☁️ Module 4: AWS Storage Services

◆ Introduction

AWS offers a wide range of **cloud storage services** designed for different purposes — such as storing files, databases, backups, or archives.
These services are **scalable, secure, and cost-efficient**, helping users store and manage data efficiently.

◆ 1. Amazon S3 (Simple Storage Service)

Definition

Amazon S3 is an **object storage service** used to store and retrieve any amount of data from anywhere on the web.
It is designed for **99.999999999% (11 nines)** durability and is one of the most popular AWS services.

🎁 Key Components

1. **Bucket:**
   - A **container** for storing objects (like a folder).
   - Each bucket has a **unique name** globally.
   - Example: `sakthi-images-bucket`
2. **Object:**
   - The **actual data** stored in a bucket (file, image, document, video, etc.).
   - Each object includes:
     - **Data** (the file content)
     - **Metadata** (file info)
     - **Unique key name**
3. **Key:**
   - The **unique identifier** for an object in a bucket (like file path).

🔐 Permissions and Access Control

Access to S3 data can be controlled in multiple ways:

- **Bucket Policy:**
  Controls access for all objects in a bucket.

- **Access Control Lists (ACLs):**
  Control permissions at object-level.
- **IAM Policies:**
  Manage user-level permissions to access S3.

Example Policy:

```
{
  "Effect": "Allow",
  "Action": "s3:*",
  "Resource": "arn:aws:s3:::sakthi-bucket/*"
}
```

⚙ Storage Classes in S3

Storage Class	Description	Use Case
S3 Standard	High availability & durability	Frequently accessed data
S3 Intelligent-Tiering	Automatically moves data to cheaper tiers	Unknown access patterns
S3 Standard-IA (Infrequent Access)	Lower cost, slower access	Backup & disaster recovery
S3 One Zone-IA	Stored in a single zone	Low-cost, non-critical data
S3 Glacier	Archival storage	Long-term data archiving
S3 Glacier Deep Archive	Lowest cost, very slow retrieval	Regulatory archives

---

🌟 Features of S3

- Virtually **unlimited storage**
- **Versioning** (track multiple versions of an object)
- **Cross-region replication**
- **Static website hosting**
- **Encryption** (server-side or client-side)
- **Lifecycle policies** (automatic data movement)

💡 Example Use Cases

- Hosting static websites.
- Backup and restore systems.
- Data lakes and analytics storage.
- Application logs and image storage.

◆ 2. Amazon EBS (Elastic Block Store)

Definition

Amazon EBS provides **block-level storage volumes** for use with EC2 instances.
Think of it as a **hard disk drive** attached to your EC2 instance.

⚙ Key Features

- **Persistent storage:** Data remains even after EC2 stops.
- **Scalable:** Increase volume size or performance anytime.
- **Encrypted by default:** Keeps data secure.
- **Snapshots:** Create backups of volumes stored in S3.

💾 EBS Volume Types

Volume Type	Description	Use Case
gp3 / gp2 (General Purpose SSD)	Balanced performance and cost	Boot volumes, web servers
io2 (Provisioned IOPS SSD)	High IOPS for heavy workloads	Databases
st1 (Throughput Optimized HDD)	High throughput	Big data, log processing
sc1 (Cold HDD)	Lowest cost, infrequently accessed	Backup storage

---

EBS Snapshots

- A **point-in-time backup** of your EBS volume.
- Stored in **Amazon S3** automatically.
- Can be used to restore or copy volumes to other regions.

💡 Example Use Case

- An EC2 instance running a database uses an EBS volume for storage.
  The admin takes daily snapshots for recovery purposes.

◆ 3. Amazon Glacier (Now part of S3 Glacier)

Definition

Amazon Glacier is a **low-cost archival storage** solution used to store data that is rarely accessed but must be retained for long periods (e.g., compliance data, backups).

⚙ Key Features

- Designed for **data archiving** and **long-term backups**.
- **Extremely low cost** compared to S3 Standard.
- **Retrieval options:**
  - **Expedited (1–5 min)**
  - **Standard (3–5 hrs)**
  - **Bulk (5–12 hrs)**

🌟 Benefits

- Highly durable and secure.
- Integration with **S3 Lifecycle Policies**.
- Ideal for compliance and record retention.

💡 Example Use Case

A financial company stores old transaction records in **S3 Glacier Deep Archive** for 10 years to meet regulatory requirements.

◆ 4. AWS Storage Gateway

Definition

AWS Storage Gateway is a **hybrid cloud storage service** that connects on-premises systems with AWS cloud storage.
It allows organizations to **seamlessly integrate local applications with AWS storage**.

⚙ Types of Storage Gateways

Gateway Type	Description	Use Case
File Gateway	Provides file-based access (NFS/SMB)	Backup, file shares
Volume Gateway	Provides block storage to on-prem apps	Disaster recovery
Tape Gateway	Replaces physical tape libraries	Archival storage

---

💡 Example Scenario

A company has an on-premises application that needs to store daily backups.
Instead of buying extra hardware, they use **AWS Storage Gateway (File Gateway)** to back up data directly to **Amazon S3**.

🌟 Benefits

- Connects existing data centers with AWS.
- Reduces hardware and maintenance costs.
- Secure and scalable cloud storage integration.

◆ 5. Data Lifecycle Management

Definition

Data Lifecycle Management helps to **automate the movement of data** between different storage classes based on **age, access frequency, or cost requirements**.

⚙ In Amazon S3: Lifecycle Policies

Lifecycle rules automatically transition data:

- From **S3 Standard → S3 IA → S3 Glacier** as data becomes older.
- Or **delete objects** after a specified number of days.

💡 Example Rule

- Files not accessed for 30 days → move to **S3 Infrequent Access**.
- Files older than 180 days → move to **S3 Glacier**.

- Files older than 365 days → **delete permanently**.

🌟 Benefits

- Reduces storage cost.
- Automates data management.
- Ensures compliance with data retention policies.

◆ 6. Summary Table

AWS Service	Type	Key Feature	Use Case
Amazon S3	Object Storage	Store any file/data	Website hosting, backups
Amazon EBS	Block Storage	Persistent EC2 volume	Databases, boot disks
Amazon Glacier	Archival Storage	Long-term, low-cost	Compliance, archives
Storage Gateway	Hybrid Storage	Connect on-prem with cloud	Backup & recovery
Lifecycle Management	Data Automation	Move data between classes	Cost optimization