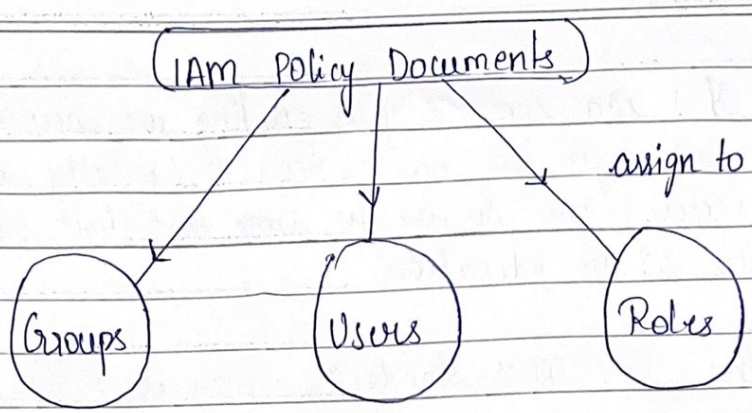
How do we control permissions using IAM?

We assign permissions using policy documents, which are made up of JSON (Javascript Object Notation)

Example:-

```
{
  "version" : "2012-10-17",
  "Statement" : [
      {
        "Effect" : "Allow",
        "Action" : "*",
        "Resource" : "*",
      }
  ]
}
```

This includes key-value pairs. The example given allows all permissions to all the resources i.e., this will give administrator access.



Best Practice :-

You shout not assign a policy document directly to a user. Instead, add user to a group and then assign a policy documents to group.

→ IAM works on global level (IAM is Uni

→ By default, aws account consists of AWS managed policies and job function policies (Database administrator, System administrator, Support User etc)

Building Blocks of IAM

1) Users : A physical person

2) Groups : Functions, such as administrator, developer, etc contains users

3) Roles : Internal usage within AWS (for resources)

best practice :-

1 physical person = 1 user

The principle of Least Privilege

Only assign a user the minimum amount of privileges they need to do their job.

→ when creating a IAM User, Access Key ID and secret access key are create which can be used for programmatic access to AWS console

→ Secret Access key is available only once i.e., during creation of IAM user.

IAM Feduation :- You can combine your existing user account with AWS. For example, when you log on to your PC (usually using Microsoft Active Directory), you can use the same credentials to log in to AWS if you set up feduation.

Identity Feduation :- Uses SAML standard, which is Active Directory.

Inline policy :- Inline policy is a policy that is just assigned to just one user or one group.

Managed policy :- is created by AWS by default.

Test Yourself :-

Q:1 What is the single best thing you can do to secure the root account in AWS?

a) Remove the root account password
b) Delete the root account
c) Enable multi-factor authentication (MFA)
d) You don't have to do anything; AWS will do it for you

Q:2 Which of the following statements describes the principle of least privilege?

a) Only assigning user the minimum amount of permissions that they need do their job
b) Giving a user full permission to most services
c) Standardizing permissions across all users in your company
d) Giving a user too few permissions

Q:3 Wat does the "EAR" in a policy document stand for?

a) Effect, Action, Resource
b) Cooks, Always, Romanticize
c) Every, Action, Reasonable
d) Effects, APIs, Roles

Q:4. True or False? An allow statement in a policy document will override an explicit deny statement.

a) True
b) False

Q:5 Which of the following is NOT part of the IAM Service?

a) Roles
b) Database passwords
c) Users
d) Groups

S3 (Simple storage Service)

→ S3 is one of the oldest services with AWS.

What is S3?

Object storage : S3 provides secure, durable, highly scalable object storage

Scalable : S3 allows you to store and retrieve any amount of data from anywhere on the web at a very low cost

Simple : Amazon S3 is easy to use, with a simple web service interface

→ S3 is object-based storage :- Manages data as objects rather than in file systems or data blocks.

→ you can upload any type of file to S3. for example photos, videos, code, documents and text files etc.

Characteristics :-

1) Unlimited Storage :- The total volume of data and the number of objects you can store is unlimited.

2) Objects upto 5TB in size :- S3 objects can range in size from a minimum of 0 bytes to a maximum of 5 terabytes.

3) S3 Buckets :- Store files in buckets (similar to folders)

Working with S3 Buckets

1) Universal NameSpace :- All AWS accounts share S3 namespace. Each S3 bucket name is globally unique.

2) Example S3 URLs
https://bucket-name.s3.Region.amazonaws.com/key-name
https://learns3.s3.us-east-1.amazonaws.com/Sudagani.jpg

3) Uploading Files
When you upload a file to S3 bucket, you will receive an HTTP 200 code if the upload was successful.

→ S3 basically works off a key-value store
Key : The name of the object (e.g., Sudagani.jpg)
Value : The data itself, which is made up of a sequence of bytes.
Version ID : Important for storing multiple versions of same object
Metadata : Data about the data you are storing
(e.g., content-type, last-modified, etc.)

S3 Availability and Durability
- Built for 99.95% - 99.99% service availability, depending on the S3 tier
- Designed for 99.999999999 (9 decimal places) durability for data stored in S3.

S3 Standard
1) High availability and durability
Data is stored redundantly across multiple devices in multiple facilities (>= 3 AZs)
2) Designed For frequent Access
Perfect for frequent accessed data
3) Suitable for Most Workloads
- The default storage class
- Use cases include websites, content distribution, mobile and gaming applications etc etc

Securing Your Data

1) Server-Side Encryption
You can set default encryption on a bucket to encrypt all new objects when they are stored in the bucket.

2) Access Control Lists (ACLs)
Define which AWS Accounts or groups are granted access and the type of access. Yours can attach S3 ACLs to individual objects within a bucket

3) Bucket Policies
S3 bucket policies specify what actions are allowed or denied.

Strong Read-After-Write Consistency

- After successful write of a new object (PUT) or an overwrite of an existing object, any subsequent read request immediately receives the latest version of the object.
- Strong consistency for list operations, so after a write, you can immediately perform a listing of the objects in a bucket with changes reflected.

Object ACLs Vs Bucket Policy

Object ACLs work on an individual object level **Vs** Bucket policies work on an entire bucket level.

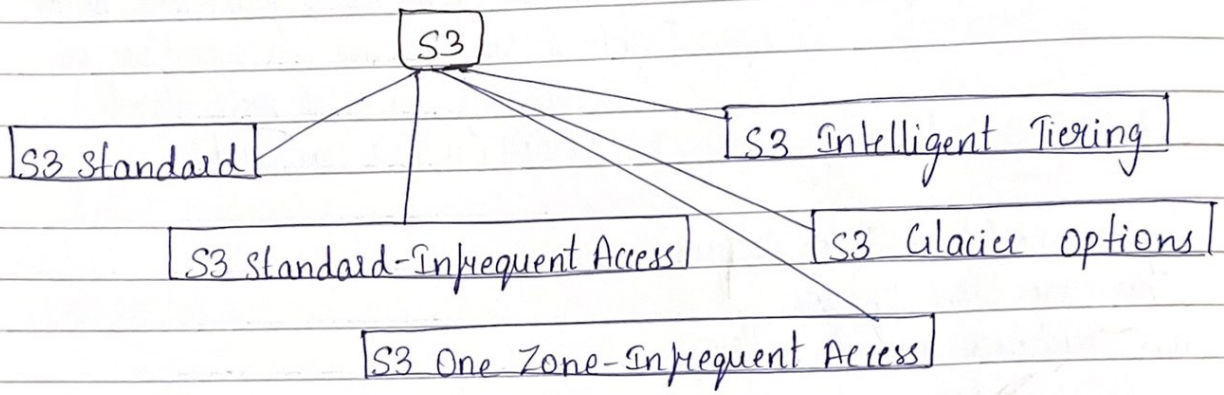
→S3 can be used to ~~store~~ host static content (not dynamic).

Versioning in S3

you can enable versioning in S3 so you can have multiple versions of an object within S3.

Advantages:-

1) All versions of an object are stored in S3. This includes all writes and even if you delete an object.
2) Can be a great backup tool
3) Once enabled, versioning cannot be disabled - only suspended.
4) Can support MFA
5) Can be integrated with lifecycle rules



1) S3 Standard

Already covered in pg: 2

Exam Tip:- When scenario is around static website which doesn't require database connections then think of S3 standard

2) S3 standard-IA

a) Rapid Access: Used for data that is accessed less frequently but requires rapid access when needed
b) You pay to access the data
c) Use Cases: Great for long-term storage, backups, and as a data store for disaster recovery files
d) 99.9% Availability; 99.999999999 (11 9's) Durability

3) S3 One Zone-Infrequent Access

just like S3 standard-IA, but data is stored redundantly within a single AZ.

- Costs 20% less than regular S3 standard-IA
- Great for long-lived, infrequently accessed, non-critical data
- 99.5% Availability, 99.999999999 (11 9's) Durability

4) S3 Intelligent-Tiering - Optimizes Cost

This can be used when you are not sure about how frequently or infrequently you access the data.

2 Tiers → Frequent, Infrequent } Automatically moves your data to the most cost-effective tier based on how frequently you access each object

- 99.99% Availability ; 99.999999999 (11 9's) Durability

5) Glacier and Glacier Deep Archive

There are three options

option a) Glacier Instant Retrieval

provides long-term data archiving with instant retrieval time for your data

option b: Glacier Flexible Retrieval

Ideal storage class for archive data that does not require immediate access but needs the flexibility to retrieve large sets of data at no cost, such as backup or recovery use cases. Can be minutes or up to 12 hours

option c: Glacier Deep Archive

cheapest storage class and designed for customers that retain data sets for 7-10 years or longer to meet customer needs and regulatory compliance requirements. The standard retrieval time is 12 hours, and the bulk retrieval time is 48 hours.

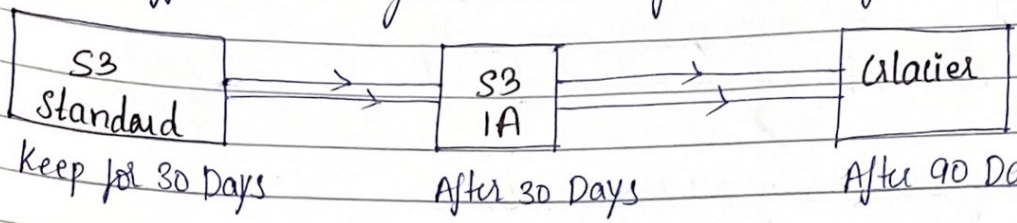
→ you pay each time you access your data
→ Use only for archiving data
→ Glacier is cheap storage
→ Optimized for data that is very infrequently accessed.
→ 99.99% Availability ; 99.999999999% (11 9's) Durability

Comparision of Use Cases of all storage classes

Storage class	AZ(s)	Use Case
S3 Standard	>=3	Suitable for most workloads (e.g., websites, mobile and gaming applications, data analytics)
S3 Standard-IA	>=3	Long-term, infrequently accessed critical data (e.g., backups, data store for disaster recovery)
S3 One Zone-IA	1	Long-term, infrequently accessed, non-critical data
S3 Intelligent-Tiering	>=3	unknown or unpredictable access patterns
S3 Glacier Instant Retrieval	>=3	provides long-term data archiving with instant retrieval time for your data
S3 Glacier Flexible Retrieval	>=3	Ideal storage class for archive data that doesn't require immediate access but should be able to retrieve large sets of data. Backup time can take upto few minutes to 12 hours
S3 Glacier Deep Archive	>=3	cheapest storage class and designed to meet customer needs and regulatory compliance requirements. Standard retrieval time is 12 hours and bulk retrieval time is 48 hours.

Lifecycle Management.

Lifecycle management automates moving your objects between the different storage tiers, thereby maximizing cost effectiveness



→ You can use lifecycle management to move different versions of objects to different storage tiers.
→ This can be applied to current versions and previous versions

S3 Object Lock

You can use S3 object Lock to store objects using write once, read many (WORM) model. It can help prevent objects from being deleted or modified for a fixed amount of time or indefinitely.

You can use S3 Object Lock to meet regulatory requirements that require WORM storage, or add an extra layer of protection.

↳ Governance mode: users can't write or delete an object version or alter its lock permissions unless they have special permissions

↳ Compliance mode: a protected version can't be overwritten or deleted by any user, including the root user in your AWS Account

Retention period:

A retention period protects an object version for a fixed amount of time. After the retention period expires, the object version can be overwritten or deleted unless you also placed a legal hold on the object version.

Legal Hold:

It prevents an object version from being overwritten or deleted. However, a legal hold doesn't have an associated retention period and remains in effect until removed.

Glacier Vault Lock

S3 Glacier Vault Lock allows you to easily deploy and enforce compliance controls for individual S3 Glacier vaults with a vault lock policy.

You can specify controls, such as WORM. Once locked, the policy can no longer be changed.

Types of Encryption in S3

1) Encryption in Transit
   - SSL/TLS
   - HTTPS
2) Encryption at Rest: Server-Side Encryption
   - SSE-S3 : S3-managed keys, using AES 256-bit encryption
   - SSE-KMS : AWS key management service-managed keys
   - SSE-C : Customer-provided keys
3) Encryption at Rest : client-side Encryption
   you encrypt the files yourself before you upload to S3.

→ Enforcing encryption from server side can be done in two Ways
   - Console : select encryption setting on your S3 bucket
   - Bucket policy : enforce using bucket policy.

S3 prefixes :-

Folders inside our buckets are called S3 prefix. prefixes

Eg:- mybucketname/folder1/subfolder1/Sudagani.jpg > /folder1/subfolder1
mybucketname/folder2/subfolder2/Sudagani.jpg > /folder2/subfolder2
mybucketname/folder3/Sudagani.jpg > /folder3
mybucketname/folder4/subfolder4/Sudagani.jpg > /folder4/subfolder4

So, in given examples, /folder1/subfolder1 ; /folder2/subfolder2 ; /folder3 ; /folder4/subfolder4 are prefixes. It doesn't include object name and the file type. It is literally just the folders within the bucket

How to improve S3 performance using S3 prefixes?

→ S3 has extremely low latency. You can achieve high number of requests : 3500 PUT/COPY/POST/DELETE and 5500 GET/HEAD requests per second, per prefix

→ you can get better performance by spreading your reads across different prefixes. For example, if you are using 2 prefixes, you can achieve 11,000 request per second.

4 prefixes → 22,000 requests per second
6 prefixes → 33,000 requests per second

S3 limitations using KMS

If you are using SSE-KMS to encrypt your object in S3

- when you upload a file, you will call GenerateDataKey in the KMS API.
- when you download a file, you will call Decrypt in the KMS API

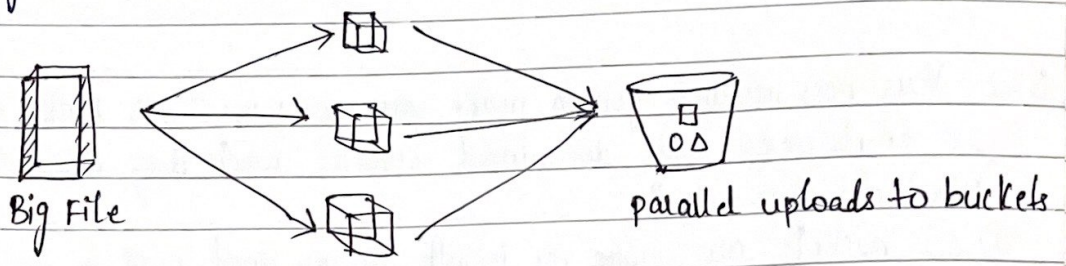
Limitation - KMS Request Rates

- Uploading/downloading will count towards the KMS quota
- Region-specific, however, it's either 5500, 10,000 or 30,000 requests per second
- Currently, you cannot requests a quota increase for KMS.

Multipart uploads :-

→ Use multipart uploads to increase perfomance when uploading files to S3.

→ This basically splits large file into multiple chunks and uploads parallely.



Split into parts

→ Use "S3 byte-range fetches" to increase perfomance when downloding files to S3. This works similar to multipart uploads i.e., spliting into several small chunks for parallel download.

S3 Replication :-

1) you can replicate objects from one bucket to another.
   versioning must be enabled on both the source and distination buckets
2) objects in an existing bucket are not replicated automatically.
   once replication is turned on, all subsequent updated objects will be automatically replicated
3) Delete markers are not replicated by default.

Test your knowledge:-

Q:1 S3 allows to host ____ website directly from the bucket with no other tools needed
a) dynamic
b) static

Q:2 Your boss instructs you to make your company's S3 buckets public so employees can download content when they are at home. What do you do?
a) S3 buckets are public by default, so you don't need to do anything to allow this
b) Tell your boss to bring floppy disk to transport the files to and from the office, instead.
c) Explain your boss that this is terrible idea and opens up huge security risks for your company.
d) Open them up! They said it was ok, right?

Q:3 Since S3 is an object-based storage solution, which type of file should never be stored in it
a) Application logs
b) Cat photos
c) Operating system's boot files
d) HTML documents

Q:4 Why is versioning not enabled by default on new S3 buckets?
a) It slows down the retrieval times in S3
b) Amazon is mean and wants us to accidentally overwrite our data
c) It's security risk, since versioning requires that your content is publicly available.
d) It costs money, as you'll be paying for every additional copy of your objects that you upload.

Q:5 What S3 Glacier storage class standard retrieval time for object is 3-5 hours, with expedited retrievals typically in 1-5 minutes.

a) S3 Intelligent-Tiering
b) S3 Glacier Flexible Retrieval
c) S3 Glacier Instant Retrieval
d) Amazon S3 Glacier Deep Archive