

# When Amazon S3 is not enough: a look into storage services on AWS

28 April 2023 - 8 min. read

[Amazon FSx for NetApp ONTAP](#)

[Amazon FSx for Windows](#)

[Amazon S3](#)

[Amazon S3 File Gateway](#)

42 is "the answer to life, the universe, and everything else", as Douglas Adams said in *The Hitchhiker's Guide to the Universe*. It goes the same way for storage needs on AWS: Amazon S3 is the answer for everything.

As we'll see, this isn't an absolute truth: sometimes we need to use something different: variety is the spice of life.

When dealing with lift and shift projects, hybrid environments, and Windows workloads, you'll always find file shares accessed by users and services. Many off-the-shelf software solutions rely on that technology; sometimes, refactoring is not an option.

As the name suggests, a file share is not equivalent to object storage: it has different properties, behaviors, and usage scenarios. In this article, we will look at the other options to migrate and adapt workloads on the Cloud.

First, let's clarify the key differences between file and object storage.

File storage is our "traditional" idea: depending on the operating system (Windows, Linux, macOS), you'll have a system that stores data in a hierarchical structure, identified by name and path. Metadata (like permissions and file properties) is stored separately, and its design depends on the filesystem in use.

Object storage is different: all data (including metadata and properties) is stored in a flat namespace, accessed using APIs by referencing an identifier. Once data is stored, you need

to write a new version to modify the object, as there's no way to append or delete portions of the data as you do in a traditional filesystem.

This different approach makes object storage more scalable and cost-effective, but you can't switch between the two technologies seamlessly.

What are the options once you have a Windows workload that needs shared file storage? As we always say, we'll see that there's no default answer.

AWS gives us three different services:

- Amazon S3 File Gateway
- Amazon FSx for Windows
- Amazon FSx for NetApp ONTAP

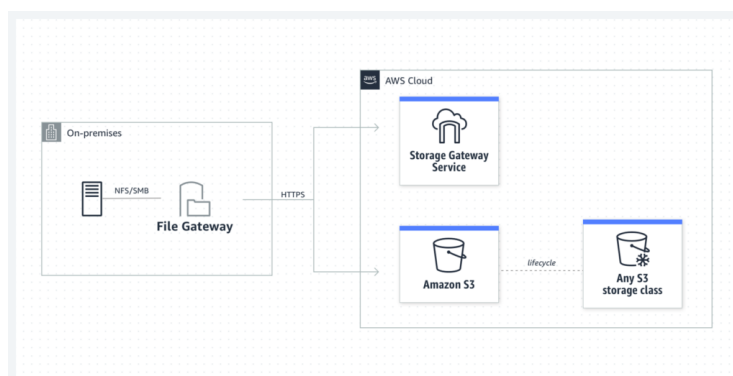
Let's see what they offer and their use case.

## Amazon S3 File Gateway

Amazon S3 File Gateway combines the world of object and file storage, so you could think it unlocks unlimited scalable storage with the famous "11 9s" of durability. Unfortunately, as we'll see, all that glitters is no gold.

With Amazon S3 File Gateway, you can use the SMB protocol to store files in S3, taking advantage of its scalability and automatically mapping API calls.

You can even use **lifecycle policies** to lower storage costs and archive or delete all files. There are no license costs, and you can deploy the solution into your on-premise environment using its internal cache to speed up access to files.



<https://aws.amazon.com/it/storagegateway/file/?nc=sn&loc=2&dn=2>

This solution has some downsides: modifying large data files will create new versions of the S3 object every time, impacting performance and costs.

There's a hard limit of 50 shares per Amazon S3 File Gateway appliance, and additionally, since its backend storage is S3, Amazon S3 File Gateway uses object metadata to map filesystems attributes. Hence, everything, including File Access Control Entries (ACE), has to be stored in metadata that is limited to 2KB in size.

The consequence is the error:

```
"ERROR 1344 (0x00000540)"
```

when using robocopy and transferring files with more than 10 Access Control Entries

(See [Troubleshooting: File Gateway issues](#))

Let's see how to overcome these limitations.

## Amazon FSx for Windows

Amazon FSx for Windows gives you a fully managed high-performance Windows fileserver compatible with standard Windows Management Instrumentation. You can deploy it in Multi-AZ mode to be highly available, even during the patch operations (managed by AWS).

It supports all the features a native solution can offer, like Shadow Copies, data deduplication, and DFS namespaces, while offloading management operations and automating encryption using AWS KMS.

Using PowerShell and ad-hoc developed commandlets, you can automate the management and even seamlessly migrate existing shares:

<https://docs.aws.amazon.com/fsx/latest/WindowsGuide/migrate-file-share-config-to-fsx.html>

In addition, you can enable automated backups and use AWS Backup for different schedules and data retention policies and have an on-premise cache appliance (Amazon Storage Gateway for FSx) to speed up access times while lowering bandwidth utilization.

As always, there's a downside: you have two different storage tiers: SSD and HDD, but you can't switch between them. You can choose the storage type only when you create an Amazon FSx for Windows filesystem.

If you need to reduce storage costs, auto-tiering is not an option in this scenario. Guess what? There's another service for that!

## **Amazon FSx for NetApp ONTAP**

Amazon FSx for NetApp ONTAP is again a fully-managed service that uses high-performance SSD storage, offering all the features an ONTAP appliance gives in an on-premise environment, such as:

- Auto-tiering: data can be automatically moved from SSD storage to HDD if accessed less frequently.
- Multi-protocol access: Besides Windows SMB Shares, it offers NFS and iSCSI volumes).
- Standard management tools used for other hardware
- Snapshots (even between regions) and file restore

In addition to all of these features on the AWS Cloud, you can deploy Multi-AZ configurations and integrate IAM resource-level permissions. Like Amazon FSx for Windows, Amazon automatically applies patches and updates during your defined maintenance window.

It seems like we found the sweet spot for hybrid cloud integrations, but:

- The minimum capacity for an Amazon FSx for NetApp ONTAP filesystem is 1 Terabyte.
- The learning curve can be steep if you have yet to gain experience with this kind of storage.
- You have to consider data transfer costs for data transfer and backup carefully.

## **Transferring your data**

You can transfer data using the good old RoboCopy or AWS DataSync. Both of them are reliable and robust file copy solutions. If you have already mastered scripting all the command line switches for RoboCopy, you can continue reusing the same knowledge you acquired over the years!

## **Choosing the right storage service**

As we said, no solution can accommodate all needs. We didn't talk about cost optimization: it's time to add this variable and see how every service fits in an example scenario, comparing costs, performance, and high availability.

Let's assume we have 5 Terabytes of data used as an archive, and only about 500 Gigabytes (10%) are accessed more frequently.

## Amazon S3 File Gateway

With Amazon S3 File Gateway, we can have a local on-premise cache, while storage is backed by S3. Since it is a single appliance, there's no high availability for the share. Data will always be safe and retrievable because Amazon S3 backs it, and we can use S3 Intelligent-Tiering

File access performance will be ok for files already stored in the cache, but when requesting files, latency will increase because you need to transfer data. You will also need at least 100 Gigabytes of storage space on-prem for caching.

Using a local on-premise appliance will not impact the AWS bill, but please remember that no hardware is TCO-free.

Using S3 Intelligent-Tiering helps automatically lower storage costs, we assumed that all data would be cold, but it gets accessed in 30 days, so it will cost about 70\$ / month.

## Amazon FSx for Windows

Amazon FSx for Windows, as we said before, doesn't let you auto-tier your data; since the access pattern is unknown, we'll assume that all our 5 TB of data will be on SSD storage.

A multi-az deployment and assuming a 50% deduplication factor could cost 687\$/month.

If we can choose to have less performance using standard HDD storage, it will cost 103\$ per month.

Like Amazon S3 File Gateway, you can have an on-premise appliance to cache files: all the previous considerations are valid, but you'll need to consider additional hourly costs, even if you host it on an on-premise server)

## Amazon FSx for NetApp ONTAP

Using the same data deduplication factor, with a Multi-AZ deployment, the total costs will be 581\$ per month.

Here is a quick recap:

Monthly cost	Description	Config Summary
687.28 USD	FSx all on SSD with 50% deduplication savings	Desired storage capacity (5 TB), Desired aggregate throughput (0 M...
581.56 USD	FSx for NetApp ONTAP	Desired aggregate throughput (1 MBps), Desired storage capacity (...)
70.20 USD	Storage Gateway with S3 Intelligent-Tiering cost	S3 INT Average Object Size (16 MB), S3 INT storage (5 TB per month)

As we see, costs change a lot, but features and performance are entirely different.

Amazon Storage Gateway is inexpensive but without high availability, while the configurations we examined for Amazon FSx for Windows and Amazon FSx for NetApp ONTAP are in a multi-AZ configuration. Auto-tiering is valuable and can help lower bills, but with NetApp ONTAP, you also pay for other valuable features like snapshots, backups, and multi-protocol sharing.

If you plan to migrate your Windows workload and refactor using cloud-native technologies, you can use Amazon S3 File Gateway. You can change your application and seamlessly use S3 without migrating data.

If you continue to need Windows native file-sharing services, Amazon FSx for Windows will enable you to lower administrative costs.

If you have a hybrid cloud scenario, Amazon for NetApp ONTAP will fit perfectly, giving you all the features you already use on-premise.

## To conclude

As you see, there is a lot of choice for storage services, and depending on your requirements, they can considerably increase your monthly bill. As always, using cloud-native solutions (like Amazon S3) helps you improve scalability, resilience, and lower costs, but you must consider your needs and find your perfect fit.

Do you already use these services? Migrating data can also be tricky. If you have any exciting stories about migrations, you can share them in the comments below!

---

## About Proud2beCloud

**Proud2beCloud** is a blog by [beSharp](#), an Italian APN Premier Consulting Partner expert in designing, implementing, and managing complex Cloud infrastructures and advanced services on AWS. Before being writers, we are Cloud Experts working daily with AWS services since 2007. We are hungry readers, innovative builders, and gem-seekers. On Proud2beCloud, we regularly share our best AWS pro tips, configuration insights, in-depth news, tips&tricks, how-tos, and many other resources. Take part in the discussion!



## **Damiano Giorgi**

Ex on-prem systems engineer, lazy and prone to automating boring tasks. In constant search of technological innovations and new exciting things to experience. And that's why I love Cloud Computing! At this moment, the only "hardware" I regularly dedicate myself to is that my bass; if you can't find me in the office or in the band room try at the pub or at some airport, then!

---

Copyright © 2011-2023 by beSharp spa - P.IVA IT02415160189