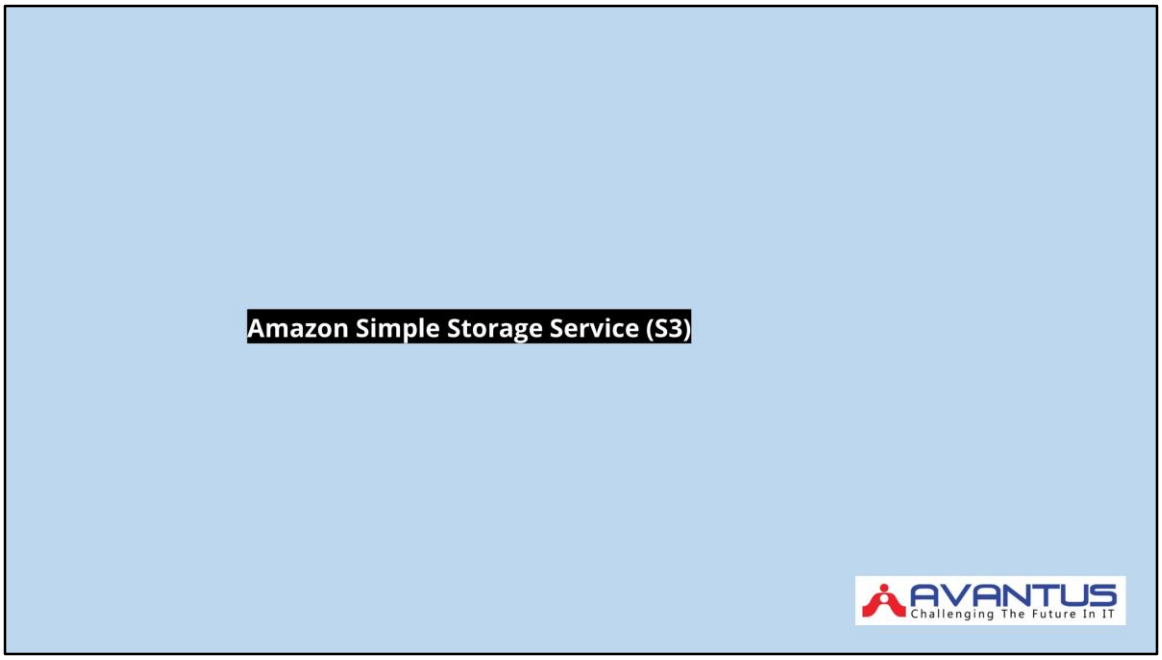
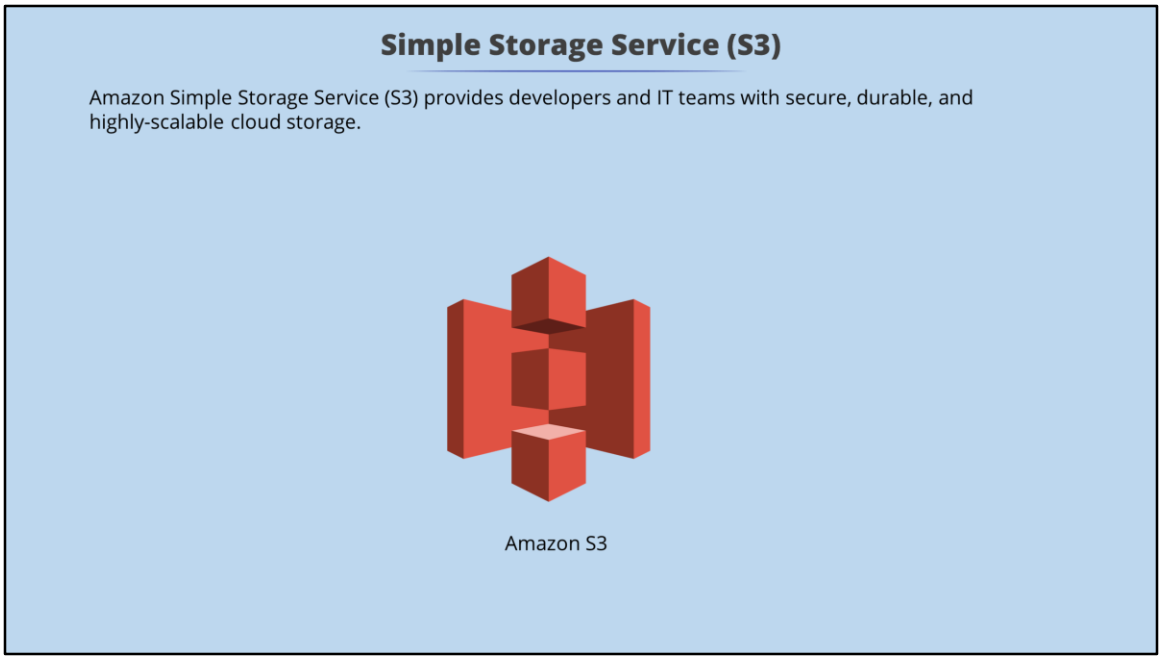


Welcome to the sixth lesson of the AWS Solutions Architect Associate level course—Amazon Simple Storage Services (S3).

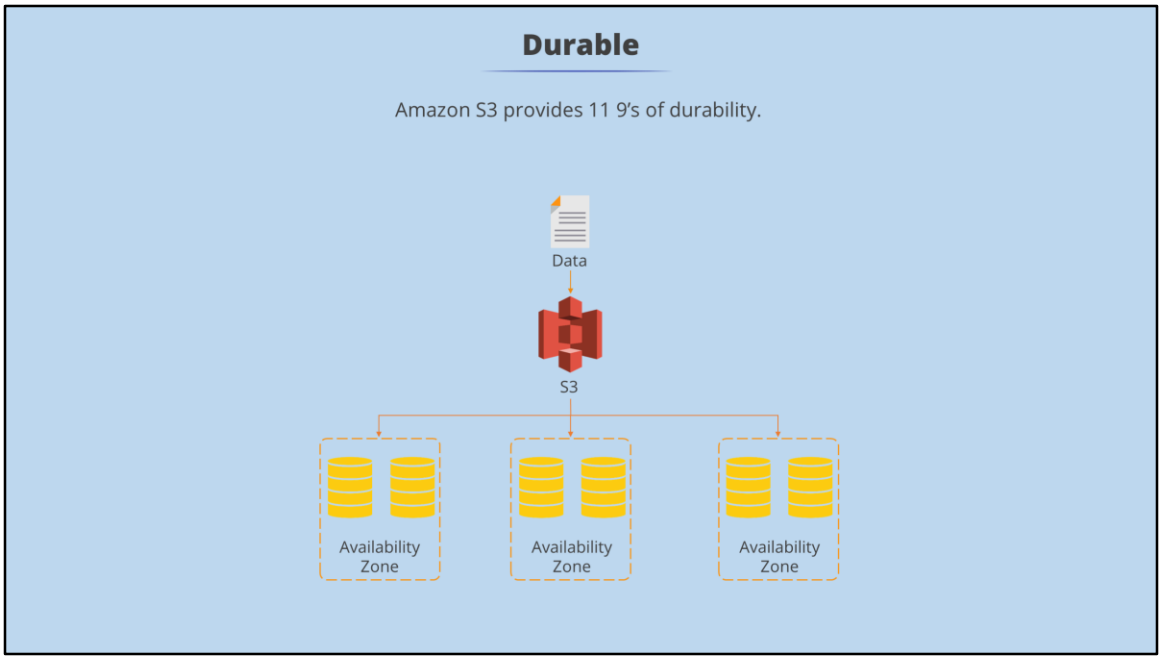
By the end of the lesson you will be able to:

- Explain S3 and its uses
- Describe Amazon S3 bucket
- List all the different storage types available
- Illustrate S3 version control and lifecycle management
- Discuss how S3 integrates with CloudFront and CDNs
- Secure and encrypt your data on S3
- Describe how to get your data in and out of S3
- Outline the AWS recommended best practices for S3
- Identify the costs associated with S3

In this section, you will learn about S3 and its benefits.

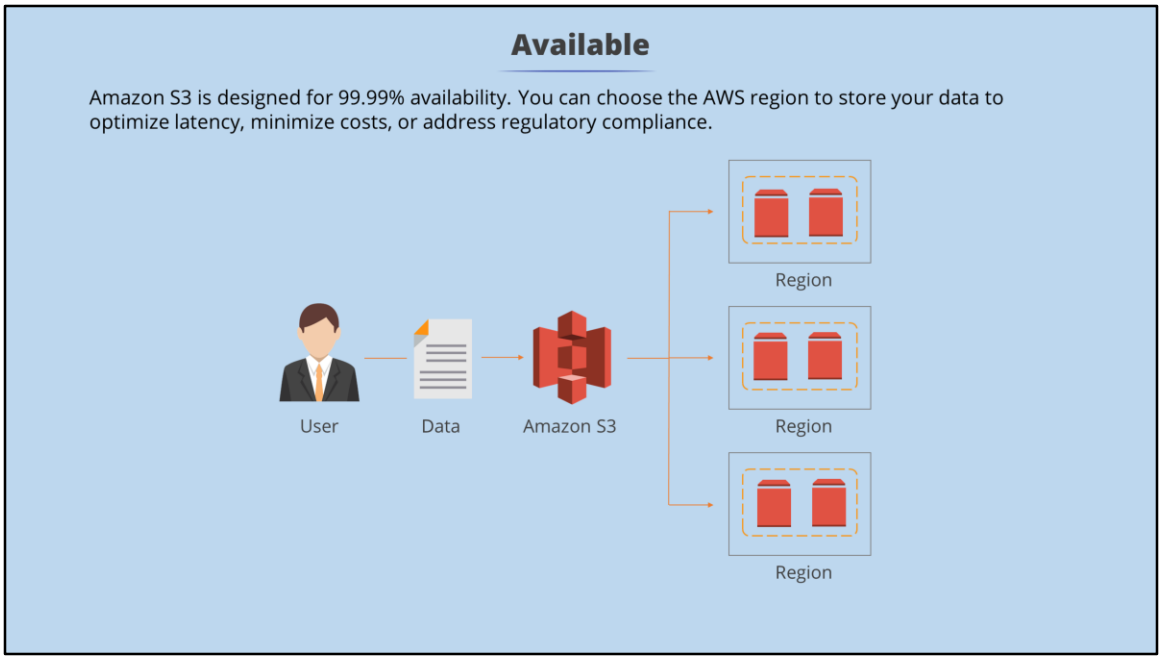


Amazon Simple Storage Service (S3) provides developers and IT teams with secure, durable, and highly-scalable cloud storage.

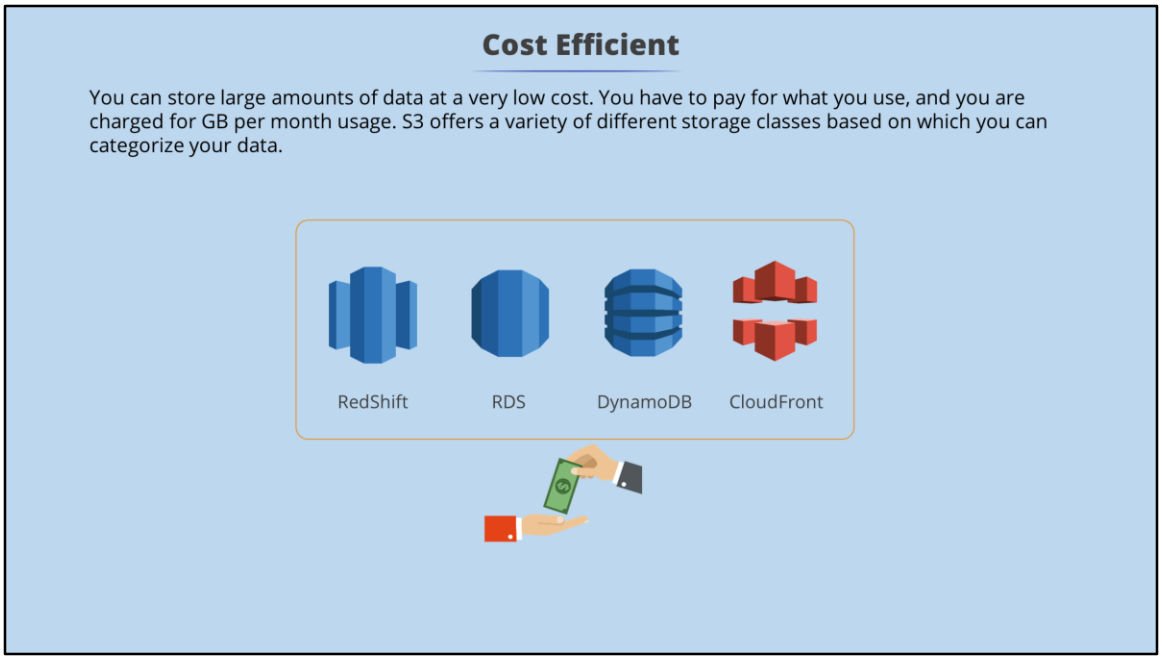


Amazon S3 is extremely durable, it provides 11 9's of durability. The data is redundantly stored across multiple facilities and multiple devices in each facility.

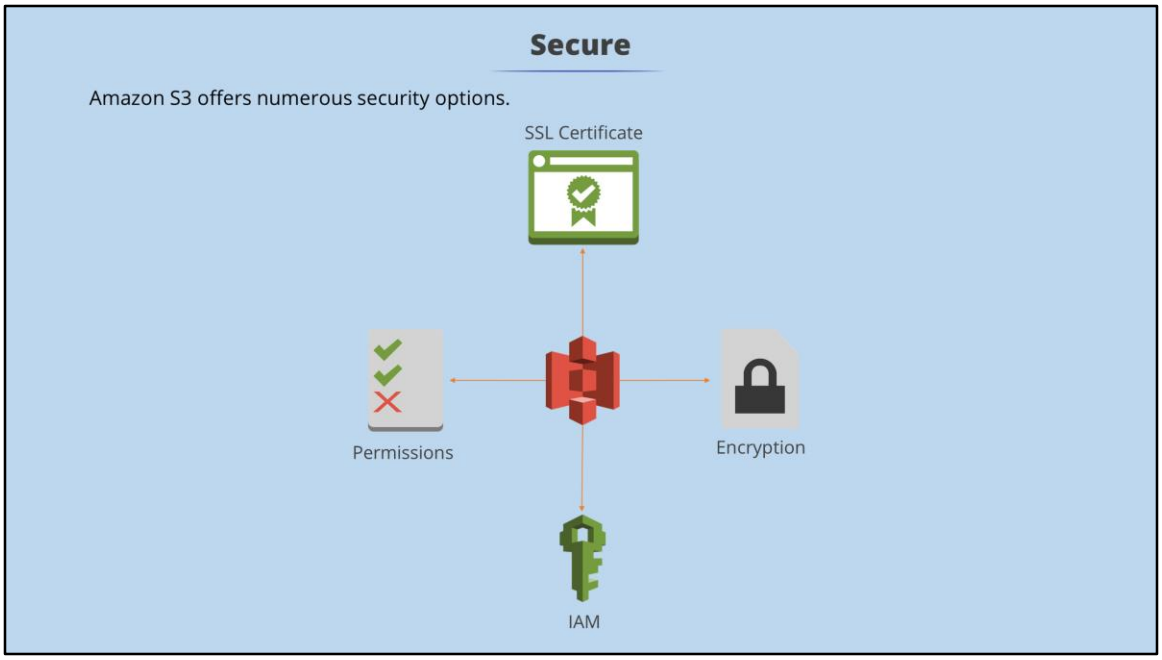
AWS S3 has 11 9's of durability, that is, if you store 10,000 objects with Amazon S3, you can, on an average, expect to incur a loss of only a single object once every 10,000,000 years.



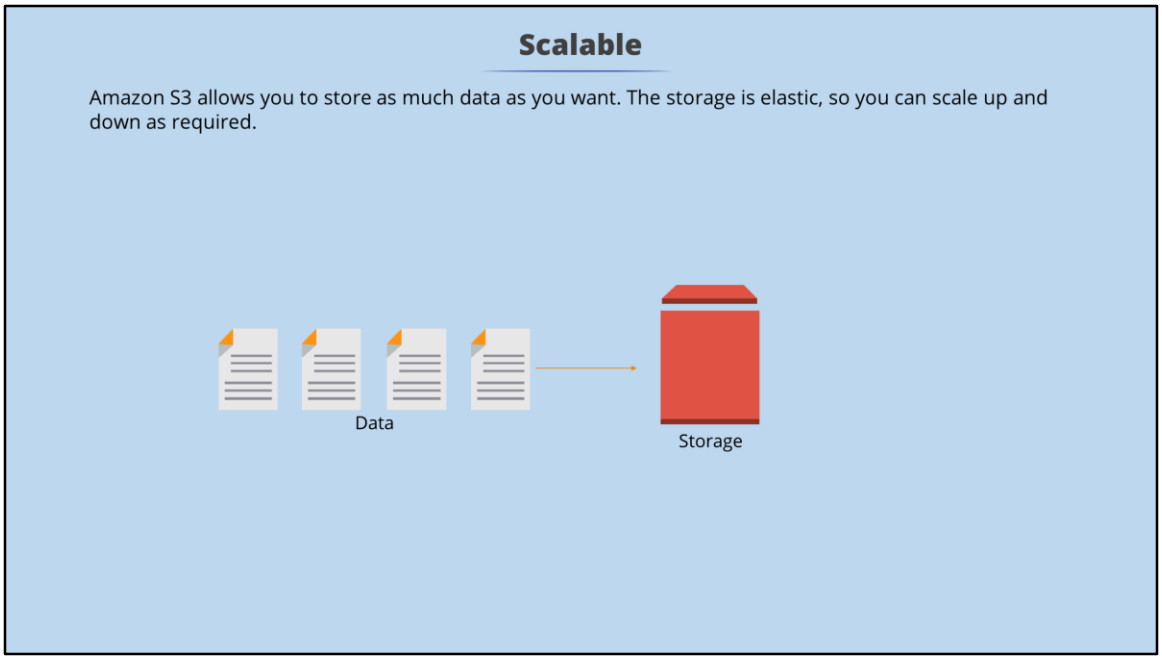
Amazon S3 is designed for 99.99% availability. You can choose the AWS region where you want to store your data to optimize latency, minimize costs, or address regulatory compliance.



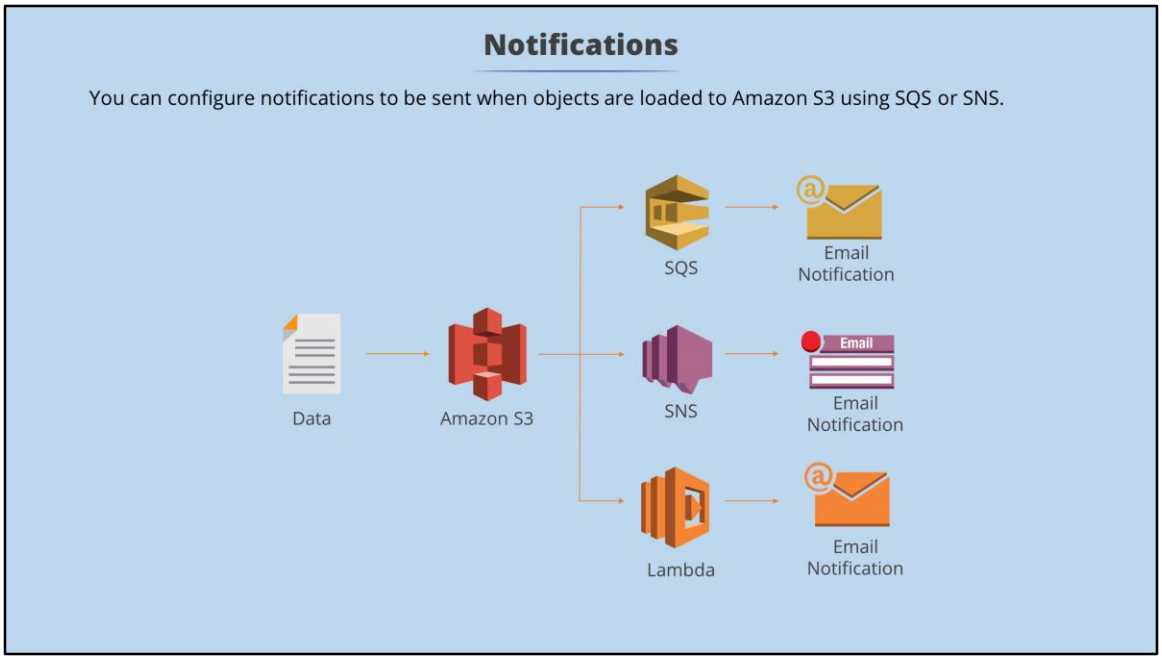
You can store large amounts of data at a very low cost. You have to pay for what you use and you are charged for GB per month usage. S3 offers a variety of different storage classes based on which you can categorize your data.



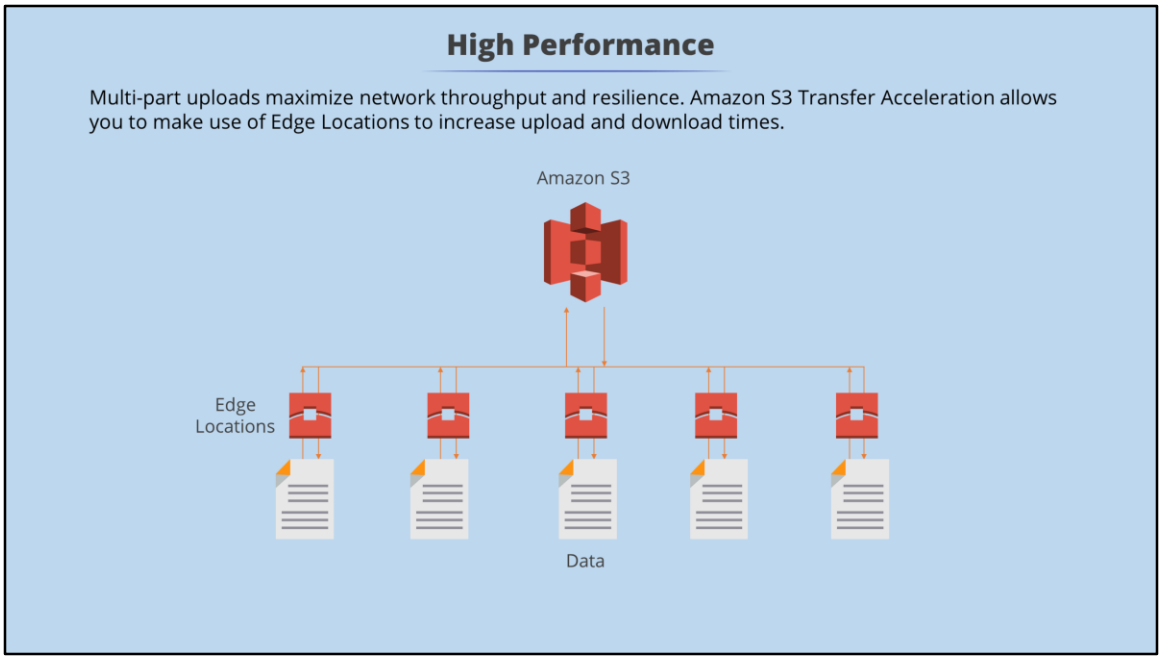
Amazon S3 supports SSL data transfer and data encryption once it is uploaded. You can control access to your data using Identity Access Management (IAM). You can also control objects' permissions using the S3 policies.



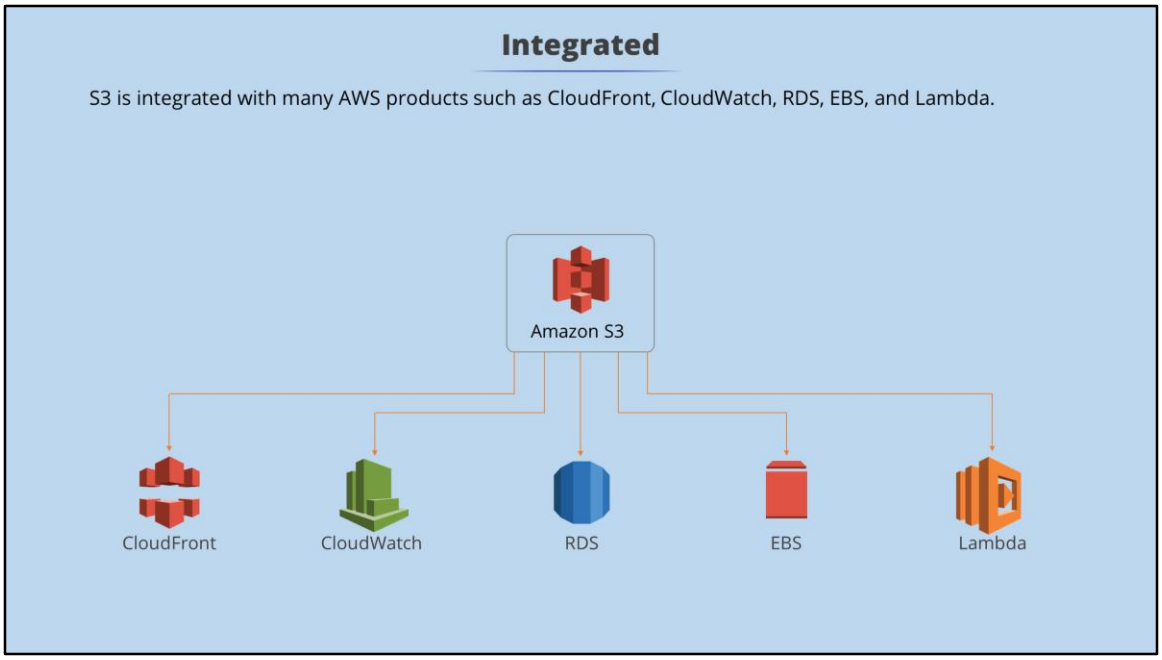
Amazon S3 allows you to store as much data as required. The storage is elastic, you can scale up and down as needed.



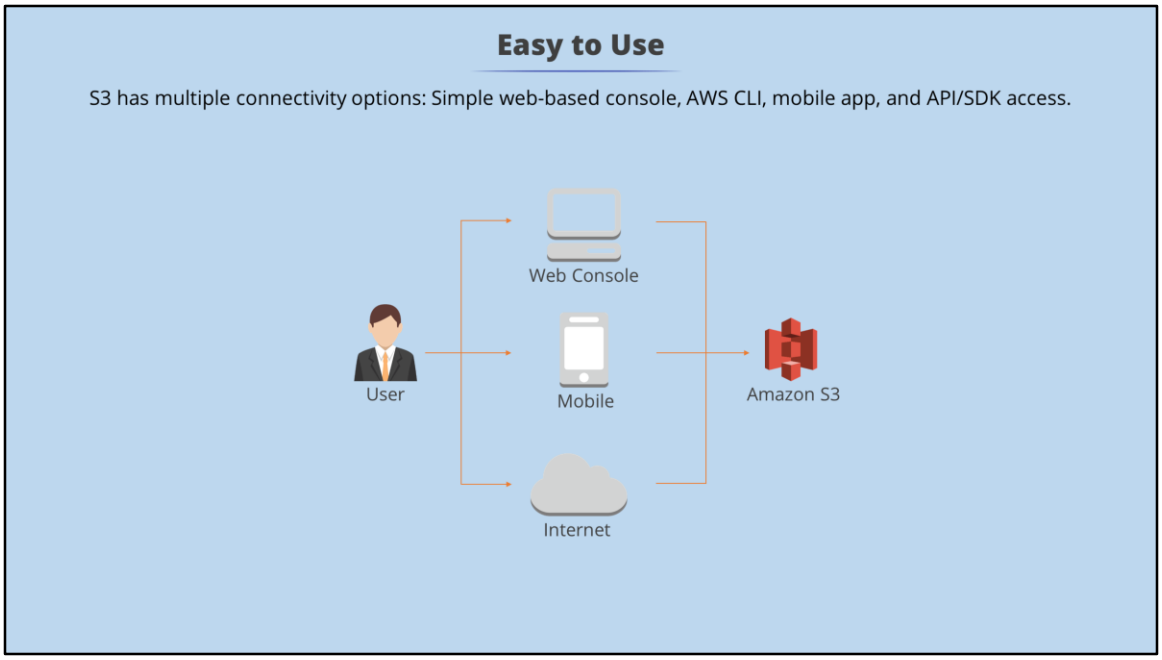
You can configure notifications to be sent when objects are loaded to Amazon S3 using SQS or SNS; this allows you to set up work-flows for your files.



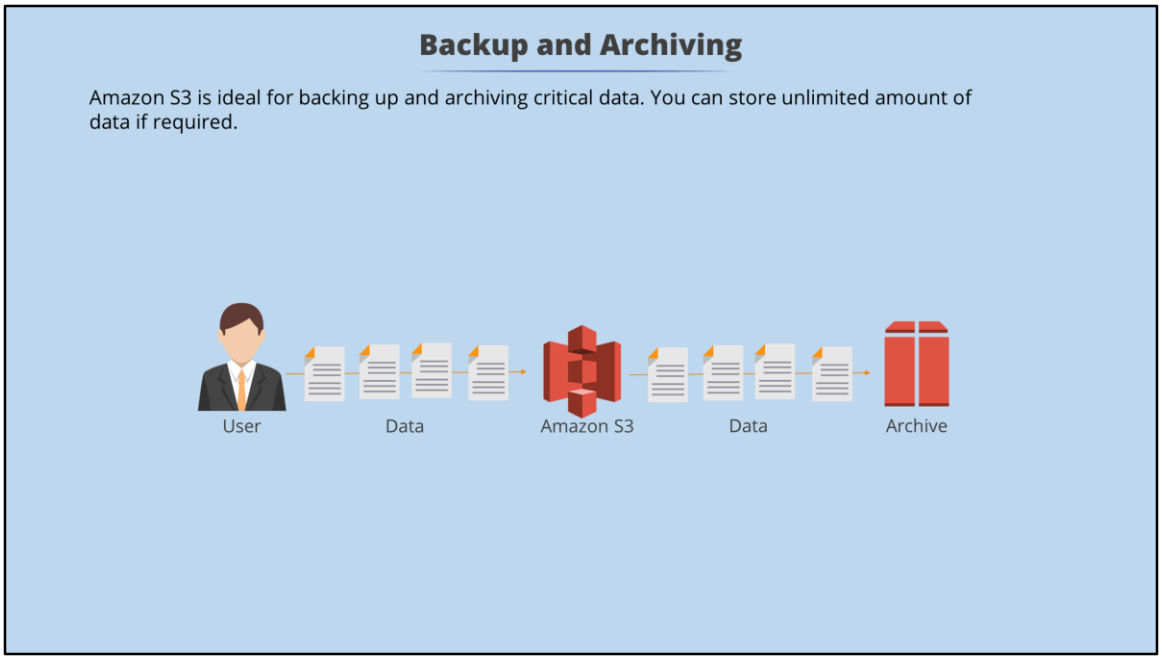
Multi-part uploads maximize network throughput and resilience. Amazon S3 Transfer Acceleration allows you to make use of edge locations to increase upload and download times.



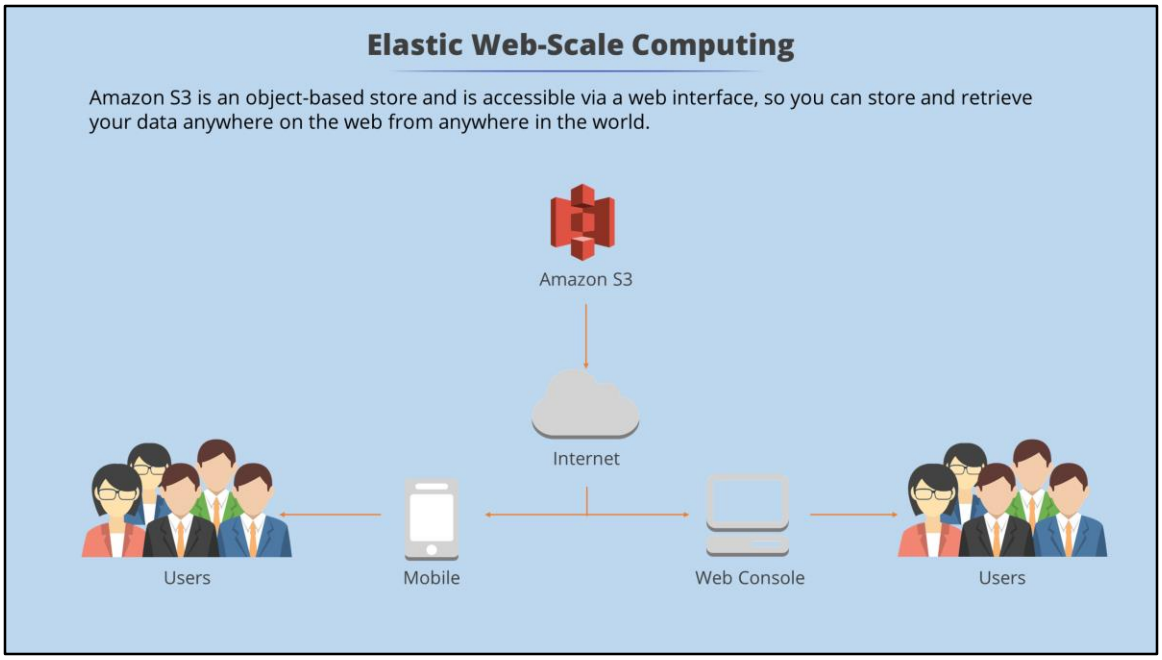
S3 is integrated with many AWS products such as CloudFront, CloudWatch, RDS, EBS, Lambda, and so on.



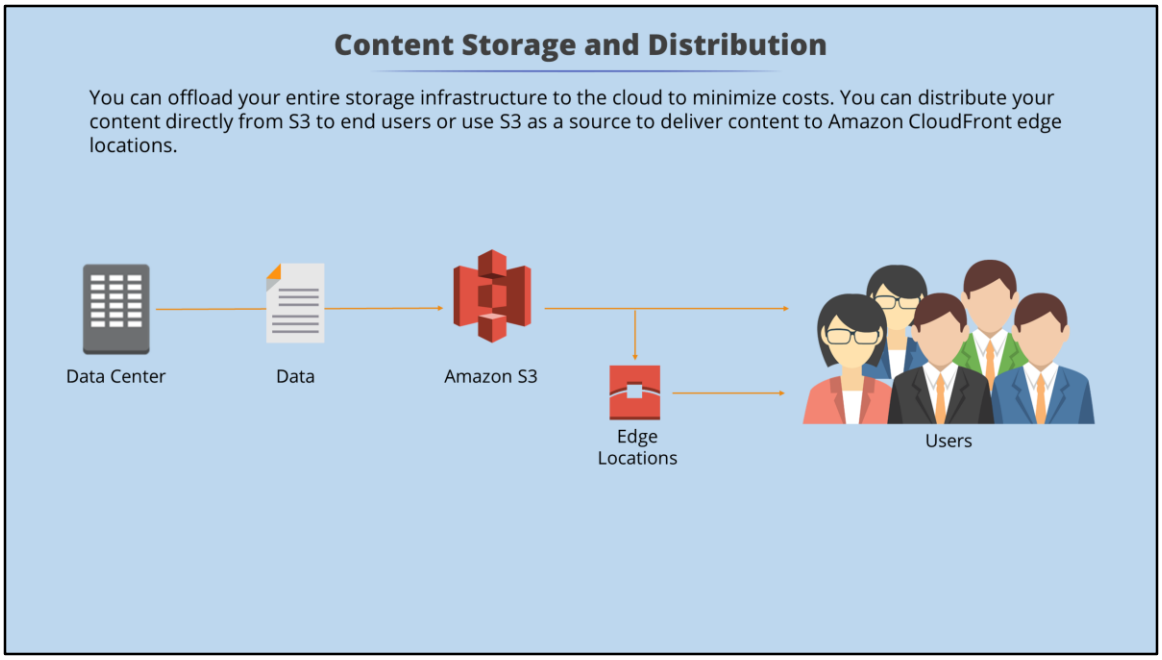
Amazon S3 is easy to use. It has multiple connectivity options: simple web based console, AWS CLI, mobile app, and API/SDK access.



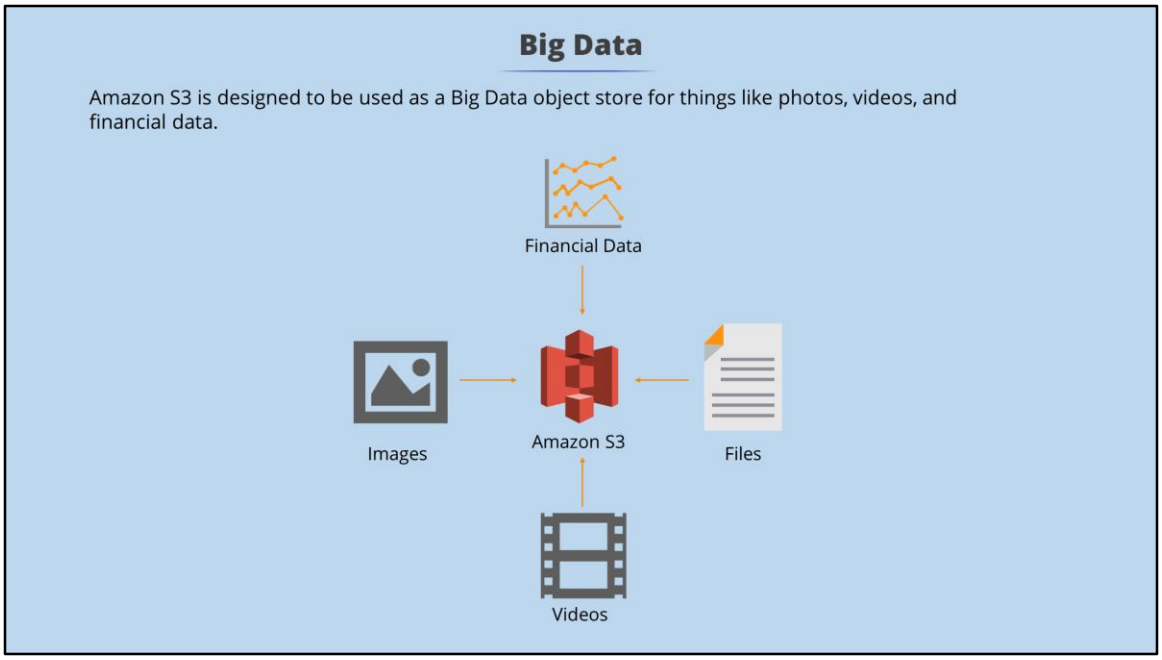
Amazon S3 is ideal for backing up and archiving critical data. You can store unlimited amount of data if required. Traditional IT infrastructure offers a finite storage capacity where you'd have to manage the backups you can retain, but with S3 you can retain as many backups as you want at a low cost.



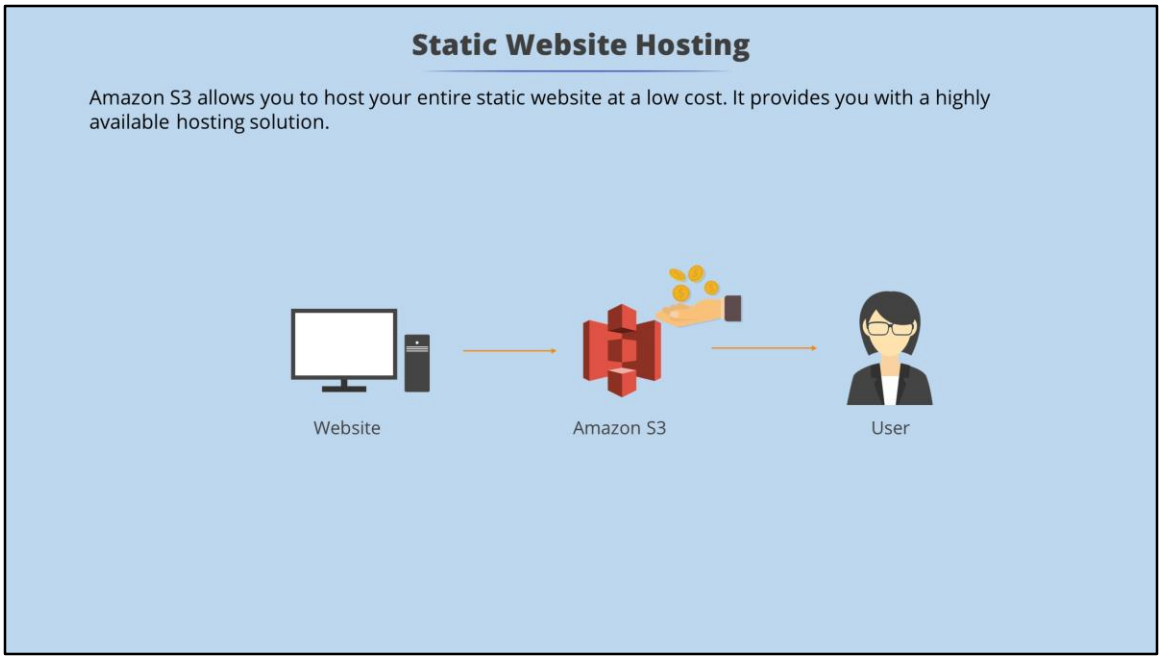
Amazon S3 is an object-based storage and is accessible via a web interface, so you can store and retrieve your data anywhere on the web from anywhere in the world.



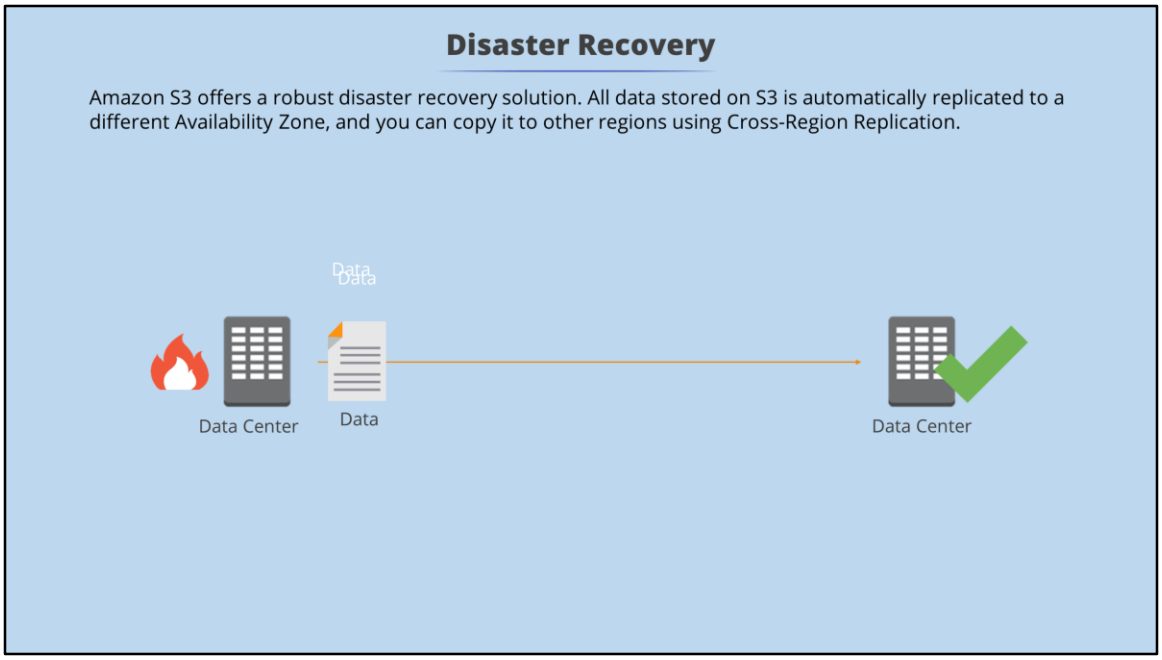
You can offload your entire storage infrastructure to the Cloud to minimize costs. You can distribute your content directly from S3 to end users, or use S3 as a source for delivering content to Amazon CloudFront edge locations.



Amazon S3 is designed to be used as a Big Data object store for objects such as photos, videos, financial data, and so on. Using AWS products and services, you can perform Big Data analytics.

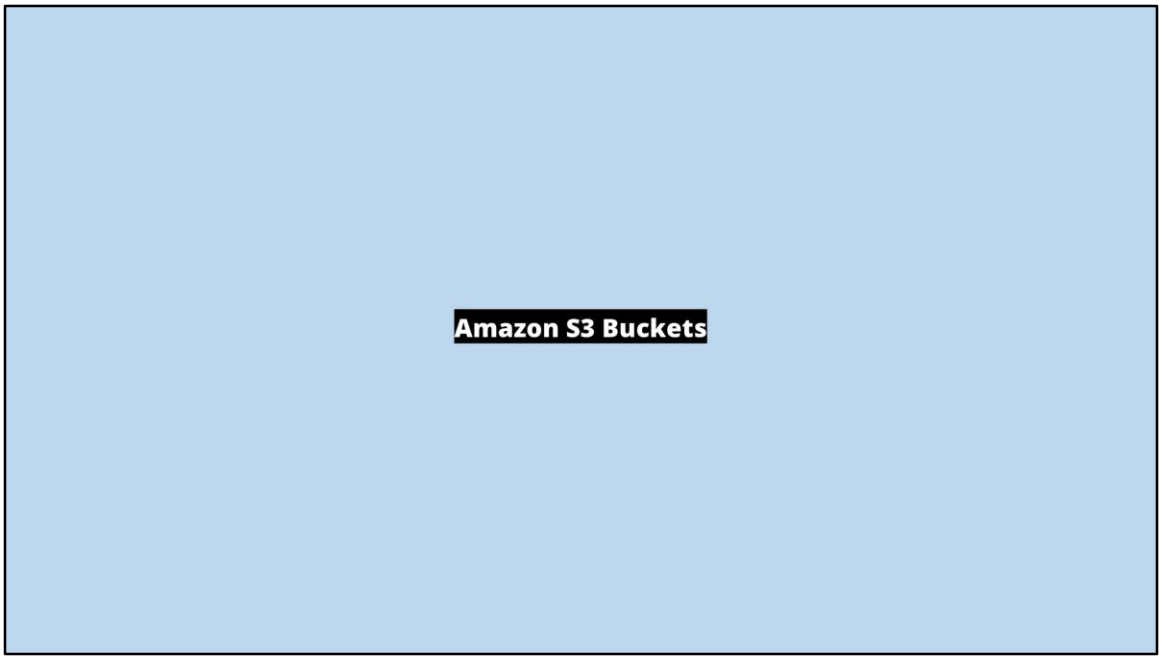


Amazon S3 allows you to host your entire static website at a low cost. It provides you with a highly available hosting solution.

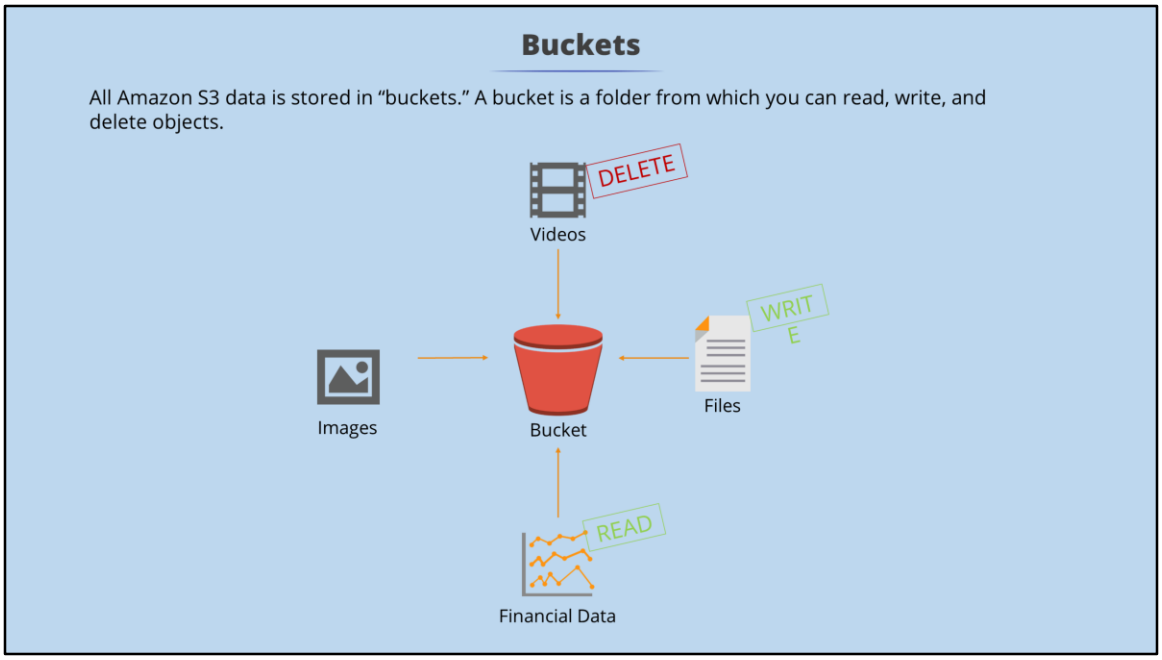


Amazon S3 offers a robust disaster recovery solution. All data stored on S3 is automatically replicated to a different Availability Zone and you can copy it to other regions using Cross-Region Replication.

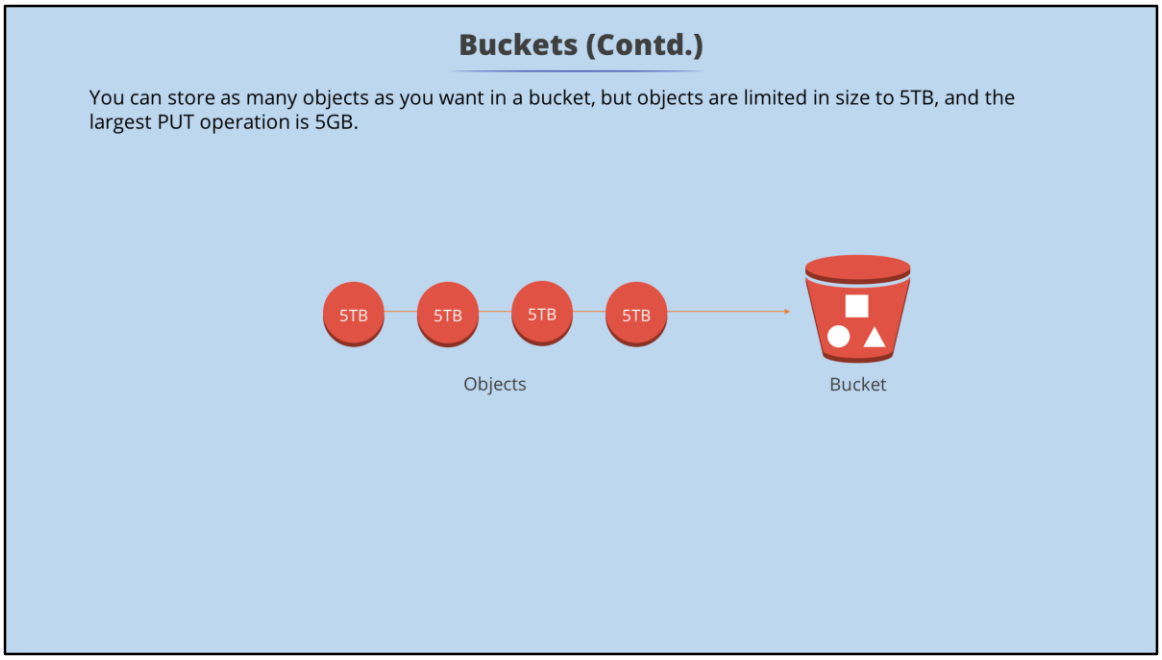
You can add further recovery options by storing multiple versions of an object for point-in-time recovery.



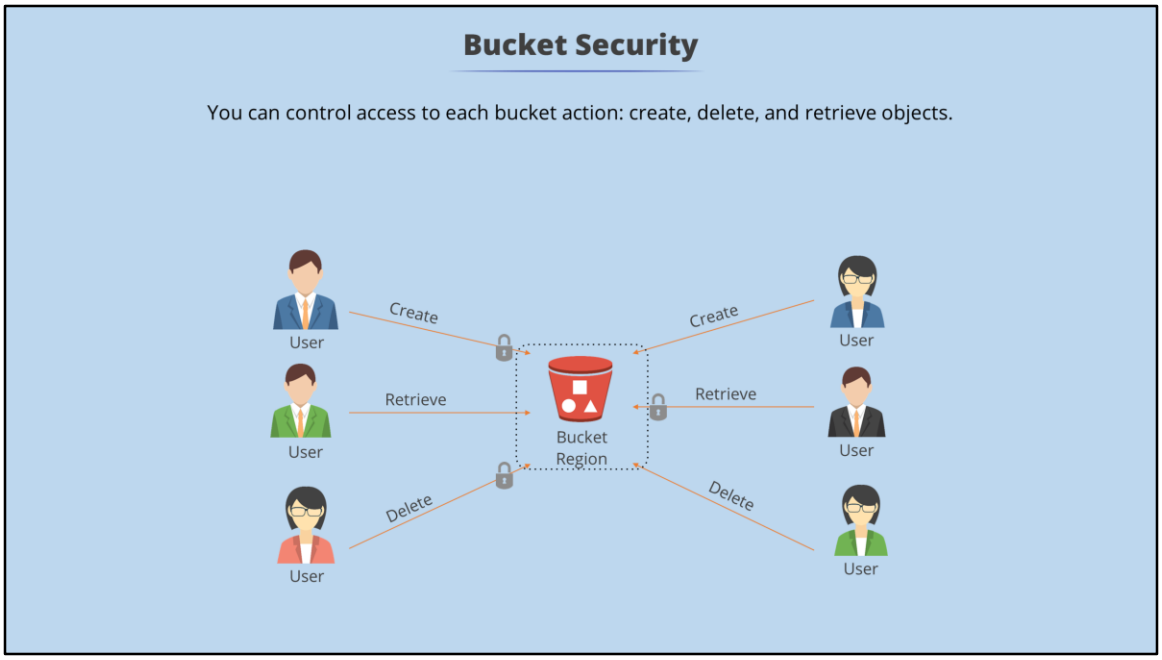
In this section, you will learn about Amazon S3 buckets.



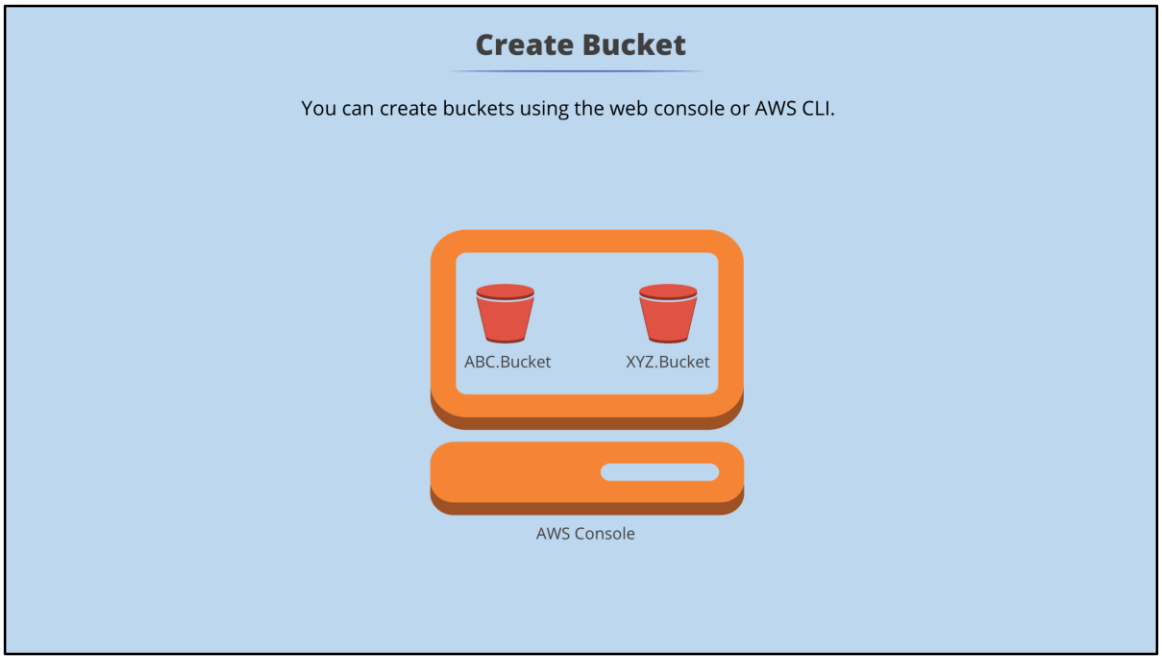
All Amazon S3 data is stored in "buckets."
A bucket is a folder from which you can read, write, and delete objects.



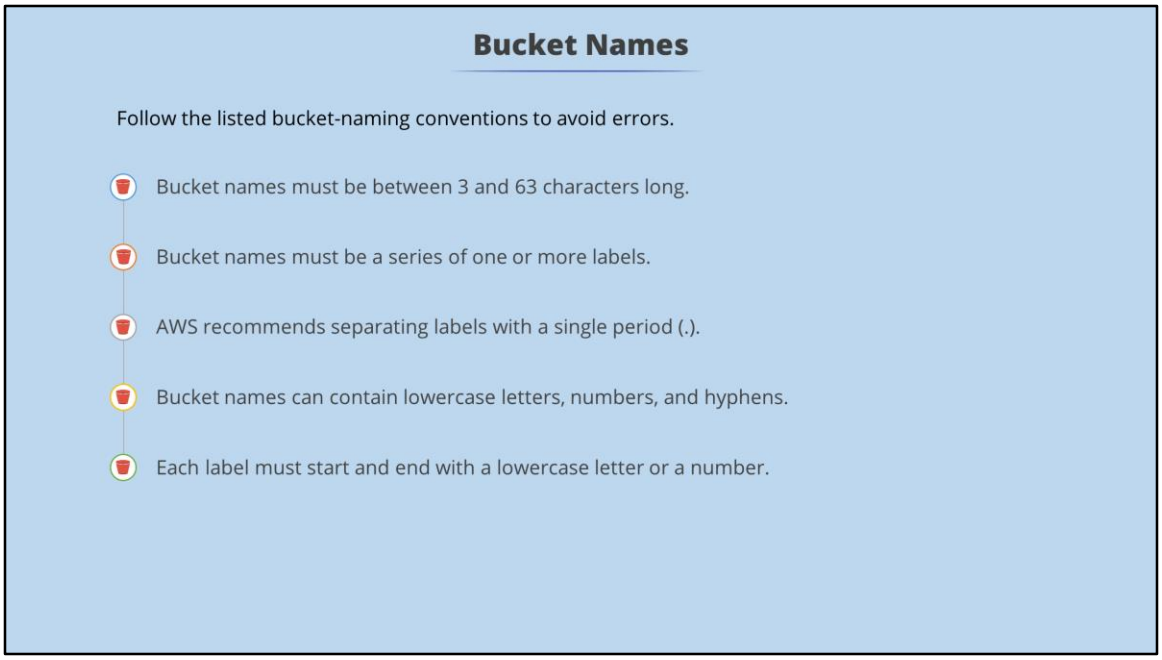
You can store as many objects as you want in a bucket, but objects are limited in size to 5TB and the largest PUT operation is 5GB.



You can control access to each bucket action: create, delete, and retrieve objects. You can control who can access bucket logs and its objects. You can choose the AWS region where a bucket is stored.



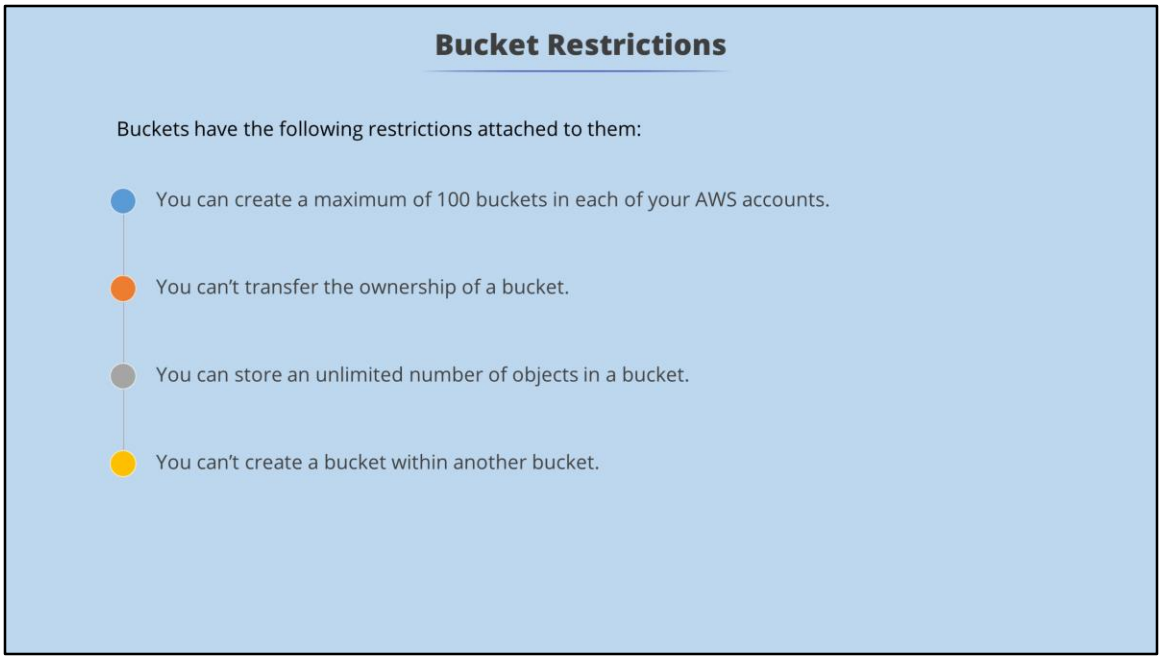
You can create buckets using the web console or AWS CLI. Ensure that the bucket name is unique across all existing bucket names in Amazon S3. AWS recommends that to ensure unique names you should add the name of your organization as a prefix to the bucket name.



The following points need to be kept in mind while creating bucket names:

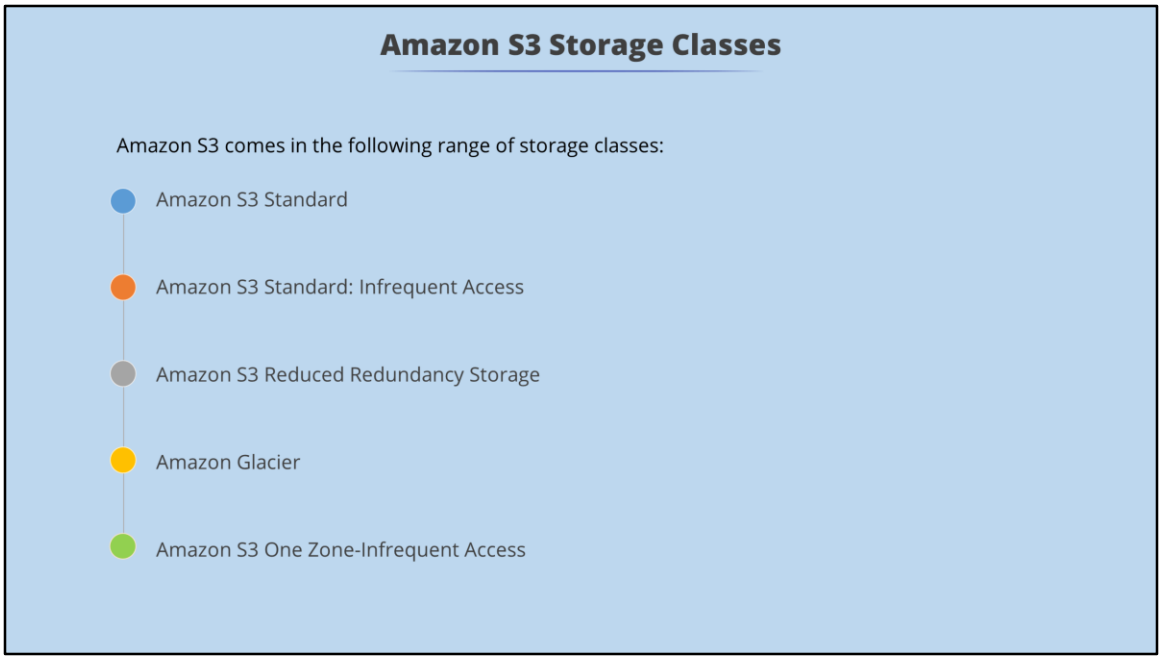
- The bucket name must be at least three characters long and can contain no more than 63 characters.
- The bucket names must be a series of one or more labels.
- AWS recommends separating labels with a single period (.).
- The bucket name can contain lowercase letters, numbers, and hyphens.
- Each label must start and end with a lowercase letter or a number.

Following are a few examples of correct bucket names: myawsbucket.1, myawsbucket, and my.aws.bucket



Bucket name ownership is not transferable; however, if a bucket is empty, you can delete it, and it will eventually become available for use.

You can create up to 100 buckets in each of your AWS accounts. There is no limit to the number of objects that can be stored in a bucket, and there is no difference in performance whether you use many buckets or just a few. However, you cannot create a bucket within another bucket.

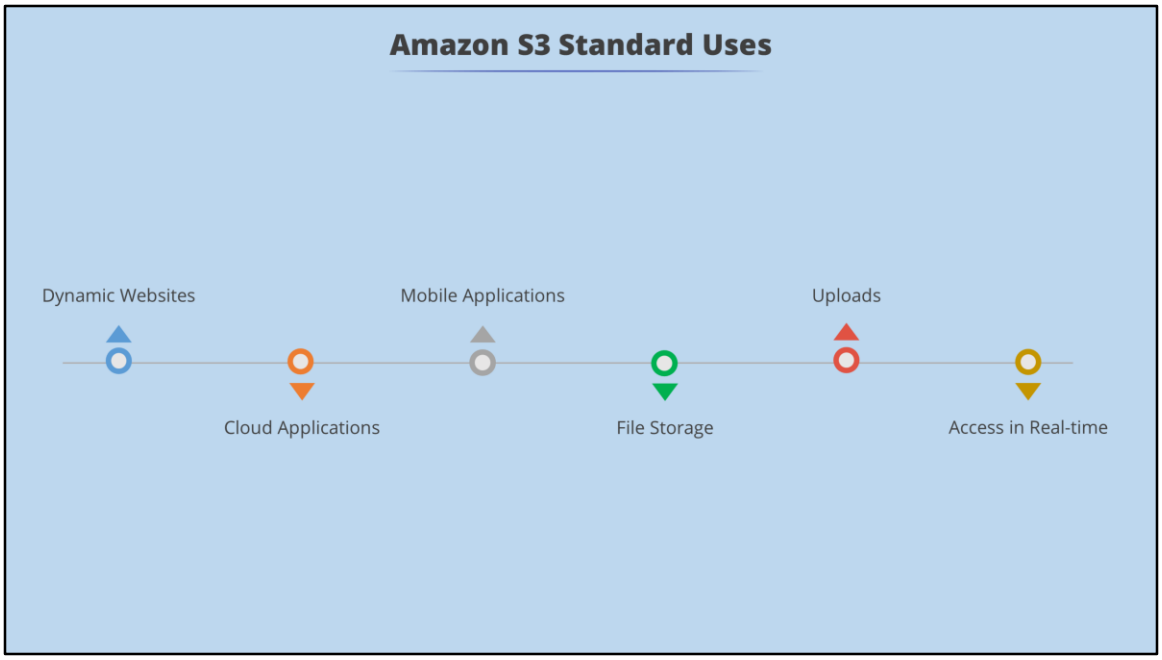


Amazon S3 comes in the following range of storage classes:

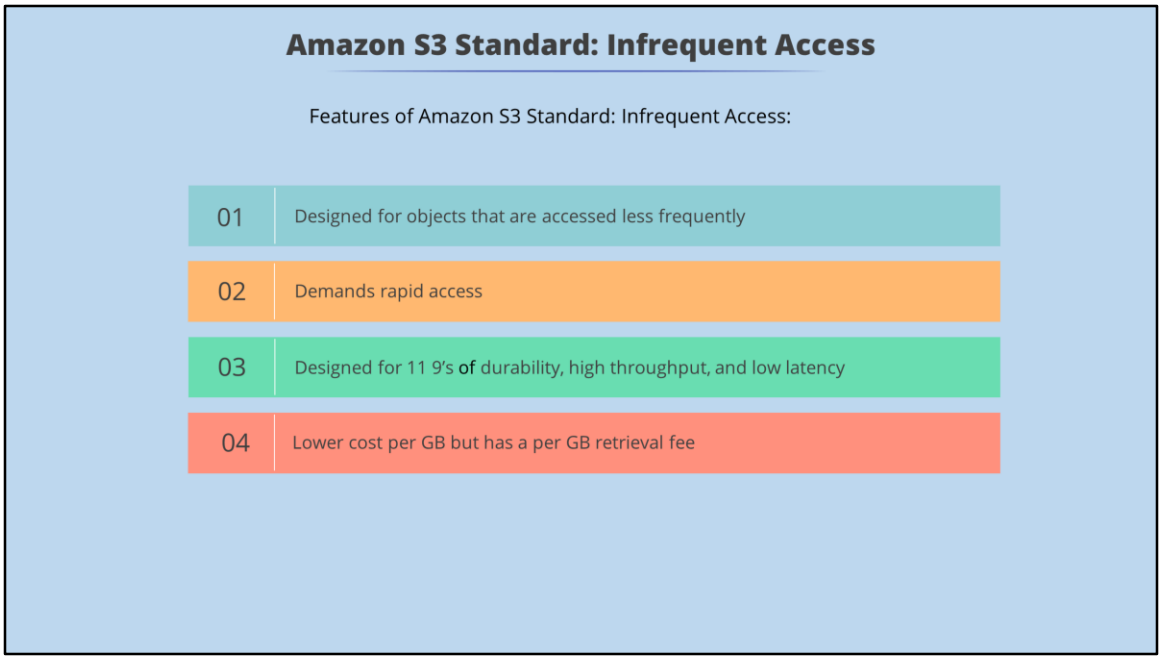
- Amazon S3 Standard
- Amazon S3 Standard—Infrequent Access
- Amazon S3 Reduced Redundancy Storage
- Amazon Glacier
- Amazon S3 One Zone-Infrequent Access

Following are the features of Amazon S3 Standard:

- High availability and durability
- Storing frequently accessed data
- 11 9's durability
- 99.99% availability
- Low latency and high throughput



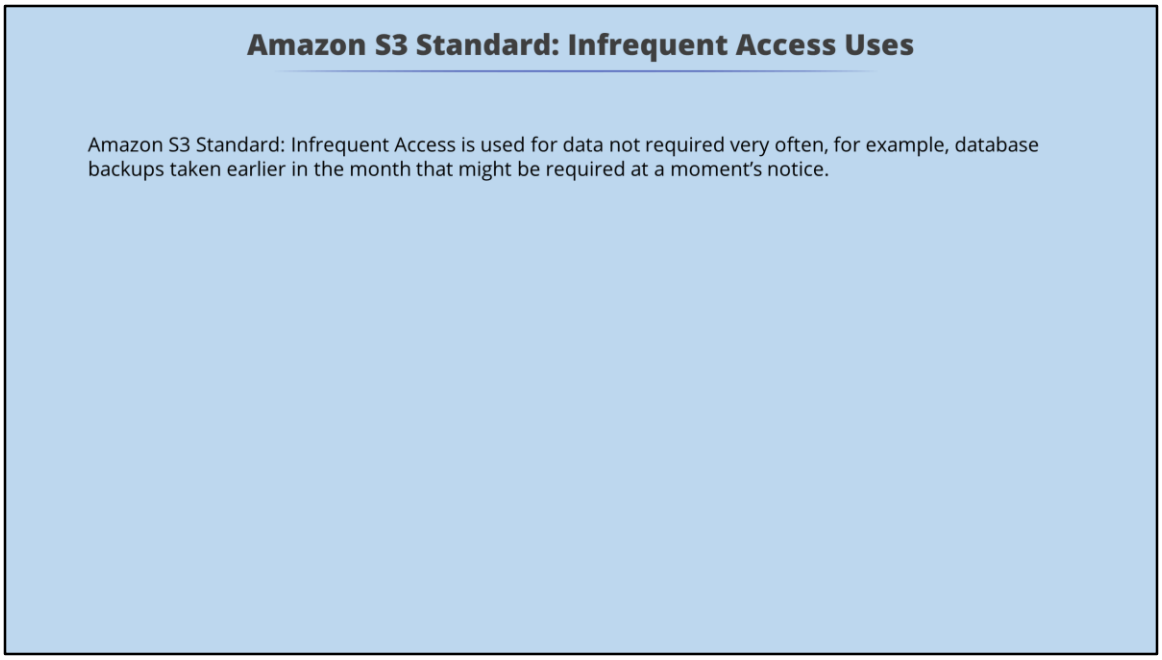
Amazon S3 Standard is ideal for dynamic websites, cloud applications, content distribution, mobile apps, and so on. It can be used as a regular file storage for your business. End users can use Amazon S3 to upload photos and videos.



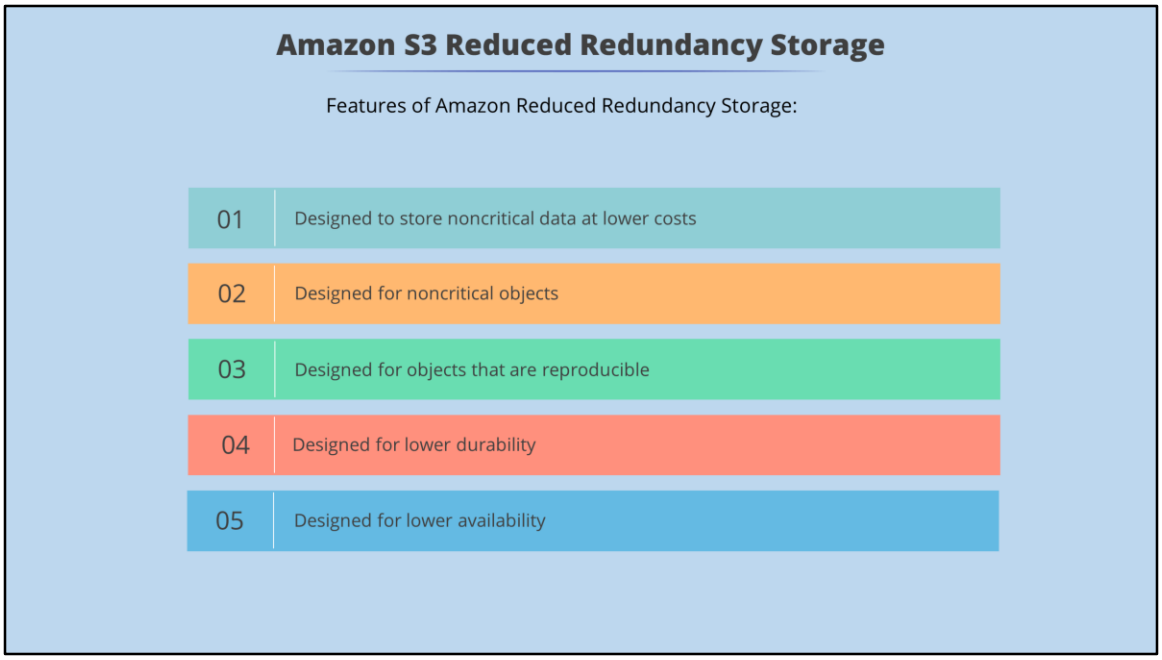
Following are the features of Amazon S3 Standard—Infrequent Access:

- It is designed for objects that are accessed less frequently, but when required, they demand fast access.
- It provides 11 9's of durability.
- It is designed for 99.9% availability.
- It offers low latency and high throughput.

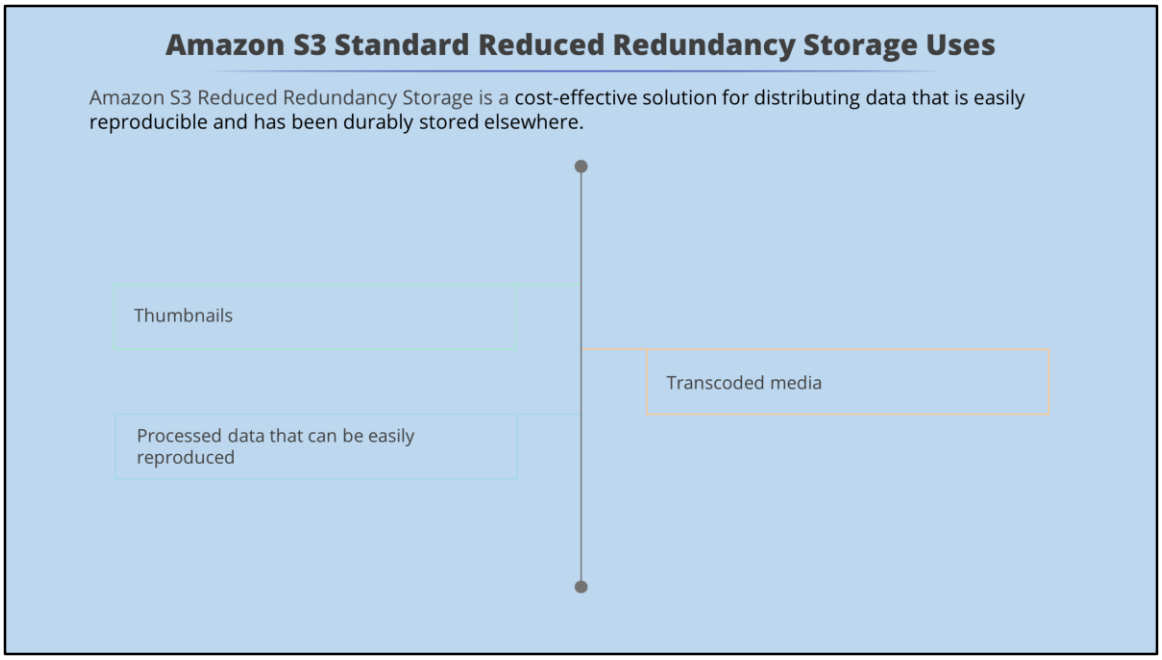
Standard—Infrequent Access offers the same durability, throughput, and low latency of Standard, but with a lower cost per GB and a per GB retrieval fee.



Amazon S3 Standard—Infrequent Access is used for data that is not often required, for example, database backups taken earlier in the month, but which might be required at a moment's notice.



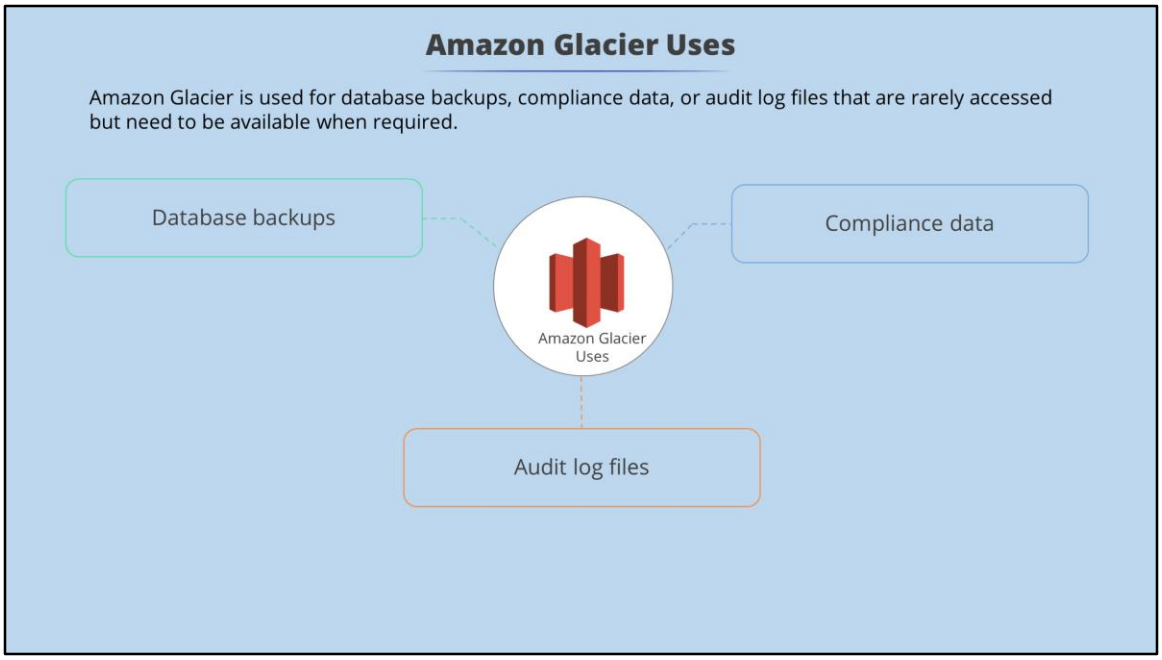
Reduced Redundancy Storage (RRS) is an Amazon S3 storage option that enables customers to reduce their costs by storing non-critical, reproducible data at lower levels of redundancy than Amazon S3's standard storage.



It provides a cost-effective solution for distributing or sharing data that has been durably stored elsewhere, or for storing thumbnails, transcoded media, or other processed data that can be easily reproduced.

Following are the features of Amazon Glacier:

- It is used to archive data that is rarely accessed, and where a retrieval time of several hours won't matter.
- It provides 11 9's of durability.
- It provides a vault lock feature that enforces compliance.
- It provides a long-term archive feature.



Amazon Glacier is used for database backups, compliance data, or audit log files that are rarely accessed but need to be available when required.

For example, database backups from months or even years ago.

Amazon Retrieval Options

The following table presents the retrieval options:

	EXPEDITED	STANDARD	BULK
Retrieval Time	1-5 minutes	3-5 hours	5-8 hours
Retrieval Requests	Charge per request	Charge per 1000 requests	Lowest charge per 1000 requests +
Data Retrieval	Charge per GB	Lower charge per GB	Lowest charge per GB

Amazon S3 storage provides four different options:

- Standard
- Standard—IA
- Glacier
- RRS

Amazon S3 Storage Comparison

The following table presents the comparisons between the four storage options:

	STANDARD	STANDARD - IA	GLACIER	RRS	S3 One Zone-IA
Durability	99.999999999 %	99.999999999 %	99.999999999 %	99.99%	99.999999999 %+
Availability	99.99%	99.99%	N/A	99.99%	99.5%
Min Storage Duration	N/A	30 days	90 days	N/A	30 days
Retrieval Fee	N/A	Per GB retrieved	Per GB retrieved	N/A	Per GB retrieved
First Byte Latency	Milliseconds	Milliseconds	Minutes-hours	Milliseconds	Milliseconds

Amazon S3 storage provides four different options:

- Standard
- Standard—IA
- Glacier
- RRS
- S3 One Zone-IA

Coming soon.

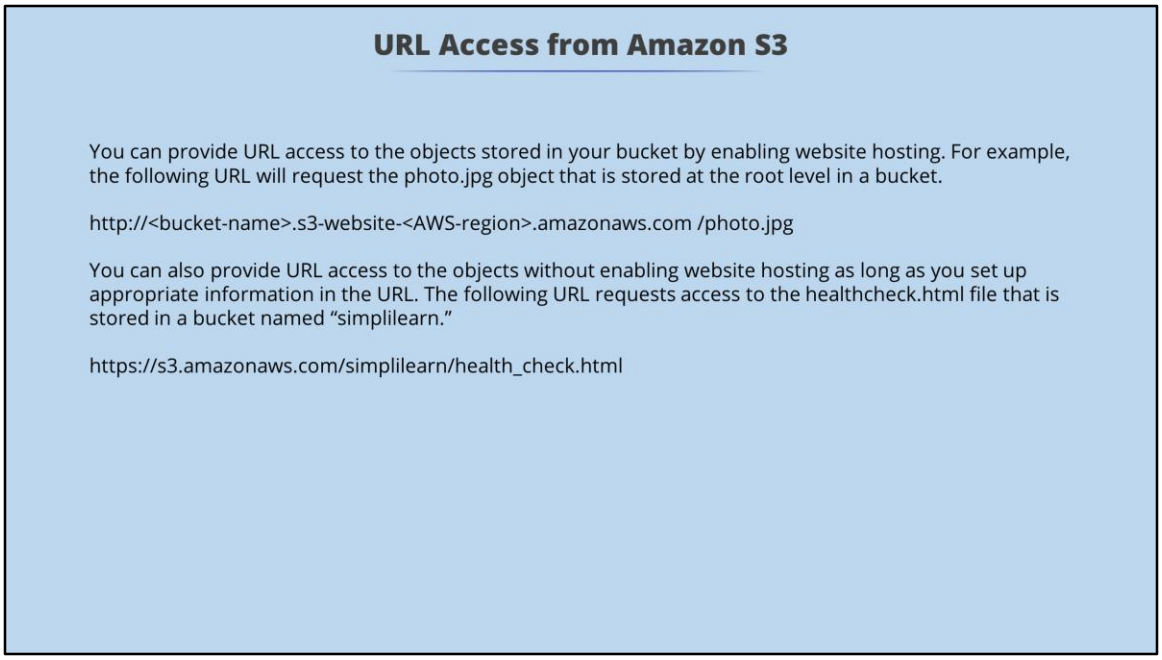
S3 Glacier Deep Archive is a new Amazon S3 storage class that provides secure, durable object storage for long-term data retention and digital preservation. S3 Glacier Deep Archive offers the lowest price of storage in AWS, and reliably stores any amount of data. It is the ideal storage class for customers who need to make archival, durable copies of data that rarely, if ever, need to be accessed.

This means that customers can eliminate the need for on-premises tape libraries, and no longer have to manage hardware refreshes and re-write data to new tapes as technologies evolve.

All data stored in S3 Glacier Deep Archive can be retrieved within 12 hours

You can easily host static websites from Amazon S3. You can configure buckets for static website hosting and then upload your website code to your bucket. It will become accessible from the URLs.

The naming convention is: <bucket-name>.s3-website-<AWS-region>.amazonaws.com

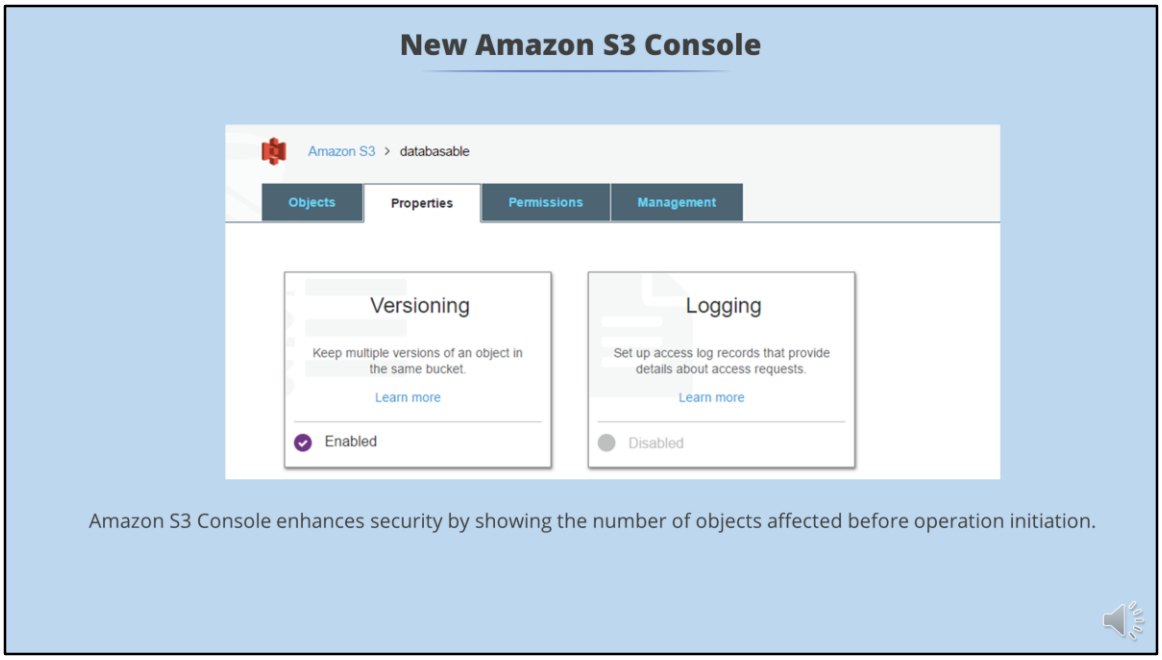


You can provide URL access to the objects stored in your bucket by enabling website hosting. For example, the following URL will request the photo.jpg object, which is stored at the root level in a bucket.

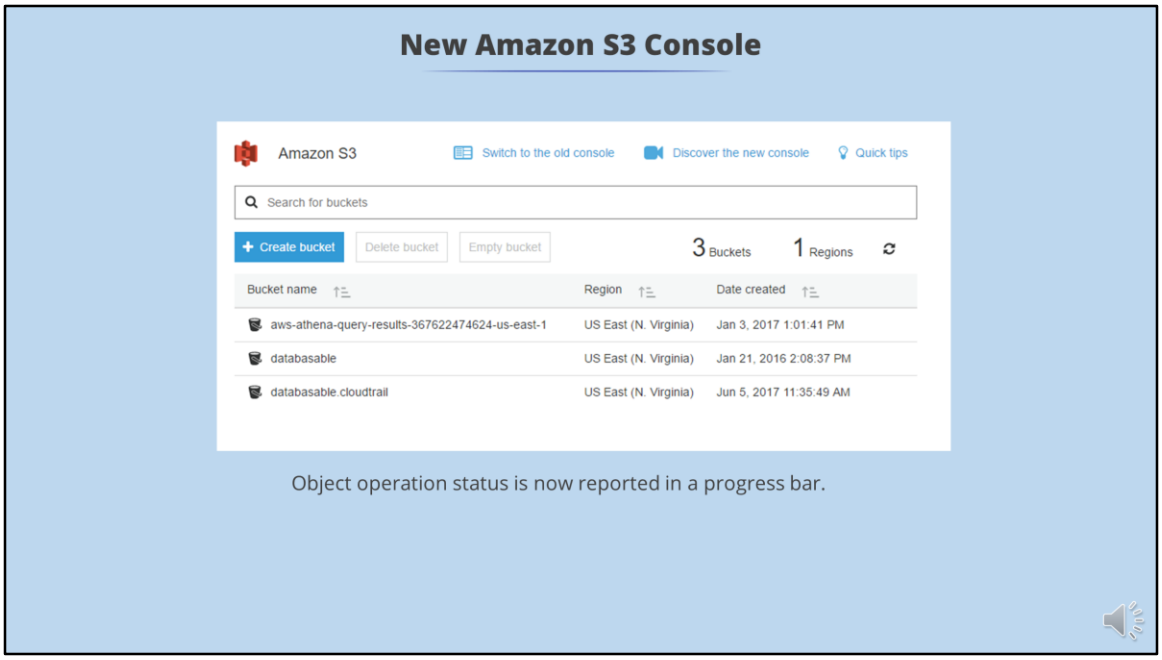
http://<bucket-name>.s3-website-<AWS-region>.amazonaws.com /photo.jpg.

You can also provide URL access to the objects without enabling website hosting as long as you set up appropriate information in the URL. The following URL requests access to the healthcheck.html file that is stored in a bucket named “simplilearn.”

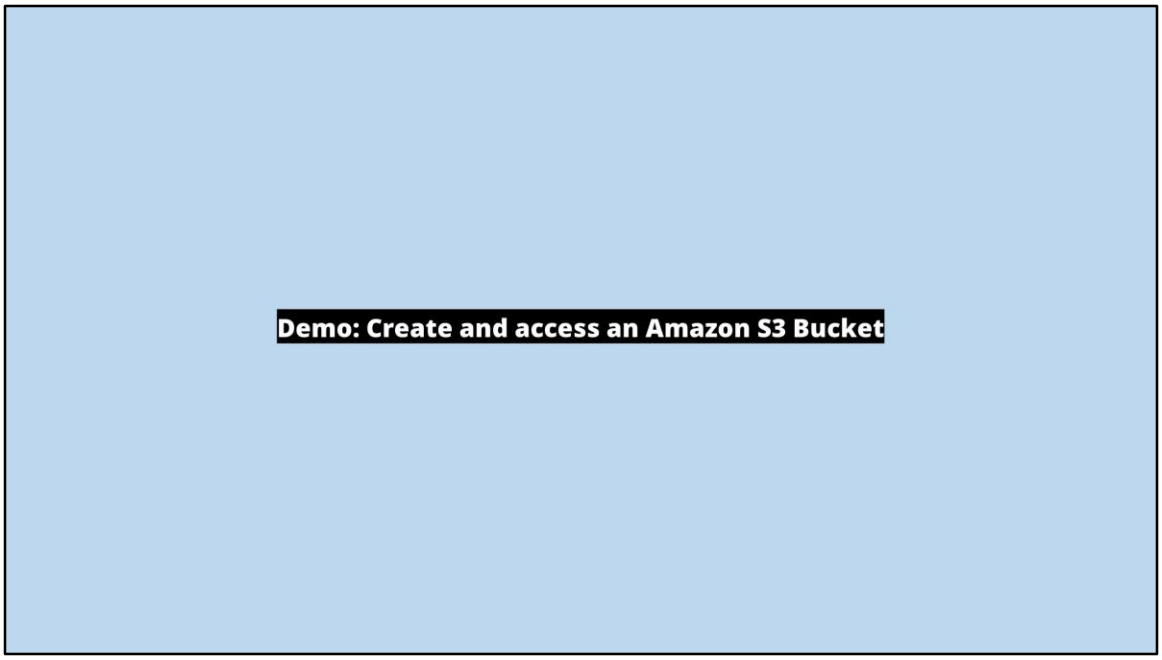
https://s3.amazonaws.com/simplilearn/health_check.html



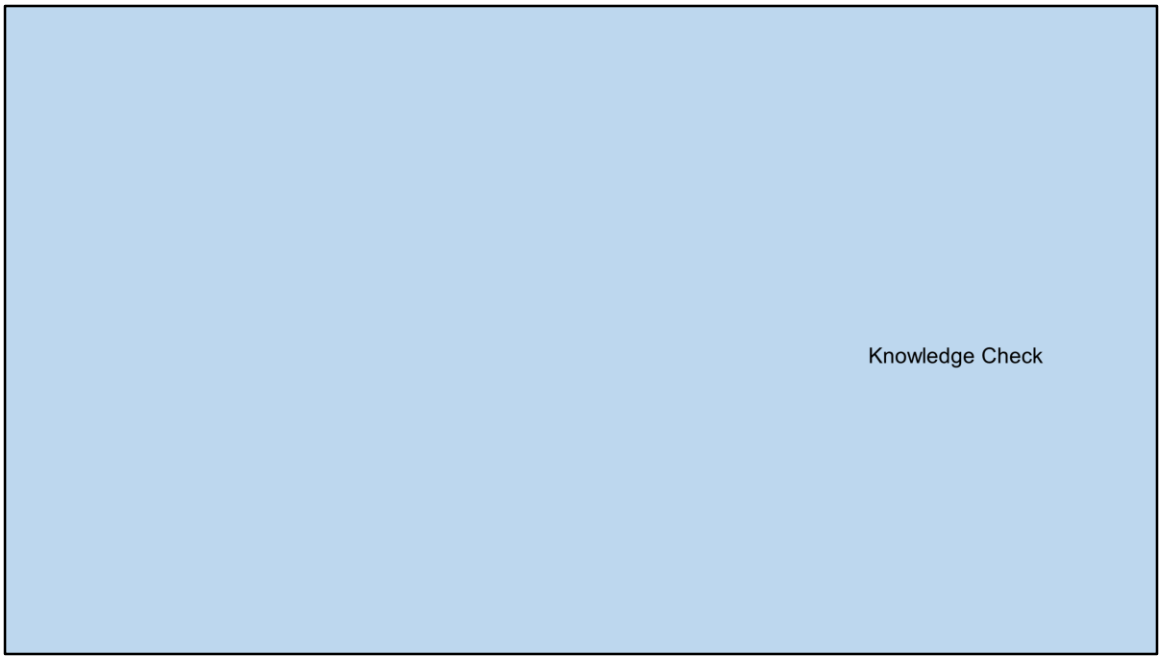
Amazon S3 storage options four different options:
Standard
Standard-IA
Glacier
RRS
And you can compare their features in the table given.



Amazon S3 storage options four different options:
Standard
Standard-IA
Glacier
RRS
And you can compare their features in the table given.



In this demonstration, you will learn how to create and access an Amazon S3 bucket.


Knowledge Check

Knowledge Check 1

Amazon S3 bucket names have to be ____.

a. unique in each region

b. unique in each Availability Zone

c. unique across all regions

d. more than 63 characters long

1

Amazon S3 bucket names have to be ____.

unique in each region

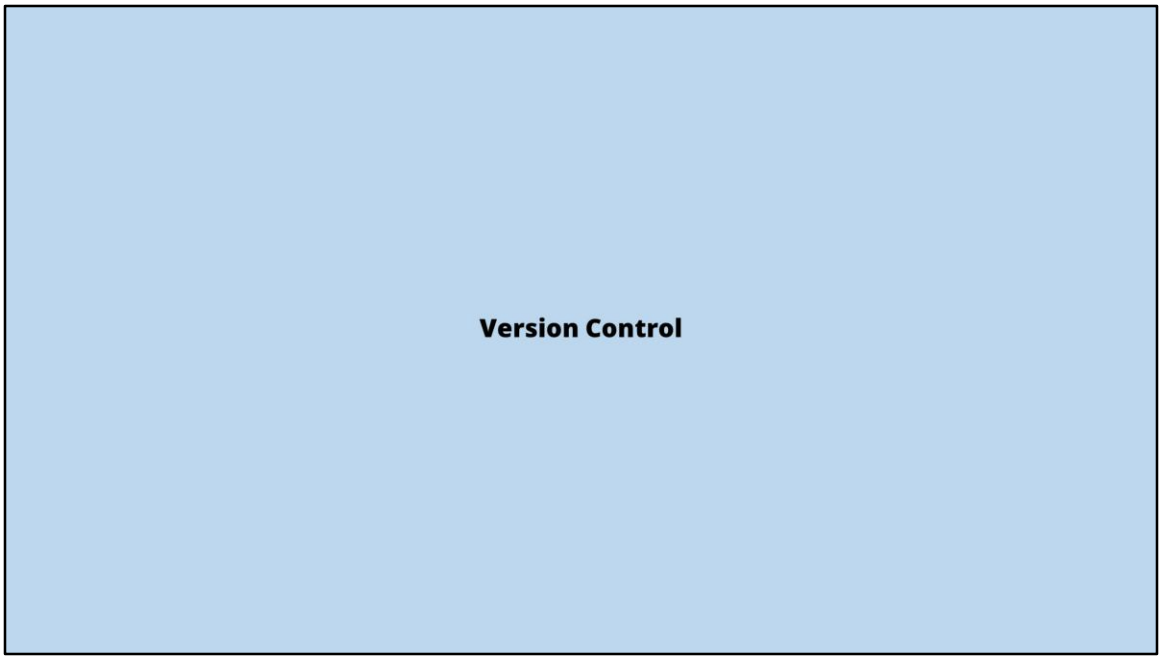
unique in each Availability Zone

unique across all regions

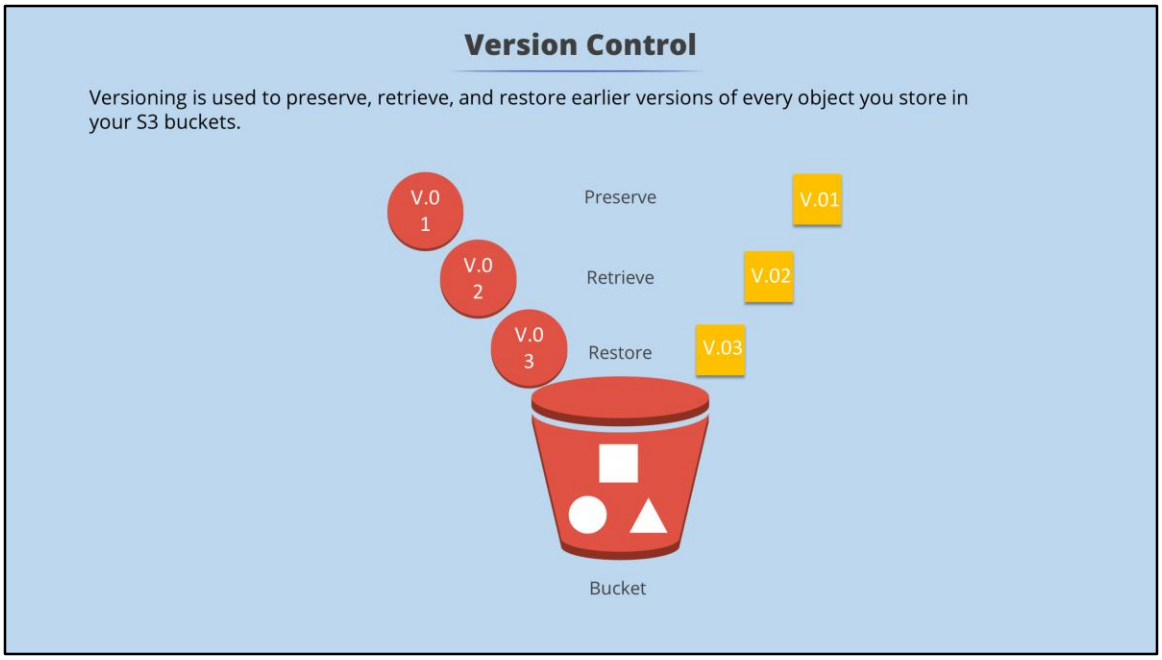
more than 63 characters long

c

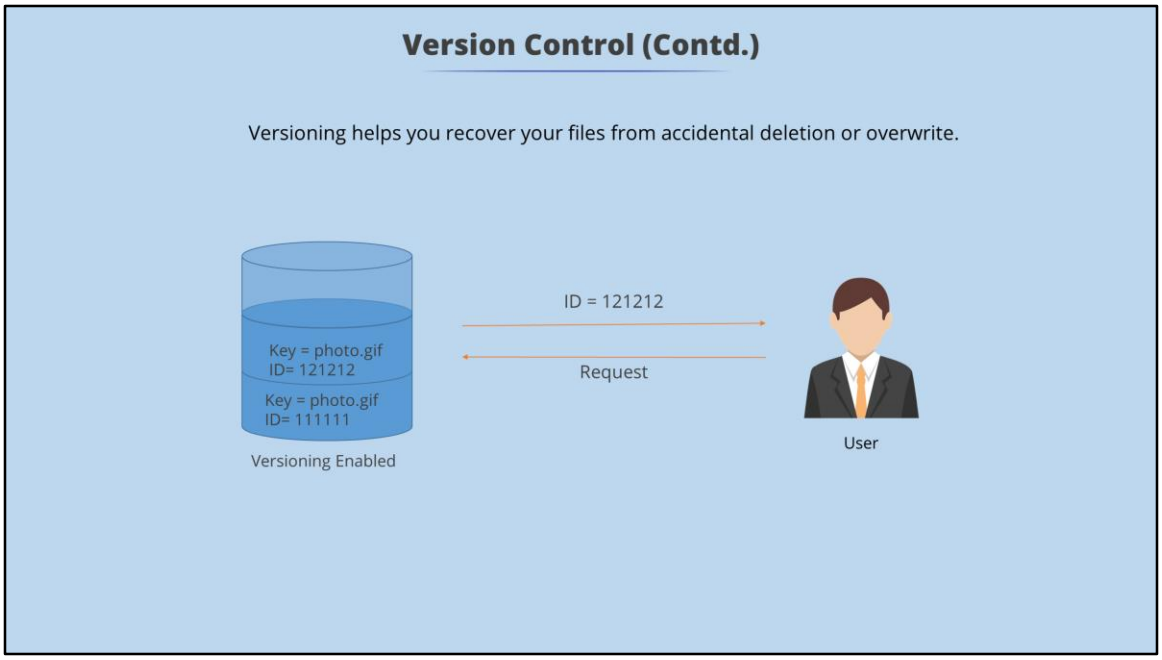
Amazon S3 bucket names have to be unique globally and between 3 and 63 characters in length.



In this section, you will learn about version control.



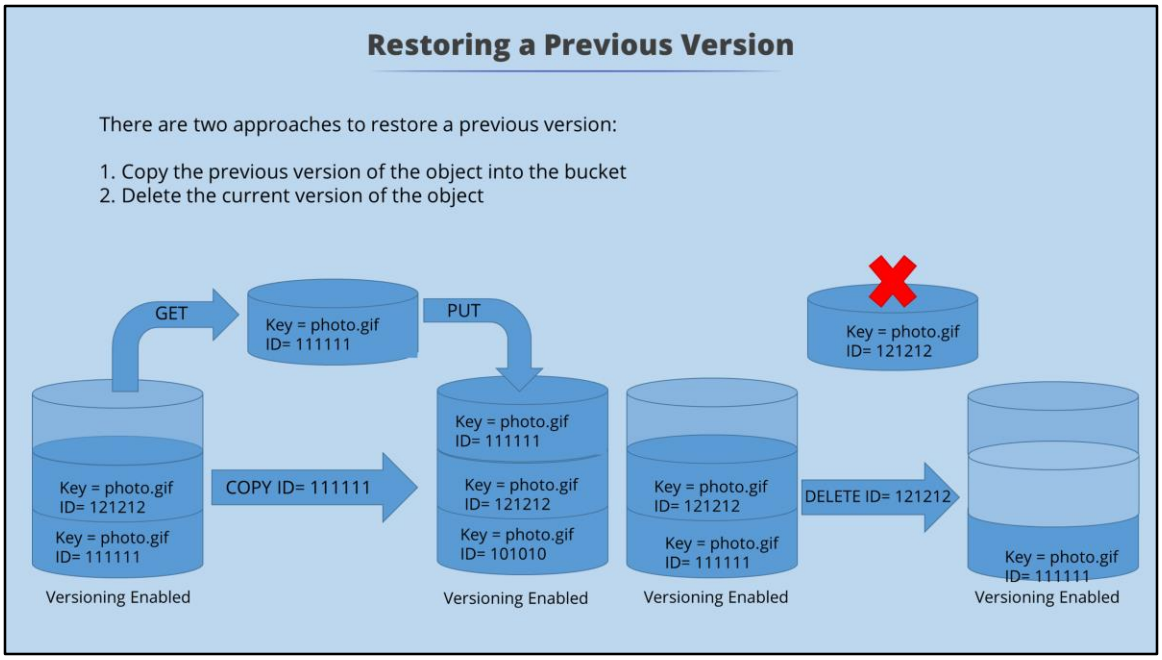
Versioning is used to preserve, retrieve, and restore earlier versions of every object you store in your S3 buckets.



Version control allows you to recover objects due to end-user mistakes and application failures by protecting against accidental deletion or overwriting.

Each time a file is updated or deleted, the previous version is stored, but with a different ID. A request for a file will, by default, restore the file with the most recent ID.

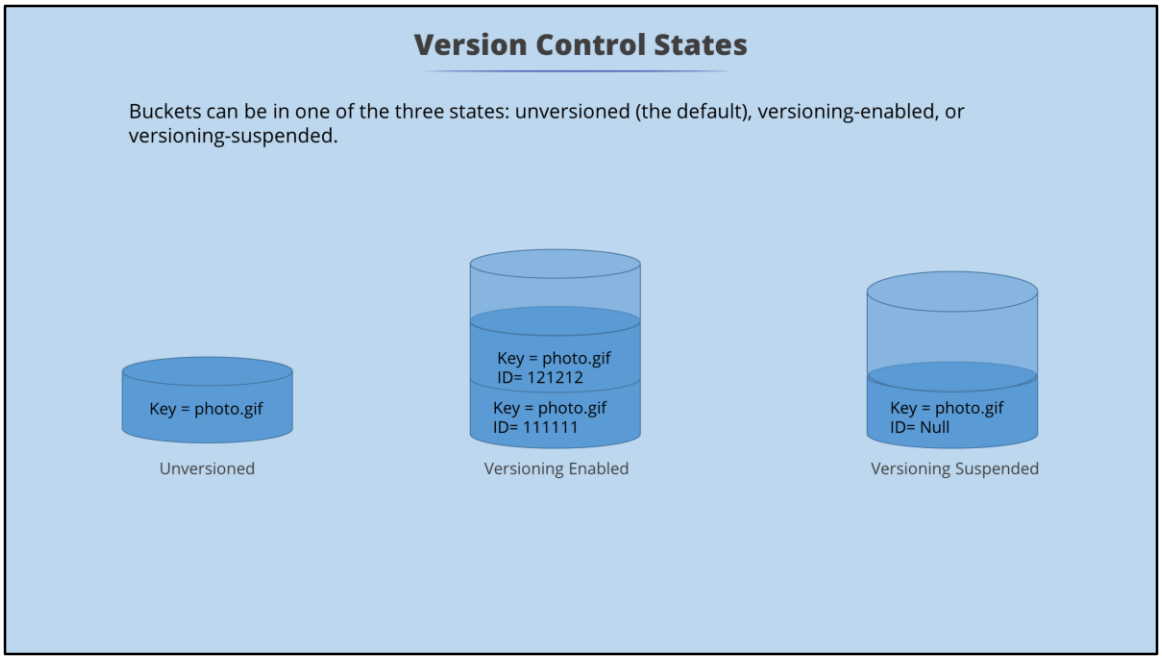
Storage rates are applied for every version that is stored.



There are two approaches to restore a previous version:

- Copy the previous version of the object into the bucket: Copied object will become the current version of that object and all previous object versions will be preserved.

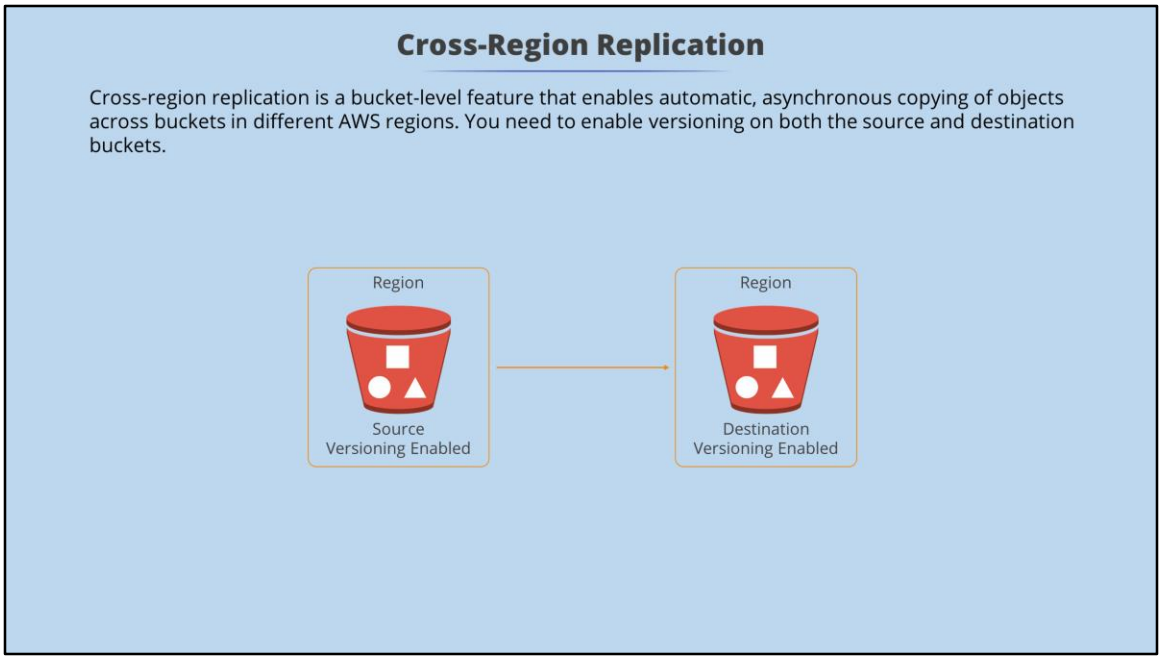
- Delete the current version of the object: Deleting the current object version will make the previous version the current version of that object.



Buckets can be in one of the three states: unversioned (the default), versioning-enabled, or versioning-suspended.

Once you version-enable a bucket, it can never return to an unversioned state. You can, however, suspend versioning on that bucket.

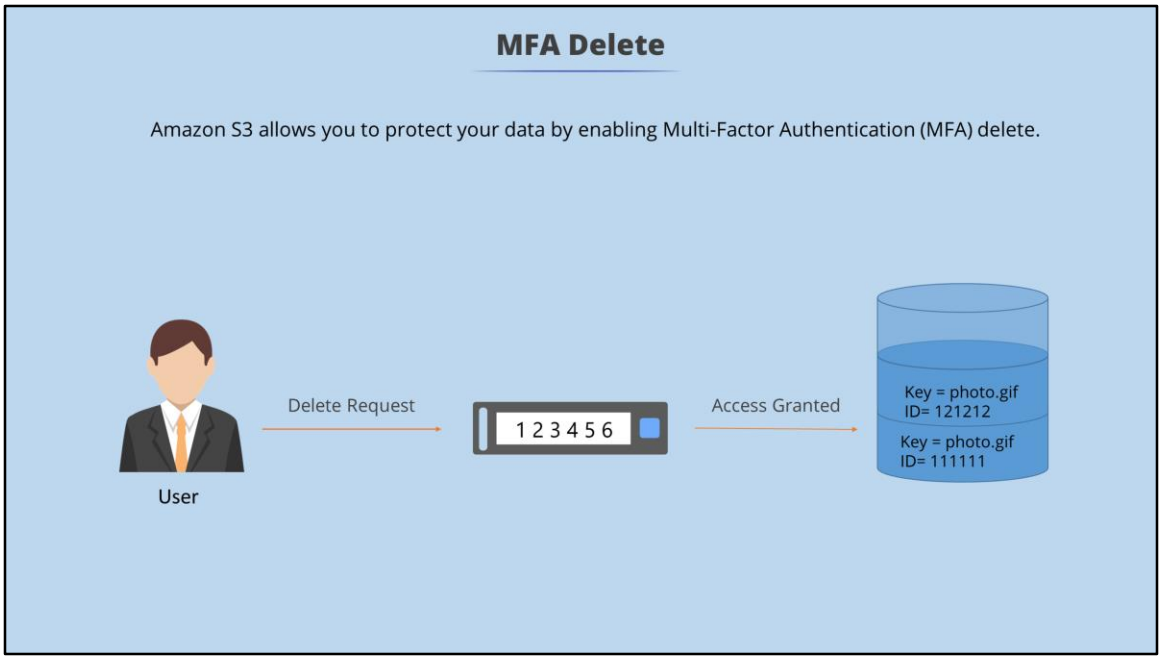
Suspended versioning means that for each file update, the ID is updated to NULL.



Cross-region replication is a bucket-level feature that enables automatic, asynchronous copying of objects across buckets in different AWS regions.

You can use this feature to provide low-latency versions of your data to users in other regions or to provide additional redundancy.

For cross-region replication to work, you need to have versioning enabled on both the source and destination buckets and you can only replicate to one destination bucket.

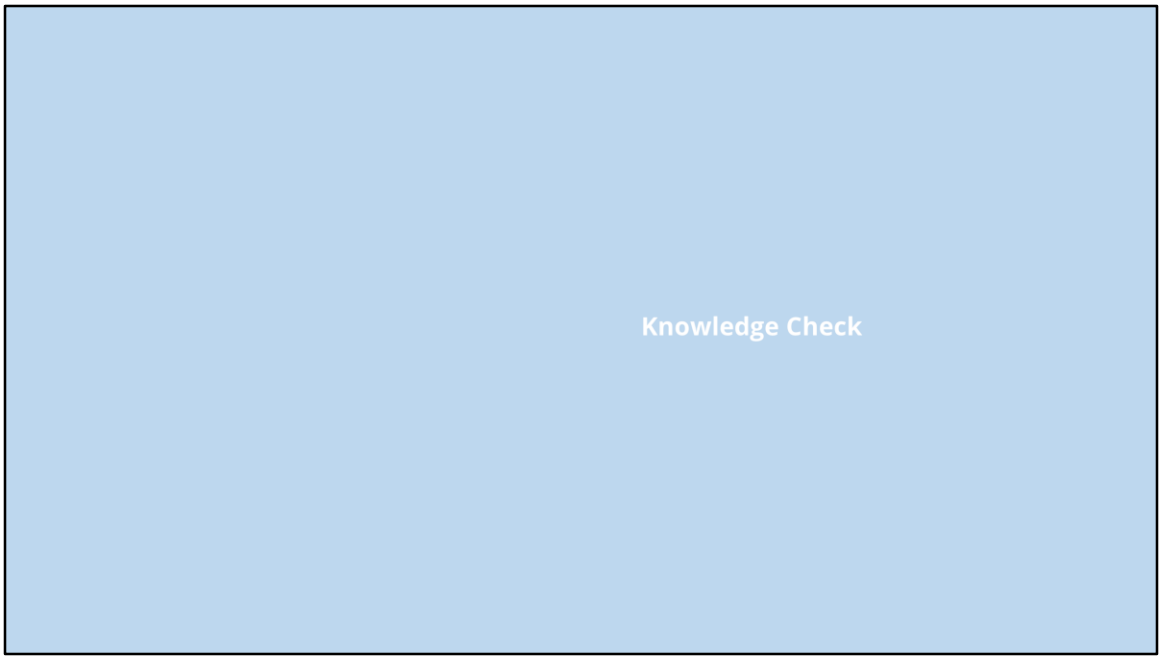


Amazon S3 allows you to protect your data by enabling MFA (Multi-Factor Authentication) delete. This provides additional authentication for the following operations:

Changing the versioning state of your bucket
Permanently deleting an object version

Demo: Amazon S3 Version Control

In this demonstration, you will learn about the version control functionality.


Knowledge Check

Knowledge Check 1

Which of the following is NOT an Amazon S3 version control state?

a. Unversioned

b. Versioning-enabled

c. Versioning-disabled

d. Versioning-suspended

simplilearn

Knowledge Check 1

Which of the following is NOT an Amazon S3 version control state?

a. Unversioned

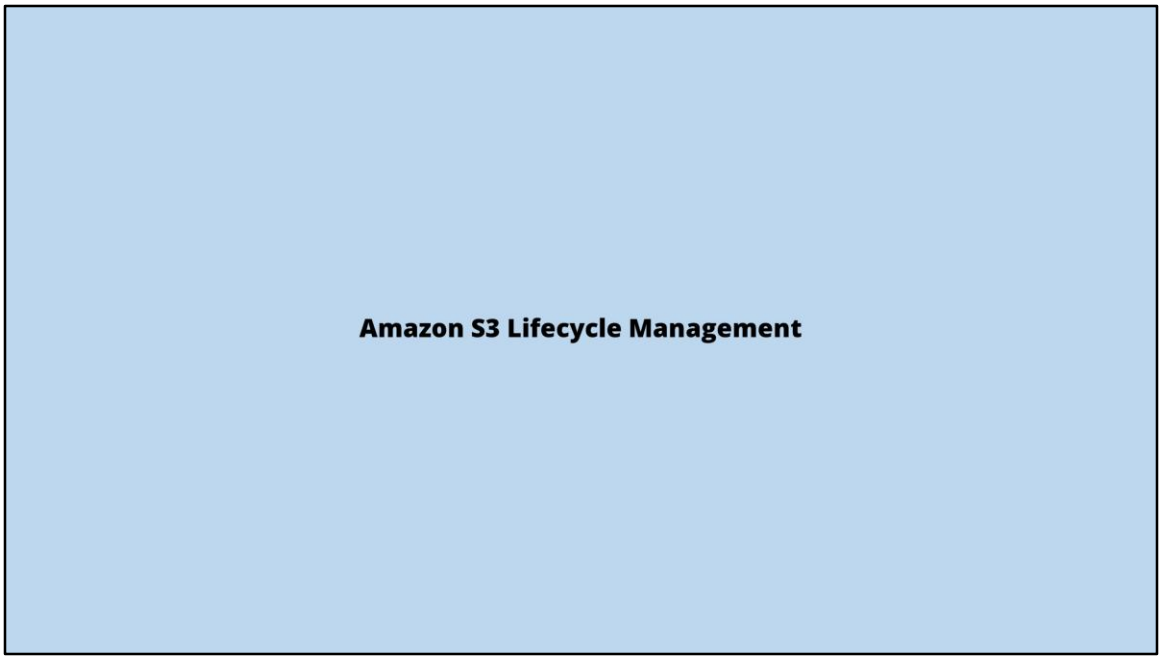
b. Versioning-enabled

c. Versioning-disabled

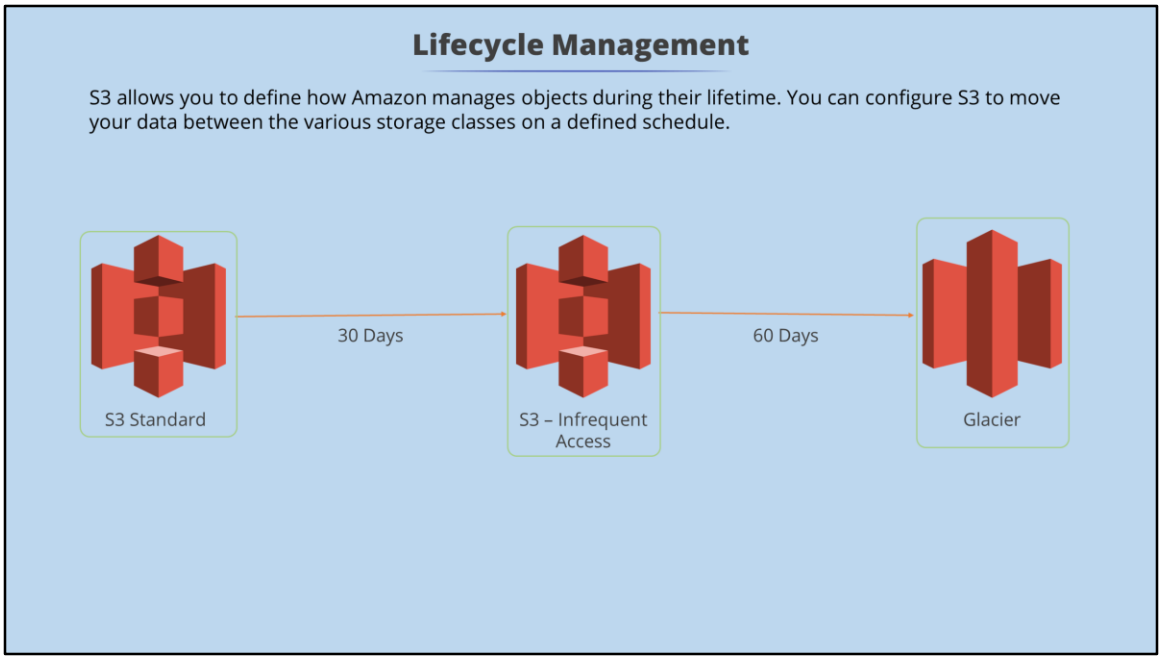
d. Versioning-suspended

The correct answer is **c**

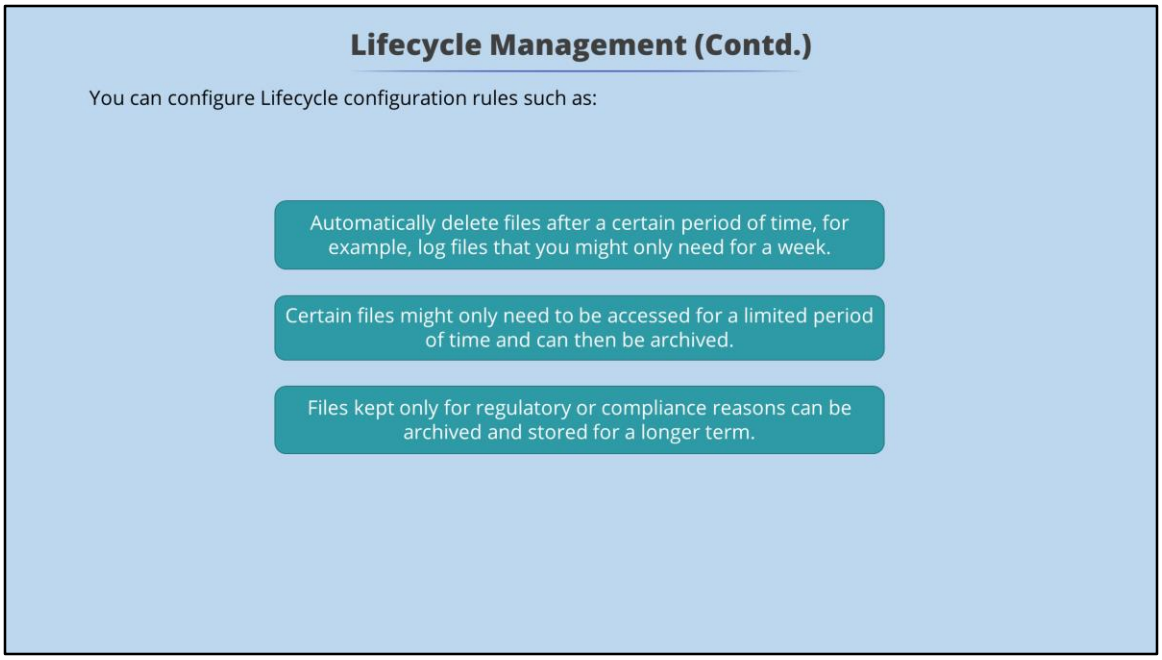
Versioning cannot be disabled; it can only be enabled or suspended.



In this section, you'll learn about Amazon S3 lifecycle management.



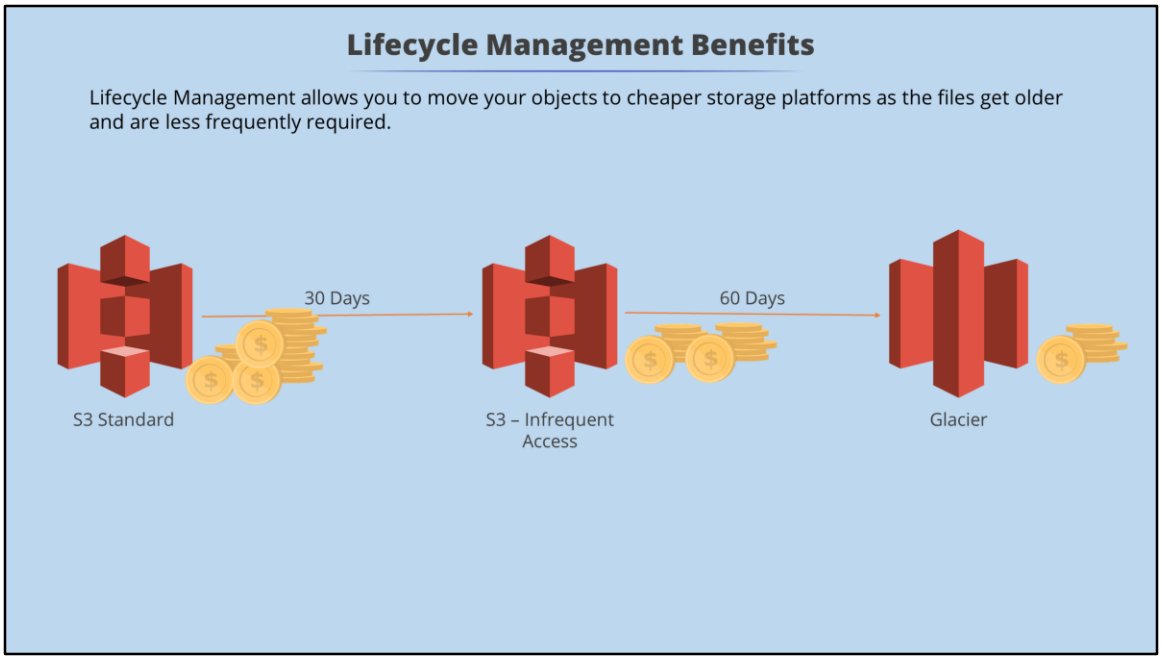
S3 allows you to define how Amazon manages objects during their lifetime.
You can configure S3 to move your data between the various storage classes on a defined schedule.



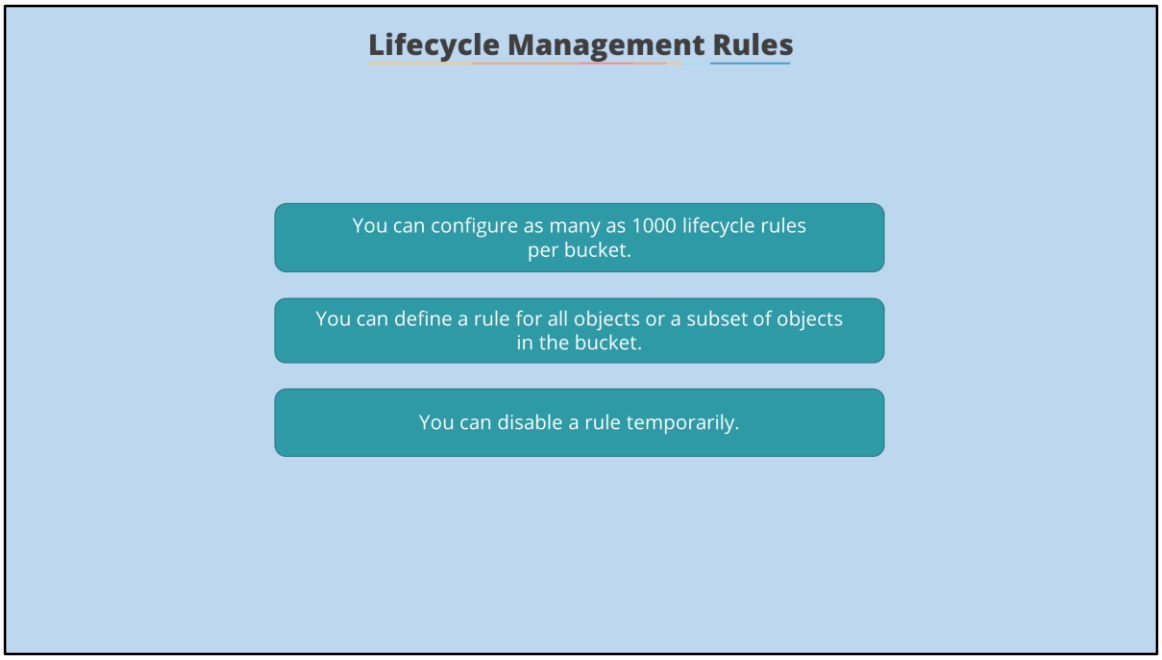
You can configure Lifecycle configuration rules such as the following:
Automatically delete files after a certain period of time, for example, log files that you might only need for a week.

Files required for a limited period of time to be archived at a pre-defined schedule.

Files kept for regulatory or compliance reasons should be archived and stored for longer terms.

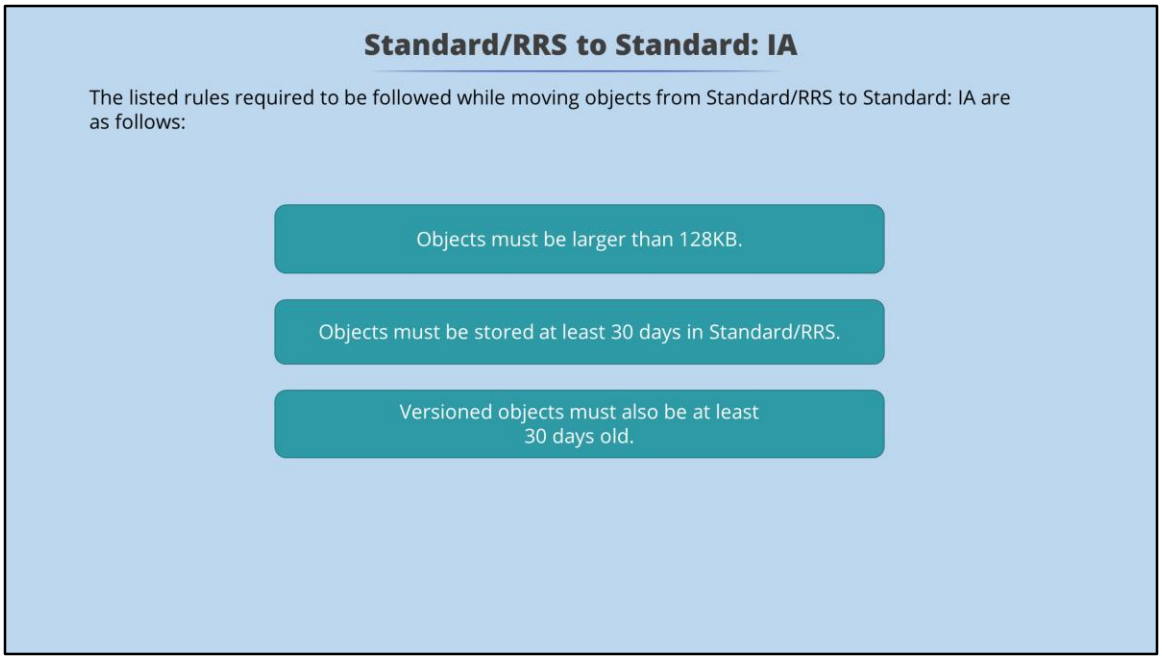


Lifecycle Management allows you to move your objects to cheaper storage platforms as the files get older and are less frequently required.



Following are the Lifecycle management rules:

- You can configure as many as 1000 lifecycle rules per bucket.
- You can define a rule for all objects or a subset of objects in the bucket.
- You can temporarily disable a rule if required.



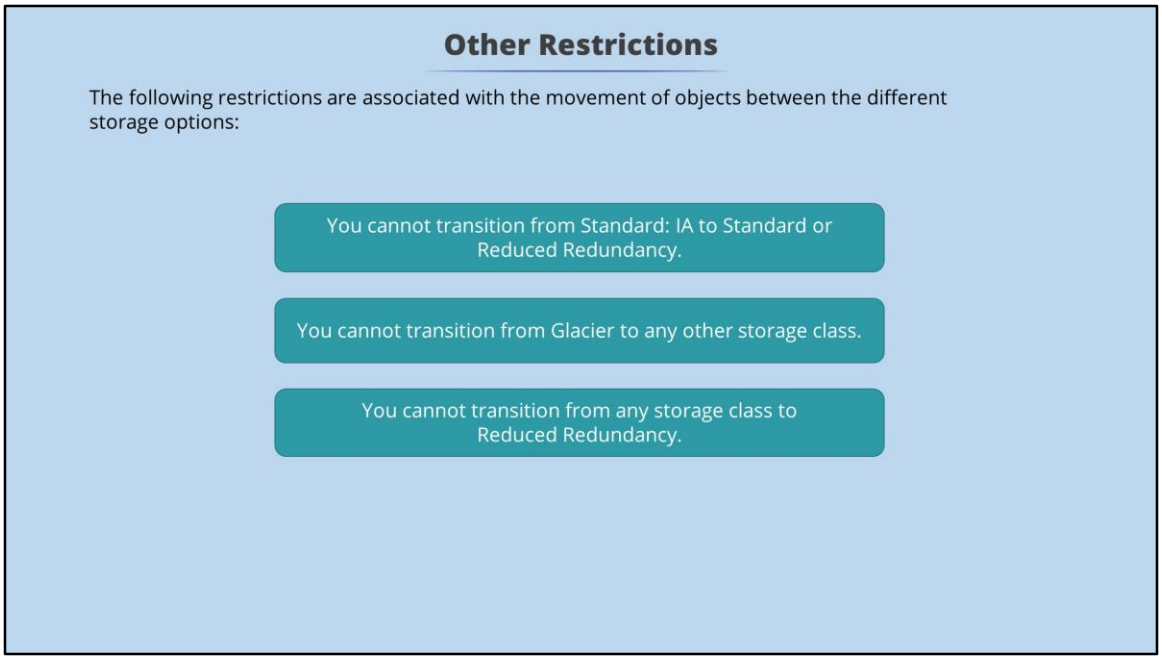
Amazon S3 does not move objects smaller than 128 kilobytes in size to the Standard—IA storage class because it is not cost effective.

Objects must be stored at least 30 days in the current storage class before you can transition them to Standard—IA.

If you are transitioning noncurrent objects, which means older object versions, you can only move objects that are >30 days old to Standard—IA.

You can transition from Standard/RRS/Standard—IA to Glacier at any time, however, the following restrictions exist with Glacier:

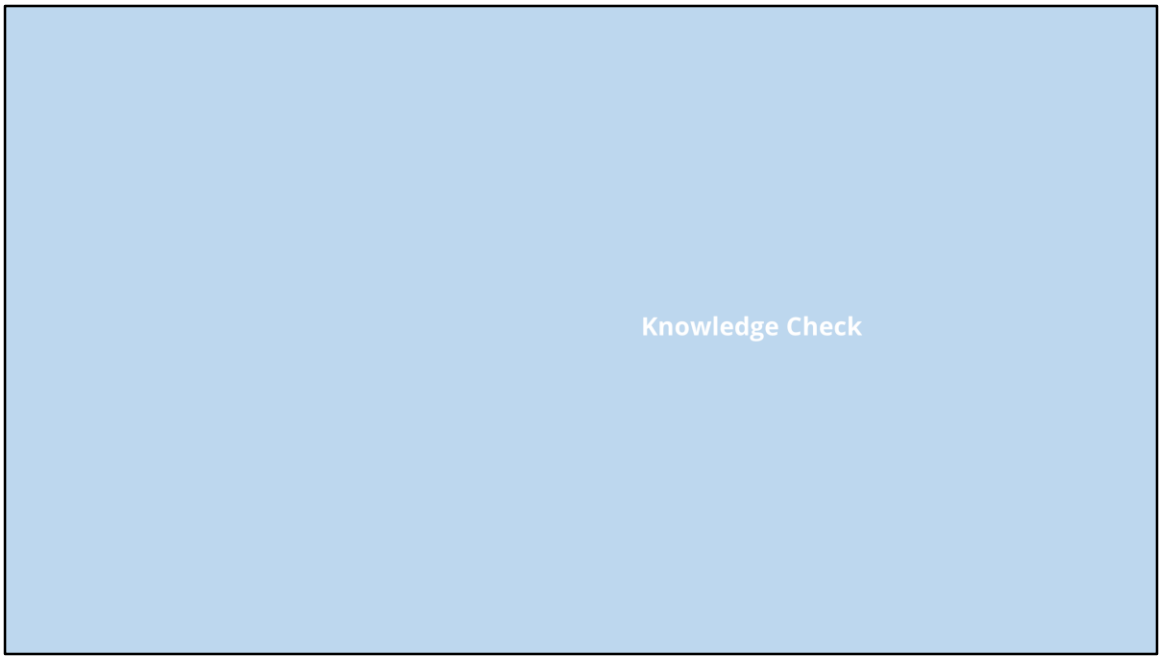
- Glacier stored objects are not available in real time.
- To access an archived object in Glacier, you first need to restore a temporary copy of it.
- The restored object is only available for the duration you specify during the "restore request."
- Glacier requests can take up to 5 hours.



You cannot transition from Standard—IA to Standard or Reduced Redundancy.
You cannot transition from Glacier to any other storage class.
You cannot transition from any storage class to Reduced Redundancy.

Demo: Amazon S3 Lifecycle Management

In this demonstration, you'll learn how to configure and use Lifecycle Management.


Knowledge Check

1

Which of the following Lifecycle Management rules is possible?

Glacier to Standard

Standard IA to Standard

Glacier to RRS

Standard to Standard IA

1

Which of the following Lifecycle Management rules is possible?

Glacier to Standard

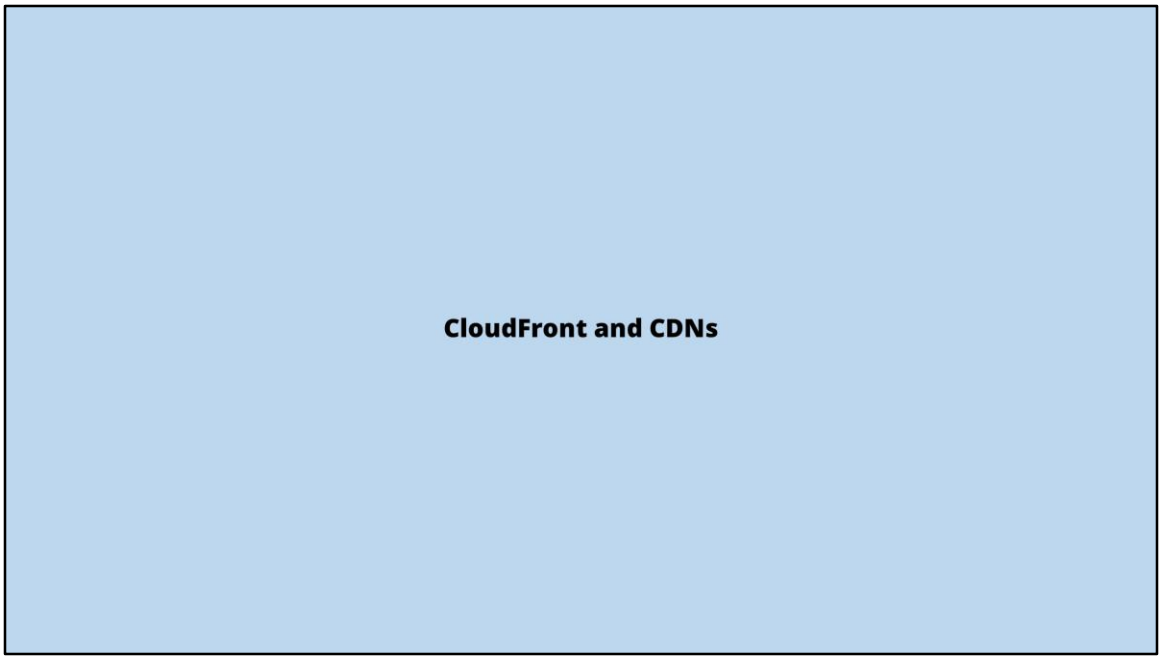
Standard IA to Standard

Glacier to RRS

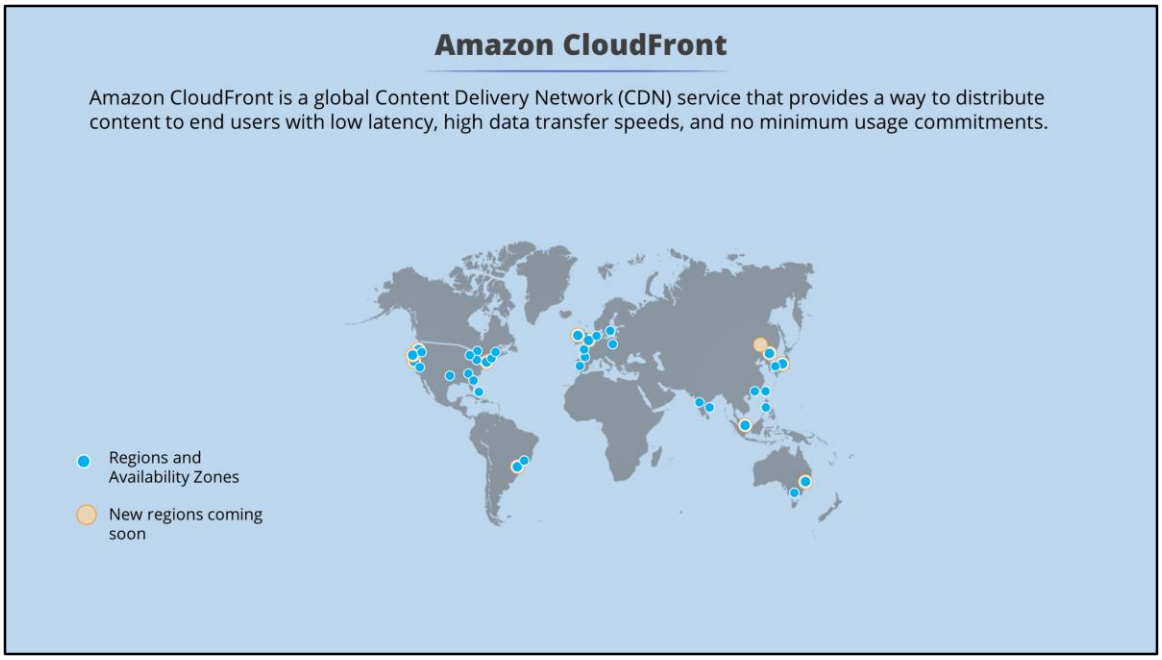
Standard to Standard IA

d

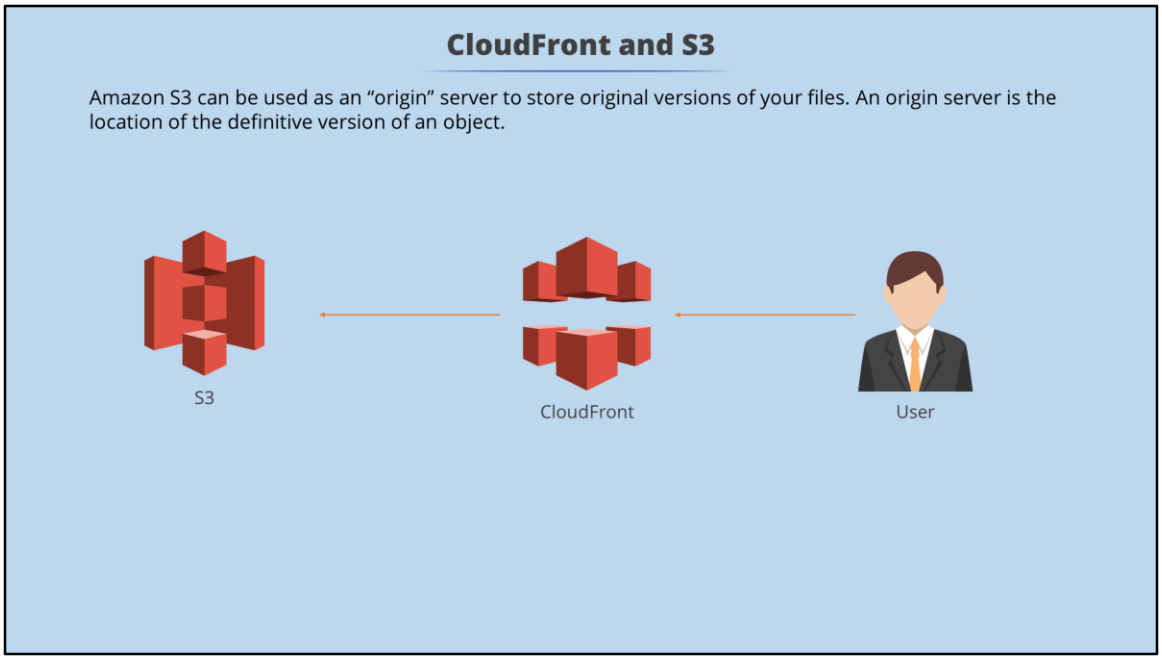
Standard to Standard IA is a supported Lifecycle Management rule, as is RRS to Standard IA, RRS to Glacier, and Standard IA to Glacier.



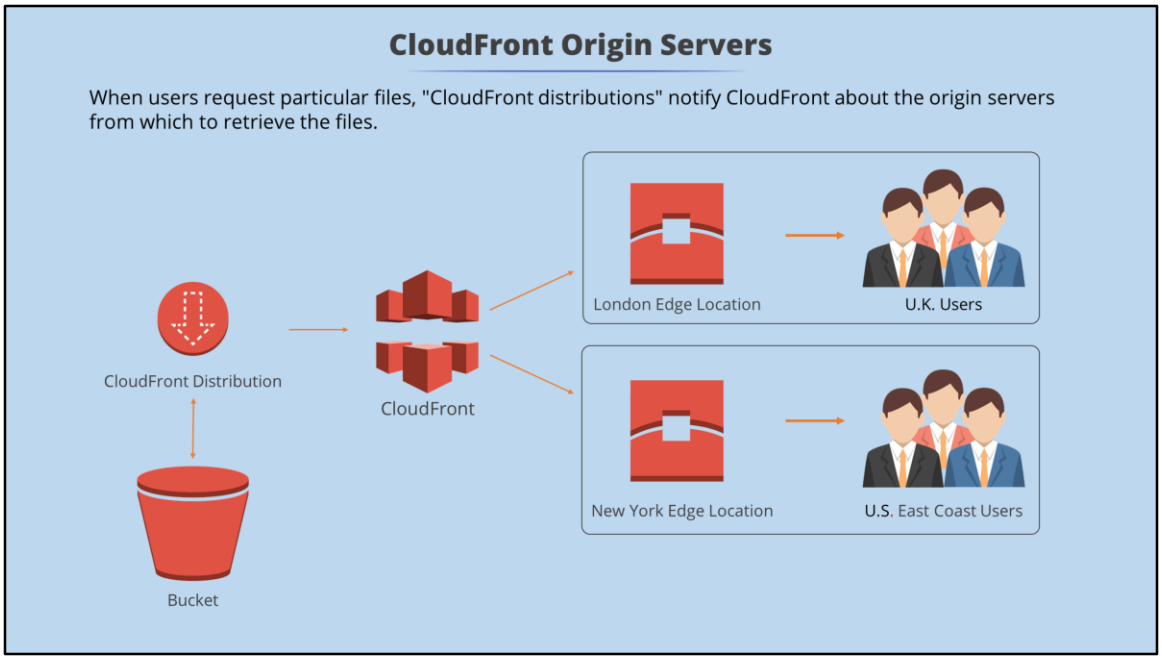
In this section, you'll learn about CloudFront and Content Delivery Networks (CDN).



Amazon CloudFront is a global Content Delivery Network (CDN) service which provides a way to distribute content to end users with low latency, high data-transfer speeds, and no minimum usage commitments.



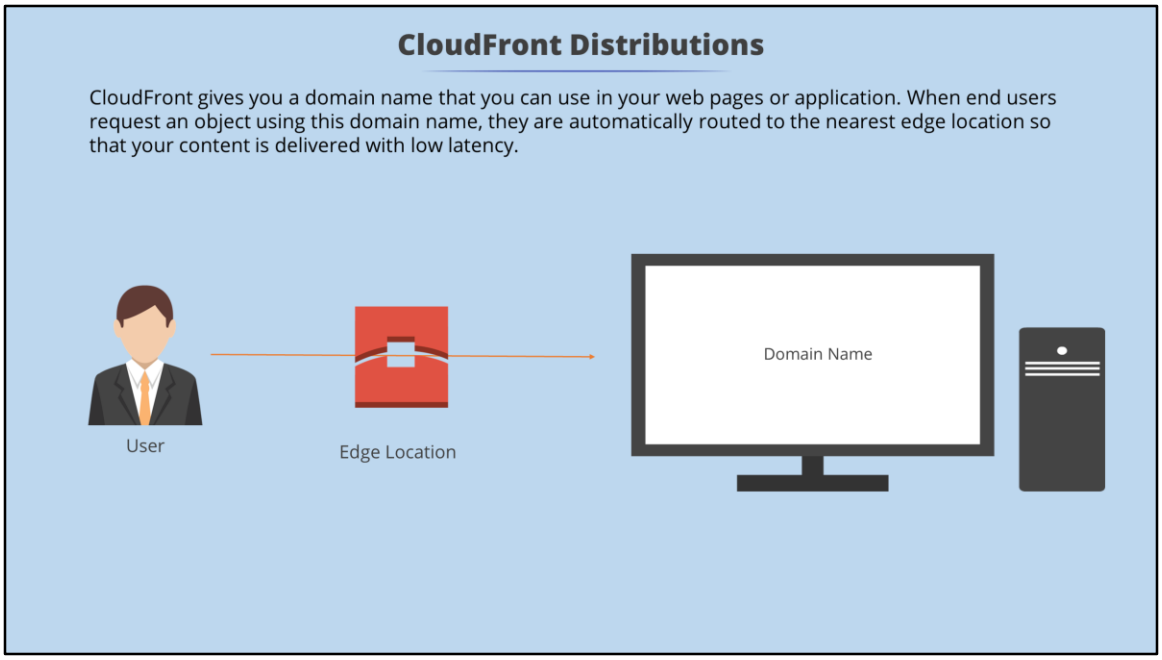
Amazon S3 can be used as an "origin" server to store original versions of your files. An origin server is the location of the definitive version of an object.



"Cloudfront distributions" notify Cloudfront about the origin servers from which to get the files from when users request them from a website or application.

You can also specify if you want CloudFront to log requests and when you want the distribution to be enabled.

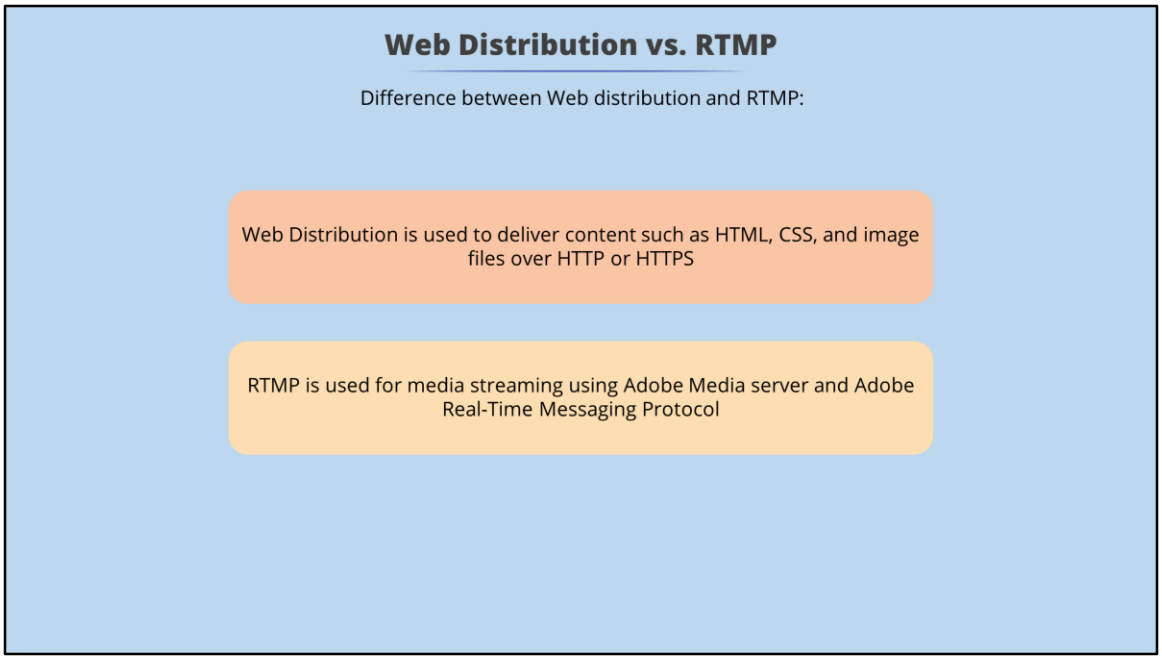
In the diagram, CloudFront specifies the bucket from which to retrieve the data. When users request access to your website, CloudFront pushes the access request directly to the edge locations closest to the end users.



Cloudfront gives you a domain name which you can use in your web pages or application.

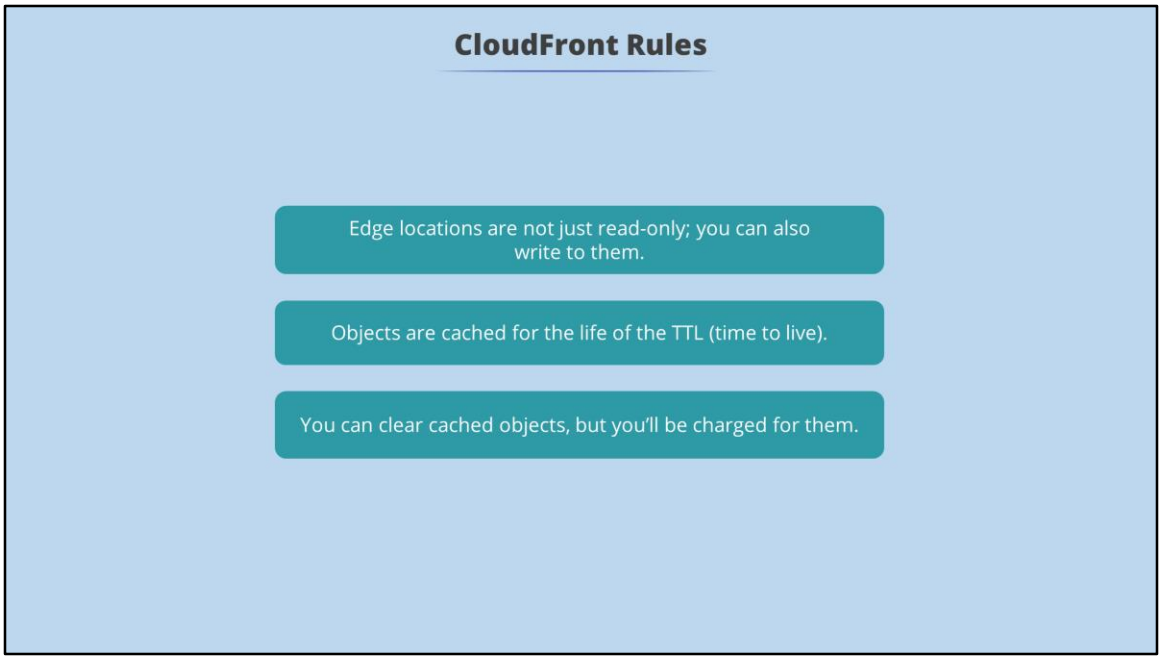
When end users request an object using this domain name, they are automatically routed to the nearest edge location so that your content is delivered with low latency.

You have to pay for the data transfer and requests that you actually use.



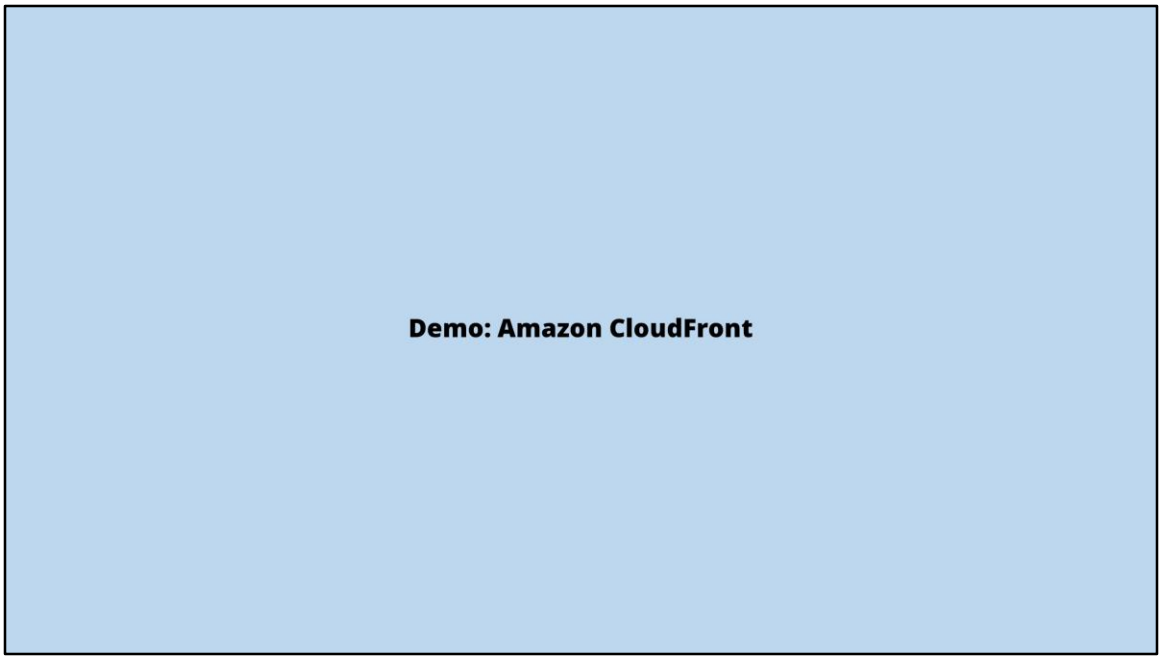
Web Distribution is used to deliver content such as html, CSS, image files, and so on over HTTP or HTTPS.

RTMP is used for media streaming through Adobe Media server and Adobe Real-Time Messaging Protocol.

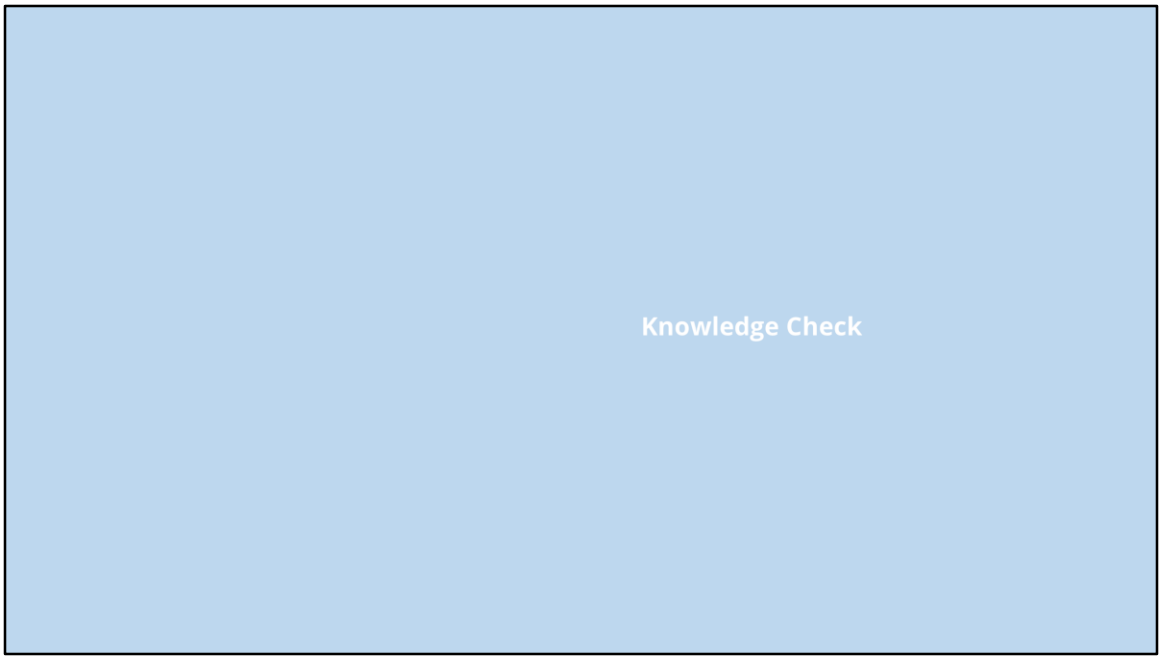


Following are the CloudFront rules:

- Edge locations are not just read-only; you can also write to them.
- Objects are cached for the life of the TTL (time to live).
- By default, each object automatically expires after 24 hours.
- You can clear cached objects if you update your origin files, but it is chargeable.



In this demonstration, you'll learn how to use Amazon CloudFront and Amazon S3


Knowledge Check

1

What is the purpose of Amazon CloudFront?

To provide disaster recovery for objects stored in your S3 buckets

To distribute content to end users with low latency

To reduce the costs of running a web site or media streaming service

To host your website source code

1

What is the purpose of Amazon CloudFront?

To provide disaster recovery for objects stored in your S3 buckets

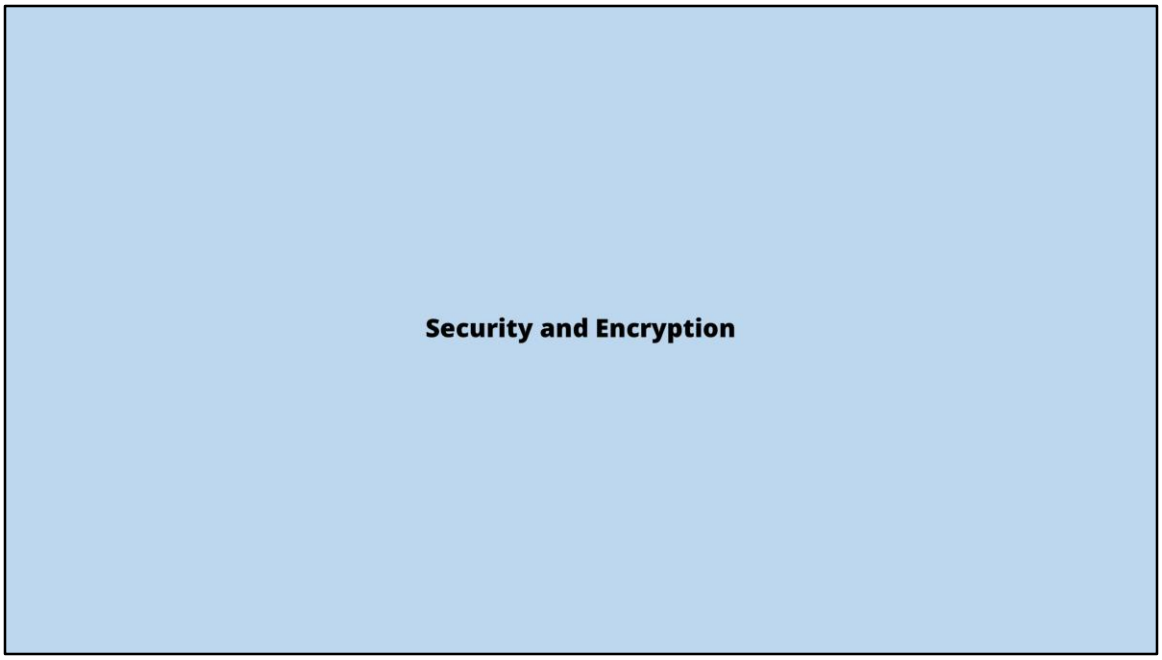
To distribute content to end users with low latency

To reduce the costs of running a web site or media streaming service

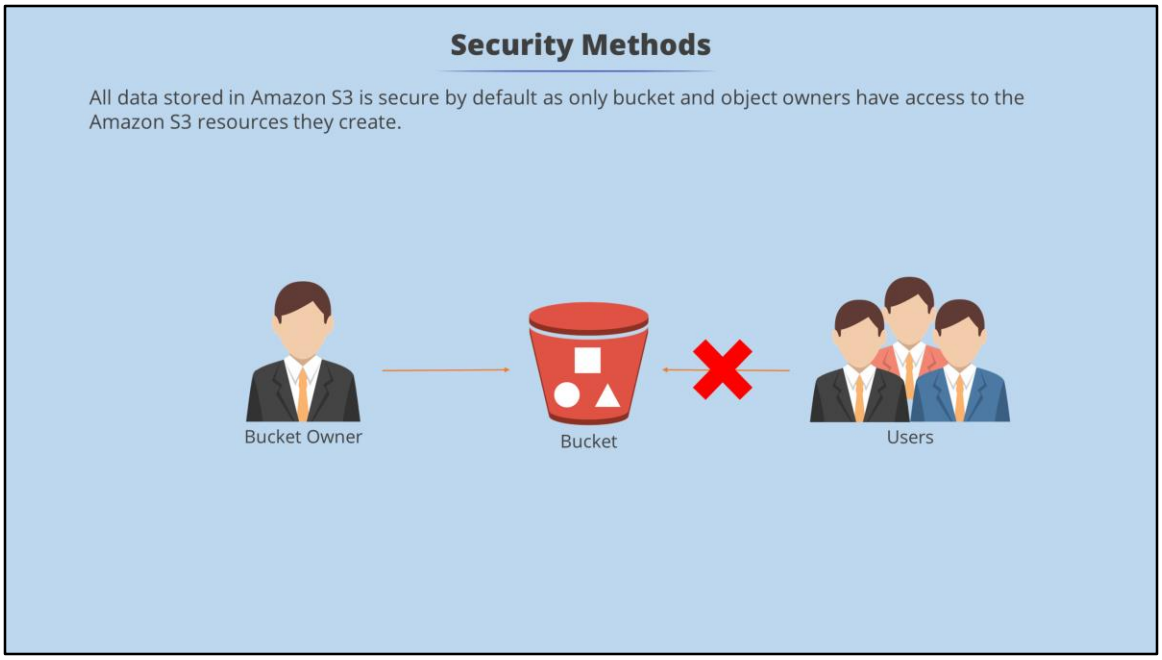
To host your website source code

b

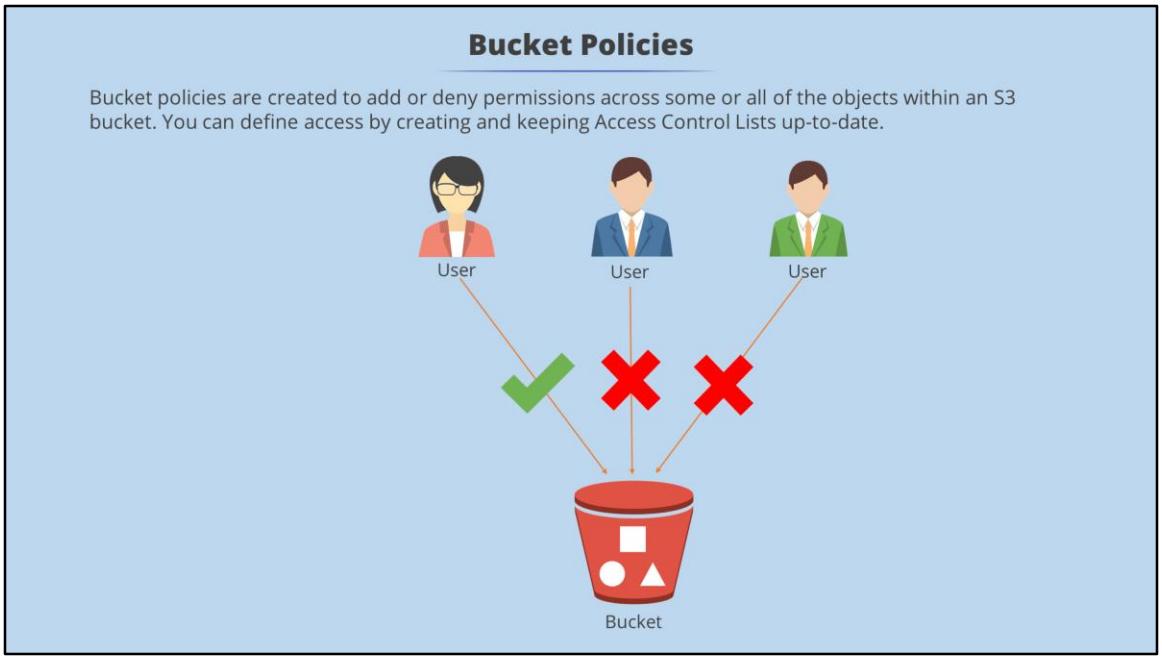
Amazon CloudFront is a way to distribute content to end users with low latency and high data transfer speeds.



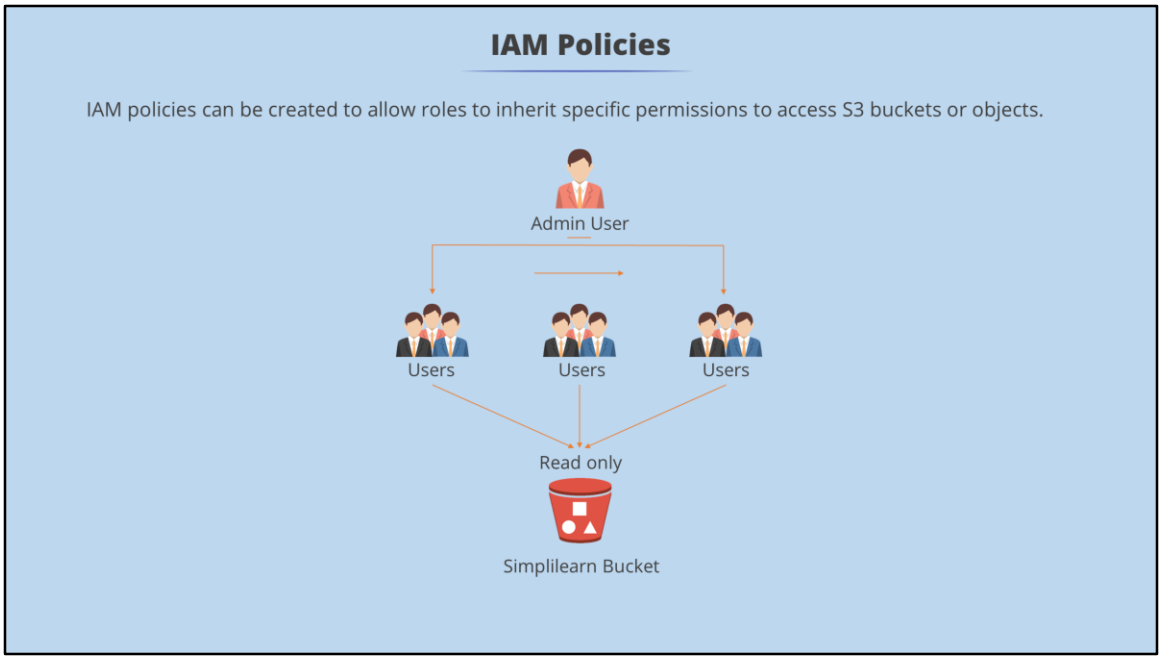
In this section, you'll learn about the security and encryption associated with S3.



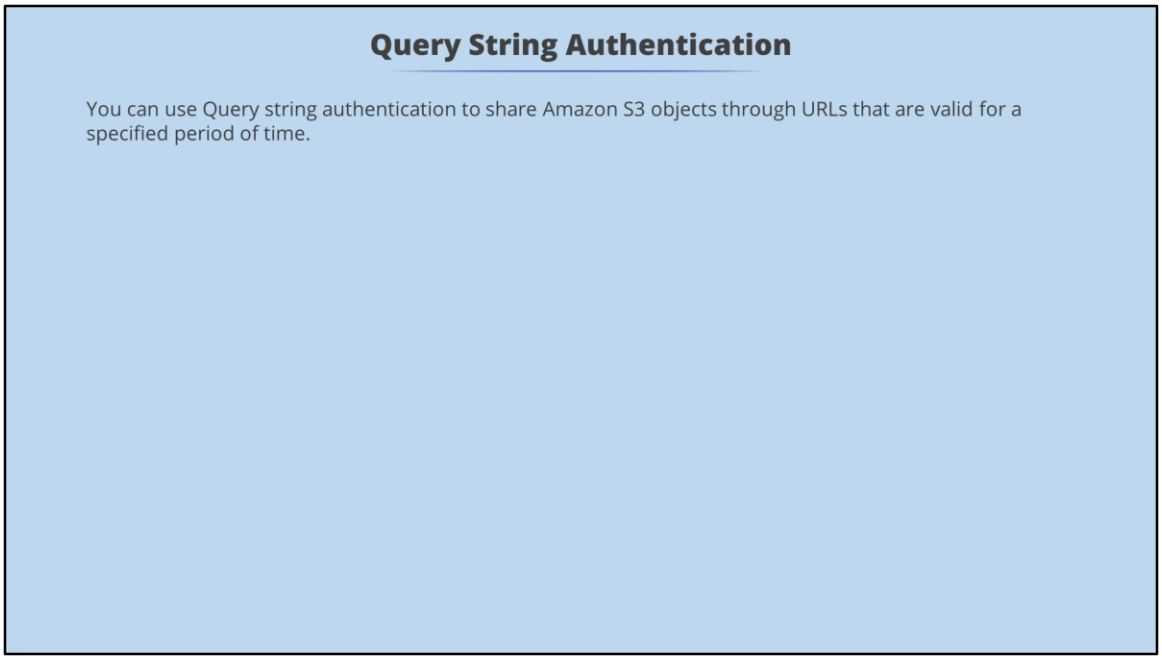
All data stored in Amazon S3 is secure by default as only the bucket and object owners have access to the Amazon S3 resources they create.



Bucket policies are created to add or deny permissions across some or all of the objects within an S3 bucket. You can define access by creating and keeping Access Control Lists up-to-date.

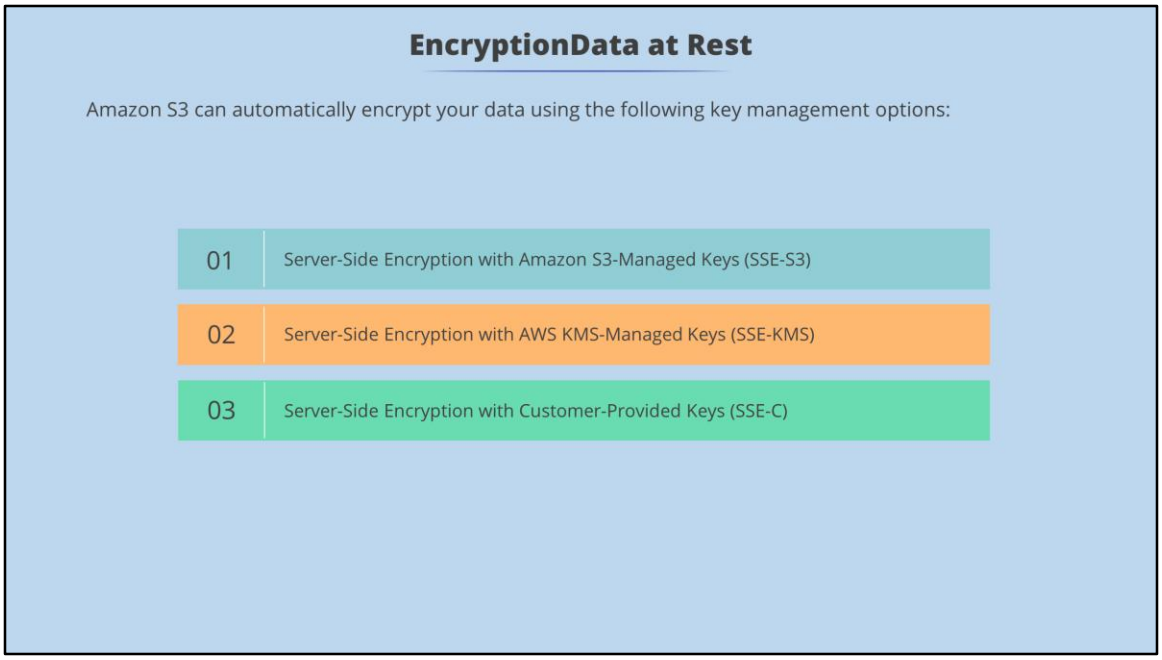


IAM policies can be created to allow specific access to S3 buckets. In the example, you can specify which users can access the Simplilearn S3 bucket.



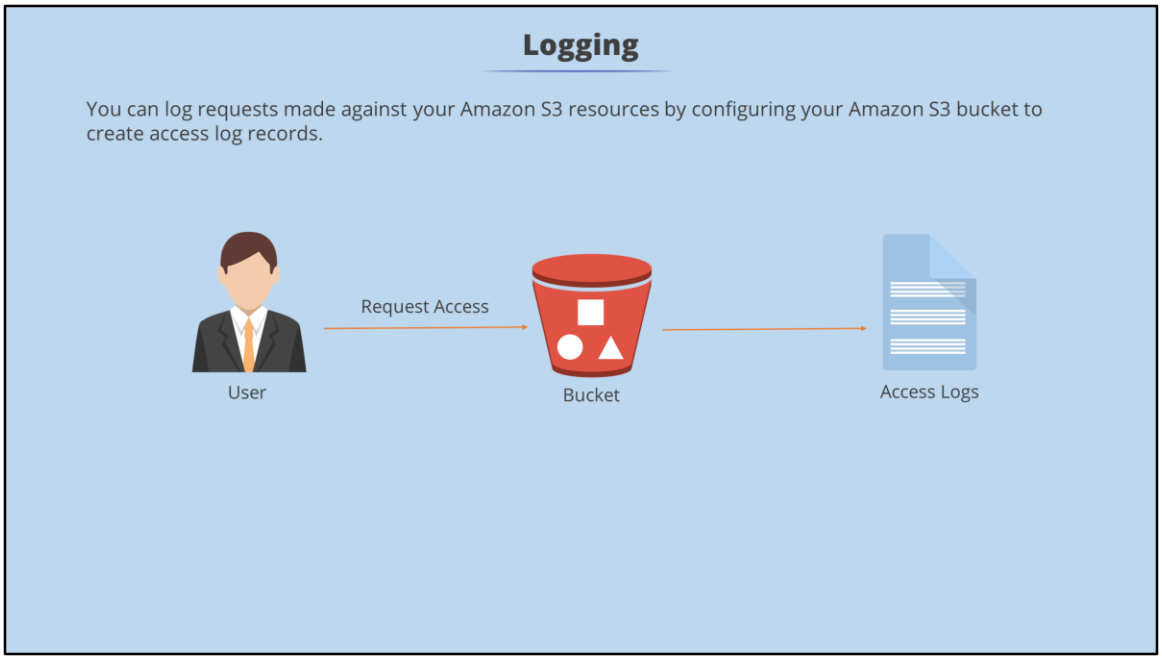
Using Query string authentication, you can share Amazon S3 objects through URLs that are valid for a specified period of time.

Using Amazon S3 SSL-encrypted endpoints that use the HTTPS protocol, you can securely upload or download your data.



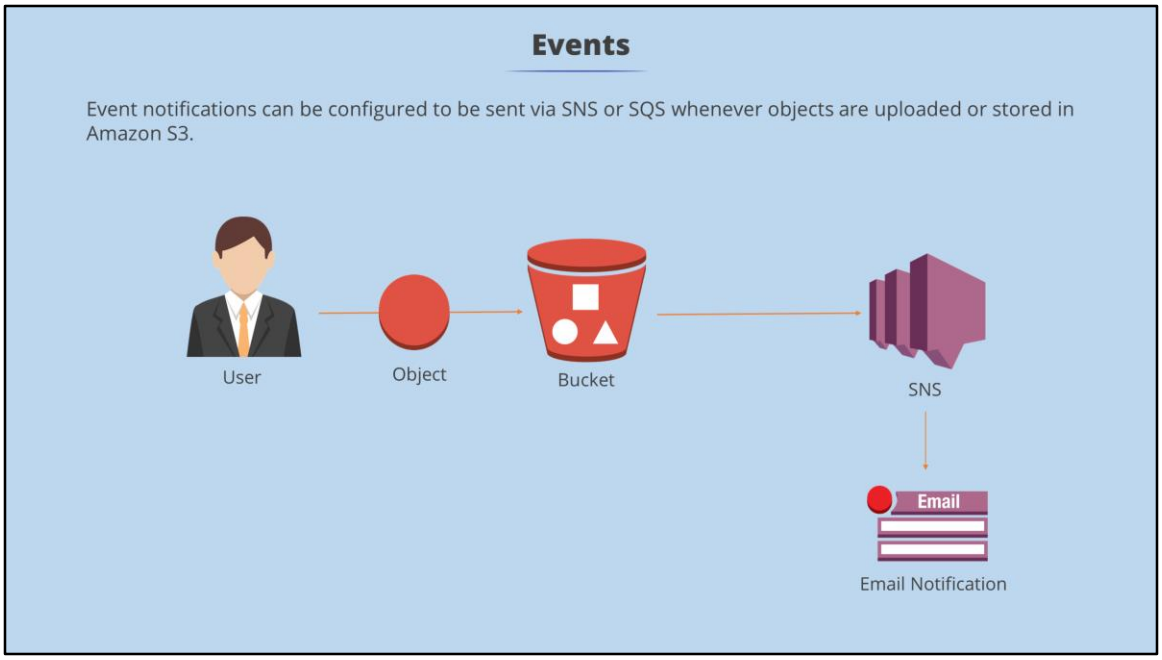
Amazon S3 can automatically encrypt your data at rest using the following key management options:

- Server-Side Encryption with Amazon S3-Managed Keys (SSE-S3):
- Each object is encrypted with a unique key employing strong multi-factor encryption. As an additional safeguard, it encrypts the key itself with a master key that it regularly rotates.
- Server-Side Encryption with AWS KMS-Managed Keys (SSE-KMS):
- They are the same as SSE-S3 but with extra features (and charges). There are separate permissions for the use of an envelope key (a key that protects your data's encryption key) that provides added protection against unauthorized access of your objects in S3. It provides an audit trail of who used your key and when.
- Server-Side Encryption with Customer-Provided Keys (SSE-C):
- In this option you manage the encryption keys and Amazon manages the encryption.



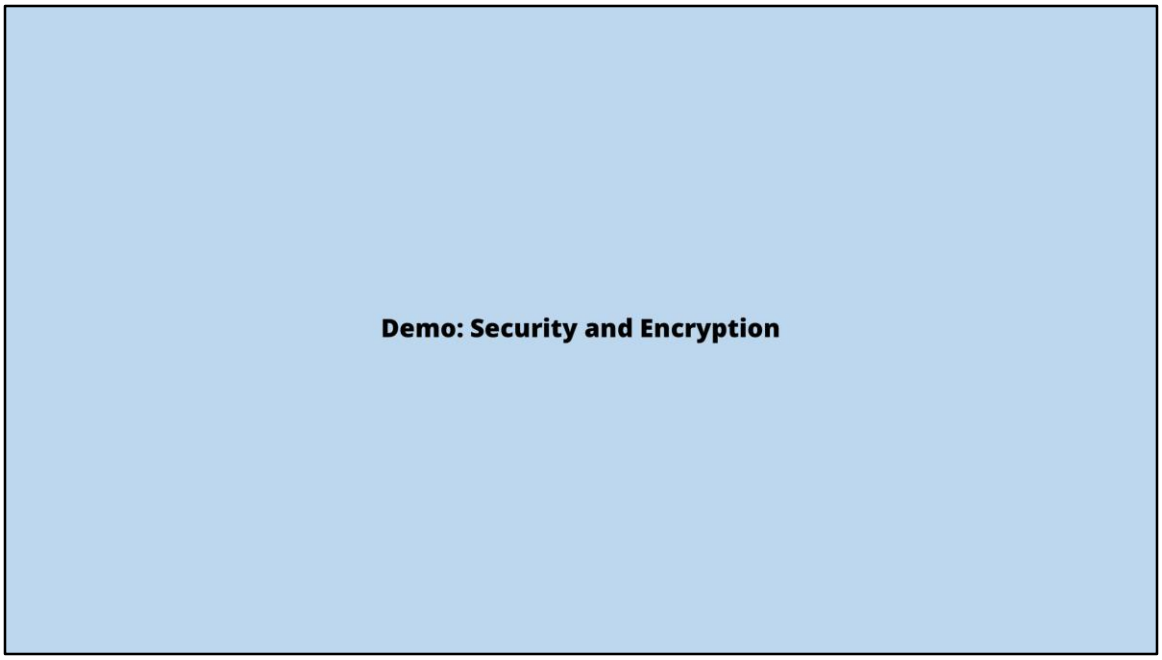
You can log requests made against your Amazon S3 resources by configuring your Amazon S3 bucket to create access log records.

The access logs capture all requests made against a bucket, or the objects in it, so you can have all the information required for auditors.

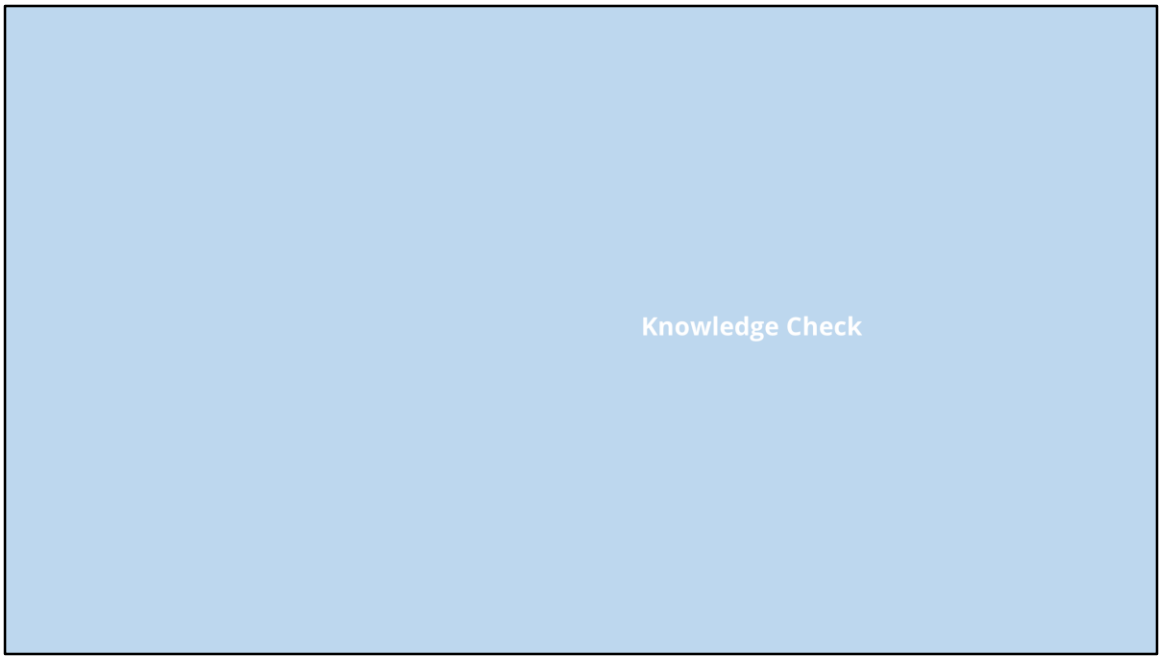


Event notifications can be configured to be sent via SNS or SQS whenever objects are uploaded or stored in Amazon S3.

You can create workflows by delivering event notifications to other AWS services such as Lambda to perform actions based on changes to your S3 objects.



In this demonstration, you'll learn about the Security and Encryption features of Amazon S3.


Knowledge Check

1 **Which of the following is NOT a method of securing access to Amazon S3 buckets?**

Query String Authentication

IAM policies

Bucket Policies

Encryption

1

Which of the following is NOT a method of securing access to Amazon S3 buckets?

Query String Authentication

IAM policies

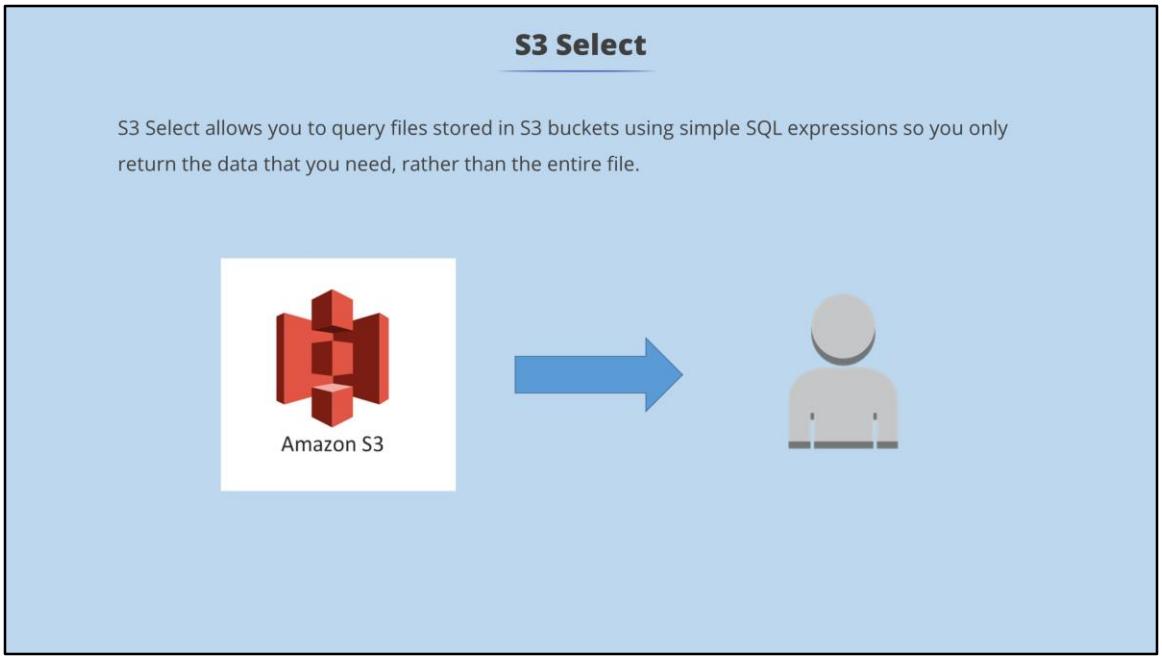
Bucket Policies

Encryption

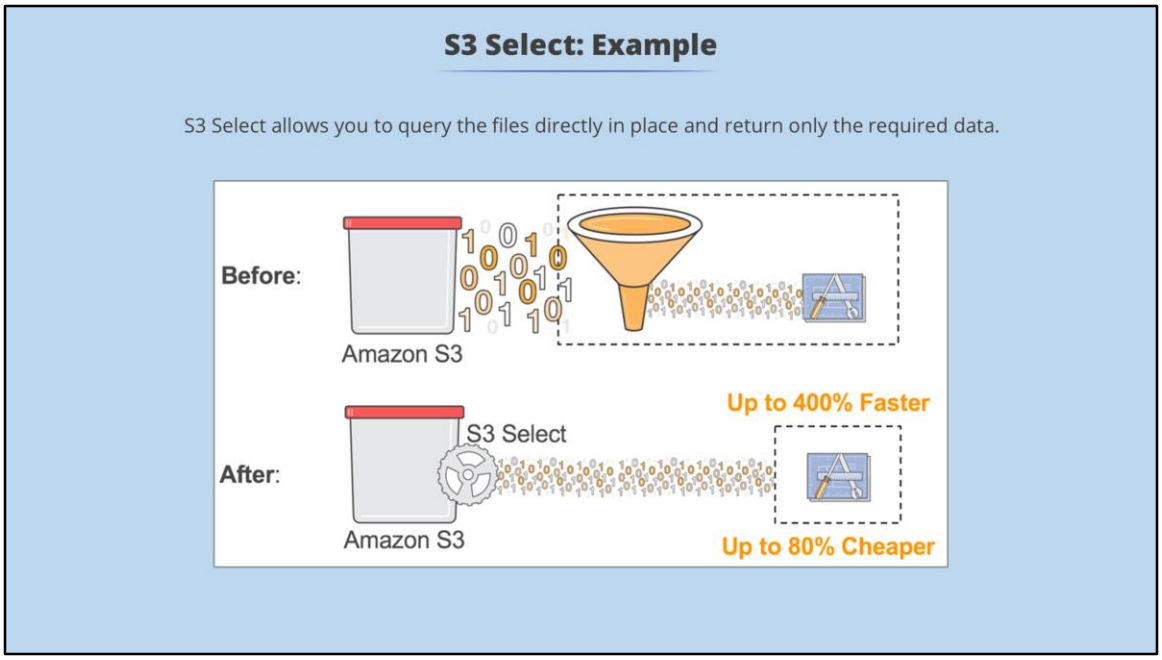
d

Encryption protects the objects stored in Amazon S3, but it doesn't provide secured access to the objects.

Demo: Amazon S3/Glacier Select



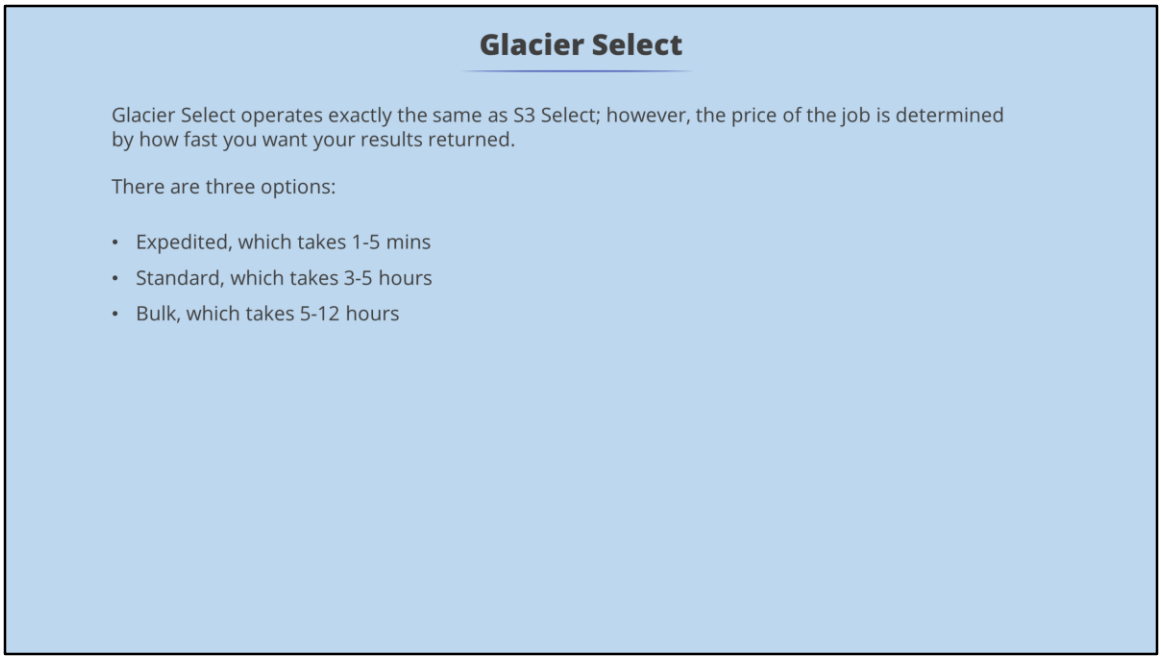
S3 Select allows you to query files stored in S3 buckets using simple SQL expressions, so you only return the data that you need, rather than the entire file.



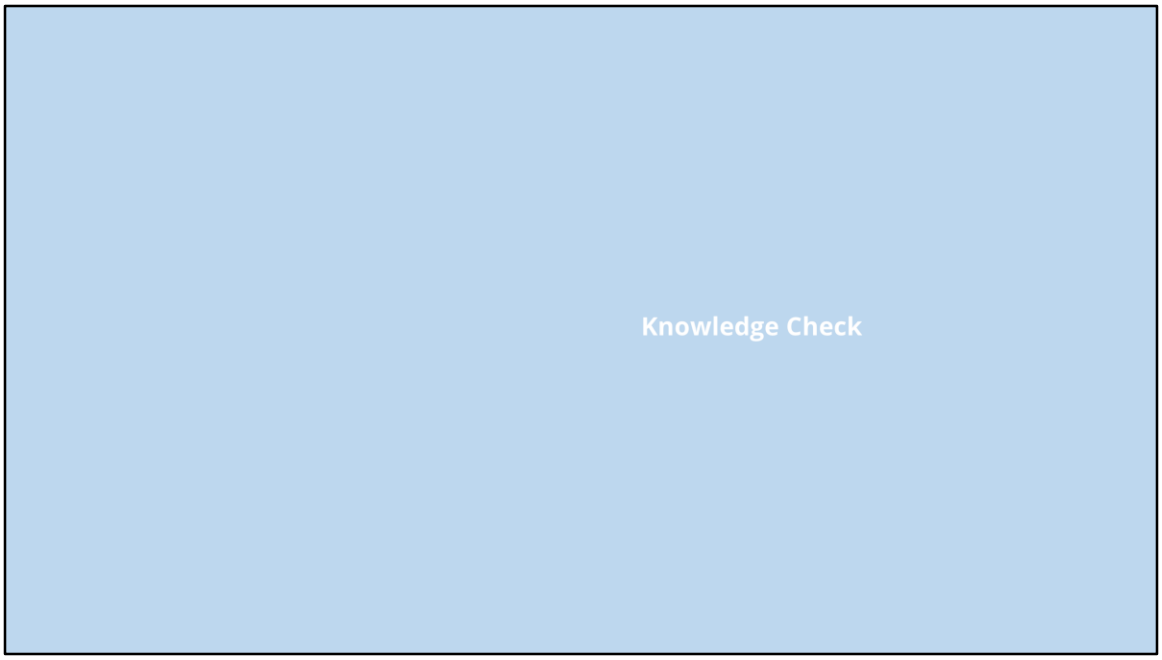
Before S3 Select you would have to download all the files that may contain useful data, extract them (if they are compressed) and then find the data you require.

With S3 Select you would just query the files directly in place and return only the data you need.

This can reduce costs by up to 80% and can be up to 400% faster.



Glacier Select operates exactly the same, however the price of the job is determined by how fast you want your results:
Expedited 1-5 mins
Standard 3-5 hours
Bulk 5-12 hours


Knowledge Check

What is the key benefit of S3 select?

1

Faster download of files stored on S3

Greater security of files stored on S3

Faster return of data from files stored on S3

Cheaper storage costs of S3 files

1

What is the key benefit of S3 select?

Faster download of files stored on S3

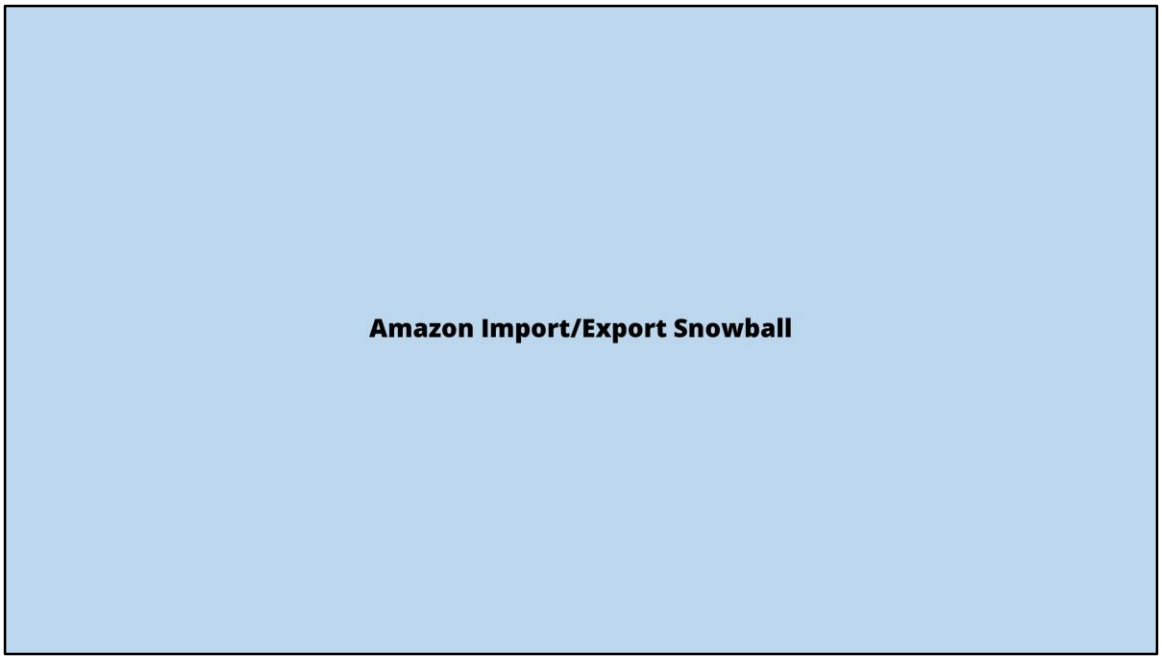
Greater security of files stored on S3

Faster return of data from files stored on S3

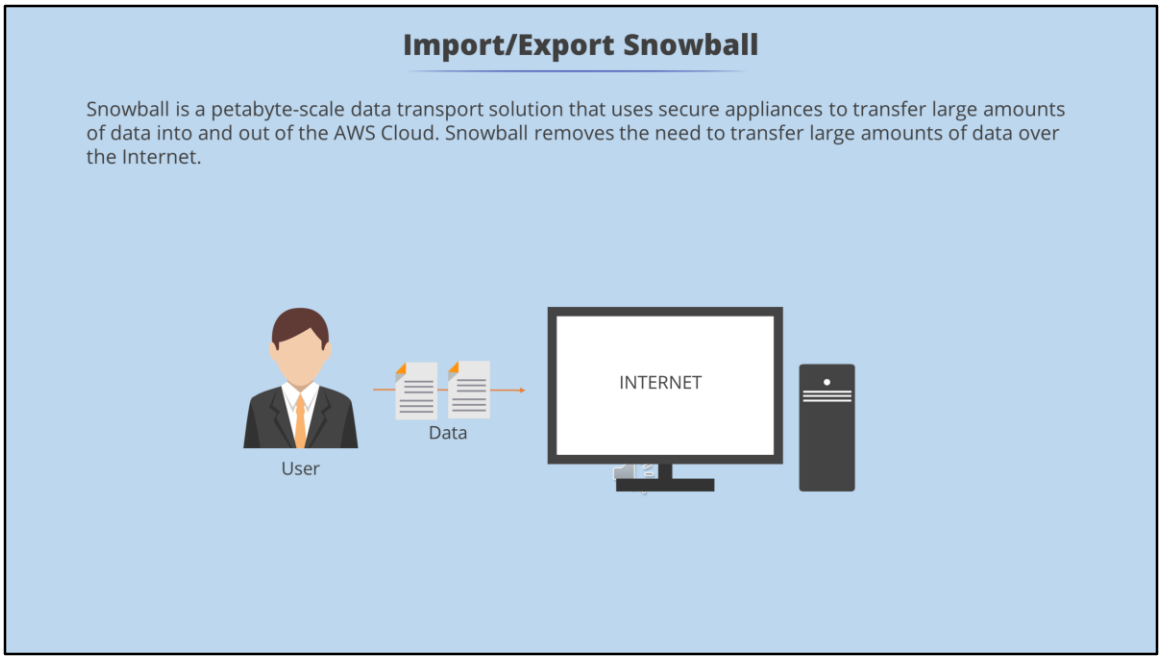
Cheaper storage costs of S3 files

c

S3 Select allows you to return data from files stored on S3 much faster than previously.



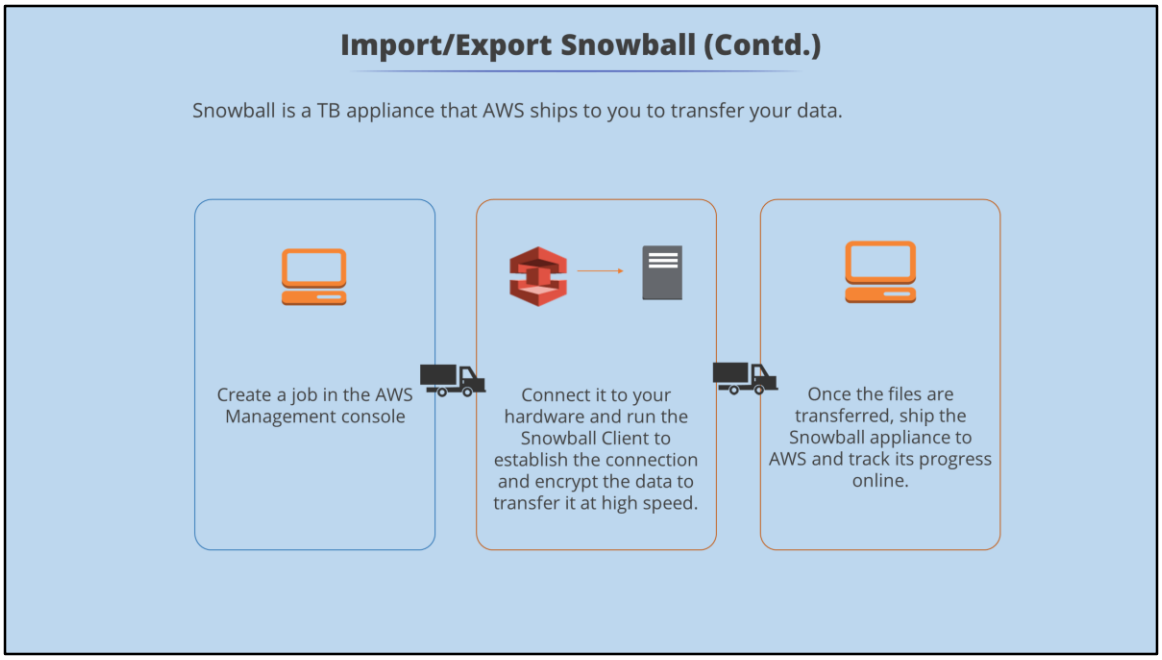
In this section, you'll learn about the features and benefits of Amazon Import/Export Snowball.



Snowball is a petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data into and out of the AWS Cloud.

Snowball removes the need to transfer large amounts of data over the Internet. Traditional transfer methods take huge amounts of time, incur high network costs, and can have security issues.

Snowball is simple, fast, and secure and can be as little as 20 percent of the cost of Internet.



Snowball is a TB appliance that AWS ship to you. You attach to your hardware and run the snowball client to establish a connection and encrypt and transfer the data onto it at high speed.

As of January 2017 it's either 50 or 80TB, depending on the region.

Once the transfer is complete the E Ink shipping label will automatically update and you simply send it back to AWS. You can track it's progress online.

With Snowball, you don't need to write any code or purchase any hardware to transfer your data. You need to create a job in the AWS Management Console and a Snowball appliance will be shipped to you. Once it arrives, attach the appliance to your local network, download and run the Snowball client to establish a connection, and then use the client to select the file directories that you want to transfer to the appliance. The client will then encrypt and transfer the files to the appliance at high speed. When the transfer is complete and the appliance is ready to be returned, the E Ink shipping label will automatically update and you can track the job status viaAmazon Simple Notification Service (SNS), text messages, or directly in the Console.

AWS Snowball Edge

01	It is an updated version of Snowball with on-board storage and compute power for select AWS capabilities.
02	It is a 100 TB device.
03	In addition to transferring day-to-day data, it can also undertake local processing and Edge computing workloads.

AWS Snowball Edge

Snowball Edge is an update version of the Snowball that has on-board storage and compute power for select AWS capabilities

As of Jan 2017 it's a 100TB device.

In addition to transferring data to AWS, Snowball Edge can undertake local processing and edge-computing workloads.

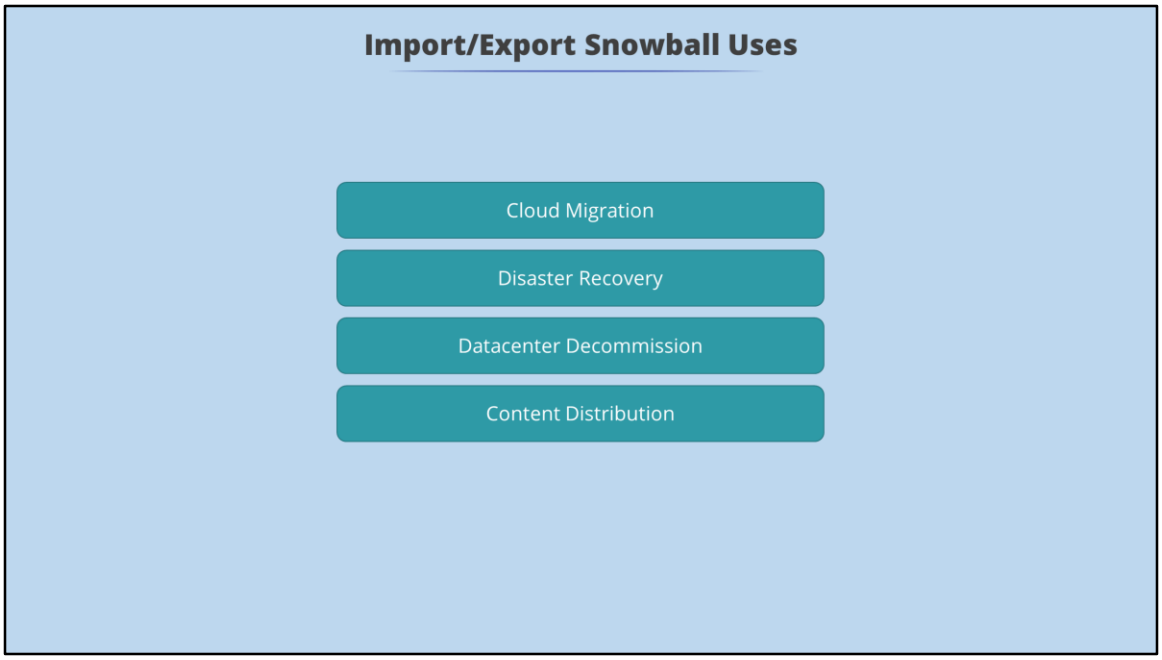
AWS Snowmobile

01	It is an Exabyte scale data transfer service used to move extremely large amounts of data to AWS.
02	It can be used to transfer up to 100PB of data.
03	It is semi-trailer truck that can used to move entire data centers to AWS.

AWS Snowmobile

AWS Snowmobile is an Exabyte-scale data transfer service used to move extremely large amounts of data to AWS.

You can transfer up to 100PB per Snowmobile, a 45-foot long ruggedized shipping container, pulled by a semi-trailer truck.



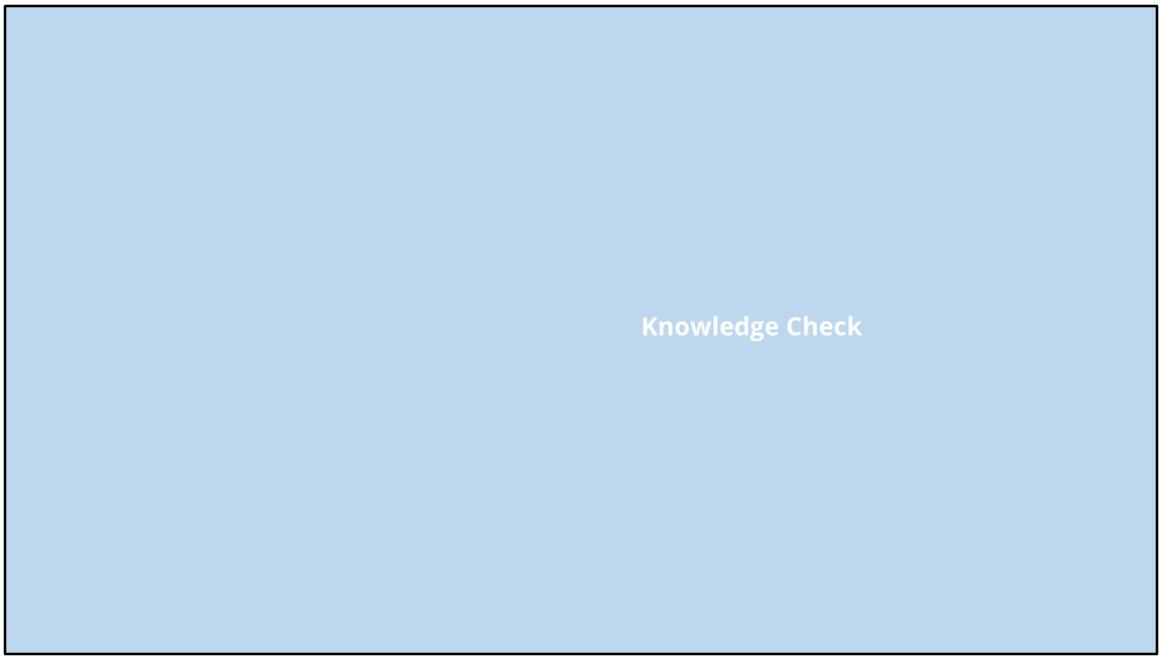
Functions of Snowball:

- Cloud Migration: If you have huge amounts of data that you need to move to AWS, then Snowball is a fast and cost effective way of achieving it.
- Disaster Recovery: If you need to quickly get hold of a large amount of data stored in S3, then you can use Snowball to transfer the data back to your site.
- Datacenter Decommission: Securely store data while you decommission your data center; the data can be moved to S3 temporarily, or permanently.
- Content Distribution: If you regularly have to send large amounts of data to clients or customers, then Snowball can be sent directly from AWS to them.

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

Transfer acceleration is achieved by copying data into S3 via the closest edge locations rather than directly over the Internet to the specified AWS location.

Transfer acceleration is higher when the source is farther from the destination, when there is more available bandwidth, and/or when the object size is bigger. This way the data travels over an optimized path. You are only charged if transfer acceleration is achieved.


Knowledge Check

1 **What is the capacity of AWS Snowball?**

70 Terabytes

50 Petabytes

80 Terabytes and 50 Terabytes

80 Terabytes and 40 Terabytes

What is the capacity of AWS Snowball?

1

70 Terabytes

50 Petabytes

80 Terabytes and 50 Terabytes

80 Terabytes and 40 Terabytes

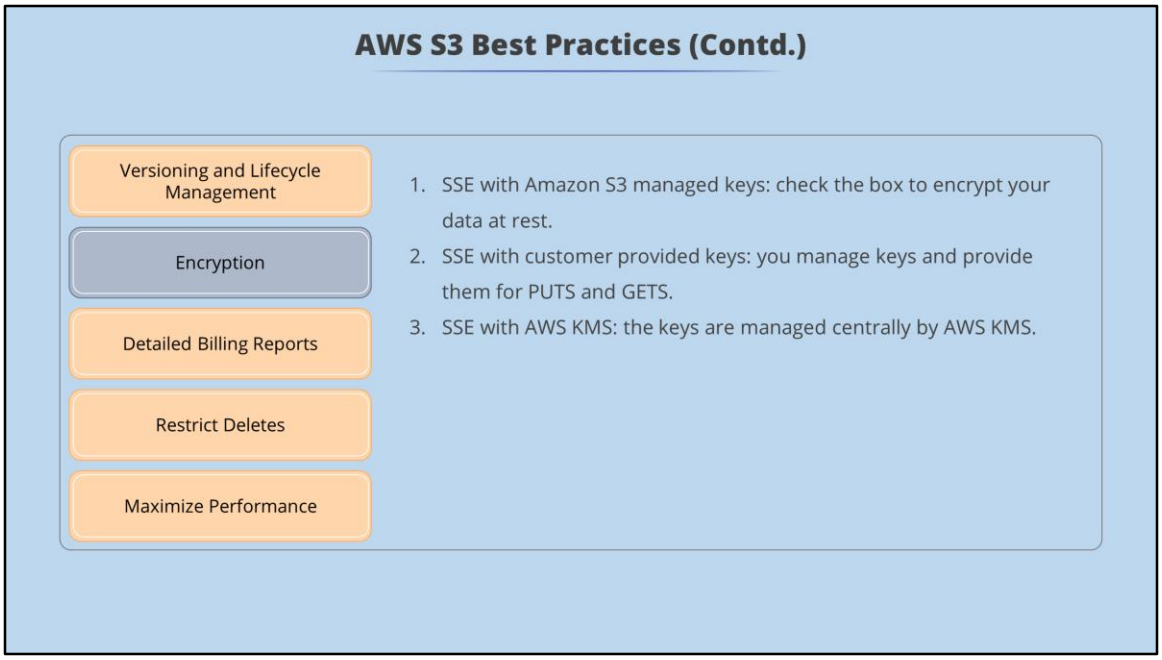
d

AWS Snowball is now available in 80 Terabytes and 50 terabytes. It was earlier available in only 80 Terabytes. To Achieve Petabyte scale transfers, you need to send multiple snowball appliances.

In this section, you'll learn about the Amazon S3 best practices.

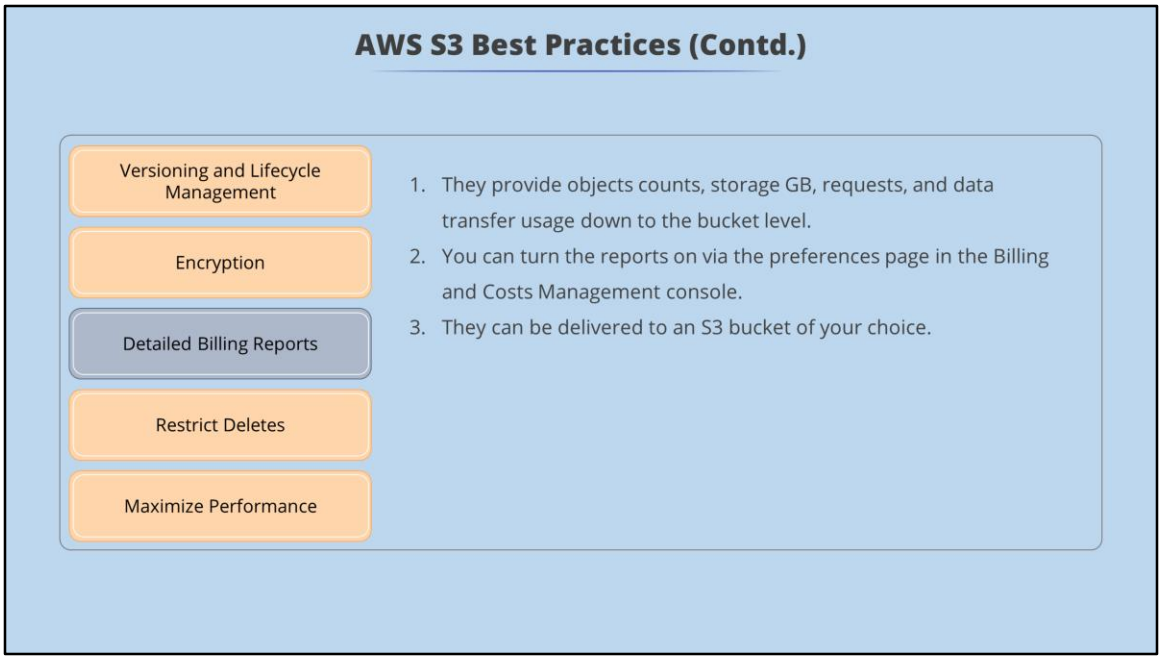
Versioning and Lifecycle Management:

- Enable versioning to protect your data and configure lifecycle policies to move your old versions to Glacier to save storage costs.
- Configure old versions to be deleted at a suitable time in the future.



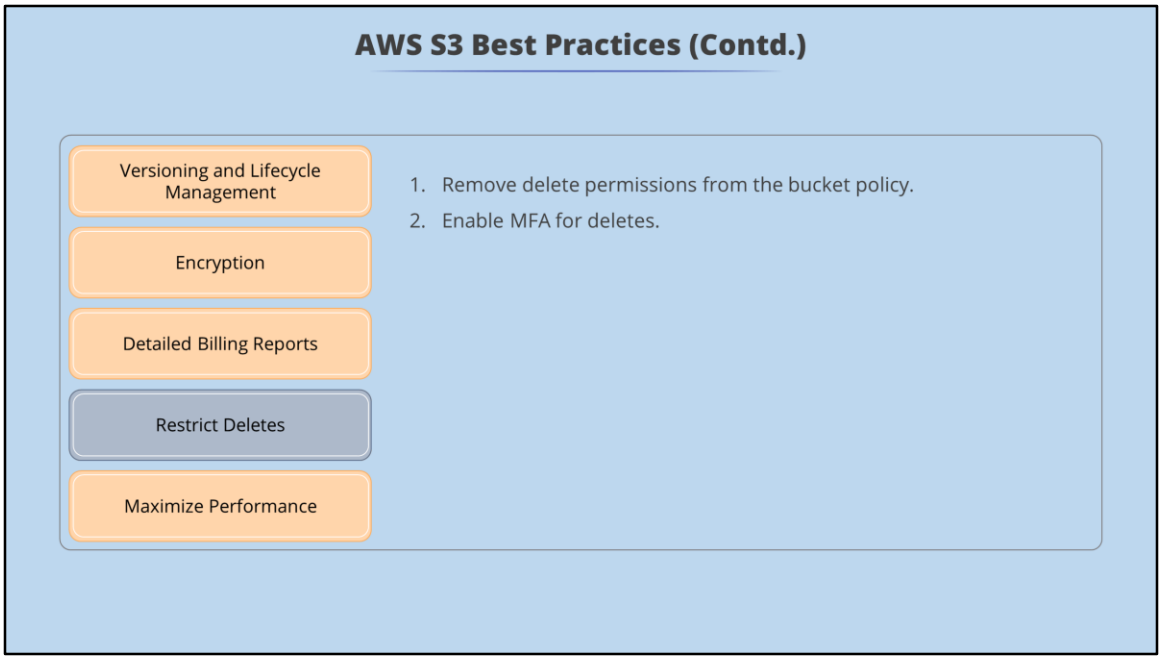
Encryption:

- SSE with Amazon S3 managed keys—check the box to encrypt your data at rest.
- SSE with customer provided keys—you manage keys and provide them for PUTS and GETS functions.
- SSE with AWS KMS—the keys are managed centrally by AWS KMS.



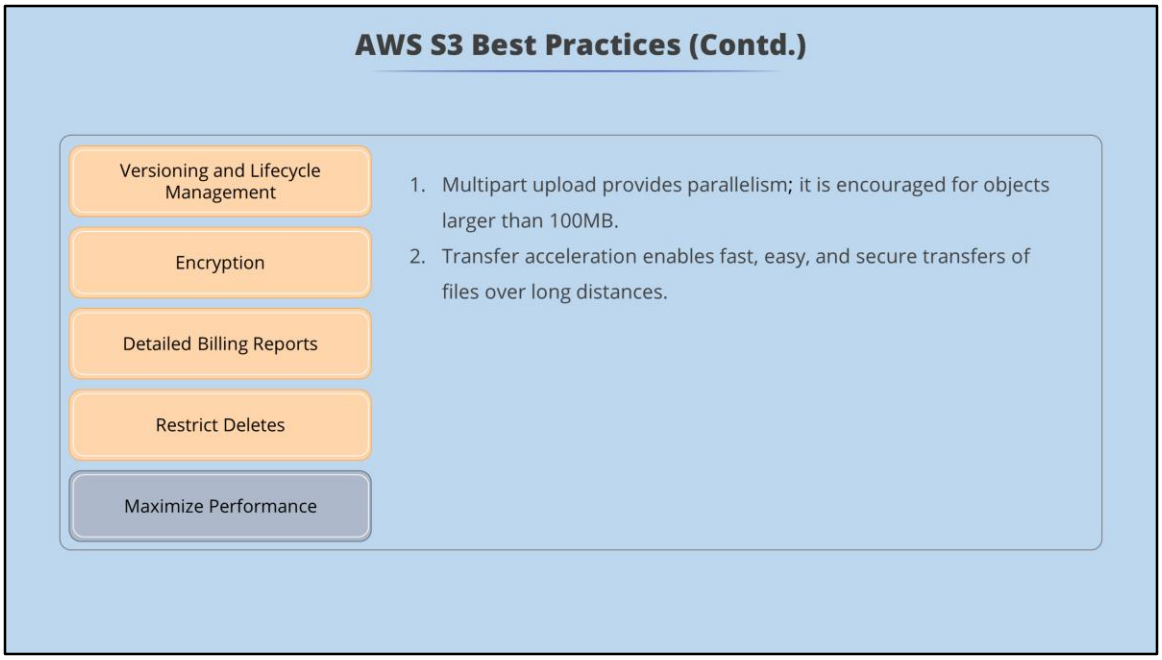
Detailed Billing Reports:

- They provide object counts, storage GB, requests, and data transfer usage down to the bucket level.
- You can turn them on via the preferences page in the Billing and Costs Management console.
- It can be delivered to the S3 bucket of your choice.



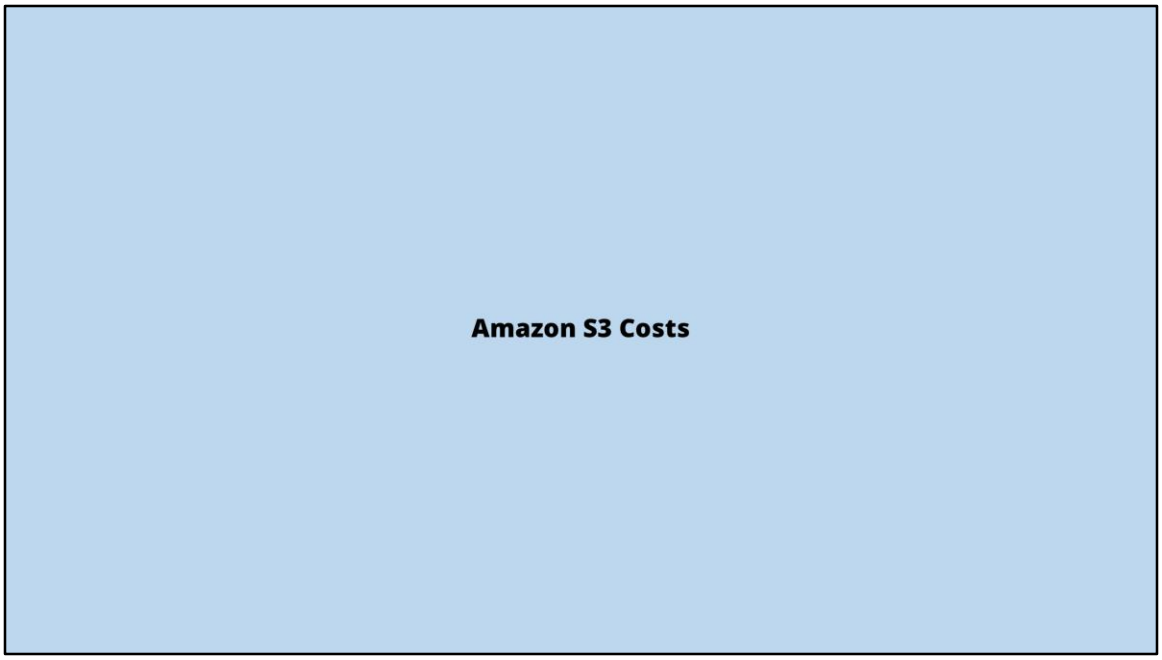
Restrict Deletes:

- Remove delete permissions from the bucket policy
- Enable MFA for deletes

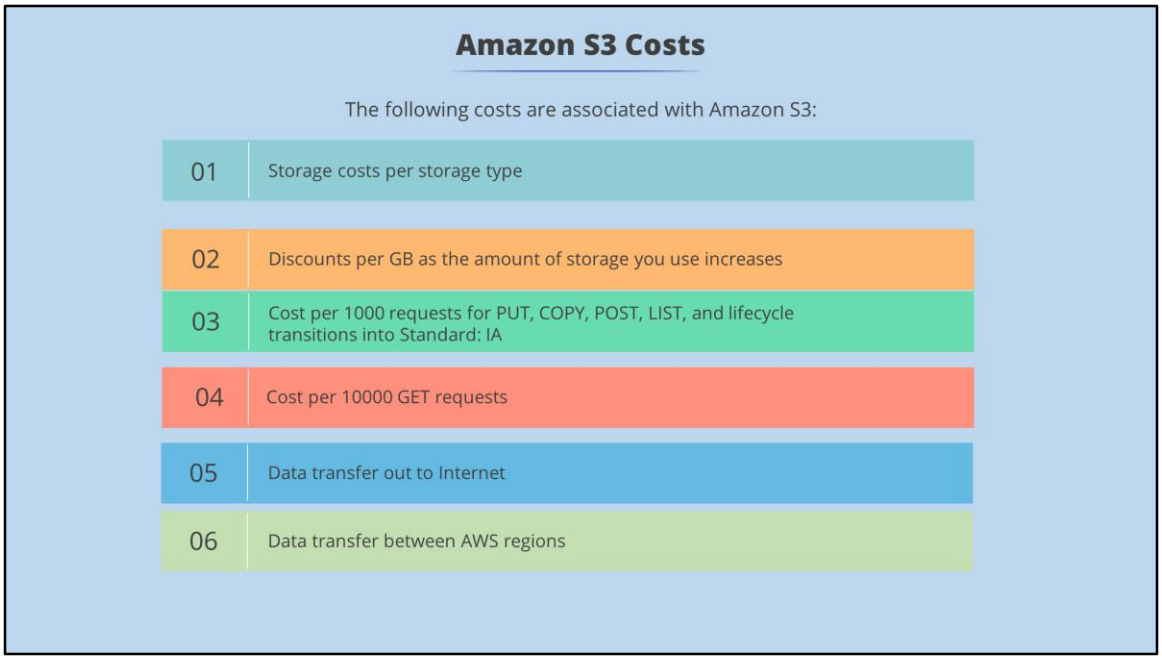


Maximize Performance:

- Multipart upload provides parallelism; it is encouraged for objects larger than 100MB.
- It supports transfer acceleration.

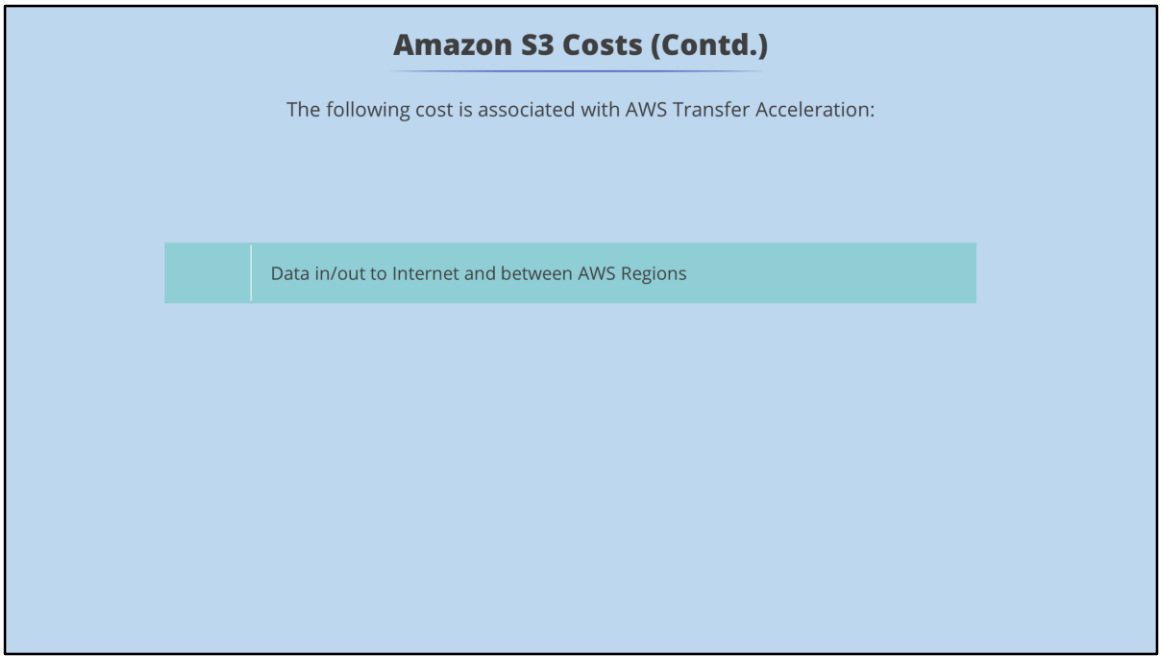


In this section, you'll learn about the costs associated with Amazon S3.



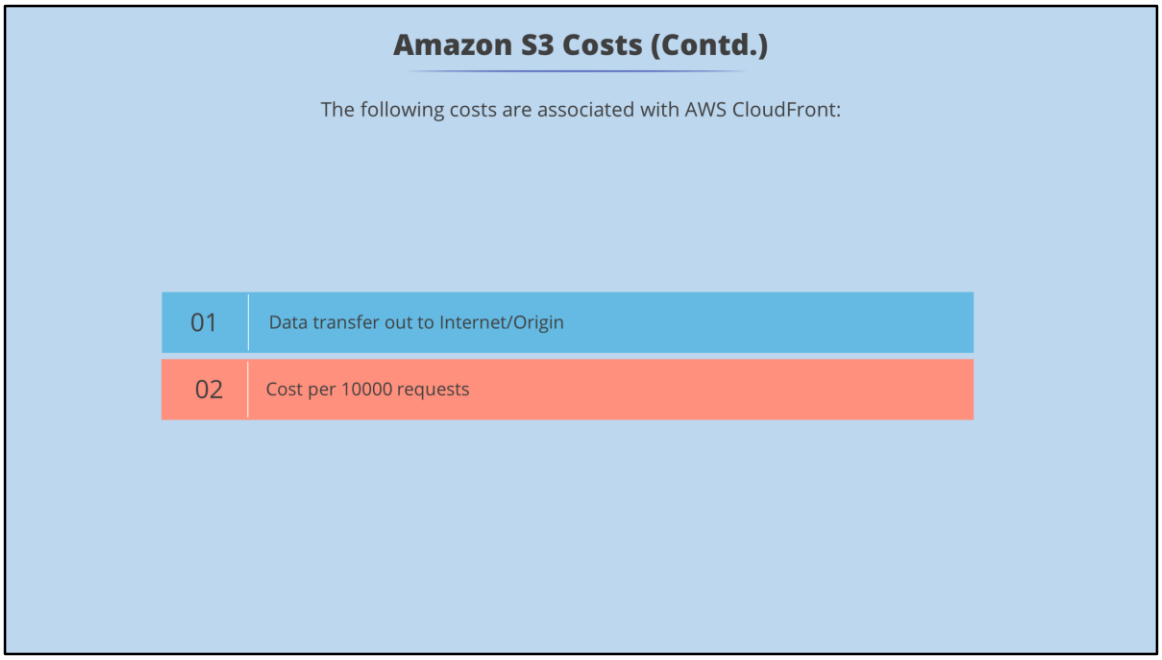
Following are the costs associated with Amazon S3:

- Storage costs per storage type
- Discounts per GB as the amount of storage you use increases
- Cost per 1000 requests for PUT, COPY, POST, LIST and lifecycle transitions into Standard—IA
- Cost per 10000 GET requests
- Data transfer out to Internet
- Data transfer between AWS regions



The following cost is associated with AWS Transfer Acceleration:

- Data transfer in/out to Internet and between AWS Regions



The following costs are associated with AWS CloudFront:

- Data transfer out to Internet/Origin
- Cost per 10000 requests

S3 vs EFS vs EBS

Feature	S3	EFS	EBS
Storage Size	No limit on the number of objects	No limitations	Max 16TiB
File Size Limitation	0 bytes – 5 TB	Max file size 52TiB	No limitation
Data Throughput	Multipart upload recommended for the objects > 100MB	Default throughput of 3GB	Variable depending on disk type
Data Access	Accessible from anywhere	Can be accessed concurrently	Limited to a single EC2 instance
Operating System	N/A	Windows not supported	All operating systems
AZ Failure	Can withstand up to two concurrent AZ failures. Can replicate to other regions.	Can survive one AZ failure	Cannot withstand AZ failure without snapshots

This table might look familiar, we saw it earlier on in the course in the compute section.

It's been modified to describe how S3 compares to EFS and EBS to help you identify when you would use each service.

Firstly, S3 is object storage, whereas EFS is filesystem storage and EBS is block storage.

Like EFS, S3 has no limit to the amount of storage you can use.

And also like EFS, S3 has a file size limitation, which is 5TB. Notice that AWS are getting fast and loose with listing sizes switching between tebibytes and terabytes. The largest object that can be uploaded in a single PUT is 5 gigabyte, after that size you should be using the multipart upload, which is recommended for objects > 100MB

S3 is accessible from absolutely anywhere as long as the correct permissions have been granted.

S3 can with withstand up to two concurrent AZ failures, but you can also replicate your buckets to other regions.

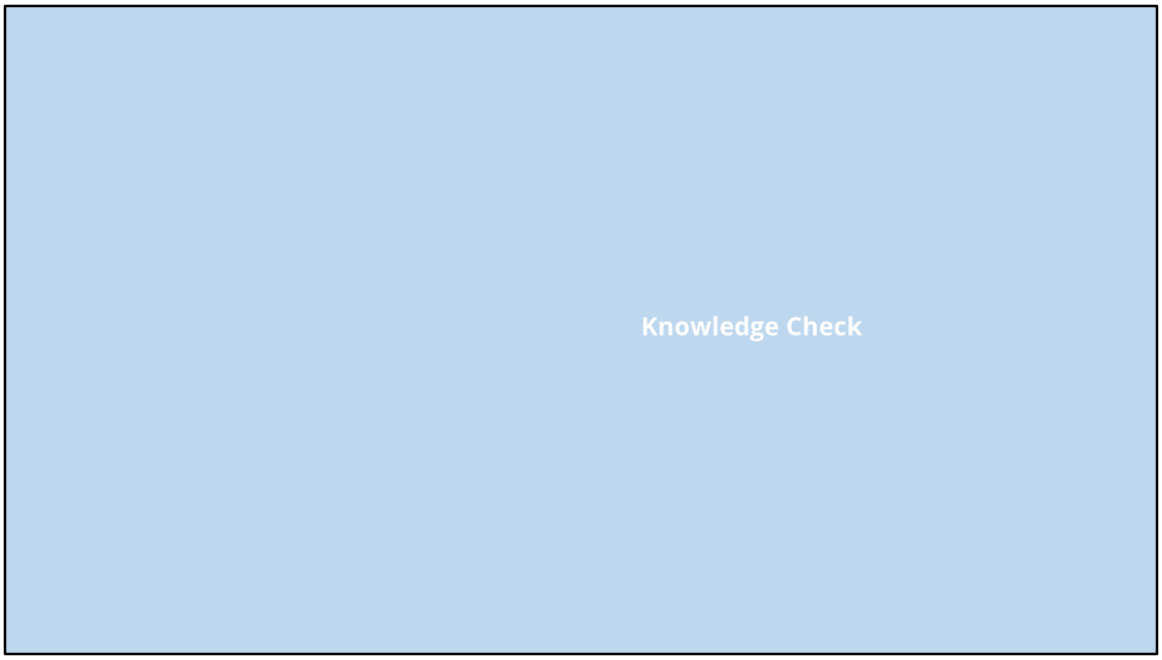
Practice Assignment: Create an Amazon S3 bucket

Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:

1. Frequently accessed critical data: This data needs to be always available and needs to be protected against accidental deletes.
2. Limited period critical data: This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days. After that, it is rarely accessed.
3. Archive data – This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.

- Amazon S3, provides developers and IT teams with secure, durable, highly-scalable cloud storage.
- Amazon S3 provides different storage options with various features.
- S3 allows you define how Amazon manages objects during their lifetime.
- All data stored in Amazon S3 is secure by default as only bucket and object owners have access to the Amazon S3 resources they create.
- Snowball is a petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data into and out of the AWS cloud.



Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.

1

Which object encryption is NOT available in Amazon S3?

Server-side encryption using customer keys

Server-side encryption using Amazon keys

Client-side encryption using customer keys

Cloud HSM

1

Which object encryption is NOT available in Amazon S3?

Server-side encryption using customer keys

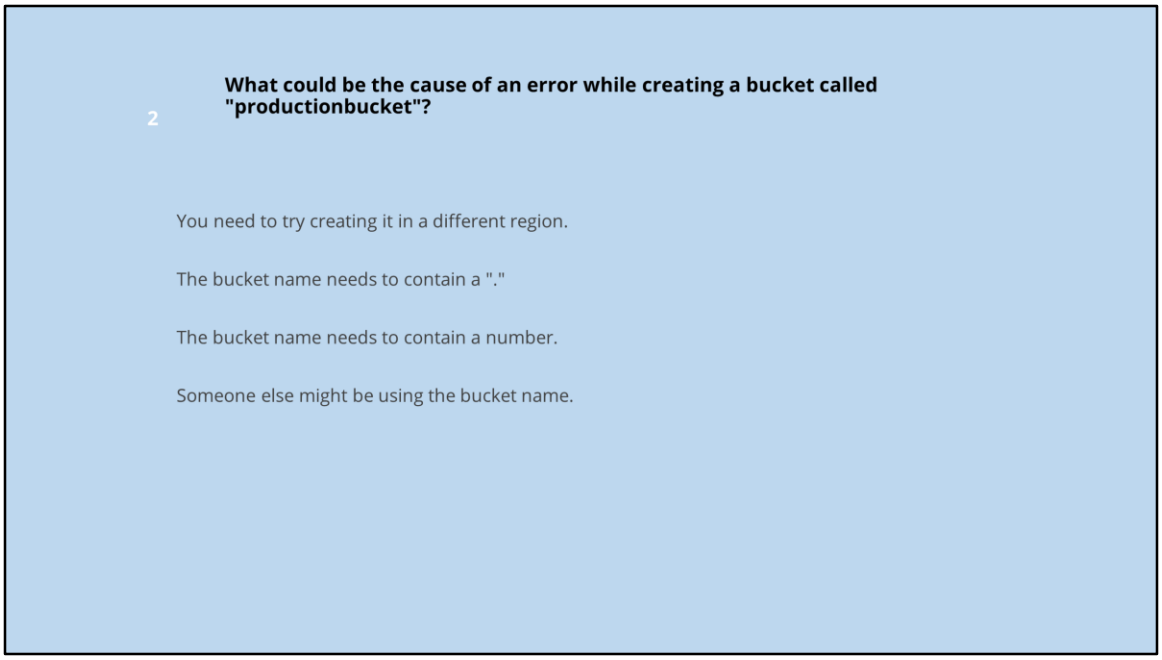
Server-side encryption using Amazon keys

Client-side encryption using customer keys

Cloud HSM

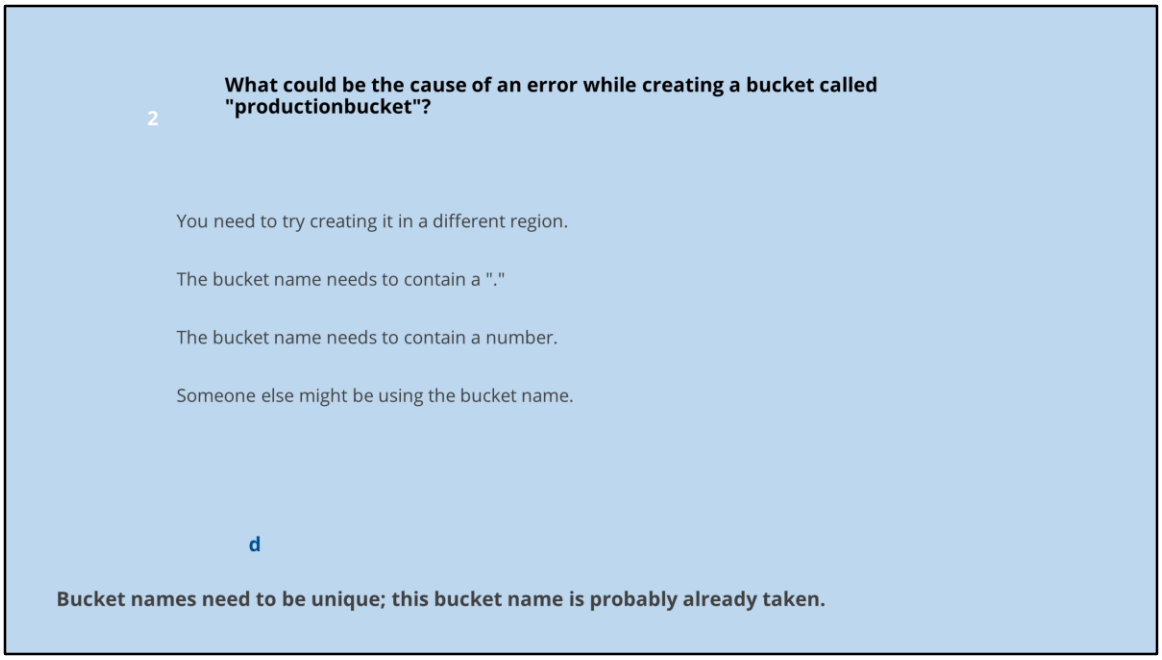
d

CloudHSM is not a valid encryption method for Amazon S3.



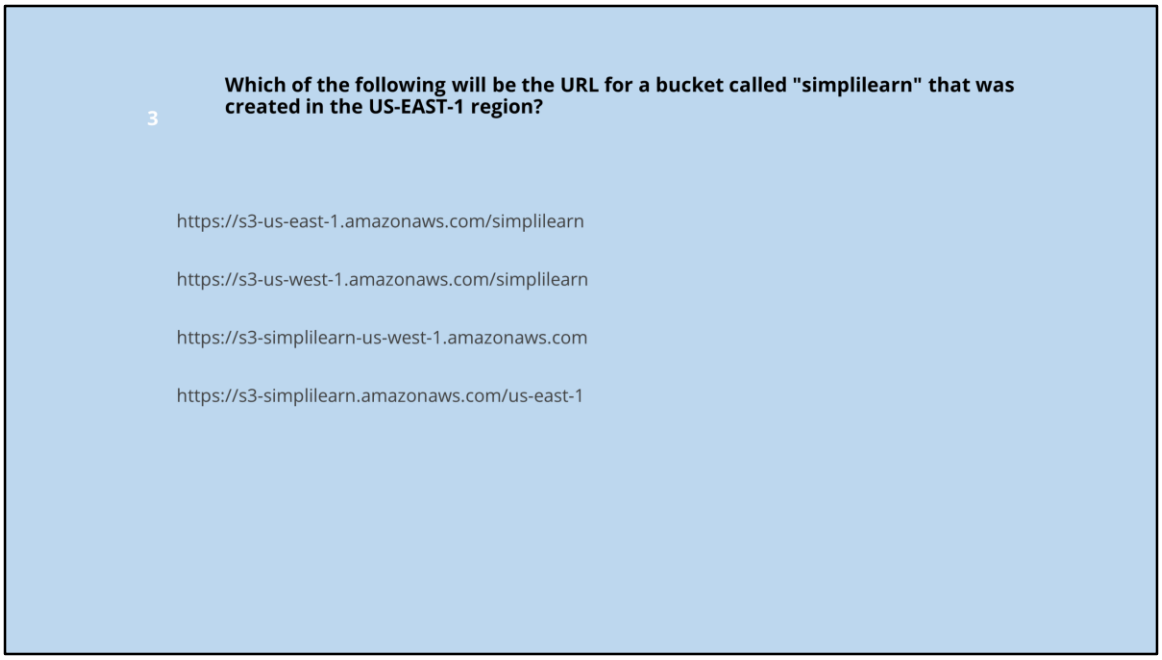
Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.



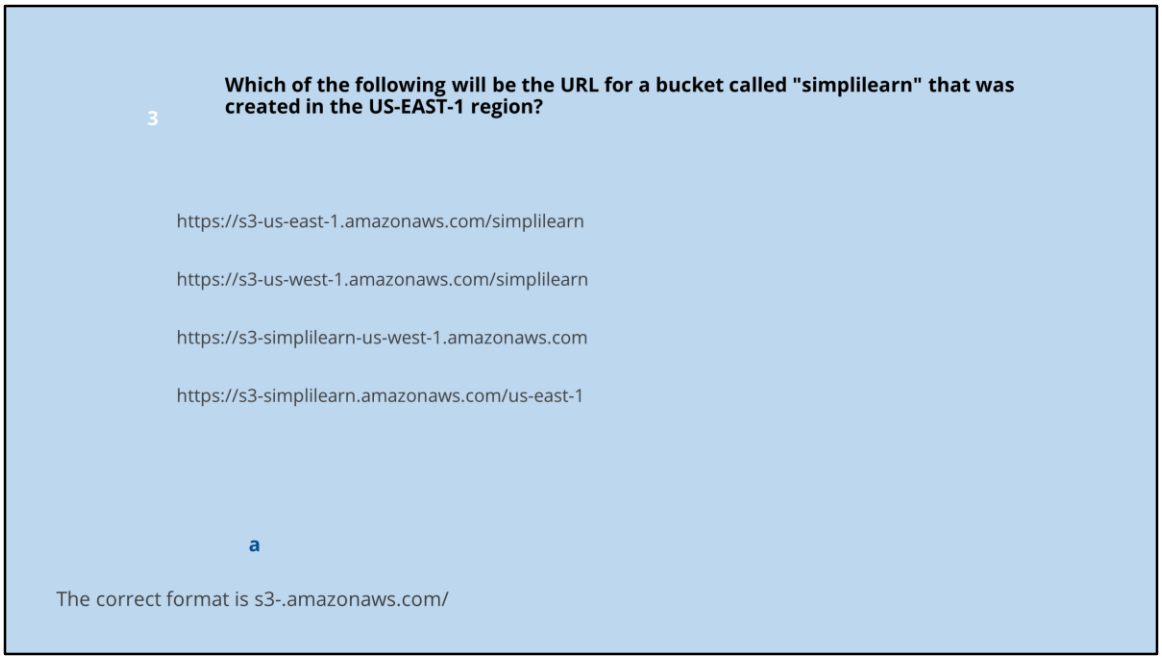
Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.



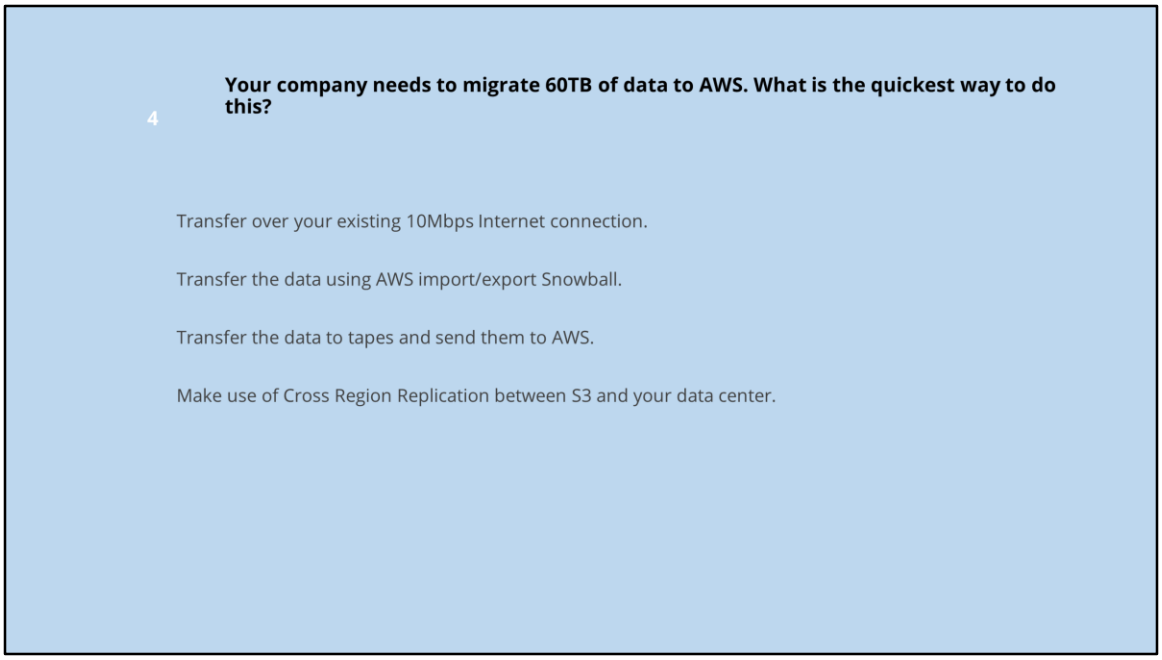
Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.



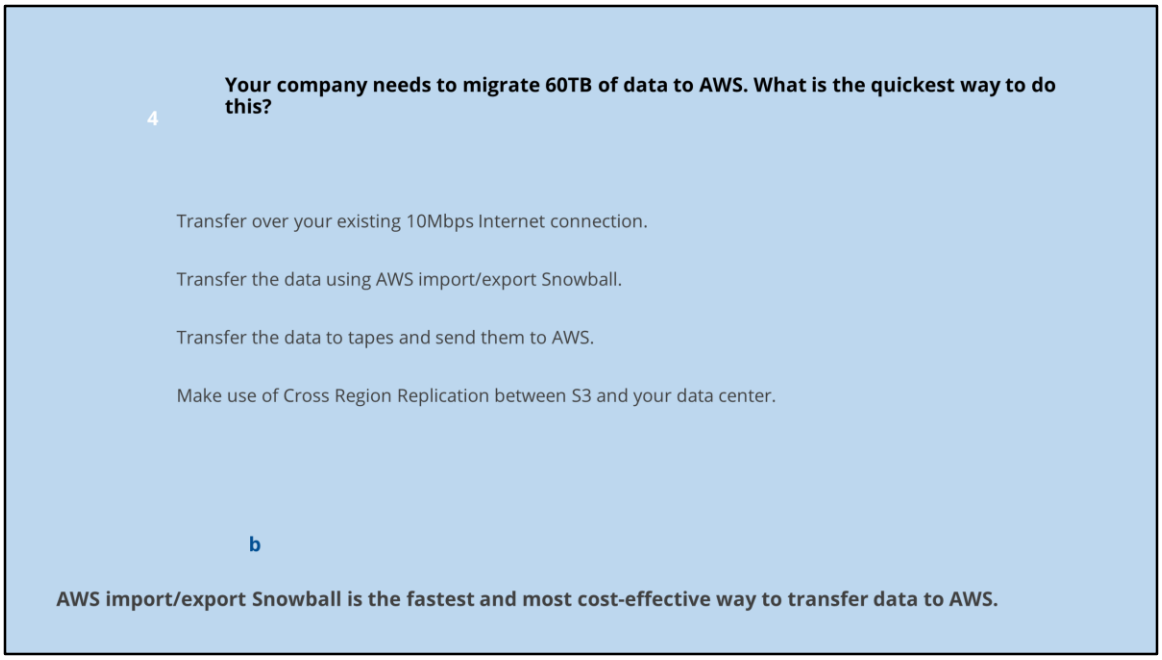
Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.



Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.

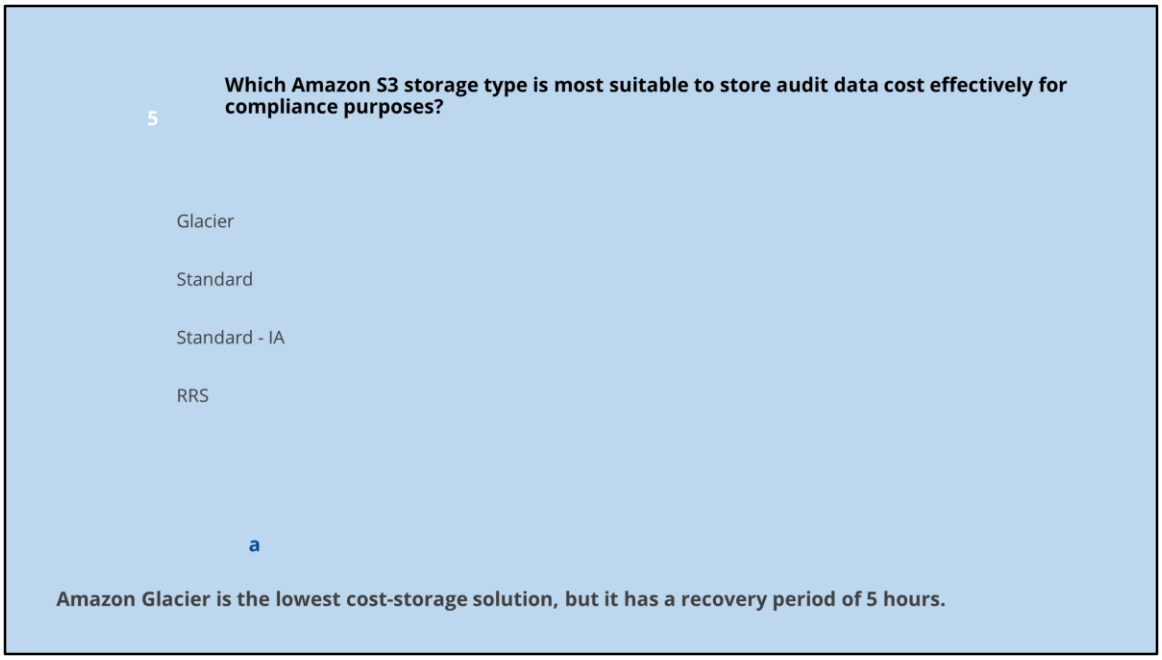


Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.

Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.



Your client wants to move all its data to Amazon S3. The company has categorized its data into the following:
1. Frequently accessed critical data—This data needs to be always available and protected against accidental deletes.
2. Limited period critical data—This data is critical only for the first 30 days, then it is only accessed periodically for another 60 days, after that it is rarely accessed.
3. Archived data—This data is older and needs to be archived for auditing purposes.

Create three buckets for each of the data categories and configure Versioning and Lifecycle Management where appropriate.