

Implementation Guide

Data Transfer from Amazon S3 Glacier Vaults to Amazon S3



Data Transfer from Amazon S3 Glacier Vaults to Amazon S3: Implementation Guide

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Solution overview	1
Features and benefits	3
Use cases	5
Concepts and definitions	5
Architecture overview	7
Architecture diagram	7
Translation of S3 Glacier vault archive descriptions to S3 object names	9
Creating custom file names for S3 objects	10
AWS Well-Architected design considerations	11
Operational excellence	11
Security	11
Reliability	11
Performance efficiency	12
Cost optimization	12
Sustainability	12
Architecture details	13
Third-party software support	13
AWS services in this solution	13
Plan your deployment	15
Cost	15
Cost table calculation	15
Sample cost tables	17
AWS CloudTrail cost	19
Potential cost savings	19
Security	19
Amazon DynamoDB	19
CloudWatch Logs	20
IAM roles	20
Supported AWS Regions	20
Quotas	21
Quotas for AWS services in this solution	21
AWS CloudFormation quotas	21
Lambda concurrent execution quota	22
Amazon S3 Glacier Initiate Job quota	22

Amazon S3 file size limit	22
Amazon S3 storage class considerations	22
Amazon S3 Glacier resource considerations	22
Amazon S3 Glacier Vault Lock policy considerations	23
Deploy the solution	24
Prerequisites	24
Deployment process overview	25
AWS CloudFormation template	25
Step 1: Launch the stack	26
Step 2: Launch the transfer workflow	29
(Optional) Download the vault inventory file	30
(Optional) Provide the vault inventory file	32
Step 3: Resume the transfer workflow	34
Monitor the solution with Service Catalog AppRegistry	36
Activate CloudWatch Application Insights	36
Confirm cost tags associated with the solution	38
Activate cost allocation tags associated with the solution	38
AWS Cost Explorer	39
Update the solution	40
Troubleshooting	41
Problem: Transfer workflow has not progressed after 14 hours	41
Resolution	41
Problem: Transfer workflow must be stopped	41
Resolution	41
Contact AWS Support	42
Create case	42
How can we help?	42
Additional information	42
Help us resolve your case faster	43
Solve now or contact us	43
Uninstall the solution	44
Using the AWS Management Console	44
Using AWS Command Line Interface	44
Deleting the S3 buckets	44
Deleting the DynamoDB tables	45
Deleting the CloudWatch Logs	45

Use the solution	47
Validate your inventory	47
Access the CloudWatch dashboard	47
Manage your Amazon S3 storage	47
Developer guide	49
Source code	49
Reference	50
Anonymized data collection	50
Contributors	51
Revisions	52
Notices	53

Automatically copy your Amazon S3 Glacier vault archives to an S3 bucket and storage classes

Publication date: *December 2023* ([last update](#): *May 2024*)

Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 is a serverless solution that automates and optimizes the restore, copy, and transfer process of [Amazon Simple Storage Service Glacier](#) (Amazon S3 Glacier) vault archives. The solution copies all of the vault's archives to a defined [Amazon Simple Storage Service](#) (Amazon S3) bucket destination and [storage class](#). Then you can attach [tags](#) to help you categorize your data, such as with data classification or cost allocation. A prebuilt [Amazon CloudWatch](#) dashboard provides a visualization of the copy operation progress.

Important

Amazon S3 and Amazon S3 Glacier are different AWS services.

Amazon S3 Glacier is an object storage service for low-cost data archiving and long-term backup. It stores *archives* in *vaults*. It doesn't offer storage classes. The Amazon S3 Glacier service provides a console. However, any archive operation, such as upload, download, or deletion, requires you to use the AWS CLI or write code. There is no console support for archive operations.

Amazon S3 is an object storage service for any type of data. It stores *objects* in *buckets*. It offers different storage classes for frequent access, infrequent access, archives, and optimized tiering. You can interact with the Amazon S3 service by using the Amazon S3 console or [AWS Command Line Interface](#) (AWS CLI).

The *S3 Glacier Instant Retrieval*, *S3 Glacier Flexible Retrieval*, and *S3 Glacier Deep Archive storage classes* are features of the Amazon S3 service. The *S3 Glacier Flexible Retrieval* storage class offers the same features as the Amazon S3 Glacier service. The Amazon S3 Glacier service doesn't offer storage classes.

For example, Saanvi works at AnyCompany Archives. Five years ago, she used the Amazon S3 Glacier service to store scanned copies of historical documents in a vault. AnyCompany just announced that they will have a different online exhibit each month, featuring documents that are stored in the S3 Glacier vault. To address this change of business:

- Saanvi wants to take advantage of the storage classes offered with the Amazon S3 service, including more flexibility in how files are stored and accessed.

- Using Data Transfer from Amazon S3 Glacier Vaults to Amazon S3, Saanvi can copy all of her document archives from her S3 Glacier vault to an S3 bucket. She can assign them to the S3 storage classes that best fit her use cases. For example, she can use the S3 Standard storage class for documents that will be featured in the first exhibit and accessed daily, and the S3 Glacier Deep Archive storage class for documents that won't be featured in any of the exhibits.
- Now that the documents are stored in the Amazon S3 service, Saanvi can also apply [S3 Lifecycle](#) configurations, tag her data, and use the Amazon S3 console.

Note

This solution doesn't delete the original archives or the source S3 Glacier vault. You must manually delete the archives and vault. For more information, refer to [Deleting an Archive in Amazon S3 Glacier](#) in the *Amazon S3 Glacier Developer Guide*.

If your source S3 Glacier vault has a [Vault Lock policy](#) that prevents deletion, you must delete this policy before deleting the original archives. However, if your Vault Lock policy is in the Locked state, you can't delete it. See [S3 Glacier Vault Lock](#) and [Abort Vault Lock \(DELETE lock-policy\)](#) in the Amazon S3 Glacier Developer Guide for more information.

This implementation guide provides an overview of the Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 solution, its reference architecture and components, considerations for planning the deployment, and configuration steps for deploying the solution to the Amazon Web Services (AWS) Cloud.

The intended audience for using this solution's features and capabilities in their environment includes solution architects, business decision makers, DevOps engineers, data scientists, and cloud professionals. Practical experience with the AWS Cloud, Amazon S3 Glacier vaults, Amazon S3 buckets, and Amazon S3 storage classes is preferred.

Use this navigation table to quickly find answers to these questions:

If you want to . . .	Read . . .
Know the cost for running this solution.	Cost
The estimated cost for running this solution in the US East (Ohio) Region is USD \$153.57	

If you want to . . .	Read . . .
to copy 100,000 S3 Glacier vault archives, totaling 100 TB of data, from an S3 Glacier vault to an S3 bucket.	
Understand the security considerations for this solution.	Security
Know how to plan for quotas for this solution. This solution uses AWS Lambda functions to transfer data. This affects your account-wide Lambda concurrency limit.	Quotas
Know which AWS Regions support this solution.	Supported AWS Regions
View or download the AWS CloudFormation template included in this solution to automatically deploy the infrastructure resources (the "stack") for this solution.	AWS CloudFormation template

Features and benefits

The solution provides the following features:

Automation

Automate the process of restoring, copying, and transferring archives from a vault in the Amazon S3 Glacier service to a bucket in the Amazon S3 service. After you move your data, you can use the Amazon S3 console. A prebuilt Amazon CloudWatch dashboard helps you monitor metrics and visualize the copy operation progress.

Ability to assign storage classes

When you use this solution to move your data from the Amazon S3 Glacier service into the Amazon S3 service, you choose a storage class to assign to all of your objects. After your data is stored in

the Amazon S3 service, you can change the storage classes to fit the use case for each file. We recommend carefully reviewing each storage class and its pricing details before deploying this solution. See [Amazon S3 storage class considerations](#) for more information.

Visibility and access to data

After the solution stores S3 Glacier archives as objects in the destination S3 bucket, you can add tags to data. Tagging offers benefits such data classification, permissions controls, object lifecycle management, and cost allocation. For more information, see [Categorizing your storage using tags](#) in the *Amazon Simple Storage Service User Guide*.

Flexibility to cancel your transfer and resume later

The solution tracks the progress of the transfer. You can stop and restart transfers without needing to retransfer existing archives. See [Problem: Transfer workflow must be stopped](#) for more information.

Cost optimization

Copy S3 Glacier vault archives to an S3 bucket and assign more [cost-effective storage classes](#), such as:

- The low-cost S3 Glacier Deep Archive storage class for data that you rarely access
- The S3 Glacier Instant Retrieval storage class if you'll need your data quarterly but within milliseconds
- The S3 Standard storage class for data you'll need daily

You can also configure and apply [S3 Lifecycles](#) to transition your objects automatically into different storage classes, based on rules you set.

Integration with Service Catalog AppRegistry and Application Manager, a capability of AWS Systems Manager

This solution includes a [Service Catalog AppRegistry](#) resource to register the solution's CloudFormation template and its underlying resources as an application in both Service Catalog AppRegistry and [Application Manager](#). With this integration, centrally manage the solution's resources and enable application search, reporting, and management actions.

Use cases

Performance and cost optimization

Balancing cloud storage performance and storage cost is crucial for organizations. You can use this solution to help you optimize your storage infrastructure with the features that the Amazon S3 service offers. With tagging, lifecycle configurations, and a variety of storage classes, you can improve your performance while minimizing costs.

Cloud archiving

Many organizations store their most fundamental asset—their data—in locations that are slow to retrieve and lack flexibility. You can use this solution to help you automate, monitor, and seamlessly move your data when and where you need it.

Concepts and definitions

This section describes key concepts and defines terminology specific to this solution:

archive

Any data stored in an Amazon S3 Glacier vault, such as a photo, video, or document. An archive is similar to an Amazon S3 object: it's the base unit of storage in the Amazon S3 Glacier service. For more information, see [Archive](#) in the *Amazon S3 Glacier Developer Guide*.

chunk

Term used to describe a *part* in a multipart upload or download for the Amazon S3 Glacier service. This solution uses multipart upload to transfer the archives. For more information, see [Uploading Large Archives in Parts \(Multipart Upload\)](#) and [Retrieving S3 Glacier Archives Using AWS Management Console](#) in the *Amazon S3 Glacier Developer Guide*.

inventory

A point in time snapshot or listing of the archives stored within an S3 Glacier vault. For more information, see [Downloading a Vault Inventory in Amazon S3 Glacier](#) in the *Amazon S3 Glacier Developer Guide*.

tag

A key-value pair used to categorize storage in the Amazon S3 service. For more information, see [Categorizing your storage using tags](#) in the *Amazon S3 User Guide*.

vault

An container in the Amazon S3 Glacier service for storing archives. An S3 Glacier vault is similar to an S3 bucket. For more information, see [Vault](#) in the *Amazon S3 Glacier Developer Guide*.

workflow_run

An identifier used to represent the transfer of an S3 Glacier vault to an S3 bucket. The solution randomly generates the `workflow_run` value on the first run (or you can choose the value). The solution uses this value when resuming a transfer.

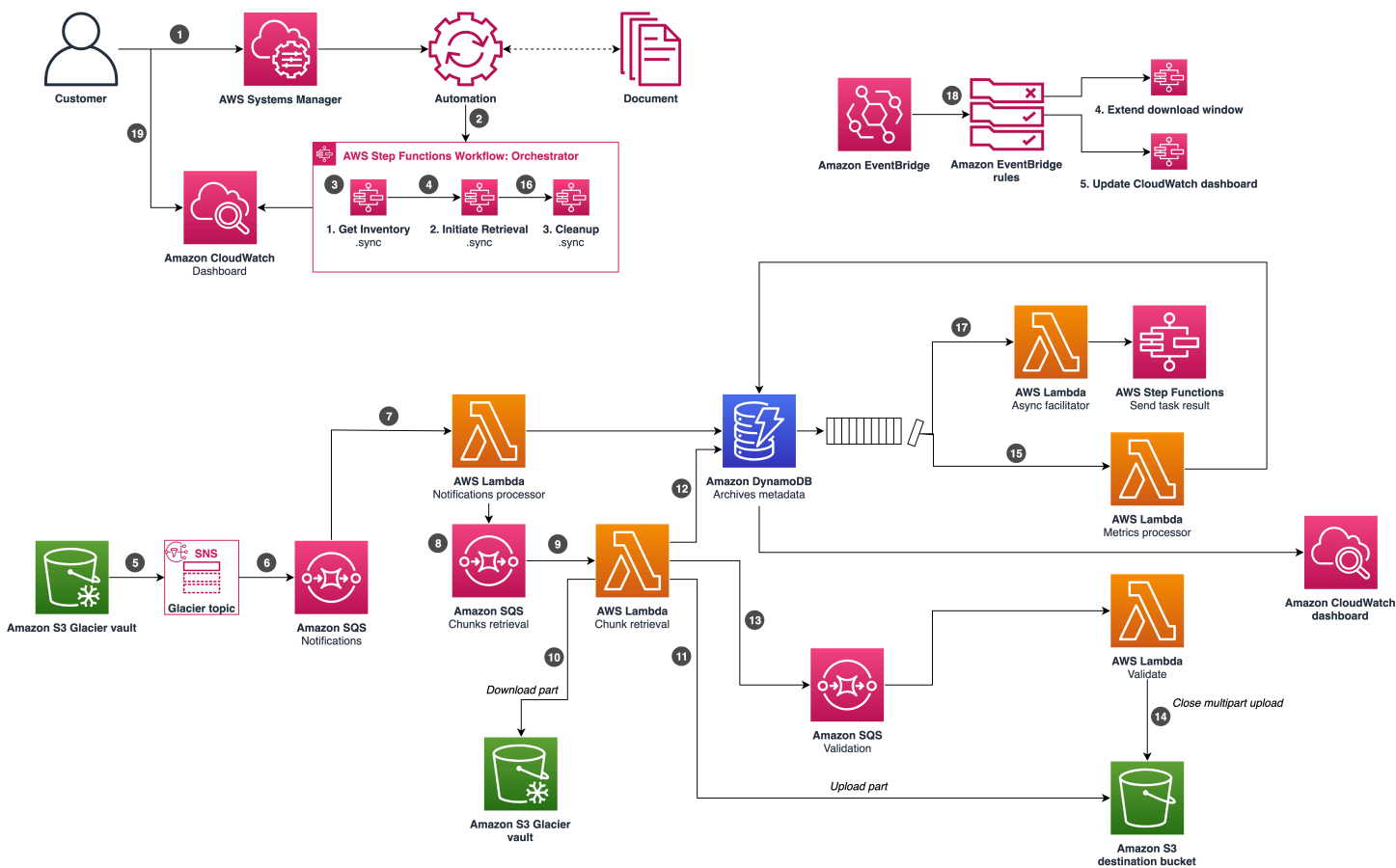
For a general reference of AWS terms, see the [AWS glossary](#) in the *AWS General Reference*.

Architecture overview

This section provides a reference implementation architecture diagram for the components deployed with this solution.

Architecture diagram

Deploying this solution with the default parameters deploys the following components in your AWS account.



Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 architecture on AWS

Note

AWS CloudFormation resources are created from [AWS Cloud Development Kit \(AWS CDK\)](#) (AWS CDK) constructs.

The high-level process flow for the solution components deployed with the [AWS CloudFormation](#) template is as follows:

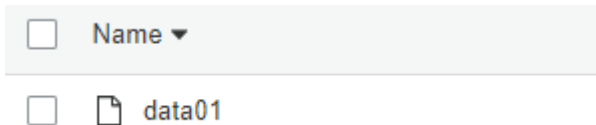
1. Customers invoke a transfer workflow by using a [Systems Manager document](#) (SSM document).
2. The SSM document starts an [AWS Step Functions](#) Orchestrator execution.
3. The Step Functions Orchestrator execution initiates a nested Step Functions Get Inventory workflow to retrieve the inventory file.
4. Upon completion of the inventory retrieval, the solution invokes the Initiate Retrieval nested Step Functions workflow.
5. When a job is ready, the [Amazon S3 Glacier](#) service sends a notification to an [Amazon Simple Notification Service](#) (Amazon SNS) topic indicating job completion.
6. The solution stores all job completion notifications in the [Amazon Simple Queue Service](#) (Amazon SQS) Notifications queue.
7. When an archive job is ready, the Amazon SQS Notifications queue invokes the [AWS Lambda](#) Notifications Processor function. This Lambda function prepares the initial steps for archive retrieval.
8. The Lambda Notifications Processor function places chunks retrieval messages in Amazon SQS Chunks Retrieval queue for chunk processing.
9. The Amazon SQS Chunks Retrieval queue invokes the Lambda Chunk Retrieval function to process each chunk.
10. The Lambda Chunk Retrieval function downloads the chunk from the Amazon S3 Glacier service.
11. The Lambda Chunk Retrieval function uploads a [multipart upload](#) part to the [Amazon S3](#) service.
12. After a new chunk is downloaded, the solution stores chunk metadata in [Amazon DynamoDB](#) (etag, checksum_sha_256, tree_checksum).
13. The Lambda Chunk Retrieval function verifies whether all chunks for that archive have been processed. If yes, it inserts an event into the Amazon SQS Validation queue to invoke the Lambda Validate function.
14. The Lambda Validate function does the following:
 - a. Performs an integrity check against the tree hash in the inventory.
 - b. Calculates a checksum and passes it to the into the close multipart upload call. If that hash is wrong, Amazon S3 rejects the request.

15. The `DynamoDB stream` invokes the `Lambda Metrics Processor` function to update the transfer process metrics in DynamoDB.
16. The `Step Functions Orchestrator` execution enters an `async wait`, pausing until the archive retrieval workflow concludes before initiating the `Step Functions Cleanup` workflow.
17. The `DynamoDB stream` invokes the `Lambda Async Facilitator` function, which unlocks asynchronous waits in Step Functions.
18. The `Amazon EventBridge` rules periodically initiate `Step Functions Extend Download Window` and `Update CloudWatch Dashboard` workflows.
19. Customers monitor the transfer progress by using the Amazon CloudWatch dashboard.

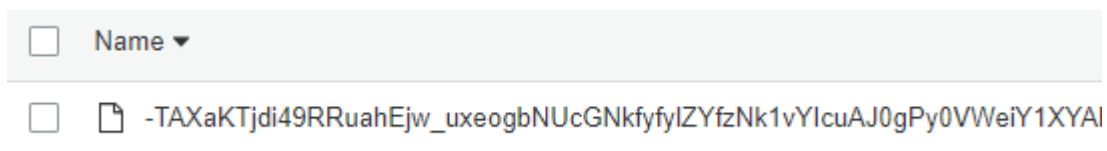
Translation of S3 Glacier vault archive descriptions to S3 object names

To create the key name for each of the new objects in the Amazon S3 service, this solution uses the `ArchiveDescription` value for each `ArchiveId` listed in the Amazon S3 Glacier inventory file. The following are examples.


- If the `ArchiveDescription` is a single string value, such as `data01`, the solution translates that value to an S3 object key name in the destination S3 bucket.





- If the `ArchiveDescription` value is blank, then the solution does the following:
 - Copies the archive.
 - Uses the `ArchiveId` as the S3 object key name.
 - Adds the prefix `00undefined` to the S3 object key names and stores the objects in the destination S3 bucket.



- If multiple `ArchiveId` entries have the same value for the `ArchiveDescription` field (for example, `duplicatefile02.txt`), then the solution appends a timestamp suffix to the name of the original file. This resolves the potential issue of having duplicate S3 object key names copied over one another. The timestamp used is the `CreationDate` of the archive.

-  duplicatefile02.txt

-  duplicatefile02.txt-2019-10-17T00:51:33Z

-  duplicatefile02.txt-2019-10-17T00:51:34Z

Creating custom file names for S3 objects

You can provide custom S3 object key names for each **ArchiveID** that's copied to your S3 bucket. To do this, provide a **NamingOverrideFile** to the solution when you [launch the transfer workflow](#), using the **NamingOverrideFile** input parameter. Use the following process.

1. Create a data file in CSV format. The file must contain only two columns: **GlacierArchiveID** and **FileName** (separated by a comma). The following table is an example.

GlacierArchiveID	FileName
WVf1rXME2KC6JIIedfadJF937412-e	Mydata.txt
yLam5H76JXYSKKIY34404D-Kwcrk	Myfolder/mydata2.txt

2. Obtain a copy of your vault inventory file for the Amazon S3 Glacier service. For more information, see [Downloading a Vault Inventory in Amazon S3 Glacier](#) in the *Amazon S3 Glacier Developer Guide*.
3. Copy all the **ArchiveID** values from your S3 Glacier vault inventory file. Paste them into the **GlacierArchiveID** column of your **NamingOverride** CSV file.
4. In the **FileName** column, for each **ArchiveID**, enter your desired S3 object key name.

Note

If you provide an empty value for the **FileName**, the solution uses the original value for **ArchiveDescription** from the S3 Glacier archive.

5. Upload the CSV file to any S3 bucket and [create a presigned URL for the file](#).
6. Use this presigned URL as the value of the **NamingOverride File** input parameter used when [launching the transfer workflow](#).

AWS Well-Architected design considerations

This solution uses the best practices from the [AWS Well-Architected Framework](#), which helps customers design and operate reliable, secure, efficient, and cost-effective workloads in the cloud.

This section describes how the design principles and best practices of the Well-Architected Framework benefit this solution.

Operational excellence

This section describes how we architected this solution using the principles and best practices of the [operational excellence pillar](#).

- This solution pushes metrics to CloudWatch at various stages to provide visibility into archive transfer progress.

Security

This section describes how we architected this solution using the principles and best practices of the [security pillar](#).

- All interservice communications use applicable [AWS Identity and Access Management \(IAM\)](#) roles.
- All roles used by the solution follow least privilege access. They only contain the minimum permissions required to accomplish the transfer.
- All data storage, including the S3 buckets, encrypts the data at rest.

Reliability

This section describes how we architected this solution using the principles and best practices of the [reliability pillar](#).

- The solution uses a serverless architecture to achieve high availability and recovery from failure.
- The solution protects against state machine definition errors through a suite of automated tests.
- Data processing uses Lambda functions. Data is stored in DynamoDB and Amazon S3, which persist in multiple Availability Zones by default.

Performance efficiency

This section describes how we architected this solution using the principles and best practices of the [performance efficiency pillar](#).

- The solution uses a serverless architecture with the ability to scale horizontally as needed.
- The solution is tested and deployed daily to achieve consistency as AWS services change.

Cost optimization

This section describes how we architected this solution using the principles and best practices of the [cost optimization pillar](#).

- The solution uses a serverless architecture that only charges customers for what they use.
- DynamoDB global secondary indexes are selected to reduce the pricing for queries.

Sustainability

This section describes how we architected this solution using the principles and best practices of the [sustainability pillar](#).

- The solution uses managed serverless services to minimize the environmental impact of the backend services compared to continually operating on-premises services.

Architecture details

This section describes the components and AWS services that make up this solution and the architecture details on how these components work together.

Third-party software support

This solution uses the value stored in the **ArchiveDescription** for each S3 Glacier vault archive (as listed in the S3 Glacier inventory file) as the key name for the new S3 object that it creates. The solution supports copying S3 Glacier vaults using either FastGlacier or CloudBerry software as follows.

- **FastGlacier (v1-v4)** – The solution extracts the value for `/ArchiveMetadata/Path` of `/m/p` from the XML metadata stored in the **ArchiveDescription** field in the S3 Glacier inventory file. It then converts that value to a string that forms the S3 object key name.
- **CloudBerry (v5.9)** – The solution extracts the value for **Path** from the JSON metadata stored in the **ArchiveDescription** field in the S3 Glacier inventory file. It then converts that value to a string that forms the S3 object key name.

AWS services in this solution

AWS service	Description
Amazon S3	Core. Stores inventory and archives transferred from the Amazon S3 Glacier service.
Amazon S3 Glacier	Core. Maintains the vault being transferred to the Amazon S3 service.
Amazon DynamoDB	Core. Stores workflow-related data, including hash validation values and progress metrics for S3 Glacier archives.
AWS Glue	Core. Reorders and splits the S3 Glacier vault inventory file into partitions for processing

AWS service	Description
	by multiple Lambda invocations. Parses file names from Inventory/OverrideNaming file.
AWS Lambda	Core. Performs workflow compute, including archive transfer and validation.
Amazon SNS	Supporting. Decouples Lambda workflows by providing communication between the solution, the Amazon S3 Glacier service, and Amazon SQS queues.
Amazon SQS	Supporting. Decouples Lambda workflows.
Amazon CloudWatch	Supporting. Stores the solution logs and metrics. Presents a custom dashboard to provide visibility of the archive copy operation progress.
Amazon EventBridge	Supporting. Invokes supporting workflows that promote the smooth operation of the Solution.
IAM	Supporting. Provides permissions for this solution's resources to perform actions.
AWS Systems Manager	Supporting. Provides a Systems Manager document (SSM document) for solutions workflow invocations.
AWS Step Functions	Supporting. Orchestrates partitioning the inventory with AWS Glue. Also orchestrates Lambda invocations to request S3 Glacier vault archive retrieval.

Plan your deployment

This section describes the [cost](#), [security](#), [Regions](#), and other considerations before deploying the solution.

Cost

You are responsible for the cost of the AWS services used while running this solution. As of this revision, the cost for running this solution with the default settings in the US East (Ohio) Region is approximately **\$153.57** for 100,000 S3 Glacier vault archives (1GB each) and **\$1,229.21** for 10,000,000 S3 Glacier vault archives (10MB each). These costs assume that the destination bucket is also in US East (Ohio) Region. Refer to Sample cost tables for more details.

Note

If the destination bucket is not in the same region as the Glacier vault, a "Data Transfer OUT From Amazon S3 Glacier" fee will be added. See [Data transfer pricing](#) for more information. This cost should be considered when planning your data storage and transfer strategies to avoid unexpected charges.

See the pricing webpage for each AWS service used in this solution. Estimated costs vary based on the number of archives processed and the total volume of data to copy from an S3 Glacier vault.

We recommend creating a [budget](#) through AWS Cost Explorer to help manage costs. Prices are subject to change. For full details, see the pricing webpage for each [AWS service used in this solution](#).

Cost table calculation

The following table shows how the sample cost tables were calculated.

Solution component	Type	[A] - Unit cost	[B] - Value	[C] - Estimated cost
Additional services	Per 1,000 requests	\$0.01	<# of S3 Glacier	$[A] \times [B] / 1,000$

Solution component	Type	[A] - Unit cost	[B] - Value	[C] - Estimated cost
			<i>vault archives></i>	
Amazon S3 multipart upload create requests	Per 1,000 requests	\$0.005	<i><# of S3 Glacier vault archives></i>	$[A] \times [B] / 1,000$
Amazon S3 multipart upload complete requests	Per 1,000 requests	\$0.03	<i><# of S3 Glacier vault archives></i>	$[A] \times [B] / 1,000$
Solution runtime (Lambda and Step Functions)	Per GB	\$0.00143	<i><Size of S3 Glacier vault in GBs></i>	$[A] \times [B]$
Solution runtime (Lambda and Step Functions)	Per 1,000 requests	\$0.0447	<i><# of S3 Glacier vault archives></i>	$[A] \times [B] / 1,000$
DynamoDB writes/reads for transfer metadata	Per 1,000 requests	\$0.02	<i><# of S3 Glacier vault archives></i>	$[A] \times [B] / 1,000$

Solution component	Type	[A] - Unit cost	[B] - Value	[C] - Estimated cost
Data Transfer OUT from S3 Glacier	Per GB	\$0.02	<i><Size of S3 Glacier vault in GBs></i>	$[A] \times [B]$
		<i>if destination bucket is in a different region than S3 Glacier vault.</i> \$0.00 <i>if destination bucket is in the same region as S3 Glacier vault.</i>		

Sample cost tables

The following tables provide two sample cost breakdowns for deploying this solution with the default parameters in the US East (Ohio) Region, with an S3 Glacier vault size of 100 TB. These cost breakdowns are based on the destination bucket is also being in the US East (Ohio) Region, the same region as the S3 Glacier vault.

Note

Costs associated with storing data in the Amazon S3 service are nearly continuous and aren't included in these estimates.

Scenario 1: 100,000 S3 Glacier vault archives

AWS service	Dimensions	Cost [USD]
Step Functions		\$0.07

AWS service	Dimensions	Cost [USD]
Lambda		\$140.00
DynamoDB		\$2.00
Amazon S3	Transfer cost	\$5.00
Additional services:		\$6.50
<ul style="list-style-type: none"> • Amazon SQS • Amazon SNS • AWS Glue • CloudWatch 		
	Total:	\$153.57 [USD]

If the destination bucket is in a different region than US East (Ohio), an additional price of \$2048 (\$0.02 x size of S3 Glacier vault in GBs) should be added to the total.

Scenario 2: 10,000,000 S3 Glacier vault archives

AWS service	Dimensions	Cost [USD]
Step Functions		\$3.21
Lambda		\$411.00
DynamoDB		\$221.00
Amazon S3	Transfer cost	\$465.00
Additional services:		\$129.00
<ul style="list-style-type: none"> • Amazon SQS • Amazon SNS • AWS Glue • CloudWatch 		

AWS service	Dimensions	Cost [USD]
	Total:	\$1,229.21 [USD]

If the destination bucket is in a different region than US East (Ohio), an additional price of \$2048 (\$0.02 x size of S3 Glacier vault in GBs) should be added to the total.

AWS CloudTrail cost

You can use [AWS CloudTrail](#) to log, continuously monitor, and retain account activity related to actions across your AWS infrastructure. When you create additional trails—for example, to capture data or insight on generated events—AWS CloudTrail charges apply. See [AWS CloudTrail pricing](#) for more information.

Potential cost savings

As of this revision, you can save \$0.00261 per GB per month by storing your archives in the Amazon S3 service with the S3 Glacier Deep Archive storage class applied. For example, if you have 100 TB stored in your S3 Glacier vault, you can save \$261.00 per month by storing that data in the Amazon S3 service with the S3 Glacier Deep Archive storage class applied.

The cost to run the solution scales with the size of the S3 Glacier vault and the number of archives. Cost savings only scale with the S3 Glacier vault size.

Security

When you build systems on AWS infrastructure, security responsibilities are shared between you and AWS. This [shared responsibility model](#) reduces your operational burden because AWS operates, manages, and controls the components including the host operating system, the virtualization layer, and the physical security of the facilities in which the services operate. For more information about AWS security, visit [AWS Cloud Security](#).

Amazon DynamoDB

All user data stored in DynamoDB is encrypted at rest using encryption keys stored in [AWS Key Management Service](#) (AWS KMS). We recommend enforcing [AWS managed keys](#) because you have permission to [audit their use](#) in AWS CloudTrail logs. Refer to [Managing encrypted tables in DynamoDB](#) for more information.

Consider enabling DynamoDB Data Plane Events for CloudTrail logging to gain insights into the data operations in DynamoDB tables, according to your use cases and your regulatory and compliance requirements. Refer to [Logging DynamoDB operations by using AWS CloudTrail](#) for more information. Additionally, consider implementing [AWS Config](#) to actively monitor DynamoDB configuration changes

CloudWatch Logs

We recommend [changing the retention period](#) of your [CloudWatch Logs](#) according to your use cases and your regulatory and compliance requirements.

IAM roles

IAM roles allow you to assign granular access policies and permissions to services and users on the AWS Cloud. This solution creates IAM roles that grant the solution's resources permission to access the S3 Glacier vault, write logs, and create EventBridge targets.

Supported AWS Regions

This solution uses AWS services that are not currently available in all AWS Regions. For the most current availability of AWS services by Region, see the [AWS Regional Services List](#).

Note

This solution launches in the US East (Ohio) Region by default. To launch the solution in a different AWS Region, use the Region selector in the console navigation bar. See [Step 1: Launch the stack](#) for more information.

This solution is NOT supported in AWS GovCloud (US) Regions.

Region name	
US East (N. Virginia)	Asia Pacific (Tokyo)
US East (Ohio)	Canada (Central)
US West (N. California)	Europe (Frankfurt)

Region name	
US West (Oregon)	Europe (Ireland)
Africa (Cape Town)	Europe (London)
Asia Pacific (Hong Kong)	Europe (Milan)
Asia Pacific (Mumbai)	Europe (Paris)
Asia Pacific (Osaka)	Europe (Stockholm)
Asia Pacific (Seoul)	Middle East (Bahrain)
Asia Pacific (Singapore)	South America (São Paulo)
Asia Pacific (Sydney)	

Quotas

Service quotas, also referred to as limits, are the maximum number of service resources or operations for your AWS account.

Quotas for AWS services in this solution

Make sure you have a sufficient quota for each of the [services implemented in this solution](#). For more information, see [AWS service quotas](#).

Use the following links to go to the page for that service. To view the service quotas for all AWS services in the documentation without switching pages, view the information in the [Service endpoints and quotas](#) page in the PDF instead.

AWS CloudFormation quotas

Your AWS account has AWS CloudFormation quotas that you should be aware of when [launching the stack](#) in this solution. By understanding these quotas, you can avoid limitation errors that would prevent you from deploying this solution successfully. For more information, see [AWS CloudFormation quotas](#) in the AWS CloudFormation User Guide.

Lambda concurrent execution quota

Your AWS account has a quota on the number of concurrent Lambda executions that can be running. For more information, see [Lambda quotas](#) in the *AWS Lambda Developer Guide*. This solution uses 230–250 concurrent Lambda executions when running at maximum capacity.

Amazon S3 Glacier Initiate Job quota

This solution optimizes your transfer by requesting archives in order. Other random restore requests can impact throughput.

The Amazon S3 Glacier service maintains a service quota of [35 random restore requests](#) per PiB stored per day. If you continue to initiate your archive retrievals as the solution runs, Amazon S3 Glacier responses might slow down. You might also see Amazon S3 Glacier [ThrottlingExceptions](#) if you initiate archive retrievals external to the solution.

Amazon S3 file size limit

The Amazon S3 service restricts file sizes to 5 TB. The solution won't transfer archives larger than 5 TB. The solution's CloudWatch dashboard indicates the number of archives that meet this condition. The solution stores inventory data for these archives in the Inventory S3 bucket under `$WORKFLOW_RUN/not_migrated/`.

Amazon S3 storage class considerations

When you deploy this solution, you must choose a storage class to apply to all of your transferred data. Before you choose this storage class, consider the availability, durability, minimum storage duration, and cost of each storage class. After your data is stored in the Amazon S3 service, you can change the storage class for each object. Some storage classes have minimum durations, so it's important to plan accordingly. For more information, see [Comparing the Amazon S3 storage classes](#) in the *Amazon Simple Storage Service User Guide* and [Amazon S3 pricing](#).

Amazon S3 Glacier resource considerations

This solution uses the entirety of your Amazon S3 Glacier service resources. You won't be able to use your Amazon S3 Glacier archive while the solution is running.

Amazon S3 Glacier Vault Lock policy considerations

This solution doesn't delete the original archives or the source S3 Glacier vault. You must manually delete the archives and vault. For more information, refer to [Deleting an Archive in Amazon S3 Glacier](#) in the *Amazon S3 Glacier Developer Guide*.

If your source S3 Glacier vault has a [Vault Lock policy](#) that prevents deletion, you must delete this policy before deleting the original archives. However, if your Vault Lock policy is in the Locked state, you can't delete it. See [S3 Glacier Vault Lock](#) and [Abort Vault Lock \(DELETE lock-policy\)](#) in the Amazon S3 Glacier Developer Guide for more information.

Deploy the solution

This solution uses [AWS CloudFormation templates and stacks](#) to automate its deployment. The CloudFormation template specifies the AWS resources included in this solution and their properties. The CloudFormation stack provisions the resources that are described in the template.

Note

This solution doesn't delete the original archives or the source S3 Glacier vault. You must manually delete the archives and vault. For more information, refer to [Deleting an Archive in Amazon S3 Glacier](#) in the *Amazon S3 Glacier Developer Guide*.

If your source S3 Glacier vault has a [Vault Lock policy](#) that prevents deletion, you must delete this policy before deleting the original archives. However, if your Vault Lock policy is in the Locked state, you can't delete it. See [S3 Glacier Vault Lock](#) and [Abort Vault Lock \(DELETE lock-policy\)](#) in the Amazon S3 Glacier Developer Guide for more information.

Prerequisites

Before deployment, ensure that there are no new uploads or deletes of archives occurring on your source Glacier vault. Your Glacier vault content must be static.

- Ensure you have reviewed the [cost section](#) before deployment.
- Create a new destination Amazon S3 bucket. This bucket will be the destination storage location for your Glacier vault archives. For more information, refer to [Creating a bucket](#) in the Amazon Simple Storage Service User Guide.
 - The new destination Amazon S3 bucket should be in the same Region as the S3 Glacier vault, otherwise an excessive additional cost of "Data Transfer OUT From Amazon S3 Glacier" will be added, see [Data transfer pricing](#).
 - It is advisable to review and modify any Service Control Policies (SCP) on the destination bucket that may block or prevent PUT operations.
 - If you are using CloudTrail on your destination Amazon S3 bucket, please review and modify the CloudTrail export configurations to prevent excessive API charges.
- Ensure that your account has permissions to deploy the CloudFormation template and create the necessary AWS IAM roles. Your account must have permissions to grant access to the source Glacier vault and destination Amazon S3 bucket.

Deployment process overview

Follow the step-by-step instructions in this section to configure and deploy the solution into your account.

Before you launch the solution, review the [cost](#), [architecture](#), [security](#), and other considerations discussed earlier in this guide.

Time to deploy: Approximately 5–10 minutes

Note

The archive transfer can take up to one day to complete. If you need to move your data faster, contact [AWS Support](#) (AWS Developer Support plan or above).

[Step 1: Launch the stack](#)

[Step 2: Launch the transfer workflow](#)

[Step 3: Resume the transfer workflow](#)

Important

This solution includes an option to send anonymized operational metrics to AWS. We use this data to better understand how customers use this solution and related services and products. AWS owns the data gathered through this survey. Data collection is subject to the [AWS Privacy Notice](#).

To opt out of this feature, download the template, modify the AWS CloudFormation mapping section, and then use the AWS CloudFormation console to upload your updated template and deploy the solution. For more information, see the [Anonymized data collection](#) section of this guide.

AWS CloudFormation template

You can download the CloudFormation template for this solution before deploying it.

[View template](#)

data-transfer-from-amazon-s3-glacier-vaults-to-amazon-s3.template – Use this template to launch the solution and all associated components. The default configuration deploys the core and supporting services found in the [AWS services in this solution](#) section, but you can customize the template to meet your specific needs.

Note

AWS CloudFormation resources are created from AWS Cloud Development Kit (AWS CDK) (AWS CDK) constructs.

This AWS CloudFormation template deploys Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 in the AWS Cloud.

Important

- **Notifications** – The Amazon S3 Glacier service sends one notification per archive to the vault Amazon SNS topic (if it exists). If you don't want subscribers to the Amazon SNS topic to receive these notifications, confirm that the S3 Glacier vault being transferred doesn't have notifications enabled. For more information, see [Configuring Vault Notifications in Amazon S3 Glacier](#) in the *Amazon S3 Glacier Developer Guide*.
- **Inventory** – This solution copies your S3 Glacier vault archives once when you launch the stack. If you make changes to your vault after launching this solution, the solution doesn't replicate those changes.
- **Simultaneous workflows** – Running multiple transfer workflows simultaneously can exceed S3 Glacier quotas and induce throttling. We recommend that you only run one transfer workflow at a time.

Step 1: Launch the stack

Follow the step-by-step instructions in this section to configure and deploy the solution into your account.

Time to deploy: Approximately 5–10 minutes

Note

The archive transfer can take up to one day to complete. If you need to move your data faster, contact [AWS Support](#) (AWS Developer Support plan or above).

1. Sign in to the [AWS Management Console](#) and select the button to launch the data-retrieval-from-amazon-s3-glacier-vaults-to-amazon-s3.template AWS CloudFormation template.

Launch solution

2. The template launches in the US East (Ohio) Region by default. To launch the solution in a different AWS Region, use the Region selector in the console navigation bar.

Note

This solution uses AWS services that are not currently available in all AWS Regions. For the most current availability by Region, see the [AWS Regional Services List](#).

3. On the **Create stack** page, verify that the correct template URL is in the **Amazon S3 URL** text box and choose **Next**.
4. On the **Specify stack details** page, assign a name to your solution stack. For information about naming character limitations, see [IAM and AWS STS quotas, name requirements, and character limits](#) in the *AWS Identity and Access Management User Guide*.
5. Under **Parameters**, review the parameters for this solution template and modify them as necessary. This solution uses the following default values.

Parameter	Default	Description
DestinationBucket	empty	The name of the destination Amazon S3 bucket.

Parameter	Default	Description
		<p>⚠ Important</p> <p>Referring a bucket in a different region than S3 Glacier vault will incur additional costs. Refer to the cost section for more details.</p>
DynamoDB Backup	false	<p>Enter</p> <p>true to enable DynamoDB table backups for tables created by the solution.</p>
Lambda Tracing	false	<p>Enter</p> <p>true to enable AWS X-Ray tracing for Lambda functions created by the solution.</p>
Step Function Logging	false	<p>Enter</p> <p>true to enable logging for Step Functions created by the solution.</p>

Parameter	Default	Description
Step Function Tracing	false	Enter true to enable X-Ray tracing for Step Functions created by the solution.

Note

It is advisable to review and modify any [Service Control Policies](#) (SCP) on the destination bucket that may block or prevent PUT operations. If you are using CloudTrail on your destination Amazon S3 bucket, please review and modify the CloudTrail export configurations to prevent excessive API charges.

6. Select **Next**.
7. On the **Configure stack options** page, choose **Next**.
8. On the **Review** page, review and confirm the settings. Select the box acknowledging that the template will create IAM resources.
9. Choose **Create stack** to deploy the stack.

You can view the status of the stack in the AWS CloudFormation console in the **Status** column. You should receive a CREATE_COMPLETE status in approximately 5–10 minutes.

Step 2: Launch the transfer workflow

Important

Avoid starting multiple transfers simultaneously. Multiple transfers can exceed S3 Glacier quotas and induce throttling. When you transfer multiple vaults, initiate transfers sequentially by starting a new transfer only after the previous one completes.

1. Sign in to the [Systems Manager console](#).

2. Under **Shared Resources**, select **Documents**.
3. On the **Documents** page, select **Owned by me**, and choose the document called Launch-Data-Retrieval-for-Glacier-S3-`<NAME_OF_STACK>`.
4. Choose **Execute Automation**.

(Optional) Download the vault inventory file


If you want to download the vault inventory file as part of this migration, follow these steps.

1. Sign in to the [Systems Manager console](#).
2. Under **Shared Resources**, select **Documents**.
3. On the **Documents** page, select **Owned by me**, and choose the document called Launch-Data-Retrieval-for-Glacier-S3-`<NAME_OF_STACK>`.
4. Choose **Execute Automation**.
5. Enter the following under **Input Parameters** on the **Execute Automation Runbook** page.

Note

To enable the data transfer to a bucket in a different region from the S3 Glacier vault, you must manually adjust the automation document's content by setting the `InputPayload` parameter `allow_cross_region_data_transfer` to `true`. By default, the solution restricts cross-region transfers to avoid potential costs.

Parameter	Value	Notes
AcknowledgeAdditionalCostForCrossRegionTransfer	NO	Select YES only if you are aware of the excessive additional cost when selecting a destination bucket in a different region than the S3 Glacier vault. See Amazon S3 Glacier data transfer pricing .

Parameter	Value	Notes
ProvidedInventory	<i><Requires input></i>	Input with two options [YES, NO] indicate if the inventory is provided.
VaultName	<i><Requires input></i>	Enter the name of your S3 Glacier vault.
WorkflowRun	<i><Requires input></i>	Input to specify the workflow identifier of the workflow that needs to be resumed. <div data-bbox="1081 751 1507 1066" style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px;"> <p> Note</p> <p>Input to specify the workflow identifier of the workflow that needs to be resumed.</p> </div>
Description	<i><Optional input></i>	Provide an extended description for this migration .
NamingOverrideFile	<i><Optional input></i>	Provide a presigned URL of the NamingOverride file and the bucket that is storing the file if you want to customize S3 object key names .
S3StorageClass	<i><Requires input></i>	Select the S3 storage class for the migrated archives. See Amazon S3 pricing .

6. Choose **Execute**.

7. It takes approximately less than 1 minute to execute the document and launch the transfer. To confirm that the transfer process started successfully, refer to the *runScript* step status under the *Executed Steps* section.

Note

To monitor the progress of the transfer after launching the workflow, please refer to the Solution's [Cloudwatch dashboard](#). Please note that it might take 5-12 hours before the dashboard *Workflow Run ID* dropdown menu entries get updated with the new entry after launching the transfer.

(Optional) Provide the vault inventory file

If you want to provide the vault inventory file, follow these steps.

1. Sign in to the [CloudFormation console](#).
 - a. On the **Stacks** page, select this solution's installation stack.
 - b. Choose the **Output** tab and find the inventory bucket name.
2. Sign in to the [Amazon S3 console](#) and choose the inventory bucket from step 1b.
 - a. Choose **Create folder**. Enter the `workflow_run` value from the **WorkflowRun** parameter input. For more information about creating and using folders, see [Organizing objects in the Amazon S3 console by using folders](#) in the *Amazon Simple Storage Service User Guide*.
 - b. Create a subfolder called `original_inventory` in this new folder.
 - c. Copy the vault inventory file into this folder. For more information, see [Downloading a Vault Inventory in Amazon S3 Glacier](#) in the *Amazon S3 Glacier Developer Guide*.
 - d. Rename the vault inventory file to `inventory.csv`.
3. Sign in to the [Systems Manager console](#).
4. Under **Shared Resources**, select **Documents**.
5. On the **Documents** page, select **Owned by me**, and choose the document called `Launch-Data-Retrieval-for-Glacier-S3-<NAME_OF_STACK>`.
6. Choose **Execute Automation**.
7. Enter the following under **Input Parameters** on the **Execute Automation Runbook** page.

Note

To enable the data transfer to a bucket in a different region from the S3 Glacier vault, you must manually adjust the automation document's content by setting the `InputPayload` parameter `allow_cross_region_data_transfer` to `true`. By default, the solution restricts cross-region transfers to avoid potential costs.

Parameter	Value	Notes
AcknowledgeAdditionalCostForCrossRegionTransfer	NO	Select YES only if you are aware of the excessive additional cost when selecting a destination bucket in a different region than the S3 Glacier vault. See Amazon S3 Glacier data transfer pricing .
ProvidedInventory	Yes	
WorkflowRun	<i><Requires input></i>	Provide the name of your workflow run, which becomes the <code>workflow_run</code> value.
Description	<i><Optional input></i>	Provide an extended description for this migration .
NamingOverrideFile	<i><Optional input></i>	Provide a presigned URL of the NamingOverride file and the bucket that is storing the file if you want to customize S3 object key names .

Parameter	Value	Notes
S3StorageClass	<i><Requires input></i>	Select the S3 storage class for the migrated archives. See Amazon S3 pricing .

- Choose **Execute** It takes approximately less than 1 minute to execute the document and launch the transfer. To confirm that the transfer process started successfully, refer to the *runScript* step status under the *Executed Steps* section.

Note

To monitor the progress of the transfer after launching the workflow, please refer to the Solution's [Cloudwatch dashboard](#). Please note that it might take 5-12 hours before the dashboard *Workflow Run ID* dropdown menu entries get updated with the new entry after launching the transfer.

Step 3: Resume the transfer workflow

- Sign in to the [Systems Manager console](#).
- Under **Shared Resources**, select **Documents**.
- On the **Documents** page, select **Owned by me**, and choose the document called `Resume-Data-Retrieval-for-Glacier-S3-<NAME_OF_STACK>`.
- Choose **Execute Automation**.
- Enter the following under **Input Parameters** on the **Execute Automation Runbook** page.

Parameter	Value	Notes
AcknowledgeAdditionalCostForCrossRegionTransfer	NO	Select YES only if you are aware of the excessive additional cost when selecting a destination bucket in a different region than the S3 Glacier vault.

Parameter	Value	Notes
		See Amazon S3 Glacier data transfer pricing .
ProvidedInventory	<i><Requires input></i>	Input with two options [YES, NO] indicate if the inventory is provided.
WorkflowRun	<i><Requires input></i>	Provide the same <code>workflow_run</code> value used in the initial execution of the transfer workflow.
Description	<i><Optional input></i>	Provide an extended description for this migration .
NamingOverrideFile	<i><Optional input></i>	Provide a presigned URL of the NamingOverride file and the bucket that is storing the file if you want to customize S3 object key names .
S3 Storage class	<i><Requires input></i>	Select the S3 storage class for the migrated archives. See Amazon S3 pricing .

 **Note**

To monitor the progress of the transfer after launching the workflow, please refer to the Solution's [Cloudwatch dashboard](#). Please note that it might take 5-12 hours before the dashboard **Workflow Run ID** dropdown menu entries get updated with the new entry after launching the transfer.

Monitor the solution with Service Catalog AppRegistry

This solution includes a Service Catalog AppRegistry resource to register the CloudFormation template and underlying resources as an application in both [Service Catalog AppRegistry](#) and [AWS Systems Manager Application Manager](#).

AWS Systems Manager Application Manager gives you an application-level view into this solution and its resources so that you can:

- Monitor its resources, costs for the deployed resources across stacks and AWS accounts, and logs associated with this solution from a central location.
- View operations data for the resources of this solution (such as deployment status, CloudWatch alarms, resource configurations, and operational issues) in the context of an application.

The following figure depicts an example of the application view for the solution stack in Application Manager.

The screenshot displays the AWS Systems Manager Application Manager console. The main heading is 'AWS-Systems-Manager-Application-Manager' with a refresh icon and a 'Start runbook' button. Below this is the 'Application information' section, which includes a 'View in AppRegistry' link. The application type is 'AWS-AppRegistry', the name is 'AWS-Systems-Manager-Application-Manager', and application monitoring is 'Not enabled'. The description reads: 'Service Catalog application to track and manage all your resources for the solution'. A navigation bar below the application information includes tabs for Overview, Resources, Instances, Compliance, Monitoring, OpsItems, Logs, Runbooks, and Cost. The 'Overview' tab is active. Below the navigation bar are two sections: 'Insights and Alarms' with a 'View all' button and a note to monitor application health with Amazon CloudWatch, and 'Cost' with a 'View all' button and a note to view resource costs per application using AWS Cost Explorer. The cost section shows a table with the header 'Cost (USD)' and a single row with a dash '-'.

Solution stack in Application Manager

Activate CloudWatch Application Insights

1. Sign in to the [Systems Manager console](#).

2. In the navigation pane, choose **Application Manager**.
3. In **Applications**, search for the application name for this solution and select it.

The application name will have App Registry in the **Application Source** column, and will have a combination of the solution name, Region, account ID, or stack name.

4. In the **Components** tree, choose the application stack you want to activate.
5. In the **Monitoring** tab, in **Application Insights**, select **Auto-configure Application Insights**.

The screenshot shows the AWS Application Insights Monitoring page. The navigation bar includes Overview, Resources, Provisioning, Compliance, Monitoring (selected), OpsItems, Logs, Runbooks, and Cost. The main content area is titled "Application Insights (0) Info" and includes a toggle for "View Ignored Problems", an "Actions" dropdown, and an "Add an application" button. Below this is a search bar with the placeholder "Find problems", a "Last 7 days" filter, a refresh button, and pagination controls showing "1" of 1 items. A table header is visible with columns: Problem su..., Status, Severity, Source, Start time, and Insights. The main content area displays a message: "Advanced monitoring is not enabled. When you onboard your first application, a service-linked role (SLR) is created in your account. The SLR is predefined by CloudWatch Application Insights and includes the permissions the service requires to monitor AWS services on your behalf." Below this message is an "Auto-configure Application Insights" button.

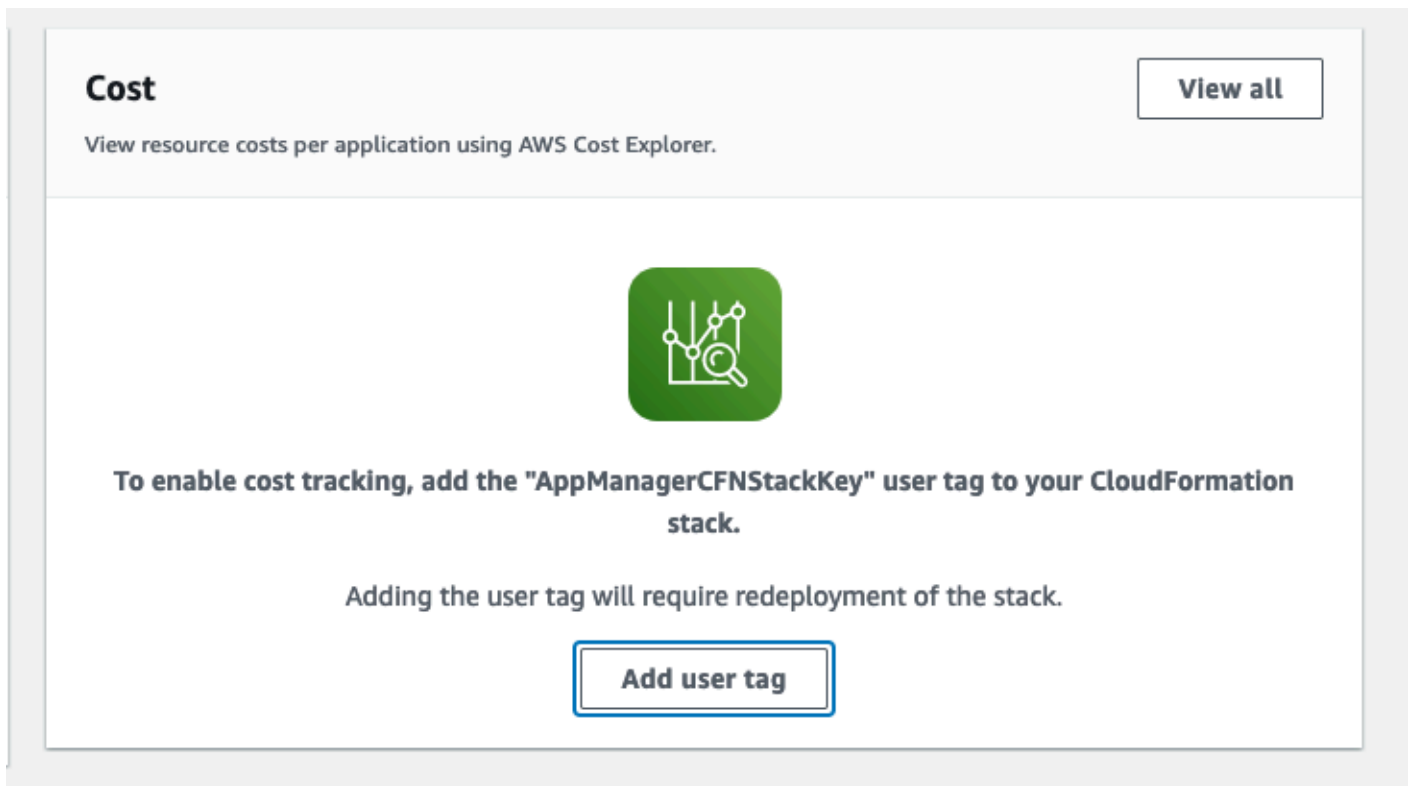
Monitoring for your applications is now activated and the following status box appears:

The screenshot shows the AWS Application Insights Monitoring page after successful activation. The navigation bar and top controls are the same as in the previous screenshot. The main content area now displays a green status box with a checkmark icon and the text: "Application monitoring has been successfully enabled. It will take some time to display any results. Please use the refresh button to view results." The rest of the page, including the search bar, filters, and table header, remains the same.

Confirm cost tags associated with the solution

After you activate cost allocation tags associated with the solution, you must confirm the cost allocation tags to see the costs for this solution. To confirm cost allocation tags:

1. Sign in to the [Systems Manager console](#).
2. In the navigation pane, choose **Application Manager**.
3. In **Applications**, choose the application name for this solution and select it.
4. In the **Overview** tab, in **Cost**, select **Add user tag**.



5. On the **Add user tag** page, enter `confirm`, then select **Add user tag**.

The activation process can take up to 24 hours to complete and the tag data to appear.

Activate cost allocation tags associated with the solution

After you confirm the cost tags associated with this solution, you must activate the cost allocation tags to see the costs for this solution. The cost allocation tags can only be activated from the management account for the organization.

To activate cost allocation tags:

1. Sign in to the [AWS Billing and Cost Management and Cost Management console](#).
2. In the navigation pane, select **Cost Allocation Tags**.
3. On the **Cost allocation tags** page, filter for the AppManagerCFNStackKey tag, then select the tag from the results shown.
4. Choose **Activate**.

AWS Cost Explorer

You can see the overview of the costs associated with the application and application components within the Application Manager console through integration with AWS Cost Explorer. Cost Explorer helps you manage costs by providing a view of your AWS resource costs and usage over time.

1. Sign in to the [AWS Cost Management console](#).
2. In the navigation menu, select **Cost Explorer** to view the solution's costs and usage over time.

Update the solution

If you have previously deployed the solution, follow this procedure to update the CloudFormation stack to get the latest version of the solution's framework.

1. Sign in to the [AWS CloudFormation console](#), select your existing Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 CloudFormation stack, and select **Update**.
2. Select **Replace current template**.
3. Under **Specify template**:
 - a. Select **Amazon S3 URL**.
 - b. Copy the link of the [latest template](#).
 - c. Paste the link in the **Amazon S3 URL** box.
 - d. Verify that the correct template URL shows in the **Amazon S3 URL** text box, and choose **Next**. Choose **Next** again.
4. Under **Parameters**, review the parameters for the template and modify them as necessary. For details about the parameters, see [Step 1: Launch the stack](#).
5. Choose **Next**.
6. On the **Configure stack options** page, choose **Next**.
7. On the **Review** page, review and confirm the settings. Select the box acknowledging that the template creates IAM resources.
8. Choose **View change set** and verify the changes.
9. Choose **Update stack** to deploy the stack.

You can view the status of the stack in the AWS CloudFormation console in the **Status** column. You should receive an UPDATE_COMPLETE status in approximately 5-10 minutes.

Troubleshooting

This section provides troubleshooting instructions for deploying and using the solution.

If these instructions don't address your issue, see the [Contact AWS Support](#) section for instructions on opening an AWS Support case for this solution.

Problem: Transfer workflow has not progressed after 14 hours

Your transfer workflow has been launched for more than 14 hours, but the **Downloaded** count on the CloudWatch dashboard has not increased.

Resolution

1. Sign in to the [Step Functions console](#).
2. Under **State machines**, select the state machine called `OrchestratorStateMachine$CFN_ID` and choose **View Details**.
3. Select the most recent execution and choose **View details**.
4. Note failures. If the workflow is still running the **InventoryRetrieval** workflow, there might be an issue with the Amazon S3 Glacier service generating the inventory file. Contact [AWS Support](#) if you have an AWS Developer Support plan or above.

Problem: Transfer workflow must be stopped

Your transfer workflow is ongoing, but you want to stop it.

Resolution

1. Follow steps 1–3 in [Problem: Transfer workflow has not progressed after 14 hours](#).
2. Under **Actions**, choose **Stop execution**.
3. If you plan to [resume the transfer workflow later](#), find the `workflow_run` value from the **Execution Input and Output** tab on this page. You need this value to resume the workflow.

Note

After you stop the execution, no new archives are requested from the Amazon S3 Glacier service. The archives that were already requested will download. These downloads take 4–8 hours to complete.

Contact AWS Support

If you have [AWS Developer Support](#), [AWS Business Support](#), or [AWS Enterprise Support](#), you can use the Support Center to get expert assistance with this solution. The following sections provide instructions.

Create case

1. Sign in to [Support Center](#).
2. Choose **Create case**.

How can we help?

1. Choose **Technical**.
2. For **Service**, select **Solutions**.
3. For **Category**, select **Other Solutions**.
4. For **Severity**, select the option that best matches your use case.
5. When you enter the **Service**, **Category**, and **Severity**, the interface populates links to common troubleshooting questions. If you can't resolve your question with these links, choose **Next step: Additional information**.

Additional information

1. For **Subject**, enter text summarizing your question or issue.
2. For **Description**, describe the issue in detail.
3. Choose **Attach files**.
4. Attach the information that AWS Support needs to process the request.

Help us resolve your case faster

1. Enter the requested information.
2. Choose **Next step: Solve now or contact us**.

Solve now or contact us

1. Review the **Solve now** solutions.
2. If you can't resolve your issue with these solutions, choose **Contact us**, enter the requested information, and choose **Submit**.

Uninstall the solution

You can uninstall the Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 solution from the AWS Management Console or by using the AWS Command Line Interface. Manually delete the following resources:

- S3 buckets (other than the output bucket if you intend to keep the transferred archives)
- DynamoDB tables
- CloudWatch Logs

AWS Solutions do not automatically delete these resources in case you have stored data to retain.

Using the AWS Management Console

1. Sign in to the [CloudFormation console](#).
2. On the **Stacks** page, select this solution's installation stack.
3. Choose **Delete**.

Using AWS Command Line Interface

Determine whether the AWS Command Line Interface (AWS CLI) is available in your environment. For installation instructions, see [What Is the AWS Command Line Interface](#) in the *AWS CLI User Guide*. After confirming that the AWS CLI is available, run the following command.

```
$ aws cloudformation delete-stack --stack-name <installation-stack-name>
```

Deleting the S3 buckets

This solution is configured to retain the solution-created S3 buckets if you decide to delete the CloudFormation stack, to prevent accidental data loss. After uninstalling the solution, you can manually delete the S3 buckets if you don't need to retain the data. Follow these steps to delete the S3 buckets.

1. Sign in to the [Amazon S3 console](#).

2. Choose **Buckets** from the left navigation pane.
3. Locate the *<stack-name>* S3 buckets.
4. Select each S3 bucket and choose **Empty**.
5. Select each S3 bucket and choose **Delete**.

To delete the S3 bucket using the AWS CLI, run the following command:

```
$ aws s3 rb s3://<bucket-name> --force
```

Deleting the DynamoDB tables

This solution is configured to retain the solution-created DynamoDB tables if you decide to delete the CloudFormation stack, to prevent accidental data loss.

1. Sign in to the [DynamoDB console](#).
2. Choose **Tables** from the left navigation pane.
3. Locate the *<stack-name>* DynamoDB tables.
4. Select each DynamoDB table and choose **Delete**.
5. Confirm the deletion.

To delete the DynamoDB table using the AWS CLI, run the following command:

```
$ aws dynamodb delete-table --table-name <table-name>
```

Deleting the CloudWatch Logs

1. Sign in to the [CloudWatch console](#).
2. Choose **Log groups** from the left navigation pane.
3. Locate the *<stack-name>* CloudWatch Logs.
4. Select each CloudWatch Log and choose **Actions**, then **Delete log groups**.
5. Confirm the deletion.

To delete the CloudWatch Logs using the AWS CLI, run the following command:

```
$ aws logs delete-log-group --log-group-name <log-group-name>
```

Use the solution

This section provides a user guide for using the AWS solution.

Validate your inventory

After you deploy this solution and transfer your archives, you can validate your Amazon S3 inventory to confirm whether all of your archives transferred. We recommend confirming that all of your archives transferred before deleting your original archive. See [Amazon S3 Inventory](#) for more information.

Access the CloudWatch dashboard

1. Sign in to the [CloudWatch console](#).
2. Choose **Dashboards** from the left navigation pane.
3. Select the dashboard that starts with `Data-Transfer-from-Amazon-S3-Glacier-to-Amazon-S3-Dashboard-`.
4. In the dashboard, you can find:
 - The total number of archives in the inventory file, and their collective size.
 - The total number of archives that you requested for download, and their collective size.
 - The total number of archives that are staged by the Amazon S3 Glacier service for download, and their collective size.
 - The total number of archives that won't be downloaded because they're larger than 5 GB in size.
 - The total number of downloaded archives.
5. When the number of downloaded archives matches the total number of requested archives, the transfer is complete.

Manage your Amazon S3 storage

After you transfer your Amazon S3 Glacier data to the Amazon S3 service, you can change your storage classes to fit your use cases. For information about how to do this, see the following resources:

- [Managing your Amazon S3 storage](#) in the *Amazon Simple Storage Service User Guide*
- [Using Amazon S3 storage classes](#) in the *Amazon Simple Storage Service User Guide*

Developer guide

This section provides the source code for the solution.

Source code

Visit our [GitHub repository](#) to download the source files for this solution and to share your customizations with others.

The AWS CDK generates the Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 templates. See the [README.md](#) file for additional information.

Reference

This section includes information about an optional feature for collecting unique metrics for this solution, and a list of builders who contributed to this solution.

Anonymized data collection

This solution includes an option to send anonymized operational metrics to AWS. We use this data to better understand how customers use this solution and related services and products. When invoked, the following information is collected and sent to AWS:

- **Solution ID** – The AWS solution identifier
- **Unique ID (UUID)** – Randomly generated, unique identifier for each Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 deployment
- **Timestamp** – Data-collection timestamp
- **Data** – The solution collects and sends the following statistics:
 - **Region** – Selected AWS Region
 - **Version** – Version of the solution deployment
 - **StorageClass** – Selected destination S3 storage class
 - **RetrievalTier** – Type of Amazon S3 Glacier retrieval (for example, Bulk)
 - **VaultSize** – Source S3 Glacier vault size
 - **ArchiveCount** – Number of S3 Glacier archives

AWS owns the data gathered through this survey. Data collection is subject to the [Privacy Notice](#). To opt out of this feature, complete the following steps before launching the AWS CloudFormation template.

1. Download the AWS CloudFormation template to your local hard drive.
2. Open the [AWS CloudFormation template](#) with a text editor.
3. Modify the Environment variable section for the SendAnonymizedStats Lambda function from:

```
"SEND_ANONYMIZED_STATS": "Yes"
```

to:

```
"SEND_ANONYMIZED_STATS": "No"
```

4. Sign in to the [AWS CloudFormation console](#).
5. Select **Create stack**.
6. On the **Create stack** page, **Specify template** section, select **Upload a template file**.
7. Under **Upload a template file**, choose **Choose file** and select the edited template from your local drive.
8. Choose **Next** and follow the steps in [Launch the stack](#) in the Deploy the solution section of this guide.

Contributors

- Bassem Wanis
- Kamyar Ziabari
- Nathaniel Schaaf
- Garvit Singh
- Simon Krol
- Evgeny Minkevich

Revisions

Date	Change
December 2023	Initial release
April 2024	Release v1.1.0: Added the ability for customers to reference an external destination bucket, added an SQS widget to the CloudWatch dashboard, added a pre-built CloudWatch Logs Insights query, and addressed bug fixes. For additional details on updates and new features, refer to the CHANGELOG.md file in the GitHub repository.
May 2024	Release v1.1.1: Upgraded Lambda runtime to Python 3.12, extended the list of supported AWS Regions, and completed other bug fixes. For additional details on updates and new features, refer to the CHANGELOG.md file in the GitHub repository.

Notices

Customers are responsible for making their own independent assessment of the information in this document. This document: (a) is for informational purposes only, (b) represents AWS current product offerings and practices, which are subject to change without notice, and (c) does not create any commitments or assurances from AWS and its affiliates, suppliers or licensors. AWS products or services are provided "as is" without warranties, representations, or conditions of any kind, whether express or implied. AWS responsibilities and liabilities to its customers are controlled by AWS agreements, and this document is not part of, nor does it modify, any agreement between AWS and its customers.

Data Transfer from Amazon S3 Glacier Vaults to Amazon S3 is licensed under the terms of the Apache License Version 2.0 available at [The Apache Software Foundation](https://www.apache.org/licenses/LICENSE-2.0).