Overview

Amazon Web Service's (AWS) Simple Storage Service (S3) is the cloud service provider's core storage offering. S3 offers scalable, low-latency storage at a minimal cost that is easily integrated with other AWS services to support numerous uses for tiered storage and delivery, including the cold storage service Glacier and the content delivery network web service CloudFront. S3 uses an object storage architecture, managing replicates of files and related metadata within regional data centers. AWS touts S3 as a backup and disaster recovery service, but its suitability for this purpose depends on the size and nature of your collection. **S3 is best suited for storage in cases where rapid access is key and data objects are relatively small.**

Simple Storage Service (S3)

SERVICE PROVIDER: Amazon Web Services

WEBSITE: aws.amazon.com/s3/

PRODUCT RELEASED: 2006

COMPLIANCE: ISO 9001, ISO 27001, FEDRamp

SERVICE: Production, streaming

INFRASTRUCTURE: Wholly owned

COST: Low

Data Management

Data is stored in a "bucket," a container for managing groups of objects. Users can create up to 100 buckets and assign them to specific storage regions within the AWS infrastructure. Buckets are configured to perform storage functions including: **lifecycle management, allowing objects to be transferred between data center locations or storage environments (to/from Glacier) at defined intervals**; versioning control; access management for some or all objects within; and logging of access and processing events. All of these functions can be implemented and managed via the web-based AWS Management Console or programmatically using S3's REST and SOAP APIs. AWS claims to use a combination of MD5 checksums and cyclic redundancy checks to monitor data integrity.

Reporting / Metadata

S3 buckets can be configured to **log all access retrieval requests, recording the ID of the requester, bucket name, event time, and action**. Logs are uploaded to the monitored bucket periodically; AWS does not specify the delivery schedule in system documentation. Notifications can be configured to alert users of object creation/upload events via push notifications to mobile devices, email, or through integration with other messaging services using AWS APIs. S3 stores basic metadata with objects such as create date, MD5 value, and object size. User-submitted metadata, up to 2KB per object, may be stored with objects in key-value pairs.

Redundancy

Data is replicated across multiple servers in geographically dispersed data centers within a single region. **Amazon claims that data will persist through the loss of two data centers.** To achieve geographical dispersal with differing disaster threats, buckets must be replicated to a second region, doubling the cost of the service and accruing additional charges for cross-region transfers. Users may also configure hierarchical storage rules to replicate or move data to Glacier for long-term high-latency storage.

Accessibility

Charges for retrieval of data from S3 buckets decrease per GB. If retrieving large amounts of data at once, AWS offers an Import/Export service for shipment of data on removable hard drives between Amazon and the customer. S3 may be integrated with CloudFront, AWS's content delivery network, to support higher delivery speeds to global destinations at a slightly higher price. AWS's CloudWatch allows clients to track use-metrics, establish alerts, and automate actions within S3 and other AWS services based on use-statistics and user-defined benchmarks.

Security

S3 provides **multiple mechanisms for managing access and security to data** using AWS APIs. Identity and Access Management is a web service for the creation and management of users and granular use permissions within an individual AWS account. Bucket policies set restrictions for users on specific buckets within an S3 instance, including read-write and IP address restrictions. The Access Control List option allows users to set read-write restrictions to S3 buckets for other AWS accounts. S3 supports server-side (AWS-managed) encryption using SSL and client-side (user-managed) encryption to protect data in transit and at rest.

End of Service

AWS provides a 30-day window for retrieval of data following termination of services whether terminated by the user or AWS. Assistance for data retrieval is available to users except in cases involving a user's breach of contract. AWS does not provide pricing for these retrieval services. **Users should monitor for additional costs as a result of retrieving data from AWS storage when terminating an account.** All outstanding charges must be paid prior to termination of service, including fees for retrieval.

Levels of Preservation

	Level 1 (Protect)	Level 2 (Know)	Level 3 (Monitor)	Level 4 (Repair)
Storage	• **Two complete copies that are not collocated**	• **At least three complete copies** • **At least one copy in a different geographic location**	• **At least one copy in a geographic location with a different disaster threat** • **Obsolescence monitoring for storage system(s)**	• At least three copies in geographic locations with different disaster threats
Data Integrity	• **Check file fixity on ingest if it has been provided with content** • **Create fixity info if it wasn't provided with the content**	• **Check fixity on all ingests** • Virus-check high risk content	• **Check fixity of content at fixed intervals** • **Maintain logs of fixity info; supply audit on demand** • **Ability to detect corrupt data** • Virus-check all content	• **Check fixity of all content in response to specific events or activities** • **Ability to replace/ repair corrupted data** • **Ensure no one person has write access to all copies**
Security	• **Identify who has read, write, move, and delete authorization to individual files** • **Restrict who has those authorizations to individual files**	• **Document access restrictions for contents**	• **Maintain logs of who performed what actions on files, including deletions and preservation actions**	• **Perform audit of logs**

Have more questions? We can help.
Contact AVPreserve at info@avpreserve.com