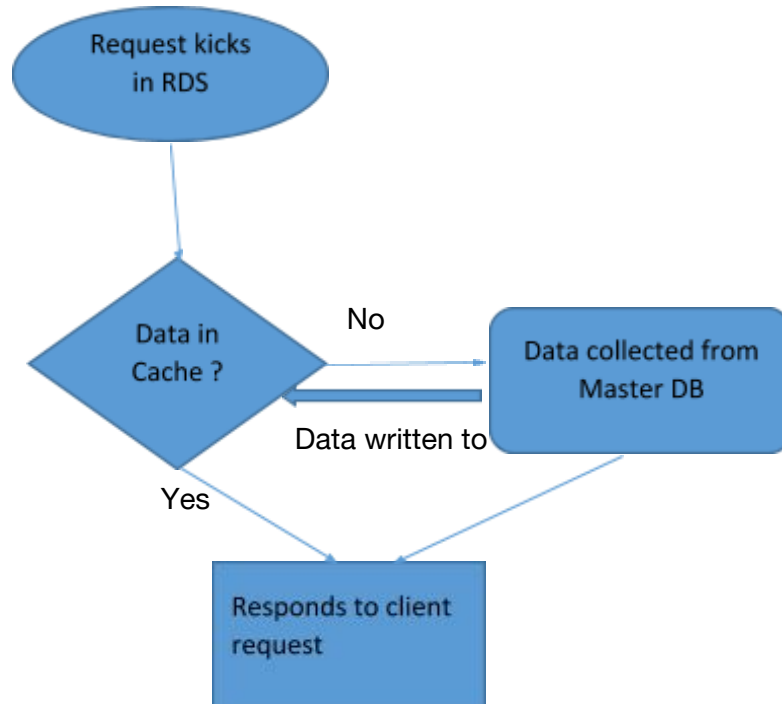


## PGP in Cloud Computing

### Managed Services on AWS

- There can be multiple databases running as a part of the **RDS**.
- In multi AZ environment, the database is automatically **replicated** to another AZ as soon as it is created.
- Sync between database in different AZ occurs with every request whenever the data is modified.
- No actions are performed directly on the secondary copy of database.
- **Read replicas** help read workloads.
- Read replicas are asynchronous and can be created in a different region.
- When a read replica is created, the primary database snapshot is taken and moved over to another region where the database is created.
- Read replicas are **eventually consistent**.  
[https://en.wikipedia.org/wiki/Eventual\\_consistency](https://en.wikipedia.org/wiki/Eventual_consistency)
- We can access the primary or the master database and the read replicas directly but not the secondary database.
- If the master db is deleted and the whole region goes down, The read replica will promote itself and act as a normal instance.
- If the master db goes down but the secondary is up, read replicas will now be created from secondary.



- Criterion to evict the cache entry
  - ❑ Time to Live (TTL) – Number of days after which the entry can be deleted.
  - ❑ Least Recently Used – Data that is not accessed for a longer time

## PGP in Cloud Computing

- **Elasticache** is the managed service for caching. It uses Redis or Memcached

[https://en.wikipedia.org/wiki/Amazon\\_ElastiCache](https://en.wikipedia.org/wiki/Amazon_ElastiCache)

- **SNS(Simple Notification Service)** is a push based notification model. Endpoints in SNS can be :

http,lambda,email,mobile devices,SMS,SQS

- SNS is a publish subscribe model. Create a topic >> all subscribers subscribe to the topic >> triggers the endpoint
- **SQS(Simple Queue Service)** is a pull based notification model. It is a queue based system and here the endpoints ASK for triggers.
- **Amazon MQ(Message Queue)** is a standard API which helps the apps integrate easily to the cloud.

<https://aws.amazon.com/about-aws/whats-new/2017/11/introducing-amazon-mq/>

- In SQS, there are producers and consumers of each queue. Maximum amount of data that can be put on the queue is limited.
- Queue types - Standard or FIFO
- A standard queue has much more throughput than FIFO.
- In FIFO, the messages are delivered in the order and exactly once whereas in standard queue, messages are delivered randomly and occasionally, they can be delivered twice.
- For idempotent operations, set a primary key to identify unique messages before sending them to the queue.
- The consumer consumes a message and acknowledges it and then the message is deleted. If the acknowledgement is not received, the delivery will be attempted again after a time interval called the **default visibility timeout**.
- SQS is ideally suited if you have small elements of data.
- **Received message wait time** is how long SQS will wait for response after polling for a message before it returns and empty set.
- **Dead letter queue** or the content redrive is the queue where the message is transmitted if there is a problem in processing the record in the mail queue. It can be used to handle exception flows.

[https://en.wikipedia.org/wiki/Dead\\_letter\\_queue](https://en.wikipedia.org/wiki/Dead_letter_queue)

- **Message fan out** is the process of sending out a message to multiple different queues for different users.
- In SNS and SQS, the users that can send and receive messages need to be given access in the policy.
- **Athena** is a service used to run SQL queries on different types of data in S3.

<https://aws.amazon.com/documentation/athena/>

- Athena is priced based on the amount of data that is spanned while running the query.
- **Partitioning** helps in reducing the data spanned for a query and thereby reduce the associated costs.

## PGP in Cloud Computing

- **AWS glue** is a central repository of metadata of different forms of data used by different managed services. A way by which you can visualize what your entire dataset is, on your managed services.

<https://aws.amazon.com/glue/>

- **Quicksight** is a visualization service. It creates various presentations of data with filters applied.

<https://aws.amazon.com/documentation/quicksight/>

- **SPICE**- Super fast, Parallel, In-memory, Calculation Engine is characteristic of Quicksight.
- **Quicksight** creates various presentations of data with filters applied.
- **Cloudwatch** is a monitoring service integrated with a lot of AWS services. All information from different sources comes in a single dashboard.

<https://aws.amazon.com/documentation/cloudwatch/>

- Logs from a variety of resources are organised into log groups. Logs can be exported from cloudwatch and stored in AWS storage and can further be used by other managed services.
- Principle of cloudwatch is to monitor CPU utilization.
- Cloudwatch also has an EC2 agent. This agent installed on EC2 instance can point to multiple log files. It pulls all those log files and pass it on to cloudwatch.
- Alerts can be set up on resources or pricing based on customised rules.
- One configured, cloudwatch metrics cannot be deleted, they expire as per their schedule

logs collected every <60sec are stored for 3 hours

logs collected every 1 min are stored for 15 days

logs collected every 5 mins are stored for 63 days

logs collected every 1 hour are stored for 15 months

- Data points are initially published with shorter period and are aggregated together for long-term storage.
- **Amazon Kinesis** is a data streaming service with two components -

Streams and Firehose

<https://aws.amazon.com/documentation/kinesis/>

- **Streams** are like pipes where at one point there are producers and the other end are the consumers.

Producers - from where the data is generated and pushed into the kinesis stream

Consumers - a firehose

## PGP in Cloud Computing

- Streams comprise of Shards

**Shard** is a sequence of data records in a stream.

Each shard can support 5 read transaction per sec upto a maximum read rate of 2MBps and 1000 records per sec for writes upto a maximum of 1MBps write rate.

[https://en.wikipedia.org/wiki/Shard\\_\(database\\_architecture\)](https://en.wikipedia.org/wiki/Shard_(database_architecture))

- A data record
  - ☐ Sequence number
  - ☐ Partition Key
  - ☐ Data Blob
- Partition key** is hashed and as per its value a shard will be chosen for data transfer. Data can be uniformly distributed across shards based on the partition key.
- Sequence numbers** are unique and auto assigned by shards and are in order.
- Firehose** can create a delivery stream to either of these:
  - ☐ Amazon S3
  - ☐ Amazon Redshift
  - ☐ Elasticsearch
  - ☐ Splunk

<https://docs.aws.amazon.com/firehose/latest/dev/what-is-this-service.html>

- Buffer conditions - Firehose waits for a buffer size or buffer interval of time to collect data before sending it out to Amazon S3.
- Firehose needs to be assigned an IAM role to access the target.
- The data stream can also be transformed before firehose delivers it to the destination.
- Kinesis provides ordering of records and real-time processing of data while Amazon SQS stores and queues messages not necessarily in any particular order.
- Kinesis is good for streaming the data (binary or text). It is not designed for being used as a messaging system between two applications.
- IaaS** key aspects
  1. Enterprise architecture
  2. Cloud hosting
  3. Virtual Data Centres
  4. Scalability
  5. No investment in hardware
  6. Pay as you go
  7. Location independence
  8. Physical security of data centre - managed by cloud provider
  9. No single point of failure

[https://en.wikipedia.org/wiki/Infrastructure\\_as\\_a\\_service](https://en.wikipedia.org/wiki/Infrastructure_as_a_service)

- PaaS** key aspects

## PGP in Cloud Computing

1. Managed service - factor in all APIs of PaaS while writing the application code.
2. Managed runtime - no need to start or stop services
3. Managed infrastructure
4. Monitoring
5. Auto Scalability
6. Billing thresholds
7. Choice of programming language

[https://en.wikipedia.org/wiki/Platform\\_as\\_a\\_service](https://en.wikipedia.org/wiki/Platform_as_a_service)

- **SaaS** key aspects
  1. Third party provider will be managing applications
  2. Multi tenant architecture
  3. Usually online application
  4. Service provider delivers as per SLA
  5. No client side installation

[https://en.wikipedia.org/wiki/Software\\_as\\_a\\_service](https://en.wikipedia.org/wiki/Software_as_a_service)

- **Stage 1**  
Deploy the application on cloud behind a load balancer  
Use IaaS
- 
- **Stage 2**  
Break your application into smaller chunks (Microservices)  
Start using the managed services provided by the cloud service provider.
- **Stage 3**  
Factor in the cost splitting into dedicated and dynamic capacities to optimize cost.

- There are **architectural templates** from AWS which should be used to start and customize it as per your use case.  
<https://aws.amazon.com/architecture/>
- In a multi tenancy system, the business logic and data access layer can be kept stateless for the system to work well.  
The Database can be partitioned logically by having multiple schema definitions for different tenants. Alternatively, split the database by having separate tables for logical partitioning.
- If collocation of data is not acceptable by customers - physical partitioning can be done in 2 ways :
  - Having separate databases
  - Having the same DB engine and storing data on different disks for different customers.
- **Sharding** is splitting data into separate components. Might be logical or physical.  
[https://en.wikipedia.org/wiki/Shard\\_\(database\\_architecture\)](https://en.wikipedia.org/wiki/Shard_(database_architecture))

## PGP in Cloud Computing

- Defining **application zones** - the combination of business functionalities is one mini application hosted independently.
- **Elastic Beanstalk** is a service which manages deploying a web application in a high availability environment using a single console.  
[https://en.wikipedia.org/wiki/AWS\\_Elastic\\_Beanstalk](https://en.wikipedia.org/wiki/AWS_Elastic_Beanstalk)  
PaaS where the infrastructure settings have to be configured by the user.
- **Serverless** - write the code and deploy it without having to worry about the hardware at all.
- **Lambda** is a managed service by AWS which runs codes written in a number of languages. There is no fixed commitment in pricing. You pay only when the lambda function is called.  
[https://en.wikipedia.org/wiki/AWS\\_Lambda](https://en.wikipedia.org/wiki/AWS_Lambda)
- It should be used for short running processes. For longer running processes, use **AWS batch**.
- Lambda function can be deployed from a multiple places:
  - Write it on the AWS console in a code editor
  - Use an IDE like eclipse and push the code to the lambda function
  - If it has dependencies, bundle the code with its dependencies using the AWS CLI and throw it on a lambda function.
  - CI/CD processes which can create an artifact which can be deployed to a lambda function.
- If any business process requires multiple lambda functions to run and the gap between the two processes is more as in a long running business process, **AWS step function** can be used to orchestrate those calls.  
[https://docs.aws.amazon.com/step-functions/index.html#lang/en\\_us](https://docs.aws.amazon.com/step-functions/index.html#lang/en_us)
- **Cognitive Services**
  - Managed Services
  - Saves businesses the trouble of implementing these from the ground up
  - IAM roles should be created with API access
  - Roles need to have policies associated with Rekognition, Polly and Comprehend.
  - Prerequisites :
    - Shell on Ubuntu, Terminal on MacOS or Git Bash on Windows
    - AWS CLI installed
    - Preferred language with supported AWS SDK (Python 3.x is used for the purposes of this demonstration)
- **Rekognition**
  - Used for text, object and facial recognition from images
  - Boto3 is the Python library used for AWS SDK

## PGP in Cloud Computing

- Confidence level shows the percentage of certainty of image recognition

<https://docs.aws.amazon.com/rekognition/latest/dg/how-it-works.html>

- **Comprehend**

- Used for natural language processing
- Can detect the following components from text
  - Dominant language
  - Entities
  - Key phrases
  - Sentiment
  - Syntax

<https://docs.aws.amazon.com/comprehend/latest/dg/getting-started.html>

- **Polly**

- Convert written text to speech
- Different accents, voices, and genders are available to use depending on the use-case.
- Default voice if not specified is “Amy”
- Pydub library can be used for voice playback

<https://docs.aws.amazon.com/rekognition/latest/dg/how-it-works.html>