Why Cloud Services & what is cloud?

- Machines
- Servers
- Networks
- Databases

To host or execute any application/program we need the infrastructure & supportive services.

Ex: Let say we want to store the data.

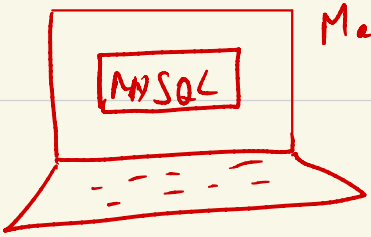
RDBMS → These are the application which helps to manipulate data.

- MySQL
- Postgres
- Oracle

Traditional ways (On Prem Solution)

Cost >> Security <<

- Infrastructure (Physical machine/servers)
- Application



Machine

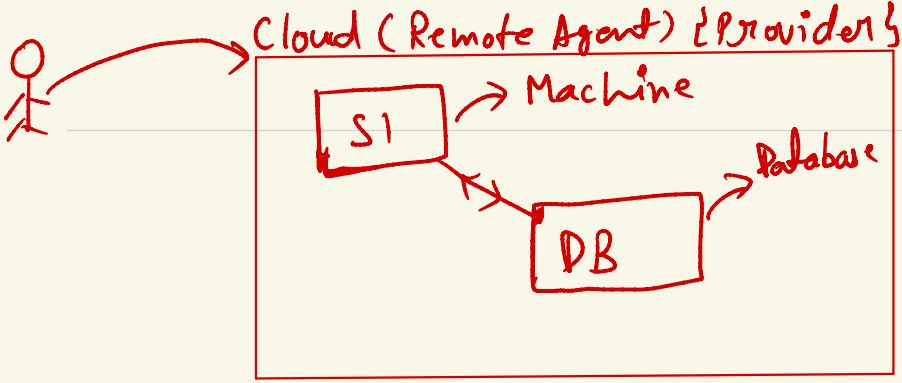
Disadvantages of Traditional Development?

- Manual setup
- Physical machine maintainence
- Cost is very high
- Security flaws will be higher

<u>Cloud</u>: A remote location or a remote agent which will provide all necessary services to build/host/maintain application

Ex: I want to build an web app then what all things we need?

- Some machines
- Database service



Why Cloud?

- ↳ Maintainence is quick
- ↳ Cost is cheap
- ↳ Security is good

Cloud Providers

- Amazon Web Services (AWS) [Amazon]
- Azure [Microsoft]
- Google Cloud platform (GCP) [Google]

- ✓ AliCloud (Alibaba)
- ✓ Oracle Cloud
- ✓ IBM Cloud

{Based on market cap}

AWS > Azure > GCP

Ex: Storage

↳ AWS S3
↳ Azure Blob
↳ GCP Storage

Common Services

How they are diff?

↳ Pricing
↳ Infra
↳ Implementation

Service or Consulting firm

✓ Cloud Data Engineer
✓ AWS Data Eng
✓ AZURE Data " }
✓ GCP Data "

Type of AWS Services

- Computing Services
  - EC2 (Elastic Cloud Compute)
  - Lambda (Serverless Service)

- Storage
  - S3 (Simple Storage Service)

- Database Services
  - AWS RDS (Relation Database Service)
  - AWS Athena (Serverless DB)

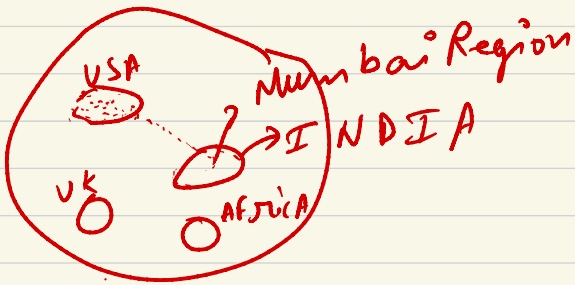
- NoSQL Database
  - AWS Dynamo (Key-Value type of DB)

- Data Warehousing
  - Redshift (Very popular)

- Analytic
  - EMR (Elastic Map Reduce)

- Glue (Distributed)
- Kinesis (for streaming)
- SQS (Simple Queue Service)
- Quick Sight (Dashboarding)

- Monitorin & Scheduling
  - CloudWatch
- Notification
  - SNS (Simple Notification Service)
- Access Management
  - IAM (Identity & Access Management)



S3 Objects

↳ uploaded or created data/files

S3 URI (Universal Resource Identifier)

- file:/// local
- hdfs:/// HDFS
- s3:// S3

$\underbrace{\text{s3://}}_{\text{file system name}}\underbrace{\text{batchJS3bucket}}_{\text{Bucketname}}\underbrace{\text{/employee_test/}}_{\text{Prefix}}\underbrace{\text{employee.csv}}_{\text{objectname}}$

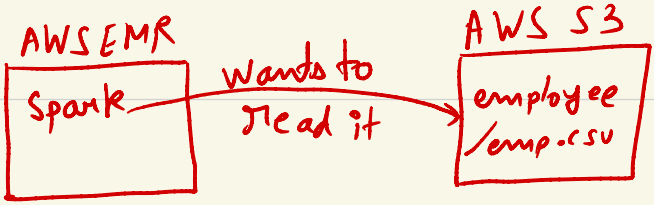
Key → prefix + object_name

Key → employee_test/employee.csv

Needed to locate each object in S3

Amazon Resource Name (ARN)

arn:aws:s3:::batchJS3bucket/employeetest/employee.csv


AWS EMR
Spark
wants to
read it
AWS S3
employee
/emp.csv