**Amazon Simple Storage Service (Amazon S3)**

**Video transcript**

–

Welcome to this video on Amazon Simple Storage Service, or S3. From the name, you've probably guessed that it is a storage service and it's, well, simple. Most businesses have data that needs to be stored somewhere. For the coffee shop, this could be receipts, images, Excel spreadsheets, employee training videos, and even text files, among others. Storing these files is where S3 comes in handy because it is a data store that allows you to store and retrieve an unlimited amount of data at any scale.

Data is stored as objects, but instead of storing them in a file directory, you store them in what we call buckets. Think of a file sitting on your hard drive, that is an object and think of a file directory, that is the bucket. The maximum object size that you can upload is five terabytes. You can also version objects to protect them from accidental deletion of an object. What this means is that you always retain the previous versions of an object, as like a paper trail. You can even create multiple buckets and store them across different classes or tiers of data. You can then create permissions to limit who can see or even access objects.

And you can even stage data between different tiers. These tiers offer mechanisms for different storage use cases such as data that needs to be accessed frequently, versus audit data that needs to be retained for several years. Let's go through the notable ones, shall we?

The first here is called S3 Standard and comes with 11 nines of durability. That means an object stored in S3 Standard has a 99.999999999 percentage – that's a lot of nines –  probability that it will remain intact after a period of one year. That's pretty high. Furthermore, data is stored in such a way that AWS can sustain the concurrent loss of data in two separate storage facilities. This is because data is stored in at least three facilities. So multiple copies reside across locations. Another useful way to use S3 is static website hosting, where a static website is a collection of HTML files and each file is akin to a physical page of the actual site. You can do this by simply uploading all your HTML, static web assets, and so forth into a bucket and then checking a box to host it as a static website. You can then enter the bucket's URL and ba-bam! Instant website. And we say static, but that doesn't mean you can't have animations and moving parts to your website. Pretty awesome way to start up that coffee blog.

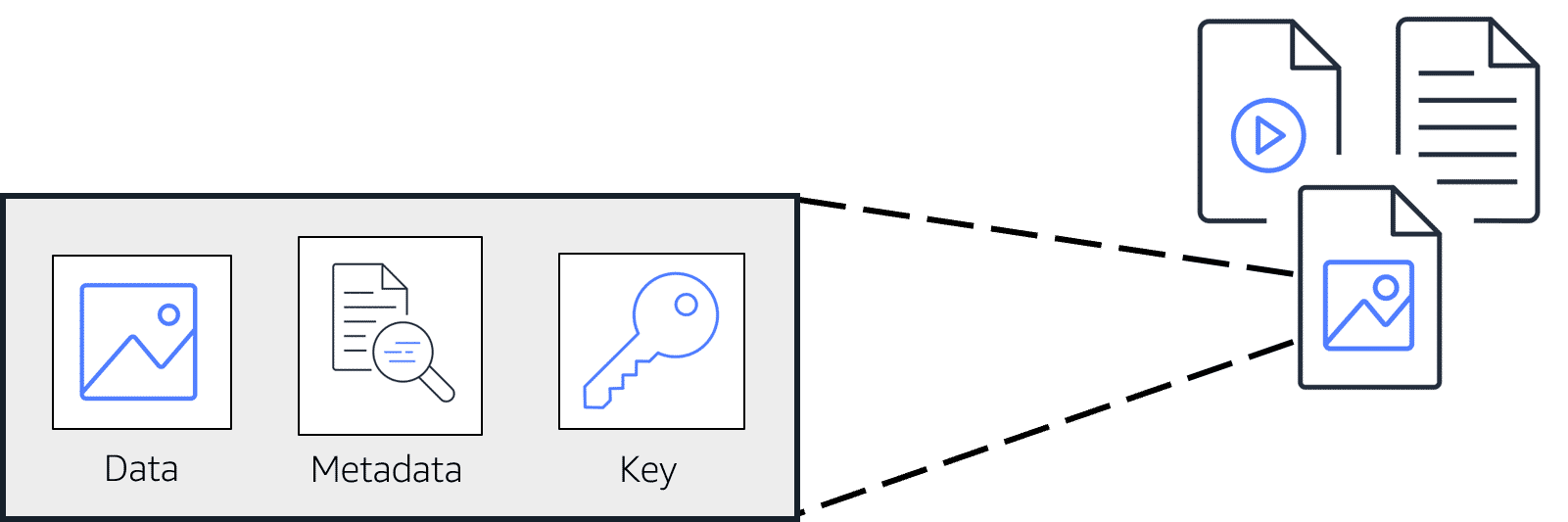
The next storage class is called S3 Infrequent Access or S3-IA. Which is used for data that is accessed less frequently but requires rapid access when needed. This means it's a perfect place to store backups, disaster recovery files, or any object that requires a long-term storage.

Another storage class or tier lends itself to that example we had earlier about audit data. Say, we need to retain data for several years for auditing purposes. And we don't need it to be retrieved very rapidly. Well, then you can use Amazon S3 Glacier to archive that data. To use Glacier, you can simply move data to it, or you can create vaults and then populate them with archives. And if you have compliance requirements around retaining data for, say, a certain period of time, you can employ an S3 Glacier vault lock policy and lock your vault. You can specify controls such as write once/ read many, or WORM, in a vault lock policy and lock the policy from future edits. Once locked, the policy can no longer be changed. You also have three options for retrieval, which range from minutes to hours, and you have the option of uploading directly to Glacier or using S3 Lifecycle policies.

Fact, I don't think we mentioned lifecycle policies up till this point. But they are policies you can create that can move data automatically between tiers. For example, say we need to keep an object in S3 Standard for 90 days, and then we want to move it to S3-IA for the next 30 days. Then after 120 days total, we want it to be moved to S3 Glacier. With Lifecycle policies, you create that configuration without changing your application code and it will perform those moves for you automatically. It's another example of a managed AWS service, helping take that burden off you, so you can focus on more of your business needs.

Just to note that there are other storage classes. Like S3 Infrequent Access One Zone and S3 Glacier Deep Archive that you can use. Happy storing.

**Object storage**

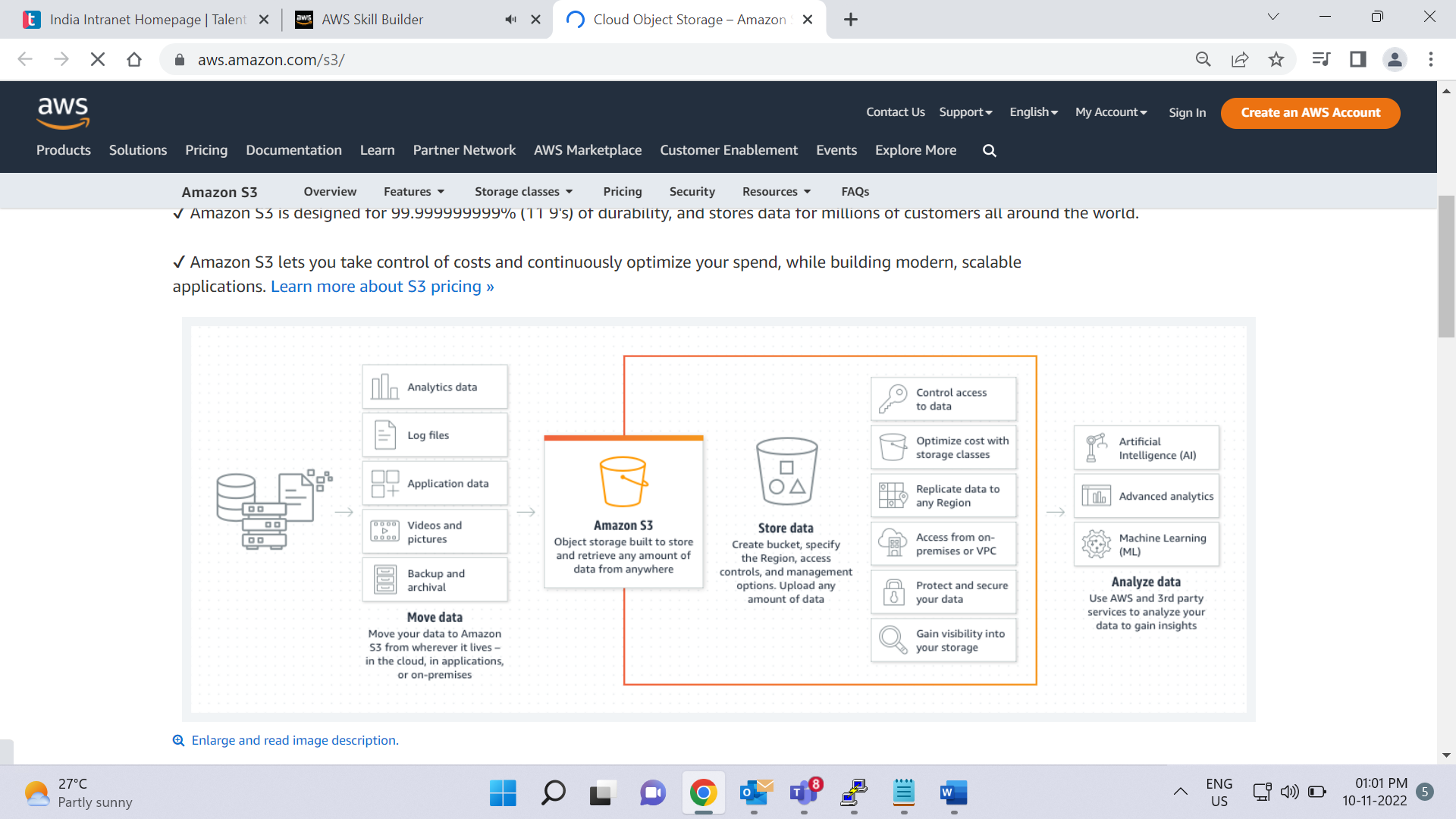


In **object storage**, each object consists of data, metadata, and a key.

The data might be an image, video, text document, or any other type of file. Metadata contains information about what the data is, how it is used, the object size, and so on. An object’s key is its unique identifier.

Recall that when you modify a file in block storage, only the pieces that are changed are updated. When a file in object storage is modified, the entire object is updated.

**Amazon Simple Storage Service (Amazon S3)**



[**Amazon Simple Storage Service (Amazon S3)**](https://aws.amazon.com/s3/) is a service that provides object-level storage. Amazon S3 stores data as objects in buckets.

You can upload any type of file to Amazon S3, such as images, videos, text files, and so on. For example, you might use Amazon S3 to store backup files, media files for a website, or archived documents. Amazon S3 offers unlimited storage space. The maximum file size for an object in Amazon S3 is 5 TB.

When you upload a file to Amazon S3, you can set permissions to control visibility and access to it. You can also use the Amazon S3 versioning feature to track changes to your objects over time.

**Amazon S3 storage classes**

With Amazon S3, you pay only for what you use. You can choose from [a range of storage classes](https://aws.amazon.com/s3/storage-classes) to select a fit for your business and cost needs. When selecting an Amazon S3 storage class, consider these two factors:

* How often you plan to retrieve your data
* How available you need your data to be

To learn more about the Amazon S3 storage classes select the **+** symbol next to each category.

**Amazon S3 Standard**

–

* Designed for frequently accessed data
* Stores data in a minimum of three Availability Zones

Amazon S3 Standard provides high availability for objects. This makes it a good choice for a wide range of use cases, such as websites, content distribution, and data analytics. Amazon S3 Standard has a higher cost than other storage classes intended for infrequently accessed data and archival storage.

**Amazon S3 Standard-Infrequent Access (S3 Standard-IA)**

–

* Ideal for infrequently accessed data
* Similar to Amazon S3 Standard but has a lower storage price and higher retrieval price

Amazon S3 Standard-IA is ideal for data infrequently accessed but requires high availability when needed. Both Amazon S3 Standard and Amazon S3 Standard-IA store data in a minimum of three Availability Zones. Amazon S3 Standard-IA provides the same level of availability as Amazon S3 Standard but with a lower storage price and a higher retrieval price.

**Amazon S3 One Zone-Infrequent Access (S3 One Zone-IA)**

–

* Stores data in a single Availability Zone
* Has a lower storage price than Amazon S3 Standard-IA

Compared to Amazon S3 Standard and Amazon S3 Standard-IA, which store data in a minimum of three Availability Zones, Amazon S3 One Zone-IA stores data in a single Availability Zone. This makes it a good storage class to consider if the following conditions apply:

* You want to save costs on storage.
* You can easily reproduce your data in the event of an Availability Zone failure.

**Amazon S3 Intelligent-Tiering**

–

* Ideal for data with unknown or changing access patterns
* Requires a small monthly monitoring and automation fee per object

In the Amazon S3 Intelligent-Tiering storage class, Amazon S3 monitors objects’ access patterns. If you haven’t accessed an object for 30 consecutive days, Amazon S3 automatically moves it to the infrequent access tier, Amazon S3 Standard-IA. If you access an object in the infrequent access tier, Amazon S3 automatically moves it to the frequent access tier, Amazon S3 Standard.

**Amazon S3 Glacier Instant Retrieval**

–

* Works well for archived data that requires immediate access
* Can retrieve objects within a few milliseconds

When you decide between the options for archival storage, consider how quickly you must retrieve the archived objects. You can retrieve objects stored in the Amazon S3 Glacier Instant Retrieval storage class within milliseconds, with the same performance as Amazon S3 Standard.

**Amazon S3 Glacier Flexible Retrieval**

–

* Low-cost storage designed for data archiving
* Able to retrieve objects within a few minutes to hours

Amazon S3 Glacier Flexible Retrieval is a low-cost storage class that is ideal for data archiving. For example, you might use this storage class to store archived customer records or older photos and video files.

**Amazon S3 Glacier Deep Archive**

–

* Lowest-cost object storage class ideal for archiving
* Able to retrieve objects within 12 hours

Amazon S3 Deep Archive supports long-term retention and digital preservation for data that might be accessed once or twice in a year. This storage class is the lowest-cost storage in the AWS Cloud, with data retrieval from 12 to 48 hours. All objects from this storage class are replicated and stored across at least three geographically dispersed Availability Zones.

**Amazon S3 Outposts**

–

* Creates S3 buckets on Amazon S3 Outposts
* Makes it easier to retrieve, store, and access data on AWS Outposts

Amazon S3 Outposts delivers object storage to your on-premises AWS Outposts environment. Amazon S3 Outposts is designed to store data durably and redundantly across multiple devices and servers on your Outposts. It works well for workloads with local data residency requirements that must satisfy demanding performance needs by keeping data close to on-premises applications.

**Comparing Amazon EBS and Amazon S3**

Play Video

**Video transcript**

–

AWS Cloud Practitioners, welcome to the clash of the storage class! In the block storage corner, weighing in at sizes up to 16 tebibytes each, with a unique ability to survive the termination of their Amazon EC2 instances, they are solid state, they are spinning platters, they are Amazon Elastic Block Storage!

In the regional object storage corner, weighing in at unlimited storage, with individual objects at 5,000 gigabytes in size, they specialize in write once/read many, they are 99 .999 999 999% durable, they are Amazon Simple Storage Service!

Head-to-head, each storage class boasts the best dynamically distributed design for different storage demands. Which storage class will ultimately be victorious in this thunderdome slugfest? To understand who wins, you need to clarify a use case.

Round one. Let's say you're running a photo analysis website where users upload a photo of themselves, and your application finds the animals that look just like them. You have potentially millions of animal pictures that all need to be indexed and possibly viewed by thousands of people at once. This is the perfect use case for S3. S3 is already web enabled. Every object already has a URL that you can control access rights to who can see or manage the image. It's regionally distributed, which means that it has 11 nines of durability, so no need to worry about backup strategies. S3 is your backup strategy. Plus the cost savings is substantial overrunning the same storage load on EBS. With the additional advantage of being serverless, no Amazon EC2 instances are needed. Sounds like S3 is the judge's winner here for this round.

But wait, round two, you have an 80-gigabyte video file that you're making edit corrections on. To know the best storage class here, we need to understand the difference between object storage and block storage. Object storage treats any file as a complete, discreet object. Now this is great for documents, and images, and video files that get uploaded and consumed as entire objects, but every time there's a change to the object, you must re-upload the entire file. There are no delta updates. Block storage breaks those files down to small component parts or blocks. This means, for that 80-gigabyte file, when you make an edit to one scene in the film and save that change, the engine only updates the blocks where those bits live. If you're making a bunch of micro edits, using EBS, elastic block storage, is the perfect use case. If you were using S3, every time you saved the changes, the system would have to upload all 80 gigabytes, the whole thing, every time. EBS clearly wins round two.

This means, if you are using complete objects or only occasional changes, S3 is victorious. If you are doing complex read, write, change functions, then, absolutely, EBS is your knockout winner. Your winner depends on your individual workload. Each service is the right service for specific needs. Once you understand what you need, you will know which service is your champion!

