AWS S3

Amazon Simple Storage Service (Amazon S3)

Introduction to Amazon S3

What is Amazon S3?

Amazon Simple Storage Service (Amazon S3) is an object storage service that provides scalability, durability, security, and high availability. It is designed for:

- **Storing and retrieving any amount of data from anywhere**.
- **Use cases like backups, archiving, data lakes, and static website hosting**.
- **Handling large-scale workloads such as big data analytics and machine learning**.

Benefits of Amazon S3

- **Highly Scalable**: Supports unlimited storage and scales automatically.
- **Secure**: Provides encryption, access control policies, and IAM-based authentication.
- **Durable & Reliable**: 99.999999999% (11 9's) durability and 99.99% availability.
- **Flexible Storage Options**: Multiple storage classes for cost optimization.
- **Easy Data Access & Management**: Simple API interface and AWS SDK support.
- **Integration**: Works with AWS services like Lambda, CloudFront, and Athena

Amazon S3 Key Concepts

Buckets and Objects

Amazon S3 stores data in **buckets**, which contain **objects** (files).

- **Bucket**: A container for objects; globally unique name required.
- **Object**: The actual file stored in a bucket, along with metadata.
- **Key**: The unique identifier for an object within a bucket.
- **Metadata**: Additional data (content type, last modified, etc.).

Amazon S3 Data Consistency

- **Strong Read-After-Write Consistency**: Immediate availability after write/update.
- **Atomic Updates**: No partial writes, ensuring data integrity.

Amazon S3 Storage Classes

Amazon S3 offers different storage classes based on data access frequency and cost.

Storage Class	Use Case	Durability	Availability	Cost
S3 Standard	Frequently accessed data	11 9's	99.99%	High
S3 Intelligent-Tiering	Automatic cost optimization	11 9's	99.9%	Moderate
S3 Standard-IA	Infrequently accessed data	11 9's	99.9%	Lower
S3 Glacier	Long-term archival	11 9's	99.9%	Very Low
S3 Glacier Deep Archive	Archival for years	11 9's	99.9%	Lowest

Choosing the Right Storage Class

- **Frequent access → S3 Standard**
- **Occasional access → S3 Standard-IA**
- **Backup/Archives → S3 Glacier & Deep Archive**

S3 Bucket Naming & Configuration

Bucket Naming Rules

- Must be globally unique.
- Can contain lowercase letters, numbers, dots, and hyphens.
- Cannot contain uppercase letters or underscores.
- Cannot be formatted as an IP address.

Bucket Configuration

- **Region Selection**: Choose an AWS region close to your users to reduce latency.
- **Public Access Settings**: By default, S3 buckets are private.
- **Versioning**: Helps track multiple versions of an object.
- **Logging**: Enables tracking access requests.

Uploading & Managing Data in S3

Uploading an Object

1. Open AWS S3 Console.
2. Click "Create Bucket" and enter details.
3. Upload files by clicking "Upload".
4. Configure permissions and storage class.
5. Click "Upload" to save the object.

Accessing an Object

- Use AWS SDK, CLI, or Console.
- Generate a pre-signed URL for temporary access.
- Restrict access via IAM policies and bucket policies.

Summary

- Amazon S3 is a scalable, durable, and secure storage solution.
- Supports multiple storage classes for cost-effective data management.
- Provides strong security with IAM, encryption, and access control policies.
- Easily integrates with AWS services for data processing and analytics.

Next Steps

- Learn about **S3 Security & Access Control**
- Implement **encryption, IAM policies, and bucket policies** for better security.
- Explore **S3 Versioning & Replication** for data protection.