

# Introduction to AWS Storage

This whitepaper introduces the storage services available from Amazon Web Services (AWS), designed to address a wide range of storage requirements for various applications and workloads. AWS offers a comprehensive suite of storage solutions that are scalable, flexible, and durability, enabling businesses to efficiently store, manage, and protect their data.

## Topics

- [Amazon Elastic Block Store](#)
- [Amazon Elastic File System](#)
- [Amazon File Cache](#)
- [Amazon FSx for Lustre](#)
- [Amazon FSx for OpenZFS](#)
- [Amazon FSx for NetApp ONTAP](#)
- [Amazon FSx for Windows File Server](#)
- [Amazon Simple Storage Service](#)
- [AWS Backup](#)
- [AWS Storage Gateway](#)

## Amazon Elastic Block Store

[Amazon Elastic Block Store](#) (Amazon EBS) provides persistent block storage volumes for use with Amazon EC2 instances in the AWS Cloud. Each Amazon EBS volume is automatically replicated within its Availability Zone to protect you from component failure, offering high availability and durability. Amazon EBS volumes offer the consistent and low-latency performance needed to run your workloads. With Amazon EBS, you can scale your usage up or down within minutes—all while paying a low price for only what you provision.

## Amazon Elastic File System

[Amazon Elastic File System \(Amazon EFS\)](#) provides a simple, scalable, elastic file system for Linux-based workloads for use with AWS Cloud services and on-premises resources. It is built to scale on demand to petabytes without disrupting applications, growing and shrinking automatically as you add and remove files, so your applications have the storage they need – when they need it. It is designed to provide massively parallel shared access to thousands of Amazon EC2 instances, enabling your applications to achieve high levels of aggregate throughput and IOPS with consistent low latencies.

Amazon EFS is a fully managed service that requires no changes to your existing applications and tools, providing access through a standard file system interface for seamless integration. Amazon EFS is a regional service storing data within and across multiple Availability Zones (AZs) for high availability and durability. You can access your file systems across Availability Zones and AWS Regions and share files between thousands of Amazon EC2 instances and on-premises servers via AWS Direct Connect or AWS VPN.

Amazon EFS is well suited to support a broad spectrum of use cases from highly parallelized, scale-out workloads that require the highest possible throughput to single-threaded, latency-sensitive workloads. Use cases such as lift-and-shift enterprise applications, big data analytics, web serving and content management, application development and testing, media and entertainment workflows, database backups, and container storage.

## Amazon File Cache

[Amazon File Cache](#) is a fully managed high-speed cache on AWS that makes it easier to process file data, regardless of where the data is stored. Amazon File Cache serves as temporary, high-performance storage for data in on-premises file systems, or in file systems or object stores on AWS. The service allows you to make dispersed datasets available to file-based applications on AWS with a unified view and high speeds. You can link the cache to multiple NFS—including on-premises and in-cloud—or [Amazon Simple Storage Service \(Amazon S3\)](#) buckets, providing a unified view of and fast access to your data spanning on-premises and multiple AWS Regions. The cache provides read and write data access to compute workloads on AWS with sub-millisecond latencies, up to hundreds of GB/s of throughput, and up to millions of [IOPS](#).

## Amazon FSx for Lustre

[Amazon FSx for Lustre](#) is a fully managed file system that is optimized for compute-intensive workloads, such as high performance computing, machine learning, and media data processing workflows. Many of these applications require the high-performance and low latencies of scale-out, parallel file systems. Operating these file systems typically requires specialized expertise and administrative overhead, requiring you to provision storage servers and tune complex performance parameters. With Amazon FSx, you can launch and run a Lustre file system that can process massive data sets at up to hundreds of gigabytes per second of throughput, millions of IOPS, and sub-millisecond latencies.

Amazon FSx for Lustre is seamlessly integrated with Amazon S3, making it easy to link your long-term data sets with your high performance file systems to run compute-intensive workloads. You can automatically copy data from S3 to Amazon FSx for Lustre, run your workloads, and then write results back to S3. Amazon FSx for Lustre also enables you to burst your compute-intensive workloads from on-premises to AWS by allowing you to access your FSx file system over Amazon Direct Connect or VPN. Amazon FSx for Lustre helps you cost-optimize your storage for compute-intensive workloads: It provides cheap and performant non-replicated storage for processing data, with your long-term data stored durably in Amazon S3 or other low-cost data stores. With Amazon FSx, you pay for only the resources you use. There are no minimum commitments, upfront hardware or software costs, or additional fees.

## Amazon FSx for OpenZFS

[Amazon FSx for OpenZFS](#) is a fully managed file storage service that lets you launch, run, and scale fully managed file systems built on the open-source OpenZFS file system. Amazon FSx for OpenZFS makes it easy to migrate your on-premises file servers—without changing your applications or how you manage data—and build new high-performance, data-driven applications in the cloud. Amazon FSx for OpenZFS offers the familiar features, performance, and capabilities of OpenZFS file systems with the agility, scalability, and simplicity of a fully managed AWS service.

## Amazon FSx for NetApp ONTAP

[Amazon FSx for NetApp ONTAP](#) offers the first complete, fully managed NetApp file system available in the cloud making it easy for you to migrate or extend existing applications to AWS without changing code or how you manage your data. Built on NetApp ONTAP, Amazon FSx for NetApp ONTAP provides the familiar features, performance, capabilities, and APIs of NetApp file systems with the agility, scalability, and simplicity of a fully managed AWS service.

Amazon FSx for NetApp ONTAP offers high-performance file storage that is broadly accessible from Linux, Windows, and macOS compute instances via the industry-standard NFS, SMB, and iSCSI protocols. With Amazon FSx for NetApp ONTAP, you get low-cost, fully elastic storage capacity with support for compression and deduplication to help you further reduce storage costs. Amazon FSx for NetApp ONTAP file systems can be deployed and managed using the AWS Management Console or NetApp Cloud Manager for seamless set up and administration.

## Amazon FSx for Windows File Server

[Amazon FSx for Windows File Server](#) provides a fully managed native Microsoft Windows file system so you can easily move your Windows-based applications that require file storage to AWS. Built on Windows Server, Amazon FSx provides shared file storage with the compatibility and features that your Windows-based applications rely on, including full support for the SMB protocol and Windows NTFS, Active Directory (AD) integration, and Distributed File System (DFS). Amazon FSx uses SSD storage to provide the fast performance your Windows applications and users expect, with high levels of throughput and IOPS, and consistent sub-millisecond latencies.

This compatibility and performance is particularly important when moving workloads that require Windows shared file storage, such as CRM, ERP, and .NET applications, as well as home directories. With Amazon FSx, you can launch highly durable and available Windows file systems that can be accessed from up to thousands of compute instances using the industry-standard SMB protocol. Amazon FSx eliminates the typical administrative overhead of managing Windows file servers. You pay for only the resources used, with no upfront costs, minimum commitments, or additional fees.

## Amazon Simple Storage Service

[Amazon Simple Storage Service](#) (Amazon S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance. This means customers of all sizes and industries can use it to store and protect any amount of data for a range of use cases, such as websites, mobile applications, backup and restore, archive, enterprise applications, IoT devices, and big data analytics. Amazon S3 provides easy-to-use management features so you can organize your data and configure finely-tuned access controls to meet your specific business, organizational, and compliance requirements. Amazon S3 is designed for 99.999999999% (11 9s) of durability, and stores data for millions of applications for companies all around the world.

[Amazon S3 storage classes](#) are a range of storage classes that you can choose from based on the data access, resiliency, and cost requirements of your workloads. S3 storage classes are purpose-built to provide the lowest cost storage for different access patterns. S3 storage classes are ideal for virtually any use case, including those with demanding performance needs, data residency requirements, unknown or changing access patterns, or archival storage.

The S3 storage classes include:

- ❑ **S3 Intelligent-Tiering** for automatic cost savings for data with unknown or changing access patterns
- ❑ **S3 Standard** for frequently accessed data
- ❑ **S3 Standard-Infrequent Access (S3 Standard-IA)** and **S3 One Zone-Infrequent Access (S3 One Zone-IA)** for less frequently accessed data
- ❑ **S3 Glacier Instant Retrieval** for archive data that needs immediate access
- ❑ **S3 Glacier Flexible Retrieval** (formerly S3 Glacier) for rarely accessed long-term data that does not require immediate access
- ❑ **Amazon S3 Glacier Deep Archive (S3 Glacier Deep Archive)** for long-term archive and digital preservation with retrieval in hours at the lowest cost storage in the cloud

If you have data residency requirements that can't be met by an existing AWS Region, you can use the **S3 Outposts** storage class to store your S3 data on premises. Amazon S3 also offers capabilities to manage your data throughout its lifecycle. Once an S3 Lifecycle policy is set, your data will automatically transfer to a different storage class without any changes to your application. For more information, refer to the [Amazon S3 storage classes overview info graphic](#). You can use [S3 Object Lock](#) to help prevent S3 objects from being deleted or overwritten for a fixed amount of time, or indefinitely. Object Lock can help you to meet regulatory requirements that require WORM (write-once-read-many) storage, or to simply add another layer of protection against object changes or deletion.

## AWS Backup

[AWS Backup](#) enables you to centralize and automate data protection across AWS services. AWS Backup offers a cost-effective, fully managed, policy-based service that further simplifies data protection at scale. AWS Backup also helps you support your regulatory compliance or business policies for data protection. Together with AWS Organizations, AWS Backup enables you to centrally deploy data protection policies to configure, manage, and govern your backup activity across your organization's AWS accounts and resources, including Amazon Elastic Compute Cloud (Amazon EC2) instances, Amazon Elastic Block Store (Amazon EBS) volumes, Amazon Relational Database Service (Amazon RDS) databases (including Amazon Aurora clusters), Amazon DynamoDB tables, Amazon Elastic File System (Amazon EFS) file systems, Amazon FSx for Lustre file systems, Amazon FSx for Windows File Server file systems, and AWS Storage Gateway volumes.

## AWS Storage Gateway

The [AWS Storage Gateway](#) is a hybrid storage service that allows your on-premises applications to seamlessly use AWS cloud storage. You can use the service for backup and archiving, disaster recovery, cloud data processing, storage tiering, and migration. Your applications connect to the service through a virtual machine or hardware gateway appliance using standard storage protocols, such as NFS, SMB and iSCSI. The gateway connects to AWS storage services, such as Amazon S3, S3 Glacier, and Amazon EBS, and Amazon FSx for Windows File Server, providing storage for files, volumes, and virtual tapes in AWS. The service includes a highly optimized data transfer mechanism, with bandwidth management, automated network resilience, and efficient data transfer, along with a local cache for low-latency on-premises access to your most active data.

Credit to: AWS Documentation