Simple Storage Service (S3)
S3 Simplified:
S3 provides developers and IT teams with secure, durable, and highly-scalable object storage.
Object storage, as opposed to block storage, is a general term that refers to data composed of three things:

1.) the data that you want to store

2.) an expandable amount of metadata

3.) a unique identifier so that the data can be retrieved

This makes it a perfect to host files or directories and a poor to host databases or operating systems. The following table highlights key differences between object and block storage:

	OBJECT STORAGE	BLOCK STORAGE
PERFORMANCE	Performs best for big content and high stream throughput	Strong performance with database and transactional data
GEOGRAPHY	Data can be stored across multiple regions	The greater the distance between storage and application, the higher the latency
SCALABILITY	Can scale infinitely to petabytes and beyond	Addressing requirements limit scalability
ANALYTICS	Customizable metadata allows data to be easily organized and retrieved	No metadata

The files uploaded into S3 have an upper-bound of 5TB per file and the number of files that can be uploaded is virtually limitless.
Each Amazon S3 object contains both data and metadata. Objects reside in containers called buckets, and each object is identified by a unique user-specified key (filename).

S3 Key Details:
Objects (regular files or directories) are stored in S3 with a key, value, version ID, and metadata.
Key: The name you assign to the object
Every object stored in an S3 bucket is identified by a unique identifier called a key. You can think of the key as a filename.
Version: It is the version Id of a specific version of a file. We will explore this in detail some of the following topics. The version helps uniquely identify a particular object.
Value and Metadata: Value is nothing but a concept we are trying to store. Whereas Metadata is the information about the data we are trying to store

The data consistency model for S3 ensures immediate read access for new objects after the initial PUT requests.
These new objects are introduced into AWS for the first time and thus do not need to be updated anywhere so they are available immediately.

The data consistency model for S3 also ensures immediate read access for PUTS and DELETES of already existing objects, since Decembre 2020.

Amazon guarantees 99.999999999% (or 11 9s) durability for all S3 storage classes except its Reduced Redundancy Storage class.

S3 comes with the following main features:

1.) tiered storage and pricing variability

2.) lifecycle management to expire older content

3.) versioning for version control

4.) encryption for privacy

5.) MFA deletes to prevent accidental or malicious removal of content

6.) access control lists & bucket policies to secure the data

S3 charges by:

1.) storage size

2.) number of requests

3.) storage management pricing (known as tiers)

4.) data transfer pricing (objects leaving/entering AWS via the internet)

5.) transfer acceleration (an optional speed increase for moving objects via Cloudfront)

6.) cross region replication (more HA than offered by default

Bucket policies secure data at the bucket level while access control lists secure data at the more granular object level.

By default, all newly created buckets are private.

S3 can be configured to create access logs which can be shipped into another bucket in the current account or even a separate account all together. This makes it easy to monitor who accesses what inside S3.

There are 3 different ways to share S3 buckets across AWS accounts:
1.) For programmatic access only, use IAM & Bucket Policies to share entire buckets
2.) For programmatic access only, use ACLs & Bucket Policies to share objects
3.) For access via the console & the terminal, use cross-account IAM roles

S3 is a great candidate for static website hosting. When you enable static website hosting for S3 you need both an index.html file and an error.html file. Static website hosting creates a website endpoint that can be accessed via the internet.

When you upload new files and have versioning enabled, they will not inherit the properties of the previous version.

S3 Storage Classes:
S3 Standard - 99.99% availability and 11 9s durability. Data in this class is stored redundantly across multiple devices in multiple facilities and is designed to withstand the failure of 2 concurrent data centers.

S3 Infrequently Accessed (IA) - For data that is needed less often, but when it is needed the data should be available quickly. The storage fee is cheaper, but you are charged for retrieval.

S3 One Zone Infrequently Accessed (an improvement of the legacy RRS / Reduced Redundancy Storage) - For when you want the lower costs of IA, but do not require high availability. This is even cheaper because of the lack of HA.

S3 Intelligent Tiering - Uses built-in ML/AI to determine the most cost-effective storage class and then automatically moves your data to the appropriate tier. It does this without operational overhead or performance impact.

S3 Glacier - low-cost storage class for data archiving. This class is for pure storage purposes where retrieval isn't needed often at all. Retrieval times range from minutes to hours. There are differing retrieval methods depending on how acceptable the default retrieval times are for you:

Expedited: 1 - 5 minutes, but this option is the most expensive.
Standard: 3 - 5 hours to restore.
Bulk: 5 - 12 hours. This option has the lowest cost and is good for a large set of data.
The Expedited duration listed above could possibly be longer during rare situations of unusually high demand across all of AWS. If it is absolutely critical to have quick access to your Glacier data under all circumstances, you must purchase Provisioned Capacity. Provisioned Capacity guarentees that Expedited retrievals always work within the time constraints of 1 to 5 minutes.

S3 Deep Glacier - The lowest cost S3 storage where retrieval can take 12 hours.

Storage Class	Designed for	Durability (designed for)	Availability (designed for)	Availability Zones	Min storage duration	Min billable object size	Other Considerations
STANDARD	Frequently accessed data	99.999999999%	99.99%	>= 3	None	None	None
STANDARD_IA	Long-lived, infrequently accessed data	99.999999999%	99.9%	>= 3	30 days	128 KB	Per GB retrieval fees apply.
INTELLIGENT_TIERING	Long-lived data with changing or unknown access patterns	99.999999999%	99.9%	>= 3	30 days	None	Monitoring and automation fees per object apply. No retrieval fees.
ONEZONE_IA	Long-lived, infrequently accessed, non-critical data	99.999999999%	99.5%	1	30 days	128 KB	Per GB retrieval fees apply. Not resilient to the loss of the Availability Zone.
GLACIER	Long-term data archiving with retrieval times ranging from minutes to hours	99.999999999%	99.99% (after you restore objects)	>= 3	90 days	40 KB	Per GB retrieval fees apply. You must first restore archived objects before you can access them. For more information, see Restoring Archived Objects.
DEEP_ARCHIVE	Archiving rarely accessed data with a default retrieval time of 12 hours	99.999999999%	99.99% (after you restore objects)	>= 3	180 days	40 KB	Per GB retrieval fees apply. You must first restore archived objects before you can access them. For more information, see Restoring Archived Objects.
RRS (Not recommended)	Frequently accessed, non-critical data	99.99%	99.99%	>= 3	None	None	None