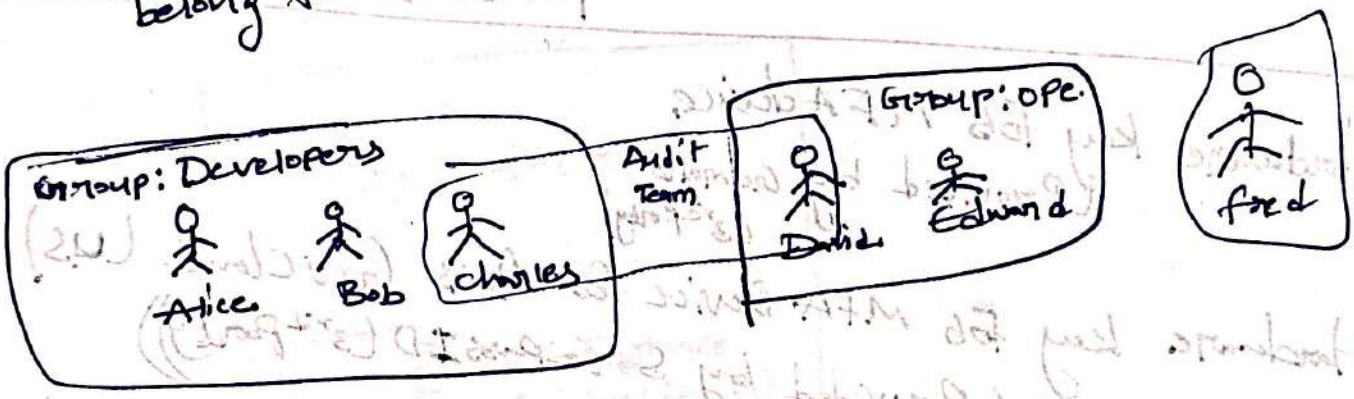
IAM - Identity & Access Management

→ Users are people within your organization, & can be grouped

→ Groups only contain users, not other groups.

→ Users don't have to belong to a group, & user can belong to multipk Groups.



IAM - Permissions

IAM Service is not region specific it is Global

Administrator Access - Json

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "*",
      "Resource": "*"
    }
  ]
}
```

IAM ReadOnly Accesy - JSON

```
{
  "version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        - - - - ,
        - - - - ,
        - - - - ,
      ]
      "Resource": "*".
    }
  ]
}
```

MFA device options in AWS

→ Virtual MFA device:

1. Google Authenticator (Phone only)
2. Authy (Phone only)

Support for multiple tokens on a single device

→ Universal 2nd Factor (U2F) Security key

yubikey by Yubico (3rd Party)

Support for multiple root & IAM users using a single security key.

→ Hardware key fob MFA device (Provided by Gemalto (3rd party))

→ Hardware key fob MFA Device for AWS GovCloud (US) (Provided by SurePassID (3rd party))

How can users access AWS)

- To access AWS, you have three options:
  - AWS Management Console (Protected by password + MFA)
  - AWS Command Line Interface (CLI): Protected by access keys
  - AWS Software Development Kit: for code: Protected by Access keys.

→ AWS CLI - Command Line Interface.

~ aws configure

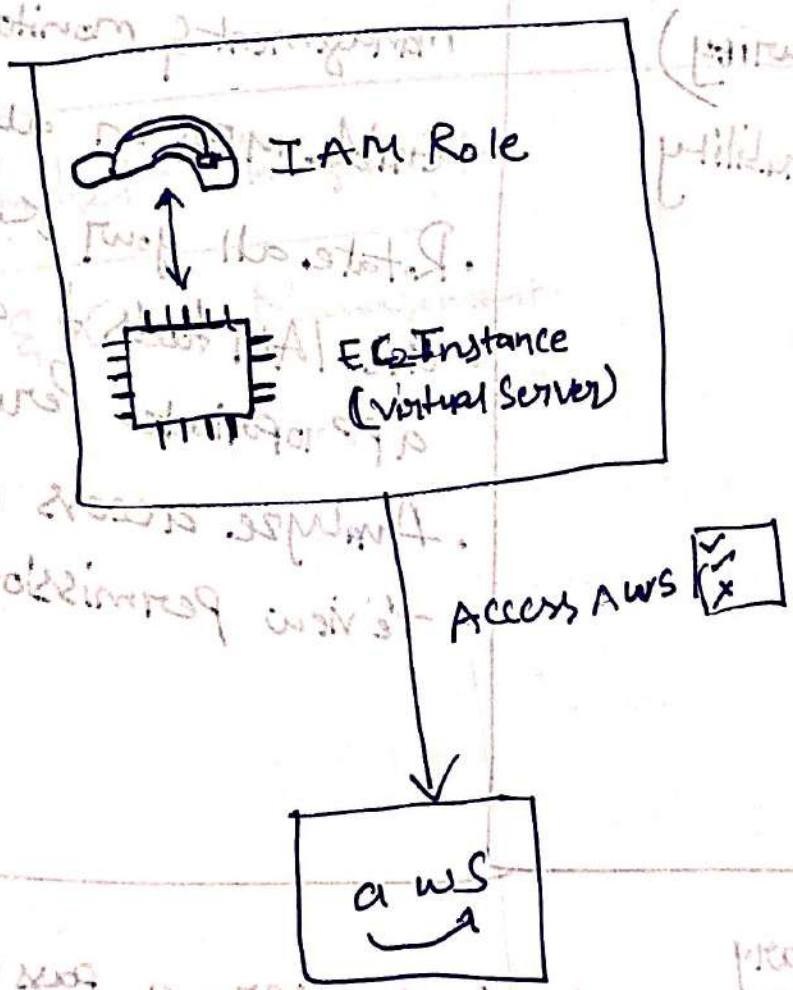
AWS Access key ID [None] : - - -

AWS Secret Access key [None] : - - -

Default region name [None] : eu-west-1

✓

IAM Roles



IAM Roles

Shared responsibility Model for IAM

Aws	You
• Infrastructure (global network Security) • Configuration & Vulnerability analysis • Compliance Validation	• Users, Groups, Roles, Policies management & monitoring • Enable MFA on all accounts • Rotate all your keys often • Use IAM tools to apply appropriate permissions • Analyze access patterns & review permissions.

IAM Section - Summary

- Users - mapped to a physical user, a password for AWS console
- Groups - Contains users only
- Policies - JSON document that outlines permission for users or groups
- Roles - for EC2 Instances or AWS Services
- Security - MFA + Password Policy

Aws CLI : Manage your Aws services using the Command-line

Aws SDK : Manage your Aws services using a Programming language

Audit: IAM Credentials Reports & IAM Access Advisor.

Aws Budget Management

Billing & Cost Management.

Amazon EC2

→ EC2 - Elastic Compute Cloud = Infrastructure as a Service

→ It mainly consists in the capability of:

- Renting Virtual Machines (EC2)
- Storing data on virtual drives (EBS)
- Distributing load across machines (ELB)
- Scaling the services using an Auto-Scaling group (ASG)

EC2 Sizing & Configuration options

- Operating System (OS): Linux, Windows or Mac OS.
- How much Compute power & Cores (CPU)
- How much Random-access Memory (RAM)
- How much Storage Space:
  - Network attached (EBS & EFS)
  - Hardware
- Network Card: Speed of the card, Public IP address.
- Firewall rules: Security Group.
- Bootstrap Script (configure at first launch): EC2 User Data

EC2 User data is used to automate boot tasks

Such as:

- Installing updates.
- Installing Software.
- Downloading Common files from the internet.
- Anything you can think of.

EC2 Instance type : example.

Instance	VCPU	Mem(GiB)	Storage	Network Perf.	EBS Bw (Mbps)
t2.micro	1	1	EBS-only	Low to mode	
t2.xlarge	4	16	EBS-only	Moderate	
C5d.4xlarge	16	32	1x400 NVMe SSD	up to 10Gbps	4,750
r5.16xlarge	64	512	EBS only	20 Gbps	13600
m5.8xlarge	32	128	EBS only	10Gbps	6,800

EC2 Instance Types

instances.vantage.sh

1. General purpose
2. Compute optimized
3. Memory optimized
4. Accelerated Computing
5. Storage optimized
6. HPC optimized
7. Instance features
8. Measuring instance performance

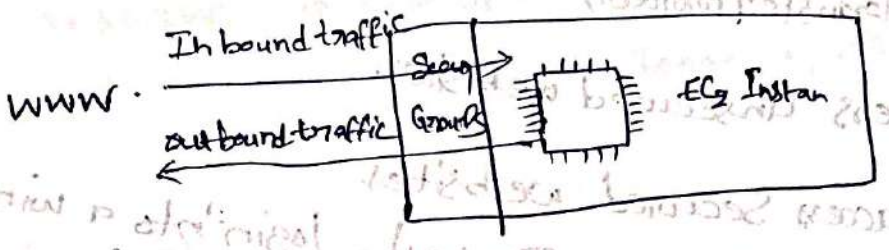
m5.2xlarge

m : instance class

5 : generation (AWS improves them over time)

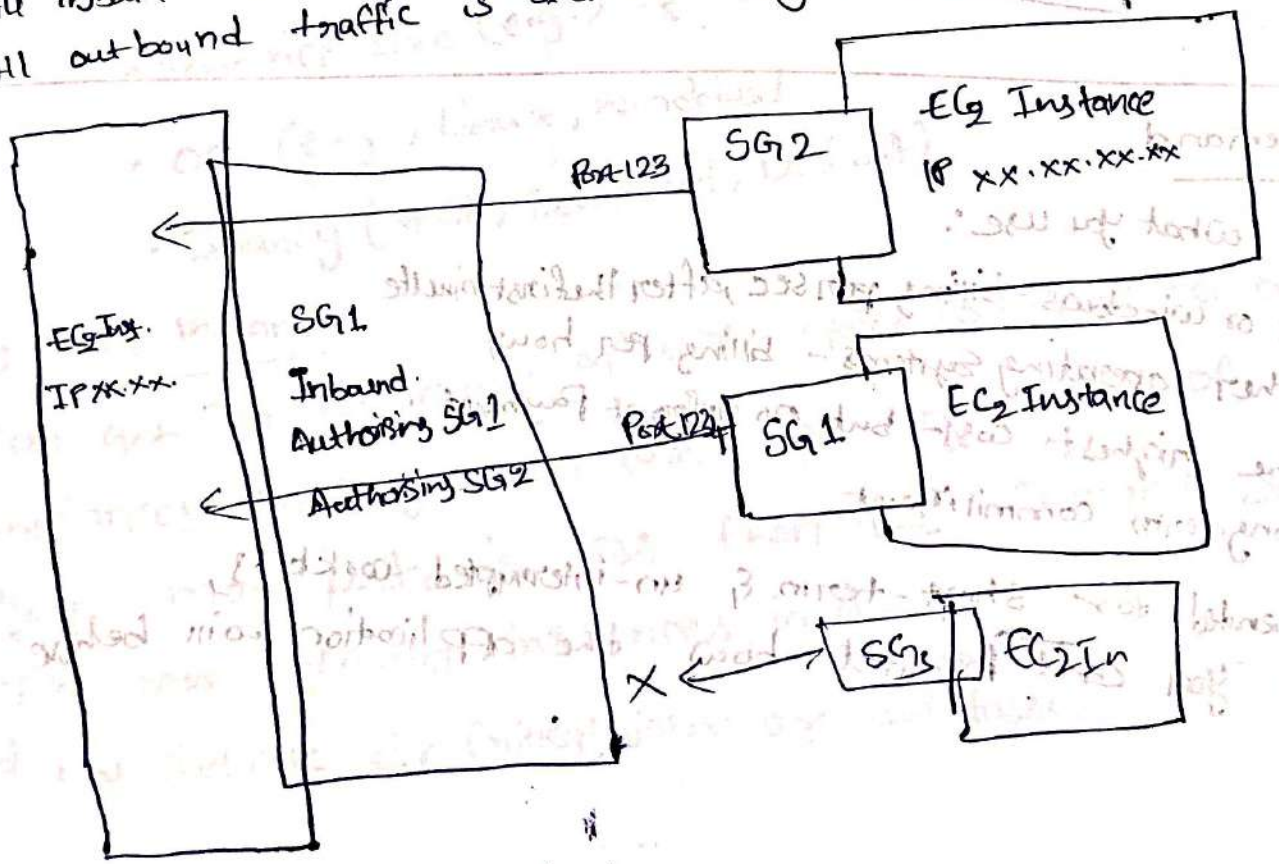
2xlarge : size within the instance class.

Security Groups



Type	Protocol	Port Range	Source	Description
Http	TCP	80	0.0.0.0/0	
SSH	TCP	22	122.149.196.85/32	
Custom TCP Rule	TCP	4567	0.0.0.0/0	

All inbound traffic is blocked by default.
All outbound traffic is authorised by default.



Classic Ports to Know

- 22 = SSH (Secure Shell) - log into a Linux instance.
- 21 = FTP (File Transfer Protocol) - upload files into a file share.
- 80 = HTTP - access unsecured websites.
- 443 = HTTPS - access Secured websites.
- 3389 = RDP (Remote Desktop Protocol) - login into a windows instance.

SSH - Secure Shell

→ SSH is one of the most important function. It allows you to control a remote machine, all using the command line.

SSH

```
ssh -i EC2Tutorial.pem ec2-user@3.250.26.200
```

EC2 On-demand

- Pay for what you use:
  - Linux or windows - billing per sec, after the first minute
  - All other operating systems - billing per hour
- Has the highest cost - but no upfront payment.
- No long-term commitment
- Recommended for short-term & un-interrupted workloads, where you can't predict how the application will behave.

EC2 Reserved instances

- Upto 72% discount compared to on-demand.
- you reserve a specific instance attributes (Instance Type, Region, Tenancy, OS)

EC2 Savings Plans

- Get a discount based on long-term usage (upto 72% - Same as RI)
- Commit to a certain type of usage ($10/hour for 1 or 3 yrs)
- Usage beyond EC2 savings plans is billed at the On-Demand price.
- Locked to a specific instance family & AWS region (e.g., M5 in us-east-1)
- Flexible across:
  - Instance size (e.g., m5.xlarge, m5.2xlarge)
  - OS (e.g., Linux, Windows)
  - Tenancy (Host, Dedicated, Default)

EC2 - Spot Instances

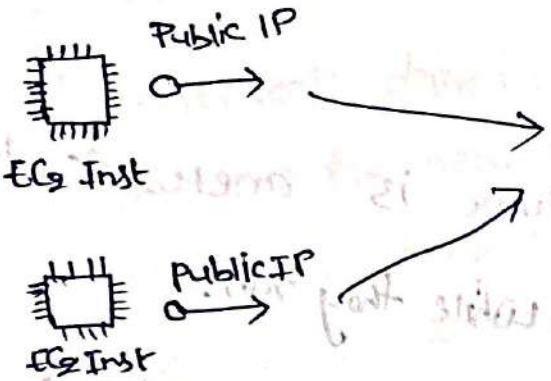
- Can get a discount of upto 90% compared to on-demand
- Instances that you can lose at any point of time if your max price is less than the current Spot time.
- The most-efficient instances in AWS.
- Not suitable for Critical jobs or databases.

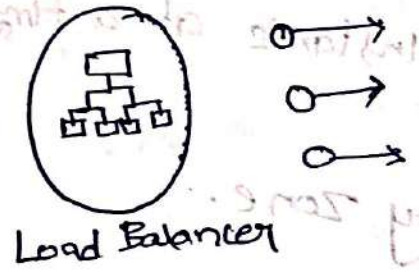
• EC2 Dedicated Hosts

- On demand: Coming & Staying in resort whenever we like, we pay the full price
- Reserved: like planning ahead & if we plan to stay for a long time, we may get a good discount.
- Savings plans: pay a certain amount per hour for certain period & stay in any room type.
- Spot instances: the hotel allows people to bid for the empty rooms & the highest bidder keeps the rooms. you can get kicked out at any time.
- Dedicated ~~instances~~ Hosts: we book an entire building of the resort
- Capacity Reservations: you book a room for a period with a full price even you don't stay in it.

Aws Charges for IPV4 addresses.

- Starting Feb 1st 2024, there's a charge for all Public IPV4 created in your account.
- $0.005 per hour of Public IPV4 (~$3.6 per month)
- for new accounts in AWS, you have a free tier for the EC2 ~~Instance~~ service: 750 hours of public IPV4 per month for the first 12 months.
- For all other services there is no free tier.





Share Responsibility model for EC2	
aws	user
• Infrastructure (global network security) • Isolation on physical hosts • Replacing faulty hardware • Compliance Validation	• Security Group rules. • operating-system patches & updates • Software & utilities installed on the EC2 instance. • IAM Roles assigned to EC2 & IAM user access management. • Data Security on your Instance.

EC2 Instance Storage

What's an EBS Volume?

- An EBS (Elastic Block Store) Volume is a network drive you can attach to your instances while they run.
- It allows your instances to persist data, even after their termination.
- They can only be mounted to one instance at a time (at the CCP level)
- They are bound to a specific availability zone.

Note: CCP - Certified Cloud practioner - one EBS can be only mounted to one EC2 instance. Associate level (solutions Architect, Developer, Sysops) 'multi-attach' feature for some EBS.

Free tier: 30 GB of free EBS storage of type General Purpose (SSD) or Magnetic per Month.

EBS (Elastic Block Store)

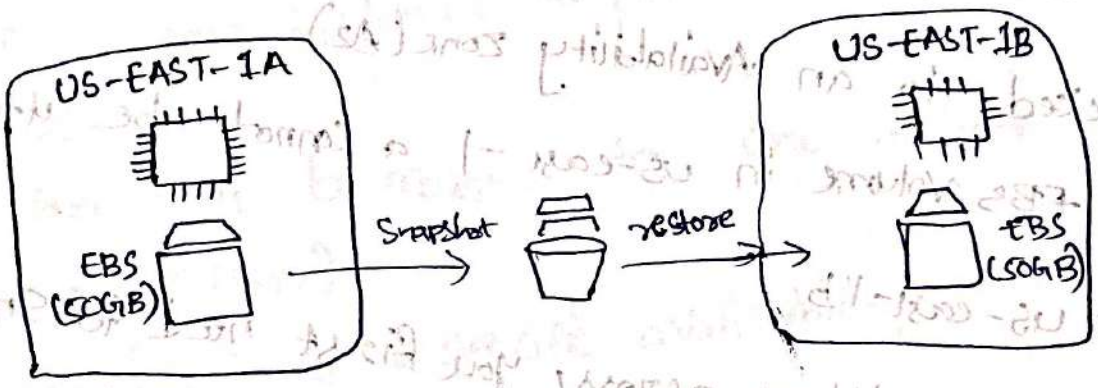
- It's a network drive (i.e. not a physical drive)
  - It uses the network to communicate the instance, which means there might be a bit of latency.
  - It can be detached from an EC2 Instance & attached to another one quickly.
- It's locked to an Availability zone (AZ)
  - An EBS volume in us-east-1 a cannot be attached to us-east-1b
  - To move a volume across, you first need to snapshot it
- Have a provisioned capacity (size in GBs & IOPS)
  - you get billed for all the provisioned capacity.

EBS (Elastic Block Store) - Delete on Termination attribute

- Controls the EBS behaviour (when an EC2 instance terminates.)
  - By default, the root EBS volume is deleted (attribute enabled)
  - By default, any other attached EBS volume is not deleted (attribute disabled)
- This can be controlled by the AWS console / AWS CLI.
- Use case: preserve root volume when instance is terminated

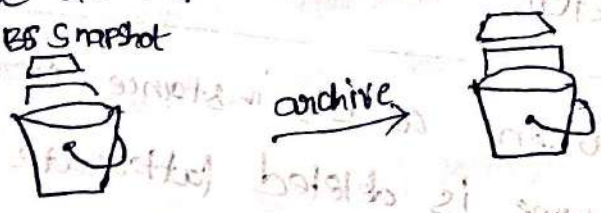
EBS Snapshots

- Make a backup (snapshot) of your EBS volume at a point in time.
- Not necessary to detach volume to do snapshot, but recommended.
- Can copy snapshots across AZ or region.



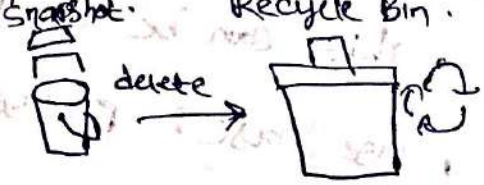
EBS (Elastic Block Store) Snapshots feature.

- EBS Snapshot Archive.
  - Move a snapshot to an archive tier that is 75% cheaper.



  - Takes within 24 to 72 hours for restoring.

- Recycle Bin for EBS Snapshots.
  - Setup rules to retain deleted snapshots, so you can recover them after an accidental deletion.
  - Specify retention (from 1 day to 1 year).



AMI - Amazon Machine Image

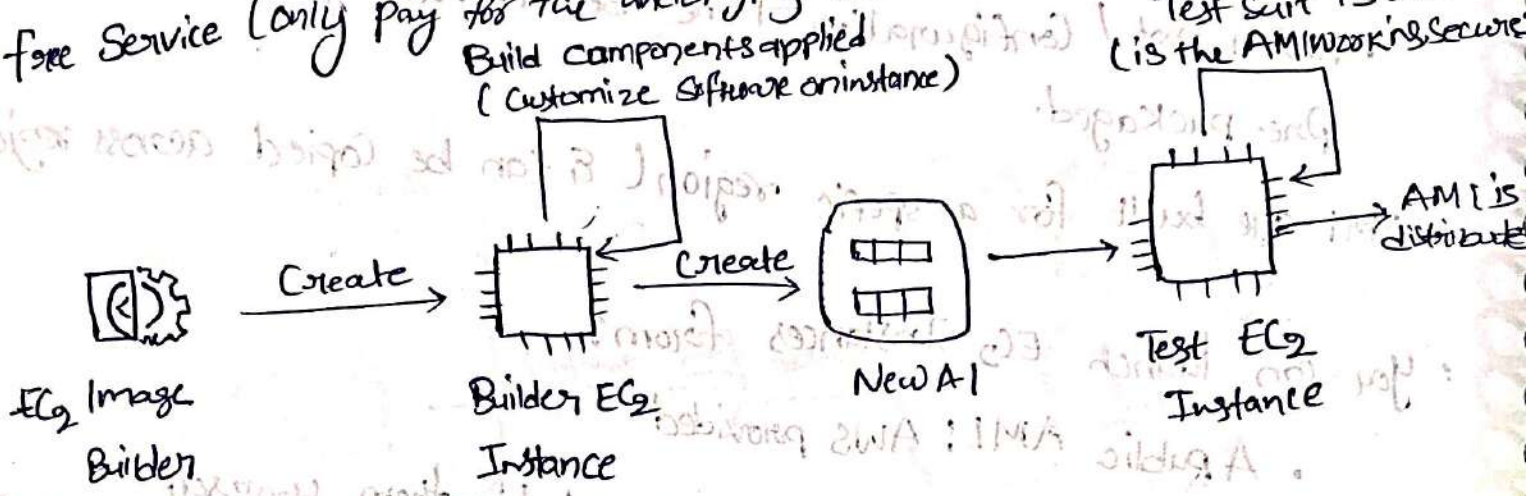
- AMI are a Customization of an EC_2 Instance.
- AMI are a Customization of an EC_2 Instance.
  - you add your own Software, Configuration, OS, monitoring...
  - Faster boot / Configuration time because all your Software is Pre-packaged.
- AMI are built for a spefic region (& can be copied across regions)
- You can launch EC_2 Instances from:
  - A public AMI: AWS provided.
  - Your own AMI: you make & maintain them yourself.
  - An AWS MarketPlace AMI: an AMI someone else made (& potentially Sells)

AMI Process (from an EC_2 Instance)

- Start an EC_2 Instance & customize it.
- Stop the Instance (for data integrity)
- Build an AMI - this will also.

EC2 Image Builder

- Used to automate the creation of Virtual Machines or Container images
- Automate the creation, maintain, validate & test EC2 AMIs
- Can be run on a schedule (weekly, whenever Packages are updated, etc)
- Free Service (only pay for the underlying resources)

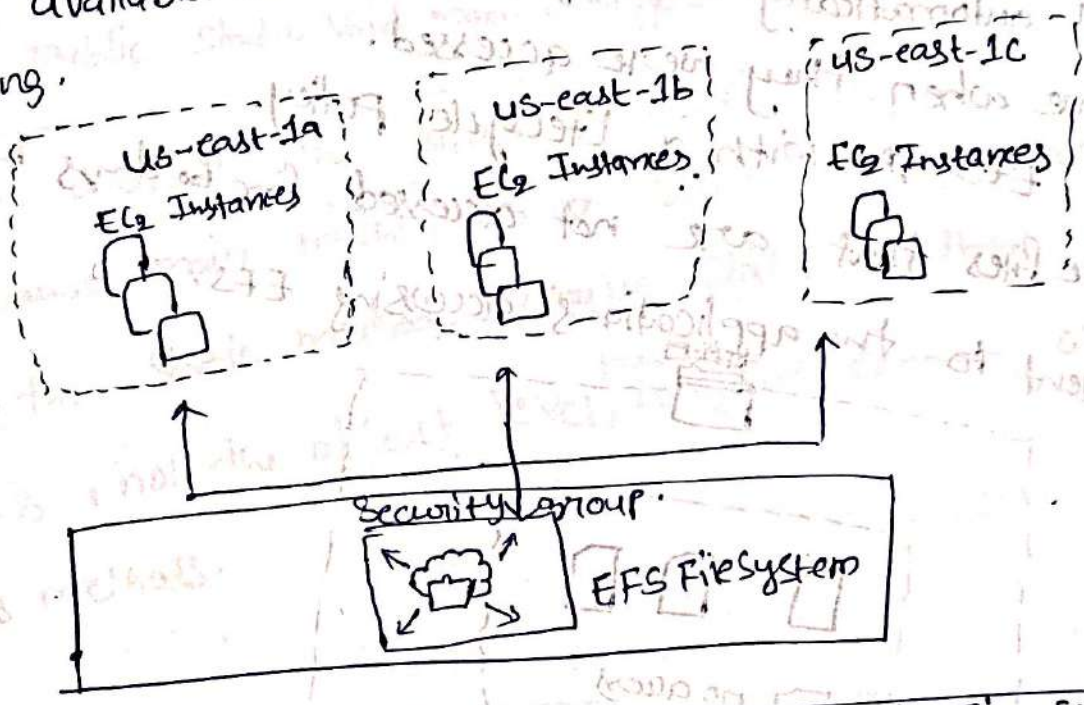


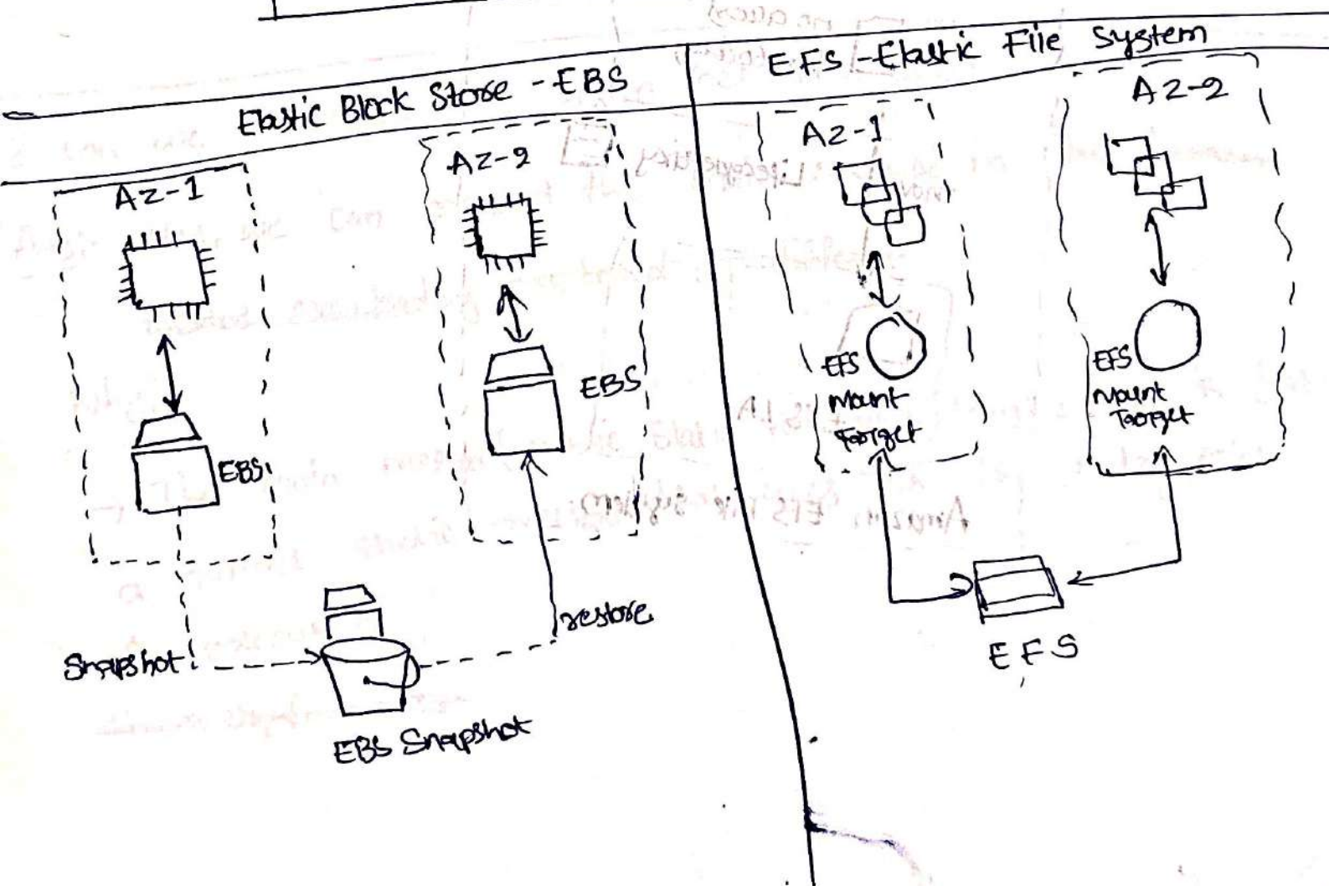
EC2 Instance Store

- EBS volumes are network drives with good but "limited" performance.
- If you need a high performance hardware disk, use EC2 Instance store.
- Better I/O performance.
- EC2 Instance store lose their storage if they're stopped (ephemeral)
- Good for buffer / cache / scratch data / temporary content.
- Risk of data loss if hardware fails.
- Backups & Replication are your responsibility.

EFS (Elastic File System)

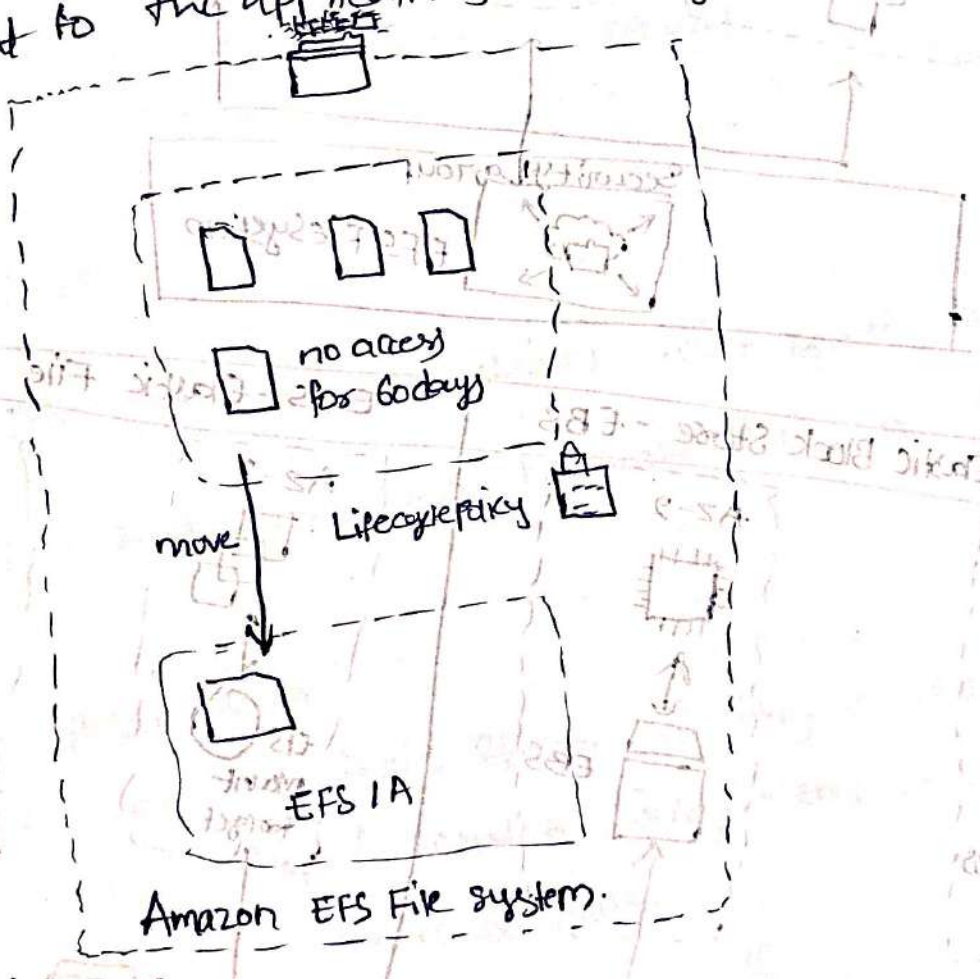
- Managed NFS (Network file System) that can be mounted on 100s of EC2
- EFS only works with Linux EC2 Instances in multiAZ.
- Highly available, scalable, expensive (3x gp2), Pay per use, no capacity planning.



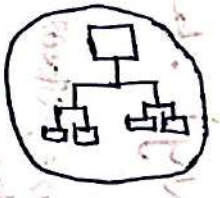


EFS Infrequent Access (EFS - IA)

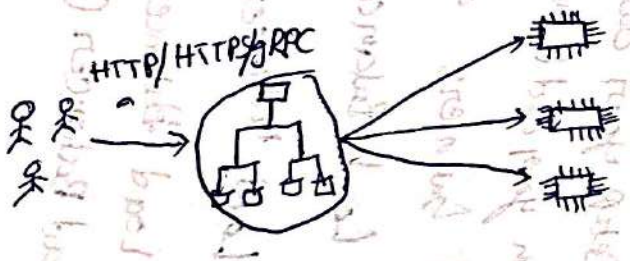
- Storage class that is cost-optimized for files not accessed every day.
- upto 92% lower cost compared to EFS Standard.
- EFS will automatically move your files to EFS-IA based on the last time when they were accessed.
- Enable EFS-IA with a Lifecycle policy.
- Eg: move files that are not accessed for 60 days to EFS-IA
- Transparent to the applications accessing EFS.



Application Load Balancer

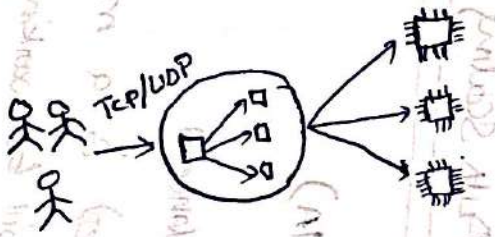


- HTTP/HTTPS / gRPC protocols (layer 7)
- HTTP Routing features.
- Static DNS (URL)



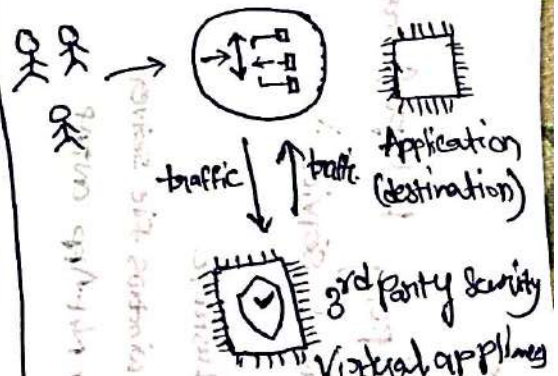
Network Load Balancer

- TCP/UDP protocols (Layer 4)
- High performance: millions of request per seconds.
- Static IP through Elastic IP.



Gateway Load Balancer

- GENEVE protocol on IP Packets (layer 3)
- Route Traffic to firewalls that you manage on EC2 Instances
- Intrusion detection.



Auto Scaling Group (ASG)

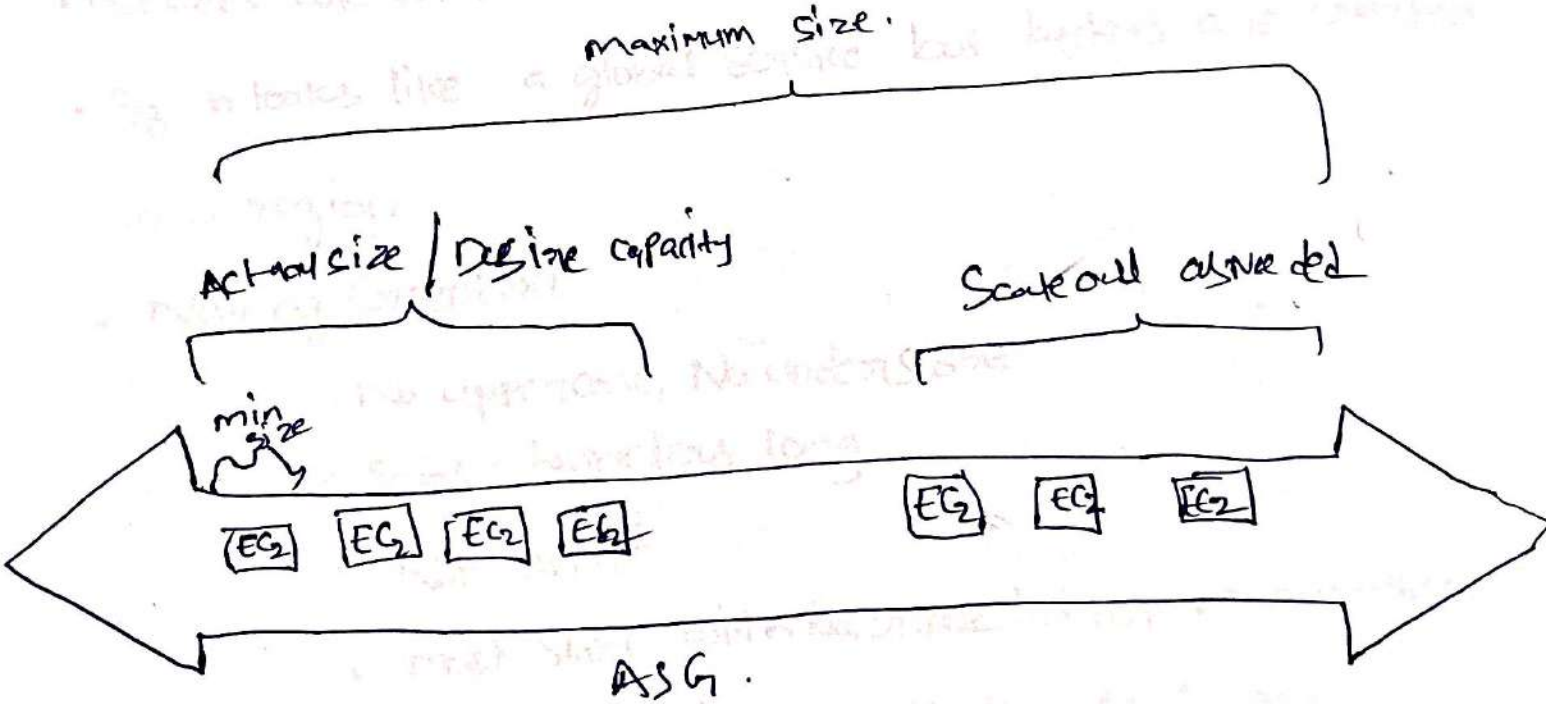
→ In real-life the load on your websites & application can change.

→ In the cloud, you can create & get rid of servers very quickly.

→ The goal of an ASG is to:

- Scale out (add EC2 instances) to match an increased load.
- Scale in (remove EC2 instances) to match a decreased load.
- Ensure we have a min & a max no. of machines running.
- Automatically register new instances to a load balancer
- Replace unhealthy instances.

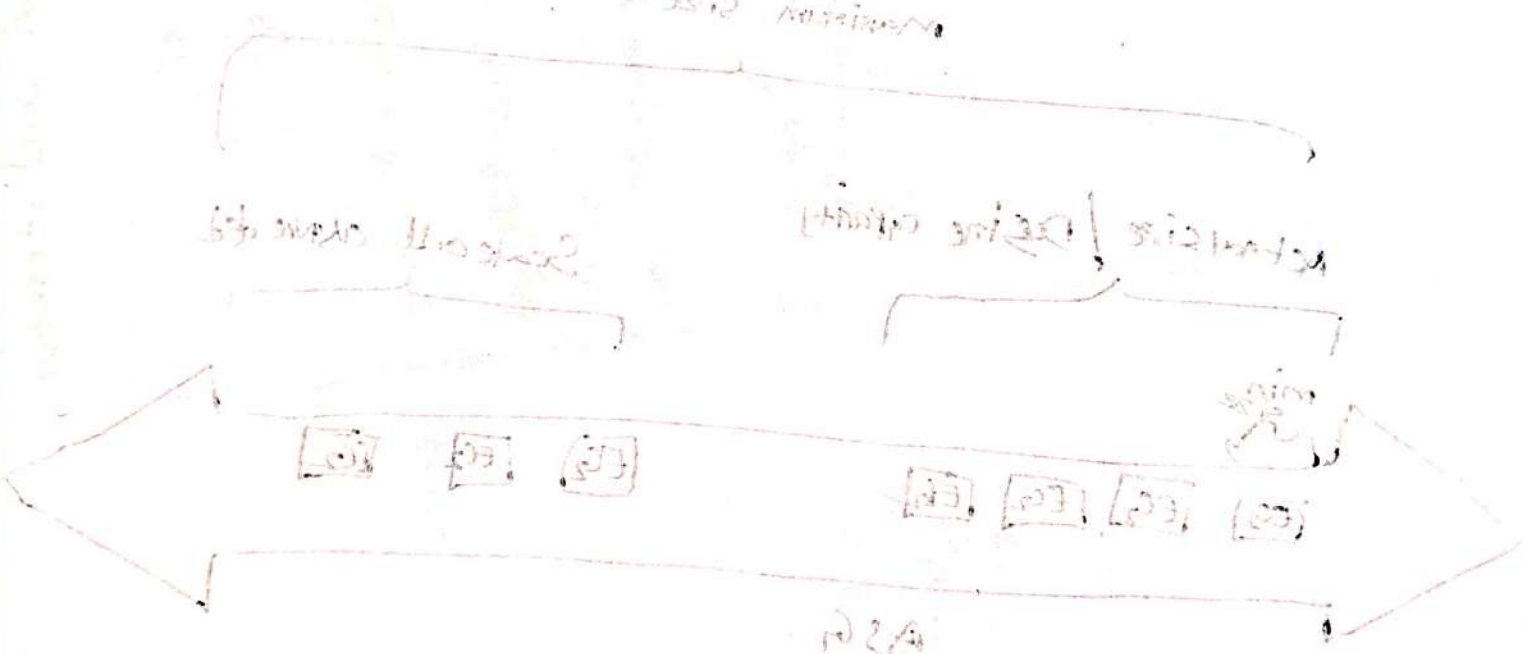
→ Cost Savings: only run at an optimal capacity (principle of the cloud).



Auto scaling Groups – Scaling Strategies.

- Manual Scaling: update the size of an ASG manually
- Dynamic Scaling: Responding to changing demand.
  - Simple / Step Scaling
    - when a cloudwatch alarm is triggered (ex: CPU > 70%) then adds 2 units
    - when a cloudwatch alarm is triggered (ex: CPU < 30%) then ~~adds~~ remove 1

- Target Tracking Scaling:
  - Eg: I want the average ASG CPU to stay at around 40%
- Scheduled Scaling:
  - Anticipate a scaling based on known usage patterns
  - Eg: increase the min capacity to 10 at 5pm on Fridays.
- Predictive Scaling:
  - Uses Machine Learning to predict future traffic ahead of time.



Amazon S3 Bucket

→ Amazon S3 is one of the main building blocks of AWS

→ It's advertised as "infinitely scaling" storage.

⇒

Amazon S3 - Buckets.

- Amazon S3 allows people to store objects (files) in "buckets" (directories)
- Buckets must have a globally unique name (across all regions all accounts)
- Buckets are defined at the region level.
- S3 is looks like a global service but buckets are created in a region.
- Naming Convention.
  - No uppercase, No underscore
  - 3-63 characters long
  - Not an IP
  - must start with lowercase letter or number.
  - must Not start with the prefix xn--
  - must Not end with the suffix -s3alias.

Amazon S3 - objects.

- object values are the content of the body:
  - Max object size is 5 TB (5000 GB)
  - If uploading more than 5GB, must use "multi-part" upload.

Amazon S3 - Security.

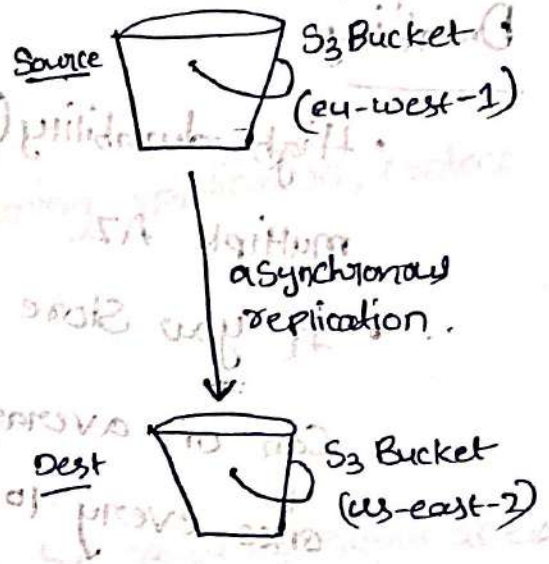
/

Amazon S3 - Versioning

- you can version your files in Amazon S3
- It is enabled at the bucket level
- Same key overwrite will change the "version": 1, 2, 3, ...
- It is best practice to version your buckets.
  - protect against unintended deletes (ability to restore a version)
  - Easy roll back to previous version
- Notes:
  - Any file that is not versioned prior to enabling versioning will have version "null".
  - Suspending versioning does not delete the previous versions.

Amazon S3 - Replication (CRR & SRR)

CRR - Cross Region Replication
SRR - Same Region Replication.



- must enable Versioning in source & destination buckets
- Cross Region Replication (CRR)
- Same Region Replication (SRP)
- Buckets can be in different AWS accounts.
- Copying is asynchronous.
- Must give proper IAM Permissions to S3.

- Use Cases:
  - CRR - compliance, lower latency access, replication across accounts.
  - SRR - log aggregation, live replication b/w prod & test accounts.

S3 - Storage Classes

- Amazon S3 Standard - General Purpose
- Amazon S3 Standard - Infrequent Access (IA)
- " " One zone - " "
- " " Glacier Instant Retrieval
- " " " Flexible Retrieval
- " " " Deep Archive
- " " Intelligent Tiering.

Note: Can move b/w classes manually or using S3 Lifecycle Configurations

S3 Durability & Availability

Durability

- High durability (99.999999999%, 11 9's) of objects across multiple AZ
- If you store 10,000,000 objects with Amazon S3, you can on average expect to incur a loss of single object once every 10,000 years.
- Same for all storage classes.

Availability

- Measures how readily available a service is
- Varies depending on storage class.
- Eg:- S3 Standard has 99.99% availability = not available S3 min a yr.

S3 Standard - General purpose

- 99.99% Availability
- Used for frequently
- Low latency & high throughput
- Sustain 2 concurrent facility failures.
- Usecases: Big Data Analytics, mobile & gaming application, content distribution.

S3 storage classes - Infrequent Access

- For data that is less frequently accessed, but requires rapid access when needed.
- Lower cost than S3 standard.
- Amazon S3 Standard - Infrequent Access (S3 Standard - IA)
  - 99.9% Availability
  - Use cases: Disaster Recovery, backup.
- Amazon S3 one zone - Infrequent Access (S3 one zone - IA)
  - High durability (99.999999%) in a single AZ; data lost when AZ is destroyed.
  - 99.5% Availability.
  - Use cases: Storing secondary backup copies of on-premise data, or data you can recreate.

Amazon S3 Glacier Storage classes.

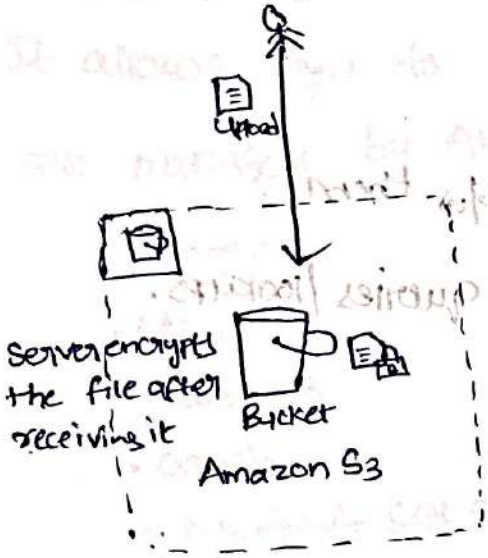
- Low-cost object storage meant for achiving / backup
- pricing: price for storage + object retrieval cost.
- **Amazon S3 Glacier Instant Retrieval**
  - millisecond retrieval, great for data accessed once a quarter
  - Minimum storage duration of 90 days.
- **Amazon S3 Glacier Flexible Retrieval** (formerly Amazon S3 Glacier):
  - Expedited (1 to 5 min), Standard (3 to 5 hrs), Bulk (5-12) - free
  - Minimum storage duration of 90 days.
- **Amazon S3 Glacier deep Archive** - for long term storage:
  - Standard (12 hrs), Bulk (48 hrs)
  - Min. storage duration of 180 days.

S3 Intelligent Tiering

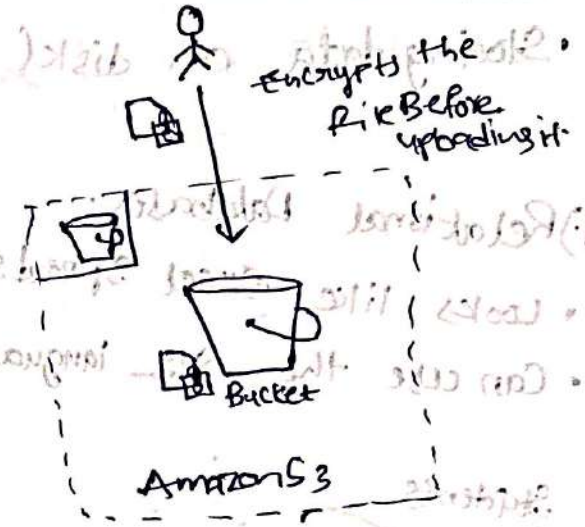
- Small monthly monitoring & auto-tiering fee.
- moves objects automatically b/w Access Tiers based on usage.
- There are no retrieval charges in S3 Intelligent-Tiering.
- Frequent Access tier (automatic): default tier
- Infrequent Access tier (automatic): objects not accessed for 30 days.
- Archive Instant Access tier (automatic): object not accessed for 90 days.
- Archive Access tier (optional): configurable from 90 days to 700+ days
- Deep Archive Access tier (optional): config from 180 days to 700+ days

S3 Encryption

Server-Side Encryption (Default)



Client-Side Encryption



IAM Access Analyzer for S3

→ Ensures that only intended people have access to your S3 buckets

→ Eg: Publicly accessible bucket, bucket shared with other AWS account

→ Evaluates S3 Bucket policies, S3 ACLs, S3 Access point policies

→ Powered by IAM Access Analyzer

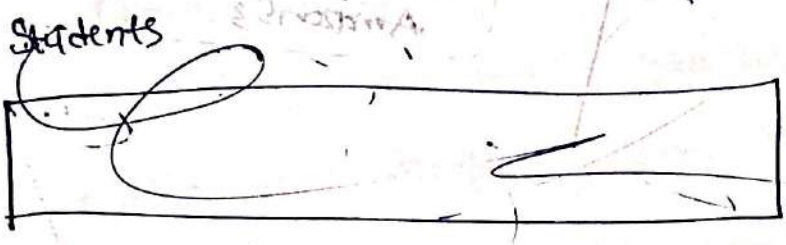
~~AWS Snowball~~

Databases & Analytics

- Storing data on disk(Efs, EBS, EC2

1) Relational Databases:

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



2) NoSQL Databases

- NoSQL = non-SQL = non relational databases.
- NoSQL databases are purpose built for specific data models & have flexible schemas for building modern applications.
- Benefits:
  - Flexibility: easy to evolve data model.
  - Scalability: designed to scale-out by using distributed clusters.
  - High-performance: optimized for a specific data model
  - Highly functional: types optimized for the data model
- Eg: key-value, document, graph, in-memory, search databases.

NoSQL data example: JSON *

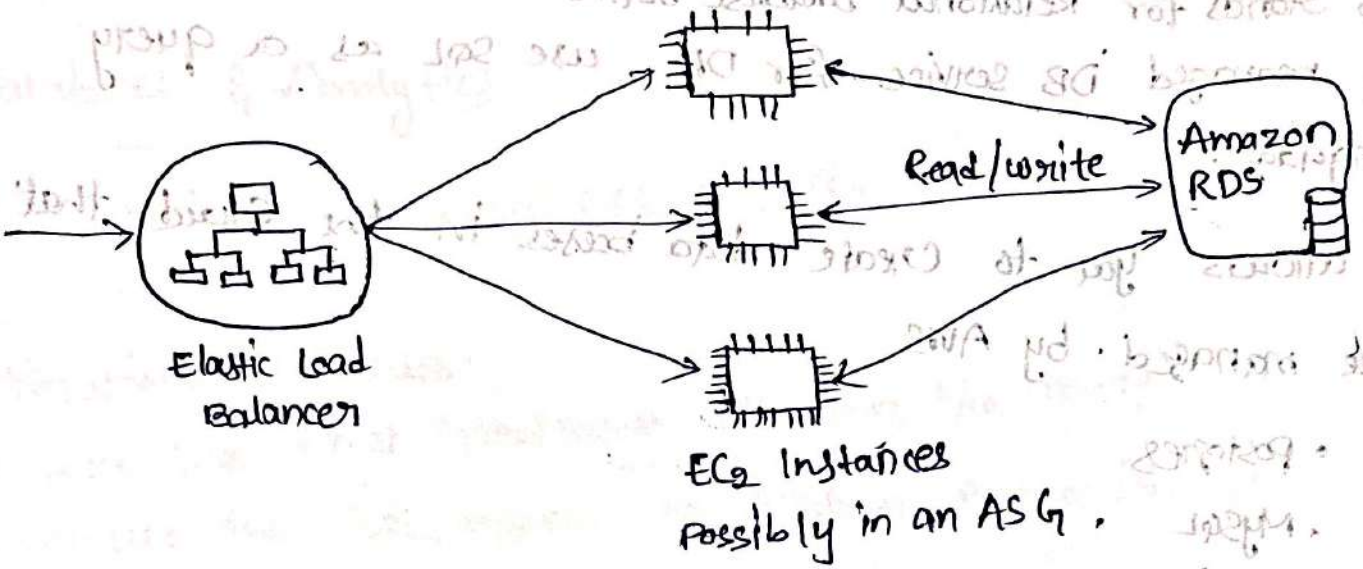
Amazon RDS

- RDS stands for Relational Database Service.
- It's managed DB service for DB use SQL as a query language.
- It allows you to create data bases in the cloud that are managed by AWS
  - Postgres
  - MySQL
  - MariaDB
  - Oracle
  - Microsoft SQL Server
  - IBM DB2.
  - Aurora (AWS proprietary Database)

Advantage over using RDS versus deploying DB on EC2

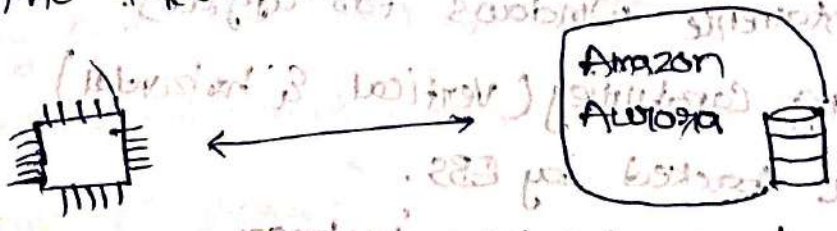
- RDS is a managed service.
  - Automated provisioning, OS patching.
  - Continous backups & restore to specific timestamp (Point in Time Restore)!
  - Monitoring dashboards.
  - Read replicas for improved read performance.
  - Multi AZ setup for DR (Disaster Recovery)
  - Maintainence windows for upgrades.
  - Scaling capabilities (vertical & horizontal)
  - Storage backed by EBS.
- But you can't SSH into your instances.

RDS (Relational Database Service)



Amazon Aurora

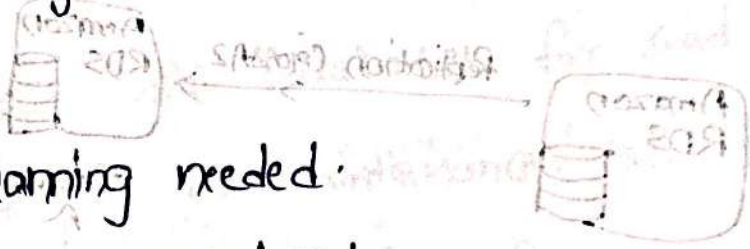
- Aurora is a proprietary technology from AWS (not open sourced)
- PostgreSQL & MySQL are both supported as AuroraDB.
- Aurora is "AWS Could optimized" & claims 5X performance improvement over mySQL on RDS, over 3x the performance of Postgres on RDS
- Aurora storage automatically grows in increments of 10GB upto 128TB.
- Aurora cost more than RDS (20% more) - but is more efficient
- Not in the free tier.



* Aurora is more cloud native, whereas RDS is going to be running the technologies

Amazon Aurora Serverless

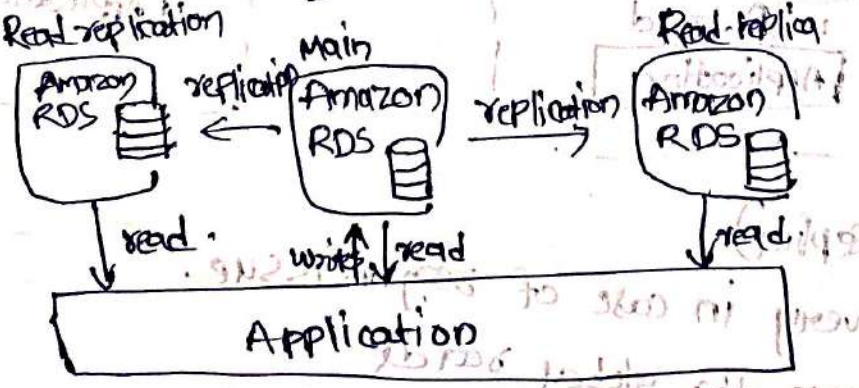
- Automated db instantiation & Auto-scaling based on actual usage.
- PostgreSQL & MySQL are both Supported as Aurora Serverless DB



- No Capacity planning needed.
- Least Management overhead.
- Pay per second, Can be more cost effective.
- Usecases: good for infrequent, intermittent or unpredictable workloads.

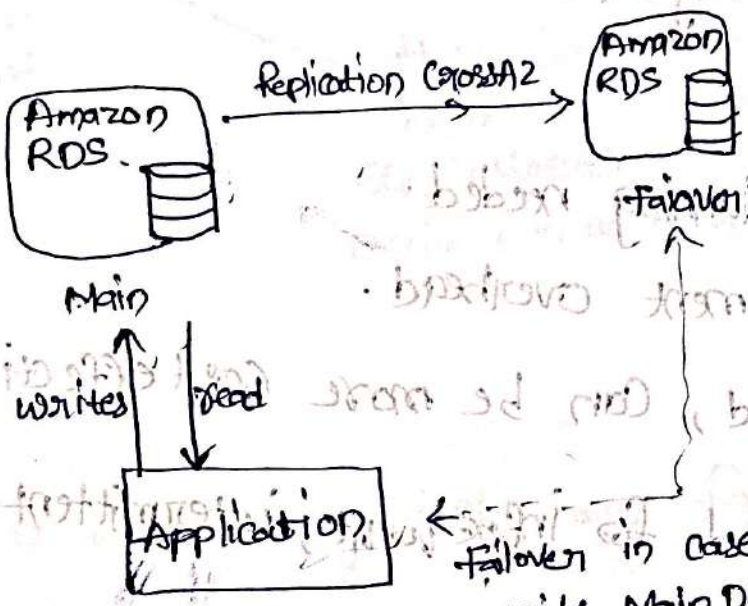
RDS Deployments: Read Replicas, Multi-AZ

- Read Replicas:
  - Scale the read workload of your DB
  - Can Create upto 15 Read Replicas
  - Data is only written to the main DB.

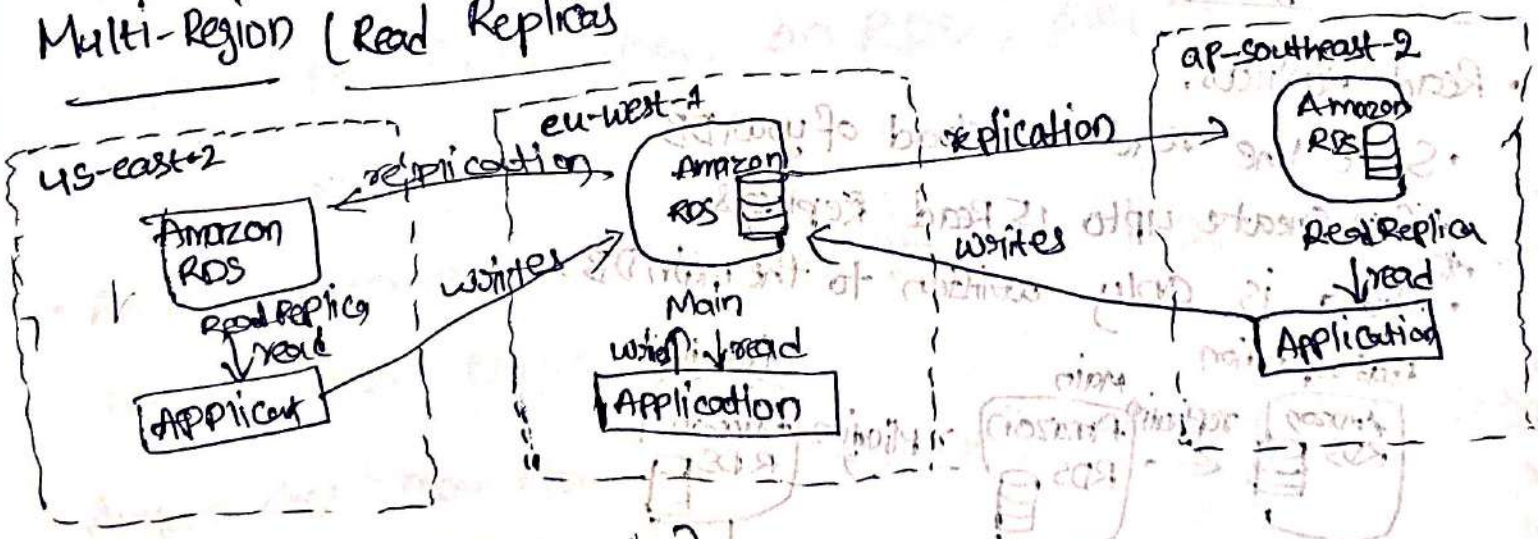


Multi-AZ:

- Failover in case of AZ outage (high availability)
- Data is only read/written to the main database
- Can only have 1 other AZ as failover.



Multi-Region (Read Replicas)



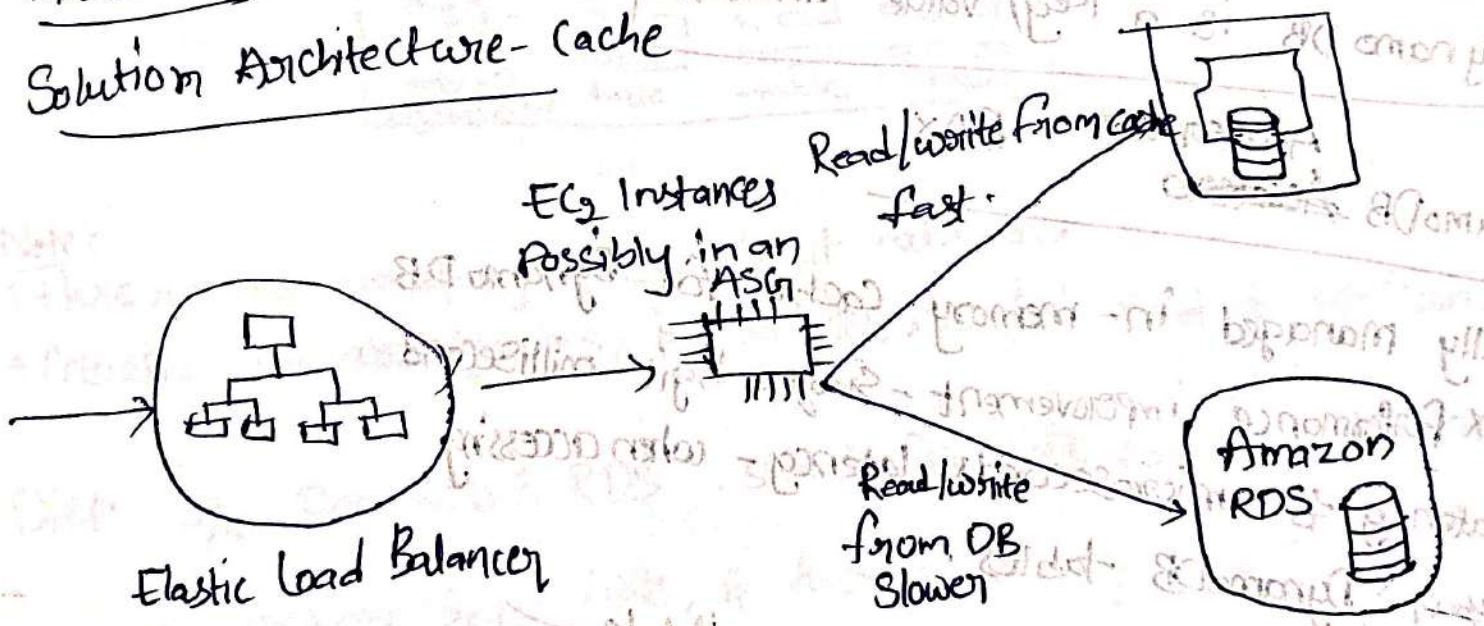
- Multi-Region (Read Replica)
  - Disaster recovery in case of region issue.
  - Local performance for global reads
  - Replication cost.

Amazon ElasticCache

- The same way RDS is to get managed Relational Databases ---
- Elastic Cache is to get managed Redis or Memcached
- Caches are in-memory databases with high performance, low latency
- Help reduce load off databases for read intensive workloads
- Aws take care of os maintenance / Patching, optimizations, Setup, Configuration, monitoring, failure recovery & backups

Elastic Cache

Solution Architecture - Cache



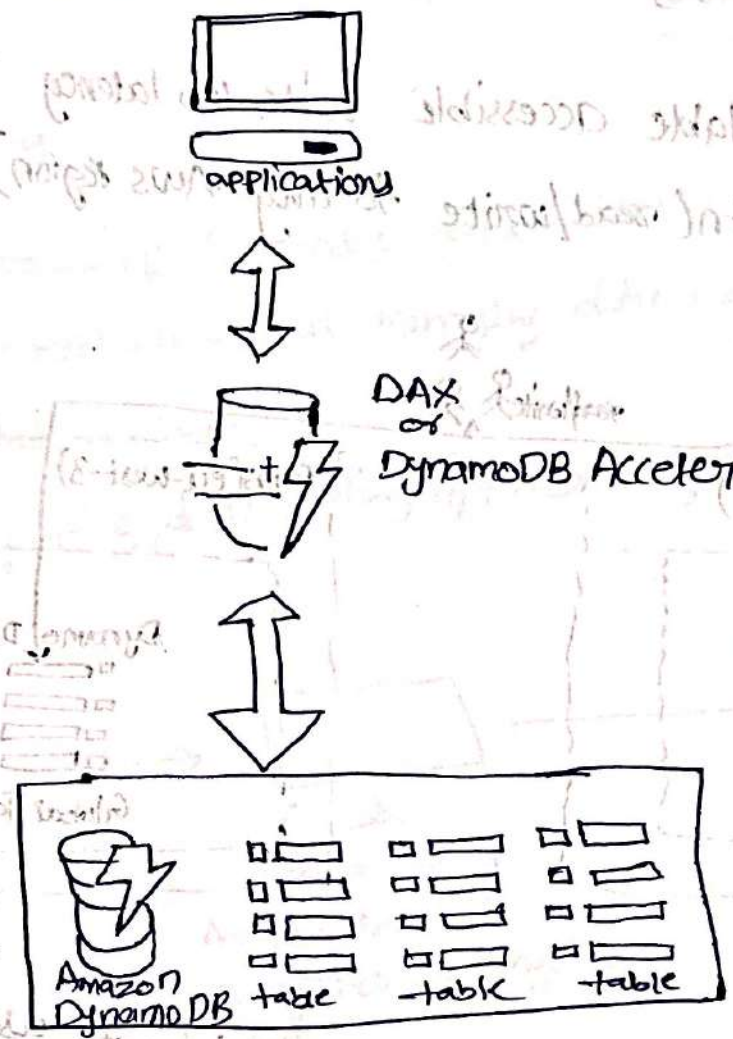
DynamoDB

- Fully managed highly available with replication across 3AZ
- NoSQL database - not a relational db
- Scales to massive workloads, distributed "serverless" database.
- Millions of requests per sec, trillions of row, 100s of TB of Storage.
- Fast & Consistent in performance.
- Single-digit millisecond latency - low latency retrieval.
- Integrated with IAM for security, authorization & administration.
- Low cost & auto scaling capabilities.
- Standard & Infrequent Access (IA) Table Class.

* • DynamoDB is a key/value database

DynamoDB ~~Accelerator~~ Accelerator - DAX

- Fully managed in-memory cache for DynamoDB.
- 10x performance improvement - Single digit millisecond latency to microseconds latency - when accessing your DynamoDB tables.
- Secure, highly scalable & highly available
- Difference with Elastic Cache at the CCP level:
  DAX is only used for & is integrated with DynamoDB, while ElasticCache can be used for other databases.

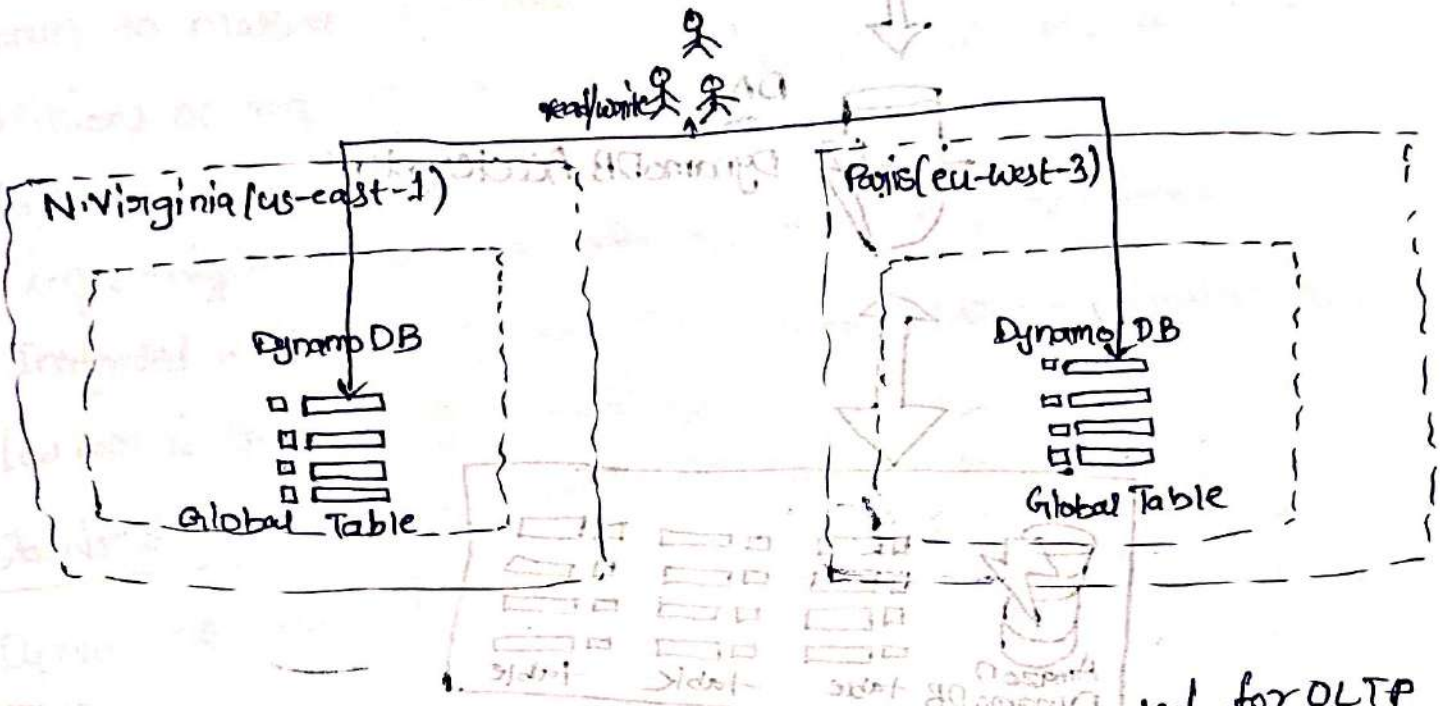


Note:
→ Here, we are creating table without database.
→ Actually, db already exists, it's serverless, we don't need to Provision servers

Diff b/w DynamoDB & RDS is that DynamoDB will have all the data living within one single table, & there's no way to join it with another table. So it's not a relational database.

DynamoDB - Global Tables

- Make a DynamoDB table accessible with low latency in multiple regions.
- Active-Active replication (read/write to any AWS region)



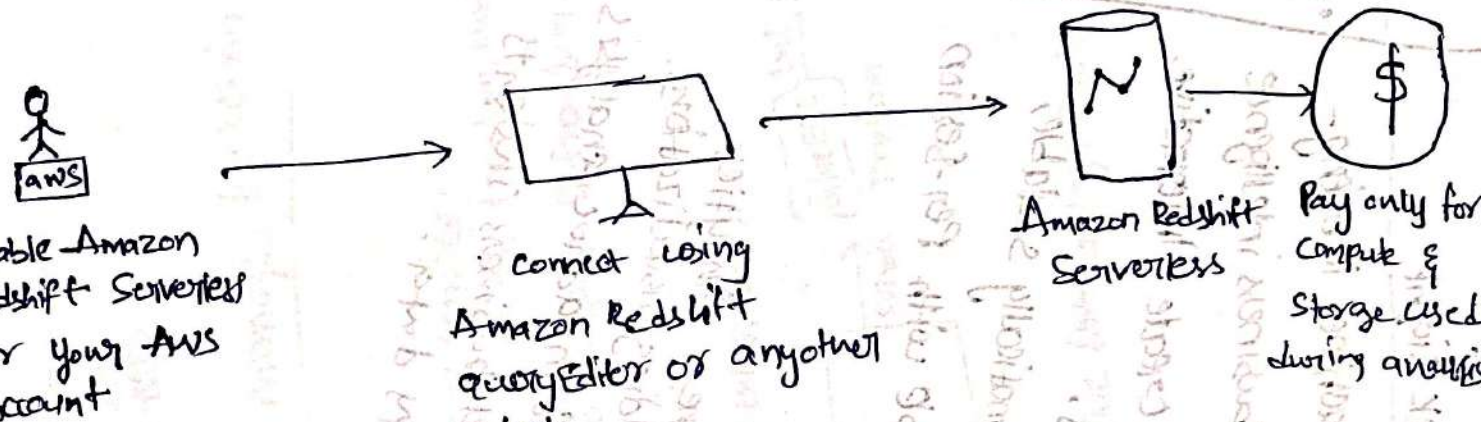
Red Shift

- Redshift is based on PostgresSQL, but it's not used for OLTP
  OLTP - Online Transaction processing.
- It's OLAP - Online Analytical processing (analytics & data warehousing)
- Load data once every hour, not every second.
- 10x better performance than other data warehouses, scale to PBs of data.
- Columnar storage of data (instead of row based)
- Massively parallel query execution (MPP), highly available.
- pay as you go based on the instances provisioned
- Has a SQL interface for performing the queries.

- BI tools such as AWS quicksight or Tableau integrate with it.

Redshift serverless

- Automatically provisions & scales data warehouse underlying capacity.
- Run analytics workloads without managing data warehouse infrastructure.
- Pay only for what you use.
- use cases: Reporting, dashboarding applications, real-time analytics.



Amazon EMR

- EMR stands for "ElasticMapReduce"
- EMR helps creating Hadoop Clusters (Big Data) to analyze & process vast amount of data.
- The Clusters can be made of hundreds of EC2 Instances.
- Also supports: Apache Spark, HBase, Presto, Flink...
- EMR takes care of all the provisioning & configuration.
- Auto-Scaling & integrated with Spot instances.
- use-cases: data processing, machine learning, web indexing, bigdata...

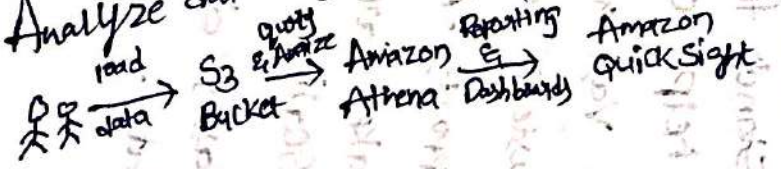
Amazon Athena

- Serverless query service to perform analytics against S3 objects.
- Uses standard SQL language to query the files.
- Supports CSV, JSON, ORC, Avro & Parquet (built on presto)
- Pricing: $5.00 per TB of data scanned
- Use compressed or columnar data for cost-savings (less scan)

Use cases: Business intelligence/analytics/ reporting, analyze & query VPC flow logs, ELB logs, Cloud Trail trails, etc...

Exam Tip:

Analyze data in S3 using serverless SQL, use Athena



Amazon QuickSight

- Serverless machine learning-Powered business intelligence service to create interactive dashboards.
- Fast, automatically scalable, embeddable with per-session pricing.
- Use cases:
  - Business Analytics
  - Building visualizations.
  - Perform adhoc analysis.
  - Get business insights using data.

Document DB

- Aurora is an AWS-implementation of PostgreSQL/MySQL
- Documentation DB is the same for MongoDB (which is a NoSQL database)
- MongoDB is used to store, query & index JSON data.
- Similar deployment concepts as Aurora.
- Fully Managed, highly available with replication across 3 AZ
- Document DB storage automatically grows in increments of 10GB.
- Automatically scales to workloads with millions of requests per seconds.

Amazon Neptune

- A fully managed graph database
- A popular graph dataset would be a Social Network.
  - Users have friends.
  - Posts have Comments
- Highly available across 3 AZ, with upto 15 read replicas.
- Build & run applications working with highly connected datasets - optimized for these complex & hard queries.
- Can Store up to billions of relations & query the graph with milliseconds latency.
- Highly available with replications across the AZs
- Great for Knowledge Graphs (wiki.), fraud detection, recommendation engines, Social networking.

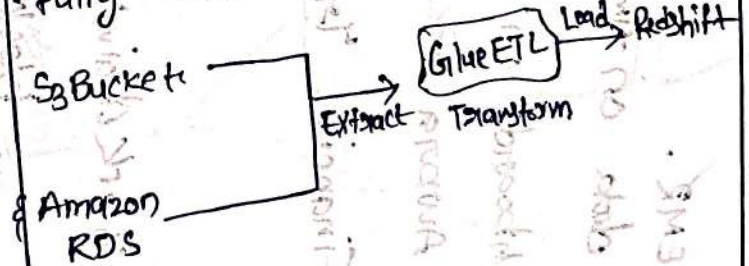
Amazon Timestream

~~Manage RDS~~

Amazon Managed Blockchain

AWS GLUE

- Managed ETL service. Extract, transform, load.
- useful to prepare & transform data for analytics.
- Fully serverless service.



Amazon Glue Data catalog: Catalog of data

DMS - Database Migration Service

How do you migrate database from one db to another db.

- Quickly & Securely migrate databases to AWS, resilient, Self healing.
- The source database remains available during the migration

Supports:

- Homogeneous migrations: Ex oracle to oracle.
- Heterogeneous migrations: Ex:- SQL server to Aurora.

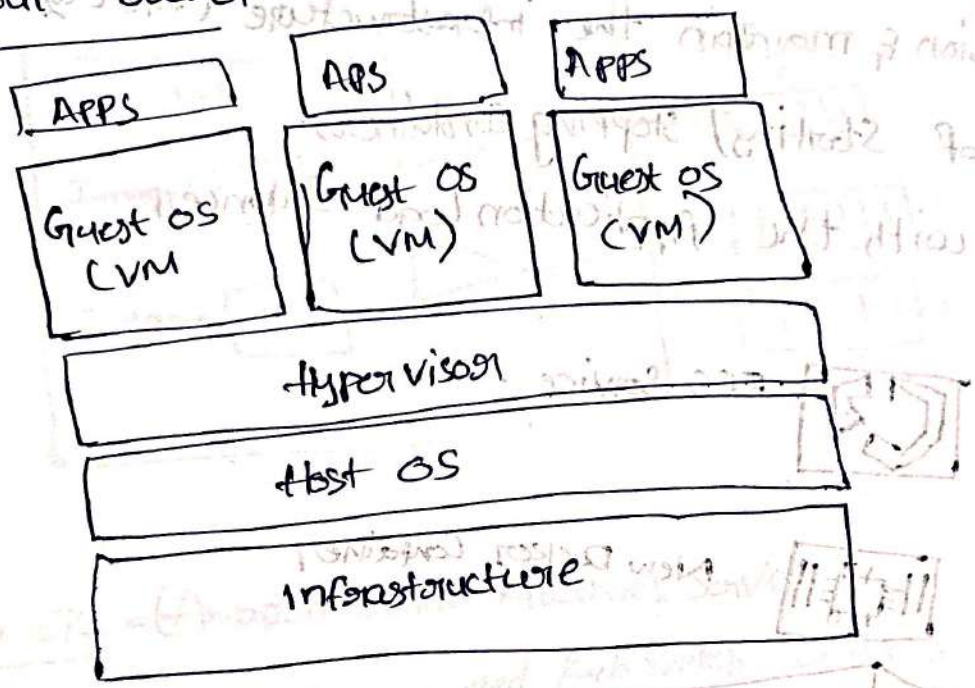
Database & Analytics Summary in Aws

- Relational Database - OLTP : RDS & Aurora (SQL)
- Diff b/w. multi-AZ, Read Replicas, Multi-Region.
- In-Memory Database : Elastic cache
- Key/value Database : DynamoDB (serverless) & DAX (cache for DynamoDB)
- Warehouse - OLAP : Redshift (SQL)
- Hadoop cluster : EMR.
- Athena : query data on Amazon S3 (serverless & SQL)
- Quicksights : dashboards on your data (serverless)
- Document DB : "Aurora for MongoDB" (JSON - NoSQL database)
- Amazon QLDB : Financial Transactions Ledger (immutable journal, cryptographically verifiable)
- Amazon Managed Blockchain : managed Hyperledger fabric & Ethereum blockchains.
- Glue : Managed ETL & Data Catlog Service.
- Database Migration : DMS
- Neptune : graph database
- Timestream : time-series database.

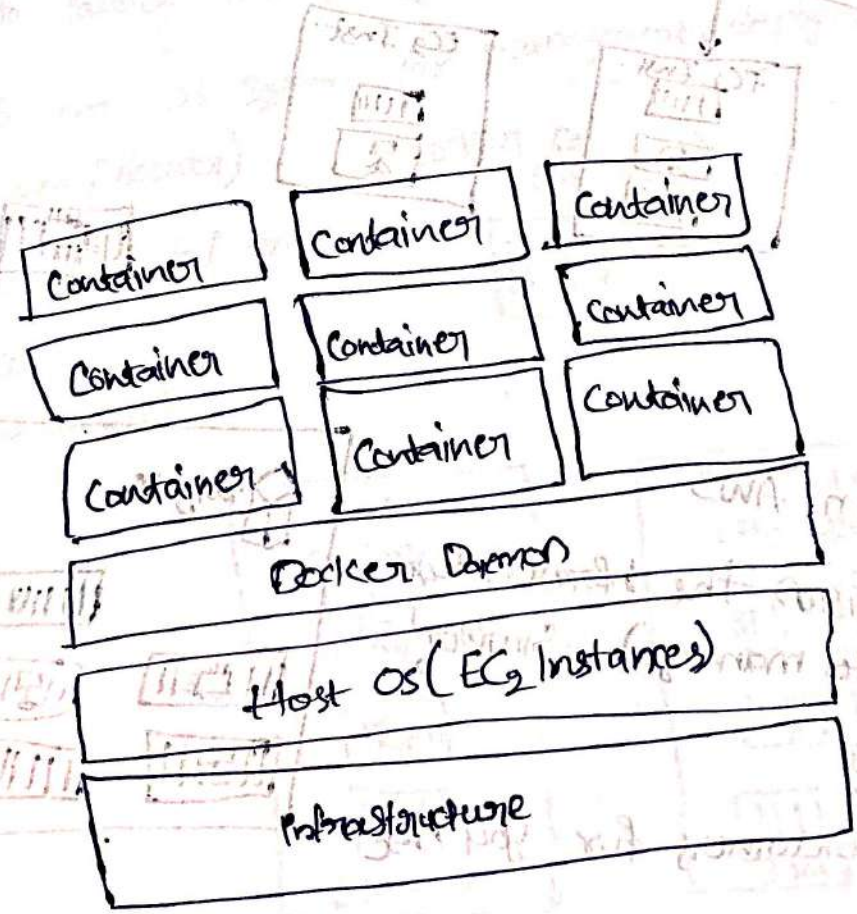
ECS Section

→ If we managed to Package our application in a Docker Container, then it will become very easy for us to run it on a EC_2 Instance.

Without Docker



With Docker



~~EC~~

ECS - Elastic Container Service

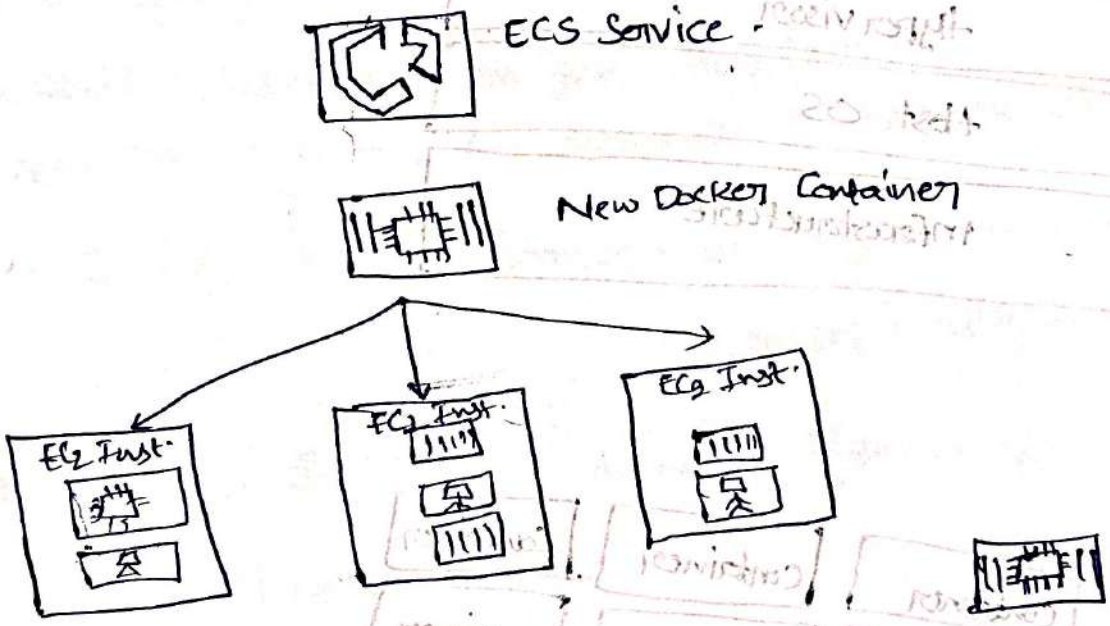
→ This is used to launch the Docker containers on AWS.

→ Launch Docker containers on AWS

→ You must provision & maintain the infrastructure (the EC2 instances)

→ AWS takes care of starting / stopping containers.

→ Has integration with the Application Load Balancer.



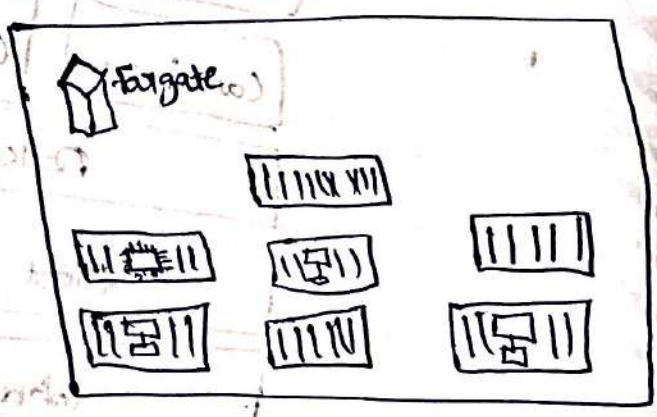
Fargate :-

→ Launch Containers on AWS

→ You do not provision the infrastructure (no EC2 instances to manage) - Simpler!

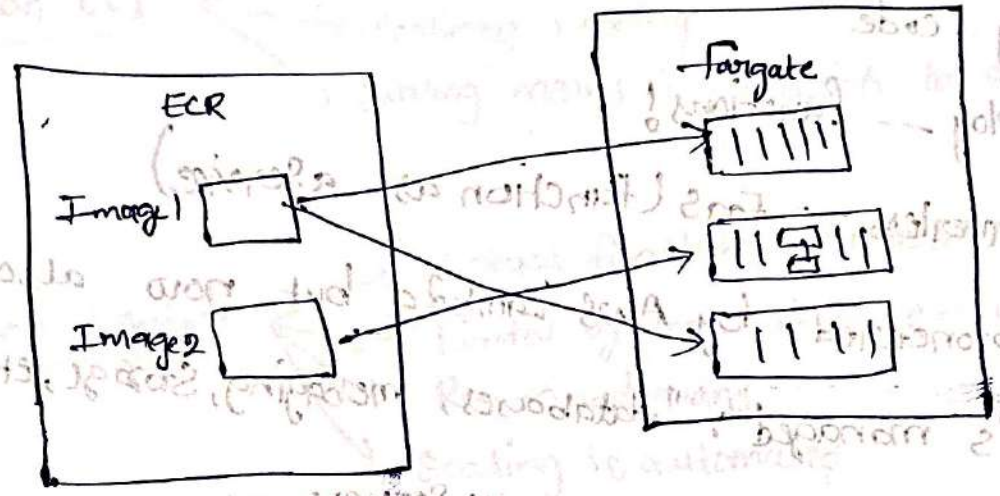
→ Serverless offering

→ AWS just runs containers for you based on the CPU / RAM you need



ECR - (Elastic Container Registry)

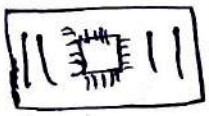
- Private Docker Registry on AWS
- This is where you store your Docker images so they can be run by ECS or Fargate.



Amazon EKS = (Amazon Elastic Kubernetes Service)

- Allows you to launch managed Kubernetes clusters on AWS.
- Kubernetes is an OS system for management, deployment & scaling of containerized apps (Docker)
- Containers can be hosted on:
  - EC2 instances.
  - Fargate (Serverless).

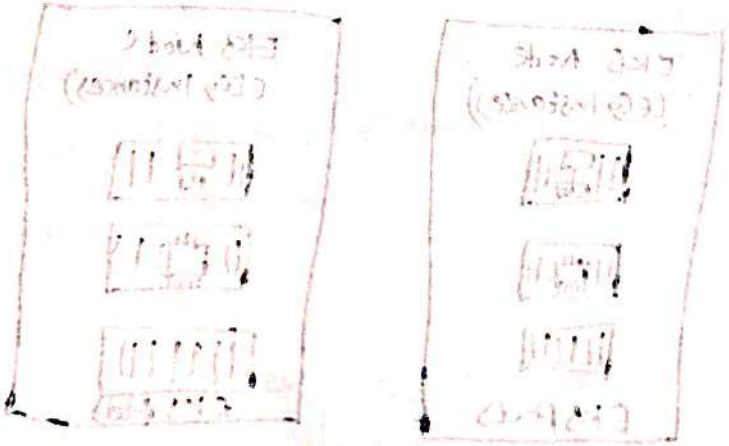
Docker Container (neo)



What's Serverless?

- Serverless is a new Paradigm in which the developers don't have to manage servers anymore
- They just deploy code
- They just deploy --- functions!
- Initially --- serverless = = Faas (function as a service)
- Serverless was pioneered by Aws Lambda but now also includes anything that's managed: "databases, messaging, storage, etc".
- Serverless doesn't mean there are no servers... it means you just

Amazon S3, DynamoDB, Fargate, Lamda are the examples for Serverless Servers.



AWS Lambda

Why!

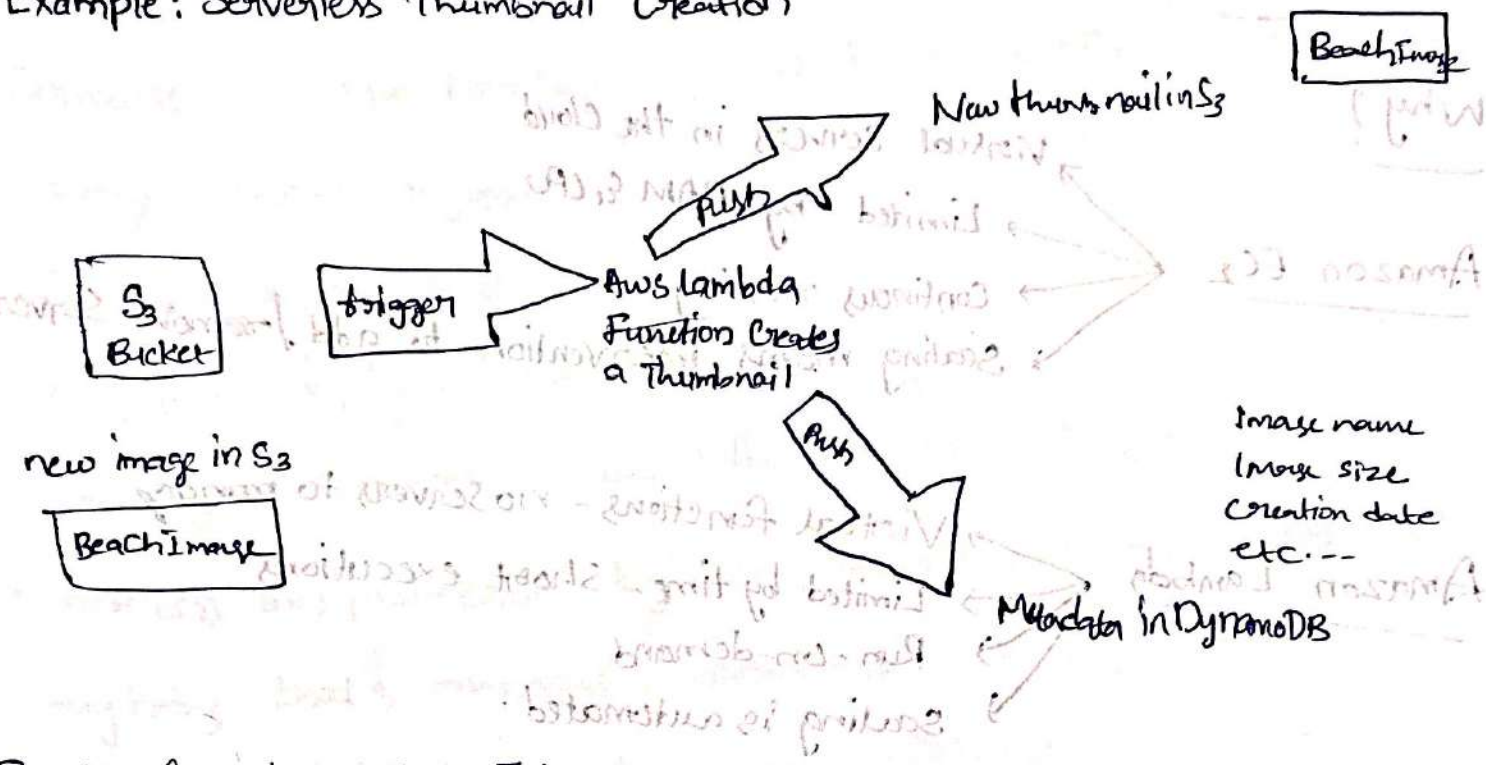
Amazon EC2
- Virtual servers in the cloud
- Limited by RAM & CPU
- Continous running
- Scaling means intervention to add/remove servers.

Amazon Lambda
- Virtual functions - no servers to manage
- Limited by time - short executions
- Run-on-demand
- Scaling is automated.

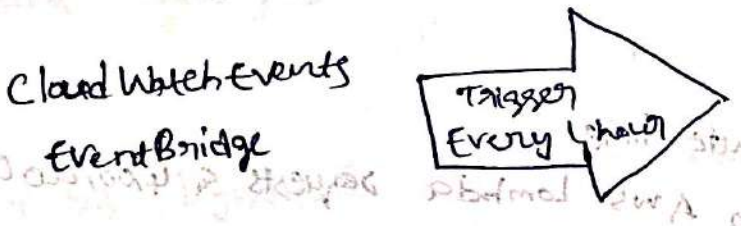
Benefits of AWS Lambda

- Easy pricing
  - Pay per request & compute time
  - Free tier of 1,000,000 AWS Lambda requests & 400,000 GBs of Compute time.
- Integrated with the whole AWS suite of services.
- Event-Drive: functions get invoked by AWS when needed (Reactive)
- Integrated with many Programming Languages
- Easy monitoring through AWS cloudwatch.
- Easy to get more resources per functions (up to 10GBs of RAM!)
- Increasing RAM will also improve CPU & network!.

Example: Serverless Thumbnail Creation



Example: Serverless CRON Job

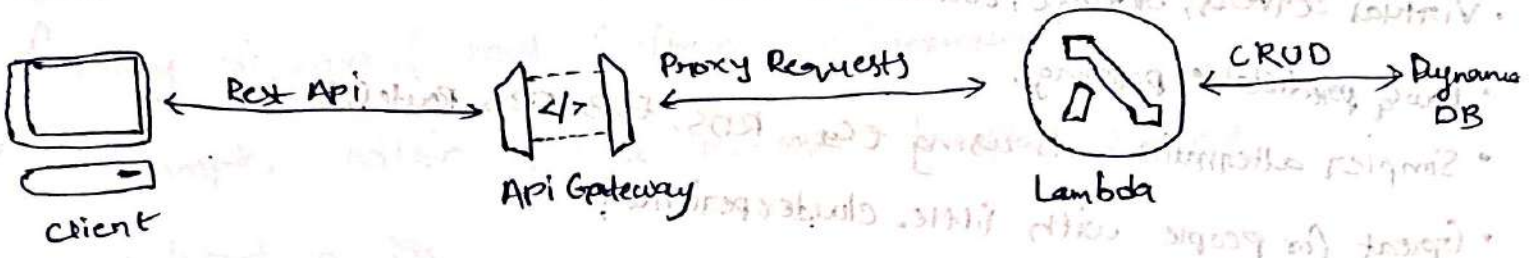


Aws Lambda Pricing:

- Pay per Calls:
  - First 1,000,000 requests are free
  - $0.20 per 1 million requests thereafter ($0.0000002 per request)
- Pay per duration: (in increment of 1 ms)
  - 400,000 GB-seconds of compute time per month if free.
  - == 400,000 sec if function is 1GB RAM.
  - == 3,200,000 sec if funct is 128 GB RAM.
  - After that $1.00 for 600,000 GB-seconds.
- It is usually very cheap to run Aws Lambda so it's very popular.

Amazon Api Gateway

- Eg: building a serverless API



- Fully managed service for developers to easily create, publish, maintain, & secure APIs.
- Serverless & scalable.
- Supports Restful APIs & webSocket APIs.
- Support for security, user authentication, API throttling, API keys, monitoring ....

AWS Batch

- Fully managed batch processing at any scale.
- Efficiently run 100,000s of computing batch jobs on AWS.
- A "batch" job is a job with a start & an end (opposed to continous)
- Batch will dynamically launch EC2 Instances or spot instances.
- AWS Batch provisions the right amount of compute / memory.
- You submit or schedule jobs & AWS Batch does the rest.
- Batch jobs are defined as Docker images & run on ECS.
- Helpful for cost optimizations & focusing less on the infrastructure.

Amazon Lightsail.

- Virtual servers, storage, database & networking.
- Low & predictable pricing.
- Simpler alternative to using EC2, RDS, ELB, EBS, Route 53 - - →
- Great for people with little cloud experience!
- Can setup notifications & monitoring of your lightsail resources.
- use cases:
  - simple web applications (has templates for LAMP, Nginx, MEAN, Node.js)
  - websites (templates for Wordpress, Magento, Plesk, Joomla)
  - Dev/Test Environment
- Has high availability but no auto-scaling, limited AWS integrations.

Deploying & Managing Infrastructure at Scale Section

- "Cloud formation" is a declarative way of outlining your AWS infra, for any resources (most of them are supported).
- for example, within a Cloud formation template, you say:
  - I want a SG
  - I want two EC2 Instances using this SG
  - I want an S3 Bucket
  - I want a load Balancer (ELB) in front of these machines.
- then Cloudformation creates those for you, in the right order, with the exact configuration that you specify.

Benefits of AWS Cloudformation (1/2)

- Infrastructure as Code
  - No resources are manually created, which is excellent for control
  - Changes to the infrastructure are reviewed through code.
- Cost
  - Each resources within the stack is tagged with an identifier
  - You can easily see how much a stack costs you.
  - you can estimate the cost of your resources using the Cloudformation template.
  - Savings strategy: In Dev, you could automation deletion of templates at 5PM & recreated at 8AM safely.

- **Productivity**
  - Ability to destroy & re-create an infrastructure on the cloud on the fly.
  - Automated generation of Diagram for your templates!
  - Declarative programming (no need to figure out ordering & orchestration).
- **Don't re-invent the wheel**
  - Leverage existing templates on the web!
  - Leverage the documentation.
- **Support (almost) all AWS resources!**
  - Everything we'll see in this course is supported.
  - you can use "custom resources" for resources that are not supported.

CLOUDFORMATION + INFRASTRUCTURE Composer

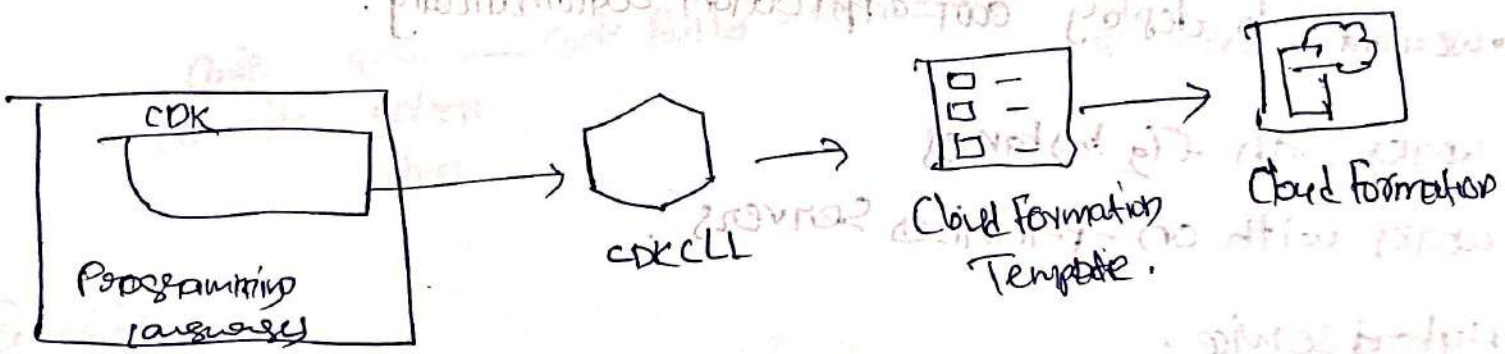
Eg: Wordpress Cloudformation Stack.

* Application Composer

Cloudformation → Infrastructure as code

AWS Cloud Development kit (CDK)

- Define your cloud infrastructure using a familiar language:
  - Javascript/Typescript, Python, Java, & .NET.
- The code is "compiled" into a Cloudformation template (Json/YAML)
- You can therefore deploy infrastructure & application runtime code together.
  - Great for Lambda functions.
  - Great for Docker containers in ECS/EKS.



BeanStalk Overview

→ Paas - platform as a service.

→ You just need to deploy the code, rest is done by BeanStalk.

Three architecture models:

- Single Instance deployment: good for dev
- LB+ASG: great for production or pre-production web applications.
- ASG only: great for non-web apps in production.

AWS CodeDeploy

- we want to deploy our application automatically.
- works with EC2 instances
- works with on-premises servers.
- Hybrid service.

AWS CodeCommit

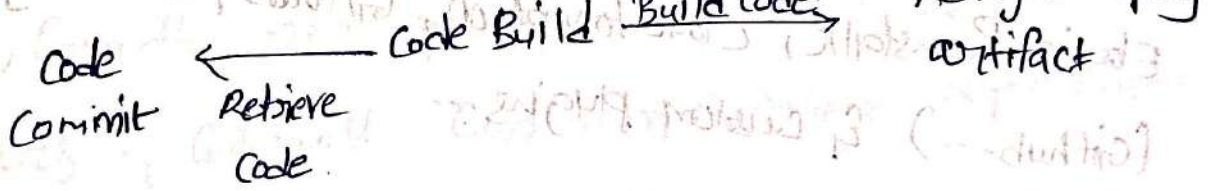
- Before pushing the application code to servers, it needs to be stored somewhere
- Developers usually store code in a repository, using the Git technology
- A famous public offering is Github, AWS competing product is code commit.
- Code Commit:
  - Source-control service that hosts Git-based repos.
  - make it easy to collaborate with others on code.

- The Code Changes are automatically versioned.

- Benefits:
  - Fully managed.
  - Scalable & highly available.
  - private, secured, integrated with AWS.

AWS Code Build

- Code building service in the cloud (name is obvious)
- Compiles source code, run tests, & produces Packages that are ready to be deployed (by code Deploy for example)



- Benefits
  - Fully managed, serverless
  - Continously scalable & highly available.
  - Secure.
  - Pay as-you-go pricing - only pay for the build time.

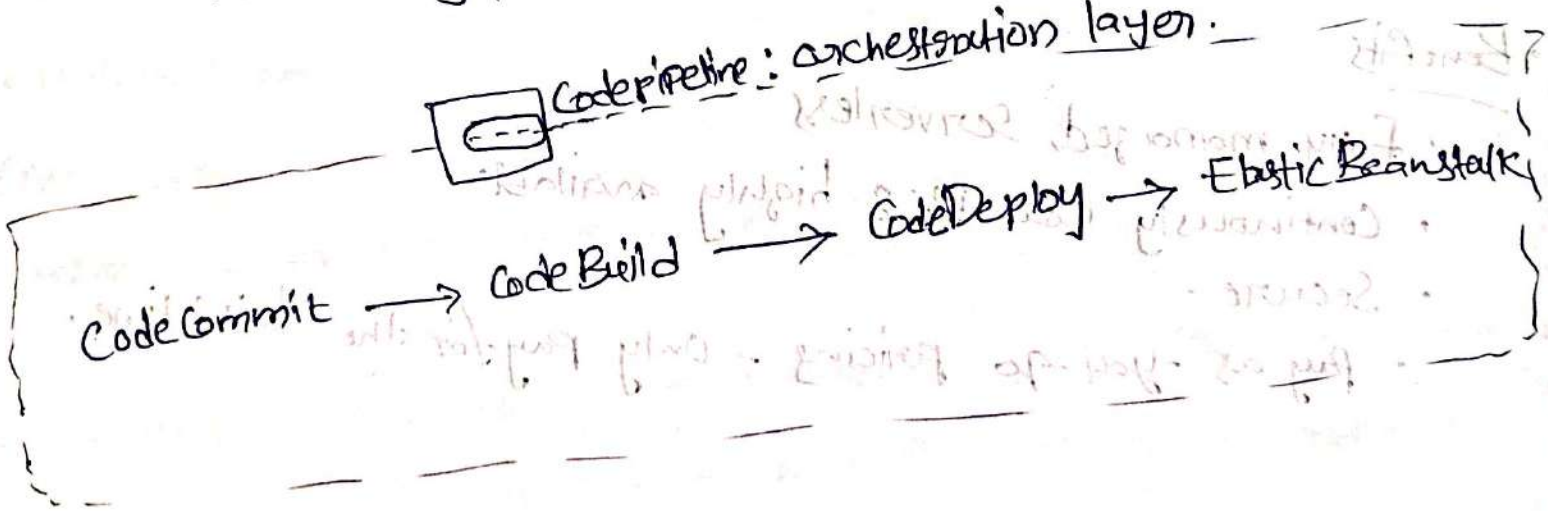
AWS CodePipeline

How do we know Code build & Code Commit are Connecte

→ we can connect them by Code Pipeline.

- Orchestrate the different steps to have the code automatically pushed to production.
  - Code ⇒ Build ⇒ Test ⇒ Provision ⇒ Deploy.
  - Basis for CICD (Continous Integration & Continous Delivery)

- Benefits:
  - Fully managed, compatible with Code commit, CodeBuild, CodeDeploy, Elastic Beanstalk, Cloudformation, Github, 3rd-party services (Github---) & custom plugins...
  - fast delivery & rapid updates.



Aws Code Artifact

- Software Packages depend one each other to be built (also called code dependencies), & new ones are created.
- Storing & retrieving these dependencies is called artifact management.
- Traditionally you need to setup your own artifact management system.
- Code artifact is a secure, scalable & cost-effective artifact management for software development.
- works with common dependency management tools such as maven, Gradle, npm, yarn, twine, pip & NuGet.
- Developers & CodeBuild can then retrieve dependencies stored from Code Artifact.

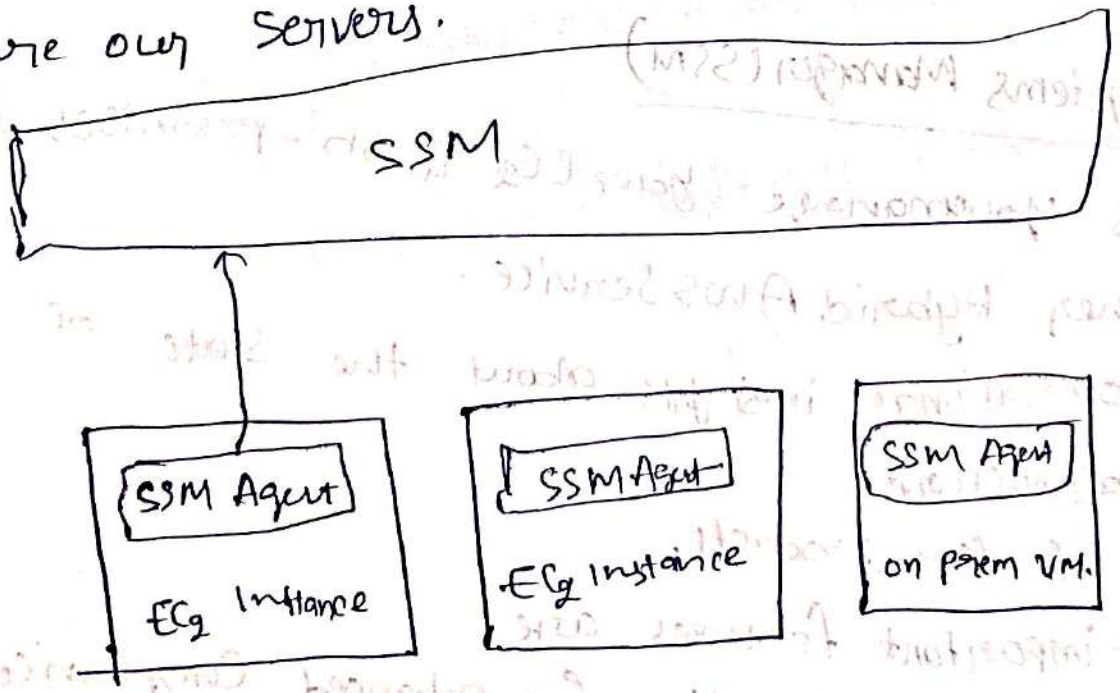
Aws Systems Manager (SSM)

- Helps you manage your EC2 & on-premises at scale.
- Another Hybrid Aws service.
- Get operational insights about the state of your infrastructure.
- Suite of 10+ products.
- Most important features are:
  - Patching automation for enhanced Compliance.
  - Run commands across an entire fleet of servers.

- Store parameter Configuration with the SSM Parameter Store.
- Works for Linux, windows, MacOS & Raspberry Pi OS (Raspbian)

How Systems Manager works

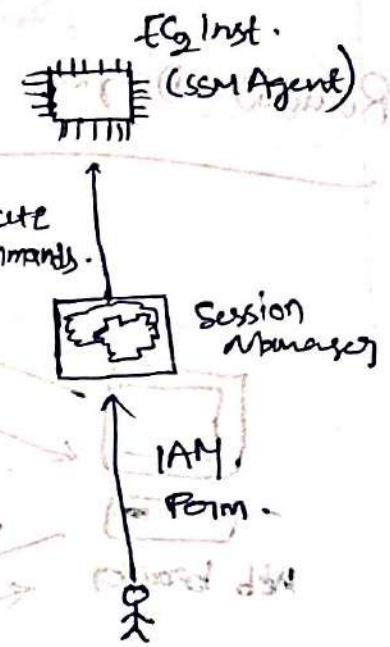
- We need to install the SSM agent onto the systems we Control.
- Installed by default on Amazon Linux AMI & Some Ubuntu AMI.
- If an Instance can't be controlled with SSM, it's probably an issue with the SSM agent.
- Thanks to the SSM agent, we can run commands, Patch & Configure our servers.



~~SSM~~

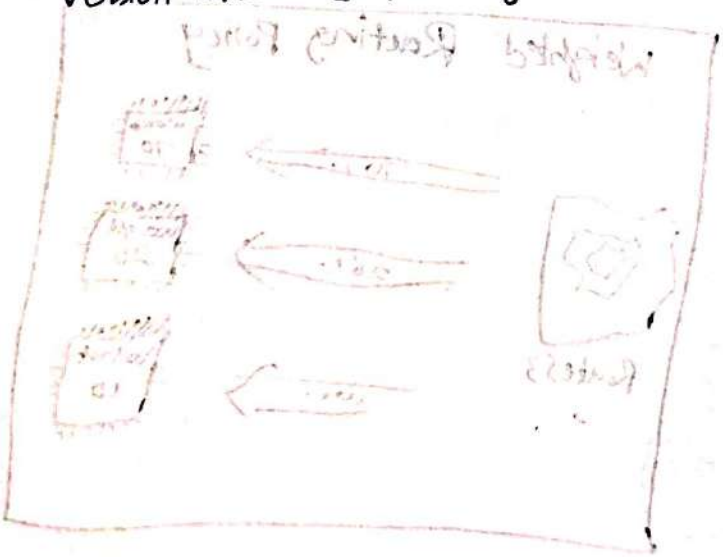
System Manager - SSM Session Manager.

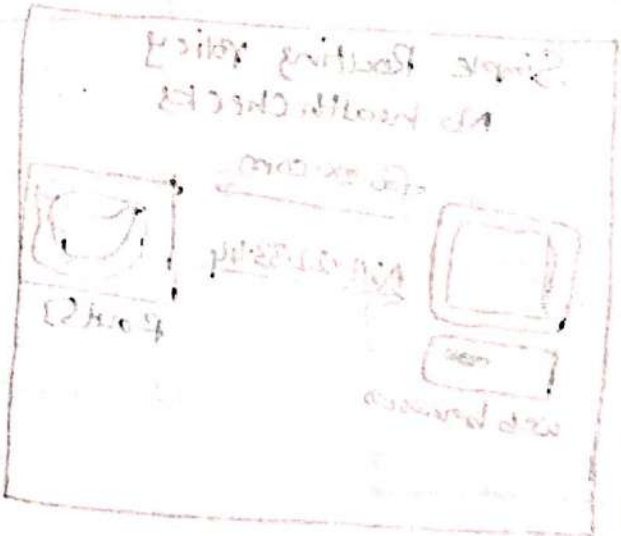
- Allows you to start a secure shell on your EC2 & on-premise servers.
- No SSH access, bastion hosts, or SSH keys needed.
- No port 22 needed (better security)
- Supports Linux, macOS & windows.
- Send session log data to S3 or cloudwatch logs.



Systems Manager Parameter Store:

- Secure Storage for Configuration & Secrets.
- Api Keys, Passwords, configurations...
- Serverless, scalable, durable, easy SDK
- Control Access Permissions using IAM
- Version tracking & encryption (optional)

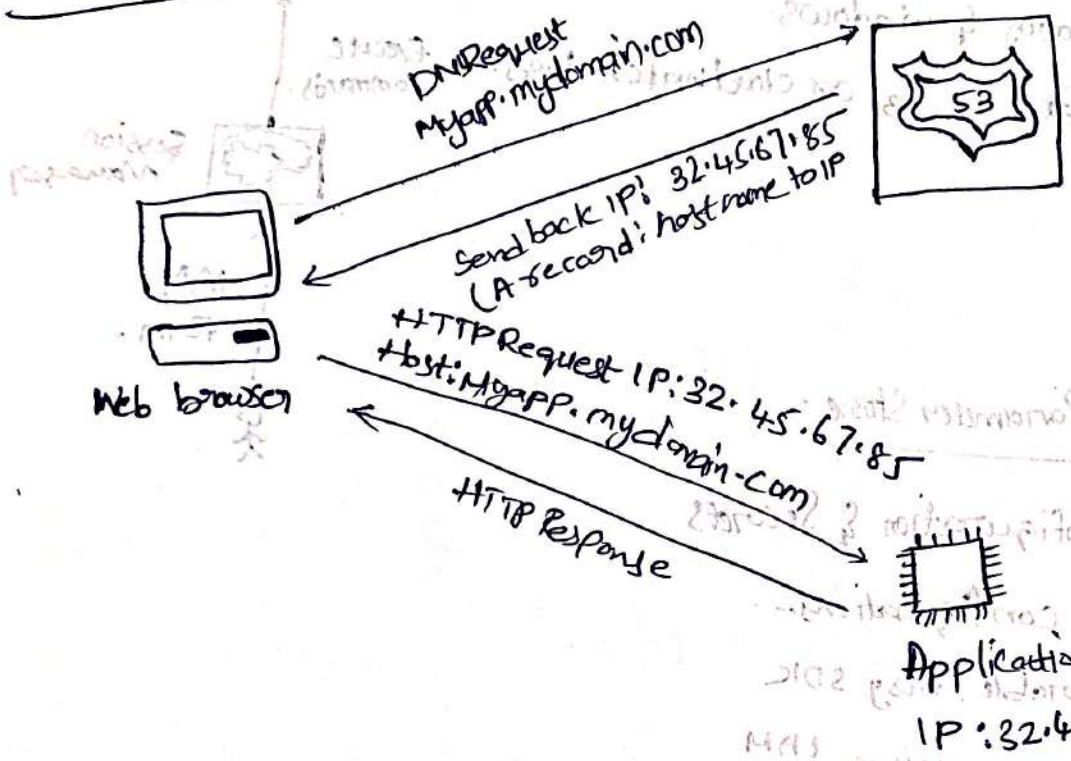




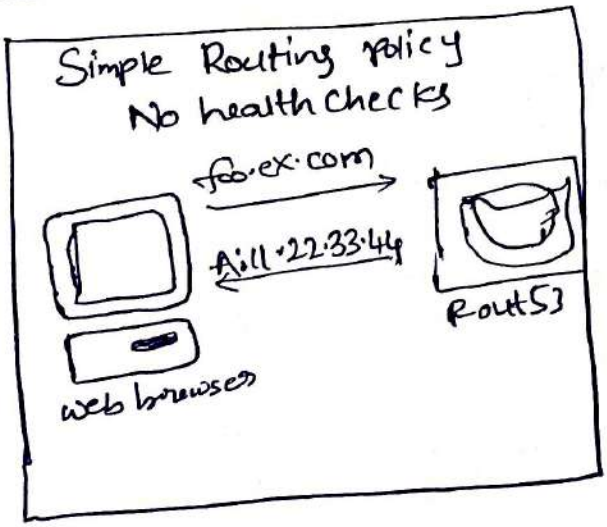
Route 53

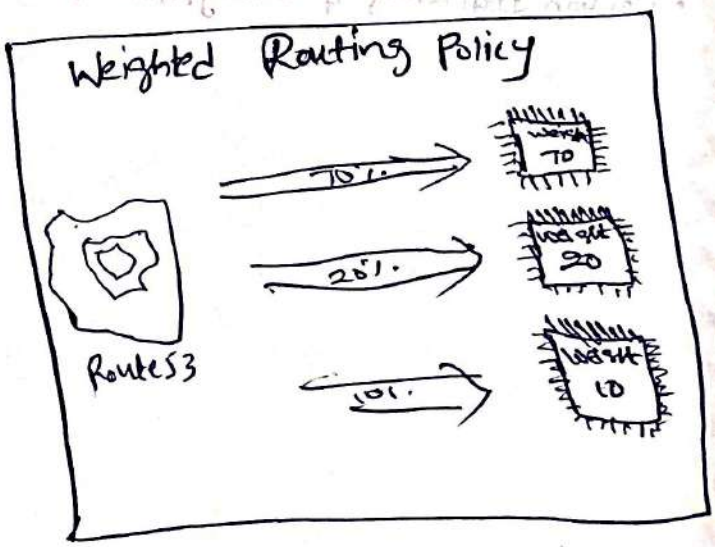
→ Route 53 is a Managed DNS (Domain Name System)

Route 53 - Diagram for A Record

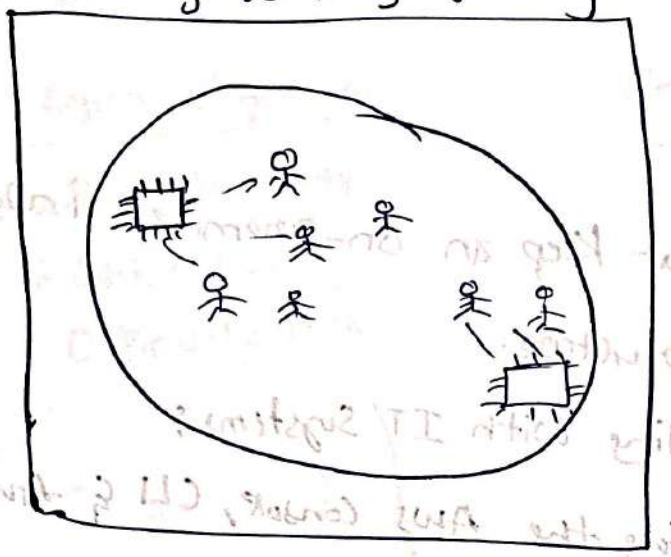


Rout 53 Routing Policies

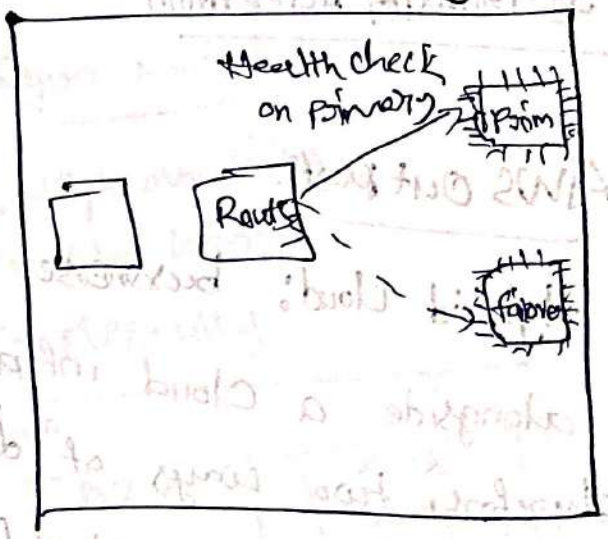




Latency routing Policy

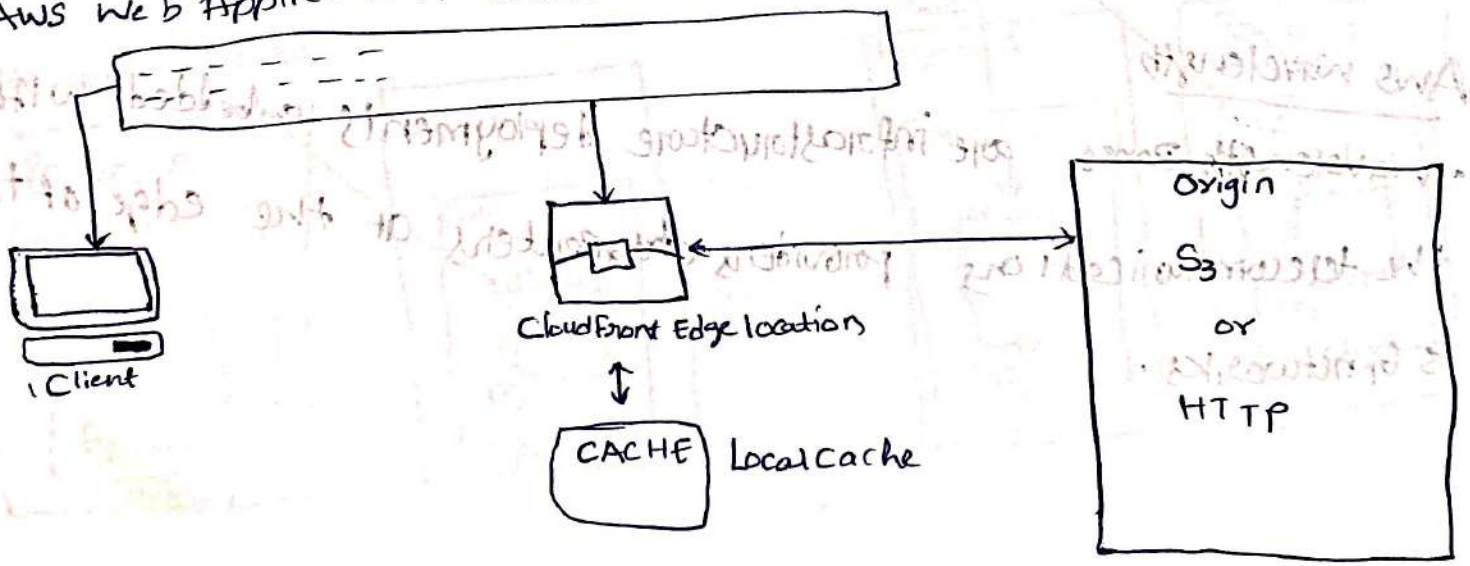


Failover Routing Policy



AWS Cloudfront

- Content Delivery Network (CDN)
- Improves read performance, content is cached at the edge.
- Improves users experience.
- 216 Point of presence globally (edge locations)
- DDoS protection (because worldwide), integration with Shield, AWS Web Application firewall



S3 Transfer Acceleration

AWS Outposts

- Hybrid Cloud: businesses that keep an on-premises infrastructure alongside a cloud infrastructure.
- Therefore, two ways of dealing with IT systems:
  - one for the AWS cloud (using the AWS Consoke, CLI & AWS APIS)
  - one for their on-premises infrastructure.
- AWS outposts are "server racks" that offers the same AWS infrastructure, services, APIS & tools to build your own applications on premises just as in the cloud.
- AWS will manage & setup the "outpost racks" within your on-premise infrastructure & you can start leveraging AWS services on-premises

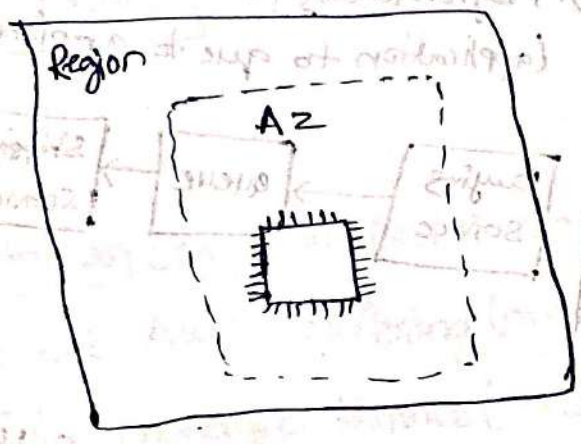
AWS wavelength

- Wavelength Zones are infrastructure deployments embedded within the telecommunications providers datacenters at the edge of the 5G networks.

Global Applications Architecture

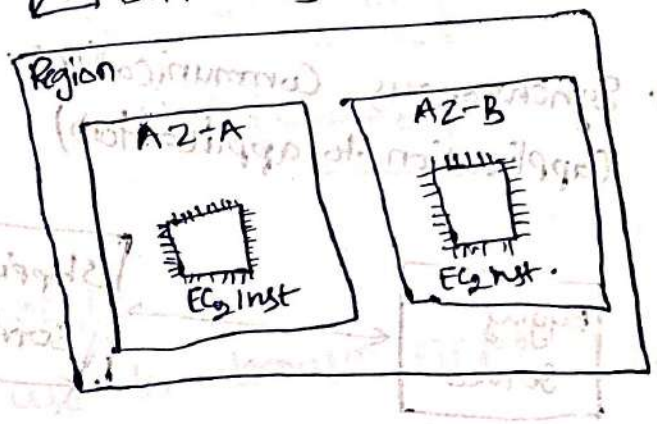
Single Region, Single AZ

- ✗ High Availability
- ✗ Global latency
- Difficulty



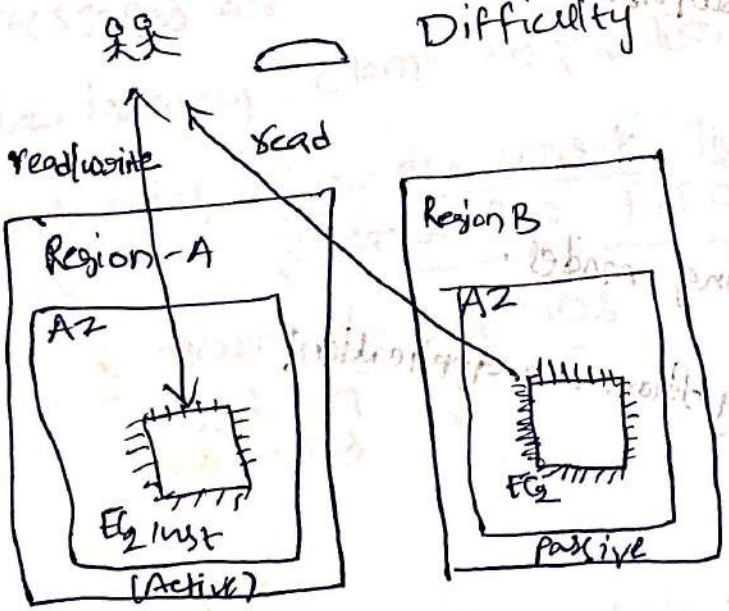
Single Region, Multi AZ

- ✓ High availability
- ✗ Global Latency
- Difficulty



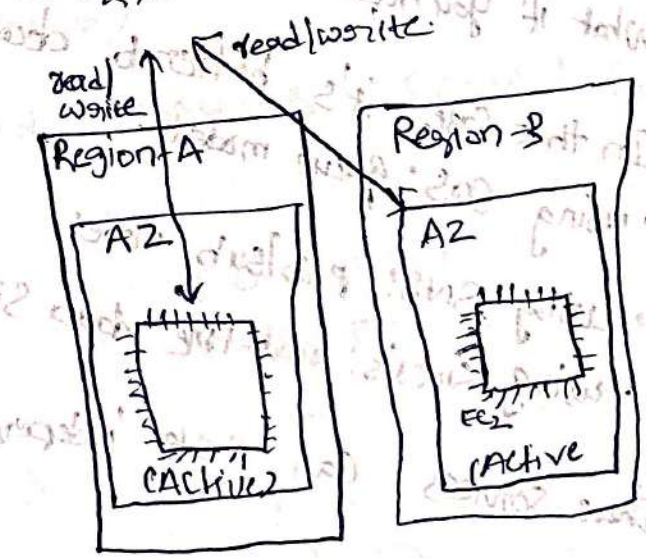
Multi Region, Active-Passive

- ✓ Global Read's Latency
- ✗ Global writes latency
- Difficulty



Multi Region, Active-Active

- ✓ Read's Latency
- ✓ writes Latency
- Difficulty

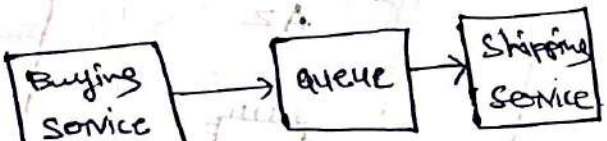


Cloud Integrations

- When we start deploying multiple applications, they will inevitably need to communicate with one another.
- There are two patterns of application communication.

1. Synchronous Communications (application to application)

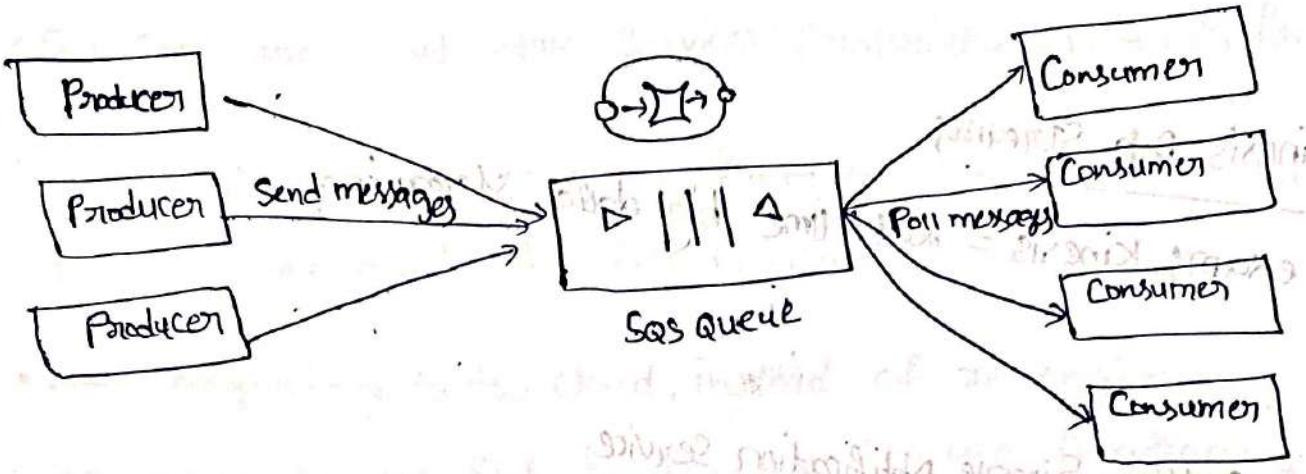
2. Asynchronous / Event based (application to que to application)



- Synchronous b/w applications can be problematic if there are sudden spikes of traffic.
- What if you need to suddenly encode 1000 videos but usually it's 10?
- In that case, it's better to decouple your applications:
  - using SQS: queue model.
  - using SNS: pub/sub model
  - using Kinesis: real-time data streaming model.

These services can scale independently from our application!

Amazon Sqs- Simple Queue Service

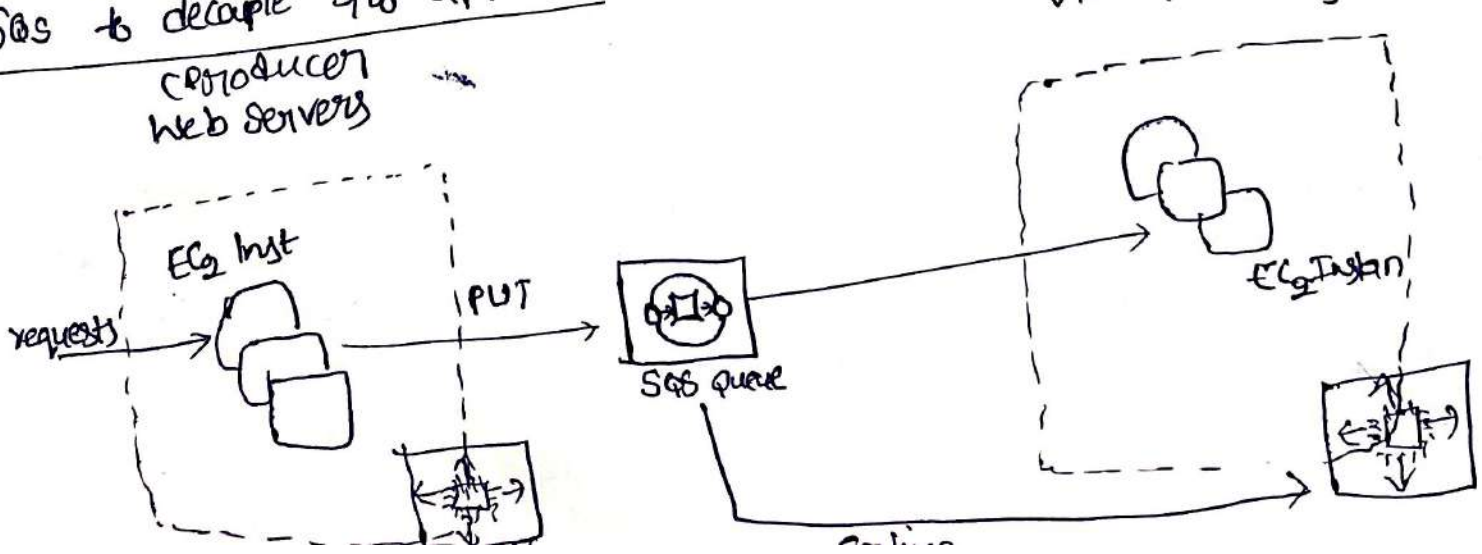


Amazon Sqs - Standard Queue.

- oldest Aws offering (over 16 years old)
- fully managed service (~ Serverless), use to decouple applications
- Scales from 1 message per second to 10,000s per second.
- Default retention of messages: 4 days, maximum of 14 days.
- No limit to how many messages can be in the queue.
- Messages are deleted after they're read by consumers.
- Low latency (<10ms on publish & receive)
- Consumers share the work to read messages & scale horizontally.

Amazon SQS - FIFO*

Sqs to decouple b/w app. tiers.



~~Amazon SQS~~

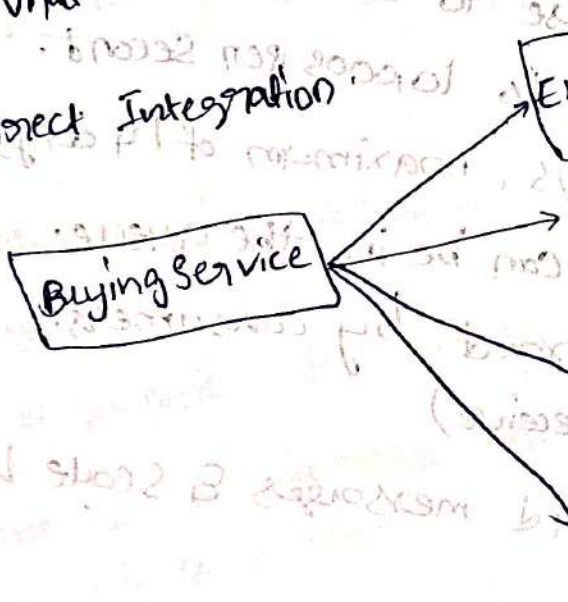
Amazon Kinesis Data Streams

- For the exam: Kinesis = "real-time" big data streaming

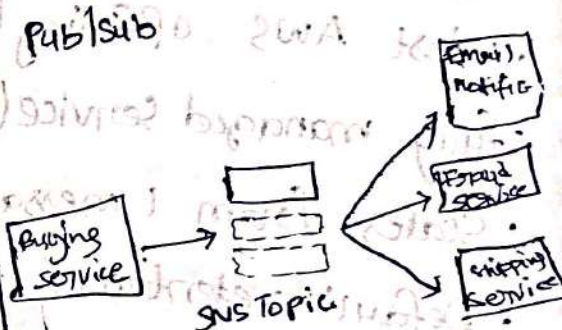
2. Amazon SNS - Simple Notification Service.

- What if you want send one message to many receivers

Direct Integration



Pub/Sub



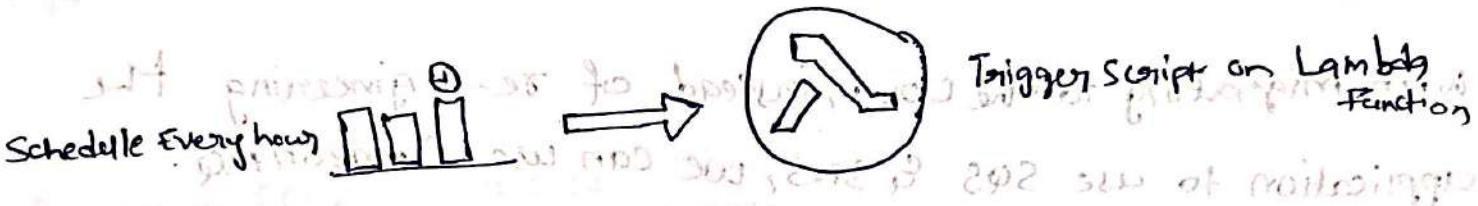
Amazon MQ

- SQS, SNS are "cloud-native" services: proprietary protocols from AWS.
- Traditional applications running from on-premises may use open protocols such as: MQTT, AMQP, STOMP, Openwire, WSS.
- When migrating to the cloud, instead of re-engineering the application to use SQS & SNS, we can use Amazon MQ.
- Amazon MQ is a managed message broker service for RabbitMQ, ActiveMQ.
- Amazon MQ doesn't "scale" as much as SQS/SNS.
- Amazon MQ runs on servers, can run in Multi AZ with failover.
- Amazon MQ has both queue feature (~ SQS) & topic features (~ SNS).

Cloud Monitoring

Amazon Event Bridge (formerly cloud watch Events)

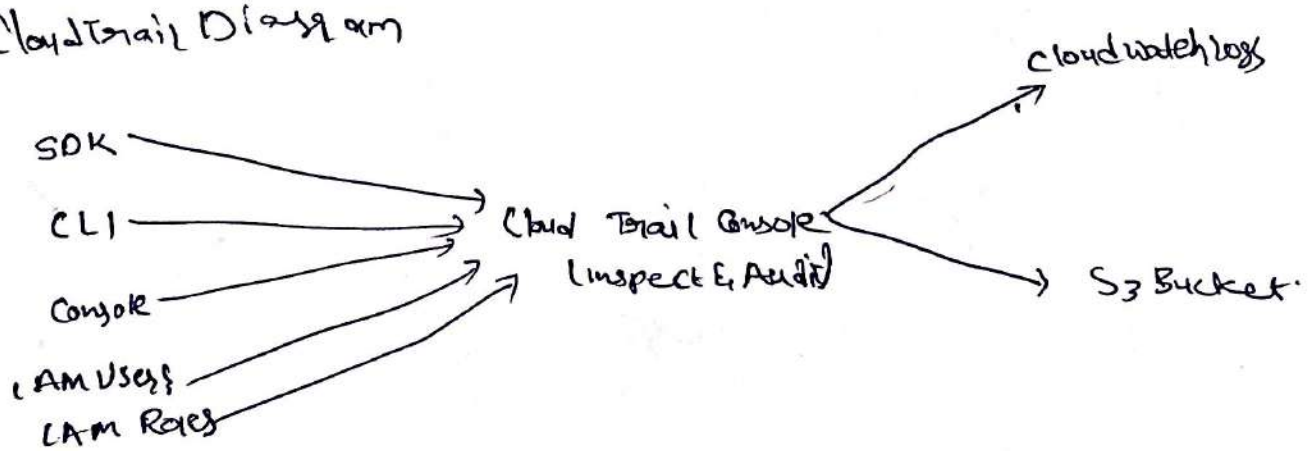
Schedule: Cron Jobs (scheduled scripts)



Cloud Trail

- Provides governance, compliance & audit for your Aws account.
- CloudTrail is enabled by default.
- Get history of events / API calls made within your Aws account by:
  - Console
  - SDK
  - CLI
  - AWS services.
- Can put logs from cloud Trail into cloudwatch logs or S3
- A trail can be applied to All regions(default) or a Single region
- If a resource is deleted in Aws, investigate CloudTrail first!.

CloudTrail Diagram



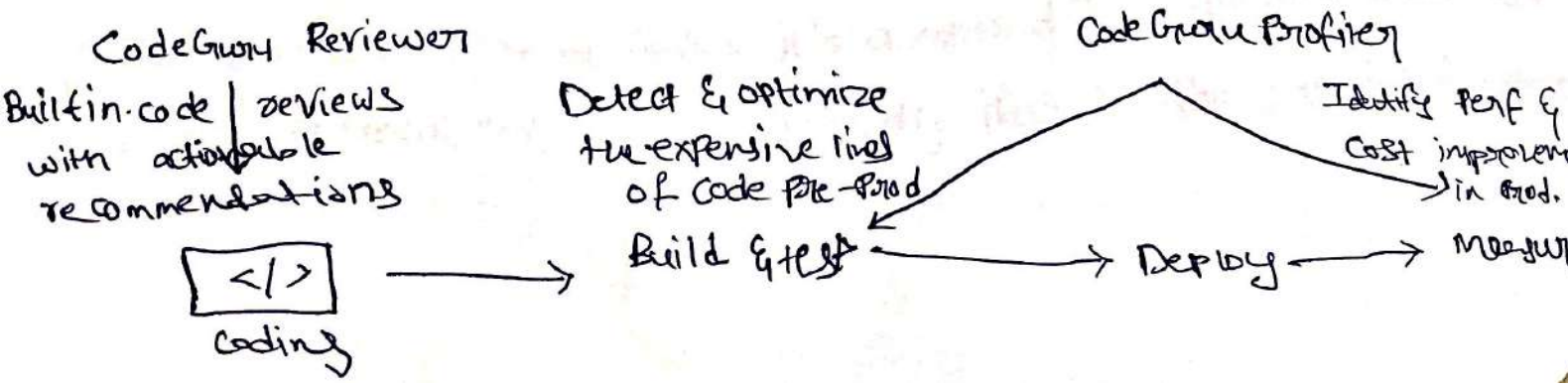
AWS x-RAY

- Debugging in production, the good old way:
  - Test locally
  - Add log statements everywhere.
  - Redeploy in statements.
- Log formats differ across applications & log analysis is hard.
- Debugging: one big monolith "easy", distributed services "hard"
- No common views of your entire architecture.
- Enter -- AWS x-Ray!

AWS X-RAY visual analysis of our applications.

Code Guru

- An ML-Powered Service for automated code reviews & application performance recommendations.
- Provides two functionalities.
  - CodeGuru Reviewer: automated code reviews for static code Analysis (development).
  - CodeGuru Profiler: visibility/recommendations about application perf. during runtime (Production)



Amazon CodeGuru profiler

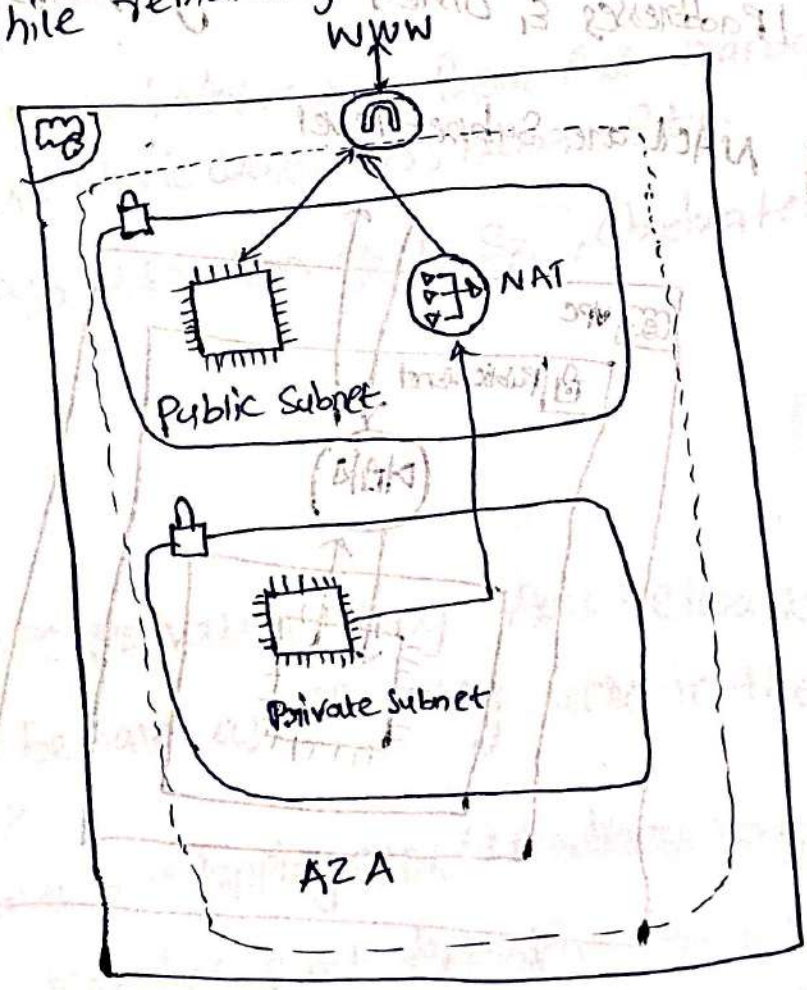
- Helps understand the runtime behaviour of your app.
- Eg: identify if your app. is consuming excessive CPU capacity on a logging routine.
- Features:
  - Identify & remove code inefficiencies.
  - improve application performance.
  - Decreas Compute costs.
  - Provide heap summary (identify which objects using up memory)
  - Anomaly Detection.
- Support applications running on Aws or on-premise.
- Minimal overhead on app.

Aws Health Dashboard - Service History & Your Account

VPC- Virtual Private Cloud

Internet Gateway (IGW)

- Internet Gateways helps our VPC instances connect with the internet.
- Public Subnets have a routte to the Internet gateway.
- NAT Gateways: (AWS-Managed) & NAT Instances (self managed) allow your instances in your private subnets to access the internet while remaining private.



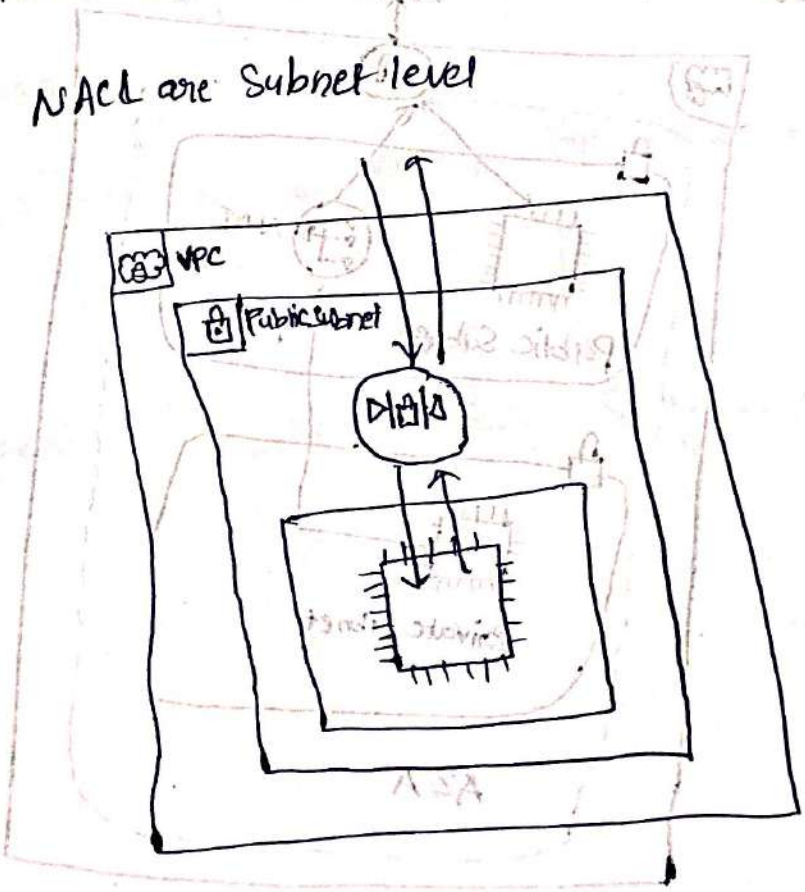
CIDR - Classless Inter-Domain Routing, it's a method for allocating IP addresses & routing network traffic more efficiently than the older class-based sys

Network ACL & Security Groups

- NACL (Network ACL)
  - A firewall which controls traffic from & to subnets.
  - Can have Allow & Deny rules.
  - Are attached at the subnet level.
  - Rules only include IP addresses.
  - Stateless

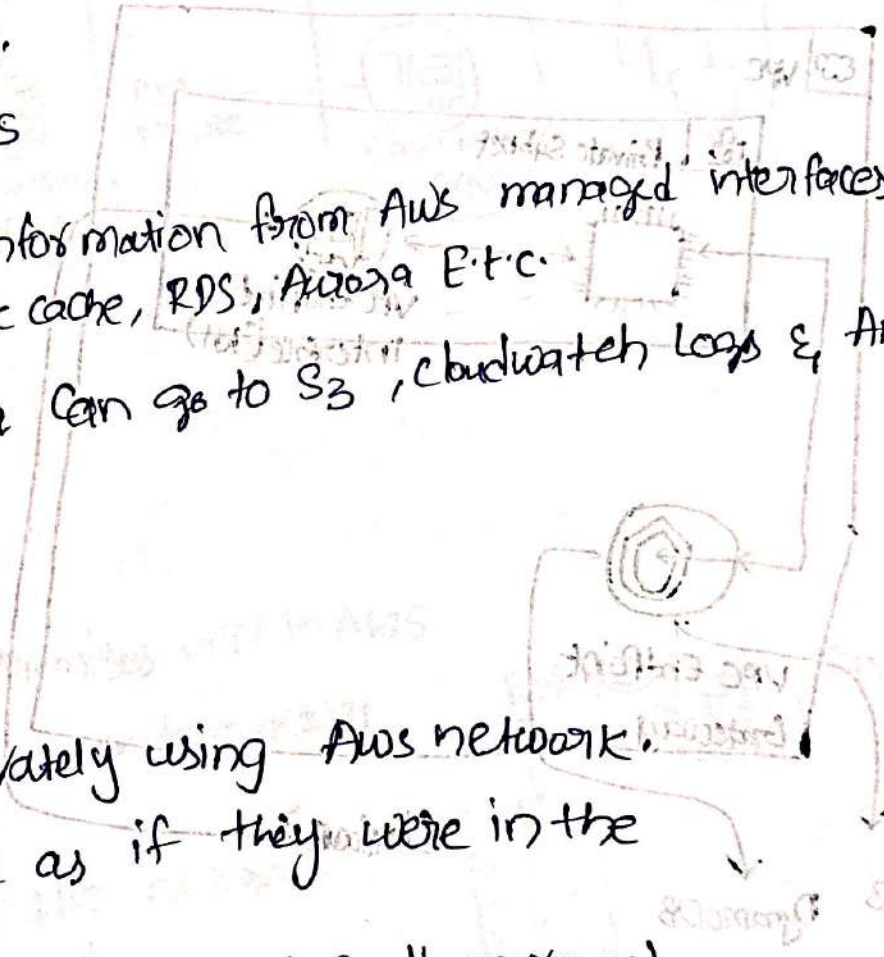
- Security Groups
  - A firewall that controls traffic to & from an EC2 Instance.
  - Can have only Allow rules.
  - Rules include IP addresses & other security groups.
  - Stateful

* SG are EC2 level & NACL are subnet level



VPC Flow Logs

- Capture Information about IP traffic going into your interfaces:
  - VPC Flow logs.
  - Subnet Flow logs.
  - Elastic Network Interface Flow Logs.
- Helps to monitor & troubleshoot Connectivity issues. Example:
  - Subnets to internet.
  - Subnets to Subnets.
  - Internet to Subnets
- Captures network information from AWS managed interfaces too: Elastic Load Balancers, Elastic cache, RDS, Aurora E.t.c.
- VPC Flow logs data can go to S3, Cloudwatch Logs & Amazon Data Firehose.



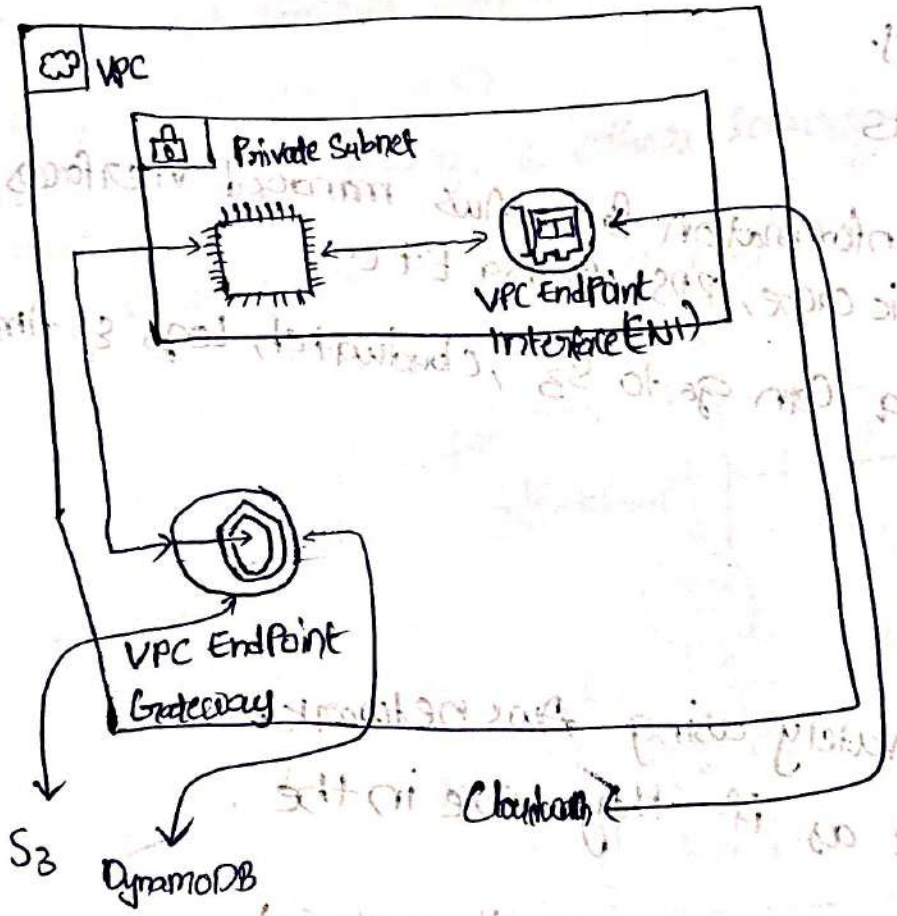
~~VPC Peerlog~~

VPC Peering

- Connect two VPC, privately using AWS network.
- Make them behave as if they were in the same network.
- Must not have overlapping CIDR (IP address range)
- VPC Peering connection is not transitive (must be established for each VPC that need to communicate with one another)

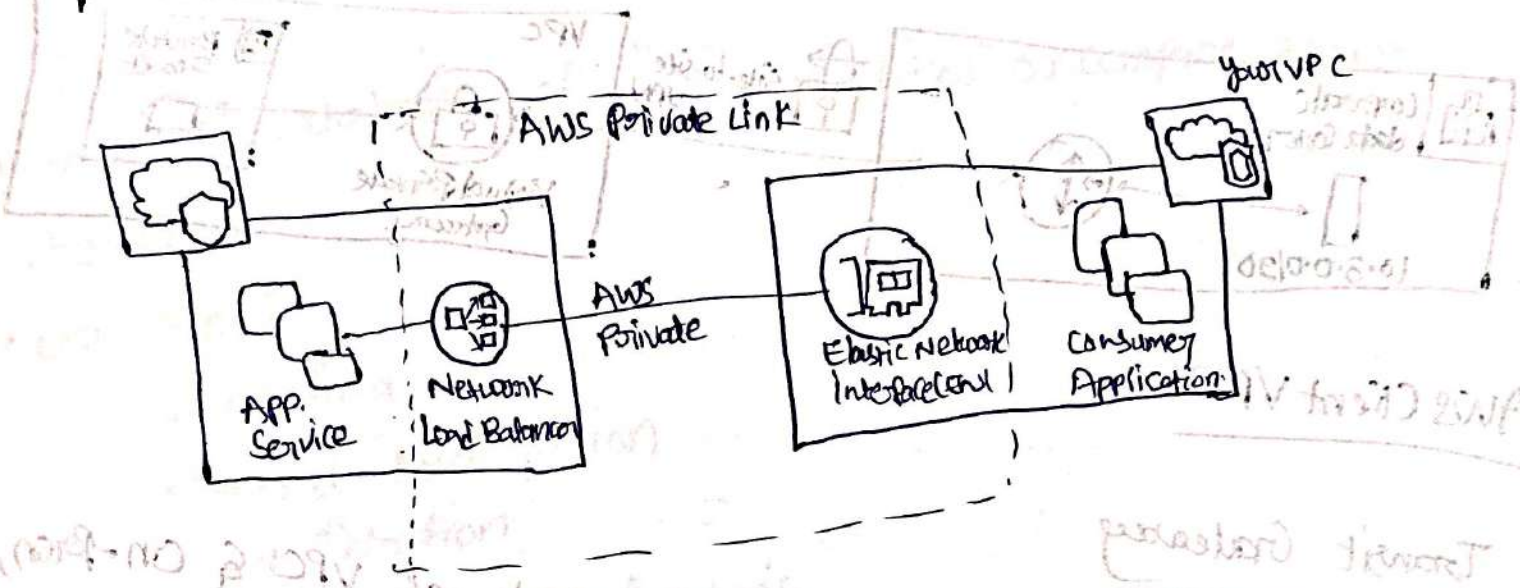
VPC Endpoints

- Endpoints allow you to Connect to AWS Services using a private network instead of the Public www network.
- This gives you enhanced security & lower latency to access AWS Services.
- VPC Endpoint Gateway: S3 & DynamoDB.
- VPC Endpoint Interface: most Services (including S3 & DynamoDB)



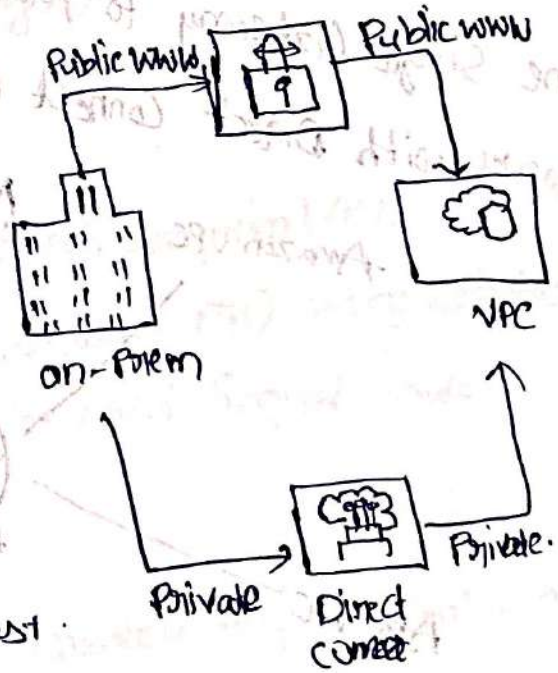
Aws Private link (VPC Endpoint Services)

- Most secure & scalable way to expose a service to 1000's of VPCs
- Does not require VPC Peering, internet gateway, NAT, route tables...
- Requires a network load balancer (service VPC) & ENI (customer VPC)



Site to site VPN

- Connect an on-premises VPN to AWS
- The connection is automatically encrypted
- Goes over the public internet

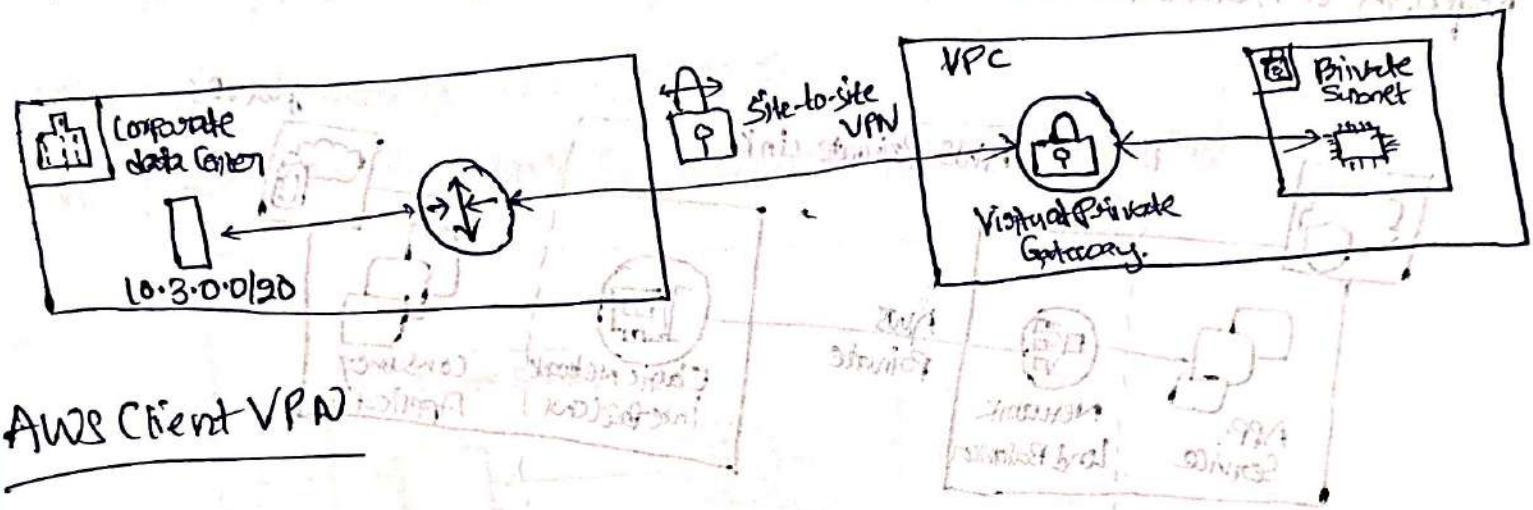


Direct Connect (DX)

- Establish a physical connection b/w on-prem & AWS.
- The connection is private, secure & fast.
- Goes over a private network.
- Takes at least a month to establish

Site to Site VPN

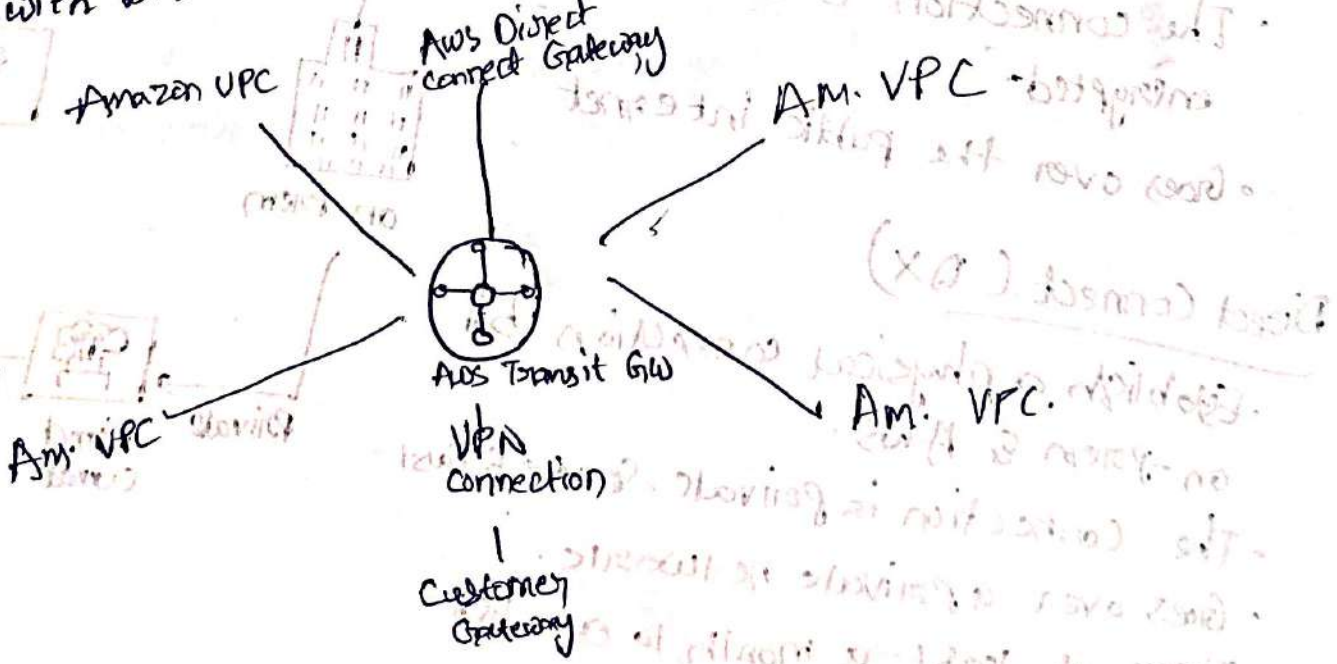
- on-prem: must use a Cust. Gateway (CGW)
- Aws: must use a Virtual Private Gateway (VGW)



AWS Client VPN

Transit Gateway

- For having transitive peering b/w thousands of VPC & on-prem, hub & spoke (star) connection.
- One single Gateway to provide this functionality.
- works with Direct Connect Gateway, VPN Connections.



Machine Learning

Amazon Rekognition

- Find objects, people, text, scenes in images & videos using ML.
- Facial analysis & facial search to do user verification, people counting.
- Create a database of "familiar faces" or compare against celebrities.
- Use Cases:
  - Labeling
  - Content Moderation
  - Text Detection
  - Face Detection & Analysis
  - Face ~~Recognition~~ Search & Verification.
  - Celebrity Recognition
  - Pathing (ex: for sports game analysis)

Amazon Transcribe

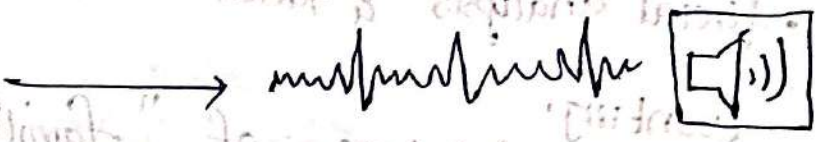
- Automatically convert speech to text.
- uses a deep learning process called automatic speech recognition (ASR) to convert speech to text quickly & accurately.
- Automatically remove personally Identifiable information (PII) using Redaction.
- Supports Automatic Language Identification for multi-lingual audio.
- Use Cases:
  - transcribe customer service calls.
  - automate closed captioning & subtitling.
  - generate metadata for media assets to create a fully searchable archive

"Hello my name is Chetna. I am enjoying the course!".

Amazon Polly

- Turn text into lifelike speech using Deep learning.
- Allowing you to create applications that talk.

Hi! my name is Stephane.
& this is a demo of Amazon Polly



Amazon Translate

- Natural & accurate language translation.
- Amazon Translate allows you to localize content - such as websites & applications - for international users, & to easily translate large volumes of text efficiently.

Source language

Hi, my name is Charan

Translate language

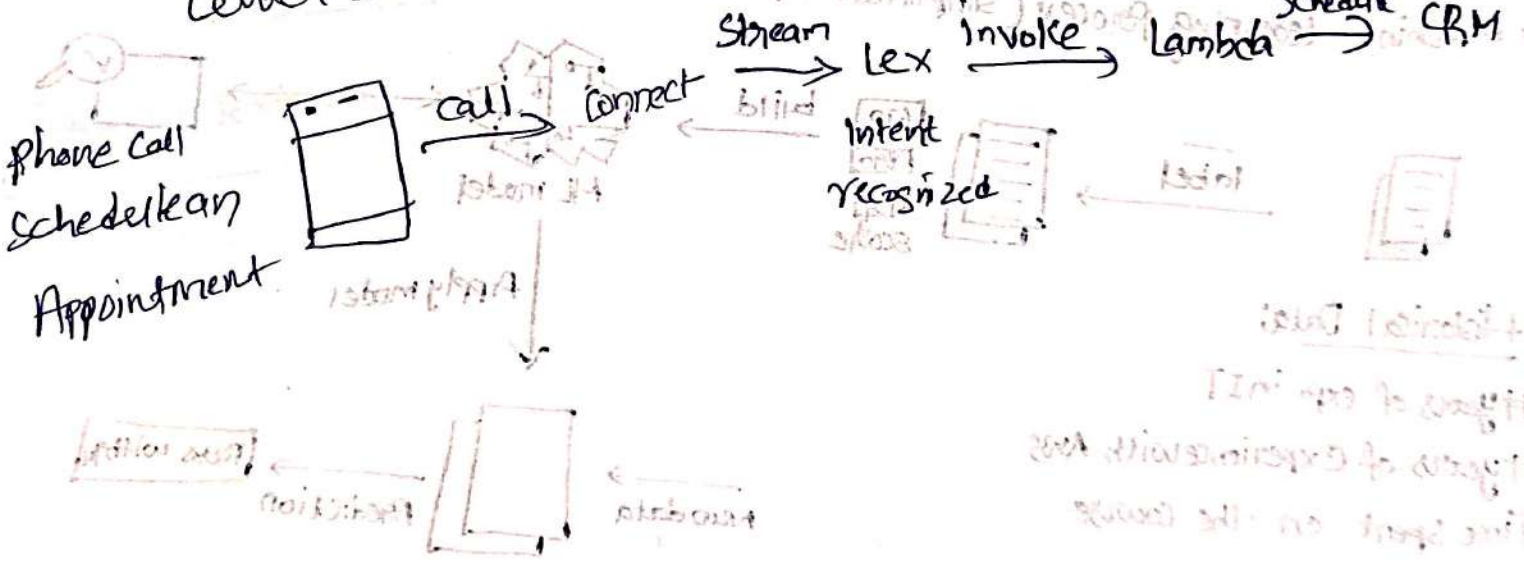
French
Bonjour, je m'appelle ~~Stéphane~~ Charan.

Portugese
Oi, meu, nome é Charan.

Hindi
हाय मेरा नाम चरण है

Amazon Lex & Connect

- **Amazon Lex:** (same technology that Powers Alexa)
  - Automatic Speech Recognition (ASR) to convert speech to text
  - Natural language understanding to recognize the intent of text, callers
  - Helps build chatbots, Call Center bots.
- **Amazon Connect**
  - Receive Calls, Create Contact flows, cloud-based virtual contact Center.
  - Can integrate with other CRM Systems or AWS.
  - No upfront Payments, 80% cheaper than traditional Contact Center Solutions.



Amazon Comprehend

- for NLP

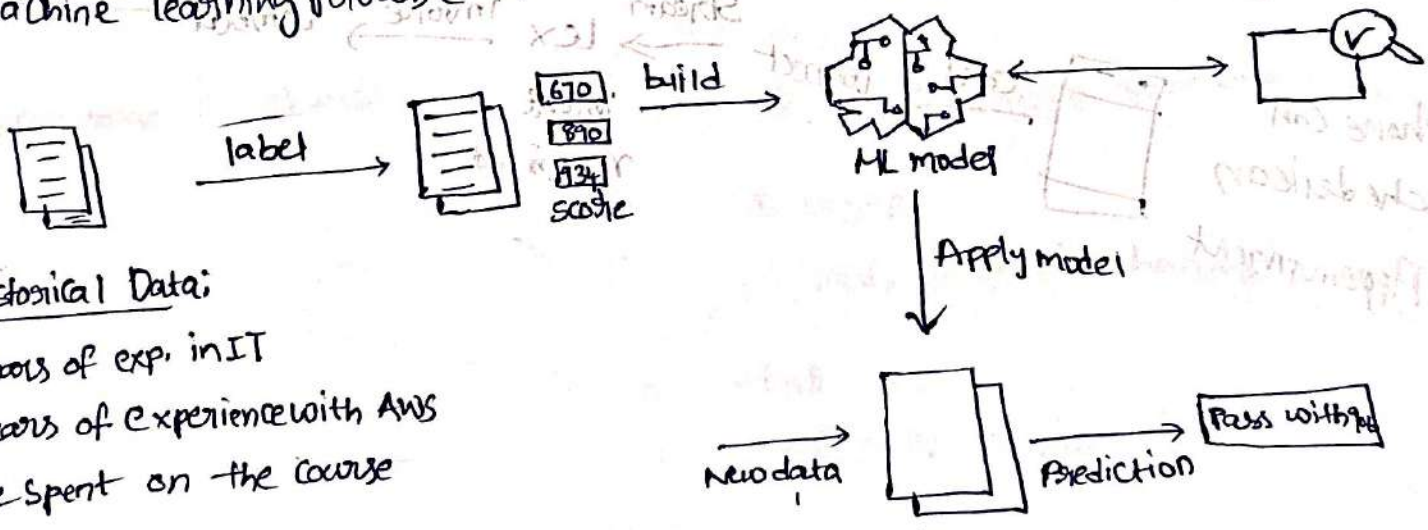
Use cases:

- Analyze customer interaction (emails) to find what leads to a positive or negative experience.
- Create & groups articles by topics that Comprehend will discover.

~~Sagemaker~~

Amazon Sagemaker

- Fully managed service for developers / data scientists to build ML models.
- Typically difficult to do all processes in one place + provision servers.
- Machine learning process (simplified): predicting your exam score.

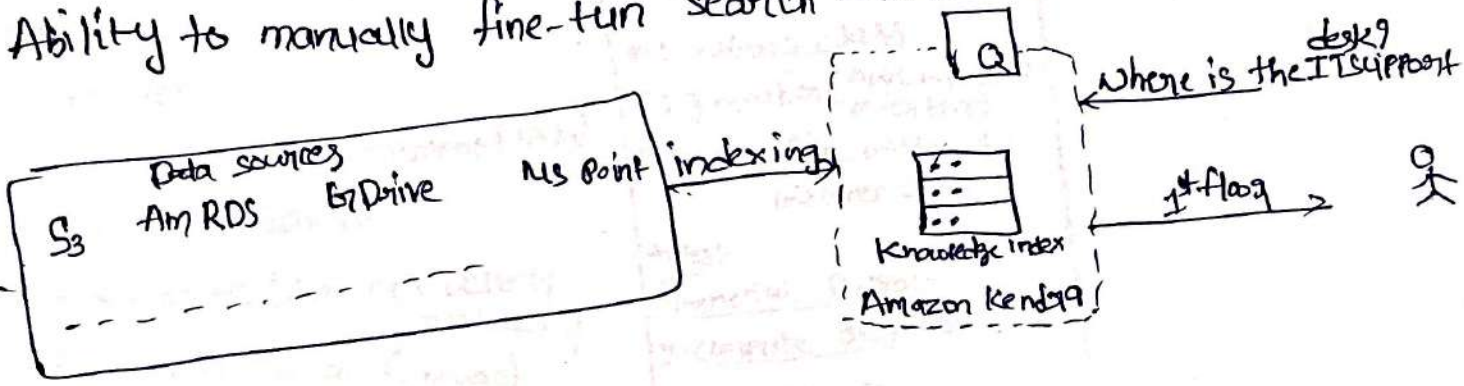


+ Historical Data:

#years of exp. in IT
#years of experience with AWS
Time spent on the course

Amazon Kendra.

- Full managed document search service powered by ML
- Extract answers from within a document (text, Pdf, Html, Powe --)
- Natural language search capabilities.
- Learn from user interactions/feedback to promote preferred results (incremental learning)
- Ability to manually fine-tun search results (importance of data, --)



Amazon personalize

Amazon Textract

AWS Cheat Sheet

AWS Cloud History

First launch - 2002
Global Launch - 2006

AWS Global Infrastructure

- AWS Regions
- AWS AZ's
- AWS Data Centers
- AWS Edge Locations / Points of Presence.

Global Services

- Identity & Access Management (IAM)
- Route 53 (DNS Service)
- CloudFront (CDN) (Content Delivery Network)
- WAF (web application firewall)

Most AWS Services are Regional Scoped

- Amazon EC2 (IaaS)
- Elastic Beanstalk (PaaS)
- Lambda (Function as a service)
- Rekognition (Software as a Service)

Identity & Access Management (IAM), Global

- In AWS, you apply the "least privilege principle": don't give permissions than a user needs.

MFA devices options in AWS

Virtual MFA device (univ. 2nd fact. security key)

Google Auth. / Authy

yubikey

Hardware Key Fob MFA device

AWS GovCloud (US)

To access AWS, you have three options:

- AWS Management Console (protected by PASS + MFA)
- AWS CLI: Protected by Access keys
- AWS SDK: Protected by Access keys

EC2 - IaaS

EC2 Instance Types

m5.2xlarge

m: instance class
5: generation (AWS improves over time)
2xlarge: Size within the instance class.

types:

1. General Purpose
2. Compute optimized
3. Memory
4. Accelerated computing
5. storage optimized
6. HPC optimized
7. Instance features
8. Measuring instance Performance.

* Compute
* Memory
* Networking

Security Groups

- They control how traffic is allowed into or out of our EC2 instances.
- Contains only allow rules.
- All inbound traffic is blocked by default
- All outbound traffic is authorized by default

Name of the Services

1. IAM
2. EC2
3. Security Groups.

Classic ports to know

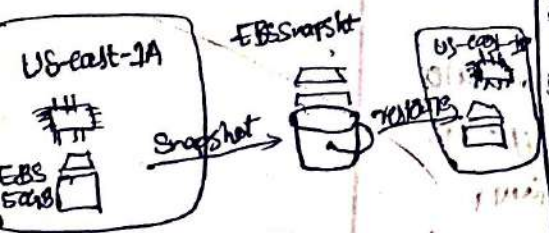
- 22 = SSH - log into a Linux instance
- 21 = FTP (File Transfer Protocol) - upload files into a file share
- 22 = SFTP (Secure File Transfer Protocol) - upload files using SSH
- 80 = HTTP - access unsecured website.
- 443 = HTTPS - access secured websites
- 3389 = RDP (Remote Desktop Protocol) - log into a windows instance

EC2 Instance Purchasing options.

- On-Demand Instances
- Reserved
- Saving plans
- Spot Instances
- Dedicated Hosts
- Dedicated Instances
- Capacity Reservation.

EC2 Instance Storage Section

EBS volume



EBS snapshot

EBS snapshot archive (75% cheaper)

EC2 Image Builder

Use to Automate the creation of virtual machines or container images.

EC2 Instance Store → Hardware

EBS →

Elastic File System

- Managed Network File System that can be mounted on 100s of EC2
- EFS - Infrequent Access

Amazon FSx

- Launch 3rd party high performance file systems on AWS
- Fully managed service.

IOPS - Input output operations per second.

Amazon FSx for Lustre
" " " windows file server

ELB (Elastic Load Balancer)

4 kinds of load balancers offered by AWS.

- Application LB (HTTP/HTTPS only) - layer 7
- Network LB (ultra-high performance, allows for TCP) - layer 4
- Gateway LB - Layer 3
- Classic LB (retired in 2023) - layer 4 & 7

7 OSI

Application layer	7
Presentation layer	6
Session layer	5
Transport layer	4
network layer	3
Data link layer	2
Physical layer.	1

Auto Scaling Groups (ASG)

- Manual scaling: update the size of an ASG manually.
- Dynamic Scaling: Respond to changing demand.
  - Simple / step Scaling.
  - Target Tracking Scaling.
  - Scheduled Scaling.
  - Predictive scaling.

Amazon S3

Buckets - directories
objects - files.

- Buckets are defined at the region level.
- S3 looks like a global service but buckets are created in a region.

Types:

- Amazon S3 Standard - General Purpose
- Standard Infrequent Access (IA)
- One zone Infrequent Access
- Glacier Instant Retrieval
- Glacier Flexible Retrieval
- Glacier Deep Archive
- Intelligent Tiering.

Can move btw classes manually or using S3 Lifecycle configurations.

S3 Durability

- High Durability (99.999999999, 11 9's)
- If your store 10,000,000 objects with Amazon S3, you can on average expect to incur a loss of a single object once every 10,000 years.

Availability

- Measure how readily available a service is.
- Varies depending on storage class.
- Eg. S3 standard has 99.99% avail = not available 53 min a year

AWS Snowball

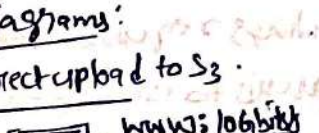
- Highly secure, portable devices to collect & process data at the edge & migrate data into & out of AWS
- Helps migrate upto petabytes of data.

+ AWS Snowball: offline devices to perform Data migrations.

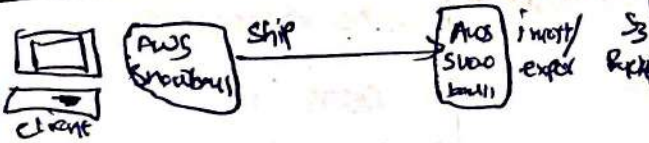
If it takes more than a week to transfer over the network, use Snowballs devices.

Diagrams:

Direct upload to S3.



with Snowball



Hybrid cloud for storage

- AWS pushing for "hybrid cloud"
  - Part of your infra. is on-premises
  - Part of your infra. is on the cloud.

S3 is a proprietary storage technology (unlike EFS/NFS), so how do you expose the S3 data on-premise?

- AWS Storage Gateway!

Types:

- File Gateway
- Volume Gateway
- Tape Gateway

Databases

Amazon RDS - SQL

Amazon Elastic Cache.

- Elastic cache is to get Managed Redis or Memcached

DynamoDB - noSQL

DynamoDB Accelerator - DAX

Diff. with Elastic cache at the CCP level: DAX is only used for & is integrated with DynamoDB, while Elastic cache can be used for other databases!!

- Active-Active replication (read/write to any AWS region)

Redshift

- Redshift is based on PostgreSQL, but it's not used for online Transaction Processing (OLTP)
- It's OLAP

Amazon EMR (Elastic Map Reduce)

ECS - Elastic Container Service.

- Launch Containers on AWS but you must Provision & Maintain the infrastructures

Fargate

- Launch Docker Containers on AWS
- You do not Provision the infra (no EC2 instances to manage)
- Serverless

ECR - Elastic Container Registry

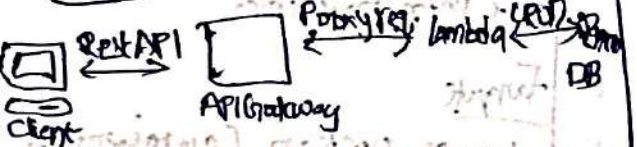
- Private Docker Registry on AWS
- This is where you store your Docker images so they can be run by ECS or Fargate.

Amazon EKS - Elastic Kubernetes Service

Aws Lambda Vs EC2

• Virtual functions no servers to manage	• Virtual servers in the Cloud
• Limited by time-short executions	• Limited by RAM & CPU
• Run-on-demand	• Continously running
• Scaling is automated	• Scaling means intervention to add/remove servers

Amazon API Gateway



LightSail

- Simpler alternative to EC2, RDS, ELB, EBS, Route 53 ---
- Has HA but no scaling, limited AWS integrations

@Cloudformation

- Cloudformation is a declarative way of outlining your AWS Infrastructure, for any resources (most of them are supported).
- Infrastructure as code.

Aws Beanstalk

- Elastic Beanstalk is a developer centric view of deploying an app on AWS.
- Beanstalk = Paas.

Aws Macie

→ Amazon Macie is a fully manage "data security & data privacy service" that uses Machine Learning & pattern to "discover & protect your sensitive data in AWS.

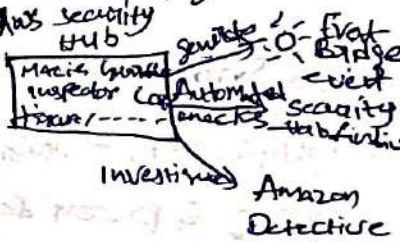
- Macie helps identify attributes to sensitive data, such as personally identifiable information (PII)

S3 → analyse → Macie → notify → Amazon Eventbridge

Aws Security Hub

central security tool to manage security across several AWS accounts & automate security checks.

- Integrated dashboards showing current security & compliance status.
- Automatically aggregates alerts in predefined or personal findings formats from various AWS services & AWS Partner tools:
  - Config
  - Guardduty
  - Inspector
  - Macie
  - IAM Access Analyzer
  - Aws Systems Manager
  - Aws firewall manager
  - Aws Health.
  - Aws Partner Network solutions.
- Must first enable the AWS Config service.



Amazon Detective

- Guardduty, Macie & Security Hub are used to identify potential security issues or findings.
- Sometimes security findings require deeper analysis to isolate the root cause & take action - it's a complex process.

Aws organizations

- Global Service
- Api is available to automate
- Aws account creation.
- Restrict account privileges using service control policies.

Multi account Strategies

- Use tagging standards for billing purposes.
- Enable CloudTrail on all accounts, send logs to central S3 acct.
- Send cloudwatch Logs to central logging account.

Aws Control Tower

- Easy way to set up & govern a secure & compliant multi-account Aws env.
- Aws Control Tower run on top of Aws organizations:
- It automatically sets up Aws Organizations to organize accounts & implement SCPs (Service Control Policies)

Aws Resource Access Manager (Aws RAM):

- Share Aws resources that you own with other Aws accounts.
- Supported resource include Aurora, VPC Subnets, Transit Gateway, Route 53, EC2 Dedicated hosts,

License Manager contd

Pricing Models in Aws:

1. Pay as you go
2. Save when you Reserve
3. Pay less by using more
4. Pay less as Aws grows

Free Services

- IAM
- VPC
- Consolidated Billing
- Elastic Beanstalk
- Cloud formation
- ASG

(You do Pay for the resources created)

Compute Pricing- Lambda & ECS

- Lambda
  - Pay per Call
  - Pay per Duration.
- ECS:
  - EC2 Launch Type Model: No additional fee, you Pay for Aws resources stored & created in your applications.
- Fargate:
  - Launch Type Model: Pay for vCPU & memory resources allocated to your applications in your containers.

Billing & costing Tools

- Estimating costs in the Cloud:
  - Pricing Calculator.
- Tracking costs in the cloud:
  - Billing Dashboard.
  - Cost Allocation Tags.
  - Cost & Usage Reports.
  - Cost Explorer.
- Monitoring against cost plans:
  - Billing Alarms.
  - Budgets.

Billing Alarms in Cloudwatch

- Billing data metric is stored in Cloudwatch us-east-1
- Billing data are for overall worldwide Aws costs.

Amazon Guardduty - Threat detection service.

" Inspector - is an automated security assessment service.

Trusted Advisor

- No need to install anything - high level Aws account assesment
- Analyze your Aws accounts & provide recomendation on 6 categories:
  - Cost optimization
  - Performance
  - Security
  - Fault tolerance
  - Service limits
  - operational Excellence.

Aws Support Plans Pricing:

- Basic support: Free
- Developer (29$ or 3% of monthly Aws charges)
- Business
- Enterprise on-Ramp
- Enterprise

Advanced Identity section

Aws Security Token Service (STS):

- Service/tool allows you to create & provide trusted users with temporary security cre. that can control access to your Aws resources.

Note:-

6 adv. of cloud computing

- Trade capital expense for variable expense.
- Benefit from massive economies of scale.
- Stop guessing capacity
- Increase speed & agility
- Stop spending money running & maintaining data centers.
- Go global in minutes.

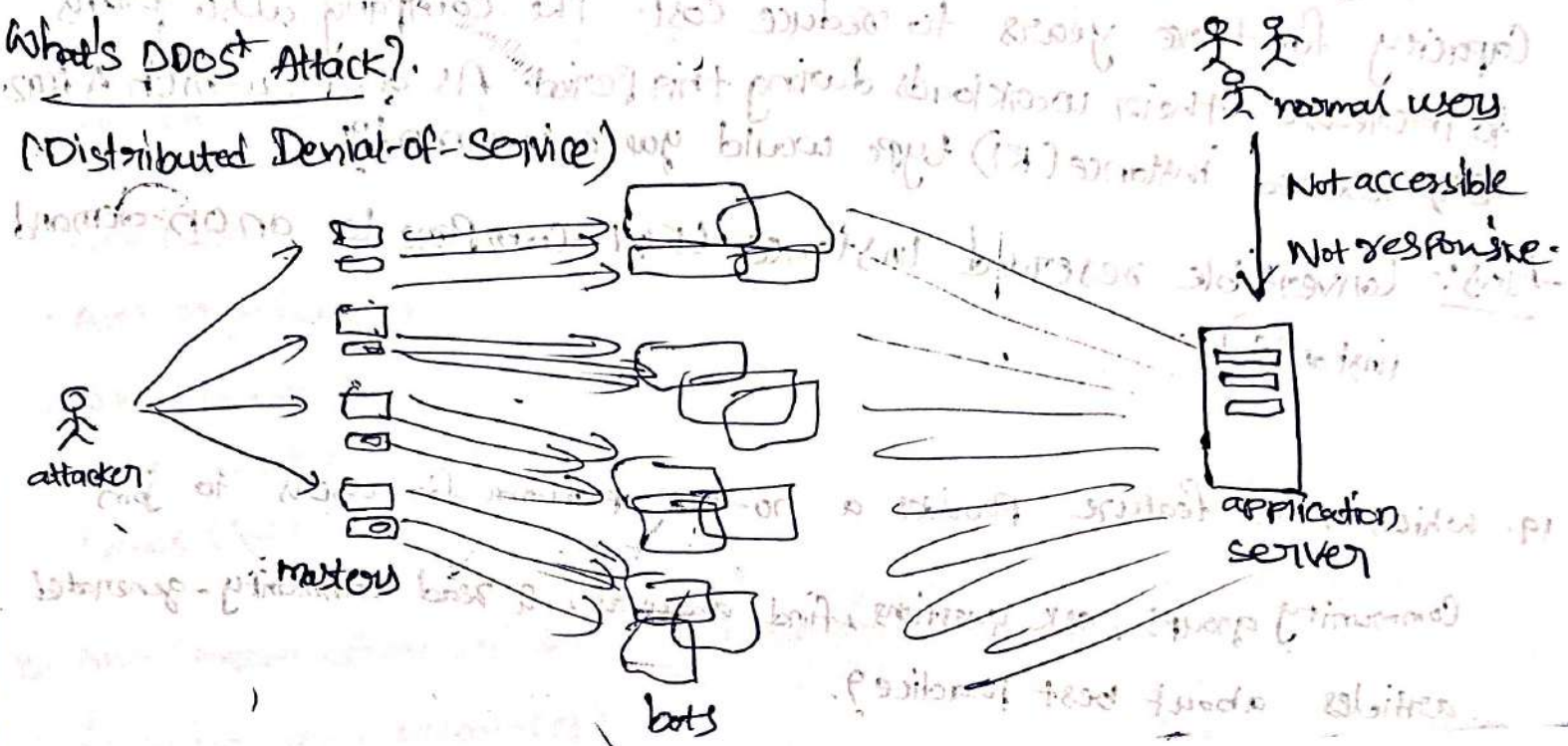
Aws-Fault Injection Service (AWS FIS)

→ Aws Fault injection service (FIS) is a fully managed service for running fault injection experiments on Aws that makes it easier to improve an app's performance.

Security & Compliance

What's DDoS Attack?

(Distributed Denial-of-Service)



DDoS protection on AWS

- Aws shield standard → free all customers.
- Aws shield Advanced: 24/7 premium DDoS Protection.
- AWS WAF: Filter specific requests based on rules. (web application firewall)
- Cloudfront & Route53:
  - Availability protection using global edge network.
  - Combined with Aws shield, provides attack mitigation at the edge.
- Be ready to Scale - leverage AWS Auto Scaling.

Aws network firewall

- Protects your entire VPC
- From layer 3 to layer 7 protection.
- Any direction, you can inspect
- VPC to VPC traffic.
- outbound to internet.
- inbound from internet.
- To / from Direct connect & site-to-site VPN.

(7) Application layer.
(6) Presentation layer.
(5) Session layer
(4) Transport layer
(3) Network layer
(2) Data link layer
(1) Physical layer

AWS Firewall Manager

- Manage security rules in all accounts of an AWS organization.
- Security policy: common set of security rules.
  - VPC security Groups for EC2, Application LB etc.
  - WAF rules
  - AWS Shield Advanced.
  - AWS network firewall
- Rules are applied to new resources as they are created (good for compliance) across all & future accounts in your organisation.

Penetration testing on Aws cloud

- Penetration testing is when you're trying to attack your own infrastructure to test the security.
- Aws customers are welcome to carry out security assessments or penetration tests against their Aws infra. without prior approval for 8 services:

Prohibited Activities:

- DNS zone walking via Amazon Route 53 Hosted Zones.
- Denial of service (DoS), Distributed Denial of service (DDoS),
- Simulated DoS, Simulated DDoS
- Port flooding
- Protocol flooding.
- Request flooding (login request flooding, API request flooding)

Account Management Billing & Support

AWS organizations

- Global Service
- Allows to manage "multiple AWS accounts.
- The main account is the master account
- Cost Benefits:
  - "Consolidated Billing" across all accounts - Single Payment method.
  - Pricing benefits from aggregated usage (Volume disc. for EC_2, S_3 ---)
  - Pooling of Reserved EC2 instances for optimal Saving
- API is available to Automate AWS account creation.
- Restrict account privileges using Service control Policies (SCP)

Service Control Policies (SCP)

- Allow or Blacklist IAM actions.
- Applied at the OU or Account level.
- Does not apply to the master account.

AWS Control Tower

AWS Resource Access Manager (AWS RAM)

- Share ~~your~~ AWS resource that you own with other AWS accounts.
- Share with any account or within your organization.
- Avoid resource duplication.
- Supported resources include Aurora, VPC Subnets, Transit Gateway, Route 53, EC2 Dedicated Hosts, License Manager configs.

AWS Service Catalog

- users that are new to AWS have too many options & may create stacks that are not compliant / inline with the rest of the organization.
- Some users just want a quick self-service portal to launch a set of authorized products pre-defined by admins.
- Include: VM's, dbs, storage options etc.

Pricing Models in AWS

- AWS has 4 Pricing Models:

- Pay as you go.
- Save when you reserve:
  - Reservations are available for EC2 Reserved Inst, DynamoDB Reserved Capacity, Elastic cache Reserved nodes, RDS Reserved Instances, Redshift Reserved Nodes
- Pay less by using more: Volume-based discounts.
- Pay less as AWS grows.

Free Services:

- IAM
- VPC
- Consolidated Billing.
- Elastic Beanstalk
- CloudFormation.
- Auto Scaling Groups (ASG)

Aws Trusted Advisor	Aws Inspector	Aws Guard Duty
• Purpose: Account-level best Pratice recommendations. • Focus Areas: • Cost optimization • Performance. • Security • Fault Tolerance. • Service limits. • Use case: Helps you optimize Aws resources & improve security & Performance. Eg:- Alerts you if an S3 bucket is publicly accessible, or if you're underutilizing EC2 instances.	• Purpose: Automated Security assessment of EC2 Instances & Container workloads. • Focus Areas: • Vulnerabilities in your app's or OS • Software dependencies • Network exposure. • Use case: Identifies security vulnerabilities & deviations from best Practices within individual resources. • Eg:- Scans an EC2 instance & flags outdated software or missing Patches.	• Purpose: Threat detection & Continous monitoring. • Focus Areas: • Suspicious activity (eg: malware, Port Scanning) • Anomalies in logs. • IAM behaviour. • Use cases: Uses machine learning to detect active threats in your Aws account. Eg:- Detects if an IAM user is accessing from an unusual location or if an EC2 Instance is communicating with a known malicious IP.

Trusted advisor = cost + Perform + Security Best Practs.

Inspector = Scan resources for known vulnerabilities.

Guard duty = Detect active threats in your Aws environment.