AWS Decision Guide

Choosing an AWS storage service

Choosing an AWS storage service: AWS Decision Guide

Table of Contents

Choosing an AWS storage service

Taking the first step

Purpose	Help determine which AWS storage service is the best fit for your organization.
Last updated	June 26, 2024
Covered services	- Amazon S3 - Amazon EBS - Amazon EFS - Amazon FSx - Amazon File Cache - AWS Backup - AWS DataSync - AWS Snow Family - AWS Storage Gateway - AWS Transfer Family

Introduction

AWS offers a broad portfolio of reliable, scalable, and secure storage services for storing, accessing, protecting, and analyzing your data. This makes it easier to match your storage methods with your needs, and provides storage options that are not easily achievable with on-premises infrastructure. When selecting a storage service, ensuring that it aligns with your access patterns will be critical to achieving the performance you want.

You can select from block, file, and object storage services as well as cloud data migration options for your workload. Choosing the right storage service for your workload requires you to make a series of decisions based on your business needs.

This decision guide will help you ask the right questions, provide a clear path for implementation, and help you migrate from your existing on-premises storage.

This six minute clip is from a 55 minute recording of a presentation by AWS senior storage solutions architects Kevin McDonald and Victor Munoz at the 2022 AWS Summit. It provides an overview of available AWS storage services.

Understand

Data is a cornerstone of successful application deployments, analytics workflows, and machine learning innovations. Well-architected systems use multiple storage services and enable different features to improve performance.

In many cases, however, choosing the right storage service will start with how well it aligns with what you're already using (or are familiar with). Working with storage services that you are familiar with will make it easier for you to get started - and can make migration of your data easier and potentially faster.
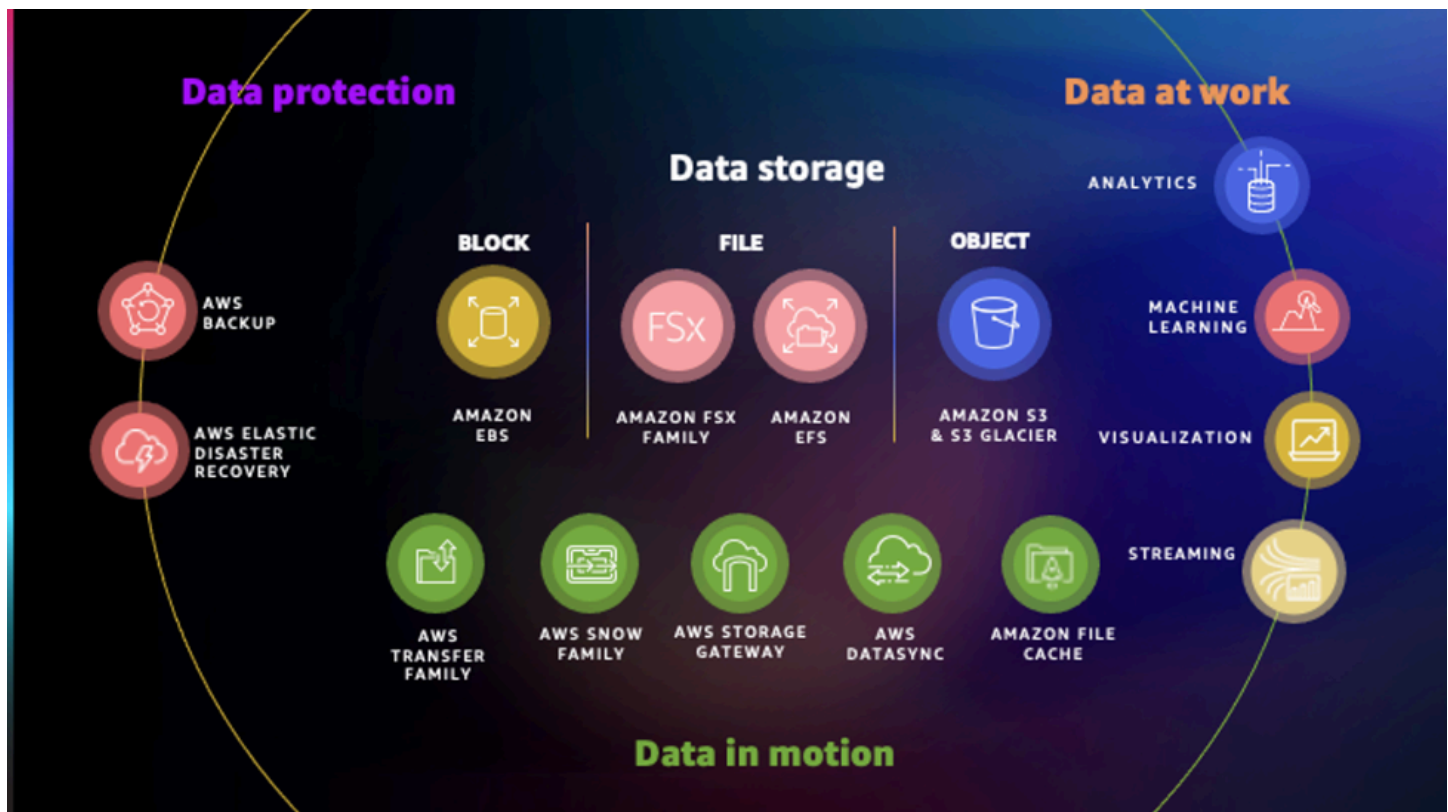
For example, services in the Amazon FSx data storage family come in four options that align to popular file systems:

- **Amazon FSx for Windows File Server** provides fully managed Microsoft Windows file servers, backed by a fully native Windows file system.

- **Amazon FSx for Lustre** allows you to launch and run the high-performance Lustre file system.

- **Amazon FSx for OpenZFS** a fully managed file storage service that enables you to move data to AWS from on-premises ZFS or other Linux-based file servers.

- **Amazon FSx for NetApp ONTAP** a fully managed service that provides highly reliable, scalable, high-performing, and feature-rich file storage built on NetApp's popular ONTAP file system.

Definitions

There are AWS service options for the following storage types:

- **Block** — Block storage is technology that controls data storage and storage devices. It takes any data, like a file or database entry, and divides it into blocks of equal sizes. The block storage system then stores the data block on underlying physical storage in a manner that is optimized for fast access and retrieval.

- **File system** — File systems store data in a hierarchical structure of files and folders. In network environments, file-based storage often uses network-attached storage (NAS) technology. NAS allows users to access network storage data in similar ways to a local hard drive. File storage is user-friendly and allows users to manage file-sharing control.

- **Object** — Object storage is a technology that stores and manages data in an unstructured format called objects. Each object is tagged with a unique identifier and contains metadata that describes the underlying content.

- **Cache** — A cache is a high-speed data storage layer used to temporarily store frequently accessed or recently used data closer to the point of access, with the aim of improving system performance and reducing latency. It serves as a buffer between the slower and larger primary storage (such as disks or remote storage) and the computing resources that need to access the data.

- **Hybrid/Edge** — Hybrid/Edge storage combines on-premises storage infrastructure with cloud storage services, allowing data mobility between the two environments based on requirements like performance, cost, and compliance. It provides benefits such as low-latency access, cost optimization, data sovereignty, cloud scalability, and business continuity.

Migration options

In addition to choosing a storage service, you will need to make choices about how you migrate your data to live within the chosen services. AWS offers several choices to migrate your data - based on whether it needs to live online or offline.

- **Online migration** involves transferring data and applications over the internet while they are still running in the on-premises data center. This approach can be more efficient than offline migration since it minimizes downtime and enables organizations to start using cloud resources sooner. However, it requires a reliable internet connection and may not be suitable for large amounts of data or mission-critical applications.

- **Offline migration** involves moving data and applications without any connection to the internet. This approach requires physically transporting the data on external hard drives or other storage media to the cloud provider's data center. This method is typically used when there are large amounts of data to transfer, limited bandwidth or connectivity, or concerns about security and privacy.

There are two key considerations:

- **Speed -** Choose online migration when speed matters. Online is measured in minutes or hours, and offline can be measured by days. If data is frequently updated and time-critical, choose online. Choose offline when it's a one-time move, and not time-critical.

- **Bandwidth -** Moving data online takes away from available bandwidth used for day-to-day. Choose offline when there are network constraints, and data can be offline while in transit without disrupting your business. AWS services in the Snow Family offer an option for offline migration.

Consider

You might be considering AWS storage services because you are migrating an existing application to the cloud or building a new application in the cloud. When moving data to the cloud, it is important for you to understand where you are moving it, the potential use cases, the type of data you are moving, and the network resources available.

Here's some of the criteria to consider when choosing an AWS storage service.

Protocol

AWS storage services offer multiple protocol options:

- **Block storage** offers high-performance storage that is direct-attached to a compute instance with low-latency access, making it suitable for applications that require fast and consistent I/O operations.

- **File-based storage** is natively mountable from virtually any operating system using industry-standard protocols like NFS and SMB. It provides simple storage for workloads that need access to shared data across multiple compute instances.

- **Object storage** provides easy access to data through an application programming interface (API) over the internet and is well-suited to read-heavy workloads (such as streaming applications and services).

Protocols play a crucial role when considering AWS storage services as they determine how data is accessed, transferred, and managed within the storage environment.

Client type

It's important to consider the operating system of the clients that will be accessing the data. Windows-based clients can use file-based storage options such as Amazon FSx for Windows File Server. It provides highly available storage to your Windows applications with full Server Message Block (SMB) support.

Amazon FSx for Lustre (for high-performance file systems) is designed for use with Unix/Linux-based file systems. FSx for Lustre is optimized for workloads where speed matters, such as machine learning, high performance computing (HPC), video processing, and financial modeling.

The choice of client type for an AWS storage service is critical to ensure easy access and sharing of data across workloads. Selecting a service that is compatible with the file systems and protocols used by your clients is key to avoiding compatibility issues and ensuring seamless data access and transfer.

Performance

Performance is a critical factor to consider when choosing an AWS storage service. There are several factors to consider when evaluating storage performance, including IOPS (input/output operations per second), access patterns, latency, and throughput or bandwidth. It is important to ask questions such as:

- Is your workload latency sensitive?

- Do other metrics (such as IOPS or throughput) dominate your applications performance profile?

- Is your workload read or write-heavy?

Migration strategy and risks

The skills of your organization are a major factor when deciding which container services you use. The approach you take can require some investment in DevOps and Site Reliability Engineer (SRE) teams. Building out an automated pipeline to deploy applications is common for most modern application development.

Some factors to consider when migrating your on-premises storage to AWS are:

- **Data transfer**: what is the most efficient method to transfer your data to AWS?

- **Compatibility**: For example, if you already leverage NetApp ONTAP appliances on-premises services (such as Amazon FSx for NetApp ONTAP) provide a seamless migration path.

- **Application integration**: Evaluate how your applications will integrate with AWS storage services. Consider any necessary modifications or configurations required to enable seamless connectivity and functionality between your applications and the AWS environment.

- **Data Management and lifecycle:** Plan for data management tasks such as backup, replication, and lifecycle management in the AWS environment. Consider AWS services and features that can help automate these tasks, such as versioning, lifecycle policies, and cross-region replication.

- **Security and compliance**: Ensure that your data remains secure during the migration process. Implement appropriate security measures, such as encryption and access controls, to protect your data both in transit and at rest.

- **Cost optimization**: Analyze the cost implications of migrating your storage solution to AWS. Consider factors such as storage pricing, data transfer costs, and any associated services or features required to optimize costs.

By carefully considering these factors, you can ensure a successful migration from an on-premises storage solution to AWS storage services, minimizing disruptions, and maximizing the benefits of cloud storage.

Backup and protection requirements

Backup and protection requirements are critical factors to consider when choosing an AWS storage service because they help ensure the availability and durability of your data.

Without adequate backup and protection measures, data can be lost due to accidental deletion, hardware failure, or natural disasters, which can have severe consequences for your business.

Familiarize yourself with services such as AWS Backup, which can backup your data on demand or automatically as part of a scheduled backup plan. AWS Backup also offers cross-region replication which can be particularly valuable if you have business continuity or compliance requirements to store backups a minimum distance away from your production data.

Disaster recovery

Disaster recovery is a critical consideration when choosing an AWS storage service because it helps ensure business continuity in the event of a disaster or outage. A disaster can be caused by various factors, such as natural disasters, human error, or cyber attacks, and can result in significant data loss and downtime.

Choosing a storage service that provides disaster recovery features, such as replication across multiple availability zones, can help minimize the impact of a disaster on your business. It's important to consider factors such as recovery time objectives (RTO) and recovery point objectives (RPO) when evaluating disaster recovery options and choose a storage service that meets your business needs.

Cost

Beyond the base storage costs, there are other factors that impact pricing such as storage capacity, data transfer, and availability that impacts the total cost of storage. The following can help you reduce cost when using an AWS storage service:

- Use the appropriate storage service for your workload type.

- Use AWS Cost Explorer and other billing tools to monitor organizational speed.

- Understand your data and how it is being used.

We also recommend that you use the AWS Pricing Calculator to estimate your cost when choosing an AWS storage service.

Security

Security at AWS is a [shared responsibility](). AWS provides a secure foundation for customers to build and deploy their applications, but customers are responsible for implementing their own security measures to protect their data, applications, and infrastructure.

You should consider aspects of security such as access control, data encryption, compliance requirements, monitoring and logging, and incident response when choosing an AWS storage service. By doing so, you can help ensure that your data is protected while using AWS services.

Choose

Now that you know the criteria you should use to evaluate your storage options, you are ready to choose which AWS storage services are right for your business needs.

The following table highlights which storage options are optimized for which circumstances. Use it to help determine the one that is the best fit for your use case.

Storage type	What is it optimized for?	Storage services or tools
Block	Applications requiring low-latency, high-performance durable storage attached to single Amazon EC2 instances or containers, such as databases and general-purpose local instance storage.	Amazon EBS Amazon EC2 instance store
File system	Applications and workloads requiring shared read and write access across multiple Amazon EC2 instances or containers, or from multiple on-prem servers, such as team file shares, highly-available enterprise applications, analytics workloads, and ML training.	Amazon EFS Amazon FSx Amazon FSx for Lustre Amazon FSx for NetApp ONTAP Amazon FSx for OpenZFS

Storage type	What is it optimized for?	Storage services or tools
		Amazon FSx for Windows File Server
		Amazon S3 File Gateway
		Amazon FSx File Gateway
Object	Read-heavy workloads such as content distribution, web hosting, big data analytics, and ML workflows. Well-suited for scenarios where data needs to be stored, accessed, and distributed globally over the internet.	Amazon S3
Cache	Fully managed, scalable, and high-speed cache on AWS for processing file data stored in disparate locations—including on premises NFS file systems, and/or in cloud file systems (Amazon FSx for OpenZFS, Amazon FSx for NetApp ONTAP), and Amazon S3.	Amazon File Cache
Hybrid/Edge	Deliver low-latency data to on-premises applications and providing on-premises applications access to cloud-backed storage.	AWS Storage Gateway Tape Gateway AWS Storage Gateway Volume Gateway

The following table provides a detailed look at your online and offline options.

Migration options	When speed is the priority	When bandwidth is important	Storage services or tools
Online	Online is optimized for frequent updates to data. Use it for time-critical or ongoing workloads.	Consider scheduling your transfer during off hours when you have sufficient bandwidth.	AWS DataSync AWS Transfer Family Amazon FSx for NetApp ONTAP SnapMirror AWS Storage Gateway
Offline	Suitable for one-time or periodic uploads - and when data can be static in transit.	This choice makes sense when you need to use only the minimum available bandwidth - and you prefer the predictability of physical moves.	AWS Snowball

Use

Now that you have determined the best protocol you need to work with your data, your performance requirements, and other criteria discussed in this guide, you should also have an understanding of which storage service would be the best fit for your needs.

To explore how to use and learn more about each of the available AWS storage services - we have provided a pathway to explore how each of the services work. The following section provides links to in-depth documentation, hands-on tutorials, and resources to get you started.

Amazon S3

- **Getting started with Amazon S3**

This guide will help you get started with Amazon S3 by working with buckets and objects. A bucket is a container for objects. An object is a file and any metadata that describes that file.

Explore the guide

- **Optimizing Amazon S3 performance**

  When building applications that upload and retrieve storage from Amazon S3, follow the AWS best practices guidelines in this paper to optimize performance.

  Read the whitepaper

- **Amazon S3 tutorials**

  The following tutorials present complete end-to-end procedures for common Amazon S3 tasks. These tutorials are intended for a lab-type environment and provide general guidance.

  Get started with the tutorials

Amazon EBS

- **Getting started with Amazon EBS**

  Amazon EBS is recommended for data that must be quickly accessible and requires long-term persistence.

  Explore the guide

- **Create an Amazon EBS volume**

  An Amazon EBS volume is a durable, block-level storage device that you can attach to your instances.

  Get started with the tutorial

- **Use Amazon EBS direct APIs to access the contents of an Amazon EBS snapshot**

  You can use the direct APIs to create Amazon EBS snapshots, write and read data on your snapshots, and identify differences.

  Explore the guide

Amazon EFS

- **Getting started with Amazon EFS**

  Learn how to create an Amazon EFS file system. You will mount your file system on an Amazon EC2 instance in your VPC, and test the end-to-end setup.

  [Get started with the tutorial](#)

- **Create a Network File System**

  Learn how to store files and create an Amazon EFS file system, launch a Linux virtual machine on Amazon EC2, mount the file system, create a file, terminate the instance, and delete the file system.

  [Get started with the tutorial](#)

- **Set up an Apache web server and serve Amazon EFS files**

  Learn how to set up an Apache web server on an Amazon EC2 instance and set up an Apache web server on multiple Amazon EC2 instances by creating an Auto Scaling group.

  [Get started with the tutorial](#)

Amazon FSx

- **Getting started with Amazon FSx**

  This getting started guide walks you through what you'll need to do to begin using Amazon FSx.

  [Explore the guide](#)

- **Getting started with Amazon FSx for Lustre**

  Learn how to use your Amazon FSx for Lustre file system to process the data in your Amazon S3 bucket with your file-based applications.

  [Explore the guide](#)

- **What is Amazon FSx for Windows File Server?**

  This guide provides an introduction to Amazon FSx for Windows File Server.

Explore the guide

- **Getting started with Amazon FSx for NetApp ONTAP**

  Learn how to get started using Amazon FSx for NetApp ONTAP.

  Get started with the tutorial

- **Learn how to get started with Amazon FSx for OpenZFS**

  This guide provides an introduction to Amazon FSx for OpenZFS.

  Get started with the tutorial

Amazon File Cache

- **Getting started with Amazon File Cache**

  Learn how to create an Amazon File Cache resource and access it from your compute instances.

  Get started with the tutorial

- **Amazon File Cache in action**

  This video shows how Amazon File Cache can be used as a temporary high performance storage location for data stored in on premises file systems.

  Watch the video

AWS Storage Gateway

- **User guide for Amazon S3 File Gateway**

  Describes Amazon S3 File Gateway concepts and provides instructions on using the various features with both the console and the API.

  Explore the guide

- **User guide for Amazon FSx File Gateway**

Describes Amazon FSx File Gateway, which provides access to in-cloud Amazon FSx for Windows File Server shares from on-premises facilities. Includes instructions on working with the console and the API.

[Explore the guide](#)

- **User guide for Tape Gateway**

Describes Tape Gateway, a durable, cost-effective tape-based solution for archiving data in the AWS cloud. Provides concepts and instructions on using the various features with both the console and the API.

[Explore the guide](#)

- **User guide for Volume Gateway**

Describes Volume Gateway concepts, including details about cached and stored volume architectures, and provides instructions on using their features with both the console and the API.

[Explore the guide](#)

AWS DataSync

- **Getting started with AWS DataSync**

This guide walks through how you can get started with AWS DataSync by using the AWS Management Console.

[Explore the guide](#)

- **Simplify multicloud data movement wherever data is stored with AWS DataSync**

AWS DataSync supports incremental transfers, integration with IAM for access control, and use cases like data migration, replication, and distribution across AWS Regions or accounts.

[Read the blog](#)

- **AWS DataSync tutorials**

These tutorials walk you through some real-world scenarios with AWS DataSync and transferring data.

[Get started with the tutorials](#)

AWS Transfer Family

- **Getting started with AWS Transfer Family**

  Learn how to create an SFTP-enabled server with publicly accessible endpoint using Amazon S3 storage, add a user with service-managed authentication, and transfer a file with Cyberduck.

  [Get started with the tutorial](#)

- **AWS Transfer Family in action**

  This video shows how the AWS Transfer Family can be used for each of the three supported protocols (SFTP, FTPS, and FTP), both over the public internet, as well as within a VPC.

  [Watch the video](#)

- **AWS Transfer Family for AS2**

  Learn how to set up an Applicability Statement 2 (AS2) configuration with AWS Transfer Family.

- **AWS Transfer Family SFTP Connectors**

  Learn how to set up an SFTP connector, and then transfer files between Amazon S3 storage and an SFTP server.

AWS Snow Family

- **Getting started with AWS Snow Family**

  These guides provide links to documentation covering all current services in the Snow Family.

  [Explore the guides](#)

- **AWS Snowball Edge developer guide**

  This guide includes guidance for local storage and compute, clustering, importing and exporting data into Amazon S3, and other features of a Snowball Edge device.

Explore the guide

Explore

- Developers
- Solution Architects
- Decision makers

- **Architecture diagrams**

  Explore reference architecture diagrams for containers on AWS.

  Explore architecture diagrams
- **Whitepapers**

  Explore whitepapers to help you get started and learn best practices.

  Explore whitepapers
- **AWS Solutions**

  Explore vetted solutions and architectural guidance for common use cases for containers.

  Explore solutions

Document history

The following table describes the important changes to this decision guide. For notifications about updates to this guide, you can subscribe to an RSS feed.

Change	Description	Date
Guide updated	Migrated to docs.aws. amazon.com, and made minor updates to Understan d, Consider, Choose, and Use sections.	June 26, 2024
Guide updated	Added AWS Copilot, AWS Batch, and AWS Outposts. Changed capacity, orchestra tion, and provisioning to compute capacity, orchestra tion, and vertical solutions. Numerous editorial changes throughout.	April 5, 2024
Initial publication	Guide first published.	April 26, 2023