

AWS S3 – Full In-Depth Notes

◆ What is Amazon S3?

Amazon S3 (Simple Storage Service) is an **object storage** service offering industry-leading scalability, availability, and performance.

Key traits:

- Stores **objects** (files) in **buckets**.
 - Unlimited storage capacity.
 - Designed for **11 nines durability** (99.999999999%).
-

◆ Core Concepts

Concept	Description
---------	-------------

Bucket	A container for storing objects (like a root folder).
---------------	---

Object	A file stored in a bucket, includes metadata and a unique key.
---------------	--

Key	The unique identifier for an object within a bucket.
------------	--

Region	Buckets are region-specific; choose based on latency and compliance.
---------------	--

Prefix	The folder-like part of the object key (folder1/file.jpg).
---------------	--

Metadata	Key-value pairs associated with the object (system-defined or custom).
-----------------	--

◆ Storage Classes

Storage Class	Use Case
S3 Standard	Frequently accessed data
S3 Intelligent-Tiering	Automatically moves data between frequent/infrequent tiers
S3 Standard-IA	Infrequent Access, still rapid access
S3 One Zone-IA	IA but stored in a single AZ
S3 Glacier	Archival storage – retrieval in minutes/hours
S3 Glacier Deep Archive	Long-term archiving – retrieval within 12 hours
S3 Express One Zone (new)	High-performance for temporary data

◆ S3 Object Features

- **Versioning:** Retains multiple versions of an object.
- **Object Lock:** Write-once-read-many (WORM) protection.
- **Tags:** Key-value pairs for managing and billing.
- **Metadata:** System-defined (e.g., Content-Type) or user-defined.
- **Encryption:**
 - **SSE-S3:** Amazon-managed keys
 - **SSE-KMS:** Customer-managed keys via KMS
 - **SSE-C:** Customer-provided keys
 - **Client-side encryption** (before uploading)

◆ Access Management

Method	Description
IAM Policies	Attach to users/roles to allow or deny access
Bucket Policies	Resource-based policies for entire buckets
ACLs (Access Control Lists)	Legacy method for per-object access
Pre-Signed URLs	Temporary access links generated via SDK or CLI

✅ Best Practice: Use **IAM + Bucket Policies**, avoid ACLs unless required.

◆ S3 URLs

- **Virtual-hosted-style:** `https://bucket-name.s3.Region.amazonaws.com/key`
- **Path-style (legacy):** `https://s3.Region.amazonaws.com/bucket-name/key`

◆ Lifecycle Management

Automate transition or deletion:

Action	Description
Transition	Move data to another storage class after N days
Expiration	Automatically delete objects after N days
Noncurrent Version Expiration	Delete old versions (when versioning is on)

Use **lifecycle rules** via console or JSON.

◆ Object Lock & Retention

S3 Object Lock lets you:

- **Prevent deletion/modification** for a defined period.
 - Use for regulatory or compliance needs (WORM).
 - Modes:
 - **Governance:** AWS account with special permission can delete.
 - **Compliance:** No one can delete—not even root.
-

◆ S3 Security & Encryption

- **Encryption at rest:** SSE-S3, SSE-KMS, SSE-C
 - **Encryption in transit:** HTTPS (TLS)
 - **Block Public Access:** Feature to prevent accidental exposure.
 - **AWS WAF & CloudFront:** Add extra layers for web-facing content.
-

◆ S3 Performance

- **Scales automatically** for millions of requests per second.
 - **Best practices:**
 - Use **random prefixes** for keys to increase performance.
 - Use **multipart upload** for files > 100MB.
 - Parallelize downloads/uploads for large files.
-

◆ Data Consistency

- **Strong consistency:** All read-after-write or overwrite operations are instantly reflected.
-

◆ S3 Event Notifications

Trigger events on object actions:

- **Put, Delete, Restore**, etc.
- Targets:
 - **SNS**

- **SQS**
- **Lambda**

Use cases: Image processing, backups, log parsing.

◆ **Hosting Static Websites**

- Enable **Static Website Hosting** on a bucket.
 - Upload HTML, CSS, JS files.
 - Configure:
 - **Index document** (e.g., index.html)
 - **Error document** (e.g., error.html)
 - Must set **bucket policy** for public read access.
-

◆ **Logging & Monitoring**

Tool	Purpose
S3 Access Logs	Log all requests to the bucket
CloudTrail	API-level logging
CloudWatch Metrics	Monitor requests, errors, data usage
AWS Config	Track changes to bucket settings

◆ **S3 Cross-Region Replication (CRR)**

- Automatically replicate objects to a bucket in another region.
 - Must enable **versioning**.
 - Common for **disaster recovery, compliance, latency** reduction.
-

◆ **S3 Transfer Acceleration**

- Uses **Amazon CloudFront** edge locations to speed up uploads.
 - Best for uploading large files **from remote locations**.
 - URL format: bucketname.s3-accelerate.amazonaws.com
-

◆ **S3 Select**

Query parts of objects using SQL (e.g., CSV, JSON, Parquet).

Benefits:

- Save bandwidth and time.
- Only fetch needed data from a large object.

Example:

```
SELECT s.name FROM S3Object s WHERE s.age > 30
```

◆ Pricing Overview (2024)

Component Pricing Model

Storage Per GB per month (varies by class)

Requests GET, PUT, DELETE priced separately

Data Transfer Outbound to the internet is billed

Replication CRR incurs additional cost

S3 Select Priced by data scanned and returned

Use the **AWS Pricing Calculator** for accurate estimates.

◆ Common Use Cases

Scenario How S3 Helps

Backup and Restore Store snapshots, logs, database exports

Data Lake Combine with Athena, EMR, Redshift

Web Hosting Serve static websites or media

Archival Store infrequently accessed logs/files in Glacier

ML/AI Store training data sets

Disaster Recovery Use CRR or versioning for resilience

◆ Best Practices

- ✓ **Enable versioning** for critical data
- ✓ Use **SSE-KMS** for sensitive data encryption
- ✓ Turn on **Block Public Access** unless needed
- ✓ Use **lifecycle rules** for cost optimization

- ✓ Enable **logging & monitoring**
- ✓ Always apply **least privilege** via IAM
- ✓ Test with **pre-signed URLs** for secure sharing
- ✓ Consider **multipart upload** for large files
- ✓ Use **Intelligent-Tiering** for unknown access patterns
- ✓ Set **retention policies** where compliance is critical