

Adding a Storage Layer with Amazon S3

OBJECT STORAGE · BUCKETS & OBJECTS · STORAGE CLASSES · LIFECYCLE · VERSIONING · ACCESS CONTROL · STUDY & REVIEW
WELL-ARCHITECTED

01 MUST KNOW

if you remember five things, remember these

- 01 Buckets hold objects: Amazon S3 is object storage** for virtually unlimited unstructured data. A single object is **0 bytes–5 TB**; the bucket name is **globally unique across every AWS account and Region**; each bucket is **Regional**. An object = **key + version ID + value + metadata + sub-resources**, and the value is immutable — to change it you re-upload.
- 02 Durability ≠ availability:** S3 Standard is designed for **11 nines (99.999999999%) of durability** — objects stored redundantly across **≥3 Availability Zones** and verified with checksums — and **4 nines (99.99%) of availability**. Durability means data is not lost; availability means you can reach it when needed.
- 03 Class + lifecycle = cost control: Match the storage class to the access pattern, then automate with lifecycle rules.** Standard (frequent) → Standard-IA / One Zone-IA (infrequent) → Glacier (archive). A lifecycle configuration is **transition actions** plus **expiration actions**; **S3 Intelligent-Tiering** moves objects for you with no retrieval fees.
- 04 Private and encrypted by default:** New buckets and objects are **private** — only the account owner has access until you explicitly grant it — and every new object is **encrypted at rest with SSE-S3** by default. **Block Public Access** overrides everything; grant least privilege with **IAM and bucket policies**.
- 05 Versioning stops data loss: Versioning protects against accidental overwrite and delete.** Enable it at the bucket level: a same-key upload keeps both versions, and a DELETE just adds a recoverable **delete marker**. Once enabled it can only be **suspended, never disabled**, and each version adds storage cost.

02 KEY TERMS

TERM	DEFINITION
Bucket	Container for objects; name is globally unique; Regional; the unit for access control, billing, and usage reporting
Object	The fundamental entity stored in S3: file data plus metadata, identified by exactly one key
Object key	The unique name that identifies an object in a bucket; includes the full path from the bucket root (S3 has no real folders)
Prefix	A shared leading string used to group objects; a GET or LIST with a prefix returns only matching objects, and the console shows prefixes as folders
Metadata	A set of name-value pairs describing an object — user-defined metadata plus S3 system metadata
Durability	The assurance that stored objects are not lost; S3 Standard = 11 nines (99.999999999%)
Availability	The ability to access data when you need it; S3 Standard = 4 nines (99.99%)
Storage class	An S3 tier that balances retrieval latency, resilience, and cost (Standard, Intelligent-Tiering, Standard-IA, One Zone-IA, three Glacier classes)
S3 Lifecycle configuration	A set of rules that automatically transitions objects between storage classes or expires (deletes) them
Versioning	A bucket feature that keeps multiple variants of an object to recover from overwrites and deletes
Delete marker	The placeholder S3 makes the current version when you delete in a versioning-enabled bucket; prior versions stay retrievable by version ID
SSE-S3	Server-side encryption with S3-managed keys — the default encryption applied to every new object at no extra cost

03 BLOCK STORAGE VS. OBJECT STORAGE

S3 is object storage – know the contrast

BLOCK STORAGE	OBJECT STORAGE (AMAZON S3)
<i>fixed-size blocks that apps and file systems manage</i>	<i>whole files stored as objects with metadata</i>
— Data is split into fixed-size blocks, each with a unique identifier — Applications and file systems control how blocks are accessed, combined, and modified — Blocks are stored wherever most efficient — across systems and different operating systems — You can modify part of a file in place	— Data is stored as objects — data + metadata + an object key — Flat namespace; no real folders — prefixes only imply structure — Updating rewrites the entire object; values are immutable — Virtually unlimited scale for unstructured data

04 S3 STORAGE CLASSES AT A GLANCE

all: 11 nines durability · lifecycle-eligible

STORAGE CLASS	AZS	MIN DURATION	BEST FOR
S3 Standard	≥3	None	Frequently accessed data — cloud apps, dynamic sites, content distribution, big data analytics
S3 Intelligent-Tiering	≥3	None	Unknown or changing access — auto-moves objects between tiers; no retrieval fees
S3 Standard-IA	≥3	30 days	Infrequent access that still needs millisecond retrieval (older images, logs); higher retrieval cost
S3 One Zone-IA	1	30 days	Infrequent, re-creatable data (secondary backups) — lower cost, single AZ
S3 Glacier Instant Retrieval	≥3	90 days	Archive rarely accessed but needing millisecond retrieval (medical images)
S3 Glacier Flexible Retrieval	≥3	90 days	Archive retrieved 1–2×/year, asynchronously in minutes; backup and DR
S3 Glacier Deep Archive	≥3	180 days	Lowest cost; 7–10-year regulatory retention; accessed 1–2×/year

05 A LIFECYCLE RULE IN ACTION

transition + expiration, no app changes

01 Upload to **S3 Standard** (frequent access) → **02 Transition** to Standard-IA after 30 days → **03 Transition** to Glacier after 1 year → **04 Expiration** deletes the object when no longer needed

06 VERSIONING: SAME KEY, TWO OUTCOMES

enable at bucket level · default is off

ACTION	VERSIONING ENABLED	DISABLED / SUSPENDED
Upload same key	Keeps the old object and adds a new version ID; both retrievable	Overwrites the original; previous object is gone
Delete object	Adds a delete marker; prior versions still retrievable by version ID	Permanently deletes the object
GET (no version ID)	Returns the most recent version; a delete marker returns 404	Returns the object, or 404 if deleted
DELETE with version ID	Permanently removes that version, no delete marker (bucket owner only)	N/A

- ✗ “A bucket name only has to be unique in my account.” — It must be **globally unique across every AWS account and Region**; buckets share one universal namespace.
- ✗ “S3 has folders.” — S3 is a **flat namespace**. Folders are just shared **prefixes** in the object key; the console displays them as folders, but the key stores the full path.
- ✗ “11 nines is the availability number.” — 11 nines (99.999999999%) is **durability** (objects are not lost). S3 Standard **availability** is 4 nines (99.99%). Do not swap them.
- ✗ “One Zone-IA is just a cheaper Standard-IA.” — One Zone-IA stores data in a **single Availability Zone**, so it is less resilient; use it only for re-creatable or secondary data.
- ✗ “Glacier is always slow.” — S3 Glacier **Instant Retrieval** returns data in milliseconds. Only **Flexible Retrieval** and **Deep Archive** retrieve asynchronously (minutes to hours).
- ✗ “I can turn versioning back off.” — After enabling versioning you can only **suspend** it, never disable it. Suspended keeps existing versions but behaves like disabled for new writes.
- ✗ “I have to switch on encryption.” — Every bucket has **SSE-S3 on by default**; new objects are encrypted at rest automatically at no extra cost. Choose SSE-KMS, DSSE-KMS, or SSE-C only if a use case needs it.
- ✗ “S3 objects are only eventually consistent.” — S3 provides **strong read-after-write consistency** for GET, LIST, PUT, and DELETE on objects in all Regions. Only bucket-level *configuration* changes are eventually consistent.

08 SELF-QUIZ

answers are upside-down below

- Q1** What is the maximum size of a single object you can store in Amazon S3?
- Q2** A bucket name must be unique across what scope?
- Q3** How many nines of durability and of availability is S3 Standard designed for?
- Q4** Which storage class stores data in a single Availability Zone?
- Q5** Which storage class automatically moves objects to the most cost-effective tier based on access, with no retrieval fees?
- Q6** Which two action types make up an S3 Lifecycle configuration?
- Q7** In a versioning-enabled bucket, what does a DELETE without a version ID actually do?
- Q8** What is the default server-side encryption applied to every new S3 object?
- Q9** Which S3 feature overrides all other policies to keep buckets from being publicly accessible?
- Q10** Sample exam: salespeople worry about accidentally deleting or overwriting documents in S3 — which single action protects against unintended user actions?

ANSWER KEY (rotate the page)

Q1 5 TB **Q2** globally — across all AWS accounts and Regions (a universal namespace) **Q3** 11 nines durability (99.999999999%) and 4 nines availability (99.99%) **Q4** S3 One Zone-IA **Q5** S3 Intelligent-Tiering **Q6** transition actions and expiration actions **Q7** inserts a delete marker as the current version; prior versions stay retrievable **Q8** SSE-S3 (server-side encryption with S3-managed keys) **Q9** Amazon S3 Block Public Access **Q10** enable versioning