

Amazon RDS

What is Amazon RDS?

- Amazon Relational Database Service (Amazon RDS) is a web service that makes it easier to set up, operate, and scale a relational database in the AWS Cloud
- It provides cost-efficient, resizable capacity for an industry-standard relational database and manages common database administration tasks.

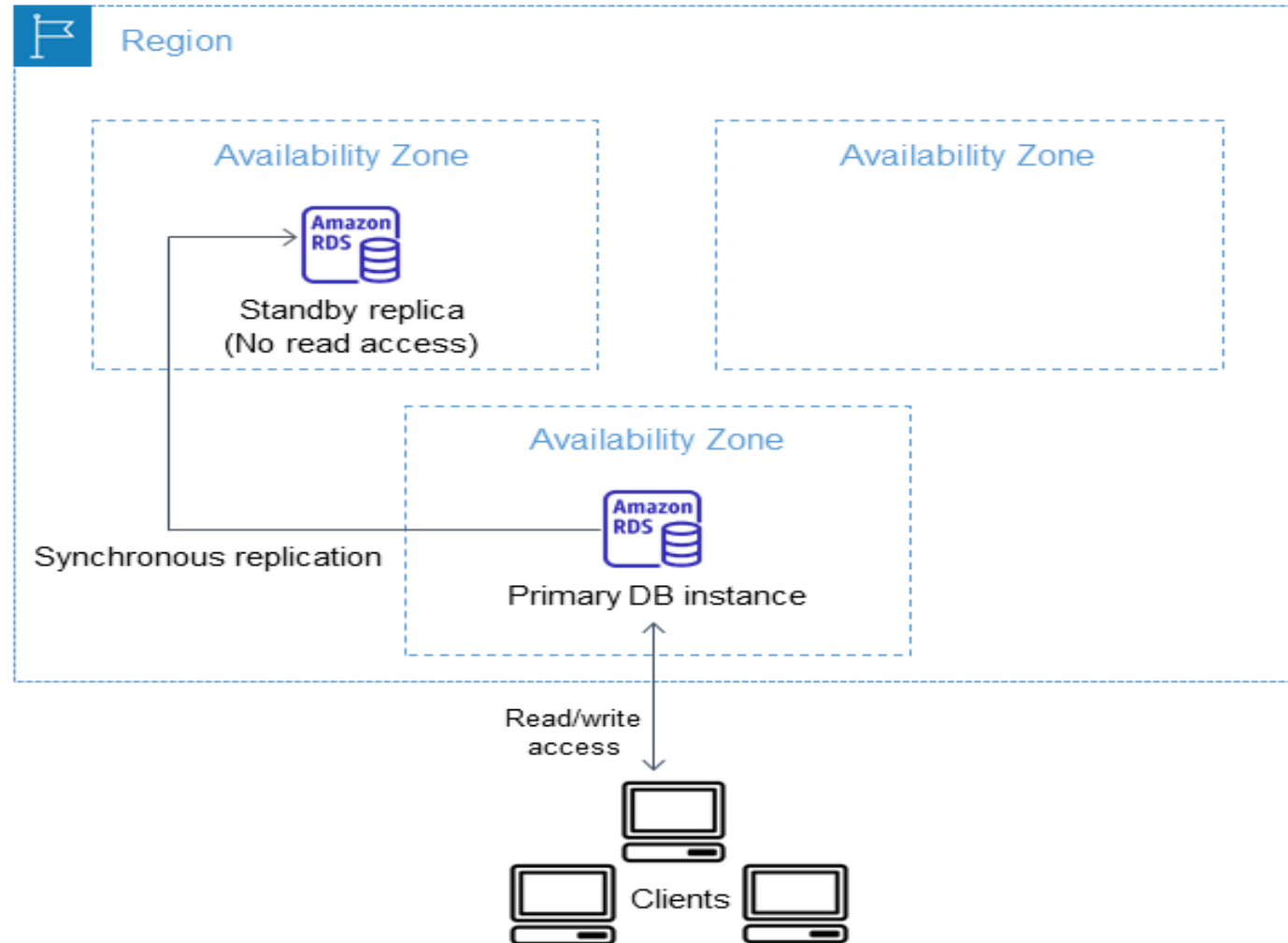
Comparison with On-Prem and RDS

Feature	On-Premise Management	Amazon EC2 Management	Amazon RDS Management
Application optimization	Customer	Customer	Customer
Scaling	Customer	Customer	AWS
High availability	Customer	Customer	AWS
Database backups	Customer	Customer	AWS
Database software patching	Customer	Customer	AWS
Database software install	Customer	Customer	AWS
Operating system (OS) patching	Customer	Customer	AWS
OS installation	Customer	Customer	AWS
Server maintenance	Customer	AWS	AWS
Hardware lifecycle	Customer	AWS	AWS
Power, network, and cooling	Customer	AWS	AWS

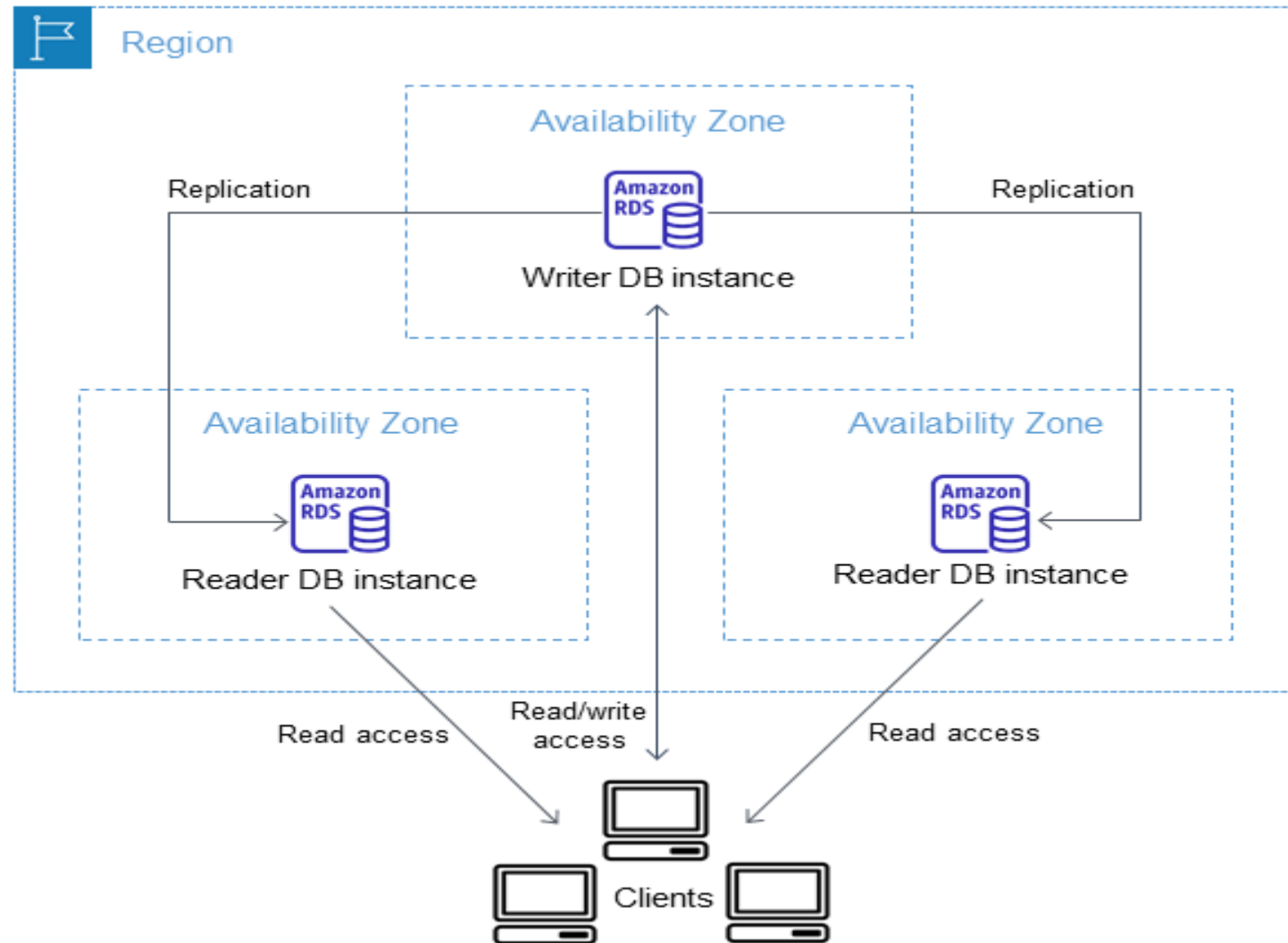
Database Engines Supported

- Db2
- MariaDB
- Microsoft SQL Server
- MySQL
- Oracle
- PostgreSQL

Multi-AZ Deployment – DB Instance



Multi-AZ Deployment – DB Cluster



Read Replicas

- RDS supports maximum 5 read replicas
- Aurora supports maximum 15 read replicase

Amazon Aurora

- Amazon Aurora (Aurora) is a fully managed relational database engine that's compatible with MySQL and PostgreSQL
- Aurora includes a high-performance storage subsystem
- Its MySQL- and PostgreSQL-compatible database engines are customized to take advantage of that fast distributed storage
- The underlying storage grows automatically as needed
- An Aurora cluster volume can grow to a maximum size of 128 terabytes (TiB)
- Aurora also automates and standardizes database clustering and replication, which are typically among the most challenging aspects of database configuration and administration

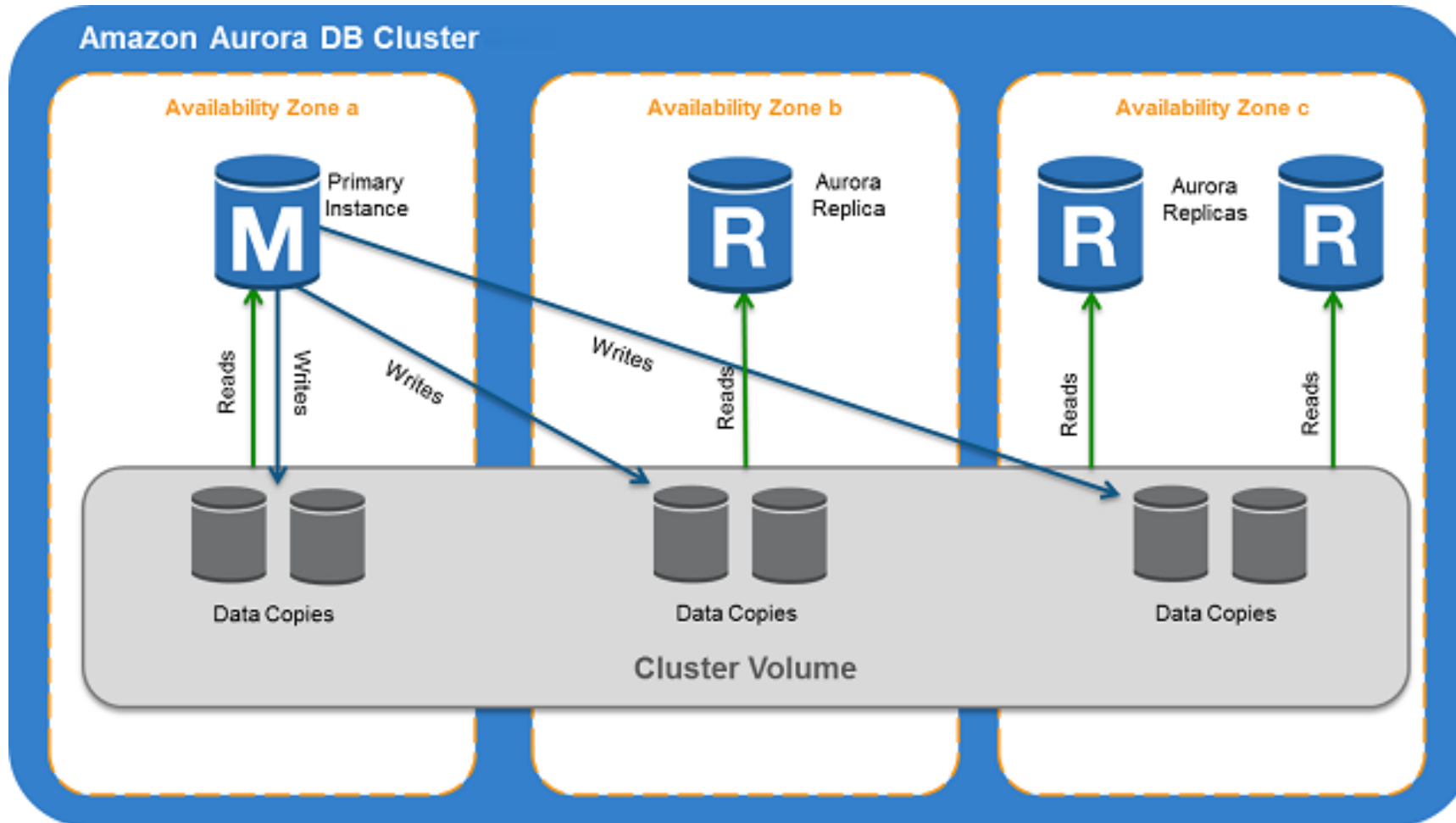
Aurora DB Clusters

- An Amazon Aurora *DB cluster* consists of one or more DB instances and a cluster volume that manages the data for those DB instances
- An Aurora *cluster volume* is a virtual database storage volume that spans multiple Availability Zones, with each Availability Zone having a copy of the DB cluster data
- Two types of DB instances make up an Aurora DB cluster
 - **Primary DB instance** – Supports read and write operations, and performs all of the data modifications to the cluster volume. Each Aurora DB cluster has one primary DB instance

Aurora DB Clusters

- Two types of DB instances make up an Aurora DB cluster
 - **Aurora Replica** – Connects to the same storage volume as the primary DB instance and supports only read operations. Each Aurora DB cluster can have up to 15 Aurora Replicas in addition to the primary DB instance. Maintain high availability by locating Aurora Replicas in separate Availability Zones. Aurora automatically fails over to an Aurora Replica in case the primary DB instance becomes unavailable. You can specify the failover priority for Aurora Replicas. Aurora Replicas can also offload read workloads from the primary DB instance

Aurora DB Clusters



Fault Tolerance for Aurora DB Clusters

- An Aurora DB cluster is fault tolerant by design.
- The cluster volume spans multiple Availability Zones (AZs) in a single AWS Region, and each Availability Zone contains a copy of the cluster volume data.
- This functionality means that your DB cluster can tolerate a failure of an Availability Zone without any loss of data and only a brief interruption of service.

Fault Tolerance for Aurora DB Clusters

- If the primary instance in a DB cluster fails, Aurora automatically fails over to a new primary instance in one of two ways:
 - By promoting an existing Aurora Replica to the new primary instance
 - By creating a new primary instance
- You can customize the order in which your Aurora Replicas are promoted to the primary instance after a failure by assigning each replica a priority.
- Priorities range from 0 for the highest priority to 15 for the lowest priority.
- If the primary instance fails, Amazon RDS promotes the Aurora Replica with the highest priority to the new primary instance.

Fault Tolerance for Aurora DB Clusters

- More than one Aurora Replica can share the same priority, resulting in promotion tiers.
- If two or more Aurora Replicas share the same priority, then Amazon RDS promotes the replica that is largest in size.
- If two or more Aurora Replicas share the same priority and size, then Amazon RDS promotes an arbitrary replica in the same promotion tier.
- The service is typically restored in less than 60 seconds, and often less than 30 seconds.
- If the DB cluster doesn't contain any Aurora Replicas, then the primary instance is recreated in the same AZ during a failure event.
- Service is restored when the new primary instance is created, which typically takes less than 10 minutes.

Aurora Global Database

- Amazon Aurora global databases span multiple AWS Regions, enabling low latency global reads and providing fast recovery from the rare outage that might affect an entire AWS Region
- An Aurora global database has a primary DB cluster in one Region, and up to five secondary DB clusters in different Regions
- An Aurora global database consists of one *primary* AWS Region where your data is written, and up to five read-only *secondary* AWS Regions
- You issue write operations directly to the primary DB cluster in the primary AWS Region. Aurora replicates data to the secondary AWS Regions using dedicated infrastructure, with latency typically under a second

Aurora Global Database

