Databases

AVANTUS
Challenging The Future In IT

Welcome to the eighth lesson of the AWS Solutions Architect Associate level course—Databases.

By the end of the lesson you will be able to:

- List the different databases that can be operated in AWS
- Explain what RDS and its uses
- Describe how to launch databases from AMIs and the self-installation options
- Discuss what DynamoDB, RedShift, Aurora, and ElastiCache are used for
- Illustrate what DMS is and how it can help with database migrations to AWS
- Outline the AWS recommended best practices for databases
- Identify about the costs associated with databases

Databases Overview

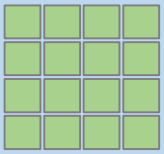
In this section you'll learn about the different databases that can be operated in AWS.

Relational Databases

The most common form of databases are relational databases or SQL databases. They are used to store structured data.

A relational database is a collection of data items organized as a set of formally described tables. It is also known as the relational model.

Structured Data



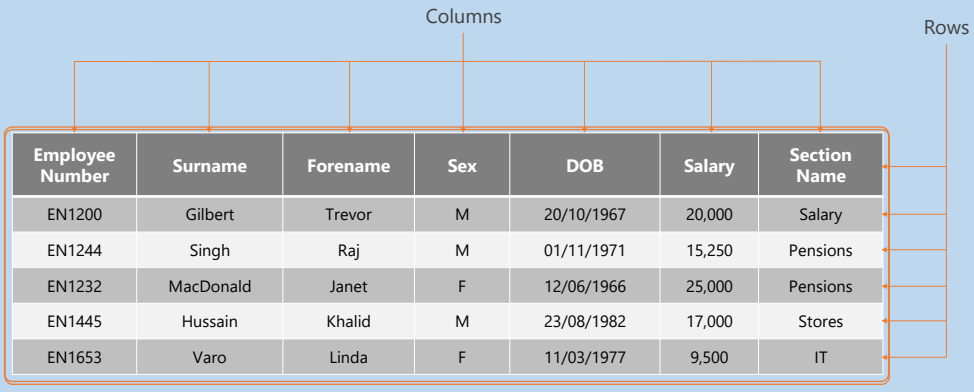
Relational Database

Unstructured Data

The most common form of databases are relational databases or SQL databases. It is used to store structured data.
It is a collection of data items organized as a set of formally described tables. It is also known as the relational model.

Relational Databases (Contd.)

Relational databases store data in rows and columns.



Employee Number	Surname	Forename	Sex	DOB	Salary	Section Name
EN1200	Gilbert	Trevor	M	20/10/1967	20,000	Salary
EN1244	Singh	Raj	M	01/11/1971	15,250	Pensions
EN1232	MacDonald	Janet	F	12/06/1966	25,000	Pensions
EN1445	Hussain	Khalid	M	23/08/1982	17,000	Stores
EN1653	Varo	Linda	F	11/03/1977	9,500	IT

Relational databases store data in rows and columns.
If you think of the Excel spread sheet as a database and a worksheet as a table, then you have the rows and columns, which actually store the data. Each column in a relational database stores the same kind of data type, which is text or integers.
Relational databases are accessed using structured query language SQL and are also referred to as an SQL database.

OLTP

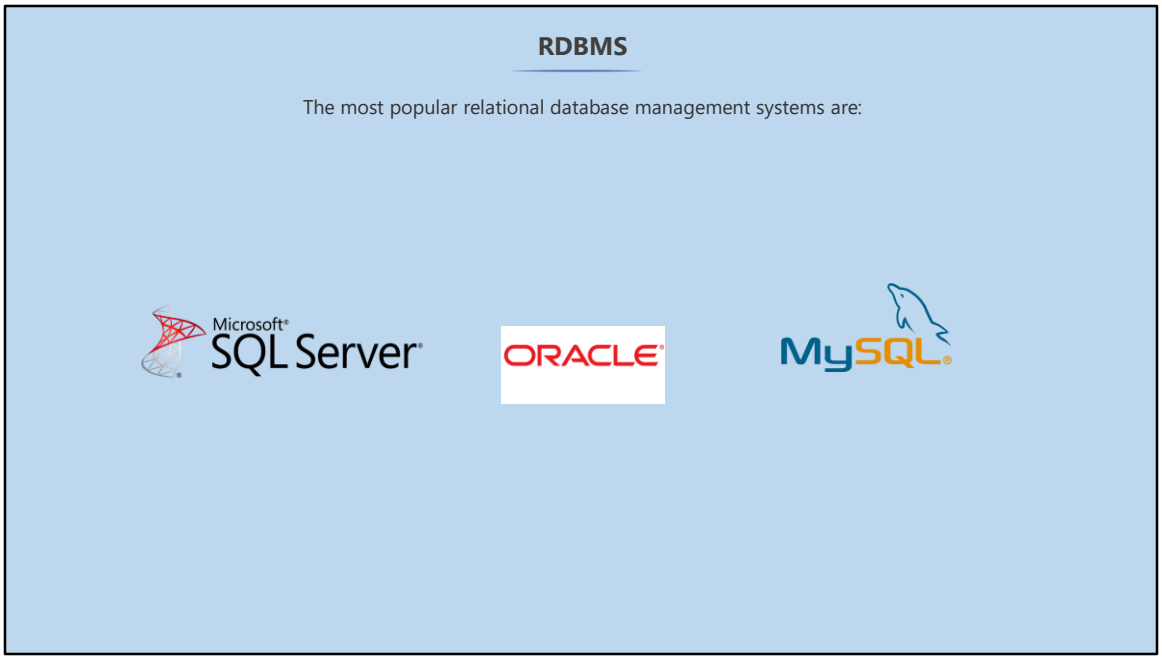
Relational databases are mainly used for OLTP (Online Transaction Processing), which involves business transactional processing using real time data.

An example of an OLTP database is a database that stores the transactions for an ATM.

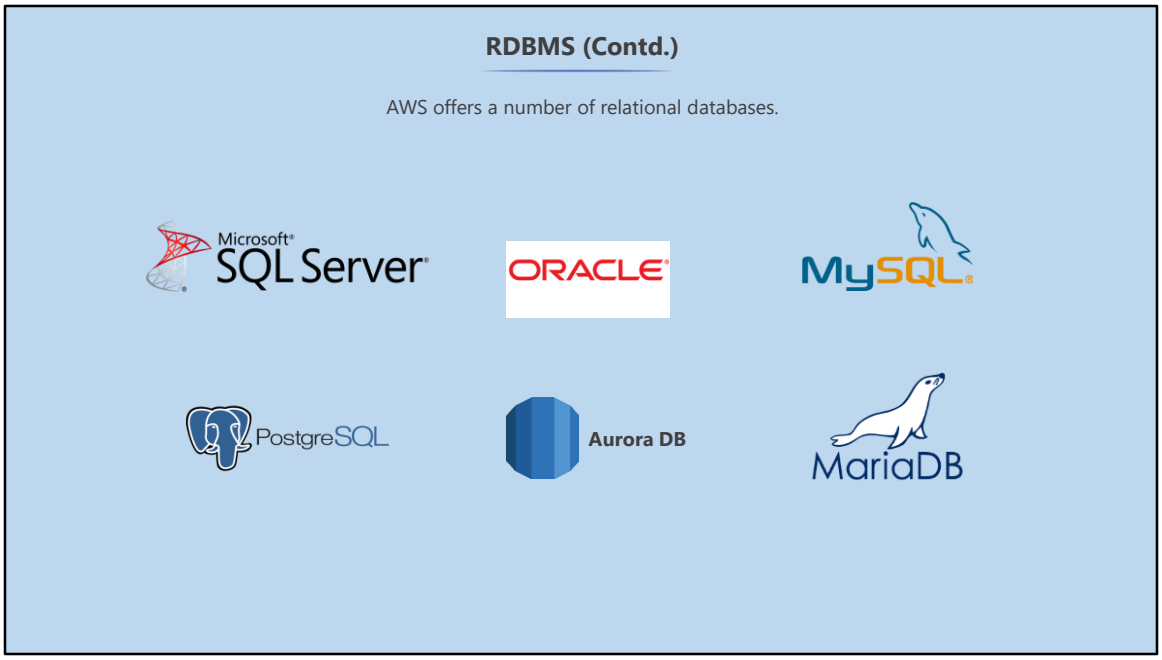
ATM

Name	Time	Amount	Account No:
David	11.30 AM	500	12345678

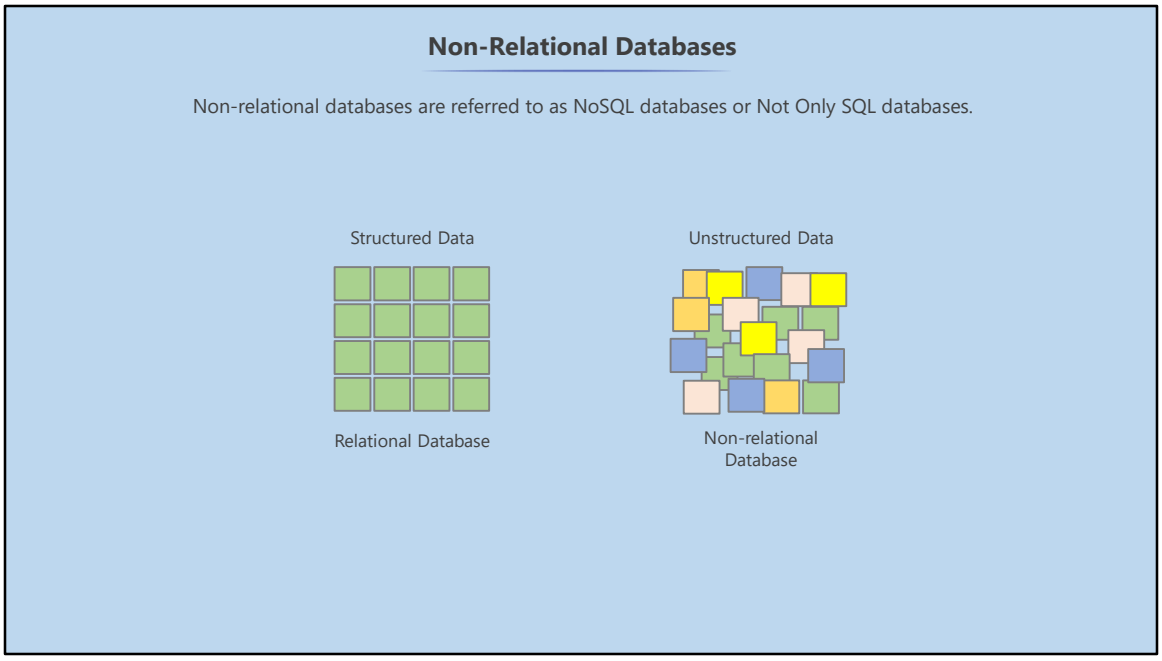
Relational databases are mainly used for Online Transaction Processing (OLTP). It involves business transactional processing using real-time or up-to-date data.
An example of an OLTP database is a database that stores the transactions for an ATM. Every time money is withdrawn, the database is updated with the details of who withdrew the money, at what time, how much they withdrew, the account number, and so on.



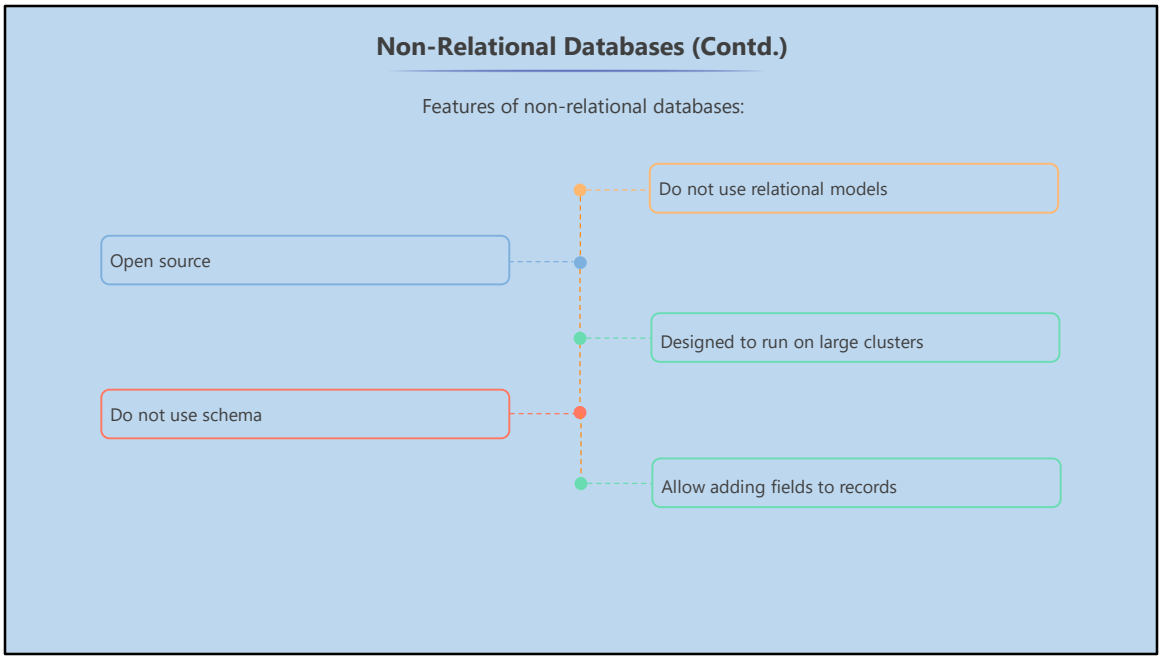
The most popular types of relational database management systems or RDBMSs are SQL Server, Oracle, and MySQL.



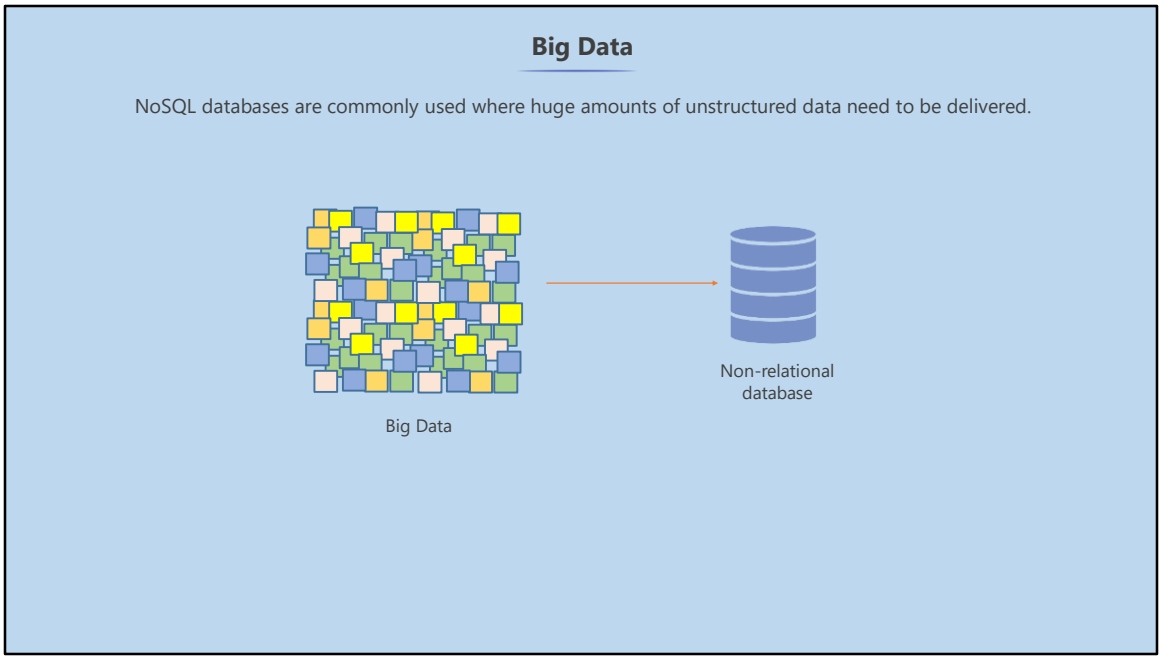
AWS offers the following RDBMSs: SQL Server, Oracle, MySQL, PostgreSQL, Aurora, and MariaDB.



Non-relational databases are also known as no SQL databases or Not Only SQL databases. It is a relatively new kind of database that can store large amounts of unstructured data. SQL databases are not designed to manage such data.



NoSQL databases do not use the relational model. They are designed to run on large clusters and do not have schema. Instead, you can add fields easily to any record whenever required.
There are different types of NoSQL databases. Some use JSON or key-value pairs. They can be column based, document based, graph based, and so on.

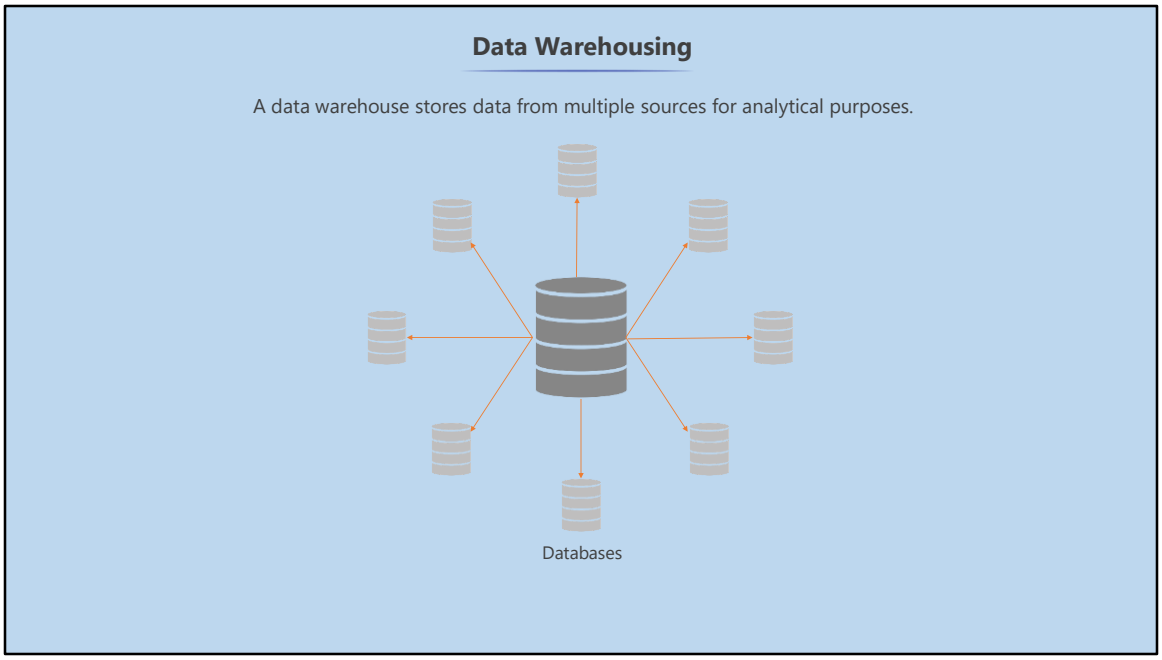


NoSQL databases are commonly used where huge amounts of data needs to be delivered. For example, Twitter has to store terabytes of data every day. Writing all these information to a single relational SQL database would be a slow process. NoSQL allows data to be written to a distributed cluster. It has a flexible schema, so it can be row-based, column-based, document oriented, and so on.

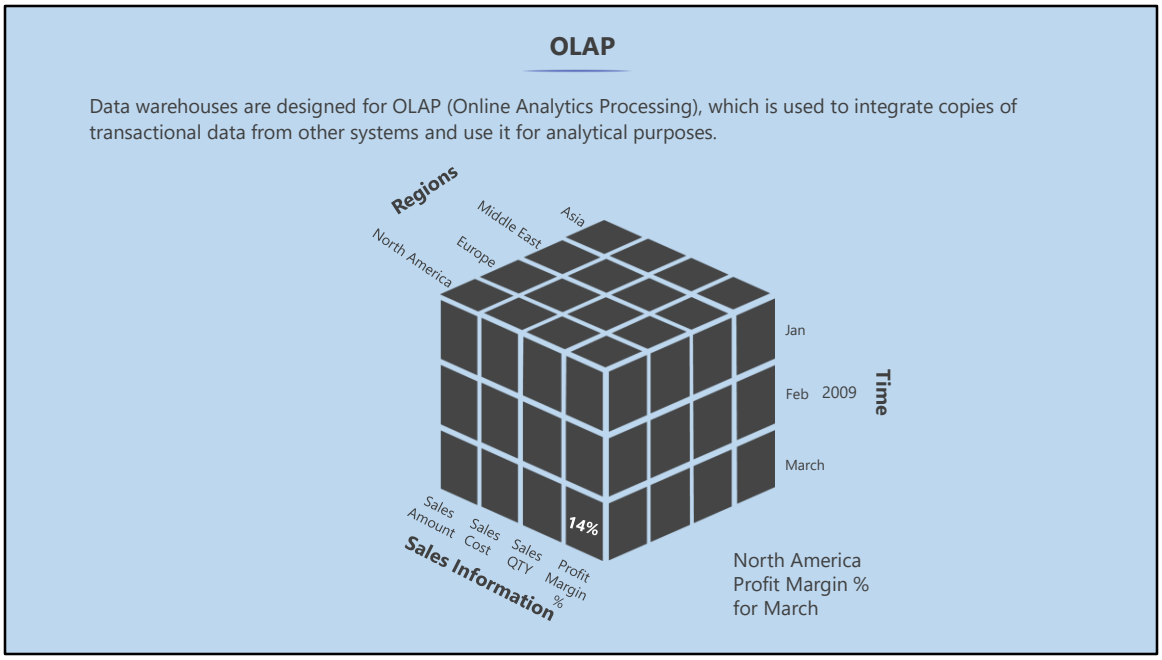
NoSQL Platforms

The most popular NoSQL databases are Cassandra, MongoDB, and CouchDB.

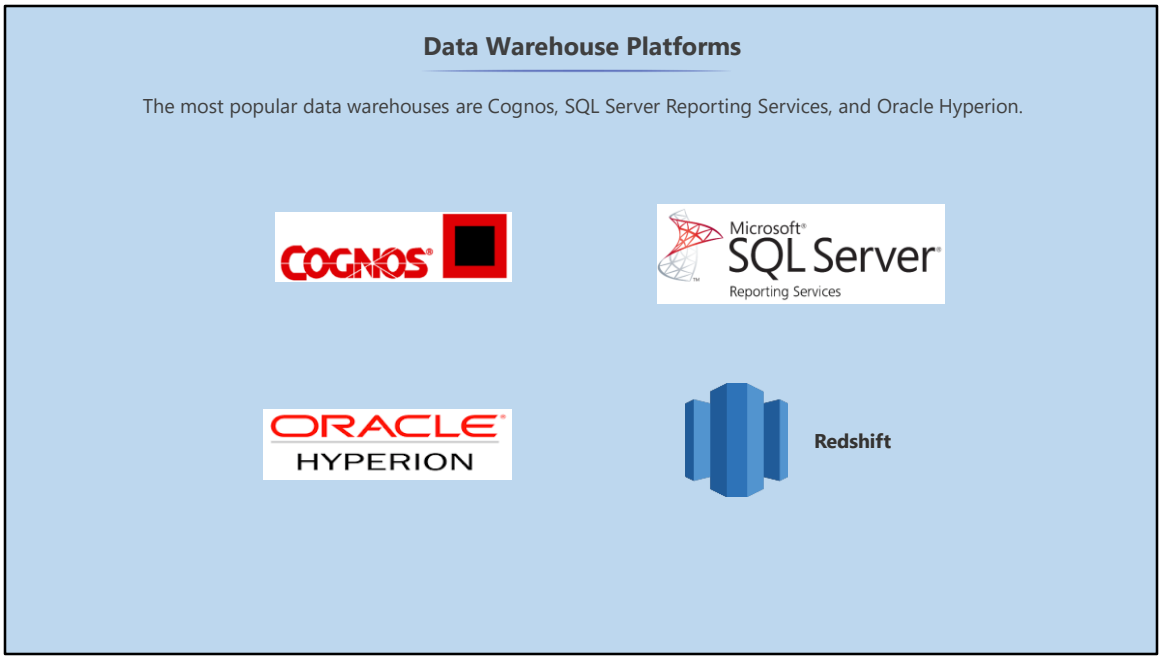
The most popular NoSQL databases are Cassandra, MongoDB, and CouchDB. There are many more as shown in the diagram. AWS offers its own NoSQL database called DynamoDB.



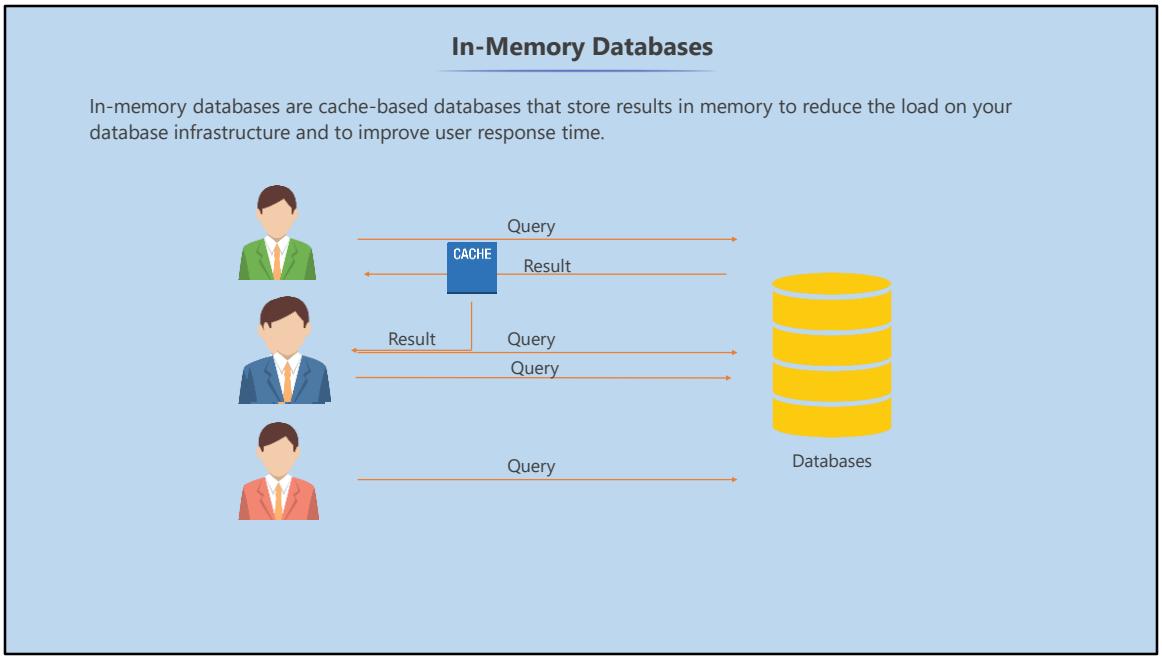
A data warehouse stores data from multiple sources for analytical purposes.



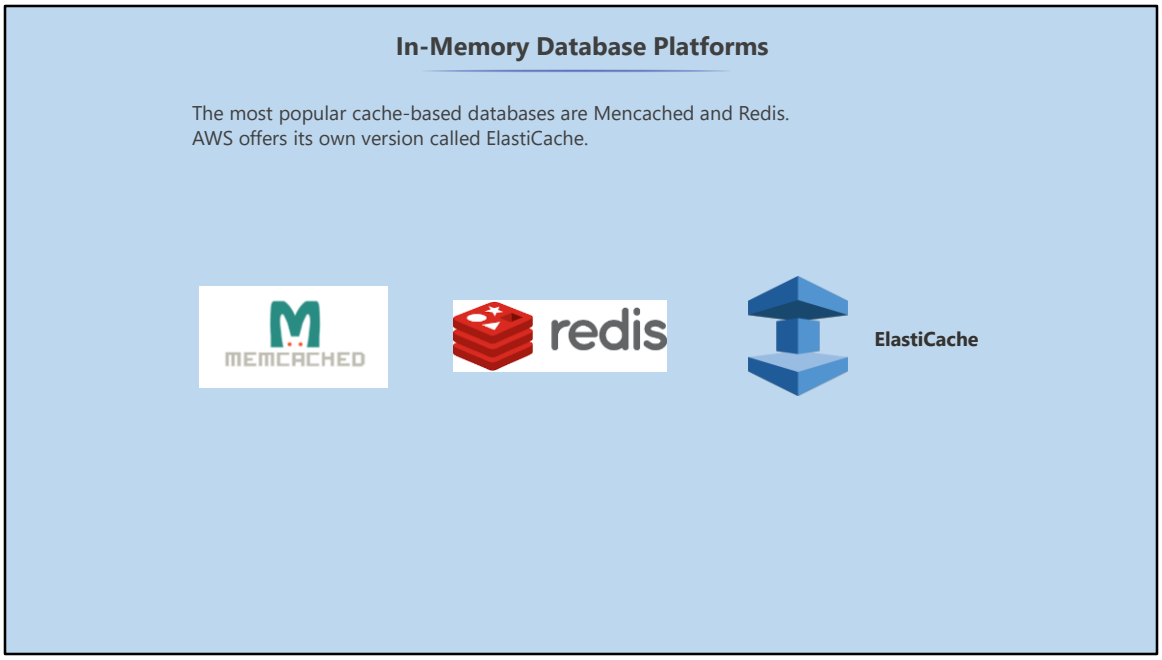
Data warehouses are designed for OLAP (Online Analytics Processing), which is used to integrate copies of transactional data from other systems and use it for analytical purposes. It is mainly used for business intelligence and has the ability to answer complex queries like how much profit a particular product made globally in the month of March.



The most popular data warehouses are Cognos, SQL Server Reporting Services, and Oracle Hyperion. AWS offers its own data warehouse called Amazon Redshift.



In-memory databases are cache-based databases. They store results in memory to reduce the load on the database infrastructure and to improve user response time. Rather than multiple users running the same queries and having to access the database every time to get the results, in In-memory databases when one person runs the query, the result gets stored in cache. When the next user runs the same query the result gets pulled straight from the cache, saving user response time.



The most popular cache-based databases are Memcached and Redis. AWS offers its own version called Amazon ElastiCache which is compatible with both Memcached and Redis.

Knowledge Check

1 **Which are the four types of database platforms in AWS?**

Relational, Non-relational, Data warehouse, and In-memory

Structured, Non structured, Data storehouse, and Cached

Relational, Non-relational, Data storehouse, and Cached

Structured, Non structured, Data Warehouse, and In-memory

1 **Which are the four types of database platforms in AWS?**

Relational, Non-relational, Data warehouse, and In-memory

Structured, Non structured, Data storehouse, and Cached

Relational, Non-relational, Data storehouse, and Cached

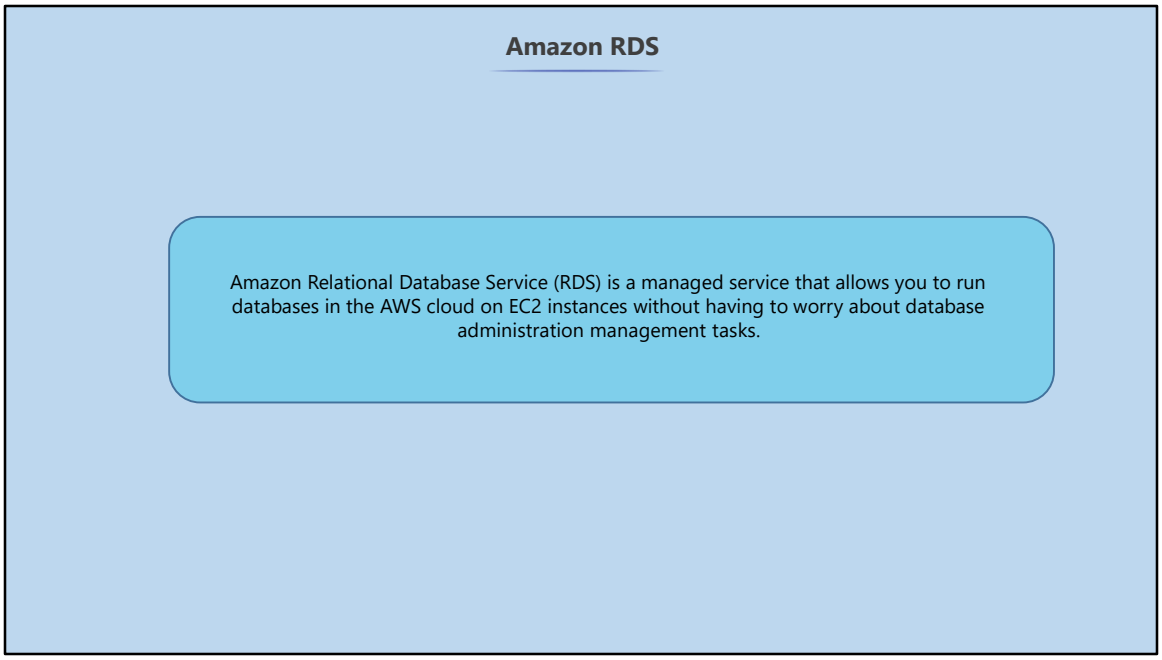
Structured, Non structured, Data Warehouse, and In-memory

a

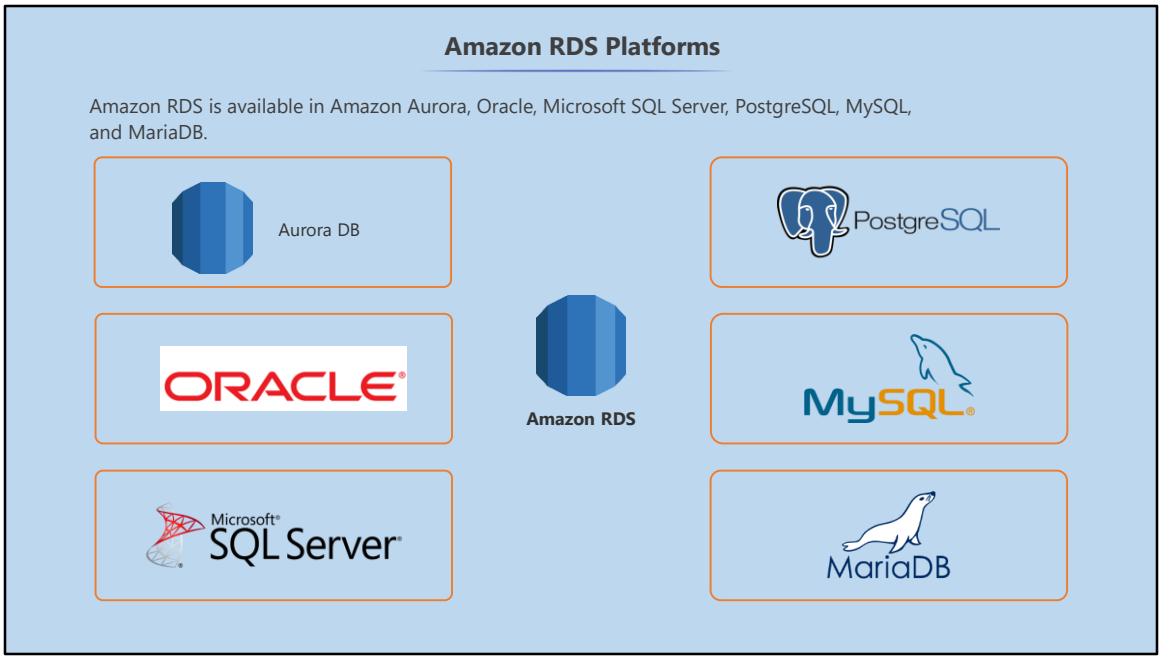
The four types of database platforms available on AWS are Relational, Non-relational, Data warehouse, and In-memory.

Amazon Relational Database Service (RDS)

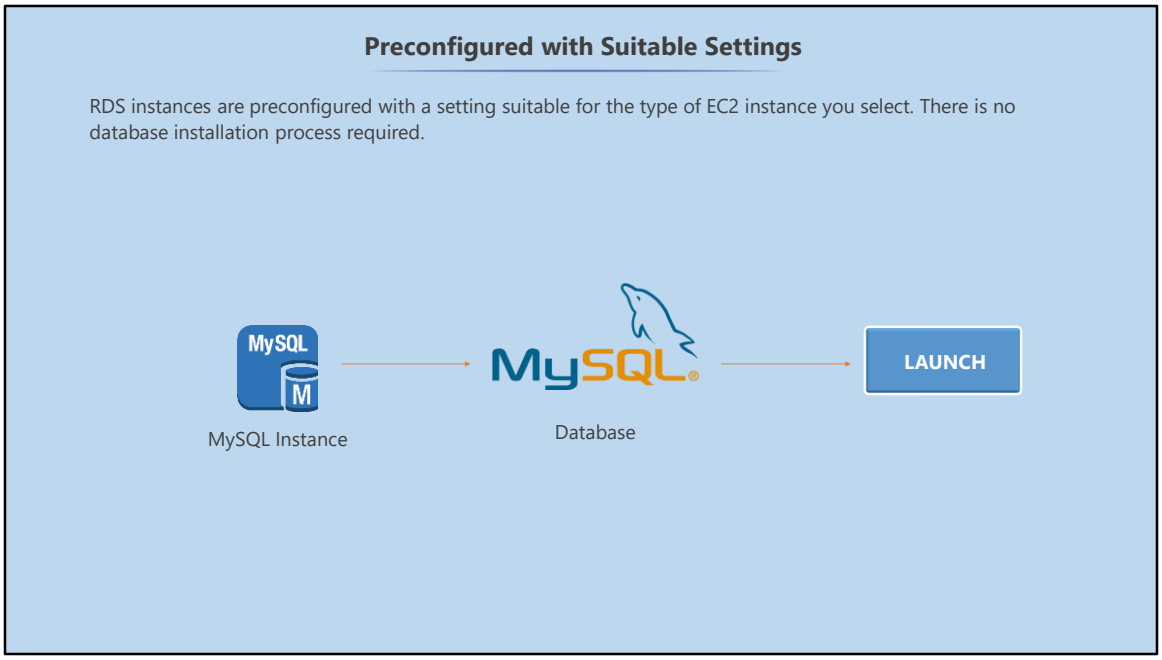
In this section you'll learn about Amazon Relational Database Service (RDS).



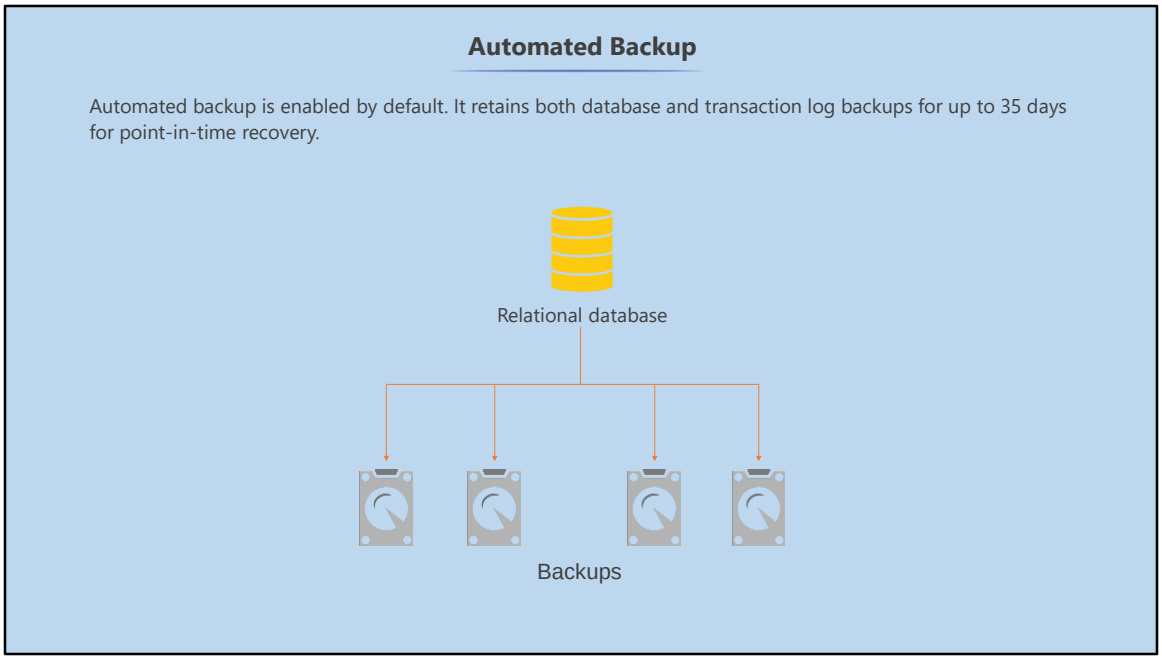
Amazon Relational Database Service is a managed service that allows you to run databases in the Cloud on EC2 instances, without worrying about database administration management tasks.



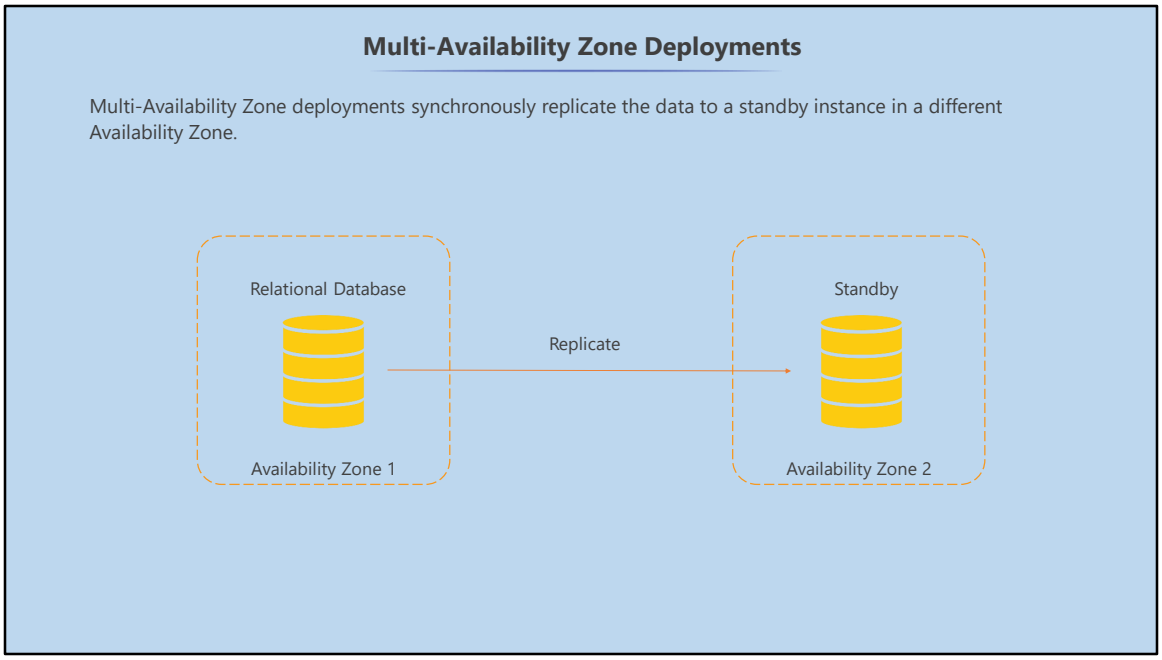
Amazon RDS is available in Amazon Aurora, Oracle, Microsoft SQL Server, PostgreSQL, MySQL, and MariaDB.



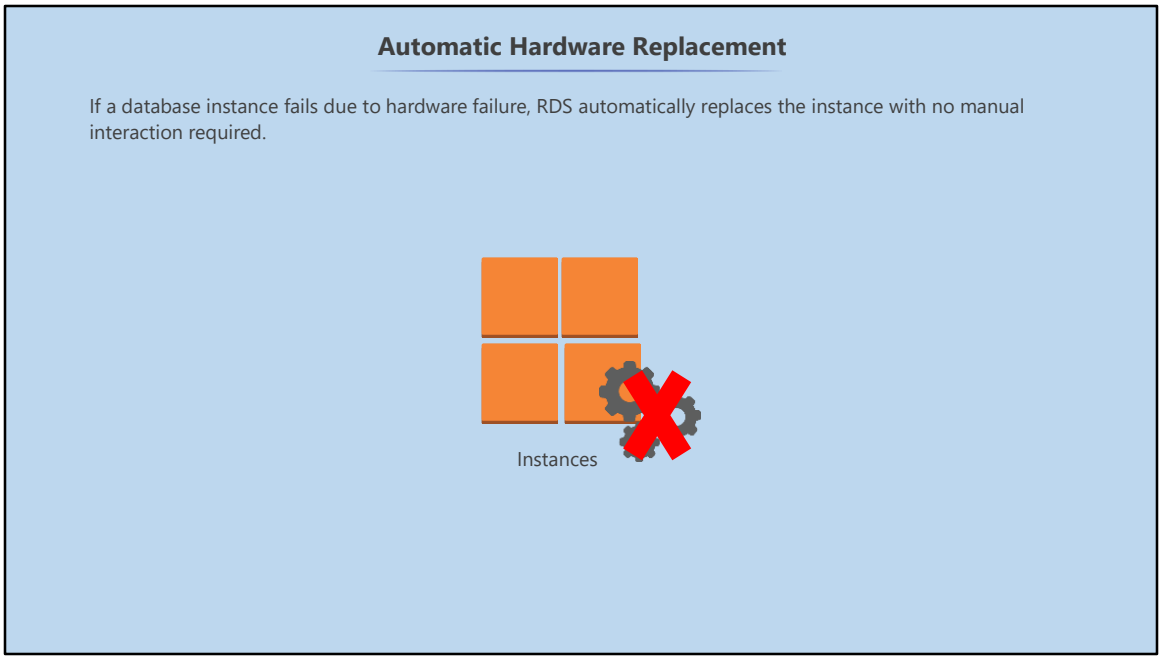
RDS instances are preconfigured with a setting suitable for the type of EC2 instance you select. There is no database installation process required— you just choose the instance you want, the type of database you require, click launch, and you have an up and running database.



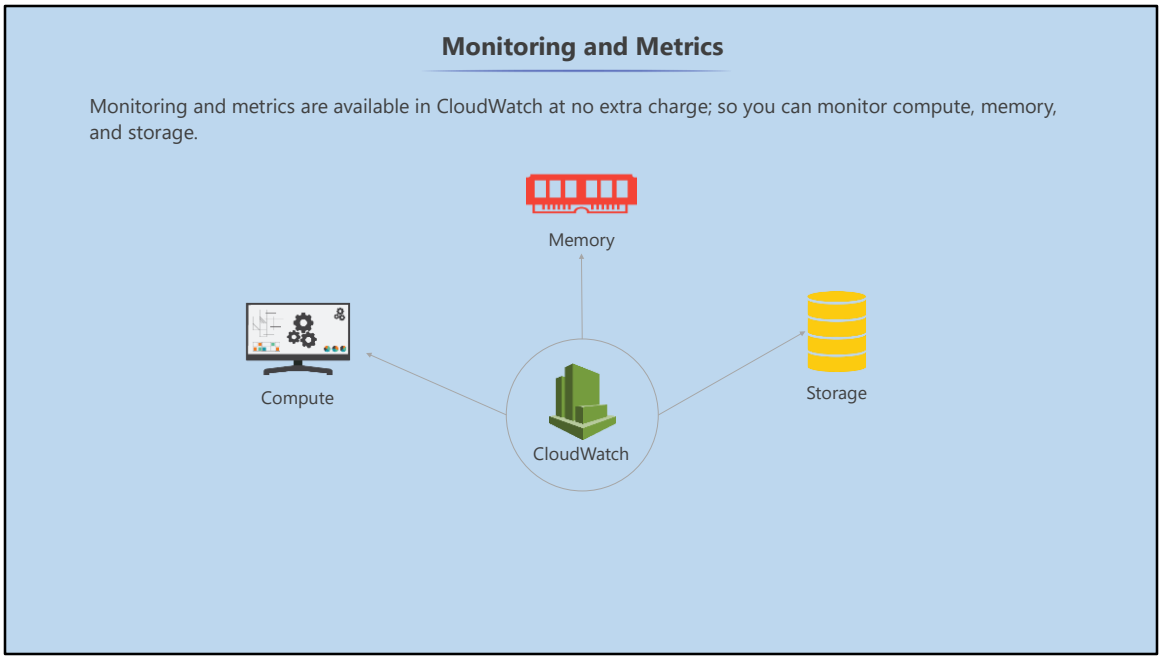
Automated backups are enabled by default. It retains both database and transaction log backups for up to 35 days for point-in-time recovery.



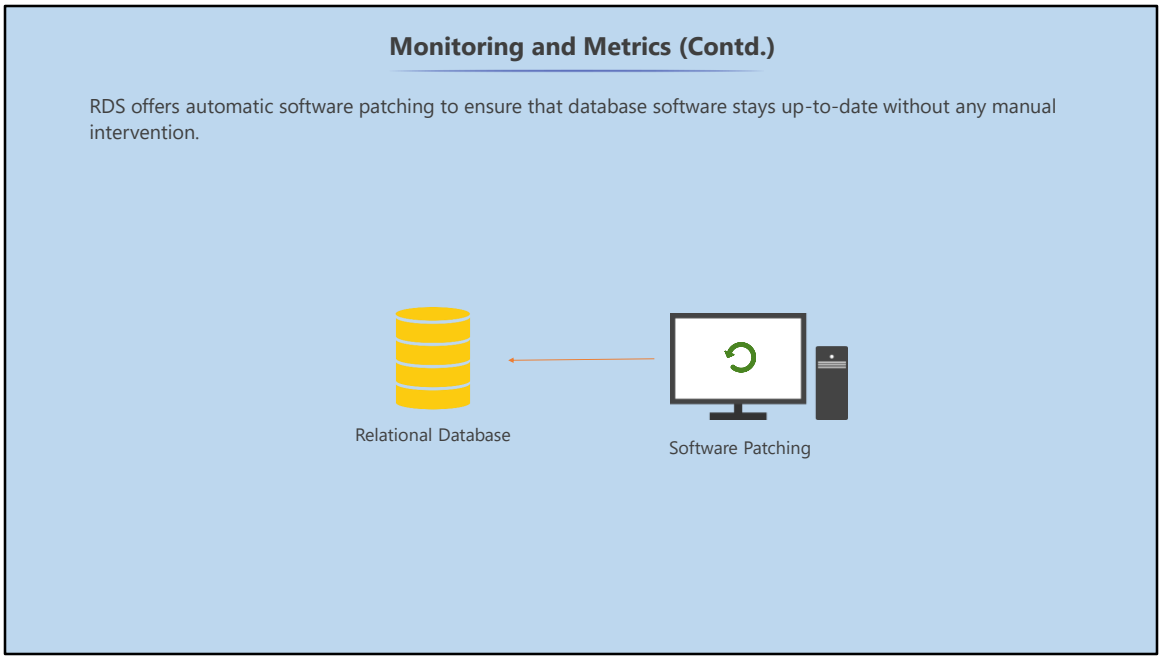
RDS provides high-availability. Multi-availability zone deployments synchronously replicate the data to a standby instance in a different availability zone. There is no need to worry about disaster recovery. It is built-in into RDS.



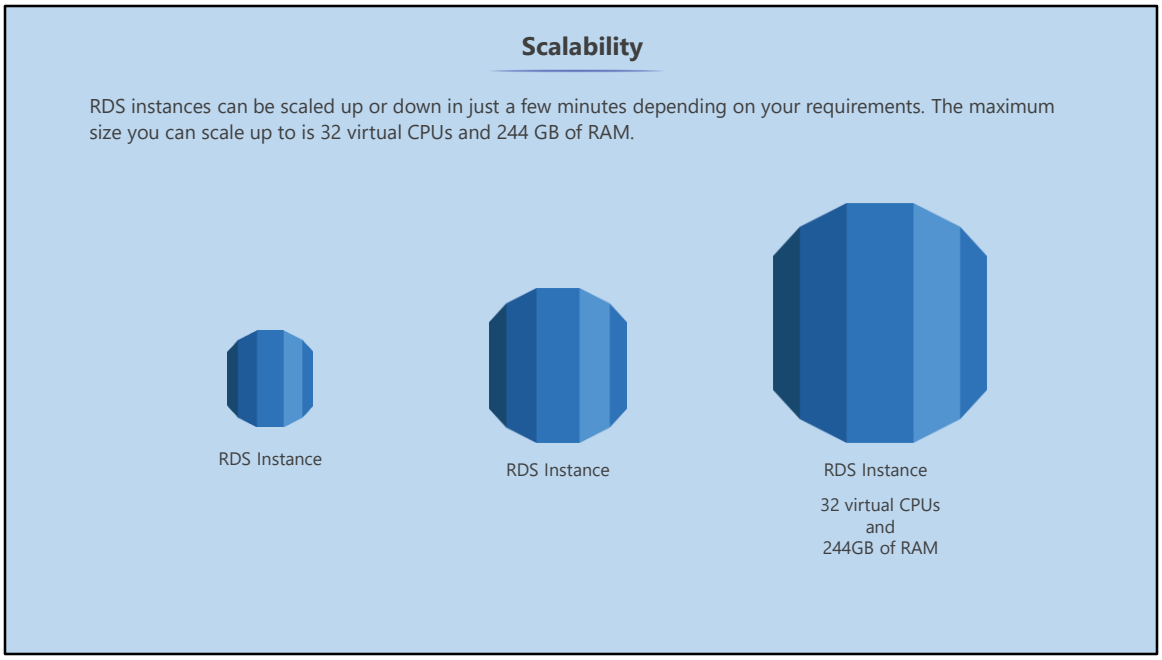
If a database instance fails due to hardware failure, RDS automatically replaces the instance with no manual interaction required.



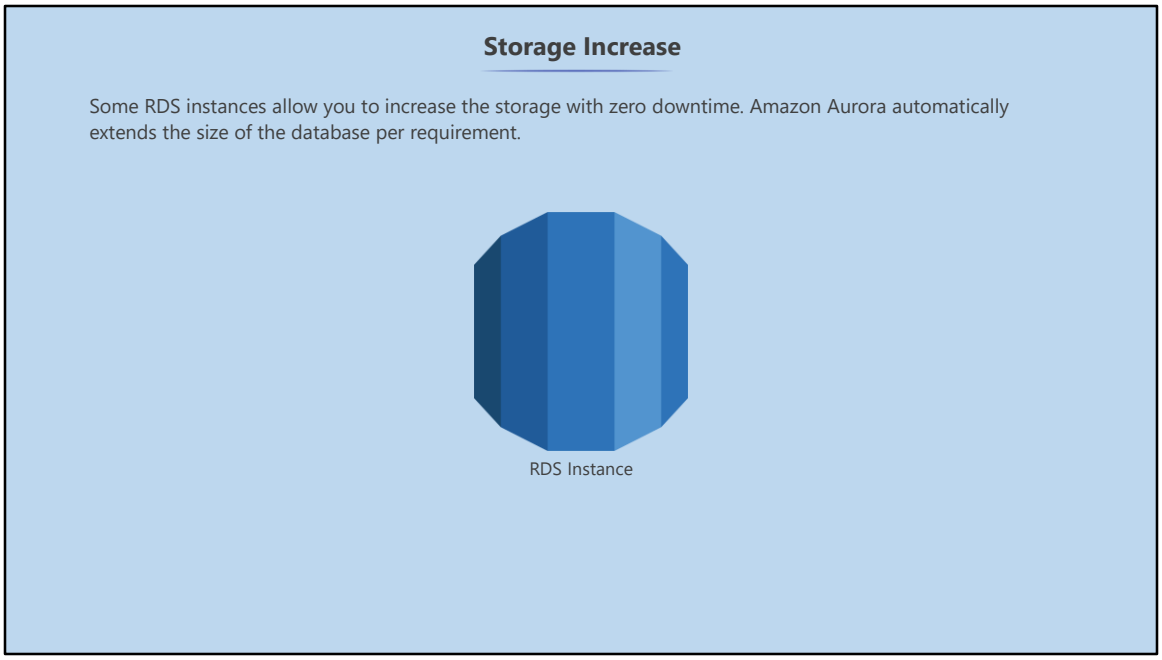
You can avail the benefits of monitoring and metrics using CloudWatch at no extra charge. You can monitor compute memory and storage capacities to make sure that you are running your RDS instance on the correct EC2 instance size.



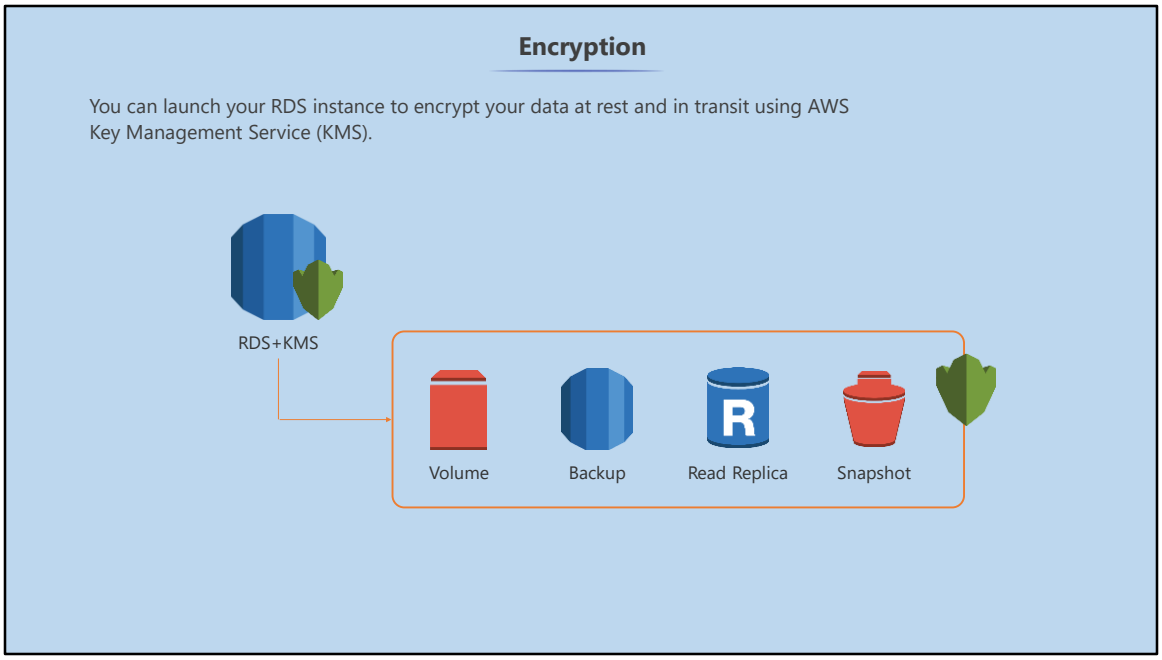
RDS offers automatic software patching to ensure that database software stays up-to-date without any manual intervention.



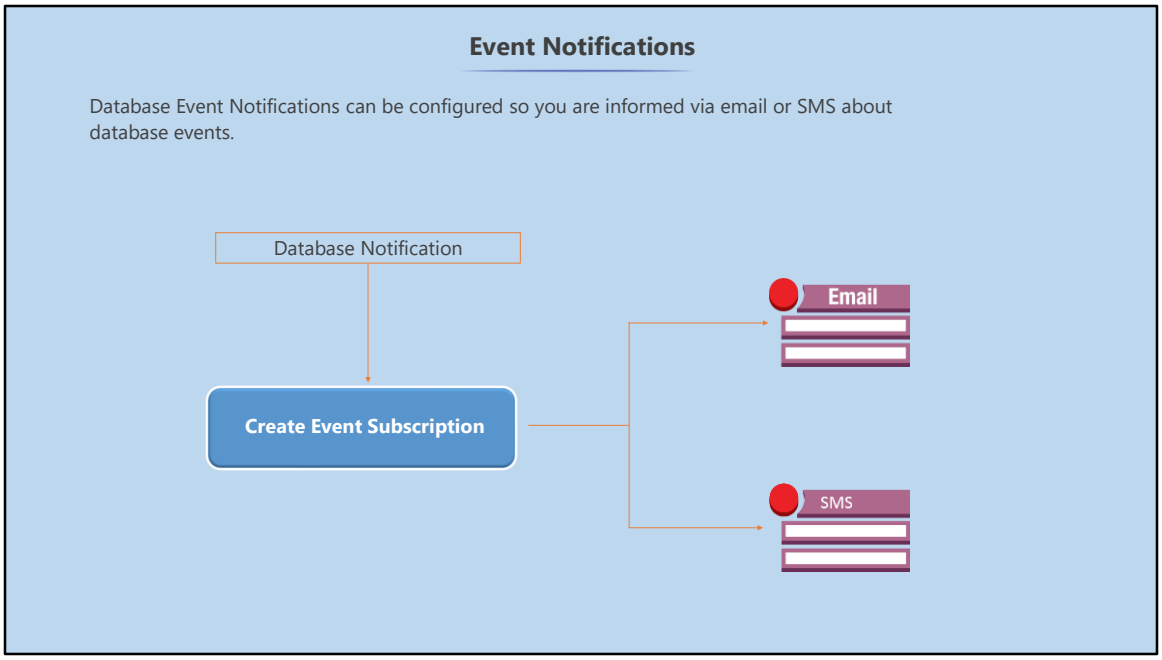
RDS instances can be scaled up or down in just a few minutes depending on your requirements. The maximum size you can scale up to is 32 virtual CPUs and 244 GB of RAM.



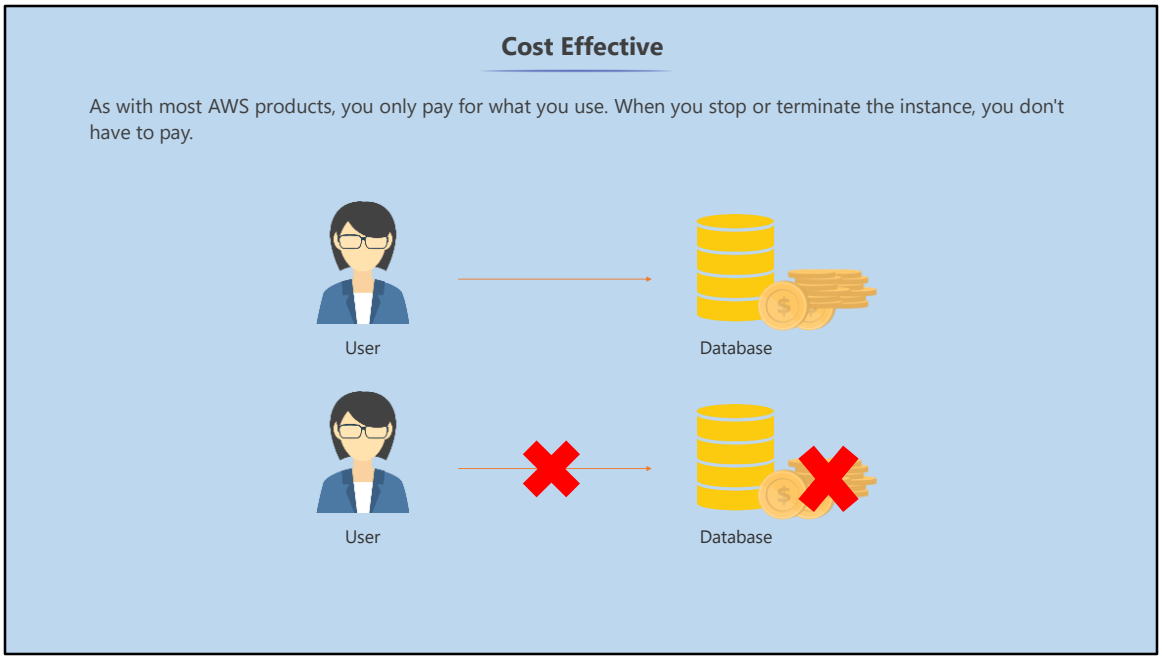
Some RDS instances allow you to increase the storage with zero downtime. Amazon Aurora automatically extends the size of the database per requirement.



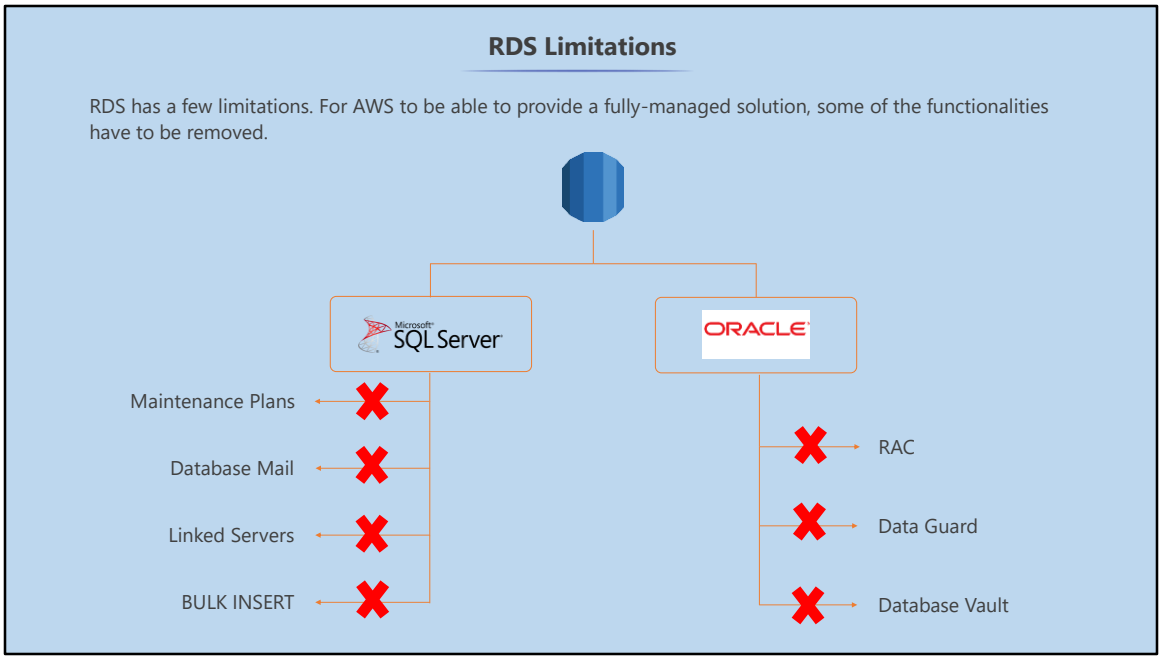
You can launch your RDS instance to encrypt your data at rest and in transit using AWS Key Management Service (KMS). Database volumes, backups, read replicas, and snapshots are all encrypted as a result.



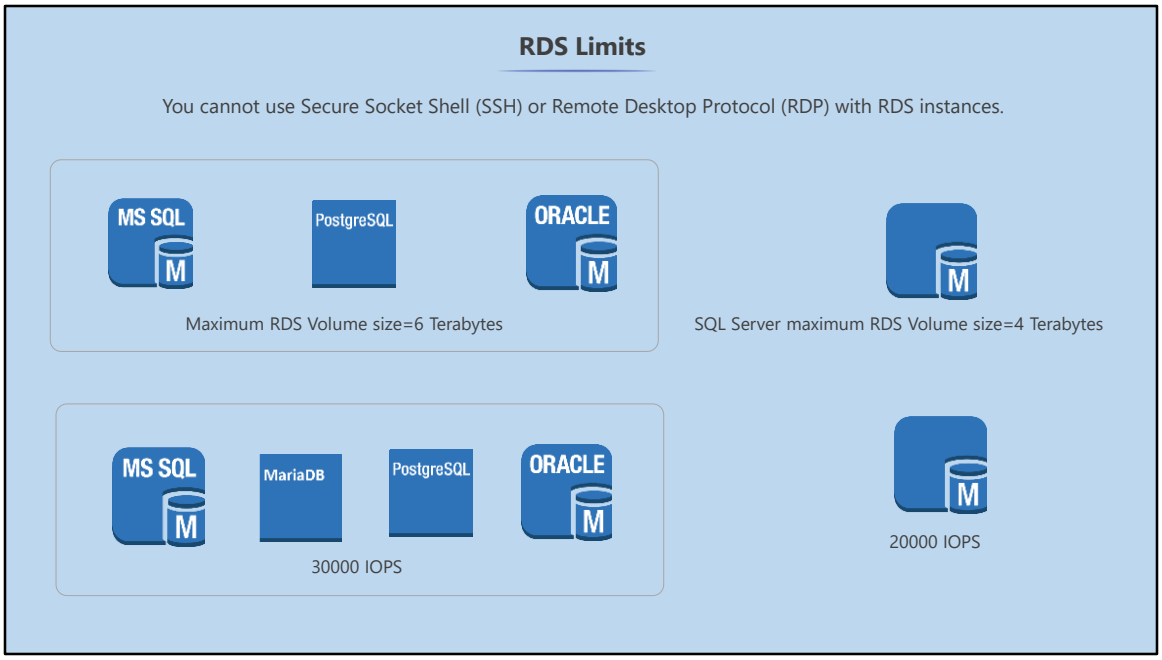
Database Event Notifications can be configured so you are informed via email or SMS about database events.



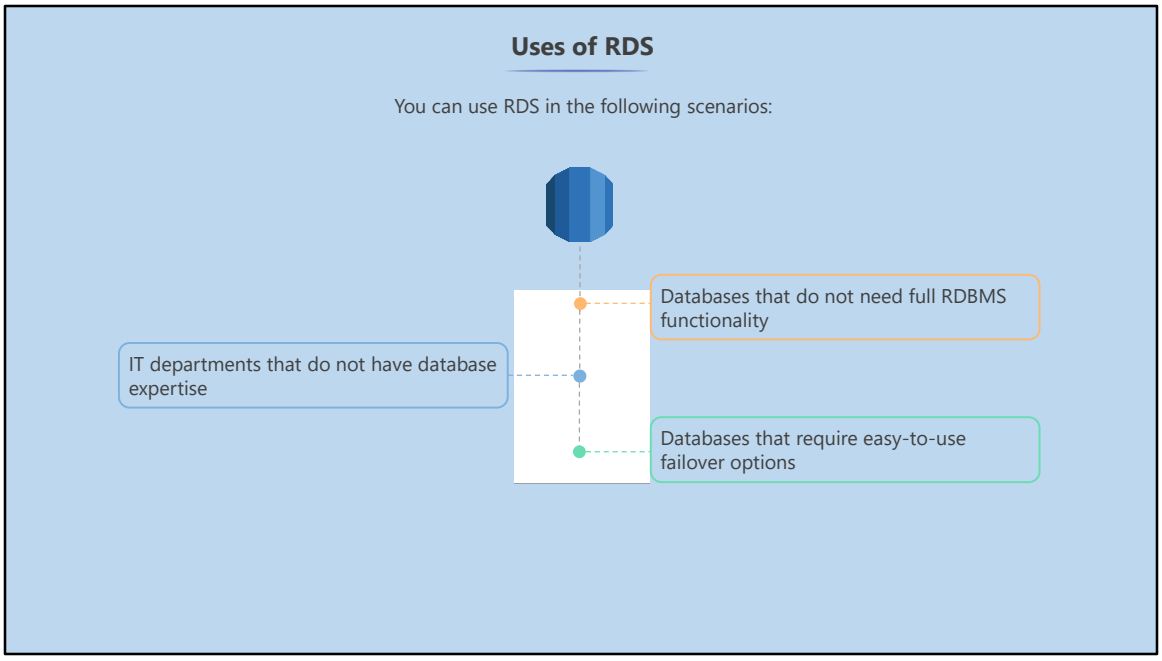
As with most AWS products, you only pay for what you use. When you stop or terminate the instance, you don't have to pay. When the RDS instance is running you are charged an hourly rate for the EC2 instance and the database license cost.



RDS has a few limitations . For AWS to be able to provide a fully managed solution, some of the functionalities have to be removed.
For example with SQL Server you cannot run maintenance plans, database mail, linked service, or BULK INSERT.
With Oracle you can't run RAC, Data Guard, or Database Vault.



The maximum RDS volume size is 6 TB for MySQL, PostgreSQL, and Oracle and it is 4 TB for SQL Server. The limit of RDS databases size for SQL Server Express is 300 GB.
If you are running MySQL, MariaDB, PostgreSQL, or Oracle, you have a limit of 30,000 IOPS for your storage and SQL Server has a limit of 20,000 IOPS.
You cannot use Secure Socket Shell (SSH) or Remote Desktop Protocol (RDP) with RDS instances.



In the following situations RDS is the perfect solution:

- When your requirement is just a database repository and you do not require any additional functionality
- If you don't have database expertise or you have no one in your team who is a database administrator
- If you just want to launch a database with failover capabilities without having to worry about administration or configuration tasks

Demo: Amazon RDS

In this demonstration, you'll learn how to launch an Amazon RDS instance.

Knowledge Check

1 **Which of the following platforms are available on RDS?**

Amazon Aurora, SAP ASE, DynamoDB, RedShift, MySQL, and MariaDB

Amazon Aurora, Oracle, Microsoft SQL Server, PostgreSQL, MySQL, and MariaDB

Amazon Aurora, Oracle, Microsoft SQL Server, RedShift, MySQL, and ElastiCache

Amazon Aurora, Oracle, DynamoDB, PostgreSQL, MySQL, and ElastiCache

1 **Which of the following platforms are available on RDS?**

Amazon Aurora, SAP ASE, DynamoDB, RedShift, MySQL, and MariaDB

Amazon Aurora, Oracle, Microsoft SQL Server, PostgreSQL, MySQL, and MariaDB

Amazon Aurora, Oracle, Microsoft SQL Server, RedShift, MySQL, and ElastiCache

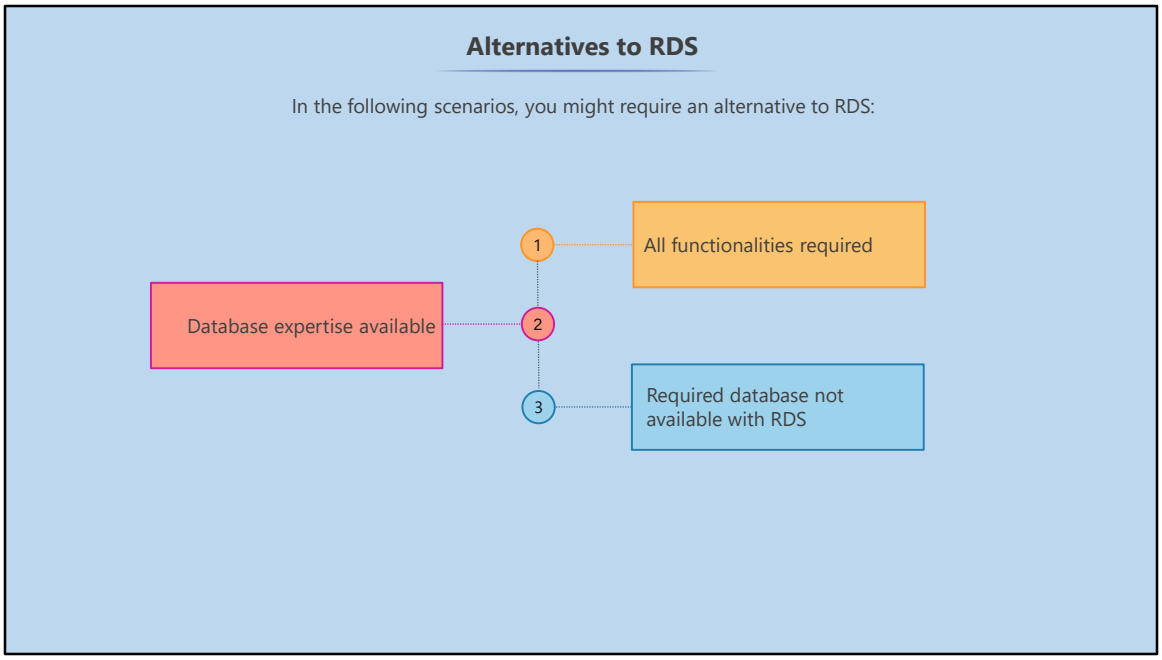
Amazon Aurora, Oracle, DynamoDB, PostgreSQL, MySQL, and ElastiCache

b

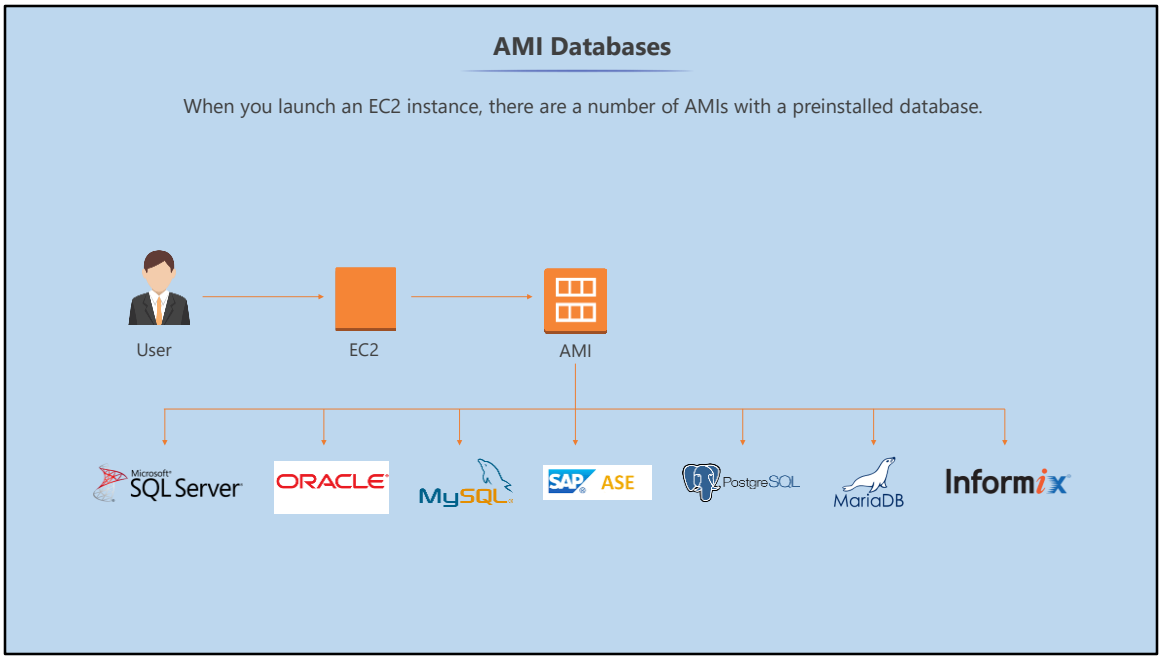
Amazon RDS is available for Amazon Aurora, Oracle, Microsoft SQL Server, PostgreSQL, MySQL, and MariaDB.

AMI Databases

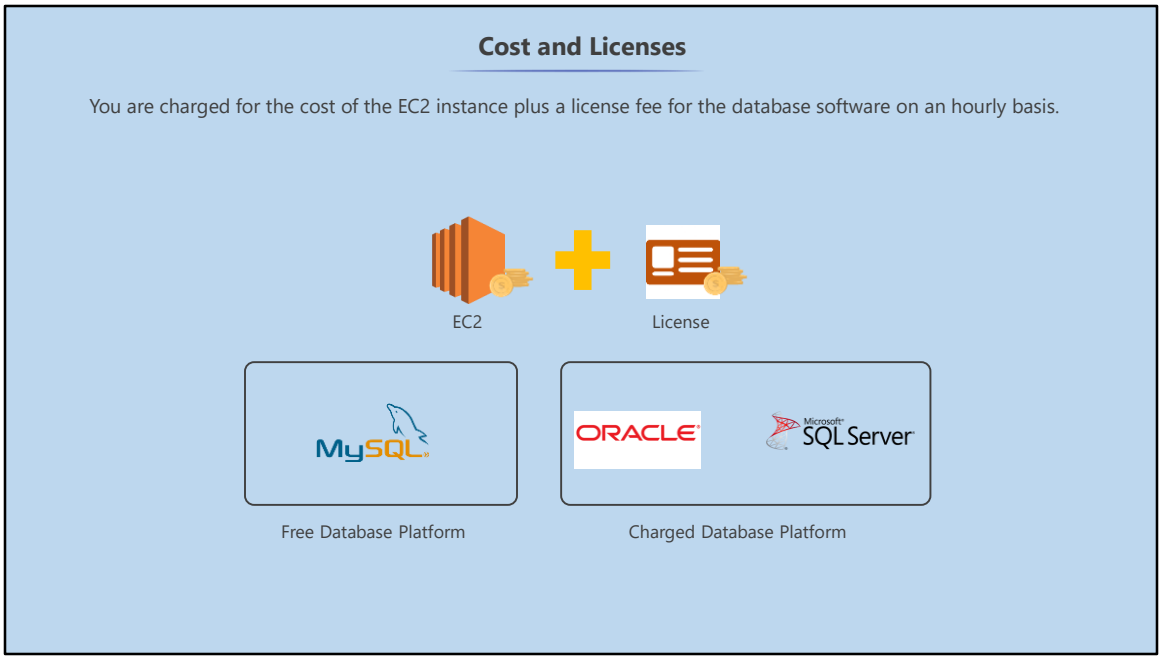
In this section, you'll learn about AMI databases and how you can launch databases from Amazon Machine Images (AMI).



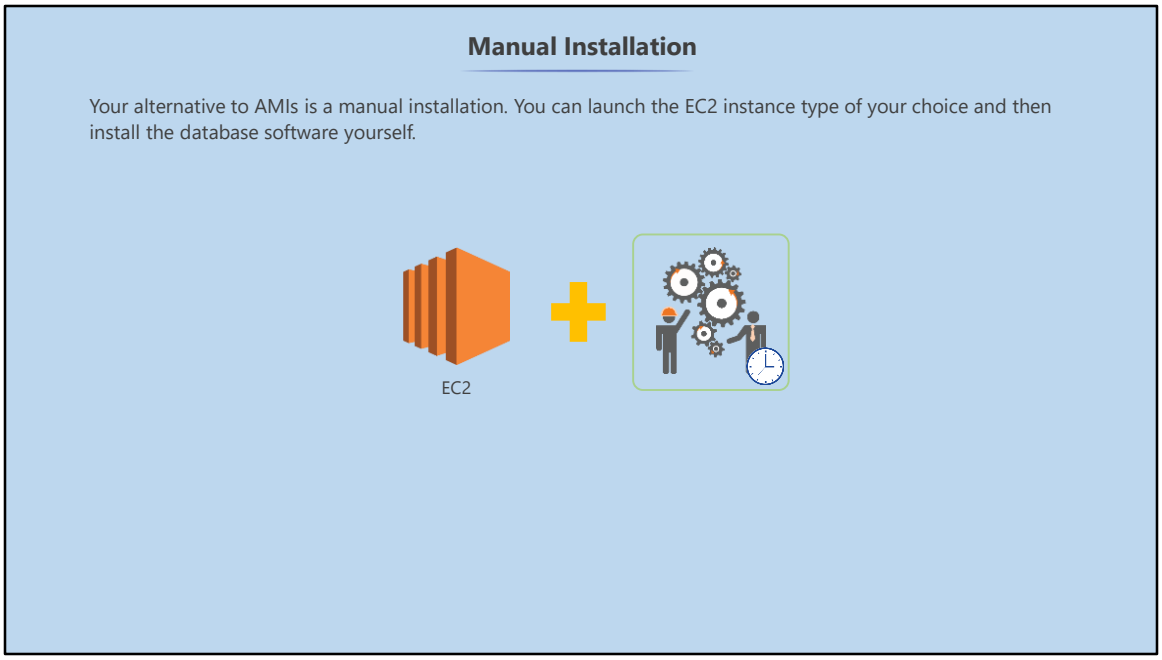
There can be situations where you need all the functionalities of SQL Server and Oracle. Or you have the database expertise to manage and administer the environment. Or perhaps the database you want to install isn't available with RDS. In such scenarios, you can either do AMI installations or manual installations.



When you launch an EC2 instance, there are a number of AMIs with preinstalled database. At Quick Start and AWS marketplace you get options such as, SQL Server, Oracle, MySQL, SAP ASC and many more AMIs that you can quickly and easily start without worrying about installing the database software.
Once the AMI is launched, you can login to find the database platform already installed and ready for you to configure for your specifications.



You are charged for the cost of the EC2 instance plus a license fee for the database software based on an hourly basis. Some database platforms are free, for example, MySQL. But with Oracle and SQL Server you have to pay the license fee.



Your alternative to AMIs is a manual installation. You can launch the EC2 instance type of your choice and then install the database software yourself. This offers the most flexibility but requires the expertise and also takes longer.
When the EC2 instance with the database is installed and configured, you can take your own AMI of it for future deployments.

Demo: Launching databases from an AMI

In this demonstration, you'll learn how to launch a database instance using an AMI.

Knowledge Check

1 **Why would you choose NOT to use RDS?**

To avoid the 12-36 month commitment AWS demands

RDS performance is significantly slower than installing your own RDBMS

Certain RDBMS functionalities are not available on RDS

RDS instances take too long to launch

1

Why would you choose NOT to use RDS?

To avoid the 12-36 month commitment AWS demands

RDS performance is significantly slower than installing your own RDBMS

Certain RDBMS functionalities are not available on RDS

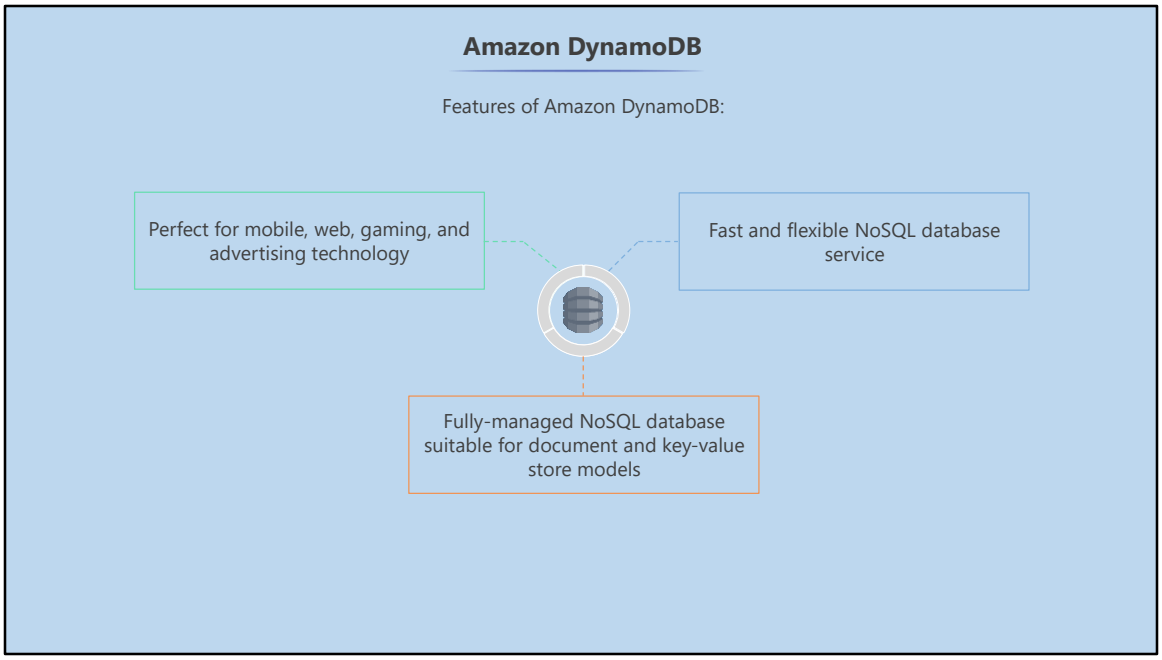
RDS instances take too long to launch

c

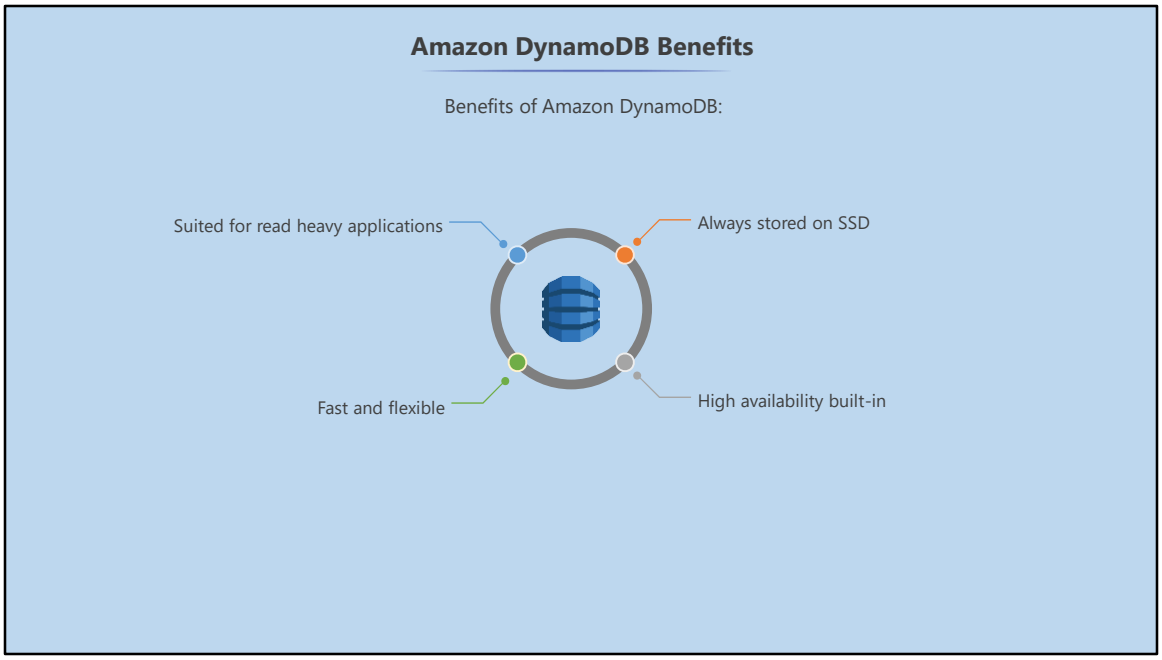
RDS has to remove certain functionalities from the RDBMS to provide a fully managed solution. If you need a particular functionality that is removed, then RDS is not suitable for you.

Amazon DynamoDB

In this section, you'll learn about Amazon DynamoDB and its features.

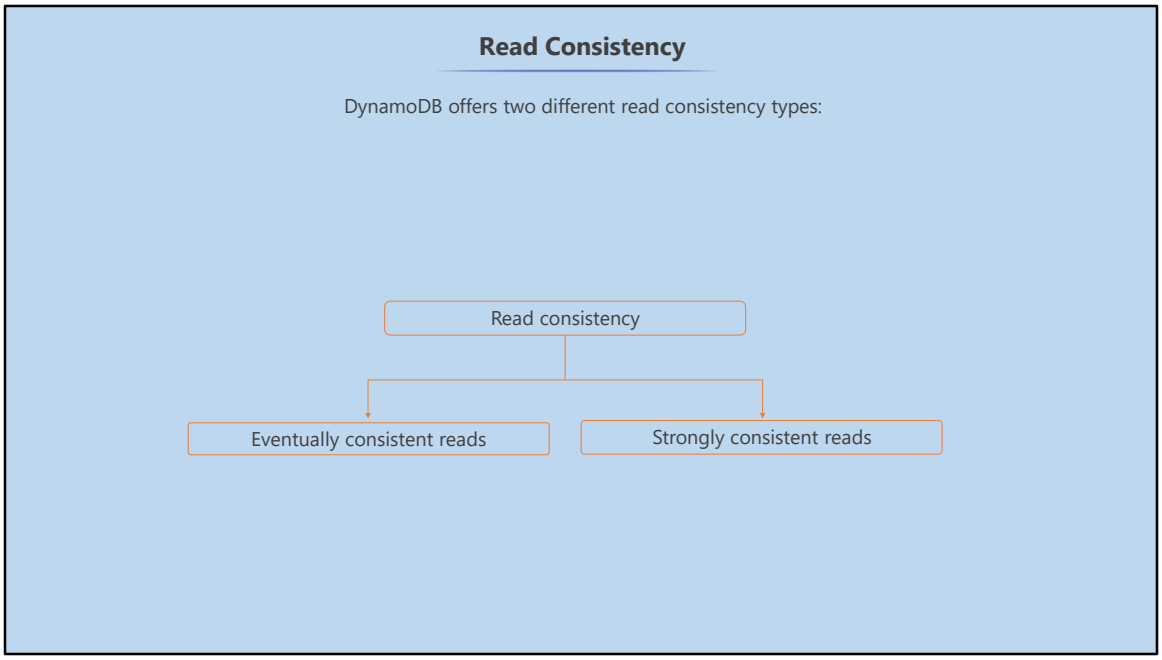


Amazon DynamoDB is a fast, flexible NoSQL database service for applications that need consistent single-digit millisecond latency at any scale. It is a fully managed NoSQL database that is suitable for document and key-value store models. It is perfect for mobile, web, gaming, advertising technology, and so on.



Following are the benefits of DynamoDB:

- It is stored on SSD storage, which provides great disk throughput.
- It has high-availability built-in. It is spread across three geographically distinct data centers.
- It is fast and flexible and it is suited for read-heavy applications.



DynamoDB offers two different read consistency types:

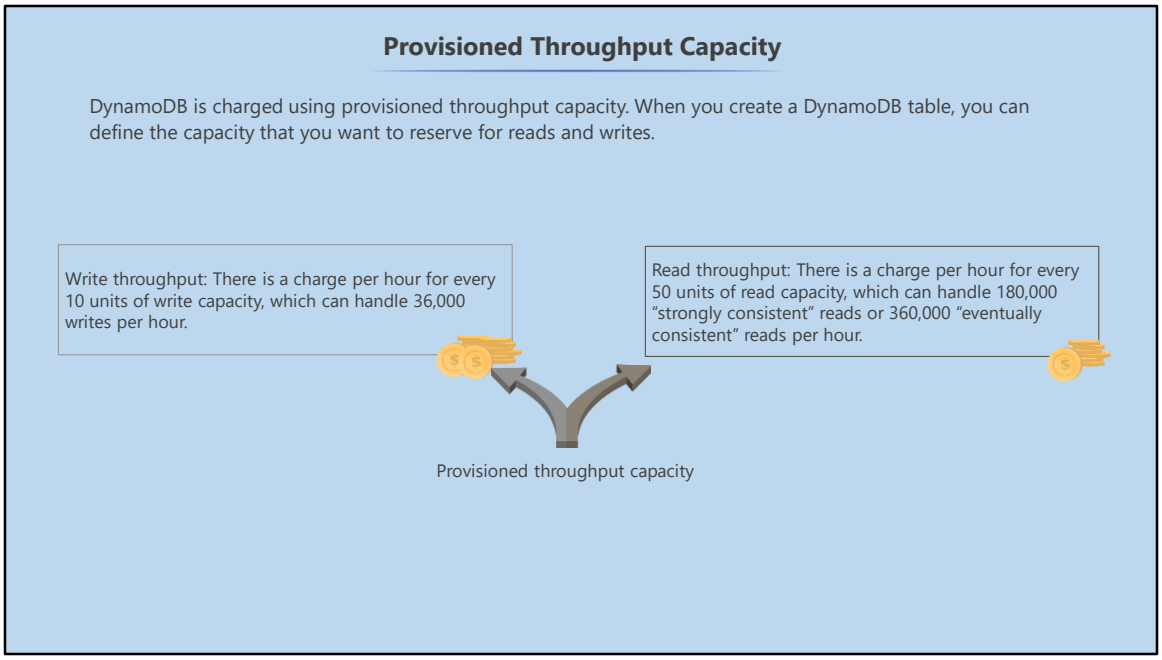
- Eventually consistent reads
- Strongly consistent reads

You can choose your read consistency based on two conditions:

- You want users to be guaranteed to have the correct data, or
- You are okay with users not getting the latest version occasionally.

“Eventually consistent reads” is the default option. The consistency is reached within one second and on repeating a read after a short time it returns the correct data. But it can throw a "dirty reads" result at times, which means you request a read but do not receive the up-to-date version.

“Strongly consistent reads” guarantees that all writes are committed prior to the read action taking place, so the applications or users receive the latest version. However, it can result in a delay in the results being returned.



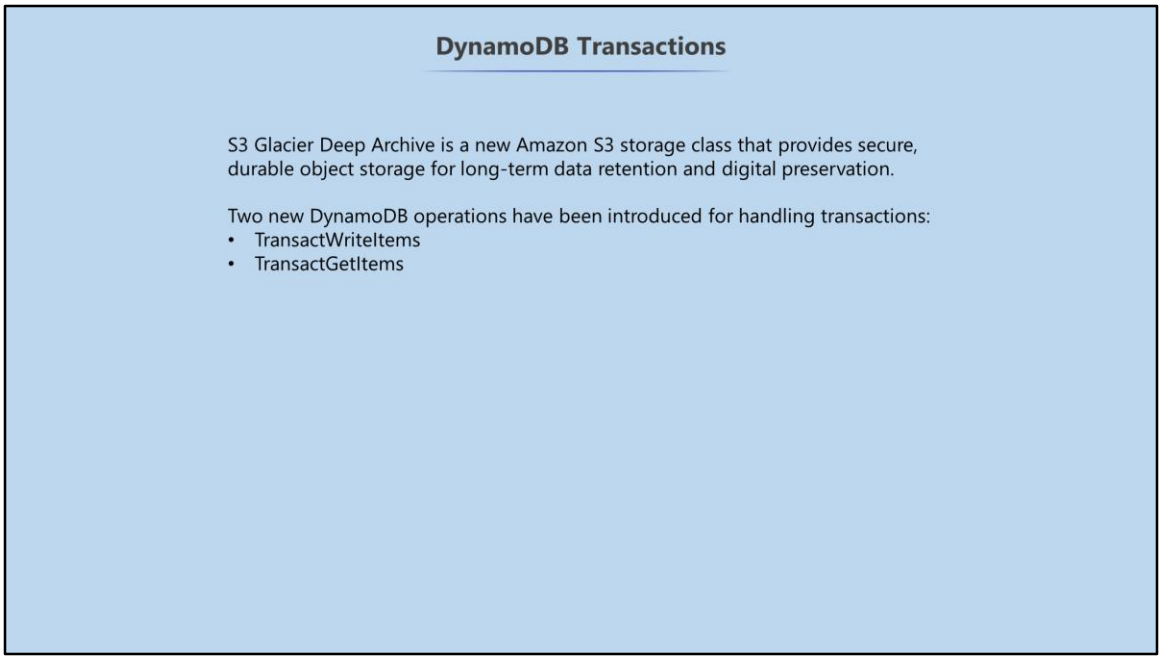
DynamoDB is charged using provisioned throughput capacity. When you create a DynamoDB table, you can define the capacity that you want to reserve for reads and writes.

Write throughput: There is a charge per hour for every 10 units of write capacity, which can handle 36,000 writes per hour.

Read throughput: There is a charge per hour for every 50 units of read capacity, which can handle 180,000 “strongly consistent” reads or 360,000 “eventually consistent” reads per hour.

DynamoDB is expensive for writes but cost effective for reads. If your application has lots of reads, little writes, and wants to be scalable, then DynamoDB is the correct solution.

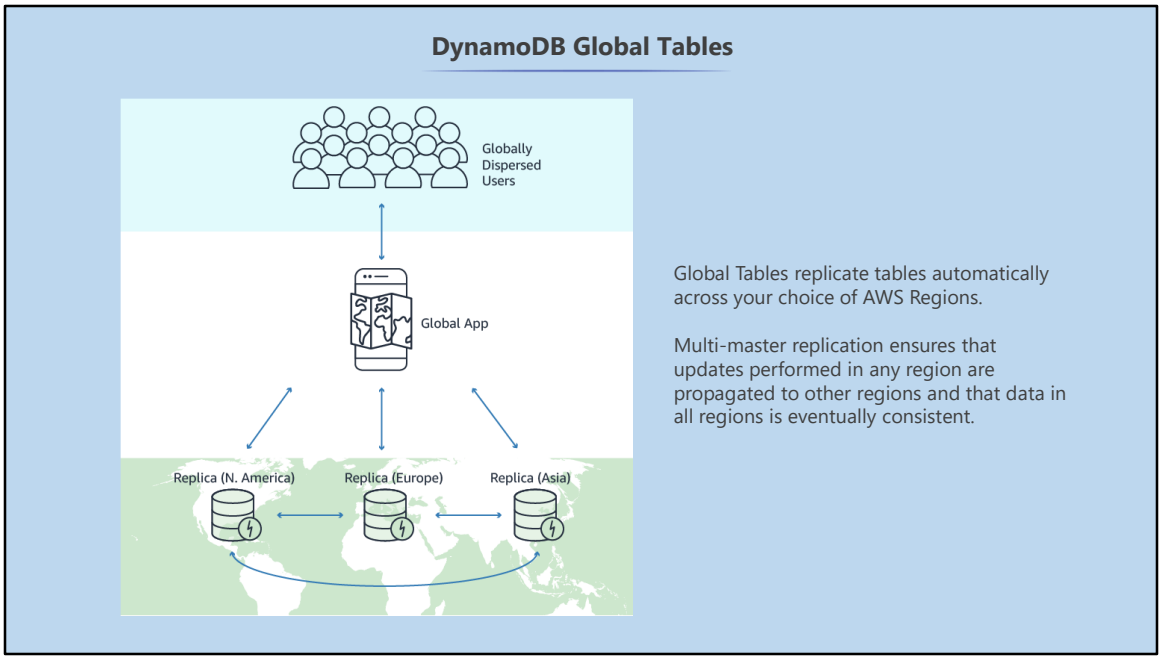
There is also a storage cost per gigabyte for the amount of data you use.



DynamoDB transactions provide developers atomicity, consistency, isolation, and durability (ACID) across one or more tables within a single AWS account and region. You can use transactions when building applications that require coordinated inserts, deletes, or updates to multiple items as part of a single logical business operation. DynamoDB is the only non-relational database that supports transactions across multiple partitions and tables.

Two new DynamoDB operations have been introduced for handling transactions: TransactWriteItems, a batch operation that contains a write set, with one or more PutItem, UpdateItem, and DeleteItem operations. TransactWriteItems can optionally check for prerequisite conditions that must be satisfied before making updates. These conditions may involve the same or different items than those in the write set. If any condition is not met, the transaction is rejected.

TransactGetItems, a batch operation that contains a read set, with one or more GetItem operations. If a TransactGetItems request is issued on an item that is part of an active write transaction, the read transaction is canceled. To get the previously committed value, you can use a standard read.

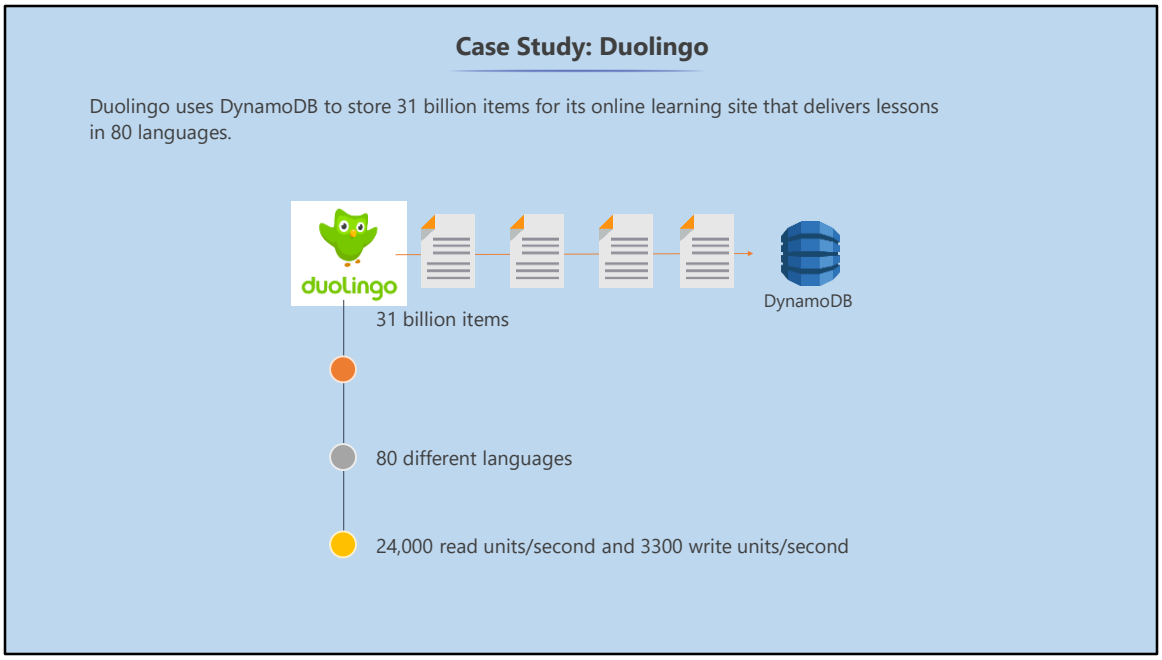


Global Tables replicates your tables automatically across your choice of AWS Regions.

Can be enabled with just a few clicks

Multi-master replication ensures that updates performed in any region are propagated to other regions, and that data in all regions are eventually consistent. This means tables accessed locally by your globally distributed application are always up to date.

Feature announced at re:Invent 2017



Duolingo uses DynamoDB to store 31 billion items for their online learning site that delivers lessons in 80 languages.

Their application hits 24,000 read units per second and 3,300 write units per second. It is a read-heavy application and it's perfect for DynamoDB.

Demo: Amazon DynamoDB

In this demonstration, you'll learn about Amazon DynamoDB and how to scale.

Knowledge Check

1 **Which application type is best suited for DynamoDB?**

Structured data with heavy write activity

Unstructured data with heavy write activity

Structured data with heavy read activity

Unstructured data with heavy read activity

1 **Which application type is best suited for DynamoDB?**

Structured data with heavy write activity

Unstructured data with heavy write activity

Structured data with heavy read activity

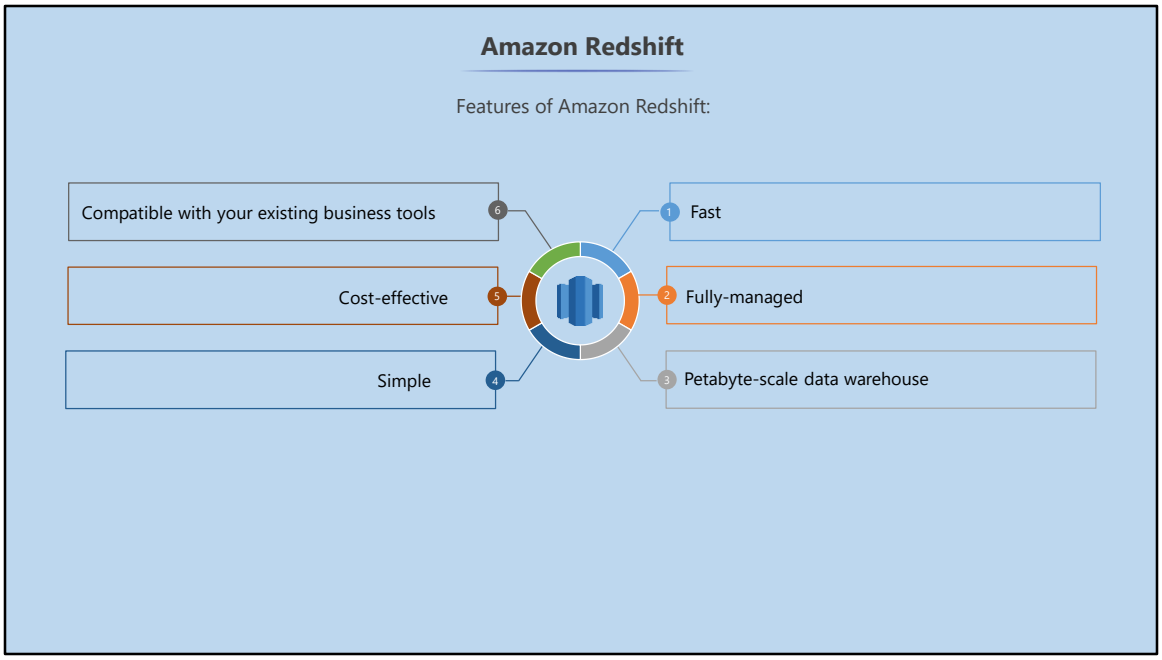
Unstructured data with heavy read activity

d

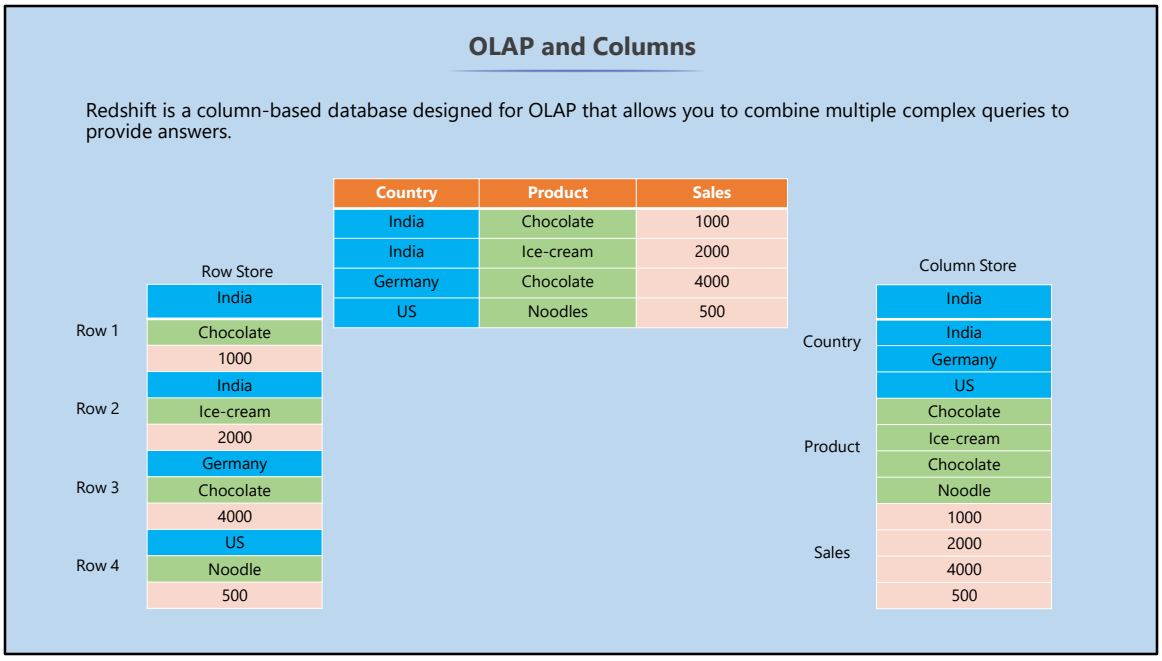
DynamoDB is a NoSQL database best suited for applications with heavy read and low write activity.

Amazon Redshift

In this section you'll learn about the overview of Amazon Redshift.

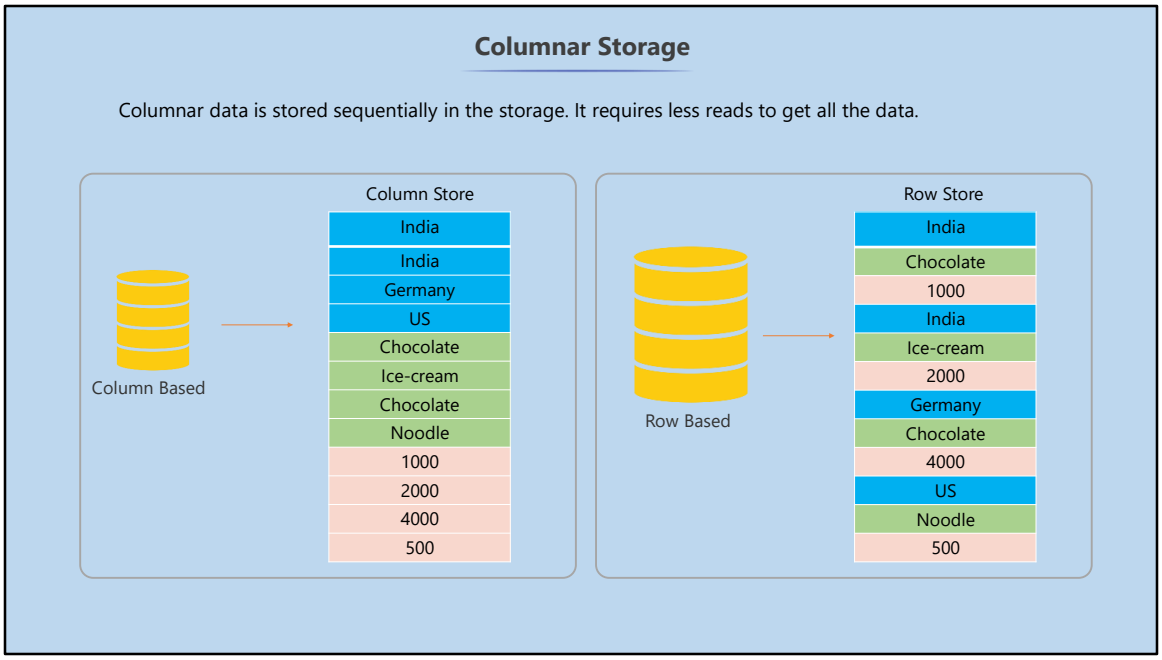


Amazon Redshift is a fast, fully-managed, petabyte scale data warehouse that makes it simple and cost-effective to analyze all your data using your existing business intelligence tools. Redshift is a great choice if your database is overloaded due to OLAP transactions.

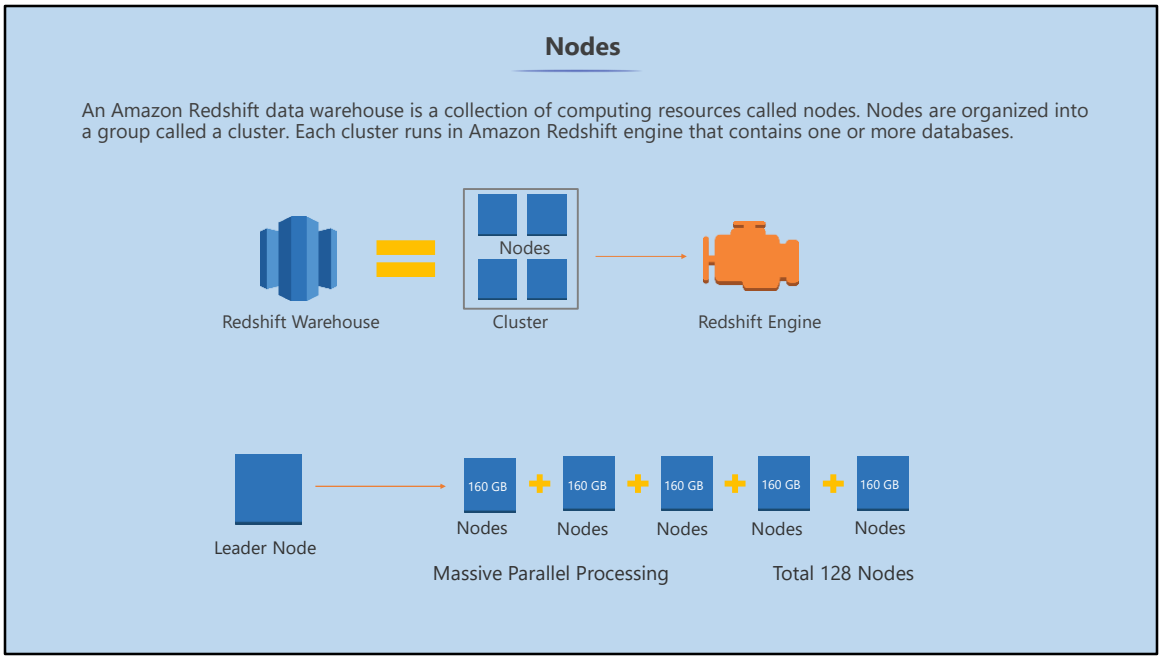


Amazon Redshift is designed for OLAP which allows you to easily combine multiple complex queries to provide answers.

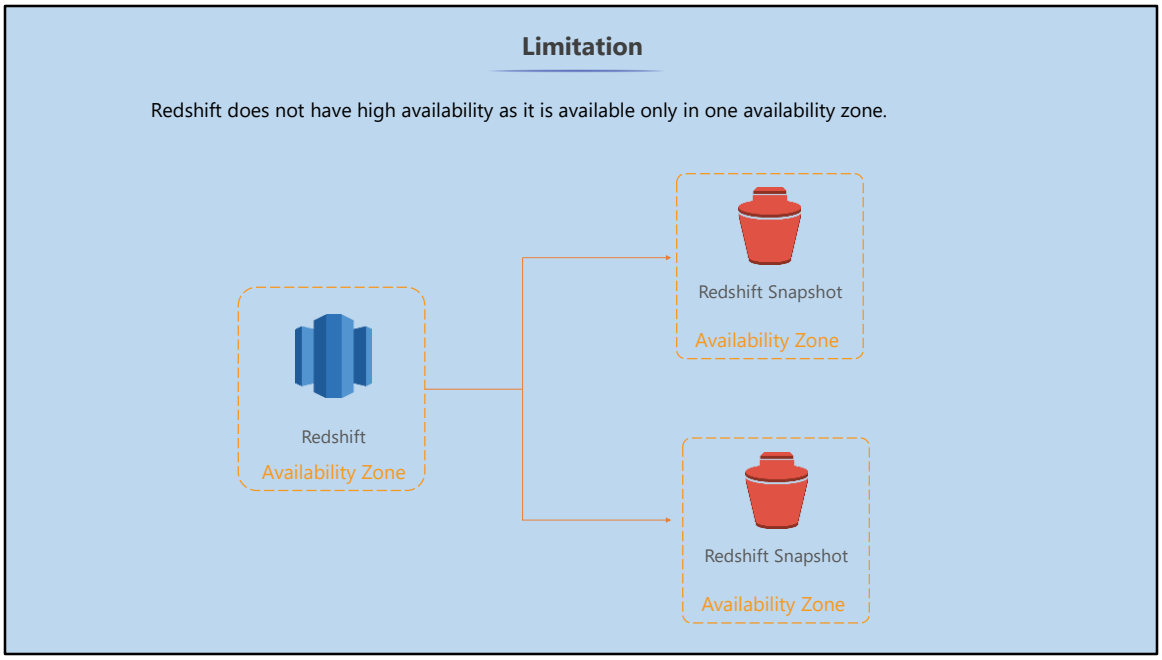
Relational or SQL databases are row-based; however, Redshift is a column-based database. The first table on the screen has three columns: Country, Product, and Sales and then the different rows presenting the products.
In a relational database, row one would be: Country is India, Product is chocolate, and Sales is 1000.
However, in a column based database, all the countries are stored together, then the products, and lastly the Sales. This is a key difference between “row store” and “column store” databases.



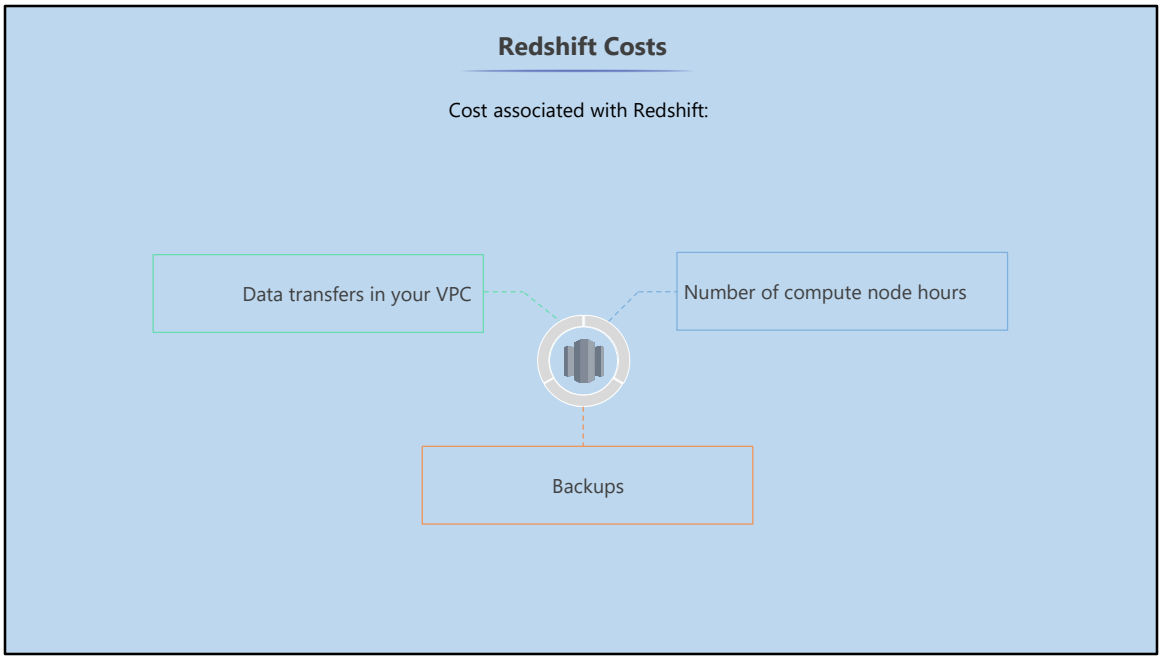
Columnar data is stored sequentially in the storage and it requires less reads to get all the data. Columnar data can be compressed much easier as all the data types are the same. But in row-based storage all the data types are different; therefore, compression is comparatively harder.



An Amazon Redshift data warehouse is a collection of computing resources called nodes. Nodes are organized into a group called a cluster. Each cluster runs in Amazon Redshift engine that contains one or more databases. When Amazon Redshift is launched, you can start with a single node that is 160 GB in size. You can add additional nodes to take advantage of massive parallel processing, and you can operate in multimode. A leader node manages all the client connections and the compute nodes. The compute nodes store data and perform queries and computations. There can be up to 128 compute nodes.



Redshift does not have high availability as it is available only in one availability zone. You can, however, restore snapshots of your Redshift databases to other availability zones.



The following are the cost associated with Redshift:
Charge for the number of compute node hours that are used (it doesn't include the leader node)
Charge for the backups that you store
Charge for any data transfers within VPC

Knowledge Check

1 **Which of the following is NOT a benefit of Redshift?**

Ability to reduce storage costs

Requires less reads to get data from storage

Provides high availability

Ability to process OLAP queries quickly and efficiently

1 **Which of the following is NOT a benefit of Redshift?**

Ability to reduce storage costs

Requires less reads to get data from storage

Provides high availability

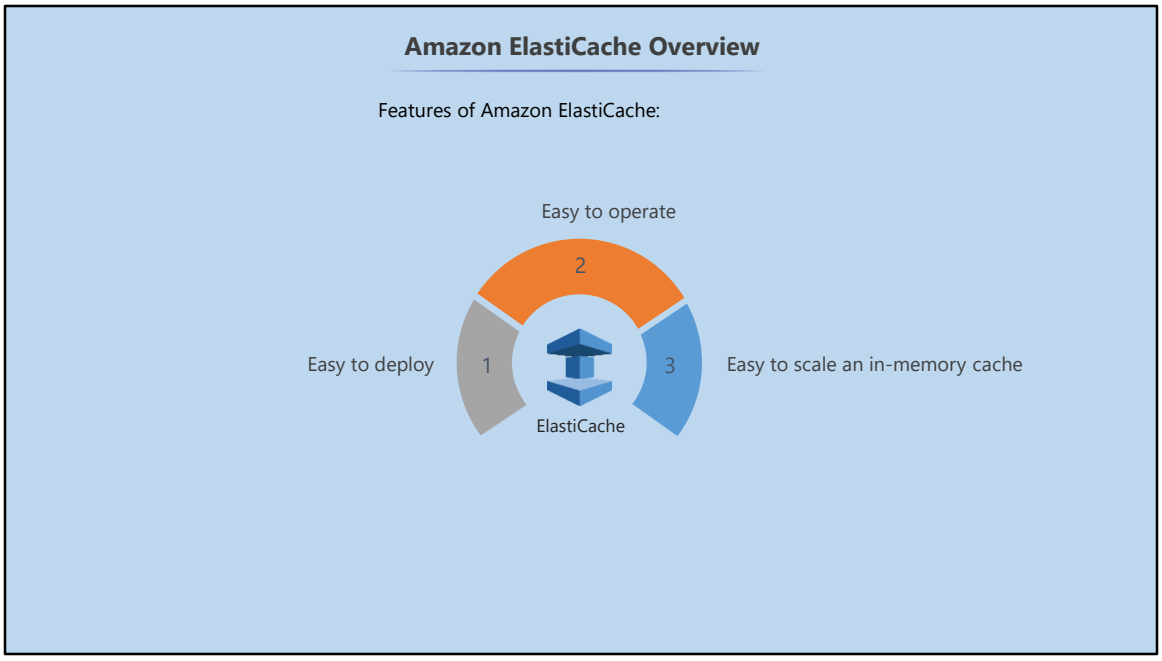
Ability to process OLAP queries quickly and efficiently

c

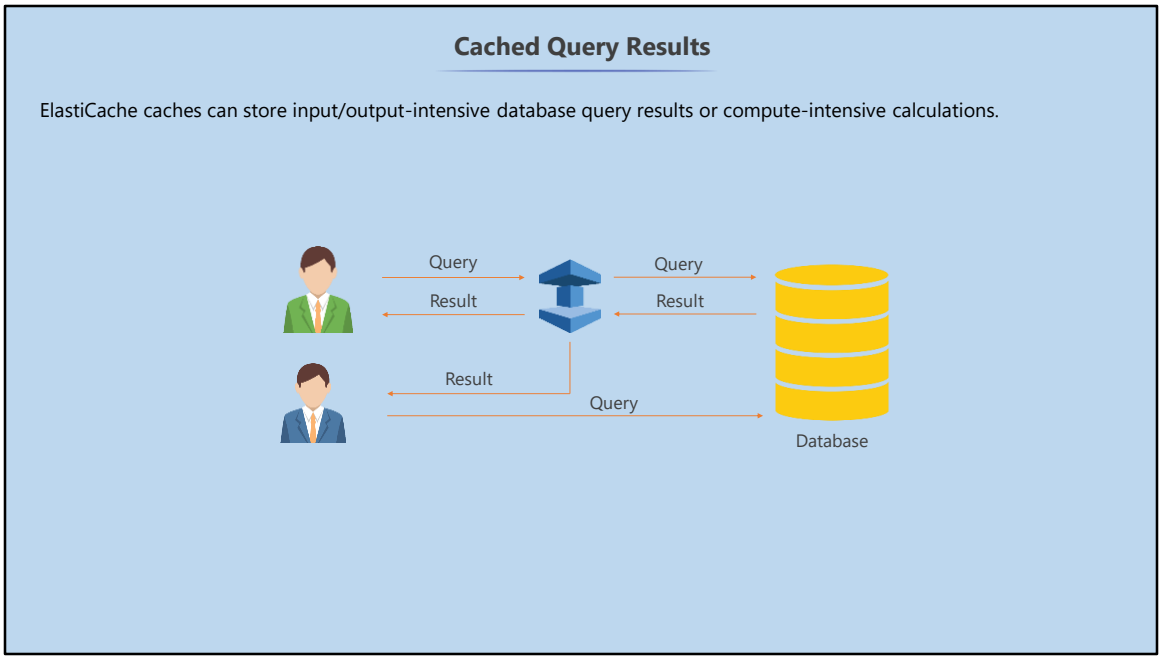
Redshift is not designed for high availability as it is only available in one Availability Zone.

Amazon ElastiCache

In this section you'll learn about Amazon ElastiCache.

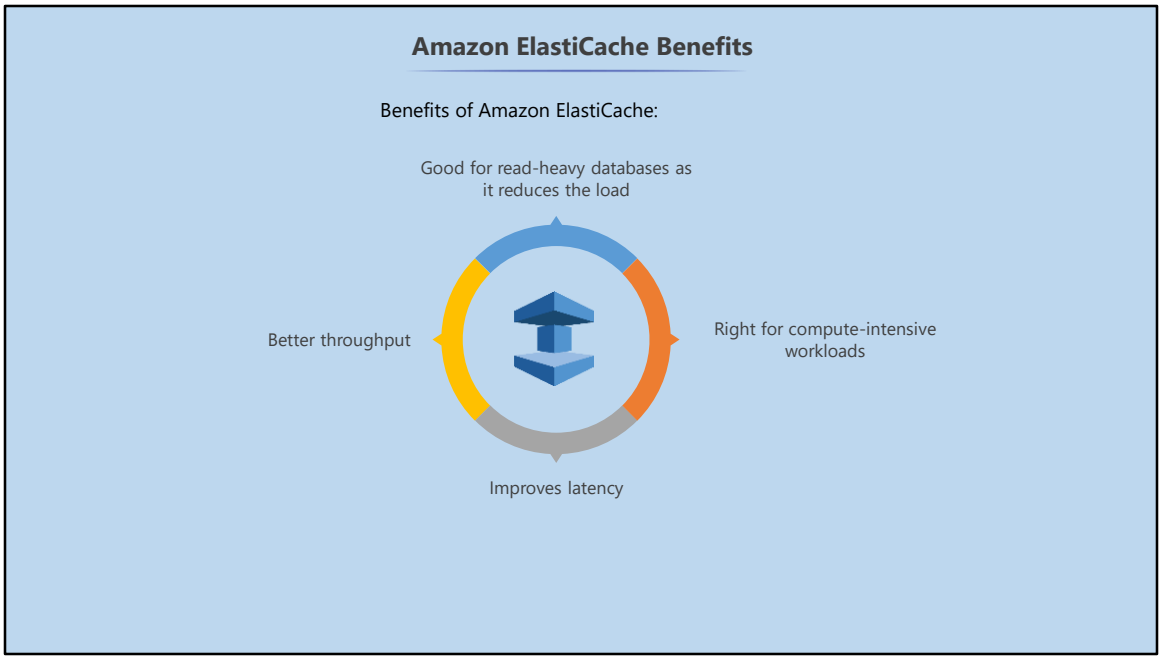


ElastiCache is an AWS service that makes it easy to deploy, operate, and scale an in-memory cache in the cloud.



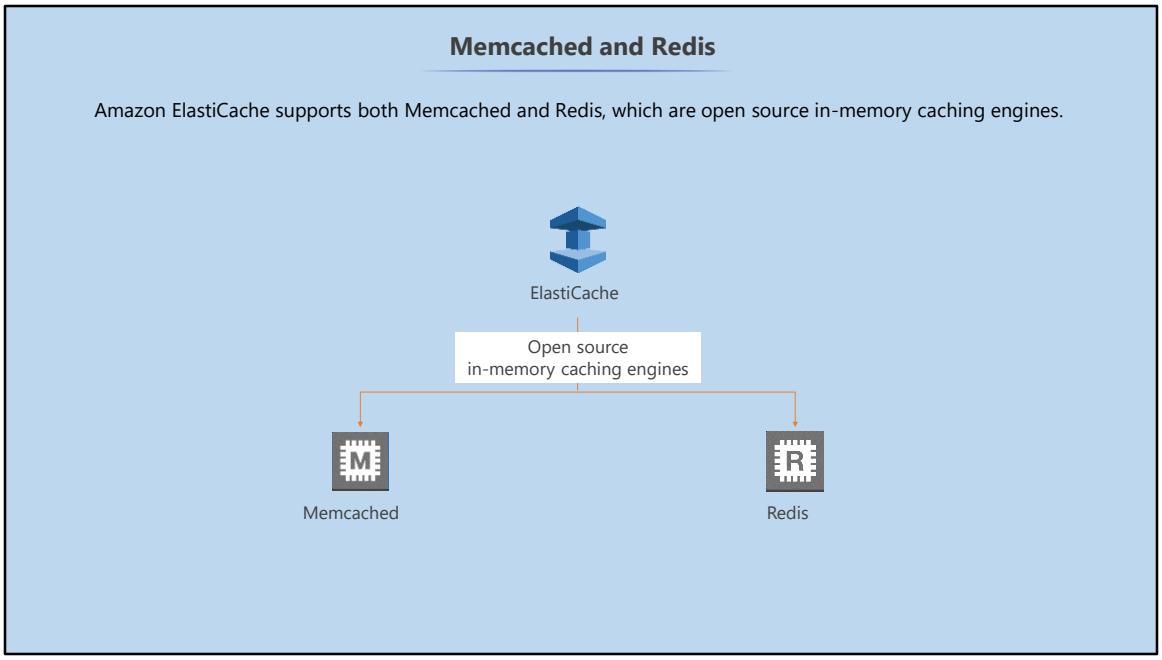
ElastiCache caches can store Input/Output intensive database query results or compute intensive calculations. ElastiCache is a good choice if you have read-heavy databases.

When users make queries from a database, ElastiCache stores the results in the cache. When the next user makes the same query the result is retrieved from the cache itself rather than going to the database.



Following are the benefits of Amazon ElastiCache:

- Good for read-heavy databases as it reduces the load on databases that serve read-intensive applications, such as social media or gaming
- Great for compute-intensive workloads
- It gives users improved latency
- It provides improved throughput



Amazon ElastiCache supports both Memcached and Redis, which are open source in-memory caching engines.
Memcached is an in-memory object caching system and Redis is an in-memory key-value store that supports sets and lists and master-slave functionality. Redis allows you to have multi-availabilities on capabilities.

Knowledge Check

1 **Which type of application is ElastiCache suitable for?**

Write heavy with low read activity

Read heavy with high write activity

Read heavy with high write activity

Write heavy with high read activity

1 **Which type of application is ElastiCache suitable for?**

Write heavy with low read activity

Read heavy with high write activity

Read heavy with high write activity

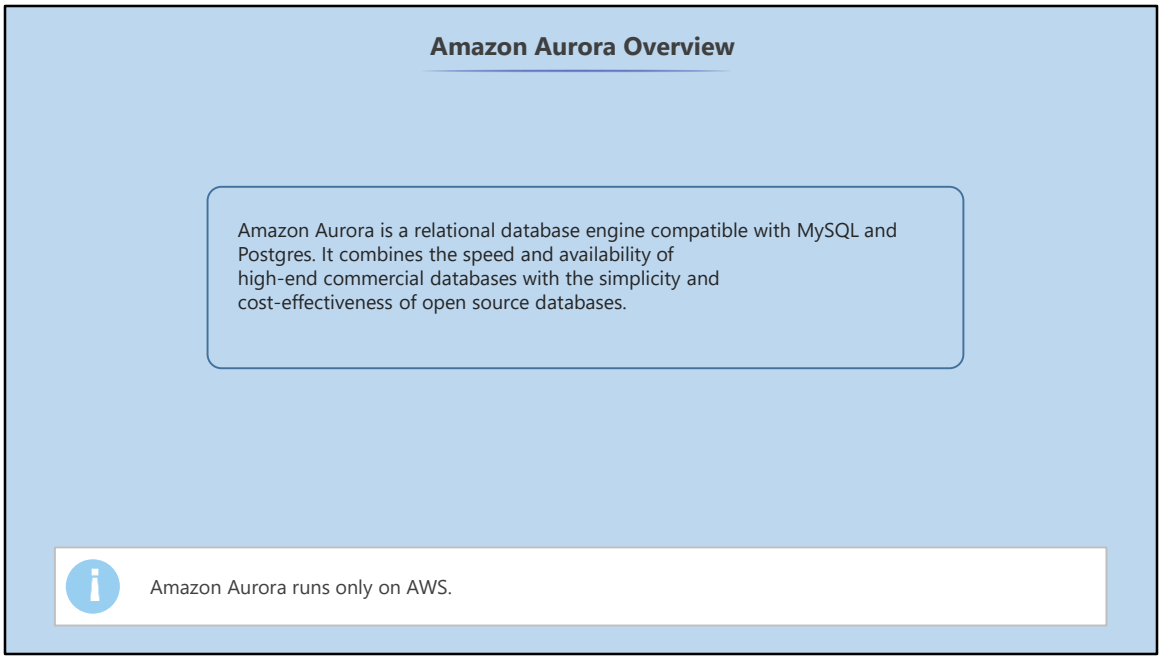
Write heavy with high read activity

b

ElastiCache is suitable for applications that are read heavy with low write activity so users' queries can be cached for future use.

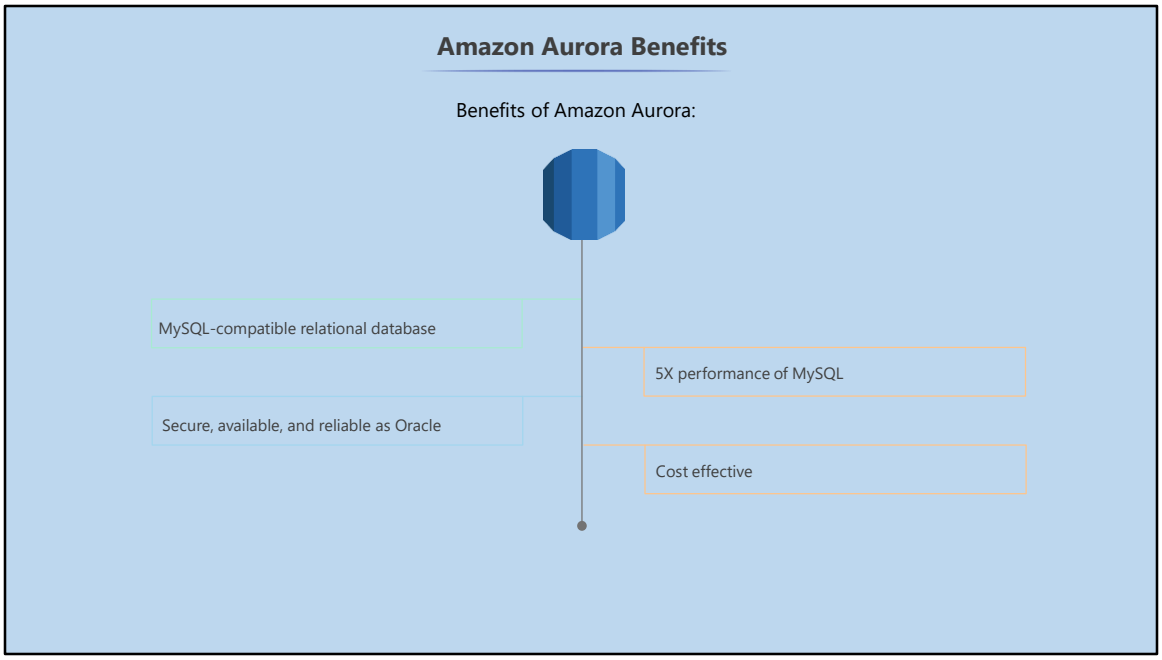
Amazon Aurora

In this section, you'll learn about Amazon Aurora and its features.

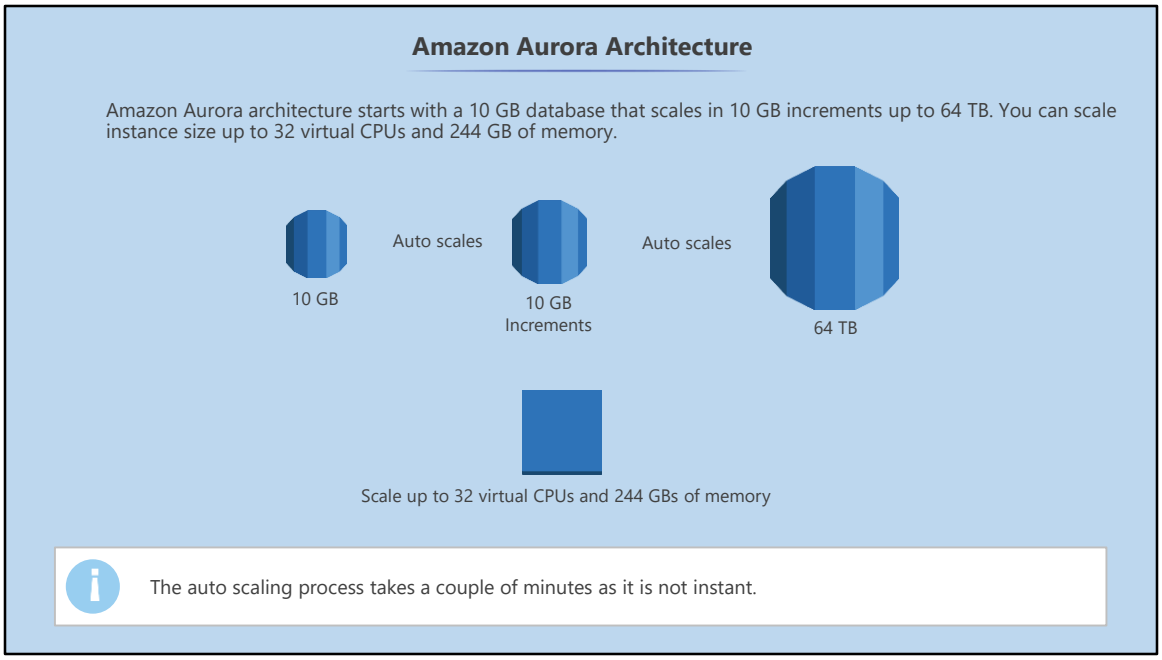


Amazon Aurora is a MySQL and Postgres compatible relational database engine that combines the speed and availability of high-end commercial databases with the simplicity and cost-effectiveness of open source databases.

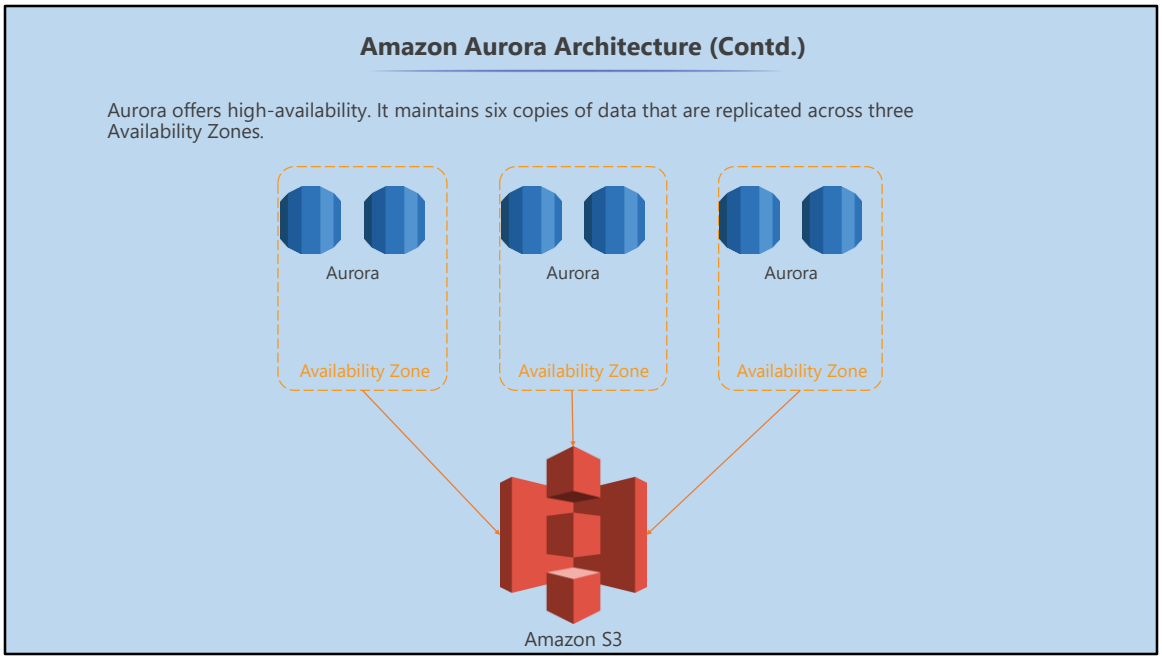
Aurora only runs on AWS.



It is MySQL compatible but it offers five times the performance of MySQL.
Aurora offers the same levels of security, availability, and reliability of commercial databases, such as Oracle but at one-tenth of the cost.



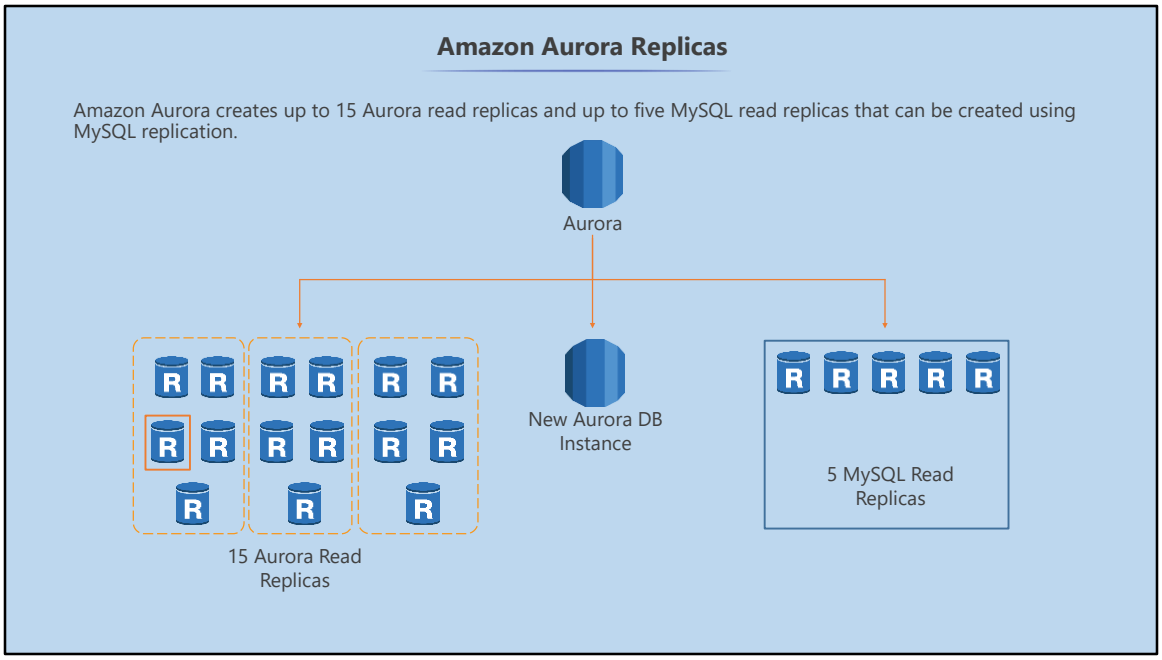
Amazon Aurora architecture starts with a 10 GB database that scales in 10 GB increment up to 64 TB. You can scale instance size up to 32 virtual CPUs and 244 GB of memory. It takes a couple of minutes to scale as it is not instant.



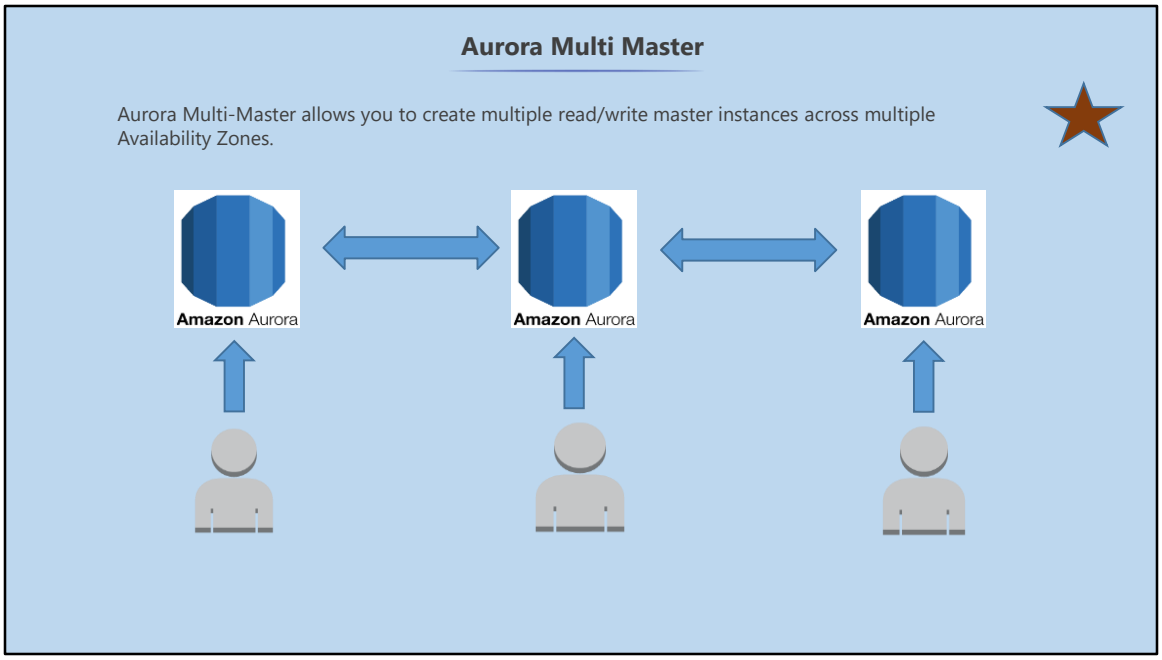
Aurora offers high-availability, it maintains six copies of data which is replicated across three availability zones.

The data results are backed up to Amazon S3. Aurora is designed to handle the loss of up to two copies of data without affecting database write availability, and you can lose up to three copies of your data without affecting read availability.

It has self-healing storage. Data blocks and disks are constantly scanned and are repaired automatically.

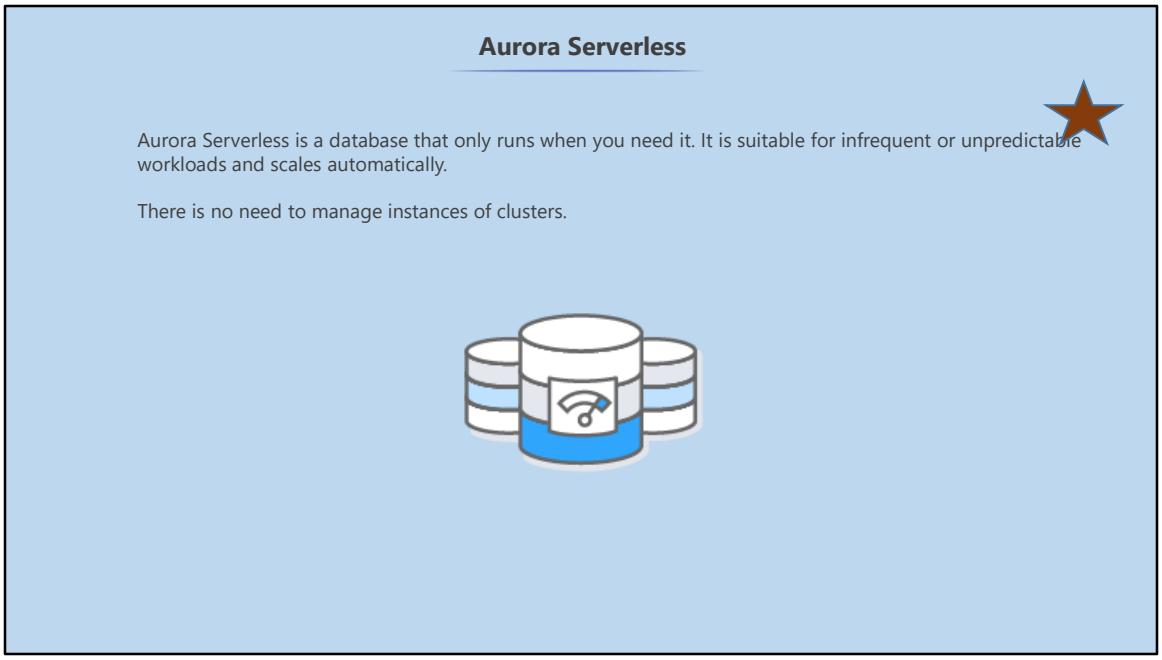


Amazon Aurora creates up to 15 Aurora read replicas and up to five MySQL read replicas which can be created using MySQL replication. If your instance fails, Amazon Aurora automatically fails over to one of the 15 Amazon Aurora replicas. You can't failover to MySQL replicas. If the user has not provisioned any Amazon Aurora replicas, then Amazon RDS automatically creates a new Amazon Aurora DB instance.



In preview in early 2018, Aurora Multi-Master allows you to create multiple read/write master instances across multiple Availability Zones.

Multi-Master clusters improve Aurora's already high availability. If one of your master instances fail, the other instances in the cluster will take over immediately, maintaining read and write availability through instance failures or even complete AZ failures, with zero application downtime.



Serverless database that only runs when you need it.

No need to manage instances of clusters.

Suitable for infrequent or unpredictable workloads.

Automatically scales.

Only pay for what you use.

Suitable for infrequently accessed websites, dev or test databases or applications or for the early stages of building a new application.

Announced at re:Invent 2017 and in preview 2018

Knowledge Check

1 **Amazon Aurora maintains:**

6 copies in 2 Availability Zones

3 copies in 3 Availability Zones

9 copies in 3 Availability Zones

6 copies in 3 Availability Zones

1 **Amazon Aurora maintains:**

6 copies in 2 Availability Zones

3 copies in 3 Availability Zones

9 copies in 3 Availability Zones

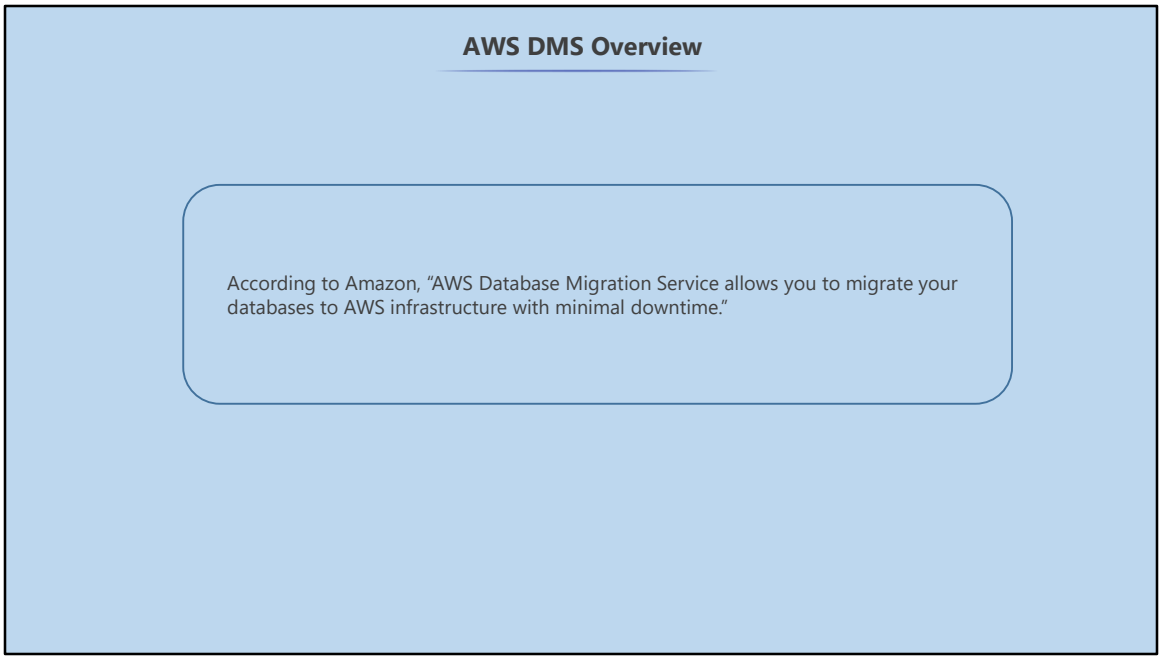
6 copies in 3 Availability Zones

d

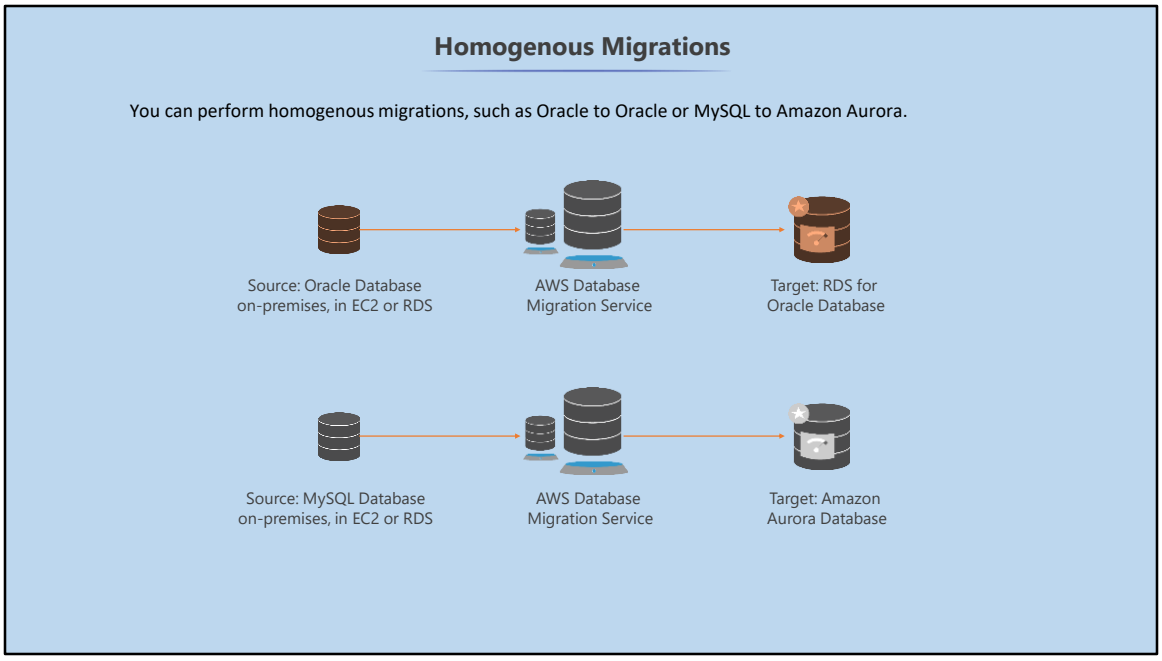
Amazon Aurora maintains 6 copies of data, which are replicated across three Availability Zones.

AWS Database Migration Services (DMS)

In this section you'll learn about AWS Database Migration Services.



According to Amazon, "AWS Database Migration Service allows you to migrate your databases to AWS infrastructure with minimal downtime."



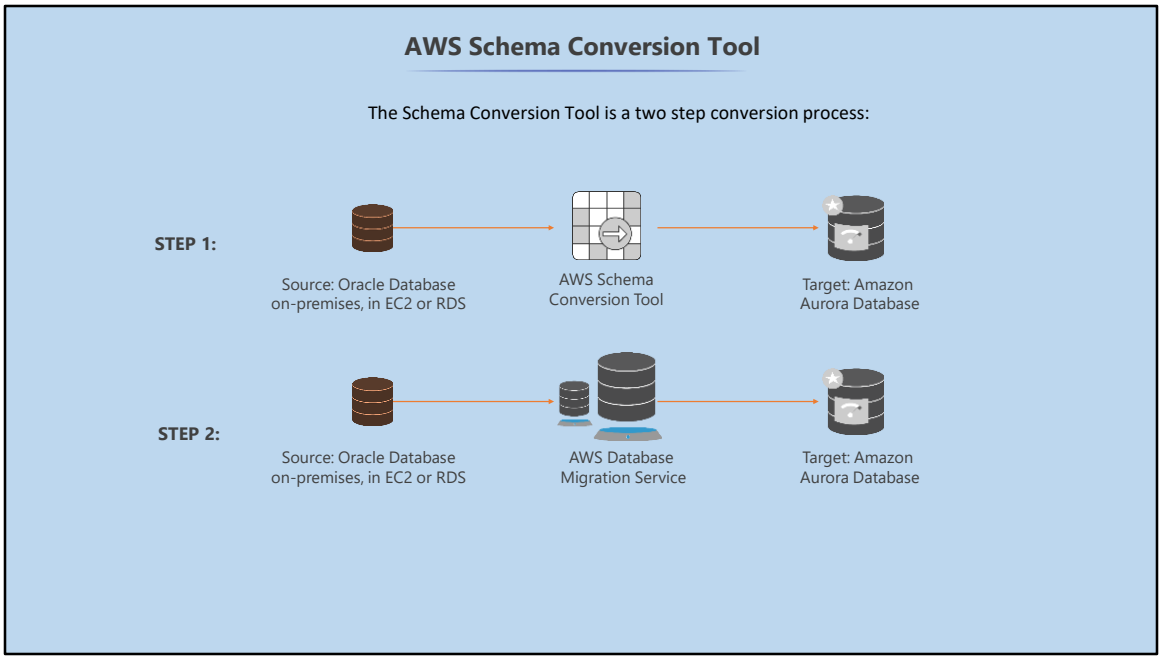
AWS DMS allows you to perform homogenous migrations, for example Oracle to Oracle or MySQL to Amazon Aurora. You can migrate any database that is in your data center in EC2 or RDS to any other database.
You can do continuous replication from your data center to the databases in AWS or you can replicate to a database in your data center from a database in AWS.

Heterogeneous Migrations

Using the AWS Schema Conversion Tool, you can perform heterogeneous migrations between different RDBMS.

Source Database	Target Database
Microsoft SQL Server	Amazon Aurora, MySQL, PostgreSQL, MariaDB
MySQL and MariaDB	PostgreSQL
Oracle	Amazon Aurora, MySQL, PostgreSQL, MariaDB
PostgreSQL	Amazon Aurora, MySQL, MariaDB
Amazon Aurora	PostgreSQL

Using the AWS Schema Conversion Tool you can perform heterogeneous migrations between different RDBMS.

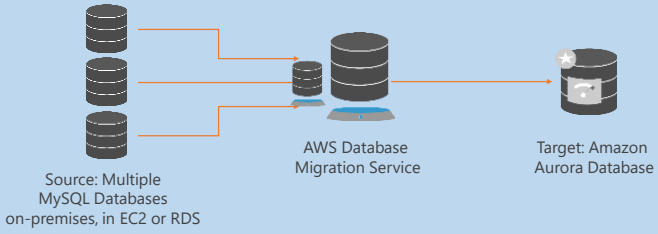


The Schema Conversion Tool is a two-step conversion process:

- You take your existing database, in this example Oracle, and you use the AWS Schema Conversion Tool to create a target database, which in this example is Amazon Aurora.
- Once the task is completed, you can set up another task using the Database Migration Service to migrate the data from your Oracle database into the new Amazon Aurora target database. Any code that cannot be automatically converted is clearly marked so that it can be manually converted.

AWS Schema Conversion Tool (Contd.)

DMS has the ability to stream data from any supported database source to Aurora or Redshift so you can easily consolidate your data or create a data warehouse.



DMS has the ability to stream data from any supported database source to Aurora or Redshift so you can easily consolidate your data or create a data warehouse.

Knowledge Check

1 **Which of the following two types of migration does AWS DMS support?**

Heterogeneous and Homogenous

Transactional and Replication

Synchronous and Asynchronous

Transient and Migratory

1 **Which of the following two types of migration does AWS DMS support?**

Heterogeneous and Homogenous

Transactional and Replication

Synchronous and Asynchronous

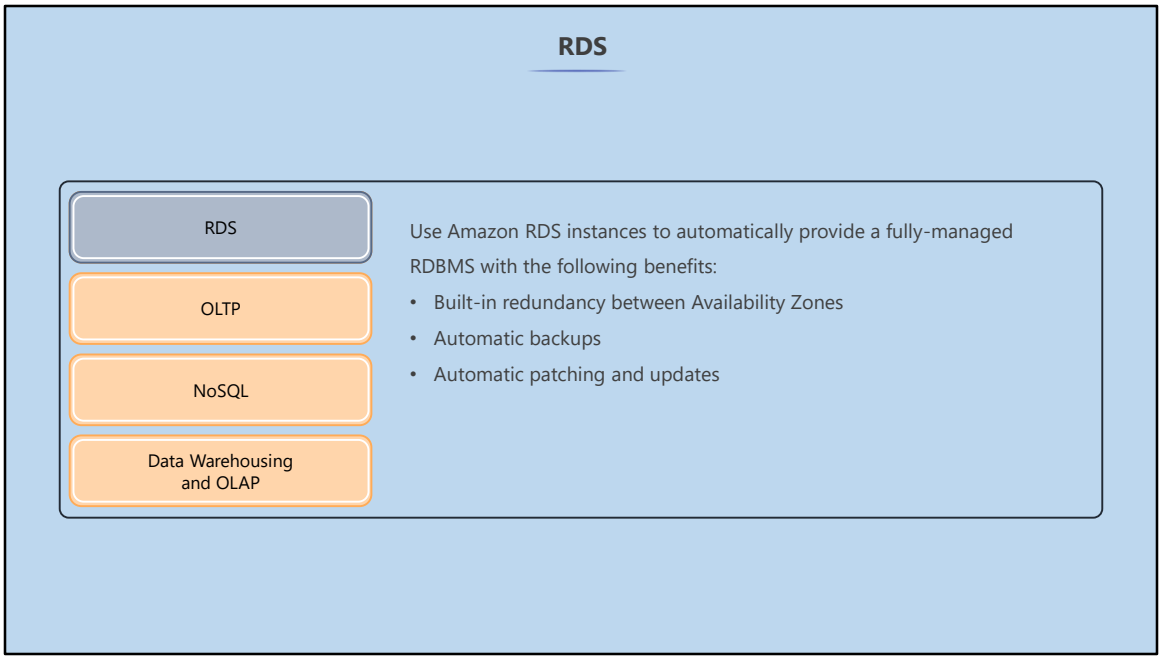
Transient and Migratory

a

AWS DMS supports Heterogeneous and Homogenous migration.

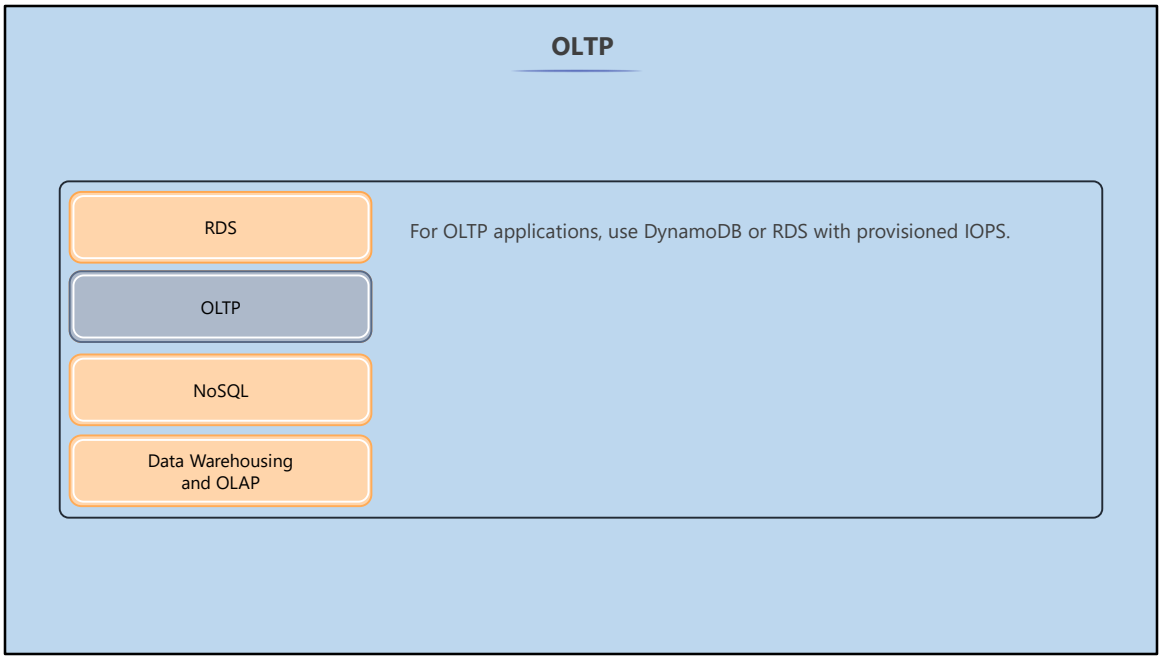
Database Best Practices

In this section you'll learn about the database best practices.

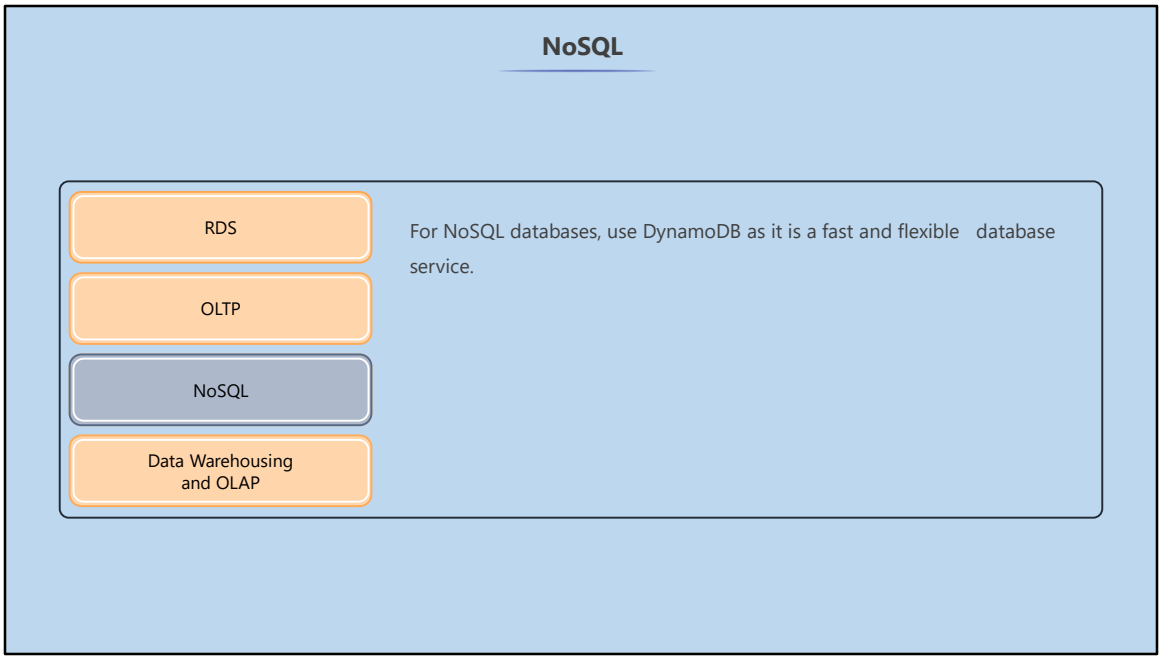


Use Amazon RDS instances to automatically provide a fully managed RDBMS with the following benefits:

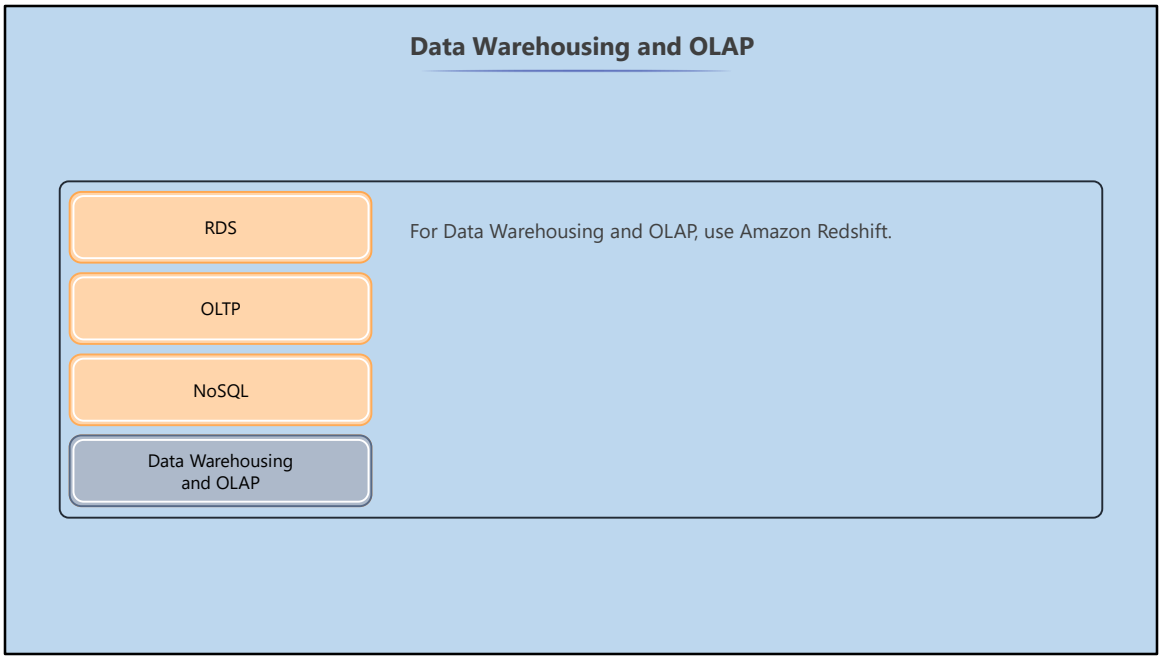
- Built-in redundancy between Availability Zones
- Automatic backups
- Automatic patching and updates



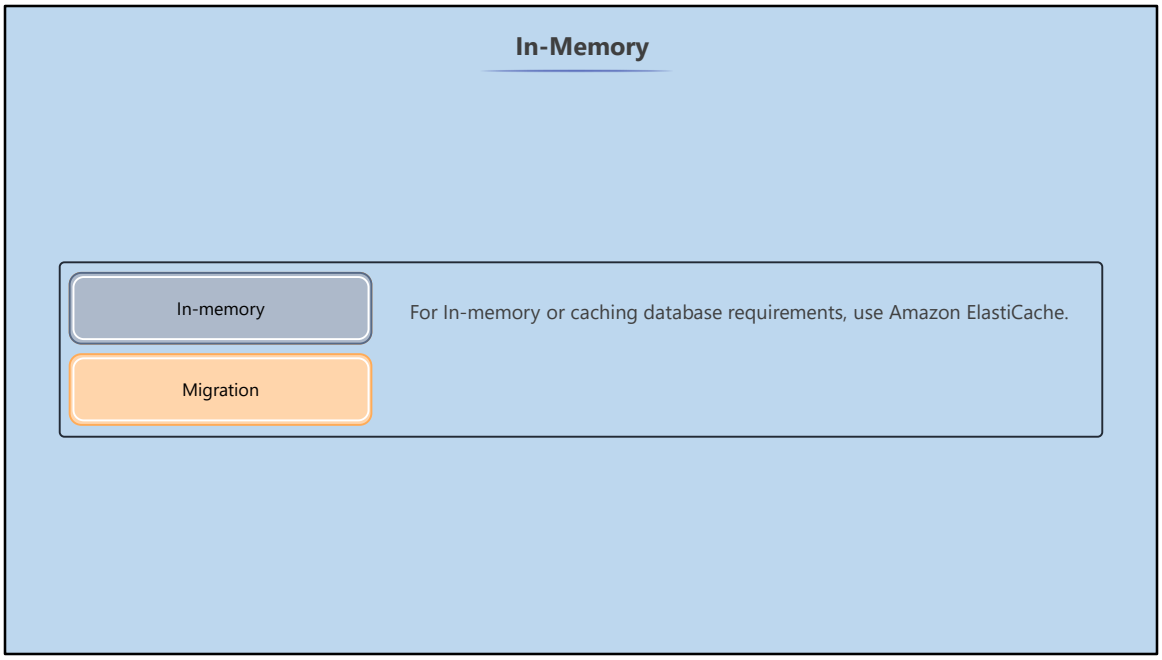
For OLTP applications use DynamoDB or RDS with provisioned IOPS.



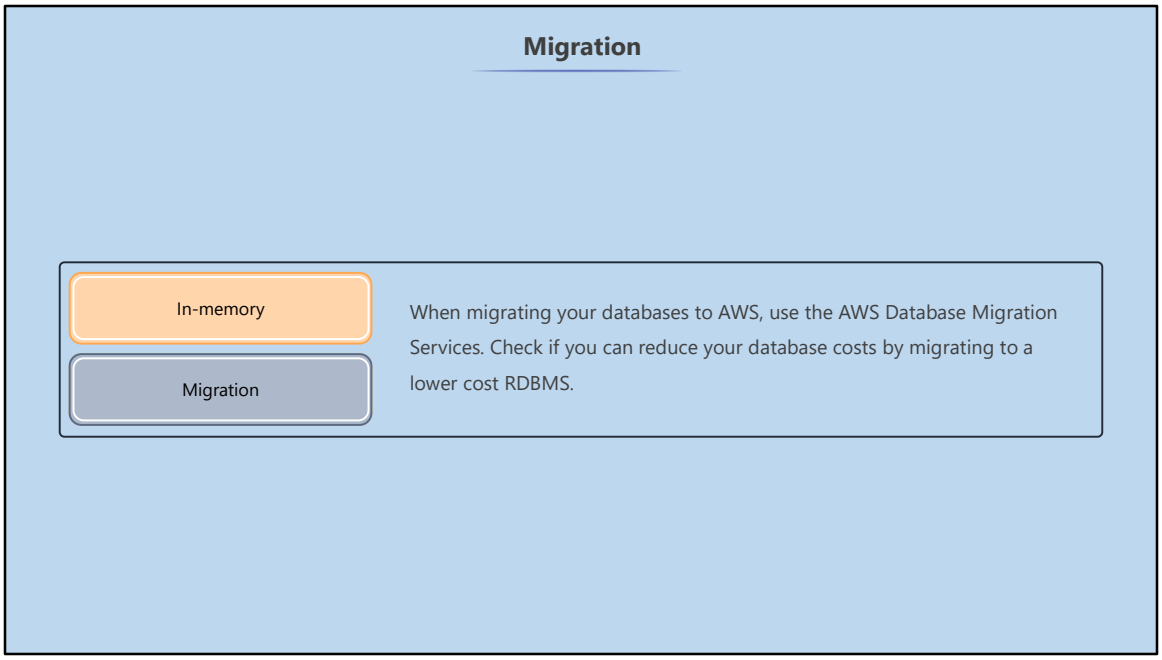
For NoSQL databases use DynamoDB as it is a fast and flexible database service.



For Data Warehousing and OLAP use Amazon Redshift.



For In-memory or caching database requirements use Amazon ElastiCache.



When migrating your databases to AWS, use the AWS Database Migration Services and check if you can reduce your database costs by migrating to a lower cost RDBMS.

Database Costs

In this section you'll learn about the database costs associated with AWS.

Database Costs

The following costs are associated with the AWS databases:

- Amazon RDS:
  - Cost per hour depending on the instance size and RDBMS licensing cost (if required)
  - Per GB cost for storage
- Amazon DynamoDB:
  - Write Throughput: fee per hour for every 10 units of Write Capacity
  - Read Throughput: fee per hour for every 50 units of Read Capacity
  - Per GB fee for indexed storage
  - Fee for DynamoDB streams
  - Data transfer out fee
- Amazon Redshift:
  - Cost per hour depending on the instance size

- Amazon RDS:
  - Cost per hour depending on the instance size and RDBMS licensing cost (if required)
  - Per GB cost for storage
- Amazon DynamoDB:
  - Write Throughput fee per hour for every 10 units of Write Capacity
  - Read Throughput fee per hour for every 50 units of Read Capacity
  - Per GB fee for indexed storage
  - Fee for DynamoDB streams
  - Data transfer out fee
- Amazon Redshift
  - Cost per hour depending on the instance size

Database Costs (Contd.)

The following costs are associated with the AWS databases:

- Amazon Elasticache:
  - Cost per hour depending on the cache size selected
  - Data transfer cost between EC2 instances and Elasticache in different AZs
- AWS Database Migration Services:
  - Cost per hour for the compute resources used during migration
  - Data transfer out to the Internet

- Amazon ElastiCache:
  - Cost per hour depending on the cache size selected
  - Data transfer cost between EC2 instances and ElastiCache in different Availability Zones
- AWS Database Migration Services
  - Cost per hour for the compute resources used during migration
  - Charge for data transfer out to the internet

- RDS allows you to run databases in the AWS cloud on EC2 instances without worrying about database administration management tasks.
- Amazon DynamoDB is a fast and flexible NoSQL database service for all applications that need consistent, single-digit millisecond latency at any scale.
- Amazon Redshift is a fast, fully-managed, petabyte-scale data warehouse that makes it simple and cost-effective to analyze all your data using your existing business intelligence tools.
- ElastiCache is an AWS service that makes it easy to deploy, operate, and scale an in-memory cache in the cloud.
- Amazon DynamoDB is a fast and flexible NoSQL database service for all applications that need consistent, single-digit millisecond latency at any scale.

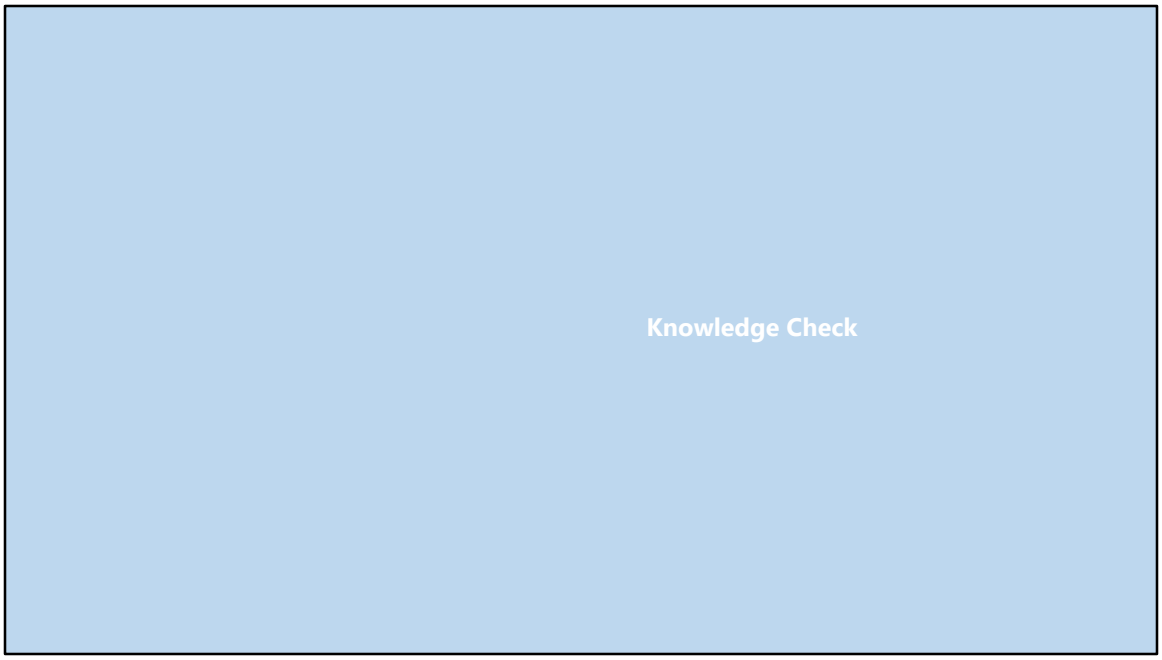
Practice Assignment: Amazon RDS

Your company wants to use an RDS instance as the data repository for their webserver.
You will need to complete the following tasks:

1. Create an additional private subnet in your custom VPC.
2. Verify the inbound rules for your DBSERVER security group.
3. Create a DB Subnet Group to host your RDS instance.
4. Launch a MySQL RDS instance into the DB subnet group.
5. Launch a new webserver into a public subnet of your custom VPC.
6. Configure connectivity between the webserver and the RDS instance.

You can use demonstration one from this lesson as a reference for this Practice Assignment.

This is the practice assignment for Amazon RDS. You have to launch an Amazon RDS instance and connect to a web server.



1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

1 **What is the charge when replicating data between your primary and secondary RDS instances?**

The AWS standard charge

Half the AWS standard charge

Depends on the amount of data transferred

No charge; it's free

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

1 **What is the charge when replicating data between your primary and secondary RDS instances?**

The AWS standard charge

Half the AWS standard charge

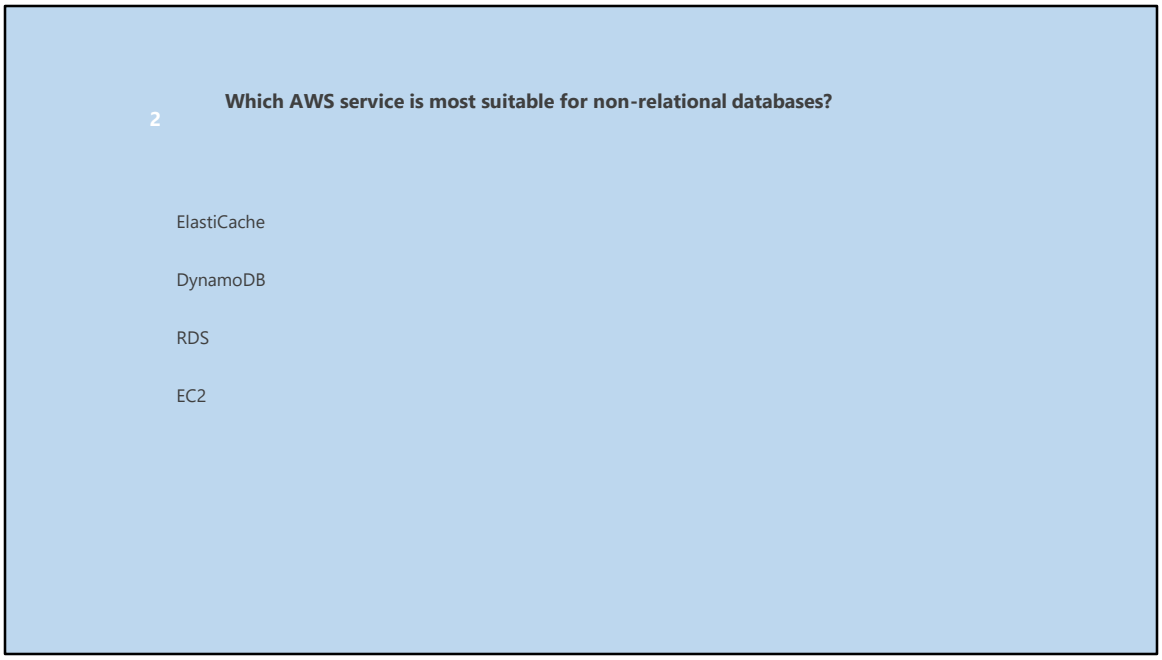
Depends on the amount of data transferred

No charge; it's free

d

RDS synchronously replicates your data to a standby instance in a different Availability zone for free.

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.



1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

2 **Which AWS service is most suitable for non-relational databases?**

ElastiCache

DynamoDB

RDS

EC2

b

DynamoDB is a NoSQL database that is best suited for non-relational databases.

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

3 What is the maximum-sized RDS volume you can have for MySQL and Oracle?

6TB

4TB

2TB

3TB

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

What is the maximum-sized RDS volume you can have for MySQL and Oracle?

3

6TB

4TB

2TB

3TB

a

MySQL, PostgreSQL, and Oracle RDS instances are limited to 6TB in size. For SQL Server, it is 4TB, and for SQL Server Express, it is 300GB.

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

4 What is the maximum IOPS capacity of an Oracle and MySQL RDS instance?

10,000 IOPS

20,000 IOPS

30,000 IOPS

50,000 IOPS

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

What is the maximum IOPS capacity of an Oracle and MySQL RDS instance?

4

10,000 IOPS

20,000 IOPS

30,000 IOPS

50,000 IOPS

c

MySQL, MariaDB, PostgreSQL, and Oracle RDS instances have a maximum of 30000 IOPS. SQL Server has a maximum of 20,000 IOPS.

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

5 **Which of the following ports needs to be open to use RDP to a SQL Server RDS instance?**

3389

1433

22

You cannot RDP to an RDS instance

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.

5 **Which of the following ports needs to be open to use RDP to a SQL Server RDS instance?**

3389

1433

22

You cannot RDP to an RDS instance

d

You cannot use RDP (or SSH) to RDS instances.

1. Amazon Web Services (AWS) is a secure cloud services platform that offers cloud-based infrastructure for compute, database storage, content delivery, and other functionalities to help businesses scale and grow.
2. AWS is truly global; it's available in 190 countries through 12 geographic Regions.
3. A region is a geographic area isolated from other Amazon regions to provide the greatest possible fault tolerance. Availability Zones are located within a region, with at least two per region, and are connected via low-latency links.
4. Edge locations are CDNs and are located all over the world in major cities. Used to provide content to end users with low latency.
5. AWS has various cloud-based products to help your business grow.