Designing Applications and Architectures in AWS

(Aligned with AWS Solution Architect Associate Certification)

simplilearn

TECHNOLOGY

Databases

```
#include <iostream>
#include <cstdlib>
using namespace std;
int main()
{
    cout << "Hello, world!" << endl;
    return 0;
}
```

A Day in the Life of Cloud Architect

You are a cloud architect in an organization and have been asked to develop a comprehensive database solution that seamlessly integrates various components for an upcoming product launch.

The marketing team needs a MySQL database to store all the customer feedback. The operations team requires a DynamoDB database for real-time product tracking across multiple regions, utilizing DynamoDB global table replicas for swift data access.

Simultaneously, the analytics team intends to leverage the Redshift database to query large datasets for insights and trends.

To achieve this, you will learn a few concepts in this lesson that will help you find a solution for the given scenario.

Learning Objectives

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

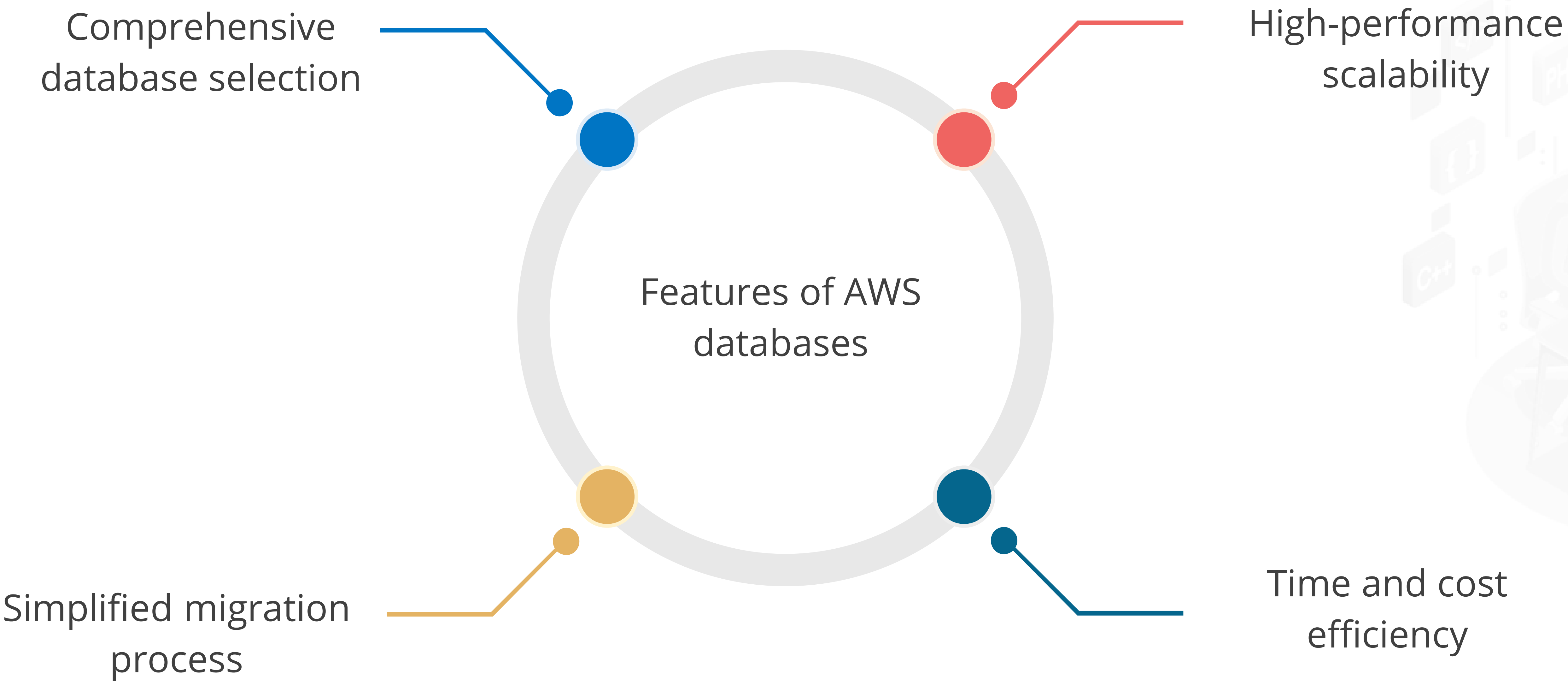
- Create and manage a MySQL database for efficient data storage and retrieval
- Set up and utilize a DynamoDB database to handle NoSQL data and achieve scalability
- Establish DynamoDB global table replicas for data redundancy, ensuring high availability and fault tolerance
- Utilize Redshift for data warehousing and perform analytical queries to extract valuable insights from your data

TECHNOLOGY

Introduction to Databases

Databases: Overview

AWS provides a broad selection of purpose-built databases to help save, grow, and build a robust cloud database.



Features of AWS databases

Comprehensive database selection

High-performance scalability

Simplified migration process

Time and cost efficiency

Types of AWS Databases

AWS databases can be categorized into:

Relational

For traditional applications, ERP, CRM, and e-commerce

Key-value

For high-traffic web applications, e-commerce, and gaming applications

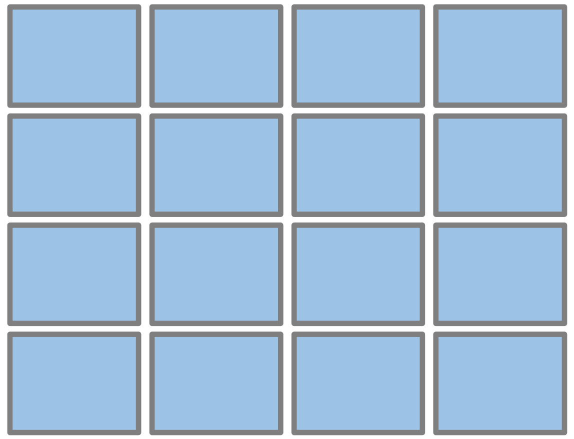
In-memory

For caching, session management, and geospatial applications

Relational Databases

Relational databases, also known as SQL databases, are the most common type of databases.

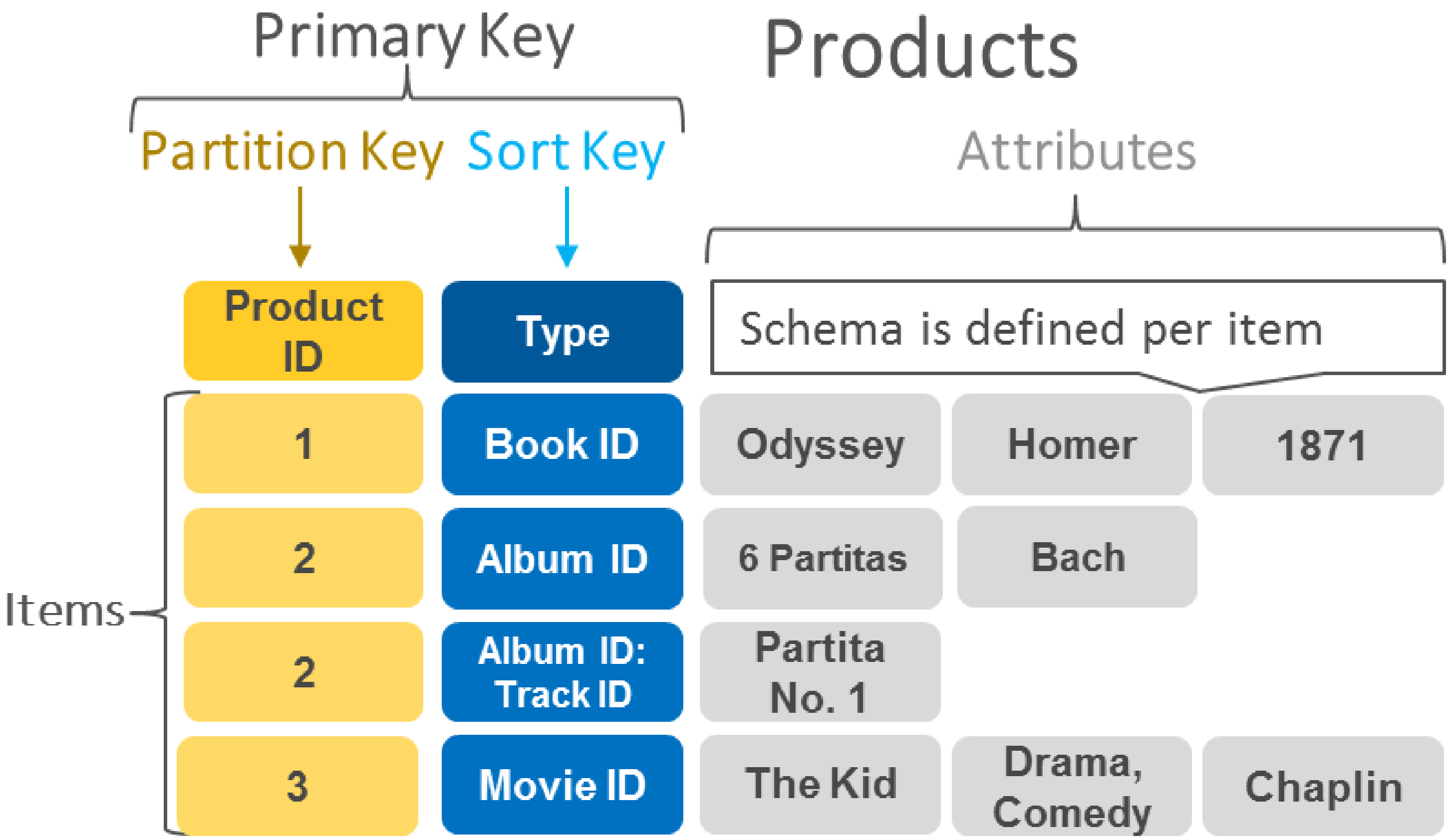
Structured data



They are used to store structured data in a set of tables that are formally described.

Key-Value Databases

Key-value databases are a type of non-relational database that utilizes a key-value method for storing data. These databases are used in high-traffic and high-computing applications.



Products

Primary Key

Partition Key Sort Key

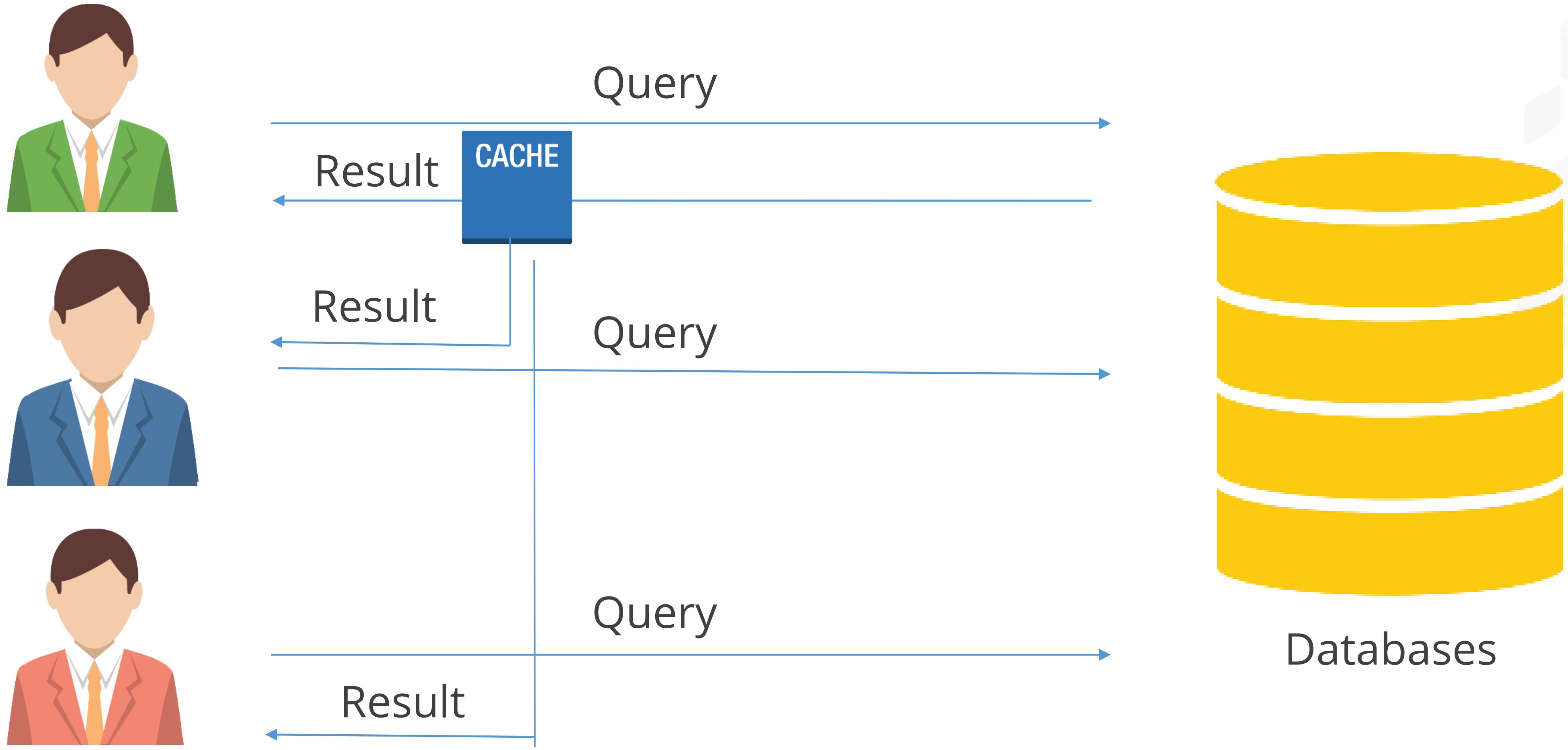
Attributes

Items

Product ID	Type	Schema is defined per item		
1	Book ID	Odyssey	Homer	1871
2	Album ID	6 Partitas	Bach	
2	Album ID: Track ID	Partita No. 1		
3	Movie ID	The Kid	Drama, Comedy	Chaplin

In-Memory Databases

In-memory databases use cache-based storage to reduce the database infrastructure load and enhance user response time.



Amazon Relational Database Service (RDS)

Amazon RDS

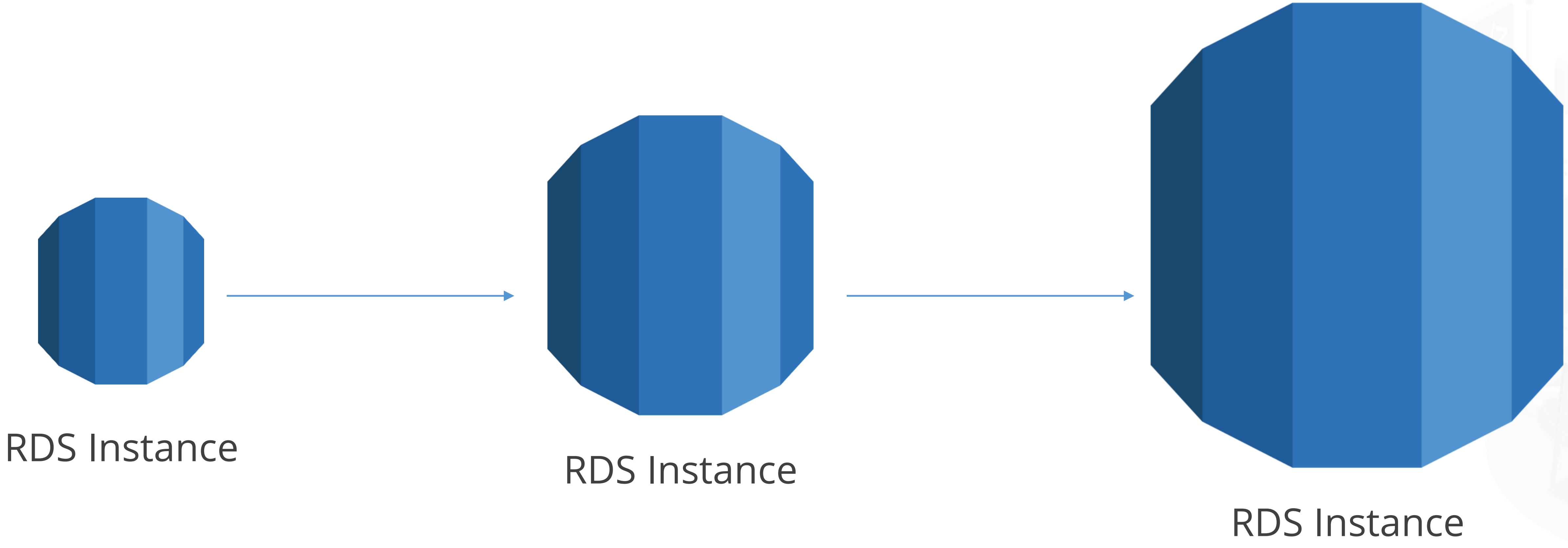
Amazon Relational Database Service (Amazon RDS) simplifies the process of setting up, operating, and scaling a relational database in the AWS cloud.

Here are some of the benefits of Amazon RDS:

- Ease of administration
- Scalability
- Security
- Cost-effectiveness
- Fast performance

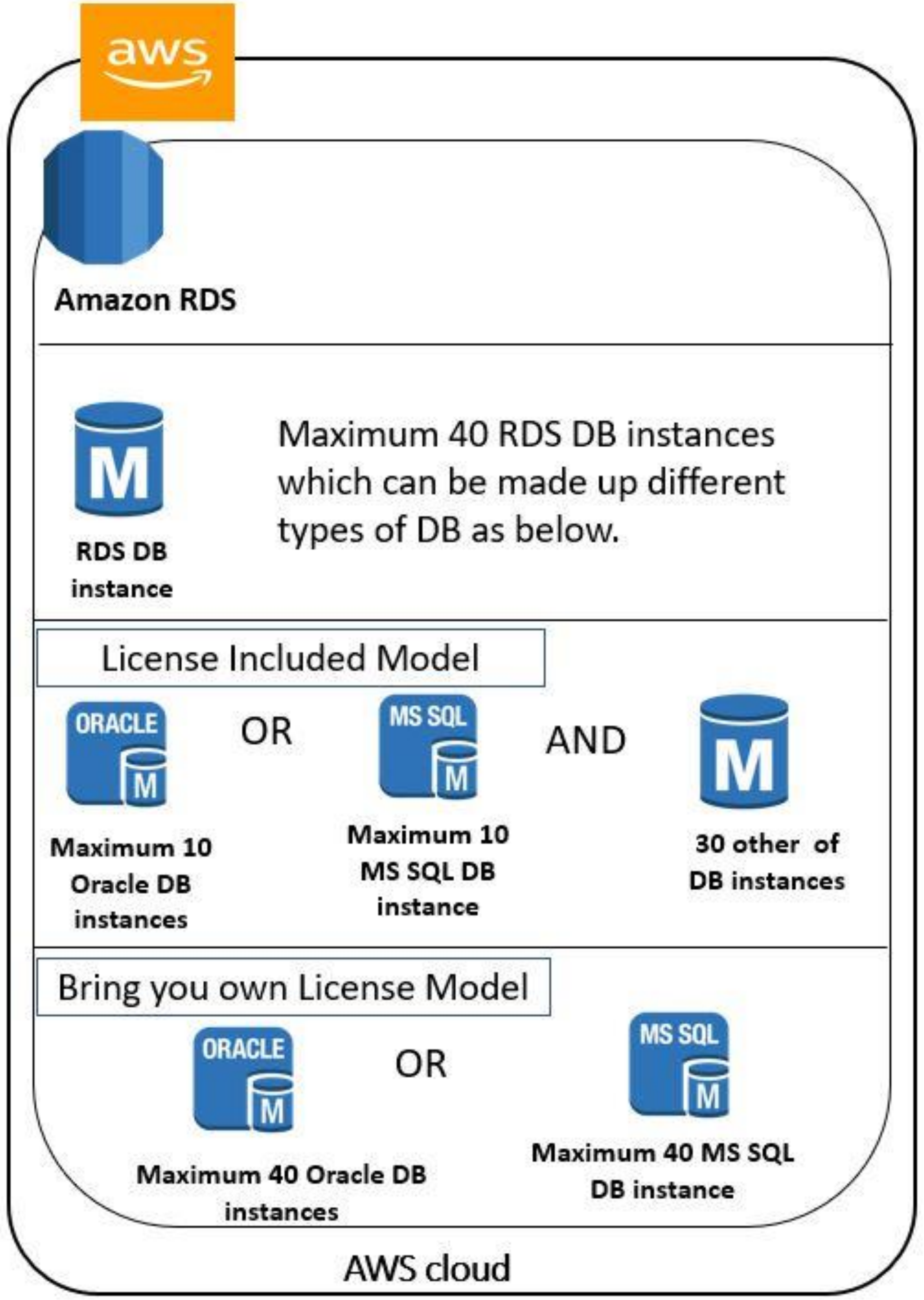
Amazon RDS

Amazon RDS enables users to scale the compute and memory resources of their deployments up to 32 vCPUs and 244 GiB of RAM.



Scaling operations usually finish within minutes.

DB Instance


aws
Amazon RDS
M
RDS DB instance
Maximum 40 RDS DB instances which can be made up different types of DB as below.
License Included Model
ORACLE
M
OR
MS SQL
M
AND
M
Maximum 10 Oracle DB instances
Maximum 10 MS SQL DB instance
30 other of DB instances
Bring you own License Model
ORACLE
M
OR
MS SQL
M
Maximum 40 Oracle DB instances
Maximum 40 MS SQL DB instance
AWS cloud

- A DB instance, the fundamental building block of Amazon RDS, is a standalone database environment that runs in the Cloud.
- Users can create multiple databases within a DB instance and access them using standard client tools and applications.

RDS classes

DB instance classes determine the computation and memory capacity of an Amazon RDS DB instance.

They can be categorized into three types:

- General-purpose
- Memory-optimized
- Burstable performance

Amazon RDS Database Engines

Amazon RDS supports several database engines:

Amazon Aurora

PostgreSQL

ORACLE®

MySQL®

Microsoft® SQL Server®

MariaDB

Limitations of Amazon RDS for Microsoft SQL

Certain server-level roles in Microsoft SQL are currently unavailable in Amazon RDS. These include:

Microsoft® SQL Server®

- bulkadmin
- dbcreator
- diskadmin
- securityadmin
- serveradmin
- sysadmin

Limitations of Amazon RDS for MySQL

Several features of MySQL are currently unavailable in Amazon RDS, such as:

- Authentication plugin
- Error logging to the system log
- Group replication plugin
- InnoDB tablespace encryption
- Password strength plugin
- Persisted system variables
- Replication filters
- Semisynchronous replication
- Transportable tablespace
- X Plugin

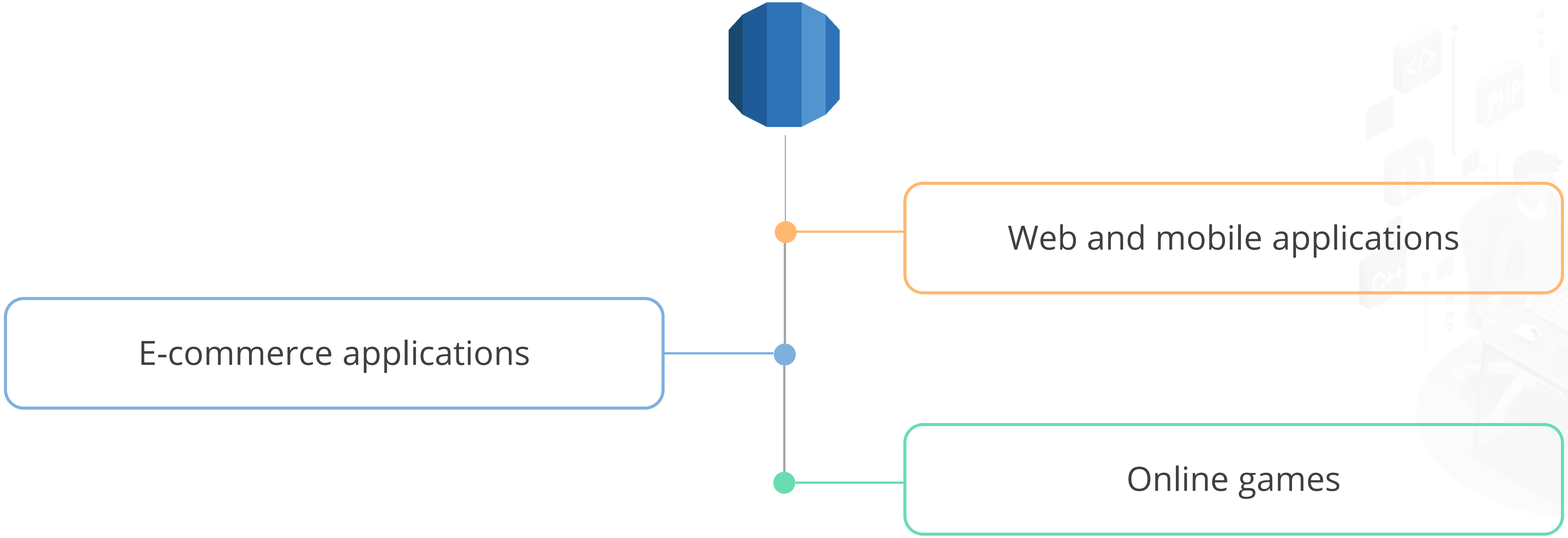
Limitations of Amazon RDS for MySQL

Certain Oracle privileges are currently unavailable in Amazon RDS, including:

- Alter database
- Alter system
- Create any directory
- Drop any directory
- Grant any privilege
- Grant any role

Use Cases of Amazon RDS

Amazon RDS is commonly used for:



Web and mobile applications

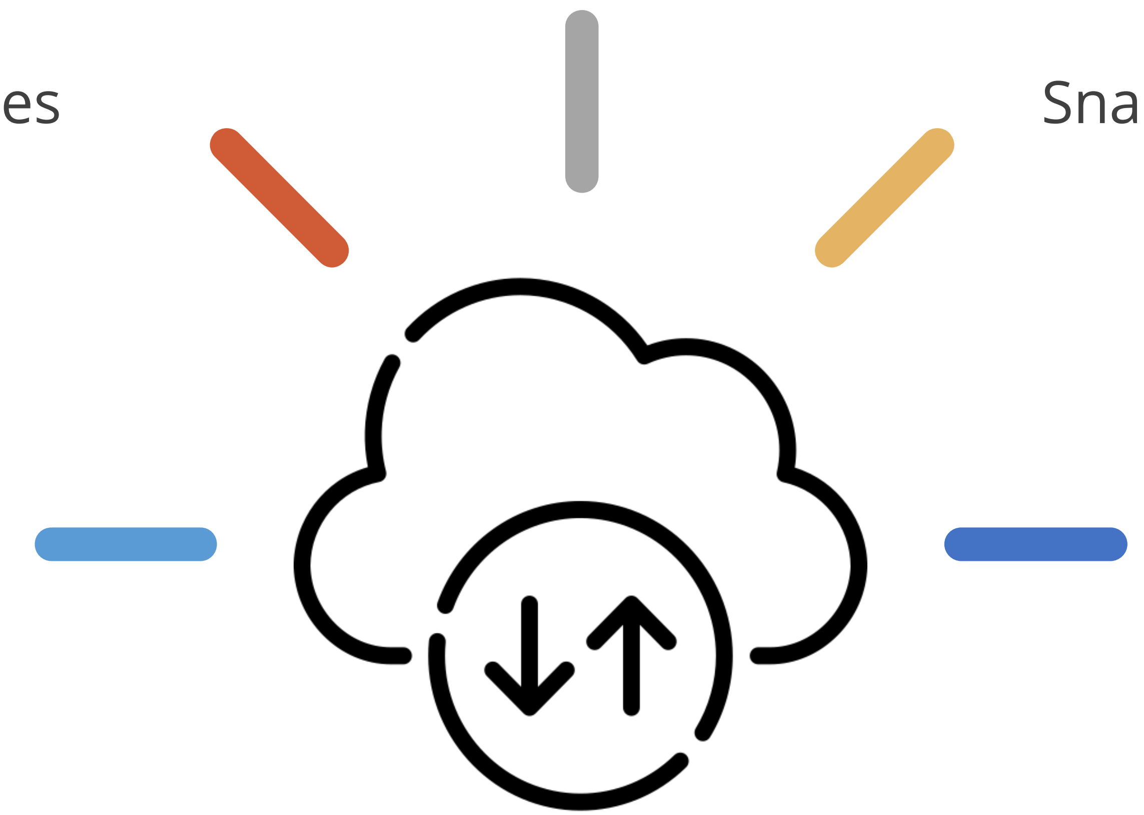
E-commerce applications

Online games

Amazon RDS Backups

Amazon RDS automatically creates and securely stores backups of DB instances in Amazon S3 for a specified retention period.

The types of Amazon RDS backups are:



Database snapshots

Point-in-time restores

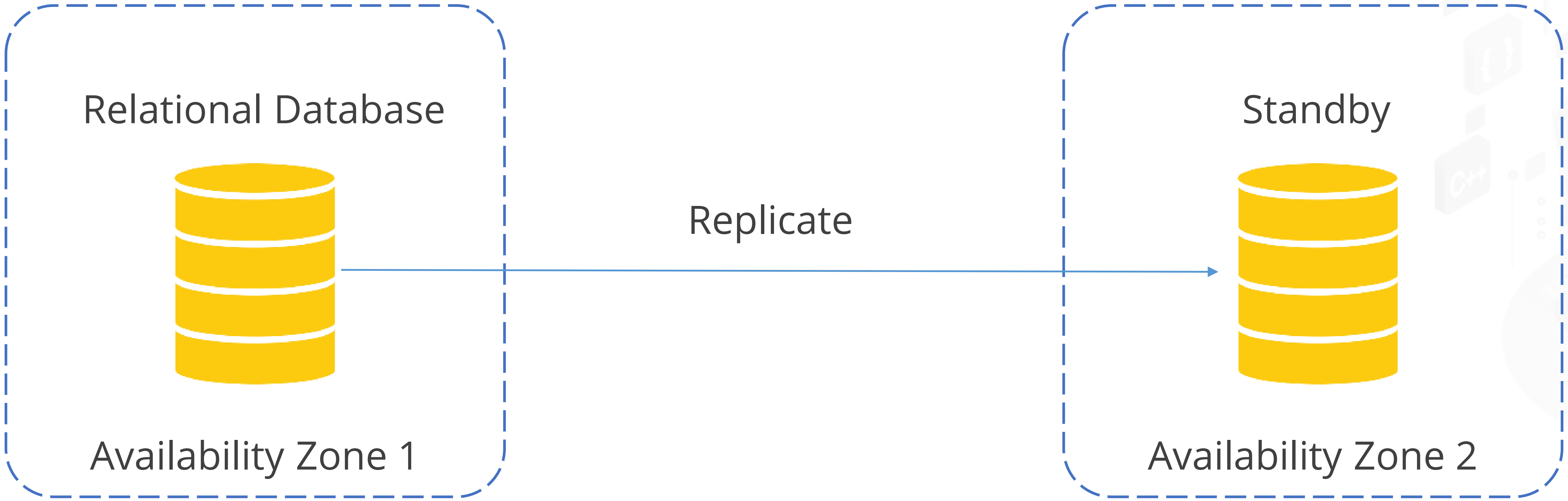
Snapshot copies

Automated backups

Snapshot sharing

Multi-Availability Zone Deployments

Multi-Availability Zone deployments synchronously replicate data to a standby instance in a different Availability Zone.

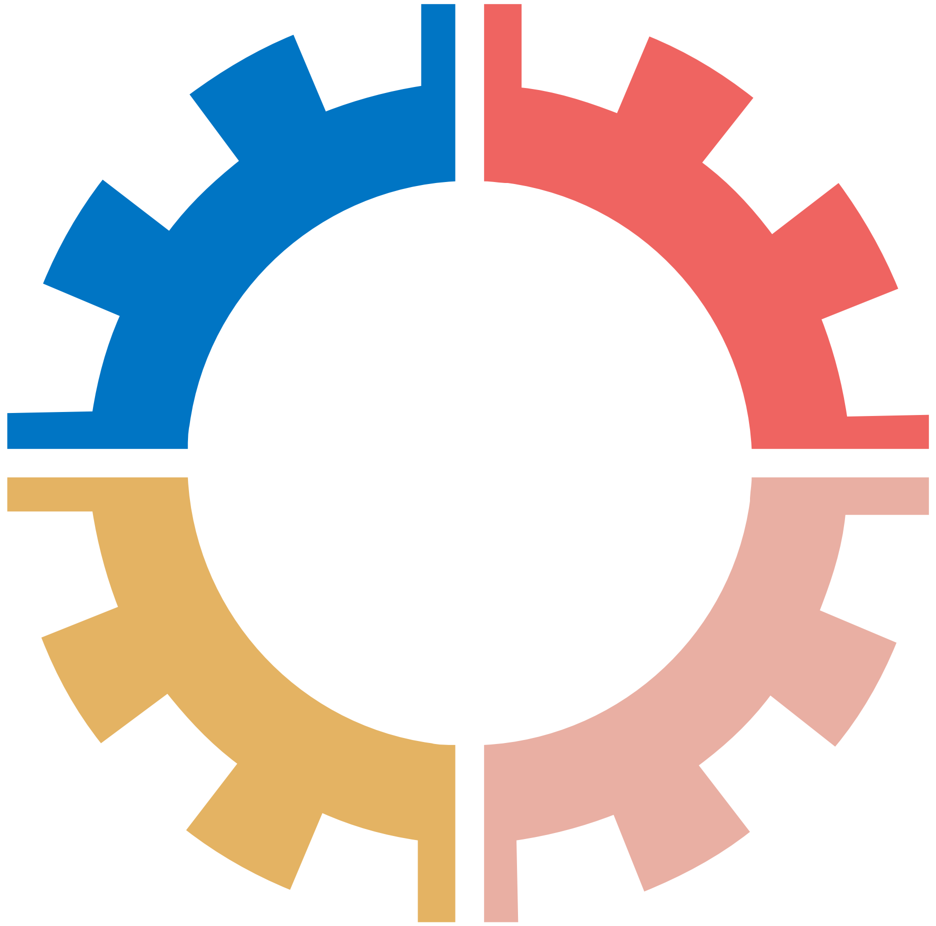


Multi-Availability Zone Deployments

Here are some benefits of Multi-Availability Zone deployments:

Enhanced durability:
The replica of the data stays up-to-date with the primary database.

Increased availability:
If an Availability Zone fails, the backup is ready in less than 30 seconds.

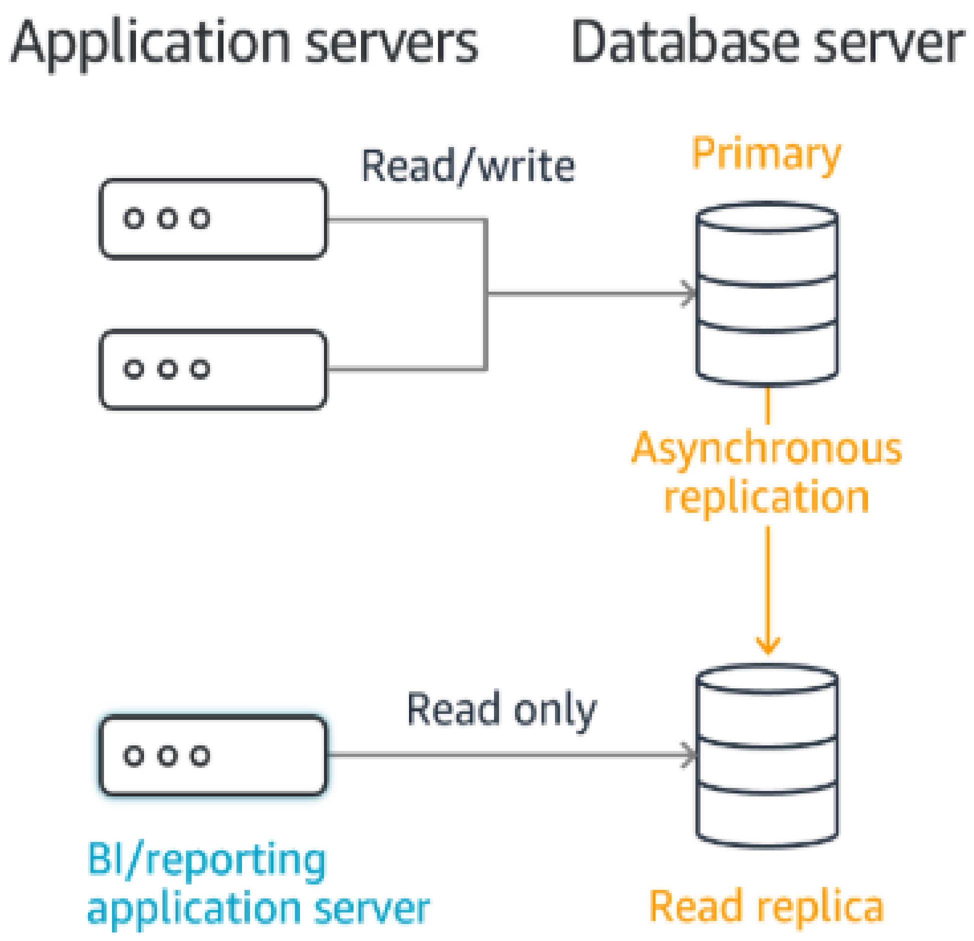


Stable performance:
Multi-AZ deployments do not suspend I/O activity on your primary during backup.

Automatic failover:
Amazon RDS automatically initiates a failover to the up-to-date standby.

Read Replicas

Amazon RDS Read Replicas enhance the performance and durability of RDS database (DB) instances.



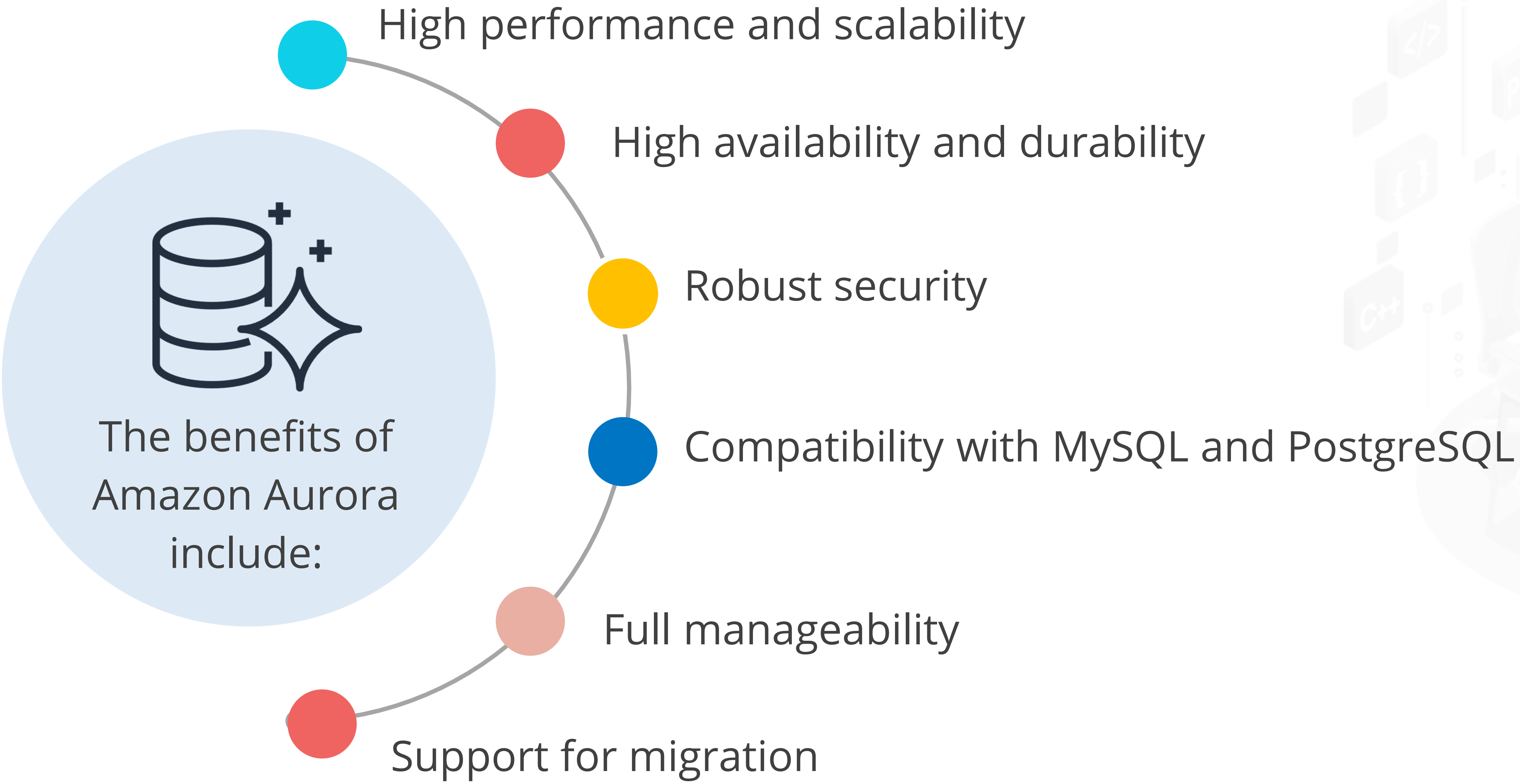
Here are some benefits of Read Replicas:

- Enhanced performance
- Increased availability
- Designed for security

Source: https://aws.amazon.com/rds/features/read-replicas/

Amazon Aurora

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

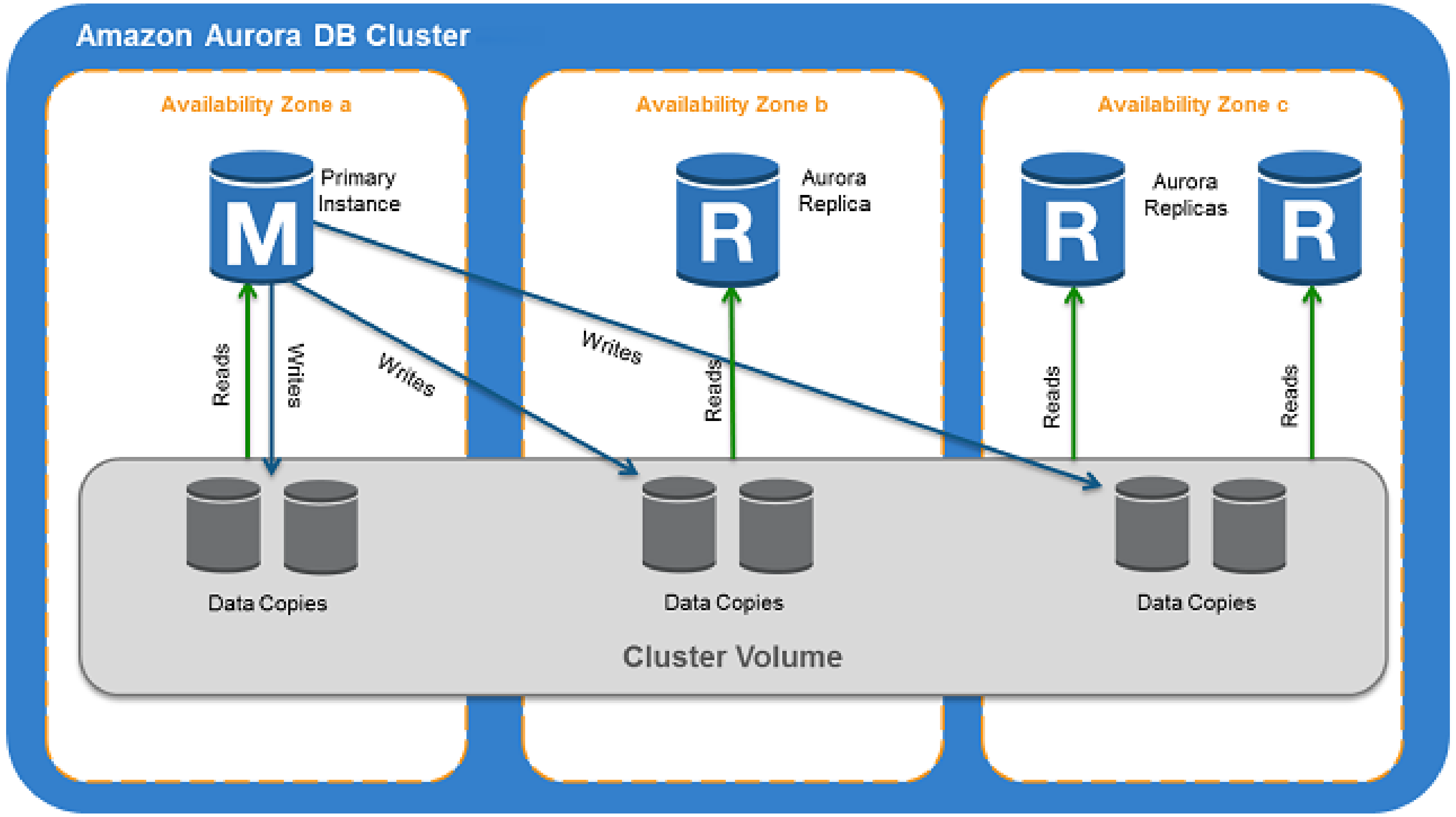


The benefits of Amazon Aurora include:

- High performance and scalability
- High availability and durability
- Robust security
- Compatibility with MySQL and PostgreSQL
- Full manageability
- Support for migration

Amazon Aurora DB Cluster

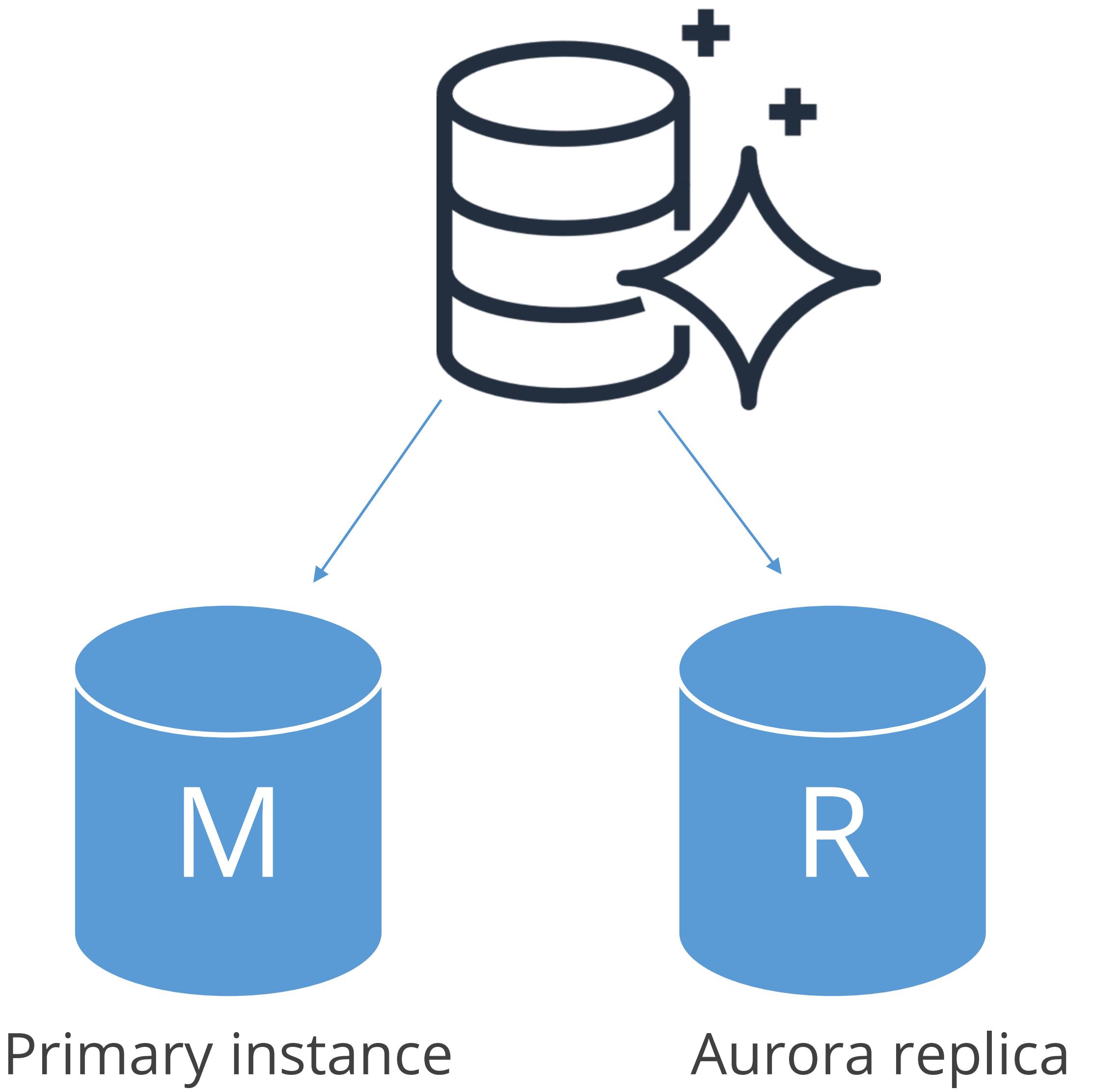
An Amazon Aurora DB cluster includes one or more DB instances and a cluster volume that manages the data for those DB instances.



Source: https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Overview.html

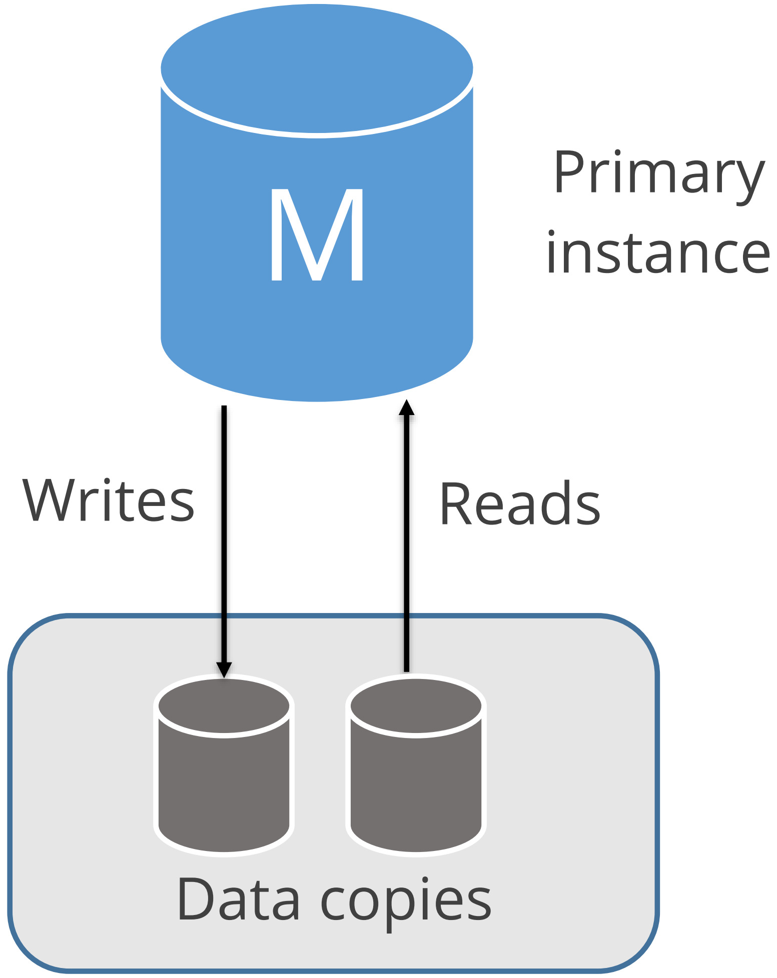
Amazon Aurora DB Cluster

An Aurora DB Cluster in Amazon consists of two types:



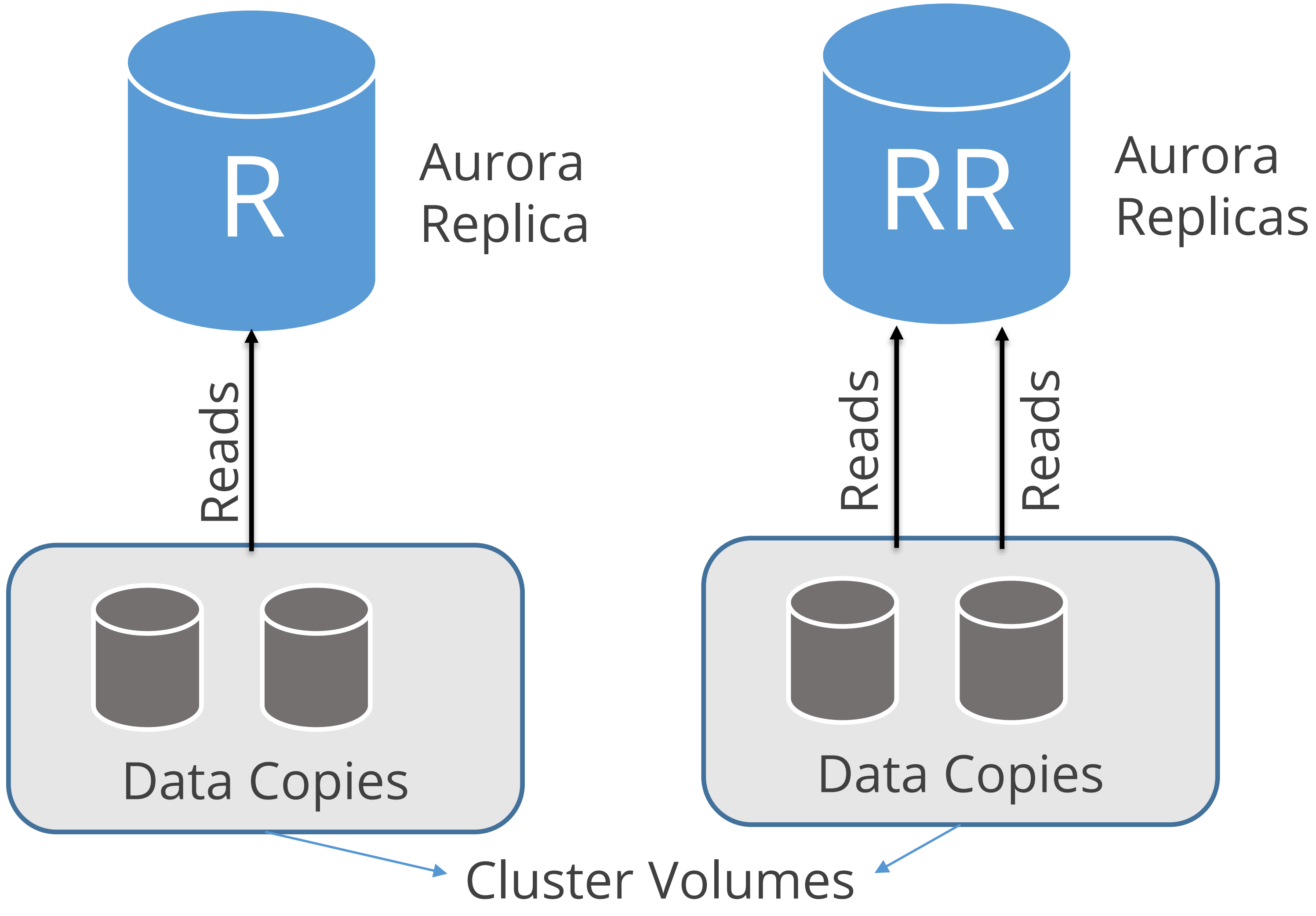
Amazon Aurora DB Cluster

In each Aurora DB cluster, the primary instance handles both read and write operations, updating the data in the cluster volume with every alteration.



Amazon Aurora DB Cluster

An Aurora replica supports only read operations and connects to the same storage disk as the primary DB instance.



Each Aurora DB cluster can contain up to 15 Aurora Replicas in addition to the primary DB instance.

Amazon Aurora Serverless

Amazon Aurora Serverless is a configuration that automatically scales. It starts, shuts down, and adjusts capacity based on the application's needs.

There are two versions of Amazon Aurora Serverless:

- Amazon Aurora Serverless v1
- Amazon Aurora Serverless v2

Amazon Aurora Serverless v1

Amazon Aurora Serverless is a configuration that automatically scales. It starts, shuts down, and adjusts capacity based on the application's needs.

- It allows the computational capacity of a DB cluster to be scaled up or down according to the application's needs.
- Unlike other versions, the capacity of DB clusters created by Aurora Serverless v1 must be managed manually.

Amazon Aurora Serverless v1

The use cases for Amazon Aurora Serverless v1 include:

Infrequently-used applications

Development and test databases

Amazon Aurora Serverless v1

Here are some benefits of Aurora Serverless v1:

Easy to use
Scalable
Cost-effective
Highly available

Amazon Aurora Serverless v2

Aurora Serverless v2 helps users to automate the procedures for workload monitoring and database capacity adjustment.

Here are some benefits of Aurora Serverless v2:

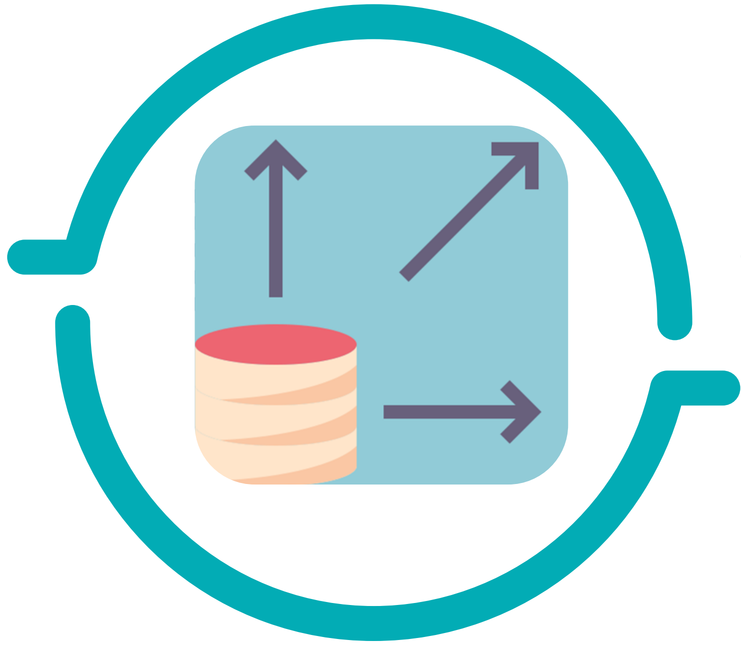
- Highly scalable
- Simple
- Highly available
- Transparent
- Cost-effective
- Durable

Amazon Aurora Serverless v2

Use cases for Amazon Aurora Serverless v2 include:

Managing enterprise database fleets

Supporting software as a service applications

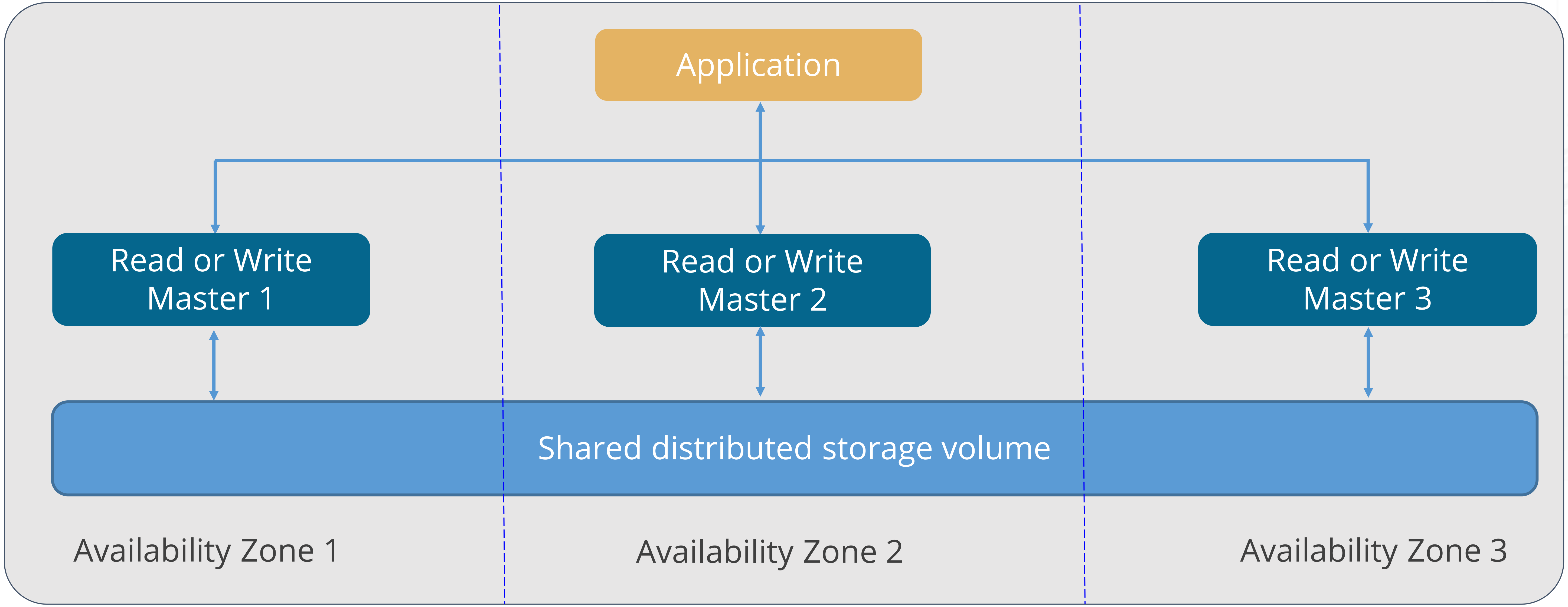


Distributing scaled-out databases across multiple servers

Handling unpredictable workloads

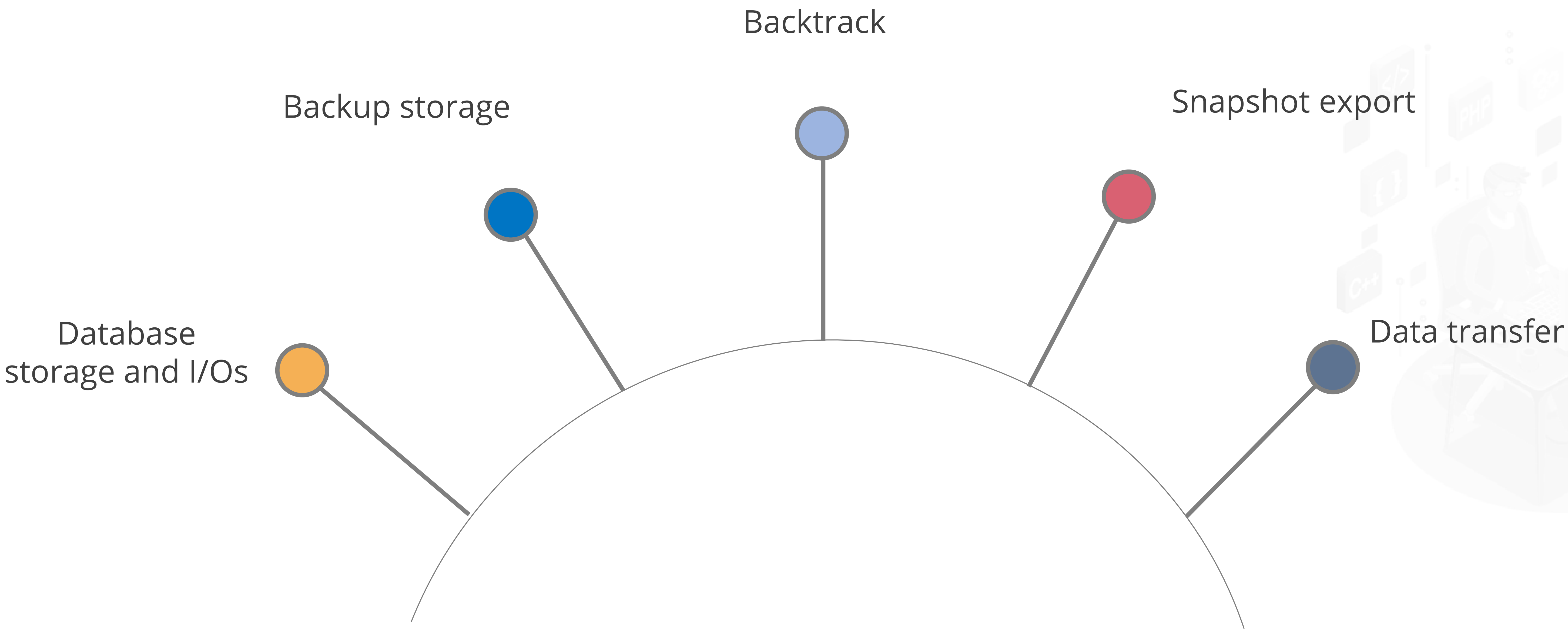
Aurora Multi-Master

Aurora Multi-Master lets users create multiple read and write master instances across several availability zones.



Amazon Aurora Costs

Costs associated with Amazon Aurora include:



Database storage and I/Os

Backup storage

Backtrack

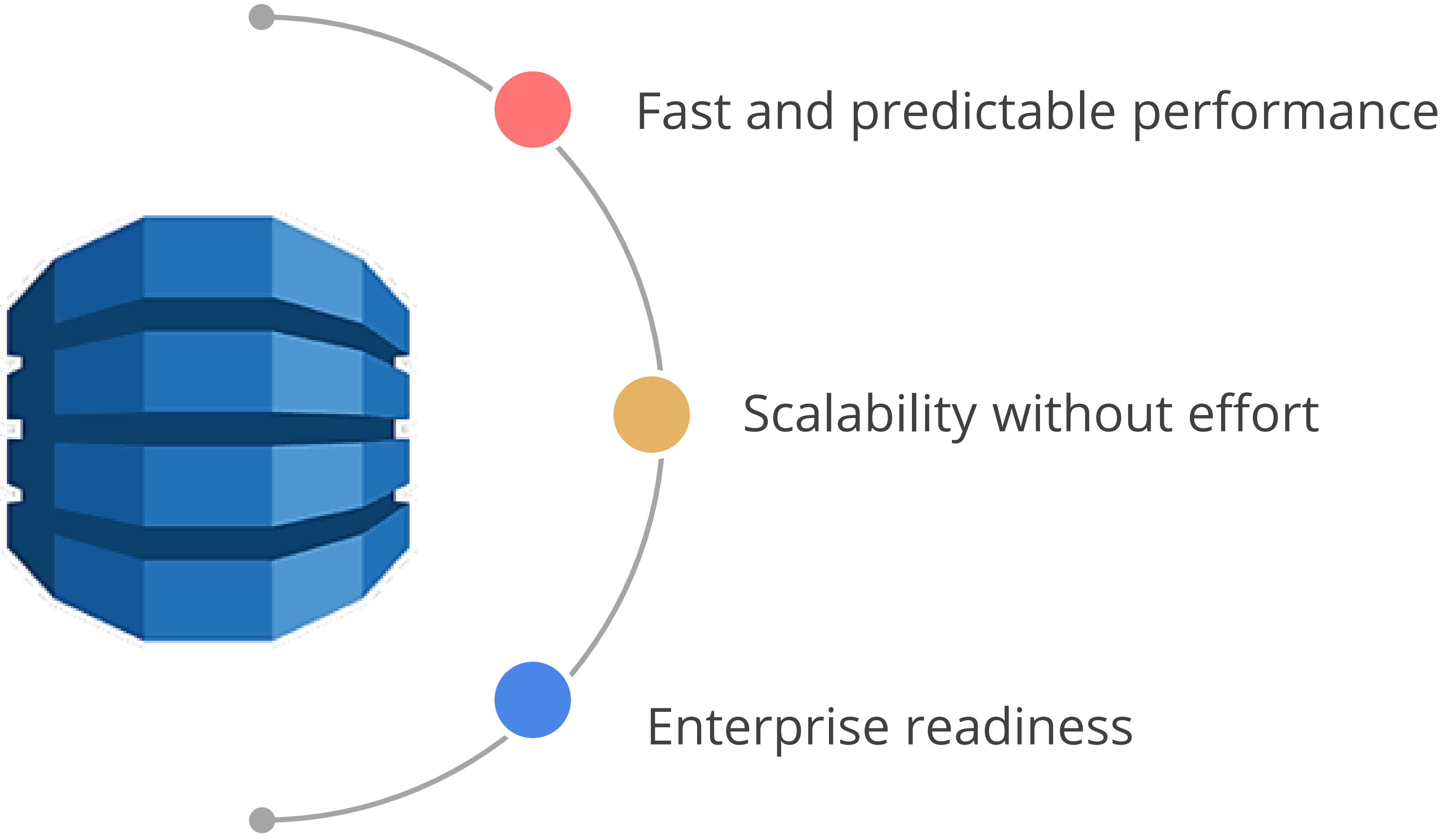
Snapshot export

Data transfer

DynamoDB

Amazon DynamoDB

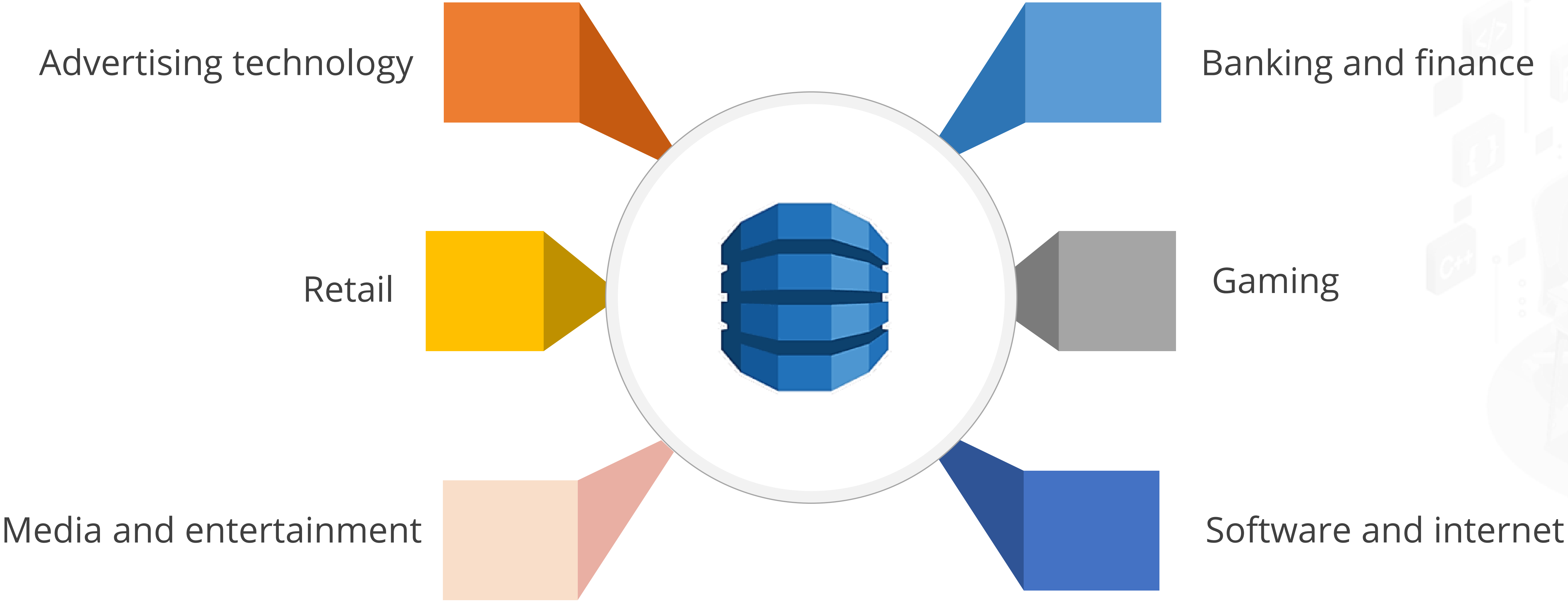
Amazon DynamoDB is a fully managed NoSQL database service. It offers:



- Fast and predictable performance
- Scalability without effort
- Enterprise readiness

Amazon DynamoDB: Use Cases

Amazon DynamoDB is widely used in:



Advertising technology

Banking and finance

Retail

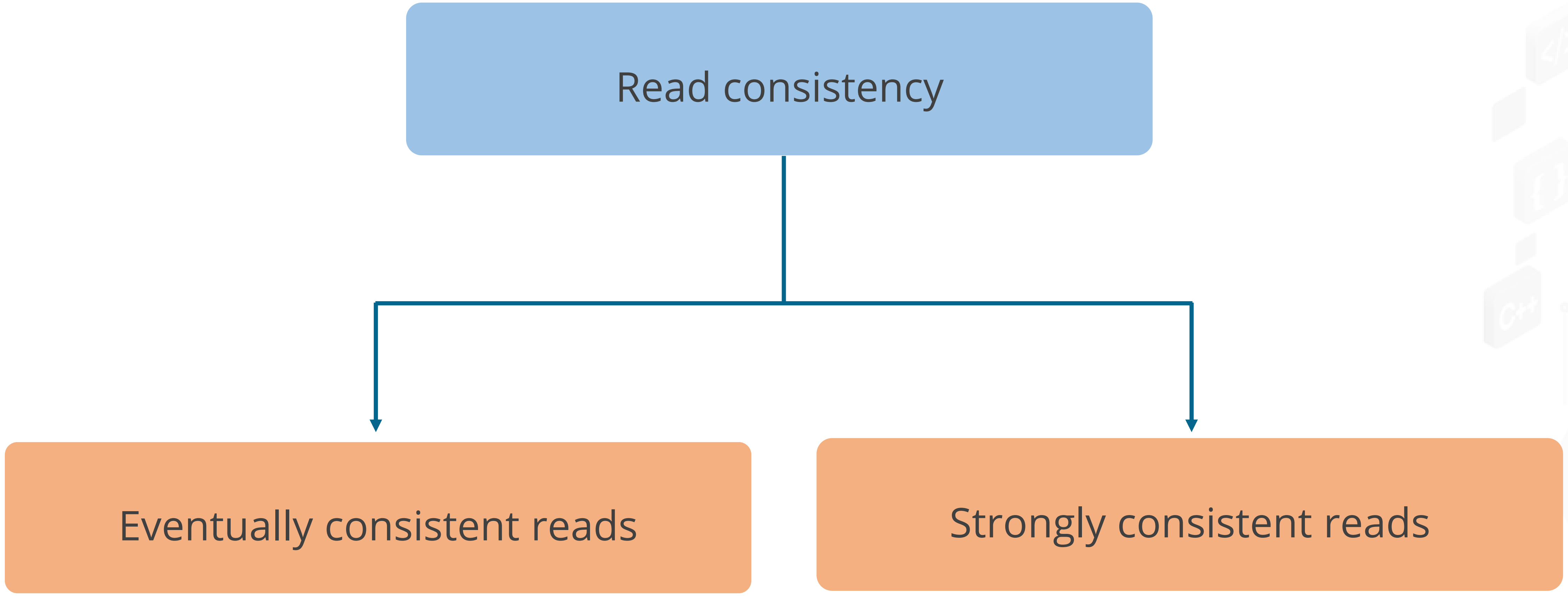
Gaming

Media and entertainment

Software and internet

Amazon DynamoDB: Read Consistency

Amazon DynamoDB provides two types of read consistency:



Amazon DynamoDB Global Tables

Global dispersed users

Global app

Replica (Europe)

Replica (Asia)

- Global tables offer a fully managed, multi-region, and multi-master database.
- Amazon ensures fast, local read-and-write performance.
- Global tables automatically replicate DynamoDB tables across the chosen AWS Regions.

Amazon DynamoDB: Reading an Item

To read an item from a DynamoDB table, use the **get-item** command

An example of a read operation:

```
aws dynamodb get-item \
   --table-name Catalog \
  --key '{"Id":{"N":"1"}}'
```

Amazon DynamoDB

To write an item in a DynamoDB table, users can use the following commands:

PutItem

This command creates a new item. If an item with the same key already exists, the new item replaces it.

DeleteItem

This command deletes the item associated with the specified key.

UpdateItem

If the specified key doesn't exist, this command adds a new item. If the key exists, it updates the item's attributes.

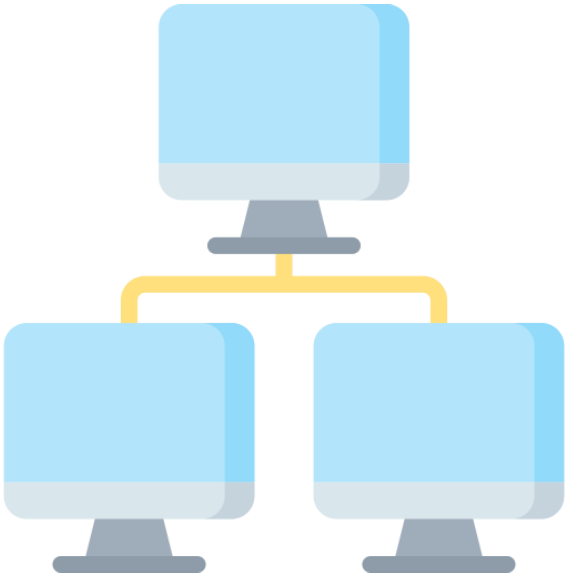
Amazon DynamoDB

In DynamoDB, a global secondary index (GSI) has a partition key and a sort key distinct from the base database's main key.

This index allows queries to access data from multiple base table partitions, making it global.

Amazon DynamoDB

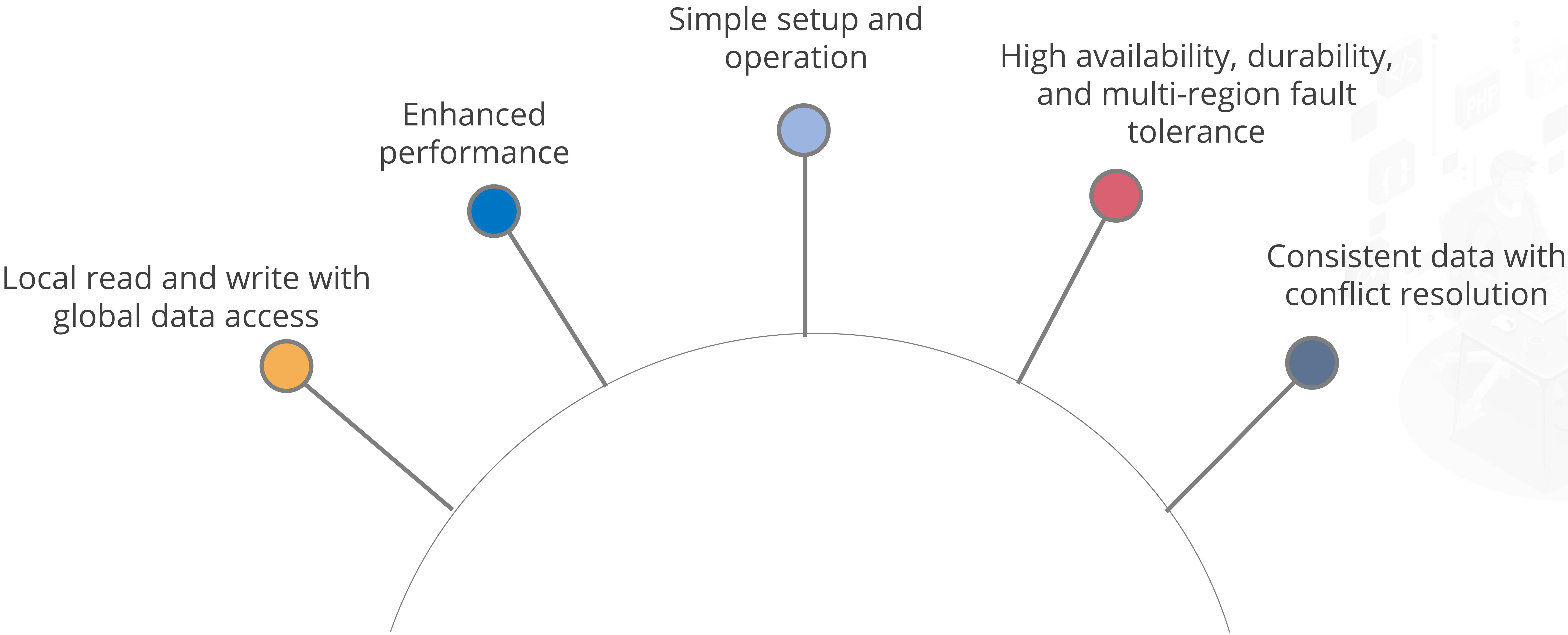
In DynamoDB, a local secondary index allows for flexible queries within a specific partition.



Unlike the global secondary index (GSI), which can span multiple partitions, the LSI operates within the constraints of a specific partition key.

DynamoDB Global Tables

The benefits of Amazon DynamoDB global tables are as follows:



Local read and write with global data access

Enhanced performance

Simple setup and operation

High availability, durability, and multi-region fault tolerance

Consistent data with conflict resolution

Case Study: Duolingo

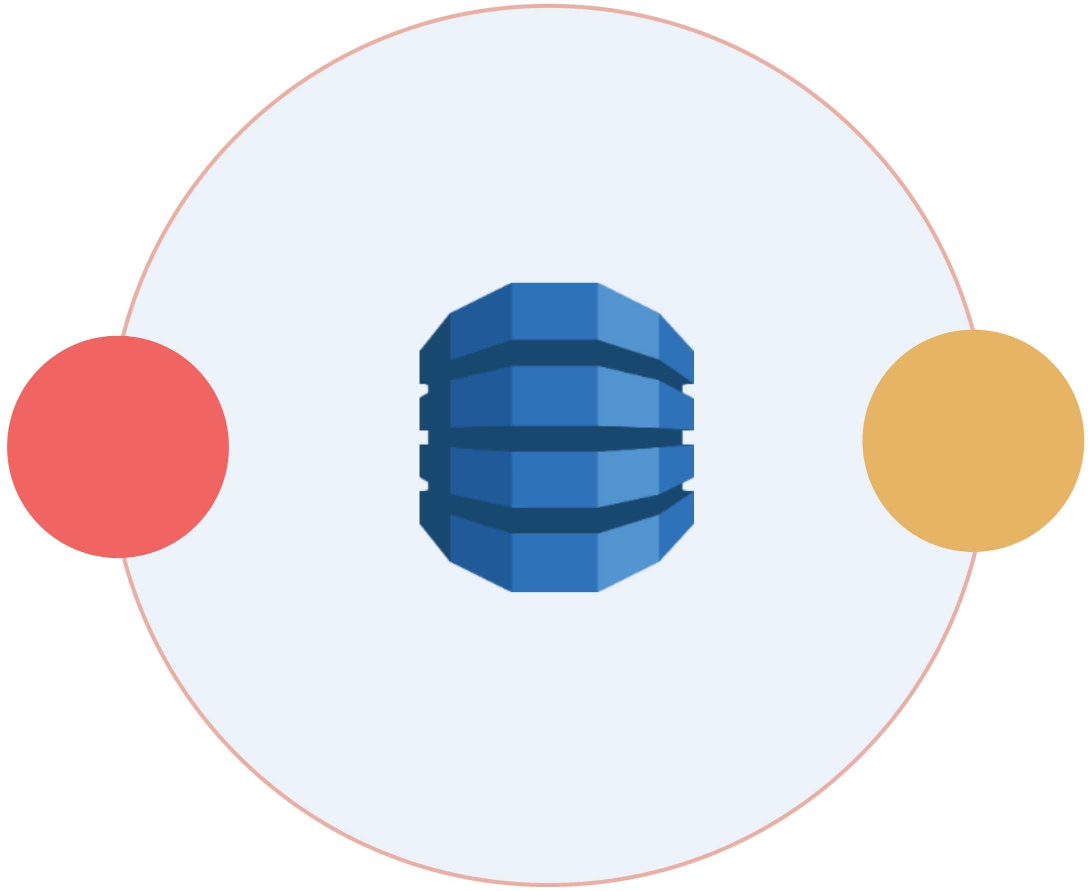
Duolingo utilizes Amazon DynamoDB to store 31 billion items, supporting its online learning platform that offers lessons in 80 languages.

duolingo
Amazon DynamoDB

- 31 billion items
- 80 different languages
- 24,000 read units per second and 3300 write units per second

Amazon DynamoDB: Costs

Amazon DynamoDB offers the following cost structures:



On-demand capacity mode

Provisioned capacity mode

Amazon DynamoDB: Features

Read Consistency

- DynamoDB supports both eventually consistent and strongly consistent reads.
- Reading data from a DynamoDB table may not immediately reflect the results of a recently completed write operation.
- DynamoDB maximizes read throughput.

Amazon DynamoDB: Features

Capacity Management

- DynamoDB manages throughput capacity for read and write operations.
- It can scale up or down to accommodate variable read/write demands.
- Developers can select the capacity mode that best aligns with their application’s requirements.

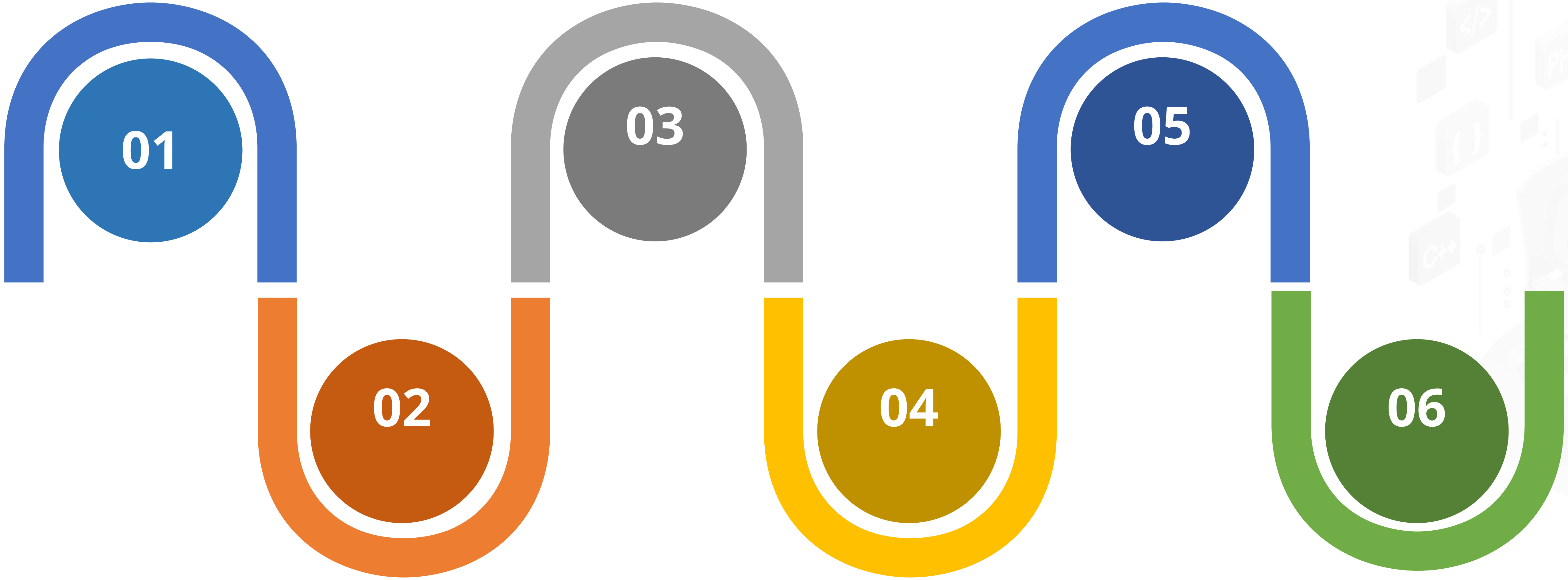
Amazon Keyspaces for Apache Cassandra

You can run Cassandra workloads on AWS using the same Cassandra application code and developer tools.

- Apache Cassandra is scalable and highly available.
- The Apache Cassandra-compatible database service is now available for preview in 18 AWS regions.

Amazon Keyspaces for Apache Cassandra

Here are some benefits of Apache Cassandra:



01 Offers high performance and scalability

02 Provides high availability

03 Ensures high security

04 Remains compatible with Apache Cassandra

05 Requires no servers to manage

06 Maintains performance at scale

Amazon Keyspaces for Apache Cassandra

The use cases for Apache Cassandra are:

Building low-latency applications

Migrating existing Cassandra workloads to the cloud

Developing applications using open-source technologies

Serving as a data store for applications

Amazon DocumentDB

Amazon DocumentDB provides powerful and intuitive APIs for flexible and agile development.

- The data is represented as a JSON document.
- The document data model is intuitive for developers.
- The data is persisted in the database using the same document model format used in the application code.

Amazon DocumentDB

Amazon DocumentDB, with its MongoDB compatibility, offers several benefits such as:

Reliability:
Amazon DocumentDB provides dependable database service with minimal downtime.

Speed:
It is fast, facilitating efficient and prompt data operations.

Management:
It is a fully managed service, removing the need for hands-on database management.

Amazon DocumentDB

Amazon DocumentDB has several use cases, including:

User profiles

Real-time big data

Content management

Database Operations in DynamoDB

Duration:15 min

Problem Statement:

You have been assigned a task to create a database table that can store and read items using the DynamoDB console.

Assisted Practice: Guidelines

Steps to be followed are:

1. Create a table
2. Store and read the items

Creating a Replica of DynamoDB

Duration:15 min

Problem Statement:

You have been assigned a task to set up a global table replica and make data available across different regions using DynamoDB.

Assisted Practice: Guidelines

Steps to be followed are:

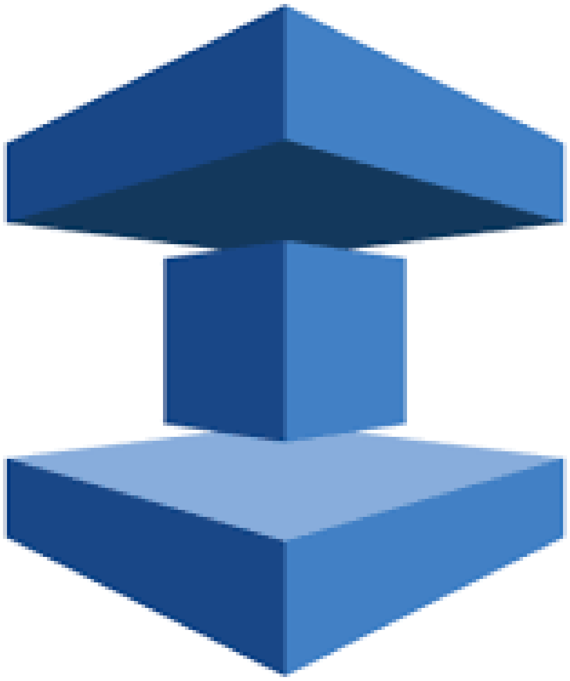
1. Create a replica in DynamoDB

TECHNOLOGY

Amazon ElastiCache

Amazon ElastiCache

Amazon ElastiCache is a web service that simplifies the deployment, operation, and scaling of an in-memory data store or cache in the cloud.

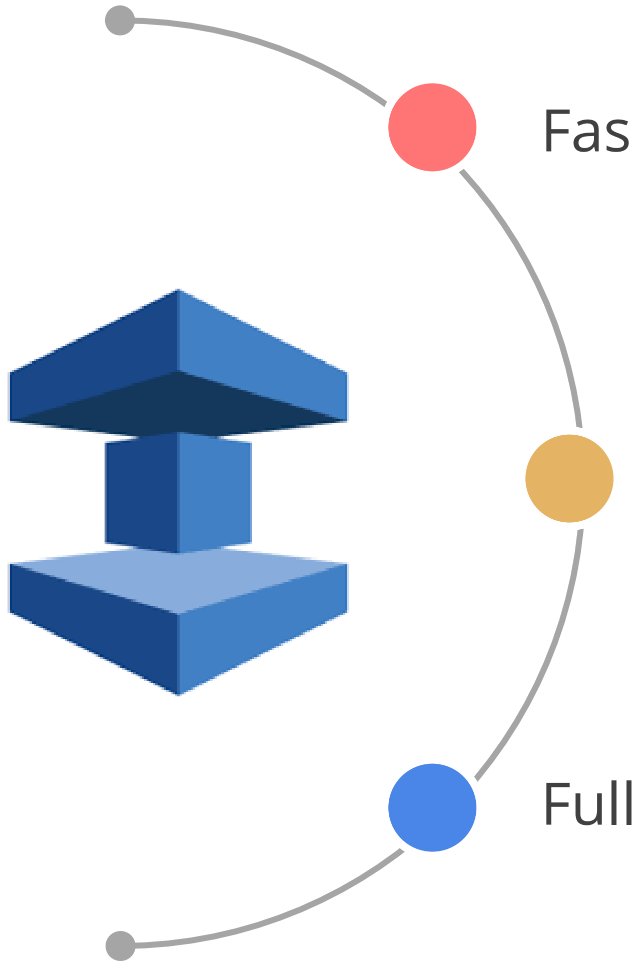


Key features of Amazon ElastiCache include:

- Easy management through the AWS management console
- Engine protocol compatibility
- Detailed engine node monitoring at no additional cost via Amazon CloudWatch
- Pay-per-use model for resource consumption

Amazon ElastiCache

Here are some benefits of Amazon ElastiCache:



Fast performance

Scalable

Fully manageable

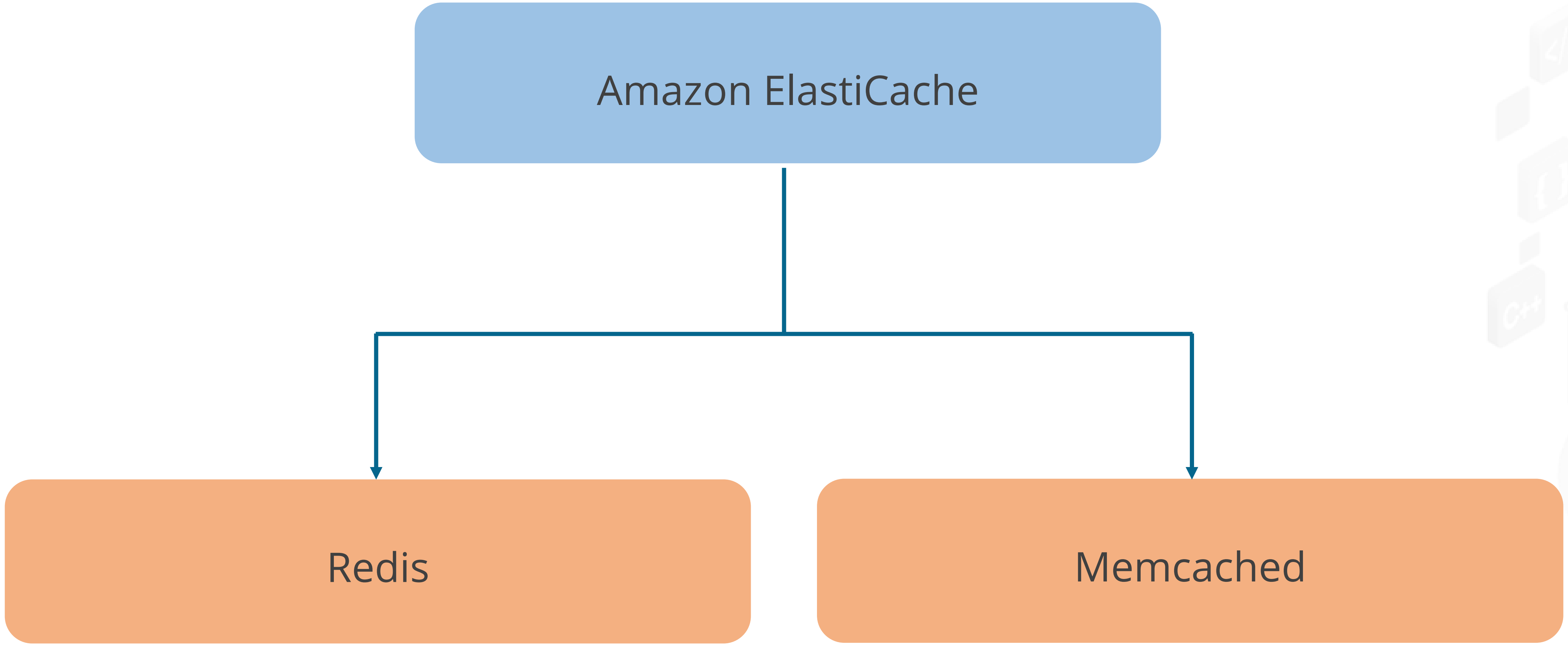
Amazon ElastiCache

Some of the use cases of Amazon ElastiCache:

- Caching
- Geospatial
- Media streaming

Amazon ElastiCache: Data Engines

Amazon ElastiCache supports two different data engines:



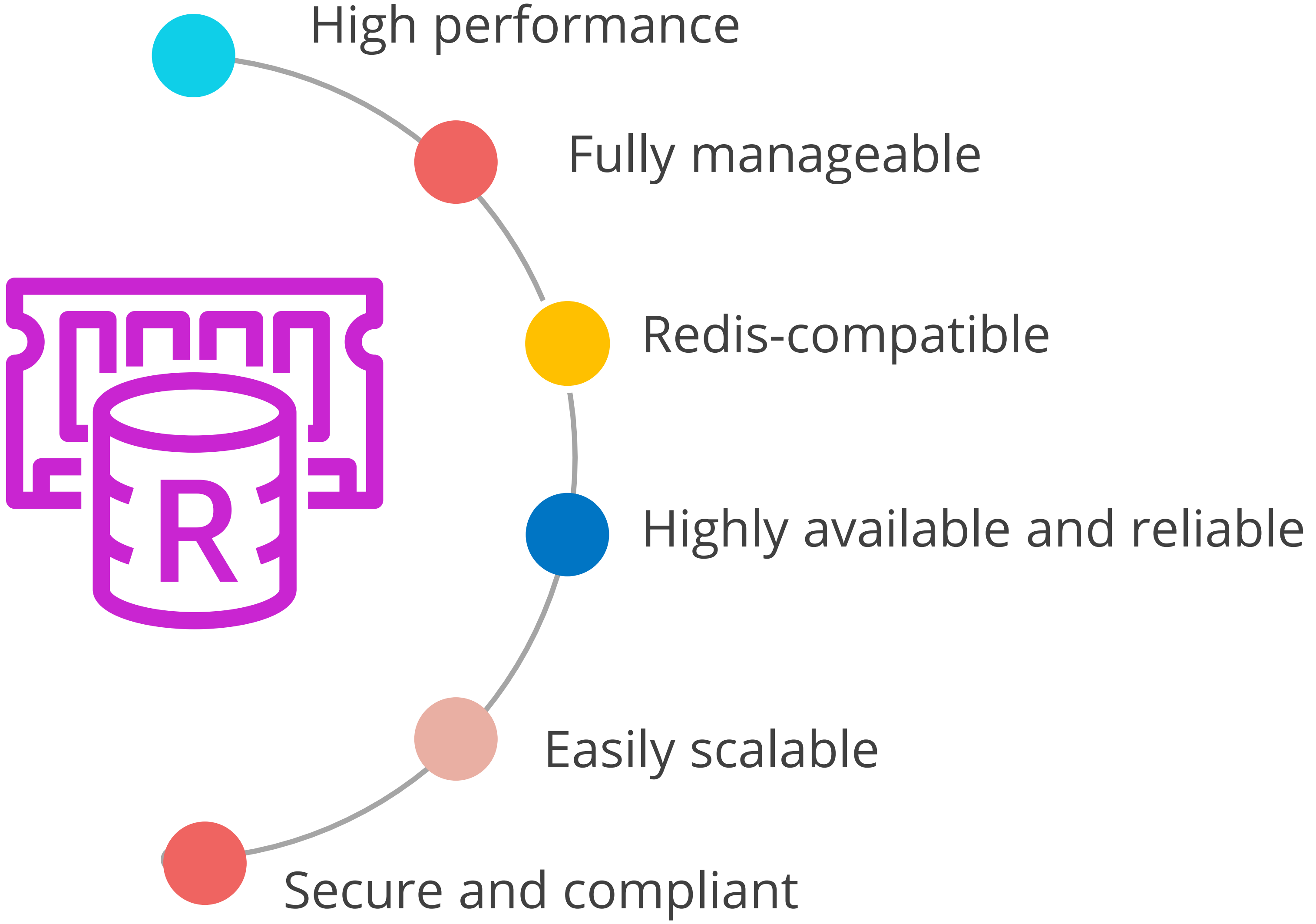
Amazon ElastiCache

Redis

Memcached

Amazon ElastiCache for Redis

Amazon ElastiCache is a Redis-compatible in-memory data store service and can support high performance since it is fully managed, scalable, and secure. The benefits are as follows:



High performance

Fully manageable

Redis-compatible

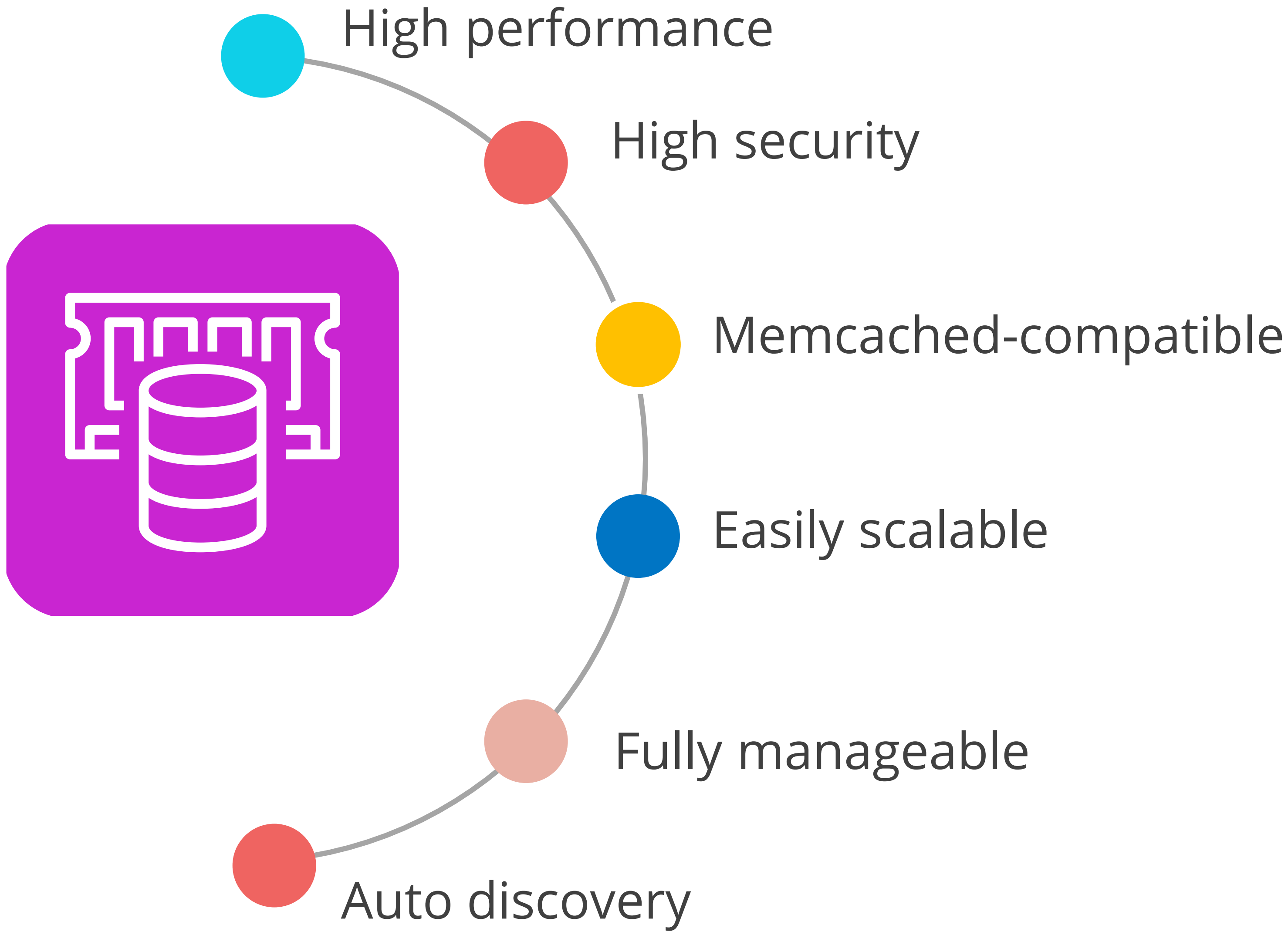
Highly available and reliable

Easily scalable

Secure and compliant

Amazon ElastiCache for Memcached

Amazon ElastiCache is a Memcached-compatible in-memory key-value store service that can be used as a cache or a data store. The benefits of Amazon ElastiCache for Memcached are as follows:



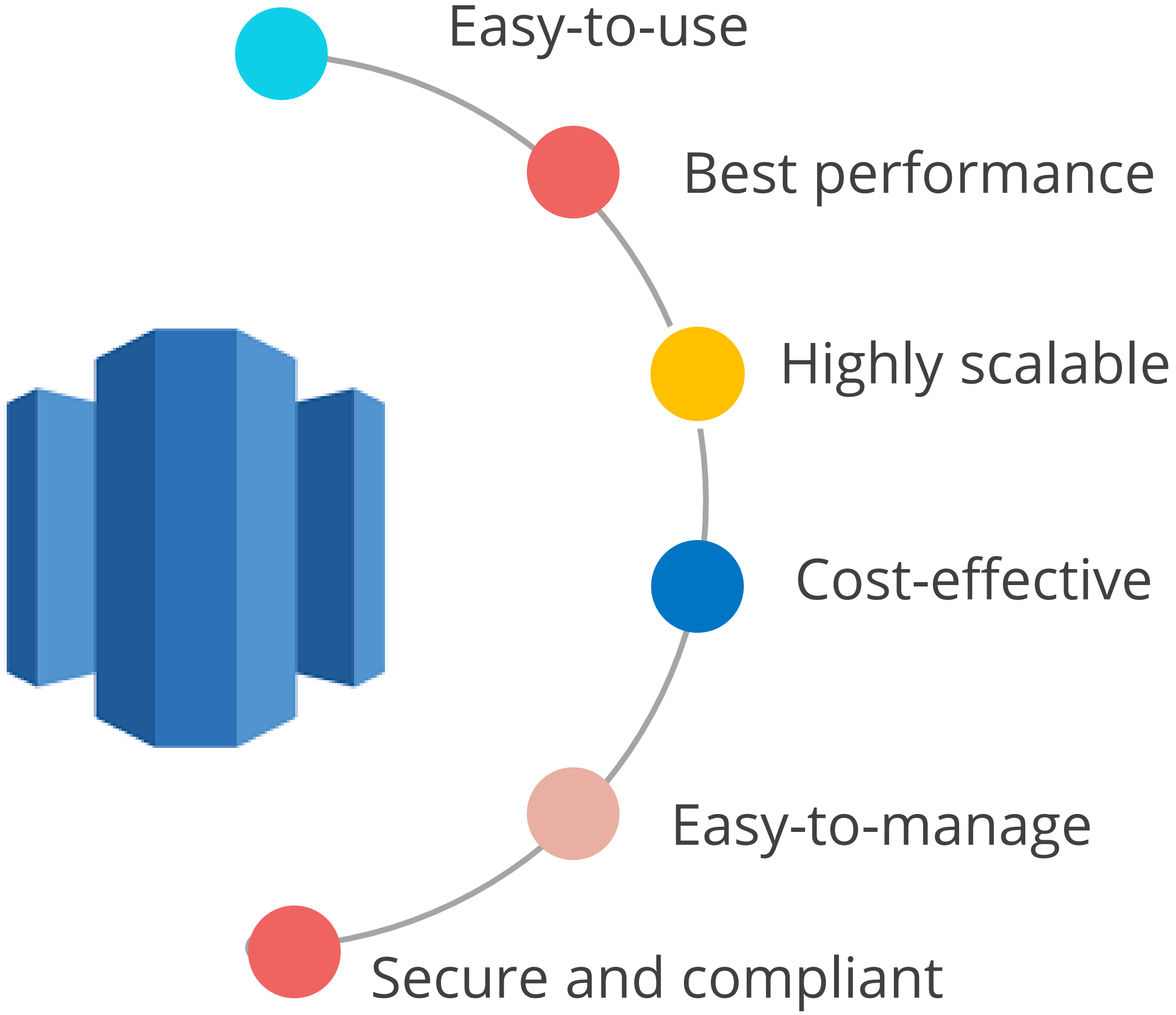
- High performance
- High security
- Memcached-compatible
- Easily scalable
- Fully manageable
- Auto discovery

TECHNOLOGY

Amazon Redshift

Amazon Redshift

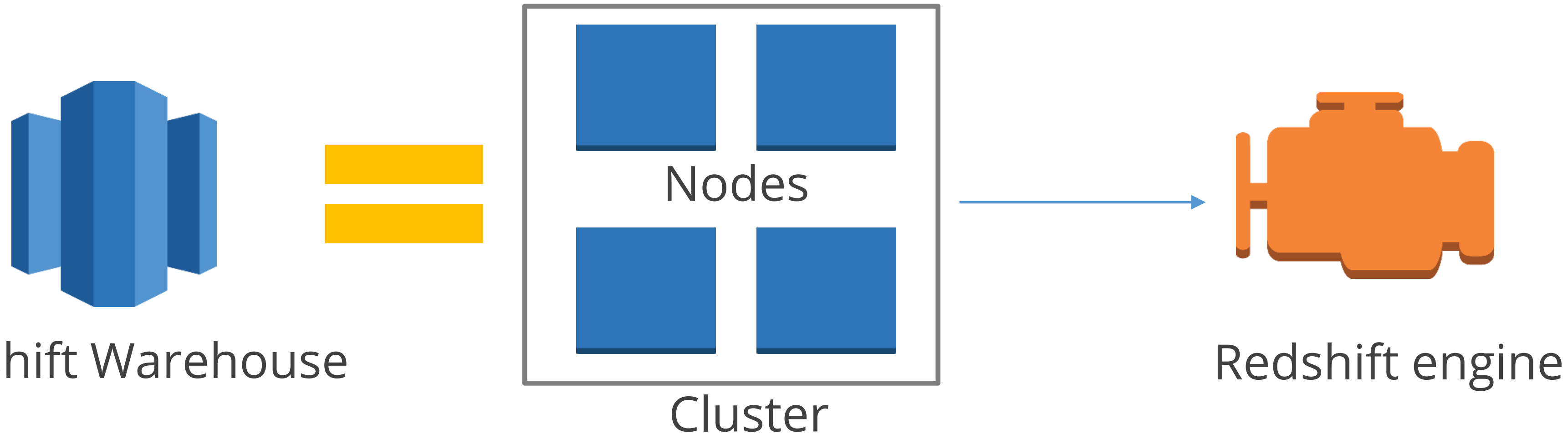
Amazon Redshift is a fully managed, petabyte-scale data warehouse service in the cloud.
The benefits of Amazon Redshift are as follows:

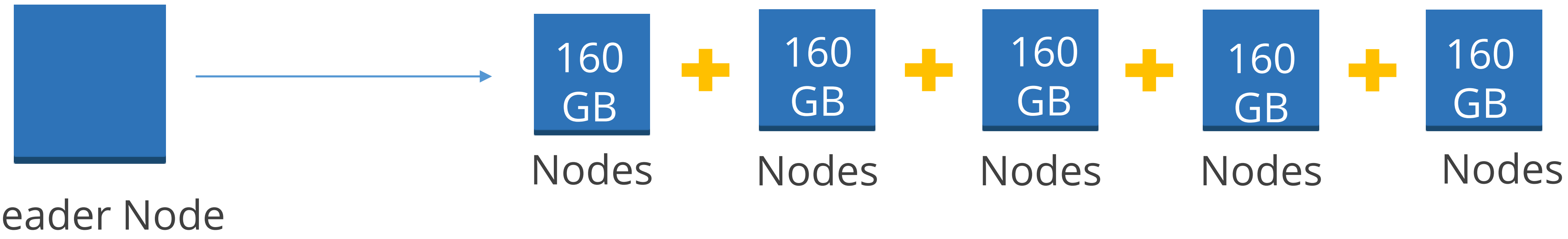


- Easy-to-use
- Best performance
- Highly scalable
- Cost-effective
- Easy-to-manage
- Secure and compliant

Amazon Redshift Clusters

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 an Amazon Redshift engine and contains one or more databases.

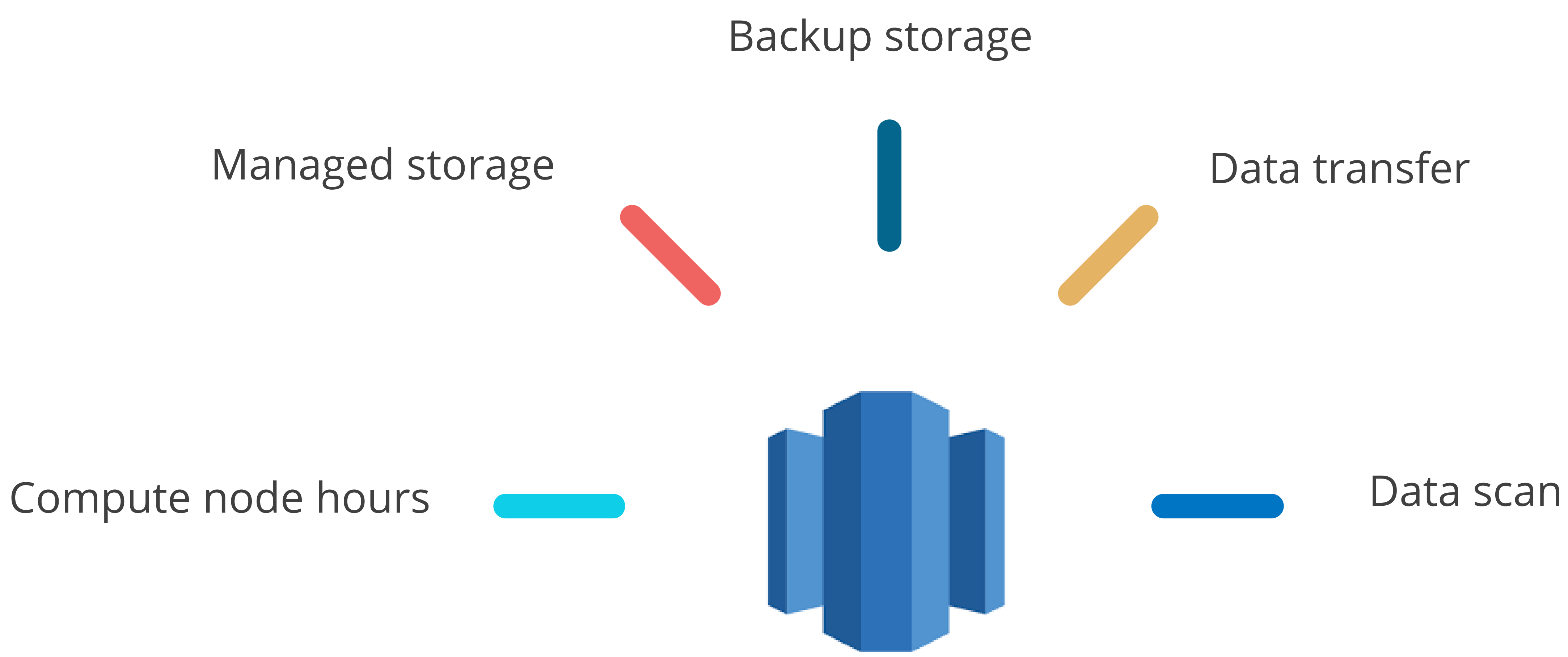




Massive Parallel Processing

Amazon Redshift Costs

The following are the costs associated with Amazon Redshift:



Backup storage

Managed storage

Data transfer

Compute node hours

Data scan

Amazon Redshift

The use cases of Amazon Redshift:

- Data analytics as a service
- Session store
- Streaming data analytics
- Real-time analytics

Configuring Query Data Using Redshift Query Editor

Duration:15 min

Problem Statement:

You have been assigned a task to create a table in Amazon Redshift and access the Redshift Query Editor.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to be followed are:

1. Create a table in Redshift

Creating a RDS MySQL Database

Duration:15 min

Problem Statement:

You have been assigned a task to demonstrate the process of creating an Amazon RDS MySQL database using the AWS Management Console.

ASSISTED PRACTICE

Assisted Practice: Guidelines

Steps to be followed are:

1. Create a RDS database

Key Takeaways

- Amazon RDS supports several database options including Amazon Aurora, Oracle, Microsoft SQL Server, PostgreSQL, MySQL, and MariaDB.

- Amazon DynamoDB is a NoSQL database service, offering fast, predictable performance and seamless scalability.

- Amazon ElastiCache is a web service that simplifies the deployment, operation, and scaling of an in-memory data store.

- Amazon DocumentDB is designed for managing and scaling JSON formatted data.

Deploying MySQL RDS Using AWS

Duration: 30 mins

LESSON-END PROJECT

Project agenda: To create and configure an RDS instance

Description: You must create an RDS database and deploy a Linux instance by creating it in EC2 and connecting an SSH client through EC2.

Perform the following:

1. Create an RDS database
2. Launch an EC2 instance
3. Create security groups
4. Connect the terminal to SSH

TECHNOLOGY

Thank You