FPT Software | Corporate Training Center

Module 6 Overview of Managed Service & Database

OPEN MIND - OPEN DOORS

1. AWS Managed Service & Database

RDS

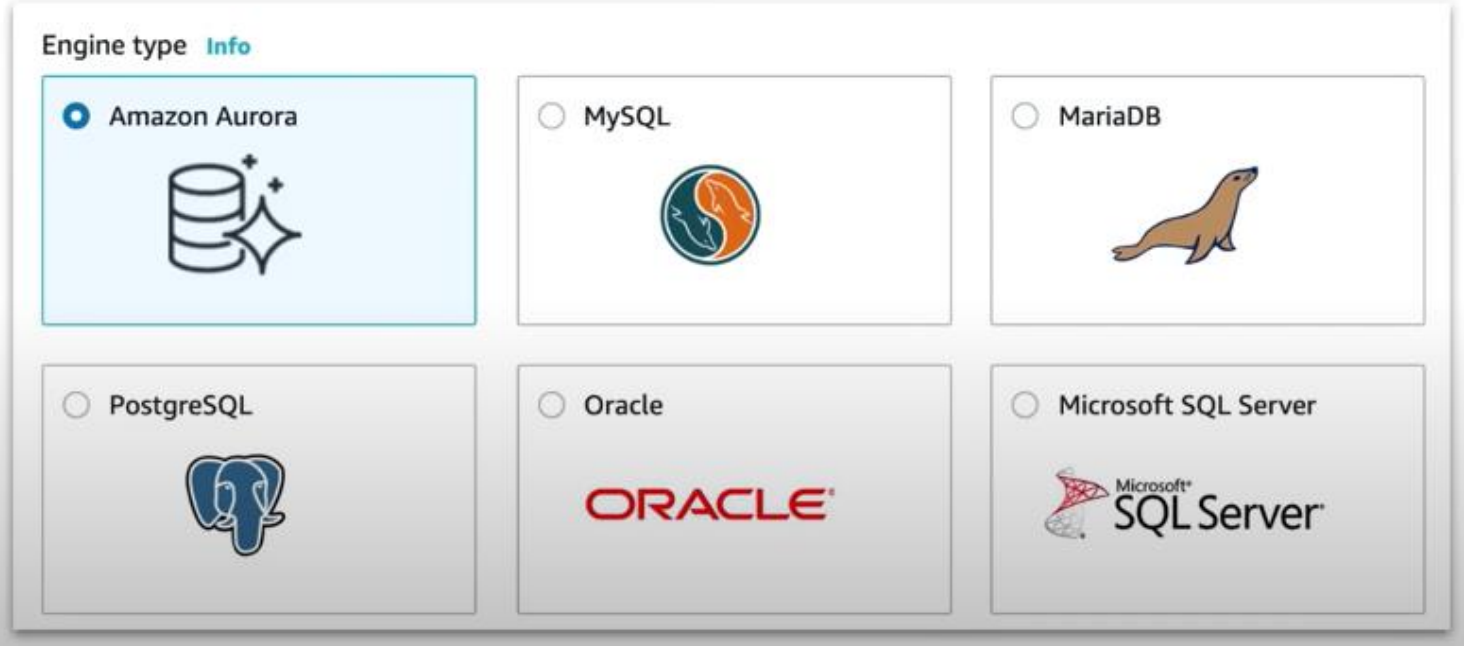
Relational Database Service Introduction

Relational Database Service (RDS)

A managed **relational database** service**.** Support multiple **SQL** engines, easy to scale, backup and secure.

Introduction to RDS

Relational Database Service (RDS) is the AWS Solution for **relational** databases.
There are **6 relational database** options currently available on AWS

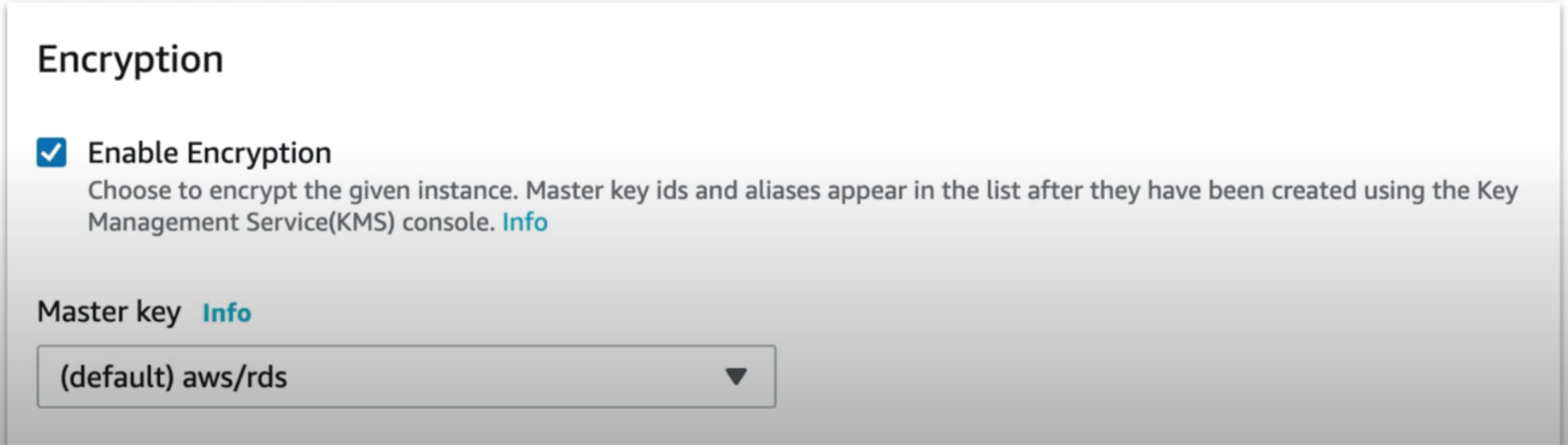



RDS
RDS Encryption

RDS - Encryption

You can turn on encryption at-rest for all RDS engines
You may not be able to turn encryption on for older versions of some engines.
It will also encrypted the automated backups, snapshots, and read replicas.

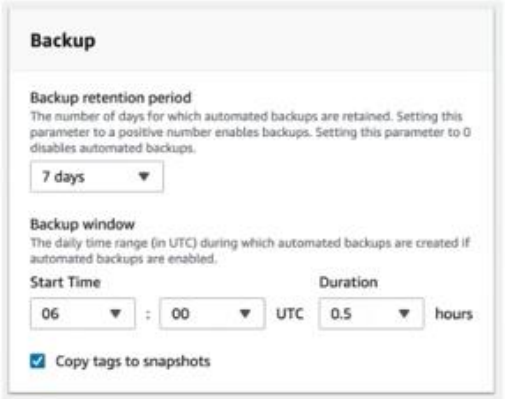
Encryption is handled using the AWS Key Management Service (KMS)




RDS
RDS Backups

RDS – Backup

There are ✌️ 2 backup solutions available for RDS



Automated Backups

Choose a Retention Period between 1 and 35 days

Stores transaction logs throughout the day

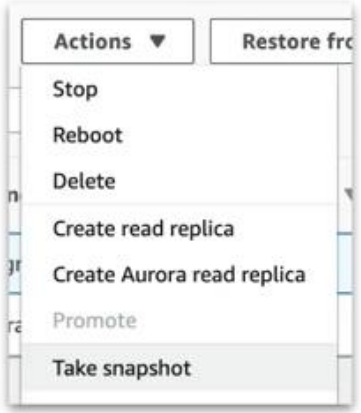
Automated backups are enabled by default

All data is stored inside S3

There is no additional charge for backup storage

You defined your backup window

Storage I/O may be suspended during backup



Manual Snapshots

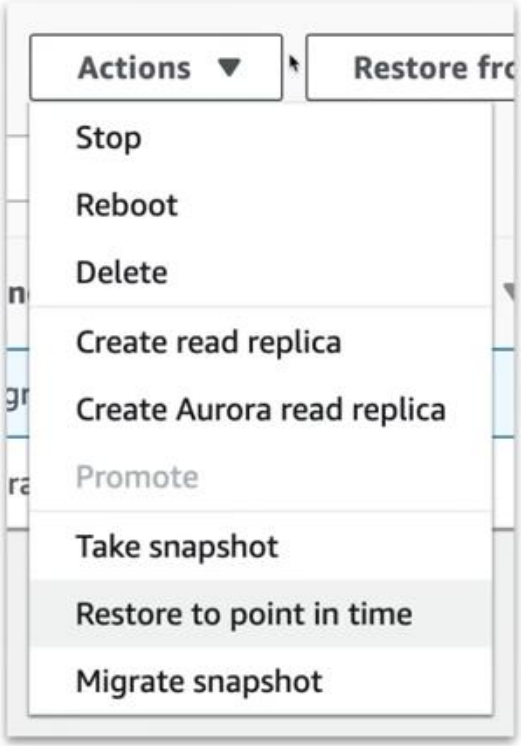
Taken manually by the user

Backups persist even if you delete the original RDS instance


RDS
RDS Restoring
Backups

RDS - Restoring Backup

When recovering AWS will take the most recent daily backup, and apply transaction log data relevant to that day. This allows point-in-time recovery down to a second inside the retention period.



Backup data is **never restored overtop** of an existing instance.

When you restore an RDS instance from Automated Backup or a Manual Snapshot a new instance is created for the restored database.

Restored RDS instances will have a new DNS endpoint.

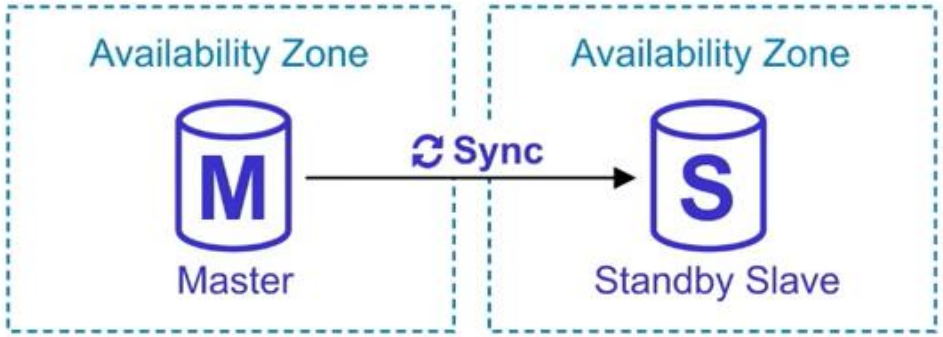

RDS
RDS Multi-AZ

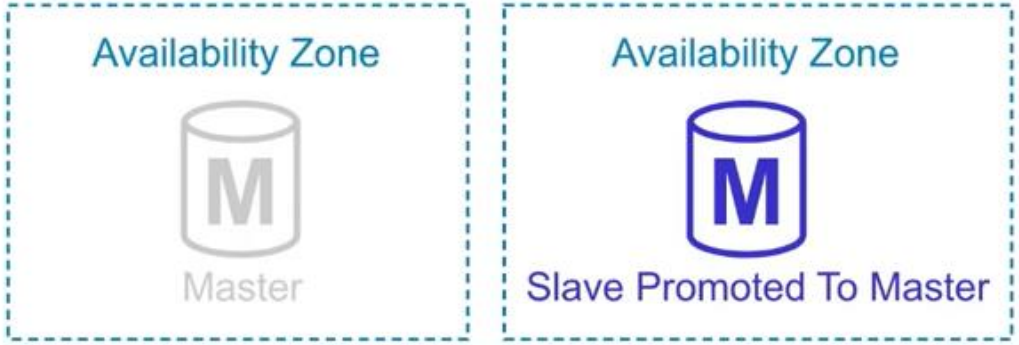
RDS - Multi AZ

Ensures database remains available if another AZ becomes unavailable

Makes an exact copy of your database in another AZ. AWS automatically **synchronizes** changes in the database over to the standby copy

Automatic Failover protection if one AZ goes down failover will occur and the standby slave will be promoted to master

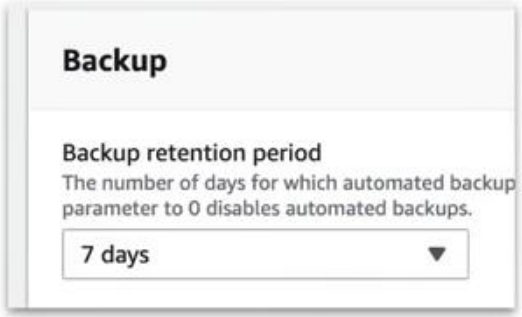





RDS
RDS Read Replicas

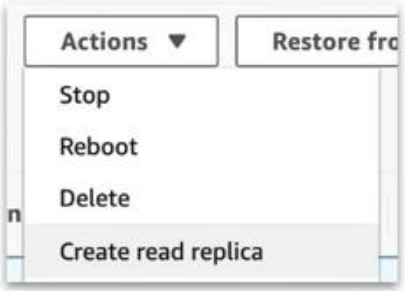
RDS – Read Replicas

Read-Replicas allow you to run **multiples copies** of your database, these copies only allows **reads** (no writes) and is intended to alleviate the workload of your primary database to improve performance

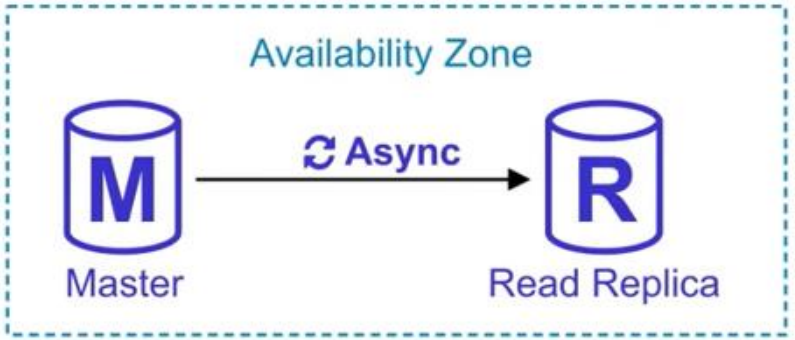


You must have automatic backups enabled to use Read Replicas

How to create a read replica:



Asynchronous replication happens between the primary RDS instance and the replicas.



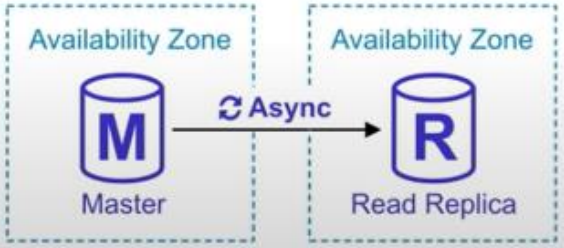
RDS – Read Replicas

You can have up to 🖐 **5 replicas** of a database

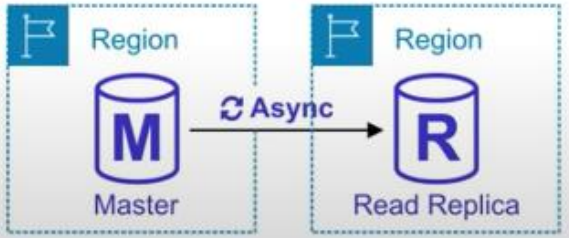
Each Read Replica will have its own DNS Endpoint

You can have Multi-AZ replicas, replicas in another region, or even replicas of other read replicas

Multi-AZ Replicas



Cross-Region Replicas



Replicas can be promoted to their own database, but this breaks replication

No automatic failover, if primary copy fails you must manually update urls to point at copy.


RDS
RDS Multi-AZ vs
Read Replicas

RDS – Multi-AZ vs Read Replicas

Multi-AZ Deployments	Read Replicas
Synchronous replication – highly durable	Asynchronous replication – highly scalable
Only database engine on primary instance is active	All read replicas are accessible and can be used for read scaling
Automated backups are taken from standby	No backups configured by default
Always span two Availability Zones within a single Region	Can be within an Availability Zone, Cross-AZ, or Cross-Region
Database engine version upgrades happen on primary	Database engine version upgrade is independent from source instance
Automatic failover to standby when a problem is detected	Can be manually promoted to a standalone database instance

DynamoDB

Introduction to DynamoDB

DynamoDB

A key-value and document database (NoSQL) which can guarantees consistent reads and writes at any scale.

Introduction to DynamoDB

What is NoSQL

NoSQL is database which is neither relational and does not use SQL to query the data for results

What is a Key/Value Store?

A form of data storage which has a key which references a value and nothing more

```
{ Title: 'S01E019 DS9 Duet' }
```

What is a Document Store?

A form of data storage which a nested data structure

```
{
  Series: 'DS9'
  Episodes: [
    {
      Season: 1,
      Epsisode: 19,
      Title: 'Duet'
    }
  ]
}
```

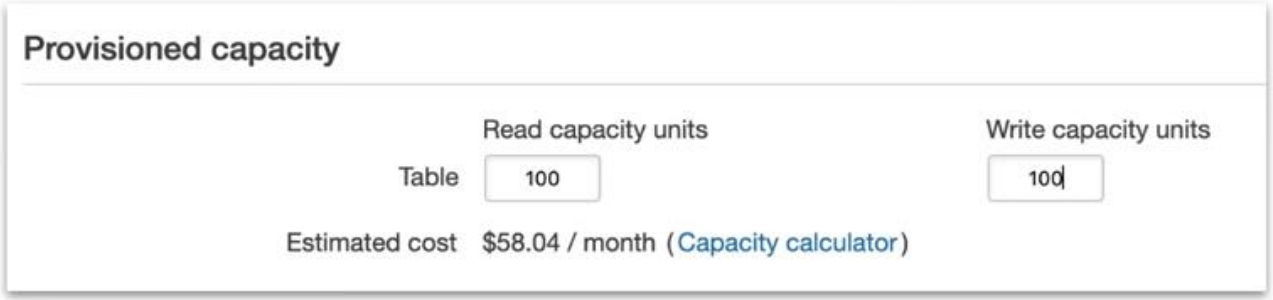
Introduction to DynamoDB

DynamoDB is a NoSQL **key/value** and **document** database for internet-scale applications.

Features

- Fully managed
- Multiregion
- Multimaster
- Durable database
- Built-in security
- Backup and restore
- In-memory caching

Specific your read and write capacity per second, it just works at **whatever capacity you need** without you tweaking anything.



Provides

- Eventual Consistent Reads (default)
- Strongly Consistent Reads

All data is stored on **SSD storage** and is spread across **3 different regions**.

DynamoDB

Table Structure

DynamoDB – Table Structure

Primary Key

Partition Key | Sort Key

IMDB ID	Year	Title	Box Office
tt0079945	1979	Star Trek: The Motion Picture	139000000
tt0084726	1982	Star Trek II: The Wrath of Khan	97000000
tt0088170	1984	Star Trek III: The Search for Spock	87000000
tt0092007	1986	Star Trek IV: The Voyage Home	133000000
tt0098382	1989	Star Trek V: The Final Frontier	63000000
tt0102975	1991	Star Trek VI: The Undiscovered Country	96900000
tt0111280	1994	Star Trek Generations	118000000
tt0117731	1996	Star Trek: First Contact	146000000
tt0120844	1998	Star Trek: Insurrection	117800000
tt0253754	2002	Star Trek: Nemesis	67300000

Item

Attribute

DynamoDB

Consistent Reads

DynamoDB – Reads

When data needs to updated it has to write updates to all copies. **It is possible for data to be inconsistent** if you are reading from a copy which has yet to be updated. You have the ability to choose the read consistency in DynamoDB to meet your needs.

Eventual Consistent Reads (DEFAULT)
When copies are being updated it is possible for you to read and be returned an inconsistent copy

Reads are fast but there is no guarantee of consistent

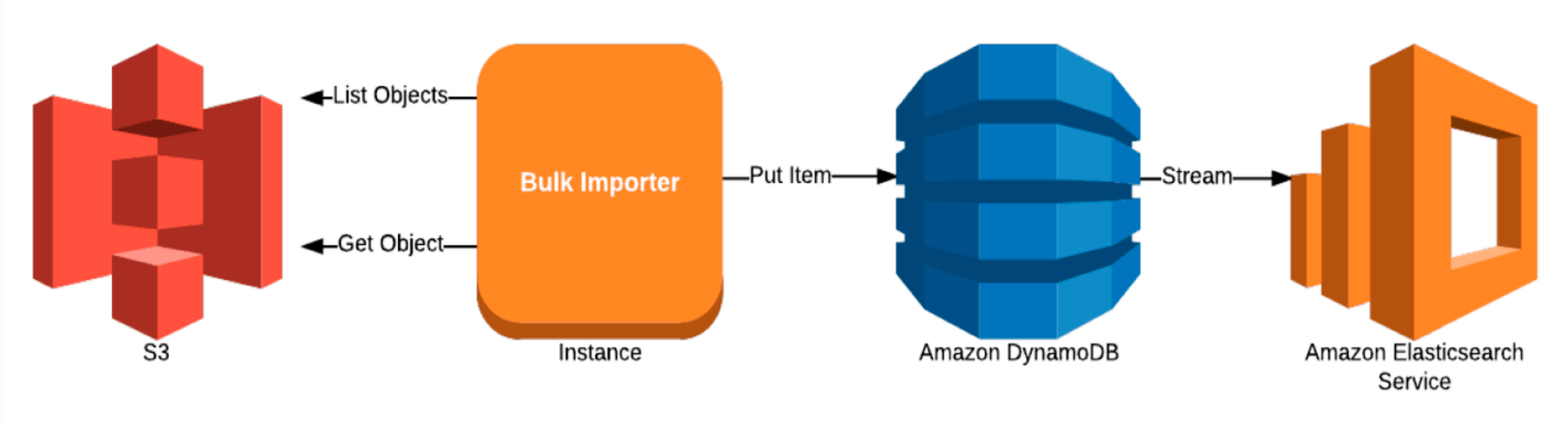
All copies of data eventually become generally consistent within a second.

Strongly Consistent Reads
When copies are being updated and you attempt to read, it will not return a result until all copies are consistent.

You have a guarantee of consistency but the trade off is higher latency (slower reads).

All copies of data will be consistent within a second

Amazon DynamoDB



FPT Software | Corporate Training Center

ElastiCache

ElastiCache Introduction

ElastiCache

Managed caching service which either runs Redis or Memcached

What is In-Memory Data Store?

Caching

Caching is the process of storing data in a cache. A cache is a **temporary storage** area. Caches are optimized for fast retrieval with the trade off that data is not durable.

In-Memory Data Store

When data is stored In-Memory (think of RAM). The trade off is high volatility (low durability, risk of data loss) but **access** to data is **very fast**.

Introduction to ElastiCache

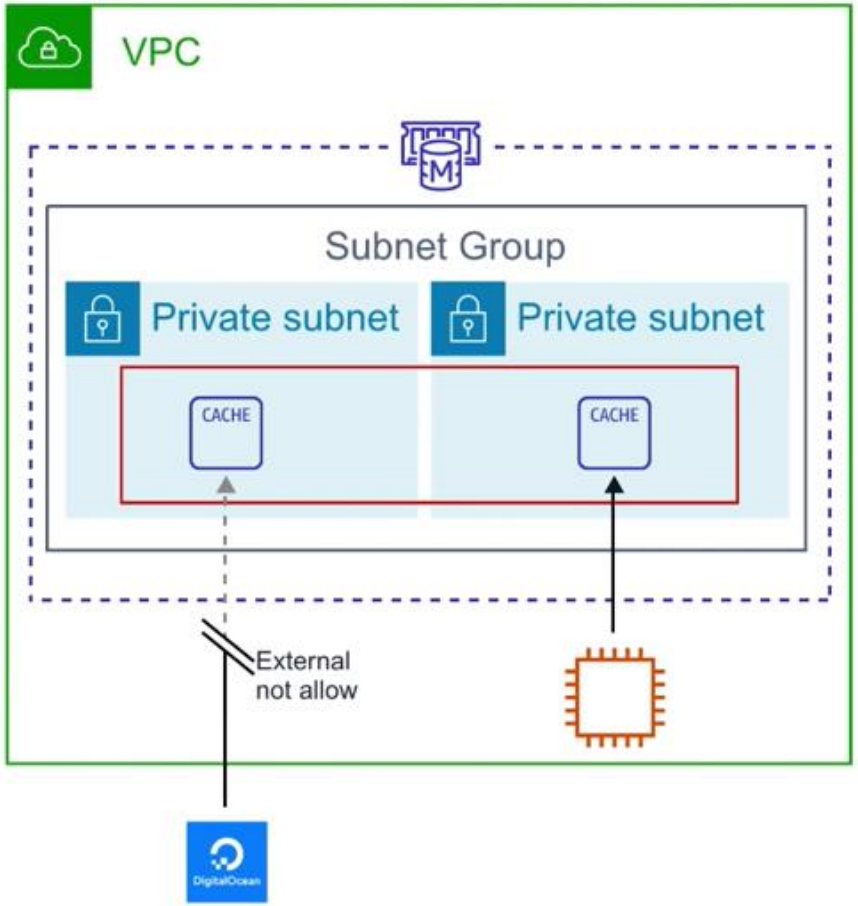
Deploy, run, and scale **popular open source compatible in-memory** data stores.

Frequently identical queries are stored in the cache.

ElastiCache is only accessible to resource operating with the same VPC to ensure low latency

ElastiCache supports 2 open-source caching engines:

1. Memcached
2. Redis



AWS Solutions Architect Associate

ElastiCache

Caching Comparison

ElastiCache – Caching Comparison

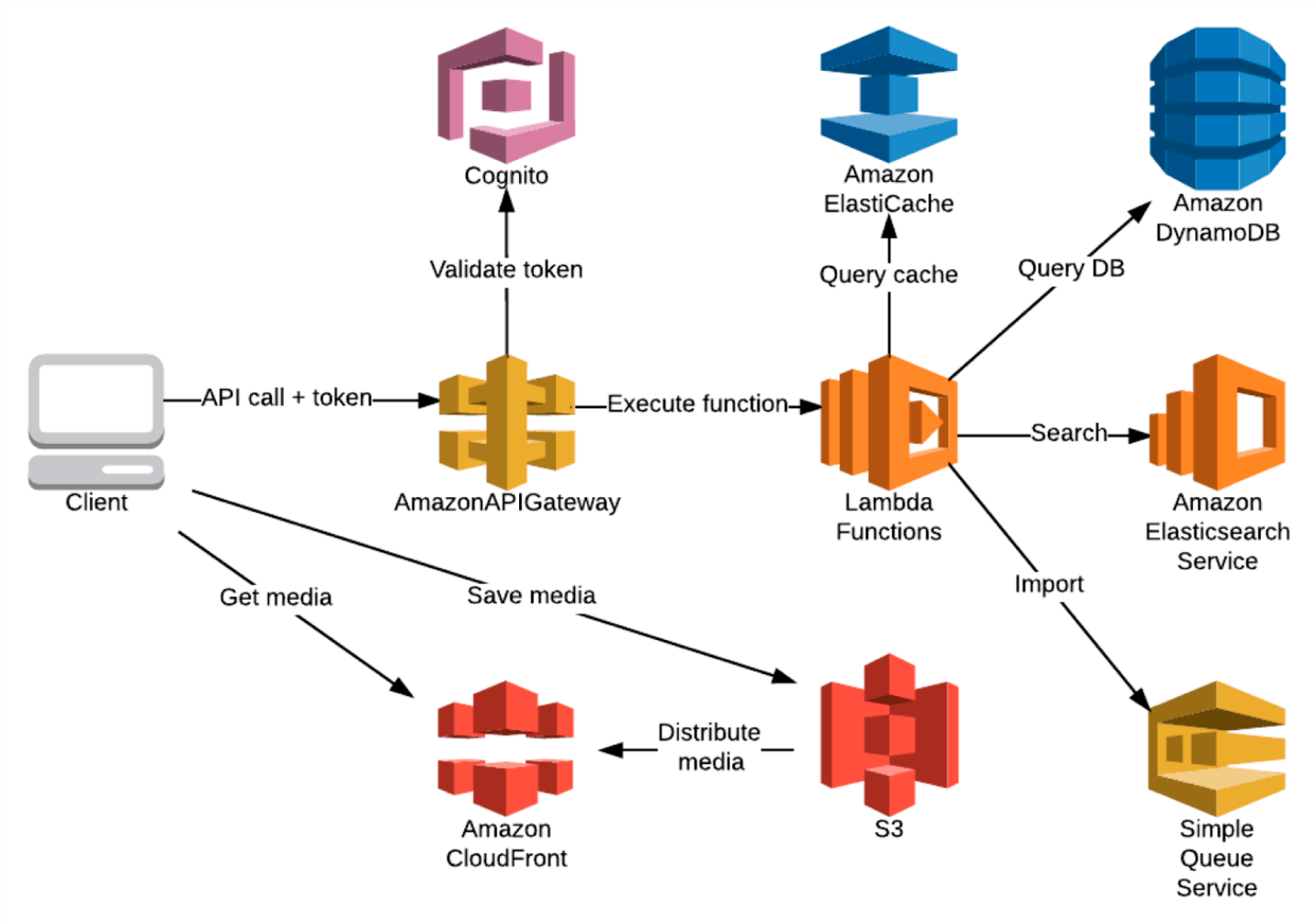
Memcached is generally preferred for caching HTML fragments. Memachce is a simple key/value store. The trade off it to being simple is that its very fast

Redis can perform many different kinds of operations on your data. It's very good for leaderboards, keep track of unread notification data. It's very fast, but **arguably** not as fast as Memcached.

Don't google "Memcache vs Redis" unless you want to read endless arguments as if people are arguing "Kirk vs Picard"

	Memcached	Redis
Sub-millisecond latency	Yes	Yes
Developer ease of use	Yes	Yes
Data partitioning	Yes	Yes
Support for a broad set of programming languages	Yes	Yes
Advanced data structures	—	Yes
Multithreaded architecture	Yes	—
Snapshots	—	Yes
Replication	—	Yes
Transactions	—	Yes
Pub/Sub	—	Yes
Lua scripting	—	Yes
Geospatial support	—	Yes

Amazon Elastic Cache



Redshift

Redshift Overview

Amazon Redshift

Fully Managed Petabyte-size Data Warehouse.
Analyze (Run complex SQL queries) on massive amounts of data
Columnar Store database.

What is a Data Warehouse?

What is a Database Transaction?
A transaction symbolizes a unit of work performed within a database management system
eg. reads and writes

Database

Online **Transaction** Processing (OLTP)

A database was built to store current transactions and enable **fast access to specific transactions** for ongoing business processes

Adding Items To Your Shopping List

Single Source

short transactions (small and simple queries) with an emphasis on writes.

VS

Data Warehouse

Online **Analytical** Processing (OLAP)

A data warehouse is built to store large quantities of historical data and **enable fast, complex queries across all the data**

Generating Reports

Multiple Sources

Long transactions (long and complex queries) with an emphasis on reads.

Introduction of Redshift

AWS Redshift is the AWS managed, petabyte-scale solution for **Data Warehousing**.

Pricing starts at just $0.25 per hour with no upfront costs or commitments.
Scale up to petabytes for $1000 per terabyte, per year.
Redshift is price is less than 1/10 cost of most similar services.

Redshift is used for Business Intelligence.

Redshift uses OLAP (Online Analytics Processing System)

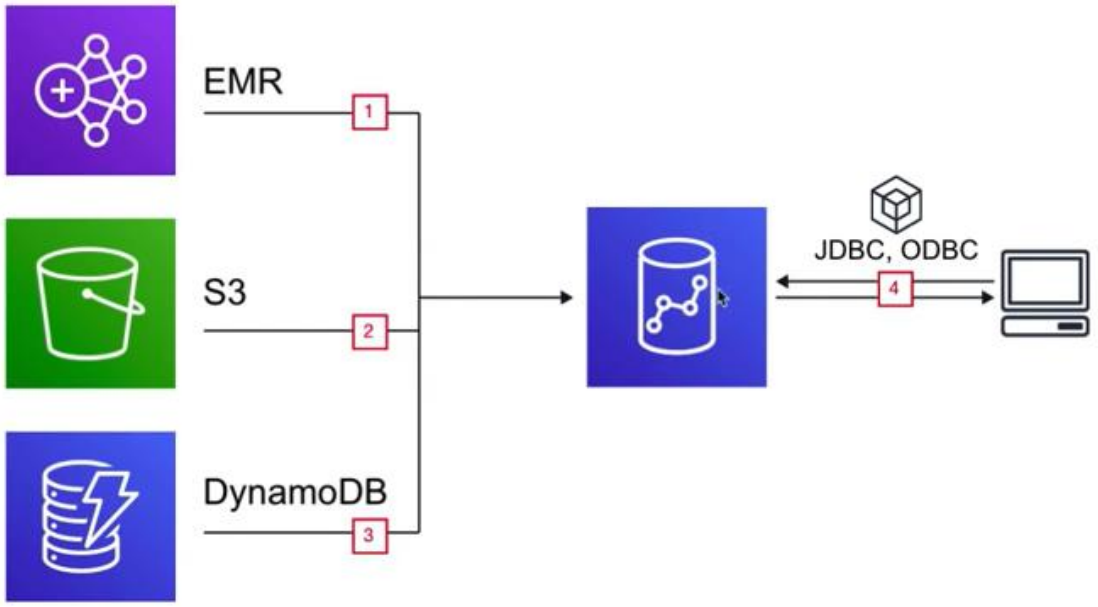
Redshift is **Columnar Storage** Database

Columnar storage for database tables is an important factor in optimizing analytic query performance because it drastically reduces the overall disk I/O requirements and reduces the amount of data you need to load from disk.

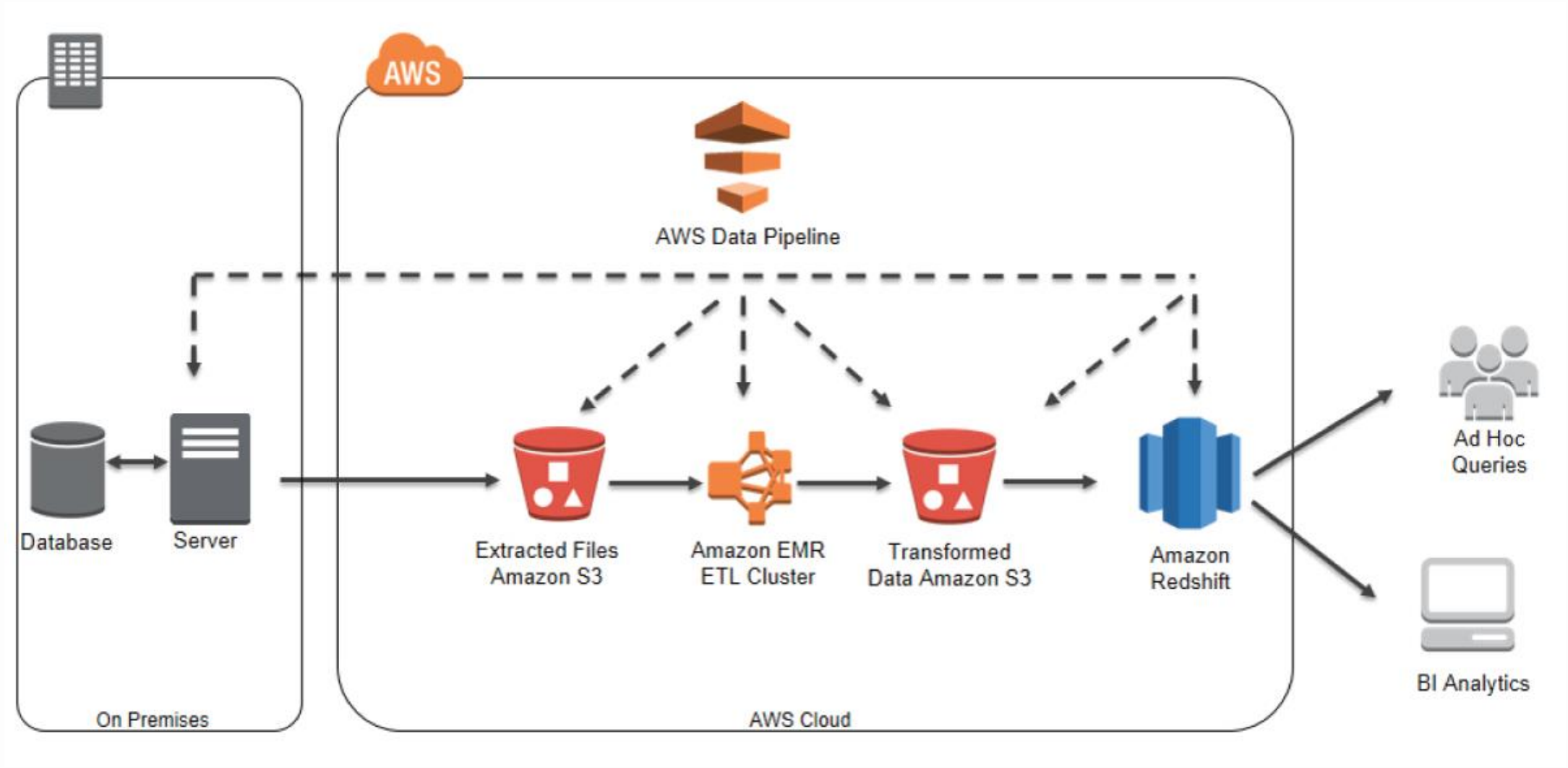
Redshift – Use Case

We want to continuously COPY data from
1. EMR,
2. S3 and
3. DynamoDB
to power a custom Business Intelligence tool.

Using a third-party library we can connect and query Redshift for data.



Amazon Redshift



FPT Software | Corporate Training Center

Thank You !

OPEN MIND - OPEN DOORS