In this article, we’ll talk about the following topics:

* What is cloud storage?
* Types of storage
* Before S3
* What is S3
* Benefits of S3
* Objects and buckets
* How does Amazon S3 work?
* Features of S3

Now, let’s jump into our first topic and learn about cloud storage in general.

What is Cloud Storage?

Cloud storage is a web service where your data can be stored, accessed, and quickly backed up by users on the internet. It is more reliable, scalable, and secure than traditional on-premises storage systems.

Cloud storage is offered in two models:

1. Pay only for what you use
2. Pay on a monthly basis

Now, let’s have a look at the different types of storage services offered by AWS.

Types of Storage

AWS offers the following services for storage purposes:

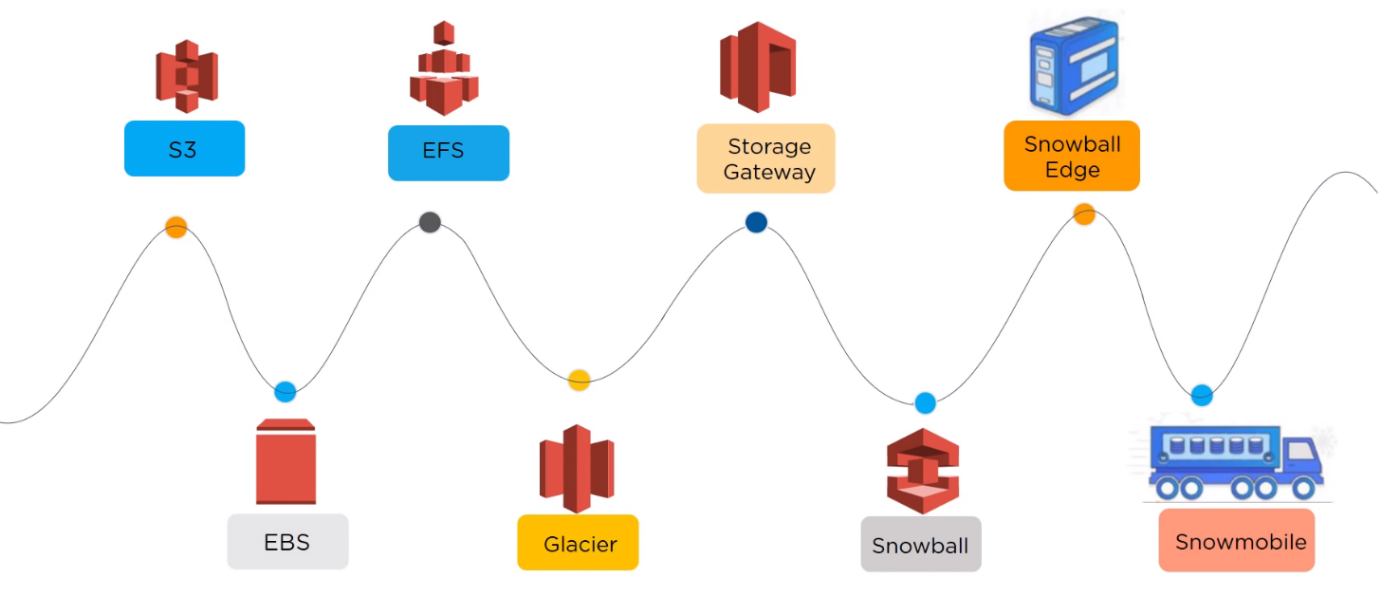


Fig: Storage services offered by Amazon

We’ll eventually take an in-depth look at the S3 service. But before we get to that, let’s have a look at how things were before we had the option of using Amazon S3.

Before S3

Organizations had a difficult time finding, storing, and managing all of your data. Not only that, running applications, delivering content to customers, hosting high traffic websites, or backing up emails and other files required a lot of storage. Maintaining the organization’s repository was also expensive and time-consuming for several reasons. Challenges included the following:

1. Having to purchase hardware and software components
2. Requiring a team of experts for maintenance
3. A lack of scalability based on your requirements
4. Data security requirements

These are the issues S3 would eventually solve. So, what exactly is S3?

What is S3?

Amazon S3 (Simple Storage Service) provides object storage, which is built for storing and recovering any amount of information or data from anywhere over the internet. It provides this storage through a web services interface. While designed for developers for easier web-scale computing, it provides 99.999999999 percent durability and 99.99 percent availability of objects. It can also store computer files up to 5 terabytes in size.

Some of the benefits of S3 are:

* Durability
* Low cost
* Scalability
* Availability
* Security
* Flexibility
* Simple data transfer

Let’s have a look at some of the major components of the S3 storage service.

Objects and Buckets

An object consists of data, key (assigned name), and metadata. A bucket is used to [store objects](https://aws.amazon.com/s3/features/). When data is added to a bucket, Amazon S3 creates a unique version ID and allocates it to the object.



Fig: Example of an object, bucket, and link address

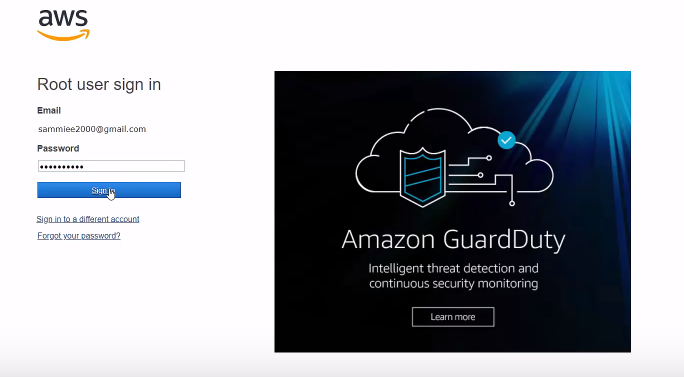


Fig: Logging into AWS

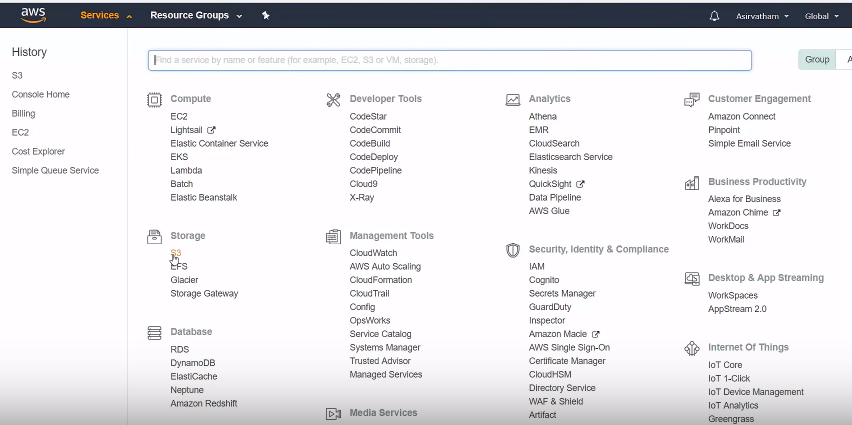


Fig: Selecting S3 from Service offerings

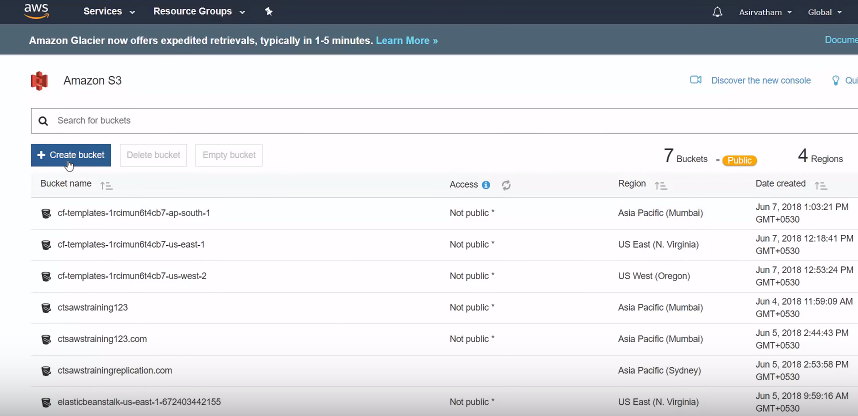


Fig: Amazon S3 bucket list (usually empty for first-time users); create a bucket by clicking on “Create bucket” button

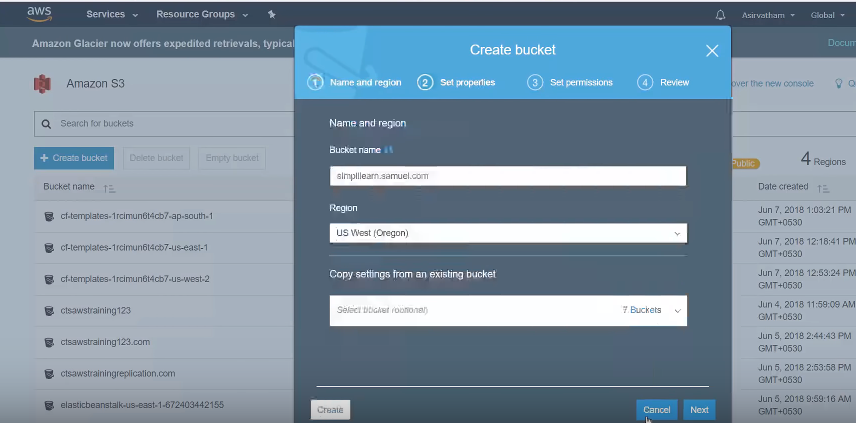


Fig: Create a bucket by setting up name, region and other options; finish off the process by pressing the “Create” button

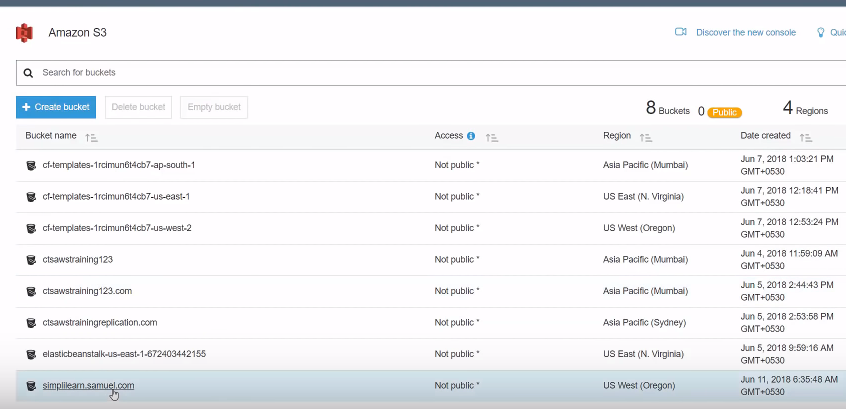


Fig: Select the created bucket

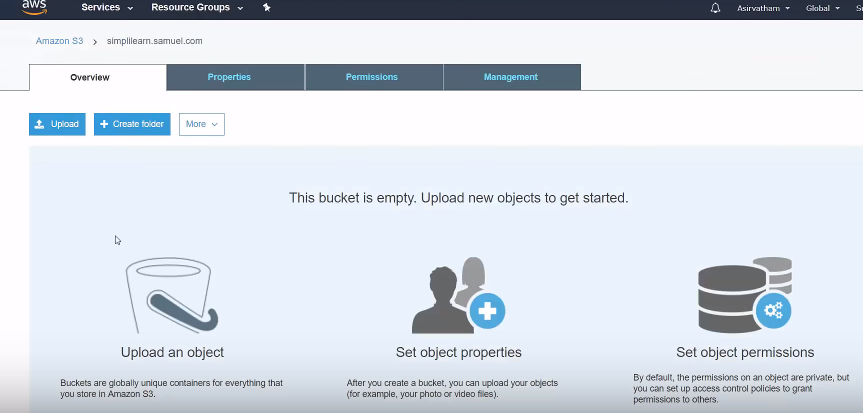


Fig: Click on upload to select a file to be added to the bucket

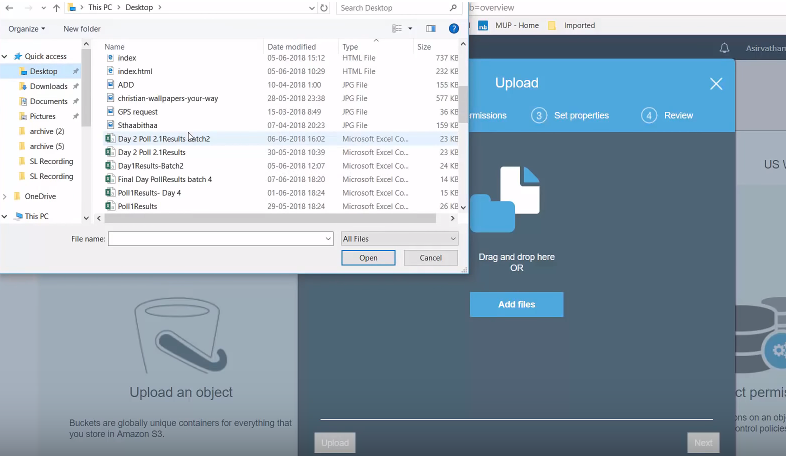


Fig: Select a file to be added

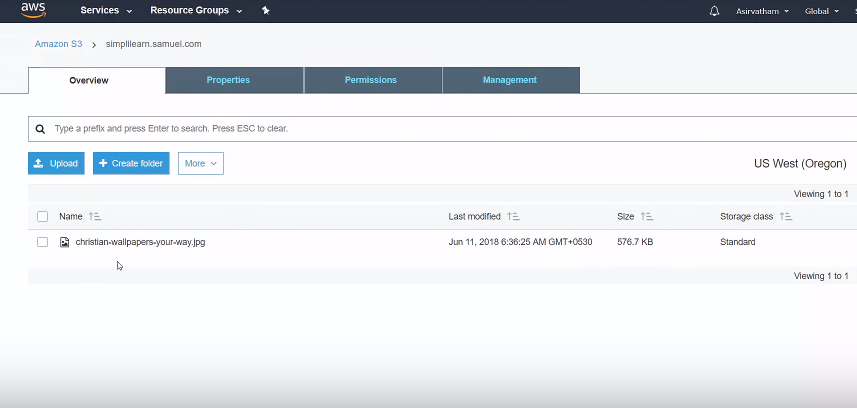


Fig: The file is now uploaded into the bucket

Let’s now have a look at how Amazon S3 works.

How Does Amazon S3 work?

Like we saw in the example above, first off, a user creates a bucket. When this [bucket](https://docs.aws.amazon.com/AmazonS3/latest/dev/UsingBucket.html) is created, the user will specify the region in which the bucket is deployed. Later, when files are uploaded to the bucket, the user will determine the type of S3 storage class to be used for those specific objects. After this, users can define features to the bucket, such as bucket policy, lifecycle policies, versioning control, etc.

Now, let’s talk about the different storage classes offered by Amazon S3.

Storage Classes of S3

Let’s have a look at the different storage classes using the example of a school:

1. Amazon S3 Standard for frequent data access: Suitable for a use case where the latency should below. *Example: Frequently accessed data will be the data of students’ attendance, which should be retrieved quickly.*
2. Amazon S3 Standard for infrequent data access: Can be used where the data is long-lived and less frequently accessed. *Example: Students’ academic records will not be needed daily, but if they have any requirement, their details should be retrieved quickly.*
3. Amazon Glacier: Can be used where the data has to be archived, and high performance is not required.*Example: Ex-student’s old record (like admission fee) will not be needed daily, and even if it is necessary, low latency is not required.*
4. One Zone-IA Storage Class: It can be used where the data is infrequently accessed and stored in a single region. *Example: Student’s report card is not used daily and stored in a single availability region (i.e., school).*
5. Amazon S3 Standard Reduced Redundancy storage: Suitable for a use case where the data is non-critical and reproduced quickly. *Example: Books in the library are non-critical data and can be replaced if lost.*

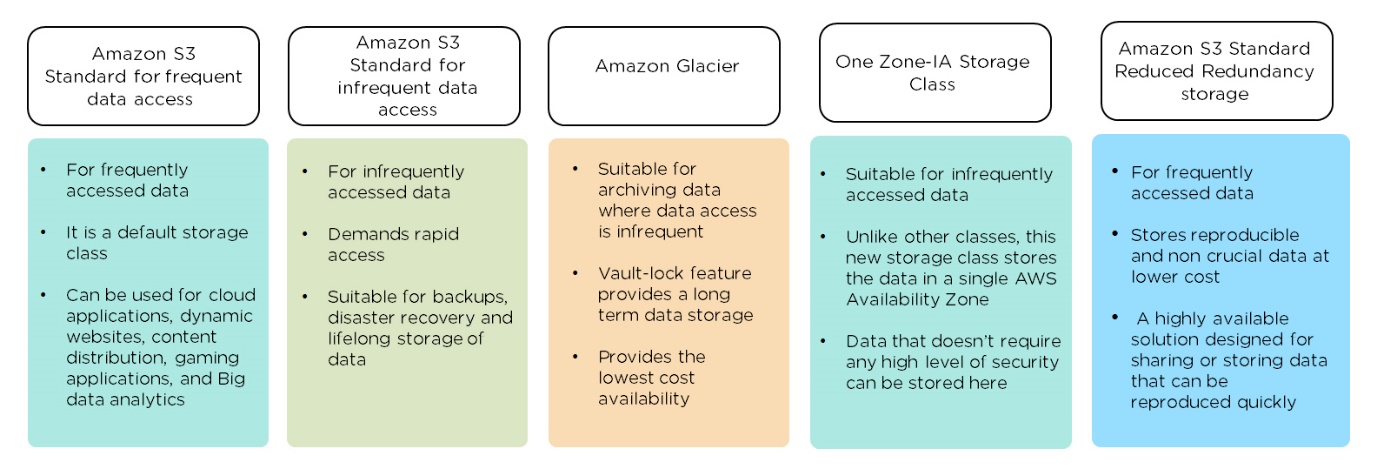


Fig: A comparison of all storage classes

**

Fig: Technical comparison between classes

Let’s now have a look at the different features offered by S3.

Features of S3

[Lifecycle Management](https://aws.amazon.com/blogs/developer/amazon-s3-lifecycle-management/)

In lifecycle management, Amazon S3 applies a set of rules that define the action to a group of objects. You can manage and store objects in a cost-effective manner. There are two types of actions:

1. Transition Action

With this action, you can choose to move objects to another storage class. With this, you can configure S3 to move your data between various storage classes on a defined schedule. Assume you’ve got some data stored in the S3 standard class. If this data is not used frequently for 30 days, it would be moved to the S3 infrequent access class. And after 60 days, it is moved to Glacier. This helps you to migrate your data to lower-cost storage as it ages automatically.

2. Expiration Actions

Here, S3 removes all objects within the bucket when a specified date or time period in the object’s lifetime is reached.

An example of how lifecycle management works:

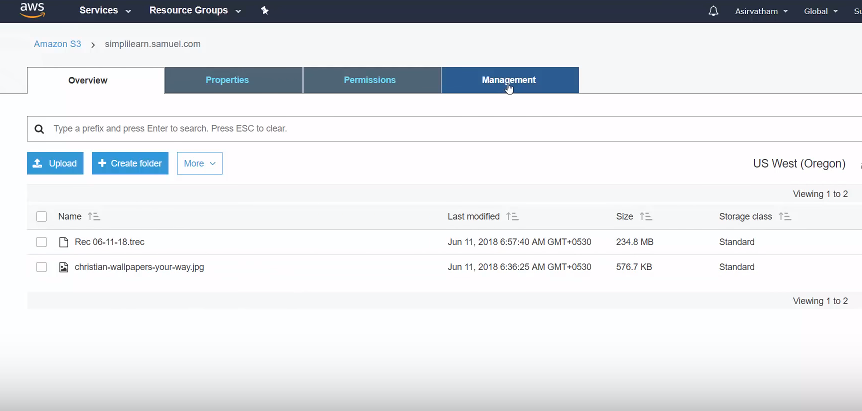


Fig: From within your bucket select management

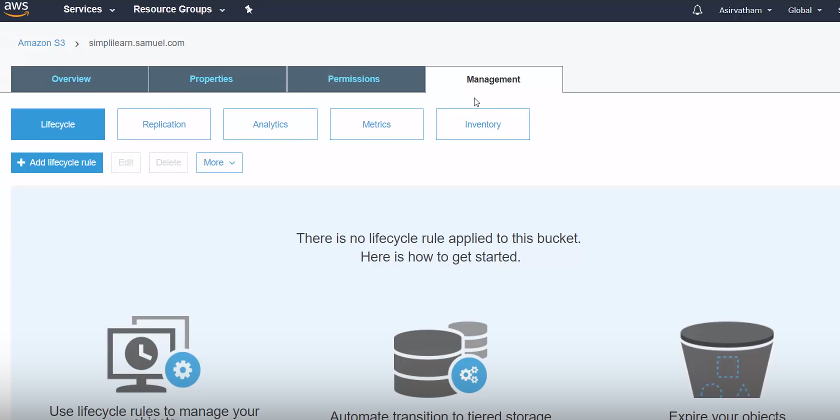


Fig: Select “Lifecycle” and then click on the “Add lifecycle rule.”

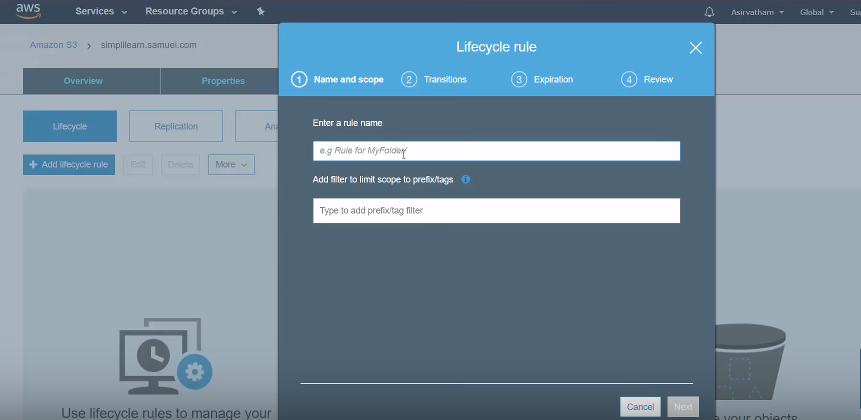


Fig: Add a rule name and scope

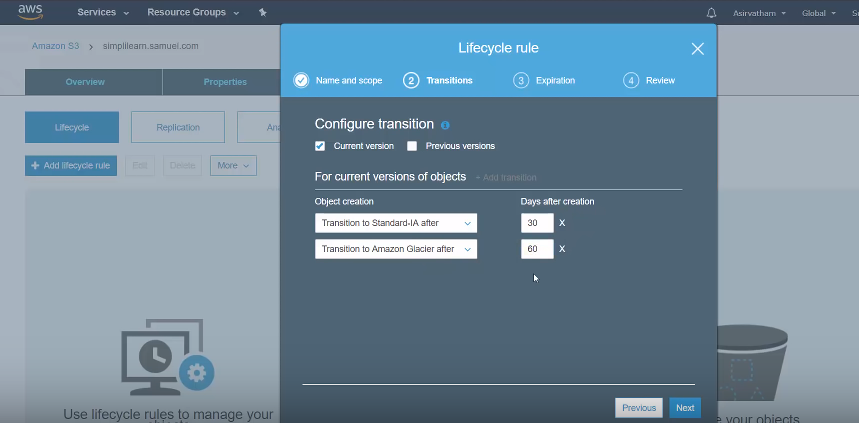


Fig: Configure transaction options

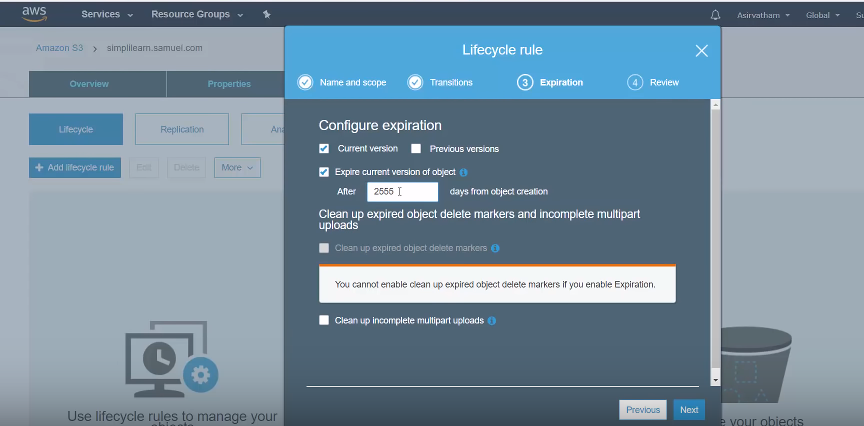


Fig: Set up expiration details

[**Bucket Policy**](https://docs.aws.amazon.com/AmazonS3/latest/dev/using-iam-policies.html)

Bucket policy is an IAM policy where you can allow or deny permission to your Amazon S3 resources. With bucket policy, you also define security rules that apply to more than one file within a bucket. For example: If you do not want a user to access the “Simplilearn” bucket, then with the help of JSON script, you can set permissions. As a result, a user would be denied access to the bucket.

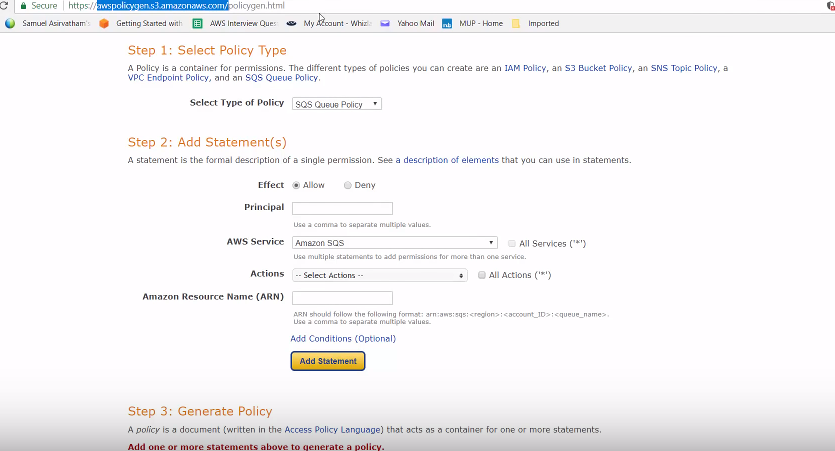


Fig: Use an online tool to generate a policy. Select the type of policy as an S3 bucket policy. Select the appropriate effect. In this case, denying access.

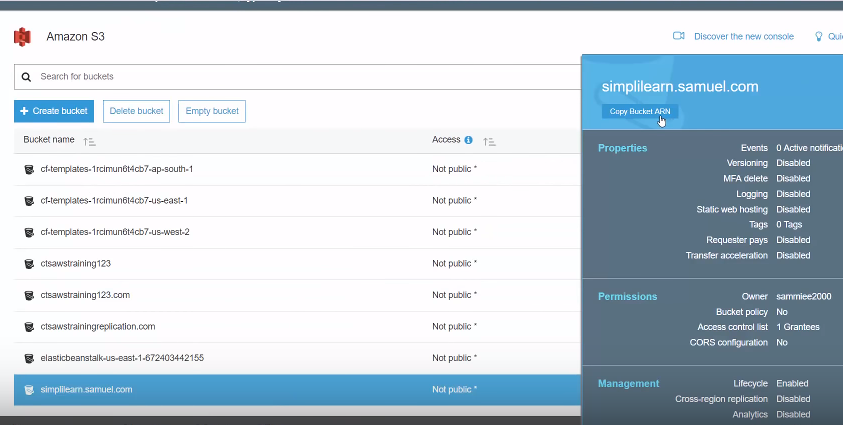


Fig: Find the ARN of the bucket

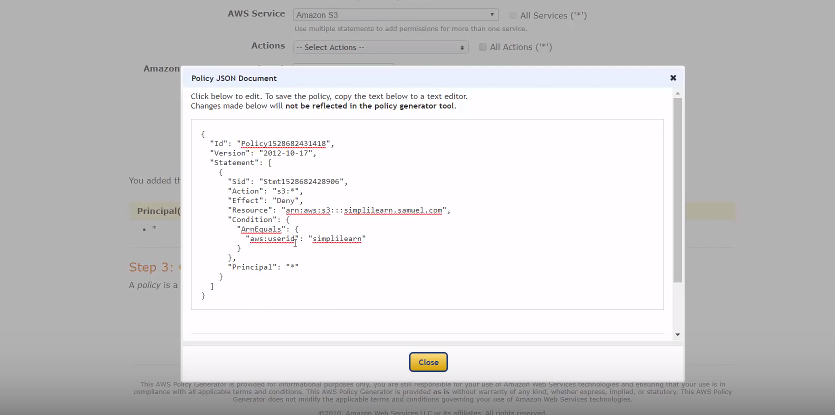


Fig: Set up additional conditions and set up a JSON script to deny access to a particular user. In this case, “simplilearn.”

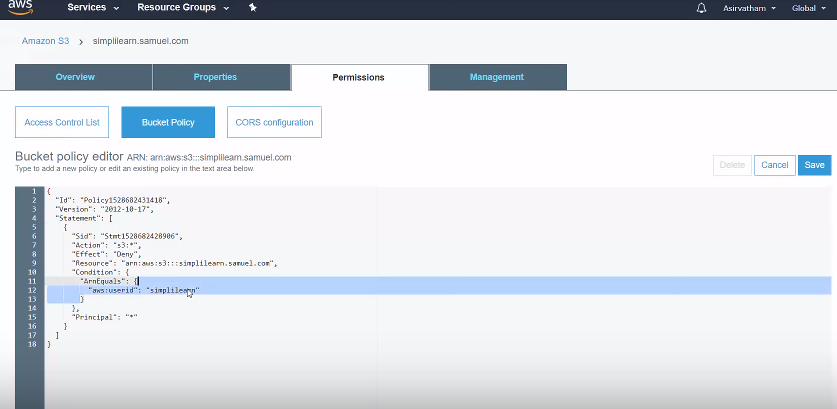


Fig: Go back to the bucket and set up a bucket policy under “Permissions.” Then click on “Save.”

Data Protection

Amazon S3 provides IT teams with a highly durable, protected, and scalable infrastructure designed for object storage.

Amazon S3 protects your data using two methods:

1. Data encryption
2. Versioning
3. Cross-region Replication
4. Transfer Acceleration

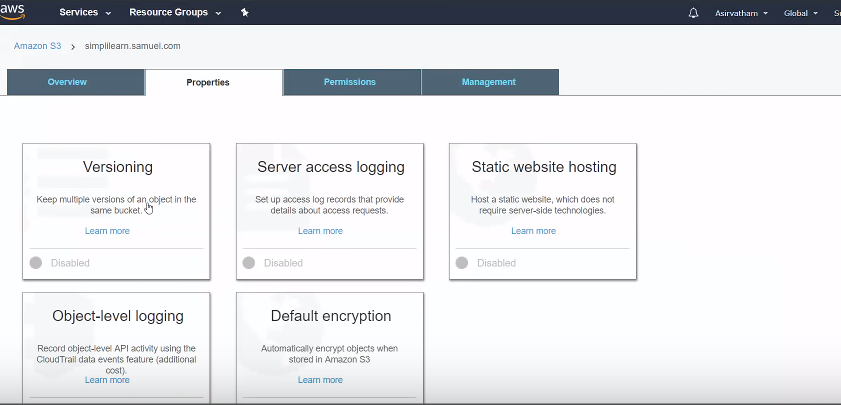


Fig: Go to your bucket, select properties, and turn on “Versioning.”

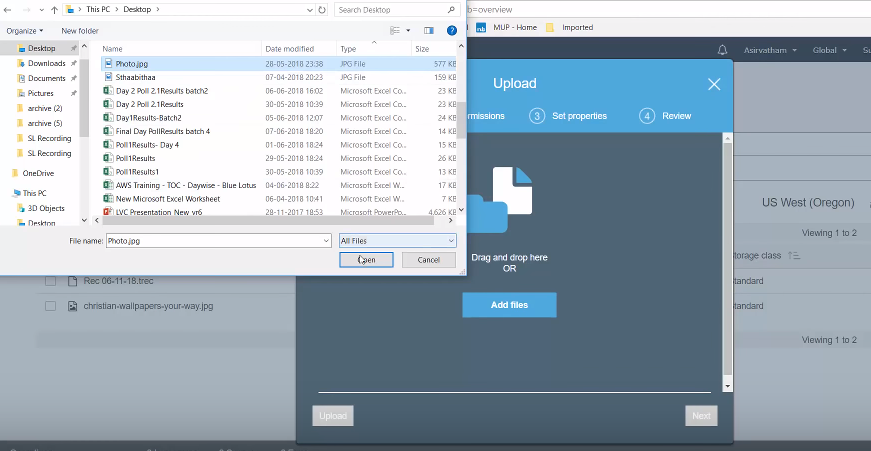


Fig: Upload an object

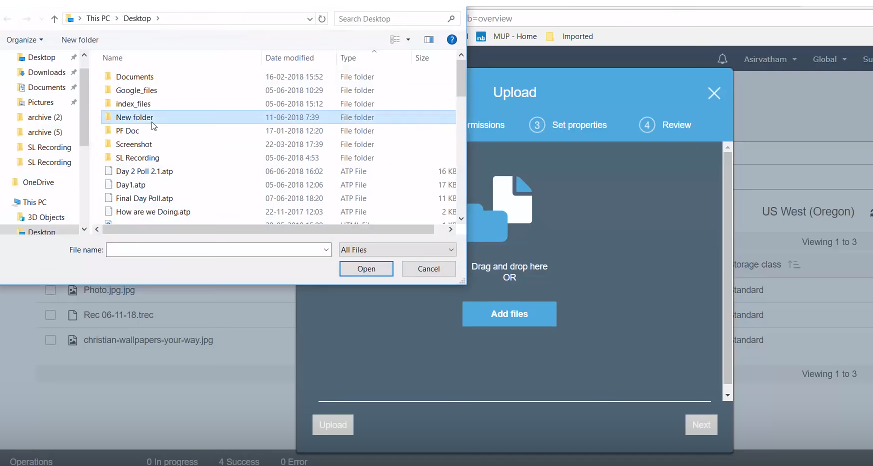


Fig: Upload another file of the same name

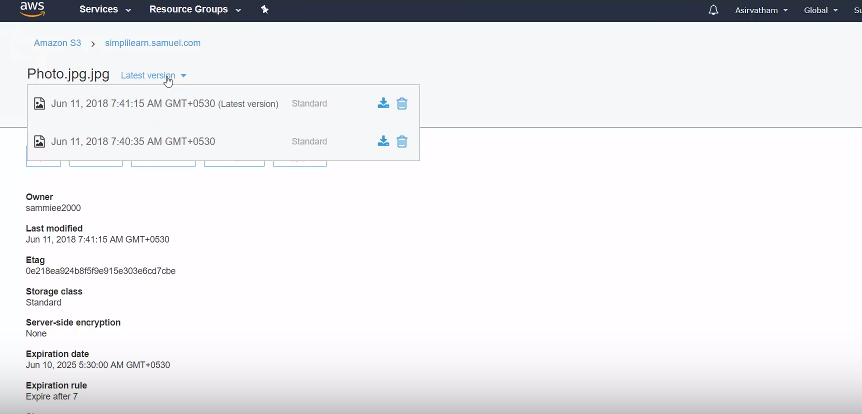


Fig: Select the file and alternate between its current and older versions

1. Data Encryption

This refers to the protection of data while it’s being transmitted and at rest. It can happen in two ways, client-side encryption (data encryption at rest) and server-side encryption (data encryption in motion).

2. Versioning

It is utilized to preserve, recover, and restore an early version of every object you store in your S3 bucket. Unintentional erases or overwriting of objects can easily be managed with versioning. For example, in a bucket, it is possible to have objects with the same key name but different version IDs.

3. Cross-region Replication

Cross-region replication provides automatic copying of every object uploaded to your buckets (source and destination bucket) in different AWS regions. Versioning needs to be turned on to enable CRR.

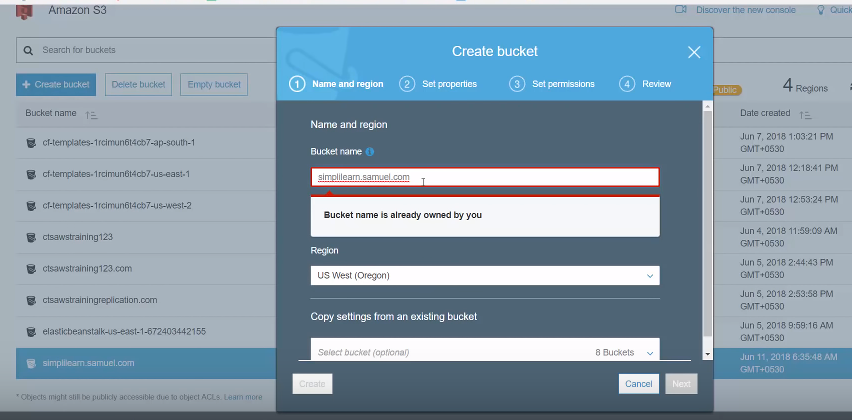


Fig: Create a new bucket in a different region

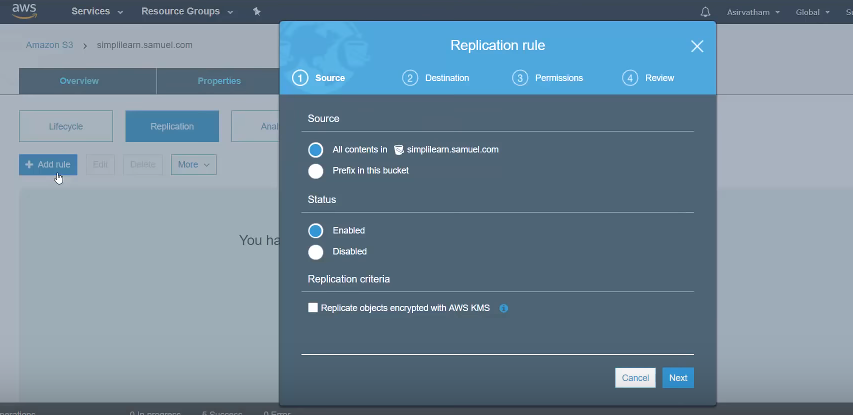


Fig: Select uploaded file, go to “Management” and then replication.

Here, click on “Add Rule.”

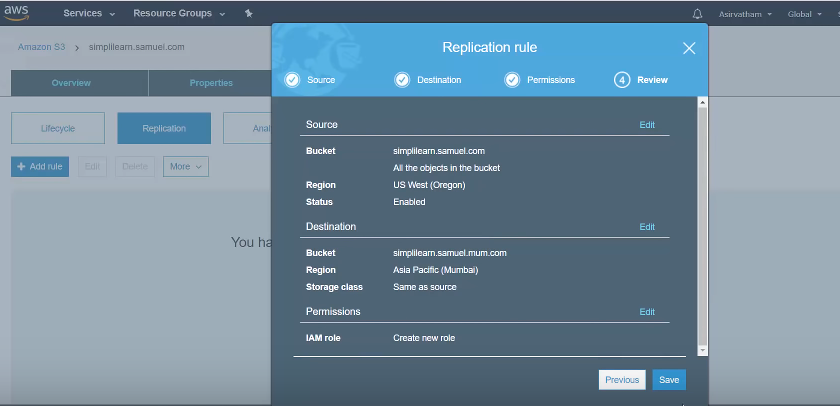


Fig: Select the source, destination and IAM rule

4. Transfer Acceleration

This enables fast, easy, and secure transfers of files over long distances between your client and S3 bucket. The edge locations around the world provided by Amazon CloudFront are taken advantage of by transfer acceleration. It works by carrying data over an optimized network bridge that keeps running between the AWS Edge Location (closest region to your clients) and your Amazon S3 bucket.

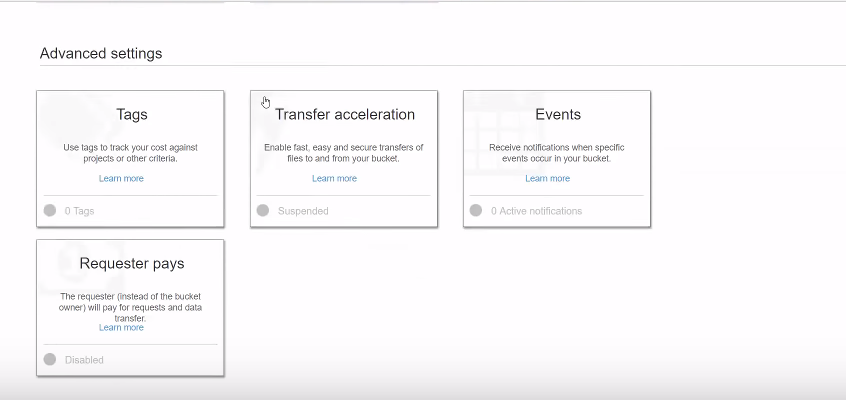


Fig: Go to properties and select transfer acceleration to enable it