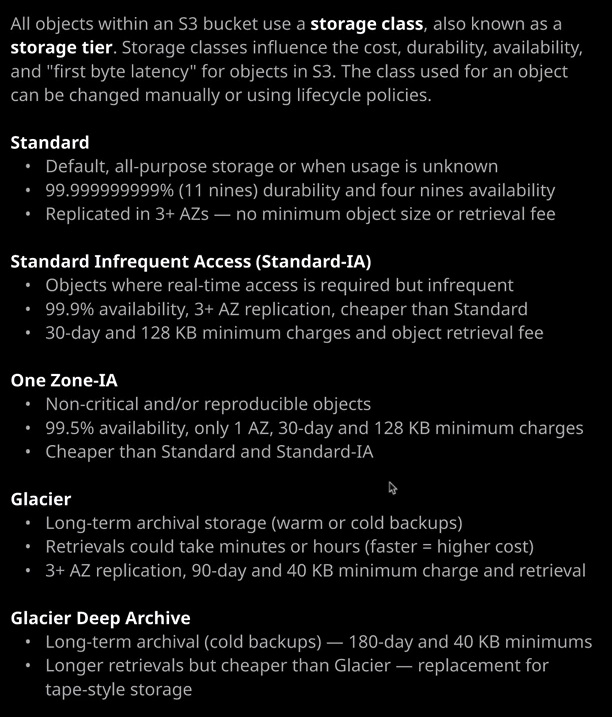
**S3 performance and high availability**



I have demonstrated how you can upload or put objects into S3 but I haven't demonstrated how there are actually a number of different performance classes or storage classes available within the product. Now from a Solutions Architect Associate perspective, what you need to be aware of for the exam is when and where you'd use these individual storage classes as well as some of the technical features and limitations of each. So let's jump in and get started.

Now, from a practical perspective, you can select the storage class to use was when you initially upload the object to S3 as well as change it with some limitations once it's in S3. So let's have a look at how that works. So I'm going to move to the S3 console. I've already got the bucket that I've been using throughout this section of the course called ac-catpics1337 so I'm going to go into that bucket I'm going to select to upload an object and then add files. Now inside the folder for this lesson. So it's called storage classes. It's an s3\_perf\_and\_ha in the 05 storage content delivery folder of the course GitHub repository. I've got four individual objects. Each is a picture of my cat, but for this lesson, it doesn't matter what the data is. I just want to use them to demonstrate uploading using the individual storage classes. So I'm going to select winky.jpeg, click on Open. I'm going to click next, next again, and it's on this screen when I'm able to select which storage class I want to use for this object. Now the default storage class is standard. Standard is used as an all purpose storage class. So whenever you don't have any specific requirements or the usage of that object is unknown then you pick standard and if you don't explicitly select one, then standard is the one that gets used. When you upload an object to S3 standard, it's replicated across three or more availability zones, and that's how it achieves this 11 Nines durability. So if any individual availability zone is damaged, then you've still got the remaining ones and your data is secure. **So S3 standard that's the storage class that you'll use whenever you don't have any specific requirements or you're unsure. Everything else is based off this standard storage class, so it's a tradeoff you're essentially getting reduced object storage fees in exchange for slightly reduced or different features and it's those features and feature differences that I want to talk about in this lesson.** So if I go ahead and click next and accept the rest of the defaults then this object will be uploaded to this bucket using the standard storage class. **Now, if I use the standard storage class, there's no minimum capacity charge per object. I can store objects inside S3 from the minimum to the maximum size of an object. There's no minimum charge to store using the standard storage class. There's also no minimum storage duration charge, and there's no retrieval fee. Now S3 uses a term known as first byte latency and that term describes the amount of time that passes from making the request to retrieve or get an object to the point of where you're delivered the first byte and using S3 standard that's measured in milliseconds. Essentially, there's no delay between when you request the object and when you receive it.** So if you're in the exam and you face any questions around which storage class to use then your default, your starting point should be standard.

The next storage class that I want to talk about is **standard infrequent access also known a standard IA.** Now, this is an interesting one because this essentiall**y reduces the costs of storing objects inside S3.** Now, you'd use this class where you **do need realtime access to data.** So if you want to get an object using this storage class, you need it to be delivered to you with the same low first byte latency as S3 standard. So essentially**, if you pick this storage class, you still get immediate access to your data and we've still got the same levels of availability and the same levels of durability because the data is replicated across all availability zones. The same availability zones that standard uses so at least three. Now with standard infrequent access, you do make a trade off. It's cheaper to store objects inside this storage class, but in order to get that, you agree to a less frequent access pattern. So there's a minimum charge of 30 days for storing objects using this storage class. So if you put an object into this storage class, you immediately get billed for 30 days of usage of that storage class. Additionally, whatever the size of the object you're billed for a 128 kb minimum charge. So if you put a very small object into standard IA you're always going to build that 128 kb minimum charge and if you ever want to retrieve that object, you pay an object retrieval fee. So you'd use this storage class when you've got data that you want to store cheaper than standard, where you know that at times you will need real time access to it but you know that it's not going to be all that frequent. So essentially you're trading a reduced cost for an agreement that you won't be accessing that data on a regular basis,** and I could select this particular object, go to actions, change the storage class, and change it for standard to standard infrequent access and remember, you've got a minimum storage duration and a minimum objects size. Now, that doesn't mean that I can't immediately delete that object. It simply means that I'm going to be billed that minimum period as though the object still existed. So even if I do delete it, I'm still going to be charged for the remainder of those 30 days.

Now, the next storage class that I want to talk about is one **zone IA or one zone infrequent access. Now this storage class is also for objects that are going to be accessed relatively infrequently so it shares a lot of the cost savings of standard IA and a lot of the trade offs of standard IA but one zone IA has another really important trade off, and that's essentially a reduction in the number of availability zones that are used for replication. So with standard and standard IA. You get three or more availability zones that that data is replicated to. With one zone IA, as the name suggests, you only get one availability zone. Now, it's cheaper than standard or standard IA but it does mean that there's a lot more risk. The likelihood of losing data in the event you've got an AZ failure or storage subsystem failure is massively increased. So you've still got the same riel time access to the data. You've got the same trade offs in terms of infrequent access but for the exam, and this is critically important, you cannot use one zone IA for important data. For any data that you can't readily reproduce, you shouldn't use one zone IA. Now from an exam perspective, one zone IA is fairly commonly used as a target for cross region replications. So if you've already got a copy of the data and you replicating it to another region, then you don't need the same level of durability in the remote region.** **So that's one use case. Another use case for one zone infrequent access is for output of data processes. So let's say that you're uploading an object into S3 using the standard storage class and then maybe running a Lambda function, which performs some manipulation, maybe adds a watermark, maybe it recreates a smaller image, say a thumbnail, where you can store the output in one zone IA. As long as you can reproduce the data, then it's safe to store in one zone IA or if it's data that isn't mission critical that you don't care about you can also achieve massive cost savings by using one zone IA.** So that's something important to keep in mind.

The next two storage classes that I want to talk about are **Glacier and Glacier deep archive. Now both of these are designed for archival storage. What this means is you pay a significantly reduced object storage fee, but you don't have that millisecond latency to first byte object retrieval. With any of these two storage classes essentially, you make a request to retrieve that data, and it isn't instant. Retrievals can take minutes or hours. You can request faster retrievals, but that comes with a higher cost. So for the exam, Glacier is designed for long term archival storage, and you can compare this to warm or cold backups. Retrievals could take minutes or hours and with Glacier, you've got the same three or more availability zone replication, but you're charged a minimum of 90 days and 40 kbs minimum charge and of course, there is a retrieval fee to get access to your data. So I want you need to think about Glacier as either archival storage on a file system or disk back ups in a traditional backup system. So it's something that you can still get access to your data relatively quickly, but you're paying reduce costs for doing so.**

**Glacier Deep Archive is specifically designed for long term archival so cold backups. You get even more reduction in object storage fees, but you've got this 180 day minimum charge and 40 kb minimum object size billing. Now retrievals using Glacier Deep Archive do take significantly longer than Glacier. With Glacier, they could take minutes or hours, depending on how what you wanted to pay, but with Glacier Deep Archive they always take hours. Essentially, for the exam I need you to think about Glacier Deep Archive in the same way that you'd think about tape storage. It's cheap, it's resilient, but it does cost significantly more and take significantly longer to retrieve data back into real time storage systems. So think about these for the exam as tiers, you start off as standard. You can go to infrequent access when you want that same real time retrieval but you won't be doing it that frequently. You could move to one zone IA if you don't need the same levels of data resilience and then if you want to trade even more reductions for longer access times, then you can move to Glacier or Glacier Deep Archive dependent on the type of data needs that you have.**

So these are the five standard storage classes available within S3. There is one more known as S3 intelligent tiering and I'm going to be talking about that in the next lesson when I'm also covering lifecycle policies. So you might face questions in the exam about these storage classes but in the Solutions Architect Associate exam, you won't be facing questions on cost. You'll generally be facing questions about appropriately using these different storage classes. So when you would and when you wouldn't use each of the individual classes. Now I've covered all that you need inside this lesson, and I'll make sure that I do include a link in the lesson description, which gives you a nice overview of all of the different storage classes together with their availability their durability, the AZ replication, any minimums in terms of object size or object duration and lots of other useful information. So I definitely check out the link in the lesson description and just use it as reference as you're revising and studying for the exam. With that being said, that's everything I wanted to cover in this lesson so go ahead mark it as complete, and when you're ready, you can move on to the next lesson where I'll be talking about intelligent tiering and lifecycle policies.

<https://aws.amazon.com/s3/storage-classes/>