**AWS Storage and Database Services**

**Introduction to AWS Storage Services**

**Types of AWS Storage Services**

AWS offers seven types of storage services with choices for back-up, archiving and recovery of lost data. Let’s see what those services are and their features:

**1. Simple Storage Service (S3)**

Amazon S3 is an object storage service that stores data of any type and size. It can store data for any business such as web applications, mobile applications, backup, archive, analytics. It also provides easy access control management for all your specific requirements and is almost 100% durable and by almost 99.(11 nines)%. It can also be used to store all kinds of file formats as you would with a dropbox. S3 also allows a simple web-based file explorer to upload files, create folders or delete them.

**2. Elastic Block Storage (EBS)**

EBS provides block storage which is similar to hard drives to store any kind of data persistently. This can be attached to any EC2 instance and used as block storage, which even allows you to install any operating system. EBS volumes are placed in availability zones so that they are replicated to prevent loss of data due to single component failures. They provide absolute low-latency performance and you can also scale up or down your resources as and when required. EBS is available in both SSD and HDD formats depending on your requirement of speed and volume.

**3. Elastic File System (EFS)**

EFS is a managed network file system that is easy to set up right from the amazon console or CLI. When you have multiple EC2 instances needed to access the same file system EFS helps in providing just that. Unlike EBS, EFS is built using the NFS4.x protocol on SSDs and have a much faster throughput. This also means that EFS is much more expensive than EBS as it can be used on very large analytical workloads. EFS scales up or down based on the size of the files you store and is also accessible from multiple availability zones. The distributed nature of the file system can tempt you to use it as a CDN. But the costs of a CDN outweigh the benefits of using EFS. Hence it is better to use a CDN and use EFS in conjunction with files that can’t be stored on a CDN.

**4. Amazon FSx for Lustre**

Luster is a file system used for compute-intensive workloads. This mainly comes into the picture when you run machine learning operations on large data sets or when you need to run media encoding workloads. Running Lustre separately requires a lot of expertise in setting it up and configuring it for the right workloads. With the help of Amazon FSx, this can be avoided and a simple interface on the console helps you to quickly get started and start working on your data. The ability to connect it seamlessly to S3 and the option of running it in VPC provides a low cost yet a performant way to achieve your compute-intensive workloads leveraging

**5. Amazon S3 Glacier**

The glacier is used mainly for archival and long-term data storage. This means that there is a low retrieval rate on this storage system due to which it is offered at an extremely cheap rate. It does also come with compliant security features to encrypt your data. Glacier allows you to run queries and analytics on it directly and you will be charged only for the few minutes or hours when you read the data. In terms of durability, it offers 99.(11 nines)% durability which is one of the highest in the industry. Glacier hopes to replace the legacy on-premise tape-based backup service with a much more cost-effective and durable solution.

**6. Amazon FSx for Windows File Server**

Whenever you need to run your windows specific software that needs to access the proprietary windows file system on the cloud, [AWS provides](https://www.educba.com/aws-features/) you with Amazon FSx to easily achieve that. Windows-based .Net applications, ERPs and CRMs require shared file storage to move workloads between them. Also, Amazon FSx provides support for all native windows based technologies such as NTFS, [SMB protocol](https://www.educba.com/what-is-smb/), Active Directory (AD) and Distributed File System (DFS). Similar to luster, Amazon FSx eliminates the administrative overhead for setting up and maintaining a windows file server and provides you with a simple cost-effective way to run your windows file server on AWS.

**7. AWS Storage Gateway**

Storage Gateway is a simple way to let your on-premise applications store, access or archive the data into the AWS cloud. This is achieved by running on a hypervisor on one of the machines in your data center which contains the storage gateway and then is available on AWS to connect to S3, Glacier or EBS. It provides a highly optimized, network resilient and low-cost way to move your data from on-prem to the cloud. Local caching is also available on your on-prem to allow for accessing the more active data. Storage gateway also supports legacy backup stores such as tapes as virtual tapes backed up directly into AWS Glacier.

**Conclusion**

So there you have it. AWS provides you with multiple storage options right from legacy backup to modern high throughput, distributed file systems. Based on your requirements, you can go with one or more storage services to achieve your business goals. AWS provides you with high availability and high durability so that you can always be assured that your data is safe, secure and readily available for all your use cases.

==================================================

## Types of Database Services Offered on AWS

AWS offers a wide range of database services for you to choose from. The service fall into two groups: relational and non-relational (NoSQL).

### AWS Relational Database Services

Relational databases store data in tabular form with columns and rows, and can be queried using the SQL query language. In these databases, columns represent attributes and rows represent records. Each field in the table is a data value.

**Use cases for relational databases on Amazon include:**

* Enterprise resource planning (ERP) apps
* Customer relationship management (CRM) apps
* Finance data
* Transactions
* Data warehousing

**The primary Amazon services providing relational databases are:**

* Amazon Aurora
* Amazon RDS
* Amazon Redshift

## AWS Databases Services

Once you understand what options are available to you for databases in AWS, you can begin narrowing them down. The following services are some of the most commonly used. Keep in mind when reviewing these descriptions that frequently, AWS customers implement multiple database types to meet their needs. You too should consider multiple options if one doesn’t meet all of your needs.

### Amazon RDS

[Amazon RDS](https://aws.amazon.com/rds/) is a managed, relational database service that includes six different database options. These include AWS Oracle, PostgreSQL, AWS MySQL, MariaDB, SQL Server, and Amazon Aurora. You can manage these database engines from a centralized management console, a command-line interface, or via API calls. When using this service, many administrative tasks are automated, including database setup, hardware provisioning, backup, and updating.

Use cases of Amazon RDS include:

* **Web and mobile applications**—provides the scalability, availability, and throughput needed for enterprise-grade applications.
* **eCommerce applications**—provides flexibility, security, and PCI compliance needed for eCommerce.
* **Mobile and online games**—provides high-throughput and availability to ensure that games remain online and responsive to players.

### Amazon Aurora

[Amazon Aurora](https://aws.amazon.com/rds/aurora/) is a fully managed relational database engine designed specifically for AWS. It is MySQL and PostgreSQL compatible with minor changes to your source database. Aurora includes features for self-healing, fault tolerance, point-in-time recovery, and continuous backup.

Use cases for Amazon Aurora include:

* **Enterprise applications**—including customer relationship management and enterprise resource planning solutions.
* **Software as a Service (SaaS) offerings**—including those requiring significant storage and compute scalability.
* **Web and mobile gaming applications**—including those requiring massive storage, high throughput, and high-availability.

### Amazon DynamoDB

[Amazon DynamoDB](https://aws.amazon.com/dynamodb/) is a fully managed, document and key-value database. It includes features for multi-master, multi-region used along with built-in security, automated backup and restoration, and in-memory caching. DynamoDB can provide support for serverless web apps, microservices, and mobile backends.

Use cases of Amazon DynamoDB include:

* **Ad tech**—including clickstreams, user events, and user profiles.
* **Gaming**—including leaderboards, player data stores, and game states.
* **Retail**—including online shopping carts, inventory tracking, and customer profiles.
* **Banking and finance**—including event-driven transaction processing, fraud detection, and change data capture.
* **Media and entertainment**—including digital rights management, user data stores, and metadata stores.
* **Software as a service (SaaS)**—including content metadata stores, metadata caches, and relationship graph data stores.

### NoSQL databases:

The table below shows the main NoSQL databases services offered by AWS.

|  |  |  |
| --- | --- | --- |
| **Type of Database** | **Use Cases** | **Amazon Services** |
| **Key-value**  Key-value databases store data as a collection of key-value pairs with the key as an ID. These databases can store various types of data, including simple and compound objects. | ● Real-time bidding  ● eCommerce shopping carts  ● Product catalogs  ● Customer preferences | ● Amazon DynamoDB |
| **Document**  Document databases store data in JSON or JSON-like documents. You can query data using the same document-model format used in programming applications. | ● Cataloging  ● Content management systems  ● Customer profiles and personalization  ● Mobile apps | ● Amazon DocumentDB |
| **In-memory**  In-memory databases store data in-memory for low-latency access. You can use these stores as a database, cache, message broker, or queue. | ● Caching  ● Session stores  ● Gaming  ● Leaderboards  ● Geospatial services  ● Pub/sub messaging  ● Real-time streaming | ● Amazon ElastiCache for Memcached  ● Amazon ElastiCache for Redis |
| **Graph**  Graph databases are a type of NoSQL (non-relational) database. This database type represents relationships directly. You can query data with specific graph languages. | ● Fraud detection  ● Social networking  ● Recommendation engines  ● Knowledge graphs  ● Data lineage | ● Amazon Neptune |
| **Time-series**  Time-series databases store data in time-order and as append-only. You can query data over various time intervals. | ● DevOps  ● Application monitoring  ● Industrial telemetry  ● IoT applications | ● Amazon Timestream |
| **Ledger**  Ledger databases store data in an immutable, transparent, and cryptographically verifiable log. This log is owned by a trusted central authority to ensure provenance. | ● Finance  ● Manufacturing  ● Insurance claims  ● HR and payroll  ● Retail inventories |  |

### Amazon DocumentDB

[Amazon DocumentDB](https://aws.amazon.com/documentdb/?nc2=type_a) is a fully managed document database service. It is scalable, highly-available, and compatible with MongoDB. With it, you can store, index, and query JSON files. With DocumentDB, you can scale your compute and storage resources separately for maximum flexibility.

Use cases of Amazon DocumentDB include:

* **Content and catalog management**—including online publications, point-of-sale terminals, and digital archives.
* **Profile management**—including user preferences, authentication profiles, and online transactions.
* **Mobile and web applications**—including applications that demand high-performance and low-latency with millions of requests per second.

### Amazon ElastiCache

[Amazon ElastiCache](https://aws.amazon.com/elasticache/) is a fully managed, in-memory data store service. It is compatible with both Redis and Memcached. ElastiCache automates setup, hardware provisioning, configuration, monitoring, updates, and backup and recovery processes. With ElastiCache you can scale both write and memory processes through sharding and data replication.

Use cases of Amazon ElastiCache include:

* **Session stores**—for web applications and sites.
* **Gaming**—including leaderboards and chats.
* **Geospatial services**—including real-time mapping and location.
* **Real-time analytics**—including Internet of things (IoT) sensor processing and AI applications.

### Amazon Neptune

[Amazon Neptune](https://aws.amazon.com/neptune/) is a fully managed graph database service. It enables you to create and run applications using highly-connected data sets. It supports the storage of massive relationship data sets with low-latency access. Neptune supports a variety of graph models and languages, including RDF, SPARQL, and Gremlin. It includes features for point-in-time recovery, read replicas, and continuous backup.

Use cases for Amazon Neptune include:

* **Social networking**—including user profiles and content prioritization.
* **Recommendation engines**—including storage of customer contacts, purchase histories, and customer preferences.
* **Fraud detection**—including fraud related to overlapping email addresses, IP addresses, or credit card numbers.
* **Knowledge graphs**—including product catalogs or wikis.
* **Life sciences**—including disease models, gene patterning, or research catalogs.
* **Network and IT operations**—including creating network visibility, monitoring, or forensic analysis.

### Amazon Timestream

[Amazon Timestream](https://aws.amazon.com/timestream/?c=db&sec=srv) is a fully managed, time-series database service. It enables you to store, process, and analyze up to 1,000X better query performance at 90% lower cost, compared to relational databases offered on AWS. Timestream provides automatic hardware provisioning, updates, setup and configuration, and data tiering.

Use cases for Amazon Timestream include:

* **DevOps**—supports performance monitoring and management, network optimization, and server monitoring.
* **IoT applications**—supports IoT analytics for the implementation of smart devices, such as thermostats or motion sensors.
* **Application monitoring**—supports clickstream monitoring and analysis.
* **Industrial telemetry**—including monitoring of industrial equipment, fleet management, trade monitoring, or route optimization.

### Amazon Quantum Ledger Database (QLDB)

[Amazon (QLDB)](https://aws.amazon.com/qldb/) is a fully managed, serverless ledger database service. You can use it to track application data changes with a verifiable history. With QLDB, you can avoid the need to build custom ledger applications and associated verification tools. You can query data in QLDB using a SQL-like API.

Use cases for Amazon QLDB include:

* **Finance**—including credit and debit transactions.
* **Insurance**—including claim transactions and auditing.
* **HR and payroll**—including employee benefits, performance histories, or certifications.
* **Retail and supply chain management**—including batch tracking, product recall processes, and shipping details.