Amazon Glacier FAQs

General

Q: What is Amazon Glacier?

Amazon Glacier is an extremely low-cost storage service that provides secure, durable, and flexible storage for data backup and archival. With Amazon Glacier, customers can reliably store their data for as little as $0.004 per gigabyte per month. Amazon Glacier enables customers to offload the administrative burdens of operating and scaling storage to AWS, so that they don't have to worry about capacity planning, hardware provisioning, data replication, hardware failure detection and repair, or time-consuming hardware migrations.

Q: How can businesses, government and other organizations benefit from Amazon Glacier?

Amazon Glacier enables any business or organization to easily and cost effectively retain data for months, years, or decades. With Amazon Glacier, customers can now cost effectively retain more of their data for future analysis or reference, and they can focus on their business rather than operating and maintaining their storage infrastructure. Customers seeking compliance storage can deploy compliance controls using Vault Lock to meet regulatory and compliance archiving requirements.

Q: How should I choose between Amazon Glacier and Amazon Simple Storage Service (Amazon S3)?

Amazon S3 is a durable, secure, simple, and fast storage service designed to make web-scale computing easier for developers. Use Amazon S3 if you need low latency or frequent access to your data. Use Amazon Glacier if low storage cost is paramount, and you do not require millisecond access to your data.

Q: What kind of data can I store?

You can store virtually any kind of data in any format. You can also deploy compliance storage controls with Vault Lock to store regulatory and compliance archives in an immutable, Write Once Read Many (WORM) format. Please refer to the Amazon Web Services Licensing Agreement for details.

Q: What does Amazon do with my data in Amazon Glacier?

Amazon will store your data and track its associated usage for billing purposes. Amazon will not otherwise access your data for any purpose outside

of the Amazon Glacier offering, except if required to do so by law. Please refer to the Amazon Web Services Licensing Agreement for details.

Q: How do I use Amazon Glacier?

Amazon Glacier provides a simple, standards-based REST web services interface as well as Java and .NET SDKs. The AWS Management console can be used to quickly set up Amazon Glacier. Data can then be uploaded and retrieved programmatically. View our documentation for more information on the Glacier APIs and SDKs.

Q: How durable is Amazon Glacier?

Amazon Glacier is designed to provide average annual durability of 99.999999999% for an archive. The service redundantly stores data in multiple facilities and on multiple devices within each facility. To increase durability, Amazon Glacier synchronously stores your data across multiple facilities before returning SUCCESS on uploading archives. Glacier performs regular, systematic data integrity checks and is built to be automatically self-healing.

Getting Started

Q: How is data within Amazon Glacier organized?

You store data in Amazon Glacier as an archive. Each archive is assigned a unique archive ID that can later be used to retrieve the data. An archive can represent a single file or you may choose to combine several files to be uploaded as a single archive. You upload archives into vaults. Vaults are collections of archives that you use to organize your data.

Q: How much data can I store?

There is no maximum limit to the total amount of data that can be stored in Amazon Glacier. Individual archives are limited to a maximum size of 40 terabytes.

Q: What is the minimum amount of data that I can store using Amazon Glacier?

There is no minimum limit to the amount of data that can be stored in Amazon Glacier and individual archives can be from 1 byte to 40 terabytes.

Q: Does the AWS Management Console support Amazon Glacier?

Yes. The AWS Management Console allows you to create and configure vaults, allowing you to easily and quickly setup Glacier. Click here to go the AWS Management Console.

Billing

Q: How much does Amazon Glacier cost?

With Amazon Glacier, storage is priced from $0.004 per gigabyte per month, and you pay for what you use. There are no setup fees, and for most archive use cases your total costs will primarily be made up of your storage cost.

Upload requests are priced from $0.05 per 1,000 requests. In addition, archives stored in Glacier have a minimum 90 days of storage, and archives deleted before 90 days incur a pro-rated charge equal to the storage charge for the remaining days. As Amazon Glacier is designed to store data that is infrequently accessed and long lived, these charges will likely not apply to most of you.

We charge less where our costs are less. Some prices vary across Amazon Glacier Regions and are based on the location of your vault. There is no Data Transfer charge for data transferred between Amazon EC2 and Amazon Glacier within the same Region. Data transferred between Amazon EC2 and Amazon Glacier across all other Regions (e.g. between the Amazon EC2 Northern California and Amazon Glacier US East North Virginia Regions) will be charged at Internet Data Transfer rates on both sides of the transfer.

To learn more about Glacier pricing, please visit the Glacier pricing page.

Q: How is my storage charge calculated?

The volume of storage billed in a month is based on the average storage used throughout the month, measured in gigabyte-months (GB-Months). The size of each of your archives is calculated as the amount of data you upload plus an additional 32 kilobytes of data for indexing and metadata (e.g. your archive description). This extra data is necessary to identify and retrieve your archive. Here is an example of how to calculate your storage costs using US East (Northern Virginia) Region pricing:

If you upload 100,000 archives that are 1 gigabyte each, your total storage would be:

1.000032 gigabytes for each archive x 100,000 archives = 100,003.20 gigabytes

If you stored the archives for 1 month, you would be charged:

100,003.20 GB-Months x $0.004 = $400.01

If you upload 200,000 archives that are 0.5 gigabytes each, your total storage would be:

0.500032 gigabytes for each archive x 200,000 archives = 100,006.40 gigabytes

If you stored the archives for 1 month, you would be charged:

100,006.40 GB-Months x $0.004 = $400.03

Your storage is measured in "TimedStorage-ByteHrs," which are added up at the end of the month to generate your monthly charges. For example, if you store an archive that is 1 gigabyte (inclusive of the 32 kilobyte overhead) for one day in the US East (Northern Virginia) Region, your storage usage would be:

1,073,741,824 bytes x 1 day x 24 hours = 25,769,803,776 Byte-Hours

Converting this to GB-Months (assuming a 30 day month) gives:

25,769,803,776 Byte-Hours x (1 GB / 1,073,741,824 bytes) x (1 month / 720 hours) = 0.03 GB-Months

So your storage charge for that day would be:

0.03 GB-Months x $0.004 = $0.00012

To learn more about Glacier pricing and view prices for other regions, please visit the Glacier pricing page.

Q: Why do prices vary depending on which Amazon Glacier Region I choose?

We charge less where our costs are less. For example, our costs are lower in the US East (North Virginia) Region than in the US West (Northern California) Region.

Q: How will I be charged and billed for my use of Amazon Glacier?

There are no setup fees to begin using the service. At the end of the month, your credit card will automatically be charged for that month's usage. You can view your charges for the current billing period at any time on the Amazon Web Services web site, by logging into your Amazon Web Services account, and clicking "Account Activity" under "Your Web Services Account".

Q: How much data can I retrieve for free?

Glacier offers a 10 GB retrieval free tier. You can retrieve 10 GB of your Amazon Glacier data per month for free. The free tier allowance can be used at any time during the month and applies to Standard retrievals.

Q: How much does it cost to retrieve data from Amazon Glacier?

There are three ways to retrieve data from Glacier and each has a different per-GB retrieval fee and per-archive request fee (i.e. requesting one archive counts as one request). Expedited retrievals cost $0.03 per GB and $0.01 per request. Standard retrievals cost $0.01 per GB and $0.05 per 1,000 requests. Bulk retrievals cost $0.0025 per GB and $0.025 per 1,000 requests.

For example, using Expedited retrievals, if you requested 10 archives with a size of 1 GB each, the cost would be 10 x $0.03 +10 x $0.01 = $0.40.

If you were using Standard retrievals to retrieve 500 archives that were 1 GB each, the cost would be 500GB x $0.01 + 500 x $0.05/1,000 = $5.025

Lastly, using Bulk retrievals, if you were to retrieve 500 archives that are 1 GB each, the cost would be 500GB x $0.0025 + 500 x $0.025/1,000 = $1.2625.

To learn more about Glacier pricing, please visit the Glacier pricing page.

Q: How will I be charged when retrieving only a range of an archive?

Range retrievals are priced in precisely the same way as regular retrievals from Amazon Glacier. You are charged a per-GB fee for only the amount of data retrieved in the range you specify.

Q: How will I be charged for deleting data that is less than 3 months old?

Amazon Glacier is designed for use cases where data is retained for months, years, or decades. Deleting data from Amazon Glacier is free if the archive being deleted has been stored for three months or longer. If an archive is deleted within three months of being uploaded, you will be charged an early deletion fee. In the US East (Northern Virginia) Region, you would be charged a prorated early deletion fee of $0.012 per gigabyte deleted within three months. So if you deleted 1 gigabyte of data 1 month after uploading it, you would be charged a $0.008 early deletion fee. If, instead you deleted 1 gigabyte after 2 months, you would be charged a $0.004 early deletion fee.

To view prices for other regions, visit the Glacier pricing page.

Q: What can I expect the total cost of ownership (TCO) to be?

Amazon Glacier is a secure, durable, and extremely low-cost cloud storage service for data archiving and long-term backup. Customers can reliably store large or small amounts of data for as little as $0.004 per gigabyte per month, a significant savings compared to on-premises solutions. To keep costs low yet suitable for varying retrieval needs, Amazon Glacier provides three options for access to archives, from a few minutes to several hours. Your total cost of ownership (TCO) for your Amazon Glacier storage will depend on your data access patterns. Below are several examples illustrating different use cases ranging from deep archives that are never retrieved to active workloads where large portions of data are accessed.

TCO example 1: Let's assume that you upload 1 PB of data into Amazon Glacier, that the average archive size is 1 GB and that you never retrieve any data. When you first upload the 1 PB, there are upload request fees of 1,048,576 GB x $0.05 / 1,000 = $52.43. Then the ongoing storage costs are 1,048,576 GB x $0.004 = $4,194.30 per month, or $50,331.65 per year.

TCO example 2: Now let's assume the same storage as example 1 and also assume that you retrieve 3 TB (3,072 GB) a day on average using Bulk retrievals and that the average archive size was 1 GB for a total of 3,072 archives. That's 90 TB retrieved per month or 8.8% of your data per month.

The total retrieval fees per day would be 3,072 x $0.0025 + 3,072 * $0.025 / 1,000 = $7.76, which equates to $232.70 per month and $2,792.45 per year. Adding storage costs, your annual TCO is $50,331.65 + $2,792.45 = $53,124.10. In this example, retrieval fees make up just 5.3% of your total Glacier fees. Your total monthly cost per GB stored including retrieval fees is $0.004222/GB.

TCO example 2: Now let's assume the same storage as example 1 and also assume that you retrieve 1 TB (1,024 GB) a day on average using Standard retrievals and that occasionally you use Expedited retrievals for urgent requests, averaging 10 GB per day. Here, we assume the average archive size is 1 GB. That's 30.3 TB per month or 3% of your data per month. The total retrieval fees per day would be (1,024 x $0.01 + 1,024 x $0.05 / 1000) + (10 x $0.03 + 10 x $0.01) = $10.69, which equates to $320.74 per month and $3,848.83 per year. Adding storage costs, your annual TCO is $50,331.65 + $3,848.83 = $54,180.48. In this example, retrieval fees make up just 7.1% of your total Glacier fees. Your total monthly cost per GB stored including retrieval fees is $0.0043/GB.

To learn more about Glacier pricing, please visit the Glacier pricing page.

Q: Do your prices include taxes?

Except as otherwise noted, our prices are exclusive of applicable taxes and duties, including VAT and applicable sales tax. For customers with a Japanese billing address, use of AWS services is subject to Japanese Consumption Tax. Learn more.

Security

Q: How do I control access to my data?

By default, only you can access your data. In addition, you can control access to your data in Amazon Glacier by using the AWS Identity and Access Management (AWS IAM) service. You simply set up an AWS IAM policy that specifies which users within an account have rights to operations on a given vault.

Q: Is my data encrypted?

Yes, all data in the service will be encrypted on the server side. Amazon Glacier handles key management and key protection for you. Amazon Glacier uses one of the strongest block ciphers available, 256-bit Advanced Encryption Standard (AES-256). 256-bit is the largest key size defined for AES. Customers wishing to manage their own keys can encrypt data prior to uploading it.

Q: Does Amazon Glacier support IAM permissions?

Yes, Glacier will support API-level permissions through AWS Identity and Access Management (IAM) service integration

For more information about IAM, go to:

- AWS Identity and Access Management
- AWS Identity and Access Management Getting Started Guide
- Using AWS Identity and Access Management

Archives and Vaults

Q: What is an archive?

An archive is a durably stored block of information. You store your data in Amazon Glacier as archives. You may upload a single file as an archive, but your costs will be lower if you aggregate your data. TAR and ZIP are common formats that customers use to aggregate multiple files into a single file before uploading to Amazon Glacier. The total volume of data and number of archives you can store are unlimited. Individual Amazon Glacier archives can range in size from 1 byte to 40 terabytes. The largest archive that can be uploaded in a single Upload request is 4 gigabytes. For items larger than 100 megabytes, customers should consider using the Multipart upload capability. Archives stored in Amazon Glacier are immutable, i.e. archives can be uploaded and deleted but cannot be edited or overwritten.

Q: How do I delete archives?

You can delete an archive at any time. You will stop being billed for your archive when your delete request succeeds at which point the archive itself will be inaccessible. Archives that are deleted within 3 months of being uploaded will be charged a deletion fee (see billing section for more details).

Q: How do I upload large archives?

When uploading large archives (100MB or larger), you can use multi-part upload to achieve higher throughput and reliability. Multi-part uploads allow you to break your large archive into smaller chunks that are uploaded individually. Once all the constituent parts are successfully uploaded, they are combined into a single archive.

Q: What is a vault?

A vault is a way to group archives together in Amazon Glacier. You organize your data in Amazon Glacier using vaults. Each archive is stored in a vault of your choice. You may control access to your data by setting vault-level access policies using the AWS Identity and Access Management (IAM) service. You can also attach notification policies to your vaults. These enable you or your application to be notified when data that you have requested for retrieval is ready for download. Click here to learn more about setting up notifications using the Amazon Simple Notification Service (Amazon SNS).

Q: How many vaults can I create?

You can create up to 1,000 vaults per account per region.

Q: How do I effectively manage my Amazon Glacier vaults?

Amazon Glacier allows you to tag your Glacier vaults for easier resource and cost management. Tags are labels that you can define and associate with your vaults, and using tags adds filtering capabilities to operations such as AWS cost reports. For example, you can use tags to allocate Glacier costs and usage across multiple departments in your organization or by any other categorization. You can tag your vaults by using the Glacier Console or the Glacier APIs. For more information see Tagging Your Amazon Glacier Vaults.

Q: How do I delete a vault?

You may delete any Glacier vault that does not contain any archives using the AWS Management Console, the Amazon Glacier APIs or the SDKs. Once a vault has been deleted, you can then re-create a vault with the same name. If your vault contains archives, you must delete all the archives before deleting the vault.

Vault Access Policies

Q: What is a vault access policy?

A vault access policy is a resource-based policy that you can attach directly to your Glacier vault (the resource) to specify who has access to the vault and what actions they can perform on it. To learn more please read Managing Vault Access Policies in the Amazon Glacier developer's guide.

Q: How are vault access policies different from access control based on AWS Identity and Access Management (IAM) policies?

Access permissions can be assigned in two ways: as user-based permissions or as resource-based permissions. Access control based on IAM policies is user-based where you would assign IAM policies to IAM users or groups to control the read, write, and delete permissions on your Glacier vaults. Access control with vault access policies is resource-based where you would attach an access policy directly on a vault to govern access to all users. Vault access policies can make certain use cases simpler. For example, to protect information in a business-critical vault from unintended deletion, you can create a vault access policy that denies delete attempts from all users. This data protection procedure can be accomplished in a matter of minutes in the AWS Management Console without having to audit and revoke delete permissions assigned to users through IAM policies.

Q: Can I use vault access policies to manage cross-account access?

Yes you can. For example, you can grant read-only access on your vault to a business partner in a different AWS account by simply adding that account to the vault's access policy and specifying that only read activities are allowed.

Q: How does billing work in a cross-account access scenario?

The vault owner's account will be billed for the charges incurred during cross-account access. For example, Alice (account A) grants Bob (account B) access to Alice's "movies" vault and allows Bob to upload data. After Bob makes 1000 requests to upload 1GB of data, Alice's account (account A) will be billed for the 1000 requests as well as the 1GB of data until the data is deleted. Bob's account (account B) will not incur these charges.

Q: How do I create and manage vault access policies?

You can create and manage vault access policies in the AWS Glacier console or use the vault access APIs in the AWS SDK. To learn more please read Managing Vault Access Policies in the Amazon Glacier developer's guide.

Q: How many vault access policies can I have?

You can set one vault access policy for each vault. The vault access policy can be used as a single location to view the list of users with vault access and the allowed actions for each user.

Vault Lock

Q: What is Vault Lock?

Vault Lock allows you to easily deploy and enforce compliance controls on individual Glacier vaults via a lockable policy (Vault Lock policy). Once locked, the Vault Lock policy becomes immutable and Glacier will enforce the prescribed controls to help achieve your compliance objectives. To learn more, please read Amazon Glacier Vault Lock in the Amazon Glacier developer's guide.

Q: What type of compliance controls can I deploy with Vault Lock?

You can deploy a variety of compliance controls in a Vault Lock policy using the AWS Identity and Access Management (IAM) policy language. For example, you can easily set up "Write Once Read Many" (WORM) or time-based records retention for regulatory archives. To learn more, please read Amazon Glacier Vault Lock in the Amazon Glacier developer's guide.

Q: How does Vault Lock enforce my compliance controls?

Vault Lock enforces your compliance controls via a lockable policy (Vault Lock policy). Once locked, the Vault Lock policy becomes immutable and Glacier will only allow operations on your data that are explicitly permitted by the compliance controls you specified. Vault Lock also ensures that a locked policy cannot be deleted or altered until there are no more archives to protect in the

vault. Learn more about Locking a Vault for compliance in the Amazon Glacier developer's guide.

Q: How is a Vault Lock policy different than a vault access policy?

Both policies govern access controls to your vault, however, a Vault Lock policy can be made immutable and provides strong enforcement for your compliance controls. You can use the Vault Lock policy to deploy regulatory and compliance controls that are typically restrictive and are "set and forget" in nature. In conjunction, you can use the vault access policy to implement access controls that are not compliance related, temporary, and subject to frequent modification. The two policies can be used in tandem to achieve governance and flexibility.

Q: What AWS electronic storage services have been assessed based on financial services regulations?

For customers in the financial services industry, Vault Lock provides added support for broker-dealers who must retain records in a non-erasable and non-rewritable format to satisfy regulatory requirements of SEC Rule 17a-4(f), FINRA Rule 4511, or CFTC Regulation 1.31. You can easily designate the records retention time frame to retain regulatory archives in the original form for the required duration, and also place legal holds to retain data indefinitely until the hold is removed.

Q: What AWS documentation supports the SEC 17a-4(f)(2)(i) and CFTC 1.31(c) requirement for notifying my regulator?

Provide notification to your regulator or "Designated Examining Authority (DEA)" of your choice to use AWS Glacier for electronic storage along with a copy of the Cohasset Assessment. For the purposes of these requirements, AWS is not a designated third party (D3P). Be sure to select a D3P and include this information in your notification to your DEA.

Q: What other controls can be applied with Amazon Glacier Vault Lock?

In certain situations, you may be faced with the need to place a legal hold on your compliance archives for an indefinite period of time. A legal hold can be initiated on a Glacier Vault by creating a vault access policy that denies the use of Glacier's Delete functions if the vault is tagged in a particular way. In addition to time-based retention and legal hold, Glacier Vault Lock can be used to implement a variety of compliance controls which can be made immutable for strong governance, such as enforcing Multifactor Authentication on all data access/read activities to a vault with classified information.

Q: How do I set up Vault Lock?

You can set up Vault Lock in the AWS Glacier console or use the Vault Lock APIs in the AWS SDK. To learn more, please read Getting Started with Amazon Glacier Vault Lock in the Amazon Glacier developer's guide.

Data Retrievals

Q: How can I retrieve data from the service?

When you make a request to retrieve data from Glacier, you initiate a retrieval job for an archive. Once the retrieval job completes, your data will be available to download or access it using Amazon Elastic Compute Cloud (Amazon EC2) for 24 hours. There are three options for retrieving data with varying access times and cost: Expedited, Standard, and Bulk retrievals.

Q: What are Standard retrievals?

Standard retrievals allow you to access any of your archives within several hours. Standard retrievals typically complete within 3 – 5 hours.

Q: How do I use Standard retrievals?

To make a Standard retrieval, set the "Tier" parameter in the InitiateJob API request to "Standard". If no tier is specified, the request will default to Standard.

Q: How much do Standard retrievals cost?

Standard retrievals are priced at a flat rate of $0.01 per GB and $0.05 per 1,000 requests. For example, retrieving 500 archives that are 1 GB each would cost 500GB x $0.01 + 500 x $0.05/1,000 = $5.025

Q: When should I use Standard retrievals?

Standard retrievals are a low-cost way to access your data within just a few hours. For example, you can use Standard retrievals to restore backup data, retrieve archived media content for same-day editing or distribution, or pull and analyze logs to drive business decisions within hours.

Q: What are Bulk retrievals?

Bulk retrievals are Glacier's lowest-cost retrieval option, enabling you to retrieve large amounts, even petabytes, of data inexpensively in a day. Bulk retrievals typically complete within 5 – 12 hours.

Q: How do I use Bulk retrievals?

To make a Bulk retrieval, set the "Tier" parameter in the InitiateJob API request to Bulk.

Q: How much do Bulk retrievals cost?

Bulk retrievals are priced at a flat rate of just $0.0025 per GB and $0.025 per 1,000 requests. For example, retrieving 500 archives that are 1 GB each would cost 500GB x $0.0025 + 500 x $0.025/1,000 = $1.2625.

Q: When should I use Bulk retrievals?

Bulk retrievals are designed to enable customers to cost-effectively pull large amounts of data for non-urgent use cases such as transcoding petabytes of raw video content or analyzing large genomics sequences.

Q: What are Expedited retrievals?

Expedited retrievals allow you to quickly access your data when occasional urgent requests for a subset of archives are required. For all but the largest archives (250MB+), data accessed using Expedited retrievals are typically made available within 1 – 5 minutes. There are two types of Expedited retrievals: On-Demand and Provisioned. On-Demand requests are like EC2 On-Demand instances and are available the vast majority of the time. Provisioned requests are guaranteed to be available when you need them.

Q: What is a Provisioned capacity unit?

Provisioned Capacity guarantees that your retrieval capacity for Expedited retrievals will be available when you need it. Each unit of capacity ensures that at least 3 expedited retrievals can be performed every 5 minutes and provides up to 150MB/s of retrieval throughput.

Q: When should I provision retrieval capacity?

Retrieval capacity can be provisioned if you have specific Expedited retrieval rate requirements that need to be met. Without provisioned capacity, Expedited retrieval requests will be accepted if capacity is available at the time the request is made.

Q: How do I purchase provisioned capacity?

You can purchase provisioned capacity using the console, SDK, or the CLI.

Q: How much does provisioned capacity cost?

Each unit of provisioned capacity costs $100 per month from the date of purchase.

Q: How do I use Expedited retrievals?

To make an Expedited retrieval, set the “Tier” parameter in the InitiateJob API request to Expedited. There is no need to designate whether an Expedited retrieval is On-Demand or Provisioned. If you have purchased provision capacity, then all Expedited retrievals will be automatically be served via your Provisioned capacity.

Q: How much do Expedited retrievals cost?

Expedited retrievals are priced at a flat rate of $0.03 per GB and $0.01 per request. For example, retrieving 10 objects with a size of 1GB each, the cost would be 10 x $0.03 +10 x $0.01 = $0.40.

Q: When should I use Expedited retrievals?

Expedited retrievals are optimized for the occasional urgent request for a small number of archives. For all but the largest archives (250MB+), data accessed using Expedited retrievals are typically made available within 1 – 5 minutes. If your application or workload requires a guarantee that your Expedited retrievals will be available when you need it, then you should consider using Provisioned capacity.

Q: Can I retrieve part of an archive?

Yes, range retrievals enable you to retrieve a specific range of an archive. Range retrievals are similar to regular retrievals in Amazon Glacier. Both require the initiation of a retrieval job (See How can I retrieve data? for more information). You can use range retrievals to reduce or eliminate your retrieval fees (See How much data can I retrieve for free?)

When initiating a retrieval job using range retrievals, you provide a byte range that can start at zero (which would be the beginning of your archive), or at any 1MB interval thereafter (e.g. 1MB, 2MB, 3MB, etc). The end of the range can either be the end of your archive or any 1MB interval greater than the beginning of your range.

Q: Why would I retrieve only a range of an archive?

There are several reasons why you might choose to perform a range retrieval. For example, you may have aggregated several files and uploaded them as a single archive. You may then need to retrieve a small selection of those files, in which case you could retrieve only the ranges of the archive that contained the required files. Another reason you could choose to perform a range retrieval is to manage how much data you download from Amazon Glacier in a given period. When you make a request to retrieve data from Glacier, you initiate a retrieval job for an archive. Once the retrieval job completes, your data will be available to download or access using Amazon Elastic Compute Cloud (Amazon EC2) for 24 hours. The data retrieved is then available for download for 24 hours. You could therefore retrieve an archive in parts in order to manage the schedule of your downloads.

Q: How do I view my jobs?

You can list your ongoing jobs for any of your vaults by calling the ListJobs API. The list of jobs provides information including the job's creation time and date and the job's status (e.g. in-progress, completed successfully, or not in which case reasons for the job not succeeding are provided). The progress of a single job can be tracked by calling the DescribeJob API and providing the corresponding job ID. The status of the job will be returned immediately.

Q: Can I be notified when a job is completed?

Yes. You can optionally configure vaults to send notifications to you or your application when jobs complete. Notifications will be delivered via the Amazon Simple Notification Service (Click here to learn more about Amazon SNS).

Data Retrieval Policies

Q: What are data retrieval policies?

Amazon Glacier data retrieval policies let you define your own data retrieval limits with a few clicks in the AWS console. You can limit retrievals to "Free Tier Only", or if you wish to retrieve more than the free tier, you can specify a "Max Retrieval Rate" to limit your retrieval speed and establish a retrieval cost ceiling. In both cases, Amazon Glacier will not accept retrieval requests that would exceed the retrieval limits you defined. Retrieval policies apply to Standard retrievals.

To learn more please read Configuring Data Retrieval Policies in the Amazon Glacier developer's guide.

Q: How do I set up data retrieval policies?

You can set up data retrieval policies in the Amazon Glacier console or via the Amazon Glacier APIs. To learn more please read Configuring Data Retrieval Policies in the Amazon Glacier developer's guide.

Q: Are data retrieval policies specific to each AWS region?

Yes. You can set one data retrieval policy for each AWS region which will govern all data retrieval activities in the region under your account. Data retrieval policies are region-specific because data retrieval costs vary across AWS regions.

Please visit Amazon Glacier Pricing for more information.

Q: Can I use data retrieval policies to "slow down" my retrievals or spread them out?

No, data retrieval policies such as "Free Tier Only" and "Max Retrieval Rate" will not accept a data retrieval request which would exceed your predefined data retrieval limit to help you manage data retrieval cost. Data retrieval policies will not change the 3 to 5 hour data retrieval latency or spread out your retrievals. You can leverage Amazon Glacier's range retrieval feature to spread out retrievals and lower the peak retrieval speed. Learn more.

Q: What impact does the change in the retrieval free tier to 10 GB per month have on my data retrieval policy?

There is no impact to your policy in terms of the rate of data retrieval. If your retrieval policy was set to the previous free tier of 5% of your average monthly storage prior to the change in the retrieval free tier on November 21, 2016, your policy will remain the same GB-per-hour retrieval rate as your previous 5% free tier rate as of November 21, 2016. For example, if on that day your average monthly storage was 14,400 GB, your retrieval rate limit was 14,400 GB x 5% / 30 day / 24 hours = 1 GB per hour. Your new policy will remain at 1 GB per hour, but will be a "Max Retrieval Rate" rather than a "Free Tier Only" policy.

Data Inventories

Q: Can I see what archives I have stored in Amazon Glacier?

Yes. Although you will need to maintain your own index of data you upload to Amazon Glacier, an inventory of all archives in each of your vaults is maintained for disaster recovery or occasional reconciliation purposes. The vault inventory is updated approximately once a day. You can request a vault inventory as either a JSON or CSV file and will contain details about the archives within your vault including the size, creation date and the archive description (if you provided one during upload). The inventory will represent the state of the vault at the time of the most recent inventory update.

Q: Can I obtain a real time list of my vaults?

Yes, you can list your vaults stored in Amazon Glacier using either the AWS Management Console or by calling the ListVaults API. As well as a list of vault names, you will also be able to see when the vault's inventory was last updated and a summary of the vault's contents at that time, as well as the vault's creation date and creator.