

AWS LAMBDA

- AWS LAMBDA is **Serverless**, compute service that lets you run code for virtually any type of application or backend service without managing the servers.
- We can trigger lambda from over 200 aws services and only pay for what you use.
- No need for any external server, it will provide server.
- for example: if you write code ,that will run for 5 min then we have pay only for that 5 min. No additional pay is required.

Use Case:

1. Cost optimizations

For example, we have extra services running, for those extra services we have pay an amount to overcome this we can use the lambda, how?

AWS services-AWS Lamba-Code-notification- cost

how means- we can write a one lamba function, in that we can write code, create one notification as these many services are running, that services we are not currently using like this, this will help to optimize the cost

2. Data engineer pipeline application

✓ ETL-AWS GLUE-LAMBDA

for example, we have a one ETL ,for that we have created AWS GLUE to trigger this glue we ca write lambda function

✓ S3 storage-new data-notification

✓ Failure

3. File Transfer

client-s3 storage-aws lambda-resize-s3 bucket

for example, client will store an image to s2 , then create a one lambda function as "the image have been uploaded", the images have been resized and then again stored into s3 bucket, then it is show on to the website

Benefits:

- 1.no need for manager server
- 2.auto scaling
- 3.pay as you go
- 4.performanace optimization

AWS EC2: elastic compute cloud

- ✓ Amazon EC2 is a web service that provides sizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers.
- ✓ Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud
- ✓ Access reliable, scalable infrastructure on demand.
- ✓ Provide a secure compute for your applications. Security is built into the foundation of Amazon EC2 with the AWS Nitro System.
- ✓ Optimize performance and cost with flexible options like AWS Graviton-based instances, Amazon EC2 Spot instances, and AWS Savings Plans.
- ✓ It will create a virtual machine to store data. or it will create a server to store data.

AWS GLUE

- ✓ It is a cloud optimized, Extract Transform and load service.
- ✓ It allows you to locate and transform all your datasets across your business, so that you can put them to use.
- ✓ why glue is different from other ETL products in three ways:
 1. it is serverless
 2. automatic schema inference
 3. auto generate etl scripts

- ✓ when should I use glue:
 1. to build data warehouses to organize, cleanse, validate and formats data.
 2. to create etl pipeline
 3. to run serverless queries against your s3 data lake.

- ✓ How will the glue architecture work?

example we have data on S3 that is like csv or json, then the in the glue, we have crawler, crawler will go to the data store to extract the schema of the csv or json then it will store in the data catalog. so data catalog contains all the schema of the data stored in the s3 using crawler, now we have schema of the data, then the jobs come into the picture, after that ETL pipelines are create then it is stored in the data target.

AWS Athena

- ❖ It is a Query service that makes it easy to analyze data directly from S3 using standard SQL.
- ❖ but it is not a database or data warehouse.
- ❖ It will allow you to write standard SQL Query on top of the S3 Data.
- ❖ **features:**
 - ✓ serverless
 - ✓ easy to query
 - ✓ flexible pricing
 - ✓ fast performance
 - ✓ secure
- ❖ **price:**
 1. We have pay only the query ,what we will run.
 2. charged for the number of bytes scanned per Query, rounded up to the nearest megabyte.,
 3. with a 10MB minimum per query.

SNS: Simple Notification Service

- ❖ Amazon Simple Notification Service (Amazon SNS) is a managed service that provides message delivery from publishers to subscribers (also known as producers and consumers).
- ❖ Publishers communicate asynchronously with subscribers by sending messages to a topic, which is a logical access point and communication channel.
- ❖ Clients can subscribe to the SNS topic and receive published messages using a supported endpoint type, such as Amazon Data Firehose, Amazon SQS, AWS Lambda, HTTP, email, mobile push notifications, and mobile text messages (SMS).

AWS Kinesis: Serverless streaming Data Service

- ✗ collect, process and analyze real-time video and data streaming.
- ✗ real time data such as video, audio, application logs, website clickstreams and IoT telemetry data for machine learning, analytics, and other applications.
- ✗ process and analyze data as it arrives and respond instantly instead of having to wait until all your data is collected before the processing can begin.

Amazon Redshift

- ✓ Amazon Redshift uses SQL to analyze structured and semi-structured data across data warehouses, operational databases, and data lakes, using AWS-designed hardware and machine learning to deliver the best price performance at any scale.
- ✓ Break through data silos and gain real time and predictive insights on all your data in a few clicks, with no data movement or data transformation.
- ✓ Gain up to 3x better price performance than any other cloud data warehouse, with performance innovation out of the box at no additional cost to you.
- ✓ Get data to insights in seconds without worrying about infrastructure management, with a secure and reliable analytics service.

