

Introduction to AWS Database

Topics

- [Compare AWS database services](#)
- [Amazon Aurora](#)
- [Amazon DynamoDB](#)
- [Amazon ElastiCache](#)
- [Amazon Keyspaces \(for Apache Cassandra\)](#)
- [Amazon MemoryDB for Redis](#)
- [Amazon Neptune](#)
- [Amazon Relational Database Service](#)
- [Amazon RDS on VMware](#)
- [Amazon Quantum Ledger Database \(Amazon QLDB\)](#)
- [Amazon Timestream](#)
- [Amazon DocumentDB \(with MongoDB compatibility\)](#)
- [Amazon Lightsail Managed Databases](#)

Compare AWS database services

Database	Use cases	AWS services
Relational	Traditional applications, enterprise resource planning (ERP), customer relationship management (CRM), e-commerce	<ul style="list-style-type: none">• Amazon Aurora — Designed for unparalleled high performance and availability at global scale with full MySQL and PostgreSQL compatibility• Amazon RDS — Set up, operate, and scale a relational database in the cloud with just a few clicks• Amazon Redshift — Accelerate your time to insights with fast, easy, and secure cloud data warehousing at scale

Database	Use cases	AWS services
Key-value	High-traffic web applications, e-commerce systems, gaming applications	<ul style="list-style-type: none"> • Amazon DynamoDB — Fast, flexible NoSQL database service for single-digit millisecond performance at any scale
In-memory	Caching, session management, gaming leaderboards, geospatial applications	<ul style="list-style-type: none"> • Amazon ElastiCache — Unlock microsecond latency and scale with in-memory caching • Amazon MemoryDB for Redis — Redis-compatible, durable, in-memory database service for ultra performance
Document	Content management, catalogs, user profiles	<ul style="list-style-type: none"> • Amazon DocumentDB (with MongoDB compatibility) — Scale JSON workloads with ease using a fully managed document database service
Wide column	High-scale industrial apps for equipment maintenance, fleet management, and route optimization	<ul style="list-style-type: none"> • Amazon Keyspaces — A scalable, highly available, and managed Apache Cassandra-compatible database service
Graph	Fraud detection, social networking, recommendation engines	<ul style="list-style-type: none"> • Amazon Neptune — Build and run graph applications with highly connected datasets
Time series	Internet of Things (IoT) applications, DevOps, industrial telemetry	<ul style="list-style-type: none"> • Amazon Timestream — Fast, scalable, serverless time series database
Ledger	Systems of record, supply chain, registrations, banking transactions	<ul style="list-style-type: none"> • Amazon Ledger Database Service (QLDB) — Maintain an immutable, cryptographically verifiable log of data changes

Amazon Aurora

[Amazon Aurora](#) is a MySQL and PostgreSQL compatible relational database engine that combines the speed and availability of high-end commercial databases with the simplicity and cost-effectiveness of open-source databases.

Amazon Aurora is up to five times faster than standard MySQL databases and three times faster than standard PostgreSQL databases. It provides the security, availability, and reliability of commercial databases at 1/10th the cost. Amazon Aurora is fully managed by Amazon Relational Database Service (Amazon RDS), which automates time-consuming administration tasks such as hardware provisioning, database setup, patching, and backups.

Amazon Aurora features a distributed, fault-tolerant, self-healing storage system that auto-scales up to 128TB per database instance. It delivers high performance and availability with up to 15 low-latency read replicas, point-in-time recovery, continuous backup to Amazon S3, and replication across three Availability Zones (AZs).

Amazon Aurora I/O-Optimized is a cluster configuration that offers improved price performance and predictable pricing for customers with I/O-intensive applications, such as e-commerce applications, payment processing systems, and financial applications. Aurora-Optimized offers improved performance, increasing throughput and reducing latency to support your most demanding workloads, with up to 40 percent cost savings when your I/O spending exceeds 25 percent of your current Aurora database spend.

Amazon Aurora MySQL zero-ETL integration with Amazon Redshift, now available in public preview, enables near real-time analytics and machine learning of data stored in Aurora MySQL-Compatible Edition. Transactional data written to Aurora is available to you in Amazon Redshift within seconds, without building and maintaining complex data pipelines.

Amazon DynamoDB

[Amazon DynamoDB](#) is a key-value and document database that delivers single-digit millisecond performance at any scale. It's a fully managed, multi-Region database with built-in security, backup and restore, and in-memory caching for internet-scale applications. DynamoDB can handle more than 10 trillion requests per day and support peaks of more than 20 million requests per second. Many of the world's fastest growing businesses such as Lyft, Airbnb, and Redfin, as well as enterprises such as Samsung, Toyota, and Capital One, depend on the scale and performance of DynamoDB to support their mission-critical workloads.

Hundreds of thousands of AWS customers have chosen DynamoDB as their key-value and document database for mobile, web, gaming, ad tech, Internet of Things (IoT), and other applications that need low-latency data access at any scale. Create a new table for your application and let DynamoDB handle the rest.

Amazon ElastiCache

[Amazon ElastiCache](#) is a web service that makes it easy to deploy, operate, and scale an in-memory cache in the cloud. The service improves the performance of web applications by allowing you to retrieve information from fast, managed, in-memory caches, instead of relying entirely on slower disk-based databases.

Amazon ElastiCache supports two open-source in-memory caching engines:

- [Redis](#) — a fast, open-source, in-memory key-value data store for use as a database, cache, message broker, and queue. [Amazon ElastiCache for Redis](#) is a Redis-compatible in-memory service that delivers the ease-of-use and power of Redis along with the availability, reliability, and performance suitable for the most demanding applications. Both single-node and up to 15- shard clusters are available, enabling scalability to up to 3.55 TiB of in-memory data. Amazon ElastiCache for Redis is fully managed, scalable, and secure. This makes it an ideal candidate to power high-performance use cases such as web, mobile apps, gaming, ad-tech, and IoT.
- [Memcached](#) — a widely adopted memory object caching system. [Amazon ElastiCache for Memcached](#) is protocol compliant with Memcached, so popular tools that you use today with existing Memcached environments will work seamlessly with the service.

Amazon Keyspaces (for Apache Cassandra)

[Amazon Keyspaces \(for Apache Cassandra\)](#) is a scalable, highly available, and managed Apache Cassandra-compatible database service. With Amazon Keyspaces, you can run your Cassandra workloads on AWS using the same Cassandra application code and developer tools that you use today.

You don't have to provision, patch, or manage servers, and you don't have to install, maintain, or operate software. Amazon Keyspaces is serverless, so you pay for only the resources you use and the service can automatically scale tables up and down in response to application traffic. You can

build applications that serve thousands of requests per second with virtually unlimited throughput and storage.

Data is encrypted by default and Amazon Keyspaces enables you to back up your table data continuously using point-in-time recovery. Amazon Keyspaces gives you the performance, elasticity, and enterprise features you need to operate business-critical Cassandra workloads at scale.

Amazon MemoryDB for Redis

[Amazon MemoryDB for Redis](#) is a Redis-compatible, durable, in-memory database service that delivers ultra-fast performance. It is purpose-built for modern applications with microservices architectures.

MemoryDB is compatible with Redis, a popular open-source data store, enabling customers to quickly build applications using the same flexible and friendly Redis data structures, APIs, and commands that they already use today. With MemoryDB, all of your data is stored in memory, which enables you to achieve microsecond read and single-digit millisecond write latency and high throughput. MemoryDB also stores data durably across multiple Availability Zones using a distributed transactional log to allow fast failover, database recovery, and node restarts.

Delivering both in-memory performance and multi-AZ durability, MemoryDB can be used as a high-performance primary database for your microservices applications eliminating the need to separately manage both a cache and durable database.

Amazon Neptune

[Amazon Neptune](#) is a fast, reliable, fully managed graph database service that makes it easy to build and run applications that work with highly connected datasets. The core of Amazon Neptune is a purpose-built, high-performance graph database engine optimized for storing billions of relationships and querying the graph with milliseconds latency. Amazon Neptune supports popular graph models Property Graph and W3C's RDF, and their respective query languages Apache TinkerPop Gremlin and SPARQL, allowing you to easily build queries that efficiently navigate highly connected datasets. Neptune powers graph use cases such as recommendation engines, fraud detection, knowledge graphs, drug discovery, and network security.

Amazon Neptune is highly available, with read replicas, point-in-time recovery, continuous backup to Amazon S3, and replication across Availability Zones. Neptune is secure with support

for encryption at rest. Neptune is fully managed, so you no longer need to worry about database management tasks such as hardware provisioning, software patching, setup, configuration, or backups.

Amazon Relational Database Service

[Amazon Relational Database Service](#) (Amazon RDS) makes it easy to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while automating time-consuming administration tasks such as hardware provisioning, database setup, patching and backups. It frees you to focus on your applications so you can give them the fast performance, high availability, security and compatibility they need.

Amazon RDS is available on several database instance types - optimized for memory, performance or I/O - and provides you with six familiar database engines to choose from, including [Amazon Aurora](#), [PostgreSQL](#), [MySQL](#), [MariaDB](#), [Oracle Database](#), and [Microsoft SQL Server](#). You can use the [AWS Database Migration Service](#) to easily migrate or replicate your existing databases to Amazon RDS.

Amazon RDS on VMware

[Amazon Relational Database Service](#) (Amazon RDS) on VMware lets you deploy managed databases in on-premises VMware environments using the Amazon RDS technology enjoyed by hundreds of thousands of AWS customers. Amazon RDS provides cost-efficient and resizable capacity while automating time-consuming administration tasks including hardware provisioning, database setup, patching, and backups, freeing you to focus on your applications. Amazon RDS on VMware brings these same benefits to your on-premises deployments, making it easy to set up, operate, and scale databases in VMware vSphere private data centers, or to migrate them to AWS.

Amazon RDS on VMware allows you to utilize the same simple interface for managing databases in on-premises VMware environments as you would use in AWS. You can easily replicate Amazon RDS on VMware databases to Amazon RDS instances in AWS, enabling low-cost hybrid deployments for disaster recovery, read replica bursting, and optional long-term backup retention in Amazon Simple Storage Service (Amazon S3).

Amazon Quantum Ledger Database (Amazon QLDB)

[Amazon QLDB](#) is a fully managed ledger database that provides a transparent, immutable, and cryptographically verifiable transaction log owned by a central trusted authority. Amazon QLDB tracks every application data change and maintains a complete and verifiable history of changes over time.

Ledgers are typically used to record a history of economic and financial activity in an organization. Many organizations build applications with ledger-like functionality because they want to maintain an accurate history of their applications' data, for example, tracking the history of credits and debits in banking transactions, verifying the data lineage of an insurance claim, or tracing movement of an item in a supply chain network. Ledger applications are often implemented using custom audit tables or audit trails created in relational databases.

However, building audit functionality with relational databases is time-consuming and prone to human error. It requires custom development, and since relational databases are not inherently immutable, any unintended changes to the data are hard to track and verify. Alternatively, blockchain frameworks, such as Hyperledger Fabric and Ethereum, can also be used as a ledger. However, this adds complexity as you need to set-up an entire blockchain network with multiple nodes, manage its infrastructure, and require the nodes to validate each transaction before it can be added to the ledger.

Amazon QLDB is a new class of database that eliminates the need to engage in the complex development effort of building your own ledger-like applications. With QLDB, your data's change history is immutable – it cannot be altered or deleted – and using cryptography, you can easily verify that there have been no unintended modifications to your application's data. QLDB uses an immutable transactional log, known as a journal, that tracks each application data change and maintains a complete and verifiable history of changes over time.

QLDB is easy to use because it provides developers with a familiar SQL-like API, a flexible document data model, and full support for transactions. QLDB is also serverless, so it automatically scales to support the demands of your application. There are no servers to manage and no read or write limits to configure. With QLDB, you only pay for what you use.

Amazon Timestream

[Amazon Timestream](#) is a fast, scalable, fully managed time series database service for IoT and operational applications that makes it easy to store and analyze trillions of events per day at 1/10th the cost of relational databases. Driven by the rise of IoT devices, IT systems, and smart industrial machines, time series data — data that measures how things change over time — is one of the fastest growing data types. Time-series data has specific characteristics such as typically arriving in time order form, data is append-only, and queries are always over a time interval. While relational databases can store this data, they are inefficient at processing this data as they lack optimizations such as storing and retrieving data by time intervals.

Timestream is a purpose-built time series database that efficiently stores and processes this data by time intervals. With Timestream, you can easily store and analyze log data for DevOps, sensor data for IoT applications, and industrial telemetry data for equipment maintenance. As your data grows over time, the Timestream adaptive query processing engine understands its location and format, making your data simpler and faster to analyze. Timestream also automates rollups, retention, tiering, and compression of data, so you can manage your data at the lowest possible cost. Timestream is serverless, so there are no servers to manage. It manages time-consuming tasks such as server provisioning, software patching, setup, configuration, or data retention and tiering, freeing you to focus on building your applications.

Amazon DocumentDB (with MongoDB compatibility)

[Amazon DocumentDB \(with MongoDB compatibility\)](#) is a fast, scalable, highly available, and fully managed document database service that supports MongoDB workloads. Amazon DocumentDB is designed from the ground-up to give you the performance, scalability, and availability you need when operating mission-critical MongoDB workloads at scale. Amazon DocumentDB implements the Apache 2.0 open-source MongoDB 3.6 and 4.0 APIs by emulating the responses that a MongoDB client expects from a MongoDB server, allowing you to use your existing MongoDB drivers and tools with Amazon DocumentDB (with MongoDB compatibility).

Amazon Lightsail Managed Databases

[Amazon Lightsail managed databases](#) are separate from compute workloads, so you can build applications and websites on Lightsail instances without interruption. Lightsail supports MySQL and PostgreSQL databases, and you can configure them for standard availability for regular workloads or high availability for critical workloads. Lightsail-managed databases bundle the underlying compute, SSD-based storage, and data transfer bandwidth into a fixed monthly price. You can manage your Lightsail-managed database by using the Lightsail console, the [AWS Command Line Interface](#) (AWS CLI), the Lightsail API, or an [AWS SDK](#).

Credit to: AWS Documentation