AWS Database Services

Module 4

AWS Database Services

Various databases that are provided by the Amazon Web Services(AWS) are :

- Amazon DynamoDB
- Amazon Aurora
- Amazon Relational Database Service
- Amazon Timestream
- Amazon Neptune
- Amazon Quantum Ledger Database (QLDB)
- Amazon RDS on VMware

Amazon DynamoDB

Amazon DynamoDB is the NoSQL database service provided by amazon which is fully managed and automated.

NoSQL means that you don’t have to write queries to create a table or retrieve the data you can do so by some clicks to create a dynamic table which means you can add any amount of attributes, columns and store the data.

The main advantage of using this is that it is **fully managed** and it **automatically handles the traffic of data on the multiple servers and gives the optimum performance** and you don’t have to lookup for the underlying hardware, setup, configuration, scaling is all managed by the amazon. It also automatically backup and restore that provides the security of data.

Amazon DynamoDB

The feature provided by the amazon which is commendable is that according to your need of data or traffic **it automatically scale up and scale down you don’t have to look upon the underlying servers or their maintenance** and according to the usage you have you used it charges accordingly and also there are no minimum charges for the usage.

According to the AWS, DynamoDB can handle 20 trillion of the request per day and also can handle the peak of the traffic up to 20 million per second which is huge and commendable.

Because of its advantageous features, many large companies Lyft, Airbnb, and Redfin as well as enterprises such as Samsung, Toyota, and Capital One transferred their workloads to DynamoDB.

Amazon Aurora

Aurora is AWS’s **managed relational database engine**. It speaks **standard SQL** and is **compatible with MySQL and PostgreSQL**, so apps, drivers, and tools that work with those engines usually work with Aurora with little or no code change.

Amazon Aurora

Amazon Aurora supports MySQL and PostgreSQL which is a relational database engine provided by the Amazon web service.

It means that code, applications, and drivers are used in the database. You have to write proper queries for creating tables and storing data in it as you do in MySQL. Aurora charges minimal or no charges for using this. It is five times faster than the actual MySQL and three times faster than the actual PostgreSQL. It can auto-scale up to 64 TB per database instance.

There are certain features of MySQL that are not provided by the Aurora such as the MyISAM storage engine. Using extensions you can communicate between two or more databases of Aurora and can move your databases across aurora and your local engine which is also a feature of this.

Amazon Aurora

The main advantage is speed, security, and availability which is done by replicating data over three availability zones.

It provides the feature of **self-healing** i.e. it performs the automatic error scanning of your data and the blocks and also provides fault tolerance i.e it provides the ability to continue operating without interruption when their ifs fault in one or more component.

It also provides **autoscaling**, according to your database size it scales out and scales down you don’t have to worry about the servers or their charges just have to pay what you are using.

Amazon RDS on VMware

- Some organizations (banks, healthcare, government) **cannot move all data to the public cloud** due to **compliance or privacy regulations**.

- RDS on VMware allows them to **enjoy AWS automation + management features** while keeping data **physically in their own private data centers**.

- Provides **hybrid cloud capability** → same AWS RDS interface both on-premises and in AWS cloud.

Amazon RDS on VMware

- Some organizations (banks, healthcare, government) **cannot move all data to the public cloud** due to **compliance or privacy regulations**.

- RDS on VMware allows them to **enjoy AWS automation + management features** while keeping data **physically in their own private data centers**.

- Provides **hybrid cloud capability** → same AWS RDS interface both on-premises and in AWS cloud.

Amazon RDS on VMware

Amazon Relational Database Service (Amazon RDS) on VMware allows you to deploy the database on the VMware environment using the Amazon RDS technology.

It allows you to work on the same simple interface, same database with the same environment on-premise you will not face any type of difference in that. It enables you to set up, configure, and operate your relational database on the cloud within a few clicks

It also supports MySQL, PostgreSQL, and the Microsoft SQL server database. It is also fully managed and doesn’t have to look upon the provisioning, the configuration of the hardware or the managements all are automated.

Amazon RDS on VMware

For security purposes and the high availability features it also backups its data.

Having RDS on-site makes it very easy for you to operate and integrate Amazon RDS on VMware within your existing VMware vSphere private data centers.

There are some perquisites to use this: you will need to have Administrative privileges on the cluster to set up RDS on VMware. You will also need to have a second set of credentials for use by RDS on VMware And the hardware that you use for the RDS should be registered in the valid VMware Hardware Compatibility Guide.

Amazon Timestream

Amazon timestream is used to handle the time series data that assesses how events change over time.

The data which changes according to time which is used by the IoT and the other operational applications gather, maintain, and query with the help of Amazon Timestream.

It is a serverless database. Like other databases, it is also automated, fully managed and you don't have to care much about the maintenance and the hardware provisioning, setup, or configuration so that you can give much time to your work. It also provides the feature of autoscaling, ensuring that you never run out of space

Amazon Timestream

Amazon timestream is used to handle the time series data that assesses how events change over time.

The data which changes according to time which is used by the IoT and the other operational applications gather, maintain, and query with the help of Amazon Timestream.

It is a serverless database. Like other databases, it is also automated, fully managed and you don’t have to care much about the maintenance and the hardware provisioning, setup, or configuration so that you can give much time to your work. It also provides the feature of autoscaling, ensuring that you never run out of space

Amazon Timestream

Amazon Timestream is optimized to assess, query, and store timestream data by storing data is set of time intervals, ranging from milliseconds, microseconds, and even nanoseconds.

Other features that provide the ability to automatically configure retention, tiering, and data compression. It is much faster than the other relational databases. It can handle the trillions of the request every single day which helps in the reduction of the cost and also it is much faster than the other databases.

All the features combined provided the cost reduction. **The data stored by the time stream database can later be used for the business intelligence tools and machine learning services.**

Amazon Neptune

Amazon Neptune supports the graph database. A graph database is used for the data which are connected, which are correlated to each other and have some relationships between them.

It is different from the SQL database. It feels like a NoSQL database with no query, no table only two entities i.e vertices and edges.

The main purpose of developing a graph database is network security, fraud detection, to understand the drug discovery, and many more.

Use cases for such highly-connected data include social networking, restaurant recommendations, etc.

Amazon Neptune

Amazon Neptune supports the graph database. A graph database is used for the data which are connected, which are correlated to each other and have some relationships between them.

It is different from the SQL database. It feels like a NoSQL database with no query, no table only two entities i.e vertices and edges.

The main purpose of developing a graph database is network security, fraud detection, to understand the drug discovery, and many more.

Use cases for such highly-connected data include social networking, restaurant recommendations, etc.

Amazon Neptune

The features provided by the Neptune is high-performance data graph and maintaining billions of the relationship of data with very ease and with very low latency.

To make the data available it continuously replicates all its data to amazon s3 and other availability zones for the security purpose also and for higher availability also.

It is also fully managed, the user doesn’t have to look upon the underlying hardware or provisioning and also provides the feature of autoscaling so that you will never run out of space and pay only for what you use.

It was developed to address the limitation of the relational database and to solve or make work more efficient for the complex data.

Amazon Quantum Ledger Database (QLDB)

Amazon Quantum Ledge Database (QLDB)was introduced alongside the Amazon-Managed Blockchain service.

It is also a fully managed and serverless database but the main function that provides is that it is a ledger database which means it is used for recording or storing the financial and economic data of an organization over some time.

It allows you to maintain the complete history of accounting and transactional data between multiple parties in an immutable, transparent, and cryptographic way through the use of the cryptographic algorithm SHA-256, making it highly secure.

It is serverless so the cost of the underlying hardware is also reduced.

Amazon Quantum Ledger Database (QLDB)

Amazon Quantum Ledge Database (QLDB)was introduced alongside the Amazon-Managed Blockchain service.

It is also a fully managed and serverless database but the main function that provides is that it is a ledger database which means it is used for recording or storing the financial and economic data of an organization over some time.

It allows you to maintain the complete history of accounting and transactional data between multiple parties in an immutable, transparent, and cryptographic way through the use of the cryptographic algorithm SHA-256, making it highly secure.

It is serverless so the cost of the underlying hardware is also reduced.

Amazon Quantum Ledger Database (QLDB)

It’s a new type of database which is provided by the amazon which focuses on the ledger so that you can focus more on the development of data rather than its managing technique.

It keeps the record at one place which is easy to retrieve and work upon and can focus more on analyzing and solving issues. it has been adopted by many of the large enterprises and businesses such as Wipro, Splunk, sage, etc.

It also cryptographically verify the data integrity as a security feature

Amazon RDS (Relational Database Service)

Amazon Relational Database Service (RDS) is a managed database service provided by Amazon Web Services (AWS).

It makes it easy to set up and operate a scalable relational database in the AWS cloud. Amazon RDS supports an array of database engines to store and organize data.

It also takes care of many database management and administration tasks, such as data migration, backup, recovery and patching, freeing up administrators to focus on other higher value tasks.

Amazon RDS facilitates the deployment and maintenance of relational databases in the AWS cloud. It can be used to set up, operate, manage and scale a relational instance of a cloud database.

Amazon RDS (Relational Database Service)

Amazon Relational Database Service (RDS) is a managed database service provided by Amazon Web Services (AWS).

It makes it easy to set up and operate a scalable relational database in the AWS cloud. Amazon RDS supports an array of database engines to store and organize data.

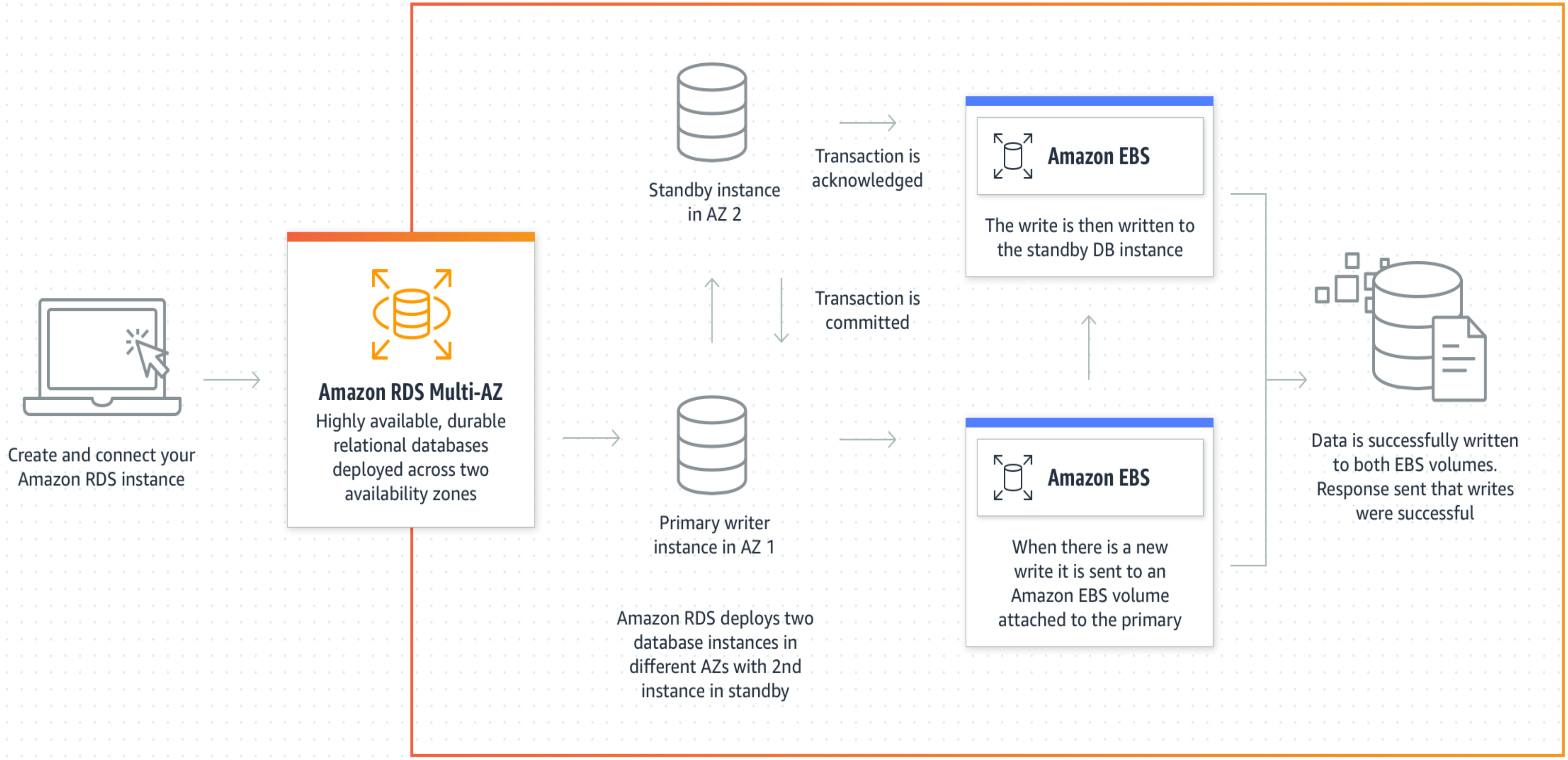
It also takes care of many database management and administration tasks, such as data migration, backup, recovery and patching, freeing up administrators to focus on other higher value tasks.

Amazon RDS facilitates the deployment and maintenance of relational databases in the AWS cloud. It can be used to set up, operate, manage and scale a relational instance of a cloud database.

RDS Multi-AZ Deployments

In an Amazon RDS Multi-AZ deployment, Amazon RDS **automatically creates a primary database (DB) instance and synchronously replicates the data to an instance in a different AZ.**

When it detects a failure, Amazon RDS automatically fails over to a standby instance without manual intervention.



Create and connect your Amazon RDS instance

Amazon RDS Multi-AZ

Highly available, durable relational databases deployed across two availability zones

Standby instance in AZ 2

Transaction is acknowledged

Transaction is committed

Primary writer instance in AZ 1

Amazon RDS deploys two database instances in different AZs with 2nd instance in standby

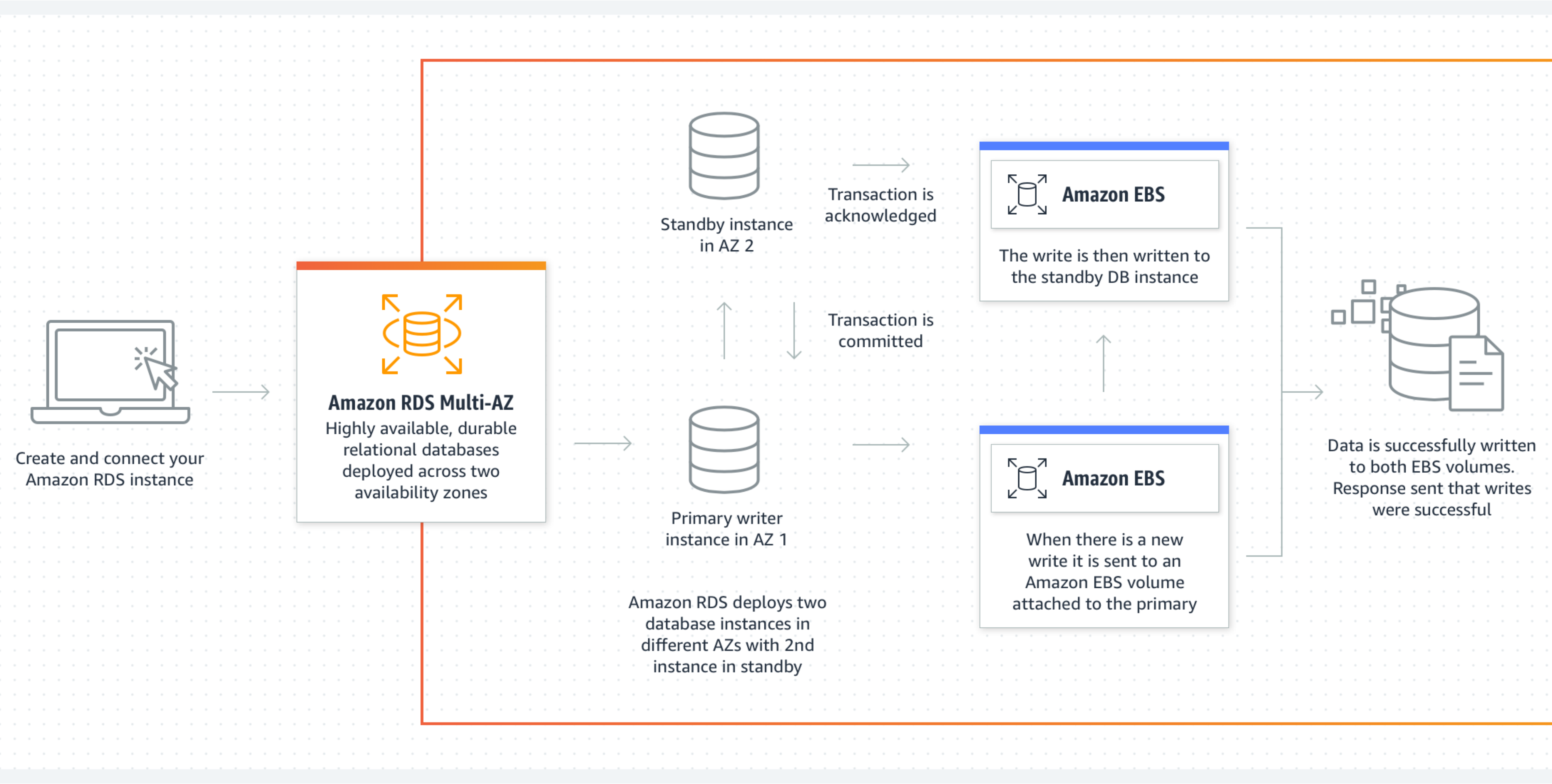
Amazon EBS

The write is then written to the standby DB instance

Amazon EBS

When there is a new write it is sent to an Amazon EBS volume attached to the primary

Data is successfully written to both EBS volumes. Response sent that writes were successful



Create and connect your Amazon RDS instance

Amazon RDS Multi-AZ

Highly available, durable relational databases deployed across two availability zones

Standby instance in AZ 2

Transaction is acknowledged

Transaction is committed

Primary writer instance in AZ 1

Amazon RDS deploys two database instances in different AZs with 2nd instance in standby

Amazon EBS

The write is then written to the standby DB instance

Amazon EBS

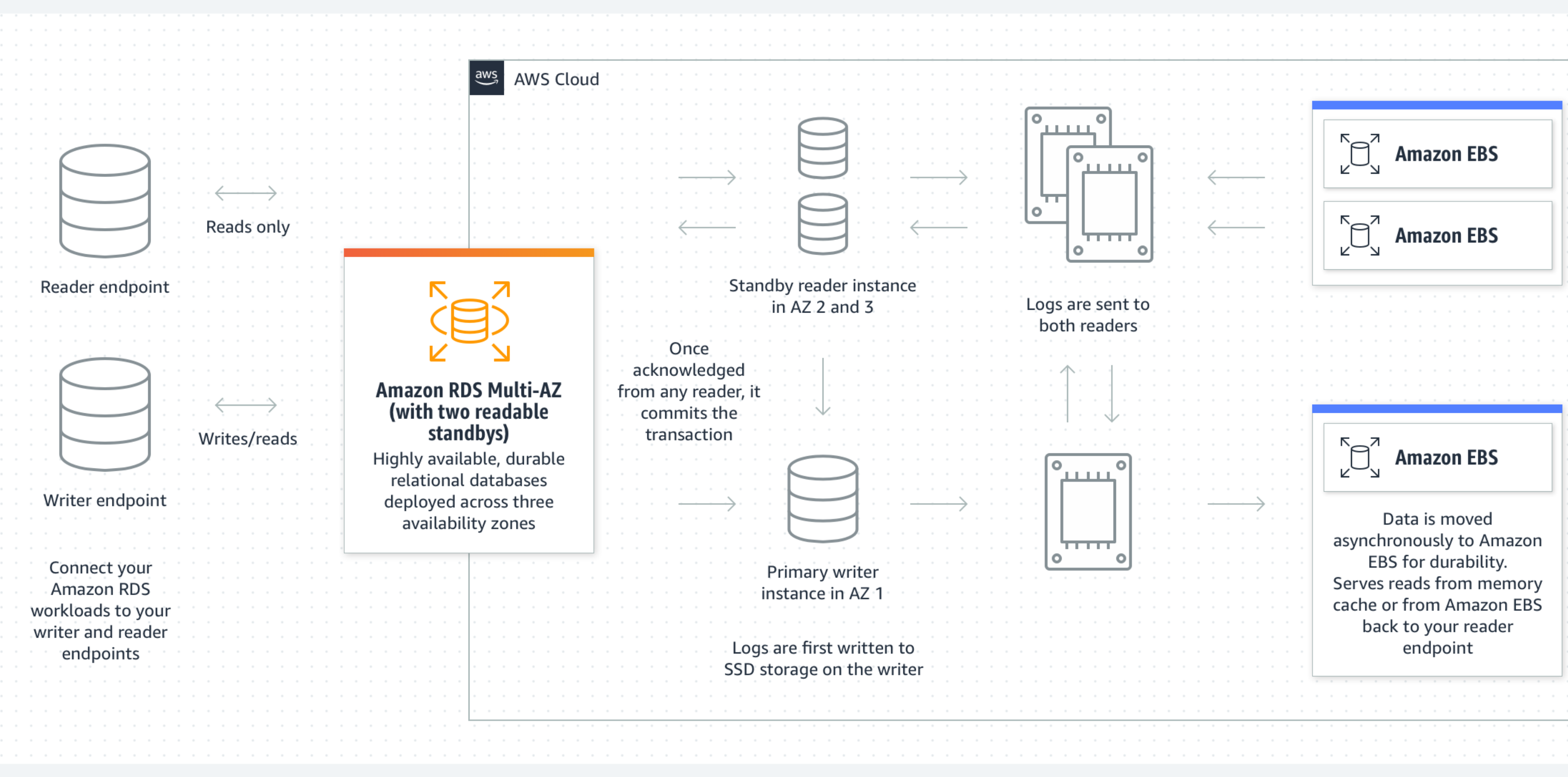
When there is a new write it is sent to an Amazon EBS volume attached to the primary

Data is successfully written to both EBS volumes. Response sent that writes were successful

Amazon RDS Multi-AZ with two readable standbys

Deploy highly available, durable MySQL or PostgreSQL databases in three AZs using Amazon RDS Multi-AZ with two readable standbys.

Gain automatic failovers in typically under 35 seconds, up to 2x faster transaction commit latency compared to Amazon RDS Multi-AZ with one standby, additional read capacity, and a choice of AWS Graviton2– or Intel–based instances for compute.



Reader endpoint

Reads only

Writer endpoint

Writes/reads

Connect your Amazon RDS workloads to your writer and reader endpoints

AWS Cloud

Amazon RDS Multi-AZ (with two readable standbys)

Highly available, durable relational databases deployed across three availability zones

Standby reader instance in AZ 2 and 3

Once acknowledged from any reader, it commits the transaction

Primary writer instance in AZ 1

Logs are first written to SSD storage on the writer

Logs are sent to both readers

Amazon EBS

Amazon EBS

Amazon EBS

Data is moved asynchronously to Amazon EBS for durability. Serves reads from memory cache or from Amazon EBS back to your reader endpoint

DynamoDB Streams

DynamoDB Streams captures a time-ordered sequence of item-level modifications in any DynamoDB table and stores this information in a log for up to 24 hours.

Applications can access this log and view the data items as they appeared before and after they were modified, in near-real time.

A DynamoDB stream is an ordered flow of information about changes to items in a DynamoDB table.

When you enable a stream on a table, DynamoDB captures information about every modification to data items in the table.

DynamoDB Streams

DynamoDB Streams captures a time-ordered sequence of item-level modifications in any DynamoDB table and stores this information in a log for up to 24 hours.

Applications can access this log and view the data items as they appeared before and after they were modified, in near-real time.

A DynamoDB stream is an ordered flow of information about changes to items in a DynamoDB table.

When you enable a stream on a table, DynamoDB captures information about every modification to data items in the table.

DynamoDB Streams

Whenever an application creates, updates, or deletes items in the table, DynamoDB Streams writes a stream record with the primary key attributes of the items that were modified. A stream record contains information about a data modification to a single item in a DynamoDB table.

DynamoDB Streams helps ensure the following:

- Each stream record appears exactly once in the stream.

- For each item that is modified in a DynamoDB table, the stream records appear in the same sequence as the actual modifications to the item.

DynamoDB Streams writes stream records in near-real time so that you can build applications that consume these streams and take action based on the contents.