

What is AWS S3?

- Amazon Simple Storage Service (S3) is 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.

Explain about AWS S3 Advantages?

1. 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.
2. 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.
3. S3 is **Highly Scalable**,
Since it automatically scales your storage according to your requirement and you only pay for the storage you use.

What kind and how much of data one can store in AWS S3?

- You can store virtually any kind of data, in any format, in S3 and when we talk about capacity, the volume and the number of objects that we can store in S3 are unlimited.
- *An object* is the fundamental entity in S3. It consists of data, key and metadata.

When we talk about data, it can be of two types-

- Data which is to be accessed frequently.
- Data which is accessed not that frequently.

Therefore, Amazon came up with 3 storage classes to provide its customers the best experience and at an affordable cost.

Let's understand the 3 storage classes with a "health-care" use case:

1. Amazon S3 Standard for frequent data access:-

- This is suitable for performance sensitive use cases where the latency should be kept low.
- e.g. in a hospital, frequently accessed data will be the data of admitted patients, which should be retrieved quickly.

2. Amazon S3 Standard for infrequent data access:-

- This is suitable for use cases where the data is long lived and less frequently accessed i.e for data archival but still expects high performance.
- e.g. in the same hospital, people who have been discharged, their records/data will not be needed on a daily basis, but if they return with any complication, their discharge summary should be retrieved quickly.

3. Amazon Glacier:-

- Suitable for use cases where the data is to be archived and high performance is not required, it has a lower cost than the other two services.
- e.g. in the hospital, patients' test reports, prescriptions, MRI, X Ray, Scan docs etc. that are older than a year will not be needed in the daily run and even if it is required, lower latency is not needed.

Specification Snapshot: Storage Classes

Characteristics	Standard	Standard - Infrequent Access	Glacier
Durability	99.99%	99.99%	99.99%
Availability	99.99%	99.90%	N/A
Minimum Object Size	No limit	128KB	No limit
Minimum Storage Duration	No minimum duration	30 Days	90 Days
First Byte Latency	milliseconds	milliseconds	4 hours
Retrieval Fee	No Fee	per GB retrieved	per GB retrieved

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.

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
- Service Availability

Let's understand this through an example:

- Suppose there is a company which has to launch these storage instances to host a website for the customers in the US and India.
- To provide the best experience, the company has to choose a region, which best fits its requirements.

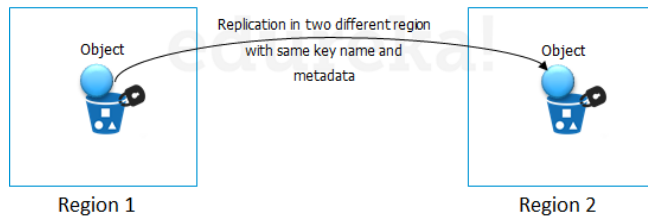
Regions	• Mumbai/N Virginia
Instance Type (Reserved Instance)	• e.g. amazon ec2- m4.4xlarge 16(vCPU), 64 GB RAM
Pricing(1 Year)	• Mumbai - \$691/monthly - \$0.9 hourly • N Virginia - \$480/monthly - \$0.6 hourly
Latency	• From USA to India - Low • From India to USA - High

- Now looking at the above parameters, we can clearly identify, that N Virginia will be the best region for this company because of the low latency and low price. Irrespective of your location, you can select any region which might suit your requirements, since you can access your S3 buckets from anywhere.
- 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.

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.

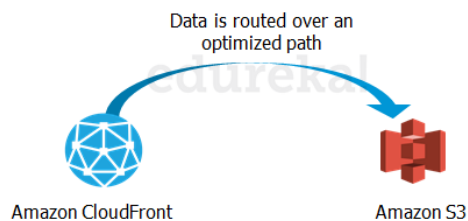
This obviously has a cost to it which has been discussed further in this article.



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



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 optimized network path.

The **Snowball** is a way of transferring your data physically. In this Amazon sends 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?"

Yes! 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:

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

Source: aws.amazon.com for North Virginia region

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

--AWS S3 Use case: 1:--

Industry “Media”

Let’s understand it through a real time use case to assimilate all what we have learnt so far: *IMDb Internet Movie Database* is a famous online database of information related to films, television programs and video games.

Let’s see how they exploit the AWS services:

- To get the lowest possible latency, all possible results for a search are pre-calculated with a document for every combination of letters in search. Each document is pushed to Amazon Simple Storage Service (S3) and thereby to **Amazon CloudFront**, putting the documents physically close to the users. The theoretical number of possible searches to calculate is mind-boggling—a 20-character search has 23 x 1030 combinations
- But in practice, using IMDb’s authority on movie and celebrity data can reduce the search space to about 150,000 documents, which Amazon S3 and **Amazon CloudFront** can distribute in just a few hours.

--AWS S3 Use case: 2:--

Project Statement – Hosting a Static Website on Amazon S3

Let’s first understand: What is a static website?

In short, it’s a website comprised of only HTML, CSS, and/or JavaScript. That means server-side scripts aren’t supported, so if you want to host a Rails or PHP app, you’ll need to look elsewhere.

For simpler purposes, welcome to the wonderful world of hosting websites on AWS S3!

Step 1: Create a bucket

To create a bucket, navigate to S3 in the AWS Management Console and hit Create Bucket. You'll be prompted to enter a name and a region.

If you plan on using your own domain/sub-domain, use that for your bucket name. For the region, pick the one closest to you and hit Create. With any luck, you'll see your new bucket appear in the console.

Create a Bucket - Select a Bucket Name and Region Cancel

A bucket is a container for objects stored in Amazon S3. When creating a bucket, you can choose a Region to optimize for latency, minimize costs, or address regulatory requirements. For more information regarding bucket naming conventions, please visit the [Amazon S3 documentation](#).

Bucket Name:





Region:

Set Up Logging > Create Cancel

Step 2: Verify the Created Bucket

Create Bucket Actions

All Buckets (6)

	Name
	aws-logs-245376966395-us-west-2
	Edureka
	edureka
	edurekaaws

Step3: Enable Website Hosting

The only thing now left to do is, to enable Static Website Hosting. Just select it from the properties panel on the right.

Static Website Hosting

You can host your static website entirely on Amazon S3. Once you enable your bucket for static website hosting, all your content is accessible to web browsers via the Amazon S3 website endpoint for your bucket.

Endpoint: edurekaaws.s3-website-us-west-2.amazonaws.com

Each bucket serves a website namespace (e.g. "www.example.com"). Requests for your host name (e.g. "example.com" or "www.example.com") can be routed to the contents in your bucket. You can also redirect requests to another host name (e.g. redirect "example.com" to "www.example.com"). See our [walkthrough](#) for how to set up an Amazon S3 static website with your host name.

☐ Do not enable website hosting

☒ Enable website hosting

Index Document:

Error Document:

Edit Redirection Rules: You can set custom rules to automatically redirect web page requests for specific content.

☐ Redirect all requests to another host name

Save Cancel

Step 4: Create a Html File

Make sure you set the Index Document to index.html. You can also set an error page if you want. When you're done, hit Save.

One nice thing about the AWS Management Console is that you can upload files to your bucket right from your browser. Let's start by creating one called **index.html**. This will be the contents of the home page:

```

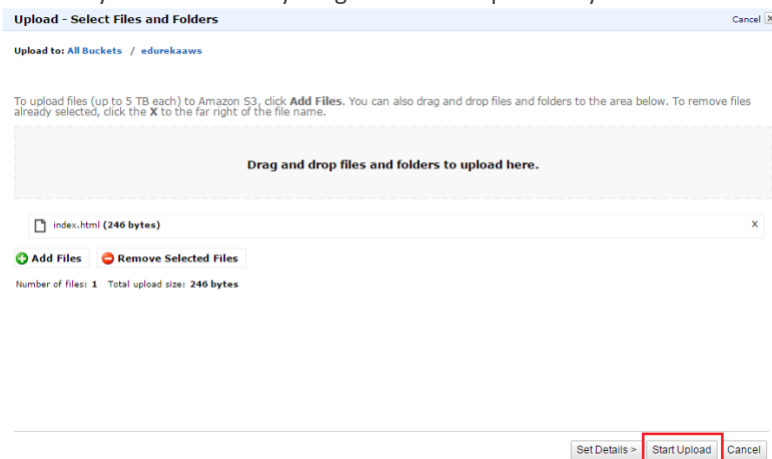
<!doctype html>
<html>
<head>
<title>
Hello, S3!
</title>
<meta name="description" content="My first S3 website">
<meta charset="utf-8">
</head>
<body>
<h2>My first S3 website</h2>
<p>I can't believe it was that easy!</p>
</body>
</html>

```

Step 5: Upload the File in a Bucket

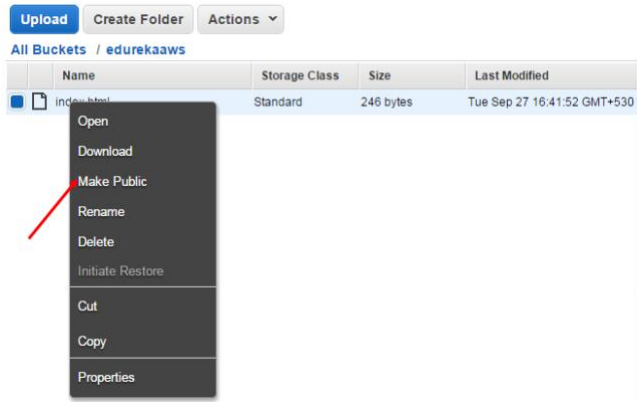
To upload the file, select your new bucket and hit Start Upload button.

Once you've uploaded index.html, it will appear in your bucket. However, you won't be able to see it in your browser yet because everything in AWS S3 is private by default.



Step 6: Make the Html File Public

i) To make index.html file public, right-click on index.html and select Make Public. (Remember to do this for any other files you upload to your website!)



Now that your homepage is visible to the world, it's time to test everything out!

ii) Now, select index.html in the console and go to the Properties tab.

Object: index.html

Bucket:	edurekaaws
Name:	index.html
Link:	https://s3-us-west-2.amazonaws.com/edurekaaws/index.html
Size:	246
Last Modified:	Tue Sep 27 13:19:30 GMT+530 2016
Owner:	support
ETag:	0e92980f3640184e681e1e5b0dda4e60
Expiry Date:	None
Expiration Rule:	N/A

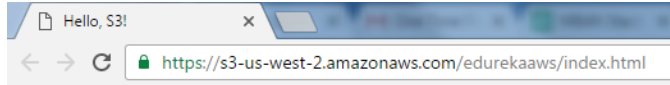
› Details

› Permissions

› Metadata

Step 7: Final Step to Verify the Result

Clicking the link will take you to your new homepage.



My first S3 website

I can't believe it was that easy!

Congratulations! You have just hosted a html website in AWS using S3.

Here's a short AWS S3 tutorial Video that explains: Traditional Storage Tiers, Disadvantages of Traditional Storage over Cloud, AWS storage options: EBS, S3, Glacier, AWS Connecting Storage: Snowball & Storage Gateway, AWS Command Line Interface (CLI), Demo etc.