Day 13/100

Databases

- Storing data on disk [EFS, EBS, EC2 Instance Store, S3] can have its limits.
- Sometimes you want to store data in a database.
- You can structure the data
- You build indexes to efficiently query/search through the data.
- You define relationships b/w your datasets.
- Databases are optimized for a purpose and come with different features, shapes and constraints.

Relational Databases

- Looks just like Excel Spreadsheet, with links b/w them
- Can use the SQL language to perform queries / lookups

NOSQL Databases

- NoSQL = non-SQL = non relational databases
- NOSQL databases are purpose built for specific models and have flexible schema for

building modern applications

- Benefits
  - Flexibility: easy to evolve data model
  - Scalability: designed to scale-out by using distributed clusters.
  - High-performance: optimized for specific data model.
  - Highly functional: types optimized for the data model.

- Example: Key-value, document, graph, in-memory, search database.

NOSQL data example: JSON

- JSON = JavaScript Object Notation
- JSON is a common form of data that fits into a NOSQL model
- Data can be nested
- fields can change over time.
- Supports for new types, arrays, etc.

```
{ "name": "John",
  "age": 30
  "cars": [
      "Ford",
      "BMW",
      "Fiat",
  ],
  "address": {
      "type" = "house",
      "number" = 23,
      "street" = "Main Road"
  }
}
```

DATE 16/12/21

Database & Share Responsibility on AWS

- AWS offers use to manage different databases.
- Benefits include:
  - Quick Provisioning, High Availability, Vertical and Horizontal Scaling.
  - Automated Backup & Restore, Operations, Upgrades.
  - Operating System Patching in handled by Aws.
  - Monitoring, alerting.
- Note: many databases technologies could be run on EC2, but you must handle yourself the resiliency, backup, patching, high availability, fault tolerance, scaling...

AWS RDS Overview

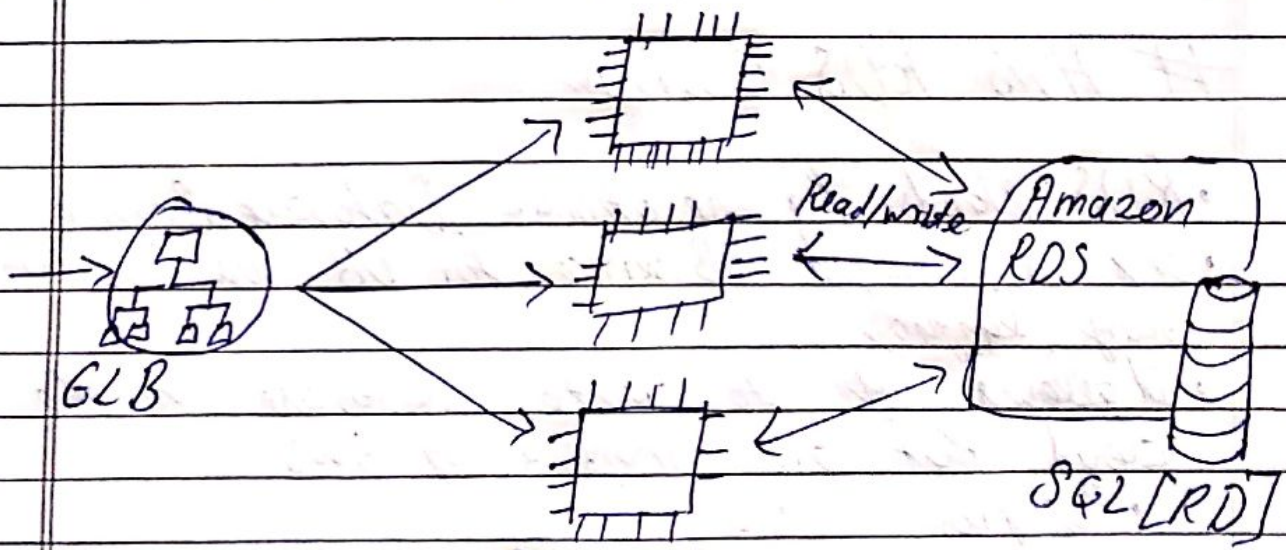
- RDS stands for Relational Database Service.
- It's a managed DB service for DB use SQL as a query language.
- It allows you to create database in the cloud that are manged by Aws
  - Postgres
  - MySQL
  - MariaDB
  - Oracle
  - Microsoft SQL Server
  - Aurora (Aws proprietary database)

⇒ Advantage over using RDS versus deploying DB on EC2

- RDS is managed service:
  - Automated provisioning, OS patching
  - Continuous backups and restore to specific timestamp (Point in Time Restore)!
  - Monitoring dashboards
  - Read replicas for improved read performance
  - Multi AZ setup for DR (Disaster Recovery)
  - Maintence Windows for upgrades
  - Scaling Capability (vertical and horizontal)
  - Storage backed by EBS (gp2 or io1)

- But you can't SSH into your instances.

RDS Solution Architecture



DATE 16/12/21

Amazon Aurora

- Aurora is a proprietary technology from AWS (not open source).
- PostgreSQL and MySQL are both supported as Aurora DB.
- Aurora is "AWS cloud optimized" and claims 5x performance improvement over MySQL on RDS, over 3x the performance of postgres on RDS
- Aurora storage automatically grows in increments of 10GB, upto 64TB.
- Aurora costs more than RDS (20% more) - but is more efficient
- Not in free tier