**AWS STORAGE:**

1. **EBS (Elastic Block Storage):** EBS provides simple, scalable, high available block storage. EBS can only attach to one EC2 system. It is a block storage.

Storage space sharing is not possible in EBS.

An EBS Snapshot refers to a point-in-time copy of an Amazon Elastic Block Store (EBS) volume. Amazon EBS provides block-level storage volumes for use with EC2 instances. EBS Snapshots are a key feature of Amazon EBS, offering data backup, disaster recovery, and data migration capabilities.

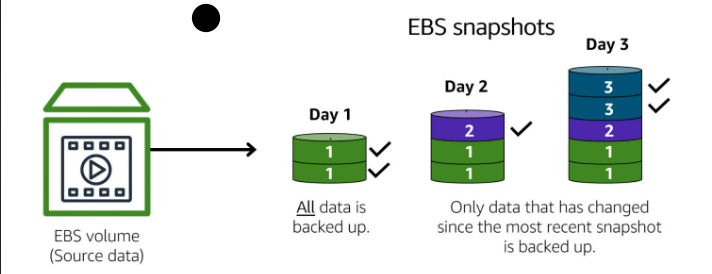
**Important 3 models of EBS:**

* Provision IOPS (64000)
* General purpose (16000)
* Magnetic (5000)

**Note: EBS can be attached to only one instance, EBS and Volume should be there in same region to avail the EBS**

**Benefits of EBS:**

* SSD storage technology (solid-state drive)
* Highly available, fast and scalable



**What is an EBS Snapshot?**

An **Amazon EBS Snapshot** is a **point-in-time backup** of an **EBS volume**, stored in **Amazon S3**. It allows you to restore data to a new volume in case of failure, migration, or scaling.

**How EBS Snapshots Work**

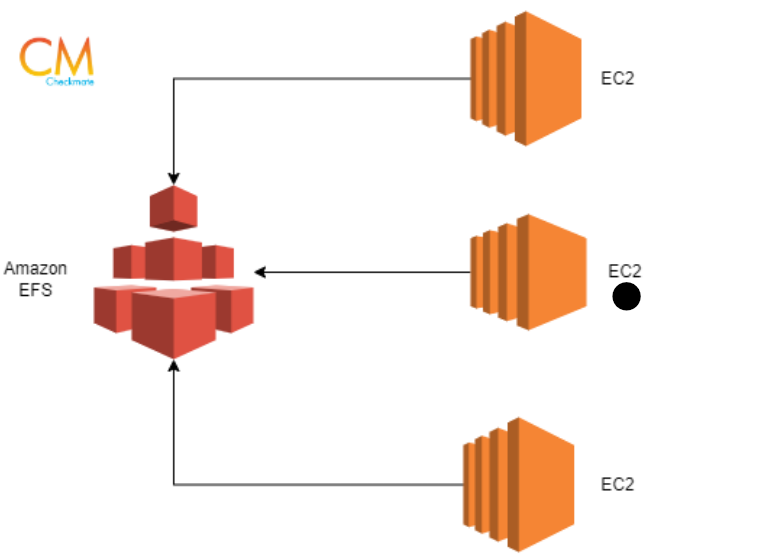
1. **First Snapshot (Full Backup)**
   * The first snapshot is a **full copy** of the EBS volume.
   * It captures all the data at that point in time.
2. **Subsequent Snapshots (Incremental Backups)**
   * AWS only saves **the changes** (delta) made since the last snapshot.
   * This **reduces storage costs** and speeds up the backup process.
3. **Restoring a Snapshot**
   * Snapshots can be used to create a new EBS volume.
   * The restored volume is independent of the original.



**2. EFS (Elastic File System):**

Amazon Elastic File System (EFS) is a cloud-based file storage service that lets you create and share files

It is similar to shared disk, EFS is more used in case of sharing the disk space. EFS storage is used in cluster management in order to have availability. Storage space sharing is possible in EFS and not possible in **EBS Eg – Cassandra cluster, Kubernetes, Machine learning and AWS lambda. Advantages of EFS: cost effective, Speed, Disk share**.



**AWS-S3-Bucket:**

It is basically an object storage. S3 cannot hold data, but it can store & hold data which are in form of objects. S3 is not region specific. We can host static website on S3.

**S3 Standard**

* High availability, low latency, and durability (99.999999999% durability).
* Ideal for frequently accessed data (e.g., websites, apps, big data).

**S3 Intelligent-Tiering**

* Automatically moves objects between access tiers based on access patterns.
* Cost-efficient for unpredictable access patterns.

**S3 Standard-IA (Infrequent Access)**

* Lower storage cost, but retrieval costs apply.
* Suitable for data that is accessed less frequently but needs fast retrieval.

**S3 One Zone-IA**

* Similar to Standard-IA but stored in a single AWS Availability Zone.
* Cheaper but less durable (data loss risk if the zone fails).

**S3 Glacier**

* Low-cost, long-term archival storage (retrieval time from minutes to hours).
* Used for compliance archives, backups, and digital preservation.

**S3 Glacier Deep Archive**

* Lowest-cost storage, retrieval time takes up to 12 hours.
* Used for data that is rarely accessed but must be retained long-term.

**S3 Versioning:**

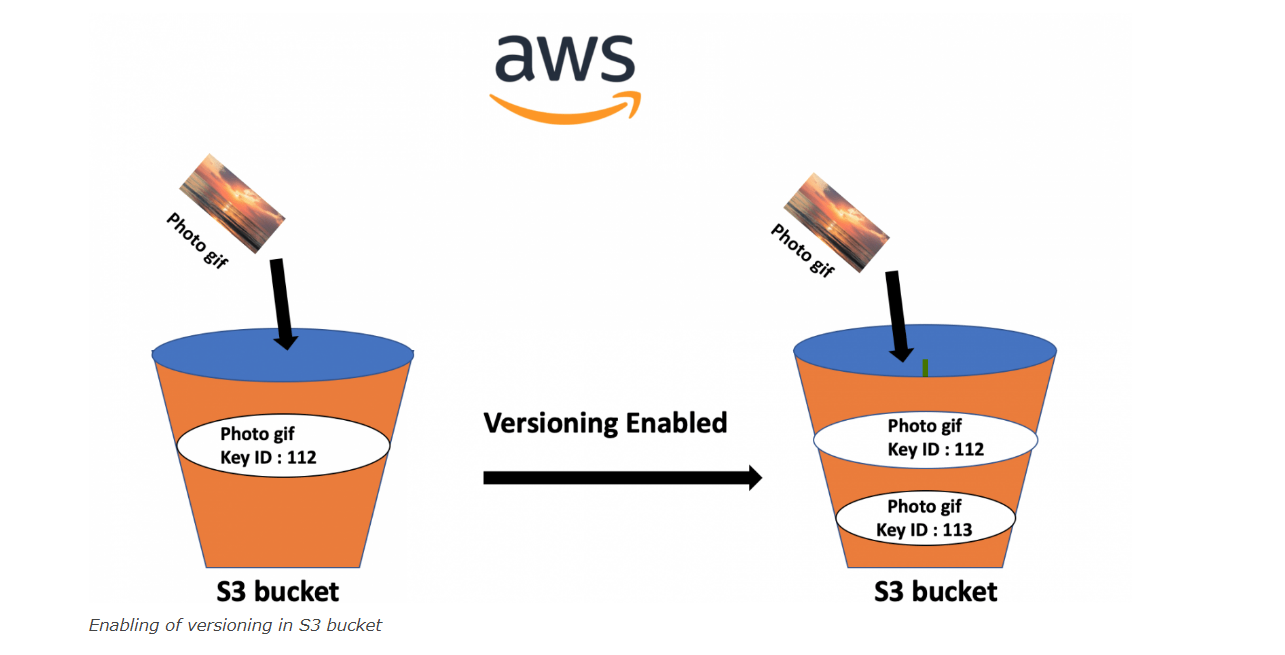
S3 **versioning** keeps multiple versions of an object, preventing accidental deletions or overwrites.

**1. How Versioning Works**

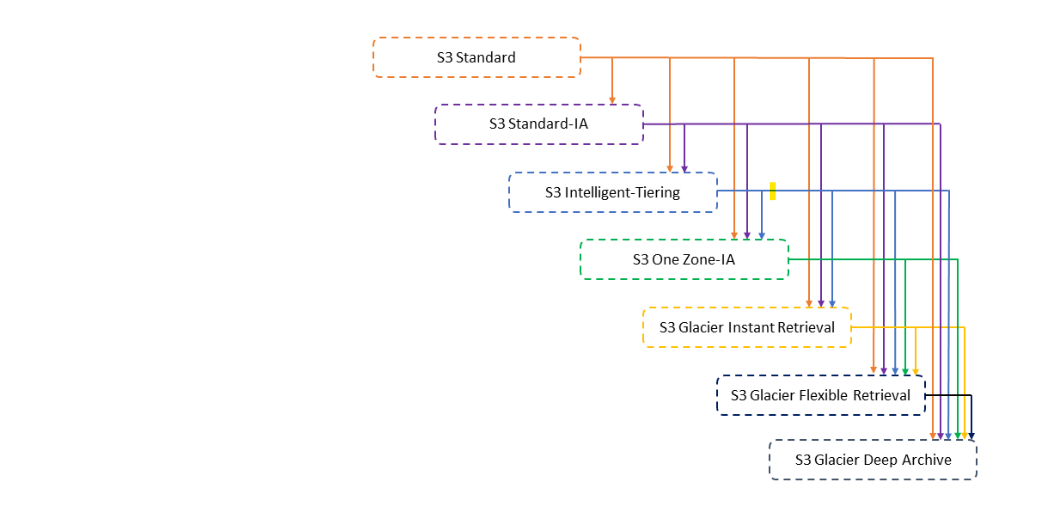
| **Action** | **Version ID** | **Object Exists?** |
| --- | --- | --- |
| Upload file.txt | v1 | ✅ Yes |
| Upload new file.txt | v2 | ✅ Yes (v1 still exists) |
| Delete file.txt | Delete marker added | ❌ (Hidden, but retrievable) |
| Remove delete marker | v2 restored | ✅ Yes |

**2. Managing Older Versions with Lifecycle Rules**

* **Keep latest X versions, delete older ones**.
* **Remove delete markers** after Y days to free space.



**S3 LIFE\_CYCLE\_MANAGEMENT:**



### **S3 Lifecycle – How Object Lifecycle Changes Over Time**

An object in an S3 bucket **changes its lifecycle** based on the defined **lifecycle rules**. The key lifecycle transitions are:

### **1. Object Creation (Day 0)**

* When an object is uploaded to an **S3 Standard** bucket, it remains there until a lifecycle rule moves or deletes it.

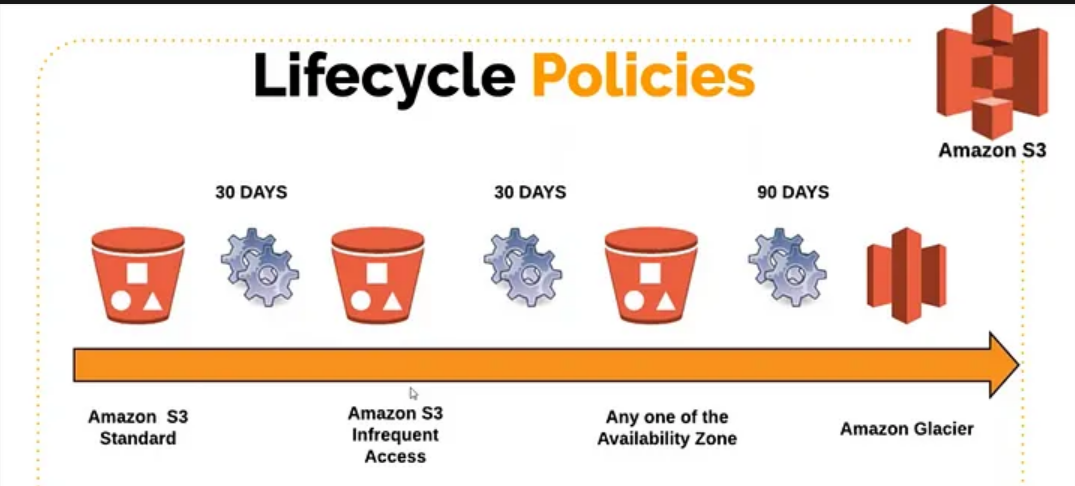
### **2. Transition Between Storage Classes**

Lifecycle policies **automatically move** objects to lower-cost storage tiers based on time.

**Example Lifecycle Rule:**  
✅ **Day 0:** Uploads object to **S3 Standard**.  
✅ **Day 30:** Moves object to **S3 Standard-IA** (cheaper, for infrequent access).  
✅ **Day 90:** Moves object to **S3 Glacier** (long-term storage, retrieval takes time).  
✅ **Day 365:** Deletes the object (if configured).

### **3. Expiration & Deletion**

* After a certain period (e.g., 1 year), objects **can be automatically deleted**.
* This helps **reduce storage costs** for temporary data (e.g., logs, backups).



|  |  |  |
| --- | --- | --- |
| Day | Action | Storage Class |
| 0 | Object uploaded | S3 Standard |
| 30 | Transition to Standard-IA | Lower-cost storage |
| 90 | Transition to Glacier | Archival storage |
| 365 | Object deletion | Storage cost removed |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Service** | **Default Storage** | **Max Size** | **Scaling** | **Best For** |
| **EBS** | Must specify | 16 TiB | Manual | EC2 block storage |
| **EFS** | 0 to Petabytes | No limit | Auto | Shared file storage |
| **S3** | Unlimited | Unlimited (5 TiB per object) | Auto | Object storage |

Real life use case of S3 Bucket

**1. Netflix – Storing and Streaming Content**

* **Use Case**: Netflix uses Amazon S3 to store and manage vast amounts of video content.
* **Real-Life Example**: When you watch a movie on Netflix, the video files are stored in S3 and delivered via a CDN (CloudFront).
* **Why S3?**: It provides scalable, reliable storage with high availability for streaming services.

**2. Airbnb – Storing Images and User Data**

* **Use Case**: Airbnb stores millions of user-uploaded images (e.g., property photos) on S3.
* **Real-Life Example**: When you browse Airbnb listings, the images you see are retrieved from an S3 bucket.
* **Why S3?**: It offers cost-effective storage with lifecycle policies for archiving old or unused data.

**3. NASA – Archiving Satellite Data**

* **Use Case**: NASA stores and shares vast amounts of satellite and space mission data.
* **Real-Life Example**: The Earth Science Data Systems program archives historical satellite imagery on S3, making it available for researchers.
* **Why S3?**: It ensures data durability, security, and global accessibility.

**4. Pinterest – Managing Millions of Images**

* **Use Case**: Pinterest relies on S3 to store and serve billions of image pins.
* **Real-Life Example**: When you save or browse pins, the images are fetched from S3.
* **Why S3?**: It allows for fast, scalable storage to handle high traffic loads.

**5. Reddit – Hosting User-Uploaded Content**

* **Use Case**: Reddit uses S3 to store user-generated media, such as images and videos.
* **Real-Life Example**: When users upload memes or videos, they are stored in an S3 bucket and served via a CDN.
* **Why S3?**: It enables efficient storage with cost control via S3 Intelligent-Tiering.

**6. Pfizer – Secure Storage for Medical Research**

* **Use Case**: Pharmaceutical companies use S3 to store clinical trial data securely.
* **Real-Life Example**: Pfizer uses AWS to store genomic and medical research data while ensuring compliance with regulatory requirements (e.g., HIPAA).
* **Why S3?**: It offers encryption, compliance, and seamless integration with analytics tools.

**7. Financial Institutions – Log and Transaction Storage**

* **Use Case**: Banks use S3 to archive logs and transaction data.
* **Real-Life Example**: A bank stores millions of daily transaction logs on S3 for auditing and fraud detection purposes.
* **Why S3?**: It allows cost-effective archival storage with lifecycle management.

**8. Tesla – Storing Autonomous Driving Data**

* **Use Case**: Tesla collects vast amounts of sensor and camera data from its fleet for AI training.
* **Real-Life Example**: Data from Tesla’s self-driving cars is uploaded to S3 for analysis and model improvement.
* **Why S3?**: It can handle petabytes of data efficiently while integrating with machine learning tools.

**9. Spotify – Storing Music Files and Metadata**

* **Use Case**: Spotify stores music files, album covers, and user-generated playlists on S3.
* **Real-Life Example**: When you stream a song, the audio file is retrieved from S3.
* **Why S3?**: It provides high-speed access to audio files for uninterrupted streaming.

**10. Government Agencies – Disaster Recovery and Archiving**

* **Use Case**: Governments store critical documents and disaster recovery backups on S3.
* **Real-Life Example**: The U.S. government archives legal documents, census data, and emergency response plans on AWS S3 Glacier.
* **Why S3?**: It ensures long-term, secure, and tamper-proof storage.