Amazon S3 is object storage built to store and retrieve any amount of data from anywhere. S3 is a simple storage service that offers industry leading durability, availability, performance, security, and virtually unlimited scalability at very low costs.

Q: What can I do with Amazon S3?

Amazon S3 provides a simple web service interface that you can use to store and retrieve any amount of data, at any time, from anywhere. Using this service, you can easily build applications that make use of cloud native storage. Since Amazon S3 is highly scalable and you only pay for what you use, you can start small and grow your application as you wish, with no compromise on performance or reliability.

Amazon S3 is also designed to be highly flexible. Store any type and amount of data that you want, read the same piece of data a million times or only for emergency disaster recovery, build a simple FTP application or a sophisticated web application such as the Amazon.com retail web site. Amazon S3 frees you to focus on innovation instead of spending time figuring out how to store your data.

Q: How can I get started using Amazon S3?

To sign up for Amazon S3, visit the S3 console. You must have an Amazon Web Services account to access this service. If you do not already have an account, you will be prompted to create one when you begin the Amazon S3 sign-up process. After signing up, refer to the Amazon S3 documentation, view the S3 getting started materials, and see the additional resources in the resource center to begin using Amazon S3.

Q: What can I do with Amazon S3 that I cannot do with an on-premises solution?

Amazon S3 lets you leverage Amazon’s own benefits of massive scale with no up-front investment or performance compromises. By using Amazon S3, it is inexpensive and simple to ensure your data is quickly accessible, always available, and secure.

Q: What kind of data can I store in Amazon S3?

You can store virtually any kind of data in any format. Refer to the Amazon Web Services Licensing Agreement for details.

Q: How much data can I store in Amazon S3?

The total volume of data and number of objects you can store in Amazon S3 are unlimited. Individual Amazon S3 objects can range in size from a minimum of 0 bytes to a maximum of 5 TB. The largest object that can be uploaded in a single PUT is 5 GB. For objects larger than 100 MB, customers should consider using the multipart upload capability.

Q: What is an S3 general purpose bucket?

A bucket is a container for objects stored in Amazon S3, and you can store any number of objects in a bucket. General purpose buckets are the original S3 bucket type, and a single general purpose bucket can contain objects stored across all storage classes except S3 Express One Zone. They are recommended for most use cases and access patterns.

Q: What is an S3 directory bucket?

A bucket is a container for objects stored in Amazon S3, and you can store any number of objects in a bucket. S3 directory buckets only allow objects stored in the S3 Express One Zone storage class, which provides faster data processing within a single Availability Zone. They are recommended for low-latency use cases. Each S3 directory bucket can support hundreds of thousands of transactions per second (TPS), independent of the number of directories within the bucket.

Q: What is the difference between a general purpose bucket and a directory bucket?

A bucket is a container for objects stored in Amazon S3, and you can store any number of objects in a bucket. General purpose buckets are the original S3 bucket type, and a single general purpose bucket can contain objects stored across all storage classes except S3 Express One Zone. They are recommended for most use cases and access patterns. S3 directory buckets only allow objects stored in the S3 Express One Zone storage class, which provides faster data processing within a single Availability Zone. They are recommended for low-latency use cases. Each S3 directory bucket can support hundreds of thousands of transactions per second (TPS), independent of the number of directories within the bucket.

Q: What does Amazon do with my data in Amazon S3?

Amazon stores your data and tracks its associated usage for billing purposes. Amazon will not otherwise access your data for any purpose outside of the Amazon S3 offering, except when required to do so by law. Refer to the Amazon Web Services Licensing Agreement for details.

Q: Does Amazon store its own data in Amazon S3?

Yes. Organizations across Amazon use Amazon S3 for a wide variety of projects. Many of these projects use Amazon S3 as their authoritative data store and rely on it for business-critical operations.

Q: How is Amazon S3 data organized?

Amazon S3 is a simple key-based object store. When you store data, you assign a unique object key that can later be used to retrieve the data. Keys can be any string, and they can be constructed to mimic hierarchical attributes. Alternatively, you can use S3 Object Tagging to organize your data across all of your S3 buckets and/or prefixes.

Q: How do I interface with Amazon S3?

Amazon S3 provides a simple, standards-based REST web services interface that is designed to work with any internet-development toolkit. The operations are intentionally made simple to make it easy to add new distribution protocols and functional layers.

Q: How reliable is Amazon S3?

Amazon S3 gives you access to the same highly scalable, highly available, fast, inexpensive data storage infrastructure that Amazon uses to run its own global network of web sites. The S3 Standard storage class is designed for 99.99% availability, the S3 Standard-IA storage class, S3 Intelligent-Tiering storage class, and the S3 Glacier Instant Retrieval storage classes are designed for 99.9% availability, the S3 One Zone-IA storage class is designed for 99.5% availability, and the S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive class are designed for 99.99% availability and an SLA of 99.9%. All of these storage classes are backed by the Amazon S3 Service Level Agreement.

Q: How will Amazon S3 perform if traffic from my application suddenly spikes?

Amazon S3 is designed from the ground up to handle traffic for any internet application. Pay-as-you-go pricing and unlimited capacity ensures that your incremental costs don’t change and that your service is not interrupted. Amazon S3’s massive scale lets you spread the load evenly, so that no individual application is affected by traffic spikes.

Q: Does Amazon S3 offer a Service Level Agreement (SLA)?

Yes. The Amazon S3 SLA provides for a service credit if a customer's monthly uptime percentage is below our service commitment in any billing cycle.

Q: What is the consistency model for Amazon S3?

Amazon S3 delivers strong read-after-write consistency automatically, without changes to performance or availability, without sacrificing regional isolation for applications, and at no additional cost.

After a successful write of a new object or an overwrite of an existing object, any subsequent read request immediately receives the latest version of the object. S3 also provides strong consistency for list operations, so after a write, you can immediately perform a listing of the objects in a bucket with any changes reflected.

Q: Why does strong read-after-write consistency help me?

Strong read-after-write consistency helps when you need to immediately read an object after a write; for example, when you often read and list immediately after writing objects. High-performance computing workloads also benefit in that when an object is overwritten and then read many times simultaneously, strong read-after-write consistency provides assurance that the latest write is read across all reads. These applications automatically and immediately benefit from strong read-after-write consistency. The strong consistency of S3 also reduces costs by removing the need for extra infrastructure to provide strong consistency.

AWS Regions

Q: Where is my data stored?

You specify an AWS Region when you create your Amazon S3 bucket. For S3 Standard, S3 Standard-IA, S3 Intelligent-Tiering, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive storage classes, your objects are automatically stored across multiple devices spanning a minimum of three Availability Zones (AZs). AZs are physically separated by a meaningful distance, many kilometers, from any other AZ, although all are within 100 km (60 miles) of each other. Objects stored in the S3 One Zone-IA storage class are stored redundantly within a single Availability Zone in the AWS Region you select. For S3 on Outposts, your data is stored in your Outpost on-premises environment, unless you manually choose to transfer it to an AWS Region. Refer to AWS regional services list for details of Amazon S3 service availability by AWS Region.

Q: What is an AWS Region?

An AWS Region is a physical location around the world where AWS cluster data centers. Each group of logical data centers within a Region is know as an Availability Zone (AZ). Each AWS Region consists of a minimum of three, isolated, and physically separate AZs within a geographic area. Unlike other cloud providers, who often define a Region as a single data center, the multiple AZ design of every AWS Region offers advantages for customers. Each AZ has independent power, cooling, and physical security and is connected via redundant, ultra-low-latency networks.

Q: What is an AWS Availability Zone (AZ)?

An Availability Zone (AZ) is one or more discrete data centers with redundant power, networking, and connectivity in an AWS Region. AZs give customers the ability to operate production applications and databases that are more highly available, fault tolerant, and scalable than would be possible from a single data center. All AZs in an AWS Region are interconnected with high-bandwidth, low-latency networking, over fully redundant, dedicated metro fiber providing high-throughput, low-latency networking between AZs.

Amazon S3 Standard, S3 Standard-Infrequent Access, S3 Intelligent-Tiering, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive storage classes replicate data across a minimum of three AZs to protect against the loss of one entire AZ. This remains true in Regions where fewer than three AZs are publicly available. Objects stored in these storage classes are available for access from all of the AZs in an AWS Region.

The Amazon S3 One Zone-IA storage class replicates data within a single AZ. The data stored in S3 One Zone-IA is not resilient to the physical loss of an Availability Zone resulting from disasters, such as earthquakes, fires, and floods.

Q: How do I decide which AWS Region to store my data in?

There are several factors to consider based on your specific application. For instance, you may want to store your data in a Region that is near your customers, your data centers, or other AWS resources to reduce data access latencies. You may also want to store your data in a Region that is remote from your other operations for geographic redundancy and disaster recovery purposes. You should also consider Regions that let you address specific legal and regulatory requirements and/or reduce your storage costs—you can choose a lower priced Region to save money. For S3 pricing information, visit the Amazon S3 pricing page.

Q: In which parts of the world is Amazon S3 available?

Amazon S3 is available in AWS Regions worldwide, and you can use Amazon S3 regardless of your location. You just have to decide which AWS Region(s) you want to store your Amazon S3 data. See the AWS regional services list for a list of AWS Regions in which S3 is available today.

Billing

Q: How much does Amazon S3 cost?

With Amazon S3, you pay only for what you use. There is no minimum charge. You can estimate your monthly bill using the AWS Pricing Calculator.

AWS charges less where our costs are less. Some prices vary across Amazon S3 Regions. Billing prices are based on the location of your S3 bucket. There is no Data Transfer charge for data transferred within an Amazon S3 Region via a COPY request. Data transferred via a COPY request between AWS Regions is charged at rates specified on the Amazon S3 pricing page. There is no Data Transfer charge for data transferred between Amazon EC2 (or any AWS service) and Amazon S3 within the same Region, for example, data transferred within the US

East (Northern Virginia) Region. However, data transferred between Amazon EC2 (or any AWS service) and Amazon S3 across all other Regions is charged at rates specified on the Amazon S3 pricing page, for example, data transferred between Amazon EC2 US East (Northern Virginia) and Amazon S3 US West (Northern California). Data transfer costs are billed to the source bucket owner.

For S3 on Outposts pricing, visit the Outposts pricing page.

Q: How will I be charged and billed for my use of Amazon S3?

There are no set up charges or commitments to begin using Amazon S3. At the end of the month, you will automatically be charged for that month’s usage. You can view your charges for the current billing period at any time by logging into your Amazon Web Services account, and selecting the 'Billing Dashboard' associated with your console profile.

With the AWS Free Usage Tier*, you can get started with Amazon S3 for free in all Regions except the AWS GovCloud Regions. Upon sign up, new AWS customers receive 5 GB of Amazon S3 Standard storage, 20,000 Get Requests, 2,000 Put Requests, and 100 GB of data transfer out (to internet, other AWS Regions, or Amazon CloudFront) each month for one year. Unused monthly usage will not roll over to the next month.

Amazon S3 charges you for the following types of usage. Note that the calculations below assume there is no AWS Free Tier in place.

Storage Used:

Amazon S3 storage pricing is summarized on the Amazon S3 pricing page.

The volume of storage billed in a month is based on the average storage used throughout the month. This includes all object data and metadata stored in buckets that you created under your AWS account. We measure your storage usage in “TimedStorage-ByteHrs,” which are added up at the end of the month to generate your monthly charges.

Storage Example:

Assume you store 100 GB (107,374,182,400 bytes) of data in Amazon S3 Standard in your bucket for 15 days in March, and 100 TB (109,951,162,777,600 bytes) of data in Amazon S3 Standard for the final 16 days in March.

At the end of March, you would have the following usage in Byte-Hours: Total Byte-Hour usage = [107,374,182,400 bytes x 15 days x (24 hours / day)] + [109,951,162,777,600 bytes x 16 days x (24 hours / day)] = 42,259,901,212,262,400 Byte-Hours. Calculate hours based on the actual number of days in a given month. For example, in our example we are using March which has 31 days or 744 hours.

Let's convert this to GB-Months: 42,259,901,212,262,400 Byte-Hours / 1,073,741,824 bytes per GB / 744 hours per month = 52,900 GB-Months

This usage volume crosses two different volume tiers. The monthly storage price is calculated below assuming the data is stored in the US East (Northern Virginia) Region: 50 TB Tier: 51,200 GB x $0.023 = $1,177.60 50 TB to 450 TB Tier: 1,700 GB x $0.022 = $37.40

Total Storage cost = $1,177.60 + $37.40 = $1,215.00

Network Data Transferred In:

Amazon S3 Data Transfer In pricing is summarized on the Amazon S3 pricing page. This represents the amount of data sent to your Amazon S3 buckets.

Network Data Transferred Out:

Amazon S3 Data Transfer Out pricing is summarized on the Amazon S3 pricing page. For Amazon S3, this charge applies whenever data is read from any of your buckets from a location outside of the given Amazon S3 Region.

Data Transfer Out pricing rate tiers take into account your aggregate Data Transfer Out from a given Region to the internet across Amazon EC2, Amazon S3, Amazon RDS, Amazon SimpleDB, Amazon SQS, Amazon SNS, and Amazon VPC. These tiers do not apply to Data Transfer Out from Amazon S3 in one AWS Region to another AWS Region.

Data Transfer Out Example:
Assume you transfer 1 TB of data out of Amazon S3 from the US East (Northern Virginia) Region to the internet every day for a given 31-day

month. Assume you also transfer 1 TB of data out of an Amazon EC2 instance from the same Region to the internet over the same 31-day month.

Your aggregate Data Transfer would be 62 TB (31 TB from Amazon S3 and 31 TB from Amazon EC2). This equates to 63,488 GB (62 TB * 1024 GB/TB).

This usage volume crosses three different volume tiers. The monthly Data Transfer Out charge is calculated below assuming the Data Transfer occurs in the US East (Northern Virginia) Region:
10 TB Tier: 10,239 GB (10×1024 GB/TB – 1 (free)) x $0.09 = $921.51
10 TB to 50 TB Tier: 40,960 GB (40×1024) x $0.085 = $3,481.60
50 TB to 150 TB Tier: 12,288 GB (remainder) x $0.070 = $860.16

Total Data Transfer Out charge = $921.51+ $3,481.60 + $860.16= $5,263.27

Data Requests:

Amazon S3 Request pricing is summarized on the Amazon S3 pricing page.

Request Example:
Assume you transfer 10,000 files into Amazon S3 and transfer 20,000 files out of Amazon S3 each day during the month of March. Then, you delete 5,000 files on March 31st.
Total PUT requests = 10,000 requests x 31 days = 310,000 requests
Total GET requests = 20,000 requests x 31 days = 620,000 requests
Total DELETE requests = 5,000×1 day = 5,000 requests

Assuming your bucket is in the US East (Northern Virginia) Region, the Request charges are calculated below:
310,000 PUT Requests: 310,000 requests x $0.005/1,000 = $1.55
620,000 GET Requests: 620,000 requests x $0.004/10,000 = $0.25
5,000 DELETE requests = 5,000 requests x $0.00 (no charge) = $0.00

Data Retrieval:

Amazon S3 data retrieval pricing applies for the S3 Standard-Infrequent Access (S3 Standard-IA) and S3 One Zone-IA storage classes and is summarized on the Amazon S3 pricing page.

Data Retrieval Example:
Assume in one month you retrieve 300 GB of S3 Standard-IA, with 100 GB going out to the internet, 100 GB going to EC2 in the same AWS Region, and 100 GB going to Amazon CloudFront in the same AWS Region.

Your data retrieval charges for the month would be calculated as 300 GB x $0.01/GB = $3.00. Note that you would also pay network data transfer charges for the portion that went out to the internet.

* * Your usage for the free tier is calculated each month across all Regions except the AWS GovCloud Region and automatically applied to your bill—unused monthly usage will not roll over. Restrictions apply. See offer terms for more details.

Q: Why do prices vary depending on which Amazon S3 Region I choose?

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

Q: How am I charged for using Versioning?

Normal Amazon S3 rates apply for every version of an object stored or requested. For example, let's look at the following scenario to illustrate storage costs when utilizing Versioning (let's assume the current month is 31 days long):

1) Day 1 of the month: You perform a PUT of 4 GB (4,294,967,296 bytes) on your bucket.
2) Day 16 of the month: You perform a PUT of 5 GB (5,368,709,120 bytes) within the same bucket using the same key as the original PUT on Day 1.

When analyzing the storage costs of the above operations, note that the 4 GB object from Day 1 is not deleted from the bucket when the 5 GB object is written on Day 15. Instead, the 4 GB object is preserved as an older version and the 5 GB object becomes the most recently written version of the object within your bucket. At the end of the month:

Total Byte-Hour usage
[4,294,967,296 bytes x 31 days x (24 hours / day)] + [5,368,709,120 bytes x 16 days x (24 hours / day)] = 5,257,039,970,304 Byte-Hours.

Conversion to Total GB-Months
5,257,039,970,304 Byte-Hours x (1 GB / 1,073,741,824 bytes) x (1 month / 744 hours) = 6.581 GB-Month

The cost is calculated based on the current rates for your Region on the Amazon S3 pricing page.

Q: How am I charged for accessing Amazon S3 through the AWS Management Console?

Normal Amazon S3 pricing applies when accessing the service through the AWS Management Console. To provide an optimized experience, the AWS Management Console may proactively execute requests. Also, some interactive operations result in more than one request to the service.

Q: How am I charged if my Amazon S3 buckets are accessed from another AWS account?

Normal Amazon S3 pricing applies when your storage is accessed by another AWS Account. Alternatively, you may choose to configure your bucket as a Requester Pays bucket, in which case the requester will pay the cost of requests and downloads of your Amazon S3 data.

You can find more information on Requester Pays bucket configurations in the Amazon S3 documentation.

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 about taxes on AWS services »]

Q: Will I incur any data transfer out to the internet charges when I move my data out of AWS?

AWS offers eligible customers free data transfer out to the internet when they move all of their data off of AWS, in accordance with the process below.

Q: I want to move my data out of AWS. How do I request free data transfer out to the internet?

Complete the following steps:

1) If you have a dedicated AWS account team, contact them first and inform them of your plans. In some cases, if you have a negotiated commitment with AWS, you'll want to discuss your options with your AWS account team.

2) Review the criteria and process described on this page.

3) Contact [AWS Customer Support] and indicate that your request is for “free data transfer to move off AWS.” AWS Customer Support will ask that you provide information, so they can review your moving plans, evaluate whether you qualify for free data transfer out, and calculate a proper credit amount.

4) If AWS Customer Support approves your move, you will receive a temporary credit for the cost of data transfer out based on the volume of all data you have stored across AWS services at the time of AWS’ calculation. AWS Customer Support will notify you if you are approved, and you will then have 60 days to complete your move off of AWS. The credit will count against data transfer out usage only, and it will not be applied to other service usage. After your move away from AWS services, within the 60-day period, you must delete all remaining data and workloads from your AWS account, or you can close your AWS account.

Free data transfers for moving IT providers are also subject to the following criteria:

a) Only customers with an active AWS account in good standing are eligible for free data transfer out.

b) If you have less than 100 GB of data stored in your AWS account you may move this data off of AWS for free under AWS's existing 100 GB monthly free tier for data transfer out. Customers with less than 100 GB of data stored in their AWS account are not eligible for additional credits.

c) AWS will provide you with free data transfer out to the internet when you move all of your data off of AWS. If you only want to move your total usage of a single service, but not everything, contact AWS Customer Support.

d) If your plans change, or you cannot complete your move off of AWS within 60 days, you must notify AWS Customer Support.

e) Standard services charges for use of AWS services are not included. Only data transfer out charges in support of your move off of AWS are eligible for credits. However, data transfer out from specialized data transfer services, such as Amazon CloudFront, AWS Direct Connect, AWS Snow Family, and AWS Global Accelerator, are not included.

f) AWS may review your service usage to verify compliance with these requirements. If we determine your use of data transfer out was for a purpose other than moving off of AWS, we may charge you for the data transfer out that had been credited.

g) AWS may make changes with respect to free data transfers out to the internet at any time.

Q: Why do I have to request AWS' pre-approval for free data transfer out to the internet before moving my data out of AWS?

AWS customers make hundreds of millions of data transfers each day, and we generally don't know the reason for any given data transfer. For example, customers may be transferring data to an end user of their application, to a visitor of their website, or to another cloud or on-premises environment for backup purposes. Accordingly, the only way we know that your data transfer is to support your move off of AWS is if you tell us beforehand.

Amazon S3 and IPv6

Q: What is IPv6?

Every server and device connected to the internet must have a unique address. Internet Protocol Version 4 (IPv4) was the original 32-bit addressing scheme. However, the continued growth of the internet means that all available IPv4 addresses will be utilized over time. Internet Protocol Version 6 (IPv6) is an addressing mechanism designed to overcome the global address limitation on IPv4.

Q: What can I do with IPv6?

Using IPv6 support for Amazon S3, applications can connect to Amazon S3 without the need for any IPv6 to IPv4 translation software or systems. You can meet compliance requirements, more easily integrate with existing IPv6-based on-premises applications, and remove the need for expensive networking equipment to handle the address translation. You can also now utilize the existing source address filtering features in IAM policies and bucket policies with IPv6 addresses, expanding your options to secure applications interacting with Amazon S3.

Q: How do I get started with IPv6 on Amazon S3?

You can get started by pointing your application to Amazon S3's "dual-stack" endpoint, which supports access over both IPv4 and IPv6. In most cases, no further configuration is required for access over IPv6, because most network clients prefer IPv6 addresses by default. Applications that are impacted by using IPv6 can switch back to the standard IPv4-only endpoints at any time. IPv6 with Amazon S3 is supported in all commercial AWS Regions, including AWS GovCloud (US) Regions, the Amazon Web Services China (Beijing) Region, operated by Sinnet, and the Amazon Web Services China (Ningxia) Region, operated by NWCD.

Q: Should I expect a change in Amazon S3 performance when using IPv6?

No, you will see the same performance when using either IPv4 or IPv6 with Amazon S3.

S3 Event Notifications

Q: What are Amazon S3 Event Notifications?

You can use the Amazon S3 Event Notifications feature to receive notifications when certain events happen in your S3 bucket, such as PUT, POST, COPY, and DELETE events. You can publish notifications to Amazon EventBridge, Amazon SNS, Amazon SQS, or directly to AWS Lambda.

Q: What can I do with Amazon S3 Event Notifications?

Amazon S3 Event Notifications let you run workflows, send alerts, or perform other actions in response to changes in your objects stored in S3. You can use S3 Event Notifications to set up triggers to perform actions including transcoding media files when they are uploaded, processing data files when they become available, and synchronizing S3 objects with other data stores. You can also set up event notifications based on object name prefixes and suffixes. For example, you can choose to receive notifications on object names that start with “images/."

Q: What is included in Amazon S3 Event Notifications?

For a detailed description of the information included in Amazon S3 Event Notification messages, refer to the configuring Amazon S3 Event Notifications documentation.

Q: How do I set up Amazon S3 Event Notifications?

For a detailed description of how to configure event notifications, refer to the configuring Amazon S3 Event Notifications documentation. You can learn more about AWS messaging services in the Amazon SNS documentation and the Amazon SQS documentation.

Q: What does it cost to use Amazon S3 Event Notifications?

There are no additional charges for using Amazon S3 for event notifications. You pay only for use of Amazon SNS or Amazon SQS to deliver event notifications, or for the cost of running an AWS Lambda function. Visit the Amazon SNS, Amazon SQS, or AWS Lambda pricing pages to view the pricing details for these services.

Amazon S3 Transfer Acceleration

Q: What is S3 Transfer Acceleration?

Amazon S3 Transfer Acceleration creates fast, easy, and secure transfers of files over long distances between your client and your Amazon S3 bucket. S3 Transfer Acceleration leverages Amazon CloudFront’s globally distributed AWS Edge locations. As data arrives at an AWS Edge Location, data is routed to your Amazon S3 bucket over an optimized network path.

Q: How do I get started with S3 Transfer Acceleration?

To get started with S3 Transfer Acceleration enable S3 Transfer Acceleration on an S3 bucket using the Amazon S3 console, the Amazon S3 API, or the AWS CLI. After S3 Transfer Acceleration is enabled, you can point your Amazon S3 PUT and GET requests to the s3-accelerate endpoint domain name. Your data transfer application must use one of the following two types of endpoints to access the bucket for faster data transfer: .s3-accelerate.amazonaws.com or .s3-accelerate.dualstack.amazonaws.com for the “dual-stack” endpoint. If you want to use standard data transfer, you can continue to use the regular endpoints.

There are certain restrictions on which buckets will support S3 Transfer Acceleration. For details, refer to the Amazon S3 documentation.

Q: How fast is S3 Transfer Acceleration?

S3 Transfer Acceleration helps you fully use your bandwidth, minimize the effect of distance on throughput, and is designed to ensure consistently fast data transfer to Amazon S3 regardless of your client’s location. The amount of acceleration primarily depends on your available bandwidth, the distance between the source and destination, and packet loss rates on the network path. Generally, you will see more acceleration when the source is farther from the destination, when there is more available bandwidth, and/or when the object size is bigger.

One customer measured a 50% reduction in their average time to ingest 300 MB files from a global user base spread across the US, Europe, and parts of Asia to a bucket in the Asia Pacific (Sydney) Region. Another customer observed cases where performance improved in excess of 500% for users in South East Asia and Australia uploading 250 MB files (in parts of 50 MB) to an S3 bucket in the US East (N. Virginia) Region.

Access the S3 Transfer Acceleration speed comparison tool to get a preview of the performance benefit from your location.

Q: Who should use S3 Transfer Acceleration?

S3 Transfer Acceleration is designed to optimize transfer speeds from across the world into S3 buckets. If you are uploading to a centralized bucket from geographically dispersed locations or if you regularly transfer GBs or TBs of data across continents, you may save hours or days of data transfer time with S3 Transfer Acceleration.

Q: How secure is S3 Transfer Acceleration?

S3 Transfer Acceleration provides the same security as regular transfers to Amazon S3. All Amazon S3 security features, such as access restriction based on a client’s IP address, are supported as well. S3 Transfer Acceleration communicates with clients over standard TCP and does not require firewall changes. No data is ever saved at AWS Edge locations.

Q: What if S3 Transfer Acceleration is not faster than a regular Amazon S3 transfer?

Each time you use S3 Transfer Acceleration to upload an object, we will check whether S3 Transfer Acceleration is likely to be faster than a regular Amazon S3 transfer. If we determine that S3 Transfer Acceleration is not likely to be faster than a regular Amazon S3 transfer of the same object to the same destination AWS Region, we will not charge for the use of S3 Transfer Acceleration for that transfer, and we may bypass the S3 Transfer Acceleration system for that upload.

Q: Can I use S3 Transfer Acceleration with multipart uploads?

Yes, S3 Transfer Acceleration supports all bucket level features including multipart uploads.

Q: How should I choose between S3 Transfer Acceleration and Amazon CloudFront’s PUT/POST?

S3 Transfer Acceleration optimizes the TCP protocol and adds additional intelligence between the client and the S3 bucket, making S3 Transfer Acceleration a better choice if a higher throughput is desired. If you have objects that are smaller than 1 GB or if the data set is less than 1 GB in size, you should consider using Amazon CloudFront's PUT/POST commands for optimal performance.

Q: How should I choose between S3 Transfer Acceleration and AWS Snow Family?

The AWS Snow Family is ideal for customers moving large batches of data at once. The AWS Snowball has a typical 5—7 days turnaround time. As a rule of thumb, S3 Transfer Acceleration over a fully-utilized 1 Gbps line can transfer up to 75 TBs in the same time period. In general, if it will take more than a week to transfer over the internet, or there are recurring transfer jobs and there is more than 25Mbps of available bandwidth, S3 Transfer Acceleration is a good option. Another option is to use both: perform initial heavy lift moves with an AWS Snowball (or series of AWS Snowballs) and then transfer incremental ongoing changes with S3 Transfer Acceleration.

Q: Can S3 Transfer Acceleration complement AWS Direct Connect?

AWS Direct Connect is a good choice for customers who have a private networking requirement or who have access to AWS Direct Connect exchanges. S3 Transfer Acceleration is best for submitting data from distributed client locations over the public internet, or where variable network conditions make throughput poor. Some AWS Direct Connect customers use S3 Transfer Acceleration to help with remote office transfers where they may suffer from poor internet performance.

Q: Can S3 Transfer Acceleration complement AWS Storage Gateway or a third-party gateway?

You can benefit from configuring the bucket destination in your third-party gateway to use an S3 Transfer Acceleration endpoint domain.

Visit this File section of the Storage Gateway FAQ to learn more about the AWS implementation.

Q: Can S3 Transfer Acceleration complement third-party integrated software?

Yes. Software packages that connect directly into Amazon S3 can take advantage of S3 Transfer Acceleration when they send their jobs to Amazon S3.

Learn more about Storage Partner Solutions »

Q: Is S3 Transfer Acceleration HIPAA eligible?

Yes, AWS has expanded its HIPAA compliance program to include S3 Transfer Acceleration as a HIPAA eligible service. If you have an executed Business Associate Agreement (BAA) with AWS, you can use S3 Transfer Acceleration to make fast, easy, and secure transfers of files, including protected health information (PHI) over long distances between your client and your Amazon S3 bucket.

Security

S3 Access Grants | S3 Access Points

Q: How secure is my data in Amazon S3?

Amazon S3 is secure by default. Upon creation, only you have access to Amazon S3 buckets that you create, and you have complete control over who has access to your data. Amazon S3 supports user authentication to control access to data. You can use access control mechanisms, such as bucket policies, to selectively grant permissions to users and groups of users. The Amazon S3 console highlights your publicly accessible buckets, indicates the source of public accessibility, and also warns you if changes to your bucket policies or bucket ACLs would make your bucket publicly accessible. You should enable Amazon S3 Block Public Access for all accounts and buckets that you do not want publicly accessible. All new buckets have Block Public Access turned on by default.

You can securely upload/download your data to Amazon S3 via SSL endpoints using the HTTPS protocol. Amazon S3 automatically encrypts all object uploads to your bucket (as of January 5, 2023). Alternatively, you can use your own encryption libraries to encrypt data before storing it in Amazon S3.

For more information on security in AWS, refer to the AWS security page, and for S3 security information, visit the S3 security page and the S3 security best practices guide.

Q: How can I control access to my data stored on Amazon S3?

Customers can use a number of mechanisms for controlling access to Amazon S3 resources, including AWS Identity and Access Management (IAM) policies, bucket policies, access point policies, access control lists (ACLs), Query String Authentication, Amazon Virtual Private Cloud (Amazon VPC) endpoint policies, service control policies (SCPs) in AWS Organizations, and Amazon S3 Block Public Access.

IAM

IAM lets organizations with multiple employees create and manage multiple users under a single AWS account. With IAM policies, customers can grant IAM users fine-grained control to their Amazon S3 bucket or objects while also retaining full control over everything the users do.

Bucket and access point policies

With bucket policies and access point policies, customers can define rules which apply broadly across all requests to their Amazon S3 resources, such as granting write privileges to a subset of Amazon S3 resources. Customers can also restrict access based on an aspect of the request, such as HTTP referrer and IP address.

ACLs

Amazon S3 supports S3's original access control method, access control lists (ACLs). With ACLs, customers can grant specific permissions (i.e. READ, WRITE, FULL_CONTROL) to specific users for an individual bucket or object. For customers who prefer to use exclusively policies for access control, Amazon S3 offers the S3 Object Ownership feature to disable ACLs. You can use S3 Inventory to review ACLs usage in your buckets before enabling S3 Object Ownership when migrating to IAM-based buckets policies.

Query String Authentication

With Query String Authentication, customers can create a URL to an Amazon S3 object which is only valid for a limited time. For more information on the various access control policies available in Amazon S3, refer to the access control documentation.

Amazon VPC

When customers create an Amazon VPC endpoint, they can attach an endpoint policy to it that controls access to the Amazon S3 resources to which they are connecting. Customers can also use Amazon S3 bucket policies to control access to buckets from specific endpoints or specific VPCs.

Service control policies

Service control policies (SCPs) are a type of AWS Organizations policy that customers can use to manage permissions in their organization. SCPs offer central control over the maximum available permissions for all accounts in an organization. With SCPs, customers can ensure their accounts stay within the organization’s access control guidelines.

S3 Block Public Access

Amazon S3 Block Public Access provides settings for access points, buckets, and accounts to help customers manage public access to Amazon S3 resources. With S3 Block Public Access, account administrators and bucket owners can easily set up centralized controls to limit public access to their Amazon S3 resources that are enforced regardless of how the resources are created. All new buckets have Block Public Access turned on by default as a security best practice.

Learn more about policies and permissions in the AWS IAM documentation.

Q: Does Amazon S3 support data access auditing?

Yes, customers can optionally configure an Amazon S3 bucket to create access log records for all requests made against it. Alternatively, customers who need to capture IAM/user identity information in their logs can configure AWS CloudTrail Data Events.

These access log records can be used for audit purposes and contain details about the request, such as the request type, the resources specified in the request, and the time and date the request was processed.

Q: What options do I have for encrypting data stored on Amazon S3?

Amazon S3 encrypts all new data uploads to any bucket. Amazon S3 applies S3-managed server-side encryption (SSE-S3) as the base level of encryption to all object uploads (as of January 5, 2023). SSE-S3 provides a fully-managed solution where Amazon handles key management and key protection using multiple layers of security. You should continue to use SSE-S3 if you prefer to have Amazon manage your keys. Additionally, you can choose to encrypt data using SSE-C, SSE-KMS, DSSE-KMS, or a client library such as the Amazon S3 Encryption Client. Each option allows you to store sensitive data encrypted at rest in Amazon S3.

SSE-C allows Amazon S3 to perform encryption and decryption of objects, while you retain control of the encryption keys. With SSE-C, you don’t need to implement or use a client-side library to perform the encryption and decryption of objects you store in Amazon S3, but you do need to manage the keys that you send to Amazon S3 to encrypt and decrypt objects. Use SSE-C if you want to maintain your own encryption keys, but don’t want to implement or leverage a client-side encryption library.

SSE-KMS lets AWS Key Management Service (AWS KMS) manage your encryption keys. Using AWS KMS to manage your keys provides several additional benefits. With AWS KMS, there are separate permissions for the use of the KMS key, providing an additional layer of control and protection against unauthorized access to your objects stored in Amazon S3. AWS KMS provides an audit trail so you can see who used your key to access which object and when, as well as view failed attempts to access data from users without permission to decrypt the data. Also, AWS KMS provides additional security controls to support customer efforts to comply with PCI-DSS, HIPAA/HITECH, and FedRAMP industry requirements.

DSSE-KMS simplifies the process of applying two layers of encryption to your data, without having to invest in infrastructure required for client-side encryption. Each layer of encryption uses a different implementation of the 256-bit Advanced Encryption Standard with Galois Counter Mode (AES-GCM) algorithm and is vetted and accepted for use on top-secret workloads. DSSE-KMS uses AWS KMS to generate data keys, and lets AWS KMS manage your encryption keys. With AWS KMS, there are separate permissions for the use of the KMS key, providing an additional layer of control and protection against unauthorized access to your objects stored in Amazon S3. AWS KMS provides an audit trail so you can see who used your key to access which object and when, as well as view failed attempts to access data from users without permission to decrypt the data. Also, AWS KMS provides additional security controls to support customer efforts to comply with PCI-DSS, HIPAA/HITECH, and FedRAMP industry requirements.

Using an encryption client library, you retain control of the keys and complete the encryption and decryption of objects client-side using an encryption library of your choice. Some customers prefer full end-to-end control of the encryption and decryption of objects; that way, only encrypted objects are transmitted over the internet to Amazon S3. Use a client-side library if you want to maintain control of your encryption keys, are able to implement or use a client-side encryption library, and need to have your objects encrypted before they are sent to Amazon S3 for storage.

For more information on using Amazon S3 SSE-S3, SSE-C, or SSE-KMS, refer to protecting data using encryption documentation.

Q: Can I comply with European data privacy regulations using Amazon S3?

Customers can choose to store all data in Europe by using the Europe (Frankfurt), Europe (Ireland), Europe (Paris), Europe (Stockholm), Europe (Milan), Europe (Spain), Europe (London), or Europe (Zurich) Region. You can also use Amazon S3 on Outposts to keep all of your data on premises on the AWS Outpost, and you may choose to transfer data between AWS Outposts or to an AWS Region. It is your responsibility to ensure that you comply with European privacy laws. View the AWS General Data Protection Regulation (GDPR) Center and AWS Data Privacy Center for more information. If you have more specific location requirements or other data privacy regulations that require you to keep data in a location where there is not an AWS Region, you can use S3 on Outposts.

Q: What is an Amazon VPC Endpoint for Amazon S3?

An Amazon VPC Endpoint for Amazon S3 is a logical entity within a VPC that allows connectivity to S3 over the AWS global network. There are two types of VPC endpoints for S3: gateway VPC endpoints and interface VPC endpoints. Gateway endpoints are a gateway that you specify in your route table to access S3 from your VPC over the AWS network. Interface endpoints extend the functionality of gateway endpoints by using private IPs to route requests to S3 from within your VPC, on-premises, or from a different AWS Region. For more information, visit the AWS PrivateLink for Amazon S3 documentation.

Q: Can I allow a specific Amazon VPC Endpoint access to my Amazon S3 bucket?

You can limit access to your bucket from a specific Amazon VPC Endpoint or a set of endpoints using Amazon S3 bucket policies. S3 bucket policies now support a condition, aws:sourceVpce, that you can use to restrict access. For more details and example policies, read the gateway endpoints for S3 documentation.

Q: What is AWS PrivateLink for Amazon S3?

AWS PrivateLink for S3 provides private connectivity between Amazon S3 and on-premises. You can provision interface VPC endpoints for S3 in your VPC to connect your on-premises applications directly to S3 over AWS Direct Connect or AWS VPN. You no longer need to use public IPs, change firewall rules, or configure an internet gateway to access S3 from on-premises. To learn more visit the AWS PrivateLink for S3 documentation.

Q: How do I get started with interface VPC endpoints for S3?

You can create an interface VPC endpoint using the AWS VPC Management Console, AWS Command Line Interface (AWS CLI), AWS SDK, or API. To learn more, visit the documentation.

Q: When should I choose gateway VPC endpoints versus AWS PrivateLink-based interface VPC endpoints?

AWS recommends that you use interface VPC endpoints to access S3 from on-premises or from a VPC in another AWS Region. For resources that are accessing S3 from VPC in the same AWS Region as S3, we recommend using gateway VPC endpoints as they are not billed. To learn more, visit the documentation.

Q: Can I use both Interface Endpoints and Gateway Endpoints for S3 in the same VPC?

Yes. If you have an existing gateway VPC endpoint, create an interface VPC endpoint in your VPC and update your client applications with the VPC endpoint specific endpoint names. For example, if your VPC endpoint id of the interface endpoint is vpce-0fe5b17a0707d6abc-29p5708s in the us-east-1 Region, then your endpoint specific DNS name will be vpce-0fe5b17a0707d6abc-29p5708s.s3.us-east-1.vpce.amazonaws.com. In this case, only the requests to the VPC endpoint specific names will route through Interface VPC endpoints to S3 while all other requests would continue to route through the gateway VPC endpoint. To learn more, visit the documentation.

Q: What is Amazon Macie and how can I use it to secure my data?

Amazon Macie is an AI-powered security service that helps you prevent data loss by automatically discovering, classifying, and protecting sensitive data stored in Amazon S3. Amazon Macie uses machine learning to recognize sensitive data such as personally identifiable information (PII) or intellectual property, assigns a business value, and provides visibility into where this data is stored and how it is being used in your organization. Amazon Macie continuously monitors data access activity for anomalies, and delivers alerts when it detects risk of unauthorized access or inadvertent data leaks.

You can use Amazon Macie to protect against security threats by continuously monitoring your data and account credentials. Amazon Macie gives you an automated and low-touch way to discover and classify your business data. It provides controls via templated Lambda functions to revoke access or trigger password reset policies upon the discovery of suspicious behavior, unauthorized data access to entities, or third-party applications. When alerts are generated, you can use Amazon Macie for incident response, using Amazon CloudWatch Events to swiftly take action to protect your data. For more information, visit the Amazon Macie documentation.

Q: What is IAM Access Analyzer for Amazon S3 and how does it work?

Access Analyzer for S3 is a feature that helps you simplify permissions management as you set, verify, and refine policies for your S3 buckets and access points. Access Analyzer for S3 monitors your existing access policies to verify that they provide only the required access to your S3 resources. Access Analyzer for S3 evaluates your bucket access policies and helps you discover and swiftly make changes to buckets that do not require access.

Access Analyzer for S3 alerts you when you have a bucket that is configured to allow access to anyone on the internet or that is shared with other AWS accounts. You receive *findings* about the source and level of public or shared access. For example, Access Analyzer for S3 will proactively inform you if unrequired read or write access was provided through an access control list or bucket policy. With these findings, you can immediately set or restore the required access policy.

When reviewing results that show potentially shared access to a bucket, you can Block Public Access to the bucket with a single click in the S3 console. You also can drill down into bucket-level permissions settings to configure granular levels of access. For auditing purposes, you can download Access Analyzer for S3 findings as a CSV report.

Additionally, the S3 console reports security warnings, errors, and suggestions from IAM Access Analyzer as you author your S3 policies. The console automatically runs more than 100 policy checks to validate your policies. These checks save you time, guide you to resolve errors, and help you apply security best practices.

For more information, visit the IAM Access Analyzer documentation.

S3 Access Grants

Q: What are Amazon S3 Access Grants?

Amazon S3 Access Grants map identities in directories such as Active Directory, or AWS Identity and Access Management (IAM) principals, to datasets in S3. This helps you manage data permissions at scale by automatically granting S3 access to end-users based on their corporate identity. Additionally, S3 Access Grants log end-user identity and the application used to access S3 data in AWS CloudTrail. This helps to provide a detailed audit history down to the end-user identity for all access to the data in your S3 buckets.

Q: Why should I use S3 Access Grants?

You should use S3 Access Grants if your S3 data is shared and accessed by many users and applications, where some of their identities are in your corporate directory such as Okta or Entra ID, and you need a scalable, simple, and auditable way to grant access to these S3 datasets at scale.

Q: How do I get started with S3 Access Grants?

You can get started with S3 Access Grants in four steps. First, configure an S3 Access Grants instance. In this step, if you want to use S3 Access Grants with users and groups in your corporate directory, enable AWS Identity Center and connect S3 Access Grants to your Identity Center instance. Second, register a location with S3 Access Grants. During this process, you give S3 Access Grants an IAM role that is used to create temporary S3 credentials that users and applications can use to access S3. Third, define permission grants that specify who can access what. Finally, at the time of access, have your application request temporary credentials from S3 Access Grants and use Access Grants-vended credentials to access S3.

Q: What types of identity are supported for S3 Access Grants permission grants?

S3 Access Grants supports two kinds of identities: enterprise user or group identities from AWS Identity Center, and AWS IAM principals including IAM users and roles. When you use S3 Access Grants with AWS Identity Center, you can define data permissions on the basis of directory group memberships. AWS Identity Center is an AWS service that connects to commonly-used identity providers, including Entra ID, Okta, Ping, and others. In addition to supporting directory identities via AWS Identity Center, S3 Access Grants also supports permission rules for AWS IAM principal including IAM users and roles. This is for use cases where you either manage a custom identity federation not through AWS Identity Center but via IAM and SAML assertion (example implementation), or manage application identities based on IAM principals, and still would like to use S3 Access Grants due to its scalability and auditability.

Q: What are the different access levels that S3 Access Grants offers?

S3 Access Grants offers three access levels: READ, WRITE, and READWRITE. READ allows you to view and retrieve objects from S3. WRITE allows you to write to and delete from S3. READWRITE allows you to do both READ and WRITE.

Q: Can I customize my access levels?

No. You can only use the three pre-defined access levels (READ/WRITE/READWRITE) that S3 Access Grants offers.

Q: Are there any limits on S3 Access Grants?

Yes. You can create up to 100,000 grants per S3 Access Grants instance, and up to 1,000 locations per S3 Access Grants instance.

Q: Is there any performance impact for data access when I use S3 Access Grants?

No. The latency for obtaining temporary credentials from S3 Access Grants is similar to obtaining temporary credentials from AWS STS today. Once you have obtained the credentials from S3 Access Grants, you can reuse unexpired credentials for subsequent requests. For these subsequent requests, there is no additional latency for requests authenticated via S3 Access Grants credentials compared to other methods.

Q: What other AWS services are required to use S3 Access Grants?

If you intend to use S3 Access Grants for directory identities, you will need to set up AWS IAM Identity Center first. AWS IAM Identity Center helps you create or connect your workforce identities, whether the identities are created and stored in Identity Center, or in an external third-party Identity Provider. Refer to the Identity Center documentation for the setup process. Once you have set up the Identity Center instance, you can connect the instance to S3 Access Grants. Thereafter, S3 Access Grants relies on Identity Center to retrieve user attributes such as group membership to evaluate requests and make authorization decisions.

Q: Does S3 Access Grants require client-side modifications?

Yes. Whereas today, you initialize your S3 client with IAM credentials associated with your application (for example, IAM role credentials for EC2 or IAM Roles Anywhere; or using long-term IAM user credentials), your application will need to instead obtain S3 Access Grants credentials first before initializing the S3 client. These S3 Access Grants credentials will be specific to the authenticated user in your application. Once the S3 client is initialized with these S3 Access Grants credentials, it can make requests for S3 data as usual using the credentials.

Q: Since client-side modifications are necessary, what AWS services and third-party applications are integrated with S3 Access Grants out-of-box today?

S3 Access Grants today already integrates with EMR and open-source Spark via the S3A connector. In addition, S3 Access Grants integrates with third-party software including Immuta and Informatica so that you can centralize permission management. And finally, S3 Access Grants supports Terraform and CloudFormation for you to programmatically provision S3 Access Grants.

Q: Is S3 Access Grants a replacement for AWS IAM?

No. S3 Access Grants does not replace IAM and in fact works well with your existing IAM-based data protection strategies (encryption, network, data-perimeter rules). S3 Access Grants is built on IAM primitives and enables you to express finer-grained S3 permissions at scale.

Q: Does S3 Access Grants work with KMS?

Yes. To utilize S3 Access Grants for objects encrypted with KMS, bucket owners include the necessary KMS permissions in the IAM role that they grant to S3 Access Grants as part of the location registration. S3 Access Grants can then subsequently utilize that IAM role to access the KMS-encrypted objects in the buckets.

Q: How do I view and manage my S3 Access Grants permission grants?

You can use either the S3 Access Grants console experience in the AWS Management Console or SDK and CLI APIs for you to view and manage your S3 Access Grants permissions.

Q: Can you grant public access to data with S3 Access Grants?

No, you cannot grant public access to data with S3 Access Grants.

Q: How can I audit requests that were authorized via S3 Access Grants?

The request by the application to initiate a data access session with S3 Access Grants will be recorded in CloudTrail. CloudTrail will distinguish the identity of the user making the request and the application identity accessing the data on the user's behalf. This helps you audit end-user identity of who accessed what data at what time.

Q: How is S3 Access Grants priced?

S3 Access Grants is charged based on the number of requests to S3 Access Grants. See the pricing page for details.

Q: What is the relationship between S3 Access Grants and Lake Formation?

AWS Lake Formation is for use cases where you need to manage access for tabular data (e.g., Glue tables), where you might want to enforce row- and column-level access. S3 Access Grants is for managing access for direct S3 permissions such as unstructured data including videos, images, logs, etc.

Q: Is S3 Access Grants integrated with IAM Access Analyzer?

No. S3 Access Grants is not integrated with IAM Access Analyzer at this time. You can't yet use IAM Access Analyzer to analyze S3 Access Grants permission grants. Customers can audit S3 Access Grants directly by going to the S3 Access Grants page in the S3 console, or programmatically using the ListAccessGrants API.

S3 Access Points

Q: What are Amazon S3 Access Points?

Today, customers manage access to their S3 buckets using a single bucket policy that controls access for hundreds of applications with different permission levels.

Amazon S3 Access Points simplifies managing data access at scale for applications using shared datasets on S3. With S3 Access Points, you can now easily create hundreds of access points per bucket, representing a new way of provisioning access to shared datasets. Access Points provide a customized path into a bucket, with a unique hostname and access policy that enforces the specific permissions and network controls for any request made through the access point. S3 Access Points can be associated with buckets in the same account or in another trusted account. Learn more by visting the S3 Access Points page and the user guide.

Q: Why should I use an access point?

S3 Access Points simplify how you manage data access to your shared datasets on S3. You no longer have to manage a single, complex bucket policy with hundreds of different permission rules that need to be written, read, tracked, and audited. With S3 Access Points, you can create access points or delegate permissions to trusted accounts to create cross-account access points on your bucket. This permits access to shared data sets with policies tailored to the specific application.

Using Access Points, you can decompose one large bucket policy into separate, discrete access point policies for each application that needs to access the shared data set. This makes it simpler to focus on building the right access policy for an application, while not having to worry about disrupting what any other application is doing within the shared data set. You can also create a Service Control Policy (SCP) and require that all access points be restricted to a Virtual Private Cloud (VPC), firewalling your data to within your private networks.

Q: How do S3 Access Points work?

Each S3 Access Point is configured with an access policy specific to a use case or application, and a bucket can have thousands of access points. For example, you can create an access point for your S3 bucket that grants access for groups of users or applications for your data lake. An Access Point can support a single user or application, or groups of users or applications within and across accounts, allowing separate management of each access point.

Additionally, you can delegate permissions to trusted accounts to create cross-account access points on your bucket. The cross-account access points don’t grant access to data until you are granted permissions from the bucket owner. The bucket owner always retains ultimate control on the data and must update the bucket policy to authorize requests from the cross-account access point. Visit the user guide for a sample bucket policy.

Each access point is associated with a single bucket and contains a network origin control, and a Block Public Access control. You can create an access point with a network origin control that only permits storage access from your Virtual Private Cloud, a logically isolated section of the AWS cloud. You can also create an access point with the access point policy configured to only allow access to objects with defined prefixes or to objects with specific tags.

You can access data in shared buckets through an access point in one of two ways. For S3 object operations, you can use the access point ARN in place of a bucket name. For requests requiring a bucket name in the standard S3 bucket name format, you can use an access point alias instead. Aliases for S3 Access Points are automatically generated and are interchangeable with S3 bucket names anywhere you use a bucket name for data access. Every time you create an access point for a bucket, S3 automatically generates a new Access Point Alias. For the full set of compatible operations and AWS services, visit the S3 documentation.

Q: Is there a quota on how many access points I can create?

By default, you can create 10,000 access points per Region per account on buckets in your account and cross-account. Unlike S3 buckets, there is no hard limit on the number of access points per AWS account. Visit AWS Service Quotas to request an increase in this quota.

Q: When using an access point, how are requests authorized?

S3 access points have their own IAM access point policy. You write access point policies like you would a bucket policy, using the access point ARN as the resource. Access point policies can grant or restrict access to the S3 data requested through the access point. Amazon S3 evaluates all the relevant policies, including those on the user, bucket, access point, VPC Endpoint, and service control policies as well as Access Control Lists, to decide whether to authorize the request.

Q: How do I write access point policies?

You can write an access point policy just like a bucket policy, using IAM rules to govern permissions and the access point ARN in the policy document.

Q: How is restricting access to specific VPCs using network origin controls on access points different from restricting access to VPCs using the bucket policy?

You can continue to use bucket policies to limit bucket access to specified VPCs. Access points provide an easier, auditable way to lock down all or a subset of data in a shared data set to VPC-only traffic for all applications in your organization using API controls. You can use an AWS Organizations Service Control Policy (SCP) to mandate that any access point created in your organization set the "network origin control" API parameter value to "vpc". Then, any new access point created automatically restricts data access to VPC-only traffic. No additional access policy is required to make sure that data requests are processed only from specified VPCs.

Q: Can I enforce a "No internet data access" policy for all access points in my organization?

Yes. To enforce a "No internet data access" policy for access points in your organization, you would want to make sure all access points enforce VPC only access. To do so, you will write an AWS SCP that only supports the value "vpc" for the "network origin control" parameter in the create_access_point() API. If you had any internet-facing access points that you created previously, they can be removed. You will also need to modify the bucket policy in each of your buckets to further restrict internet access directly to your bucket through the bucket hostname. Since other AWS services may be directly accessing your bucket, make sure you set up access to allow the AWS services you want by modifying the policy to permit these AWS services. Refer to the S3 documentation for examples of how to do this.

Q: Can I completely disable direct access to a bucket using the bucket hostname?

Not currently, but you can attach a bucket policy that rejects requests not made using an access point. Refer to the S3 documentation for more details.

Q: Can I replace or remove an access point from a bucket?

Yes. When you remove an access point, any access to the associated bucket through other access points, and through the bucket hostname, will not be disrupted.

Q: What is the cost of Amazon S3 Access Points?

There is no additional charge for access points or buckets that use access points. Usual Amazon S3 request rates apply.

Q: How do I get started with S3 Access Points?

You can start creating S3 Access Points on new buckets as well as existing buckets through the AWS Management Console, the AWS Command Line Interface (CLI), the Application Programming Interface (API), and the AWS Software Development Kit (SDK) client. To learn more about S3 Access Points, visit the user guide.

Durability & Data Protection

S3 Object Lock

Q: How durable is Amazon S3?

Amazon S3 provides the most durable storage in the cloud. Based on its unique architecture, S3 is designed to provide 99.999999999% (11 nines) data durability. Additionally, S3 stores data redundantly across a minimum of 3 Availability Zones (AZ) by default, providing built-in resilience against widespread disaster. Customers can store data in a single AZ to minimize storage cost or latency, in multiple AZs for resilience against the permanent loss of an entire data center, or in multiple AWS Regions to meet geographic resilience requirements.

Q: How is Amazon S3 designed for 99.999999999% durability?

Amazon S3's designed for durability is a function of storage device failure rates and the rate at which S3 can detect failure, and then re-replicate data on those devices. S3 has end-to-end integrity checking on every object upload and verifies that all data is correctly and redundantly stored across multiple storage devices before it considers your upload to be successful. Once your data is stored in S3, S3 continuously monitors data durability over time with periodic integrity checks of all data at rest. S3 also actively monitors the redundancy of your data to help verify that your objects are able to tolerate the concurrent failure of multiple storage devices.

Q: Is data stored in a One Zone storage class protected against damage or loss of the Availability Zone?

In the unlikely case of the loss or damage to all or part of an AWS Availability Zone, data in a One Zone storage class may be lost. For example, events like fire and water damage could result in data loss. Apart from these types of events, One Zone storage classes use similar engineering designs as Regional storage classes to protect objects from independent disk, host, and rack-level failures, and each are designed to deliver 99.999999999% data durability.

Q: How does Amazon S3 go beyond 99.999999999% durability?

Amazon S3 has a strong durability culture, and durability best practices are designed into our systems and software from the ground up. AWS has more experience operating high-durability storage than any other cloud provider, and we use this experience to mitigate durability risk and to incorporate durability safeguards into everything we do.

Q: With such high durability, do I still need to back up my critical data?

Yes. Amazon S3’s durability system does not protect against accidental or malicious deletes. S3 relies on customers to decide what data they want to keep, what data they want to get rid of, and what optional controls they need to protect against deletes that are incorrect, either due to accidents or malice. When you tell Amazon S3 to delete data, that data is immediately deleted, and it cannot be recovered by AWS. Honoring a delete request in this way is an important characteristic of the service.

Q: What capabilities does Amazon S3 provide to protect my data against accidental or malicious deletes?

S3 Object Versioning, S3 Replication, and S3 Object Lock are all optional features that you can use to add additional data protection, beyond the durability that S3 automatically provides. In addition, you can use a backup application to back up all or part of the data in your S3 buckets.

Q: What checksums does Amazon S3 support for data integrity checking?

Amazon S3 uses a combination of Content-MD5 checksums, secure hash algorithms (SHAs), and cyclic redundancy checks (CRCs) to verify data integrity. Amazon S3 performs these checksums on data at rest and repairs any disparity using redundant data. In addition, S3 calculates checksums on all network traffic to detect alterations of data packets when storing or retrieving data. You can choose from four supported checksum algorithms for data integrity checking on your upload and download requests. You can choose a SHA-1, SHA-256, CRC32, or CRC32C checksum algorithm, depending on your application needs. You can automatically calculate and verify checksums as you store or retrieve data from S3, and can access the checksum information at any time using the GetObjectAttributes S3 API or an S3 Inventory report. Calculating a checksum as you stream data into S3 saves you time as you're able to both verify and transmit your data in a single pass, instead of as two sequential operations. Using checksums for data validation is a best practice for data durability, and these capabilities increase the performance and reduce the cost to do so.

Q: What is Versioning?

Versioning allows you to preserve, retrieve, and restore every version of every object stored in an Amazon S3 bucket. Once you enable Versioning for a bucket, Amazon S3 preserves existing objects anytime you perform a PUT, POST, COPY, or DELETE operation on them. By default, GET requests will retrieve the most recently written version. Older versions of an overwritten or deleted object can be retrieved by specifying a version in the request.

Q: Why should I use Versioning?

Amazon S3 provides customers with a highly durable storage infrastructure. Versioning offers an additional level of protection by providing a means of recovery when customers accidentally overwrite or delete objects. This allows you to easily recover from unintended user actions and application failures. You can also use Versioning for data retention and archiving.

Q: How do I start using Versioning?

You can start using Versioning by enabling a setting on your Amazon S3 bucket. For more information on how to enable Versioning, refer to the Amazon S3 documentation.

Q: How does Versioning protect me from accidental deletion of my objects?

When a user performs a DELETE operation on an object, subsequent simple (un-versioned) requests will no longer retrieve the object. However, all versions of that object will continue to be preserved in your Amazon S3 bucket and can be retrieved or restored. Only the owner of an Amazon S3 bucket can permanently delete a version. You can set Lifecycle rules to manage the lifetime and the cost of storing multiple versions of your objects.

Q: Can I set up a trash, recycle bin, or rollback window on my Amazon S3 objects to recover from deletes and overwrites?

You can use Amazon S3 Lifecycle rules along with S3 Versioning to implement a rollback window for your S3 objects. For example, with your versioning-enabled bucket, you can set up a rule that archives all of your previous versions to the lower-cost S3 Glacier Flexible Retrieval storage class and deletes them after 100 days, giving you a 100-day window to roll back any changes on your data while lowering your storage costs. Additionally, you can save costs by deleting old (noncurrent) versions of an object after five days and when there are at least two newer versions of the object. You can change the number of days or the number of newer versions based on your cost optimization needs. This allows you to retain additional versions of your objects when needed, but saves you cost by transitioning or removing them after a period of time.

Q: How can I ensure maximum protection of my preserved versions?

Versioning's Multi-Factor Authentication (MFA) Delete capability can be used to provide an additional layer of security. By default, all requests to your Amazon S3 bucket require your AWS account credentials. If you enable Versioning with MFA Delete on your Amazon S3 bucket, two forms of authentication are required to permanently delete a version of an object: your AWS account credentials and a valid six-digit code and serial number from an authentication device in your physical possession. To learn more about enabling Versioning with MFA Delete, including how to purchase and activate an authentication device, refer to the Amazon S3 documentation.

Q: How am I charged for using Versioning?

Normal Amazon S3 rates apply for every version of an object stored or requested. For example, let's look at the following scenario to illustrate storage costs when utilizing Versioning (let's assume the current month is 31 days long):

1) Day 1 of the month: You perform a PUT of 4 GB (4,294,967,296 bytes) on your bucket.
2) Day 16 of the month: You perform a PUT of 5 GB (5,368,709,120 bytes) within the same bucket using the same key as the original PUT on Day 1.

When analyzing the storage costs of the above operations, note that the 4 GB object from Day 1 is not deleted from the bucket when the 5 GB object is written on Day 15. Instead, the 4 GB object is preserved as an older version and the 5 GB object becomes the most recently written version of the object within your bucket. At the end of the month:

Total Byte-Hour usage
[4,294,967,296 bytes x 31 days x (24 hours / day)] + [5,368,709,120 bytes x 16 days x (24 hours / day)] = 5,257,039,970,304 Byte-Hours.

Conversion to Total GB-Months
5,257,039,970,304 Byte-Hours x (1 GB / 1,073,741,824 bytes) x (1 month / 744 hours) = 6.581 GB-Month

The cost is calculated based on the current rates for your region on the Amazon S3 pricing page.

S3 Object Lock

Q: What is Amazon S3 Object Lock?

Amazon S3 Object Lock is an Amazon S3 feature that prevents an object version from being deleted or overwritten for a fixed amount of time or indefinitely, so that you can enforce retention policies as an added layer of data protection or for regulatory compliance. You can migrate workloads from existing write-once-read-many (WORM) systems into Amazon S3, and configure S3 Object Lock at the object- and bucket-level to prevent object version deletions prior to pre-defined Retain Until Dates or indefinitely (Legal Hold Dates). S3 Object Lock protection is maintained regardless of which storage class the object version resides in and throughout S3 Lifecycle transitions between storage classes.

You should use S3 Object Lock if you have regulatory requirements that specify that data must be WORM protected, or if you want to add an additional layer of protection to data in Amazon S3. S3 Object Lock can help you to meet regulatory requirements that specify that data should be stored in an immutable format, and also can protect against accidental or malicious deletion for data in Amazon S3.

Learn more by visiting the S3 Object Lock user guide.

Q: How does Amazon S3 Object Lock work?

Amazon S3 Object Lock prevents deletion of an object version for the duration of a specified retention period or indefinitely until a legal hold is removed. With S3 Object Lock, you're able to ensure that an object version remains immutable for as long as WORM protection is applied. You can apply WORM protection by either assigning a Retain Until Date or a Legal Hold to an object version using the AWS SDK, CLI, REST API, or the S3 Management Console. You can apply retention settings within a PUT request, or apply them to an existing object after it has been created.

The Retain Until Date defines the length of time for which an object version will remain immutable. Once a Retain Until Date has been assigned to an object, that object version cannot be modified or deleted until the Retain Until Date has passed. If a user attempts to delete an object before the Retain Until Date, the operation will be denied.

Alternatively, you can make an object immutable by applying a Legal Hold. A Legal Hold prevents an object version from being modified or deleted indefinitely until it is explicitly removed. In order to place and remove Legal Holds, your AWS account must have write permission for the PutObjectLegalHold action. Legal Hold can be applied to any object in an S3 Object Lock enabled bucket, whether or not that object is currently WORM-protected by a retention period.

S3 Object Lock can be configured in one of two Modes. When deployed in Governance Mode, AWS accounts with specific IAM permissions are able to remove WORM protection from an object version. If you require stronger immutability in order to comply with regulations, you can use Compliance Mode. In Compliance Mode, WORM protection cannot be removed by any user, including the root account.

How do I enable Amazon S3 Object Lock on a bucket?

You can use the Amazon S3 console, AWS API, or AWS CLI to enable S3 Object Lock while creating a new bucket or to configure S3 Object Lock on existing buckets. To enable S3 Object Lock on existing buckets, you can use the Amazon S3 console to edit S3 Object Lock settings in the bucket Properties tab the *PutObjectLockConfiguration* AWS API, or the AWS CLI. Once S3 Object Lock is enabled, you can set a default bucket level retention mode and time that will be applicable to all new objects uploaded to the bucket. For more information see the documentation on configuring S3 Object Lock using the S3 console, using the AWS API, and using the AWS CLI.

Q. How does enabling S3 Object Lock on existing buckets impact new objects uploaded to the buckets?

After enabling S3 Object Lock on existing buckets, you have the option to set a default S3 Object Lock retention mode and period for all new objects uploaded to the bucket. On the S3 console, you can do this by using the **Properties** tab for a selected bucket and enabling default retention mode and retention period for all new objects uploaded to the bucket. Alternatively, you can choose to not configure any bucket-level Object Lock settings, which will keep the new objects in the buckets unlocked. You can then lock individual objects by editing your S3 Object Lock settings in the **Object Properties** tab or use S3 Batch Operations to lock objects in bulk. For more information see the documentation on default retention settings.

Q: How does enabling S3 Object Lock for existing buckets impact the objects already existing in the buckets?

After enabling S3 Object Lock on existing buckets, the retention settings only apply to new objects uploaded to the buckets. To lock objects already existing in the buckets, you can choose to alter individual object-level retention properties using the Amazon S3 console, AWS CLI, or AWS API. On the S3 console, you can do this in the Properties tab for the object and editing the **Object Lock legal hold** or the **Object Lock retention** settings. Alternatively, you can use S3 Batch Operations to manage retention or enable a legal hold for multiple objects at once. For more information see the documentation on enabling S3 Object Lock using S3 Batch Operations.

Q: Can I disable S3 Object Lock after I have enabled it?

No, you cannot disable S3 Object Lock or S3 Versioning for buckets once S3 Object Lock is enabled.

Q: How do I get started with replicating objects from buckets with S3 Object Lock enabled?

To start replicating objects with S3 Replication from buckets with S3 Object Lock enabled , you can add a replication configuration on your source bucket by specifying a destination bucket in the same or different AWS Region and in the same or different AWS account. You can choose to replicate all objects at the S3 bucket level, or filter objects on a shared prefix level, or an object level using S3 object tags. You will also need to specify an AWS Identity and Access Management (IAM) role with the required permissions to perform the replication operation. You can use the S3 console, AWS API, AWS CLI, AWS SDKs, or AWS CloudFormation to enable replication and must have S3 Versioning enabled for both the source and destination buckets. Additionally, to replicate objects from S3 Object Lock enabled buckets, your destination bucket must also have S3 Object Lock enabled. For more information see the documentation on setting up S3 Replication and using S3 Object Lock with S3 Replication.

Q: Do I need additional permissions to replicate objects from buckets with S3 Object Lock enabled?

Yes, to replicate objects from S3 Object Lock enabled buckets you need to grant two new permissions, s3:GetObjectRetention and s3:GetObjectLegalHold, on the source bucket in the IAM role that you use to set up replication. Alternatively, if the IAM role has an s3:Get* permission, it satisfies the requirement. For more information see the documentation on using S3 Object Lock with S3 Replication.

Q: Are there any limitations for using S3 Replication while replicating from S3 Object Lock buckets?

No, all features of S3 Replication, such as S3 Same-Region Replication (S3 SRR), S3 Cross-Region Replication (S3 CRR), S3 Replication metrics to track progress, S3 Replication Time Control (S3 RTC), and S3 Batch Replication, are supported while replicating from S3 Object Lock buckets.

Q: How can I replicate existing objects from S3 Object Lock enabled buckets?

You can use S3 Batch Replication to replicate existing objects from S3 Object Lock enabled buckets. For more information on replicating existing objects, see the documentation on S3 Batch Replication.

Q: What is the retention status of the replicas of source objects protected with S3 Object Lock?

The replicas of objects protected with S3 Object Lock follow the same retention policy as the source objects. You can use the GET Object or HEAD Object commands to view the Object Lock status of the replica objects. Both commands return the **Retention mode**, **Retain until date**, and the legal hold status for the specified object version. You can also configure Amazon S3 Inventory reports on your buckets to include the

Retain until date, **Retention mode**, and legal hold status for all objects in a bucket. For more information see the documentation on viewing the S3 Object Lock information for an object and configuring Amazon S3 Inventory.

Storage Classes

S3 Intelligent-Tiering | S3 Standard | S3 Express One Zone | S3 Standard-Infrequent Access | S3 One Zone-Infrequent Access | Amazon S3 Glacier Instant Retrieval | Amazon S3 Glacier Flexible Retrieval | Amazon S3 Glacier Deep Archive | S3 on Outposts

Q: What are the Amazon S3 storage classes?

Amazon S3 offers a range of storage classes that you can choose from based on the data access, resiliency, and cost requirements of your workloads. S3 storage classes are purpose-built to provide the lowest cost storage for different access patterns. S3 storage classes are ideal for virtually any use case, including those with demanding performance needs, data residency requirements, unknown or changing access patterns, or archival storage. Each S3 storage class charges a fee to store data and fees to access data. In deciding which S3 storage class best fits your workload, consider the access patterns and retention time of your data to optimize for the lowest total cost over the lifetime of your data.

S3 Storage Classes can be configured at the object level and a single bucket can contain objects stored across all of the storage classes. You can also use S3 Lifecycle policies to automatically transition objects between storage classes without any application changes.

Q: How do I decide which S3 storage class to use?

In deciding which S3 storage class best fits your workload, consider the access patterns and retention time of your data to optimize for the lowest total cost over the lifetime of your data. Many workloads have changing (user-generated content), unpredictable (analytics, data lakes), or unknown (new applications) access patterns, and that is why S3 Intelligent-Tiering should be the default storage class to automatically save on storage costs. If you know the access patterns of your data, you can follow this guidance. The S3 Standard storage class is ideal for frequently accessed data; this is the best choice if you access data more than once a month. S3 Standard-Infrequent Access is ideal for data retained for at least a month and accessed once every month or two.

The Amazon S3 Glacier storage classes are purpose-built for data archiving, providing you with the highest performance, most retrieval flexibility, and the lowest cost archive storage in the cloud. You can now choose from three archive storage classes optimized for different access patterns and storage duration. For archive data that needs immediate access, such as medical images, news media assets, or genomics data, choose the S3 Glacier Instant Retrieval storage class, an archive storage class that delivers the lowest cost storage with milliseconds retrieval. For archive data that does not require immediate access but needs the flexibility to retrieve large sets of data at no cost, such as backup or disaster recovery use cases, choose S3 Glacier Flexible Retrieval, with retrieval in minutes or free bulk retrievals in 5—12 hours. To save even more on long-lived archive storage such as compliance archives and digital media preservation, choose S3 Glacier Deep Archive, the lowest cost storage in the cloud with data retrieval within 12 hours. All these storage classes provide multi-Availability Zone (AZ) resiliency by redundantly storing data on multiple devices and physically separated AWS Availability Zones in an AWS Region.

For data that has a lower resiliency requirement, you can reduce costs by selecting a single-AZ storage class, like S3 One Zone-Infrequent Access. If you have data residency or latency requirements that can’t be met by an existing AWS Region, you can choose S3 on Outposts to store data on-premises.

You can learn more about these storage classes on the Amazon S3 Storage Classes page.

S3 Intelligent-Tiering

Q: What is S3 Intelligent-Tiering?

S3 Intelligent-Tiering is the first cloud storage that automatically reduces your storage costs on a granular object level by automatically moving data to the most cost-effective access tier based on access frequency, without performance impact, retrieval fees, or operational overhead. S3 Intelligent-Tiering delivers milliseconds latency and high throughput performance for frequently, infrequently, and rarely accessed data in the Frequent, Infrequent, and Archive Instant Access tiers. For a small monthly object monitoring and automation charge, S3

Intelligent-Tiering monitors the access patterns and moves the objects automatically from one tier to another. There are no retrieval charges in S3 Intelligent-Tiering, so you won't see unexpected increases in storage bills when access patterns change.

You can use S3 Intelligent-Tiering as the default storage class for virtually any workload, especially data lakes, data analytics, machine learning, new applications, and user-generated content.

Q: How does S3 Intelligent-Tiering work?

The Amazon S3 Intelligent-Tiering storage class is designed to optimize storage costs by automatically moving data to the most cost-effective access tier when access patterns change. For a low monthly object monitoring and automation charge, S3 Intelligent-Tiering monitors access patterns and automatically moves objects that have not been accessed for 30 consecutive days to the Infrequent Access tier to save up to 40% on storage costs. After 90 days consecutive days of no access, objects are moved to the Archive Instant Access tier to save up to 68% on storage costs. There is no impact on performance and there are no retrieval charges in S3 Intelligent-Tiering. If an object in the Infrequent Access tier or Archive Instant Access tier is accessed later, it is automatically moved back to the Frequent Access tier.

To get the lowest storage cost on data that can be accessed asynchronously, you can choose to activate additional archiving capabilities. Once you enable one or both of the asynchronous archive access tiers, S3 Intelligent-Tiering will move objects that have not been accessed for a minimum of 90 days to the Archive Access tier to save up to 71% and after 180 days of no access to the Deep Archive Access tier to save up to 95% for rarely accessed objects. If an object in the optional Archive or Deep Access tiers is restored later, it is moved back to the Frequent Access tier, and before you can retrieve the object you must first restore the object using RestoreObject. For information about restoring archived objects, see Restoring Archived Objects. There are no retrieval charges in S3 Intelligent-Tiering. No additional tiering or lifecycle charges apply when objects are moved between access tiers within the S3 Intelligent-Tiering storage class.

There is no minimum object size for S3 Intelligent-Tiering, but objects smaller than 128KB are not eligible for auto-tiering. These smaller objects may be stored in S3 Intelligent-Tiering, but will always be charged at the Frequent Access tier rates, and are not charged the monitoring and automation charge.

If you would like to standardize on S3 Intelligent-Tiering as the default storage class for newly created data, you can modify your applications by specifying INTELLIGENT-TIERING on your S3 PUT API request header. S3 Intelligent-Tiering is designed for 99.9% availability and

99.999999999% durability, and automatically offers the same low latency and high throughput performance of S3 Standard. You can use AWS Cost Explorer to measure the additional savings from the Archive Instant Access tier.

Q: Why would I choose to use S3 Intelligent-Tiering?

You can use S3 Intelligent-Tiering as the default storage class for virtually any workload, especially data lakes, data analytics, machine learning, new applications, and user-generated content. S3 Intelligent-Tiering is the first cloud storage that automatically reduces your storage costs on a granular object level by automatically moving data to the most cost-effective access tier based on access frequency, without performance impact, retrieval fees, or operational overhead. If you have data with unknown or changing access patterns, including data lakes, data analytics, and new applications, we recommend using S3 Intelligent-Tiering. If you have data that does not require immediate retrieval, we recommend activating the Deep Archive Access tier where you pay as little as $1 per TB per month for data that may become rarely accessed over long periods of time. S3 Intelligent-Tiering is for data with unknown or changing access patterns. There are no retrieval fees when using the S3 Intelligent-Tiering storage class.

Q: What performance does S3 Intelligent-Tiering offer?

S3 Intelligent-Tiering automatically optimizes your storage costs without an impact to your performance. The S3 Intelligent-Tiering Frequent, Infrequent, and Archive Instant Access tiers provide milliseconds latency and high throughput performance.

Q: What performance do the optional Archive Access and Deep Archive Access tiers provide?

For data that can be accessed asynchronously, the optional Archive Access tier has the same performance as S3 Glacier Flexible Retrieval, and the Deep Archive Access tier has the same performance as the S3 Glacier Deep Archive storage class. You should only activate the asynchronous archive capabilities if your application can wait minutes to hours. If the object you are retrieving is stored in the Archive or Deep Archive Access tiers, before you can retrieve the object you must first restore an object using RestoreObject. For information about restoring archived objects, see Restoring Archived Objects. Objects in the Archive Access tier are moved to the Frequent Access tier in 3—5 hours and within 12 hours if they are in the Deep Archive Access tier. If you need faster access to an object in the Archive Access tier, you can pay for faster retrieval by using the console to select the expedited retrieval speed option.

Q: How durable and available is S3 Intelligent-Tiering?

S3 Intelligent-Tiering is designed for the same 99.999999999% durability as the S3 Standard storage class. S3 Intelligent-Tiering is designed for 99.9% availability, and carries a service level agreement providing service credits if availability is less than our service commitment in any billing cycle.

Q: How do I get my data into S3 Intelligent-Tiering?

There are two ways to get data into S3 Intelligent-Tiering. You can directly PUT into S3 Intelligent-Tiering by specifying INTELLIGENT_TIERING in the x-amz-storage-class header or set lifecycle policies to transition objects from S3 Standard or S3 Standard-IA to S3 INTELLIGENT_TIERING.

Q: How am I charged for S3 Intelligent-Tiering?

S3 Intelligent-Tiering charges you for monthly storage, requests, and data transfer, and charges a small monthly charge for monitoring and automation per object. The S3 Intelligent-Tiering storage class automatically stores objects in three access tiers: a Frequent Access tier priced at S3 Standard storage rates, an Infrequent Access tier priced at S3 Standard-Infrequent Access storage rates, and an Archive Instant Access tier priced at the S3 Glacier Instant Retrieval storage rates. S3 Intelligent-Tiering also has two optional archive tiers designed for asynchronous access, an Archive Access tier priced at S3 Glacier Flexible Retrieval storage rates, and a Deep Archive Access tier priced at S3 Glacier Deep Archive storage rates.

For a small monitoring and automation fee, S3 Intelligent-Tiering monitors access patterns and automatically moves objects through low latency and high throughput access tiers, as well as two opt in asynchronous archive access tiers where customers get the lowest storage costs in the cloud for data that can be accessed asynchronously.

There is no minimum billable object size in S3 Intelligent-Tiering, but objects smaller than 128KB are not eligible for auto-tiering. These small objects will not be monitored and will always be charged at the Frequent Access tier rates, with no monitoring and automation charge. For each object archived to the Archive Access tier or Deep Archive Access tier in S3 Intelligent-Tiering, Amazon S3 uses 8 KB of storage for the name of the object and other metadata (billed at S3 Standard storage rates) and 32 KB of storage for index and related metadata (billed at S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive storage rates).

Q: Is there a charge to retrieve data from S3 Intelligent-Tiering?

No. There are no retrieval fees for S3 Intelligent-Tiering. S3 Intelligent-Tiering monitors the access patterns of your data and if you access an object in the Infrequent Access, Archive Instant Access, or the asynchronous archive tiers, S3 Intelligent-Tiering automatically moves that object to the Frequent Access tier.

Q: How do I activate S3 Intelligent-Tiering archive access tiers?

You can activate the Archive Access tier and Deep Archive Access tier by creating a bucket, prefix, or object tag level configuration using the Amazon S3 API, CLI, or S3 management console. You should only activate one or both of the archive access tiers if your objects can be accessed asynchronously by your application.

Q: Can I extend the time before objects are archived within S3 Intelligent-Tiering storage class?

Yes. In the bucket, prefix, or object tag level configuration, you can extend the last access time for archiving objects in S3 Intelligent-Tiering. When enabled, by default objects that haven't been accessed for a minimum of 90 consecutive days automatically move to the Archive Access tier, skipping the Archive Instant Access tier. Objects that haven't been accessed for a minimum of 180 consecutive days automatically move to the Deep Archive Access tier. The default configuration for the consecutive days since last access before automatic archiving in S3 Intelligent-Tiering can be extended for up to 2 years.

Q: How do I access an object from the Archive Access or Deep Archive Access tiers in the S3 Intelligent-Tiering storage class?

To access an object in the Archive or Deep Archive Access tiers, you need to issue a Restore request and the object will begin moving back to the Frequent Access tier, all within the S3 Intelligent-Tiering storage class. Objects in the Archive Access Tier are moved to the Frequent Access tier in 3-5 hours, objects in the Deep Archive Access tier are moved to the Frequent Access tier within 12 hours. Once the object is in the Frequent Access tier, you can issue a GET request to retrieve the object.

Q: How do I know in which S3 Intelligent-Tiering access tier my objects are stored in?

You can use Amazon S3 Inventory to report the access tier of objects stored in the S3 Intelligent-Tiering storage class. Amazon S3 Inventory provides CSV, ORC, or Parquet output files that list your objects and their corresponding metadata on a daily or weekly basis for an S3 bucket or a shared prefix. You can also make a HEAD request on your objects to report the S3 Intelligent-Tiering archive access tiers.

Q: Can I lifecycle objects from S3 Intelligent-Tiering to another storage class?

Yes. You can lifecycle objects from S3 Intelligent-Tiering Frequent Access, Infrequent, and Archive Instant Access tiers to S3 One-Zone Infrequent Access, S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive. In addition, you can lifecycle objects from the S3 Intelligent-Tiering optional archive access tiers to S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive, and from the S3 Intelligent-Tiering Deep Archive Access tier to S3 Glacier Deep Archive.

Q: Is there a minimum duration for S3 Intelligent-Tiering?

No. The S3 Intelligent-Tiering storage class has no minimum storage duration.

Q: Is there a minimum billable object size for S3 Intelligent-Tiering?

No. The S3 Intelligent-Tiering storage class has no minimum billable object size, but objects smaller than 128KB are not eligible for auto-tiering. These smaller objects will always be charged at the Frequent Access tier rates, with no monitoring and automation charge. For each object archived to the opt-in Archive Access tier or Deep Archive Access tier in S3 Intelligent-Tiering, Amazon S3 uses 8 KB of storage for the name of the object and other metadata (billed at S3 Standard storage rates) and 32 KB of storage for index and related metadata (billed at S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive storage rates). For more details, visit the Amazon S3 pricing page.

S3 Standard

Q: What is S3 Standard?

Amazon S3 Standard delivers durable storage with millisecond access latency and high throughput performance for frequently accessed data, typically more than once per month. S3 Standard is designed for performance-sensitive uses cases, such as data lakes, cloud-native applications, dynamic websites, content distribution, mobile and gaming applications, analytics, and machine learning models. S3 Standard is designed for 99.99% data availability and durability of 99.999999999% of objects across multiple Availability Zones in a given year. You can use S3 Lifecycle policies to control exactly when data is transitioned between S3 Standard and lower costs storage classes without any application changes.

Q: Why would I choose to use S3 Standard?

S3 Standard is ideal for your most frequently accessed or modified data that requires access in milliseconds and high throughput performance. S3 Standard is ideal for data that is read or written very often, as there are no retrieval charges. S3 Standard is optimized for a wide variety of use cases, including data lakes, cloud native applications, dynamic websites, content distribution, mobile and gaming applications, and analytics.

S3 Express One Zone

Q: What is the Amazon S3 Express One Zone storage class?

Amazon S3 Express One Zone is a high-performance, single-Availability Zone Amazon S3 storage class purpose-built to deliver consistent single-digit millisecond data access for customers’ most latency-sensitive applications. Amazon S3 Express One Zone is the lowest latency cloud object storage class available today, with data access speed up to 10x faster and with request costs 50% lower than Amazon S3 Standard. With S3 Express One Zone, you can select a specific AWS Availability Zone within an AWS Region to store your data. You can choose to co-locate your storage and compute resources in the same Availability Zone to further optimize performance.

Q: Why would I choose to use the Amazon S3 Express One Zone storage class?

S3 Express One Zone is the ideal storage class for applications that need the fastest data access speed and highest performance for latency-sensitive applications. S3 Express One Zone is the best storage class for request-intensive operations such as machine learning (ML) training and inference, interactive analytics, and media content creation.

Q: How do I get started with the Amazon S3 Express One Zone storage class?

You can get started by creating an S3 directory bucket in an AWS Availability Zone (AZ) of your choosing. You can choose to co-locate your storage and compute resources in the same AZ to further optimize performance. Directory buckets have S3 Block Public Access on by default. After creating the directory bucket, you can directly upload objects to the S3 Express One Zone storage class or copy objects from existing S3 storage classes into S3 Express One Zone. You can also import data with a single click in the AWS Management Console into S3 Express One Zone or use S3 Batch Operations to copy an entire bucket, prefix, or subsets of data from an existing S3 storage class into S3 Express One Zone.

Q: How can I import data into the Amazon S3 Express One Zone storage class?

You can import data from within the same AWS Region into the S3 Express One Zone storage class via the S3 console by using the **Import** option after you create a directory bucket. Import simplifies copying data into S3 directory buckets by letting you choose a prefix or bucket to import data from without having to specify all of the objects to copy individually. S3 Batch Operations copies the objects in the selected prefix or general purpose bucket and you can monitor the progress of the import copy job through the S3 Batch Operations job details page.

Q: How many Availability Zones are Amazon S3 Express One Zone objects stored in?

S3 Express One Zone objects are stored in a single AWS Availability Zone (AZ) that you choose. Storing objects in one zone gives you the ability to store your data local to your compute to minimize latency. You can access data from across Availability Zones, although latency will increase.

Q: What performance does the Amazon S3 Express One Zone storage class provide?

S3 Express One Zone provides similar performance elasticity as other S3 storage classes, but with consistent single-digit millisecond first-byte read and write latency request latencies—up to 10x faster than existing S3 storage classes. With S3 Express One Zone, customers don't need to plan or provision capacity or throughput requirements in advance, and benefit immediately from requests completing up to an order of magnitude faster. S3 Express One Zone is ideal for analytics jobs where storage latency speeds job completion times and reduces overall TCO. It's also ideal for interactive workloads, like video editing, where creative professionals need the most responsive possible access to their S3 data.

Q: How does the Amazon S3 Express One Zone storage class achieve high performance?

S3 Express One Zone uses a unique architecture to optimize for performance and deliver consistently low request latency. S3 Express One Zone stores data on high-performance hardware and its object protocol has been enhanced to streamline authentication and metadata overheads. Additionally, to further increase access speed and support hundreds of thousands of requests per second, data is stored in a new bucket type—an Amazon S3 directory bucket. With S3 Express One Zone, you can select a specific AWS Availability Zone within an AWS Region to store your data. You can choose to co-locate your storage and compute resources in the same Availability Zone to further optimize performance.

Q: What request rate performance does an S3 directory bucket support?

Each S3 directory bucket can support hundreds of thousands of transactions per second (TPS), independent of the number of directories within the bucket.

Q: What happens to an S3 directory bucket with no request activity for an extended period of time?

S3 directory buckets that have no request activity for a period of at least 3 months will transition to an inactive state. While in an inactive state, a directory bucket is temporarily inaccessible for reads and writes. Inactive buckets retain all storage, object metadata, and bucket metadata. Existing storage charges will apply to inactive buckets. On an access request to an inactive bucket, the bucket will transition to an active state, typically within a few minutes. During this transition period, reads and writes will return a **503 SlowDown** error code.

Q: How should I plan for my application's throughput needs with the S3 Express One Zone storage class?

S3 Express One Zone provides similar high, elastic throughput as other Amazon S3 storage classes. S3 Express One Zone is designed from the ground up to allow individual customers to burst throughput to very high aggregate levels. For example, machine learning model training applications can train against millions of objects and petabytes of data. You can achieve the highest performance by spreading these requests over separate connections to maximize the accessible bandwidth.

Q: How is request authorization different with Amazon S3 Express One Zone compared to other S3 storage classes?

With S3 Express One Zone, you authenticate and authorize requests through a new session-based mechanism, S3 CreateSession, which is optimized to provide the lowest latency. You can use CreateSession to request temporary credentials that provide low latency access to your bucket. These temporary credentials are scoped to a specific S3 directory bucket. For more information on this session-based model, refer to S3 Create Session in the developer guide.

Q: How reliable is the Amazon S3 Express One Zone storage class?

S3 Express One Zone is designed to deliver 99.95% availability within a single Availability Zone, with an availability SLA of 99.9%.

Q: How is the Amazon S3 Express One Zones storage class designed to provide 99.95% availability?

With S3 Express One Zone, your data is redundantly stored on multiple devices within a single AZ. S3 Express One Zone is designed to sustain concurrent device failures by quickly detecting and repairing any lost redundancy. This means that S3 Express One Zone automatically shifts requests to new devices within an AZ if the existing device encounters a failure. This redundancy gives you uninterrupted access to your data within an AZ.

Q: How am I charged for Amazon S3 Express One Zone?

There are no set up charges or commitments to begin using S3 Express One Zone. S3 Express One Zone charges you for storage and requests. The volume of storage billed in a month is accrued based on total storage used per hour, measured in gigabyte per month (GB-Month). You are also charged a per request fee for access based on the request type—such as PUTs and GETs. You will pay an additional per-GB fee for the portion of the request size exceeding 512 KB.

Example 1:

Assume you store 10 GB of data in S3 Express One Zone for 30 days, making a total of 1,000,000 writes and 9,000,000 reads, accessing with Athena with a request size of 10 KB. Then, you delete 1,000,000 files by the end of 30 days. Assuming your bucket is in the US East (Northern Virginia) Region, the storage and request charges are calculated below:

Storage Charges

Total Byte-Hour usage = 10 GB-Month

Total Storage cost = 10 GB-Month x $0.16 = $1.6

Request Charges

1,000,000 PUT Requests: 1,000,000 requests x $0.0025/1,000 = $2.5

9,000,000 GET Requests: 9,000,000 requests x $0.0002/1,000 = $1.8

1,000,000 DELETE requests = 1,000,000 requests x $0.00 (no charge) = $0

Total Charges = $1.6+$2.5+$1.8 = $5.9

Example 2:

Assume you store 10 TB of data for machine learning training for an 8-hour workload every day, and then delete it. During the 8-hour workload you make 5,242,880 writes and 10,485,760 reads for a 2 MB request size. Assume you do this for 30 days (a month).

Storage Charges

Total Byte-Hour usage = [10,995,116,277,760 bytes x 30 days x (8 hours / day)] = 2,638,827,906,662,400 Byte-Hours = 3303.77 GB-Month

Total Storage cost = 3303.77 GB x $0.16 = $528.51

Request Charges

5,242,880 PUT Requests/day: 5,242,880 requests x 30 x $0.0025/1,000 = $393.22

10,485,760 GET Requests/day: 10,485,760 requests x 30 x $0.0002/1,000 = $62.91

5,242,880 DELETE requests/day: 5,242,880 requests x $0.00 (no charge) = $0

Per request additional bandwidth Charge will be applied on : 1.5 MB (2-0.5 MB) = 0.001465 GB
PUT Bandwidth Charge : 0.001465 GB x 5,242,880 x 30 x $0.008 = $1843.2
GET Bandwidth Charge : 0.001465 GB x 10,485,760 x 30 x $0.015 = $691.2

Total Charges = $528.51+$393.22+$62.91+$1843.2+$691.2= $3519.05

Q: Are there any additional Data Transfer charges for using the Amazon S3 Express One Zone storage class within the same Region?

The request charges to access data in S3 Express One Zone includes costs to transfer data within the AWS network in a Region, and there is no additional Data Transfer charge for data transferred between Amazon EC2 (or any AWS service) and S3 Express One Zone within the same Region, for example, data transferred within the US East (Northern Virginia) Region.

Q: Are there any additional networking charges for using Gateway VPC endpoints with the Amazon S3 Express One Zone storage class?

The request charges to access data in S3 Express One Zone includes costs to use Gateway VPC endpoints, and there is no additional charge for using Gateway endpoints with S3 Express One Zone.

S3 Standard-Infrequent Access (S3 Standard-IA)

Q: What is S3 Standard-Infrequent Access?

Amazon S3 Standard-Infrequent Access (S3 Standard-IA) is an Amazon S3 storage class for data that is accessed less frequently but requires rapid access when needed. S3 Standard-IA offers the high durability, throughput, and low latency of the Amazon S3 Standard storage class, with a low per-GB storage price and per-GB retrieval charge. This combination of low cost and high performance make S3 Standard-IA ideal for long-term storage, backups, and as a data store for disaster recovery. The S3 Standard-IA storage class is set at the object level and can exist in the same bucket as the S3 Standard or S3 One Zone-IA storage classes, allowing you to use S3 Lifecycle policies to automatically transition objects between storage classes without any application changes.

Q: Why would I choose to use S3 Standard-IA?

S3 Standard-IA is ideal for data that is accessed less frequently, but requires rapid access when needed. S3 Standard-IA is ideally suited for long-term file storage, older sync and share storage, and other aging data.

Q: What performance does S3 Standard-IA offer?

S3 Standard-IA provides the same milliseconds latency and high throughput performance as the S3 Standard storage class.

Q: How do I get my data into S3 Standard-IA?

There are two ways to get data into S3 Standard-IA. You can directly PUT into S3 Standard-IA by specifying STANDARD_IA in the x-amz-storage-class header. You can also set Lifecycle policies to transition objects from the S3 Standard to the S3 Standard-IA storage class.

Q: What charges will I incur if I change the storage class of an object from S3 Standard-IA to S3 Standard with a COPY request?

You will incur charges for an S3 Standard (destination storage class) COPY request and an S3 Standard-IA (source storage class) data retrieval. For more information, visit the Amazon S3 pricing page.

Q: Is there a minimum storage duration charge for S3 Standard-IA?

S3 Standard-IA is designed for long-lived, infrequently accessed data that is retained for months or years. Data that is deleted from S3 Standard-IA within 30 days will be charged for a full 30 days. See the Amazon S3 pricing page for information about S3 Standard-IA pricing.

Q: Is there a minimum object storage charge for S3 Standard-IA?

S3 Standard-IA is designed for larger objects and has a minimum object storage charge of 128KB. Objects smaller than 128KB in size will incur storage charges as if the object were 128KB. For example, a 6KB object in S3 Standard-IA will incur S3 Standard-IA storage charges for 6KB and an additional minimum object size charge equivalent to 122KB at the S3 Standard-IA storage price. See the Amazon S3 pricing page for information about S3 Standard-IA pricing.

Q: Can I tier objects from S3 Standard-IA to S3 One Zone-IA or to the S3 Glacier Flexible Retrieval storage class?

Yes. In addition to using Lifecycle policies to migrate objects from S3 Standard to S3 Standard-IA, you can also set up Lifecycle policies to tier objects from S3 Standard-IA to S3 One Zone-IA, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, and the S3 Glacier Deep Archive storage class.

S3 One Zone-Infrequent Access (S3 One Zone-IA)

Q: What is S3 One Zone-IA storage class?

S3 One Zone-IA storage class is an Amazon S3 storage class that customers can choose to store objects in a single availability zone. S3 One Zone-IA storage redundantly stores data within that single Availability Zone to deliver storage at 20% less cost than geographically redundant S3 Standard-IA storage, which stores data redundantly across multiple geographically separate Availability Zones.

S3 One Zone-IA offers a 99% available SLA and is also designed for eleven 9's of durability within the Availability Zone. However, the data in the S3 One Zone-IA storage class is not resilient to the physical loss of an entire Availability Zone.

S3 One Zone-IA storage offers the same Amazon S3 features as S3 Standard and S3 Standard-IA and is used through the Amazon S3 API, CLI and console. S3 One Zone-IA storage class is set at the object level and can exist in the same bucket as S3 Standard and S3 Standard-IA storage classes. You can use S3 Lifecycle policies to automatically transition objects between storage classes without any application changes.

Q: What use cases are best suited for S3 One Zone-IA storage class?

Customers can use S3 One Zone-IA for infrequently-accessed storage, like backup copies, disaster recovery copies, or other easily re-creatable data.

Q: What performance does S3 One Zone-IA storage offer?

S3 One Zone-IA storage class offers the same latency and throughput performance as the S3 Standard and S3 Standard-Infrequent Access storage classes.

Q: How durable is the S3 One Zone-IA storage class?

S3 One Zone-IA storage class is designed for 99.999999999% of durability within an Availability Zone. However, data in the S3 One Zone-IA storage class is not resilient to the loss of availability or physical loss of an Availability Zone. In contrast, S3 Standard, S3 Intelligent-Tiering, S3 Standard-Infrequent Access, and the S3 Glacier storage classes are designed to withstand loss of availability or the destruction of an Availability Zone. S3 One Zone-IA can deliver the same or better durability and availability than most modern, physical data centers, while providing the added benefit of elasticity of storage and the Amazon S3 feature set.

Q: Is an S3 One Zone-IA “Zone” the same thing as an AWS Availability Zone?

Yes. Each AWS Region is a separate geographic area. Each Region has multiple, isolated locations known as Availability Zones. The Amazon S3 One Zone-IA storage class uses an individual AWS Availability Zone within the Region.

Q: How much disaster recovery protection do I forgo by using S3 One Zone-IA?

Each Availability Zone uses redundant power and networking. Within an AWS Region, Availability Zones are on different flood plains, earthquake fault zones, and geographically separated for fire protection. S3 Standard and S3 Standard-IA storage classes offer protection against these sorts of disasters by storing your data redundantly in multiple Availability Zones. S3 One Zone-IA offers protection against equipment failure within an Availability Zone, but the data is not resilient to the physical loss of the Availability Zone resulting from disasters, such as earthquakes and floods. Using S3 One Zone-IA, S3 Standard, and S3 Standard-IA options, you can choose the storage class that best fits the durability and availability needs of your storage.

Amazon S3 Glacier Instant Retrieval storage class

Q: What is the S3 Glacier Instant Retrieval storage class?

The S3 Glacier Instant Retrieval storage class delivers the lowest cost storage for long-lived data that is rarely accessed and requires milliseconds retrieval. S3 Glacier Instant Retrieval delivers the fastest access to archive storage, with the same throughput and milliseconds access as S3 Standard and S3 Standard-IA storage classes. S3 Glacier Instant Retrieval is designed for 99.999999999% (11 9s) of data durability and 99.9% availability by redundantly storing data across a minimum of three physically separated AWS Availability Zones.

Q: Why would I choose to use S3 Glacier Instant Retrieval?

S3 Glacier Instant Retrieval is ideal if you have data that is rarely accessed (once a quarter) and requires milliseconds retrieval times. It’s the ideal storage class if you want the same low latency and high throughput performance as S3 Standard-IA, but store data that is accessed less frequently than S3 Standard-IA, with a lower storage price and slightly higher data access costs.

Q: How available and durable is S3 Glacier Instant Retrieval?

S3 Glacier Instant Retrieval is designed for 99.999999999% (11 9s) of durability and 99.9% availability, the same as S3 Standard-IA, and carries a service level agreement providing service credits if availability is less than 99% in any billing cycle.

Q: What performance does S3 Glacier Instant Retrieval offer?

S3 Glacier Instant Retrieval provides the same milliseconds latency and high throughput performance as the S3 Standard and S3 Standard-IA storage classes. Unlike the S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive storage classes, which are designed for asynchronous access, you do not need to issue a Restore request before accessing an object stored in S3 Glacier Instant Retrieval.

Q: How do I get my data into S3 Glacier Instant Retrieval?

There are two ways to get data into S3 Glacier Instant Retrieval. You can directly PUT into S3 Glacier Instant retrieval by specifying GLACIER_IR in the x-amz-storage-class header or set S3 Lifecycle policies to transition objects from S3 Standard or S3 Standard-IA to S3 Glacier Instant Retrieval.

Q: Is there a minimum storage duration charge for Amazon S3 Glacier Instant Retrieval?

S3 Glacier Instant Retrieval is designed for long-lived, rarely accessed data that is retained for months or years. Objects that are archived to S3 Glacier Instant Retrieval have a minimum of 90 days of storage, and objects deleted, overwritten, or transitioned before 90 days incur a pro-rated charge equal to the storage charge for the remaining days. View the Amazon S3 pricing page for information about Amazon S3 Glacier Instant Retrieval pricing.

Q: Is there a minimum object size charge for Amazon S3 Glacier Instant Retrieval?

S3 Glacier Instant Retrieval is designed for larger objects and has a minimum object storage charge of 128KB. Objects smaller than 128KB in size will incur storage charges as if the object were 128KB. For example, a 6KB object in S3 Glacier Instant Retrieval will incur S3 Glacier Instant Retrieval storage charges for 6KB and an additional minimum object size charge equivalent to 122KB at the S3 Glacier Instant Retrieval storage price. View the Amazon S3 pricing page for information about Amazon S3 Glacier Instant Retrieval pricing.

Q: How am I charged for S3 Glacier Instant Retrieval?

S3 Glacier Instant Retrieval charges you for monthly storage, requests based on the request type, and data retrievals. The volume of storage billed in a month is based on average storage used throughout the month, measured in gigabyte per month (GB-Month). You are charged for requests based on the request type—such as PUTs, COPYs, and GETs. You also pay a per GB fee for every gigabyte of data returned to you.

Amazon S3 Glacier Flexible Retrieval storage class

Q: What is the S3 Glacier Flexible Retrieval storage class?

The S3 Glacier Flexible Retrieval storage class delivers low-cost storage, up to 10% lower cost (than S3 Glacier Instant Retrieval), for archive data that is accessed 1-2 times per year and is retrieved asynchronously, with free bulk retrievals. For archive data that does not require immediate access but needs the flexibility to retrieve large sets of data at no cost, such as backup or disaster recovery use cases, S3 Glacier Flexible Retrieval is the ideal storage class. S3 Glacier Flexible Retrieval delivers the most flexible retrieval options that balance cost with access times ranging from minutes to hours and with free bulk retrievals. It is an ideal solution for backup, disaster recovery, offsite data storage needs, and for when some data needs to occasionally retrieved in minutes, and you don’t want to worry about costs. S3 Glacier Flexible Retrieval is designed for 99.999999999% (11 9s) of data durability and 99.99% availability by redundantly storing data across multiple physically separated AWS Availability Zones in a given year.

Q: Why would I choose to use S3 Glacier Flexible Retrieval storage class?

For archive data that does not require immediate access but needs the flexibility to retrieve large sets of data at no cost, such as backup or disaster recovery use cases, S3 Glacier Flexible Retrieval is the ideal storage class. S3 Glacier Flexible Retrieval delivers the most flexible retrieval options that balance cost with access times ranging from minutes to hours and with free bulk retrievals. It is an ideal solution for backup, disaster recovery, offsite data storage needs, and for when some data needs to occasionally retrieved in minutes, and you don’t want to worry about costs to retrieve the data.

Q: How do I get my into S3 Glacier Flexible Retrieval?

There are two ways to get data into S3 Glacier Flexible Retrieval. You can directly PUT into S3 Glacier Flexible Retrieval by specifying GLACIER in the x-amz-storage-class header. You can also use S3 Lifecycle rules to transition objects from any of the S3 storage classes for active data (S3 Standard, S3 Intelligent-Tiering, S3 Standard-IA, S3 One Zone-IA, and S3 Glacier Instant Retrieval) to Amazon S3 Glacier Flexible Retrieval based on object age. Use the Amazon S3 Management Console, the AWS SDKs, or the Amazon S3 APIs to directly PUT into Amazon S3 Glacier or define rules for archival.

Note: S3 Glacier Flexible Retrieval is also available through the original direct Glacier APIs and through the Amazon S3 Glacier Management Console. For an enhanced experience complete with access to the full S3 feature set including lifecycle management, S3 Replication, S3 Storage Lens, and more, we recommend using S3 APIs and the S3 Management Console to use S3 Glacier features.

Q: How can I retrieve my objects that are archived in S3 Glacier Flexible Retrieval and will I be notified when the object is restored?

Objects that are archived in S3 Glacier Flexible Retrieval are accessed asynchronously. To retrieve data stored in S3 Glacier Flexible Retrieval, initiate a retrieval request using the Amazon S3 APIs or the Amazon S3 console. The retrieval request creates a temporary copy of your data in the S3 Standard storage class while leaving the archived data intact in S3 Glacier Flexible Retrieval. You can specify the amount of time in days for which the temporary copy is stored in Amazon S3. You can then access your temporary copy from S3 through an Amazon S3 GET request on the archived object. In AWS Regions where Reduced Redundancy Storage is a lower price than S3 Standard, temporarily available data is billed as Reduced Redundancy Storage. However, the Reduced Redundancy billing storage class doesn't reflect how the data is stored.

With restore notifications, you can now be notified with an S3 Event Notification when an object has successfully restored from S3 Glacier Flexible Retrieval and the temporary copy is made available to you. The bucket owner (or others, as permitted by an IAM policy) can arrange for notifications to be issued to Amazon Simple Queue Service (SQS) or Amazon Simple Notification Service (SNS). Notifications can also be delivered to AWS Lambda for processing by a Lambda function.

Q: How long will it take to restore my objects archived in Amazon S3 Glacier Flexible Retrieval?

When processing a retrieval job, Amazon S3 first retrieves the requested data from S3 Glacier Flexible Retrieval, and then creates a temporary copy of the requested data in Amazon S3. This typically takes a few minutes. The access time of your request depends on the retrieval option you choose: Expedited, Standard, or Bulk retrievals. For all but the largest objects (250MB+), data accessed using Expedited retrievals are typically made available within 1-5 minutes. Objects retrieved using Standard retrievals typically complete between 3-5 hours. Standard retreivals typically start in minutes when initiated using S3 Batch Operations. Bulk retrievals typically complete within 5—12 hours, and are free of charge. For more information about the S3 Glacier Flexible Retrieval options, refer to restoring an archived object in the S3 user guide.

With S3 Glacier storage class provisioned capacity units, you can pay a fixed upfront fee for a given month to ensure the availability of retrieval capacity for expedited retrievals from S3 Glacier Flexible Retrieval. You can purchase two provisioned capacity units per month to increase the amount of data you can retrieve. Each unit of capacity ensures that at least three expedited retrievals can be performed every five minutes, and it provides up to 150 MB/s of retrieval throughput. If your workload requires highly reliable and predictable access to a subset of your data in minutes, you should purchase provisioned retrieval capacity. Without provisioned capacity, expedited retrievals might not be accepted during periods of high demand. If you require access to expedited retrievals under any circumstance, we recommend that you purchase provisioned retrieval capacity.

You can purchase provisioned capacity using the Amazon S3 console, the purchase provisioned capacity REST API, the AWS SDKs, or the AWS CLI. A provisioned capacity unit lasts for one month starting at the date and time of purchase, which is the start date. The unit expires on the expiration date, which is exactly one month after the start date to the nearest second. For provisioned capacity pricing information, see Amazon S3 pricing.

Q: How is my storage charge calculated for Amazon S3 objects archived to S3 Glacier Flexible Retrieval?

The volume of storage billed in a month is based on average storage used throughout the month, measured in gigabyte-months (GB-Months). Amazon S3 calculates the object size as the amount of data you stored, plus an additional 32 KB of S3 Glacier data, plus an additional 8 KB of Amazon S3 Standard storage class data. S3 Glacier Flexible Retrieval requires an additional 32 KB of data per object for S3 Glacier's index and metadata so you can identify and retrieve your data. Amazon S3 requires 8 KB to store and maintain the user-defined name and metadata for objects archived to S3 Glacier Flexible Retrieval. This enables you to get a real-time list of all of your Amazon S3 objects, including those stored using S3 Glacier Flexible Retrieval, using the Amazon S3 LIST API, or the S3 inventory report.

For example, if you have archived 100,000 objects that are 1 GB each, your billable storage would be:
1.000032 gigabytes for each object x 100,000 objects = 100,003.2 gigabytes of S3 Glacier storage.
0.000008 gigabytes for each object x 100,000 objects = 0.8 gigabytes of S3 Standard storage.

The fee is calculated based on the current rates for your AWS Region on the Amazon S3 pricing page. For additional Amazon S3 pricing examples, go to the S3 billing FAQs or use the AWS pricing calculator.

Q: Are there minimum storage duration and minimum object storage charges for Amazon S3 Glacier Flexible Retrieval?

Objects archived to S3 Glacier Flexible Retrieval have a minimum of 90 days of storage. If an object is deleted, overwritten, or transitioned before 90 days, a pro-rated charge equal to the storage charge for the remaining days will be incurred.

S3 Glacier Flexible Retrieval also requires 40 KB of additional metadata for each archived object. This includes 32 KB of metadata charged at the S3 Glacier Flexible Retrieval rate required to identify and retrieve your data. And, an additional 8 KB data charged at the S3 Standard rate which is required to maintain the user-defined name and metadata for objects archived to S3 Glacier Flexible Retrieval. This allows you to get a real-time list of all of your S3 objects using the S3 LIST API or the S3 Inventory report. View the Amazon S3 pricing page for information about Amazon S3 Glacier Flexible Retrieval pricing.

Q: How much does it cost to retrieve data from Amazon S3 Glacier Flexible Retrieval?

There are three ways to retrieve data from S3 Glacier Flexible Retrieval: Expedited, Standard, and Bulk Retrievals. Expedited and Standard have a per-GB retrieval fee and per-request fee (i.e., you pay for requests made against your Amazon S3 objects). Bulk Retrievals from S3 Glacier Flexible Retrieval are free. For detailed S3 Glacier pricing by AWS Region, visit the Amazon S3 pricing page.

Q: Does Amazon S3 provide capabilities for archiving objects to lower cost storage classes?

The Amazon S3 Glacier storage classes are purpose-built for data archiving, providing you with the highest performance, most retrieval flexibility, and the lowest cost archive storage in the cloud. You can now choose from three archive storage classes optimized for different access patterns and storage duration. For archive data that needs immediate access, such as medical images, news media assets, or genomics data, choose the S3 Glacier Instant Retrieval storage class, an archive storage class that delivers the lowest cost storage with milliseconds retrieval. For archive data that does not require immediate access but needs the flexibility to retrieve large sets of data at no cost, such as backup or disaster recovery use cases, choose S3 Glacier Flexible Retrieval, with retrieval in minutes or free bulk retrievals in 5—12 hours. To save even more on long-lived archive storage such as compliance archives and digital media preservation, choose S3 Glacier Deep Archive, the lowest cost storage in the cloud with data retrieval within 12 hours.

Q: What is the backend infrastructure supporting the S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive storage class?

We prefer to focus on the customer outcomes of performance, durability, availability, and security. However, this question is often asked by our customers. We use a number of different technologies which allow us to offer the prices we do to our customers. Our services are built using common data storage technologies specifically assembled into purpose-built, cost-optimized systems using AWS-developed software. The S3 Glacier storage classes benefit from our ability to optimize the sequence of inputs and outputs to maximize efficiency accessing the underlying storage.

Amazon S3 Glacier Deep Archive

Q: What is the Amazon S3 Glacier Deep Archive storage class?

S3 Glacier Deep Archive is an Amazon S3 storage class that provides secure and durable object storage for long-term retention of data that is accessed once or twice in a year. From just $0.00099 per GB-month (less than one-tenth of one cent, or about $1 per TB-month), S3 Glacier Deep Archive offers the lowest cost storage in the cloud, at prices significantly lower than storing and maintaining data in on-premises magnetic tape libraries or archiving data off-site.

Q: What use cases are best suited for the S3 Glacier Deep Archive storage class?

S3 Glacier Deep Archive is an ideal storage class to provide offline protection of your company’s most important data assets, or when long-term data retention is required for corporate policy, contractual, or regulatory compliance requirements. Customers find S3 Glacier Deep Archive to be a compelling choice to protect core intellectual property, financial and medical records, research results, legal documents, seismic exploration studies, and long-term backups, especially in highly regulated industries, such as Financial Services, Healthcare, Oil & Gas, and Public Sectors. In addition, there are organizations, such as media and entertainment companies, that want to keep a backup copy of core intellectual property. Frequently, customers using S3 Glacier Deep Archive can reduce or discontinue the use of on-premises magnetic tape libraries and off-premises tape archival services.

Q: How does the S3 Glacier Deep Archive storage class differ from the S3 Glacier Instant Retrieval, and S3 Glacier Flexible Retrieval storage classes?

S3 Glacier Deep Archive expands our data archiving offerings, enabling you to select the optimal storage class based on storage and retrieval costs, and retrieval times. Choose the S3 Glacier Instant Retrieval storage class when you need milliseconds access to low cost archive data. For archive data that does not require immediate access but needs the flexibility to retrieve large sets of data at no cost, such as backup or disaster recovery use cases, choose S3 Glacier Flexible Retrieval, with retrieval in minutes or free bulk retrievals in 5-12 hours. S3 Glacier Deep Archive, in contrast, is designed for colder data that is very unlikely to be accessed, but still requires long-term, durable storage. S3 Glacier Deep Archive is up to 75% less expensive than S3 Glacier Flexible Retrieval and provides retrieval within 12 hours using the Standard retrieval tier. Standard retreivals typically start within 9 hours when initiated using S3 Batch Operations. You may also reduce retrieval costs by selecting Bulk retrieval, which will return data within 48 hours.

Q: How do I get started using S3 Glacier Deep Archive?

The easiest way to store data in S3 Glacier Deep Archive is to use the S3 API to upload data directly. Just specify “S3 Glacier Deep Archive” as the storage class. You can accomplish this using the AWS Management Console, S3 REST API, AWS SDKs, or AWS Command Line Interface.

You can also begin using S3 Glacier Deep Archive by creating policies to migrate data using S3 Lifecycle, which provides the ability to define the lifecycle of your object and reduce your cost of storage. These policies can be set to migrate objects to S3 Glacier Deep Archive based on the age of the object. You can specify the policy for an S3 bucket, or for specific prefixes. Lifecycle transitions are billed at the S3 Glacier Deep Archive Upload price.

Tape Gateway, a cloud-based virtual tape library feature of AWS Storage Gateway, now integrates with S3 Glacier Deep Archive, enabling you to store your virtual tape-based, long-term backups and archives in S3 Glacier Deep Archive, thereby providing the lowest cost storage for this data in the cloud. To get started, create a new virtual tape using AWS Storage Gateway Console or API, and set the archival storage target either to S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive. When your backup application ejects the tape, the tape will be archived to your selected storage target.

Q: How do you recommend migrating data from my existing tape archives to S3 Glacier Deep Archive?

There are multiple ways to migrate data from existing tape archives to S3 Glacier Deep Archive. You can use the AWS Tape Gateway to integrate with existing backup applications using a virtual tape library (VTL) interface. This interface presents virtual tapes to the backup application. These can be immediately used to store data in Amazon S3, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive.

You can also use AWS Snowball to migrate data. Snowball accelerates moving terabytes to petabytes of data into and out of AWS using physical storage devices designed to be secure for transport. Using Snowball helps to eliminate challenges that can be encountered with large-scale data transfers including high network costs, long transfer times, and security concerns.

Finally, you can use AWS Direct Connect to establish dedicated network connections from your premises to AWS. In many cases, Direct Connect can reduce your network costs, increase bandwidth throughput, and provide a more consistent network experience than internet-based connections.

Q: How can I retrieve my objects stored in S3 Glacier Deep Archive?

To retrieve data stored in S3 Glacier Deep Archive, initiate a “Restore” request using the Amazon S3 APIs or the Amazon S3 Management Console. The Restore creates a temporary copy of your data in the S3 Standard storage class while leaving the archived data intact in S3 Glacier Deep Archive. You can specify the amount of time in days for which the temporary copy is stored in S3. You can then access your temporary copy from S3 through an Amazon S3 GET request on the archived object.

When restoring an archived object, you can specify one of the following options in the Tier element of the request body: Standard is the default tier and lets you access any of your archived objects within 12 hours, with retrievals typically starting within 9 hours when initiated using S3 Batch Operations. Bulk lets you retrieve large amounts of data, even petabytes of data, inexpensively and typically completes within 48 hours.

Q: How am I charged for using S3 Glacier Deep Archive?

S3 Glacier Deep Archive storage is priced based on the amount of data you store in GBs, the number of PUT/lifecycle transition requests, retrievals in GBs, and number of restore requests. This pricing model is similar to S3 Glacier Flexible Retrieval. See the [Amazon S3 pricing page]() for information about S3 Glacier Deep Archive pricing.

Q: How will S3 Glacier Deep Archive usage show up on my AWS bill and in the AWS Cost Management tool?

S3 Glacier Deep Archive usage and cost will show up as an independent service line item on your monthly AWS bill, separate from your Amazon S3 usage and costs. However, if you are using the AWS Cost Management tool, S3 Glacier Deep Archive usage and cost will be included under the Amazon S3 usage and cost in your detailed monthly spend reports, and not broken out as a separate service line item.

Q: Are there minimum storage duration and minimum object storage charges for S3 Glacier Deep Archive?

Objects that are archived to S3 Glacier Deep Archive have a minimum of 180 days of storage. If an object is deleted, overwritten, or transitioned before 180 days, a pro-rated charge equal to the storage charge for the remaining days will be incurred.

S3 Glacier Deep Archive also requires 40 KB of additional metadata for each archived object. This includes 32 KB of metadata charged at the S3 Glacier Deep Archive rate required to identify and retrieve your data. And, an additional 8 KB data charged at the S3 Standard rate which is required to maintain the user-defined name and metadata for objects archived to S3 Glacier Deep Archive. This allows you to get a real-time list of all of your S3 objects using the S3 LIST API or the S3 Inventory report. View the Amazon S3 pricing page for information about S3 Glacier Deep Archive pricing.

Q: How does S3 Glacier Deep Archive integrate with other AWS Services?

S3 Glacier Deep Archive is integrated with Amazon S3 features, including S3 Object Tagging, S3 Lifecycle policies, S3 Object Lock, and S3 Replication. With S3 storage management features, you can use a single Amazon S3 bucket to store a mixture of S3 Glacier Deep Archive, S3 Standard, S3 Standard-IA, S3 One Zone-IA, and S3 Glacier Flexible Retrieval data. This allows storage administrators to make decisions based on the nature of the data and data access patterns. Customers can use Amazon S3 Lifecycle policies to automatically migrate data to lower-cost storage classes as the data ages, or S3 Cross-Region Replication or Same-Region Replication policies to replicate data to the same or a different region.

AWS Storage Gateway service integrates Tape Gateway with S3 Glacier Deep Archive storage class, allowing you to store virtual tapes in the lowest-cost Amazon S3 storage class, reducing the monthly cost to store your long-term data in the cloud by 75%. With this feature, Tape Gateway supports archiving your new virtual tapes directly to S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive, helping you meet your backup, archive, and recovery requirements. Tape Gateway helps you move tape-based backups to AWS without making any changes to your existing backup workflows. Tape Gateway supports most of the leading backup applications such as Veritas, Veeam, Commvault, Dell EMC NetWorker, IBM Spectrum Protect (on Windows OS), and Microsoft Data Protection Manager.

S3 on Outposts

Q: What is Amazon S3 on Outposts?

Amazon S3 on Outposts delivers object storage in your on-premises environment, using the S3 APIs and capabilities that you use in AWS today. AWS Outposts is a fully managed service that extends AWS infrastructure, AWS services, APIs, and tools to virtually any datacenter, co-location space, or on-premises facility. Using S3 on Outposts, you can securely process and store customer data generated on-premises before moving it to an AWS Region, access data locally for applications that run on-premises, or store data on your Outpost for companies in locations with data residency requirements, and or those in regulated industries. To learn more about S3 on Outposts, visit the overview page.

Storage Management

S3 Object Tags | S3 Inventory | S3 Batch Operations | S3 CloudWatch Metrics | S3 Lifecycle Management

S3 Object Tags

Q: What are S3 Object Tags?

S3 Object Tags are key-value pairs applied to S3 objects which can be created, updated or deleted at any time during the lifetime of the object. With these, you have the ability to create Identity and Access Management (IAM) policies, set up S3 Lifecycle policies, and customize storage metrics. These object-level tags can then manage transitions between storage classes and expire objects in the background. You can add tags to new objects when you upload them or you can add them to existing objects. Up to ten tags can be added to each S3 object and you can use either the AWS Management Console, the REST API, the AWS CLI, or the AWS SDKs to add object tags.

Learn more by visiting the S3 Object Tags user guide.

Q: Why should I use object tags?

Object tags are a tool you can use to enable simple management of your S3 storage. With the ability to create, update, and delete tags at any time during the lifetime of your object, your storage can adapt to the needs of your business. These tags allow you to control access to objects tagged with specific key-value pairs, allowing you to further secure confidential data for only a select group or user. Object tags can also be used to label objects that belong to a specific project or business unit, which could be used in conjunction with S3 Lifecycle policies to manage transitions to other storage classes (S3 Standard-IA, S3 One Zone-IA, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, and S3 Glacier Deep Archive) or with S3 Replication to selectively replicate data between AWS Regions.

Q: How can I update the object tags on my objects?

Object tags can be changed at any time during the lifetime of your S3 object, you can use either the AWS Management Console, the REST API, the AWS CLI, or the AWS SDKs to change your object tags. Note that all changes to tags outside of the AWS Management Console are made to the full tag set. If you have five tags attached to a particular object and want to add a sixth, you need to include the original five tags in that request.

Q: How much do object tags cost?

Object tags are priced based on the quantity of tags and a request cost for adding tags. The requests associated with adding and updating Object Tags are priced the same as existing request prices. See the Amazon S3 pricing page for more information.

Q: How do I get started with Storage Class Analysis?

You can use the AWS Management Console or the S3 PUT Bucket Analytics API to configure a Storage Class Analysis policy to identify infrequently accessed storage that can be transitioned to the S3 Standard-IA or S3 One Zone-IA storage class or archived to the S3 Glacier storage classes. You can navigate to the “Management” tab in the S3 console to manage Storage Class Analysis, S3 Inventory, and S3 CloudWatch metrics.

S3 Inventory

Q: What is S3 Inventory?

The S3 Inventory report provides a scheduled alternative to Amazon S3’s synchronous List API. You can configure S3 Inventory to provide a CSV, ORC, or Parquet file output of your objects and their corresponding metadata on a daily or weekly basis for an S3 bucket or prefix. You can simplify and speed up business workflows and big data jobs with S3 Inventory. You can also use S3 inventory to verify encryption and replication status of your objects to meet business, compliance, and regulatory needs. Learn more at the Amazon S3 Inventory user guide.

Q: How do I get started with S3 Inventory?

You can use the AWS Management Console or the PUT Bucket Inventory Configuration API to configure a daily or weekly inventory report for all the objects within your S3 bucket or a subset of the objects under a shared prefix. As part of the configuration, you can specify a destination S3 bucket for your S3 Inventory report, the output file format (CSV, ORC, or Parquet), and specific object metadata necessary for your business application, such as object name, size, last modified date, storage class, version ID, delete marker, non-current version flag, multipart upload flag, replication status, or encryption status. You can use S3 Inventory as a direct input into your application workflows or Big Data jobs. You can also query S3 Inventory using Standard SQL language with Amazon Athena, Amazon Redshift Spectrum, and other tools such as Presto, Hive, and Spark.

Learn more at the Amazon S3 Inventory user guide.

Q: How am I charged for using S3 Inventory?

See the Amazon S3 pricing page for S3 Inventory pricing. Once you configure encryption using SSE-KMS, you will incur KMS charges for encryption, refer to the KMS pricing page for detail.

S3 Batch Operations

Q: What is S3 Batch Operations?

S3 Batch Operations is a feature that you can use to automate the execution of a single operation (like copying an object, or executing an AWS Lambda function) across many objects. With S3 Batch Operations, you can, with a few clicks in the S3 console or a single API request, make a change to billions of objects without having to write custom application code or run compute clusters for storage management applications. Not only does S3 Batch Operations administer your storage operation across many objects, S3 Batch Operations manages retries, displays progress, delivers notifications, provides a completion report, and sends events to AWS CloudTrail for all operations performed on your target objects. S3 Batch Operations can be used from the S3 console, or through the AWS CLI and SDK.

To learn more, visit the S3 Batch Operations page, or the user guide.

Q: How do I get started with S3 Batch Operations?

You can get started with S3 Batch Operations by going into the Amazon S3 console or using the AWS CLI or SDK to create your first S3 Batch Operations job. A S3 Batch Operations job consists of the list of objects to act upon and the type of operation to be performed (see the full list of available operations). Start by selecting an S3 Inventory report or providing your own custom list of objects for S3 Batch Operations to act upon. An S3 Inventory report is a file listing all objects stored in an S3 bucket or prefix. Next, you choose from a set of S3 operations supported by S3 Batch Operations, such as replacing tag sets, changing ACLs, copying storage from one bucket to another, or initiating a restore from S3 Glacier Flexible Retrieval to S3 Standard storage class. You can then customize your S3 Batch Operations jobs with specific parameters such as tag values, ACL grantees, and restoration duration. To further customize your storage actions, you can write your own Lambda function and invoke that code through S3 Batch Operations.

Once you create your S3 Batch Operations job, S3 Batch Operations will process your list of objects and send the job to the “awaiting confirmation” state if required. After you confirm the job details, S3 Batch Operations will begin executing the operation you specified. You can view your job’s progress programmatically or through the S3 console, receive notifications on completion, and review a completion report that itemizes the changes made to your storage.

If you are interested in learning more about S3 Batch Operations watch the tutorials videos and visit the documentation.

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

For customers in the financial services industry, S3 Object 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 Amazon S3 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.

S3 CloudWatch Metrics

Q: How do I get started with S3 CloudWatch Metrics?

You can use the AWS Management Console to enable the generation of one-minute CloudWatch request metrics for your S3 bucket or configure filters for the metrics using a prefix or object tag, or access point. Alternatively, you can call the S3 PUT Bucket Metrics API to enable and configure publication of S3 storage metrics. CloudWatch Request Metrics will be available in CloudWatch within 15 minutes after they are enabled. CloudWatch Storage Metrics are enabled by default for all buckets, and reported once per day. Learn more about CloudWatch metrics for Amazon S3.

Q: What alarms can I set on my storage metrics?

You can use CloudWatch to set thresholds on any of the storage metrics counts, timers, or rates and trigger an action when the threshold is breached. For example, you can set a threshold on the percentage of 4xx Error Responses and when at least three data points are above the threshold trigger a CloudWatch alarm to alert a DevOps engineer.

Q: How am I charged for using S3 CloudWatch Metrics?

CloudWatch storage metrics are provided free. Cloudwatch request metrics are priced as custom metrics for Amazon CloudWatch. See the Amazon CloudWatch pricing page for general information about S3 CloudWatch metrics pricing.

S3 Lifecycle Management

Q: What is S3 Lifecycle management?

S3 Lifecycle management provides the ability to define the lifecycle of your object with a predefined policy and reduce your cost of storage. You can set a lifecycle transition policy to automatically migrate objects stored in the S3 Standard storage class to the S3 Standard-IA, S3 One Zone-IA, and/or S3 Glacier storage classes based on the age of the data. You can also set lifecycle expiration policies to automatically remove objects based on the age of the object. You can set a policy for multipart upload expiration, which expires incomplete multipart uploads based on the age of the upload.

Learn more by visiting the S3 Lifecycle user guide.

Q: How do I set up an S3 Lifecycle management policy?

You can set up and manage Lifecycle policies in the AWS Management Console, S3 REST API, AWS SDKs, or AWS Command Line Interface (CLI). You can specify the policy at the prefix or at the bucket level.

Q: How can I use Amazon S3 Lifecycle management to help lower my Amazon S3 storage costs?

With Amazon S3 Lifecycle policies, you can configure your objects to be migrated from the S3 Standard storage class to S3 Standard-IA or S3 One Zone-IA and/or archived to S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, or S3 Glacier Deep Archive storage classes.

You can also specify an S3 Lifecycle policy to delete objects after a specific period of time. You can use this policy-driven automation to quickly and easily reduce storage costs as well as save time. In each rule you can specify a prefix, a time period, a transition to S3 Standard-IA, S3 One Zone-IA, S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, and/or an expiration. For example, you could create a rule that archives into S3 Glacier Flexible Retrieval all objects with the common prefix “logs/” 30 days from creation and expires these objects after 365 days from creation.

You can also create a separate rule that only expires all objects with the prefix “backups/” 90 days from creation. S3 Lifecycle policies apply to both existing and new S3 objects, helping you optimize storage and maximize cost savings for all current data and any new data placed in S3 without time-consuming manual data review and migration.

Within a lifecycle rule, the prefix field identifies the objects subject to the rule. To apply the rule to an individual object, specify the key name. To apply the rule to a set of objects, specify their common prefix (e.g. “logs/”). You can specify a transition action to have your objects archived and an expiration action to have your objects removed. For time period, provide the creation date (e.g. January 31, 2015) or the number of days from creation date (e.g. 30 days) after which you want your objects to be archived or removed. You may create multiple rules for different prefixes.

Q: How much does it cost to use S3 Lifecycle management?

There is no additional cost to set up and apply Lifecycle policies. A transition request is charged per object when an object becomes eligible for transition according to the Lifecycle rule. Refer to the Amazon S3 pricing page for pricing information.

Q: Why would I use an S3 Lifecycle policy to expire incomplete multipart uploads?

The S3 Lifecycle policy that expires incomplete multipart uploads allows you to save on costs by limiting the time non-completed multipart uploads are stored. For example, if your application uploads several multipart object parts, but never commits them, you will still be charged for that storage. This policy can lower your S3 storage bill by automatically removing incomplete multipart uploads and the associated storage after a predefined number of days.

Learn more about using S3 Lifecycle to expire incomplete multipart uploads »

Q: Can I set up Amazon S3 Event Notifications to send notifications when S3 Lifecycle transitions or expires objects?

Yes, you can set up Amazon S3 Event Notifications to notify you when S3 Lifecycle transitions or expires objects. For example, you can send S3 Event Notifications to an Amazon SNS topic, Amazon SQS queue, or AWS Lambda function when S3 Lifecycle moves objects to a different S3 storage class or expires objects.

Storage Analytics & Insights

Q: What features are available to analyze my storage usage on Amazon S3?

S3 Storage Lens delivers organization-wide visibility into object storage usage, activity trends, and makes actionable recommendations to optimize costs and apply data protection best practices. S3 Storage Class Analysis enables you to monitor access patterns across objects to help you decide when to transition data to the right storage class to optimize costs. You can then use this information to configure an S3 Lifecycle policy that makes the data transfer. Amazon S3 Inventory provides a report of your objects and their corresponding metadata on a daily or weekly basis for an S3 bucket or prefix. This report can be used to help meet business, compliance, and regulatory needs by verifying the encryption, and replication status of your objects.

Q: What is Amazon S3 Storage Lens?

Amazon S3 Storage Lens provides organization-wide visibility into object storage usage and activity trends, as well as actionable recommendations to optimize costs and apply data protection best practices. Storage Lens offers an interactive dashboard containing a single view of your object storage usage and activity across tens or hundreds of accounts in your organization, with drill-downs to generate insights at multiple aggregation levels. This includes metrics like bytes, object counts, and requests, as well as metrics detailing S3 feature utilization, such as encrypted object counts and S3 Lifecycle rule counts. S3 Storage Lens also delivers contextual recommendations to find ways for you to reduce storage costs and apply best practices on data protection across tens or hundreds of accounts and buckets. S3 Storage Lens free metrics are enabled by default for all Amazon S3 users. If you want to get more out of S3 Storage Lens, you can activate advanced metrics and recommendations. Learn more by visiting the S3 Storage Lens user guide.

Q: How does S3 Storage Lens work?

S3 Storage Lens aggregates your storage usage and activity metrics on a daily basis to be visualized in the S3 Storage Lens interactive dashboard, or available as a metrics export in CSV or Parquet file format. A default dashboard is created for you automatically at the account level, and you have the option to create additional custom dashboards. S3 Storage Lens dashboards can be scoped to your AWS organization or specific accounts, Regions, buckets, or even prefix level (available with S3 Storage Lens advanced metrics). You can also use S3 Storage Lens groups to aggregate metrics using custom filters based on object metadata like object tag, size, and age. While configuring your dashboard you can use the default metrics selection, or upgrade to receive 35 additional metrics and prefix-level aggregations for an additional cost. Also, S3 Storage Lens provides recommendations contextually with storage metrics in the dashboard, so you can take action to optimize your storage based on the metrics.

Q: What are the key questions that can be answered using S3 Storage Lens metrics?

The S3 Storage Lens dashboard is organized around four main types of questions that can be answered about your storage. With the Summary filter, top-level questions related to overall storage usage and activity trends can be explored. For example, “How rapidly is my overall byte count and request count increasing over time?” With the Cost Optimization filter, you can explore questions related to storage cost reduction, for example, “Is it possible for me to save money by retaining fewer non-current versions?” With the Data Protection and Access Management filters you can answer questions about securing your data, for example, “Is my storage protected from accidental or intentional deletion?” Finally, with the Performance and Events filters you can explore ways to improve performance of workflows. Each of these questions represent a first layer of inquiry that would likely lead to drill-down analysis.

Q: What metrics are available in S3 Storage Lens?

S3 Storage Lens contains more than 60 metrics, grouped into free metrics and advanced metrics (available for an additional cost). Within free metrics, you receive metrics to analyze usage (based on a daily snapshot of your objects), which are organized into the categories of cost optimization, data protection, access management, performance, and events. Within advanced metrics, you receive metrics related to activity (such as request counts), deeper cost optimization (such as S3 Lifecycle rule counts), additional data protection (such as S3 Replication rule counts), and detailed status codes (such as 403 authorization errors). In addition, derived metrics are also provided by combining any base metrics. For example, “Retrieval Rate" is a metric calculated by dividing the "Bytes Downloaded Count" by the "Total Storage.” To view the complete list of metrics, visit the S3 Storage Lens documentation.

Q: What are my dashboard configuration options?

A default dashboard is configured automatically provided for your entire account, and you have the option to create additional custom dashboards that can be scoped to your AWS organization, specific regions, or buckets within an account. You can set up multiple custom dashboards, which can be useful if you require some logical separation in your storage analysis, such as segmenting on buckets to represent various internal teams. By default, your dashboard will receive the S3 Storage Lens free metrics, but you have the option to upgrade to receive S3 Storage Lens advanced metrics and recommendations (for an additional cost). S3 Storage Lens advanced metrics have 7 distinct options: Activity metrics, Advanced Cost Optimization metrics, Advanced Data Protection metrics, Detailed Status Code metrics, Prefix aggregation, CloudWatch publishing, and Storage Lens groups aggregation. Additionally, for each dashboard you can enable metrics export, with additional options to specify destination bucket and encryption type.

Q: How much historical data is available in S3 Storage Lens?

For metrics displayed in the interactive dashboard, Storage Lens free metrics retains 14 days of historical data, and Storage Lens advanced metrics (for an additional cost) retains 15 months of historical data. For the optional metrics export, you can configure any retention period you wish, and standard S3 storage charges will apply.

Q: How will I be charged for S3 Storage Lens?

S3 Storage Lens is available in two tiers of metrics. The free metrics are enabled by default and available at no additional charge to all S3 customers. The S3 Storage Lens advanced metrics and recommendations pricing details are available on the S3 pricing page. With S3 Storage Lens free metrics you receive 28 metrics at the bucket level, and can access 14 days of historical data in the dashboard. With S3 Storage Lens advanced metrics and recommendations you receive 35 additional metrics, prefix-level aggregation, CloudWatch metrics support, custom object metadata filtering with S3 Storage Lens groups, and can access 15 months of historical data in the dashboard.

Q: What is the difference between S3 Storage Lens and S3 Inventory?

S3 Inventory provides a list of your objects and their corresponding metadata for an S3 bucket or a shared prefix, which can be used to perform object-level analysis of your storage. S3 Storage Lens provides metrics that can be aggregated by organization, account, region, storage class, bucket, prefix, and S3 Storage Lens group levels, which improve organization-wide visibility of your storage.

Q: What is the difference between S3 Storage Lens and S3 Storage Class Analysis (SCA)?

S3 Storage Class Analysis provides recommendations for an optimal storage class by creating object age groups based on object-level access patterns within an individual bucket/prefix/tag for the previous 30-90 days. S3 Storage Lens provides daily organization level recommendations on ways to improve cost efficiency and apply data protection best practices, with additional granular recommendations by account, region, storage class, bucket, S3 Storage Lens group, or prefix (available with S3 Storage Lens advanced metrics). You can also use custom filters with S3 Storage Lens groups to visualize your storage based on object age and inform your storage archival strategy.

Storage Class Analysis

Q: What is Storage Class Analysis?

With Storage Class Analysis, you can analyze storage access patterns to determine the optimal storage class for your storage. This S3 feature automatically identifies infrequent access patterns to help you transition storage to S3 Standard-IA. You can configure a Storage Class Analysis policy to monitor an entire bucket, prefix, or object tag. Once an infrequent access pattern is observed, you can easily create a new S3 Lifecycle age policy based on the results. Storage Class Analysis also provides daily visualizations of your storage usage on the AWS Management Console and you can also enable an export report to an S3 bucket to analyze using business intelligence tools of your choice such as Amazon QuickSight.

Learn more and get started by visiting the S3 Storage Class Analysis user guide.

Q: How often is the Storage Class Analysis updated?

Storage Class Analysis is updated on a daily basis in the S3 Management Console, but initial recommendations for storage class transitions are provided after 30 days.

Query in Place

Q: What is "Query in Place" functionality?

Amazon S3 allows customers to run sophisticated queries against data stored without the need to move data into a separate analytics platform. The ability to query this data in place on Amazon S3 can significantly increase performance and reduce cost for analytics solutions leveraging S3 as a data lake. S3 offers multiple query in place options, including S3 Select, Amazon Athena, and Amazon Redshift Spectrum, allowing you to choose one that best fits your use case. You can even use Amazon S3 Select with AWS Lambda to build serverless apps that can take advantage of the in-place processing capabilities provided by S3 Select.

Q: What is S3 Select?

S3 Select is an Amazon S3 feature that makes it easy to retrieve specific data from the contents of an object using simple SQL expressions without having to retrieve the entire object. S3 Select simplifies and improves the performance of scanning and filtering the contents of objects into a smaller, targeted dataset by up to 400%. With S3 Select, you can also perform operational investigations on log files in Amazon S3 without the need to operate or manage a compute cluster.

You can use S3 Select to retrieve a subset of data using SQL clauses, like SELECT and WHERE, from objects stored in CSV, JSON, or Apache Parquet format. It also works with objects that are compressed with GZIP or BZIP2 (for CSV and JSON objects only), and server-side encrypted objects.

You can use S3 Select with AWS Lambda to build serverless applications that use S3 Select to efficiently and easily retrieve data from Amazon S3 instead of retrieving and processing entire object. You can also use S3 Select with Big Data frameworks, such as Presto, Apache Hive, and Apache Spark to scan and filter the data in Amazon S3.

Learn more by visiting the S3 Select user guide.

Q: What is Amazon Athena?

Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL queries. Athena is serverless, so there is no infrastructure to set up or manage, and you can start analyzing data immediately. You don’t even need to load your data into Athena; it works directly with data stored in any S3 storage class. To get started, just log into the Athena Management Console, define your schema, and start querying. Amazon Athena uses Presto with full standard SQL support and works with a variety of standard data formats, including CSV, JSON, ORC, Apache Parquet and Avro. While Athena is ideal for quick, ad-hoc querying and integrates with Amazon QuickSight for easy visualization, it can also handle complex analysis, including large joins, window functions, and arrays.

Q: What is Amazon Redshift Spectrum?

Amazon Redshift Spectrum is a feature of Amazon Redshift that lets you run queries against exabytes of unstructured data in Amazon S3 with no loading or ETL required. When you issue a query, it goes to the Amazon Redshift SQL endpoint, which generates and optimizes a query plan. Amazon Redshift determines what data is local and what is in Amazon S3, generates a plan to minimize the amount of Amazon S3 data that needs to be read, and requests Redshift Spectrum workers out of a shared resource pool to read and process data from Amazon S3.

Redshift Spectrum scales out to thousands of instances if needed, so queries run quickly regardless of data size. And, you can use the exact same SQL for Amazon S3 data as you do for your Amazon Redshift queries today and connect to the same Amazon Redshift endpoint using the same business intelligence tools. Redshift Spectrum lets you separate storage and compute, allowing you to scale each independently. You can set up as many Amazon Redshift clusters as you need to query your Amazon S3 data lake, providing high availability and limitless concurrency. Redshift Spectrum gives you the freedom to store your data where you want, in the format you want, and have it available for processing when you need it.

Replication

S3 Replication Time Control | S3 Multi-Region Access Points

Q: What is Amazon S3 Replication?

Amazon S3 Replication enables automatic, asynchronous copying of objects across Amazon S3 buckets. Buckets that are configured for object replication can be owned by the same AWS account or by different accounts. You can replicate new objects written to the bucket to one or more destination buckets between different AWS Regions (S3 Cross-Region Replication), or within the same AWS Region (S3 Same-Region Replication). You can also replicate existing bucket contents (S3 Batch Replication), including existing objects, objects that previously failed to replicate, and objects replicated from another source. Learn more by visiting the S3 Replication user guide.

Q: What is Amazon S3 Cross-Region Replication (CRR)?

CRR is an Amazon S3 feature that automatically replicates data between buckets across different AWS Regions. With CRR, you can set up replication at a bucket level, a shared prefix level, or an object level using S3 object tags. You can use CRR to provide lower-latency data access in different geographic regions. CRR can also help if you have a compliance requirement to store copies of data hundreds of miles apart. You can use CRR to change account ownership for the replicated objects to protect data from accidental deletion. To learn more visit the S3 CRR user guide.

Q: What is Amazon S3 Same-Region Replication (SRR)?

SRR is an Amazon S3 feature that automatically replicates data between buckets within the same AWS Region. With SRR, you can set up replication at a bucket level, a shared prefix level, or an object level using S3 object tags. You can use SRR to create one or more copies of your data in the same AWS Region. SRR helps you address data sovereignty and compliance requirements by keeping a copy of your data in a separate AWS account in the same region as the original. You can use SRR to change account ownership for the replicated objects to protect data from accidental deletion. You can also use SRR to easily aggregate logs from different S3 buckets for in-region processing, or to configure live replication between test and development environments. To learn more visit the S3 SRR user guide.

Q: What is Amazon S3 Batch Replication?

Amazon S3 Batch Replication replicates existing objects between buckets. You can use S3 Batch Replication to backfill a newly created bucket with existing objects, retry objects that were previously unable to replicate, migrate data across accounts, or add new buckets to your data lake. You can get started with S3 Batch Replication with just a few clicks in the S3 console or a single API request. To learn more visit the S3 Batch Replication user guide.

Q: How do I enable Amazon S3 Replication (Cross-Region Replication and Same-Region Replication)?

Amazon S3 Replication (CRR and SRR) is configured at the S3 bucket level, a shared prefix level, or an object level using S3 object tags. You add a replication configuration on your source bucket by specifying a destination bucket in the same or different AWS Region for replication.

You can use the S3 console, API, the AWS CLI, AWS SDKs, or AWS CloudFormation to enable replication. Versioning must be enabled for both the source and destination buckets to enable replication. To learn more, visit overview of setting up S3 Replication in the Amazon S3 documentation.

Q: How do I use S3 Batch Replication?

You would first need to enable S3 Replication at the bucket level. See the previous question for how you can do so. You may then initiate an S3 Batch Replication job in the S3 console after creating a new replication configuration, changing a replication destination in a replication rule from the replication configuration page, or from the S3 Batch Operations Create Job page. Alternatively, you can initiate an S3 Batch Replication jobs via the AWS CLI or SDKs. To learn more, visit S3 Replication in the Amazon S3 documentation.

Q: Can I use S3 Replication with S3 Lifecycle rules?

With S3 Replication, you can establish replication rules to make copies of your objects into another storage class, in the same or a different region. Lifecycle actions are not replicated, and if you want the same lifecycle configuration applied to both source and destination buckets, enable the same lifecycle configuration on both.

For example, you can configure a lifecycle rule to migrate data from the S3 Standard storage class to the S3 Standard-IA or S3 One Zone-IA storage class or archive data to a S3 Glacier storage class on the destination bucket.

If you have S3 Lifecycle configured for your destination bucket, we recommend disabling Lifecycle rules while the Batch Replication job is active to maintain parity between noncurrent and current versions of objects in the source and destination buckets.

You can find more information about lifecycle configuration and replication in the S3 Replication documentation.

Q: Can I use S3 Replication to replicate to more than one destination bucket?

Yes. S3 Replication allows customers to replicate their data to multiple destination buckets in the same, or different AWS Regions. When setting up, you simply specify the new destination bucket in your existing replication configuration or create a new replication configuration with multiple destination buckets. For each new destination you specify, you have the flexibility to choose storage class of destination bucket, encryption type, replication metrics and notifications, Replication Time Control (RTC), and other properties.

Q: Can I use S3 Replication to set up two-way replication between S3 buckets?

Yes. To set up two-way replication, you create a replicate rule from bucket A to bucket B and set up another replication rule from bucket B to bucket A. Make sure to enable replica modification sync on both buckets A and B to replicate replica metadata changes like object access control lists (ACLs), object tags, or object locks on the replicated objects.

Q: Can I use replication across AWS accounts to protect against malicious or accidental deletion?

Yes, for CRR and SRR, you can set up replication across AWS accounts to store your replicated data in a different account in the target region. You can use Ownership Overwrite in your replication configuration to maintain a distinct ownership stack between source and destination, and grant destination account ownership to the replicated storage.

Q: Will my object tags be replicated if I use Cross-Region Replication?

Object tags can be replicated across AWS Regions using Cross-Region Replication. For customers with Cross-Region Replication already enabled, new permissions are required in order for tags to replicate. For more information about setting up Cross-Region Replication, visit How to Set Up Cross-Region Replication in the Amazon S3 documentation.

Q: Can I replicate delete markers from one bucket to another?

Yes, you can replicate delete markers from source to destination if you have delete marker replication enabled in your replication configuration. When you replicate delete markers, Amazon S3 will behave as if the object was deleted in both buckets. You can enable delete marker replication for a new or existing replication rule. You can apply delete marker replication to the entire bucket or to Amazon S3 objects that have a specific prefix, with prefix based replication rules. Amazon S3 Replication does not support delete marker replication for object tag based replication rules. To learn more about enabling delete marker replication see Replicating delete markers from one bucket to another.

Q: Can I replicate data from other AWS Regions to China? Can a customer replicate from one China Region bucket outside of China Regions?

No, Amazon S3 Replication is not available between AWS China Regions and AWS Regions outside of China. You are only able to replicate within the China regions.

Q: Can I replicate existing objects?

Yes. You can use S3 Batch Replication to replicate existing objects between buckets. To learn more, visit the S3 User Guide.

Q: Can I re-try replication if object fail to replicate initially?

Yes. You can use S3 Batch Replication to re-replicate objects that fail to replicate initially. To learn more, visit the S3 User Guide.

Q: What encryption types does S3 Replication support?

S3 Replication supports all encryption types that S3 offers. S3 offers both server-side encryption and client-side encryption – the former requests S3 to encrypt the objects for you, and the latter is for you to encrypt data on the client-side before uploading it to S3. For server-side encryption, S3 offers server-side encryption with Amazon S3-managed keys (SSE-S3), server-side encryption with KMS keys stored in AWS Key Management Service (SSE-KMS), and server-side encryption with customer-provided keys (SSE-C). For further details on these encryption types and how they work, visit the S3 documentation on using encryption.

Q: What is the pricing for cross account data replication?

With S3 Replication, you can configure cross account replication where the source and destination buckets are owned by different AWS accounts. Excluding S3 storage and applicable retrieval charges, customers pay for replication PUT requests and inter-region Data Transfer OUT from S3 to your destination region when using S3 Replication. If you have S3 Replication Time Control (S3 RTC) enabled on your replication rules, you will see a different Data Transfer OUT and replication PUT request charges specific to S3 RTC. For cross account replication, the source account pays for all data transfer (S3 RTC and S3 CRR) and the destination account pays for the replication PUT requests. Data transfer charges only apply for S3 Cross Region Replication (S3 CRR) and S3 Replication Time Control (S3 RTC), there are no data transfer charges for S3 Same Region Replication (S3 SRR).

If you are using S3 Batch Replication to replicate objects across accounts, you will incur the S3 Batch Operations charges, in addition to the replication PUT requests and Data Transfer OUT charges (note that S3 RTC is not applicable to Batch Replication). The Batch Operations charges include the Job and Object charges, which are respectively based on the number of jobs and number of objects processed. Additionally, if you opt in for the Amazon S3 generated manifest, you will incur a manifest charge based on the number of objects in the source bucket.

Visit the Amazon S3 pricing page for more details on S3 Replication pricing.

S3 Replication Time Control

Q: What is Amazon S3 Replication Time Control?

Amazon S3 Replication Time Control provides predictable replication performance and helps you meet compliance or business requirements. S3 Replication Time Control is designed to replicate most objects in seconds, and 99.99% of objects within 15 minutes. S3 Replication Time Control is backed by a Service Level Agreement (SLA) commitment that 99.9% of objects will be replicated in 15 minutes for each replication region pair during any billing month. Replication Time works with all S3 Replication features. To learn more, visit the replication documentation.

Q: How do I enable Amazon S3 Replication Time Control?

Amazon S3 Replication Time Control is enabled as an option for each replication rule. You can create a new S3 Replication policy with S3 Replication Time Control, or enable the feature on an existing policy. You can use either the S3 console, API, AWS CLI, AWS SDKs, or AWS CloudFormation to configure replication. To learn more, please visit overview of setting up Replication in the Amazon S3 developer guide.

Q: Can I use S3 Replication Time Control to replicate data within and between China Regions?

Yes, you can enable Amazon S3 Replication Time Control to replicate data within and between the AWS China (Ningxia) and China (Beijing) Regions.

Q: What are Amazon S3 Replication metrics and events?

Amazon S3 Replication provides four detailed metrics in the Amazon S3 console and in Amazon CloudWatch: operations pending, bytes pending, replication latency, and operations failed replication. You can use these metrics to monitor the total number of operations and size of objects that are pending to replicate, the replication latency between source and destination buckets, and the number of operations that did not replicate successfully for each replication rule. Additionally, you can set up Amazon S3 Event Notifications of *s3:Replication* type to get more information about objects that failed to replicate and the reason behind the failures. We recommend using Amazon S3 replication failure reasons to diagnose the errors quickly and fix them before re-replicating the failed objects with S3 Batch Replication. Finally, if you have S3 Replication Time Control (S3 RTC) enabled you will receive an S3 Event Notification when an object takes more than 15 minutes to replicate, and another when that object replicates successfully to the destination.

Q: How do I enable Amazon S3 Replication metrics and events?

Amazon S3 Replication metrics and events can be enabled for each new or existing replication rules, and are enabled by default for S3 Replication Time Control enabled rules. You can access S3 Replication metrics through the Amazon S3 console and Amazon CloudWatch. Like other Amazon S3 events, S3 Replication events are available through Amazon Simple Queue Service (Amazon SQS), Amazon Simple Notification Service (Amazon SNS), or AWS Lambda. To learn more, please visit Monitoring progress with replication metrics and Amazon S3 Event Notifications in the Amazon S3 developer guide.

Q. What information does the operations failed replication metric show?

The operations failed replication metric will show the total number of operations failing replication per minute for a specific replication rule. The metric will refresh every minute to emit +1 for each failed operation, 0 for successful operations, and nothing when there are no replication operations carried out for the minute. This metric is emitted every time an operation does not replicate successfully.

Q: Can I use Amazon S3 Replication metrics and events to track S3 Batch Replication?

You cannot use metrics like bytes pending, operations pending, and replication latency to track S3 Batch Replication progress. However, you can use the operations failed replication metric to monitor existing objects that do not replicate successfully with S3 Batch Replication. Additionally, you can also use S3 Batch Operations completion reports to keep track of objects replicating with S3 Batch Replication.

Q. Where are Amazon S3 Replication metrics published?

The bytes pending, operations pending, and replication latency metrics are published in the source AWS account and destination AWS Region. However, the operations failed replication metric is published in the source AWS account and source AWS Region instead of the destination AWS Region. There are two primary reasons for this. First, if the operations failed replication metric is published in the destination Region the customer will not see the metric when the destination bucket is erroneously configured. For example, if the customer has mis-typed the destination bucket name in the replication configuration and replication is unsuccessful because the destination bucket is not found, the customer will not be able to see any value for this metric because the destination Region will be unknown when the destination bucket is not found. Second, if the customer is replicating to an opt-in destination Region like Hong Kong or Bahrain, in the event of replication failures the customer will not see any metric if the source account has not opted-in for the destination Region.

Q: What is the Amazon S3 Replication Time Control Service Level Agreement (SLA)?

Amazon S3 Replication Time Control is designed to replicate 99.99% of your objects within 15 minutes, and is backed by a service level agreement. If fewer than 99.9% of your objects are replicated in 15 minutes for each replication region pair during a monthly billing cycle, the S3 RTC SLA provides a service credit on any object that takes longer than 15 minutes to replicate. The service credit covers a percentage of all replication-related charges associated with the objects that did not meet the SLA, including the RTC charge, replication bandwidth and request charges, and the cost associated with storing your replica in the destination region in the monthly billing cycle affected. To learn more, read the S3 Replication Time Control SLA.

Q: What is the pricing for S3 Replication and S3 Replication Time Control?

For S3 Replication (Cross-Region Replication and Same Region Replication), you pay the S3 charges for storage in the selected destination S3 storage classes, the storage charges for the primary copy, replication PUT requests, and applicable infrequent access storage retrieval charges. For CRR, you also pay for inter-region Data Transfer OUT From S3 to your destination region. S3 Replication Metrics are billed at the same rate as Amazon CloudWatch custom metrics. Additionally, when you use S3 Replication Time Control, you also pay a Replication Time Control Data Transfer charge. For more information, visit the Amazon S3 pricing page.

If the source object is uploaded using the multipart upload feature, then it is replicated using the same number of parts and part size. For example, a 100 GB object uploaded using the multipart upload feature (800 parts of 128 MB each) will incur request cost associated with 802 requests (800 Upload Part requests + 1 Initiate Multipart Upload request + 1 Complete Multipart Upload request) when replicated. You will incur a request charge of $0.00401 (802 requests x $0.005 per 1,000 requests) and (if the replication was between different AWS Regions) a charge of $2.00 ($0.020 per GB transferred x 100 GB) for inter-region data transfer. After replication, the 100 GB will incur storage charges based on the destination region.

Q. How am I charged for S3 Replication metrics on Amazon CloudWatch?

All S3 Replication metrics, including bytes pending, operations pending, replication latency, and operations failed replication, are billed at the same rate as Amazon CloudWatch Custom metrics: $0.30 per metric per month for the first 10K metrics, $0.10 per metric per month for the next 240K metrics, $0.05 per metric per month for the next 750K metrics, and $0.02 per metric per month for over 1M metrics.

For example, if your S3 bucket has 100 replication rules with Replication Metrics enabled for each rule, you will see a monthly Amazon CloudWatch charge for 400 replication metrics (100 replication rules x 4 metrics per replication rule). The monthly prorated charge for these 400 metrics will be $120 (400 replication metrics x $0.30 per metric (for the first 10K metrics)). For information on Amazon CloudWatch billing, see the Amazon CloudWatch Pricing page.

S3 Multi-Region Access Points

Q: What are S3 Multi-Region Access Points?

Amazon S3 Multi-Region Access Points accelerate performance by up to 60% when accessing data sets that are replicated across multiple AWS Regions. Based on AWS Global Accelerator, S3 Multi-Region Access Points consider factors like network congestion and the location of the requesting application to dynamically route your requests over the AWS network to the lowest latency copy of your data. This automatic routing allows you to take advantage of the global infrastructure of AWS while maintaining a simple application architecture.

Q: Why should I use S3 Multi-Region Access Points?

S3 Multi-Region Access Points accelerate and simplify storage for your multi-region applications. By dynamically routing S3 requests made to a replicated data set, S3 Multi-Region Access Points reduce request latency, so that applications run up to 60% faster. S3 Multi-Region Access Points can also help you build resilient, multi-region and multi-account applications that are more protected against accidental or unauthorized data deletion. With S3 Multi-Region Access Points, you are able to take advantage of the global infrastructure of AWS while maintaining a simple region-agnostic architecture for your applications.

Q: How do S3 Multi-Region Access Points work?

Multi-Region Access Points dynamically route client requests to one or more underlying S3 buckets. You can configure your Multi-Region Access Point to route across one bucket per AWS Region, in up to 20 AWS Regions. When you create a Multi-Region Access Point, S3 automatically generates a DNS-compatible name. This name is used as a global endpoint that can be used by your clients. When your clients make requests to this endpoint, S3 will dynamically route those requests to one of the underlying buckets that are specified in the configuration of your Multi-Region Access Point. Internet-based requests are onboarded to the AWS global network to avoid congested network segments on the internet, which reduces network latency and jitter while improving performance. Based on AWS Global Accelerator, applications that access S3 over the internet can see performance further improved up to 60% by S3 Multi-Region Access Points.

To directly control this routing, you can operate S3 Multi-Region Access Points in an active-active or active-passive configuration. In an active-passive configuration, you can use S3 Multi-Region Access Points failover controls to initiate a failover to shift S3 data access request traffic to the chosen alternate AWS Region and account within minutes.

In an active-active configuration, S3 Multi-Region Access Points consider factors like network congestion and the location of the requesting application to dynamically route your requests over the AWS network to the closest copy of your data. S3 Multi-Region Access Points route your requests through the closest AWS location to your client, and then over the global private AWS network to S3.

In either configuration, S3 Multi-Region Access Points allow you to take advantage of the global infrastructure of AWS while maintaining a simple application architecture.

Q. How do S3 Multi-Region Access Points failover controls work?

By default, S3 Multi-Region Access Points route requests to the underlying bucket closest to the client, based on network latency in an active-active configuration. For example, you can configure a Multi-Region Access Point with underlying buckets in US East (N. Virginia) and in Asia Pacific (Mumbai). With this configuration, your clients in North America route to US East (N. Virginia), while your clients in Asia route to Asia Pacific (Mumbai). This lowers latency for your requests made to S3, improving the performance of your application. If you prefer an active-passive configuration, all S3 data request traffic can be routed through the S3 Multi-Region Access Point to US East (N. Virginia) as the active Region and no traffic will be routed to Asia Pacific (Mumbai). If there is a planned or unplanned need to failover all of the S3 data request traffic to Asia Pacific (Mumbai), you can initiate a failover to switch to Asia Pacific (Mumbai) as the new active Region within minutes. Any existing uploads or downloads in progress in US East (N. Virginia) continue to completion and all new S3 data request traffic through the S3 Multi-Region Access Point is routed to Asia Pacific (Mumbai).

Q. Can S3 Multi-Region Access Points work with buckets owned by different AWS accounts?

Yes, you can add buckets in multiple AWS accounts to a new S3 Multi-Region Access Point by entering the account IDs that own the buckets at the time of creation. If the buckets are not already configured for cross-account replication, you can then configure S3 Cross-Region Replication rules to synchronize the contents of the buckets across AWS accounts and Regions. Your applications will be then able to request or write data through the Multi-Region Access Point global endpoint across AWS accounts and Regions.

Q. How do Block Public Access settings work for Multi-Region Access Points that span multiple AWS accounts?

Each S3 Multi-Region Access Point has distinct settings for Amazon S3 Block Public Access. These settings operate in conjunction with the Block Public Access settings for the buckets that underlie the Multi-Region Access Point, the Block Public Access settings for the AWS accounts that owns the Multi-Region Access Point, and the Block Public Access settings for the AWS accounts that own underlying buckets.

When Amazon S3 authorizes a request, it applies the most restrictive combination of these settings. If the Block Public Access settings for any of these resources (the Multi-Region Access Point, the underlying bucket, the Multi-Region Access Point owner account, or the bucket owner account) block access for the requested action or resource, Amazon S3 rejects the request.

This behavior is consistent with cross-account S3 Access Points. The same authorization logic is applied in when serving requests for cross-account S3 Access Points and cross-account S3 Multi-Region Points.

Q: What is the difference between S3 Cross-Region Replication (S3 CRR) and S3 Multi-Region Access Points?

S3 CRR and S3 Multi-Region Access Points are complementary features that work together to replicate data across AWS Regions and then to automatically route requests to the replicated copy with the lowest latency. S3 Multi-Region Access Points help you to manage requests across AWS Regions, while CRR allows you to move data across AWS Regions to create isolated replicas. You use S3 Multi-Region Access Points and CRR together to create a replicated multi-Region dataset that is addressable by a single global endpoint.

Q: How much do S3 Multi-Region Access Points cost?

When you use an S3 Multi-Region Access Point to route requests within AWS, you pay a low per-GB data routing charge for each GB processed, as well as standard charges for S3 requests, storage, data transfer, and replication. If your application runs outside of AWS and accesses S3 over the internet, S3 Multi-Region Access Points increase performance by automatically routing your requests through an AWS edge location, over the global private AWS network, to the closest copy of your data based on access latency. When you accelerate requests made over the internet, you pay the data routing charge and an internet acceleration charge. S3 Multi-Region Access Points internet acceleration pricing varies based on whether the source client is in the same or in a different location as the destination AWS Region, and is in addition to standard S3 data transfer pricing. To use S3 Multi-Region Access Points failover controls, you are only charged for standard S3 API costs to view the current routing control status of each Region and submit any routing control changes for initiating a failover. See the Amazon S3 pricing page and the data transfer tab for more pricing information.

Q: Can I use Requester Pays with S3 Multi-Region Access Points?

Yes, you can configure the underlying buckets of the S3 Multi-Region Access Point to be Requester Pays buckets. With Requester Pays, the requester pays all of the cost associated to the endpoint usage, including the cost for requests and data transfer cost associated with both the bucket and the Multi-Region Access Point. Typically, you want to configure your buckets as Requester Pays buckets if you wish to share data but not incur charges associated with others accessing the data. In general, bucket owners pay for all Amazon S3 storage associated with their bucket. To learn more, please visit S3 Requester Pays.

Q: How is S3 Transfer Acceleration different than S3 Multi-Region Access Points?

S3 Multi-Region Access Points and S3 Transfer Acceleration provide similar performance benefits. You can use S3 Transfer Acceleration to speed up content transfers to and from Amazon S3 using the AWS global network. S3 Transfer Accelerator can help accelerate long-distance transfers of larger objects to and from a single Amazon S3 bucket. With S3 Multi-Region Access Points, you can perform similar accelerated transfers using the AWS global network, but across many S3 buckets in multiple AWS Regions for internet-based, VPC-based, and on-premises requests to and from S3. When you combine S3 Multi-Region Access Points with S3 Cross Replication, you provide the capability for S3 Multi-Region Access Points to dynamically route your requests to the lowest latency copy of your data for applications from clients in multiple locations.

Q: How do I get started with S3 Multi-Region Access Points and failover controls?

The S3 console provides a simple guided workflow to quickly set up everything you need to run multi-Region storage on S3 in just three simple steps. First, create an Amazon S3 Multi-Region Access Point endpoint and specify the AWS Regions you want to replicate and failover between. You can add buckets in multiple AWS accounts to a new S3 Multi-Region Access Point by entering the account IDs that own the buckets at the time of creation. Second, for each AWS Region and S3 bucket behind your S3 Multi-Region Access Point endpoint, specify whether their routing status is active or passive, where active AWS Regions accept S3 data request traffic, and passive Regions are not be routed to until you initiate a failover. Third, configure your S3 Cross-Region Replication rules to synchronize your data in S3 between the Regions and/or accounts. You can then initiate a failover at any time between the AWS Regions within minutes to shift your S3 data requests and monitor the shift of your S3 traffic to your new active AWS Region in Amazon CloudWatch. Alternatively, you can use AWS CloudFormation to automate your multi-Region storage configuration. All of the building blocks required to set up multi-Region storage on S3, including S3 Multi-Region Access Points, are supported by CloudFormation, allowing you to automate a repeatable setup process outside of the S3 console.

Data processing

Object Lambda

Q: What is S3 Object Lambda?

S3 Object Lambda allows you to add your own code to S3 GET, LIST, and HEAD requests to modify and process data as it is returned to an application. You can use custom code to modify the data returned by S3 GET requests to filter rows, dynamically resize images, redact confidential data, and much more. You can also use S3 Object Lambda to modify the output of S3 LIST requests to create a custom view of objects in a bucket and S3 HEAD requests to modify object metadata like object name and size. S3 Object Lambda helps you to easily meet the unique data format requirements of any application without having to build and operate additional infrastructure, such as a proxy layer, or having to create and maintain multiple derivative copies of your data. S3 Object Lambda uses AWS Lambda functions to automatically process the output of a standard S3 GET, LIST, or HEAD request. AWS Lambda is a serverless compute service that runs customer-defined code without requiring management of underlying compute resources.

With just a few clicks in the AWS Management Console, you can configure a Lambda function and attach it to an S3 Object Lambda service Access Point. From that point forward, S3 will automatically call your Lambda function to process any data retrieved through the S3 Object Lambda endpoint, returning a transformed result back to the application. You can author and execute your own custom Lambda functions, tailoring S3 Object Lambda's data transformation to your specific use case.

To get started with S3 Object Lambda, you can use the S3 Management Console, SDK, or API. Learn more on the S3 Object Lambda page, or the S3 Object Lambda user guide.

Q: Why should I use S3 Object Lambda?

You should use S3 Object Lambda if you want to process data inline with an S3 GET, LIST, or HEAD request. You can use S3 Object Lambda to share a single copy of your data across many applications, avoiding the need to build and operate custom processing infrastructure or to store derivative copies of your data. For example, by using S3 Object Lambda to process S3 GET requests, you can mask sensitive data for compliance purposes, restructure raw data for the purpose of making it compatible with machine learning applications, filter data to restrict access to specific content within an S3 object, or to address a wide range of additional use cases. You can use S3 Object Lambda to enrich your object lists by querying an external index that contains additional object metadata, filter and mask your object lists to only include objects with a specific object tag, or add a file extension to all the object names in your object lists. For example, if you have an S3 bucket with multiple discrete data sets, you can use S3 Object Lambda to filter an S3 LIST response depending on the requester.

S3 Object Lambda can be set up with just a few clicks in the Amazon S3 Management Console. Read the user guide to learn more.

Q: How does S3 Object Lambda work?

S3 Object Lambda uses Lambda functions specified by you to process the output of GET, LIST, and HEAD requests. Once you have defined a Lambda function to process requested data, you can attach that function to an S3 Object Lambda Access Point. GET, LIST, and HEAD requests made through an S3 Object Lambda Access Point will now invoke the specified Lambda function. Lambda will then fetch the S3 object requested by the client and process that object. Once processing has completed, Lambda will stream the processed object back to the calling client. Read the S3 Object Lambda user guide to learn more.

Q: How do I get started with S3 Object Lambda?

S3 Object Lambda can be set up in multiple ways. You can set up S3 Object Lambda in the S3 console by navigating to the Object Lambda Access Point tab. Next, create an S3 Object Lambda Access Point, the Lambda function that you would like S3 to execute against your GET, LIST, and HEAD requests, and a supporting S3 Access Point. Grant permissions to all resources to interact with Object Lambda. Lastly, update your SDK and application to use the new S3 Object Lambda Access Point to retrieve data from S3 using the language SDK of your choice. You can use an S3 Object Lambda Access Point alias when making requests. Aliases for S3 Object Lambda Access Points are automatically generated and are interchangeable with S3 bucket names for data accessed through S3 Object Lambda. For existing S3 Object Lambda Access Points, aliases are automatically assigned and ready for use. There are example Lambda function implementations in the AWS documentation to help you get started.

You can also use AWS CloudFormation to automate your S3 Object Lambda configuration. When you use the AWS CloudFormation template, the Lambda function that is deployed in your account will pass S3 objects back to your requesting client or application without any changes.

You can add custom code to modify and process data as it is returned to an application. To learn more, visit the S3 Object Lambda User Guide.

Q: What kinds of operations can I perform with S3 Object Lambda?

Any operation supported in a Lambda function is supported with S3 Object Lambda. This gives you a wide range of available options for processing your requests. You supply your own Lambda function to run custom computations against GET, LIST, and HEAD requests, giving you the flexibility to process data according to the needs of your application. Lambda processing time is limited to a maximum of 60 seconds. For more details, see the S3 Object Lambda documentation.

Q: Which S3 request types does S3 Object Lambda support?

S3 Object Lambda supports GET, LIST and HEAD requests. Any other S3 API calls made to an S3 Object Lambda Access Point will return the standard S3 API response. Learn more about S3 Object Lambda in the user guide.

Q: What will happen when a S3 Object Lambda function fails?

When a S3 Object Lambda function fails, you will receive a request response detailing the failure. Like other invocations of Lambda functions, AWS also automatically monitors functions on your behalf, reporting metrics through Amazon CloudWatch. To help you troubleshoot failures, Lambda logs all requests processed by your function and automatically stores logs generated by your code with Amazon CloudWatch Logs. For more information about accessing CloudWatch logs for AWS Lambda, visit CloudWatch documentation.

Q: Does S3 Object Lambda affect the S3 availability SLA or S3 durability?

S3 Object Lambda connects Amazon S3, AWS Lambda, and optionally, other AWS services of your choosing to deliver objects relevant to requesting applications. All AWS services used in connection with S3 Object Lambda will continue to be governed by their respective Service Level Agreements (SLA). For example, in the event that any AWS Service does not meet its Service Commitment, you will be eligible to receive a Service Credit as documented in that service’s SLA. Creating an S3 Object Lambda Access Point does not impact the durability of your objects. However, S3 Object Lambda invokes your specified AWS Lambda function and you must ensure your specified Lambda function is intended and correct. See the latest Amazon S3 SLA here.

Q: How much does S3 Object Lambda cost?

When you use S3 Object Lambda, you pay a per GB charge for every gigabyte of data returned to you through S3 Object Lambda. You are also charged for requests based on the request type (GET, LIST, and HEAD requests) and AWS Lambda compute charges for the time your specified function is running to process the requested data. To see pricing details and an example, read the S3 pricing page.

Data Access

Mountpoint for Amazon S3

Q: What is Mountpoint for Amazon S3?

Mountpoint for Amazon S3 is an open source file client that you can use to mount an S3 bucket on your compute instance and access it as a local file system. Mountpoint for Amazon S3 translates local file system operations to REST API calls on objects stored in Amazon S3. With Mountpoint for Amazon S3, you can achieve high single-instance throughput to finish jobs faster. Mountpoint for Amazon S3 is backed by AWS Support. Customers with access to AWS Enterprise Support get 24x7 technical support from Amazon support engineers and architectural guidance delivered in the context of their use cases. Mountpoint for Amazon S3 works with the Linux operating system and AWS compute services such as Amazon Elastic Compute Cloud (EC2). Learn more on the Mountpoint for Amazon S3 page or the user guide.

Q: When should I use Mountpoint for Amazon S3?

Mountpoint for Amazon S3 is ideal for read-heavy data lake workloads that process petabytes of data using random and sequential read operations on existing files and sequential write operations for creating new files. These workloads write from a single node and do not modify existing data in Amazon S3. Common use cases include petabyte-scale autonomous vehicle simulation, machine learning training, genomics analysis, and image rendering. These workloads scale up and down quickly, and rely on Amazon S3's elasticity to minimize underutilized capacity and avoid the cost of over-provisioning throughput. You can save on compute costs with Mountpoint for Amazon S3 by efficiently utilizing the network bandwidth use of your compute instances, and reliably scale to thousands of compute instances for petabyte-scale data lake workloads.

Q: What file system operations does Mountpoint for Amazon S3 support?

Mountpoint for Amazon S3 supports basic file system operations such as reading files up to 5TB in size, writing new files, listing existing files, and creating and listing directories. Mountpoint for Amazon S3 does not support modifying existing files or deleting existing directories. With these operations, Mountpoint for Amazon S3 is ideal for applications that read and write data at high throughput in Amazon S3 data lakes. It is not suitable for applications that need collaboration and coordination across multiple compute instances or users. These applications typically need shared file system features like appending to existing files and file locking. You can use Amazon FSx for Lustre for data lake applications that need POSIX semantics and shared file system features.

Q: How do I get started with Mountpoint for Amazon S3?

You can get started with Mountpoint for Amazon S3 by mounting an S3 bucket at a local directory on your compute instance using the instructions provided in the documentation. Once you mount the S3 bucket at a local directory, your applications can access S3 objects as files available locally on their compute instance. Mountpoint for Amazon S3 supports sequential and random read operations on existing Amazon S3 objects, and supports sequential writes for new objects. You should read the semantics documentation for Mountpoint for Amazon S3 for more details on supported file system operations. You can use Mountpoint for Amazon S3 to access objects in all S3 storage classes, excluding objects in S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, and objects in the Archive Access tier and Deep Archive Access tier in S3 Intelligent-Tiering.

Q: How am I charged for Mountpoint for Amazon S3?

There is no additional charge for using Mountpoint for Amazon S3. You pay for S3 API requests such as GET, PUT, and LIST requests made by Mountpoint for Amazon S3 when you run file system operations such as file-read, file-write, and directory-listing operations. For S3 request pricing, please visit the pricing page.

Q: What performance can I expect from Mountpoint for Amazon S3?

Mountpoint for Amazon S3 delivers the same performance as the AWS SDKs. This means data lake applications achieve high single-instance transfer rates, efficiently utilizing the available network bandwidth on their Amazon EC2 instance. To achieve even higher throughput, these applications can aggregate throughput across multiple instances to get multiple Tb/s.

Q: How can I control access to my data when using Mountpoint for Amazon S3?

When using Mountpoint for Amazon S3, you can control access to your data using Amazon S3's existing access control mechanisms, including bucket policies and AWS Identity and Access Management (IAM) policies. Mountpoint for Amazon S3 translates file system operations like read and write into object API requests made to your S3 bucket. Afterwards, Amazon S3 evaluates all the relevant policies, such as those on the user and bucket, to decide whether to authorize the request. Mountpoint for Amazon S3 does not introduce new access control mechanisms.

Q: Does Mountpoint for Amazon S3 support POSIX-style metadata, such as user ID, group ID, and permission fields?

Mountpoint for Amazon S3 does not support reading or writing POSIX-style metadata, such as user ID, group ID, and permission fields. You can use Amazon FSx for Lustre with Amazon S3 or AWS DataSync to store POSIX-style metadata for S3 objects.

Q: Does Mountpoint for Amazon S3 support access over AWS PrivateLink?

Yes, Mountpoint for Amazon S3 supports access over AWS PrivateLink. AWS PrivateLink for S3 provides private connectivity between Amazon S3 and on-premises. You can provision interface VPC endpoints for S3 in your VPC to connect your on-premises applications directly to S3 over AWS Direct Connect or AWS VPN.

Q: Does Mountpoint for Amazon S3 support access over gateway VPC endpoints?

Yes, Mountpoint for Amazon S3 supports access over gateway VPC endpoints. We recommend that you use AWS PrivateLink-based interface VPC endpoints to access S3 from on-premises or from a VPC in another AWS Region. For resources that access S3 from a VPC in the same AWS Region as your S3 bucket, we recommend using gateway VPC endpoints as they are not billed.