Amazon Simple Storage Service (S3) FAQs

General

Q: What is Amazon S3?

Amazon S3 is storage for the Internet. It's a simple storage service that offers software developers a highly-scalable, reliable, and low-latency data storage infrastructure 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 on the web. Using this web service, developers can easily build applications that make use of Internet storage. Since Amazon S3 is highly scalable and you only pay for what you use, developers can start small and grow their application as they 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 developers to focus on innovation, not figuring out how to store their data.

Q: How can I get started using Amazon S3?

To sign up for Amazon S3, click the "Sign up for This Web Service" button on the Amazon S3 detail page. You must have an Amazon Web Services account to access this service; if you do not already have one, you will be

prompted to create one when you begin the Amazon S3 sign-up process. After signing up, please refer to the Amazon S3 documentation and sample code in the Resource Center to begin using Amazon S3.

Q: What are the technical benefits of Amazon S3?

Amazon S3 was carefully engineered to meet the requirements for scalability, reliability, speed, low-cost, and simplicity that must be met for Amazon's internal developers. Amazon S3 passes these same benefits onto any external developer. More information about the Amazon S3 design requirements is available on the Amazon S3 detail page.

Q: What can developers do now that they could not before?

Until now, a sophisticated and scalable data storage infrastructure like Amazon's has been beyond the reach of small developers. Amazon S3 enables any developer to leverage Amazon's own benefits of massive scale with no up-front investment or performance compromises. Developers are now free to innovate knowing that no matter how successful their businesses become, it will be inexpensive and simple to ensure their data is quickly accessible, always available, and secure.

Q: What kind of data can I store?

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

Q: How much data can I store?

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

Q: What storage classes does Amazon S3 offer?

Amazon S3 offers a range of storage classes designed for different use cases. There are three highly durable storage classes including Amazon S3 Standard for general-purpose storage of frequently accessed data, Amazon S3 Standard - Infrequent Access for long-lived, but less frequently accessed data, and Amazon Glacier for long-term archive. You can learn more about those three storage classes on the Amazon S3 Storage Classes page.

Reduced Redundancy Storage (RRS) is an Amazon S3 storage option that enables customers to reduce their costs by storing noncritical, reproducible data at lower levels of redundancy than Amazon S3's standard storage. You can learn more about Reduced Redundancy Storage on the Reduced Redundancy detail page.

Q: How can I delete large numbers of objects?

You can use Multi-Object Delete to delete large numbers of objects from Amazon S3. This feature allows you to send multiple object keys in a single request to speed up your deletes. Amazon does not charge you for using Multi-Object Delete.

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

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

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

Yes. Developers within 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 can be constructed to mimic hierarchical attributes.

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 any developer access to the same highly scalable, reliable, fast, inexpensive data storage infrastructure that Amazon uses to run its own global network of web sites. S3 Standard is designed for 99.99% availability and Standard - IA is designed for 99.9% availability. Both are backed by the Amazon S3 Service Level Agreement.

Q: What data consistency model does Amazon S3 employ?

Amazon S3 buckets in all Regions provide read-after-write consistency for PUTS of new objects and eventual consistency for overwrite PUTS and DELETES.

Learn more

Q: What happens if traffic from my application suddenly spikes?

Amazon S3 was 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 enables us to spread load evenly, so that no individual application is affected by traffic spikes.

Q: What is the BitTorrent™ protocol, and how do I use it with Amazon S3?

BitTorrent is an open source Internet distribution protocol. Amazon S3's bandwidth rates are inexpensive, but BitTorrent allows developers to further save on bandwidth costs for a popular piece of data by letting users

download from Amazon and other users simultaneously. Any publicly available data in Amazon S3 can be downloaded via the BitTorrent protocol, in addition to the default client/server delivery mechanism. Simply add the ?torrent parameter at the end of your GET request in the REST API.

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: How can I Increase the number of Amazon S3 buckets that I can provision?

By default, customers can provision up to 100 buckets per AWS account. However, you can increase your Amazon S3 bucket limit by visiting AWS Service Limits.

Regions

Q: Where is my data stored?

You specify a region when you create your Amazon S3 bucket. Within that region, your objects are redundantly stored on multiple devices across multiple facilities. Please refer to Regional Products and Services for details of Amazon S3 service availability by region.

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

There are several factors to consider based on your specific application. You may want to store your data in a region that…

- …is near to your customers, your data centers, or your other AWS resources in order to reduce data access latencies.
- …is remote from your other operations for geographic redundancy and disaster recovery purposes.
- …enables you to address specific legal and regulatory requirements.
- …allows you to reduce storage costs. You can choose a lower priced region to save money. For S3 pricing information, please visit the S3 pricing page.

Q: I'm not in the US or Europe; can I use Amazon S3?

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.

Q. Wasn't there a US Standard region?

We renamed the US Standard Region to US East (Northern Virginia) Region to be consistent with AWS regional naming conventions. There is no change to the endpoint and you do not need to make any changes to your

application.

Billing

Q: How much does Amazon S3 cost?

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

We charge less where our costs are less. Some prices vary across Amazon S3 Regions and are based on the location of your 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 Regions is charged at rates specified on the pricing section of the Amazon S3 detail page. There is no Data Transfer charge for data transferred between Amazon EC2 and Amazon S3 within the same Region or for data transferred between the Amazon EC2 Northern Virginia Region and the Amazon S3 US East (Northern Virginia) Region. Data transferred between Amazon EC2 and Amazon S3 across all other Regions (i.e. between the Amazon EC2 Northern California and Amazon S3 US East (Northern Virginia) Regions is charged at rates specified on the pricing section of the Amazon S3 detail page.

For Amazon S3 pricing information, please visit the pricing page.

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

We charge 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, please 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 fee is calculated based on the current rates for your region on the Amazon S3 Pricing Page.

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

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

With the AWS Free Usage Tier*, you can get started with Amazon S3 for free in all regions except the AWS GovCloud Region. Upon sign-up, new AWS customers receive 5 GB of Amazon S3 standard storage, 20,000 Get Requests, 2,000 Put Requests, 15GB of data transfer in, and 15GB of data transfer out each month for one year.

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 Chart.

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 100GB (107,374,182,400 bytes) of standard Amazon S3 storage data in your bucket for 15 days in March, and 100TB (109,951,162,777,600 bytes) of standard Amazon S3 storage data 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.

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 Fee = $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 Chart.

This represents the amount of data sent to your Amazon S3 buckets. Data Transfer is $0.000 per GB for buckets in the US East (Northern Virginia), US West (Oregon), US West (Northern California), EU (Ireland), EU (Frankfurt),

Asia Pacific (Singapore), Asia Pacific (Tokyo), Asia Pacific (Sydney), South America (Sao Paulo), and AWS GovCloud (US) Regions.

Network Data Transferred Out:

Amazon S3 Data Transfer Out pricing is summarized on the Amazon S3 Pricing Chart. 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 1TB 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 1TB 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 fee 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 Fee = $921.51+ $3,481.60 + $860.16= $5,263.27

Data Requests:

Amazon S3 Request pricing is summarized on the Amazon S3 Pricing Chart.

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 fees 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 Standard – Infrequent Access (Standard - IA) storage class and is summarized on the Amazon S3 Pricing Chart.

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

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

Please see here for details on billing of objects archived to Amazon Glacier.

* * 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: 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 for accessing Amazon S3 from another AWS Account?

Normal Amazon S3 pricing applies when your storage is accessed by another AWS Account. 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.

Security

Q: How secure is my data?

Amazon S3 is secure by default. Only the bucket and object owners originally have access to Amazon S3 resources they create. Amazon S3 supports user authentication to control access to data. You can use access control mechanisms such as bucket policies and Access Control Lists (ACLs) to selectively grant permissions to users and groups of users. 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 can securely upload/download your data to Amazon S3 via SSL endpoints using the HTTPS protocol. If you need extra security you can use the Server Side Encryption (SSE) option to encrypt data stored at rest. You can

configure your Amazon S3 buckets to automatically encrypt objects before storing them if the incoming storage requests do not have any encryption information. Alternatively you can use your own encryption libraries to encrypt data before storing it in Amazon S3.

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

Customers may use four mechanisms for controlling access to Amazon S3 resources: Identity and Access Management (IAM) policies, bucket policies, Access Control Lists (ACLs) and query string authentication. IAM enables organizations with multiple employees to create and manage multiple users under a single AWS account. With IAM policies, companies can grant IAM users fine-grained control to their Amazon S3 bucket or objects while also retaining full control over everything the users do. With bucket policies, companies 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. With ACLs, customers can grant specific permissions (i.e. READ, WRITE, FULL_CONTROL) to specific users for an individual bucket or object. 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, please refer to the Access Control topic in the Amazon S3 Developer Guide.

Q: Does Amazon S3 support data access auditing?

Yes, customers can optionally configure Amazon S3 buckets to create access log records for all requests made against it. 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?

You can choose to encrypt data using SSE-S3, SSE-C, SSE-KMS, or a client library such as the Amazon S3 Encryption Client. All four enable you to store sensitive data encrypted at rest in Amazon S3.

SSE-S3 provides an integrated solution where Amazon handles key management and key protection using multiple layers of security. You should choose SSE-S3 if you prefer to have Amazon manage your keys.

SSE-C enables you to leverage Amazon S3 to perform the encryption and decryption of your objects while retaining control of the keys used to encrypt objects. 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 enables you to use AWS Key Management Service (AWS KMS) to 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 master key, providing an additional layer of control as well as 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, such as the Amazon S3 Encryption Client, 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, please refer to the topic on Using Encryption in the Amazon S3 Developer Guide.

Q: How does Amazon protect SSE encryption keys?

With SSE, every protected object is encrypted with a unique key. This object key is itself encrypted by a separate master key. A new master key is issued at least monthly. Encrypted data, encryption keys and master keys are stored and secured on separate hosts for multiple layers of protection.

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

Customers can choose to store all data in the EU by using the EU (Ireland) or EU (Frankfurt) region. It is your responsibility to ensure that you comply with EU privacy laws.

Q: Where can I find more information about security on AWS?

For more information on security on AWS please refer to our Amazon Web Services: Overview of Security Processes document.

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 only to S3. The VPC Endpoint routes requests to S3 and routes responses back to the VPC. For more information about VPC Endpoints, read Using VPC Endpoints.

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 Using VPC Endpoints.

Q: What is Amazon Macie?

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.

Q: What can I do with Amazon Macie?

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 or 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.

Q: How does Amazon Macie secure your data?

As part of the data classification process, Amazon Macie identifies customers' objects in their S3 buckets, and streams the object contents into memory for analysis. When deeper analysis is required for complex file formats, Amazon Macie will download a full copy of the object, only keeping it for the short time it takes to fully analyze the object. Immediately after Amazon Macie has analyzed the file content for data classification, it deletes the stored content and only retains the metadata required for future analysis. At any time, customers can revoke Amazon Macie access to data in the Amazon S3 bucket. For more information, go to the Amazon Macie User Guide.

Data Protection

Q: How durable is Amazon S3?

Amazon S3 Standard and Standard - IA are designed to provide 99.999999999% durability of objects over a given year. This durability level corresponds to an average annual expected loss of 0.000000001% of objects. For example, if you store 10,000 objects with Amazon S3, you can on average expect to incur a loss of a single object once every 10,000,000 years. In addition, Amazon S3 is designed to sustain the concurrent loss of data in two facilities.

As with any environments, the best practice is to have a backup and to put in place safeguards against malicious or accidental users errors. For S3 data, that best practice includes secure access permissions, Cross-Region Replication, versioning and a functioning, regularly tested backup.

Q: How is Amazon S3 designed to achieve 99.999999999% durability?

Amazon S3 Standard and Standard - IA redundantly stores your objects on multiple devices across multiple facilities in an Amazon S3 Region. The service is designed to sustain concurrent device failures by quickly detecting and repairing any lost redundancy. When processing a request to store data, the service will

redundantly store your object across multiple facilities before returning SUCCESS. Amazon S3 also regularly verifies the integrity of your data using checksums.

Q: What checksums does Amazon S3 employ to detect data corruption?

Amazon S3 uses a combination of Content-MD5 checksums and cyclic redundancy checks (CRCs) to detect data corruption. Amazon S3 performs these checksums on data at rest and repairs any corruption using redundant data. In addition, the service calculates checksums on all network traffic to detect corruption of data packets when storing or retrieving data.

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, please refer to the Amazon S3 Technical 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 setup a trash, recycle bin, or rollback window on my Amazon S3 objects to recover from deletes and overwrites?

You can use Lifecycle rules along with Versioning to implement a rollback window for your Amazon 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 Glacier 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.

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

Versioning's MFA Delete capability, which uses multi-factor authentication, 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, please refer to the Amazon S3 Technical 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, please 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 fee is calculated based on the current rates for your region on the Amazon S3 Pricing Page.

S3 Standard - Infrequent Access

Q: What is S3 Standard - Infrequent Access?

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

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

Standard - IA is ideal for data that is accessed less frequently, but requires rapid access when needed. Standard - IA is ideally suited for long-term file storage, older data from sync and share, backup data, and disaster recovery files.

Q: What performance does S3 Standard - Infrequent Access offer?

S3 Standard - Infrequent Access provide the same performance as S3 Standard storage.

Q: How durable is Standard - IA?

S3 Standard - IA is designed for the same 99.999999999% durability as Standard and Amazon Glacier. Standard - IA 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 available is Standard - IA?

Designed for 99.9% availability, Standard - IA has a thinner front end that provides nine one-hundredths of a percent less availability than S3 Standard. Standard - IA 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 Standard - IA?

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

Q: Are my Standard - IA objects backed with the Amazon S3 Service Level Agreement?

Yes, Standard - IA is backed with the Amazon S3 Service Level Agreement, and customers are eligible for service credits if availability is less than our service commitment in any billing cycle.

Q: How will my latency and throughput performance be impacted as a result of using Standard - IA?

You should expect the same latency and throughput performance as Amazon S3 Standard when using Standard - IA.

Q: How am I charged for using Standard - IA?

Please see the Amazon S3 pricing page for general information about Standard - IA pricing.

Q. What charges will I incur if I change storage class of an object from Standard-IA to Standard with a copy request?

You will incur charges for an Standard-IA copy request and a Standard-IA data retrieval.

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

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

Q: Is there a minimum object size for Standard - IA?

Standard - IA is designed for larger objects and has a minimum object size 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 fee equivalent to 122KB at the S3 Standard - IA storage price. Please see the Amazon S3 pricing page for information about Standard - IA pricing.

Q: Can I tier objects from Standard - IA to Amazon Glacier?

Yes. In addition to using lifecycle policies to migrate objects from Standard to Standard - IA, you can also set up lifecycle policies to tier objects from Standard - IA to Amazon Glacier.

Query in Place

Q1: What is "Query in Place" functionality?

Amazon S3 allows customers to run sophisticated queries against data stored without the need to extract, transform, and load (ETL) 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.

Q2: What is S3 Select?

S3 Select is an Amazon S3 feature (currently in Preview) 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. You can use S3 Select to retrieve a subset of data using SQL clauses, like SELECT and WHERE, from delimited text files and JSON objects in Amazon S3.

Q3: What can I do with S3 Select?

You can use S3 Select to retrieve a smaller, targeted data set from an object using simple SQL statements. 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.

Q4: Why should I use S3 Select?

S3 Select provides a new way to retrieve specific data using SQL statements from the contents of an object stored in Amazon S3 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.

Q5: How do I get started with S3 Select?

Amazon S3 Select is currently available in Limited Preview. To apply for access to this Preview, complete the Amazon S3 Select Preview Application Form. During the Preview, you can use Amazon S3 Select through the available Presto connector, with AWS Lambda, or from any other application using the S3 Select SDK for Java or Python.

Q6: What is Amazon Athena?

Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL. Athena is serverless, so there is no infrastructure to setup 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 S3. 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 Amazon 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.

Q7: What is Amazon Redshift Spectrum?

Amazon Redshift Spectrum is a feature of Amazon Redshift that enables you to 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, 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 your same BI tools. Redshift Spectrum lets you separate storage and compute, allowing you to scale each independently. You can setup 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.

Amazon Glacier

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

Yes, Amazon S3 enables you to utilize Amazon Glacier's extremely low-cost storage service as storage for data archival. Amazon Glacier stores data for as little as $0.004 per gigabyte per month. To keep costs low yet suitable for varying retrieval needs, Amazon Glacier provides three options for access to archives, from a few minutes to several hours. Some examples of archive uses cases include digital media archives, financial and healthcare records, raw genomic sequence data, long-term database backups, and data that must be retained for regulatory compliance.

Q: How can I store my data using the Amazon Glacier option?

You can use Lifecycle rules to automatically archive sets of Amazon S3 objects to Amazon Glacier based on lifetime. Use the Amazon S3 Management Console, the AWS SDKs or the Amazon S3 APIs to define rules for archival. Rules specify a prefix and time period. The prefix (e.g. "logs/") identifies the object(s) subject to the rule. The time period specifies either the number of days from object creation date (e.g. 180 days) or the specified date after which the object(s) should be archived. Any Amazon S3 Standard or Amazon S3 Standard - IA objects which have names beginning with the specified prefix and which have aged past the specified time period are archived to Amazon Glacier. To retrieve Amazon S3 data stored in Amazon Glacier, initiate a retrieval job via the Amazon S3 APIs or Management Console. Once the job is complete, you can access your data through an Amazon S3 GET object request.

For more information on using Lifecycle rules for archival, please refer to the Object Archival topic in the Amazon S3 Developer Guide.

Q: Can I use the Amazon S3 APIs or Management Console to list objects that I've archived to Amazon Glacier?

Yes, like Amazon S3's other storage options (Standard or Standard - IA), Amazon Glacier objects stored using Amazon S3's APIs or Management Console have an associated user-defined name. You can get a real-time list of all of your Amazon S3 object names, including those stored using the Amazon Glacier option, using the Amazon S3 LIST API.

Q: Can I use Amazon Glacier APIs to access objects that I've archived to Amazon Glacier?

Because Amazon S3 maintains the mapping between your user-defined object name and Amazon Glacier's system-defined identifier, Amazon S3 objects that are stored using the Amazon Glacier option are only accessible through the Amazon S3 APIs or the Amazon S3 Management Console.

Q: How can I retrieve my objects that are archived in Amazon Glacier?

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

Q: How long will it take to retrieve my objects archived in Amazon Glacier?

When processing a retrieval job, Amazon S3 first retrieves the requested data from Amazon Glacier, and then creates a temporary copy of the requested data in RRS (which typically takes on the order of 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. Lastly, Bulk retrievals typically complete within 5 – 12 hours. For more information about the retrieval options, please refer to the Glacier FAQ.

Q: What am I charged for archiving objects in Amazon Glacier?

Amazon Glacier storage is priced from $0.004 per gigabyte per month. Lifecycle transition requests into Amazon Glacier cost $0.05 per 1,000 requests. Objects that are archived to Glacier have a minimum of 90 days of storage, and objects deleted before 90 days incur a pro-rated charge equal to the storage charge for the remaining.

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

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 kilobytes of Glacier data plus an additional 8 KB of S3 standard storage data. Amazon Glacier requires an additional 32 KB of data per object for Glacier's index and metadata so you can identify and retrieve your data. Amazon S3 requires 8KB to store and maintain the user-defined name and metadata for objects archived to Amazon Glacier. This enables you to get a real-time list of all of your Amazon S3 objects, including those stored using the Amazon Glacier option, using the Amazon S3 LIST API. For example, if you have archived 100,000 objects that are 1GB each, your billable storage would be:

1.000032 gigabytes for each object x 100,000 objects = 100,003.2 gigabytes of Amazon Glacier storage.
0.000008 gigabytes for each object x 100,000 objects = 0.8 gigabytes of Amazon S3 Standard storage.

The fee is calculated based on the current rates for your region on the Amazon S3 Pricing Page.

Q: How much data can I retrieve for free?

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

Q: How am I charged for deleting objects from Amazon Glacier that are less than 3 months old?

Amazon Glacier is designed for use cases where data is retained for months, years, or decades. Deleting data that is archived to Amazon Glacier is free if the objects being deleted have been archived in Amazon Glacier for three months or longer. If an object archived in Amazon Glacier is deleted or overwritten within three months of being archived then there will be an early deletion fee. This fee is prorated. If you delete 1GB of data 1 month after uploading it, you will be charged an early deletion fee for 2 months of Amazon Glacier storage. If you delete 1 GB after 2 months, you will be charged for 1 month of Amazon Glacier storage.

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

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

For example, using Expedited retrievals in the US East (Northern Virginia) region, if you requested 10 archives with a size of 1 GB each, the cost would be 10 x $0.03 +10 x $0.01 = $0.40.

If you were using Standard retrievals in the US East (Northern Virginia) region to retrieve 500 archives that were 1 GB each, the cost would be 500GB x $0.01 + 500 x $0.05/1,000 = $5.25

Lastly, using Bulk retrievals in the US East (Northern Virginia) region, if you were to retrieve 500 archives that are 1 GB each, the cost would be 500GB x $0.0025 + 500 x $0.025/1,000 = $1.2625.

When an archived object is retrieved, it resides in both RRS and Glacier.

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

Event Notification

Q1: What are Amazon S3 event notifications?

Amazon S3 event notifications can be sent in response to actions in Amazon S3 like PUTs, POSTs, COPYs, or DELETEs. Notification messages can be sent through either Amazon SNS, Amazon SQS, or directly to AWS Lambda.

Q2: What can I do with Amazon S3 event notifications?

Amazon S3 event notifications enable you to run workflows, send alerts, or perform other actions in response to changes in your objects stored in Amazon S3. You can use Amazon 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 Amazon 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/."

Q3: What is included in an Amazon S3 event notification?

For a detailed description of the information included in Amazon S3 event notification messages, please refer to the Configuring Amazon S3 event notifications topic in the Amazon S3 Developer Guide.

Q4: How do I set up Amazon S3 event notifications?

For a detailed description of how to configure event notifications, please refer to the Configuring Amazon S3 event notifications topic in the Amazon S3 Developer Guide. You can learn more about the AWS messaging services in the Amazon SNS Documentation and the Amazon SQS Documentation.

Q5: What does it cost to use Amazon S3 event notifications?

There are no additional charges from 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 the AWS Lambda function. Visit the Amazon SNS, Amazon SQS, or AWS Lambda pricing pages to view the pricing details for these services.

Static Website Hosting

Q: Can I host my static website on Amazon S3?

Yes, you can host your entire static website on Amazon S3 for an inexpensive, highly available hosting solution that scales automatically to meet traffic demands. Amazon S3 gives you access to the same highly scalable, reliable, fast, inexpensive infrastructure that Amazon uses to run its own global network of web sites. Service availability corresponds to storage class and the service level agreement provides service credits if a customer's availability falls below our service commitment in any billing cycle. To learn more about hosting your website on Amazon S3, please see our walkthrough on setting up an Amazon S3 hosted website.

Q: What kinds of websites should I host using Amazon S3 static website hosting?

Amazon S3 is ideal for hosting websites that contain only static content, including html files, images, videos, and client-side scripts such as JavaScript. Amazon EC2 is recommended for websites with server-side scripting and database interaction.

Q: Can I use my own host name with my Amazon S3 hosted website?

Yes, you can easily and durably store your content in an Amazon S3 bucket and map your domain name (e.g. "example.com") to this bucket. Visitors to your website can then access this content by typing in your website's URL (e.g., "http://example.com") in their browser.

Q: Does Amazon S3 support website redirects?

Yes, Amazon S3 provides multiple ways to enable redirection of web content for your static websites. Redirects enable you to change the Uniform Resource Locator (URL) of a web page on your Amazon S3 hosted website (e.g. from www.example.com/oldpage to www.example.com/newpage) without breaking links or bookmarks pointing to the old URL. You can set rules on your bucket to enable automatic redirection. You can also configure a redirect on an individual S3 object.

Q: Is there an additional charge for hosting static websites on Amazon S3?

There is no additional charge for hosting static websites on Amazon S3. The same pricing dimensions of storage, requests, and data transfer apply to your website objects.

Refer to the S3 Pricing page for more information.

Storage Management

S3 Object Tagging

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'll have the ability to create Identity and Access Management (IAM) policies, setup S3 Lifecycle policies, and customize storage metrics. These object-level tags can then manage transitions between storage classes and expire objects in the background.

Q. How do I apply Object Tags to my objects?

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.

Q. Why should I use Object Tags?

Object Tags are a new 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 lifecycle policies to manage transitions to the S3 Standard – Infrequent Access and Glacier storage tiers.

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. Will my Object Tags be replicated if I use Cross-Region Replication?

Object Tags can be replicated across regions using Cross-Region Replication. For more information about setting up Cross-Region Replication, please visit How to Set Up Cross-Region Replication in the Amazon S3 Developer Guide.

For customers with Cross-Region Replication already enabled, new permissions are required in order for tags to replicate. For more information on the policies required, please visit How to Set Up Cross-Region Replication in the Amazon S3 Developer Guide.

Q. How much do Object Tags cost?

Object Tags are priced at $0.01 per 10,000 tags per month. The requests associated with adding and updating Object Tags are priced the same as existing request prices, please see the Amazon S3 pricing page for more

information.

S3 Analytics - Storage Class Analysis

Q. What is S3 Analytics – Storage Class Analysis?

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

Q. How do I get started with S3 Analytics – 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 Standard-IA or archived to Glacier. You can navigate to the “Management” tab in the S3 Console to manage S3 Analytics, S3 Inventory, and S3 CloudWatch metrics.

Q. How am I charged for using S3 Analytics – Storage Class Analysis?

Please see the Amazon S3 pricing page for general information about S3 Analytics – Storage Class Analysis pricing.

Q. How often is the Storage Class Analysis updated?

Storage Class Analysis is updated on a daily basis on the S3 Management Console. Additionally, you can configure S3 Analytics to export your daily storage class analysis to a S3 bucket of your choice.

S3 Inventory

Q. What is S3 Inventory?

S3 Inventory provides a scheduled alternative to Amazon S3’s synchronous List API. You can configure S3 Inventory to provide a CSV or ORC 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 use S3 inventory to verify encryption and replication status of your objects to meet business, compliance, and regulatory needs.

Q. How do I get started with S3 Inventory?

You can use the AWS Management Console or the PUT Bucket Inventory API to configure a daily or weekly inventory 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 inventory, the output file format (CSV or ORC),

and specific object metadata necessary for your business application, such as object name, size, last modified date, storage class, version ID, delete marker, noncurrent version flag, multipart upload flag, replication status, or encryption status.

Q. Can files written by S3 Inventory be encrypted?

Yes, you can configure to encrypt all files written by S3 inventory to be encrypted by SSE-S3 or SSE-KMS. For more information, refer to the user guide.

Q. How do I use S3 Inventory?

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 about querying Inventory with Athena.

Q. How am I charged for using S3 Inventory?

Please 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 KMS pricing page for detail.

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 1-minute CloudWatch metrics for your S3 bucket or configure filters for the metrics using a prefix or object tag. Alternately, you can call the S3 PUT Bucket Metrics API to enable and configure publication of S3 storage metrics. Storage metrics will be available in CloudWatch within 15 minutes of being enabled.

Q. Can I align storage metrics to my applications or business organizations?

Yes, you can configure S3 CloudWatch metrics to generate metrics for your S3 bucket or configure filters for the metrics using a prefix or object tag. For example, you can monitor a spark application that accesses data under the prefix “/Bucket01/BigData/SparkCluster” as metrics filter 1 and define a second metrics filter with the tag “Dept, 1234” as metrics filter 2. An object can be a member of multiple filters, e.g., an object within the prefix “/Bucket01/BigData/SparkCluster” and with the tag “Dept,1234” will be in both metrics filter 1 and 2. In this way, metrics filters can be aligned to business applications, team structures or organizational budgets, allowing you to monitor and alert on multiple workloads separately within the same S3 bucket.

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 fire 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 3 data points are above the threshold fire a CloudWatch alarm to alert a Dev Ops engineer.

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

S3 CloudWatch Metrics are priced as custom metrics for Amazon CloudWatch. Please see Amazon CloudWatch pricing page for general information about S3 CloudWatch metrics pricing.

Lifecycle Management Policies

Q. What is 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 lifecycle transition policy to automatically migrate Amazon S3 objects to Standard - Infrequent Access (Standard - IA) and/or Amazon Glacier 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 upload based on the age of the upload.

Q. How do I set up a 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 much does it cost to use 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 S3 Pricing page for pricing information.

Q. What can I do with Lifecycle Management Policies?

As data matures, it can become less critical, less valuable and/or subject to compliance requirements. Amazon S3 includes an extensive library of policies that help you automate data migration processes. For example, you can set infrequently accessed objects to move into lower cost storage tier (like Standard-Infrequent Access) after a period of time. After another period, it can be moved into Amazon Glacier for archive and compliance, and eventually deleted. These rules can invisibly lower storage costs and simplify management efforts, and may be leveraged across the Amazon family of storage services. These policies also include good stewardship practices to remove objects and attributes that are no longer needed to manage cost and optimize performance.

Q: How can I use Amazon S3's lifecycle policy to help lower my Amazon S3 storage costs?

With Amazon S3's lifecycle policies, you can configure your objects to be migrated to Standard - Infrequent Access (Standard - IA), archived to Amazon Glacier, or deleted 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 Standard - IA or Amazon Glacier, and/or an expiration. For example, you could create a rule that archives into Amazon Glacier 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. 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.

Learn more.

Q: How can I configure my objects to be deleted after a specific time period?

You can set a lifecycle expiration policy to remove objects from your buckets after a specified number of days. You can define the expiration rules for a set of objects in your bucket through the Lifecycle Configuration policy that you apply to the bucket. Each Object Expiration rule allows you to specify a prefix and an expiration period. 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/”). For expiration period, provide the number of days from creation date (i.e. age) after which you want your objects removed. You may create multiple rules for different prefixes. For example, you could create a rule that removes all objects with the prefix “logs/” 30 days from creation, and a separate rule that removes all objects with the prefix “backups/” 90 days from creation.

After an Object Expiration rule is added, the rule is applied to objects that already exist in the bucket as well as new objects added to the bucket. Once objects are past their expiration date, they are identified and queued for removal. You will not be billed for storage for objects on or after their expiration date, though you may still be able to access those objects while they are in queue before they are removed. As with standard delete requests, Amazon S3 doesn’t charge you for removing objects using Object Expiration. You can set Expiration rules for your versioning-enabled or versioning-suspended buckets as well.

Learn more.

Q. Why would I use a lifecycle policy to expire incomplete multipart uploads?

The 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.

Cross-Region Replication

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

CRR is an Amazon S3 feature that automatically replicates data across AWS regions. With CRR, every object uploaded to an S3 bucket is automatically replicated to a destination bucket in a different AWS region that you choose. 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.

Q: How do I enable CRR?

CRR is a bucket-level configuration. You enable a CRR configuration on your source bucket by specifying a destination bucket in a different region for replication. You can use either the AWS Management Console, the REST API, the AWS CLI, or the AWS SDKs to enable CRR. Versioning must be turned on for both the source and destination buckets to enable CRR. To learn more, please visit How to Set Up Cross-Region Replication in the Amazon S3 Developer Guide.

Q: What does CRR replicate to the target bucket?

CRR replicates every object-level upload that you directly make to your source bucket. The metadata and ACLs associated with the object are also part of the replication. Any change to the underlying data, metadata, or ACLs on the object would trigger a new replication to the destination bucket. You can either choose to replicate all objects uploaded to a source bucket or just a subset of objects uploaded by specifying prefixes. Existing data in the bucket prior to enabling CRR is not replicated. You can use S3's COPY API to copy the existing data into your destination bucket. To learn more about CRR please visit How to Set Up Cross-Region Replication in the Amazon S3 Developer Guide.

Q: Can I use CRR with lifecycle rules?

Yes, you can configure separate lifecycle rules on the source and destination buckets. For example, you can configure a lifecycle rule to migrate data from Standard to Standard - IA on the destination bucket or configure a lifecycle rule to archive data into Amazon Glacier.

Q: Can I use CRR with objects encrypted by AWS KMS?

Yes, you can replicate KMS-encrypted objects by providing destination KMS key in your replication configuration. Learn more.

Q. Does enabling AWS KMS support for Cross-Region Replication affect KMS API rate?

Yes, AWS KMS support for CRR will increase KMS API rate for your account. Specifically, CRR will double the S3-related KMS API rate in the source region and increase by the same increment in the destination region. We recommend requesting an increase in your KMS API rate limit by creating a case in the AWS support center. There is no additional cost for KMS API rate limit increase.

Q: Are objects securely transferred and encrypted throughout replication process?

Yes, objects remain encrypted throughout the CRR process. The encrypted objects are transmitted securely via SSL from the source region to the destination region.

Q: Can I use CRR across accounts?

Yes, you can set up CRR across account 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: What is the pricing for CRR?

You pay the Amazon S3 charges for storage, copy requests, and inter-region data transfer for the replicated copy of data. Copy requests and inter-region data transfer are charged based on the source region. Storage for replicated data is charged based on the target region. For more information, please visit the 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 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.

Amazon S3 Transfer Acceleration

Q. What is Transfer Acceleration?

Amazon S3 Transfer Acceleration enables fast, easy, and secure transfers of files over long distances between your client and your Amazon S3 bucket. 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 Transfer Acceleration?

It's easy to get started with Transfer Acceleration. First, enable Transfer Acceleration on an S3 bucket using the Amazon S3 console, the Amazon S3 API, or the AWS CLI. After 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: <bucketname>.s3-accelerate.amazonaws.com or <bucketname>.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 bucket will work with transfer acceleration. For details, please refer the Amazon S3 developer guide here.

Q. How fast is Transfer Acceleration?

Transfer Acceleration helps you fully utilize 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. 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 50MB) to an S3 bucket in the US East (N. Virginia) region.

Try the speed comparison tool to get a preview of the performance benefit from your location!

Q. Who should use Transfer Acceleration?

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.

Q. How secure is Transfer Acceleration?

Transfer Acceleration provides the same security as regular transfers to Amazon S3. All Amazon S3 security features, such as restricting access based on a client's IP address, are supported as well. 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 Transfer Acceleration isn't faster?

Each time you use Transfer Acceleration to upload an object, we will check whether Transfer Acceleration is likely to be faster than a regular Amazon S3 transfer. If we determine that 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 that use of Transfer Acceleration for that transfer, and may bypass the Transfer Acceleration system for that upload.

Q. Can I use Transfer Acceleration with multipart uploads?

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

Q. How should I choose between Transfer Acceleration and Amazon CloudFront's PUT/POST?

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

Q. How should I choose between Transfer Acceleration and AWS Snowball?

The AWS Import/Export Snowball 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, Transfer Acceleration over a fully-utilized 1 Gbps line can transfer up to 75 TBs in the same time. 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, 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 Transfer Acceleration.

Q. Can Transfer Acceleration complement AWS Direct Connect?

AWS Direct Connect is a good choice for customers with a private networking requirement or have access to AWS Direct Connect exchanges. 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 Transfer Acceleration to help with remote office transfers, where they may suffer from poor Internet performance.

Q. Can Transfer Acceleration complement the AWS Storage Gateway or a 3rd party gateway?

If you can configure the bucket destination in your 3rd party gateway to use an S3 Transfer Acceleration endpoint domain name you will see the benefit.

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

Q. Can Transfer Acceleration complement 3rd party integrated software?

Yes. Software packages that connect directly into Amazon S3 (read more about storage partner solutions here) can take advantage of Transfer Acceleration when they send their jobs to Amazon S3.

Q: Is Transfer Acceleration HIPAA eligible?

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

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 the new 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 needing 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 new "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. Your applications may continue to access data through the existing APIs and virtual hosted style (e.g. http://bucket.s3.dualstack.aws-region.amazonaws.com) or path style (e.g. http://s3.dualstack.aws-region.amazonaws.com/bucket) URLs without code changes. When using Amazon S3 Transfer Acceleration, the "dual-stack" endpoint must be of the form http(s)://bucket.s3-accelerate.dualstack.amazonaws.com. However, you must also evaluate your bucket and Identity and Access Management (IAM) policies to ensure you have the appropriate access configured for your new IPv6 addresses. For more information about getting started accessing Amazon S3 over IPv6, see Making Requests to Amazon S3 over IPv6.

Q. If I point to Amazon S3's "dual-stack" endpoint, will I still be able to access Amazon S3's APIs over IPv4?

Yes, you can continue to access Amazon S3 APIs using both IPv6 and IPv4 addresses when connecting to the Amazon S3 “dual-stack” endpoints. You will need to configure your client to prefer IPv4 addresses, which can be an application-level or host-level configuration option for many application runtime languages. Please consult the documentation for the language you are using for your runtime platform for the specific configuration option that prefers IPv4 connections.

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.

Q. Will existing VPC Endpoints continue to work if I point to Amazon S3's "dual-stack" endpoint?

Yes, you can continue using VPC Endpoint to access Amazon S3 over IPv4. If you use the dual-stack endpoint in an IPv4-only VPC, the VPC instances will drop the AAAA record and always access Amazon S3 over IPv4.

Q. If I enable IPv6, will the IPv6 address appear in the Server Access Log?

Yes, IPv6 addresses will now be shown in the Server Access logs if you have the Amazon S3 Server Access logs feature enabled. Any customer tool or software that parses the logs should be updated to handle the new IPv6 address format. Please contact Developer Support if you have any issues with IPv6 traffic impacting your tool or software’s ability to handle IPv6 addresses in Server Access logs.

Q. Do I need to update my bucket and IAM policies?

Yes, if you use policies to grant or restrict access via IP addresses, you will need to update those policies to include the associated IPv6 ranges before you switch to the “dual-stack” endpoint. If your bucket grants or restricts access to specific IAM users, you will also need to have the IAM policy administrator review those users’ IAM policies to ensure they have appropriate access to the associated IPv6 ranges before you switch to the “dual-stack” endpoint. Failure to do so may result in clients incorrectly losing or gaining access to the bucket when they start using IPv6.

Q: What can I do if my clients are impacted by policy, network, or other restrictions in using IPv6 for Amazon S3?

Applications that are impacted by using IPv6 can switch back to the standard IPv4-only endpoints at any time.

Q: Can I use IPv6 with all Amazon S3 features?

No, IPv6 support is not currently available when using Website Hosting and access via BitTorrent. All other features should work as expected when accessing Amazon S3 using IPv6.

Q: Is IPv6 supported in all regions?

You can use IPv6 with Amazon S3 in all commercial AWS Regions except China (Beijing). You can also use IPv6 in the AWS GovCloud (US) region.