

Move your archive to the cloud

Let's explore AWS Glacier



David Evans
Sr Cloud Architect
PCM

Introduction

- Cloud migrations
- Education and training
- Automation
- Developing cloud agnostic tools and processes



What is AWS Glacier?

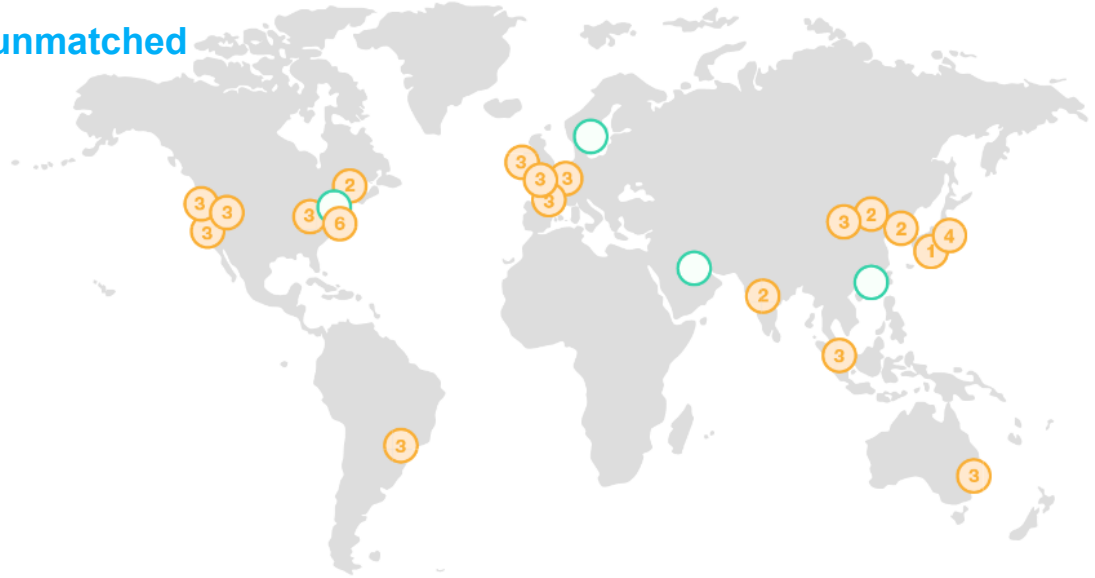
AWS Glacier is Amazon's long term storage option.

- Scalable
- Secure
- Designed for 99.999999999% durability
- Provides compliance capabilities
- Low cost, as low as \$0.004 per gigabyte per month

Glacier Scalability

AWS Glacier's scalability is unmatched

- 18 geographic regions
- 55 Availability Zones



Glacier Security

Glacier protects your data using AWS's comprehensive security.

- Data is encrypted by default
- AES 256 symmetric keys
- Granular user access controls using IAM
- MFA control for physical access
- Individuals with physical access do not have data access

Glacier Durability

Data durability and reliability

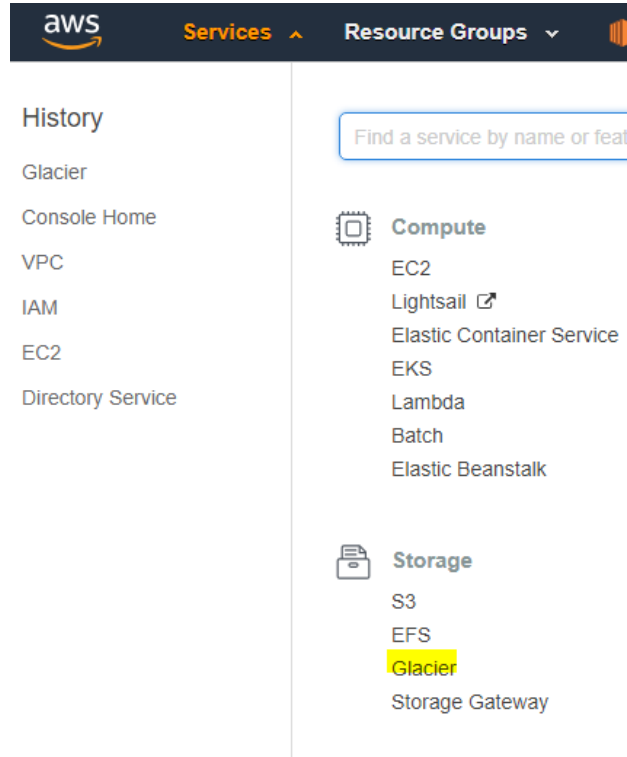
- Average annual durability of 99.999999999%
- Stored in multiple Availability Zones on multiple devices within each AZ
- Checksum is uploaded during data upload to prevent corruption
- Validates data against checksum during retrieval as well

Compliance

Glacier supports many compliance certifications.

- SEC Rule 17a-4
- PCI-DSS
- HIPAA/HITECH
- FedRAMP
- EU GDPR
- FISMA

How to setup Glacier



The screenshot shows the AWS Management Console interface. At the top, there is a dark navigation bar with the AWS logo, the word "Services" with an upward arrow, and "Resource Groups" with a downward arrow. Below this, a search bar contains the text "Find a service by name or feat". The main content area is divided into two columns. The left column is a sidebar menu with the following items: History, Glacier, Console Home, VPC, IAM, EC2, and Directory Service. The right column displays a list of services under two categories: "Compute" and "Storage". Under "Compute", the services listed are EC2, Lightsail (with an external link icon), Elastic Container Service, EKS, Lambda, Batch, and Elastic Beanstalk. Under "Storage", the services listed are S3, EFS, Glacier (highlighted in yellow), and Storage Gateway.



Amazon Glacier

Amazon Glacier is an extremely low-cost storage service that provides secure, durable, and flexible storage for data backup and archival.

[Create Vault](#)

[Getting started guide](#)



Create Vaults

A vault is a container for storing archives. An archive is any object, such as a photo, video, or document, which you store in a vault.

[Learn more](#)



Set data retrieval policies

Set data retrieval limits such as "Free Tier Only" or "Max Retrieval Rate" to manage retrieval costs with a few clicks in the AWS console.

[Learn more](#)



Set event notifications

You can configure vaults to send notifications to you or your application when retrieving data. Notifications are delivered by using the Amazon Simple Notification Service (Amazon SNS).

[Learn more](#)

Create Vault

Step 1: Vault Name

Step 2: Event Notifications

Step 3: Event Notification Details

Step 4: Review

Welcome to Amazon Glacier

Data is stored in Amazon Glacier in "archives." An archive can be any data such as a photo, video, or document. You can upload a single file as an archive or aggregate multiple files into a TAR or ZIP file and upload as one archive.

A single archive can be as large as 40 terabytes. You can store an unlimited number of archives and an unlimited amount of data in Amazon Glacier. Each archive is assigned a unique archive ID at the time of creation, and the content of the archive is immutable, meaning that after an archive is created it cannot be updated.

Vaults allow you to organize your archives and set access policies and notification policies. Get started by giving your vault a name. You can then create your vault now or click **Next Step** to set up your vault's properties.

Region

US East (N. Virginia)



Vault Name*

DavesVault

Cancel

Next Step

Create Vault

Step 1: Vault Name

Step 2: Event Notifications

Step 3: Event Notification Details

Step 4: Review

Set Event Notifications

You can choose to have notifications sent to you or your application whenever certain Amazon Glacier jobs complete. Notification messages are sent by the Amazon Simple Notifications Service (SNS).

You specify the Amazon SNS topic to use for job completion notifications by using the Amazon Resource Name (ARN) of the topic. You then select which types of Amazon Glacier jobs can trigger the sending of a notification upon completion of the job. Applications or users that subscribe to the Amazon SNS topic receive a notification message when a job of the type you select completes.

Do not enable notifications

You can enable, set up, and change your notification settings later.

Enable notifications and create a new SNS topic

Enable notifications and create a new Amazon SNS topic to send the notifications.

Enable notifications and use an existing SNS topic

Enable notifications and enter an existing SNS topic to send the notifications.

Cancel

Previous

Next Step

Create Vault

Step 1: Vault Name

Step 2: Event Notifications

Step 3: Event Notification Details

Step 4: Review

Review

Make sure the following information is correct before you choose **Submit**. To go back and make changes, choose **Previous**.

Region	US East (N. Virginia)
Vault Name	DavesVault

Cancel

Previous

Submit

Amazon Glacier Vaults

[Create Vault](#) [Delete Vault](#) [Settings](#) ↻ ?


Filter By Name:

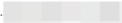
Name ▲	Inventory Last Updated	Size (as of last inventory)	# of Archives (as of last inventory)
DavesVault	Not updated yet	--	--

Vault Name: DavesVault ✕

[Details](#) [Notifications](#) [Permissions](#) [Vault Lock](#) [Tags](#)

Region: ? US East (N. Virginia)

Created on: ? 

Vault ARN: ? arn:aws:glacier:us-east-1: vaults/DavesVault

Inventory Last Updated: ? Not updated yet

Vault Details as of the last inventory update:

Size: ? --

of Archives: ? --

Moving data in

Glacier has a variety of ways to import data

Moving data in

Uploading archives

- Archives are imported programmatically
- Smaller archives can be uploaded in a single operation
- Larger archives can use a multipart upload
 - Once completed, AWS concatenates all of the parts
- Multipart maximum size – 10,000 x 4 GB
- Can be uploaded using Java, .NET, REST API, 3rd party tools

Moving data in

S3 lifecycle

Moving data in

S3 lifecycle




Amazon S3

+ Create bucket

Delete bucket

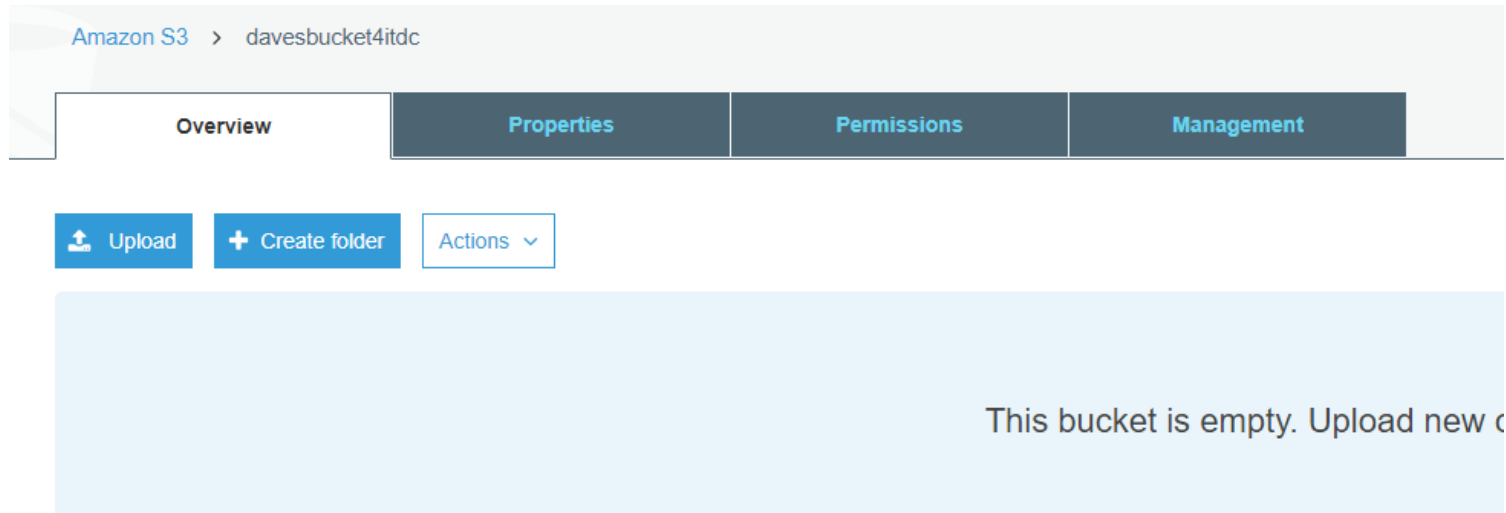
Empty bucket

Bucket name ↑☰

 davesbucket4itdc

Moving data in

S3 lifecycle



Amazon S3 > davesbucket4itdc

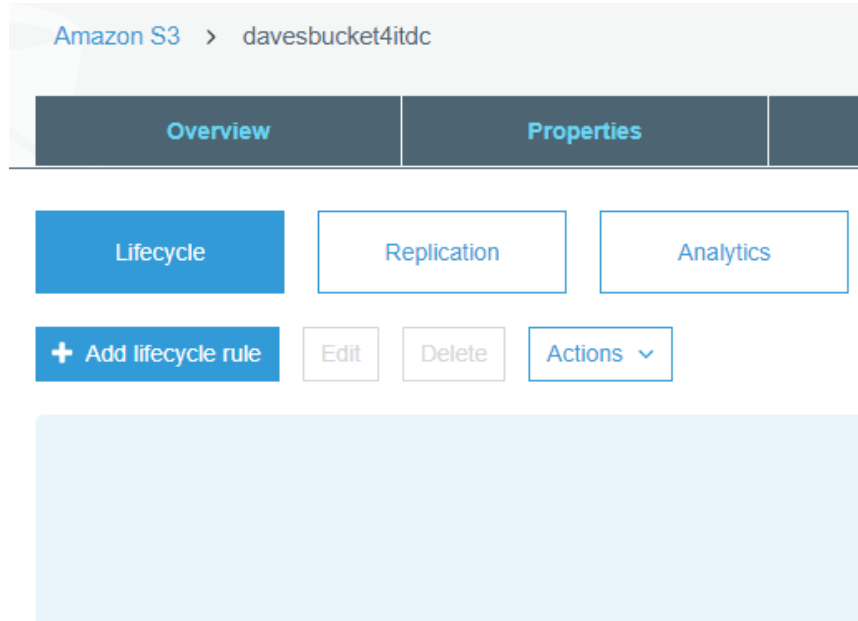
Overview Properties Permissions Management

Upload Create folder Actions

This bucket is empty. Upload new c

Moving data in

S3 lifecycle



Amazon S3 > davesbucket4itdc

Overview Properties

Lifecycle Replication Analytics

+ Add lifecycle rule Edit Delete Actions ▾

The screenshot shows the Amazon S3 console interface for a bucket named 'davesbucket4itdc'. The breadcrumb navigation indicates the path 'Amazon S3 > davesbucket4itdc'. Below the breadcrumb, there are two tabs: 'Overview' and 'Properties'. Underneath the tabs, there are three main action buttons: 'Lifecycle' (highlighted in blue), 'Replication', and 'Analytics'. Below these buttons, there is a row of four smaller buttons: '+ Add lifecycle rule' (highlighted in blue), 'Edit', 'Delete', and 'Actions ▾' (with a dropdown arrow).

Moving data in

S3 lifecycle

Lifecycle rule

1 Name and scope 2 Transitions 3 Expiration 4 Review

Enter a rule name

Add filter to limit scope to prefix/tags ⓘ

Cancel Next

Moving data in

S3 lifecycle

Lifecycle rule

1 Name and scope 2 **Transitions** 3 Expiration 4 Review

Configure transition ⓘ

Current version Previous versions

Previous Next

Moving data in

S3 lifecycle

Lifecycle rule

☑ Name and scope ☑ Transitions ③ Expiration ④ Review

Configure expiration

Current version Previous versions

Expire current version of object ⓘ

After days from object creation

Clean up expired object delete markers and incomplete multipart uploads

Clean up expired object delete markers ⓘ

You cannot enable clean up expired object delete markers if you enable Expiration.

Clean up incomplete multipart uploads ⓘ

After Days from start of upload

[Previous](#) [Next](#)

Moving data in

AWS Snowball

- 80 TB (72 TB usable)
- Fast Data Transfer
- Encrypted
- Rugged
- Tamper Resistant
- Secure Erasure



Moving data in

AWS Snowmobile

- 100 PB
- 45' shipping container
- GPS tracking
- Video surveillance
- Optional security escort



Getting data out

Retrieving an archive is an asynchronous operation.

- Initiate the retrieval job
- Download once the job has completed
 - You have 24 hours to download once a job has completed

Getting data out

AWS Glacier 3 separate retrieval plans.

- Expedited
- Standard
- Bulk

Getting data out

Amazon Glacier Select

- Archive objects must be uncompressed CSV files
- Input or output can not exceed 1MB
- SQL expression is limited to 128 KB
- Performed as an initiate job
- Results are stored as text files in S3

Demo

Questions?

David Evans

dave@daveinthecloud.com

[@letsScriptIt](#)