Overview

Amazon Web Services' Glacier provides long-term, high-latency storage, positioned as affordable deep storage for large or growing collections and a replacement for tape libraries and their associated infrastructure. **Glacier's pay-as-you-go pricing model allows users to scale from gigabytes to petabytes.** Amazon claims it has eliminated the risk involved in regular hardware migration and will last for years, even decades. The underlying architecture of the services is a mystery, as Amazon has not released the details behind the infrastructure supporting the service. The company is less transparent than others on the market, providing little detail on their service's operations and hardware.

Glacier

SERVICE PROVIDER: Amazon Web Servicest

WEBSITE: aws.amazon.com/glacier

PRODUCT RELEASED: 2012

COMPLIANCE: ISO/IEC 27001:2013; FedRAMP(SM);

HIPAA; MPAA best practices

SERVICE: Deep storage

INFRASTRUCTURE: Wholly owned

COST: Low

Data Management

Glacier stores content in two levels: the archive and the vault. "Archives" are one or more digital objects grouped in a package similar to a zip or tar file, capped at a maximum size of 40 terabytes, and collected in "vaults." **Amazon does not currently supply an interface for uploading data to Glacier.** Submission of data to the service requires establishment of a private dedicated network connection (AWS Direct Connect) or delivery by portable storage drives (AWS Import/Export). Amazon claims the service performs regular data integrity checks on all objects in storage, but **the fixity checking method and outcomes of these checks are not available to clients.**

Reporting / Metadata

Glacier does not support management of client-side metadata within the AWS Management Console. **Daily-generated inventories document the archives within a vault, including the total size of the archives, the creation date, and any client-supplied description of the archive.** Inventories are available as either JSON or CSV files. No additional reports are provided.

Redundancy

All data stored in Glacier is redundant across multiple Amazon data centers and on multiple storage devices within each. **Customers may specify the locations where their data is stored to confirm geographical separation.** Amazon Web Services offers lifecycle management services to establish rules and processes to automatically transfer data from Amazon S3 storage to Glacier for long-term archiving, to ensure older or low-usage data is stored at the most cost-effective storage level.

Accessibility

Glacier's low cost ($.01 per gigabyte per month for storage) is predicated on low amounts of access to data once ingested into the system. Therefore, performance speeds for delivery of data from Glacier are slow. **Glacier estimates archives retrieved will be available in approximately 24 hours.** Glacier's cost model for retrieval is complex (a pricing calculator is available for estimating costs), and retrieving large amounts in a short time window or initiating many retrieval jobs can drive up the monthly cost of the service. Speed and regularity can be adjusted to keep retrieval below the allotted service levels and avoid extra charges.

Security

Glacier supports access management at the vault level. Users may define policies specifying what user accounts may view, add to, and arrange archives within a vault. Data within Glacier is encrypted upon ingest, using the 256-bit Advanced Encryption Standard. Amazon, not the client, manages the encryption keys.

End of Service

Amazon's customer agreement provides them the right to suspend or terminate service for delinquent payments, security risks, and/or end of customer operations. Amazon requires 30 days notice prior to end of service by the customer and provides "data retrieval assistance." Given the major cost implications associated with retrieval and data transfer out of Glacier it is extremely advantageous to plan far in advance in order to dramatically reduce costs in the instance of migration out of Glacier.

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 authorization 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