AWS S3 (Simple Storage Service)

Amazon Simple Storage Service (Amazon S3) is a widely used cloud storage service provided by Amazon Web Services (AWS). It is designed to store and retrieve any amount of data at anytime from anywhere on the web. S3 is known for its scalability, durability, security, and ease of use. It is commonly used to store a wide variety of data, including documents, images, videos, backups, log files, and more.

**Key features of AWS S3:**

1. **Buckets:** S3 uses the concept of "buckets" to organize and store data. A bucket is similar to a directory or folder and serves as a container for objects (files).
2. **Objects:** Objects are the data items stored in S3 buckets. An object typically consists of the data itself, metadata, and a unique identifier. Each object is accessible via a unique URL, making it easy to access and share data over the internet.
3. **Scalability:** S3 can scale to accommodate any amount of data. It automatically scales its infrastructure to handle growing amounts of data without any upfront provisioning.
4. **Durability:** S3 is designed to provide high durability for data. It stores multiple copies of data across different facilities within a region to ensure data resilience.
5. **Availability:** S3 offers high availability, ensuring that data is accessible and retrievable at any time with low latency.
6. **Security:** S3 provides several security mechanisms to protect data, including server-side encryption, client-side encryption, access control lists (ACLs), and bucket policies.
7. **Lifecycle Policies:** You can define lifecycle policies for objects in S3 to automatically transition them between storage classes or delete them after a certain period. This helps in optimizing storage costs.
8. **Versioning:** S3 supports versioning, allowing you to keep multiple versions of an object in the same bucket. This helps protect against accidental data deletion or overwrites.

**Note:**

* Amazon S3 : Object storage built to store & retrieve any amount of data from anywhere
* 5 GB of S3 standard storage for 12 months with the AWS Free Tier.
* S3 is a storage service in AWS cloud.
* S3 is unlimited storage.
* S3 is object based storage.
* In s3 we can store all flat files (any type of file).
* We can upload, download and access files from S3 at any point of time
* The files in s3 can't be executed / edited.
* We can't install OS, DB etc in S3.
* We can't attach S3 to EC2 instance but we can access s3 buckets data from EC2 instance
* S3 supports static website hosting.
* In S3 we will store data in buckets. Bucket is a container & bucket contains objects
  + object = file
  + key is name of the object
* S3 is global but buckets are regional.
* Bucket names are universal or unique.
* Always create a bucket with your company name or project name.
* We can't create one bucket inside another bucket.
* We can create multiple buckets in multiple regions.
* Max no. of buckets you can create in S3 is 100 (soft limit)
* By default buckets are private, if required we can make it public.
* create bucket inside that create folder called photos inside that upload puppy.jpg
* Every object will have its own URL/endpoint.
  + Ex: http://7ambukcet.s3.amazonaws.com/photos/puppy.jpg
  + bucketname+ domain + object name
* S3 uses WORM model (Write Once and Read Many)

**Hosting a static website using AWS S3:**

**1. Create an S3 Bucket:** Sign in to the AWS Management Console, navigate to the S3 service, and click the "Create bucket" button. Give your bucket a unique name (bucket names must be globally unique across all of AWS) and choose the region where you want to host the website.

**2. Enable Static Website Hosting:** After creating the bucket, go to the bucket ***properties***, and click on the "Static website hosting" card. Select the option "Use this bucket to host a website." Enter the index document (e.g., ***index.html***) and, optionally, the error document (e.g., error.html) for handling 404 errors.

**3. Upload Your Website Content:** Upload your static website files (HTML, CSS, JavaScript, images, etc.) to the S3 bucket using the AWS Management Console or any S3-compatible client tool.

**4. Set Permissions:** Ensure that the objects (files) in your bucket have appropriate permissions to be publicly accessible. You can either apply a bucket policy to allow public read access or configure ***Access Control Lists (ACLs)*** on individual objects.

**5. Enable Website Access:** Once your files are uploaded and public access is set, your website is ready to be accessed. In the "Static website hosting" section, you will find the endpoint URL for your website (e.g., `http://your-bucket-name.s3-website-your-region.amazonaws.com`). Users can now access your website using this URL.

**Example:** Website Code GitHub Repo: https://github.com/ashokitschool/s3\_static\_website\_hosting.git

1) Create a bucket in S3

- Enter unique name for bucket

- uncheck block public access

2) Upload Website content files in bucket (assets folder, index.html and error.html) make sure you given public access for files which we are uploading.

3) Go to Bucket Properties tab -> Enable Static Website hosting and configure index and error pages

index.html for main content, error.html for wrong URL.

4) It will display URL, access that URL.

5) Go to Bucket Permissions tab -> Object Ownership -> Edit -> ACLs enabled -> Save Changes

6) Come back to Bucket files -> select all files -> Click on Action Drop Down -> Choose Make Public using ACL. Now you try to access URL.

**S3 Versioning:**

Simple Storage Service (S3) supports versioning, which is a feature that allows you to keep multiple versions of an object in the same bucket. This means that when you update an object in an S3 bucket, instead of overwriting the existing object, S3 creates a new version of the object, preserving the previous versions.

**Versioning can be beneficial for various reasons:**

**1. Data Protection:** Versioning helps protect against accidental deletions or overwrites. If an object is deleted, the previous version remains accessible, reducing the risk of data loss.

**2. Point-in-Time Recovery:** Versioning allows you to restore objects to specific points in time. This can be useful in case of data corruption or if you need to revert to an earlier version of the object.

**3. Compliance and Retention Policies:** Some industries or regulations may require data retention for a specific period. Versioning enables you to meet these requirements by retaining object versions for the desired duration.

- Navigate to the Amazon S3 section in the AWS Management Console.

- Select the bucket for which you want to enable versioning.

- Click on the "Properties" tab.

- Under the "Advanced settings" section, click on "Versioning."

- Choose the "Enable" option and click "Save changes."

Note:

* Versioning is like a backup tool.
* By default versioning is not enabled on the bucket.
* Versioning is enabled on bucket level but applied on object level.
* When we upload same file multiple times then versions will be created.
  + corejava.jpg (v1, v2, v3) - v3 latest version.
  + advjava.jpg (v1, v2, v3) - v3 is latest version.
* If somebody deleted my original object by mistake, for latest version delete marker label will be applied.
* Version ID is always unique.
* Versioning files we can download at any time.
* Once you have enabled the versioning, you can't disable it. You can only suspend it.
* AWS charges for Versioning, be careful while you enable versioning for huge files.
* S3 is unlimited storage
  + Min object size = 0 Bytes
  + Max object size = 5 TB
* We can have unlimited no. of 5TB objects in a single bucket.

***Storage classes***

***https://aws.amazon.com/s3/pricing/***

**1. S3 Standard:** This is the default storage class for S3 and provides high durability, availability, and performance. It is suitable for frequently accessed data, such as web hosting content, big data analytics, and dynamic websites.

**2. S3 Intelligent-Tiering:** This storage class uses machine learning to automatically move objects between two access tiers: frequent access and infrequent access. It is well-suited for data with changing access patterns, as it optimizes costs by moving data to the most cost-effective tier.

**3. S3 Standard-IA (Infrequent Access):** Ideal for data that is accessed less frequently but still requires quick access when needed. It is a cost-effective choice for long-term storage, backups, and disaster recovery data.

**4. S3 One Zone-IA:** Similar to Standard-IA but stores data in a single Availability Zone, providing a lower-cost option for infrequently accessed data with lower redundancy.

**5. S3 Glacier:** This storage class is designed for long-term data archival. It offers lower costs but longer retrieval times (minutes to hours) compared to the previous classes. Suitable for regulatory compliance data, financial records, and other archives.

**6. S3 Glacier Deep Archive:** The most cost-effective storage class in S3, intended for data that is rarely accessed and can tolerate longer retrieval times (hours). Best suited for digital preservation, compliance archives, and cold data storage.

**Amazon S3 Lifecycle Management:**

Cost of data (Objects) storage is important to consider when making decisions to use right storage technology. Amazon S3 Life cycle allows to configure a lifecycle for stored objects on S3, ***to optimize the cost***. A Lifecycle configuration is a set of rules that define actions applied to a group of objects.

**Types of S3 Lifecycle configuration Actions:**

In Amazon S3 Lifecycle, there are two types of actions: transitions and expiration actions.

***Transition actions:*** It allows to transit of the objects from one storage class to another storage class.

***Expiration actions:*** It allows you to delete expired objects automatically on your behalf.

**Example of S3 lifecycle:**

Learn how to configure Amazon S3 Lifecycle for improved data storage and management. Example shown: S3 Lifecycle Configuration.


Application server, database logs are stored in s3 but logs may not require after a few weeks or months, in this case, you can delete the objects automatically by applying an expiration action.

Frequency of access requirement of an organization’s documents (financial, media, employee data) some documents are frequently accessed, but after a few days or months, they are infrequently accessed.

After some time, organization may need to archive them as documents are not used anymore but must be retained for regulatory compliance, in this case, you can use transition action.

**Supported lifecycle transitions**

Amazon S3 supports the following lifecycle transitions between storage classes using an S3 Lifecycle configuration.

**You *can transition* from the following:**


                    Amazon S3 storage class waterfall graphic.
                

* The S3 Standard storage class to any other storage class.
* The S3 Standard-IA storage class to the S3 Intelligent-Tiering, S3 One Zone-IA, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, or S3 Glacier Deep Archive storage classes.
* The S3 Intelligent-Tiering storage class to the S3 One Zone-IA, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, or S3 Glacier Deep Archive storage classes.
* The S3 One Zone-IA storage class to the S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive storage classes.
* The S3 Glacier Instant Retrieval storage class to the S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive storage classes.
* The S3 Glacier Flexible Retrieval storage class to the S3 Glacier Deep Archive storage class.
* Any storage class to the S3 Glacier Deep Archive storage class.