



AWS Academy Cloud Architecting  
Module 04 Student Guide  
Version 3.0.0

200-ACACAD-30-EN-SG

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

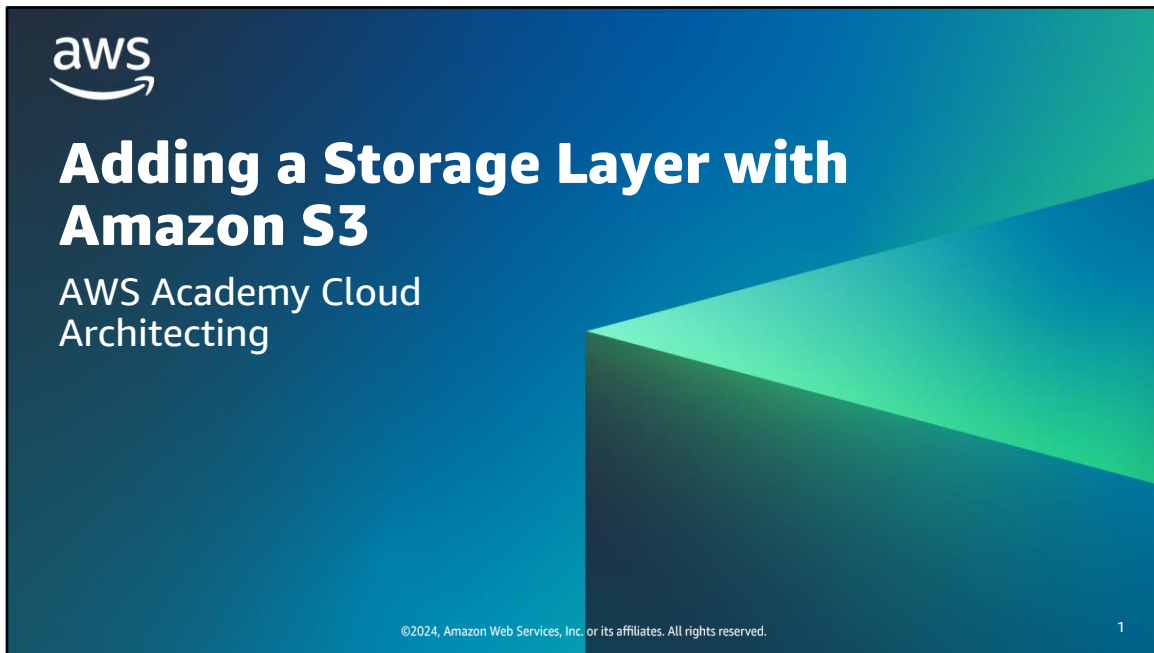
This work may not be reproduced or redistributed, in whole or in part,  
without prior written permission from Amazon Web Services, Inc.  
Commercial copying, lending, or selling is prohibited.

All trademarks are the property of their owners.

# Contents

[Module 4: Adding a Storage Layer with Amazon S3](#)

4





This introduction section describes the content of this module.

## Module objectives



This module prepares you to do the following:

- Define Amazon Simple Storage Service (Amazon S3) and how it works.
- Recognize the problems that Amazon S3 can solve.
- Describe how to move data to and from Amazon S3.
- Manage the storage of content efficiently by using Amazon S3.
- Recommend the appropriate use of Amazon S3 based on requirements.
- Configure a static website on Amazon S3.
- Use the AWS Well-Architected Framework principles when designing a storage layer with Amazon S3.

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

3

## Module overview

### Presentation sections

- Defining Amazon S3
- Using Amazon S3
- Moving data to and from Amazon S3
- Storing content with Amazon S3
- Designing with Amazon S3
- Applying the AWS Well-Architected Framework principles to storage

### Demos

- Amazon S3 Transfer Acceleration
- Managing Lifecycles in Amazon S3
- Amazon S3 Versioning

### Activity

- Designing with Amazon S3

### Knowledge checks

- 10-question knowledge check
- Sample exam question



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

4

The objectives of this module are presented across multiple sections.

You will also participate in an activity on how to setup Amazon S3 for a particular use case.

Your instructor might also demonstrate the features listed. The module wraps up with a 10-question knowledge check delivered in the online course and a sample exam question to discuss in class.

The next slide describes the lab in this module.

## Hands-on lab in this module

### Challenge (Café) lab



- Creating a Static Website for the Café




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

5

This module includes the café lab listed. Additional information about this lab is included in the student guide where the lab takes place, and the lab environment provides detailed instructions.

**As a cloud architect working with Amazon S3:**



- I need to consider access patterns and use cases of the business so that I can choose Amazon S3 configuration options that optimize cost while supporting performance and compliance requirements.
- I need to apply security best practices so that the storage layer is protected against unwanted access and accidental data loss.

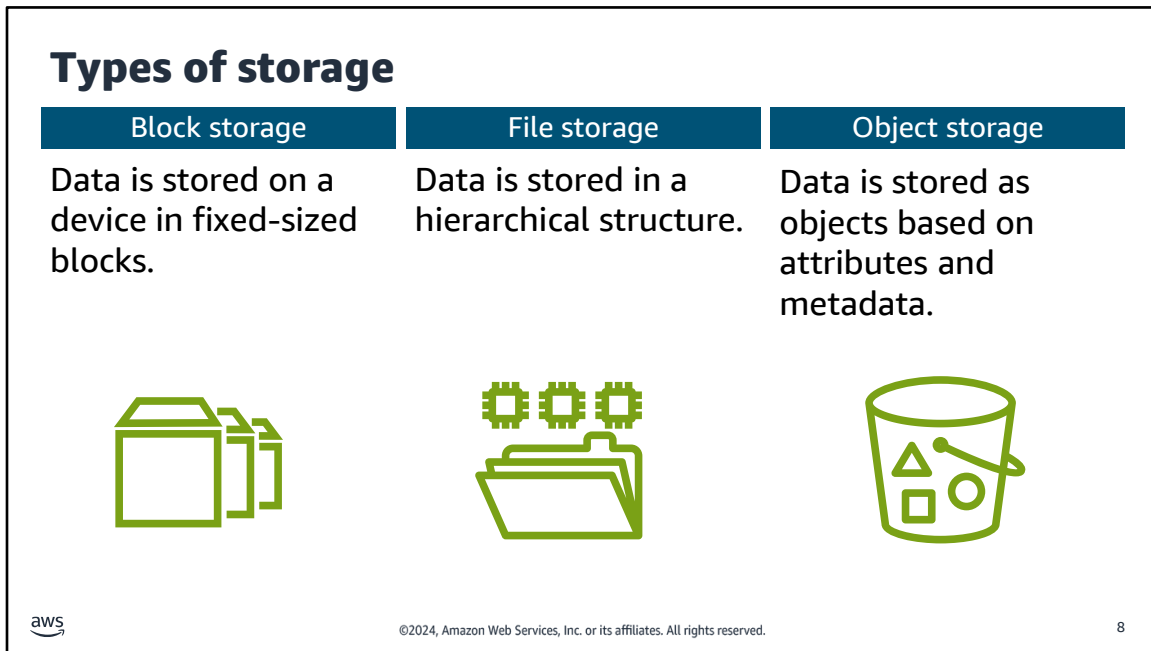
©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

6

This slide asks you to take the perspective of a cloud architect as you think about how to approach cloud network design. Keep these considerations in mind as you progress through this module. Also remember that the cloud architect should work backward from the business need to design the best architecture for a specific use case. As you progress through the module, consider the café scenario presented in the course as an example business need, and think about how you would use Amazon Simple Storage Service (Amazon S3) storage for the café.



This section describes Amazon S3.




Storage comes in three basic types: block storage, file storage, and object storage. These types of storage each have a particular use, and it is likely that you will work with all three types.

Block storage data is stored on a device in fixed-sized blocks. Applications and file systems regulate how blocks are accessed, combined, and modified. Block storage breaks up data into blocks and then stores those blocks as separate pieces, each with a unique identifier. These blocks are stored wherever it is most efficient. Thus, blocks can be stored across different systems, and each block can be configured to work with different operating systems.

File storage is a methodology that helps users, applications, and services access data in a shared file system. Data is stored in a hierarchical structure. This structure is similar to a centralized shared network drive in a company where employees store and access files.

Object storage files are stored as objects based on attributes and metadata. Each object consists of data, metadata, and an object key. The metadata has information about the data (object size, object purpose, and more), and the object key is the unique identifier of the object. When you update files in object storage, the entire file object is updated instead of a piece of a file, as in block storage.

## Amazon S3



Amazon S3

- Amazon S3 stores massive (unlimited) amounts of unstructured data.
- Amazon S3 stores data files as objects in a bucket that you define.
- Five TB is the maximum file size of a single object.
- Objects have a globally unique URL (universal namespace).
- Objects have a key, version ID, value, metadata, and sub-resources.

aws

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

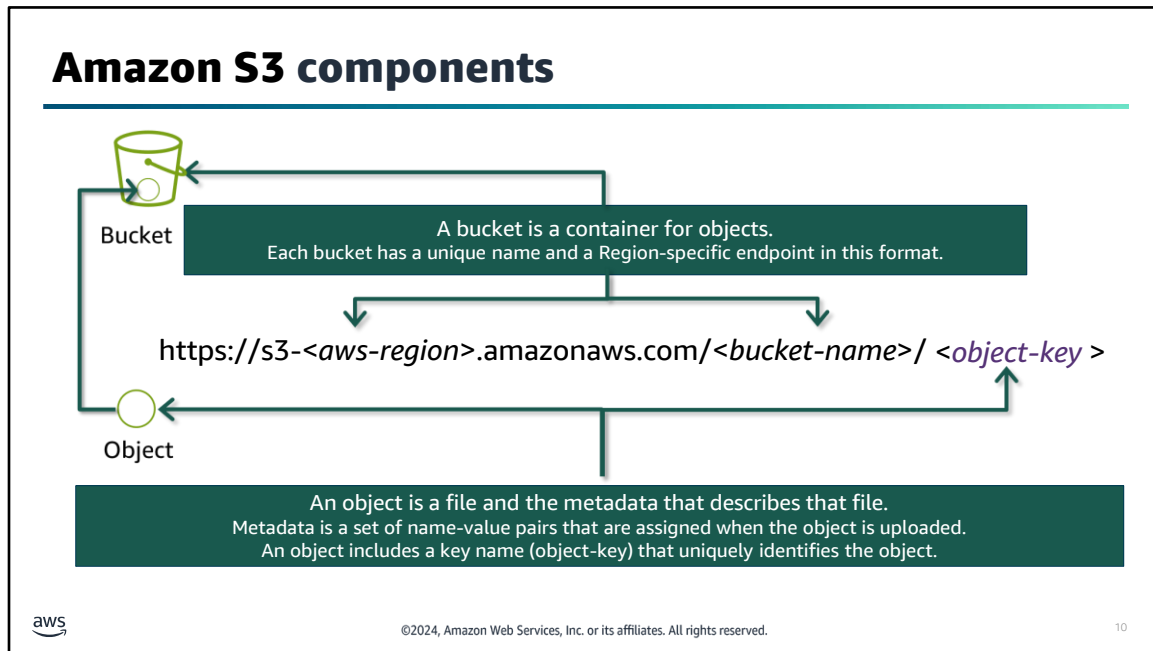
9

- Amazon S3 is an object storage service. You can use it to store virtually unlimited amounts of data.
- Data files are stored as objects. You place objects in a bucket, which you define. Every bucket must have a name that is globally unique across Regions, which means that the bucket name must be unique across all AWS customer accounts.
- The objects that you store can vary in size from 0 bytes to 5 TB. Though individual objects cannot be larger than 5 TB, you can store as much total data as you need.
- Objects have a globally unique URL (universal namespace).

Each object has five consistent characteristics:

- First, it has a *key*, which is the name that you assign to an object. You use the object key to retrieve the object. In the AWS Management Console, you can create a directory inside a bucket and upload an object to that directory. However, in reality, Amazon S3 does not know about directories, so the key value includes the full path relative to the bucket root.
- Objects also include a version ID. In a bucket, a key and version ID uniquely identify an object. You will learn more about versioning later in this module.
- The *value* of the object is the actual content that you store. It can be any sequence of bytes. Object values are immutable, which means that after you upload an object, you cannot modify the value. If you want to modify the object, you must make a change outside of Amazon S3 and then reupload the object.
- Objects also include metadata, which is a set of name-value pairs that you can use to store information about the object. You can assign metadata, which is referred to as *user-defined metadata*, to your objects in Amazon S3. Amazon S3 also assigns system metadata to these objects, which it uses for managing objects.
- Finally, Amazon S3 also uses sub-resources to store additional object-specific information.

For more information on Amazon S3 see the link provided on the course resources page.



Remember, objects are placed in buckets. Here is some high-level information about the essential components of Amazon S3:

- A bucket is a container for objects that are stored in Amazon S3. Buckets serve several purposes. They organize the Amazon S3 namespace at the highest level, and they identify the account responsible for storage and data transfer charges. They also play a role in access control, and they serve as the unit of aggregation for usage reporting.
- An object is the fundamental entity that is stored in Amazon S3. An object can be any kind of file, such as text, video, photo, or another binary format. Objects consist of object data and metadata. The metadata is a set of name-value pairs that describe the object. For example, metadata includes a content type (such as an image or video) that is included in the response header of a browser request so that browser knows how to render the file.
- An object key uniquely identifies the object in a bucket. Each object in a bucket has exactly one key.

Each bucket is Regional. You can choose the AWS Region where Amazon S3 will store the buckets that you create. Objects stored in an AWS Region never leave the Region unless you explicitly transfer them to another Region. The Region code indicates the Region. For example, the Region code for a bucket that's created in the US Oregon Region is `us-west-2`.

See the link provided on the course resources page for more information on Amazon S3 concepts and how it works.

## Use prefixes to imply a folder structure in an S3 bucket

The following objects are in a bucket named **graphics-bucket**.

```
photos/2022/catpiano.jpg
photos/2022/catonphone.jpg
photos/2022/ninepuppies.png
photos/2021/lakefront.png
photos/2021/coveredbridge.png
photos/2021/openairmarket.jpg
video-source/9984.mp4
video-source/9918.mp4
video-source/18446.mp4
```

A **GET** query with the prefix **photos/2022** returns the following objects:

```
graphics-bucket/photos/2022/catpiano.jpg
graphics-bucket/photos/2022/catonphone.jpg
graphics-bucket/photos/2022/ninepuppies.png
```






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


11

In Amazon S3, buckets and objects are the primary resources, and objects are stored in buckets. For organizational simplicity, Amazon S3 supports the folder concept to group objects. Amazon S3 uses a shared name prefix for objects (that is, objects that have names that begin with a common string). When you create a folder, the Amazon S3 console creates an object with the name followed by a slash (/). The Amazon S3 console then displays that object as a folder.

When an S3 bucket is queried, a prefix will limit results to only those objects that begin with that prefix. In the example, the bucket contains objects that are photos organized by year. To retrieve 2022 photos, specify the prefix `photos/2022`.

For more information, see the link titled “Organizing Objects Using Prefixes” in your course resources.

Amazon S3 benefits	
Benefit	Description
 Durability	<ul style="list-style-type: none"><li>Helps ensure that data is not lost</li><li>Provides S3 Standard storage with 11 nines (or 99.999999999 percent) of durability</li></ul>
 Availability	<ul style="list-style-type: none"><li>Provides access to data when needed</li><li>Includes unlimited capacity to store data</li><li>Provides S3 Standard storage with 4 nines (or 99.99 percent) of availability</li></ul>
 High performance	<ul style="list-style-type: none"><li>Achieves thousands of transactions each second when uploading and retrieving storage</li><li>Automatically scales to high request rates</li></ul>

 ©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved. 12



This slide describes three key benefits of Amazon S3 for cloud storage.

First, it provides *durability*, which describes the average annual expected loss of objects. The term 11 nines of durability means that every year, there is a 0.000000001 percent chance of losing an object. For example, if you store 10,000 objects on Amazon S3, you can expect to incur a loss of a single object once every 10,000,000 years on average. Amazon S3 redundantly stores your objects on multiple devices across multiple facilities in the Amazon S3 Region you designate. Amazon S3 is designed to sustain concurrent device failures by quickly detecting and repairing any lost redundancy. Amazon S3 also regularly verifies the integrity of your data by using checksums.

Amazon S3 also provides 4 nines (or 99.99 percent) of *availability*. Availability refers to your ability to access your data quickly when you want it. It also provides a virtually unlimited capacity to store your data, so it is *scalable*. Amazon S3 has robust *security* settings. It provides many ways to control access to the data that you store and also gives you the ability to encrypt your data.

Finally, Amazon S3 is *high performing*. Your applications can achieve thousands of transactions per second when uploading and retrieving storage from Amazon S3. Amazon S3 automatically scales to high request rates.

## Key takeaways: Defining Amazon S3



- Storage comes in three basic types: block storage, file storage, and object storage.
- Amazon S3 is an object storage service.
- Amazon S3 stores massive amounts of unstructured data.
- A bucket is a container for objects that are stored in Amazon S3.
- An object is the fundamental entity that is stored in Amazon S3.
- The key benefits of Amazon S3 include durability, availability, and high performance.

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





13


Here are a few key points to summarize this section.



This section describes different use cases that Amazon S3 supports for customers' solutions.

## How customers use Amazon S3

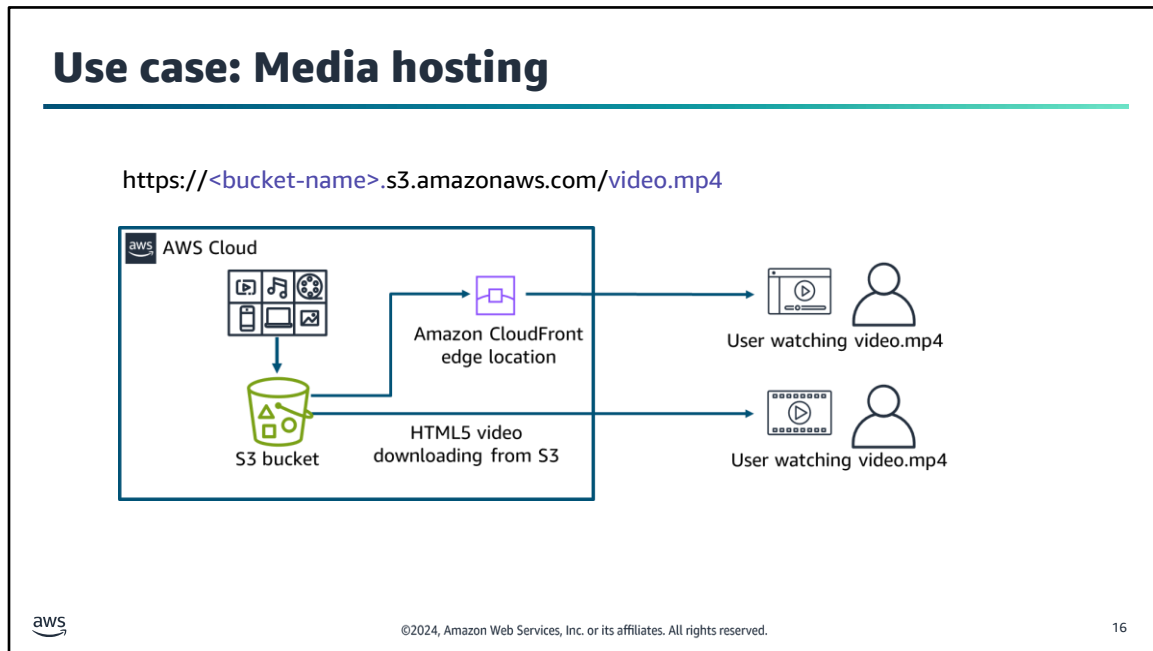
			
<b>Spikes in demand</b>	<b>Static site</b>	<b>Financial analysis</b>	<b>Disaster recovery</b>
Host web content that needs bandwidth to address extreme spikes in demand.	Host a static site that consists of HTML files, images, and videos.	Store data that other services can use for analysis.	Support disaster recovery or data backup solutions.

 ©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved. 15

Now that you know about Amazon S3 features, how can customers use these features to address their needs?

In this section of the module, you will learn about common use cases and see how they relate to the issues addressed on this slide.

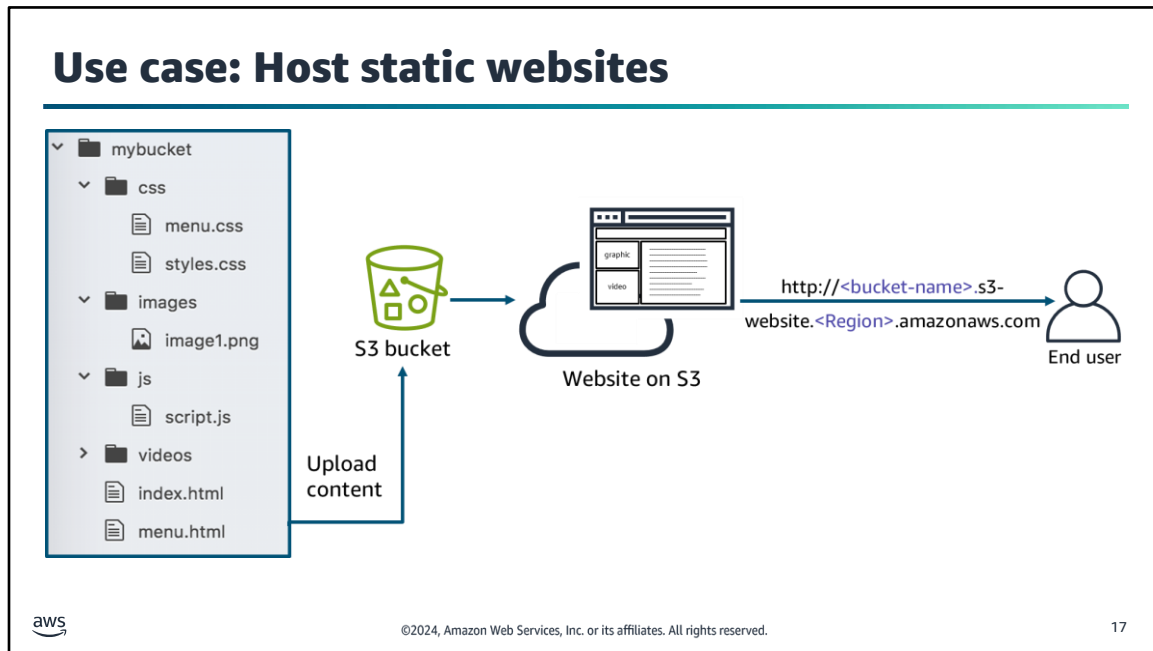
These use cases use Amazon S3 as an essential part of a robust architectural solution.



**Image description:** Diagram of an S3 bucket with video content and two users streaming content from it. One user receives the stream directly, and the other receives it through a CloudFront edge location. **End description.**

One common use case scenario for Amazon S3 is to use it for *media hosting*. In this use case, Amazon S3 is used to store and distribute videos, photos, music files, and other media. This content can be delivered directly from Amazon S3 because each object in Amazon S3 has a unique HTTP URL.

Alternatively, Amazon S3 can serve as an origin store for a content delivery network (CDN), such as Amazon CloudFront. The elasticity of Amazon S3 makes it well-suited for hosting web content that needs bandwidth to address extreme spikes in demand. Also, because you do not need to provision storage for Amazon S3, it works well for fast-growing websites that host data-intensive, user-generated content, such as video- and photo-sharing sites.



**Image description:** Diagram shows example objects stored in an S3 bucket configured to host a static website. The bucket contents include HTML, .js, and .css files, and an end-user is able to view the website through a URL in the form of `http://bucket-name.s3-website.Region.amazonaws.com`. **End description.**

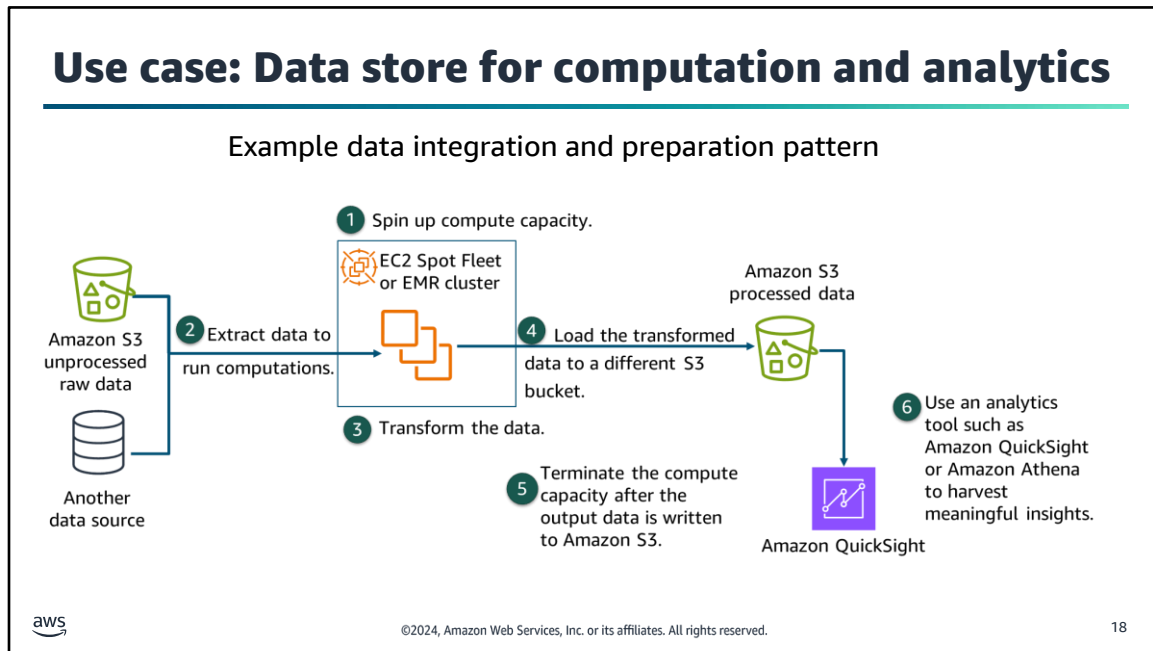
A second Amazon S3 use case is to use the service to host a static website. On a static website, individual webpages include static content. They might also contain client-side scripts.

By contrast, a *dynamic* website relies on server-side processing, which might involve database queries that run in response to server-side scripts, such as PHP, JSP, or ASP.NET. Amazon S3 does not support server-side scripting. However, AWS offers other services that you can use to host dynamic websites.

To host a *static* website, configure an S3 bucket for website hosting, and attach a bucket policy that allows access to the objects. Then, upload your website content to the bucket.

The example shows that the static site might consist of HTML files, images, videos, and client-side scripts in formats such as JavaScript.

With this approach, you do not need to run a virtual machine that hosts a web server. In fact, you do not need to run a server. However, you can still host a website. Amazon S3 provides a low-cost solution for web hosting that includes high performance, scalability, and availability.

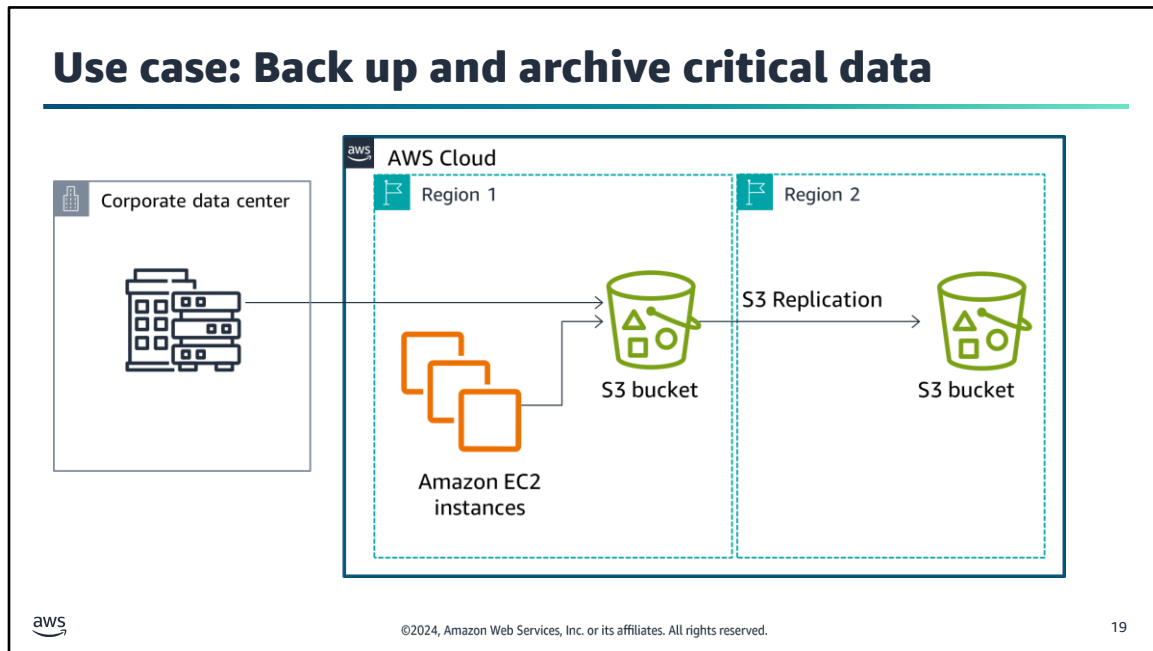


**Image description** Graphic shows a compute cluster spun up temporarily. It extracts data from an S3 bucket and another undefined source, transforms the data, and then loads the transformed data to a processed data S3 bucket where Amazon QuickSight or another analytics tool can access the transformed data. **End Description**

You can also use Amazon S3 as a data store for computation or large-scale analytics, such as financial transaction analysis, clickstream analytics, and media transcoding. Amazon S3 can support these workloads because of its horizontal scaling ability, which provides for multiple concurrent transactions.

The following steps describe this example:

1. An Amazon Elastic Compute Cloud (Amazon EC2) Spot Fleet is spun up when the bid price for Spot Instances is low or when an Amazon EMR cluster is spun up.
2. Regardless, after the compute capacity is available, *raw unprocessed* data is extracted from Amazon S3 and also from another data source.
3. The data is run through compute algorithms that integrate and transform it.
4. The resulting *processed* data is loaded into a different S3 bucket.
5. Now that the data has been processed, the compute capacity is terminated to save on costs.
6. Finally, an analytics tool, such as Amazon QuickSight, might be used to harvest meaningful insights from the processed data. This is just one example scenario of how Amazon S3 can play an essential role for data storage in a large-scale analytics solutions architecture.



**Image description:** Graphic shows data from a corporate data center being uploaded to an S3 bucket. Other data from Amazon EC2 instances is also shown as being backed up to the same bucket. **End description.**



In the fourth and final use case discussed in this module, Amazon S3 is used as a data backup solution. Because of its highly durable and scalable nature, Amazon S3 works well as a data backup and archival tool.

In the scenario, data is backed up from an on-premises corporate data center and also from a large number of Amazon EC2 instances. These instances run applications that generate data.

Another Amazon S3 option that you can configure on your buckets—to achieve even higher levels of durability—is cross-Region replication. In cross-Region replication, objects that are uploaded to a bucket in one Region will be automatically copied to other S3 buckets in other Regions. Note that this is an asynchronous process.

Additionally, you can move long-term data from Amazon S3 standard storage to Amazon Simple Storage Service Glacier (Amazon S3 Glacier). This module discusses this process in more detail later.

## Key takeaways: Using Amazon S3



Amazon S3 is often used to do the following:

- Store and distribute videos, photos, music files, and other media.
- Support static content, including HTML files, images, videos.
- Store data for computation and large-scale analytics.
- Provide a data backup solution.

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

20

Here are a few key points to summarize this section.



This section describes how to move data to and from Amazon S3.

## Storing objects in Amazon S3

- There is no limit to the number of objects in a bucket.
- Uploading an object requires write permission to the bucket.
- Objects are encrypted by default.
  - During upload, objects are automatically encrypted by using server-side encryption.
  - During download, objects are decrypted.




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

22

When you upload a file to Amazon S3, it is stored as an S3 object. You can have an unlimited number of objects in a bucket. Before you can upload files to an S3 bucket, you need write permissions for the bucket.

You can upload any file type, including images, backups, data, and movies, into an S3 bucket. When you upload an object, the object is automatically encrypted by using server-side encryption with Amazon S3 managed keys (SSE-S3) by default. When you download an object, the object is decrypted.

Options for uploading objects to Amazon S3	
Uploading Amazon S3	Description
AWS Management Console	Use a wizard-based approach to move data into or out of Amazon S3, including the option to drag and drop files. (maximum size is 160 GB).
AWS Command Line Interface (AWS CLI)	Upload or download from a terminal command prompt or in a call from a script.
AWS SDKs	Use AWS SDKs to upload objects programmatically.
Amazon S3 REST API	Send a PUT request to upload data in a single operation.

 ©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved. 23

The AWS Management Console offers a way to move data into or out of Amazon S3. It offers a wizard-based approach, including the option to drag and drop files that you want to copy into a bucket. The maximum size of a file that you can upload by using the Amazon S3 Management Console is 160 GB.

To upload a file larger than 160 GB, use the AWS Command Line Interface (AWS CLI), AWS SDKs, or Amazon S3 REST API. With these three methods, you can upload an object in parts by using the multipart upload API operation. You can upload a single large object up to 5 TB in size. The following slide covers multipart uploads in more detail.

You can use the AWS CLI to upload or download from a terminal command prompt or in a call from a script.

Use the AWS SDKs to upload objects in Amazon S3. The SDKs provide wrapper libraries for you to upload data. See the link provided on the course resources page for a list of supported SDKs..

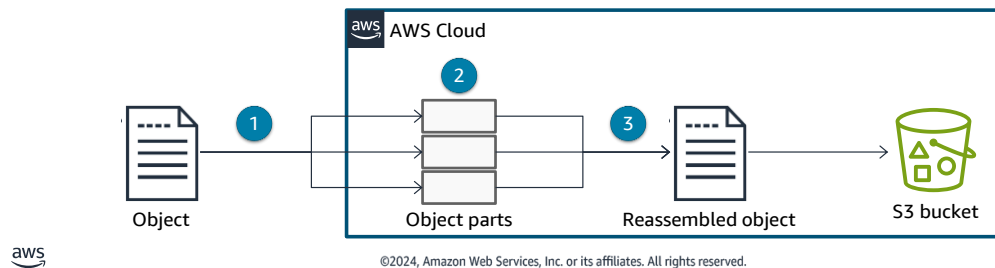
Send Amazon S3 REST API requests to upload an object. You can send a PUT request to upload data in a single operation. A PUT request adds an object to a bucket. You must have write permissions on a bucket to add an object to it. See the link provided on the course resources page for more information PUT objects.

See the link provided on the course resources page for more information about S3 AWS CLI Command reference.

## Amazon S3 feature: Multipart upload

Multipart uploads have the following advantages:

- Improve throughput.
- Recover quickly from any network issues.
- Pause and resume object uploads.
- Begin an upload before you know the final object size.



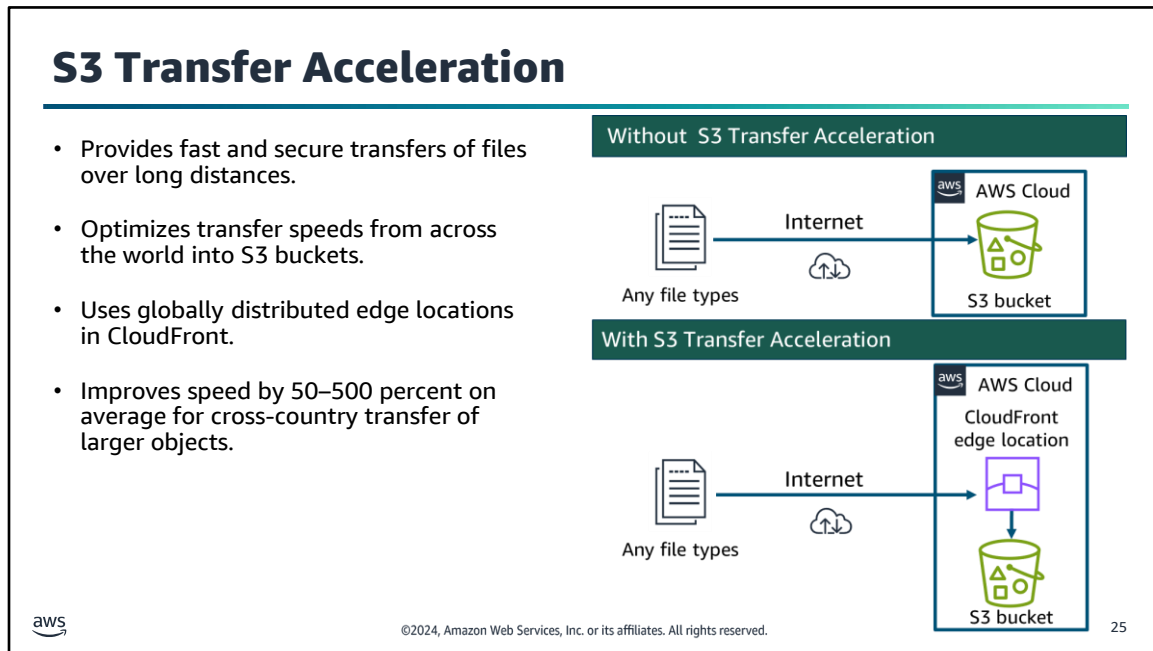
24

You can use multipart upload to upload a single object as a set of parts. Each part is a contiguous portion of the object's data. You can upload these object parts independently and in any order. If transmission of any part fails, you can retransmit that part without affecting other parts. After all parts of your object are uploaded, Amazon S3 assembles these parts and creates the object. Amazon S3 carries out this process behind the scenes whenever you upload files.

The following are advantages to multipart uploads:

- Improved throughput: Parts are uploaded in parallel to improve throughput.
- Recover quickly from any network issues: Smaller part size minimizes the impact of restarting a failed upload due to a network error.
- Pause and resume object uploads: You can upload object parts over time. After you initiate a multipart upload, there is no expiry; you must explicitly complete or stop the multipart upload.
- Begin an upload before you know the final object size: You can upload an object as you are creating it.

See the link provided on the course resources page for more information about multipart upload.



**Image description:** Diagram compares a standard Amazon S3 upload that is uploaded directly to an S3 bucket with an S3 Transfer Acceleration upload that is uploaded to a CloudFront edge location before it is then transferred to the S3 bucket. **End description**

Now that you have learned about the advantages of multipart upload, you will look at how can you improve transfer speeds when uploading files.


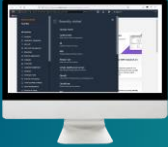
S3 Transfer Acceleration is a bucket-level feature that enables fast and secure transfers of files over long distances between your client and an S3 bucket. Transfer Acceleration is designed to optimize transfer speeds from across the world into S3 buckets. Transfer Acceleration takes advantage of the globally distributed edge locations in CloudFront. As the data arrives at an edge location, the data is routed to Amazon S3 over an optimized network path.

You might want to use Transfer Acceleration on a bucket for various reasons:

- Your customers upload to a centralized bucket from all over the world.
- You transfer gigabytes to terabytes of data on a regular basis across continents.
- You can't use all of your available bandwidth over the internet when uploading to Amazon S3.

See the link provided on the course resources page to compare accelerated and non-accelerated upload speeds across Amazon S3 Regions.

## Demo: Amazon S3 Transfer Acceleration



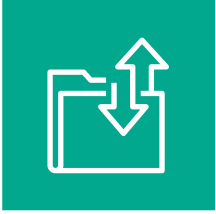
- This demo uses Amazon S3 Transfer Acceleration.
- In this demonstration, you will see how to enable transfer acceleration for an S3 bucket and access accelerated data transfers.

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

26

Find this recorded demo in your course as part of this module.

## AWS Transfer Family



AWS Transfer Family

- Is a fully managed AWS service
- Is used to transfer files into and out of Amazon S3 storage or Amazon Elastic File System (Amazon EFS) file systems over the following protocols:
  - Secure Shell (SSH) File Transfer Protocol (SFTP) version 3
  - File Transfer Protocol Secure (FTPS)
  - File Transfer Protocol (FTP)
  - Applicability Statement 2 (AS2)

aws

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

27

AWS Transfer Family is a secure transfer service that you can use to transfer files into and out of AWS storage services. Transfer Family is part of the AWS Cloud platform.

Transfer Family supports transferring data from or to the following AWS storage services.

- Amazon S3 storage
- Amazon Elastic File System (Amazon EFS) Network File System (NFS) file systems

Transfer Family supports transferring data over the following protocols:

- Secure Shell (SSH) File Transfer Protocol (SFTP) version 3
- File Transfer Protocol Secure (FTPS)
- File Transfer Protocol (FTP)
- Applicability Statement 2 (AS2)

For more information about Transfer Family, see the AWS Transfer Family user guide linked in your course resources.

## Transfer Family benefits

- Transfer Family is a managed service that scales in real time to meet your needs.
- You don't need to modify your applications or run any file transfer protocol infrastructure.
- With Transfer Family, you use native AWS services for processing, analytics, reporting, auditing, and archival functions with your data in durable Amazon S3 storage.
- Transfer Family is a managed elastic file system (with Amazon EFS) for use with AWS Cloud services and on-premises resources.
- Transfer Family is a managed, serverless file transfer workflow service that you can use to set up, run, automate, and monitor file uploads.
- You pay for only the use of the service, and there are no upfront costs.



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

28

Transfer Family provides the following benefits:

- It is a fully managed service that scales in real time to meet your needs.
- You don't need to modify your applications or run any file transfer protocol infrastructure.
- With your data in durable Amazon S3 storage, you can use native AWS services for processing, analytics, reporting, auditing, and archival functions.
- With Amazon EFS as your data store, you get a fully managed elastic file system for use with AWS Cloud services and on-premises resources. Amazon EFS is built to scale on demand to petabytes without disrupting applications. It grows and shrinks automatically as you add and remove files to help eliminate the need to provision and manage capacity to accommodate growth.
- It is a fully managed, serverless file transfer workflow service that you can use to set up, run, automate, and monitor file uploads.
- There are no upfront costs, and you pay for only the use of the service.

## Use cases for Transfer Family

### Amazon S3

- Data lakes in AWS for uploads from third parties
- Subscription-based data distribution with customers
- Internal transfers within your organization

### Amazon EFS

- Data distribution
- Supply chain
- Content management
- Web-serving applications



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

29

The following are some common use cases for using Transfer Family with Amazon S3:


- Data lakes in AWS for uploads from third parties, such as vendors and partners
- Subscription-based data distribution with your customers
- Internal transfers within your organization

The following are some common use cases for using Transfer Family with Amazon EFS:

- Data distribution
- Supply chain
- Content management
- Web serving applications

See the link provided on the course resources page for more information about AWS Transfer Family.

**Key takeaways:  
Moving data to and  
from Amazon S3**



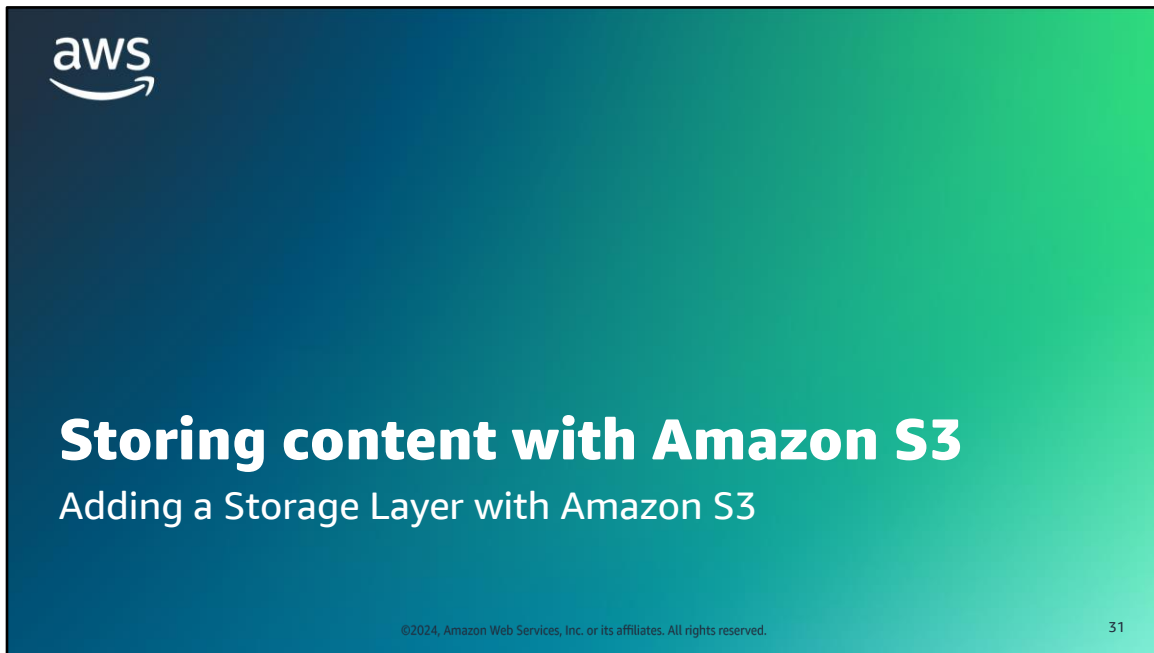
aws

- Upload objects to Amazon S3 by using the AWS Management Console, AWS CLI, AWS SDKs, or Amazon S3 REST API.
- Use multipart upload to upload a single object as a set of parts when the object size reaches 100 MB or greater.
- Use S3 Transfer Acceleration on an S3 bucket to provide fast and secure transfers of files over long distances between your client and an S3 bucket.
- Use Transfer Family to provide a secure transfer of files into and out of AWS storage services.

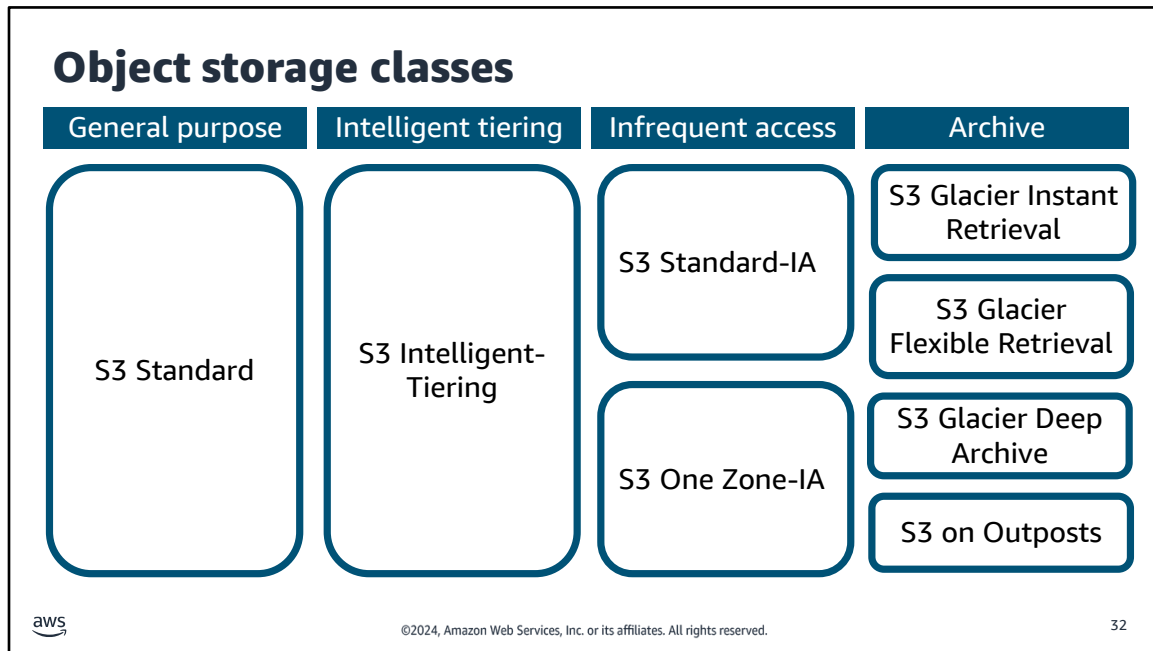
©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

30

Here are a few key points to summarize this section.



This section describes storing content with Amazon S3.



Amazon S3 provides different storage classes aligned to different customer requirements. Here is a list of each storage type with a description:

- General purpose
  - S3 Standard offers high durability, availability, and high-performing object storage for frequently accessed data. Because it delivers low latency and high throughput, S3 Standard is appropriate for a variety of use cases, including cloud applications, dynamic websites, content distribution, mobile and gaming applications, and big data analytics. It provides durability across at least three Availability Zones.
- Intelligent tiering
  - S3 Intelligent-Tiering is the first cloud storage that automatically reduces your storage costs on a granular object level by automatically moving data to the most cost-effective access tier based on access frequency, without performance impact, retrieval fees, or operational overhead. S3 Intelligent-Tiering delivers milliseconds latency and high throughput performance for frequently, infrequently, and rarely accessed data in the Frequent, Infrequent, and Archive Instant Access tiers. You can use S3 Intelligent-Tiering as the default storage class for virtually any workload, especially data lakes, data analytics, new applications, and user-generated content.
- Infrequent access
  - S3 Standard-IA (Infrequent Access) offers all the benefits of Amazon S3 Standard, but it runs on a different cost model to store infrequently accessed data, such as older digital images or older log files. There is a 30-day minimum storage fee applied to any data placed in it, and the cost is higher to retrieve data from S3 Standard-IA than from S3 Standard storage.
  - S3 One Zone-IA stores data in a single Availability Zone. It is ideal for customers who want a lower-cost option and who do not need the availability and resilience of S3 Standard or S3 Standard-IA. It's a good choice for storing secondary backup copies of on-premises data or data that you can recreate. You can also use it as cost-effective storage for data that is replicated from another AWS Region.

- Archive
  - S3 Glacier Instant Retrieval is an archive storage class that delivers the lowest-cost storage for long-lived data that is rarely accessed and requires retrieval in milliseconds. S3 Glacier Instant Retrieval is ideal for archive data that needs immediate access, such as medical images, news media assets, or user-generated content archives. You can upload objects directly to S3 Glacier Instant Retrieval or use S3 lifecycle policies to transfer data from the Amazon S3 storage classes.
  - S3 Glacier Flexible Retrieval (formerly S3 Glacier) is for data that does not require immediate access but needs the flexibility to retrieve large sets of data 1-2 times per year and is retrieved asynchronously at no cost. It is an ideal solution for backup, disaster recovery, and offsite data storage needs. It is also a good solution when some data occasionally needs to be retrieved in minutes, and you don't want to worry about costs.
  - S3 Glacier Deep Archive is the lowest-cost storage class in Amazon S3. This storage class supports long-term retention and digital preservation for data that may be accessed once or twice in a year. It is designed for customers—particularly those in highly regulated industries, such as financial services, healthcare, and public sectors—that retain data sets for 7-10 years or longer to meet regulatory compliance requirements.
  - S3 on Outposts delivers object storage to your on-premises AWS Outposts environment and uses the S3 APIs and features available in AWS Regions today. You can use S3 on Outposts to store and retrieve data on your Outpost and to secure the data, control access, tag the data, and report on the data. S3 on Outposts provides a single Amazon S3 storage class, named OUTPOSTS, that uses the Amazon S3 APIs and is designed to durably and redundantly store data across multiple devices and servers on your Outposts. The S3 on Outposts storage class is ideal for workloads with local data residency requirements and satisfies demanding performance needs by keeping data close to on-premises applications. This module includes information about this storage class for your awareness, but this training does not discuss this storage class further.

See the link titled Amazon S3 Storage Classes in the course resources page for more information.

## S3 storage classes breakdown

	S3 Standard	S3 Intelligent-Tiering	S3 Standard-IA	S3 One Zone-IA	S3 Glacier Instant Retrieval	S3 Glacier Flexible Retrieval	S3 Glacier Deep Archive
Availability Zones	≥3	≥3	≥3	1	≥3	≥3	≥3
Minimum capacity charge for each object	N/A	N/A	128 KB	128 KB	128 KB	N/A	N/A
Minimum storage duration charge	N/A	N/A	30 days	30 days	90 days	90 days	180 days
Retrieval charge	N/A	N/A	Per GB retrieved	Per GB retrieved	Per GB retrieved	Per GB retrieved	Per GB retrieved



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

33

All storage class are designed for the following:

- Durability at 99.99999999 percent (11 nines)
- Availability at 99.9 percent with S3 One Zone-IA at 99.5 percent
- Available service-level agreement (SLA) at 99 percent
- Millisecond latency for the first byte of data
- Object storage type
- Lifecycle transitions

Some of the key differences include the following:

- The number of Availability Zones is greater than or equal to three for all storage classes except for S3 One Zone-IA, which is one Availability Zone.
- The minimum capacity charge for each object is 128 KB for S3 Standard-IA, S3 One Zone-IA, and S3 Glacier Instant Retrieval.
- The following are the minimum storage duration charges:
  - 30 days: S3 Standard-IA and S3 One Zone-IA
  - 90 days: S3 Glacier Instant Retrieval and S3 Glacier Flexible Retrieval
  - 180 days: S3 Glacier Deep Archive
- The retrieval charge is for each gigabyte retrieved for all storage classes except S3 Standard and S3 Intelligent-Tiering, which have no retrieval charge.

There may be a reason to have an object stored in a different class for different reasons, which leads to the lifecycle policy.

## Configuring an Amazon S3 Lifecycle

Amazon S3 lifecycle configurations are a set of rules that define actions that Amazon S3 applies to a group of objects.

- Transition actions transition to another storage class.
- Expiration actions define when objects expire.



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

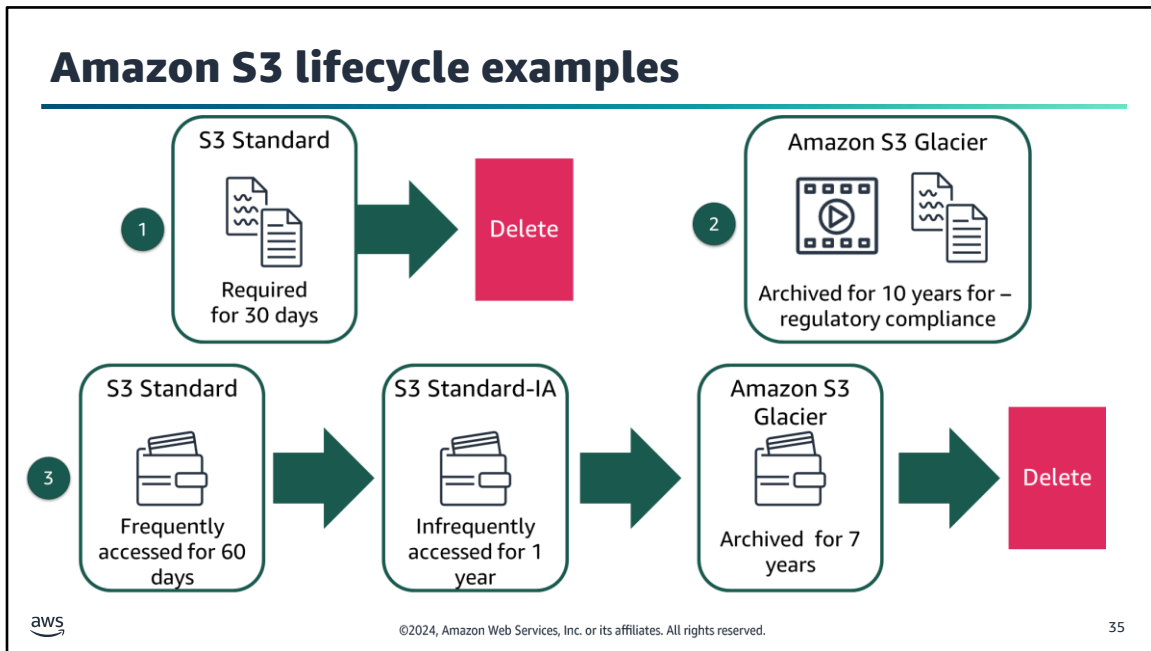
34

An S3 lifecycle configuration is a set of rules that defines the actions that Amazon S3 applies to a group of objects. There are two types of actions:

- **Transition actions:** These actions define when objects transition to another storage class. For example, you might choose to transition objects to the S3 Standard-IA storage class 30 days after creating them or to archive objects to the S3 Glacier Flexible Retrieval storage class 1 year after creating them. There are costs associated with lifecycle transition requests. For more information about Amazon S3 pricing, see the Amazon S3 pricing page linked in course resources.
- **Expiration actions:** These actions define when objects expire. Amazon S3 deletes expired objects on your behalf. Lifecycle expiration costs depend on when you choose for objects to expire. For more information, see the Expiring Objects page in the Amazon S3 User Guide. Your course resources include a direct link.

After an S3 lifecycle policy is set, your data will automatically transfer to a different storage class without any changes to your application.

By using lifecycle policies, you can cycle data at regular intervals among different Amazon S3 storage types. This cycling reduces your overall cost because you pay less for data as it becomes less important over time. In addition to being able to set lifecycle rules per object, you can also set lifecycle rules per bucket.


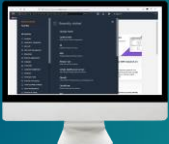


Define S3 lifecycle configuration rules for objects that have a well-defined lifecycle. The following are examples:

1. If you upload periodic logs to a bucket, your application might need them for a week or a month. After that, you might want to delete them.
2. You might upload some types of data to Amazon S3 primarily for archival purposes. For example, you might archive digital media, financial and healthcare records, raw genomics sequence data, long-term database backups, and data that must be retained for regulatory compliance.
3. Some documents are frequently accessed for a limited period of time. After that, they are infrequently accessed. At some point, you might not need real-time access to them, but your organization or regulations might require you to archive them for a specific period. After that, you can delete them.

With S3 Lifecycle configuration rules, you can tell Amazon S3 to transition objects to less-expensive storage classes, archive objects, or delete objects.

## Demo: Managing Lifecycles in Amazon S3



- This demo uses Amazon S3.
- In this demonstration, you will see how to create a lifecycle rule for Amazon S3 and apply transition actions to objects.

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

36

Find this recorded demo in your course as part of this module.

Amazon S3 versioning		
Amazon S3 versioning protects objects from accidental overwrites and deletes.		
Action	Versioning Enabled	Versioning Disabled or Versioning Suspended
Upload an object with the same key	Creates a new object with a different version ID, and both are retrievable by the version ID.	Overwrites the original object, and the previous object is no longer retrievable.
Delete	Adds a delete marker, but the object is still retrievable by the version ID.	Deletes the object, and it is no longer retrievable.

Amazon S3 provides a versioning feature to protect objects from accidental overwrites and deletes. You can use versioning to recover from both unintended user actions and application failures.

You enable versioning at the bucket level. Each object in a bucket has a version ID, and when versioning is disabled, the bucket's value is set to null. When versioning is enabled, Amazon S3 creates a new object and assigns a unique value to its version ID (increments it) every time it is uploaded.

By default, S3 Versioning is disabled on buckets, and you must explicitly enable it. When versioning is not enabled, the ID of the object is null.

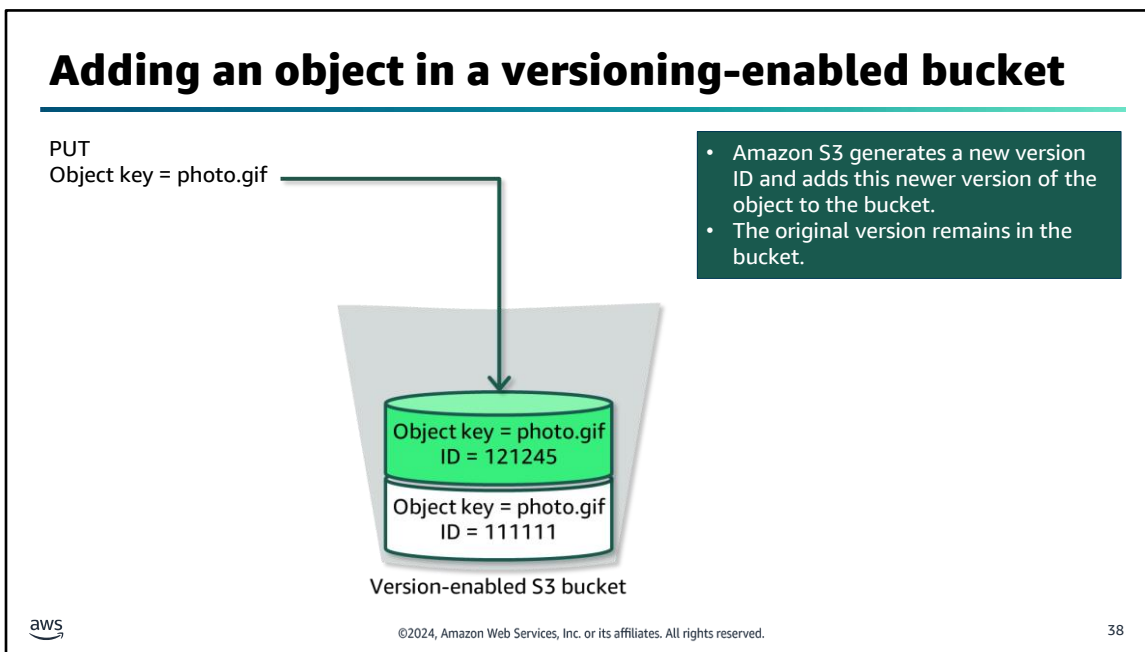
An S3 bucket can be in one of three versioning states:

- **Versioning enabled:** If you upload an object with the same key, Amazon S3 keeps the old object and creates a new object with a new version ID. If you delete an object, Amazon S3 logically deletes it and adds a marker, but the old version is retrievable by its version ID.
- **Versioning disabled:** This is the default setting. If you upload an object to a bucket with the same key, it overwrites the old version. If you delete an object, it is permanently deleted.
- **Versioning suspended:** The versions of existing objects are maintained, but the bucket temporarily behaves as if versioning were disabled.

When versioning has been enabled on a bucket, it cannot be disabled and can only be suspended.

There is no cost for using versioning, but because each version is a copy of an object, each version contributes to storage costs.

For more information about versioning in S3 buckets, see the page [Using versioning in S3 buckets](#) in the Amazon S3 user guide. Your course resources include a direct link.

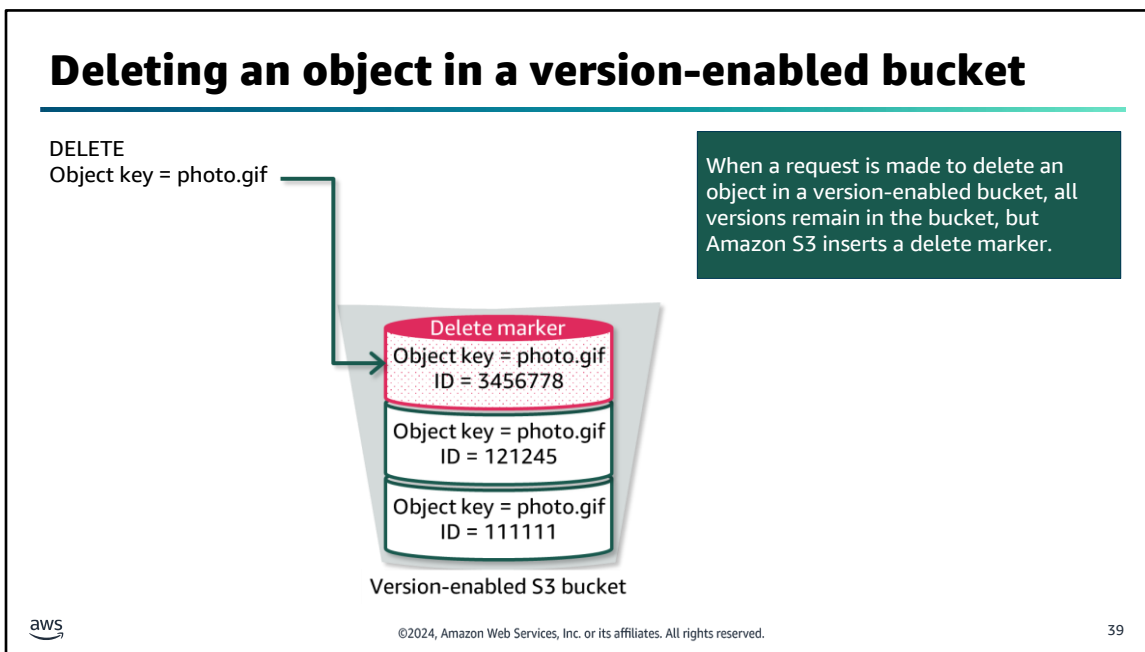


Now you will look at some examples of how versioning works.

When you use a PUT request to add an object into a versioning-enabled bucket, the noncurrent version is not overwritten. As the diagram shows, when a new version of photo.gif is PUT into a bucket that already contains an object with the same name, the following behavior occurs:

The original object (ID = 111111) remains in the bucket.

Amazon S3 generates a new version ID (121212) and adds this newer version of the object to the bucket.



With versioning enabled, you can retrieve a previous version of an object if an object has been accidentally overwritten or deleted.

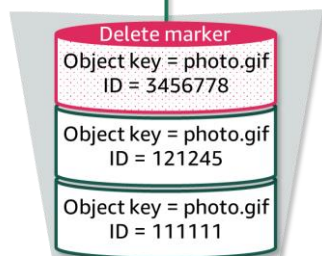
When you use the DELETE command to remove an object, all versions remain in the bucket, and Amazon S3 inserts a delete marker, as the diagram shows.

In this example, you can still retrieve the prior versions with the ID 121212 or 111111.

## Retrieving the most recently stored version

GET  
Object key = photo.gif

404 no object found



Version-enabled S3 bucket

- Requests for an object key return the most recent version.
- If the most recent version is a delete marker, the request is not successful.



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

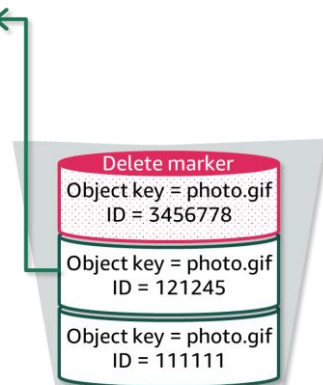
40

The delete marker becomes the current version of the object. By default, GET requests retrieve the most recently stored version. Performing a GET object request when the current version is a delete marker returns a 404 Not Found error, as the diagram shows.

## Retrieving an object with its specific ID

GET  
Object key = photo.gif  
with version ID = 121245

Requests for an object with its version ID will successfully return that version of the object.



Version-enabled S3 bucket



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

41

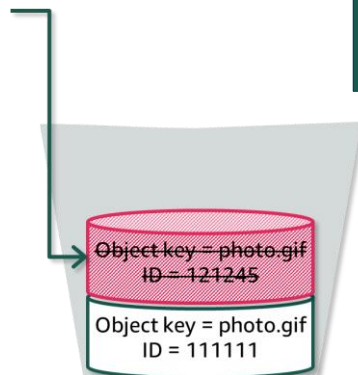
However, you can GET a noncurrent version of an object by specifying its version ID. In the diagram, you use a GET request for a specific object version, 111111. Amazon S3 returns that object version even though it's not the current version.

## Permanently delete an object

DELETE

Object key = photo.gif

With version id = 121245



Version-enabled S3 bucket

- Owners of the bucket can permanently delete an object by using delete with the version ID.
- In this case, no delete marker is added, and the specified version is not recoverable.




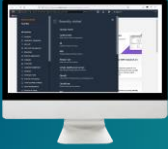
©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

42

Lastly, you can permanently delete an object by specifying the version that you want to delete. Only the owner of an S3 bucket can permanently delete a version.

If your delete operation specifies the version ID, that object version is permanently deleted, and Amazon S3 doesn't insert a delete marker.

## Demo: Amazon S3 Versioning

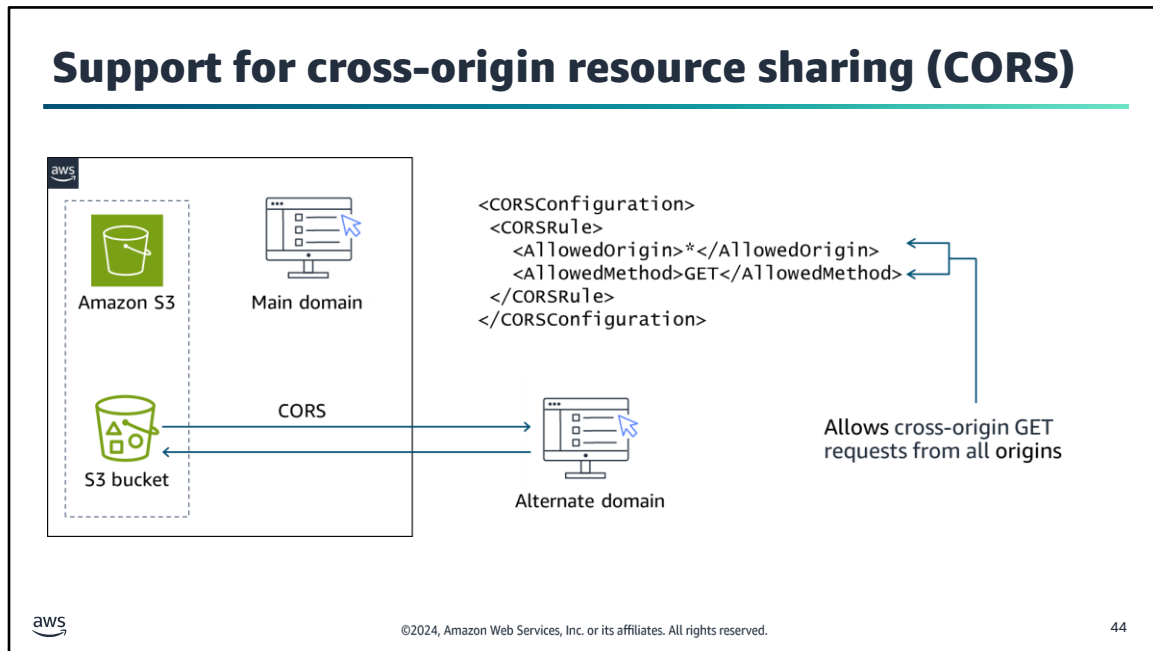


- This demo uses Amazon S3.
- In this demonstration, you will see how to enable Amazon S3 versioning on an S3 bucket.

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

43

Find this recorded demo in your course as part of this module.



Cross-origin resource sharing (CORS) defines a way for client web applications that are loaded in one domain to interact with resources in a different domain. With CORS support, you can build rich client-side web applications with Amazon S3 and selectively allow cross-origin access to your Amazon S3 resources.

To configure your bucket to allow cross-origin requests, you create a CORS configuration. A CORS configuration is an XML document with rules that identify the following

- The origins that you will allow to access your bucket.
- The operations (HTTP methods) that will support for each origin. In this example, GET requests are allowed from all origins.
- Other operation-specific information.

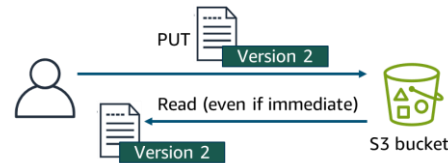
When Amazon S3 receives a preflight request from a browser, it evaluates the CORS configuration for the bucket and uses a CORS rule that matches the incoming browser request to allow a cross-origin request.

This slide shows one example: You want to host a web font in your S3 bucket. A webpage in an alternate domain may try to use this web font. Before the browser loads this webpage, it performs a CORS check to make sure that the domain from which the page is being loaded is allowed to access S3 bucket resources.

For more information about using cross-origin resource sharing (CORS), see the CORS page in the Amazon S3 user guide. Your course resources include a direct link.

## Amazon S3 data consistency model

- Is consistent for all new and existing objects in all Regions
- Provides read-after-write consistency for all GET, LIST, PUT, and DELETE operations on objects in S3 buckets
- Offers an advantage for big data workloads
- Simplifies the migration of on-premises analytics workloads



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

45

Along with versioning, consistency of data is important. Now you review the Amazon S3 consistency model.

**Image description:** Graphic shows a new object uploaded to S3 (as version 1). Any immediate read back of that object will consistently be v1 (which confirms read-after-write consistency for PUTS of new objects). **End description**

Amazon S3 pioneered object storage in the cloud with high availability, performance, and virtually unlimited scalability, with eventual consistency. Without strong consistency, you would insert custom code into these applications or provision databases to keep objects consistent with any changes in Amazon S3 across millions or billions of objects.

Amazon S3 delivers strong read-after-write and list consistency automatically for all applications. With strong consistency, Amazon S3 simplifies the migration of on-premises analytics workloads by removing the need to make changes to applications. Amazon S3 also reduces costs by removing the need for extra infrastructure to provide strong consistency.

Amazon S3 is strongly consistent for all new and existing S3 objects in all AWS Regions.



Amazon S3 achieves high availability by replicating data across multiple servers within AWS data centers. If a PUT request is successful, the data is safely stored. Any read (GET or LIST) that is initiated following a successful PUT response will return the data written by the PUT. This strong read-after-write consistency exists automatically for all applications, without changes to performance or availability.

Strong consistency simplifies the migration of on-premises analytics workloads by removing the need to make changes to support applications. It also removes the need for extra infrastructure, such as S3Guard, to provide strong consistency.

Although objects are strongly consistent, S3 bucket configurations have an eventual consistency model. For example, if you delete a bucket and immediately list all buckets, the deleted bucket might still appear in the list. However, within a short period of time, if you run the list bucket command again, the deleted bucket will no longer appear in the results for the list bucket.

See the link provided on the course resources page for more information on Amazon S3 strong consistency.

## Key takeaways: Storing content with Amazon S3



- S3 Standard storage is appropriate for cloud applications, dynamic websites, content distribution, mobile and gaming applications, and big data analytics.
- Set an S3 lifecycle policy and your data will automatically transfer to a different storage class without any changes to your application.
- Recover from both unintended user actions and application failures by using versioning.
- CORS defines a way for client web applications that are loaded in one domain to interact with resources in a different domain.
- The Amazon S3 data consistency model simplifies the migration of on-premises analytics workloads by removing the need to make changes to support applications.

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

46

Here are a few key points to summarize this section.



This section describes designing with Amazon S3.

## Amazon S3 default security configurations

- S3 buckets and objects created are private and protected by default.
- S3 buckets have encryption configured by default.
- Server-side encryption with Amazon S3 managed keys (SSE-S3) is the default encryption.

When use cases must share Amazon S3 data, do the following:

- Manage and control the data access.
- Follow the principle of least privilege.



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

48

When your objective is to protect digital data, data encryption is an essential tool. Data encryption takes data that is legible and encodes it. Encrypted data is unreadable to anyone who does not have access to the secret key that can be used to decode it. Thus, even if an attacker gains access to your data, they cannot make sense of it. Optionally, use AWS Key Management Service (AWS KMS) to manage secret keys.

Protecting data is very important. By default, all S3 buckets are private, and only users who are explicitly granted access to S3 buckets can access those buckets. It is essential that you manage and control access to Amazon S3 data.

For more information, see [Data protection in Amazon S3](#) on the content resources page of your online course.

## Encrypting objects in Amazon S3

Encryption encodes data with a secret key, which makes it unreadable without a key.

- Server-side encryption
  - Amazon S3 encrypts objects before it saves the objects to disk and decrypts the objects when you download them.
  - Enable this feature by selecting the default encryption option on the bucket.
- Client-side encryption
  - Encrypt data on the client side and upload the encrypted data to Amazon S3.
  - In this case, you manage the encryption process.



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

49

All S3 buckets have encryption configured by default, and all new objects that are uploaded to an S3 bucket are automatically encrypted at rest. SSE-S3 is the default encryption configuration for every bucket in Amazon S3. To use a different type of encryption, you can either specify the type of server-side encryption to use in your S3 PUT requests, or you can set the default encryption configuration in the destination bucket.

If you want to specify a different encryption type in your PUT requests, you can use server-side encryption with AWS Key Management Service (AWS KMS) keys (SSE-KMS), dual-layer server-side encryption with AWS KMS keys (DSSE-KMS), or server-side encryption with customer-provided keys (SSE-C).

If you want to set a different default encryption configuration in the destination bucket, you can use SSE-KMS or DSSE-KMS.

Client-side encryption is another option. When you use this approach, you encrypt the data on the client side before you upload it to Amazon S3. In this case, you manage the encryption process, the encryption keys, and related tools. Like server-side encryption, client-side encryption can reduce risk by encrypting the data with a key that is stored in a different mechanism than the mechanism that stores the data itself.

For more information, see [Protecting data with encryption](#) on the content resources page of your online course.

## Amazon S3 tools for protecting buckets and objects

Tool	Description
Block Public Access feature	Makes buckets inaccessible to the public
AWS Identity and Access Management (IAM) policies	Authenticates users by using IAM
Bucket policies	Defines access based on specific written rules
Access control lists (ACLs)	Sets rules for access to buckets and objects (bucket policies are the preferred method for controlling bucket access)
Amazon S3 access points	Configures access with names and permissions specific to each application
Preassigned URLs	Grants time-limited access to others with temporary URLs
AWS Trusted Advisor	Provides a bucket permission check



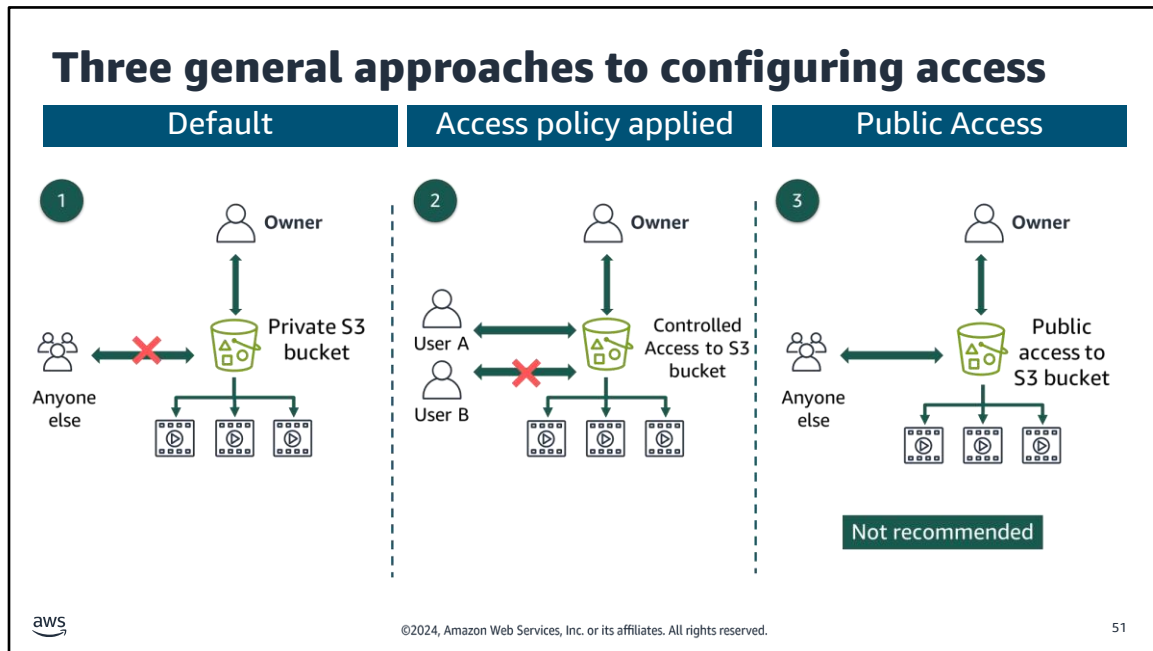
©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

50

AWS provides many tools and options for controlling access to your S3 buckets or objects, such as the following:

- The Amazon S3 Block Public Access feature overrides any other policies or object permissions. Enable Block Public Access for all buckets that you don't want to be publicly accessible. This feature provides a straightforward method for avoiding unintended exposure of Amazon S3 data.
- Write AWS Identity and Access Management (IAM) policies that specify the users or roles that can access specific buckets and objects.
- You typically use bucket policies to define access to specific buckets or objects when the user or system cannot authenticate by using IAM. You can configure bucket policies to grant access across AWS accounts or to grant public or anonymous access to Amazon S3 data. If you use bucket policies, you should carefully write them and fully test them. You can specify a deny statement in a bucket policy to restrict access. Access will be restricted even if the users have permissions that are granted in an identity-based policy that is attached to the users.
- Set access control lists (ACLs) on your buckets and objects that are less commonly used (ACLs predate IAM). If you use ACLs, do not set access that is too open or permissive.
- S3 access points are named network endpoints that are attached to buckets. You can use access endpoints to perform S3 object operations. Customers with shared datasets can scale access for many applications by creating individualized access points with names and permissions that are customized for each application.
- Preassigned URLs grant time-limited access to others with temporary URLs.
- AWS Trusted Advisor provides a bucket permission check feature. It is a useful tool for discovering if any of the buckets in your account have permissions that grant global access.

See the link titled Data protection in Amazon S3 in the course resources page for more information.



Here are three different general approaches to configuring access to objects in an S3 bucket:

- Option 1 shows the default security settings for Amazon S3. By default, all S3 buckets and the objects stored in them are private (protected). The only entities with access to a newly created, unmodified bucket are the account administrator and the AWS account root user. The resource owner can grant specific access permissions to others, but anyone without those permissions will not have access.
- Option 2 shows a case where Amazon S3 was configured to provide controlled access. User A was granted access to the objects in the bucket, but User B was denied access. Controlled access scenarios are common. The bucket owner can configure these scenarios by using one or more of the tools or options for controlling access to Amazon S3 data discussed earlier.
- Option 3 shows an occasion where S3 security settings have been disabled and anyone can publicly access the objects stored in the bucket.

### Not recommended

Using an S3 bucket to host a static website is an example of setting up an AWS architecture quickly. However, for most Amazon S3 use cases, you would not want to grant public access to Amazon S3. Most use cases do not require public access. More often, you use Amazon S3 to store data that is used by an application that runs outside of Amazon S3 or to back up sensitive data. For these common use cases, you should never grant public access to buckets that hold data.

## Considerations when choosing a Region

Considerations	Details
Data privacy laws and regulatory compliance	<ul style="list-style-type: none"><li>• Are there relevant Region data privacy laws?</li><li>• Can customer data be stored outside the country?</li><li>• Can you meet your governance obligation?</li></ul>
Proximity of users to data	<ul style="list-style-type: none"><li>• Small differences in latency can impact customer experience.</li><li>• Choose the Region closest to your users.</li></ul>
Availability service and feature	<ul style="list-style-type: none"><li>• Not all AWS services are available in all Regions.</li><li>• Services expand to new Regions regularly.</li><li>• Use some services cross-Region but at increased latency.</li></ul>
Cost-effectiveness	<ul style="list-style-type: none"><li>• Costs vary by Region.</li><li>• Some services such as Amazon S3 have costs for transferring data out.</li><li>• Consider the cost-effectiveness of replicating the entire environment in another Region.</li></ul>



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

52

Now that you have learned about protecting and encrypting your data, there are many considerations when deciding what Region to host your data in:

- You should consider data privacy laws and your regulatory compliance requirements. Data that you store on AWS is subject to the laws of the country and locality where it is stored. In addition, some laws dictate that if you are operating your business in their jurisdiction, you cannot store that data anywhere else. Similarly, compliance standards (such as the US Health Insurance Portability and Accountability Act, or HIPAA) have strict guidelines on how and where data can be stored.
- Proximity is an important factor when choosing your Region, especially when latency is a critical factor. In most cases, the latency difference between using the closest Region and the farthest Region is relatively small, but even small differences in latency can impact customer experience. Customers expect responsive environments, and as time passes and technology becomes more and more powerful, those expectations also rise.
- Availability of AWS services and features is an important consideration. Though AWS strives to make services and features available everywhere, the complications that arise from having a global reach make it challenging to accomplish that goal. Instead of waiting until a service is available everywhere before launching it, services are released when they are ready. Service availability is then expanded as soon as possible.
- Cost is an important consideration when choosing a Region. Service costs can differ depending on which Region they are used in. For example, an EC2 instance in the us-east-1 Region might not cost the same as if it ran in the eu-west-1 Region. Typically, the difference in cost might not be enough to supersede the other three considerations. However, in cases where the latency, compliance, and service availability differences between Regions are minimal, you might be able to save by using the lower-cost Region for your environment.

- Finally, in circumstances where your customers are in different areas of the world, consider optimizing their experience by replicating your environment in multiple Regions that are closer to them. Because you would then be distributing your load across multiple environments, your costs for components in each environment might go down even as you add more infrastructure. For example, by adding a second application environment, you might be able to cut your processing and storage capacity requirements in half in each environment. Because AWS is designed to provide that kind of flexibility and because you pay for only what you use, you could scale your existing environment down as a way to mitigate the cost of adding another environment. The downside to this approach is that you now have two environments to manage. Also, not all of your components will scale down enough to mitigate all the costs of the new components. Additionally, you might need to maintain one single storage source of truth in one Region, such as a primary Amazon Relational Database Service (Amazon RDS) instance. Your secondary Region would need to communicate with the storage instance, which might increase latency and cost for those operations.

See the link titled [AWS Services by Region](#) in the course resources page for more information.

## Amazon S3 Inventory

- Use Amazon S3 Inventory to help manage your storage.
- Use it to audit and report on the replication and encryption status of your objects for business, compliance, and regulatory needs.
- Speed up business workflows and big data jobs by using Amazon S3 Inventory.
- Provide a scheduled alternative to the Amazon S3 synchronous List API operations.



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

53

Now that you understand the importance of encrypting objects, configuring access, and choosing a Region, this next feature, Amazon S3 Inventory, helps you manage your storage.

For example, you can use it to audit and report on the replication and encryption status of your objects for business, compliance, and regulatory needs. You can also simplify and speed up business workflows and big data jobs by using Amazon S3 Inventory, which provides a scheduled alternative to the Amazon S3 synchronous List API operations.

Amazon S3 Inventory provides comma-separated values (CSV), Apache optimized row columnar (ORC) files, or Apache Parquet output files that list your objects and their corresponding metadata on a daily or weekly basis for an S3 bucket or objects with a shared prefix (that is, objects that have names that begin with a common string). If you set up a weekly inventory, a report is generated every Sunday (UTC time zone) after the initial report.

When you're configuring an inventory list, you can specify the following:

- What object metadata to include in the inventory
- Whether to list all object versions or only current versions
- Where to store the inventory list file output
- Whether to generate the inventory on a daily or weekly basis
- Whether to encrypt the inventory list file

You can query Amazon S3 Inventory with standard SQL queries by using Amazon Athena, Amazon Redshift Spectrum, and other tools, such as Presto, Apache Hive, and Apache Spark.

The bucket that the inventory lists objects for is called the *source bucket*. The bucket where the inventory list file is stored is called the *destination bucket*.

For more information, see Amazon S3 Inventory on the content resources page of your online course.

## Amazon S3 costs

### Pay for only what you use:

- Gigabytes of objects stored (per month) with different pricing for each Region and each storage class
- PUT, COPY, POST, LIST or lifecycle transition to move data into any Amazon S3 storage class

### No charge for data transferred:

- Out to the internet for the first 100 GB per month
- In from the internet
- Between S3 buckets or to any service in the same AWS Region
- From an S3 bucket to any AWS service or services within the same AWS Region as the S3 bucket
- Out to CloudFront



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

54

With Amazon S3, you pay for only what you use. There is no minimum fee. You pay for storing objects in your S3 buckets. The rate you're charged depends on the size of your object, the amount of time that you stored the objects during the month, and the storage class. You pay a monthly monitoring and automation charge for each object stored in the S3 Intelligent-Tiering storage class to monitor access patterns and move objects between access tiers. In S3 Intelligent-Tiering, there are no retrieval charges, and no additional tiering charges apply when objects are moved between access tiers.

For more information see AWS Pricing Calculator on the content resources page of your online course.

There are ingest charges for each request when using PUT, COPY, POST, or LIST requests or when using lifecycle rules to move data into any Amazon S3 storage class.

For more information see Requests & data retrievals on the content resources page of your online course.

### Encryption

Amazon S3 automatically applies SSE-S3 as a base layer of encryption to all new objects added to Amazon S3 at no additional cost and with no impact on performance. SSE-C also does not incur any additional Amazon S3 charges. For SSE-KMS, you pay AWS KMS charges to generate or retrieve the data key used for encryption and decryption. For DSSE-KMS, in addition to the charges for AWS KMS, you pay an additional encryption fee for each gigabyte for the second layer of encryption and decryption of data.

For more information see S3 encryption on the content resources page of your online course.

### AWS Free Tier

As part of the AWS Free Tier, you can get started with Amazon S3 for free. Upon sign-up, new AWS customers receive 5 GB of Amazon S3 storage in the S3 Standard storage class; 20,000 GET requests; 2,000 PUT, COPY, POST, or LIST requests; and 100 GB of data transfer out each month. Your usage for the free tier is calculated each month across all AWS Regions except the AWS GovCloud Region and automatically applied to your bill; unused monthly usage will not roll over.

## Activity: Designing with Amazon S3



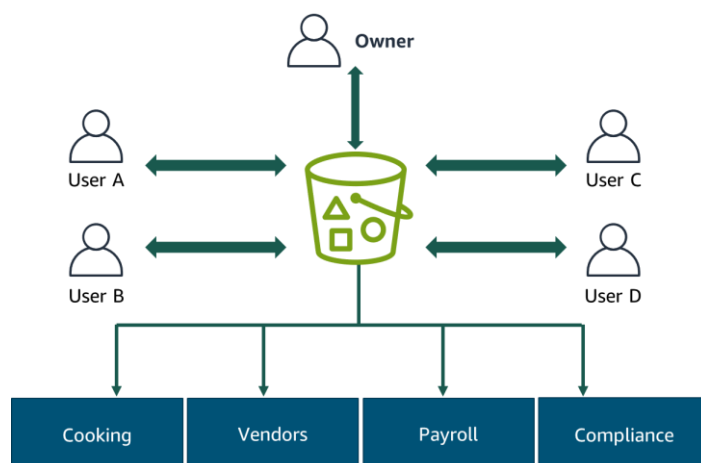
- The café owner has successfully created a static website but now wants to make documents more accessible to employees. They would like to house content (cooking instructions, vendor information, payroll, and compliance training) that only employees can access.
- How could you set this up to meet the owner's needs?

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

55

Read the activity and discuss how you might set this up.

## Activity solution: Created controlled access for each user



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

56

Because you do not want to set this up for public access, setting up controlled access meets the customer's need. First, create a folder for each type of content (cooking, vendors, payroll, and compliance). Then create IAM policies (group or individual) to allow or deny users access to specific information based on their employee level at the café.

## Key takeaways: Designing with Amazon S3

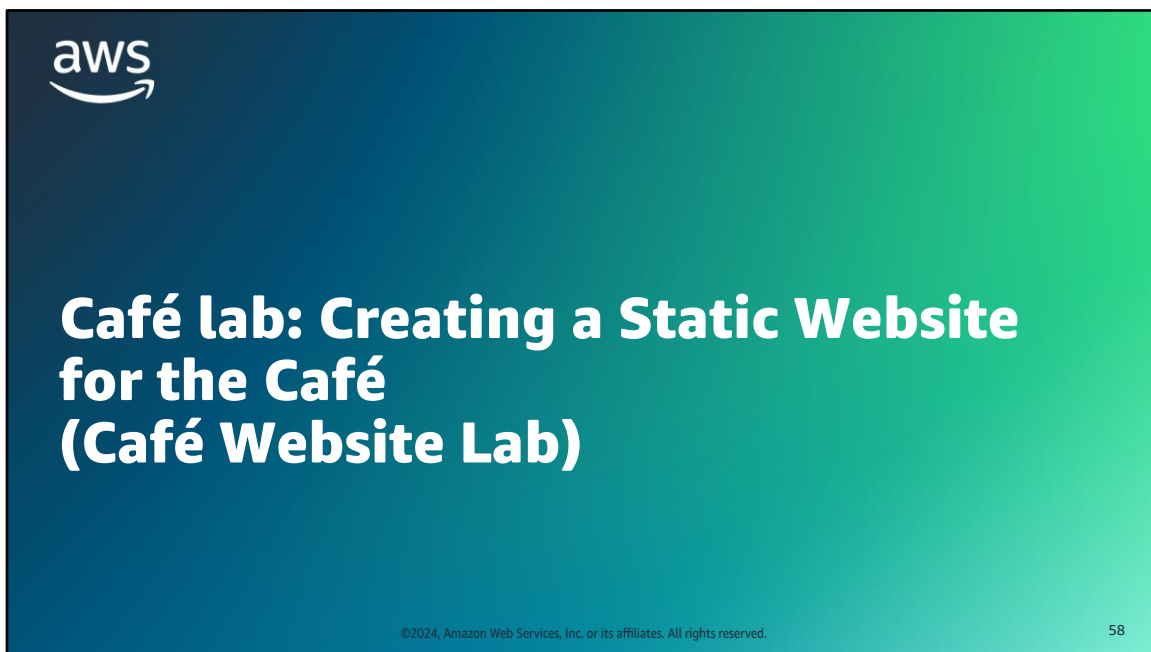


- S3 buckets are private and can be accessed only by users who are explicitly granted access.
- SSE-S3 is the default encryption configuration for every bucket in Amazon S3.
- Considerations when choosing a Region include data privacy laws and regulatory compliance, proximity of users to data, availability service and features, and cost-effectiveness.
- Amazon costs are based on your objects' size, the amount of time you stored the objects during the month, and the storage class.

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

57

Here are a few key points to summarize this section.



You will now complete a lab. The next slides summarize what you will do in the lab, and you will find the detailed instructions in the lab environment.

## The evolving café architecture: version 1

Architecture Version	Business Reason for the Update	Technical Requirements and Architecture Update
V1	Create a static website for a small business.	Host the website on Amazon S3.
V2	Add online ordering.	Deploy the web application and database on Amazon EC2.
V3	Reduce the effort to maintain the database and secure its data.	Separate web and database layers. Migrate the database to Amazon Relational Database Service (Amazon RDS) on a private subnet.
V4	Enhance the security of the web application.	Use Amazon Virtual Private Cloud (Amazon VPC) features to configure and secure public and private subnets.
V5	Create separate access mechanisms based on role.	Add IAM groups and attach resource policies to application resources. Add IAM users to groups based on role.
V6	Help ensure that the website can handle an expected increase in traffic.	Add a load balancer, implement auto scaling on the EC2 instances, and distribute compute and database instances across two Availability Zones.





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

59

Sofía mentioned to Nikhil that she would like the café to have a website that will showcase the café visually through images. The website should also provide customers with business details, such as the location of the store, the business hours, and telephone number.

Nikhil is pleased to create the first website for the café. During this activity, you will take on the role of Nikhil and work on producing the results that everyone at the café hopes that you can deliver. Perhaps you can even exceed their expectations!

## Café Website Lab tasks



- In this lab, you will do the following:
- Create an S3 bucket to host your static website and upload content to your S3 bucket.
- Create a bucket policy to grant public read access.
- Enable versioning on the S3 bucket.
- Set lifecycle policies.
- Enable cross-Region replication.

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

60

Access the lab environment through your online course to get additional details and complete the lab.

## Debrief: Café Website Lab

---

- How did you protect the café website from accidental overwrite and deletion?
- Which strategy did you use to help save on costs as the size of the website grows through the use of versioning?



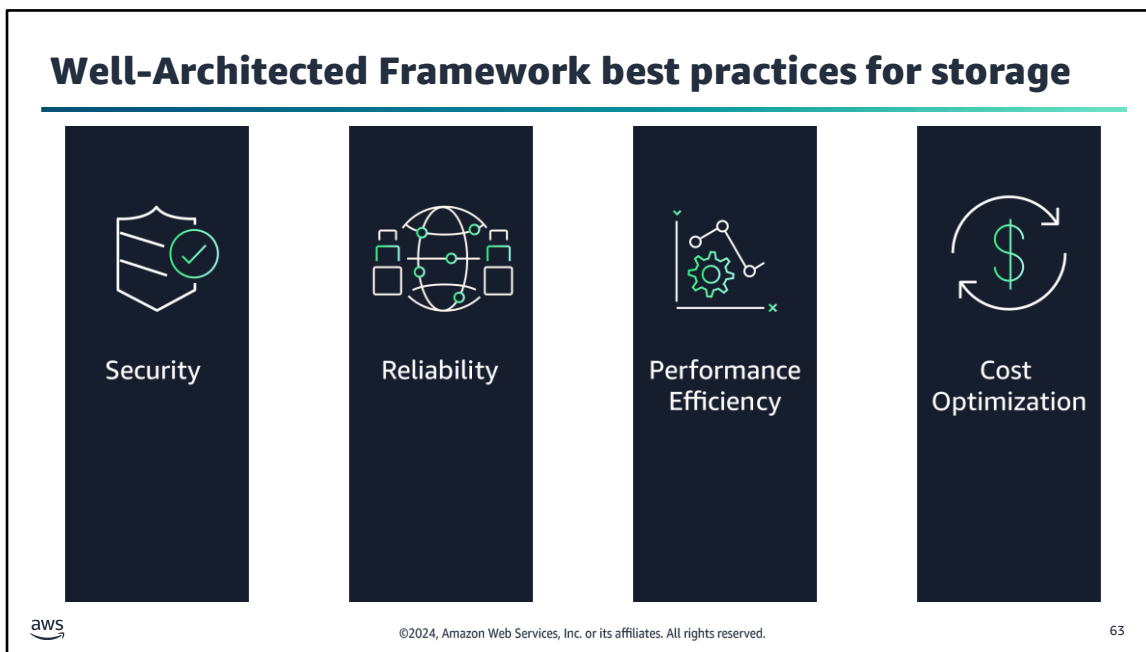


# Applying the AWS Well-Architected Framework principles to storage

## Adding a Storage Layer with Amazon S3

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

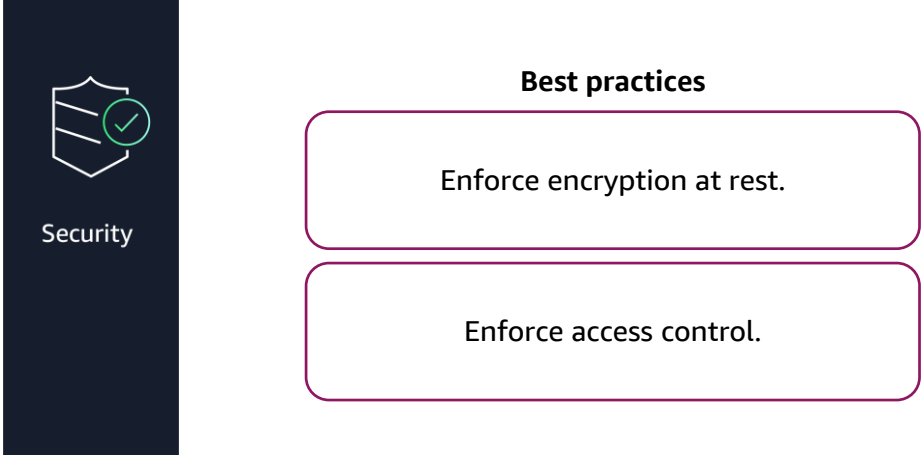
62



The AWS Well-Architected Framework has six pillars, and each pillar includes best practices and a set of questions that you should consider when you architect cloud solutions. This section highlights a few best practices from the pillars that are most relevant to this module. This includes Security, Reliability, Performance Efficiency, and Cost Optimization.

For more information, see the AWS Well-Architected Framework on the content resources page of your online course.

## Best practice approach: Data protection – protecting data at rest



The diagram illustrates the best practice approach for data protection at rest. On the left, a dark blue vertical bar contains a shield icon with a green checkmark and the word "Security". To the right, under the heading "Best practices", are two rounded rectangular boxes: "Enforce encryption at rest." and "Enforce access control.".

**Security**

**Best practices**

- Enforce encryption at rest.
- Enforce access control.

aws

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

64

Data protection is an important best practice under the security pillar. Before architecting any workload, foundational practices that influence security should be in place. The methods described for protecting data are important because they support objectives such as preventing mishandling or complying with regulatory obligations. Protecting data at rest is an important part of protecting data. Two important best practices that relate to architecting with Amazon S3 as your storage layer are enforcing encryption at rest and enforcing access control.

**Enforce encryption at rest:** Maintain the confidentiality of sensitive data in the event of unauthorized access or accidental disclosure. Private data should be encrypted by default when at rest. Data that is encrypted cannot be read or accessed without first unencrypting the data. Any data stored unencrypted should be inventoried and controlled.


**Enforce access control:** Use mechanisms such as isolation and versioning, and apply the principle of least privilege. Prevent the granting of public access to your data. Verify that only authorized users can access data on a need-to-know basis. Protect your data with regular backups and versioning to prevent against intentional or inadvertent modification or deletion of data. Isolate critical data from other data to protect its confidentiality and data integrity.

In this, module you've learned about a number of Amazon S3 features and configuration choices that support these best practices, including the following:

- S3 buckets and objects are encrypted by default, and you have options for how you manage the security keys for that encryption.
- Buckets and objects are private by default, so you must explicitly grant access. You should grant access to only those individuals and resources that require it. You can use IAM policies and bucket policies to limit access to your Amazon S3 resources by both user and resource.
- You can use versioning to protect objects in your S3 buckets against the loss of data due to accidental deletions or overwrites.
- The Block Public Access feature makes it difficult for you to inadvertently leave objects or buckets open for public access.

For more information, see the security pillar whitepaper which is linked in your course resources. Also see Data protection which is linked in your course resources.

## Best practice approach: Architecture selection



Performance Efficiency

### Best practices

- Learn about and understand available cloud services and features
- Factor cost into architectural decisions

aws

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

65

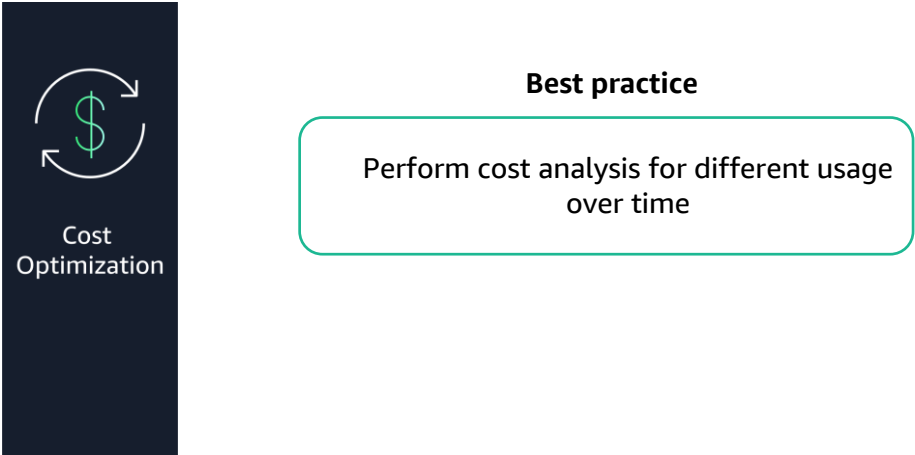
The optimal solution for a particular workload varies, and solutions often combine multiple approaches. Well-Architected workloads use multiple solutions and allow different features to improve performance. An important best practice for architecture selection is to understand what's available. It is also important to consider cost when selecting your storage.

Topics in this module that relate to these best practices include the following:

- Amazon S3 is a good choice for object storage and can store massive amounts of unstructured data.
- Available configurations that help with performance include S3 Transfer Acceleration for moving files over a long distance between a client and a bucket, and multipart upload, which improves throughput when uploading very large files.
- Amazon S3 provides different storage classes that you can use to store objects based on expected access patterns. S3 Intelligent-Tiering automatically moves objects based on access patterns to the storage tier that is most cost-effective.

For more information, see the performance efficiency pillar whitepaper linked in your course resources.

## Best practice approach: Cost-effective resources – evaluate cost when selecting services



Cost Optimization

**Best practice**

Perform cost analysis for different usage over time

aws

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

66


Cost-effective resources is an important best practice under the cost optimization pillar. Using the appropriate services, resources, and configurations for your workloads is key to cost savings. Workloads can change over time. Some services or features are more cost effective at different usage levels. By performing the analysis on each component over time and at projected usage, the workload remains cost-effective over its lifetime.

Topics in this module that relate to these best practices include the following:

- Amazon S3 provides different storage classes, and you can create lifecycle rules to automatically move data to a more cost-effective class. S3 Intelligent-Tiering automatically moves objects based on access patterns to the storage tier that is most cost-effective. Note how these features support both performance and cost-optimization best practices.
- Use Amazon S3 Inventory to audit how Amazon S3 is being used to help make cost-effective choices about how your organization is using Amazon S3.

For more information, see the cost optimization pillar whitepaper linked in your course resources.


## Best practice approach: Failure management - Use fault isolation to protect your workload



Reliability

### Best practice

Select the appropriate locations for your multi-location deployment



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

67

Failure management is an important best practice under the reliability pillar. Werner Vogels, CTO of Amazon.com, is credited with this quote: “Failures are a given, and everything will eventually fail over time.” For resilience, you should use an approach that builds layers of defense. For high availability, always (when possible) deploy your workload components to multiple Availability Zones (AZs).

In this module, you've learned how Amazon S3 is designed to withstand failures so that you as an architect do not have to think about incorporating these best practices. Here are some characteristics of Amazon S3 that support this best practice:

- Amazon S3 is designed for 11 nines of durability to help ensure that data is not lost and is designed for 4 nines of availability so that you can rely on your data being accessible when needed.
- Amazon S3 redundantly stores your objects on multiple Availability Zones in the Amazon S3 Region you designate.
- Amazon S3 regularly verifies the integrity of your data by using checksums.

For more information, see the reliability pillar whitepaper linked in your course resources.

**Key takeaways:  
Applying the AWS  
Well-Architected  
Framework  
principles to storage**

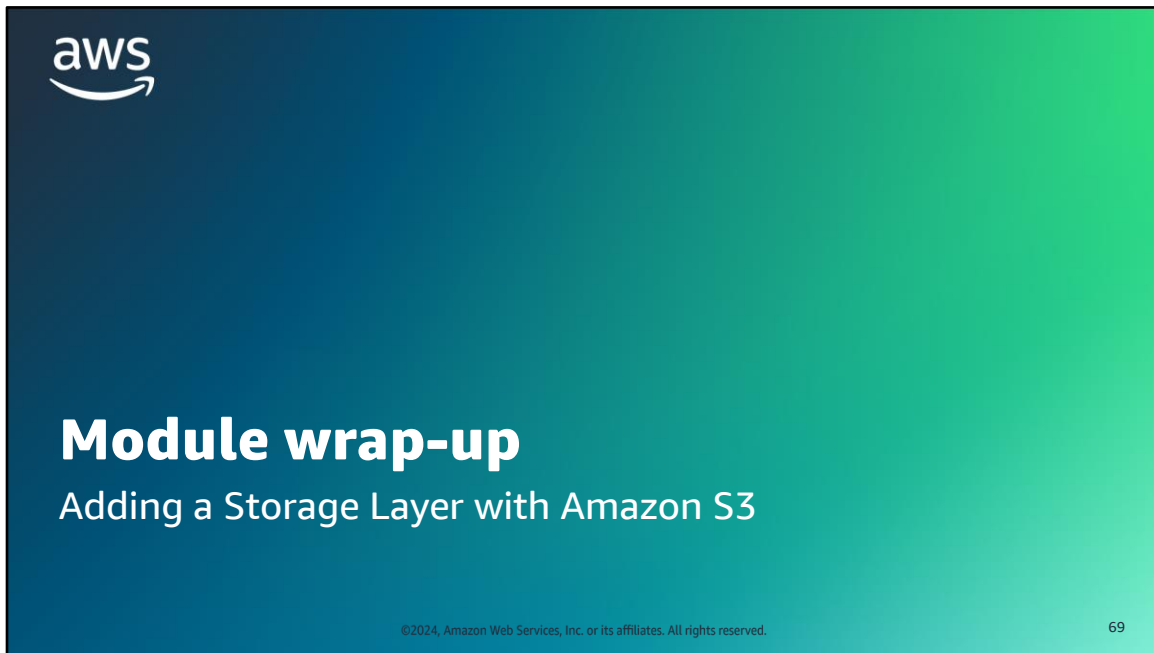


- Protecting data is a security best practice that Amazon S3 supports through these default configurations: encrypting objects, making objects private, blocking public access.
- You can protect data in Amazon S3 by limiting access through IAM policies and enabling versioning.
- Selecting an architecture is a performance efficiency best practice that Amazon S3 supports through its ability to store massive amounts of unstructured data.
- Amazon S3 includes performance-improving options such as S3 Transfer Acceleration and multipart upload.
- Selecting cost-effective resources is a cost-optimization best practice that Amazon S3 supports through features such as lifecycle policies, intelligent tiering, and Amazon S3 Inventory.
- Failure management is a reliability best practice that Amazon S3 has been designed for through its durability and availability features.
- You can use Amazon S3 for backing up data to improve failure management of your applications and data.

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

68

Here are a few key points to summarize this section.



This section summarizes what you have learned and brings the module to a close.

## Module summary

---

This module prepared you to do the following:

- Define Amazon S3 and how it works.
- Recognize the problems that Amazon S3 can solve.
- Describe how to move data to and from Amazon S3.
- Manage the storage of content efficiently by using Amazon S3.
- Recommend the appropriate use of Amazon S3 based on requirements.
- Configure a static website on Amazon S3.
- Use the Well-Architected Framework principles when designing a storage layer with Amazon S3.



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

70

## Considerations for the café

---



- Discuss how the café lab in this module addressed the cloud architect's key concerns presented at the start of this module.



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

71

## Module knowledge check



- The knowledge check is delivered online within your course.
- The knowledge check includes 10 questions based on material presented on the slides and in the slide notes.
- You can retake the knowledge check as many times as you like.

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

72

Use your online course to access the knowledge check for this module.

## Sample exam question

Company salespeople upload their sales figures daily to Amazon S3, but their solutions architect is concerned that users might accidentally delete or overwrite important documents.

Which action will protect against unintended user actions?

Identify the key words and phrases before continuing.

The following are the key words and phrases:

- Amazon S3
- Accidentally delete or overwrite
- Protect against unintended



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


73

## Sample exam question: Response choices

Company salespeople upload their sales figures daily to **Amazon S3**, but their solutions architect is concerned that users might **accidentally delete or overwrite** important documents.

Which action will **protect against unintended** user actions?

Choice	Response
A	Store data in two S3 buckets in different AWS Regions.
B	Enable versioning on the S3 bucket where files are stored.
C	Move uploaded data to an Amazon S3 Infrequent Access storage class at the end of each week.
D	Use Amazon S3 Inventory to audit the status of objects in the S3 bucket where files are stored.


 ©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved. 74

Use the key words that you identified on the previous slide, and review each of the possible responses to determine which one best addresses the question.

## Sample exam question: Answer

The answer is B.

Choice	Response
A	
B	Enable versioning on the S3 bucket where files are stored.
C	
D	

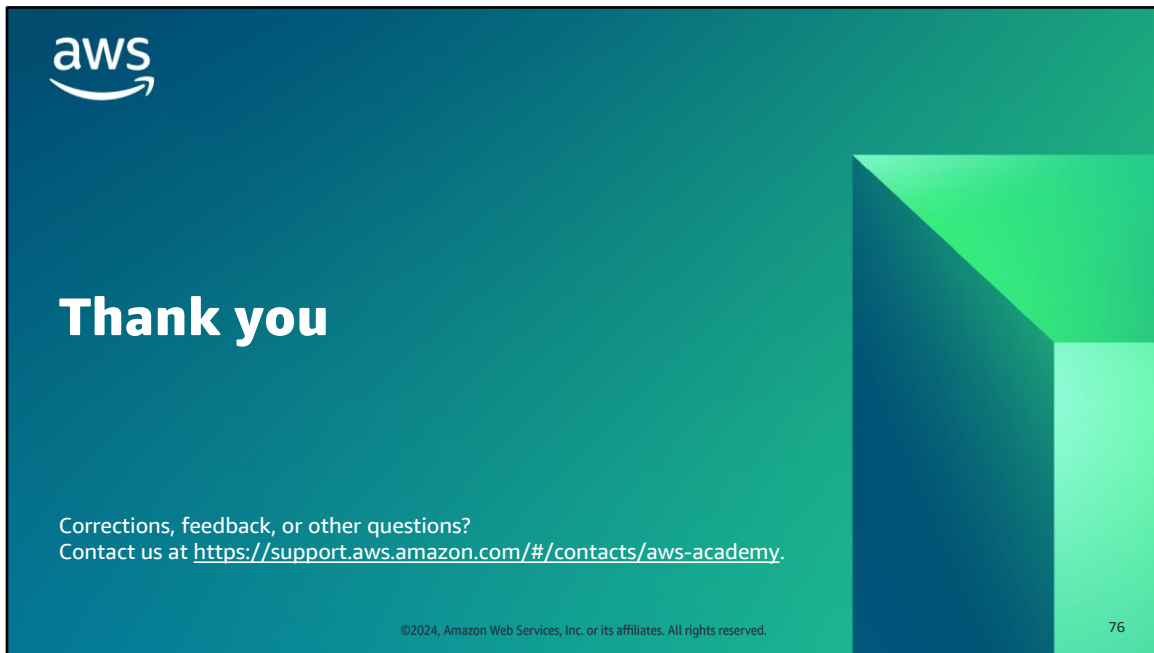
 ©2024, Amazon Web Services, Inc. or its affiliates. All rights reserved. 75

Choice A (Store data in two S3 buckets) would provide some limited protection against data loss due to an outage in one of the Regions, but it would not address the issue of someone accidentally deleting or overwriting a file.

Choice C (Move data to Amazon S3 Infrequent Access) would be helpful if the question asked about reducing costs for files that were not accessed often.

Choice D (Use Amazon S3 Inventory) might help the salespeople understand the activity on the objects in their bucket, but it wouldn't prevent files from being deleted.

Choice B (Enable versioning) would provide a means of protecting files from being deleted or overwritten. When versioning is enabled, the prior version of the file will still be available.



That concludes this module. The Content Resources page of your course includes links to additional resources that are related to this module.