S3 (Simple Storage service)

AWS offers a complete range of cloud services to support both application and archival compliance requirements. Select from the objects, files, and block storage services as well as cloud data migration to start designing the foundation of your cloud IT environments.

❖ Types of storage:

AWS offers 8 types of storage services such as:

1. Simple Storage Service (S3)
2. Elastic File System (EFS)
3. Elastic Block Store (EBS)
4. Glacier
5. Snowball
6. EC2 Instance Storage
7. AWS storage Gateway
8. CloudFront

	Amazon Simple Storage Service (Amazon S3)	A service that provides scalable and highly durable object storage in the cloud.
	Amazon Glacier	A service that provides low-cost highly durable archive storage in the cloud.
	Amazon Elastic File System (Amazon EFS)	A service that provides scalable network file storage for Amazon EC2 instances.
	Amazon Elastic Block Store (Amazon EBS)	A service that provides block storage volumes for Amazon EC2 instances.
	Amazon EC2 Instance Storage	Temporary block storage volumes for Amazon EC2 instances.
	AWS Storage Gateway	An on-premises storage appliance that integrates with cloud storage.
	AWS Snowball	A service that transports large amounts of data to and from the cloud.
	Amazon CloudFront	A service that provides a global content delivery network (CDN).

❖ Difference between Object Storage and Block Storage:

- **Block Storage:**

  - Block storage is suitable for transitional databases, random read/write loads and structured database storage.
  - Block storage divides the data to be stored in evenly sized blocks called data chunks for instance, a file can be split into evenly sized blocks before it is stored.
  - Data blocks stored in block storage would not contain metadata. (Data created, data modified, content type etc.)
  - Block storage only keeps the address (index number) where the data blocks are stored, it does not care what is in that block, just how to retrieve it when required.

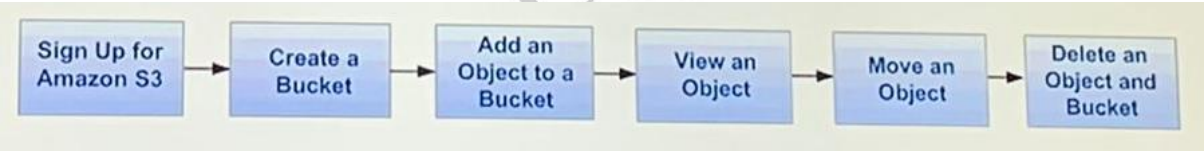
- **Object Storage:**

  - Object storage stores the files as a whole and does not divide them.
  - In object storage an object is: the file/ data itself, its Meta data, object global unique ID.
  - The object global unique ID is a unique identifier for the object (can be the object name itself) and it must be unique such that it can be retrieved disregarding where it's physical storage location is.
  - Object storage cannot be mounted as a drive.
  - Example of object storage solutions are Dropbox, AWS S3, Facebook.

	Performance	Availability and Accessibility	Access Control	Storage and File Size Limits	Cost
Amazon S3	- Supports 3500 PUT / LIST / DELETE requests per second - Scalable to 5500 GET requests per second	- Usually 99.9% available - If lower, returns 10-100% of cost as service credits - Accessible via Internet using APIs	- Access is based on IAM - Uses bucket policies and user policies - Public access via Block Public Access	- No limit on quantity of objects - Individual objects up to 5TB	- **Free tier:** 5GB - **First 50 TB/month:** $0.023 per GB - **Next 450 TB/month:** $0.022 per GB - **Over 500 TB/month:** $0.021 per GB
AWS EBS	- HDD volumes: 250-500 IOPS/volume depending on volume type - SSD volumes: 16-64K IOPS/volume	- 99.99% available - Accessible via single EC2 instance	- Security groups - User-based authentication (IAM)	- Max storage size of 16TB - No file size limit on disk	- **Free tier:** 30GB - **General Purpose:** $0.045 per GB/month - **Provisioned SSD:** $0.125 per GB/month, $0.065 per IOPS/month
AWS EFS	- 3GB/s baseline performance - Up to 10GB/s - Up to 7K IOPS	- No publicly available SLA - Up to 1,000 concurrent EC2 instances - Accessible from any AZ or region	- IAM user-based authentication - Security groups	- 16TB per volume - 52TB maximum for individual files	- **Standard storage:** $0.30-$0.39 per GB-month depending on region - **Infrequent storage:** $0.025-$0.03 per GB-month - **Provisioned throughput:** $6 per MB/s-month

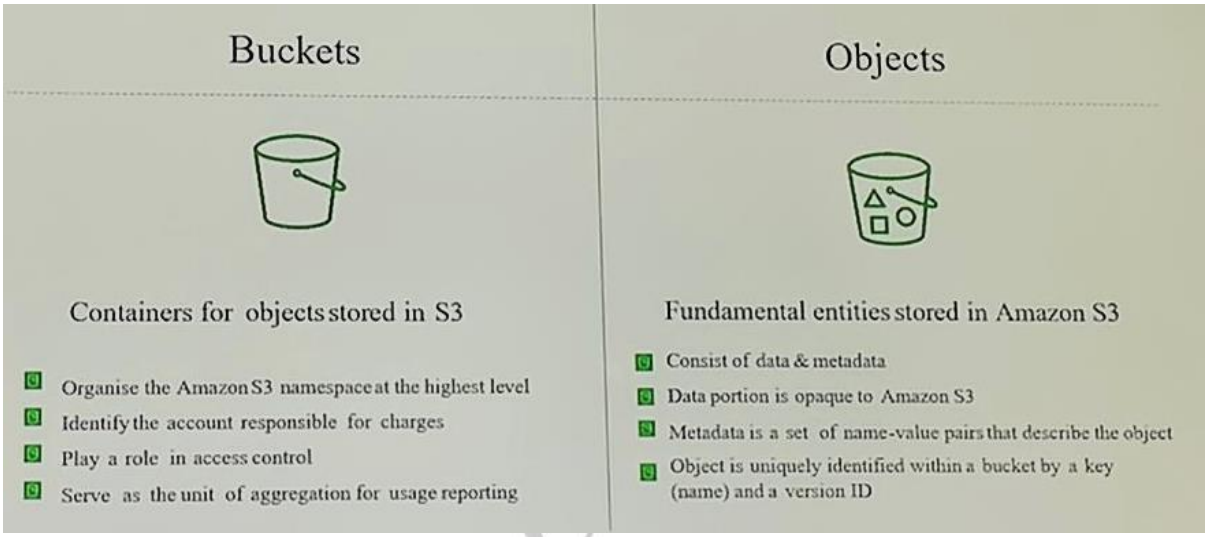
❖ Simple Storage Service (S3):

- S3 is a storage for the internet.
- It has a simple web service interface for simplestoring and retrieving of any amount of data, anytime from anywhere on the internet.
- S3 is object-based storage and NOT a block storage (audio, video, snapshot etc.)
- You cannot install operating system on S3.
- S3 has a distributed data store architecture where objects are redundantly storedin multiple locations. (Minimum 3 locations in same region)
- Data is stored in bucket.
- A bucket is a flat container of objects.
- A bucket can be viewed as a container for objects (default its id private)
- It does not provide any hierarchical of objects (actual folders)
- An Object size Stored Maximum capacity of a bucket is 5TB.
- There is unlimited storage
- You can create folders in your bucket (available through console)
- You can name (object key) for folders in a bucket when using the AWS console.
- You cannot create nested buckets (a bucket inside another).
- Bucket ownership is non transferrable.
- S3 bucket is region specific.
- Cross region replication can be done.
- You can have up to 100 buckets per account. (May expand on request)
- S3 has 99.99% Availability
- Durability 99.999999999% (11 times)



❖ S3 Bucket Naming Rules

- S3 bucket names (keys) are globally unique across all AWS regions.
- Bucket names cannot be change after they are created.
- If bucket is deleted its name become available again to you or other account to use.
- Bucket names must be at least 3 and no more than 63 characters long.
- Bucket names are part of URL used to access a bucket.
- Bucket name must be a series of one or more labels (xyz bucket)
- Bucket names can contain lowercase, numbers and hyphen but cannot use uppercase letters.
- Bucket name should not be an IP address.
- Each label must start and end with a lowercase letter or a number.
- By default, buckets and its objects are private, and by default only owner can access the bucket.

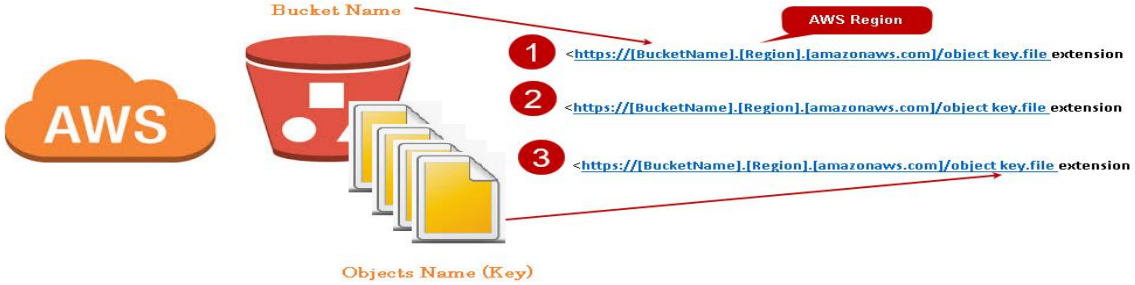


❖ S3 Bucket Sub-Resources:

Sub-resources of S3 bucket includes:

- **Lifecycle:** to decide on objects lifecycle management.
- **Website:** to hold configurations related to static website hosted in S3 buckets.
- **Versioning:** keep objects versions as it changes (set updated)
- **Access Control List**: bucket policies
- **The name is simply two parts**: bucket region's end point / bucket name

Example: for S3 bucket named mybucket in Europe west region is
https://s3-eu-west1.amazonaws.com/mybucket

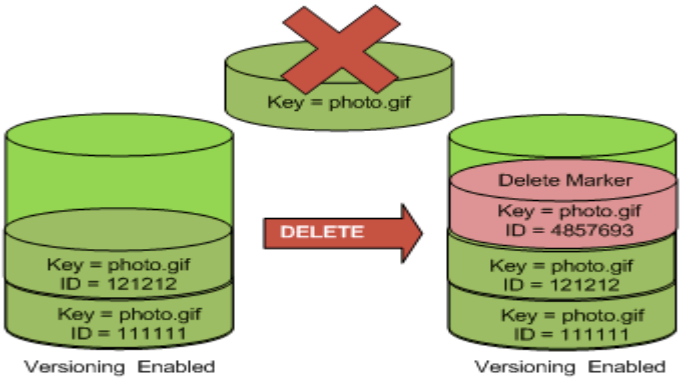


❖ S3 Objects:

- An object size stored in an S3 bucket can be 0 byte to 5TB.
- Each object is stored and retrieve by unique key. (ID or name)
- An object in AWS S3 is uniquely identified and addressed through:
  - service endpoint (Region)
  - bucket name
  - object key (name of object)
  - optionally object version
- Object stored in a S3 bucket in a region will never leave that region unless you specifically move them to another region or CRR.
- A bucket owner can grant cross account permissions to another AWS account (or users in another account) to upload objects.
- You can grant S3 bucket / object permission to:
  - Individual users
  - AWS account
  - Make the resource public
  - To all authenticate user

❖ S3 Bucket Versioning:

- Bucket versioning is a S3 bucket sub resource used to protect against accidental object/data deletion or overwrites.
- Versioning can also be used for data retention and archive.
- Once you enable versioning on a bucket it cannot be disabled however it can be suspended.
- When enable, bucket versioning will protect existing and new objects and maintains their versions as they are updated.
- Updating objects refers to PUT, POST, COPY, DELETE actions on objects.
- When versioning is enabled and you try to delete an object a delete marker is placed on the object.
- You can still view the object and delete the marker.
- If you reconsider deleting the objects you can delete the delete marker and the object will be enabled again.
- You will be charged for all S3 storage cost for all object versions stored.
- You can use versioning with S3 lifecycle policies to delete older version or you can move them to a cheaper S3 storage (Glacier.)



- Bucket version state: -
  - Enabled
  - Suspended
  - Un-versioned
- Versioning applies to all objects in a bucket and not partially applied.
- Object existing before enable versioning will have a version ID or NULL.
- If you have a bucket that is already versioned then you suspended versioning existing objects and their versions remain as it is.
- However, they will not be updated/ version further with future updates while the bucket versioning is suspended.
- New objects (uploaded after suspension) they will have a version ID "null" if the same key (name) is used to stone other objects it will override the existing one.
- An object deletion in a suspended versioning buckets will only delete the objects with ID "null".

❖ S3 -Consistency Level

- **Read After – Write**
  - (Immediate or strong) consistency of PUTs for new object loads to S3
  - A PUT is an HTTP request to store the data
- **Eventually Consistency**
  - It is for overwrite PUTs and DELETE's (for changes/updates to existing objects in S3)

❖ S3 Bucket Versioning-MFA Delete:

- Multifactor authentication delete is a versioning capacity that adds another level of security in case your account is compromised.
- This adds another layer of security for the following:
  - Changing your bucket's versioning state.
  - Permanently deleting on objects version.
- **MFA delete requires:**
  - Your security credentials.
  - The code displayed on an approved physical or s/w-based authentication device.

❖ S3 Multipart Upload:

- It is used to upload an object in parts.
- Parts are uploaded independently and in parallel in any order.
- It is recommended for objects sizes of 100MB or larger.
- You must use it for objects larger than 5GB.
- This is done though S3 multipart upload API.

❖ Copying S3 Objects:

- The copy operation creates a copy of an objects that is already stored in Amazon S3.
- You can create a copy of your object up to 5GB in size a single atomic operation.
- However, to copy an object greater then 5GB you must use the multipart upload API.
- Incur charges if copy to another region.

❖ Use the copy operation to:

- Generate additional copies of the subjects.
- Renaming object (copy to a new name)
- Changing the copy's storage class or encrypt it at rest.
- Move object across AWS location/region.
- Change object metadata.

❖ STORAGE CLASSES OF AMAZON S3:

	S3 Standard	S3 Intelligent-Tiering*	S3 Standard-IA	S3 One Zone-IA†	S3 Glacier	S3 Glacier Deep Archive
Designed for durability	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)
Designed for availability	99.99%	99.9%	99.9%	99.5%	99.99%	99.99%
Availability SLA	99.9%	99%	99%	99%	99.9%	99.9%
Availability Zones	≥3	≥3	≥3	1	≥3	≥3
Minimum capacity charge per object	N/A	N/A	128KB	128KB	40KB	40KB
Minimum storage duration charge	N/A	30 days	30 days	30 days	90 days	180 days
Retrieval fee	N/A	N/A	per GB retrieved	per GB retrieved	per GB retrieved	per GB retrieved
First byte latency	milliseconds	milliseconds	milliseconds	milliseconds	select minutes or hours	select hours
Storage type	Object	Object	Object	Object	Object	Object
Lifecycle transitions	Yes	Yes	Yes	Yes	Yes	Yes

There are 6 types of storage classes of Amazon S3 is available such as:

1. Amazon S3 Standard
2. Amazon S3 Glacier Deep Archive
3. Amazon Glacier
4. Amazon S3 Standard Infrequent Access
5. Amazon S3 one-zone-IA
6. Amazon S3 Intelligent Tiering

1. Amazon S3 Standard:

- S3 standard offers high durability, availability and performance object storagefor frequently accessed data.
- Durability is 99.999999999%.
- Designed for 99.99% availability over a given year.
- Supports SSL for data in transit and encryption of data at rest.
- The storage cost for the object is fairly high but there is very less charge foraccessing the objects.
- Largest object that can be uploaded in a single PUT in 5GB.

2. Amazon S3 IA (standard):

- S3-IA is for data that is accessed less frequently but requires rapid access whenneeded.
- The storage cost is much cheaper than S3-standard almost half the price, butyou are charged more heavily for accessing your objects.
- Durability is 99.999999999%.
- Resilient against events that impact an entire AZ.
- Availability is 99.9% in a year.
- Supports SSL for data in transit and encryption of data at rest.
- Data that is deleted from S3-IA within 30 days will be charged for a full 30days.
- Backed with the Amazon S3 service level agreement for availability.

3. Amazon S3 Intelligent Tiering:

- The S3 intelligent tiering storage class is designed to optimize cost byautomatically moving data to the most cost-effective access tier.
- It works by storing objects in two access tiers.
- If an object in the frequent access tier is accessed it is automatically moved backto the frequent access tier.
- There are no retrieval fees when using the S3 intelligent tiering storage class andno additional tiering fees when objects are moved between access tiers.
- Same low latency and high performance of S3 standard.
- Objects less than 128kb cannot move to IA.
- Durability is99.999999999%.
- Availability is 99.9%.

4. Amazon One-Zone IA

- S3 one zone IA is for data that is accessed less frequently but requires rapidaccess when needed.
- Data store is single AZ.
- Ideal for those who want lower cost option of IA data.
- It is good choice for storing secondary backup copies of on-premise data ofeasily re-creatable data.
- You can use S3 lifecycle policies.
- Durability is 99.999999999%.
- Availability is 99.5%.
- Because S3 one zone IA stores data in a single AZ, data stored in this storageclass will be lost in the event of AZ destruction.

5. Amazon S3 Glacier:

- S3 glacier is a secure, durable, low-cost storage class for data archiving.
- To keep cost low yet suitable for varying needs S3 glacier provides threeretrieval options that ranges from a few minutes to hours.
- You can upload object directly to glacier or use lifecycle policies.
- Durability is 99.999999999%.
- Data is resilient in the event of one entire AZ destruction.
- Supports SSL for data in transit and encryption data at rest.
- You can retrieve 10GB of your amazon S3 glacier data per month for free withfree tier account.

6. Amazon S3 Glacier Deep Archive:

- S3 glacier deep archive is amazon S3 cheapest storage.
- Design to retain data for long period even if for 10 years.
- All objects stored in S3 glacier deep archive are replicated and stored across atleast at three geographically AZ.
- Durability is 99.999999999%.
- Ideal alternative to magnetic tape libraries.
- Retrieval time within 12 hours.
- Storage cost is up to 75% less than for the existing S3 glacier storage class.
- Availability is 99.9%.