S3 – Simple Storage Service

It is one of the webservice from Amazon use for storage purposes.

You Can store and retrieve any amount of data at any time.

Important terminologies which will go through:

1. Buckets
2. Objects
3. Permissions
4. Type of Access
5. Keys
6. Regions

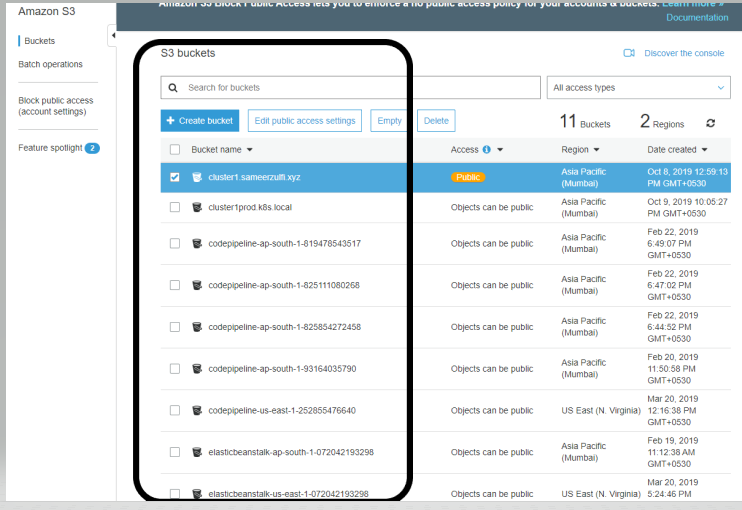
What will we do with Amazon S3 Service?

No Rocket science. Only we store and retrieve the data. Few tasks which we will try to explain here are the following.

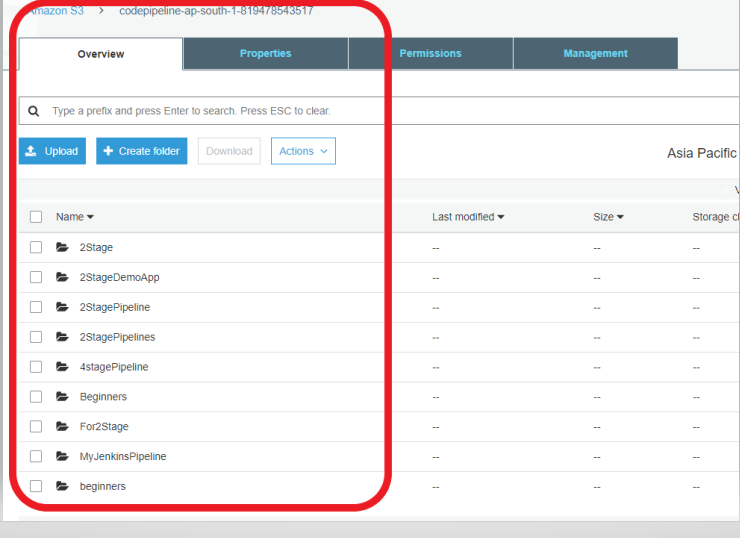
1. Creating buckets
2. Storing data
3. Downloading data
4. Permissions

What is Bucket?

It is a box/container for objects stored in Amazon S3. Every object is contained in a bucket.

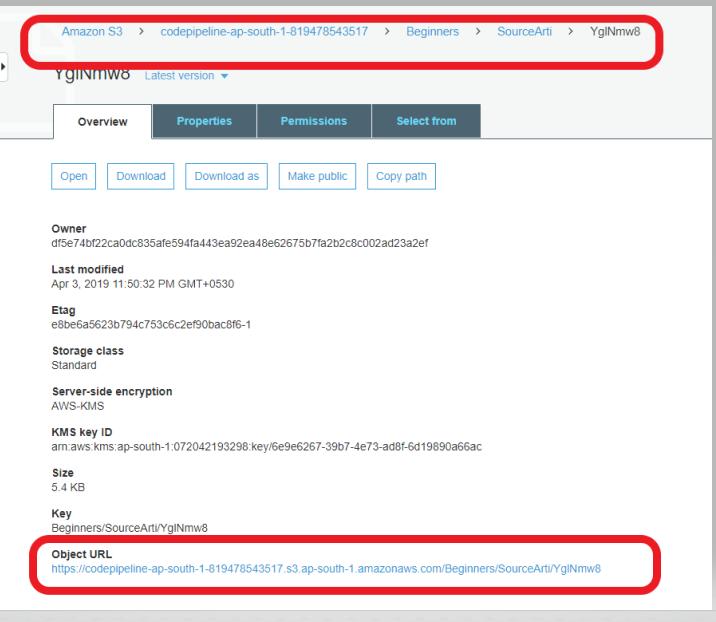


If you see the above image, we have lot of buckets. And the below image has lot of objects inside one of the buckets.



There will be a URL assigned to each of your object.

See the below image



Now you must have understood what exactly are we talking about?

Next is about the Data Consistency Model.

Remember these words “ PUTS – DELETES “.

For any new object/new file which you are uploading in S3 will have “Read After Write consistency.

i.e., any PUTS of new objects will have RAWc

which means “ As Soon as you uploaded the data, it will be available to everyone”

Next is about “ Overwrite PUTS & DELETES “.

For any old objects needs to be updated / deleted, those changes will not be available immediately. This is opposite to the earlier one which we discussed.

if a process replaces an existing object and immediately attempts to read it. Until the change is fully propagated, S3 might return the prior data

if a process deletes an existing object and immediately attempts to read it. Until the deletion is fully propagated, S3 might return the deleted data

if a process deletes an existing object and immediately lists keys within its bucket. Until the deletion is fully propagated, S3 might list the deleted object.

Let’s talk about S3 Classes:

Amazon offers various types of storage classes for various use cases.

1. S3 Standard ( mainly used for frequently accessed data )
2. S3 Standard\_IA ( mainly use for non-frequent and long-lived data )
3. Glacier ( Long-term archive )

Let’s talk about Amazon S3 BUCKETS:

Couple of terms need to be known here.

Resources means ( either we can refer it to objects or buckets ).

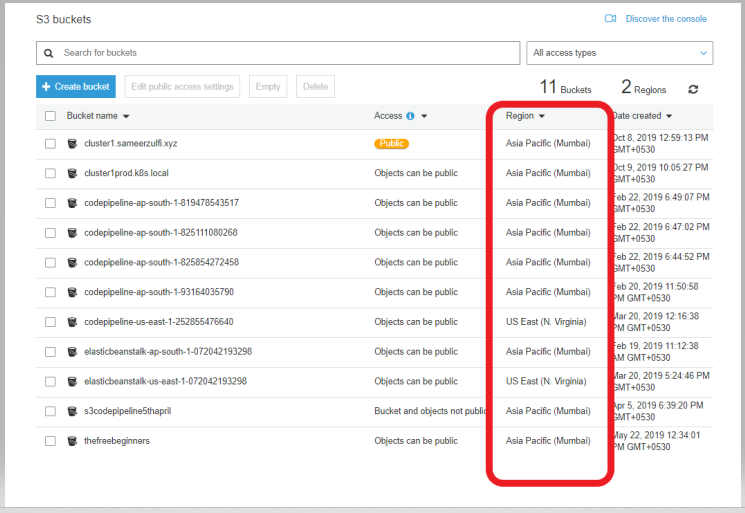
Now S3 buckets name is unique. There is no two buckets having the same name. i.e, It acts like DNS names.

Just like [www.google.com](http://www.google.com) is unique, you can’t register a new domain name with the same [www.google.com](http://www.google.com)

This means that after a bucket is created, the name of that bucket cannot be used by another AWS account in any AWS Region until the bucket is deleted

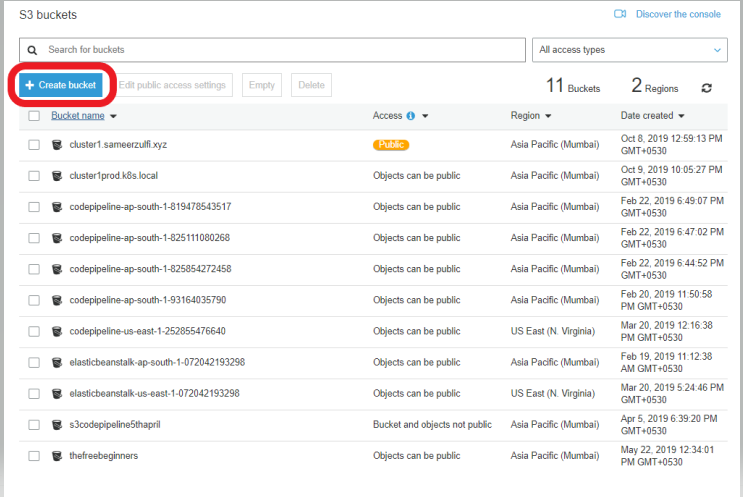
Ok. Now Make sure you create a bucket where you are based out of to avoid any latency and other issues.

Please see the below diagram

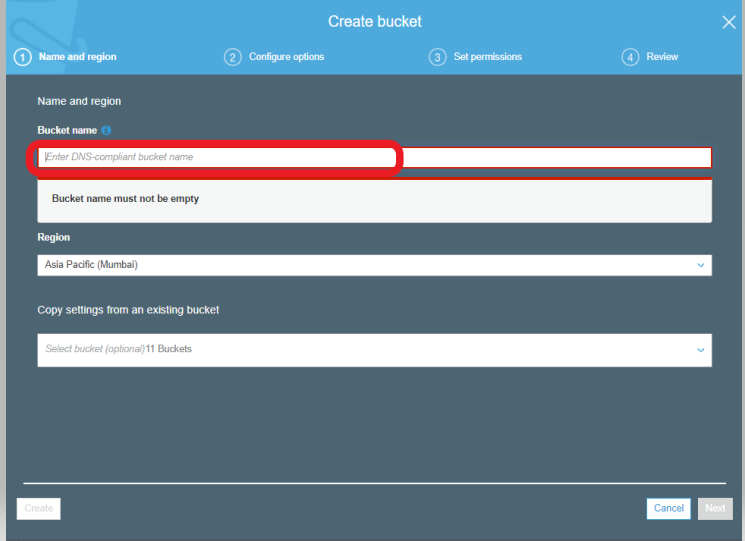


Let us show you HOW TO CREATE A BUCKET.

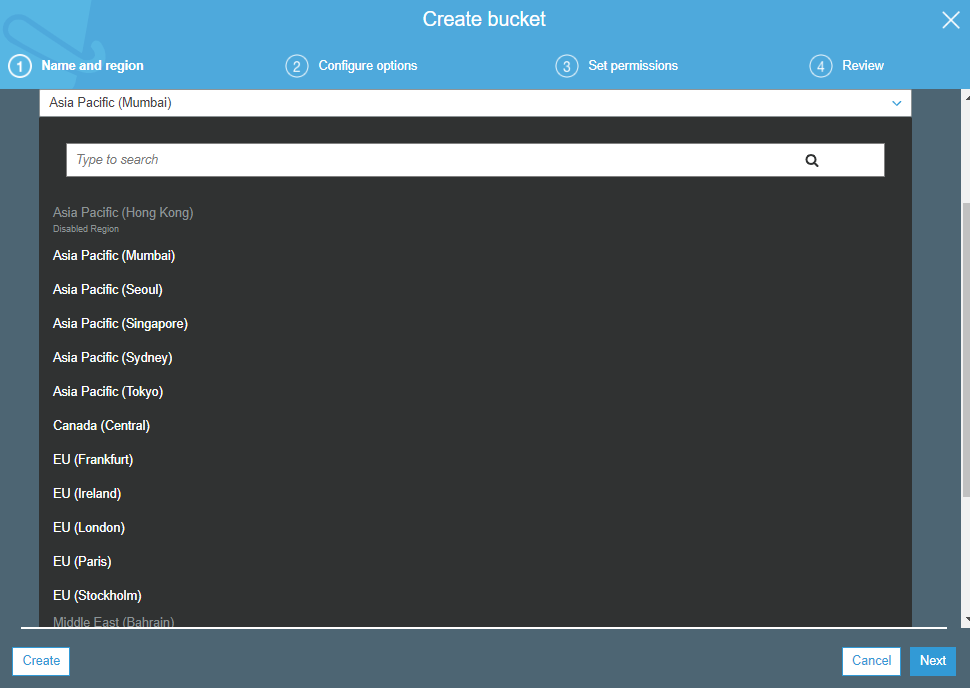
Click on Create bucket



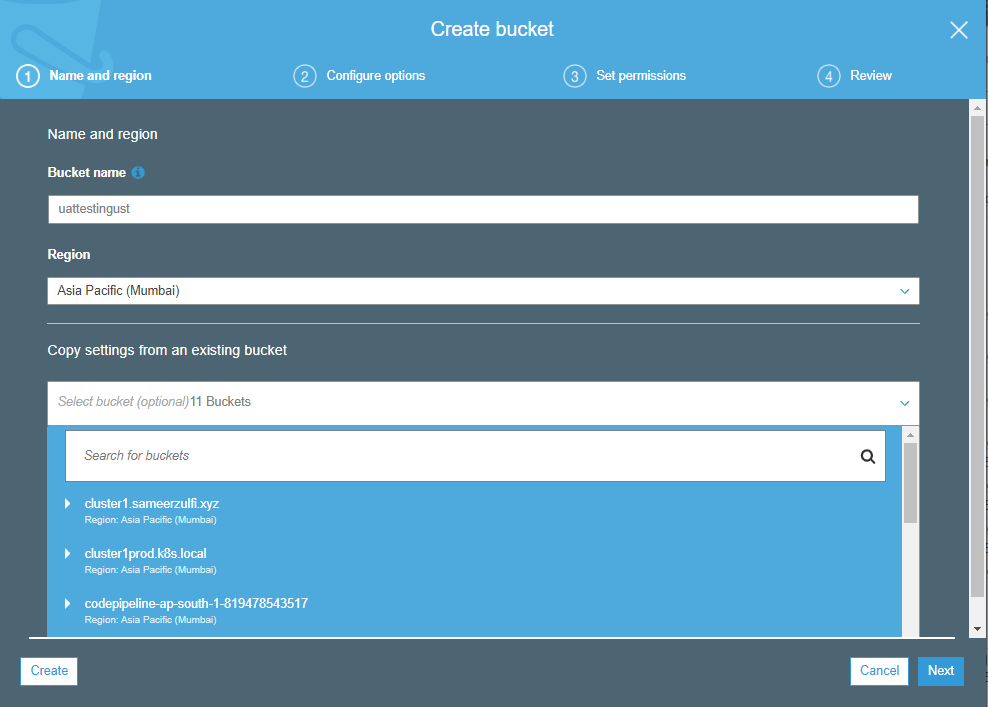
The Bucket name should be DNS-complaint. In simple terms, it should be unique across different accounts. It’s not within your account. Other Organization/Person shouldn’t carry the same name.



After typing your unique bucket name, choose the region which is closest to you. We are choosing Mumbai



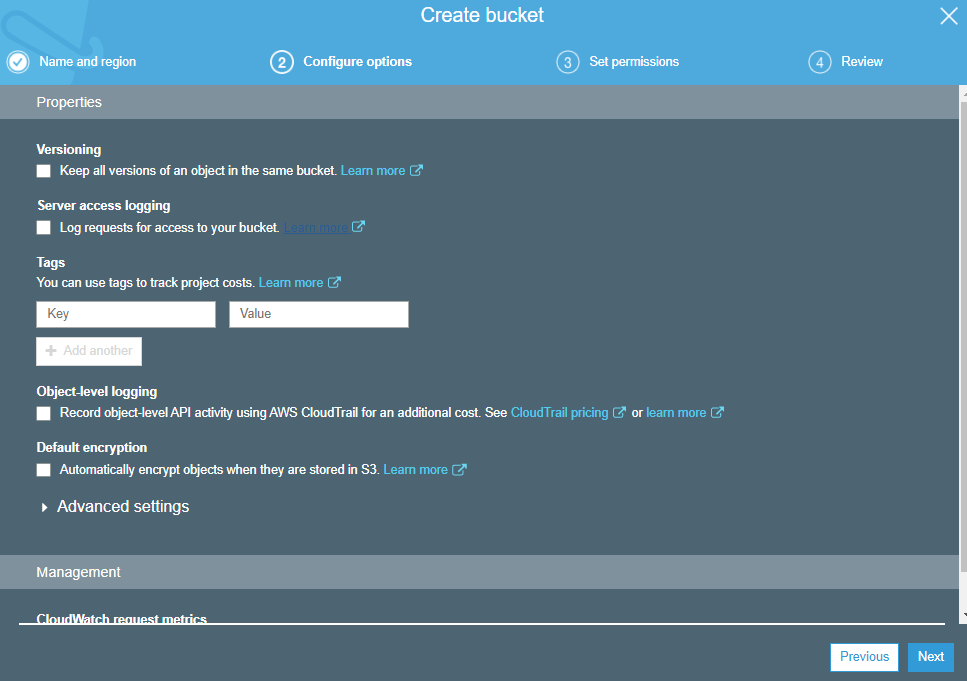
You Can copy the settings from an existing bucket or can leave blank.



Click Next.

So we have chosen the name of the bucket, which region you want to create it. Next will be the configuration part.

If you understand the below screen, it will be more than enough for he basics of S3.



Versioning ? What is it ?

Let’s see the first check box. Versioning enables you to keep multiple versions of an object in one bucket

Server Access Logging :

Server access logging provides detailed records for the requests that are made to a bucket. There is no extra charge for enabling server access logging on an Amazon S3 bucket

Tags:

To track the storage cost or other criteria for individual projects or groups of projects,label your Amazon S3 buckets using cost allocation tags. A cost allocationtag is a key-value pair that you associate with an S3 bucket.

Object Level Logging:

Amazon S3 is integrated with AWS CloudTrail, a service that provides a record of actionstaken by a user, role, or an AWS service in Amazon S3.

If you create a trail, you can enablecontinuous delivery of CloudTrail events to an Amazon S3 bucket, including eventsforAmazon S3. If you don't configure a trail, you can still view the most recent eventsin the CloudTrail console in Event history. Using the informationcollected by CloudTrail, you can determine the request that was made to Amazon S3,the IPaddress from which the request was made, who made the request, when it was made, andadditional details.

Default Encryption:

Data protection refers to protecting data while in-transit (as it travels to and fromAmazon S3)and at rest (while it is stored on disks in Amazon S3 data centers). You can protectdata intransit using Secure Sockets Layer (SSL) or client-side encryption. You have thefollowing options for protecting data at rest in Amazon S3:

•Server-Side Encryption – Request Amazon S3 to encrypty our object before saving it on disks in its data centers and then decrypt itwhen you download the objects.

•Client-Side Encryption – Encrypt data client-sideand upload the encrypted data to Amazon S3. In this case, you manage the encryption process, the encryption keys, and related tools.

Object Lock:

Locking objects means you can store objects using a write-once-read-many (WORM) model. You can use it to prevent anobject from being deleted or overwritten for a fixed amount of time or indefinitely. For this to be turned ON, you need to make sure Versioning is enabled as the first option.

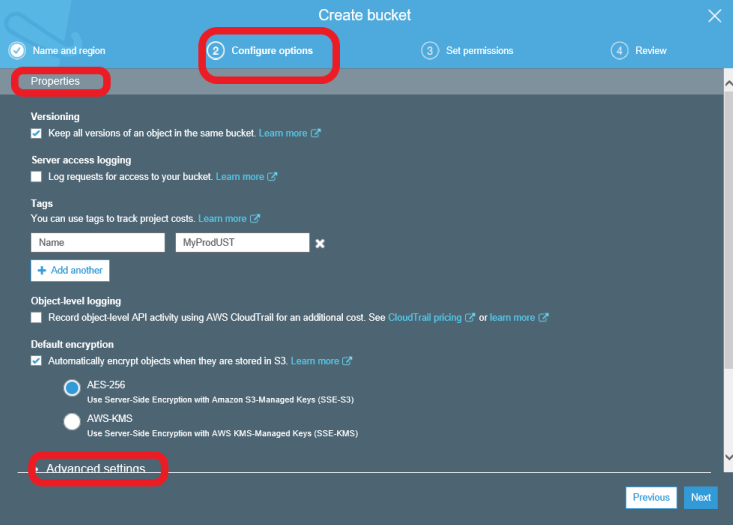
Let’s talk with screenshots.

Under Configure Options, there are two sections classified :

1. Properties
2. Advanced Settings
3. Management

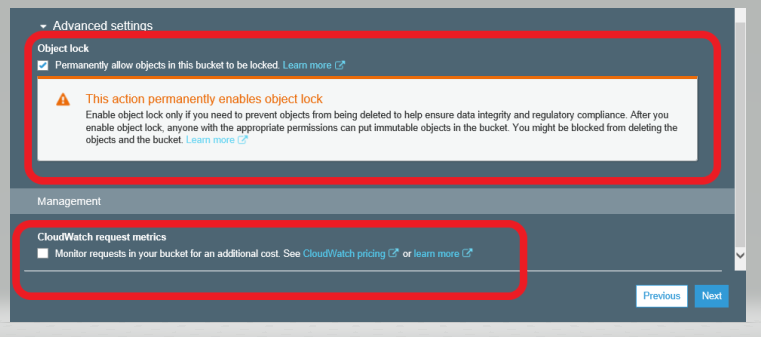
Under Properties, you have many options which I have discussed earlier.

Under Advanced settings, you have



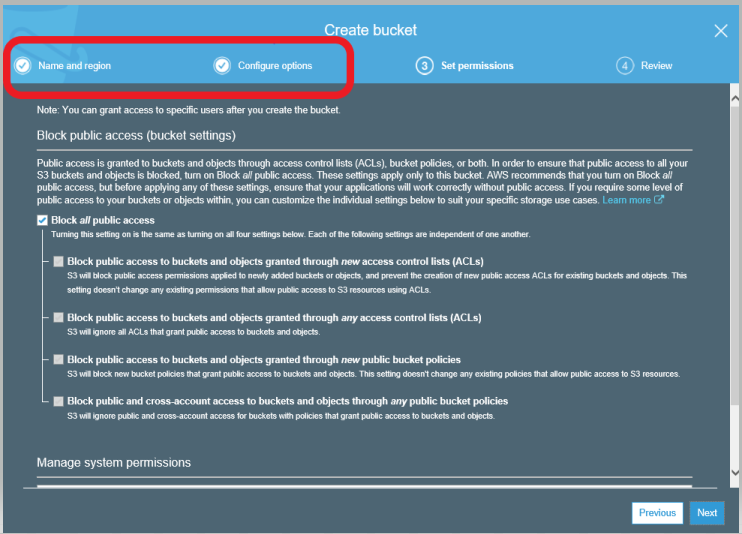
Object Lock, which we have described earlier.

Under Management, you can enable the cloudwatch metrics for additional cost.



Click Next,

If you see the below screen, the first two options are having a tick mark means, it’s completed, now we are on “Set Permissions” screen.



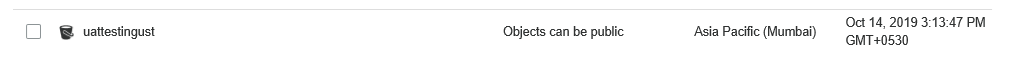
Permissions is all about different ways of giving access to different users

Please check the below screenshot to under the public access settings. It’s newly updated by Amazon.

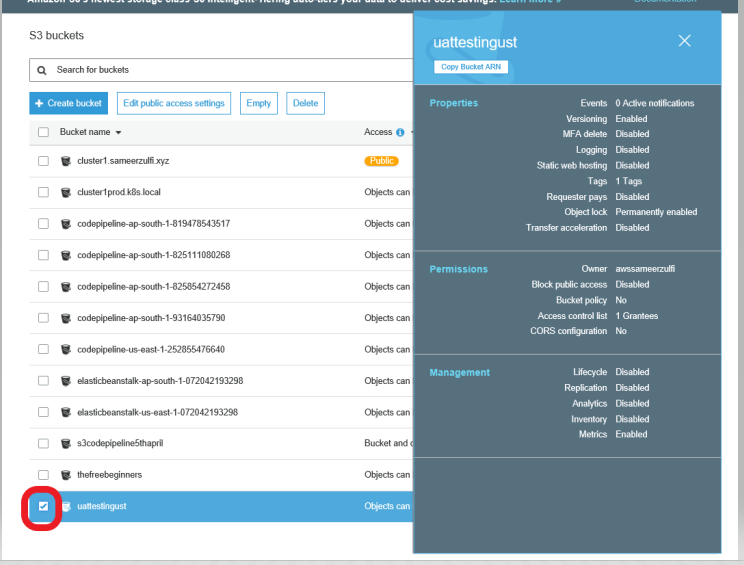


After you hit Next, it will give you the Review page, then click “Create bucket”.

Your bucket looks like this.



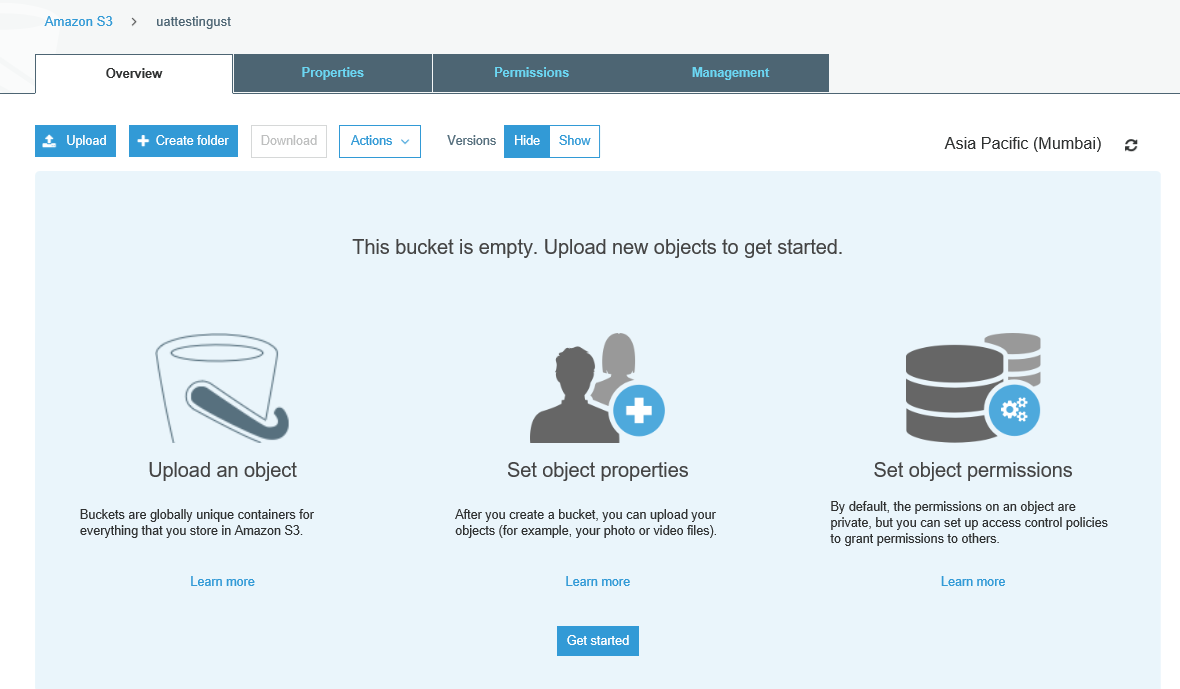
When you put a checkmark, on the right you will get the information page, which will explain you about the properties, permissions and management whatever we had set it up.



Let’s click on the bucket,

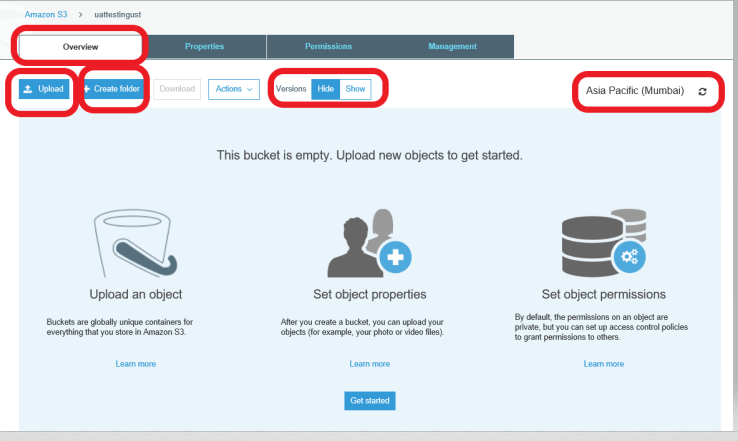


You will get the below page,



Initially it will be empty. Please upload new objects to get started.

On the below screen, you could see the below screen, please see the red marked items.



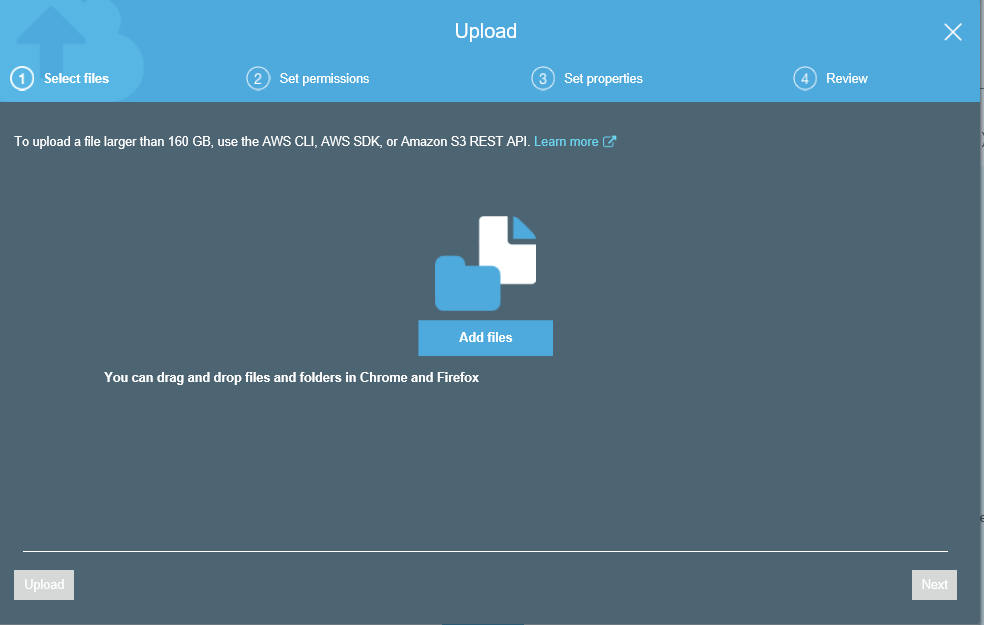
Upload = to add a file to this bucket

Create Folder = create a folder inside the bucket

Versions = Since we have enabled the version, we could see show or hide it

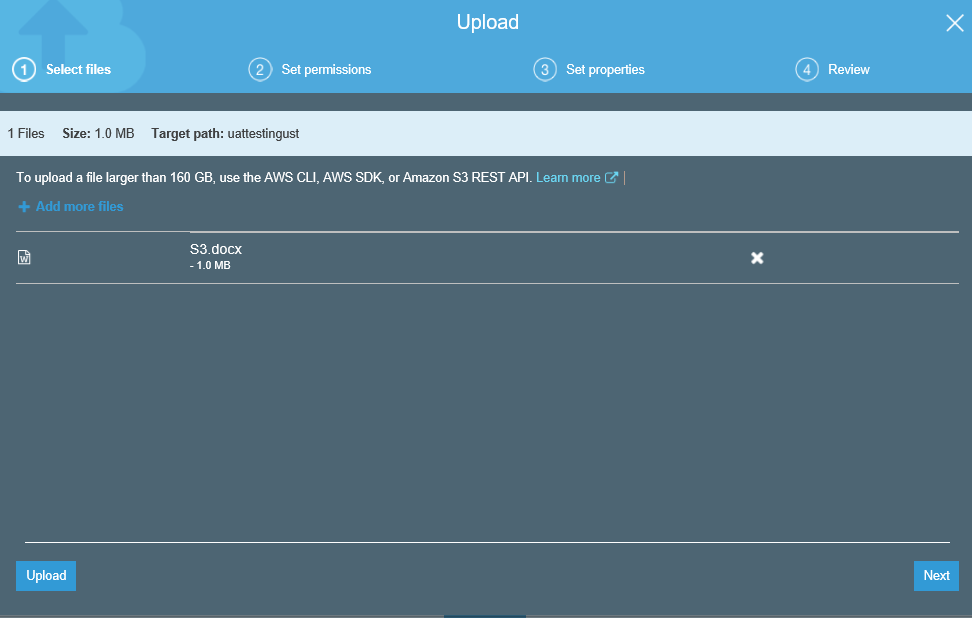
Region = Asia Pacific

Let’s try to upload a file by clicking on Upload button.



You will get the above screen, If you see it closely, if you upload a larger file which is more than 160 GB, use the AWS CLI, AWS SDK or S3 REST API.

Once you added the file, it will automatically tells you the size

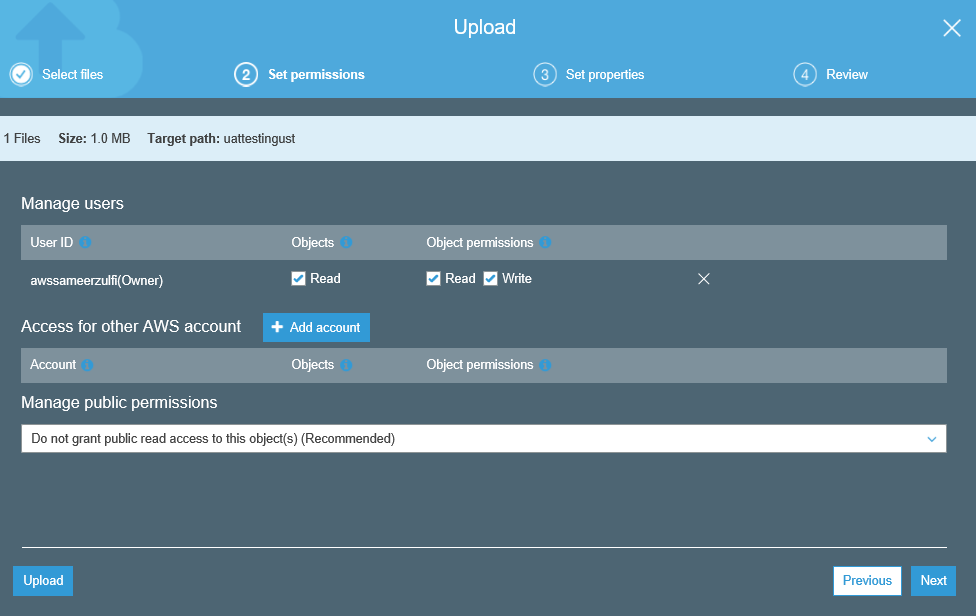


Hit Next,

On the below screen, you would get the Users who will be able to write or read the objects.

Deault will have the owner ID who has created it.

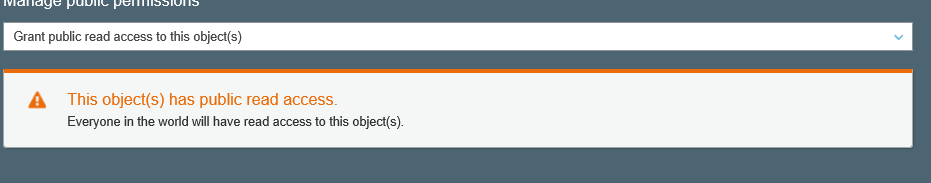
Under Access for other AWS Account. Please click on Add Account and provide the Account ID.



Under Manage Public permissions, you can make it public or not public

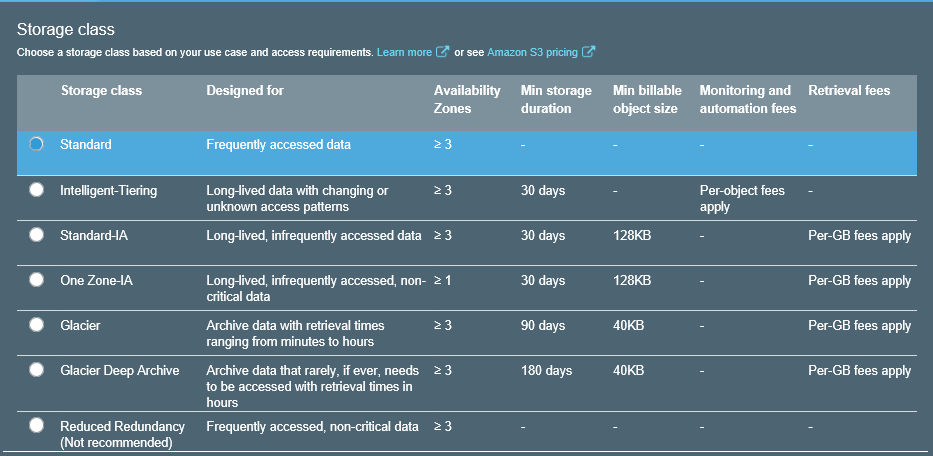


When you choose grant public, it will give you a disclaimer like below



Click Next

You will get the Storage class page,



Standard - Frequently accessed data

Intelligent Tiering - Long-lived data

Standard IA - Long lived data

One Zone IA - Long Lived data

Glacier - Archived data

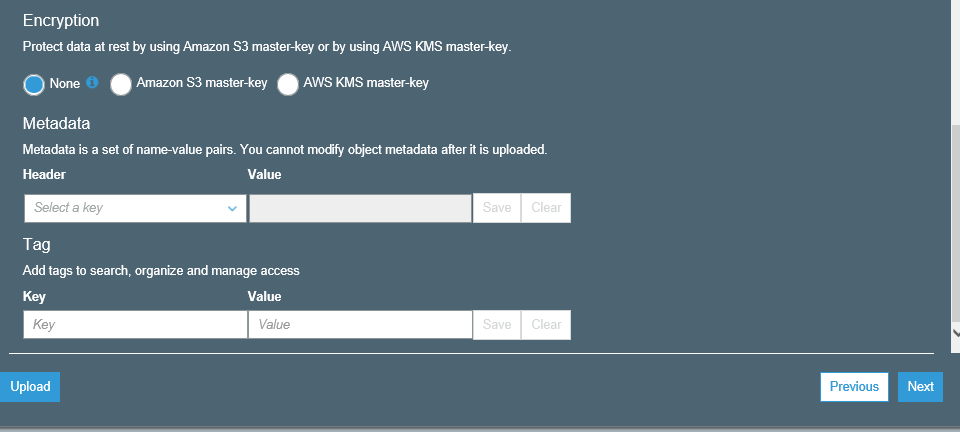
Glacier Deep Archive - Archived data

Reduced Redundancy - Frequently accessed, non-critical data

Please see the above diagram for each description and the AZ.

There is one storage class which has less AZ is “One Zone-IA”

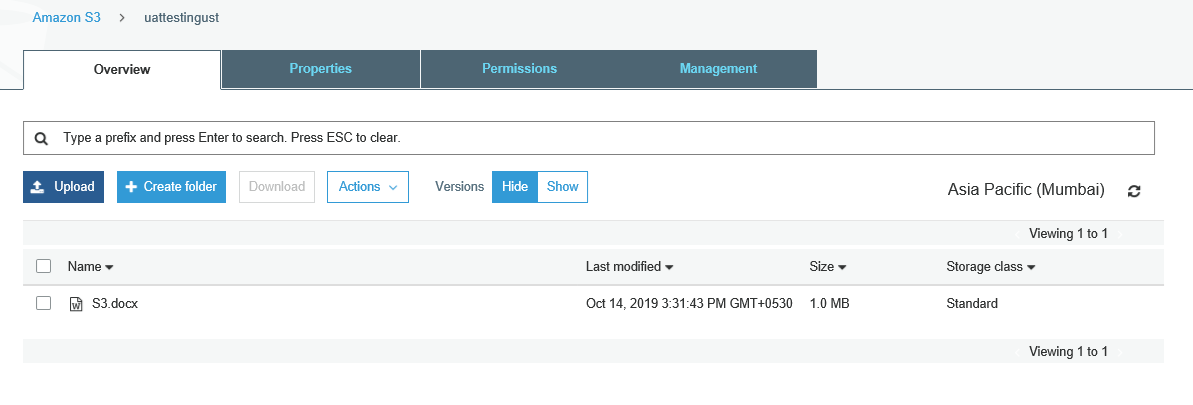
If you scroll down the screen, you would get the below options



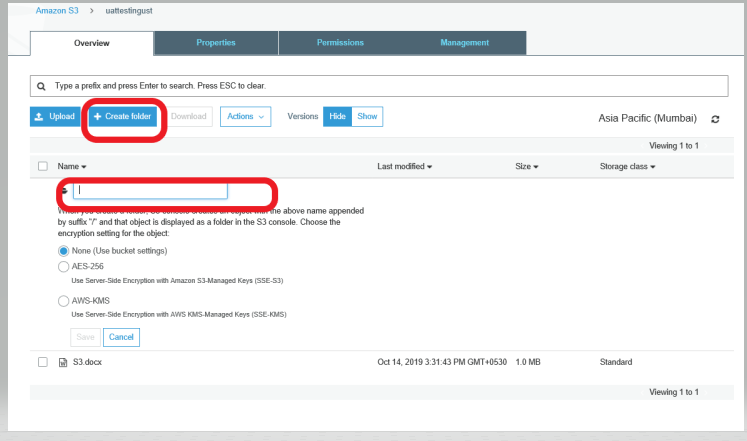
You can choose the Encryption, Metadata, Tags as per the screenshot.

Click Next and then you will get a Review.

Once you hit Upload, you will get the below page.

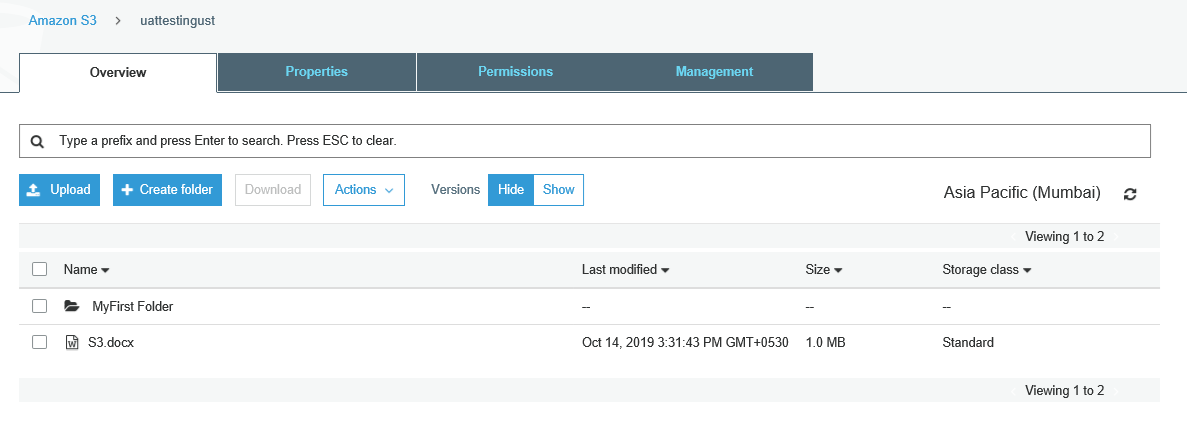


We have seen how to upload the file, let’s see how to create the folder.

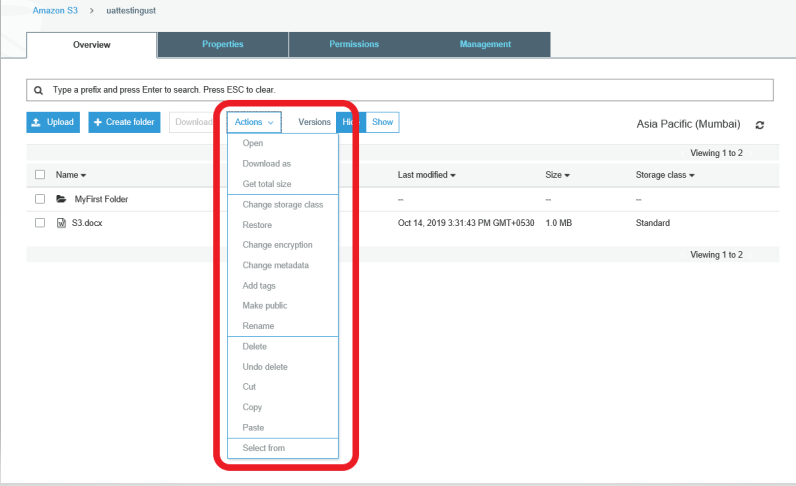


You can choose the encryption for the folder

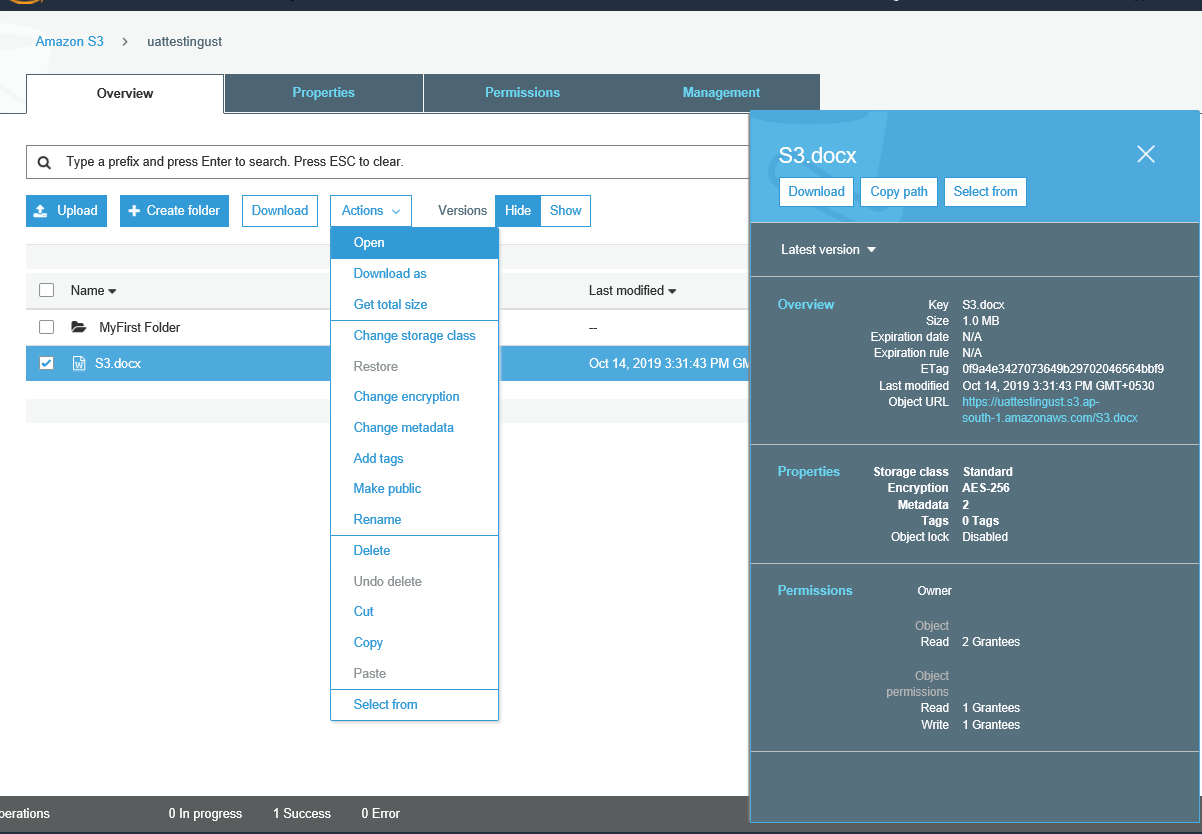
I have a Folder and a file.



Without selecting any file or folder, you would get all the options grayed out under Actions

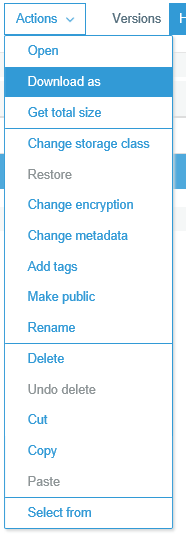


If you select a file, you would get the options enabled.

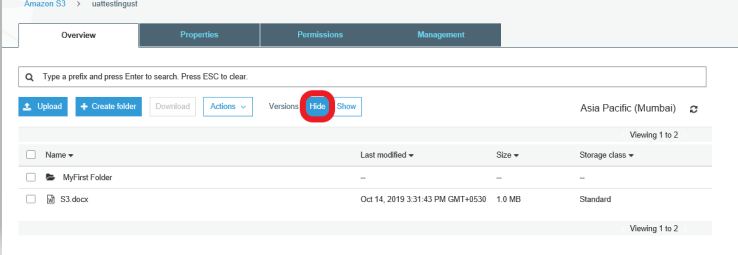


You can Open, Download, Download As and important is you can change the storage class.

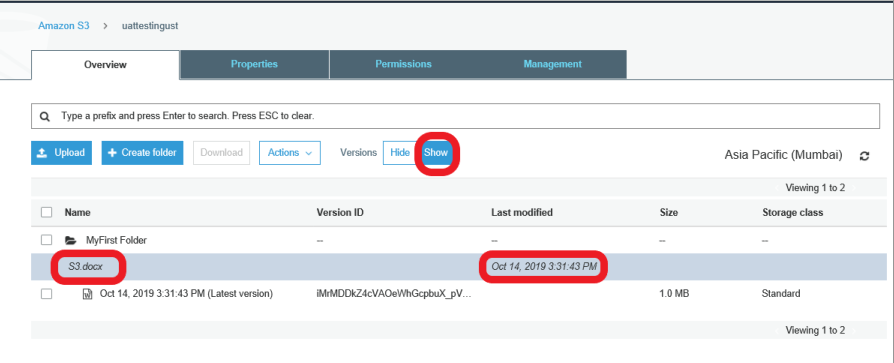
There are lot of important options, which you have to look for



Now the version shows Hide



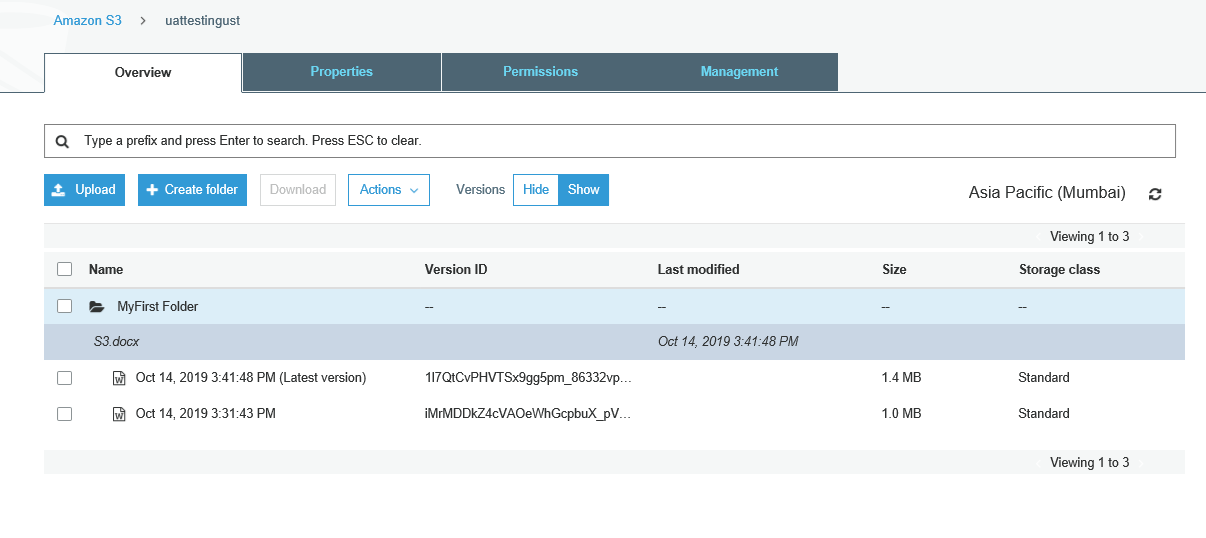
Let’s turn on the SHOW and see what happens.



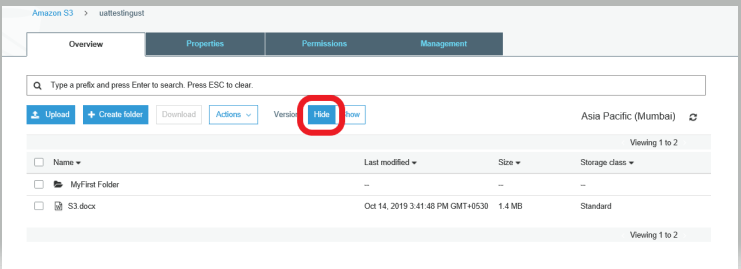
You would get the option expanded. Everytime you keep modifying the S3.docx, it keeps adding the entry.

Assume I have modified the file and uploading it again.

See the below figure, you have two versions under S3.docx. The first version size is 1 MB and the 2nd is 1.4 MB



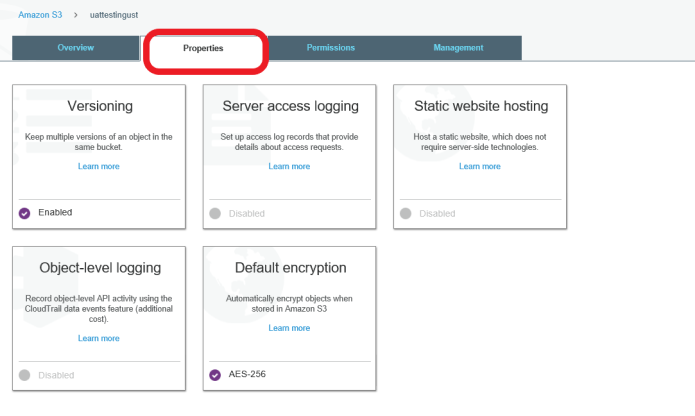
Now if you hide it,



Now see the difference, it will not show you entire versions. It just gives you the latest file. If you want the older version, you need to Show it and choose it.

This applies to Folder as well.

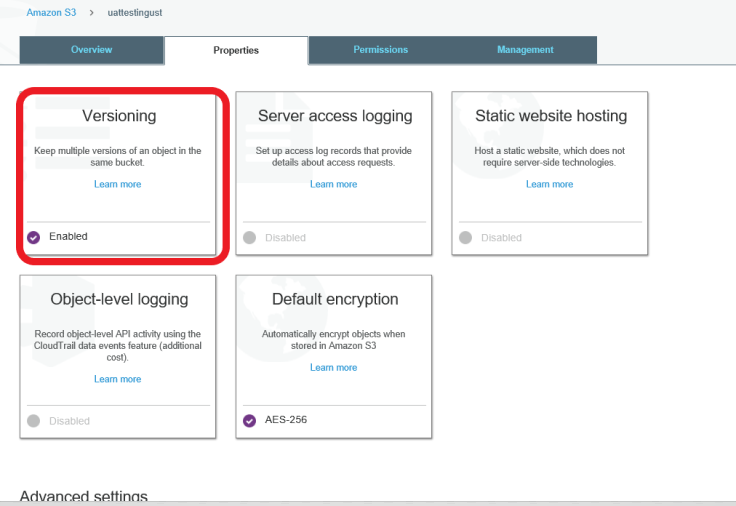
Let’s go to Properties :



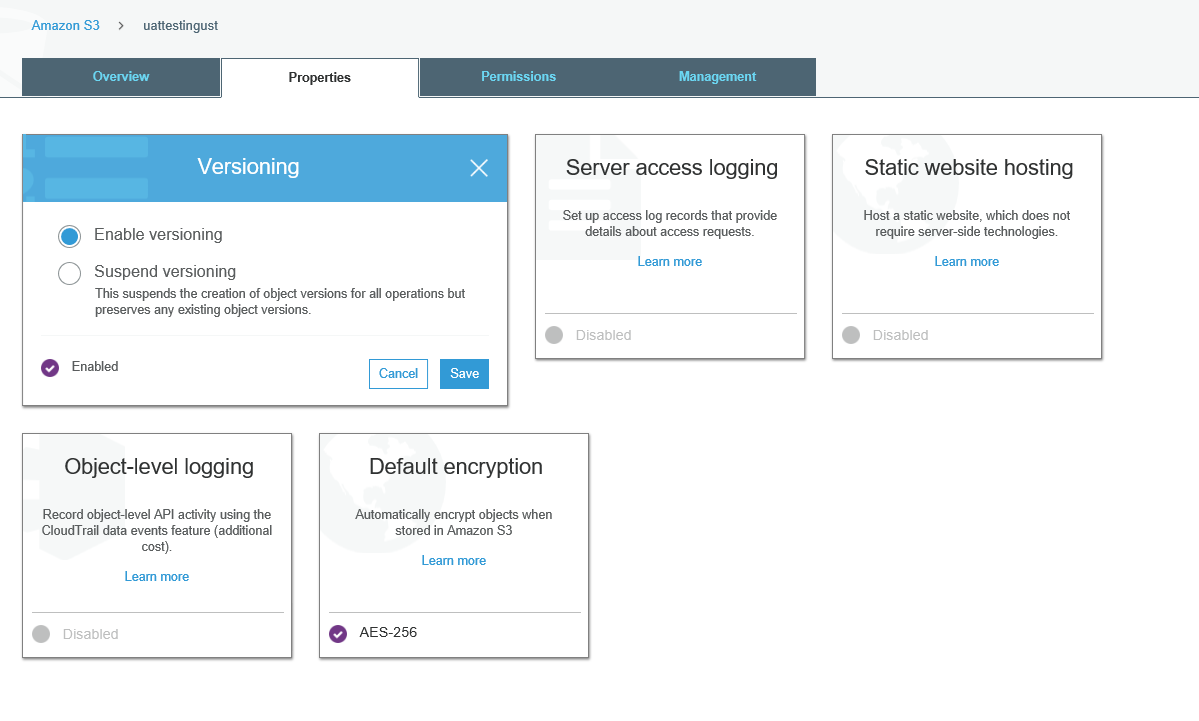
If you see, versioning is enabled and Default encryption as well.

These are the properties settings which we have given when we created the file/folder.

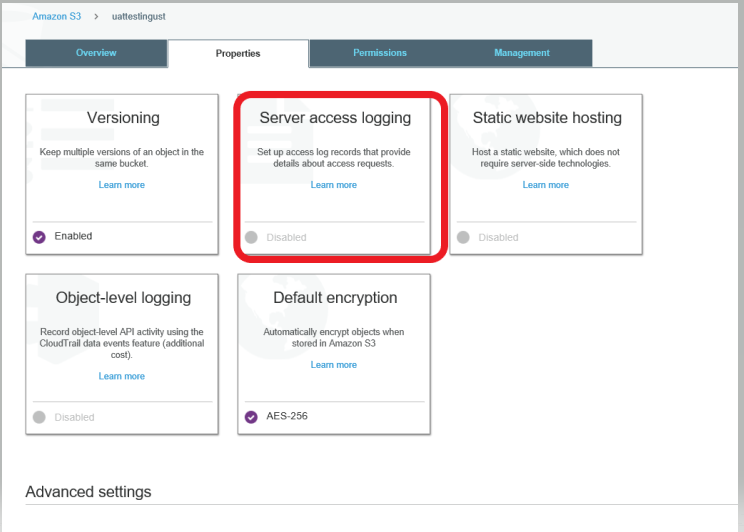
Let’s click on Versioning



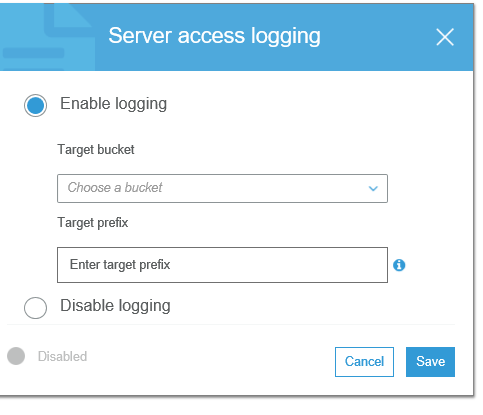
It’s already enabled, there is no option to disable. Only you can suspend it. But the old objects will remain versioning even though if you do suspend.



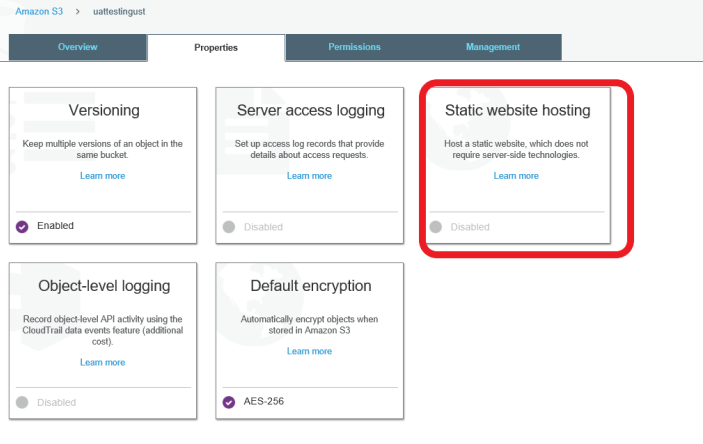
Click on Server Access Logging



You Can enable by clicking the below option

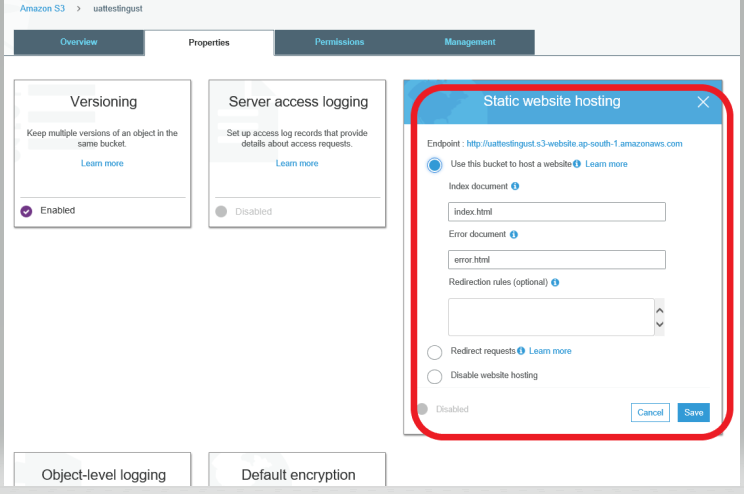


You can host a static website which doesn’t require any PHP or server technologies.

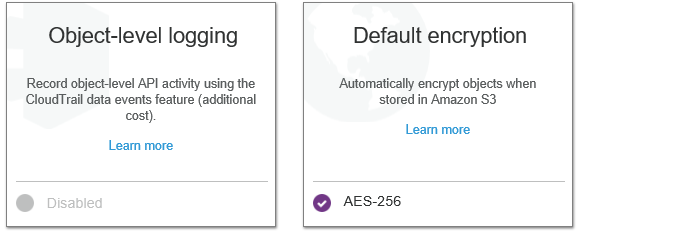


Make sure you upload your “index.html” and “error.html” inside this bucket.

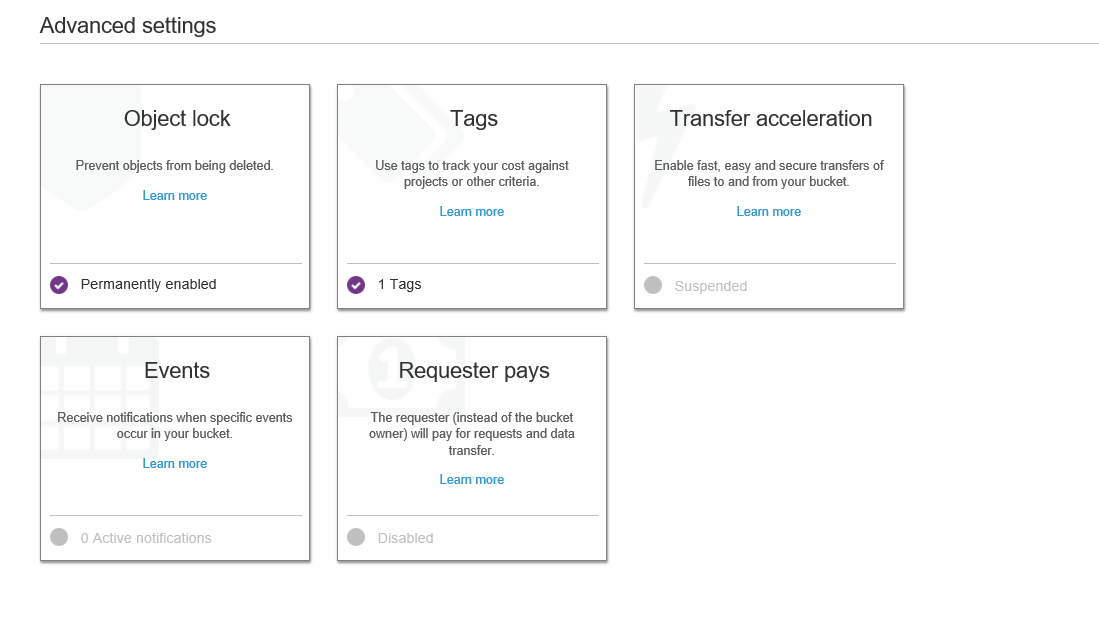
And you have the endpoint URL as well.



These two objects are already we had set it up when we created the file/folder



Let’s talk about Advanced settings :

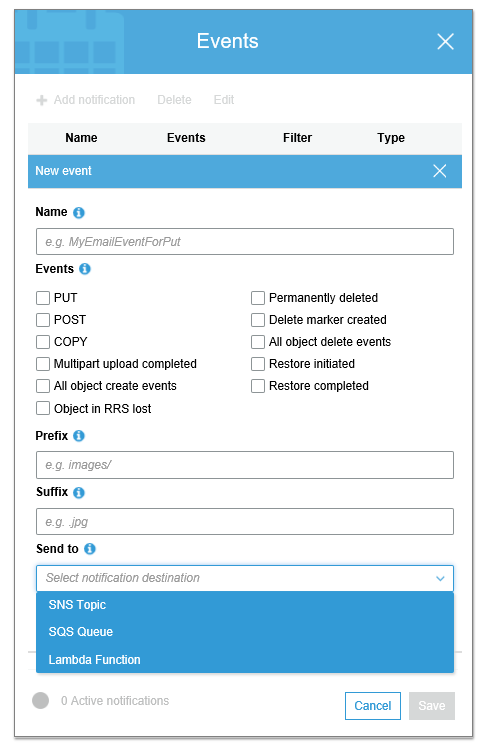


Object lock will help you prevent from being deleted.

Tags to track you cost against porjects

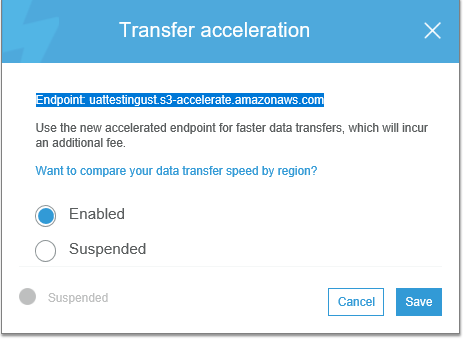
Important is Events.

Under Event, you have lot of options which would primarirly for the amazon s3 to send the notification if someone has uploaded/deleted/modified etc., It’s based on the events which will have the trigger to SNS Topic, SNS queue or Lambda Function



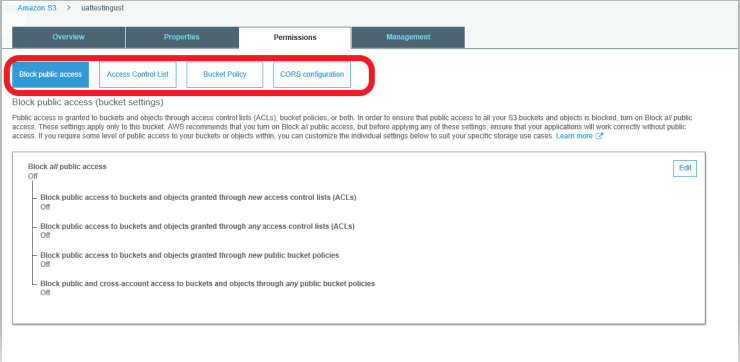
Transfer Acceleration :

◦Instead of directly uploading the file to S3 bucket, you will get a distinct URL that will upload the data to the nearest edge location which in turn transfer the file to S3 bucket. The distinct URL would look like:



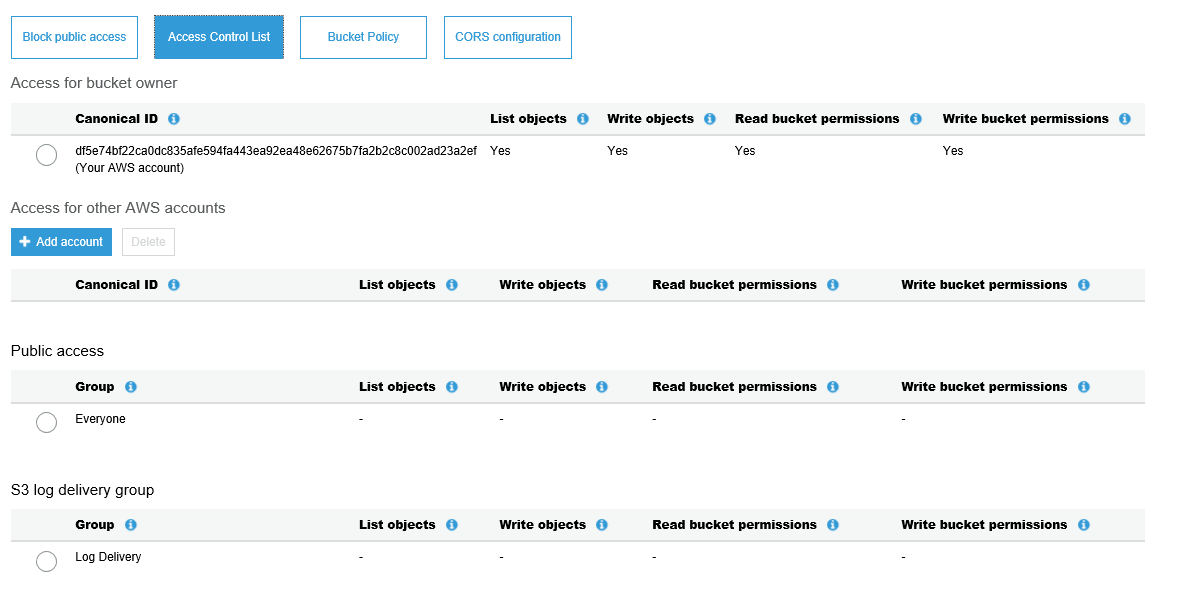
Let’s go to Permissions:

You need to be very careful with the red marked items.



Public access is granted to buckets and objects through access control lists (ACLs), bucket policies, or both. In order to ensure that public access to all your S3 buckets and objects is blocked, turn on Block all public access. These settings apply only to this bucket. AWS recommends that you turn on Block all public access, but before applying any of these settings, ensure that your applications will work correctly without public access. If you require some level of public access to your buckets or objects within, you can customize the individual settings below to suit your specific storage use cases

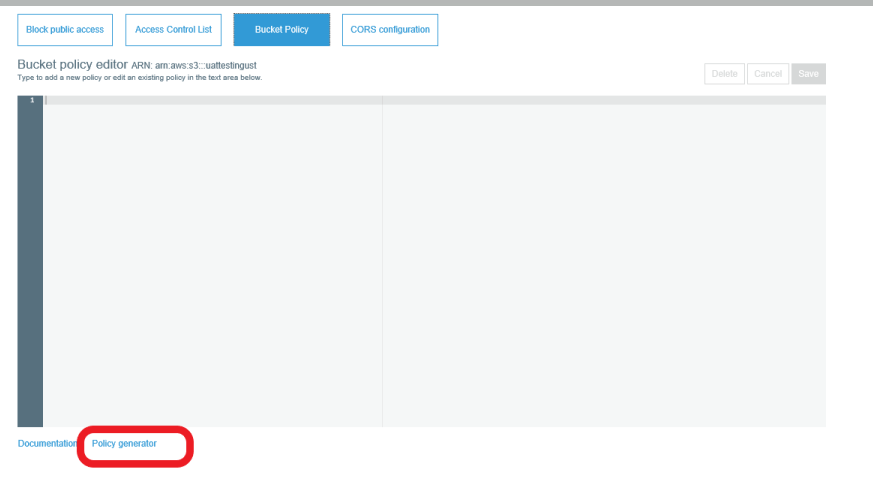
Let’s talk about Access Control List :



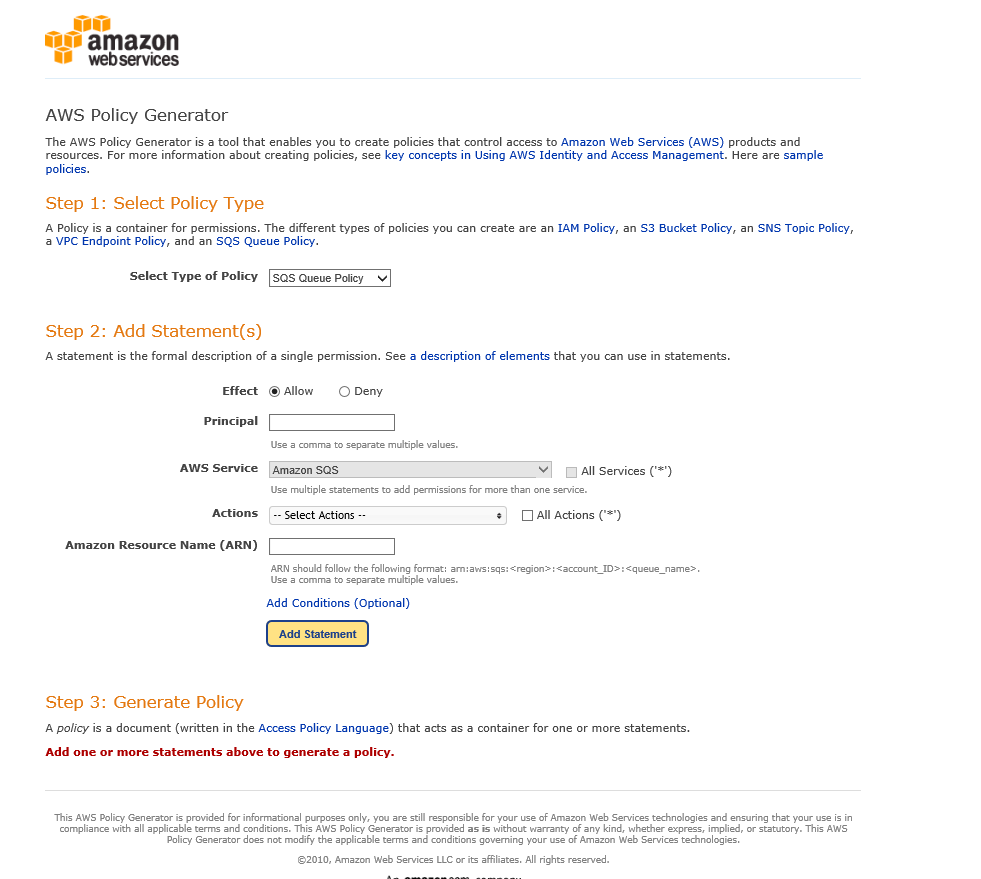
In this one, Bucket Owner Access is defined, Also For Other AWS Accounts, also for Public Access and also for S3 log delivery group.

This is ACL.

Let’s talk about Bucket Policy which is applicable to the entire bucket.

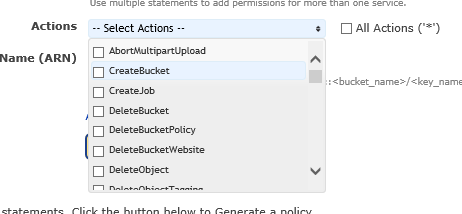


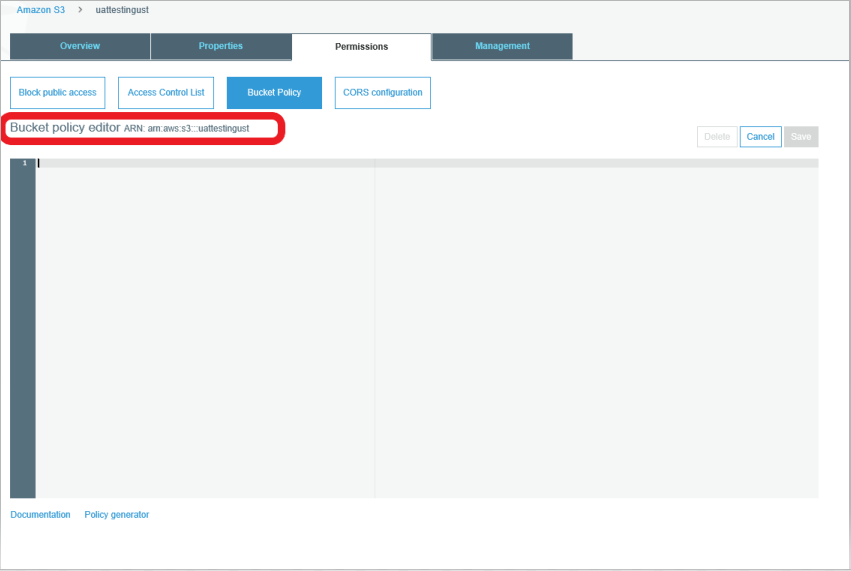
We have a policy generator, where it will open up a new page



Will have detail instructions on how the bucket objects should react.

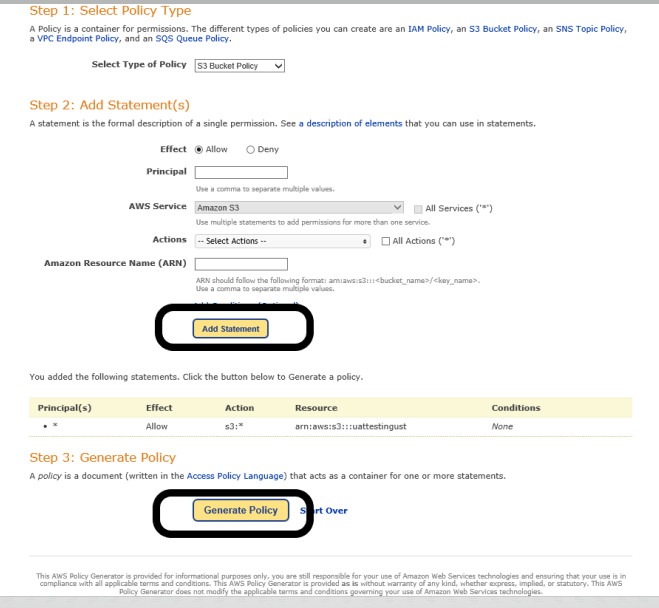
Bucket Policy is a container for permissions. We have different policies like IAM Policy, S3 bucket policy, SNS topic policy, VPC endpoint policy, SQS queue policy.

1. Select the type of Policy as “S3 bucket Policy”
2. Effect can be Allow or Deny
3. Principal means if you choose “\*” , open to public. You can restrict access to particular to particular group or aws accounts etc.,
4. AWS Service, already S3 would have been selected
5. Check the Actions, and choose which operations you want to be selected
6. The ARN – Amazon Resource Name. You can get it on the top

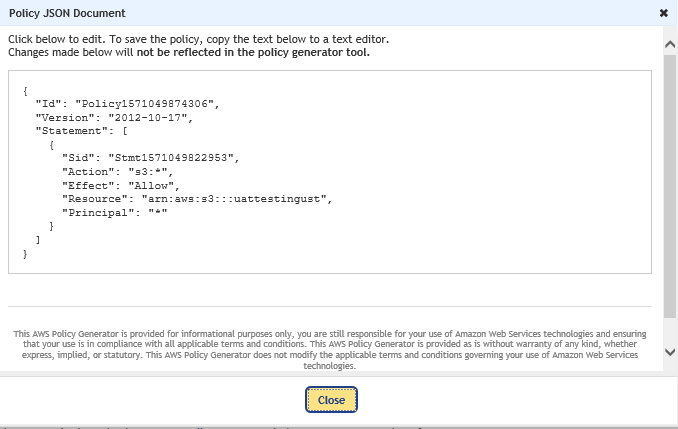


1. For Ex : arn:aws:s3:::uattestingust

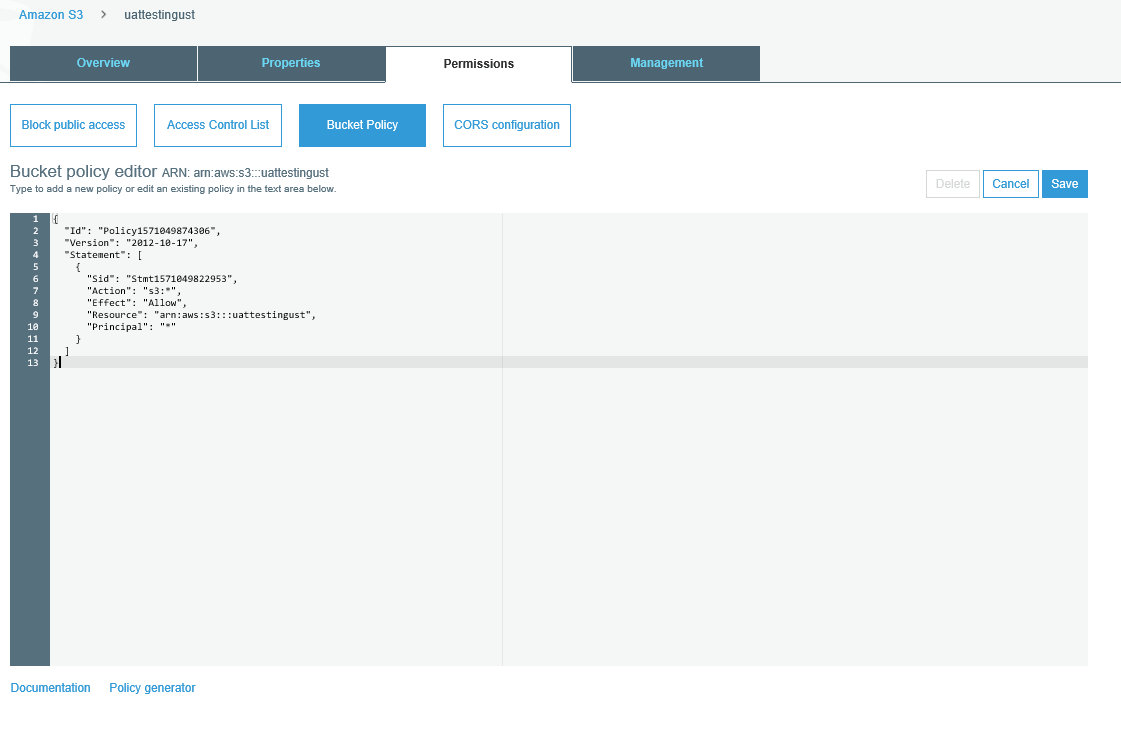
Click on Add Statement



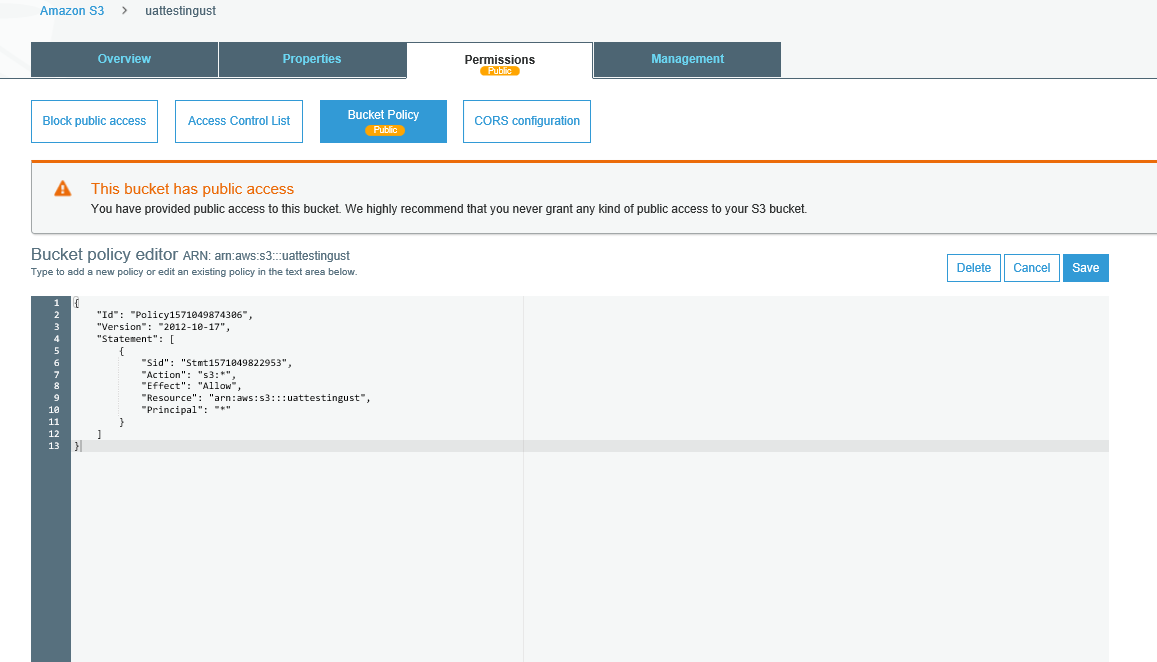
Generate Policy will give the below screen



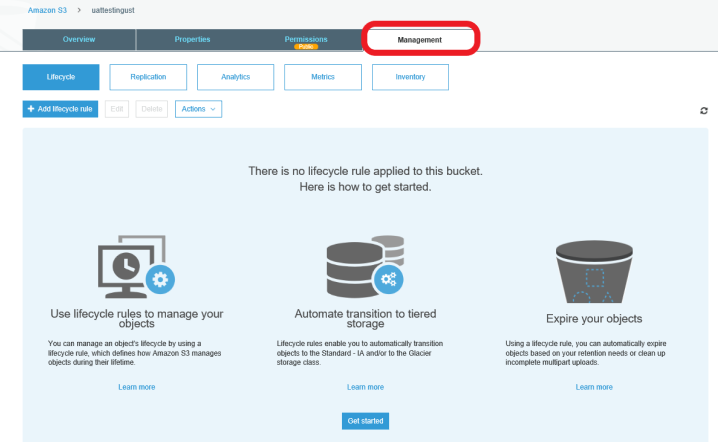
Now Copy the above json statement and paste it here



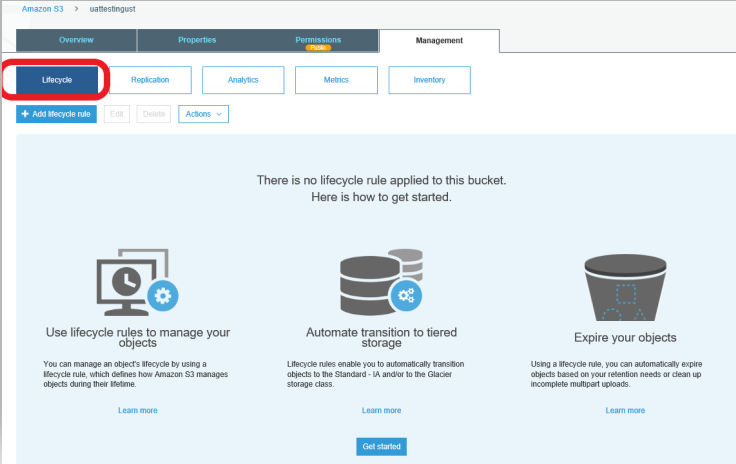
When you click Save, it gives you the below warning which says your bucket has public access.



Now, let’s go to the last tab. Management



By default, you will be at Lifecycle button which is blue color marked highlighted



What do you mean by LifeCycle ?

It is a group of rules which will manage the lifecycle of an object which you have uploaded.

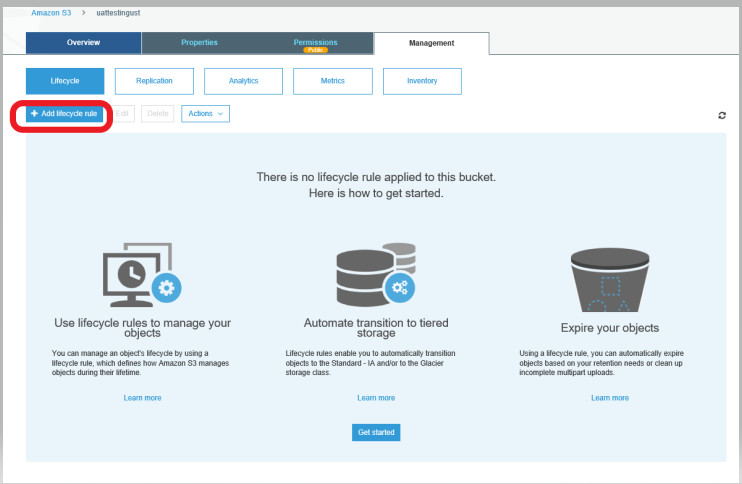
When will you use Lifecycle ?

For ex, you need some data for an application for 30 days and later it can be deleted. You can apply this rule.

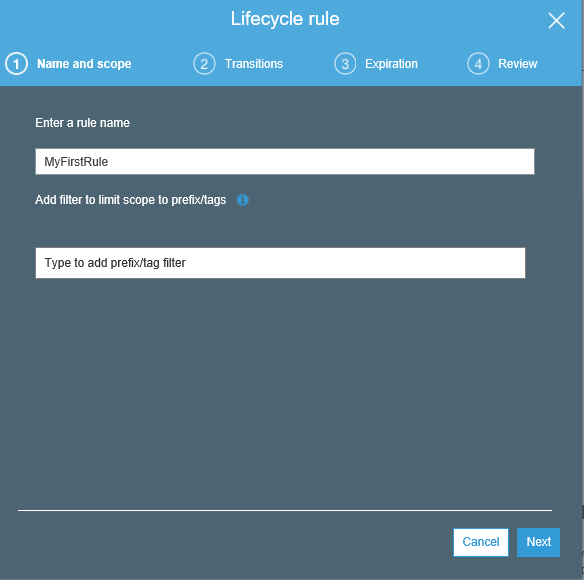
Some documents needs to be accessed frequently for certain days.

We can automatically transition the data from one storage class to other storage class after certain period of time.

Let’s see the options under “Add LifeCycle Rule”

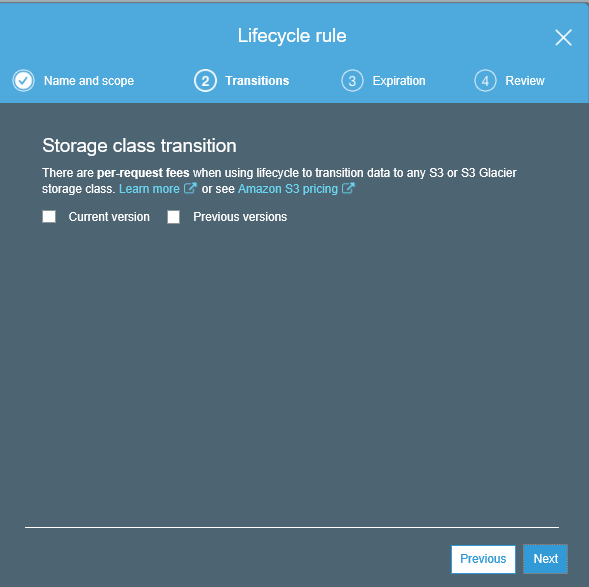


Type the Rule name you want to give

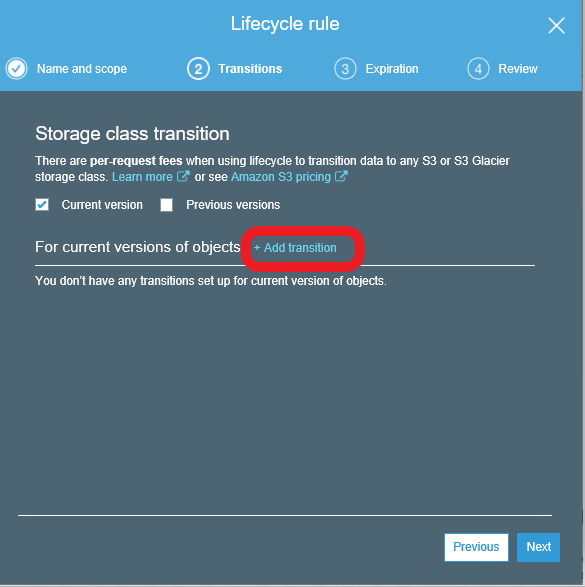


Next would be #2 which says Transitions.

You can choose which version of object need to be applied.



If you select current version only for testing purpose, it would give you add the Transition rule for current version of the objects.



Now if you see the transition list here, it says you can transition the current version of any objects to

Standard-IA

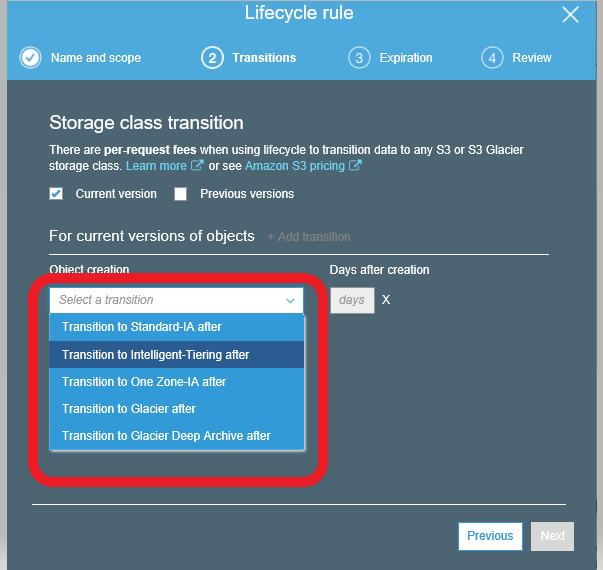
Intelligent-Tiering

One-Zone IA

Glacier

Glacier Deep Archive

After certain # of days of creation



Now the below rule simply says that move all the current versions to Standard-IA after 30 days of creation.

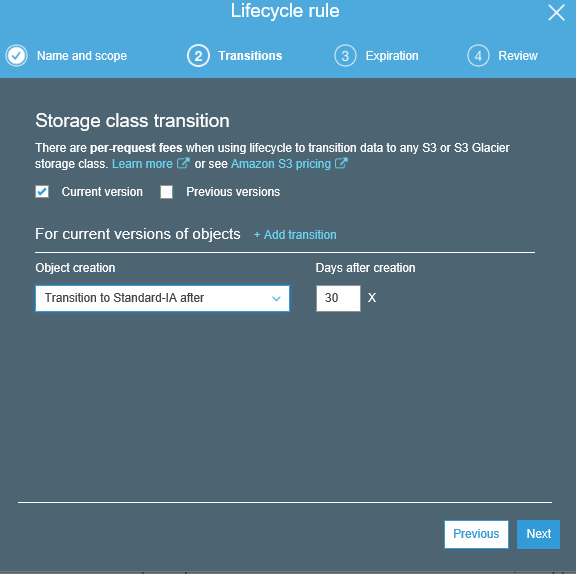
For ex :

FileName : FileA

Created : 15th Oct 2019

Storage : Standard

This file will be moved to IA on 15th Nov 2019



This is only for current version. If you want something to be done on Previous version also, you need to put a checkbox.