S3 (Simple Storage Service)

1. To upload your data (photos, vedioes, documents etc.) to Amazon s3, you must create an s3 bucket in one of AWS region. You can then upload any number of objects to the bucket.
2. You can optionally choose other storage management option for the bucket. After you create a bucket, you cannot change the bucket name & region.
3. S3 is storage for the internet. It has a simple web service interface for simple storage & retrieving of any amount of data anytime from anywhere on the internet. If we given the permission.
4. S3 is object based storage & you cannot install operating system on s3.
5. You cannot create nested (bucket inside bucket) buckets.
6. By default you can create up to 100 buckets in each of your AWS account. If you need more buckets you can increase your account bucket limits to maximum of 1000 buckets by summiting a service increase limit.
7. Uploading Objects- when you upload a file to Amazon s3, it is stored as an s3 object. You can upload any type of files- images, backups, movie, data etc. into an s3 bucket. Object size maximum is 5 TB.
8. The maximum size of a file that you can upload by using Amazon s3 console is 160 GB.
9. Depending upon the size of data you are uploading , following option given by Amazon –
10. AWS SDKs , REST API , AWS CLI – you can upload single object up to 5 GB in size single put operation.
11. Amazon s3 console- 160 GB upload single object.
12. AWS SDKs , REST API , AWS CLI – using the multipart upload API, you can upload a single large object greater than 5 gb in size.
13. S3 multipart upload-
14. It is used to upload an object in parts.
15. Parts are uploaded independently and in parallel, in any order.
16. It is recommended for objects size of 100 MB or larger.
17. You must use it for objects larger than 5 GB.
18. This is done through s3 multipart upload API.
19. Buckets naming rules:-
20. Buckets name between 3 to 63 characters long.
21. Buckets name can consist only lowercase letter, numbers, dots (.) & hyphens (-)
22. Buckets name must begin & end with letter or number
23. Buckets name must not be formatted as IP address.
24. S3 bucket versioning
25. Used to protect against accidental object/data deleted or overwrites.
26. Once you enable versioning on a bucket it cannot be disable however it can be suspended.
27. When versioning is enabled and you try to delete an object a delete marker placed on the object.
28. If you reconsider deleting the object you can delete the “delete marker” and object will be available again.
29. You will be charged for all s3 storage cost for all object version stored.
30. You can use versioning with s3 lifecycle rules to delete the older version or you can move them to cheaper storage.
31. Versioning applies to all objects in a bucket and not partially applied.
32. Storage classes of Amazon s3

S3 standard –

1. S3 standard offers high durability, availability and performance object storage for frequently accessed data.
2. Durability is 11 times 9 (99.999999999)
3. Designed for 99.99% availably over a given year
4. The storage cost of the object is fairly high but there is very less charges for retrieving the object.
5. S3 infrequent access –
6. S3-IA is for data that is accessed less frequently but require rapid access when needed.
7. The storage cost is much cheaper than s3 standard almost half the price. But you are charged more heavily for accessing your object.
8. Durability is 11 times 9
9. Availability is 99.9% in a year
10. Data is deleted from s3-IA within 30 days will be charged for full 30 days.

S3- Intelligent Tering –

1. The s3 intelligent tering storage class is designed to optimize cost by automatically moving data to the most cost effective access tier.
2. If an object in the infrequent access tier is accessed then it is automatically moved to the frequent access tier.
3. There are no retrieve fees when using the s3 intelligent tering storage class and no additional charges when objects move between access tier. object monitoring fees will be charge.
4. Durability is 99.999999999
5. Availability is 99.9 over a year

S3 Glacier Instant Retrieval-

1. Use for archiving data that is rarely accessed and requires milliseconds retrieval.
2. Durability is 11 times 9
3. Availability is 99.9 over a year
4. If you have deleted, overwritten, or transitioned to a different storage class an object before the 90-day minimum, you are charged for 90 days.
5. S3 Glacier Instant Retrieval has a minimum billable object size of 128 KB. Smaller objects may be stored but will be charged for 128 KB.
6. Per GB retrieval fees apply.

S3 Glacier Flexible Retrieval-

1. Use for archives where portions of the data might need to be retrieved in minutes.
2. Data can be accessed in as little as 1-5 minutes using expedited retrieval. The retrieval time is flexible, and you can request free bulk retrievals in up to 5-12 hours.
3. Durability is 11 times 9
4. Availability is 99.99 over a year
5. If you have deleted, overwritten, or transitioned to a different storage class an object before the 90-day minimum, you are charged for 90 days.
6. S3 Glacier Flexible Retrieval has a minimum billable object size of 40KB.
7. Per GB retrieval fees apply.

S3 Glacier Deep Archive-

1. Use for archiving data that rarely needs to be accessed and a default retrieval time of 12 hours.
2. Durability is 11 times 9
3. Availability is 99.99 over a year
4. If you have deleted, overwritten, or transitioned to a different storage class an object before the 180-day minimum, you are charged for 180 days.
5. S3 Glacier Deep Archive has a minimum billable object size of 40KB.
6. Per GB retrieval fees apply.

S3 One Zone-IA-

1. Amazon S3 stores the object data in only one Availability Zone, which makes it less expensive than S3 Standard-IA.
2. Data is not resilient to the physical loss of the Availability Zone resulting from disasters, such as earthquakes and floods.
3. Durability is 11 times 9
4. Availability is 99.5 over a year
5. If you delete an object before the end of the 30-day minimum storage duration period, you are charged for 30 days.
6. If an object is less than 128 KB, Amazon S3 charges you for 128 KB.
7. Per GB retrieval fees apply.

S3 Object Lock-

Before you lock any objects, you have to enable a bucket to use S3 Object Lock. You enable Object Lock when you create a bucket.

After you enable Object Lock on a bucket, you can lock objects in that bucket.

When you create a bucket with Object Lock enabled, you can't disable Object Lock or suspend versioning for that bucket.

With S3 Object Lock, you can store objects in Amazon S3 using a write-once-read-many (WORM) model. You can use S3 Object Lock to prevent an object from being deleted or overwritten for a fixed amount of time or indefinitely.

Object Lock provides two ways to manage object retention

In governance mode, users can't overwrite or delete an object version or alter its lock settings unless they have special permissions.

With governance mode, you protect objects against being delete by most users, but you can still grant some users permission to alter the retention settings or delete the object if necessary.

In compliance mode, a protected object version can't be overwritten or deleted by any user, including the root user in your AWS account.

S3 data protection-

Data protection refers to protecting data while in-transit (as it travels to and from Amazon S3) and at rest (while it is stored on disks in Amazon S3 data centers).

Server-Side Encryption – Request Amazon S3 to encrypt your object before saving it on disks in its data centers and then decrypt it when you download the objects.

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

Server-Side Encryption with Amazon S3-Managed Keys (SSE-S3)-

Each object is encrypted with a unique key.

It encrypts the key itself with a root key that it regularly rotates.

It uses one of the strongest block ciphers available, 256-bit Advanced Encryption Standard (AES-256), to encrypt your data.

Server-Side Encryption with KMS keys Stored in AWS Key Management Service (SSE-KMS)

Server-Side Encryption with AWS KMS keys (SSE-KMS) is similar to SSE-S3, but with some additional benefits and charges for using this service. There are separate permissions for the use of a KMS key that provides added protection against unauthorized access of your objects in Amazon S3.

Server-Side Encryption with Customer-Provided Keys (SSE-C)-

You manage the encryption keys and Amazon S3 manages the encryption, as it writes to disks, and decryption, when you access your objects.

Client side encryption-

Client-side encryption is the act of encrypting your data locally to ensure its security as it passes to the Amazon S3 service. The Amazon S3 service receives your encrypted data; it does not play a role in encrypting or decrypting it.

To enable client-side encryption, you have the following options:

Use a key stored in AWS Key Management Service (AWS KMS).

Use a key that you store within your application.

Using a KMS key stored in AWS KMS-

When uploading an object — using the KMS key ID, the client first sends a request to AWS KMS for a new symmetric key that it can use to encrypt their object data. AWS KMS returns two versions of a randomly generated data key:

A plaintext version of the data key that the client uses to encrypt the object data.

A cipher blob of the same data key that the client uploads to Amazon S3 as object metadata.

When downloading an object — the client downloads the encrypted object from Amazon S3 along with the cipher blob version of the data key stored as object metadata. The client then sends the cipher blob to AWS KMS to get the plaintext version of the data key so that it can decrypt the object data.

KMS key-

AWS Key Management Service (AWS KMS) is a managed service that makes it easy for you to create and control the cryptographic keys that are used to protect your data.

AWS KMS uses hardware security modules (HSM) to protect and validate your AWS KMS keys under the FIPS 140-2 Cryptographic Module Validation Program, except in the China (Beijing) and China (Ningxia) Regions.

AWS supports the key types-

Symmetric CMK- represents a 256-bit key that is used for encryption & decryption.

Asymmetric CMK- represents an RSA key pair that is used for encryption & decryption or signing & verification (but not both)

Data Import/Export (snow family)

Physical data transport solution that helps moving TB’s or PB’s of data in or out of AWS.

1. Snow cone- is small rugged, edge compute and data storage product. You can use Snowcone to collect, process, and transfer data to AWS, either offline by shipping the device, or online with AWS DataSync.

Running applications in non-data center edge environments or where there is lack of consistent network connectivity or low bandwidth can be challenging because these locations often lack the space, power, and cooling needed for data center IT equipment.

Snowcone provides 8 TB of available storage. Snowcone performs best when large files are transferred. Snowcone is not optimized for transferring many smaller files with a size of 5 MB or below.

Snowcone performs best when it is used for write workloads (copying data into Snowcone) or read workloads (copying data from Snowcone), and not both. Performance cannot be guaranteed in a mixed read-write scenario.

AWS Snowcone is currently available in the US East (Northern Virginia), US East (Ohio), US West (San Francisco), US West (Oregon), South America (Sao Paulo), Europe (Ireland), Europe (Frankfurt), Asia Pacific (Sydney), Asia Pacific (Tokyo), Asia Pacific (Singapore), and Canada (Central) Regions. AWS Regions with availability planned in additional AWS Regions in the coming months.

AWS Snowcone is designed to be requested and used with a single AWS Region. It may not be requested from one region and returned to another.

All data on AWS Snowcone is always automatically encrypted using 256-bit keys that you manage by using the AWS Key Management Service (KMS). The encryption keys are never stored on the device. This helps ensure that your data stays secure during device transit.

Snowball/ Edge-AWS Snowball is a service that provides secure, rugged devices, so you can bring AWS computing and storage capabilities to your edge environments, and transfer data into and out of AWS.

Those rugged devices are commonly referred to as AWS Snowball or AWS Snowball Edge devices.

Previously, AWS Snowball referred specifically to an early hardware version of these devices; however that model has been replaced by updated hardware. Now the AWS Snowball service operates with Snowball Edge devices, which include on-board computing capabilities as well as storage.

Snow Edge working-You start by requesting one or more Snowball Edge Compute Optimized or Snowball Edge Storage Optimized devices in the AWS Management Console based on how much data you need to transfer and the compute needed for local processing.

The buckets, data, Amazon EC2 AMIs, and Lambda functions you select are automatically configured, encrypted, and preinstalled on your devices before they are shipped to you.

Once a device arrives, you connect it to your local network and set the IP address either manually or automatically with DHCP. Then use the Snowball Edge client software, job manifest, and unlock code to verify the integrity of the Snowball Edge device or cluster, and unlock it for use.

The manifest and unlock code are uniquely generated and crypto-logically bound to your account and the Snowball Edge shipped to you, and cannot be used with any other devices. Data copied to Snowball Edge is automatically encrypted and stored in the buckets you specify.

Snow Edge device available in size of 50TB or 80 TB.

Snowmobile- AWS Snowmobile is the first Exabyte-scale data migration service that allows you to move very large datasets from on-premises to AWS.

Each Snowmobile is a secured data truck with up to 100PB storage capacity that can be dispatched to your site and connected directly to your network backbone to perform high-speed data migration.

You can quickly migrate an Exabyte of data with ten Snowmobiles in parallel from a single location or multiple data centers. Snowmobile is offered by AWS as a managed service.

Each Snowmobile has a total capacity of up to 100 petabytes and multiple Snowmobiles can be used in parallel to transfer Exabyte of data.

The Snowmobile is designed to transfer data at a rate up to 1 Tb/s, which means you could fill a 100PB Snowmobile in less than 10 days. The actual transfer speed may vary depending on the available local network capacity at your site and the speed of your on-premises storage devices.

Snowmobile can be made available for use with AWS services in specific AWS regions.

To migrate large datasets of 10PB or more in a single location, you should use Snowmobile. For datasets less than 10PB or distributed in multiple locations, you should use Snowball. In addition, you should evaluate the amount of available bandwidth in your network backbone.

S3 transfer acceleration- is a bucket-level feature that enables fast, easy, and secure transfers of files over long distances between your client and an S3 bucket. Transfer Acceleration takes advantage of the globally distributed edge locations in Amazon Cloud Front. As the data arrives at an edge location, the data is routed to Amazon S3 over an optimized network path.

You can use the Amazon S3 Transfer Acceleration Speed Comparison tool to compare accelerated and non-accelerated upload speeds across Amazon S3 Regions. The Speed Comparison tool uses multipart uploads to transfer a file from your browser to various Amazon S3 Regions with and without using Transfer Acceleration.

Storage Gateway- is a hybrid cloud storage service that gives you on-premises access to virtually unlimited cloud storage. Storage Gateway provides a standard set of storage protocols such as iSCSI, SMB, and NFS, which allow you to use AWS storage without rewriting your existing applications. It provides low-latency performance by caching frequently accessed data on premises, while storing data securely and durably in Amazon cloud storage services.

Amazon S3 File Gateway- presents a file-based interface to Amazon S3, which appears as a network file share. It enables you to store and retrieve Amazon S3 objects through standard file storage protocols. File Gateway allows your existing file-based applications or devices to use secure and durable cloud storage without needing to be modified. With S3 File Gateway, your configured S3 buckets will be available as Network File System (NFS) mount points or Server Message Block (SMB) file shares. Your applications read and write files and directories over NFS or SMB, interfacing to the gateway as a file server.

Tape Gateway- is a cloud-based Virtual Tape Library (VTL). It presents your backup application with a VTL interface, consisting of a media changer and tape drives. You can create virtual tapes in your virtual tape library using the AWS Management Console. Your backup application can read data from or write data to virtual tapes by mounting them to virtual tape drives using the virtual media changer. Virtual tapes are discovered by your backup application using its standard media inventory procedure. Virtual tapes are available for immediate access and are backed by Amazon S3. You can also archive tapes. Archived tapes are stored in Amazon S3 Glacier or Amazon S3 Glacier Deep Archive.

Volume Gateway- provides an iSCSI target, which enables you to create block storage volumes and mount them as iSCSI devices from your on-premises or EC2 application servers. The Volume Gateway runs in either a cached or stored mode.

In the cached mode, your primary data is written to S3, while retaining your frequently accessed data locally in a cache for low-latency access.

In the stored mode, your primary data is stored locally and your entire dataset is available for low-latency access while asynchronously backed up to AWS.

In either mode, you can take point-in-time snapshots of your volumes, which are stored as Amazon EBS Snapshots in AWS, enabling you to make space-efficient versioned copies of your volumes for data protection, recovery, migration and various other copy data needs.