

AWS S3 Storage Classes & Data Transfer Methods Cheat Sheet

S3 Storage Classes

Amazon S3 offers multiple storage classes optimized for different use cases.

Storage Class	Use Case	Availability	Durability	Key Features
S3 Standard	Frequently accessed data	99.99%	99.999999999% (11 9s)	High redundancy, stored across 3+ AZs
S3 Intelligent-Tiering	Unpredictable access patterns	High	11 9s	Auto-moves between Standard & IA, small monitoring fee
S3 Standard-IA (Infrequent Access)	Less frequently accessed data	99.9%	11 9s	Lower storage cost, higher retrieval fee
S3 One Zone-IA	Backup, non-critical data	99.5%	11 9s	Stored in a single AZ, lower redundancy, lower cost
S3 Glacier	Long-term archival storage	Varies	11 9s	Retrieval: Expedited (1-5 min), Standard (3-5 hrs), Bulk (5-12 hrs)
S3 Glacier Deep Archive	Very rarely accessed data	Varies	11 9s	Lowest-cost storage, retrieval: Standard (12 hrs), Bulk (48 hrs)

Data Transfer Methods

Method	Best For	Features
AWS APIs (SDKs, CLI, REST API)	Developers integrating S3 into applications	Programmatic access, automation
Amazon Direct Connect	Secure, high-speed transfer from on-premises to AWS	Dedicated private link, low latency

Method	Best For	Features
AWS Storage Gateway	Hybrid cloud storage with local caching	Supports file, volume, and tape gateways
Amazon Kinesis Firehose	Streaming large-scale data analytics to S3	Real-time data streaming
S3 Transfer Acceleration	Faster uploads from global locations	Uses CloudFront edge locations for speed optimization
AWS Snow Family	Large data volume transfers (physical)	Snowball (TBs), Snowball Edge (compute + storage), Snowmobile (Exabytes)

Core S3 Concepts

- **Buckets:** Containers for storing objects (default limit: 100 per account).
- **Regions:** Buckets exist in specific AWS regions for performance & compliance.
- **Objects:** Actual stored data (similar to files).
- **Keys:** Unique identifiers for objects in a bucket.
- **Prefixes & Delimiters:** Simulate folder structure (e.g., marketing/reports/file.pdf).
- **Object URLs:** Web addresses for accessing objects.

Consistency Model

- **Eventual Consistency:** Changes (writes, updates, deletes) take time to propagate.
- **Strong Consistency:** Not guaranteed in S3; eventual consistency applies.

S3 Operations

- **Bucket Management:** Create, delete, and organize buckets.
- **Object Operations:** Write, read, delete, and manage object properties.
- **Listing Objects:** Uses keys, not folder structures.

S3 as a Static Website Host

- Can host static websites (HTML, CSS, JS).
 - Requires making objects public and configuring bucket policy.
-

S3 REST API & HTTP Methods

Operation	HTTP Method	Purpose
Create	PUT/POST	Upload an object
Read	GET	Retrieve an object
Update	PUT/POST	Modify an object
Delete	DELETE	Remove an object

Advanced S3 Features

1. Object Lifecycle Management

- Automates moving objects between storage classes based on age.
- Example: S3 Standard → S3 IA (after 60 days) → Glacier (after 120 days).

2. Encryption Options

- **Server-Side Encryption (SSE):**
 - **SSE-S3:** AWS-managed encryption keys.
 - **SSE-KMS:** Uses AWS KMS for key management.
 - **SSE-C:** You provide and manage encryption keys.
- **Client-Side Encryption:** Data is encrypted before upload (you manage keys).

3. Versioning

- Stores multiple versions of an object.
- Once enabled, cannot be disabled (can only be suspended).

4. Security & Access

- **MFA Delete:** Requires MFA to delete an object (extra security).
- **Access Logs:** Tracks all bucket activities.

5. Performance & Data Management

- **Multipart Upload:** Splits large files into parts for faster uploads.
- **Range GETs:** Retrieve specific byte ranges from an object.
- **Cross-Region Replication (CRR):** Automatically copies objects to another region.
 - Only replicates new objects, not existing ones!
- **Event Notifications:** Triggers alerts (SNS, SQS, Lambda) when changes occur.

Key Takeaways

- ✓ **S3 is object storage, not traditional file storage.**
- ✓ **Data is distributed across multiple AZs (except for One Zone-IA).**
- ✓ **Choose the right storage class based on cost, availability, and retrieval needs.**
- ✓ **Select an optimal data transfer method based on data volume & speed requirements.**
- ✓ **Highly scalable & durable (objects replicated across AZs).**
- ✓ **Not suitable for transactional databases (use EBS for block storage).**

Amazon S3 Glacier - Exam Cheat Sheet

1. What is Glacier?

- **Low-cost archival storage** for rarely accessed data.
- **AES-256 encryption enabled by default.**
- Available as **standalone service** or **S3 storage class (Glacier, Glacier Deep Archive).**

2. Glacier Retrieval Tiers

Tier	Time	Cost
Expedited	1-5 min	\$\$\$

Tier	Time	Cost
Standard	3-5 hrs	\$\$
Bulk	5-12 hrs	\$

- **5% of data retrieval per month is free** (non-rollover).

3. Key Glacier Concepts

- **Archive** → Stored data object (like S3 object).
- **Vault** → Container for archives (like an S3 bucket).
- **Vault Lock** → Enforces security policies to control access.

4. Glacier Integrations

- **S3 Lifecycle Policies** → Auto-move cold data to Glacier.
- **AWS Snow Family** → Transfer large volumes of data.
- **Storage Gateway** → Hybrid storage & backups to Glacier.

5. Best Practices

- ✓ **Limit retrieval access** to reduce costs.
- ✓ **Use lifecycle policies** for automatic transitions.
- ✓ **Plan retrievals** to avoid high-cost Expedited tier.

S3 Enhanced Features - Exam Cheat Sheet

1. Intelligent Tiering (Cost Optimization)

- Uses **Lifecycle Rules** to automatically move objects between storage tiers.
- Example: Move objects to **Glacier** after **90 days** of inactivity.
- **Caution:** Small objects may incur **higher transaction costs**.

2. Object Locking (Write Once, Read Many - WORM)

- Prevents modification or deletion of objects.
- **Enabled only at bucket creation** (requires **Versioning** to be enabled).
- Two modes:

- **Governance Mode** → Protects against accidental deletion (admins can override).
- **Compliance Mode** → No one, including root, can modify/delete until retention period expires.

3. Batch Operations (Automated Bulk Actions)

- Uses a **manifest file** (CSV/JSON) to perform bulk operations.
- Actions include:
 - ✓ Copy/Move/Delete objects
 - ✓ Modify object properties (e.g., storage class)
 - ✓ Restore objects from Glacier

Elastic Block Store (EBS) - Exam Cheat Sheet

1. What is EBS?

- **Block-level storage** for EC2 instances (like a hard drive in the cloud).
- **Durable storage** (persists even if the instance stops).
- Can be attached/detached to/from instances.

2. EBS Volume Types

Volume Type	Performance	Use Case
Magnetic (Standard HDD)	Slowest, Cheapest	Low-cost storage, rarely used
General Purpose SSD (gp3, gp2)	Baseline IOPS (up to 16,000)	Default for most workloads
Provisioned IOPS SSD (io1, io2)	High IOPS (10,000+), Guaranteed	High-performance databases
Cold HDD (sc1)	Lowest-cost, Low IOPS	Large, rarely accessed workloads
Throughput Optimized HDD (st1)	Higher throughput	Big data, logs, streaming

💡 Use EBS-optimized instances for SSD performance!

3. EBS Snapshots (Backup & Recovery)

- **Point-in-time backup** of EBS volumes.
- Used for **recovery**, **cloning instances**, or **moving across regions**.
- **Incremental** (only changes from the last snapshot are stored).

4. Volume Recovery & Encryption

- **Attach EBS Volume** to another instance for recovery.
- **EBS Encryption** → Encrypts the entire volume (including snapshots).
- **Regulatory Compliance** → Encryption might be required for compliance (e.g., HIPAA, GDPR)

Elastic File System (EFS) - Exam Cheat Sheet

1. What is EFS?

- **Fully managed, shared file storage** for Linux-based workloads.
- **Multi-instance access** → Many EC2 instances across multiple AZs can access it.
- **Uses NFSv4 protocol** for file system mounting.
- **Scales automatically** to petabytes with no provisioning required.

2. EFS vs. EBS vs. S3

Feature	EFS	EBS	S3
Storage Type	File System	Block Storage	Object Storage
Accessibility	Multi-instance, multi-AZ	Single instance, single AZ	Accessible via web/API
Performance	Low-latency, scalable	Lowest latency	High throughput, variable latency
Use Cases	Web servers, CMS, home directories	Boot volumes, DB storage	Static websites, backups, data lakes

Feature	EFS	EBS	S3
Redundancy	Multi-AZ replication	Single AZ replication	Multi-AZ replication
Scaling	Automatic	Manual (increase size)	Automatic

💡 **EFS = Shared Storage | EBS = Dedicated Instance Storage | S3 = Object Storage**

3. Performance Considerations

- **Latency:**
 - **EBS < EFS < S3**
- **Throughput:**
 - **EFS scales dynamically** based on demand.
 - **Provisioned Throughput** available for high-performance workloads.
- **Not supported on Windows** (Linux only!).

4. EFS Use Cases

- **Shared file storage** for multiple EC2 instances.
- **Content management systems (CMS)** like WordPress.
- **Big data processing** (logs, analytics).
- **Home directories & Dev environments.**
- **Container storage** (ECS, Kubernetes).

EFS & PrivateLink – Key Takeaways for AWS Exam

1. What is AWS PrivateLink?

- PrivateLink **secures access to AWS services** over **private VPC endpoints**, avoiding exposure to the public internet.
- Uses **Elastic Network Interfaces (ENI)** with fixed **private IPs** in your VPC.
- Reduces **latency, security risks, and data transfer costs.**

2. Why Use PrivateLink with EFS?

- **Cross-account access:** Share EFS storage with multiple AWS accounts securely.
- **Private connectivity:** Avoid public internet exposure while accessing EFS.
- **Lower data transfer costs:** Data stays within AWS's private network.
- **Improved security & compliance:** Useful for industries with strict security policies.

3. How to Set Up PrivateLink for EFS?

1. **Go to AWS VPC Console** → Select **"Endpoints"**.
2. Click **"Create Endpoint"** → Choose **Elastic File System (EFS)**.
3. **Select VPC, Subnets & Security Groups** (ensure they allow NFS traffic).
4. Create an **Interface Endpoint** (powered by PrivateLink).
5. Use the **provided private IP address** to mount EFS securely.

4. Key Considerations

- Each **VPC endpoint** has a **fixed private IP**, which incurs a **monthly charge**.
- Only **Linux instances** can use EFS (not Windows).
- **Multi-AZ architecture** is supported for high availability.

💡 **Exam Tip:** PrivateLink secures **cross-account** access to EFS, while still keeping traffic private inside AWS's backbone network! 🚀