What is Amazon S3?


iCloud
Dropbox
OneDrive
Google Drive

Amazon Simple Storage Service (S3)

Amazon S3 (Simple Storage Service) is a cloud storage service offered by AWS. It is a scalable, durable, high-speed, web-based storage service. As per AWS, S3 is designed for 99.99% of durability.

Many Websites use Amazon S3 as a backbone
Many AWS Services use Amazon S3 as integration as well

What is Cloud Storage?

Before we start talking about Amazon S3, we should first understand what is cloud storage. Cloud storage is a cloud computing model where storage is provided as a service by a cloud provider like AWS. The pricing model used is pay-as-you-go. So, if you need to suddenly store terabytes of data, the option is available by using one of the services like Amazon S3. This removes the need for buying and managing a storage infrastructure on your own and provides easy to go option while you are building your products.

Buckets: An Amazon S3 bucket is a cloud storage resource. It is similar to a folder on your local machine but stored in the cloud.

Objects: An Amazon S3 object is the actual data you are storing in the bucket. The value of the object is the data that you have pushed for long or short-term storage.

Console: Amazon S3 console is the UI provided by AWS to access and manage all the S3 buckets you have created for your data storage purpose.

Amazon S3 – Buckets

Amazon S3 allows people to store objects (files) in “buckets”

Buckets must have a globally unique name

Naming Convention:

1. No upper case
2. No underscore
3. 3-63 characters long
4. Not an IP
5. Must start with a lowercase letter or number

Amazon S3 - objects

Objects (files) have a key

The key is the full path:

S3://my-bucket/my_file.txt

S3://my_bucket/my_folder1/another_folder/my_file.txt

/my_folder1/another_folder – prefix

my_file.txt – object name

1. Max object size is 5TB
2. if uploading more than 5TB, must use “Multi-part upload”

Benefits of Amazon S3

Cost-Effective: S3 offers different storage classes which can be utilized to reduce cost while storing a huge amount of data. The S3 Storage Class Analysis can be used to figure out which data can be moved to a different storage class to save cost.

Durable and Scalable: We have already mentioned that as per AWS S3 is designed for 99.999999999% of durability. To achieve this, AWS stores all the data in S3 in different Availability Zones in a region. Also, S3 is a very scalable service. We don't have to buy anything in advance and can use it if we need a higher amount of storage.

Benefits of Amazon S3

Security, Compliance, and Audit Capabilities: S3 has built-in encryption and access management tools. You can encrypt all the data in S3 using Amazon KMS. There are bucket policies available to block public access to buckets and block any request which is not required.

Easy to Manage: S3 can be easily managed using the console provided by AWS. There are other services like Replication, Batch operations, access points, etc. which help in managing data stored in S3 for various use cases.

Benefits of Amazon S3

Easy to Process:

With so many features packed into one service, S3 became a choice for storing and processing a huge amount of data in the cloud.

Use cases of Amazon S3

Backup and Restore:

S3 is also utilized in storing the backup of databases deployed for the application. Amazon RDS also provides you to create and store a backup directly into an S3 bucket and use them to restore it later directly from S3.

S3 is also used as an archive service where a huge amount of data that is not utilized frequently can be saved on the cloud safely and at a low cost. The S3 Glacier storage class allows you to store archival data at a cheap cost.

Use cases of Amazon S3

Disaster Recovery: With the S3 replication feature, we can store mission-critical data in multiple locations so that we can recover it in case of any disaster

Storage for Applications: Some of the applications require storage options. S3 is an obvious choice for this. With all its features we can create an application that can store user and application data safely.

Use S3 Buckets for Policy To:

1. Grant public access to the bucket
2. Force object to be encrypted at upload

S3 Bucket Policies:

1. JSON based Policies:
1. Resources: buckets and objects
2. Action: set of API to allow or deny
3. Effect: Allow/Deny

S3 Websites:

1. S3 can host static websites and have accessible on the www
2. The website URL will be:
<bucket_name>.s3-websites-<aws-region>.amazonaws.com
3. If you got a 403 (forbidden) error, make sure the bucket policy allows public reads.!!

Amazon S3 – Versioning:

You can version your files in Amazon S3

It is enabled at the bucket level

The same key overwrites will increment the version: 1, 2, 3…

It is best practice to version your buckets:

1. Protect against unintended deletes (ability to restore a version)
2. Easy rollback to provide a version

Notes:

1. Any file that is versioned prior to enabling versioning will have version “null”

S3 Storage Classes

1. Amazon S3 Standard – General Purpose
2. Amazon S3 Standard – Infrequent Access (IA)
3. Amazon S3 One Zone-Infrequent access
4. Amazon S3 Intelligent Tiering
5. Amazon Glacier
6. Amazon Glacier Deep Archive

S3 Storage Classes

1. Amazon S3 Standard – General Purpose

- ✓ More than 3 replication (AZ)
- ✓ Highly durable
- ✓ Highly Available
- ✓ Fast Retrieval
- ✓ Costly

S3 Storage Classes

2. Amazon S3 Standard – Infrequent Access (IA)

- ✓ More than 3 replication (AZ)
- ✓ Highly durable
- ✓ Highly Available
- ✓ Fast Retrieval
- ✓ Less costly compared to standard
- ✓ Minimum size of the file is 128KB
- ✓ Object should be there at least 30days in standard to move to IA

S3 Storage Classes

3. Amazon S3 One Zone-Infrequent access

- ✓ No Replication
- ✓ Less durable
- ✓ Less available
- ✓ Lower cost
- ✓ Less critical data
- ✓ Data should be there for 30 days in S3 standard

S3 Storage Classes

4. Amazon S3 Intelligent Tiering

- ✓ Designed to optimize cost by automatically moving data to the most cost-effective access tier, without performance impact or operational overhead
- ✓ Automatically move the data to the storage classes depending on the usage

S3 Storage Classes

5. Amazon Glacier

- ✓ It is low-cost storage class
- ✓ Data archiving
- ✓ Retrieval times configurable from minutes to hours
- ✓ Infrequently access

S3 Storage Classes

6. Amazon Glacier Deep Archive

- ✓ It is cheapest deep archival storage class
- ✓ Retrieval time of 12 hours is acceptable
- ✓ Infrequently access

S3 Storage Classes

S3 Pricing Model

1 Storage

- ✓ How much data we have stored

2 Requests

- ✓ How many peoples are accessing

3 Data Transfer Pricing

- ✓ How much data we are uploading and downloading

S3 Storage Classes

S3 Encryptions for Objects:

There are 4 methods of encrypting objects in S3

- ✓ 1. SSE-S3 (Server-side encryption): Encrypts S3 objects using key handled & managed by AWS
- ✓ 2. SSE-KMS: Leverage AWS key management service to manage encryption keys
- ✓ 3. SSE-C: When you want to manage your own encryption key
- ✓ 4. Client-Side Encryption: To encrypt your objects before you send them to Amazon S3