
amazon
web services


S3

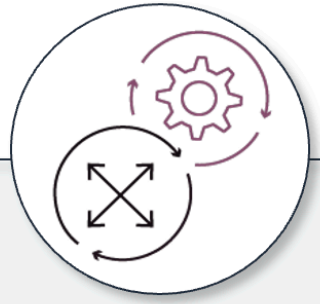
Simple Storage Service - S3

- Amazon S3 (Simple Storage Service) is an object storage provided by AWS that enables you to store and retrieve any amount of data at any time from anywhere on the web.
- It is designed to provide scalable storage with high availability, durability, and security.

Amazon S3 (Simple Storage Service)

Amazon Simple Storage Service (Amazon S3) is a durable object storage solution.

Accelerate innovation



Increase agility

Reduce cost

Strengthen security

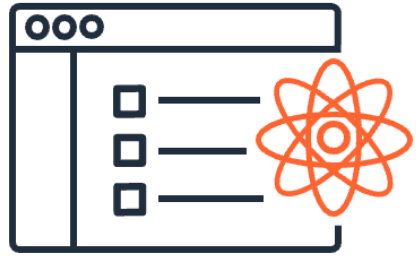
- **Accelerate innovation** – Integrate S3 buckets as storage solutions for static files and rely less on traditional file systems.
- **Increase agility** – With hosted object storage, you won't need to expand your storage as the quantity and size of data grows. Individual objects cannot be larger than 5 TB; however, you can store as much total data as you need.
- **Reduce cost** – Use the variety of storage tiers in Amazon S3 to spend less on infrequently accessed data. Archive data in S3 for your long-term storage needs.
- **Strengthen security** – Store your data in Amazon S3 and secure it from unauthorized access with encryption features and access management tools. S3 maintains compliance programs, such as PCI-DSS, HIPAA/HITECH, FedRAMP, EU Data Protection Directive, and FISMA, to help you meet regulatory requirements.

Amazon S3 use cases

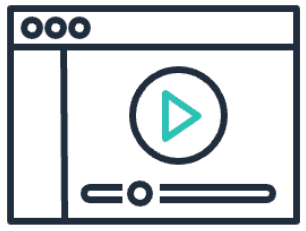
Use Amazon S3 when you have:

- Large number of users accessing your content
- Growing data sets
- Data you will write once and read many times

Backup and restore

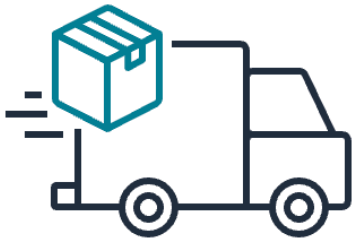


Data lakes for analytics



Media storage and streaming

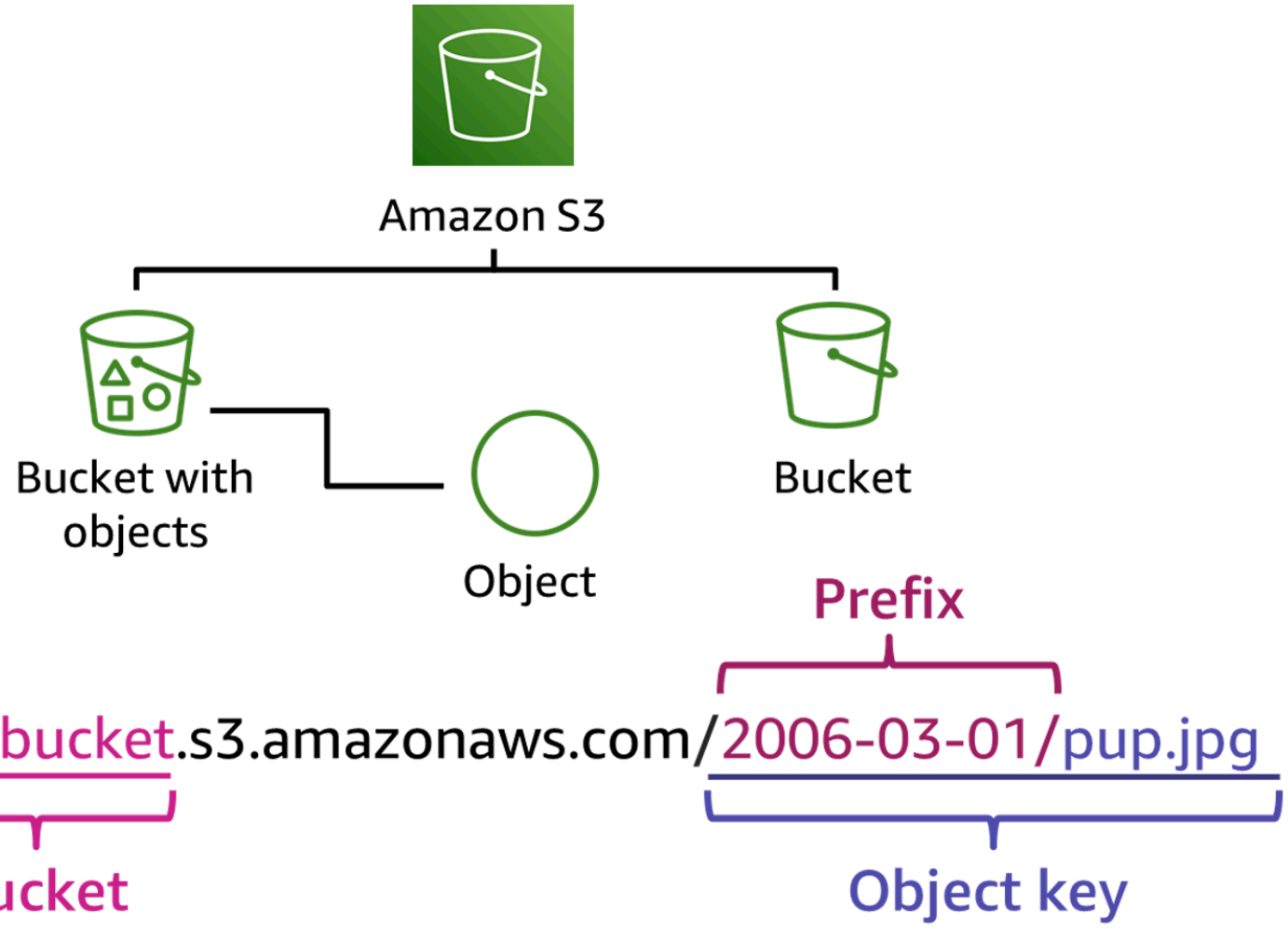
Static website



Archiving and compliance

Buckets & Objects

- Amazon S3 stores data as objects within buckets.
- An object includes a file and any metadata that describes the file.
- You can control access to the bucket and its objects.



Amazon S3 Storage Classes

Higher cost, frequent access ⟷ **Lower cost, infrequent access**

S3 Standard

- Active, frequently accessed data
- Milliseconds to access

S3 Standard-IA

- Infrequently accessed objects
- Milliseconds to access

S3 One Zone-IA

- Re-creatable, less accessed data
- Milliseconds to access

S3 Glacier Instant Retrieval

- Archived data that needs fast restore times
- Milliseconds to restore

S3 Glacier Flexible Retrieval

- Objects with unpredictable restore needs
- Minutes to hours to restore

S3 Glacier Deep Archive

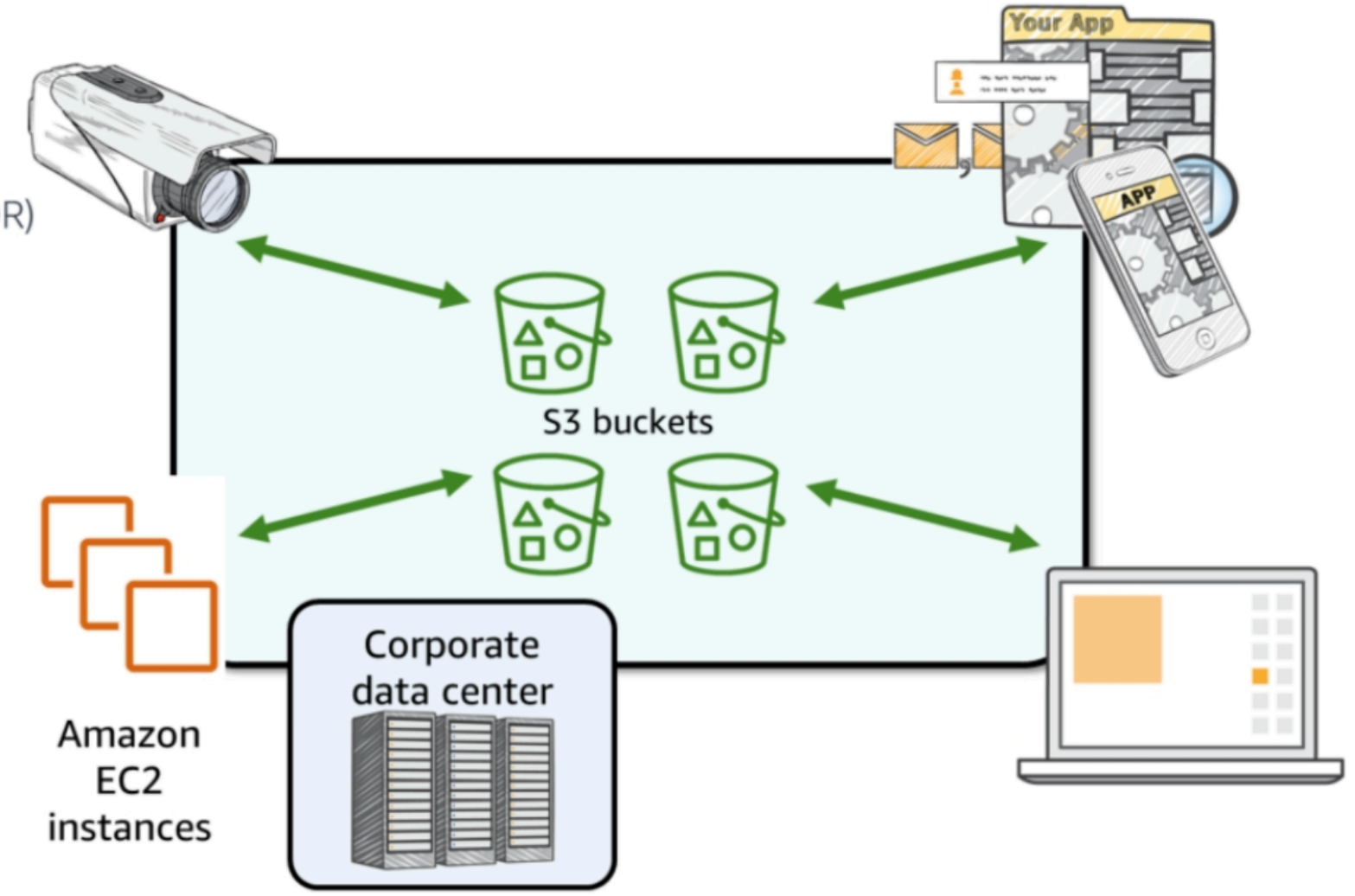
- Archive data not likely to be restored
- 12 hours or less to restore

S3 Intelligent-Tiering – Data with unknown or changing access patterns. Milliseconds to access.

- **S3 Standard** for general-purpose storage of frequently accessed data.
- **S3 Standard-Infrequent Access (S3 Standard-IA)** for long-lived, but less frequently accessed data.
- **S3 One Zone-Infrequent Access (S3 One Zone-IA)** for long-lived, less frequently accessed data that can be stored in a single Availability Zone.
- **S3 Glacier Instant Retrieval** for archive data that is rarely accessed but requires a restore in milliseconds.
- **S3 Glacier Flexible Retrieval** for the most flexible retrieval options that balance cost with access times ranging from minutes to hours. Your retrieval options permit you to access all the archives you need, when you need them, for one low storage price. This storage class comes with multiple retrieval options: expedited retrievals (restore in 1–5 minutes), standard retrievals (restore in 3–5 hours), or bulk retrievals (restore in 5–12 hours). Bulk retrievals are available at no additional charge.
- **S3 Glacier Deep Archive** for long-term cold storage archive and digital preservation. Your objects can be restored in 12 hours or less.

Common use cases for Amazon S3

- Storing application assets
- Static web hosting
- Backup and disaster recovery (DR)
- Staging area for big data
- Many more….



Amazon S3 pricing

- **Pay Only for what you use :**
  - GBs per month
  - Transfer OUT to other Regions
  - PUT, COPY, POST, LIST, and GET requests

- **You do NOT have to pay for :**
  - Transfer IN to Amazon S3
  - Transfer OUT form Amazon S3 to Amazon CloudFront or Amazon EC2 in the same Region

Bucket Versioning

- You can use S3 Versioning to keep multiple versions of an object in one bucket so that you can restore objects that are accidentally deleted or overwritten. For example, if you apply S3 Versioning to a bucket, the following changes occur:

- If you delete an object, instead of removing the object permanently, Amazon S3 inserts a delete marker, which becomes the current object version. You can then restore the previous version.
- If you overwrite an object, Amazon S3 adds a new object version in the bucket. The previous version remains in the bucket and becomes a noncurrent version. You can restore the previous version.

Simple site hosting with Amazon S3 and HTTP

S3 lifecycle rules

An S3 lifecycle rule is a set of configurations that defines actions you want AWS S3 to take at a specified time after an object's creation. The main actions include:

1. **Transition Actions**: Moving objects between storage classes, like from **S3 Standard** to **S3 Infrequent Access (IA)** or **S3 Glacier** for cost savings.
2. **Expiration Actions**: Permanently deleting objects after a certain period.

S3 lifecycle rules - example

Real-Time Example:

Imagine you manage a company's data for processing daily transaction logs. The logs are needed in real-time for analysis for the first 30 days, after which they are accessed occasionally for audits or reference. After a year, they no longer hold value, so you want to delete them to avoid unnecessary storage costs.

Here's how you could set up lifecycle rules for these logs in S3:

1. **Transition to S3 IA (Infrequent Access)**: After 30 days, a lifecycle rule could move the logs from **S3 Standard** to **S3 IA**, as they are accessed less frequently. This reduces storage costs because S3 IA is cheaper than S3 Standard for infrequent access data.
2. **Transition to S3 Glacier**: After 90 days, you could set another rule to move the logs from **S3 IA** to **S3 Glacier** for archival storage, as these logs will be rarely accessed but need to be kept for audit purposes.
3. **Expiration**: After 365 days, a rule could delete the logs, as they're no longer required.

This setup would allow you to minimize costs by aligning the storage class with the access pattern over the data's lifecycle

S3 Replication rules

AWS S3 replication rules are used to automatically copy objects across different S3 buckets, which can be in the same or different AWS regions. Replication rules are useful for creating backups, disaster recovery, data compliance, or reducing data access latency in various geographical locations.

S3 Replication rules - example

Real-Time Example:

Imagine you run a global e-commerce site where users upload images, and you want those images to be available quickly to customers in different countries. Additionally, you need to keep copies of some data in certain locations for compliance with local laws.

1- **Copy for Faster Access in Different Regions**: If your main storage is in the U.S., you can set up a rule to automatically copy images to a European storage location. This makes the images load faster for customers in Europe and gives you a backup in case something happens to the U.S. data.

2- **Backup Important Data**: For important records, like financial logs, you can set up a rule to create a backup in the same region but in a different storage location. This way, even if something goes wrong, you have a safe copy of each version.

Choosing a storage class – case 1

A user uploads a video to your application and your application generates a thumbnail preview of the video. Your user agreement permits a user to access the video thumbnail for one year. However, your usage data indicates that most thumbnail previews are not often accessed after 60 days but must be accessed quickly.

Which storage class might you recommend? *Choose the storage class icon.*

That's correct! Amazon S3 – Infrequent Access is the best choice for this situation. Using Amazon S3 - IA gives users quick access to the thumbnail but saves costs. You could consider using a lifecycle policy to move the thumbnail into archive storage after 60 days.

Amazon S3 Standard

Amazon S3 Standard-IA

Amazon S3 One Zone-IA

Amazon S3 Glacier Flexible Retrieval

Amazon S3 Glacier Deep Archive

Amazon S3 Intelligent-Tiering

Choosing a storage class – case 2

A business wants to implement stronger business continuity practices and keep backups of their data for a long period of time. They do not anticipate needing to access the data often. But they want to be able to have the data within a few hours in the event of disaster recovery.

Which storage class might you recommend? *Choose the storage class icon.*

That's correct! Amazon S3 Glacier Flexible Retrieval is a low-cost option for long-term storage. Data can be retrieved in a number of hours with standard retrieval or 1-5 minutes with expedited retrieval.

Amazon S3 Standard

Amazon S3 Standard-IA

Amazon S3 One Zone-IA

Amazon S3 Glacier Flexible Retrieval

Amazon S3 Glacier Deep Archive

Amazon S3 Intelligent-Tiering

Choosing a storage class – case 3

A company builds a cloud-native application with Amazon S3 storage for school exams. The application gets some use during the school term but then is used very heavily during the exam season near the end of the term. The application has unpredictable use patterns, and the company is trying to control costs.

Which storage class might you recommend? *Choose the storage class icon.*

That's correct! Amazon S3 Intelligent-Tiering should be used when workloads are unpredictable. Amazon S3 Intelligent-Tiering will optimize storage costs by automatically moving data to the most cost-effective access tier when access patterns change.

Amazon S3 Standard

Amazon S3 Standard-IA

Amazon S3 One Zone-IA

Amazon S3 Glacier Flexible Retrieval

Amazon S3 Glacier Deep Archive

Amazon S3 Intelligent-Tiering