

Module-2 : S3 (Simple Storage Service)

Note

*** 5GB is free for free tier account - for all region except AWS GovCloud region.**

***20,000 GET(data retrieval) request and it is charged per GB & 2,000 PUT(changes or modifies) request**

***15 GB of data transfer in & 15 GB of data transfer out /month/year**

What is S3???

- (Simple Storage Service)
- S3 provides developers and IT Teams with Secure, Scalable, Durable object storage.
- Amazon S3 is easy to use with a simple web services interfaces to store and retrieve any amount of data from anywhere on the web.
- S3 is a safe place to store your files.
- It is object based storage.
- The Data is spread across multiple devices and facilities.

Difference between data types ?

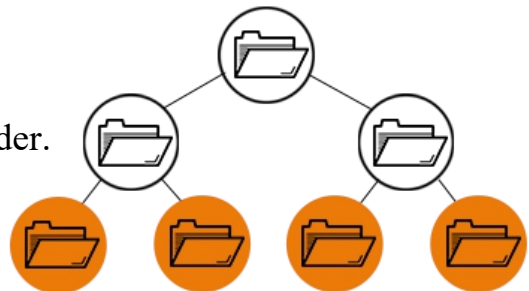
1 - File Storage

2 - Block Storage

3 - Object Oriented Storage

1 - File storage

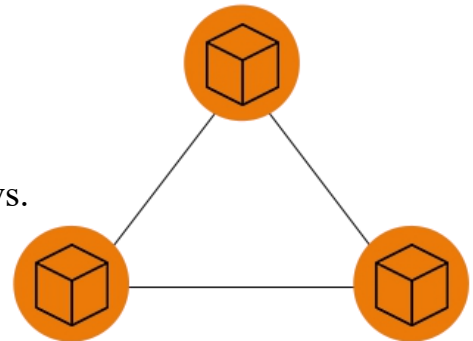
- Data is stored as a single piece of information inside a folder, just like you'd organize pieces of paper inside a regular folder.
- When you need to access that piece of data, your computer needs to know the path to find it.
- It's like a library card catalog for data files.



- It is hierarchy storage form & works as (NAS)network attached storage.

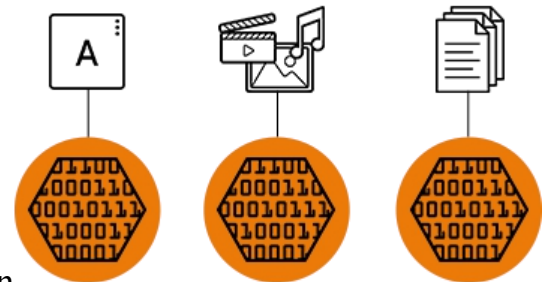
2 - Block Storage

- Block storage cuts data into separate pieces & each block of data is given as unique identifier.
- some blocks can be stored in linux as well as windows.
- Software resembles data whenever require & it is basically deployed in storage-area network(SAN) environment



3 - Object oriented Storage

- Object storage, also known as object-based storage, Is a flat structure in which files are broken into pieces and spread out among hardware.
- In object storage, the data is broken into discrete units called objects and is kept in a single repository, instead of being kept as files in folders or as blocks on servers.
- To retrieve the data, the storage operating system uses the metadata and identifiers, which distributes the load better and lets administrators apply policies that perform more robust searches.
- Object storage requires a simple HTTP application programming interface (API), which is used by most clients in all languages.
- Object storage is cost efficient: you only pay for what you use



Basics of S3 are as below

- SE is Object-Based i.e. allows you to upload files.
- Files can be from 0 bytes to 5TB.

- There is unlimited storage.
- Files are stored in Buckets.
- * What is Bucket - Just like folder , you can upload any file???
- S3 is *universal namespace.*. That means names must be unique globally.
- <https://dineshnetleap.s3.amazonaws.com/>
- <https://dineshbetleap.eu-west-1.amazonaws.com/>
- Bucket name must be unique because it works on dns name
- When you upload a file to SE, you will receive a HTTP 200 code if the upload was successful.

S3 is object based. Think of Objects just as files.

*Objects consist of the following.:

- Key (Thus is simply the name of the object)
- Value (This is simply the data and is made up of a sequence of bytes).
- Version ID (Important for versioning)
- Metadata (Data about data you are storing)

Subresources:

Access Control Lists

Torrent

How does data consistency work for S3 ?

- Read after write consistency for PUTS of new object.
- Eventual consistency for overwrite PUTS and DELETES (can take some time to propagate)

In other words

- if you write a new file and read it immediately afterwards, you will be able to view that data.
- If you update AN EXISTING file or delete a file and read it immediately, you may get the older version, or you may not. Basically changes to objects can take a little bit of time to propagate.

S3 Guarantees

- Built for 99.99% availability for the S3 platform.
- Amazon guarantee 99.9% availability
- Amazon guarantee 99.9999999999% durability for S3 info.

S3 Features

- Tiered Storage Available.
- Lifecycle Management.
- Versioning.
- Encryption.
- MFA Delete.
- Secure your data using Access Control Lists (ACL) permissions.

S3 Storage Classes

1 - S3 Standard

- 99.99% availability
- 99.9999999999% durability,
- stored redundantly across multiple devices in multiple facilities.

2 - S3-IA (Infrequently Accessed)

- For data that is accessed less frequently, but requires rapid access when needed.
- Lower fee than S3, but you are charged a retrieval fee.

3 - S3 One Zone - IA

- For where you want a lower cost option for infrequently accessed data, but do not require the multiple Availability zone data resilience.

4 -S3 - Intelligent Tiering

- Designed to optimized costs by automatically moving data to the most cost- effective access tier, without performance impact or operational overhead.

5 - S3 Glacier

- S3 glacier is secure, durable and low cost storage for data archiving. You can reliably store any amount of data at costs that are competitive with or cheaper than on-premises solutions. Retrieval times configurable from minutes to hours.

6 - S3 Glacier Deep Archival

- S3 glacier deep archival is amazon S3's lower cost storage class where a retrieval time is acceptable.

*****S3 Charges*****

-Storage

-Requests

-Storage Management Pricing

- Data Transfer pricing

-Transfer Acceleration

-Cross Region Replication Pricing.

*****S3 Transfer Acceleration*****

- Amazon S3 transfer Acceleration enables fast, easy and secure transfers of files over long distances between your end users and an S3 Bucket.
- Transfer Acceleration takes advantage of Amazon CloudFront's Globally distributed edge locations. As the data arrives at an edge locations, data is routed to amazon S3 over an optimized network path.

*****Encryption types*****

- 1- Server side encry - AWS S3
- 2- Client side encryp
- 3- transition side encryp - web browser - http:\\
SSL certificate

S3 Versioning

- Stores all versions of an object (including all writes and even if you delete an object)
- Great backup tool.
- Once enabled, Versioning cannot be disabled, only suspended.
- Integrates with Lifecycle rules.
- Versioning's MFA Delete capability, which uses multifactor authentication, can be used to provide an additional layer of security.

Lifecycle Management with S3

- Automates moving your objects between the different storage tiers.
- Can be used in conjunction with versioning,
- Can be applied to current versions and previous versions.

Different ways to share S3 Buckets across accounts

- Using bucket Policies & IAM (applies across the entire bucket).
- Using Bucket ACLs & IAM (Individual objects). Programmatic Access only.
- Cross-account IAM Roles. Programmatic AND Console Access.

Cross Region Replication

- Versioning must be enabled on both the source and destination buckets.
- Files in an existing bucket are not replicated automatically.
- Delete markers are not replicated.
- Deleting individual versions or delete markers will not be replicated.
- Understand what Cross Region Replication is at a high level.

What is CloudFront ?

- CDN - A content delivery Network (CDN) is a system of distributed servers (network) that deliver webpages and other web content to a user based on the geographic locations of the user, origin of the webpage, and a content delivery server.

Key Points to remember of CloudFront

Edge Location

- This is the location where content will be cached. This is separate to an AWS Region / AZ.

Origin

- This is the origin of all the files that the CDN will distribute. This can be either an S3 Bucket, an EC2 Instance, an Elastic Load Balancer, or Route 53.

Distribution

- This is the name given the CDN which consist of a collection of Edge Locations.

Web Distribution

- Typically used for Websites.

RTMP

- Used for Media Streaming.

Edge Locations are not just READ only - you can write to them too. (i.e. put an object on to them).

Objects are cached for the life of the TTL (Time To Live).

You can clear cached objects, but you will be charged.

What is Snowball ?

- Snowball is a petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data into and out of AWS.
- Using Snowball addresses common challenges with large-scale data transfers including high network costs, long transfer times, and security concerns.
- Transferring data with Snowball is simple, fast, secure and can be as little as one-fifth the cost of high speed internet.

Snowball-

- **Snowball comes in either a 50TB or 80TB size**
- **Snowball uses multiple layers of security designed to protect your data including tamper-resistant enclosures, 256 bit encryption and an industry-standard Trusted Platform Module (TPM) designed to ensure both security and full chain-of-custody of your data.**
- **Once the data transfer job has been processed and verified, AWS performs a software erasure of the Snowball appliance.**

What is Snowball Edge ?

- **AWS Snowball Edge is a 100TB data transfer device with on-board storage and compute capabilities.**
- **You can use Snowball Edge to move large amounts of data into out of AWS, as a temporary storage tier for large local datasets, or to support local workloads in remote or offline locations.**
- **Snowball Edge connects to your existing applications and infrastructure using standard storage interfaces, streamlining the data transfer process and minimizing setup and integration.**
- **Snowball Edge can cluster together to form a local storage tier and process your data on-premises, helping ensure your applications continue to run even when they are not able to access the cloud.**

What is Snowmobile ?

- **AWS Snowmobile is an Exabyte-Scale data transfer service used to move extremely large amounts of data to AWS.**
- **You can transfer upto 100PB per Snowmobile, a 45 foot long shipping container, pulled by semi - trailer truck.**
- **Snowmobile makes it easy to move massive volumes of data to the cloud including video libraries, image repositories, or even a**

complete data center migration.

- Transferring data with Snowmobile is secure, fast and cost effective.

Remember to Snowball

-Snowball Can

- Import to S3

- Export from S3

*****S3 Exam Tips*****

- Remember that S3 is Object - Based : I.e. allows you to upload files.
- Files can be from 0 Bytes to 5 TB.
- There is unlimited storage.
- Files are stored in buckets.
- S3 is universal namespace. That is, name must be unique globally.
- <https://netleapdinesh.s3.amazonaws.com/>
- <https://netleapdinesh.eu-west-1.amazonaws.com/>
- Not suitable to install an operating system on.
- Successful uploads will generate a HTTP 200 status code.
- You can turn on MFA delete.