S3 Simple Storage Service

Introduction:

S3 (Simple Storage Service) is a scalable, high-speed, low-cost web-based service designed for online backup and archiving of data and application programs. It allows to upload, store, and download any type of files up to 5 TB in size. This service allows the subscribers to access the same systems that Amazon uses to run its own web sites. The subscriber has control over the accessibility of data, i.e. privately/publicly accessible.

What is Amazaon S3?

Amazon Simple Storage Service is storage for the Internet. It is designed to make web-scale computing easier for developers.

Amazon Simple Storage Service (S3) is a storage for the internet. It is designed for large-capacity, low-cost storage provision across multiple geographical regions. Amazon S3 provides developers and IT teams with **Secure**, **Durable** and **Highly Scalable** object storage.

a) S3 is **Secure** because AWS provides:

- Encryption to the data that you store. It can happen in two ways:
    - Client Side Encryption
    - Server Side Encryption
- Multiple copies are maintained to enable regeneration of data in case of data corruption
- *Versioning,* wherein each edit is archived for a potential retrieval.

b) S3 is **Durable** because:

- It regularly verifies the integrity of data stored using checksums e.g. if S3 detects there is any corruption in data, it is immediately repaired with the help of replicated data.
- Even while storing or retrieving data, it checks incoming network traffic for any corrupted data packets.

c) S3 is **Highly Scalable**, since it automatically scales your storage according to your requirement and you only pay for the storage you use.

1) Overview of Amazon S3:

Amazon S3 has a simple web services interface that you can use to store and retrieve any amount of data, at any time, from anywhere on the web.

In This describes how you send requests to create buckets, store and retrieve your objects, and manage permissions on your resources and also describes access control and the authentication process. Access control defines who can access objects and buckets within Amazon S3, **and** the type of access (e.g., READ and WRITE). The authentication process verifies the identity of a user who is trying to access Amazon Web Services (AWS).

1.1 How is data organized in S3?

Data in S3 is organized in the form of buckets.

- A Bucket is a logical unit of storage in S3.
- A Bucket contains objects which contain the data and metadata.

Before adding any data in S3 the user has to create a bucket which will be used to store objects.

1.2 Where is your data stored geographically?

You can self-choose where or in which region your data should be stored. Making a decision for the region is important and therefore it should be planned well.

These are the 4 parameters to choose the optimal region –

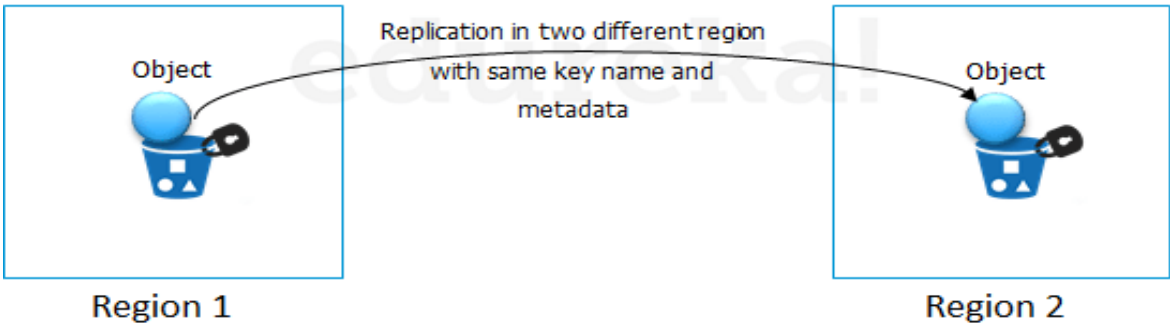
- Pricing
- User/Customer Location
- Latency

Talking about regions, let's see about the possibility of having a backup in some other availability region or you may want to move your data to some other region. Thankfully, this feature has been recently added to the AWS S3 system and is pretty easy to use.

1.3 Cross-region Replication

As the name suggests, Cross-region Replication enables user to either replicate or transfer data to some other location without any hassle.

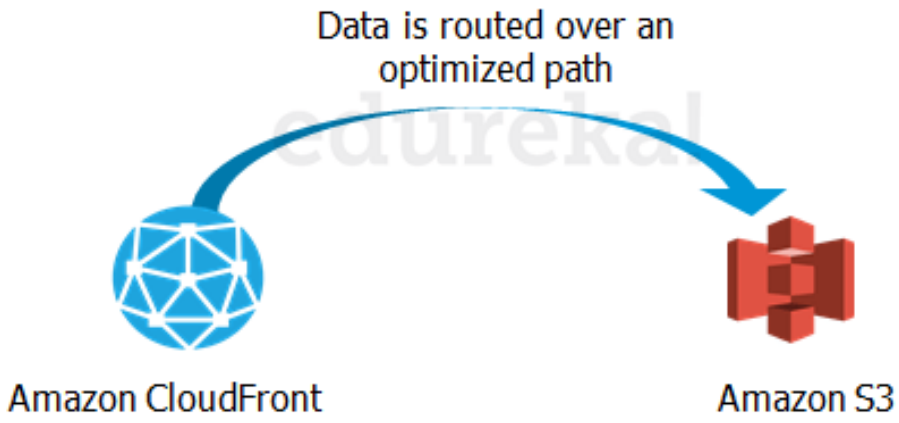
This obviously has a cost to it which has been discussed further in this article.



How is the data transferred?

Besides traditional transfer practices that is over the internet, AWS has 2 more ways to provide data transfer securely and at a faster rate:

- Transfer Acceleration
- Snowball



a) Transfer Acceleration enables fast, easy and secure transfers over long distances by exploiting Amazon's CloudFront edge technology.

CloudFront is a caching service by AWS, in which the data from client site gets transferred to the nearest edge location and from there the data is routed to your AWS S3 bucket over an optimised network path.

b) The **Snowball** is a way of transferring your data physically. In this Amazon sends an equipment to your premises, on which you can load the data. It has a kindle attached to it which has your shipping address when it is shipped from Amazon. When data transfer is complete on the Snowball, kindle changes the shipping address back to the AWS headquarters where the Snowball has to be sent.

The Snowball is ideal for customers who have large batches of data move. The average turnaround time for Snowball is 5-7 days, in the same time Transfer Acceleration can transfer up to 75 TB of data on a dedicated 1Gbps line. So depending on the use case, a customer can decide.

Obviously, there will be some cost around it, let's look at the overall costing around S3.

Pricing

"**Isn't anything free on AWS?"**

As a part of the AWS Free Usage Tier, you can get started with AWS S3 for free. Upon sign up, new AWS customers receive 5 GB of Amazon S3 standard storage, 20,000 Get-Requests, 2,000 Put-Requests, and 15GB of data transfer-out each month for one year.

Over this limit, there is a cost attached, let's understand how amazon charges you:

i. How is S3 billed?

Though having so many features, AWS S3 is affordable and flexible in its costing. It works on **Pay Per Use,** meaning, you only pay what you use. The table below is an example for pricing of S3 for a specific region:

Storage/month	Standard Storage	Standard – Infrequent Access Storage	Glacier Storage
First 1 TB / month	$0.0300 per GB	$0.0125 per GB	$0.007 per GB
Next 49 TB / month	$0.0295 per GB	$0.0125 per GB	$0.007 per GB
Next 450 TB / month	$0.0295 per GB	$0.0125 per GB	$0.007 per GB
Next 500 TB / month	$0.0285 per GB	$0.0125 per GB	$0.007 per GB

Cross Region Replication is billed in the following way:

If you replicate 1,000 1 GB objects (1,000 GB) between regions you will incur a request charge of $0.005 (1,000 requests x $0.005 per 1,000 requests) for replicating 1,000 objects and a charge of $20 ($0.020 per GB transferred x 1,000 GB) for inter-region data transfer. After replication, the 1,000 GB will incur storage charges based on the destination region.

Snowball, there are 2 variants:

- Snowball 50 TB : 200$
- Snowball 80 TB: 250$

This is the fixed service fee that they charge.

Apart from this there are on-site, charges which are exclusive of shipping days, the shipping days are free.

The first 10 on-site days are also free, meaning when the Snowball reaches your premises from then, till the day it is shipped back, they are the on-site days. The day it arrives, and the day it is shipped gets counted as shipping days, therefore are free.

Transfer Acceleration pricing is shown in the following table:

Data Transfer IN to Amazon S3 from the Internet:	
Accelerated by AWS Edge Locations in the United States, Europe, and Japan	$0.04/GB
Accelerated by all other AWS Edge Locations	$0.08/GB
Data Transfer OUT from Amazon S3 to the Internet:	
Accelerated by any AWS Edge Location	$0.04/GB
Data Transfer between Amazon S3 and another AWS region:	
Accelerated by any AWS Edge Location	$0.04/GB

2) Advantages to Amazon S3:

Amazon S3 is intentionally built with a minimal feature set that focuses on simplicity and robustness. Following are some of advantages of the Amazon S3 service:

- **Create Buckets** :– Create and name a bucket that stores data. Buckets are the fundamental container in Amazon S3 for data storage.
- **Store data in Buckets** :– Store an infinite amount of data in a bucket. Upload as many objects as you like into an Amazon S3 bucket. Each object can contain up to 5 TB of data. Each object is stored and retrieved using a unique developer-assigned key.
- **Download data** :– Download your data or enable others to do so. Download your data any time you like or allow others to do the same.
- **Permissions** :– Grant or deny access to others who want to upload or download data into your Amazon S3 bucket. Grant upload and download permissions to three types of users. Authentication mechanisms can help keep data secure from unauthorized access.
- **Standard interfaces** :– Use standards-based REST and SOAP interfaces designed to work with any Internet-development toolkit.

3) Features

- **Low cost and Easy to Use:**− Using Amazon S3, the user can store a large amount of data at very low charges.

- **Secure:**− Amazon S3 supports data transfer over SSL and the data gets encrypted automatically once it is uploaded. The user has complete control over their data by configuring bucket policies using AWS IAM.

- **Scalable:**− Using Amazon S3, there need not be any worry about storage concerns. We can store as much data as we have and access it anytime.

- **Higher performance:**− Amazon S3 is integrated with Amazon CloudFront, that distributes content to the end users with low latency and provides high data transfer speeds without any minimum usage commitments.

- **Integrated with AWS services:**− Amazon S3 integrated with AWS services include Amazon CloudFront, Amazon CLoudWatch, Amazon Kinesis, Amazon RDS, Amazon Route 53, Amazon VPC, AWS Lambda, Amazon EBS, Amazon Dynamo DB, etc.

4) Amazon S3 Concepts:

This section describes key concepts and terminology you need to understand to use Amazon S3 effectively.

(a) Buckets (b) Objects (c) Keys (d) Regions (e) Amazon S3 Data Consistency Model

(a) Buckets:

A bucket is a container for objects stored in Amazon S3. Every object is contained in a bucket.

Buckets serve several purposes: they organize the Amazon S3 namespace at the highest level, they identify the account responsible for storage and data transfer charges, they play a role in access control, and they serve as the unit of aggregation for usage reporting.

Working with Amazon S3 Buckets

Amazon S3 is cloud storage for the internet. To upload your data (photos, videos, documents etc.), you first create a bucket in one of the AWS Regions. You can then upload any number of objects to the bucket.

An Amazon S3 bucket name is globally unique, and the namespace is shared by all AWS accounts. You should not depend on specific bucket naming conventions for availability or security verification purposes. For bucket naming guidelines, see Bucket Restrictions and Limitations.

Amazon S3 creates buckets in a Region you specify. To optimize latency, minimize costs, or address regulatory requirements, choose any AWS Region that is geographically close to you. For example, if you reside in Europe, you might find it advantageous to create buckets in the EU (Ireland) or EU (Frankfurt) Regions.

Creating a Bucket

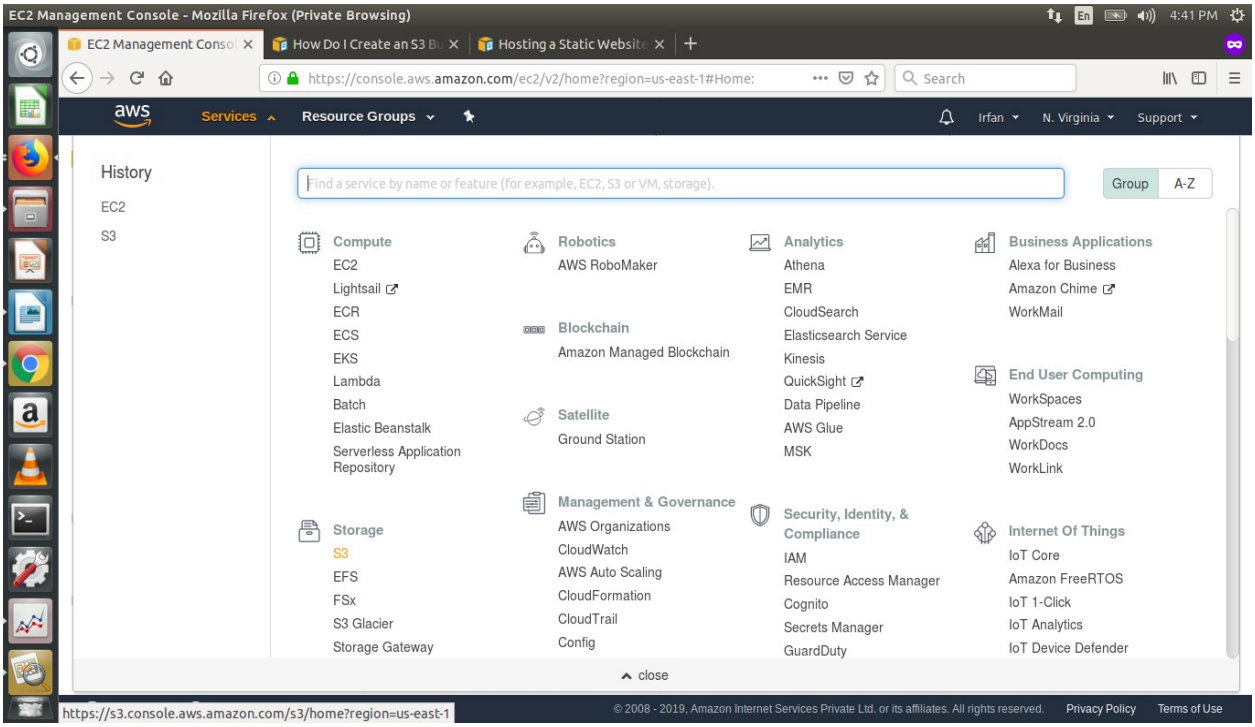
Amazon S3 provides APIs for creating and managing buckets. By default, you can create up to 100 buckets in each of your AWS accounts. If you need more buckets, you can increase your account bucket limit to a maximum of 1,000 buckets by submitting a service limit increase. T

When you create a bucket, you provide a name and the AWS Region where you want to create the bucket. For information about naming buckets, see Rules for Bucket Naming.

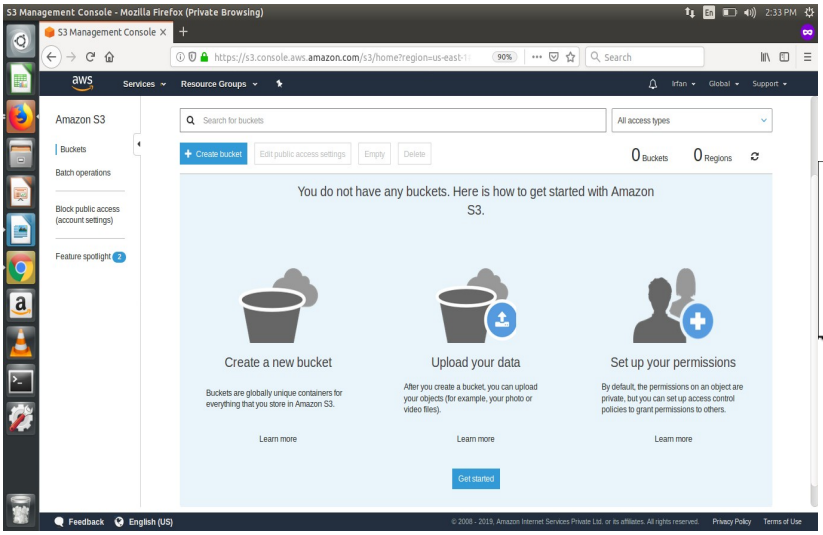
You can store any number of objects in a bucket.

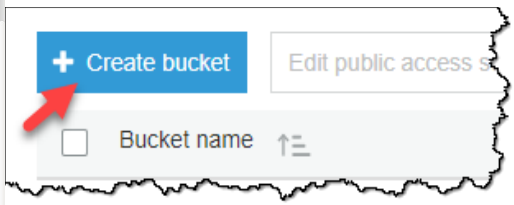
You can create a bucket using the following steps:

1. Sign in to the AWS Management Console and open the Amazon S3 console at https://console.aws.amazon.com/s3/.



2. Choose **Create bucket.**





3. On the **Name and region** page, type a name for your bucket and choose the AWS Region where you want the bucket to reside. Complete the fields on this page as follows:

(a) For **Bucket name**, type a unique DNS-compliant name for your new bucket. Follow these naming guidelines:

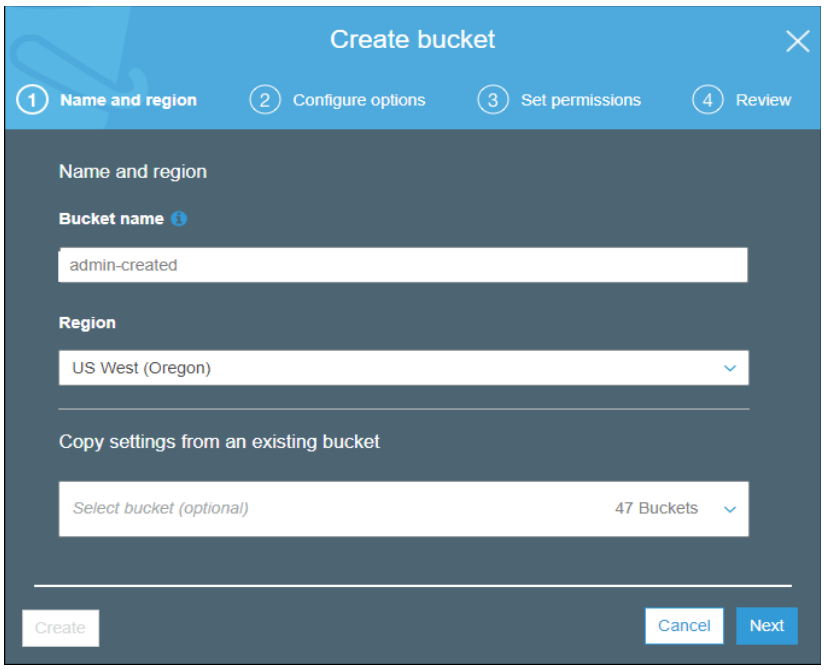
- The name must be unique across all existing bucket names in Amazon S3.
- The name must not contain uppercase characters.
- The name must start with a lowercase letter or number.
- The name must be between 3 and 63 characters long.
- After you create the bucket you cannot change the name, so choose wisely.
- Choose a bucket name that reflects the objects in the bucket because the bucket name is visible in the URL that points to the objects that you're going to put in your bucket.

(b) For **Region**, choose the AWS Region where you want the bucket to reside. Choose a Region close to you to minimize latency and costs, or to address regulatory requirements. Objects stored in a Region never leave that Region unless you explicitly transfer them to another Region.

The settings for the following bucket properties are copied: versioning, tags, and logging.

Do one of the following:

- If you copied settings from another bucket, choose **Create**. You're done, so skip the following steps.
- If not, choose **Next**.



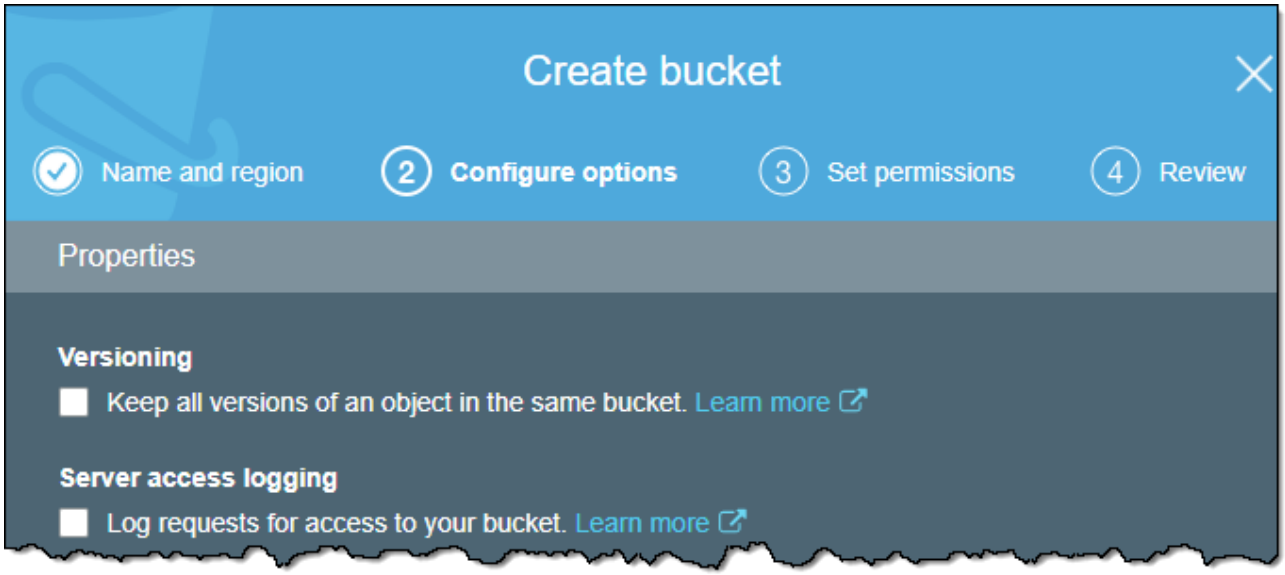
3. On the **Configure options** page, you can configure the following properties and Amazon CloudWatch metrics for the bucket. Or, you can configure these properties and CloudWatch metrics later, after you create the bucket.

(a) Versioning

Select **Keep all versions of an object in the same bucket.** to enable object versioning for the bucket. For more information on enabling versioning.

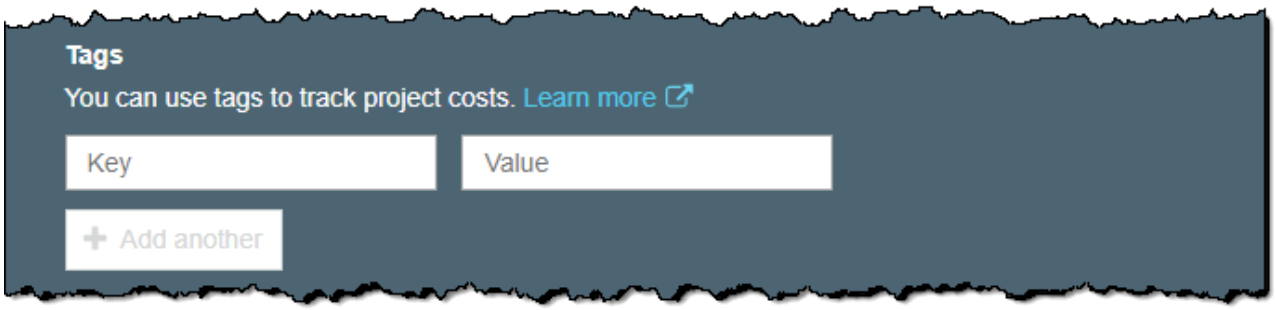
(b) Server access logging

Select **Log requests for access to your bucket.** to enable server access logging on the bucket. Server access logging provides detailed records for the requests that are made to your bucket.



(c) Tags

You can use cost allocation bucket tags to annotate billing for your use of a bucket. Each tag is a key-value pair that represents a label that you assign to a bucket.



(d) Object-level logging

Select **Record object-level API activity by using CloudTrail for an additional cost** to enable object-level logging with CloudTrail.

(e) Default encryption

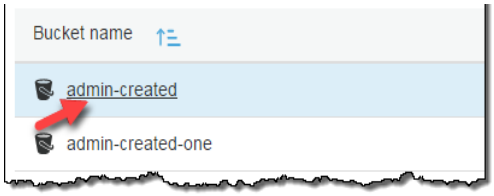
Select **Automatically encrypt objects when they are stored in S3** to

enable default encryption for the bucket. You can enable default encryption for a bucket so that all objects are encrypted when they are stored in the bucket.

To enable default encryption on an S3 bucket

1. Sign in to the AWS Management Console and open the Amazon S3 console athttps://console.aws.amazon.com/s3/.

2. In the **Bucket name** list, choose the name of the bucket that you want.

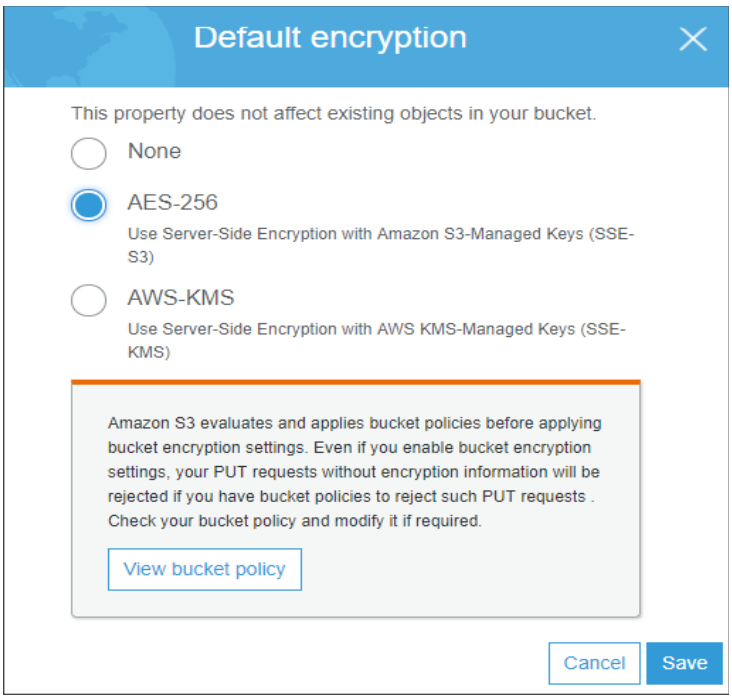


3.Choose **Properties**.

4.Choose **Default encryption**.

5.Choose **AES-256** or **AWS-KMS**.

Do use keys that are managed by Amazon S3 for default encryption, choose**AES-256**. For more information about using Amazon S3 server-side encryption to encrypt your data, see Protecting Data with Amazon S3-Managed Encryption Keys in the Amazon Simple Storage Service Developer Guide.



6. Choose **Save**.

(f) Object lock

Select **Permanently allow objects in this bucket to be locked** if you want to be able to lock objects in the bucket. Object lock requires that you enable versioning on the bucket.

Locking Objects Using Amazon S3 Object Lock

With Amazon S3 Object Lock, you can store objects using a *write-once-read-many* (WORM) model. You can use it to prevent an object from being deleted or overwritten for a fixed amount of time or indefinitely. Amazon S3 Object Lock helps you meet regulatory requirements that require WORM storage, or simply add another layer of protection against object changes and deletion.

To use Amazon S3 Object Lock, follow these basic steps:

1. Create a new bucket with Amazon S3 Object Lock enabled.

2. (Optional) Configure a default retention period for objects placed in the

bucket.

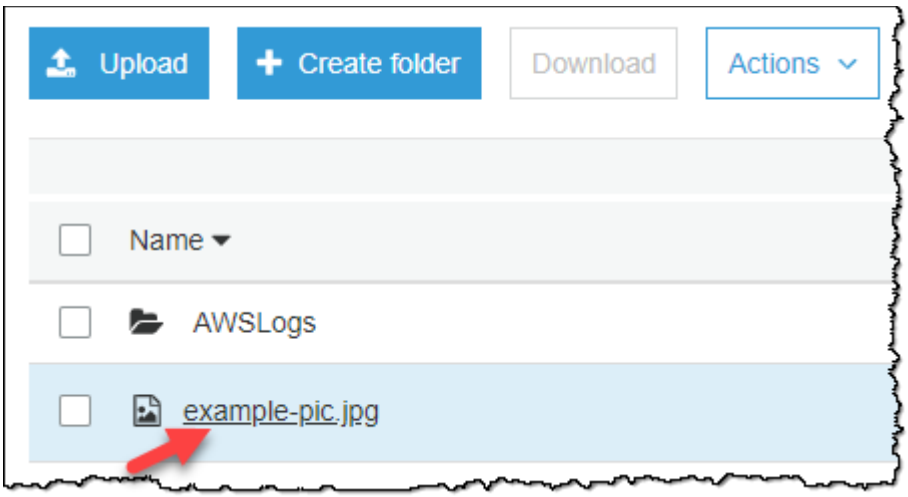
3. Place the objects that you want to lock in the bucket.
4. Apply a retention period, a legal hold, or both, to the objects that you want to protect.

To lock an Amazon S3 object

1. Sign in to the AWS Management Console and open the Amazon S3 console at https://console.aws.amazon.com/s3/.
2. In the **Bucket name** list, choose the name of the bucket that you want.

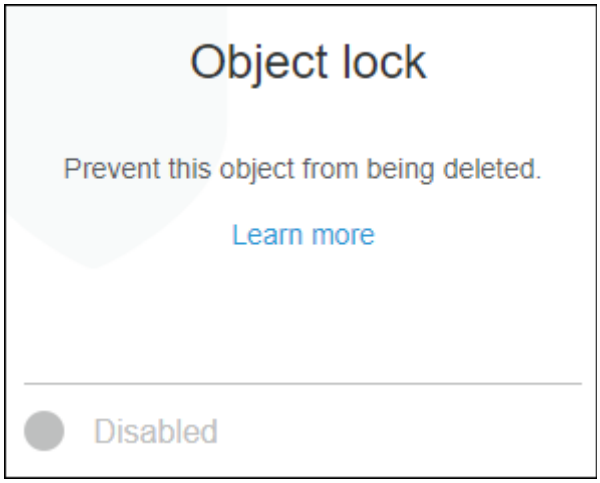
3. In the **Name** list, choose the name of the object that you want to lock.

4. Choose

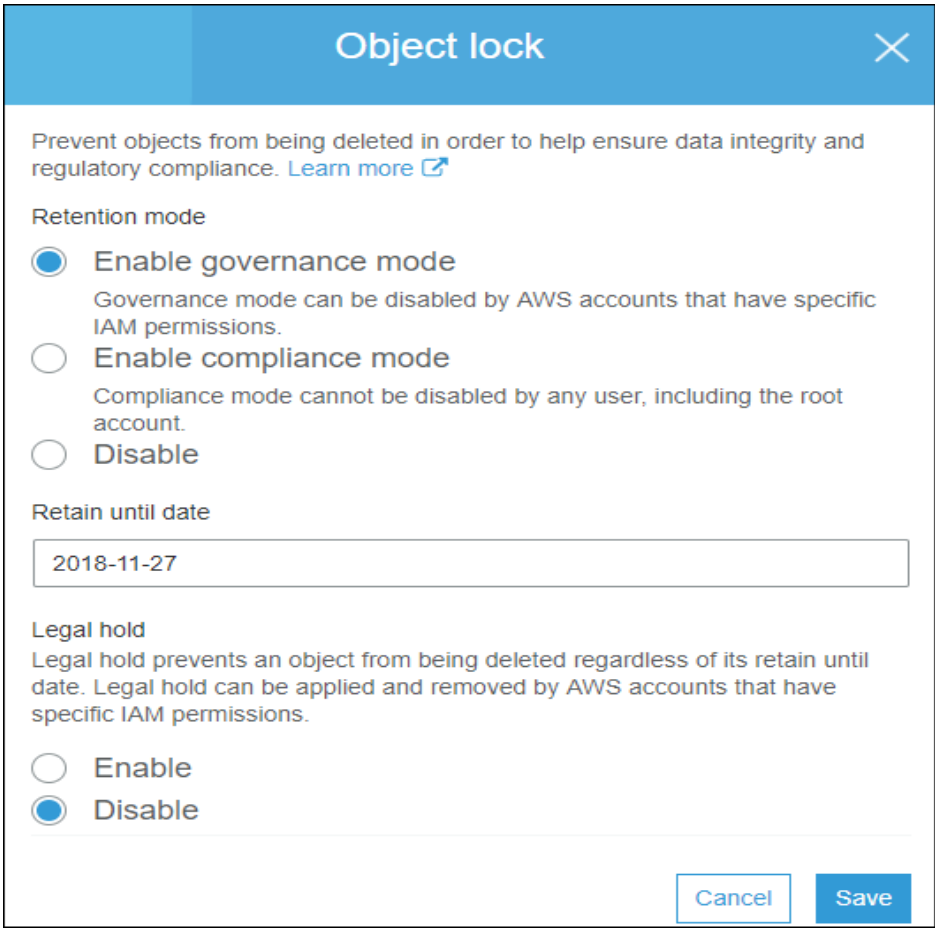


Properties.

5. Choose **Object lock**.



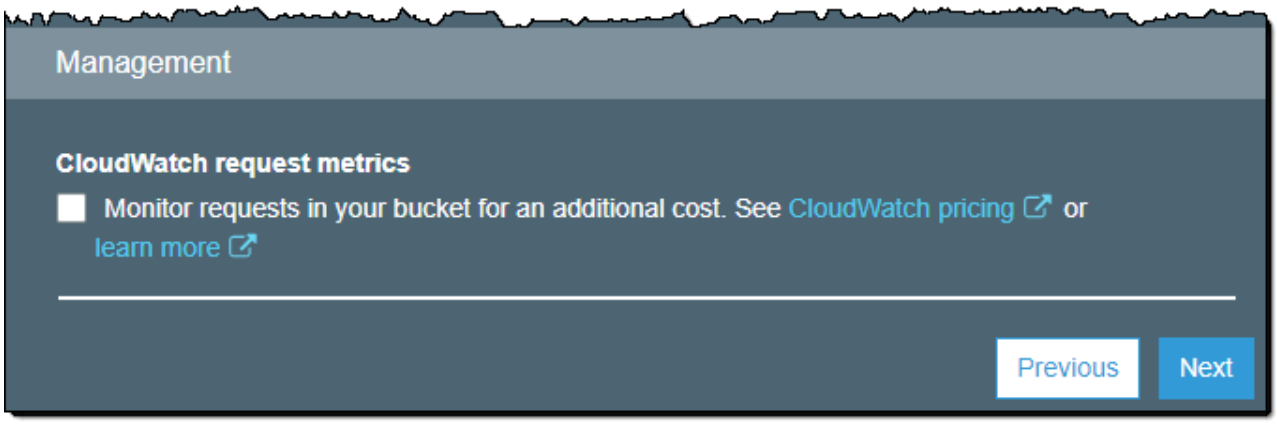
6. Choose a retention mode. You can change the **Retain until date**. You can also choose to enable a legal hold.



7. Choose **Save**.

(g) CloudWatch request metrics

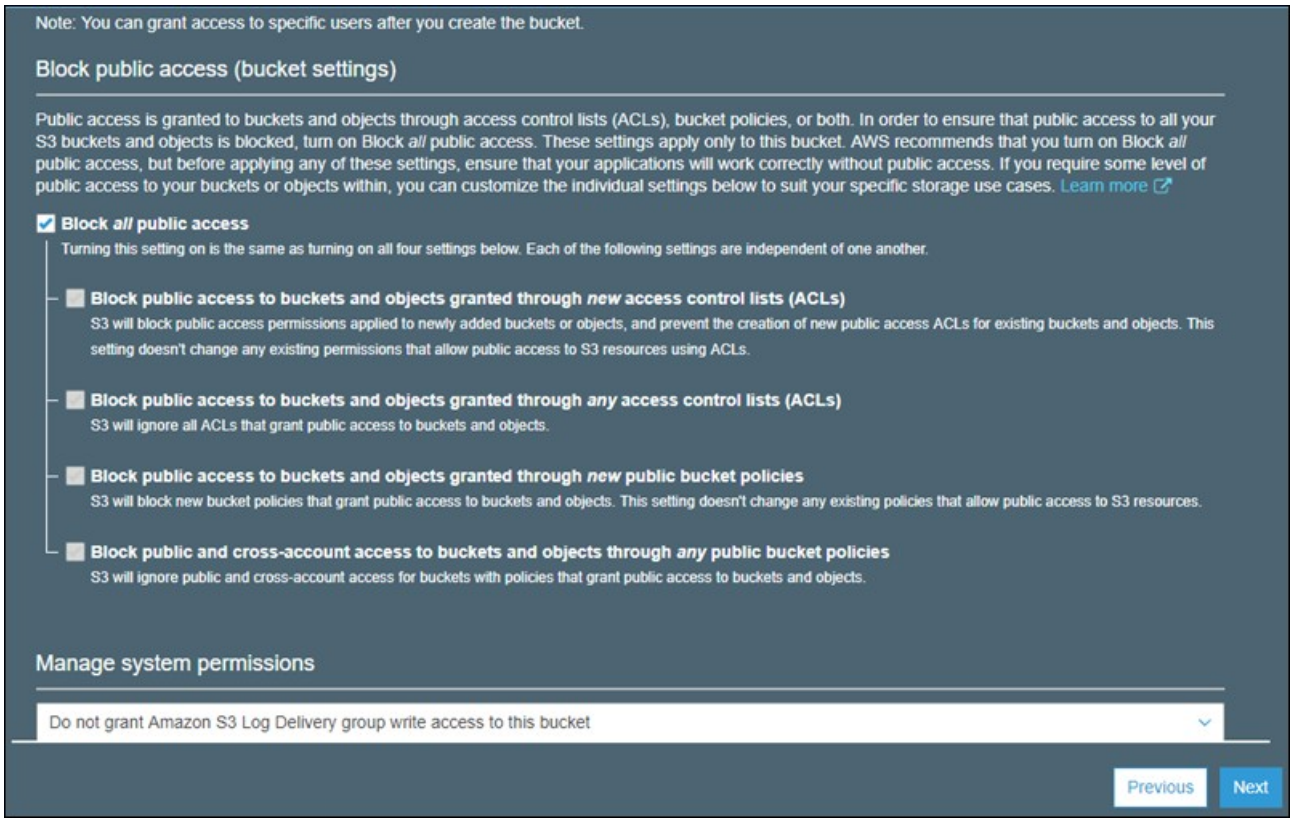
Select **Monitor requests in your bucket for an additional cost.** to configure CloudWatch request metrics for the bucket.



5. Choose **Next**.

6. On the **Set permissions** page, you manage the permissions that are set on the bucket that you are creating.Under **Block public access (bucket settings)**, we recommend that you do not change the default settings that are listed under **Block all public access**. You can change the permissions after you create the bucket.

If you intend to use the bucket to store Amazon S3 server access logs, in the **Manage system permissions** list, choose **Grant Amazon S3 Log Delivery group write access to this bucket**



When you're done configuring permissions on the bucket, choose **Next**.

7. On the **Review** page, verify the settings. If you want to change something,

choose **Edit**. If your current settings are correct, choose **Create bucket**.

Deleting a Bucket

1. You can delete an empty bucket, and when you're using the AWS Management Console, you can delete a bucket that contains objects. If you delete a bucket that contains objects, all the objects in the bucket are permanently deleted.

Before deleting a bucket, consider the following:

- Bucket names are unique. If you delete a bucket, another AWS user can use the name.
- When you delete a bucket that contains objects, all the objects in the bucket are permanently deleted, including objects that transitioned to the Amazon S3 Glacier storage class.
- If the bucket hosts a static website, and you created and configured an Amazon Route53 hosted zone as described in Create and Configure Amazon Route 53 Hosted Zone: You must clean up the Route 53 hosted zone settings that are related to the bucket as described in Delete the Route 53 Hosted Zone.
- If the bucket receives log data from Elastic Load Balancing (ELB): We recommend that you stop the delivery of ELB logs to the bucket before deleting it. After you delete the bucket, if another user creates a bucket using the same name, your log data could potentially be delivered to that bucket.

  Important

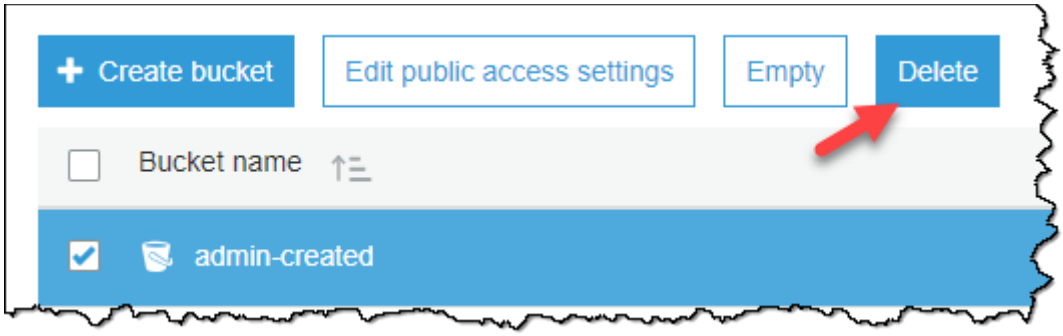
  If you want to continue to use the same bucket name, don't delete the bucket. We recommend that you empty the bucket and keep it. After a bucket is

deleted, the name becomes available to reuse, but the name might not be available for you to reuse for various reasons. For example, it might take some time before the name can be reused, and some other account could create a bucket with that name before you do.

To delete an S3 bucket

1. Sign in to the AWS Management Console and open the Amazon S3 console at https://console.aws.amazon.com/s3/.

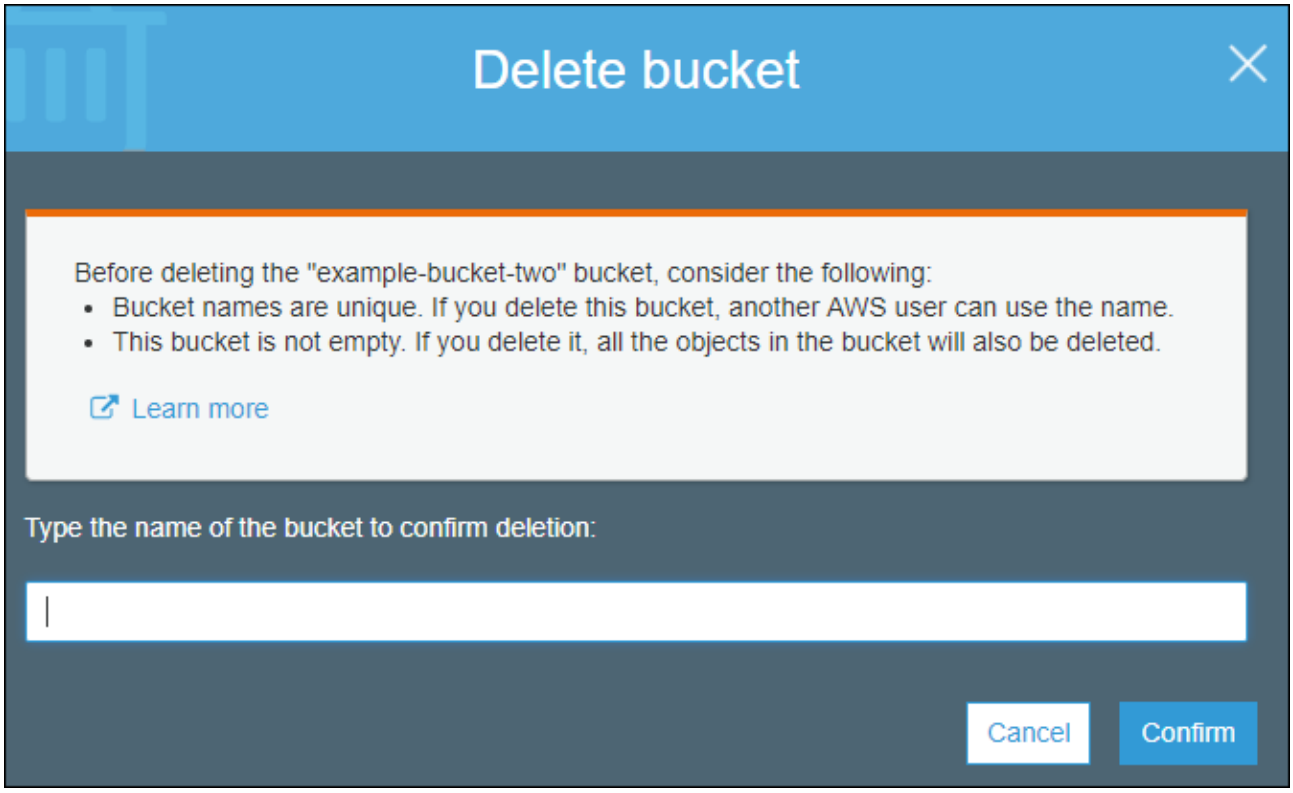
2. In the **Bucket name** list, choose the bucket icon next to the name of the bucket that you want to delete and then choose **Delete bucket**.



3. In the **Delete bucket** dialog box, type the name of the bucket that you want to delete for confirmation, and then choose **Confirm**.

Note

The text in the dialog box changes depending on whether the bucket is empty, is used for a static website, or is used for ELB access logs.

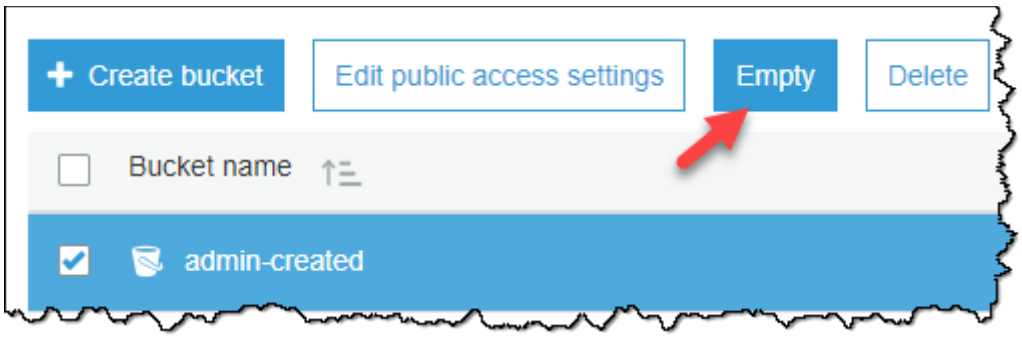


How Do I Empty an S3 Bucket?

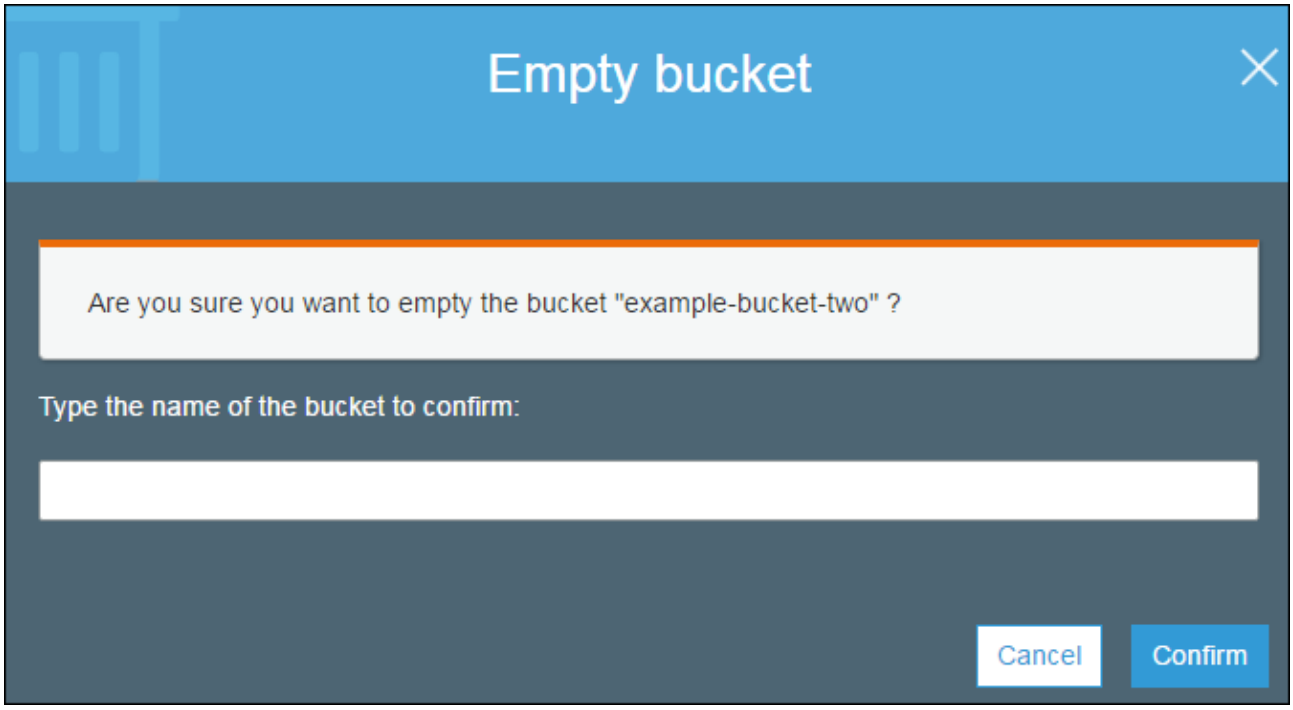
You can empty a bucket, which deletes all of the objects in the bucket without deleting the bucket. When you empty a bucket with versioning enabled, all versions of all the objects in the bucket are deleted.

To empty an S3 bucket

1. Sign in to the AWS Management Console and open the Amazon S3 console at https://console.aws.amazon.com/s3/.
2. In the **Bucket name** list, choose the name of the bucket that you want to empty and then choose **Empty**.



3. In the **Empty bucket** dialog box, type the name of the bucket you want to empty for confirmation and then choose **Confirm**.



https://docs.aws.amazon.com/AmazonS3/latest/user-guide/create-bucket.html