

Root

nasirzaheb@gmail.com

Date: _____

NASA

Page: _____

Neil Goh

nasirzaheb@gmail.com

AWS - IAM - Services.

Q1 what is IAM

It is an AWS function help you control access to resources.

Resources are the entities that you create in AWS - ie bucket, or object.

Q2 what is the difference b/w root user A/C and IAM user Account.

→ root user :- It the main account through which we can create n number of IAM user and has full access of all the resources.

→ IAM user :- It the A/C which is created by the root user user with a defined sets of permissions.



Scanned with OKEN Scanner

EC2 →

Q1 what is the full form of EC2?

~~Elastic Compute~~

Elastic Compute Cloud.

Q2 what is EC2?

It is an Amazon Service that let you run virtual computer in the cloud to run your application. You can easily scale up and scale down based on your need and you can only pay for what you use.

Q3 what are the benefits of EC2?

- Scalability :- Easy scale up & scale down based on the need.
- Cost-Effectiveness :- Pay for what you use.
- Flexibility :- Choose instance type, operating system, & storage options.
- Elasticity :- Automatic adjust computing capacity due to auto scaling.

Q4 Tell me 5 different EC2 instance that are available and when to use what instance give example.

① General purpose (t3, t4g, m5)

use for - Balanced CPU, memory.

CPU

② Compute optimized :- (C6g, E5)

Used :- Application that required high compute power.

③ Memory optimized :- (R5, X1e)

Use :- Application that require high compute. (RAn)

④ Storage optimized :- (I3, d2)

use :- large local storage.

⑤ GPU storage (P4, g4)

Machine learning, deep learning, & video rendering.

Q5 How to calculate the cost of EC2 instance. how is it charged:-

EC2 cost depends on instance type, running time
& storage - EBS (on left hand on EC2 instance)

It is calculated through AWS Pricing calculator.

90% off

what are spot instance and when to use them?

26

Spot instance are the instance available on AWS cloud at a discounted price compare to on demand instance.

Is mai Bid Karke hui spot instance ke lege-

Used :- Cost Savings

- Flexible workload
- Testing Purpose.

5 Linux command

Implementation

Q2 • df - h — available and used disk space.

• uptime — how long instance running

• top — ~~shows~~ real time memory Processing \rightarrow ^{CPU, memory usage} usage

• ls - la — list all files & directory

• sudo yum update -y \rightarrow install all update.

Q3

Using AWS CLI connect to EC2 instance.

- First download and install AWS CLI.
- aws --version C to check the version
- pip install awscli
- aws configure
- aws ec2 describe-instances. --query "Reservations[*].Instances[*].PublicIPAddresses"
 - Note down the Public IP address.
- ssh -i C:/Path of .pem file in download@^{Kali Linux} ec2 user @ < instance Public IP

Linux = ec2-user

Ubuntu = ubuntu

RHEL = ec2-user

CentOS = centos

Debian = admin

⇒ Then you are inside the instance
run the Linux com

df -h

top

ipconfig -----.

3 [ec2-user@ip-xxx-xxx-xx] ⇒ This confirm you
are inside the instance

Q4 How to use Putty to connect to EC2 instance , after connecting to EC2 instance using AWS CLI create a sample S3 bucket.

Already alone in Prevere doorway put
down Put.

- convert Pem to PPK using Putty ~~for gen~~
 - open up Public key and upload PPK.
 - Open.

Q5 Insert a sample text file in the S3 bucket using AWS CLI

aws S3 CP Path S3://mybucket/nafin.

3 | litgaws S3 1S

aws s3 sync ss://mybucket

remove ~~add~~ file from the bullet

aws ss tcb 83:11 ——— lemmore bucket,

S3

Q1 What is S3?

It is simple storage service where we can store structured, un-structured & semi-structured data. You can create 100 bucket in one A/C.

Bucket is the folder.

Each file is called object.

Q2 What is the full form of S3?

Simple Storage Service (S3)

Q3 What are the different storage classification available in S3?

- S3 Standard (General)
- S3 Standard (Infrequent)
- S3 Intelligent-Tiering \rightarrow unknown changing access pattern
- S3 One Zone - IA \rightarrow store in 1 AZ, infrequent access, recreate if AZ fails.
- S3 Glacier :- Archive storage for long term backup.
- S3 Glacier Deep Archive :- Archive data for least frequent access & with lowest cost storage.

Q4 How to enable versioning in S3?

To enable the version B

- First go to the bucket.
- Then go to the security and enable the version selection and save.

Q6 How to enable security in S3?

In bucket \Rightarrow Permission \Rightarrow Manage all access to the Public.

by using custom policy \Rightarrow Permission

\Downarrow
control of the user from the user

Q6 when to use S3 bucket, give 5 different example where S3 bucket usage would be more suitable.

- Backup files :- Stores important files safely
- Big data storage :- Keep large data files for analysis.
- Host a website :-
- Store Media file storage :-
- Log storage :- keep log from App or servers for future clustering.

Q7 what is sigma registry?

Sigma registry is a tool to ~~key~~ track the data format.

It ensures everyone is using the same format for reading and writing. or use the same schema when exchanging the data.

Commonly use in Apache Kafka.

Q3 Create bucket

import boto3

S3-client = boto3.
client('S3')

Response = S3-client.create_bucket(

Bucket = " ", "

CreateBucketConfiguration = {

'LocationConstraint': 'ap-southeast-1'

},

)
Print(response)

upload

S3-client = boto3.client('S3')

S3-client.upload_file(Filename = ' ', Bucket = ' ', LocalFile = ' ', BucketFile = ' ')

file name
uploaded

Response = S3-client.delete_bucket(

Bucket = " ", "

,

)

Print(response)

Q How to host the S3 bucket created in the web.

Hosting means making the content available in your S3 bucket on internet.

- Create bucket.
- unblock all Public access in Permission.
- ~~Block~~ Enable the static website hosting in Properties.
- Edit Public Policy in.

The is available on internet using Static website hosting link.

and change by using Cloud Access Test.

Q How to use KMS Service S3

key Management service . encryption & decryption

- Create the user. attach the Policy.
- Create the KMS
- Create bucket use one of the KMS Key.

They start working

Athena

Q1 what is Athena?

Athena help you to analyze structured, unstructured semi str. data store in Amazon S3 by using SQL.

Q2

when to use Athena, which scenario is useful.

- Query large data set quickly from S3 without loading it into the data base.
- Query structured, unstructured data from S3
- Pay only for the query you run.

Q3

How to integrate Athena to S3?

- upload to S3
- Create output bucket
- go to Athena console
- Select the output bucket
- create table.

Lambda

Date: _____
Page: _____

Q1 what is Lambda?

Lambda is a serverless compute service that let you run code without provisioning or managing servers.

It automatically execute your code in response to an event.

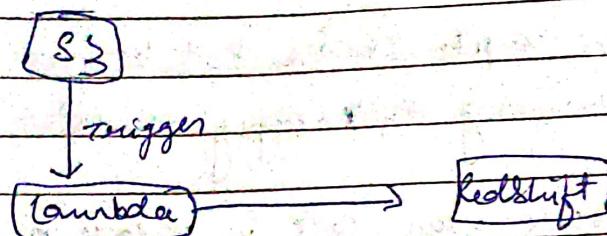
Q2 what are the real life use cases of Lambda, give three examples.

- Real File Processing :- Automatically process the data when file upload the S3.
- Real-Time Notifications :- send real time notification (email or SMS) when an event occur.
- Scheduled Task :- Automate regular task, deleting old data, generate daily report.

Implementation

Q-3

Write a lambda code to read data from S3 & load to redshift table.



- create the lambda function \Rightarrow sample , Python 3.8.
 - while basic version auto.
- Create a bucket - \circ inside the bucket. (Properties) \Rightarrow Event notification.

select - put

- Post

destination \Rightarrow Lambda function.

function specific Lambda function \Rightarrow 1. Lambda function created above.
choose

- Create Redshift cluster \Rightarrow provide user name and password
go to query editor v2 \Rightarrow Enter the credentials.
 \rightarrow Create table (optional ; not for cluster)
- cluster \Rightarrow properties.

\Rightarrow create the VPC connect b/w Lambda & Redshift

- Go to Lambda VPC \Rightarrow create common all details found
the Redshift properties.

DMS

Date: _____

Page: _____

Q8 what is DMS services.

It is a managed service that help you to migrate the database ~~to cloud storage~~ from one location to another location.

It support homogenous as well as heterogeneous migration.
If same type *If different type*

Q2 what are the types of data Migration on cloud.

① Storage Migration :- Moving the data from one storage system (on-premises storage) to cloud storage (like S3)

② Database Migration :- Moving the data from one location to another.

③ Application Migration :- Moving the entire application, along with their data, to the cloud.

Q3 what is slicing and dicing migration approach.

Slicing and Dicing Migration is an approach where data is divided into smaller, manageable chunks or "slices" before migration. These chunks are migrated and analysed separately and reduce the risk of data loss.

Slicing :- Breaking the large data set into smaller part.

Dicing :- Analyzing and validating each part after migrate

Q 4

what is lift and shift Migration?

Lift and Shift Migration is a cloud migration approach where application and data are moved to the cloud without any changes to their architecture or code.

Lift :- Moving the application from its current environment (on-prem)

Shift :- Rehosting it in the cloud like AWS, Azure.



Step functions

Date: _____

Page: _____

Q. What are step functions?

Step function are the serverless service that let you to coordinate multiple AWS service into workflows.

It enable you to define workflow with task that run in sequence and handle errors and manage retries.

It automate the process like ETL jobs, etc.

Q. How to enable the step function.

- 1) Go to AWS console
- 2) Stepfunction navigate
- 3) Create state machine.
- 4) Choose a template

Q. What is SQS and SNS service in AWS and how do you use it.

Amazon SQS (Simple Queue Service) :- It send messages without requiring item to be connected directly.

It has a decoupling component means one part of the system can send message without waiting for the other part.

2) SNS (Simple notification service) is able to send message to large number of subscribers.

Q4 Type of triggers in AWS? which trigger used in the project.

- ① Cloud Watch Event
- ② S3 event
- ③ SNS
- ④ Lambda

→ Lambda is the common choice.

Costing in AWS

Q1

How to calculate the AWS cost.

1. AWS Pricing calculator :- use it to estimate cost based on service
2. AWS budget :- set up budget to track.
3. Billing & cost :- check the actual usage and charges.

Q2

How to Optimize the Costing in AWS?

- Use the right instance
- Auto Scaling
- Use S3 infrequent Standard :-
- Stop the resources not in use.

Q3

How to Pricing can be checked

same as 1

Aurora

Q1 what is Aurora.

Aurora is relational database service provided by AWS.
It is compatible with MySQL & PostgreSQL.

Q2 what is the Architecture of Aurora. X not done.

Q3 what is the difference between RDS and Aurora?

RDS is a broader service support multiple database engine (MySQL, PostgreSQL, MariaDB, Oracle)

- Aurora is a high performance database with RDS
it has auto scaling storage.

5x faster for MySQL 2x faster than PostgreSQL

than traditional

- It work of MySQL and PostgreSQL.

Redshift

Q1 what is Redshift.

A Redshift is a cloud based service or a data warehouse.

that let you to store & analyze large amount of data

gives SQL to process and analyze big data on cloud.

Q2 Explain the Redshift Architecture?

- Leader node :- Leader node consider as Manager. It manages the load across the multiple compute node.
- compute node :- Each node has its own CPU and memory and divided into slices. and have multiple compute node.
- Storage node :- The compute node divided into slices. It is responsible for processing a portion of data in HDFS.

Q3 what is the difference b/w redshift and Hive Architecture.

Hive architecture work on Hadoop framework.

① Data storage

Redshift \Rightarrow It is use columnar storage format

Hive \Rightarrow It is HDFS.

Query Processing :-

- Redshift :- Massively parallel processing for executing SQL query.
- Hives :- It uses MapReduce to process query.

Q 4 How to calculate the cost of Redshift services.

Ans To calculate the cost of Amazon Redshift consider the cluster type, number.

Q 5 what are the optimization techniques of Redshift? Give 3-4 optimization techniques.

Q 6

what is the difference between AWS Redshift vs Athena?

Redshift :- Redshift is a data warehousing service that let you to store and analyze large amount of data. It uses SQL to process & analyze data.

Answer :- It is a fully managed service we can analyze data in real time, extract, transform, and load data directly from S3.

Q1 How to create RDS connection with Redshift.

- Create an RDS.
- Connect to SQL workbench DB
- Create Parameter group
- Select the RDS modify attach the parameter group.
- Create workspace & Namespace.
- Create role & go to workgroup. > Select ~~role~~ ->
Zero ETL Integration

Q2 What command can we use to connect the on-premises data to Redshift?

- Setup AWS DMS:- Create Replication instance.
- Create Source & Target Endpoint.
- Create a Replication task.

Q3 If query is taking more time to run in Redshift - how to solve the issue?

- Analyze the Query execution
- Table distribution
- Data compression
- Rewrite the query
- Workload Management.

Q How to copy the table in AWS Redshift?

→ ~~copy both data and schema~~
create new table AS
Select * from existing

→ copy schema.

Create Table new-table (like new_existing_table)

Q How to Merge the table in AWS Redshift?

here we use Merge to Join the table.

Merge into custo AS t.

Using new-cust AS a.

on t.customer_id = a.customer_id.

when Matched Then.

update set t.customer_name = a.customer_name
t.customer_email = a.customer_email

when Not matched Then

insert (customer_id, customer_name, customer_email)

Value (s.customer_id, s.customer_name, s.customer_email)

Q what is the base Server / database of Redshift?

Amazon Redshift based on PostgreSQL SQL

what types of table available in Redshift.

Regular table :- These are traditional table where the data are stored permanently.

Temporary table :- These are session specific. drop when the session end.

view table :- It is virtual table.

Q. what SCD2 apply using query in Redshift database?

Ans. Redshift involve SCD query to manage historical change, inserting new record with current date and update the old record with start & end date.