

## Module 4: Adding a Storage Layer with Amazon S3

Monday, October 6, 2025 1:59 PM

### A. Defining Amazon S3:

- Types of storage:
  - + **Block Storage**: Data is stored on a device in fix-sized blocks
  - + **File Storage**: Data is stored in a hierarchical structure.
  - + **Object Storage**: Data is stored as objects based on attributes and metadata

**NOTE:** Image the data is the content of the book and metadata is the cover of the book

#### - Amazon S3:

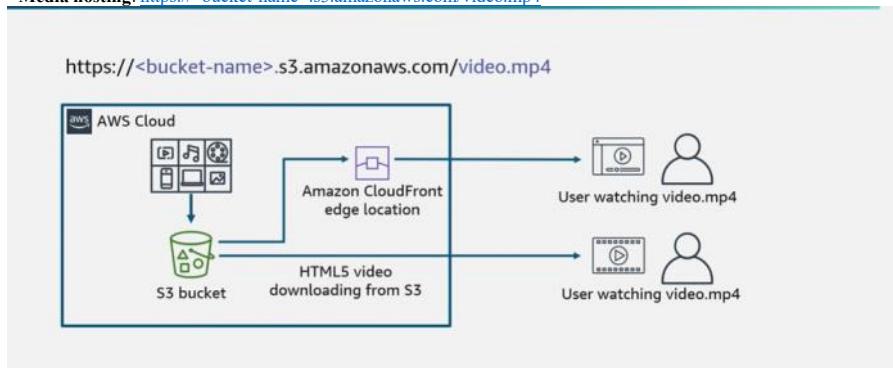
- + Amazon S3 stores **massive (unlimited)** amounts of unstructured data
- + Amazon S3 stores data file as objects in a bucket that you define
- + Five TB is the maximum file size of a single object
- + Objects have a globally unique URL (universal)
- + Objects have a key, version ID, value, metadata,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

- Prefixes to imply a folder structure in an S3 bucket
- Amazon S3 benefits:

Benefit	Description
Durability	<ul style="list-style-type: none"><li>• Helps ensure that <b>data is not lost</b></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>

### B. Using Amazon S3:

- **Spikes in demand**: Host web content that needs bandwidth to address extreme spikes in demand
- **Static site**: Host a static site that consists of HTML, files, images, and videos
- **Financial analysis**: Store data that other services can use for analysis.
- **Disaster recovery**: Support disaster recovery or data backup solutions.
- **Media hosting**: <https://<bucket-name>.s3.amazonaws.com/video.mp4>



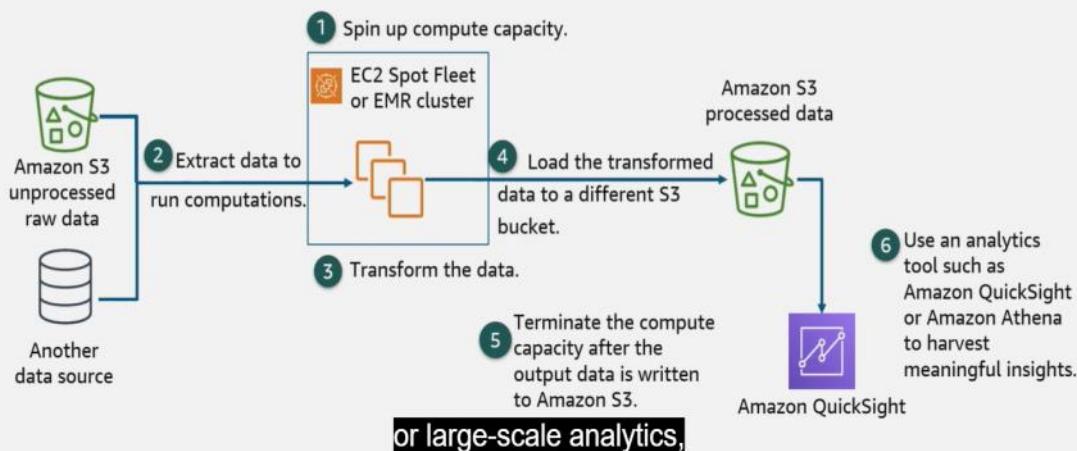
### Use case: Host static websites

The diagram shows the process of hosting a static website using an S3 bucket. On the left, there is a file structure for 'mybucket' containing 'css', 'images', 'js', 'videos', and 'index.html'. An 'Upload content' arrow points from this structure to an 'S3 bucket' icon. From the S3 bucket, an arrow points to a 'Website on S3' icon, which then points to an 'End user' icon. The URL 'http://<bucket-name>.s3-website.<Region>.amazonaws.com' is shown between the website icon and the end user.

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# Use case: Data store for computation and analytics

## Example data integration and preparation pattern



### C. Moving 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

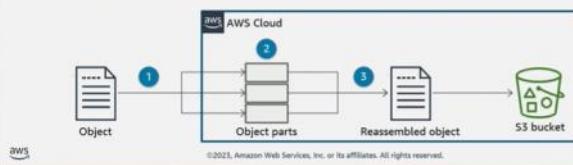
- Options for uploading objects to Amazon S3

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 (max is 160GB)
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

### Amazon S3 feature: Multipart upload

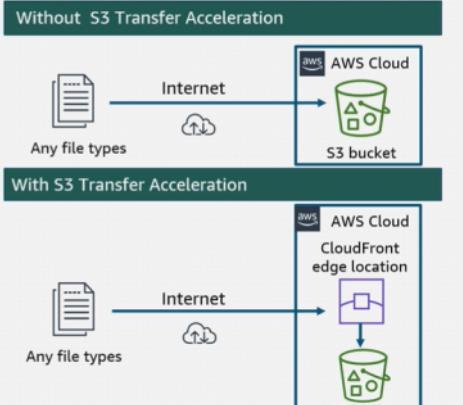
Multipart uploads have the following advantages:

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



## S3 Transfer Acceleration

- Provides fast and secure transfers of files over long distances.
- Optimizes transfer speeds from across the world into S3 buckets.
- Uses globally distributed edge locations in CloudFront.
- Improves speed by 50–500 percent on average for cross-country transfer of larger objects.



- 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)

### 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.

AWS

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

### D. Storing Content with Amazon S3

	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

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

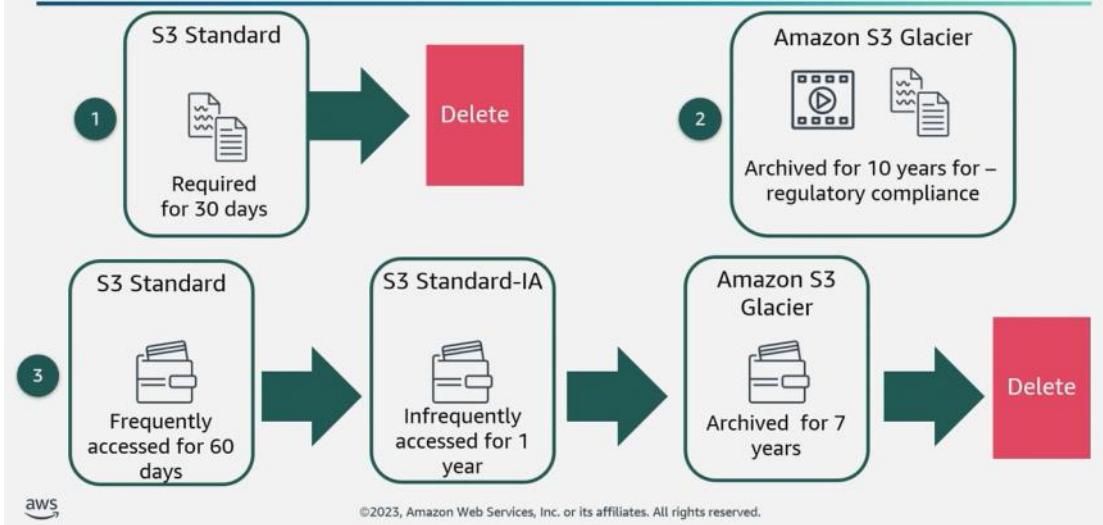
- + Transition actions transition to another storage class
- + Expiration actions define when objects expire

Set an S3 lifecycle policy.



Data will automatically transfer to a different storage class without any changes to your application.

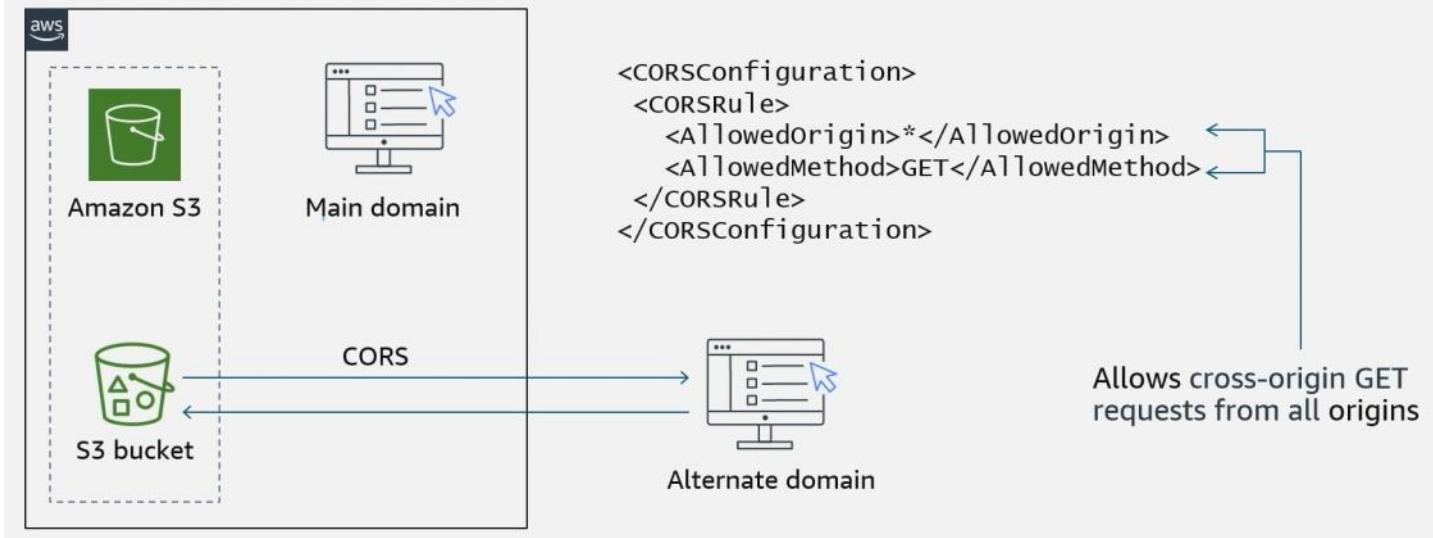
## Amazon S3 lifecycle examples



- Amazon S3 versioning: Amazon S3 versioning protects 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

## Support for cross-origin resource sharing (CORS)



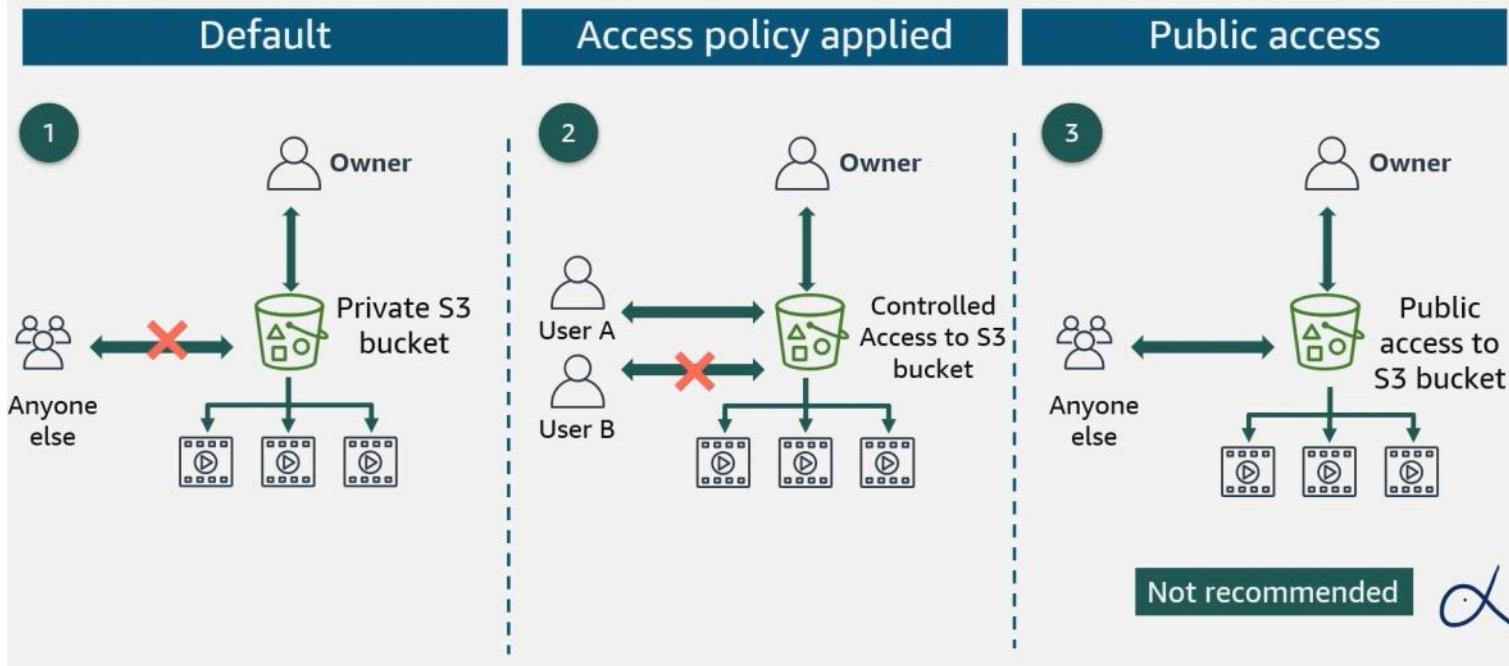
### E. Designing with Amazon S3\*

- 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
- 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 users download them
    - \* Enable this feature by selecting the default encryption on the bucket
    - + Client-side encryption:
      - \* Encrypt data on the client side and upload the encrypted data to Amazon S3
      - \* In this case, users manage the encryption process

- Amazon S3 tools for protecting buckets and objects

Tool	Description
Block Public Access feature	Makes bucket 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 list (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

## Three general approaches to configuring access



- 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>

- 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

# 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

There are costs associated with moving data in and out of Amazon S3 in some scenarios. See the Amazon S3 Pricing page linked in your course resources for details.

### **F. Applying the AWS Well-Architected Framework principles to storage**

#### *1. Best practice approach for Security pillar:*

- Enforce encryption at rest
- Encrypt objects
- Make private and block public access
- Enforce access control

#### *2. Best practice approach for Performance Efficiency*

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

#### *3. Best practice approach for Cost Optimization:*

Perform cost analysis for different usage over time

#### *4. Best practice approach for Reliability*

Select the appropriate locations for user's multi-location deployment