Cognizant Technology Solutions

A Project Report on

"EMPLOYEE PERFORMANCE TRACKING"

Prepared By

Ankita Shivane	2320390
Bhavesh Uprikar	2320803
Harshit Yadav	2320382
Mimansa Yadav	2320956
Nandhini T	2320539
Rajanya Kamilya	2320828
Shambhavi Shukla	2320529

REQUIREMENTS

Employee Performance Tracking:

- Import employee performance data into Amazon Redshift.
- Track employee performance metrics such as sales targets, customer satisfaction ratings, and productivity.
- Identify top-performing employees by analyzing performance metrics and comparing them to predefined targets or benchmarks.
- Identify areas for improvement by analyzing performance gaps and providing targeted training or support.
- Visualize employee performance data using charts and graphs (e.g., bar charts, radar charts) to identify trends and patterns.

INDEX

TOPIC

PURPOSE:

- **Data Analysis:** Perform complex analytics on large datasets stored in S3 using the powerful querying capabilities of Redshift.

- **Business Intelligence (BI):** Create reports, dashboards, and visualizations from data stored in S3 by loading it into Redshift for BI purposes.

- **Data Warehousing:** Store and organize large volumes of structured and semi-structured data from S3 into Redshift for centralized data warehousing.

- **Data Integration:** Integrate data from multiple sources stored in S3 into a single data warehouse in Redshift for comprehensive analysis and reporting.

- **Data Migration:** Migrate data from on-premises data warehouses or other cloud storage solutions to Redshift for better scalability, performance, and cost-effectiveness.

- **ETL (Extract, Transform, Load):** Extract data from S3, transform it as needed, and load it into Redshift for further analysis and reporting.

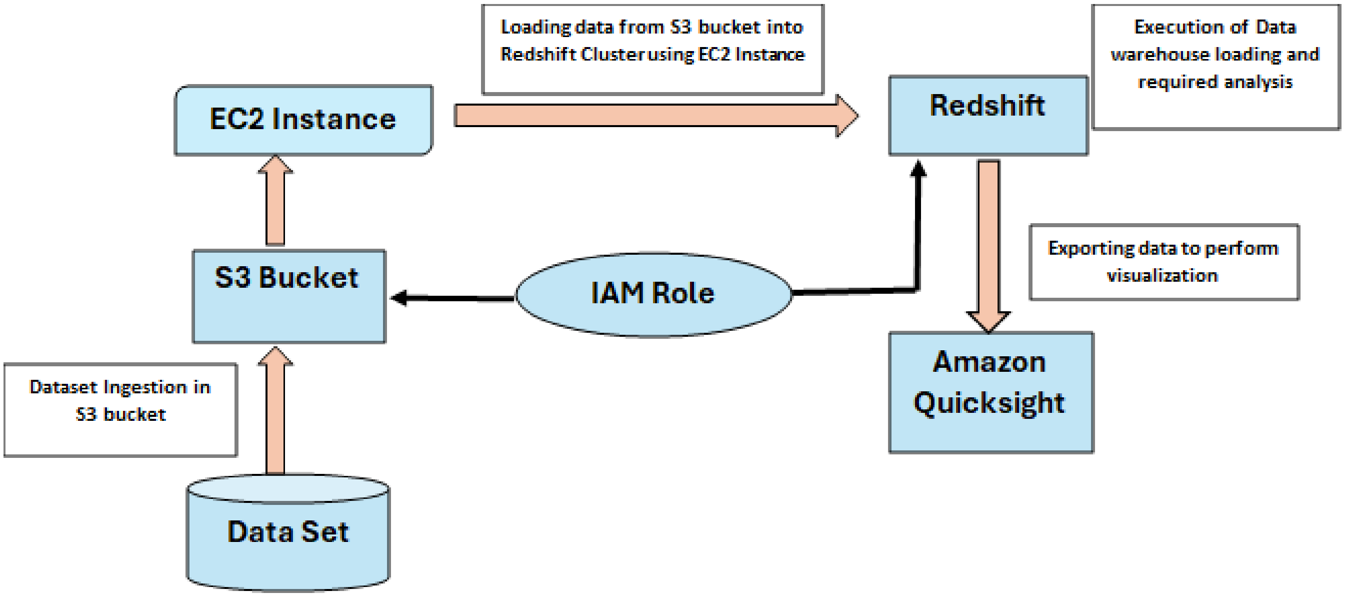
- **Real-time Analytics:** Continuously load streaming data from S3 into Redshift to perform real-time analytics and gain insights into rapidly changing data.

- **Cost Optimization:** Optimize costs by storing raw or historical data in S3's cost-effective storage and transferring only relevant data into Redshift for analysis.

- **Data Archiving:** Archive historical data from S3 into Redshift for long-term storage and analysis while keeping it easily accessible for future reference.

- **Data Governance and Security**: Centralize sensitive data stored in S3 into Redshift for better governance, access control, and security management, ensuring compliance with regulatory requirements.

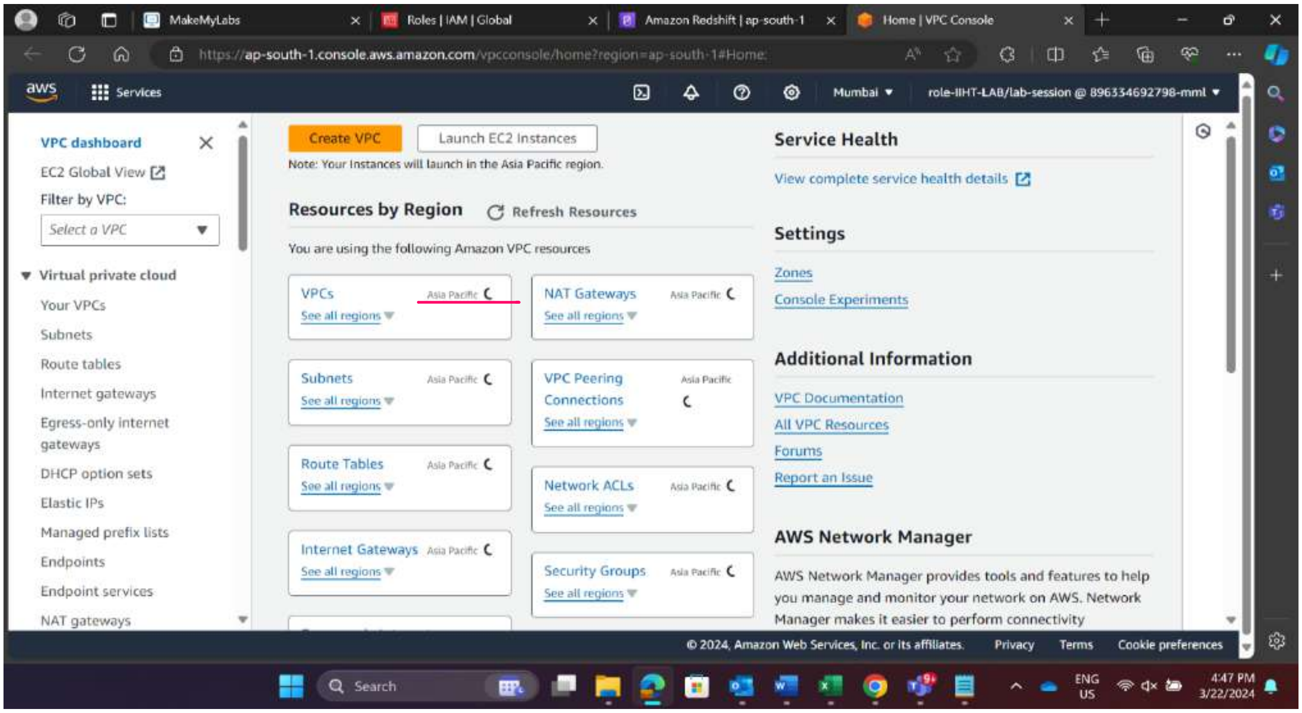
PICTORIAL FLOWCHART


Loading data from S3 bucket into Redshift Cluster using EC2 Instance
Execution of Data warehouse loading and required analysis
EC2 Instance
Redshift
S3 Bucket
IAM Role
Exporting data to perform visualization
Amazon Quicksight
Dataset Ingestion in S3 bucket
Data Set

DESIGN AND ARCHITECTURE

VPC CREATION

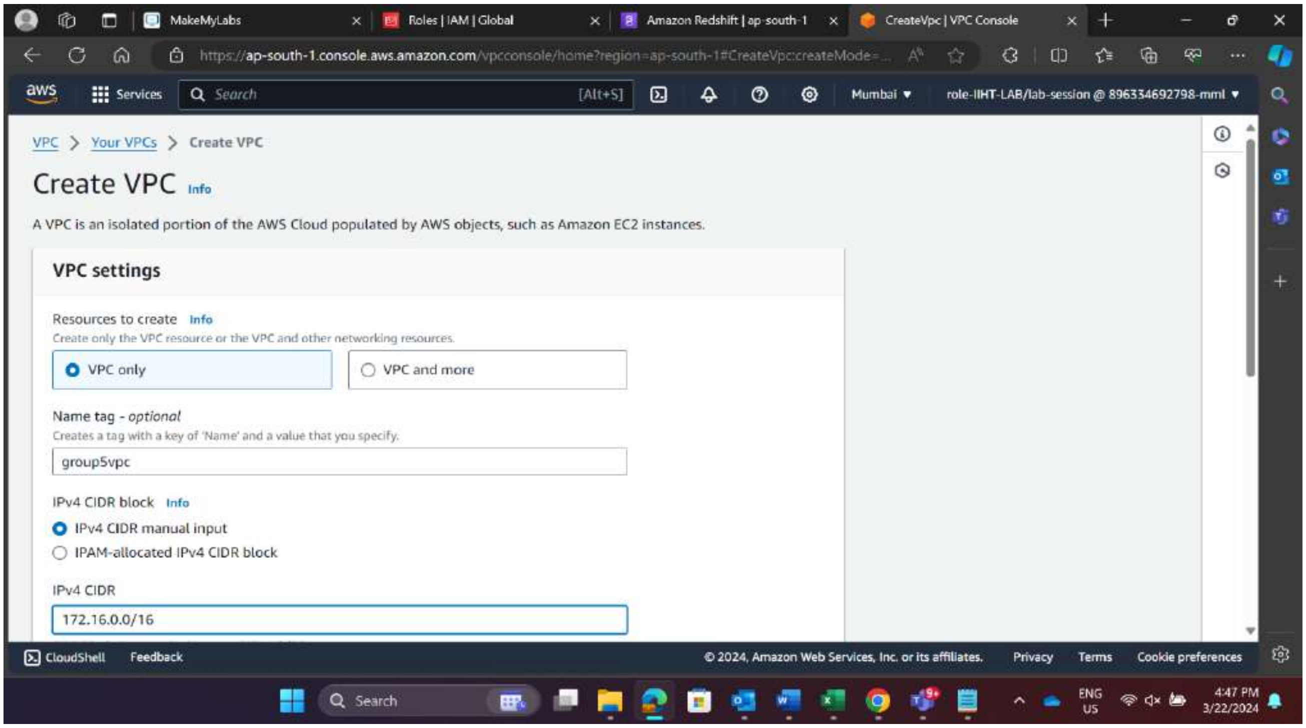
Step 1: Initiate the VPC creation process, provide the necessary details for your VPC, including the VPC name, CIDR block (the range of IP addresses for your VPC).



Note: Your Instances will launch in the Asia Pacific region.

You are using the following Amazon VPC resources

AWS Network Manager provides tools and features to help you manage and monitor your network on AWS. Network Manager makes it easier to perform connectivity


MakeMyLabs
Roles | IAM | Global
Amazon Redshift | ap-south-1
CreateVpc | VPC Console
https://ap-south-1.console.aws.amazon.com/vpcconsole/home?region=ap-south-1#CreateVpc:createMode=...
aws
Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
VPC
Your VPCs
Create VPC
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG
US
4:47 PM
3/22/2024

Create VPC Info

A VPC is an isolated portion of the AWS Cloud populated by AWS objects, such as Amazon EC2 instances.

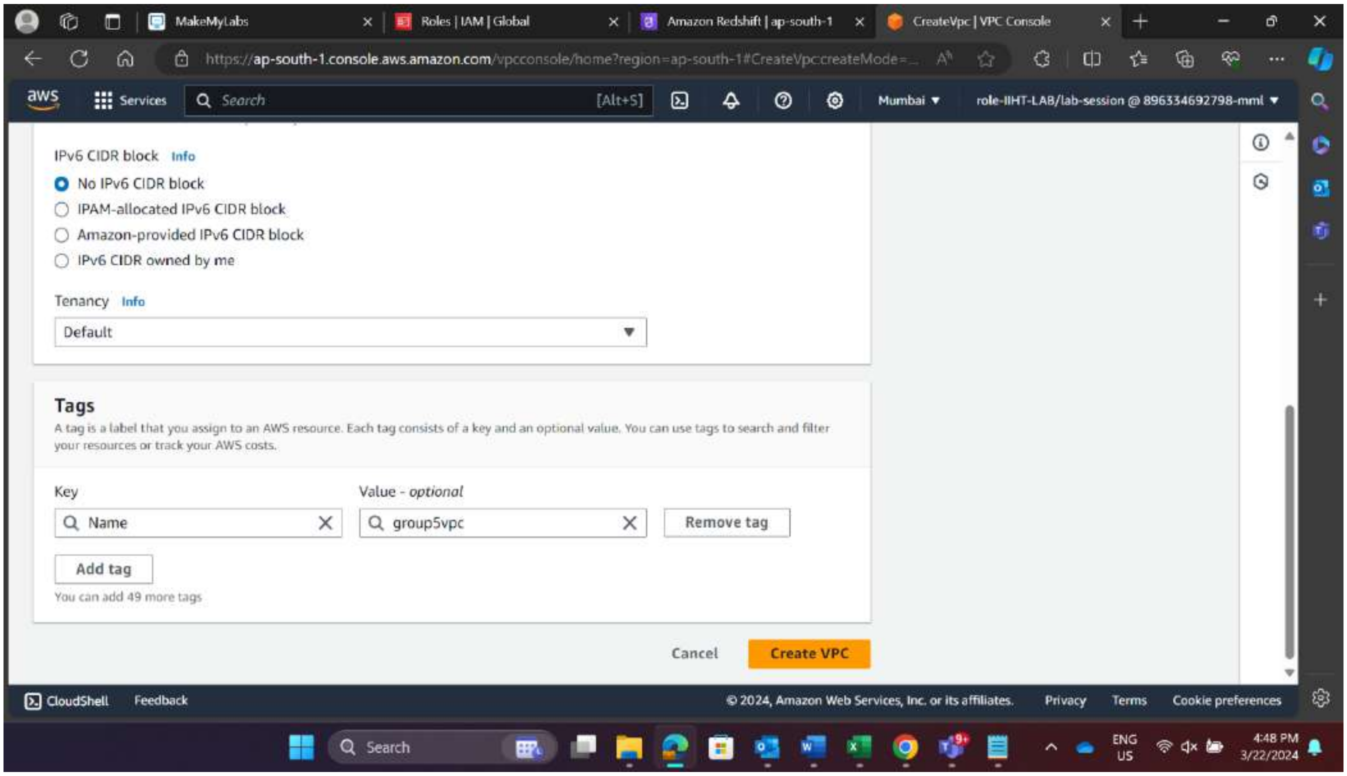
VPC settings

Resources to create Info
Create only the VPC resource or the VPC and other networking resources.
● VPC only ○ VPC and more

Name tag - *optional*
Creates a tag with a key of 'Name' and a value that you specify.
group5vpc

IPv4 CIDR block Info
● IPv4 CIDR manual input
○ IPAM-allocated IPv4 CIDR block

IPv4 CIDR
172.16.0.0/16



IPv6 CIDR block Info

● No IPv6 CIDR block
○ IPAM-allocated IPv6 CIDR block
○ Amazon-provided IPv6 CIDR block
○ IPv6 CIDR owned by me

Tenancy Info

Default

Tags

A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

Key	Value - *optional*
Name	group5vpc

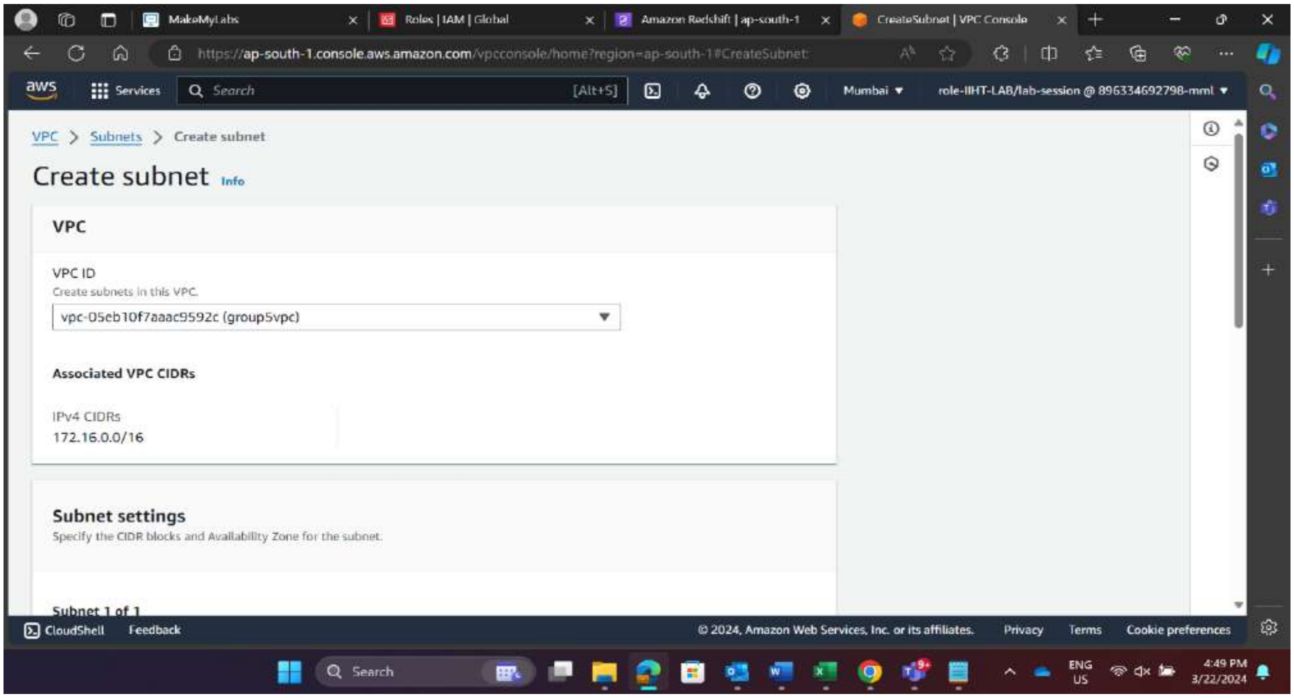
Remove tag

Add tag

You can add 49 more tags

Cancel Create VPC

Step 2: Define the subnets within your VPC and specify the CIDR block for subnet.



VPC > Subnets > Create subnet

Create subnet Info

VPC

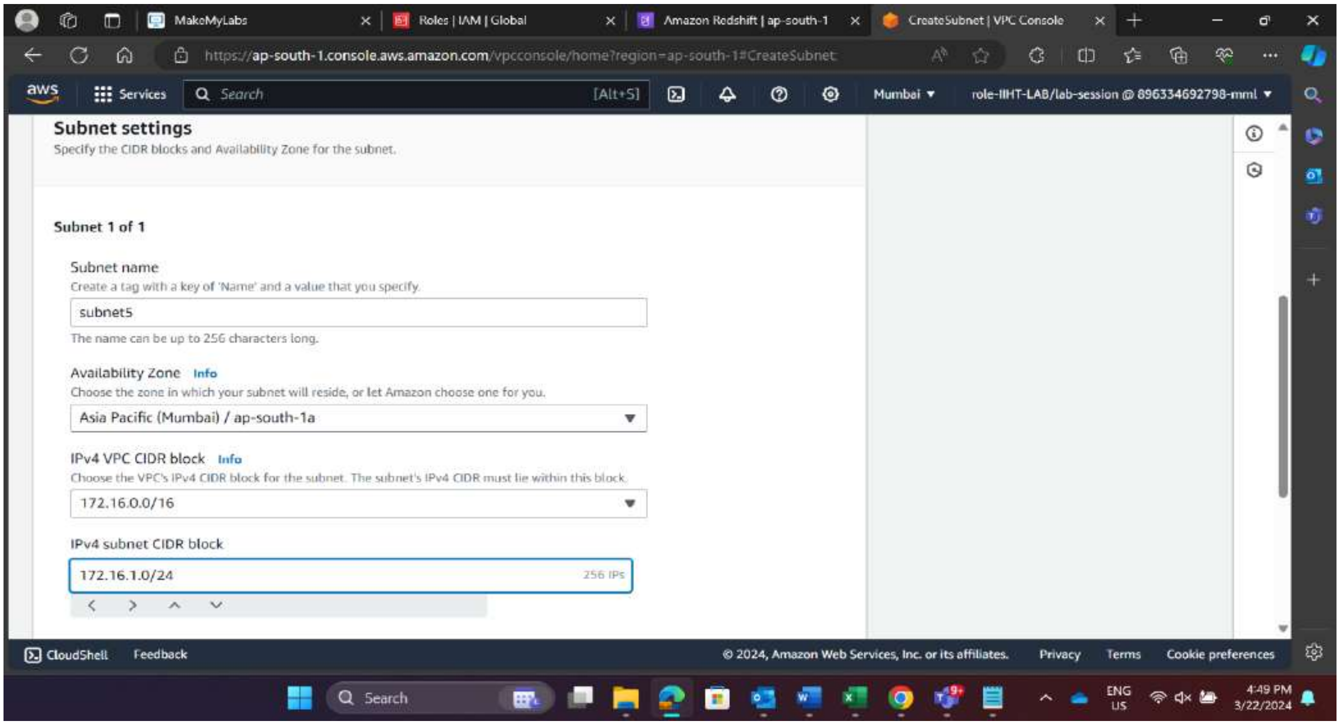
VPC ID
Create subnets in this VPC.
vpc-05eb10f7aaac9592c (group5vpc)

Associated VPC CIDRs

IPv4 CIDRs
172.16.0.0/16

Subnet settings
Specify the CIDR blocks and Availability Zone for the subnet.

Subnet 1 of 1



Subnet settings

Specify the CIDR blocks and Availability Zone for the subnet.

Subnet 1 of 1

Subnet name

Create a tag with a key of 'Name' and a value that you specify.

subnet5

The name can be up to 256 characters long.

Availability Zone Info

Choose the zone in which your subnet will reside, or let Amazon choose one for you.

Asia Pacific (Mumbai) / ap-south-1a

IPv4 VPC CIDR block Info

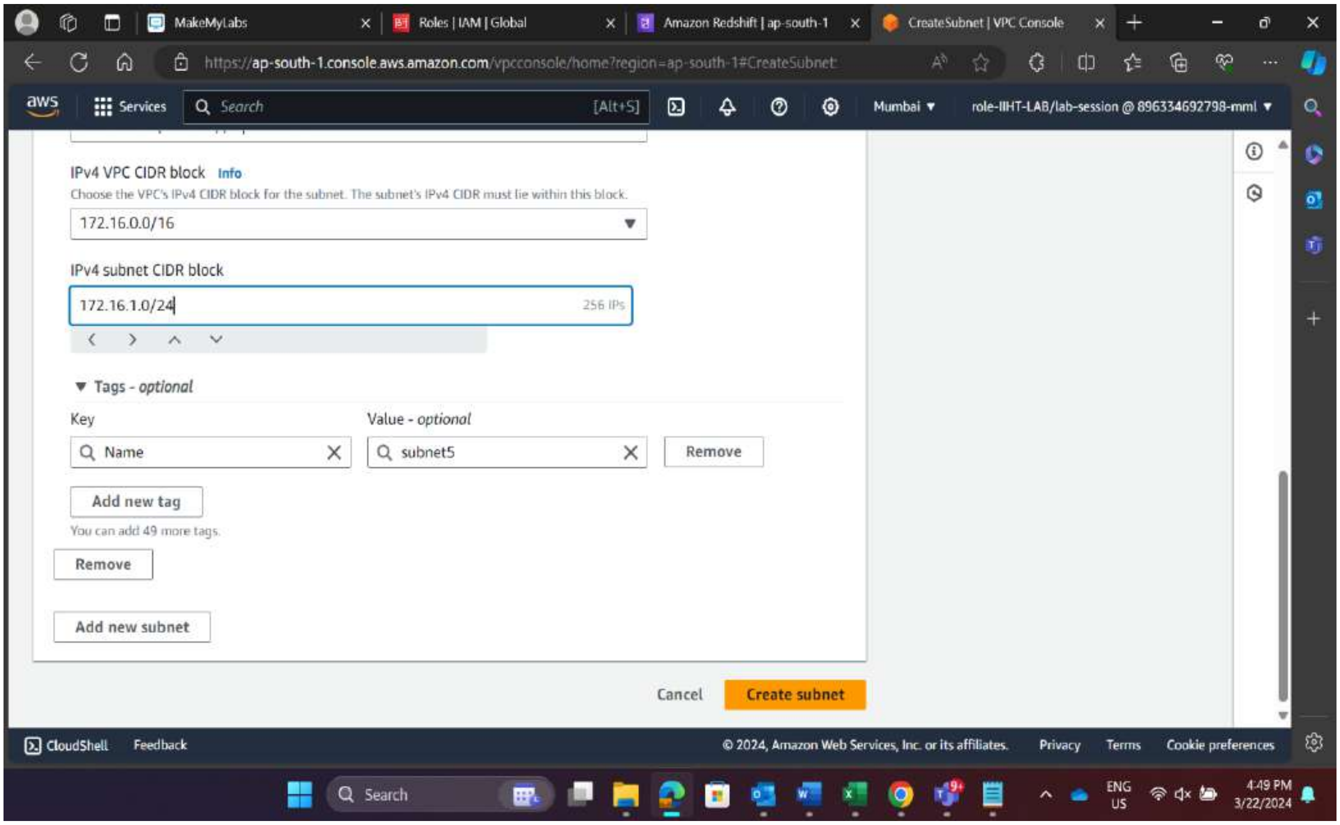
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.

172.16.0.0/16

IPv4 subnet CIDR block

172.16.1.0/24 256 IPs

MakeMyLabs
Roles | IAM | Global
Amazon Redshift | ap-south-1
CreateSubnet | VPC Console
https://ap-south-1.console.aws.amazon.com/vpcconsole/home?region=ap-south-1#CreateSubnet:
aws
Services
Search
[Alt+S]
Mumbai ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:49 PM 3/22/2024



IPv4 VPC CIDR block Info

Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.

172.16.0.0/16

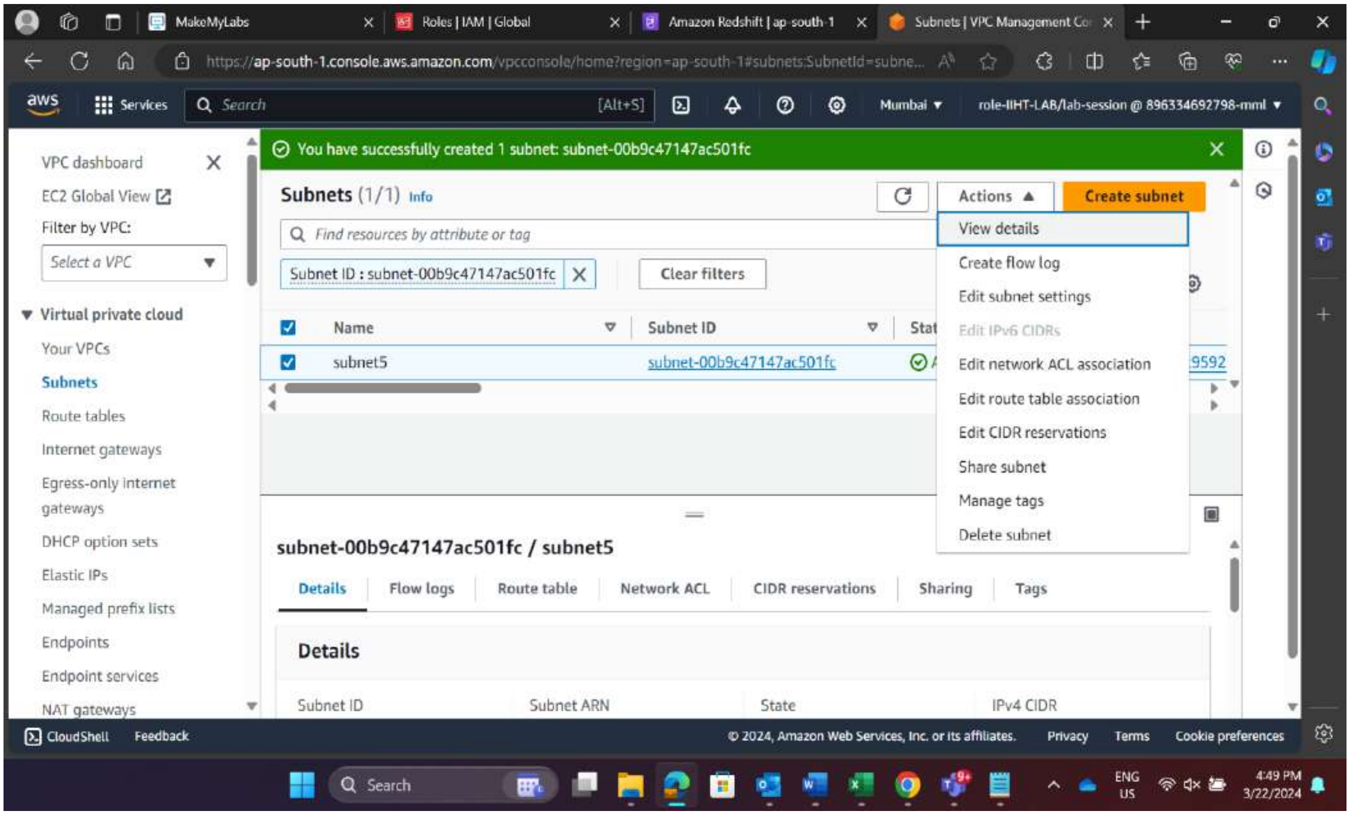
IPv4 subnet CIDR block

172.16.1.0/24 256 IPs

▼ Tags - *optional*

Key	Value - *optional*
Name	subnet5

Remove
Add new tag
You can add 49 more tags.
Remove
Add new subnet
Cancel
Create subnet
MakeMyLabs
Roles | IAM | Global
Amazon Redshift | ap-south-1
CreateSubnet | VPC Console
https://ap-south-1.console.aws.amazon.com/vpcconsole/home?region=ap-south-1#CreateSubnet:
aws
Services
Search
[Alt+S]
Mumbai ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:49 PM 3/22/2024



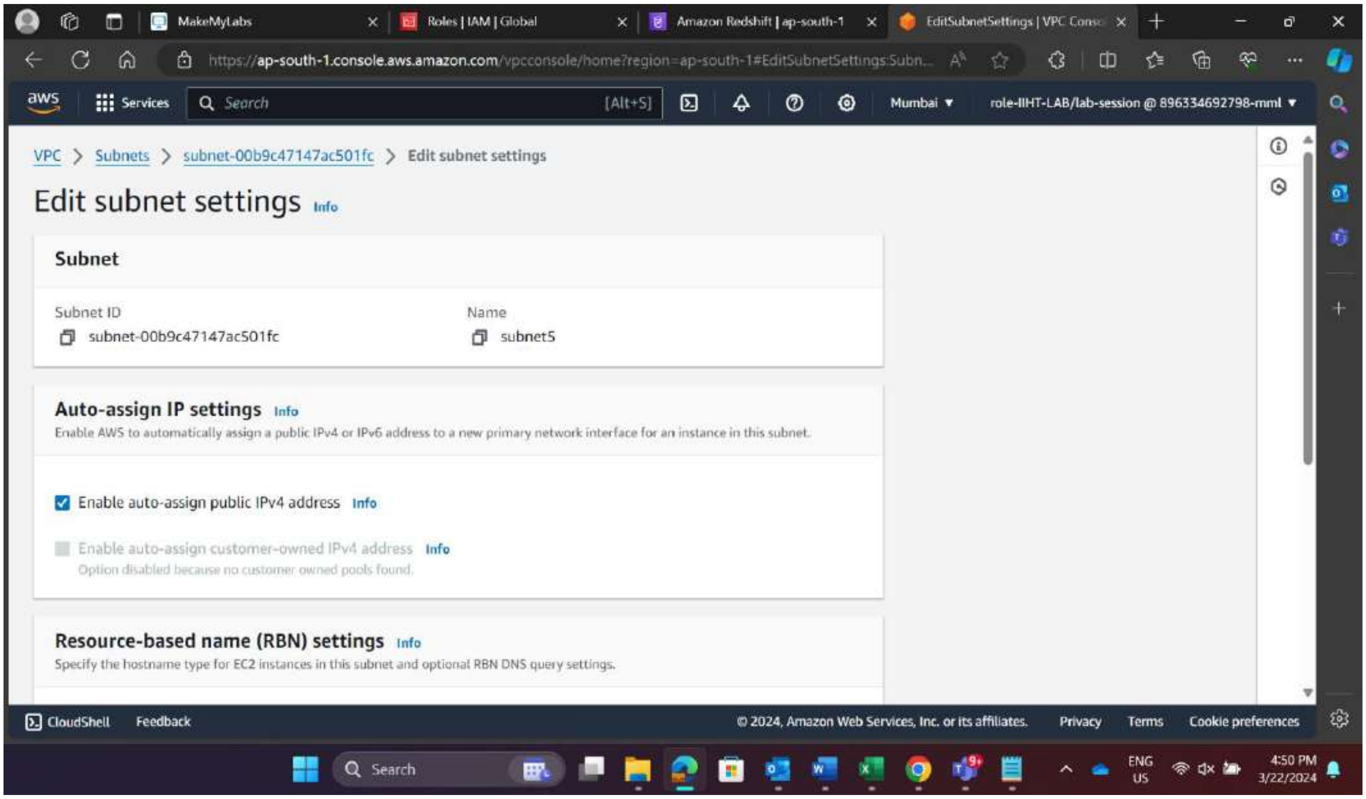
You have successfully created 1 subnet: subnet-00b9c47147ac501fc

Subnets (1/1) Info

Subnet ID : subnet-00b9c47147ac501fc

Name	Subnet ID	Stat
subnet5	subnet-00b9c47147ac501fc	A

subnet-00b9c47147ac501fc / subnet5



VPC > Subnets > subnet-00b9c47147ac501fc > Edit subnet settings

Edit subnet settings Info

Subnet

Subnet ID: subnet-00b9c47147ac501fc

Name: subnet5

Auto-assign IP settings Info

Enable AWS to automatically assign a public IPv4 or IPv6 address to a new primary network interface for an instance in this subnet.

☑ Enable auto-assign public IPv4 address Info

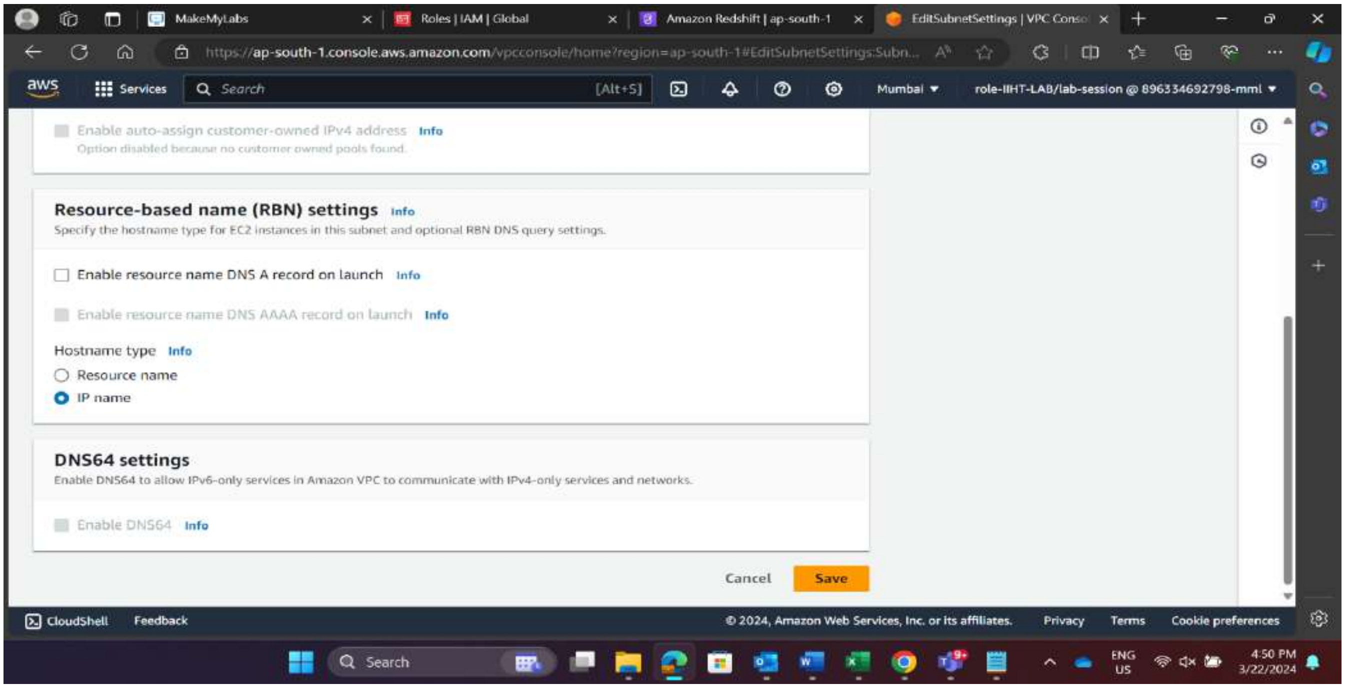
□ Enable auto-assign customer-owned IPv4 address Info

Option disabled because no customer owned pools found.

Resource-based name (RBN) settings Info

Specify the hostname type for EC2 instances in this subnet and optional RBN DNS query settings.

MakeMyLabs
Roles | IAM | Global
Amazon Redshift | ap-south-1
EditSubnetSettings | VPC Conso
https://ap-south-1.console.aws.amazon.com/vpcconsole/home?region=ap-south-1#EditSubnetSettings:Subn...
aws
Services
Search
[Alt+S]
Mumbai ▼
role-IIHT-LAB/lab-session @ 896334692798-mml ▼
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:50 PM
3/22/2024



□ Enable auto-assign customer-owned IPv4 address Info
Option disabled because no customer owned pools found.

Resource-based name (RBN) settings Info
Specify the hostname type for EC2 instances in this subnet and optional RBN DNS query settings.

□ Enable resource name DNS A record on launch Info

□ Enable resource name DNS AAAA record on launch Info

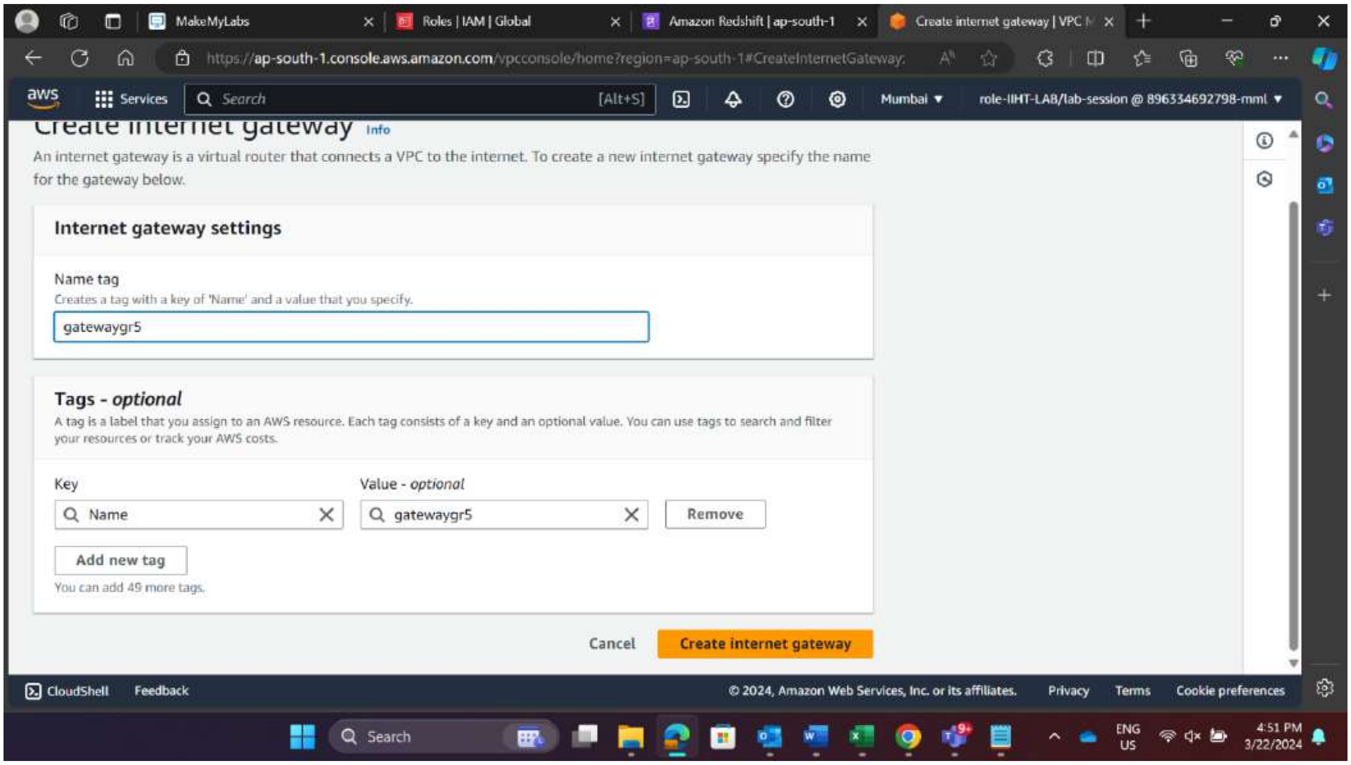
Hostname type Info
○ Resource name
● IP name

DNS64 settings
Enable DNS64 to allow IPv6-only services in Amazon VPC to communicate with IPv4-only services and networks.

□ Enable DNS64 Info

Cancel Save

Step 3: Configure Internet Gateway: create a new internet gateway, give it a name, and attach it to your VPC.



Create internet gateway Info
An internet gateway is a virtual router that connects a VPC to the internet. To create a new internet gateway specify the name for the gateway below.

Internet gateway settings

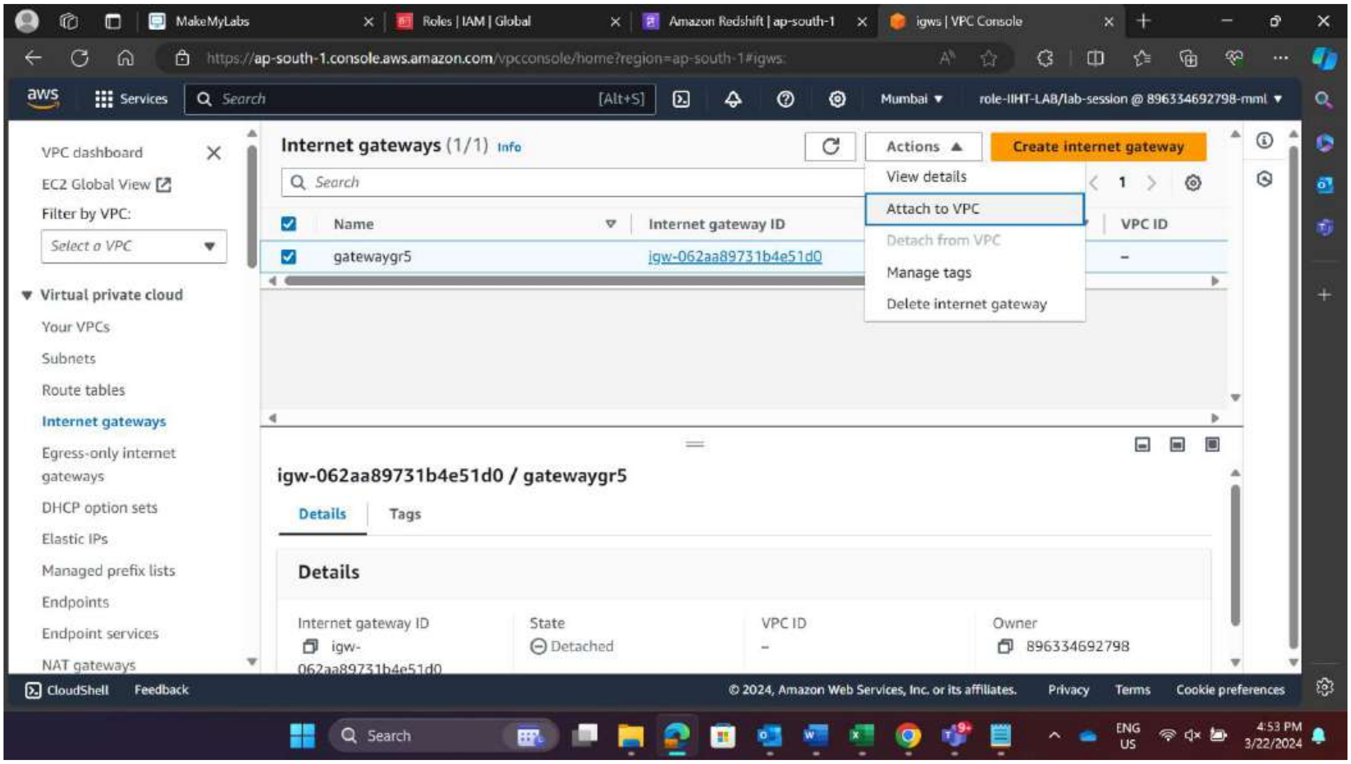
Name tag
Creates a tag with a key of 'Name' and a value that you specify.
gatewaygr5

Tags - *optional*
A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

Key	Value - *optional*	
Name	gatewaygr5	Remove

Add new tag
You can add 49 more tags.

Cancel Create internet gateway

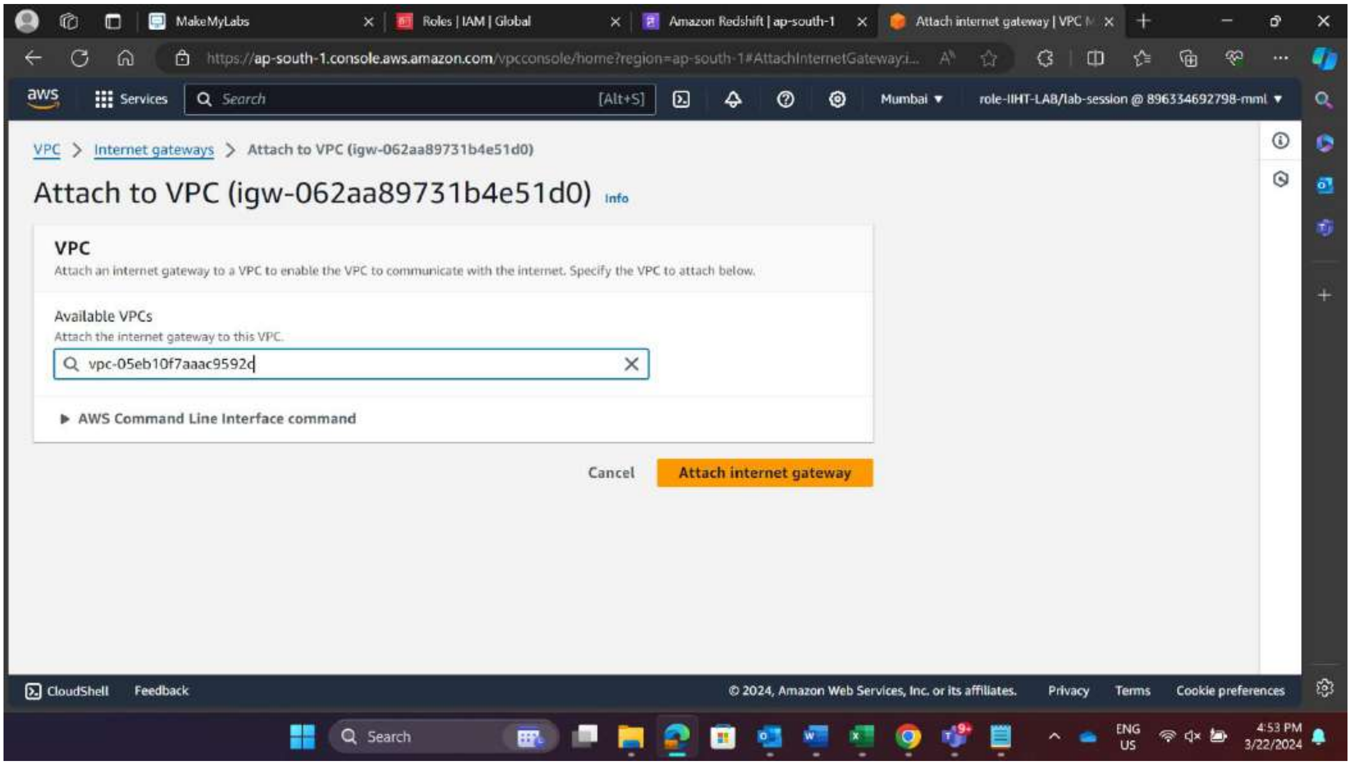


Name	Internet gateway ID	VPC ID
gatewaygr5	igw-062aa89731b4e51d0	–

igw-062aa89731b4e51d0 / gatewaygr5

Details

Internet gateway ID	State	VPC ID	Owner
igw-062aa89731b4e51d0	Detached	–	896334692798



VPC > Internet gateways > Attach to VPC (igw-062aa89731b4e51d0)

Attach to VPC (igw-062aa89731b4e51d0)

VPC

Attach an internet gateway to a VPC to enable the VPC to communicate with the internet. Specify the VPC to attach below.

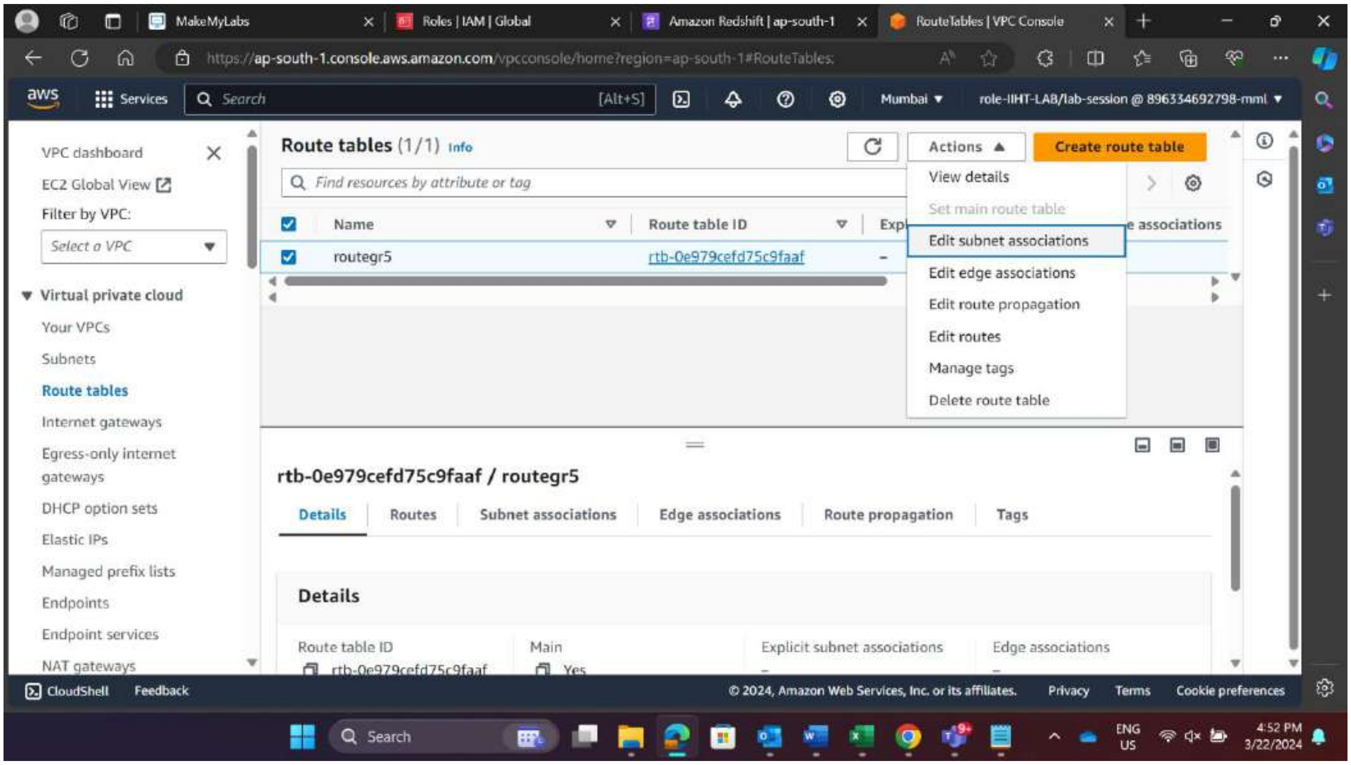
Available VPCs

Attach the internet gateway to this VPC.

vpc-05eb10f7aaac9592c

AWS Command Line Interface command
Cancel
Attach internet gateway

Step 4: Configure Route Table: create a route table, give it a name, and attach it to internet gateway.



Route tables (1/1) Info

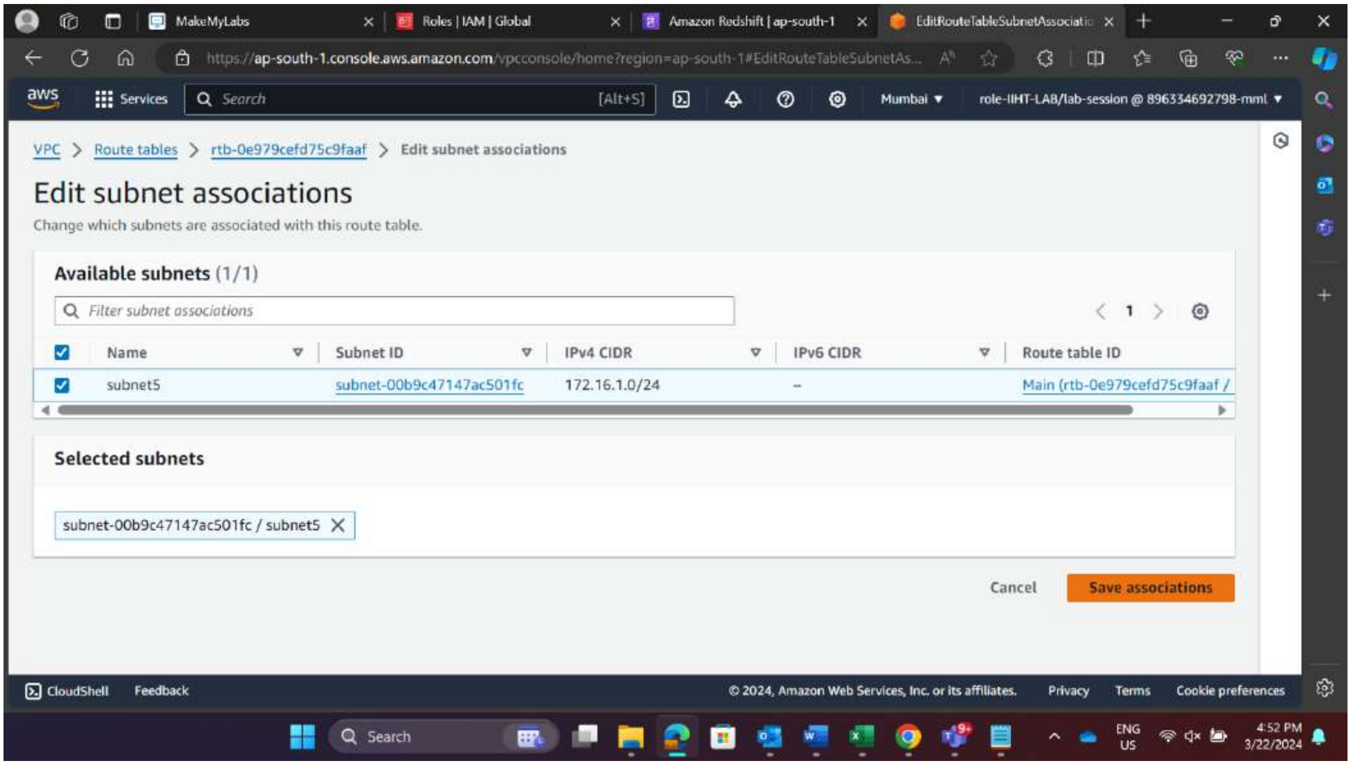
Find resources by attribute or tag

☑	Name	Route table ID	Expl...	...e associations
☑	routegr5	rtb-0e979cefd75c9faaf	–	

rtb-0e979cefd75c9faaf / routegr5

Details

Route table ID	Main	Explicit subnet associations	Edge associations
rtb-0e979cefd75c9faaf	Yes	–	–



Edit subnet associations

Change which subnets are associated with this route table.

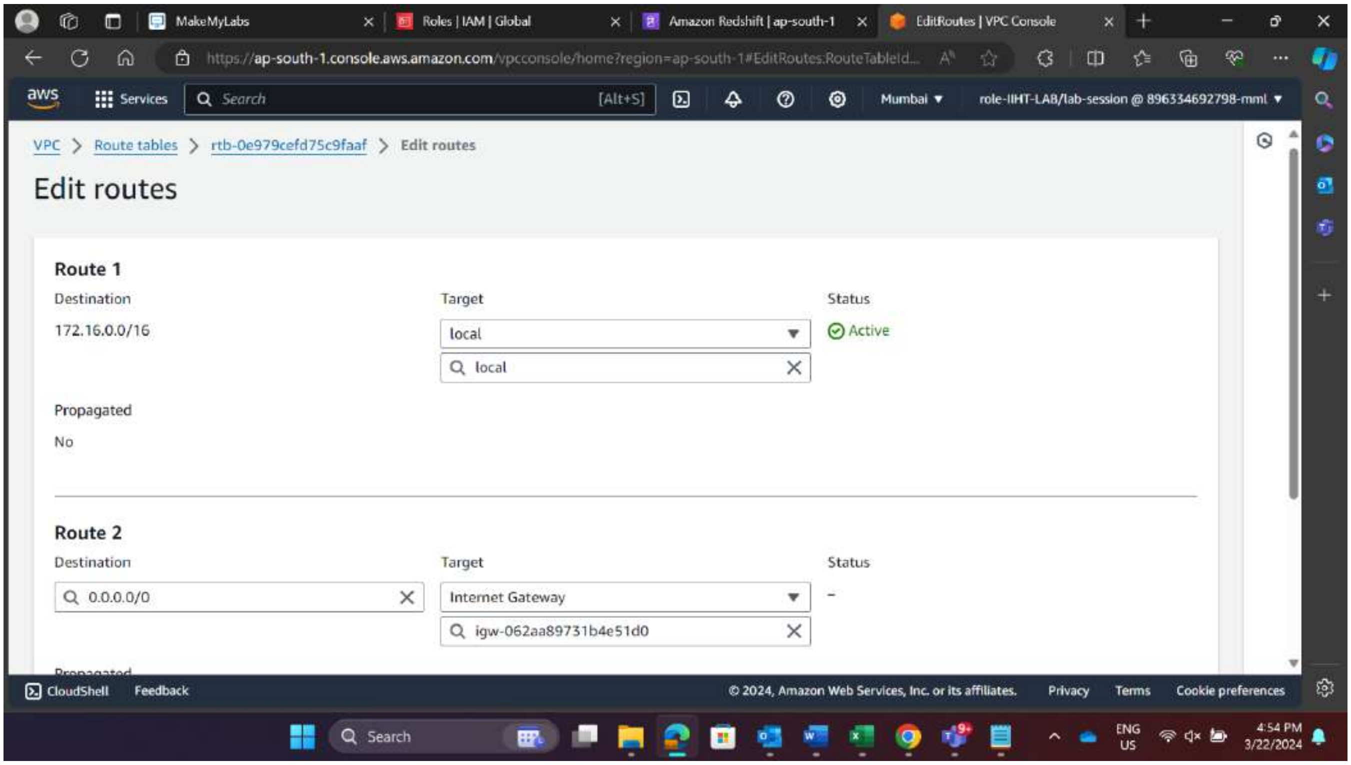
Available subnets (1/1)

Filter subnet associations

☑	Name	Subnet ID	IPv4 CIDR	IPv6 CIDR	Route table ID
☑	subnet5	subnet-00b9c47147ac501fc	172.16.1.0/24	–	Main (rtb-0e979cefd75c9faaf /

Selected subnets

subnet-00b9c47147ac501fc / subnet5



VPC > Route tables > rtb-0e979cefd75c9faaf > Edit routes

Edit routes

Route 1

Destination: 172.16.0.0/16

Target: local

local

Status: Active

Propagated: No

Route 2

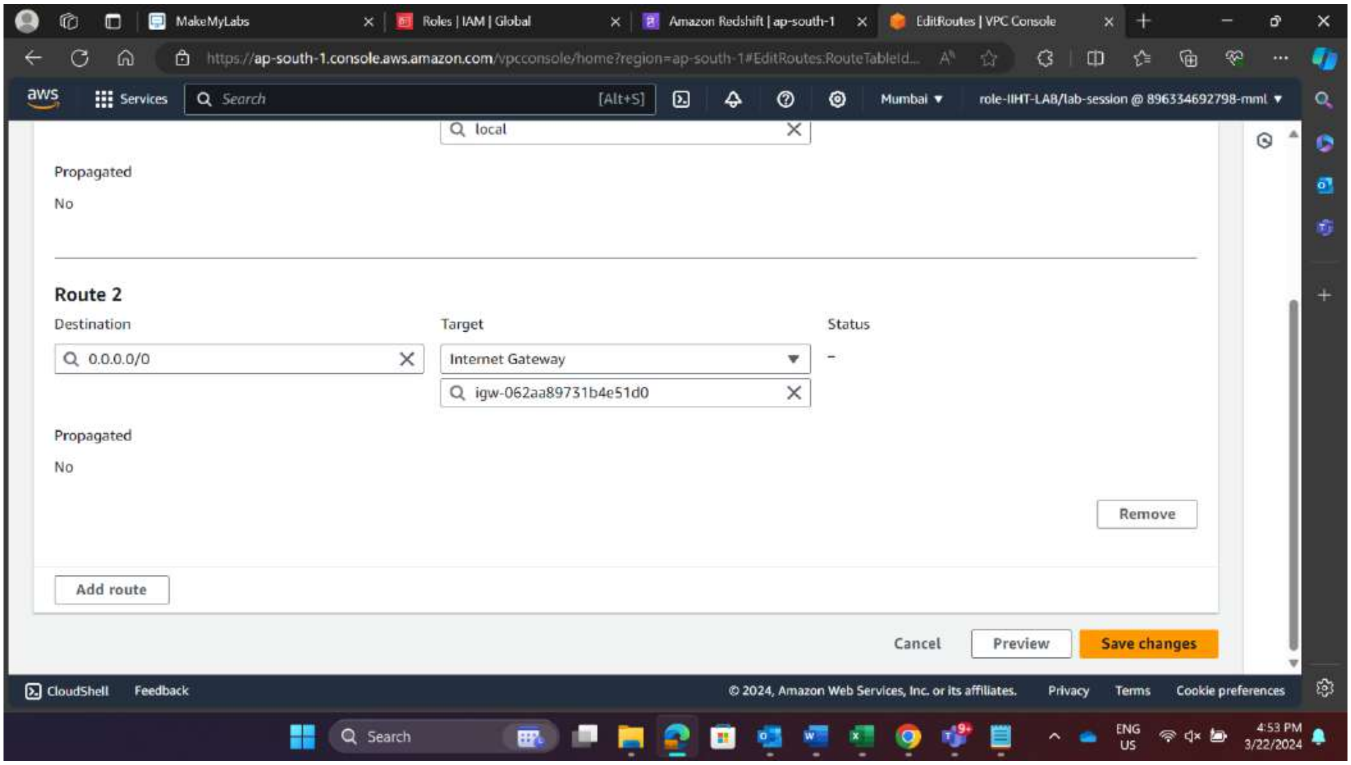
Destination: 0.0.0.0/0

Target: Internet Gateway

igw-062aa89731b4e51d0

Status: -

Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
4:54 PM
3/22/2024



local

Propagated: No

Route 2

Destination: 0.0.0.0/0

Target: Internet Gateway

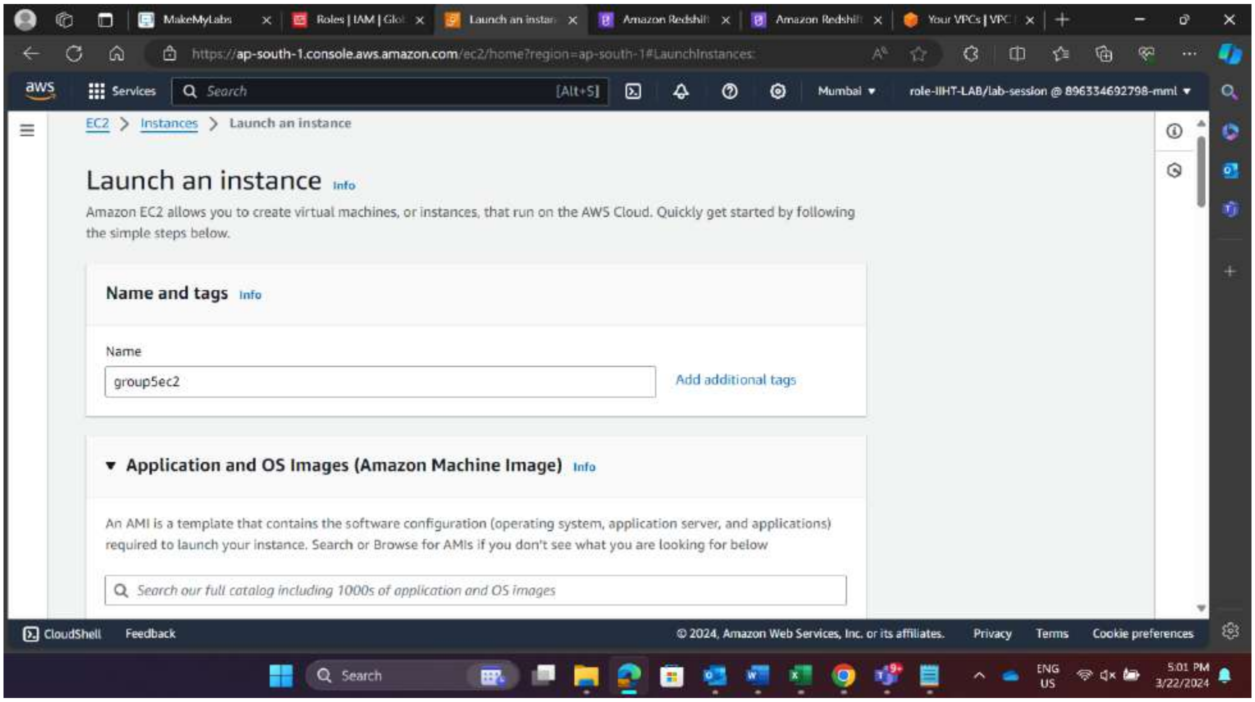
igw-062aa89731b4e51d0

Status: -

Propagated: No

EC2 INSTANCE

Step 1: On the EC2 dashboard, click the **"Launch Instance"** button. This will start the instance creation wizard.



Launch an instance Info

Amazon EC2 allows you to create virtual machines, or instances, that run on the AWS Cloud. Quickly get started by following the simple steps below.

Name and tags Info

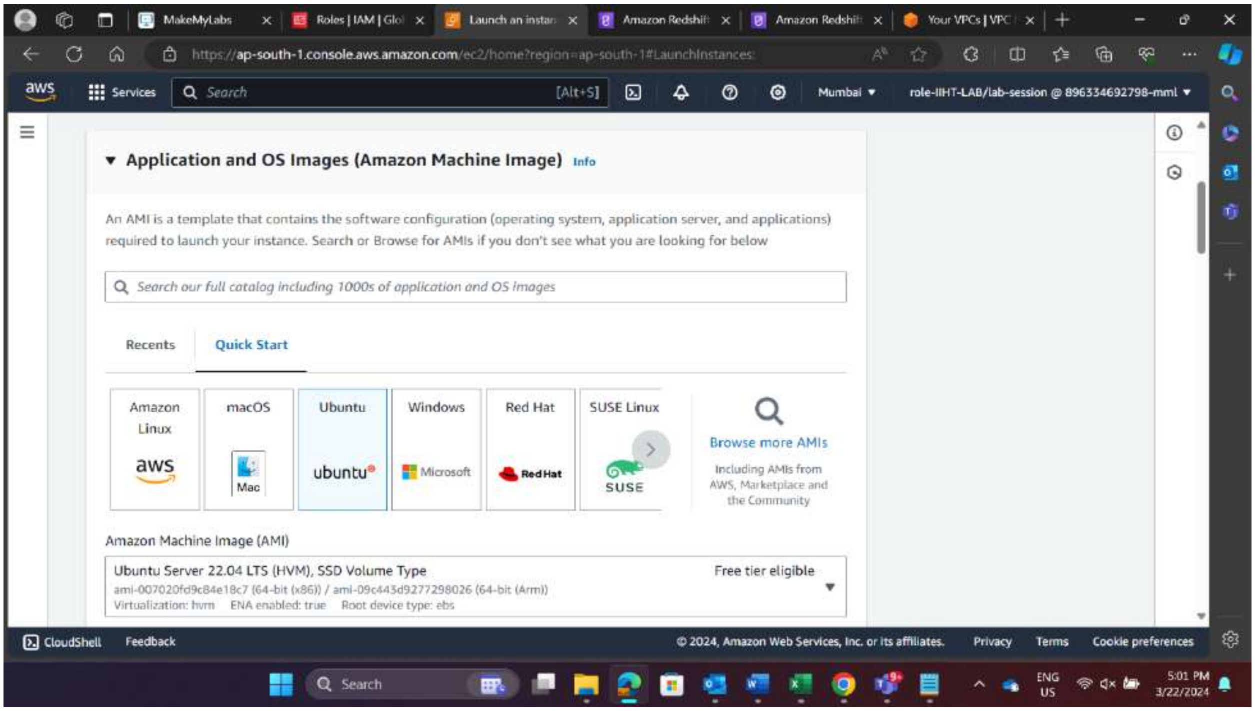
Name: group5ec2

Application and OS Images (Amazon Machine Image) Info

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below

Search our full catalog including 1000s of application and OS images
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Step 2: Choose an Amazon Machine Image (AMI). It is a template for instance's operating system.



Application and OS Images (Amazon Machine Image) Info

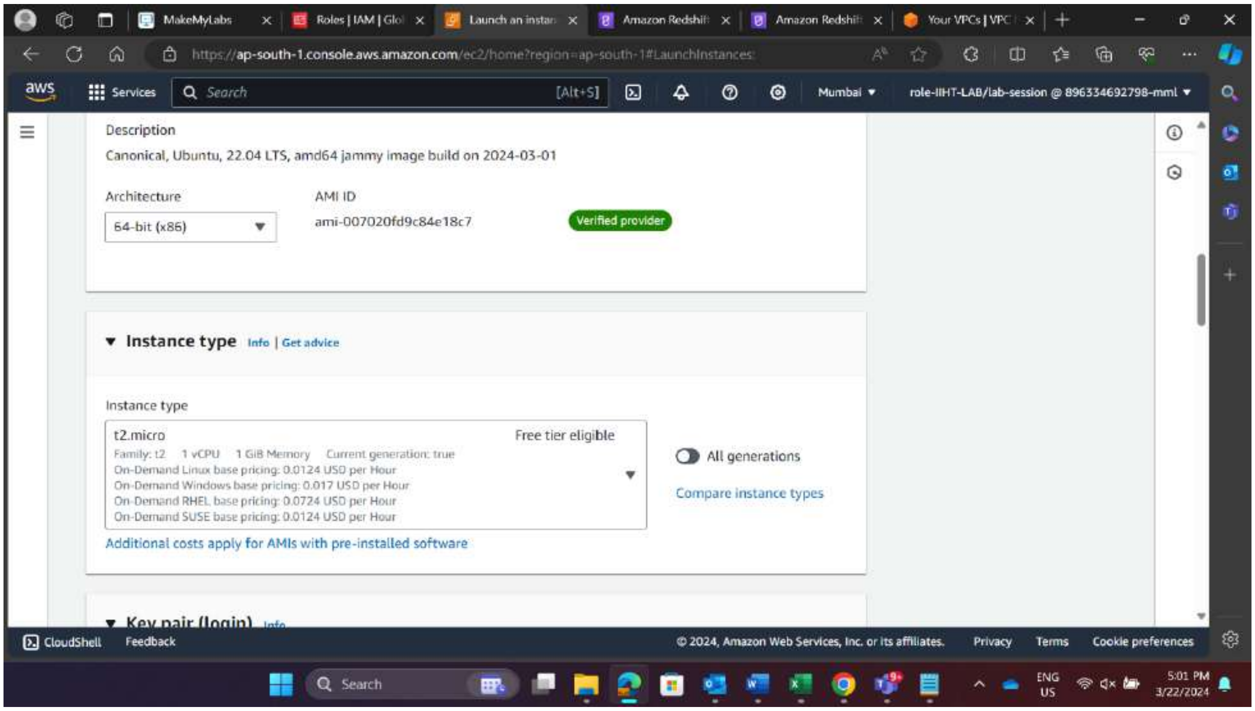
An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below

Search our full catalog including 1000s of application and OS images
Recents
Quick Start
Amazon Linux
macOS
Ubuntu
Windows
Red Hat
SUSE Linux
Browse more AMIs
Including AMIs from AWS, Marketplace and the Community

Amazon Machine Image (AMI)

Ubuntu Server 22.04 LTS (HVM), SSD Volume Type — Free tier eligible
ami-007020fd9c84e18c7 (64-bit (x86)) / ami-09c443d9277298026 (64-bit (Arm))
Virtualization: hvm ENA enabled: true Root device type: ebs

Step 3: In this step, select the type of instance based on your business requirements.



Description
Canonical, Ubuntu, 22.04 LTS, amd64 jammy image build on 2024-03-01

Architecture: 64-bit (x86)
AMI ID: ami-007020fd9c84e18c7 — Verified provider

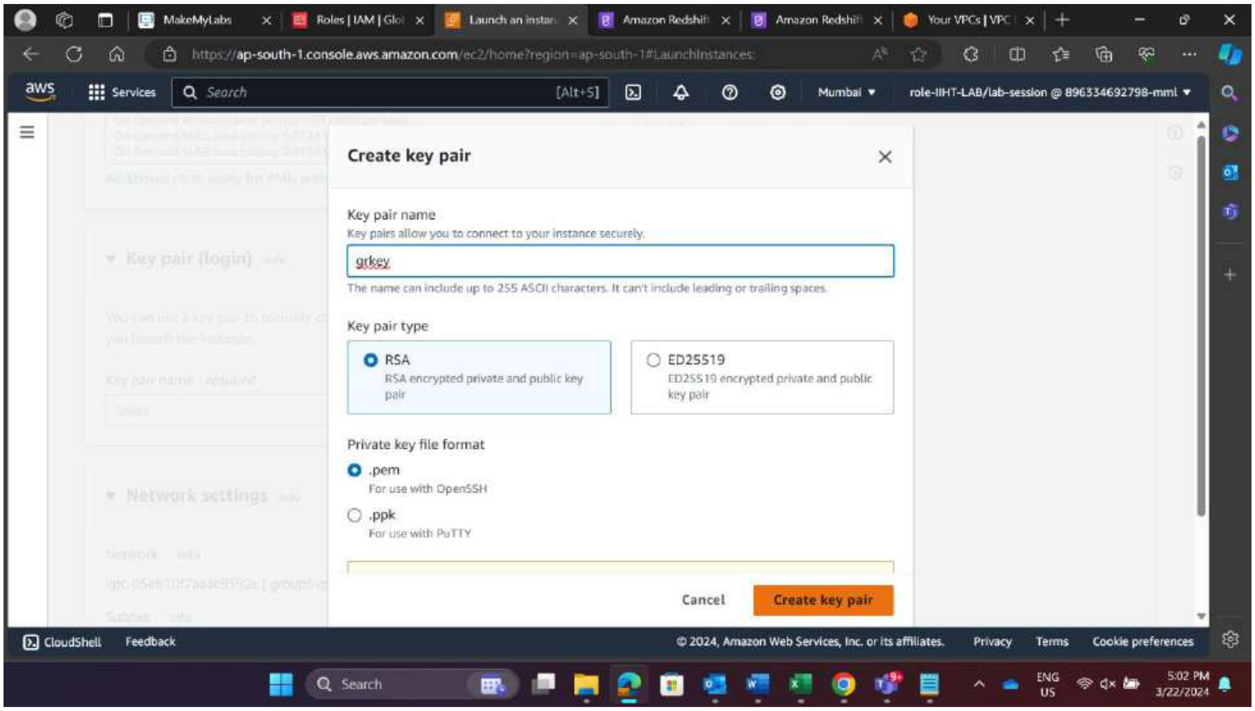
Instance type

t2.micro — Free tier eligible
Family: t2 1 vCPU 1 GiB Memory Current generation: true
On-Demand Linux base pricing: 0.0124 USD per Hour
On-Demand Windows base pricing: 0.017 USD per Hour
On-Demand RHEL base pricing: 0.0724 USD per Hour
On-Demand SUSE base pricing: 0.0124 USD per Hour

Additional costs apply for AMIs with pre-installed software

https://ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LaunchInstances:
Services
Search
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
Info | Get advice
All generations
Compare instance types
Key pair (login)
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
5:01 PM
3/22/2024

Step 4: If you don't have an existing key pair, create a new one. Download the private key file (.pem) and keep it secure.



Create key pair

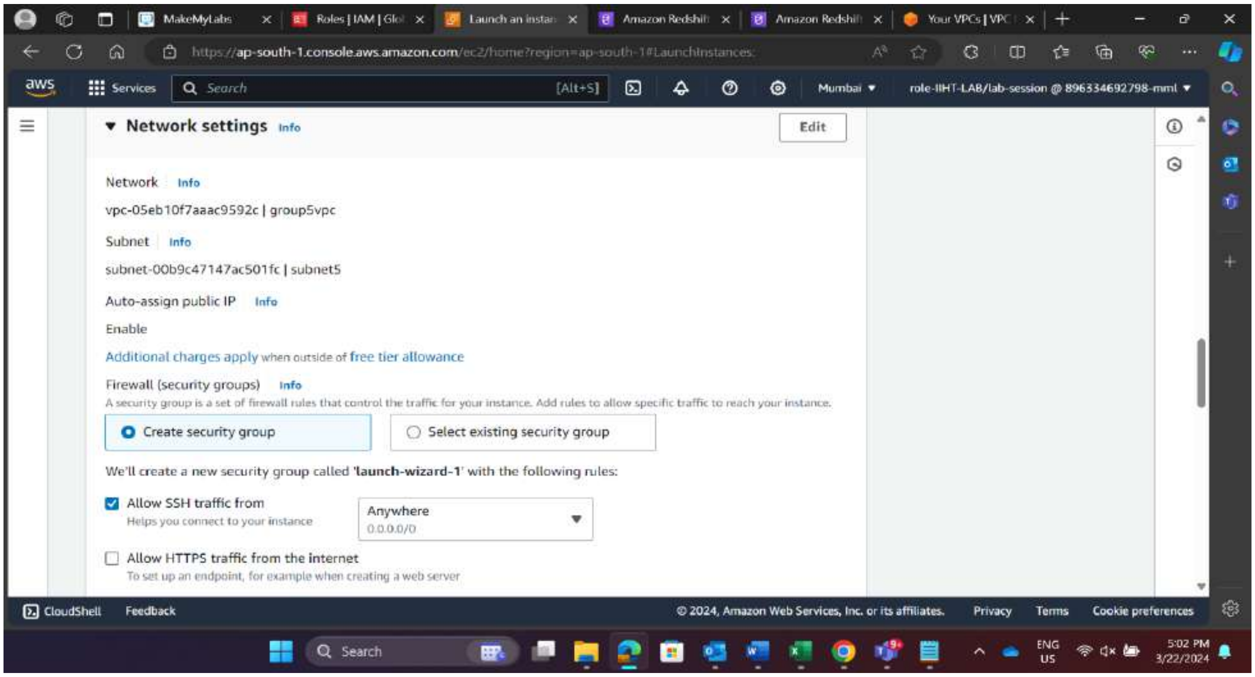
Key pair name
Key pairs allow you to connect to your instance securely.
grkey
The name can include up to 255 ASCII characters. It can't include leading or trailing spaces.

Key pair type
● RSA — RSA encrypted private and public key pair
○ ED25519 — ED25519 encrypted private and public key pair

Private key file format
● .pem — For use with OpenSSH
○ .ppk — For use with PuTTY

Cancel
Create key pair
https://ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LaunchInstances:
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
5:02 PM
3/22/2024

Step 5: Set up networking details like VPC, subnet, and security groups.



▾ Network settings Info

Network Info
vpc-05eb10f7aaac9592c | group5vpc

Subnet Info
subnet-00b9c47147ac501fc | subnet5

Auto-assign public IP Info
Enable

Additional charges apply when outside of free tier allowance

Firewall (security groups) Info
A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.

● Create security group ○ Select existing security group

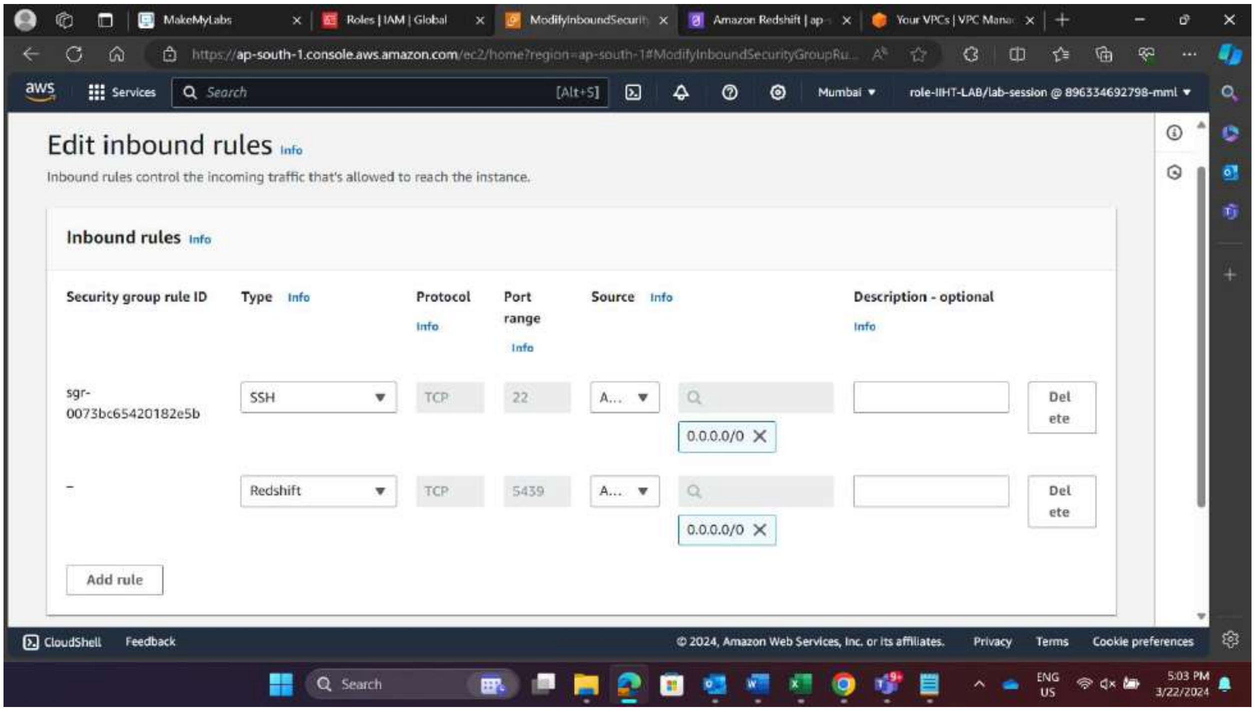
We'll create a new security group called **'launch-wizard-1'** with the following rules:

☑ Allow SSH traffic from — Anywhere 0.0.0.0/0
Helps you connect to your instance

□ Allow HTTPS traffic from the internet
To set up an endpoint, for example when creating a web server

https://ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LaunchInstances:
Edit
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
5:02 PM
3/22/2024

Step 6: Add rules for SSH and Redshift in inbound rule for allowing the traffic to redshift cluster.



Edit inbound rules Info

Inbound rules control the incoming traffic that's allowed to reach the instance.

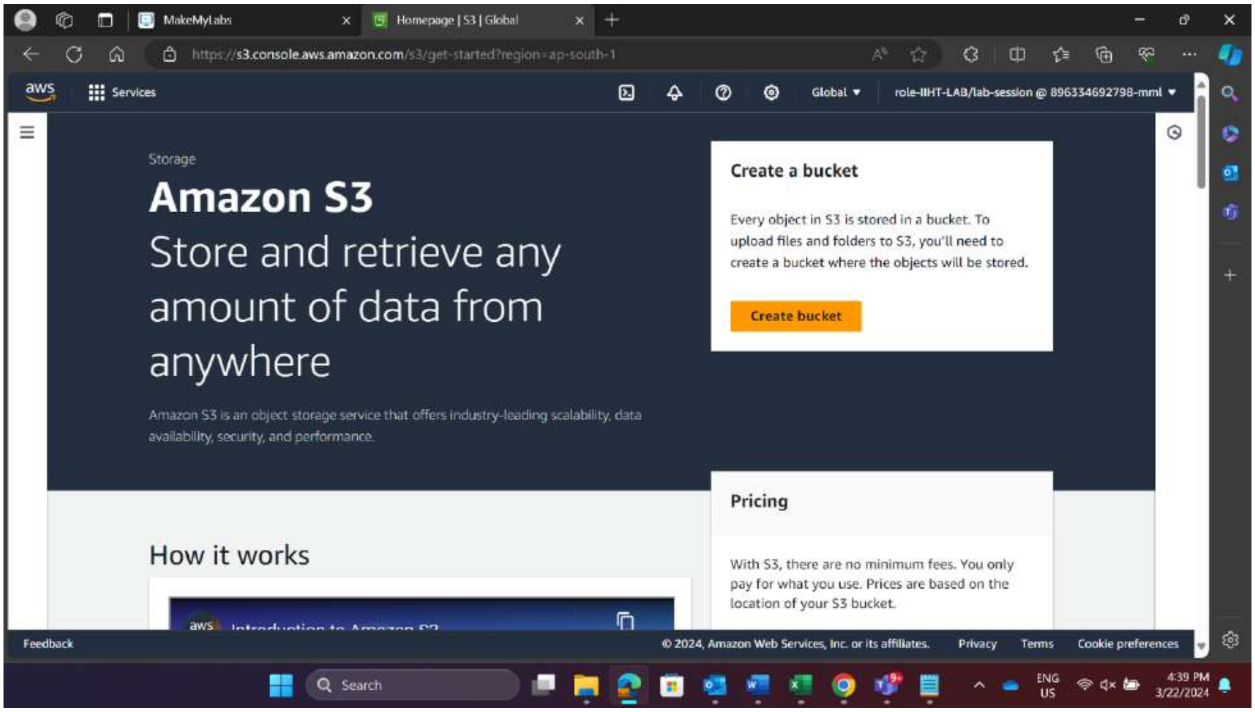
Inbound rules Info

Security group rule ID	Type	Protocol	Port range	Source		Description - optional
sgr-0073bc65420182e5b	SSH	TCP	22	A...	0.0.0.0/0	
-	Redshift	TCP	5439	A...	0.0.0.0/0	

https://ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#ModifyInboundSecurityGroupRu...
Delete
Add rule
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
5:03 PM
3/22/2024

S3 BUCKET

Step 1: Create bucket by giving unique name and select the AWS region where you want to create bucket. Configure bucket permissions whether to keep the bucket private or public. After creating bucket add objects.


MakeMyLabs
Homepage | S3 | Global
https://s3.console.aws.amazon.com/s3/get-started?region=ap-south-1
aws
Services
Global ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:39 PM 3/22/2024

Storage

Amazon S3

Store and retrieve any amount of data from anywhere

Amazon S3 is an object storage service that offers industry-leading scalability, data availability, security, and performance.

Create a bucket

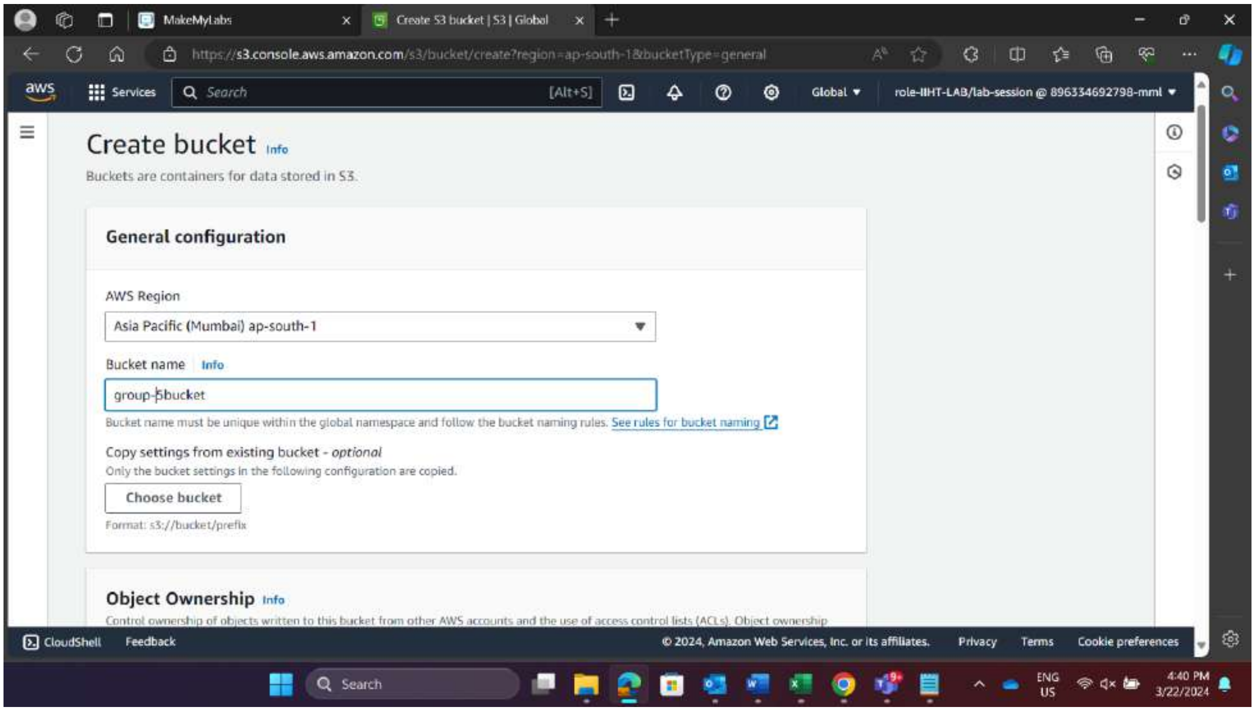
Every object in S3 is stored in a bucket. To upload files and folders to S3, you'll need to create a bucket where the objects will be stored.

Create bucket

Pricing

With S3, there are no minimum fees. You only pay for what you use. Prices are based on the location of your S3 bucket.

How it works


MakeMyLabs
Create S3 bucket | S3 | Global
https://s3.console.aws.amazon.com/s3/bucket/create?region=ap-south-1&bucketType=general
aws
Services
Search
[Alt+S]
Global ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:40 PM 3/22/2024

Create bucket Info

Buckets are containers for data stored in S3.

General configuration

AWS Region: Asia Pacific (Mumbai) ap-south-1

Bucket name Info: group-5bucket

Bucket name must be unique within the global namespace and follow the bucket naming rules. See rules for bucket naming

Copy settings from existing bucket - *optional*

Only the bucket settings in the following configuration are copied.

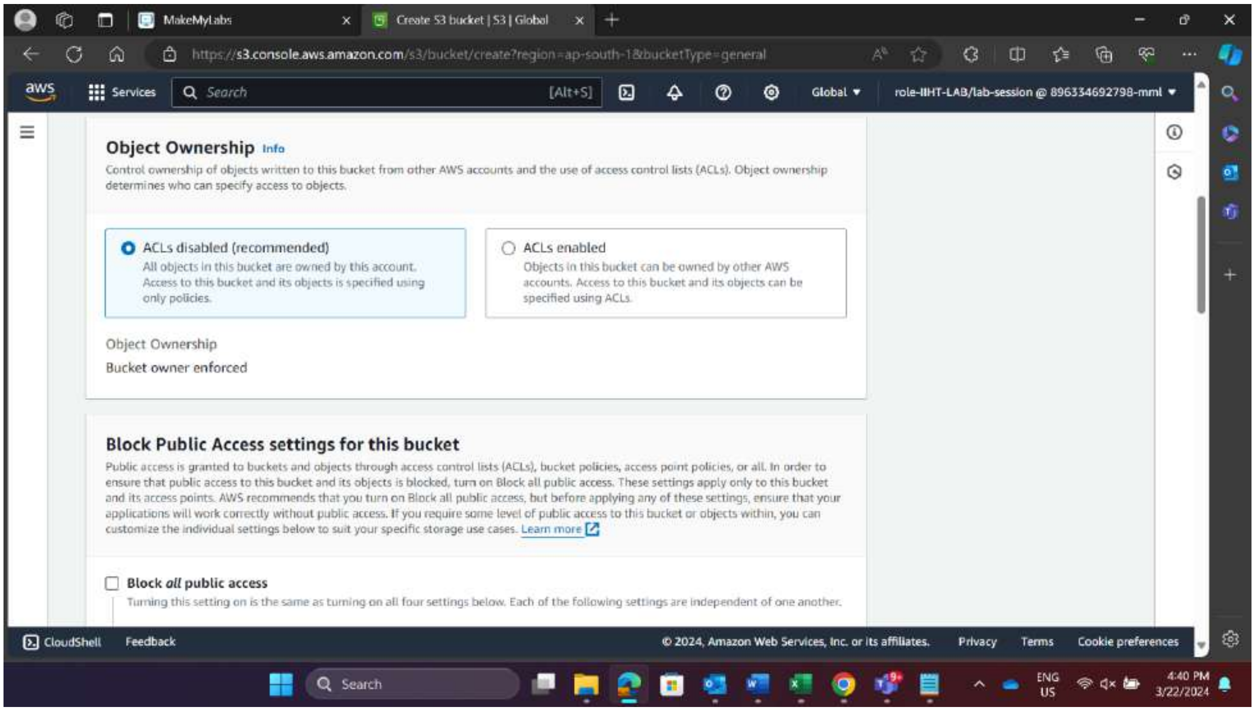
Choose bucket

Format: s3://bucket/prefix

Object Ownership Info

Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership

Step 2: Configure whether **ACLs (Access Control Lists) are enabled** for the bucket. You can make the bucket accessible to all or restrict access based on your requirements.


MakeMyLabs
Create S3 bucket | S3 | Global
https://s3.console.aws.amazon.com/s3/bucket/create?region=ap-south-1&bucketType=general
aws
Services
Search [Alt+S]
Global
role-IIHT-LAB/lab-session @ 896334692798-mml
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:40 PM 3/22/2024

Object Ownership Info

Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.

● ACLs disabled (recommended)
All objects in this bucket are owned by this account. Access to this bucket and its objects is specified using only policies.

○ ACLs enabled
Objects in this bucket can be owned by other AWS accounts. Access to this bucket and its objects can be specified using ACLs.

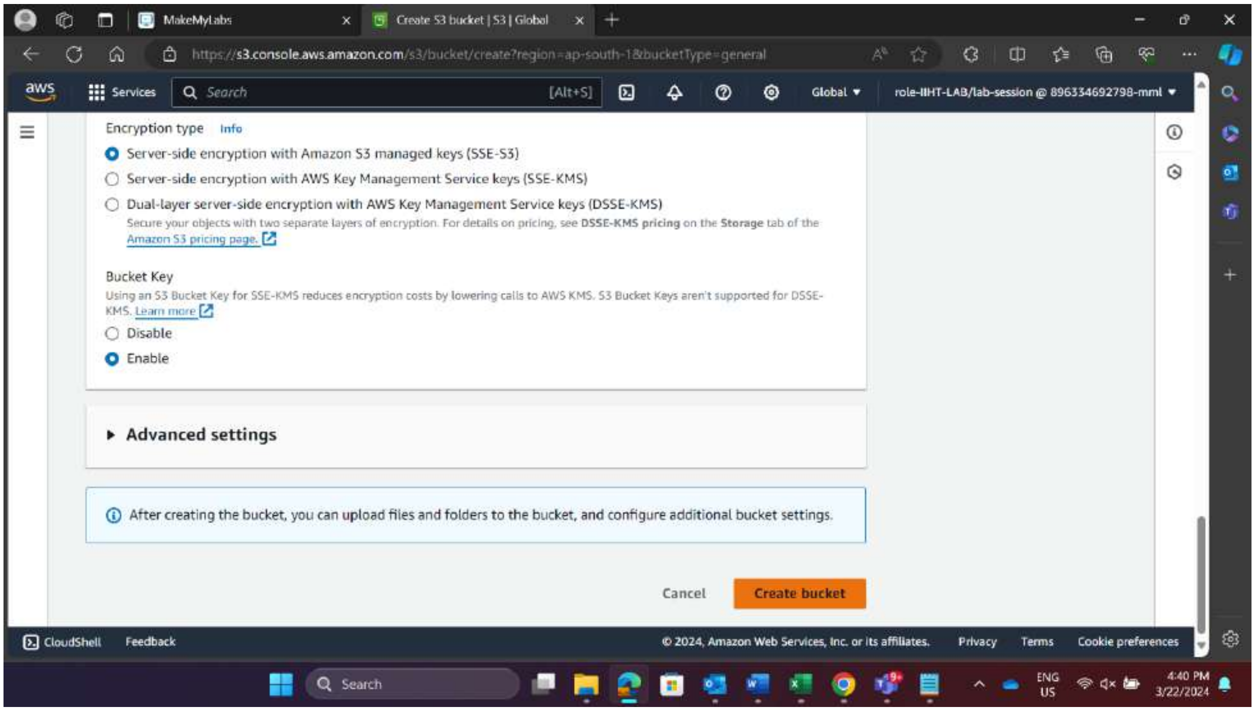
Object Ownership
Bucket owner enforced

Block Public Access settings for this bucket

Public access is granted to buckets and objects through access control lists (ACLs), bucket policies, access point policies, or all. In order to ensure that public access to this bucket and its objects is blocked, turn on Block all public access. These settings apply only to this bucket and its access points. AWS recommends that you turn on Block all public access, but before applying any of these settings, ensure that your applications will work correctly without public access. If you require some level of public access to this bucket or objects within, you can customize the individual settings below to suit your specific storage use cases. Learn more

□ **Block *all* public access**
Turning this setting on is the same as turning on all four settings below. Each of the following settings are independent of one another.

Step 3: Selecting SSE with Amazon S3 managed keys and Enabling bucket key.


MakeMyLabs
Create S3 bucket | S3 | Global
https://s3.console.aws.amazon.com/s3/bucket/create?region=ap-south-1&bucketType=general
aws
Services
Search [Alt+S]
Global
role-IIHT-LAB/lab-session @ 896334692798-mml
Cancel
Create bucket
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:40 PM 3/22/2024

Encryption type Info

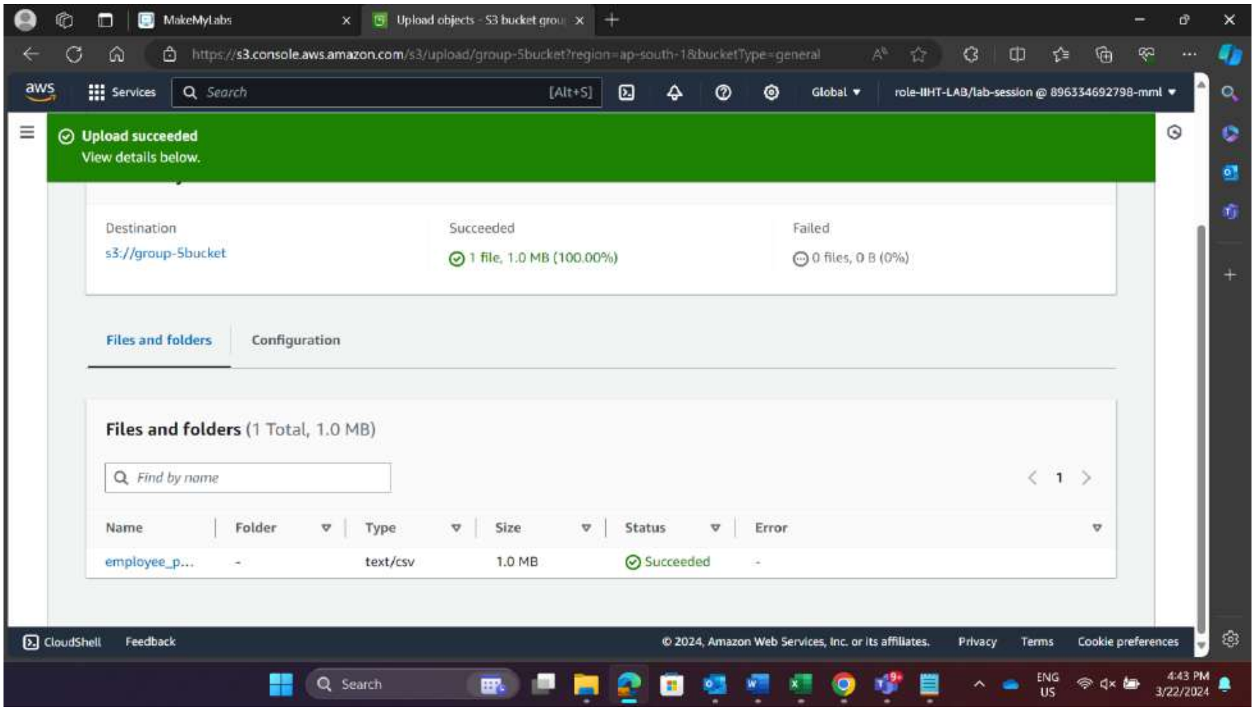
● Server-side encryption with Amazon S3 managed keys (SSE-S3)
○ Server-side encryption with AWS Key Management Service keys (SSE-KMS)
○ Dual-layer server-side encryption with AWS Key Management Service keys (DSSE-KMS)
Secure your objects with two separate layers of encryption. For details on pricing, see **DSSE-KMS pricing** on the **Storage** tab of the Amazon S3 pricing page.

Bucket Key
Using an S3 Bucket Key for SSE-KMS reduces encryption costs by lowering calls to AWS KMS. S3 Bucket Keys aren't supported for DSSE-KMS. Learn more
○ Disable
● Enable

▸ **Advanced settings**

After creating the bucket, you can upload files and folders to the bucket, and configure additional bucket settings.

Step 4: After creating bucket, upload the file in bucket.



Upload succeeded
View details below.

Destination
s3://group-5bucket

Succeeded
1 file, 1.0 MB (100.00%)

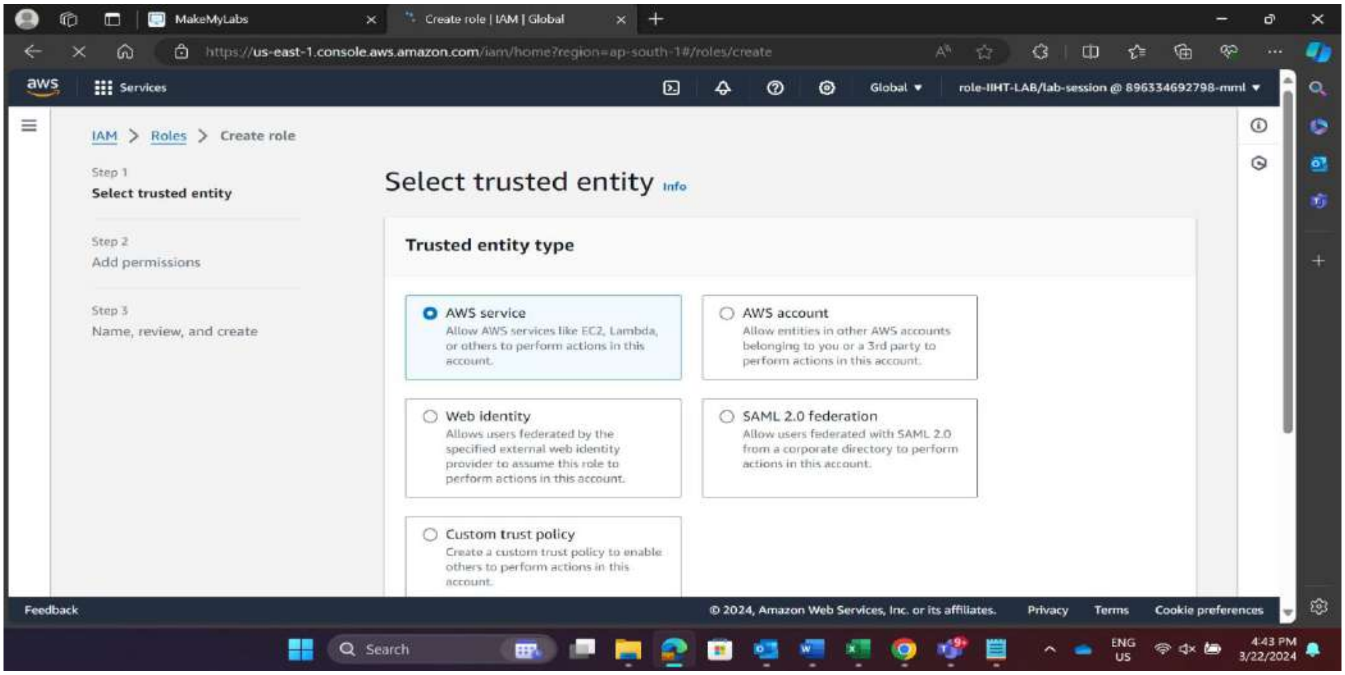
Failed
0 files, 0 B (0%)

Files and folders (1 Total, 1.0 MB)

Name	Folder	Type	Size	Status	Error
employee_p...	-	text/csv	1.0 MB	Succeeded	-

IAM ROLE ASSOCIATION

Step 1: Navigate the IAM console and create a role, choose redshift that will assume the role and then choose S3 as the service that will use this role.


MakeMyLabs
Create role | IAM | Global
https://us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles/create
aws
Services
Global
role-IIHT-LAB/lab-session @ 896334692798-mml
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:43 PM 3/22/2024

IAM > Roles > Create role

Step 1
Select trusted entity

Step 2
Add permissions

Step 3
Name, review, and create

Select trusted entity Info

Trusted entity type

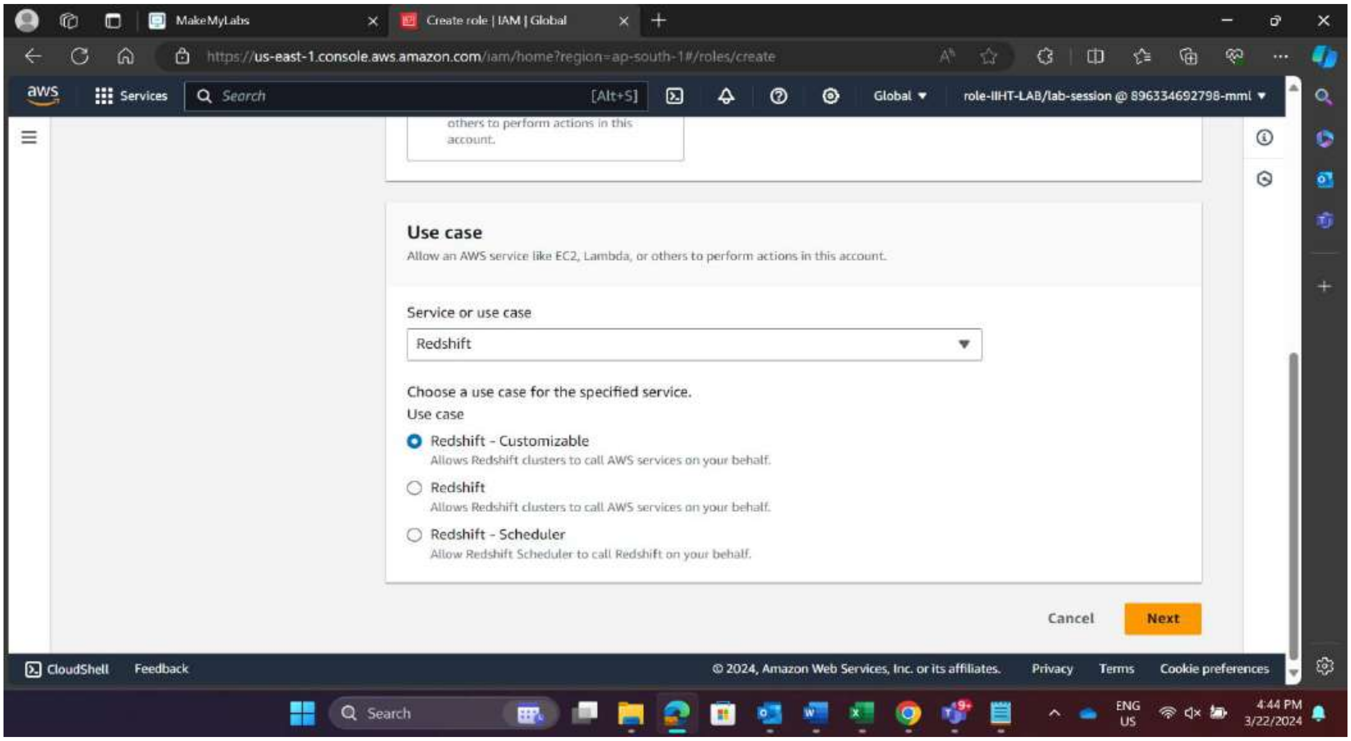
● AWS service
Allow AWS services like EC2, Lambda, or others to perform actions in this account.

○ AWS account
Allow entities in other AWS accounts belonging to you or a 3rd party to perform actions in this account.

○ Web identity
Allows users federated by the specified external web identity provider to assume this role to perform actions in this account.

○ SAML 2.0 federation
Allow users federated with SAML 2.0 from a corporate directory to perform actions in this account.

○ Custom trust policy
Create a custom trust policy to enable others to perform actions in this account.


MakeMyLabs
Create role | IAM | Global
https://us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles/create
aws
Services
Search
[Alt+S]
Global
role-IIHT-LAB/lab-session @ 896334692798-mml
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:44 PM 3/22/2024

others to perform actions in this account.

Use case
Allow an AWS service like EC2, Lambda, or others to perform actions in this account.

Service or use case
Redshift

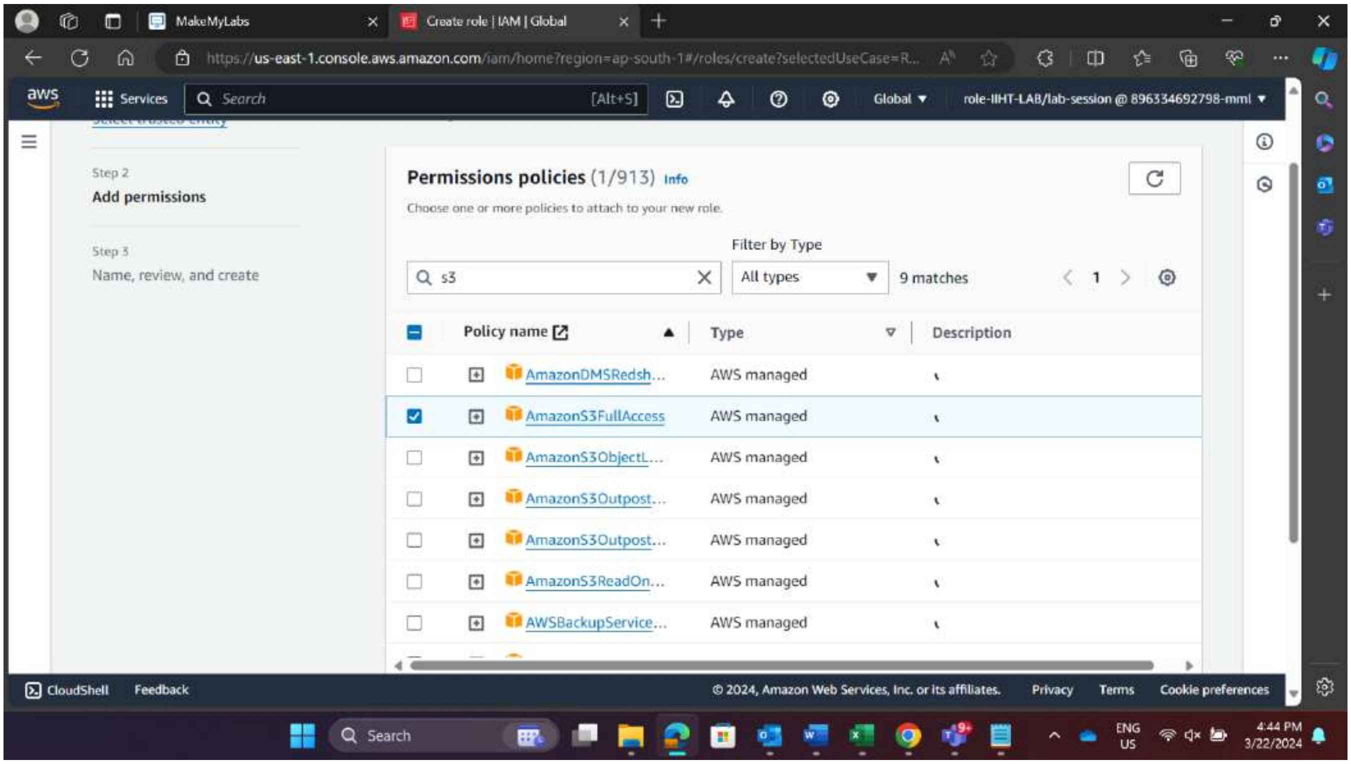
Choose a use case for the specified service.
Use case

● Redshift - Customizable
Allows Redshift clusters to call AWS services on your behalf.

○ Redshift
Allows Redshift clusters to call AWS services on your behalf.

○ Redshift - Scheduler
Allow Redshift Scheduler to call Redshift on your behalf.

Cancel Next


MakeMyLabs
Create role | IAM | Global
https://us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles/create?selectedUseCase=R...
Services
Search
[Alt+S]
Global ▼
role-IIHT-LAB/lab-session @ 896334692798-mml ▼
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
4:44 PM
3/22/2024

Step 2
Add permissions

Step 3
Name, review, and create

Permissions policies (1/913) Info

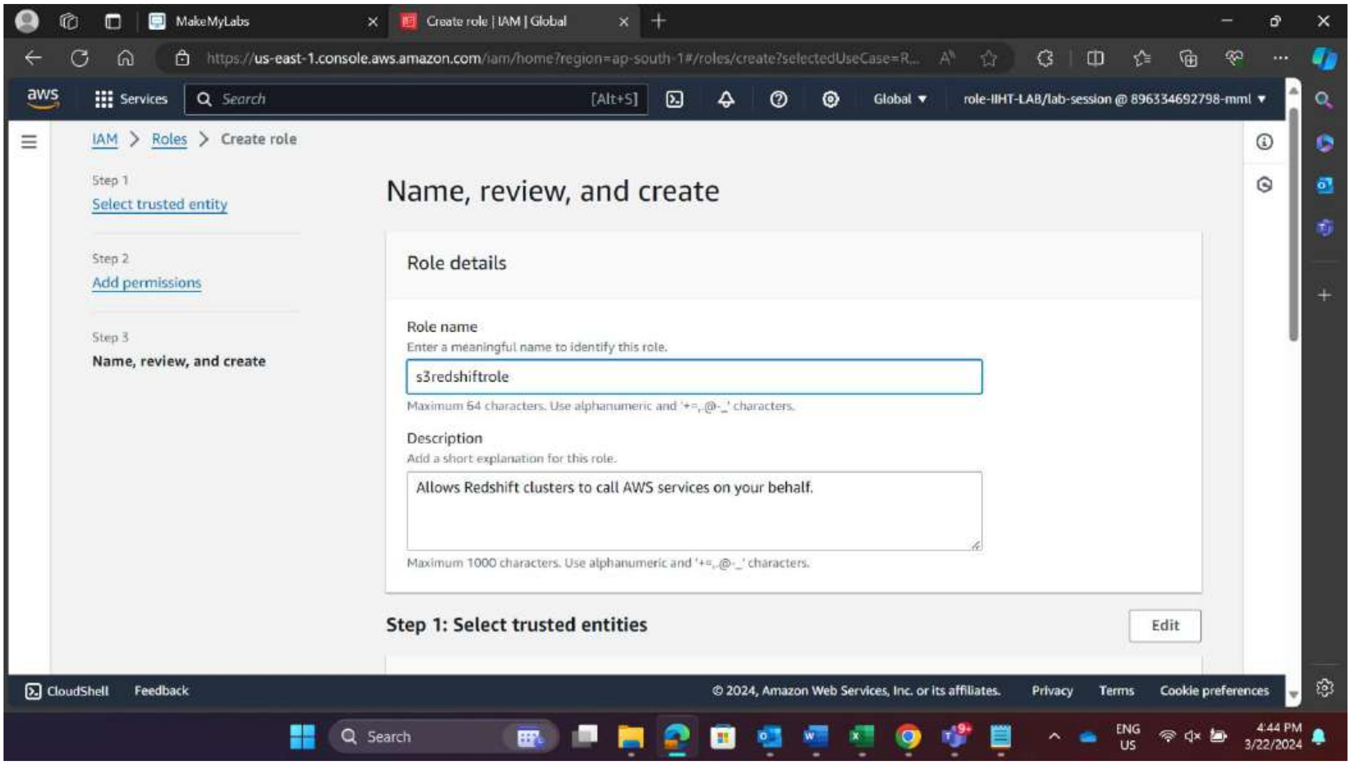
Choose one or more policies to attach to your new role.

Filter by Type

s3 | All types ▼ | 9 matches

	Policy name	Type	Description
□	AmazonDMSRedsh...	AWS managed	
☑	AmazonS3FullAccess	AWS managed	
□	AmazonS3ObjectL...	AWS managed	
□	AmazonS3Outpost...	AWS managed	
□	AmazonS3Outpost...	AWS managed	
□	AmazonS3ReadOn...	AWS managed	
□	AWSBackupService...	AWS managed	

Step 2: Provide the details and finally create the role.


MakeMyLabs
Create role | IAM | Global
https://us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles/create?selectedUseCase=R...
Services
Search
[Alt+S]
Global ▼
role-IIHT-LAB/lab-session @ 896334692798-mml ▼
IAM > Roles > Create role
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
4:44 PM
3/22/2024

Step 1
Select trusted entity

Step 2
Add permissions

Step 3
Name, review, and create

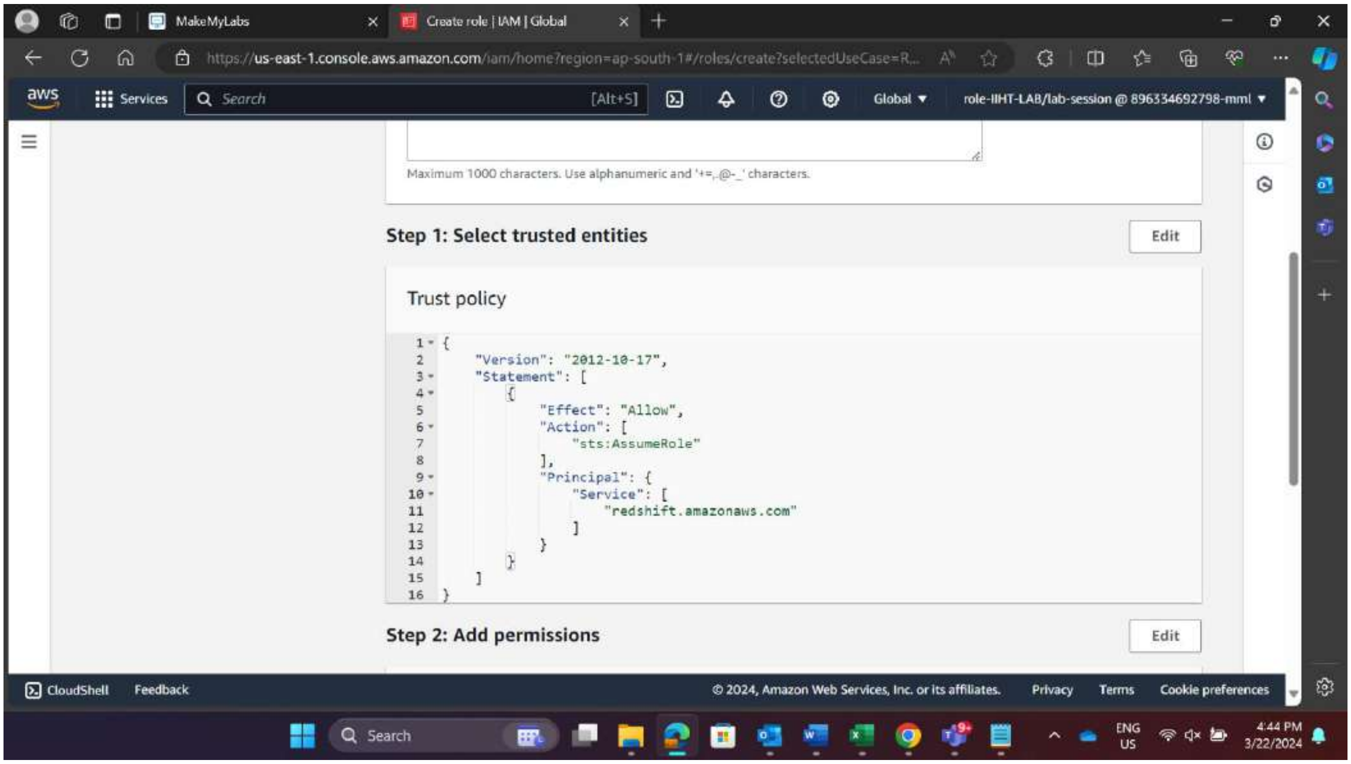
Name, review, and create

Role details

Role name
Enter a meaningful name to identify this role.
s3redshiftrole
Maximum 64 characters. Use alphanumeric and '+=,.@-_' characters.

Description
Add a short explanation for this role.
Allows Redshift clusters to call AWS services on your behalf.
Maximum 1000 characters. Use alphanumeric and '+=,.@-_' characters.

Step 1: Select trusted entities Edit


MakeMyLabs
Create role | IAM | Global
https://us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles/create?selectedUseCase=R...
aws
Services
Search
[Alt+S]
Global ▼
role-IIHT-LAB/lab-session @ 896334692798-mml ▼
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:44 PM 3/22/2024

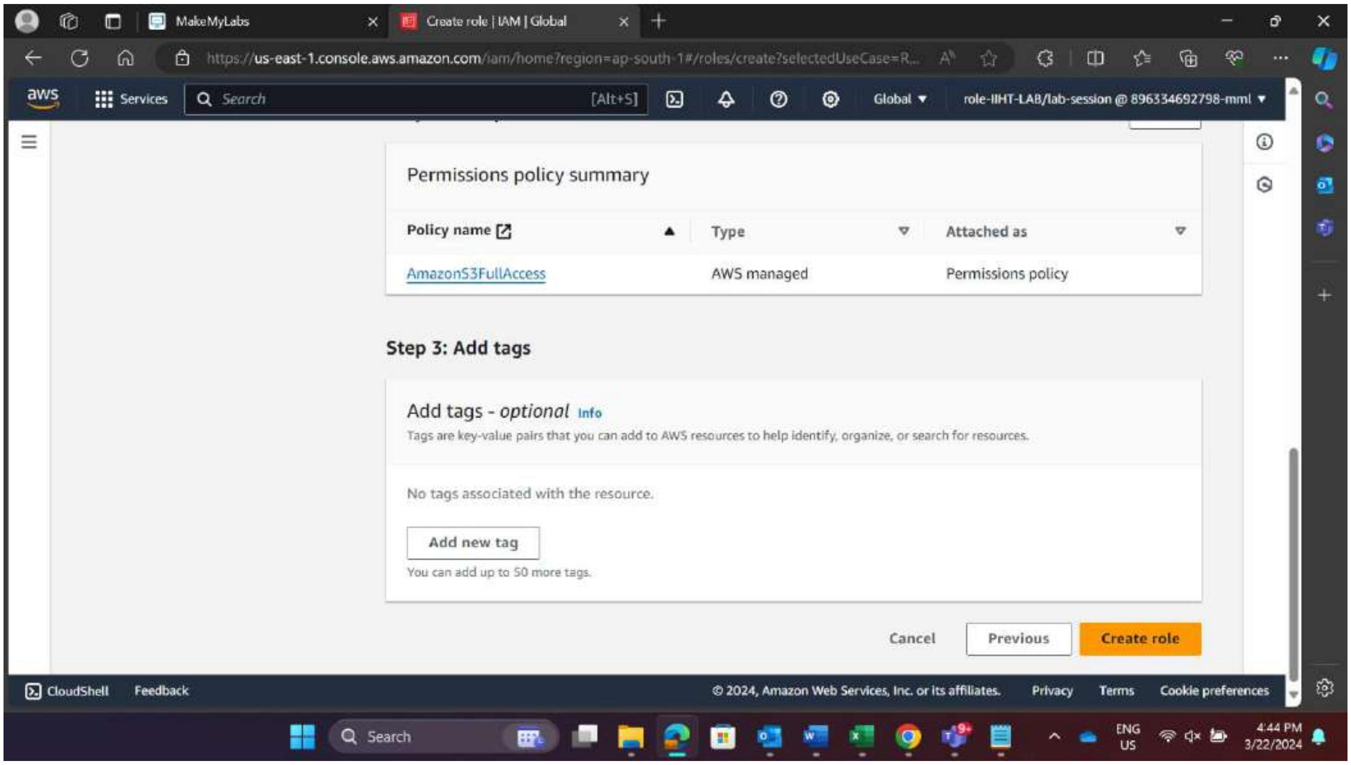
Maximum 1000 characters. Use alphanumeric and '+=,.@-_' characters.

Step 1: Select trusted entities Edit

Trust policy

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "sts:AssumeRole"
            ],
            "Principal": {
                "Service": [
                    "redshift.amazonaws.com"
                ]
            }
        }
    ]
}
```

Step 2: Add permissions Edit


MakeMyLabs
Create role | IAM | Global
https://us-east-1.console.aws.amazon.com/iam/home?region=ap-south-1#/roles/create?selectedUseCase=R...
aws
Services
Search
[Alt+S]
Global ▼
role-IIHT-LAB/lab-session @ 896334692798-mml ▼
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Search
ENG US
4:44 PM 3/22/2024

Permissions policy summary

Policy name	Type	Attached as
AmazonS3FullAccess	AWS managed	Permissions policy

Step 3: Add tags

Add tags - *optional* Info

Tags are key-value pairs that you can add to AWS resources to help identify, organize, or search for resources.

No tags associated with the resource.

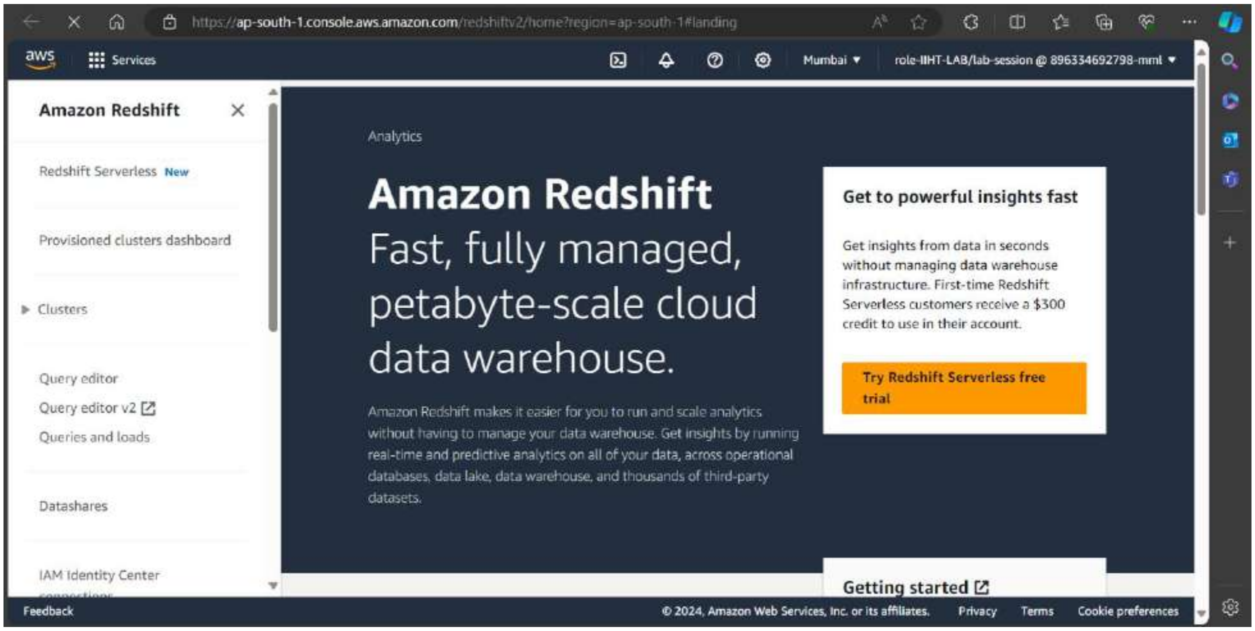
Add new tag

You can add up to 50 more tags.

Cancel Previous Create role

REDSHIFT CREATION

Step1: Sign in to the AWS Management Console and open the Amazon Redshift console and create the cluster. At upper right, choose the AWS Region where you want to create the cluster.


https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#landing
aws
Services
Mumbai ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
Amazon Redshift
Redshift Serverless New
Provisioned clusters dashboard
▸ Clusters
Query editor
Query editor v2
Queries and loads
Datashares
IAM Identity Center
Analytics
Try Redshift Serverless free trial
Getting started
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Amazon Redshift

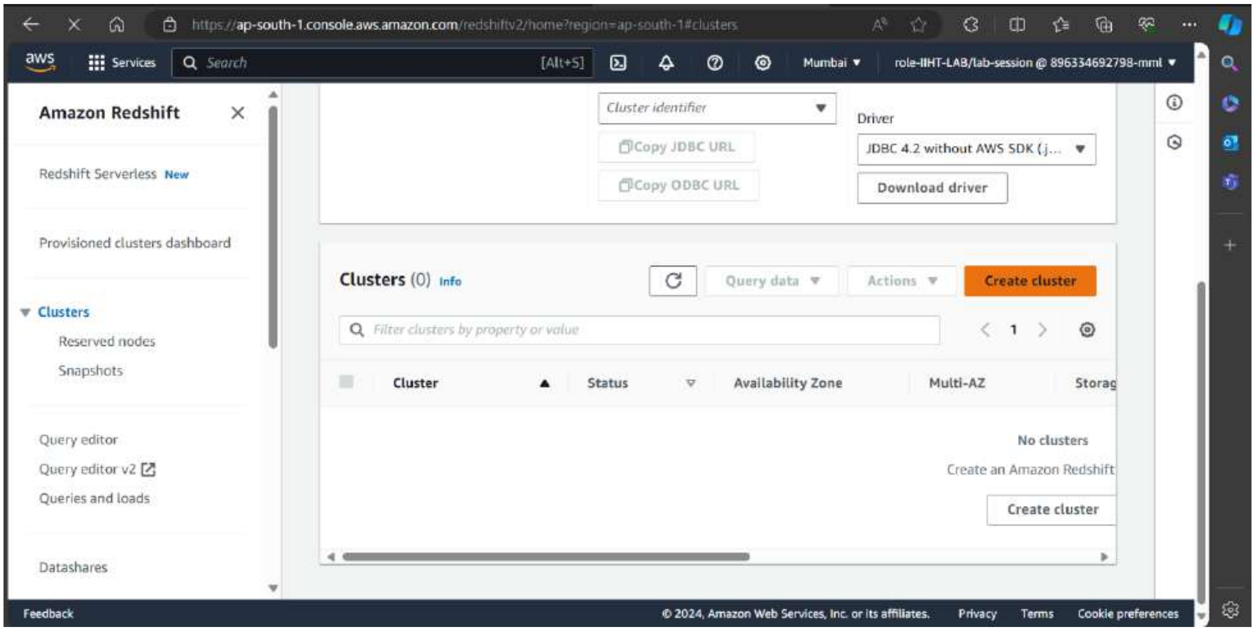
Fast, fully managed, petabyte-scale cloud data warehouse.

Amazon Redshift makes it easier for you to run and scale analytics without having to manage your data warehouse. Get insights by running real-time and predictive analytics on all of your data, across operational databases, data lake, data warehouse, and thousands of third-party datasets.

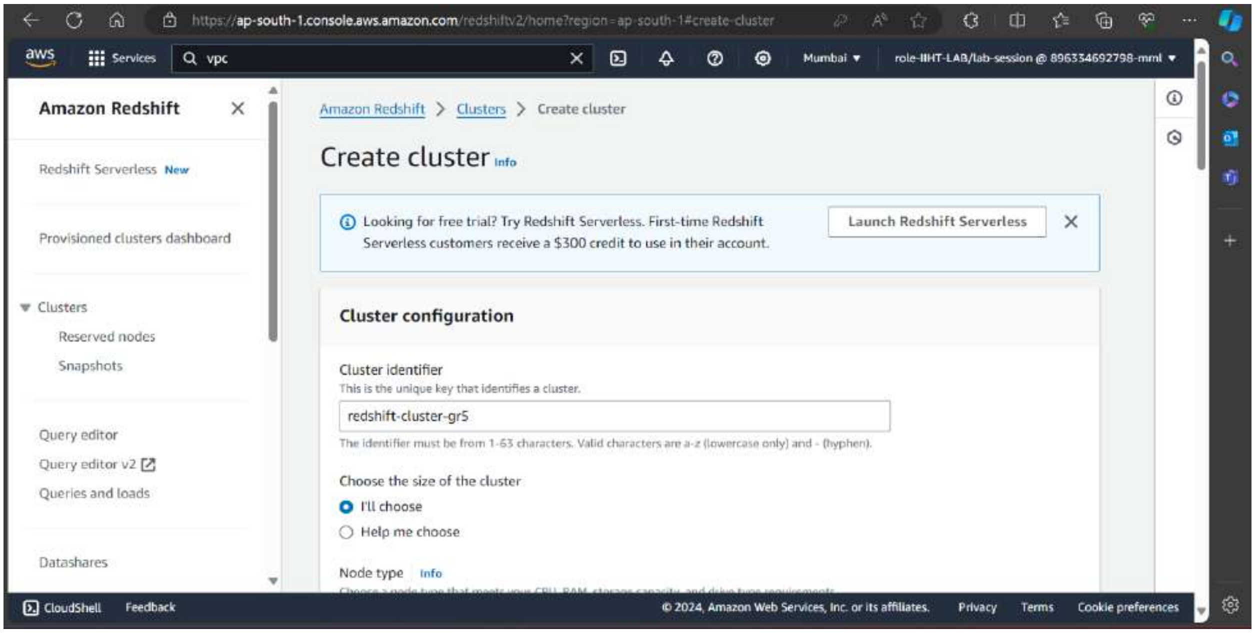
Get to powerful insights fast

Get insights from data in seconds without managing data warehouse infrastructure. First-time Redshift Serverless customers receive a $300 credit to use in their account.

Choose **Clusters**, then choose **Create cluster** and start creating the cluster.



Step 2: In the Cluster configuration section, specify values for Cluster identifier and size of the cluster.


https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-cluster
aws
Services
vpc
Mumbai ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
Amazon Redshift
Redshift Serverless New
Provisioned clusters dashboard
Clusters
Reserved nodes
Snapshots
Query editor
Query editor v2
Queries and loads
Datashares
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Amazon Redshift > Clusters > Create cluster

Create cluster Info

Looking for free trial? Try Redshift Serverless. First-time Redshift Serverless customers receive a $300 credit to use in their account. Launch Redshift Serverless

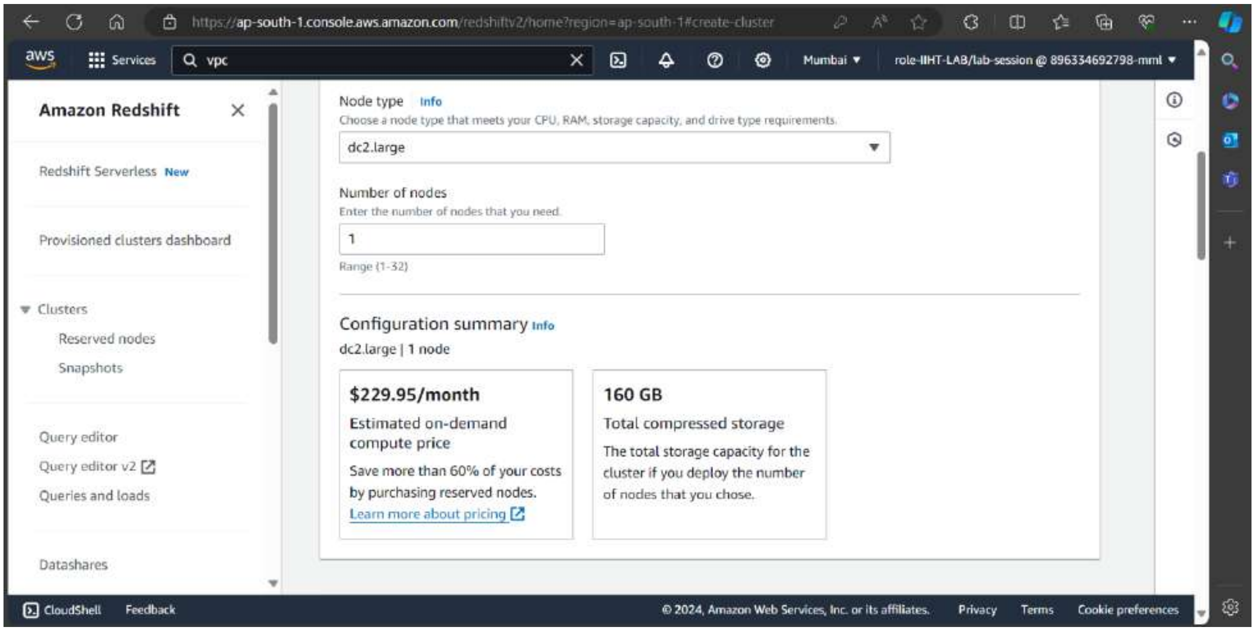
Cluster configuration

Cluster identifier
This is the unique key that identifies a cluster.
redshift-cluster-gr5
The identifier must be from 1-63 characters. Valid characters are a-z (lowercase only) and - (hyphen).

Choose the size of the cluster
● I'll choose
○ Help me choose

Node type Info

Then choose the **Node type** and number of **Nodes** to size your cluster. Choose dc2.large Node type and 1 for Nodes.


https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-cluster
aws
Services
vpc
Mumbai ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
Amazon Redshift
Redshift Serverless New
Provisioned clusters dashboard
Clusters
Reserved nodes
Snapshots
Query editor
Query editor v2
Queries and loads
Datashares
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Node type Info
Choose a node type that meets your CPU, RAM, storage capacity, and drive type requirements.
dc2.large ▾

Number of nodes
Enter the number of nodes that you need.
1
Range (1-32)

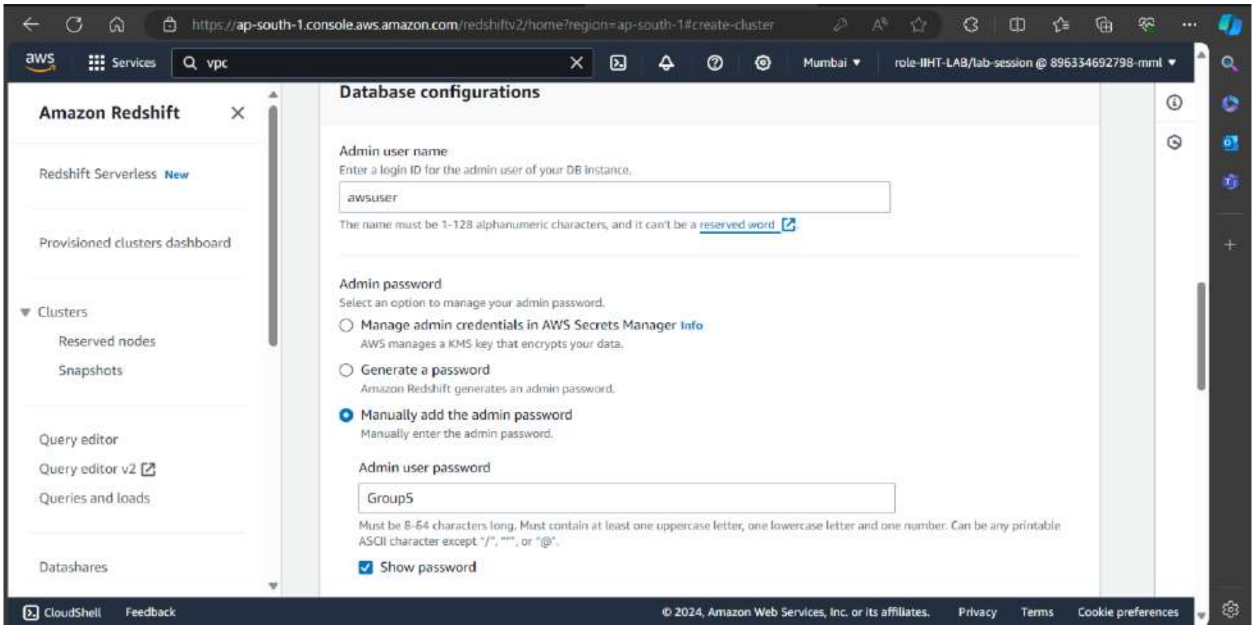
Configuration summary Info
dc2.large | 1 node

$229.95/month
Estimated on-demand compute price
Save more than 60% of your costs by purchasing reserved nodes.
Learn more about pricing

160 GB
Total compressed storage
The total storage capacity for the cluster if you deploy the number of nodes that you chose.

Step 3: In the Database configuration section, specify a value for **Admin username**. Choose Admin password as follows:

Manually add an admin password (Use your own password).


https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-cluster
aws
Services
vpc
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
Amazon Redshift
Redshift Serverless New
Provisioned clusters dashboard
Clusters
Reserved nodes
Snapshots
Query editor
Query editor v2
Queries and loads
Datashares
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Database configurations

Admin user name
Enter a login ID for the admin user of your DB instance.
awsuser
The name must be 1-128 alphanumeric characters, and it can't be a reserved word.

Admin password
Select an option to manage your admin password.

○ Manage admin credentials in AWS Secrets Manager Info
AWS manages a KMS key that encrypts your data.

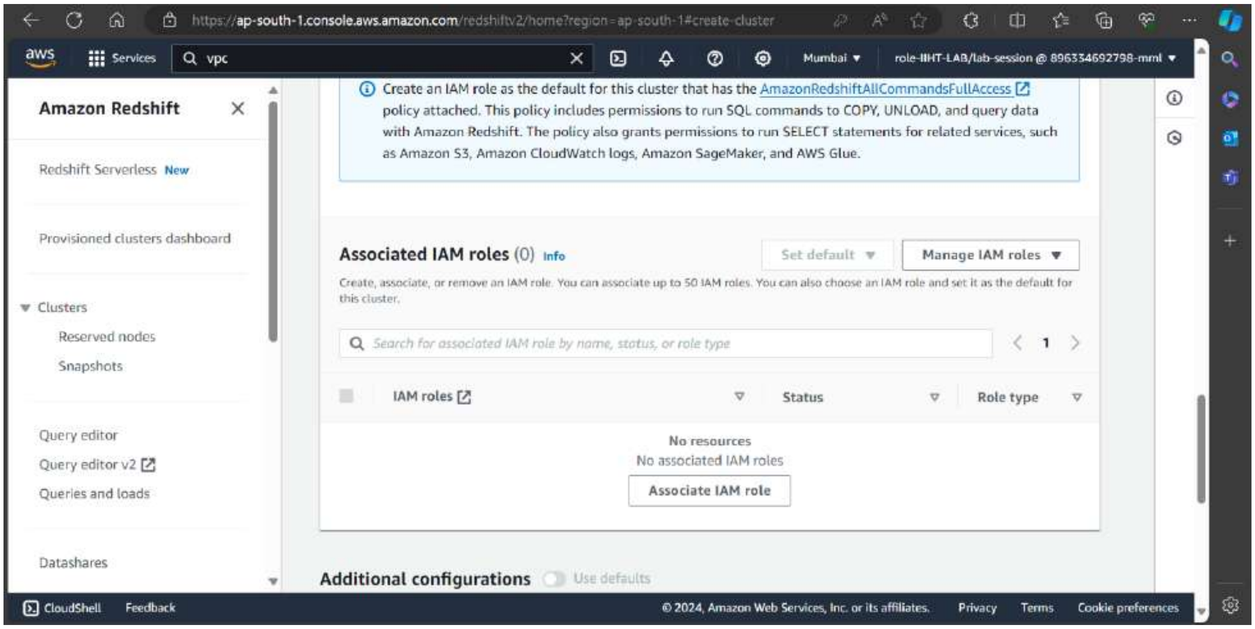
○ Generate a password
Amazon Redshift generates an admin password.

● Manually add the admin password
Manually enter the admin password.

Admin user password
Group5
Must be 8-64 characters long. Must contain at least one uppercase letter, one lowercase letter and one number. Can be any printable ASCII character except "/", """, or "@".

☑ Show password

Step 4: Create an IAM role and choose the one you have created for your cluster. Under Cluster permissions, for Manage IAM roles.



ⓘ Create an IAM role as the default for this cluster that has the AmazonRedshiftAllCommandsFullAccess policy attached. This policy includes permissions to run SQL commands to COPY, UNLOAD, and query data with Amazon Redshift. The policy also grants permissions to run SELECT statements for related services, such as Amazon S3, Amazon CloudWatch logs, Amazon SageMaker, and AWS Glue.

Associated IAM roles (0) Info

Create, associate, or remove an IAM role. You can associate up to 50 IAM roles. You can also choose an IAM role and set it as the default for this cluster.

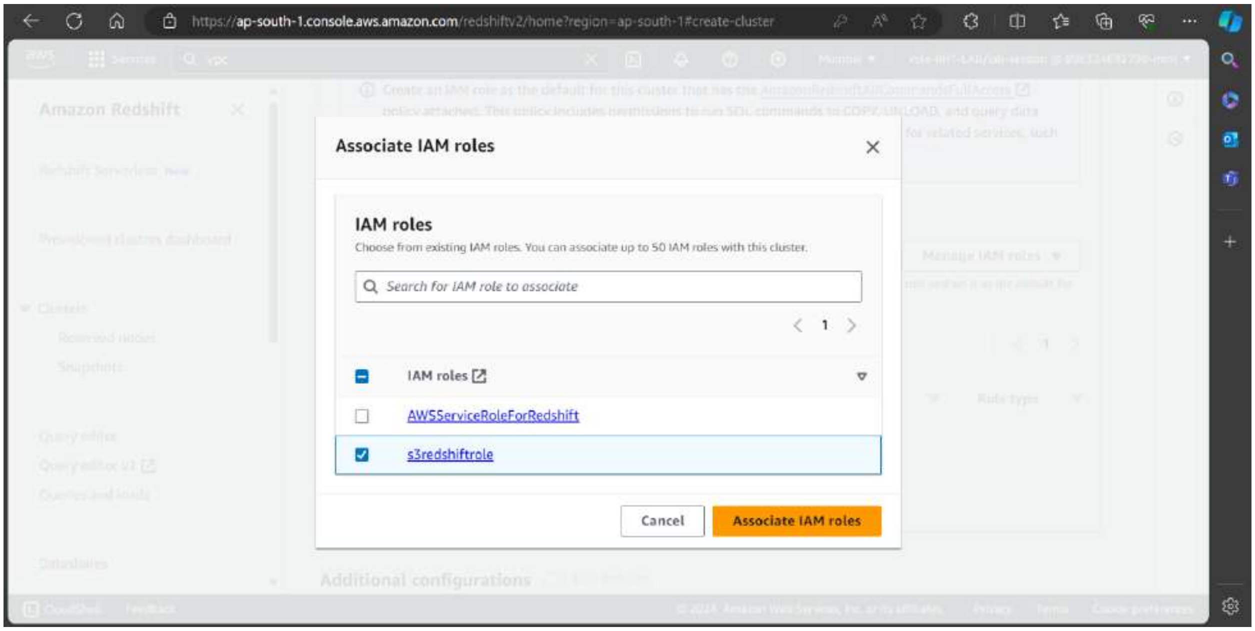
Search for associated IAM role by name, status, or role type

IAM roles	Status	Role type

No resources
No associated IAM roles

Additional configurations Use defaults

Specify an Amazon S3 bucket for the IAM role to access then Associate IAM role.



https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-cluster

Associate IAM roles

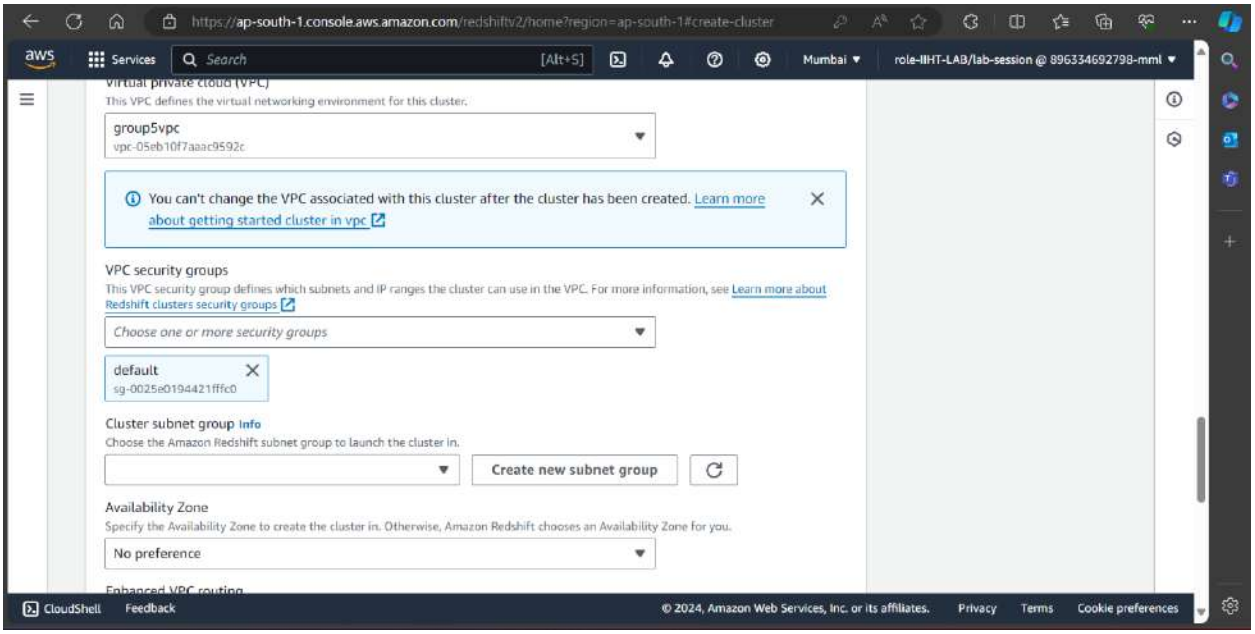
IAM roles

Choose from existing IAM roles. You can associate up to 50 IAM roles with this cluster.

Search for IAM role to associate

	IAM roles
□	AWSServiceRoleForRedshift
☑	s3redshiftrole

Step 5: Choose the VPC that you have created previously and set the VPC security groups as default.



https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-cluster

Virtual private cloud (VPC)

This VPC defines the virtual networking environment for this cluster.

group5vpc
vpc-05eb10f7aaac9592c

You can't change the VPC associated with this cluster after the cluster has been created. Learn more about getting started cluster in vpc

VPC security groups

This VPC security group defines which subnets and IP ranges the cluster can use in the VPC. For more information, see Learn more about Redshift clusters security groups

Choose one or more security groups

default
sg-0025e0194421fffc0

Cluster subnet group Info

Choose the Amazon Redshift subnet group to launch the cluster in.

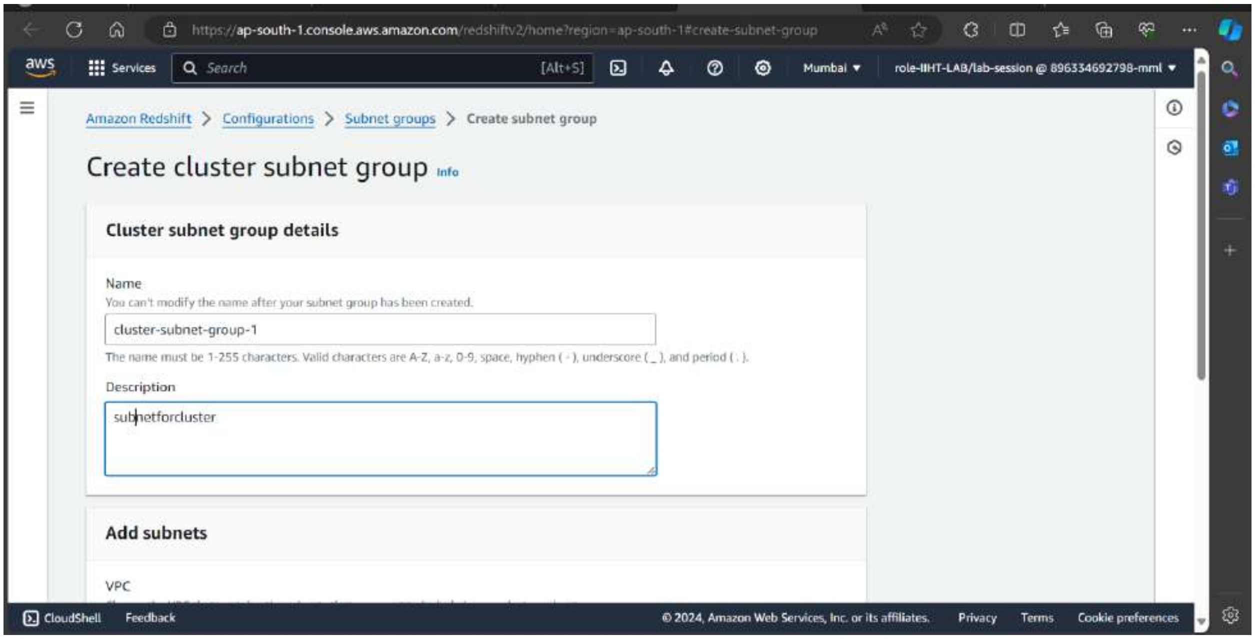
Availability Zone

Specify the Availability Zone to create the cluster in. Otherwise, Amazon Redshift chooses an Availability Zone for you.

No preference

Enhanced VPC routing

Step 6: Create a cluster subnet group. A cluster subnet group allows you to specify a set of subnets in your VPC.



https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-subnet-group
Services
Search
[Alt+S]
Mumbai ▼
role-IIHT-LAB/lab-session @ 896334692798-mml ▼
Amazon Redshift > Configurations > Subnet groups > Create subnet group
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Create cluster subnet group Info

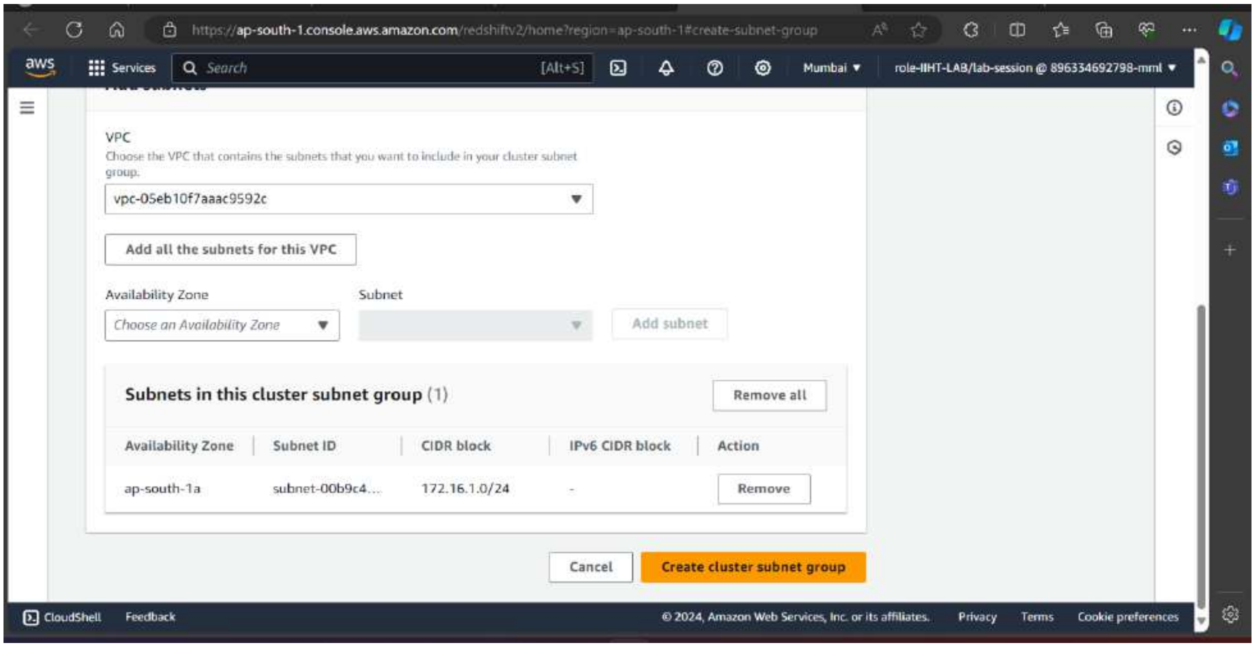
Cluster subnet group details

Name
You can't modify the name after your subnet group has been created.
cluster-subnet-group-1
The name must be 1-255 characters. Valid characters are A-Z, a-z, 0-9, space, hyphen (-), underscore (_), and period (.).

Description
subnetforcluster

Add subnets

VPC



https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-subnet-group
Services
Search
[Alt+S]
Mumbai ▼
role-IIHT-LAB/lab-session @ 896334692798-mml ▼
Add all the subnets for this VPC
Add subnet
Remove all
Cancel
Create cluster subnet group
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

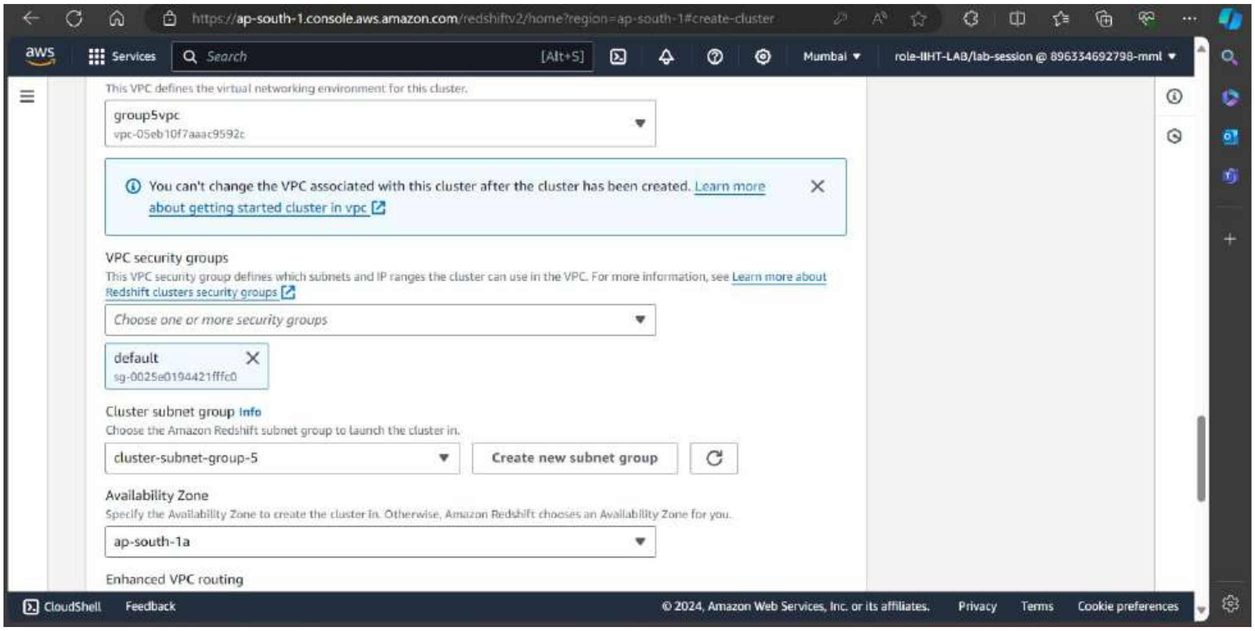
VPC
Choose the VPC that contains the subnets that you want to include in your cluster subnet group.
vpc-05eb10f7aaac9592c

Availability Zone: Choose an Availability Zone
Subnet:

Subnets in this cluster subnet group (1)

Availability Zone	Subnet ID	CIDR block	IPv6 CIDR block	Action
ap-south-1a	subnet-00b9c4...	172.16.1.0/24	-	Remove

Step 7: Choose the Availability Zone in which location should subnet run.



This VPC defines the virtual networking environment for this cluster.

group5vpc
vpc-05eb10f7aaac9592c

You can't change the VPC associated with this cluster after the cluster has been created. Learn more about getting started cluster in vpc

VPC security groups
This VPC security group defines which subnets and IP ranges the cluster can use in the VPC. For more information, see Learn more about Redshift clusters security groups

Choose one or more security groups

default
sg-0025e0194421fffc0

Cluster subnet group Info
Choose the Amazon Redshift subnet group to launch the cluster in.

cluster-subnet-group-5

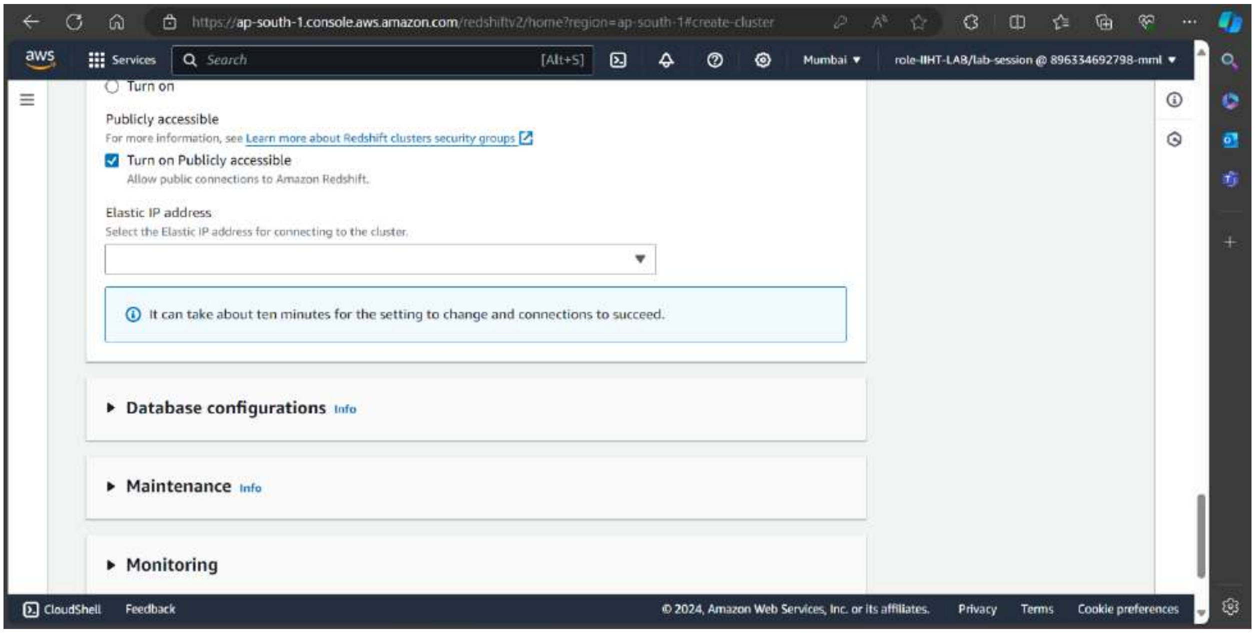
Availability Zone
Specify the Availability Zone to create the cluster in. Otherwise, Amazon Redshift chooses an Availability Zone for you.

ap-south-1a

Enhanced VPC routing

https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-cluster
aws
Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
Create new subnet group
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Turn on **"publicly accessible"**. This allows public connection to redshift.



○ Turn on

Publicly accessible
For more information, see Learn more about Redshift clusters security groups

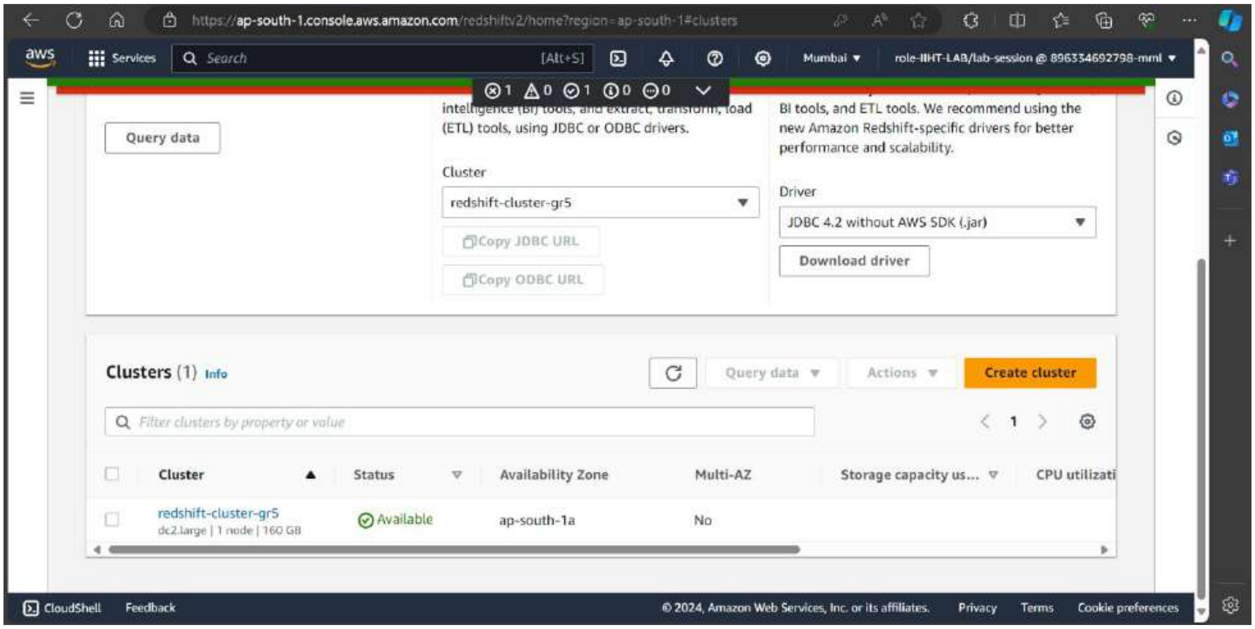
☑ Turn on Publicly accessible
Allow public connections to Amazon Redshift.

Elastic IP address
Select the Elastic IP address for connecting to the cluster.

It can take about ten minutes for the setting to change and connections to succeed.

https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#create-cluster
aws
Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
Database configurations Info
Maintenance Info
Monitoring
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

After completing the configuration setup "Create the cluster"


https://ap-south-1.console.aws.amazon.com/redshiftv2/home?region=ap-south-1#clusters
aws
Services
Search
[Alt+S]
Mumbai ▾
role-IIHT-LAB/lab-session @ 896334692798-mml ▾
⊗1 ⚠0 ⊘1 ⓘ0 ⊝0
Query data
Copy JDBC URL
Copy ODBC URL
Download driver
Query data ▾
Actions ▾
Create cluster
Filter clusters by property or value
1
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

intelligence (BI) tools, and extract, transform, load (ETL) tools, using JDBC or ODBC drivers.

BI tools, and ETL tools. We recommend using the new Amazon Redshift-specific drivers for better performance and scalability.

Cluster: redshift-cluster-gr5

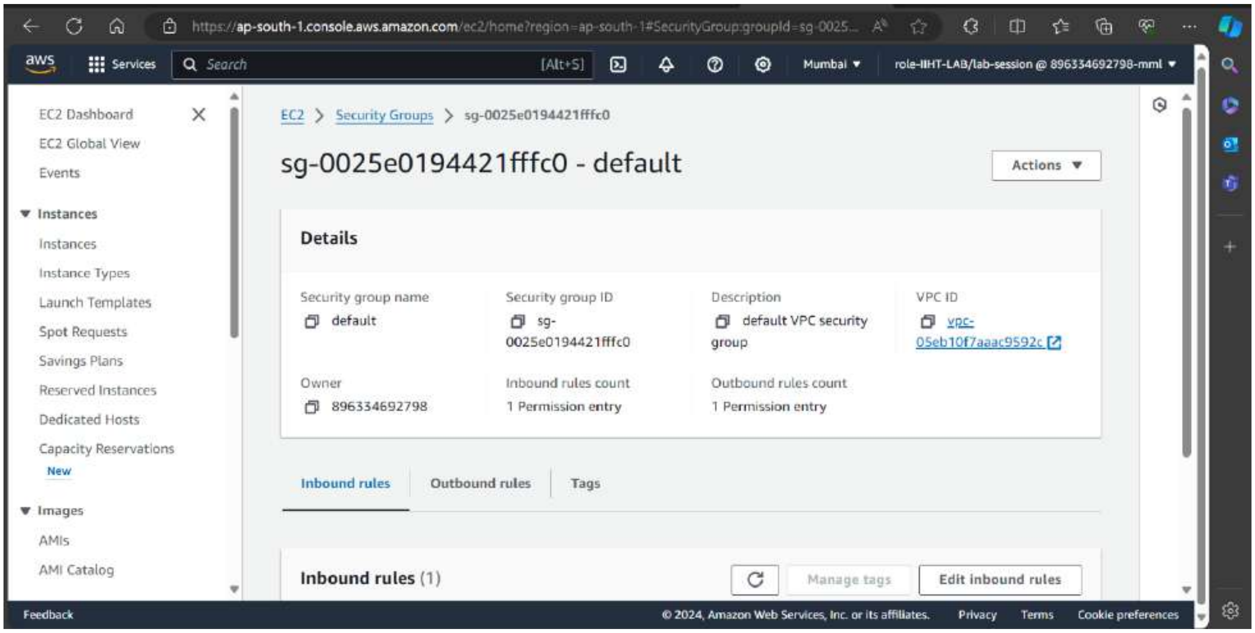
Driver: JDBC 4.2 without AWS SDK (.jar)

Clusters (1) Info

Cluster	Status	Availability Zone	Multi-AZ	Storage capacity us...	CPU utilizati
redshift-cluster-gr5 dc2.large | 1 node | 160 GB	✓ Available	ap-south-1a	No		

Step 8: Edit inbound rules under security groups and grant inbound access to cluster.

To access from an Amazon EC2 external, add a rule to the security group attached to your cluster that allows inbound traffic.



sg-0025e0194421fffc0 - default

Details

Security group name: default

Security group ID: sg-0025e0194421fffc0

Description: default VPC security group

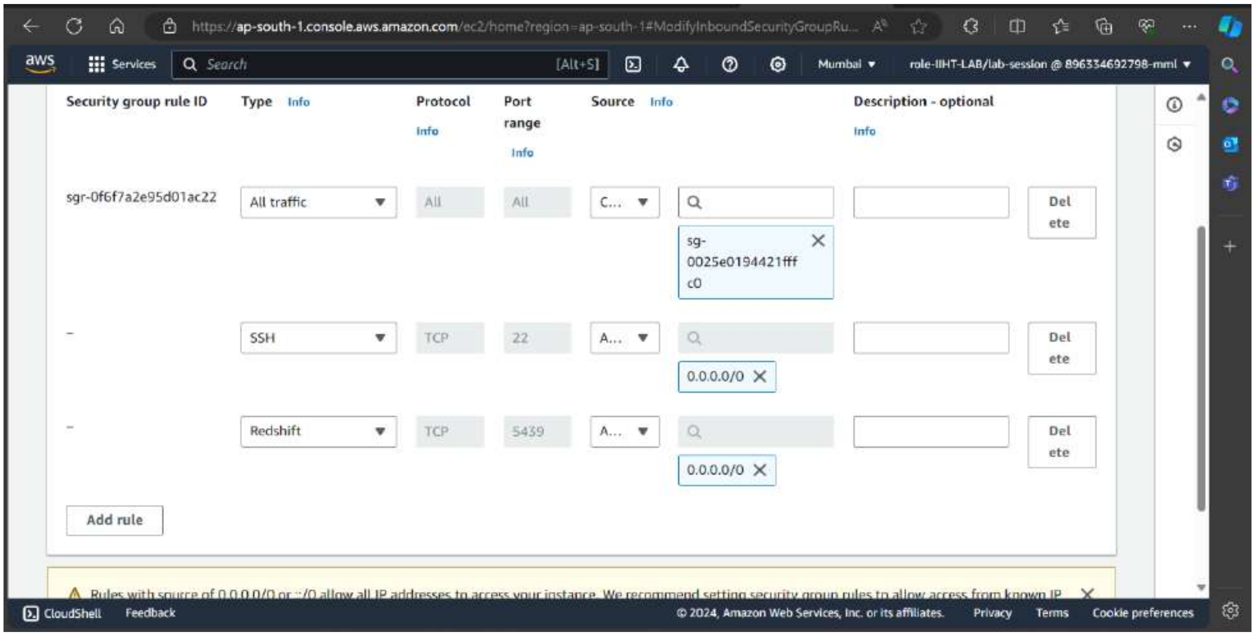
VPC ID: vpc-05eb10f7aaac9592c

Owner: 896334692798

Inbound rules count: 1 Permission entry

Outbound rules count: 1 Permission entry

Inbound rules (1)


https://ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#ModifyInboundSecurityGroupRu...
aws
Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 896334692798-mml
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Security group rule ID	Type Info	Protocol Info	Port range Info	Source Info		Description - optional Info	
sgr-0f6f7a2e95d01ac22	All traffic	All	All	C...	sg-0025e0194421fffc0		Delete
-	SSH	TCP	22	A...	0.0.0.0/0		Delete
-	Redshift	TCP	5439	A...	0.0.0.0/0		Delete

Add rule

⚠ Rules with source of 0.0.0.0/0 or ::/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP

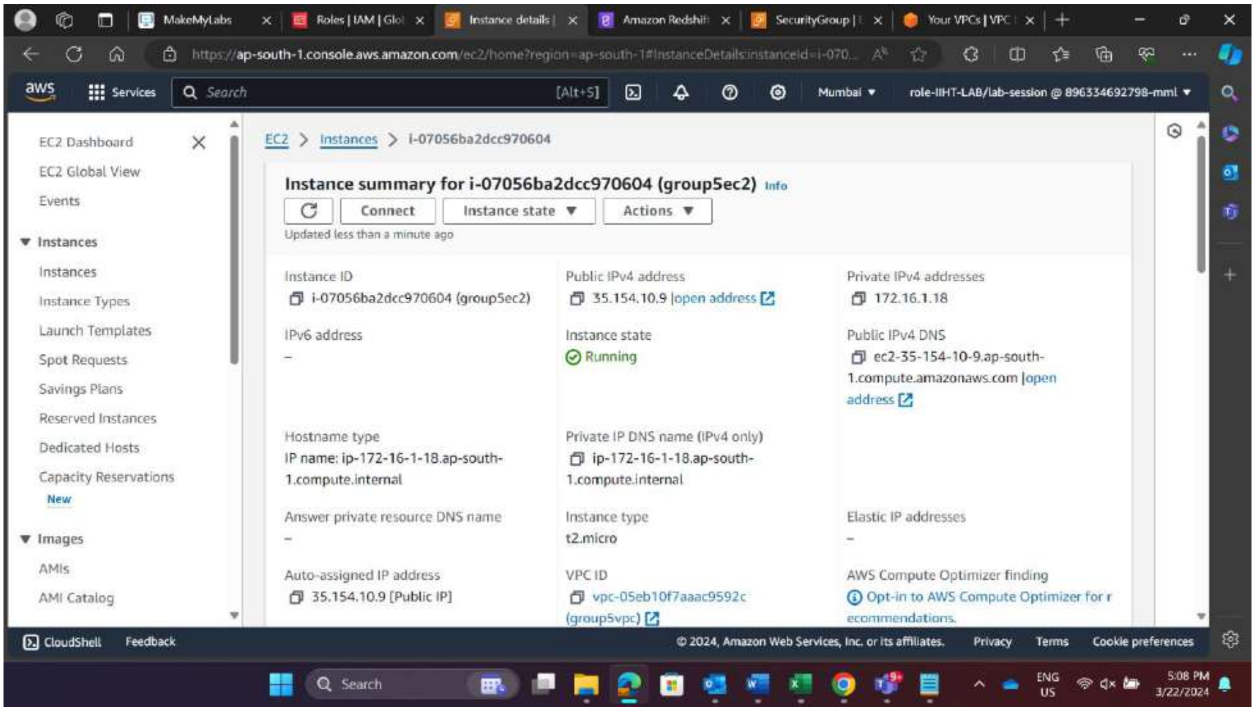
Step 9: After creating your cluster, you can immediately run queries using the Amazon Redshift console.

To query databases through Amazon Redshift cluster,

- Connect to your cluster and run queries on the AWS Management Console with one of the query editors.
- Connect to your cluster through an SQL client tool, such as SQL Workbench

Connection to the EC2 instance

Step 1: Goto your EC2 dashboard and select the instance and click on Connect.



https://ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#InstanceDetails:instanceId=i-070...

Mumbai ▼ role-IIHT-LAB/lab-session @ 896334692798-mml ▼

Instance summary for i-07056ba2dcc970604 (group5ec2) Info

Updated less than a minute ago

Instance ID: i-07056ba2dcc970604 (group5ec2)

Public IPv4 address: 35.154.10.9 | open address

Private IPv4 addresses: 172.16.1.18

IPv6 address: –

Instance state: Running

Public IPv4 DNS: ec2-35-154-10-9.ap-south-1.compute.amazonaws.com | open address

Hostname type: IP name: ip-172-16-1-18.ap-south-1.compute.internal

Private IP DNS name (IPv4 only): ip-172-16-1-18.ap-south-1.compute.internal

Answer private resource DNS name: –

Instance type: t2.micro

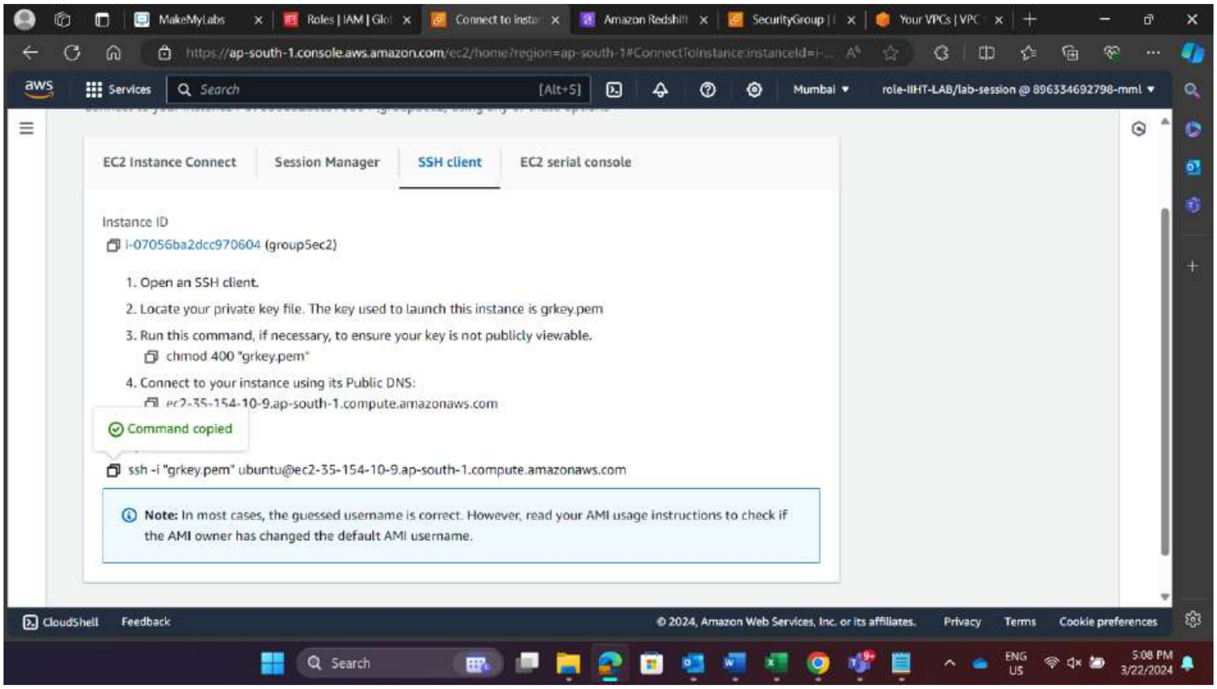
Elastic IP addresses: –

Auto-assigned IP address: 35.154.10.9 [Public IP]

VPC ID: vpc-05eb10f7aaac9592c (group5vpc)

AWS Compute Optimizer finding: Opt-in to AWS Compute Optimizer for recommendations.

Step 2: Use SSH Client option to connect to your instance and copy the command given as an example.



https://ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#ConnectToInstance:instanceId=i-...

Instance ID: i-07056ba2dcc970604 (group5ec2)

1. Open an SSH client.
2. Locate your private key file. The key used to launch this instance is grkey.pem
3. Run this command, if necessary, to ensure your key is not publicly viewable.

```
chmod 400 "grkey.pem"
```

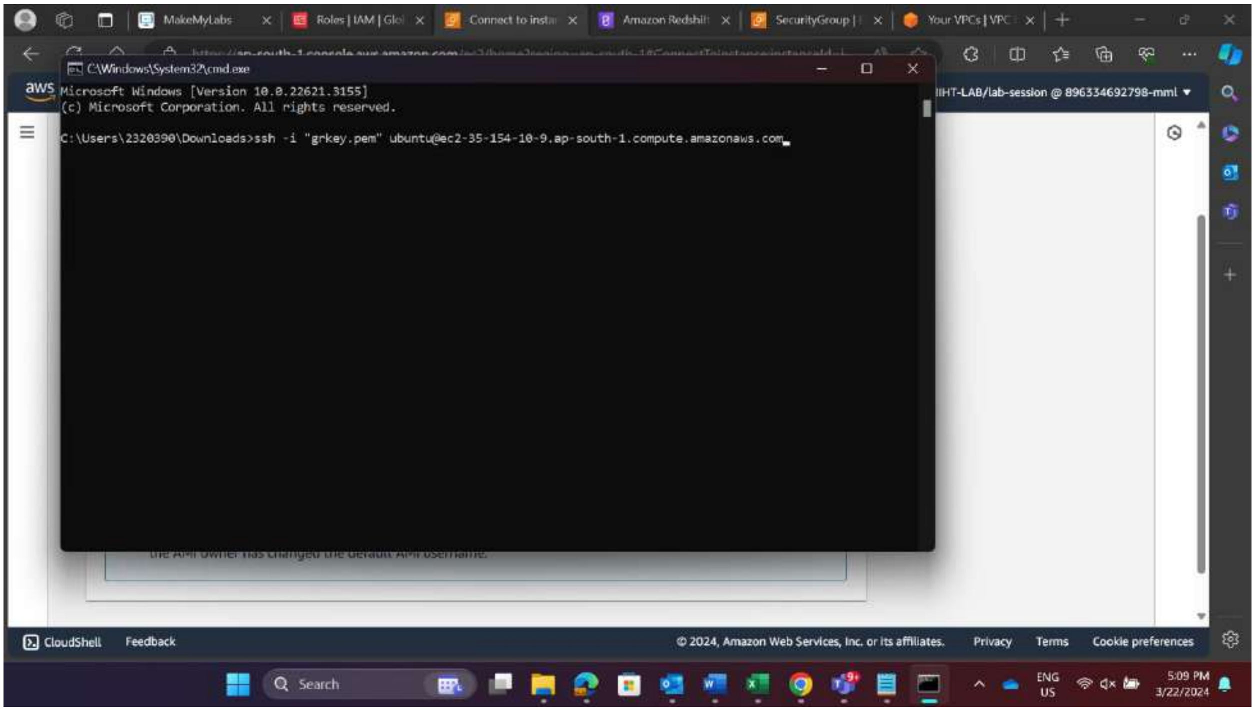
4. Connect to your instance using its Public DNS:

```
ec2-35-154-10-9.ap-south-1.compute.amazonaws.com
```

```
ssh -i "grkey.pem" ubuntu@ec2-35-154-10-9.ap-south-1.compute.amazonaws.com
```

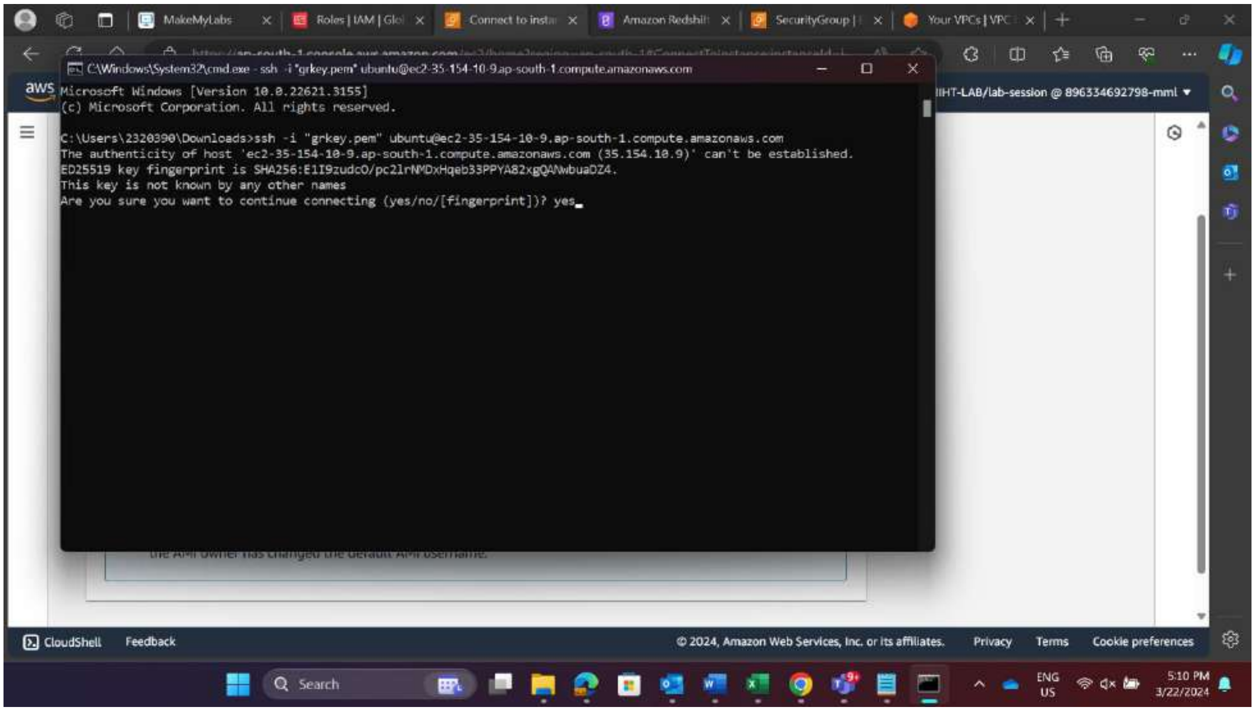
Note: In most cases, the guessed username is correct. However, read your AMI usage instructions to check if the AMI owner has changed the default AMI username.

Step 3: Open a terminal or the command prompt on the local machine and paste the command copied in the previous step. Make sure that you are inside the directory where your key pair generated during the instance creation is stored before executing the command.



C:\Windows\System32\cmd.exe

```
Microsoft Windows [Version 10.0.22621.3155]
(c) Microsoft Corporation. All rights reserved.

C:\Users\2320390\Downloads>ssh -i "grkey.pem" ubuntu@ec2-35-154-10-9.ap-south-1.compute.amazonaws.com
```

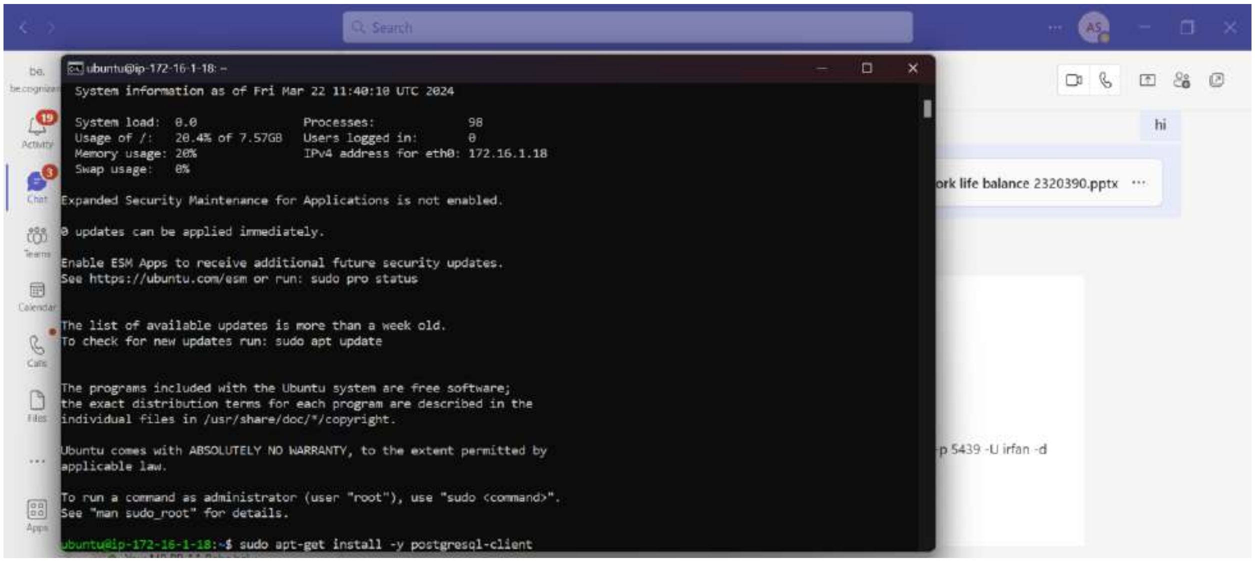


C:\Windows\System32\cmd.exe - ssh -i "grkey.pem" ubuntu@ec2-35-154-10-9.ap-south-1.compute.amazonaws.com

```
Microsoft Windows [Version 10.0.22621.3155]
(c) Microsoft Corporation. All rights reserved.

C:\Users\2320390\Downloads>ssh -i "grkey.pem" ubuntu@ec2-35-154-10-9.ap-south-1.compute.amazonaws.com
The authenticity of host 'ec2-35-154-10-9.ap-south-1.compute.amazonaws.com (35.154.10.9)' can't be established.
ED25519 key fingerprint is SHA256:E1I9zudcO/pc2lrNMDxHqeb33PPYA82xgQANwbuaDZ4.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
```

Step 4: Once the above command is successfully executed you will be logged into your ubuntu server. The next step is to install PostgreSQL on your virtual server by running command '**sudo apt-get install -y PostgreSQL-client'**

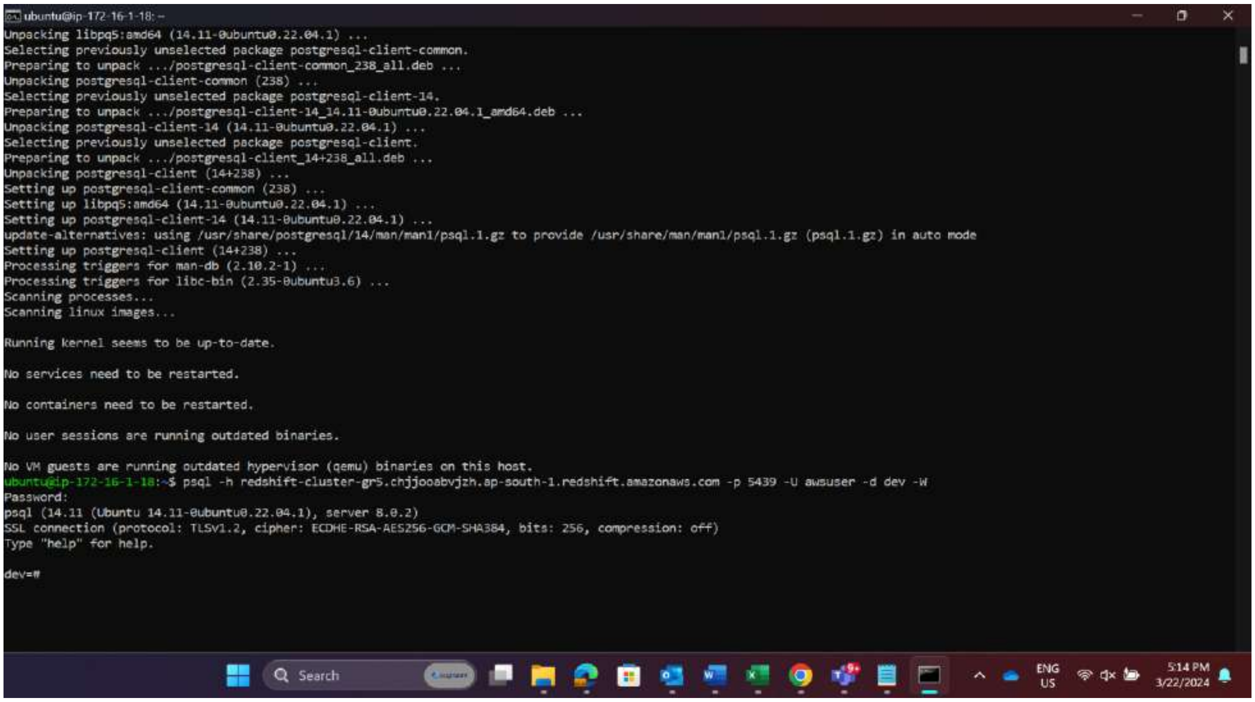
PostgreSQL is installed successfully on your ubuntu server.



```
 System information as of Fri Mar 22 11:40:10 UTC 2024

 System load:  0.0              Processes:             98
 Usage of /:   20.4% of 7.57GB  Users logged in:       0
 Memory usage: 20%              IPv4 address for eth0: 172.16.1.18
 Swap usage:   0%

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status


The list of available updates is more than a week old.
To check for new updates run: sudo apt update


The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

ubuntu@ip-172-16-1-18:~$ sudo apt-get install -y postgresql-client
```

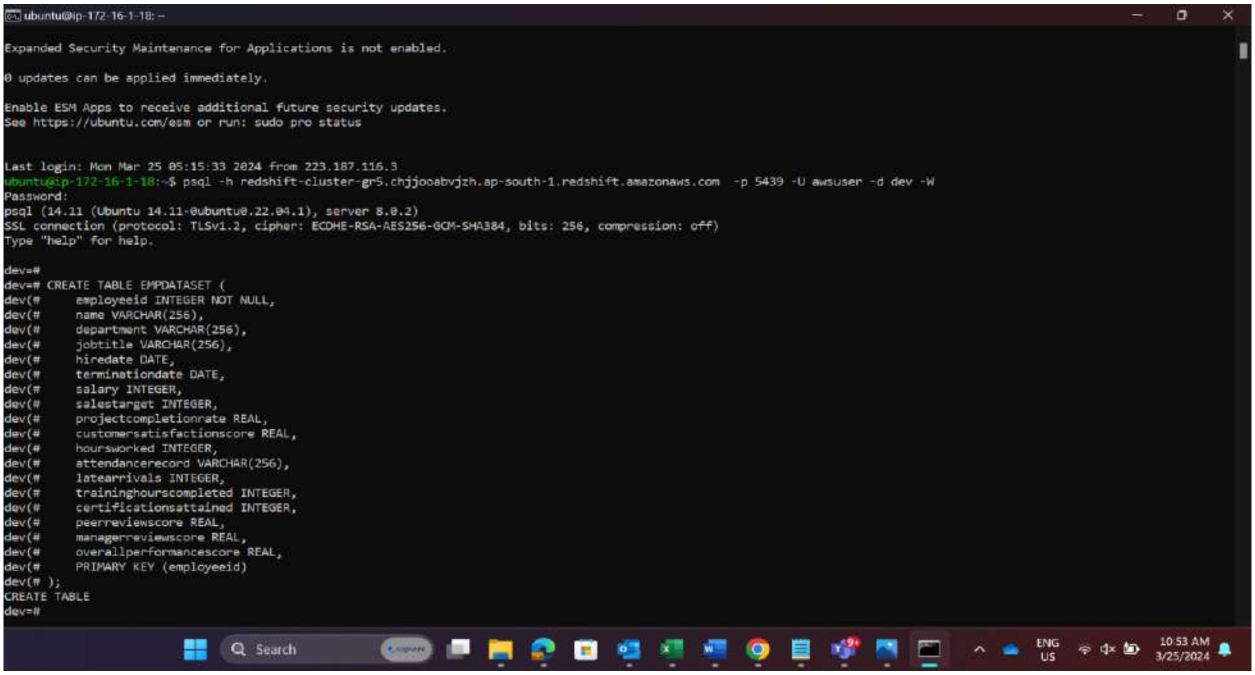
Step 5: After successfully installing the PostgreSQL on your server, the next step is connecting to the redshift cluster. This is done by using the following command.

'psql -h redshift-cluster-1.cackughiw5hj.ap-south-1.redshift.amazonaws.com -p 5439 -U awsuser -d dev –W'



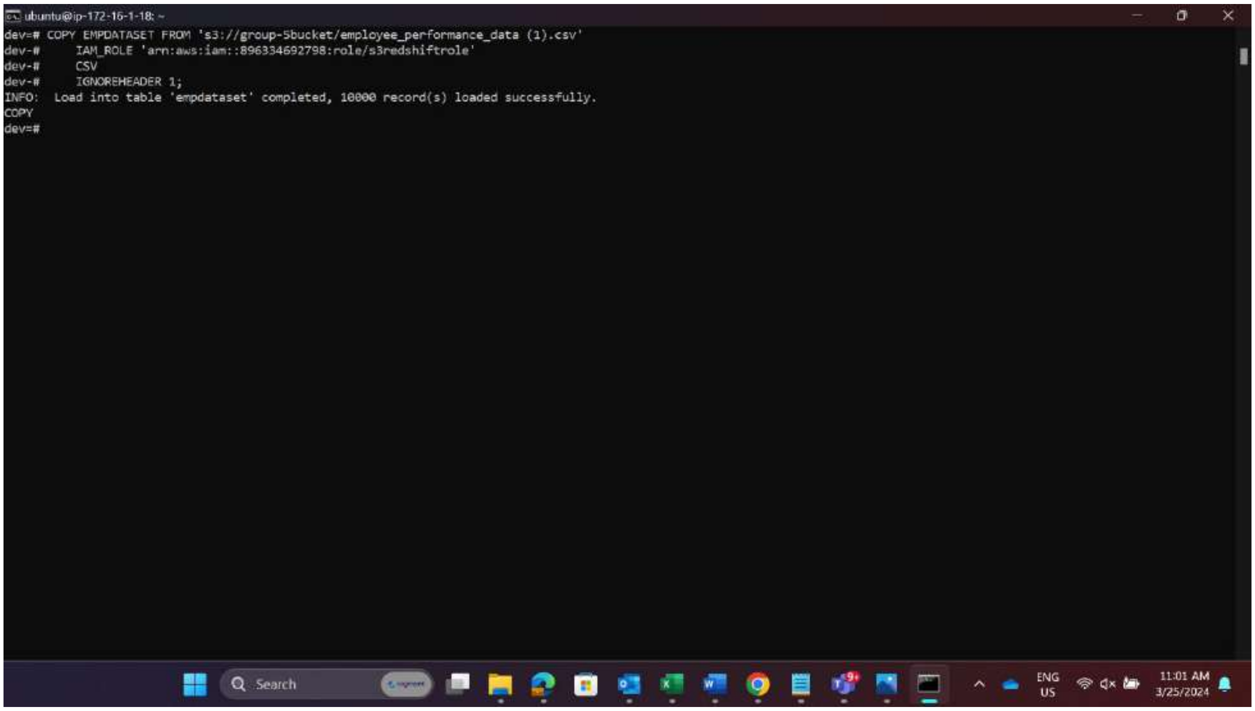
```
Unpacking libpq5:amd64 (14.11-0ubuntu0.22.04.1) ...
Selecting previously unselected package postgresql-client-common.
Preparing to unpack .../postgresql-client-common_238_all.deb ...
Unpacking postgresql-client-common (238) ...
Selecting previously unselected package postgresql-client-14.
Preparing to unpack .../postgresql-client-14_14.11-0ubuntu0.22.04.1_amd64.deb ...
Unpacking postgresql-client-14 (14.11-0ubuntu0.22.04.1) ...
Selecting previously unselected package postgresql-client.
Preparing to unpack .../postgresql-client_14+238_all.deb ...
Unpacking postgresql-client (14+238) ...
Setting up postgresql-client-common (238) ...
Setting up libpq5:amd64 (14.11-0ubuntu0.22.04.1) ...
Setting up postgresql-client-14 (14.11-0ubuntu0.22.04.1) ...
update-alternatives: using /usr/share/postgresql/14/man/man1/psql.1.gz to provide /usr/share/man/man1/psql.1.gz (psql.1.gz) in auto mode
Setting up postgresql-client (14+238) ...
Processing triggers for man-db (2.10.2-1) ...
Processing triggers for libc-bin (2.35-0ubuntu3.6) ...
Scanning processes...
Scanning linux images...

Running kernel seems to be up-to-date.

No services need to be restarted.

No containers need to be restarted.

No user sessions are running outdated binaries.

No VM guests are running outdated hypervisor (qemu) binaries on this host.
ubuntu@ip-172-16-1-18:~$ psql -h redshift-cluster-gr5.chjjooabvjzh.ap-south-1.redshift.amazonaws.com -p 5439 -U awsuser -d dev -W
Password:
psql (14.11 (Ubuntu 14.11-0ubuntu0.22.04.1), server 8.0.2)
SSL connection (protocol: TLSv1.2, cipher: ECDHE-RSA-AES256-GCM-SHA384, bits: 256, compression: off)
Type "help" for help.

dev=#
```

Step 6: Once the above command is successfully executed, you can access your database. Now, create the required table (EMPDATASET) using the CREATE statement and define the column names and its attributes.



```
Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status


Last login: Mon Mar 25 05:15:33 2024 from 223.187.116.3
ubuntu@ip-172-16-1-18:~$ psql -h redshift-cluster-gr5.chjjooabvjzh.ap-south-1.redshift.amazonaws.com  -p 5439 -U awsuser -d dev -W
Password:
psql (14.11 (Ubuntu 14.11-0ubuntu0.22.04.1), server 8.0.2)
SSL connection (protocol: TLSv1.2, cipher: ECDHE-RSA-AES256-GCM-SHA384, bits: 256, compression: off)
Type "help" for help.

dev=#
dev=# CREATE TABLE EMPDATASET (
dev(#     employeeid INTEGER NOT NULL,
dev(#     name VARCHAR(256),
dev(#     department VARCHAR(256),
dev(#     jobtitle VARCHAR(256),
dev(#     hiredate DATE,
dev(#     terminationdate DATE,
dev(#     salary INTEGER,
dev(#     salestarget INTEGER,
dev(#     projectcompletionrate REAL,
dev(#     customersatisfactionscore REAL,
dev(#     hoursworked INTEGER,
dev(#     attendancerecord VARCHAR(256),
dev(#     latearrivals INTEGER,
dev(#     traininghourscompleted INTEGER,
dev(#     certificationsattained INTEGER,
dev(#     peerreviewscore REAL,
dev(#     managerreviewscore REAL,
dev(#     overallperformancescore REAL,
dev(#     PRIMARY KEY (employeeid)
dev(# );
CREATE TABLE
dev=#
```

Step 7: The table has been created successfully. Now the next step is to load the data present in your S3 bucket into the table you have created in the redshift cluster. This is done by using the copy command.



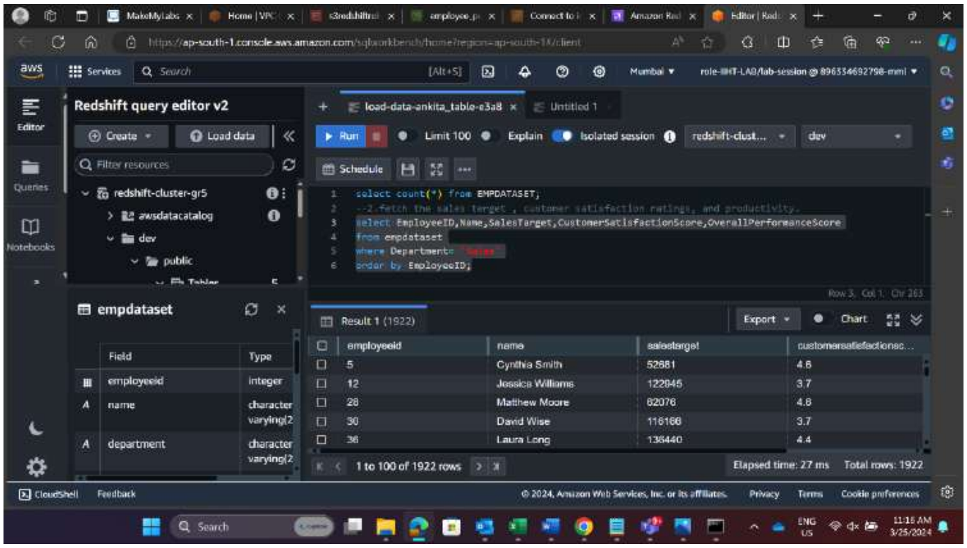
```
dev=# COPY EMPDATASET FROM 's3://group-5bucket/employee_performance_data (1).csv'
dev-#     IAM_ROLE 'arn:aws:iam::896334692798:role/s3redshiftrole'
dev-#     CSV
dev-#     IGNOREHEADER 1;
INFO:  Load into table 'empdataset' completed, 10000 record(s) loaded successfully.
COPY
dev=#
```

The data has been successfully loaded into the table.

Redshift Queries Generation and Visualization

1. Fetch the sales target, customer satisfaction ratings, and productivity for all the employees of the sales department.

a. Query execution.



```
select count(*) from EMPDATASET;
--2.fetch the sales target , customer satisfaction ratings, and productivity.
select EmployeeID,Name,SalesTarget,CustomerSatisfactionScore,OverallPerformanceScore
from empdataset
where Department= 'Sales'
order by EmployeeID;
```

empdataset

Field	Type
employeeid	integer
name	character varying(2
department	character varying(2

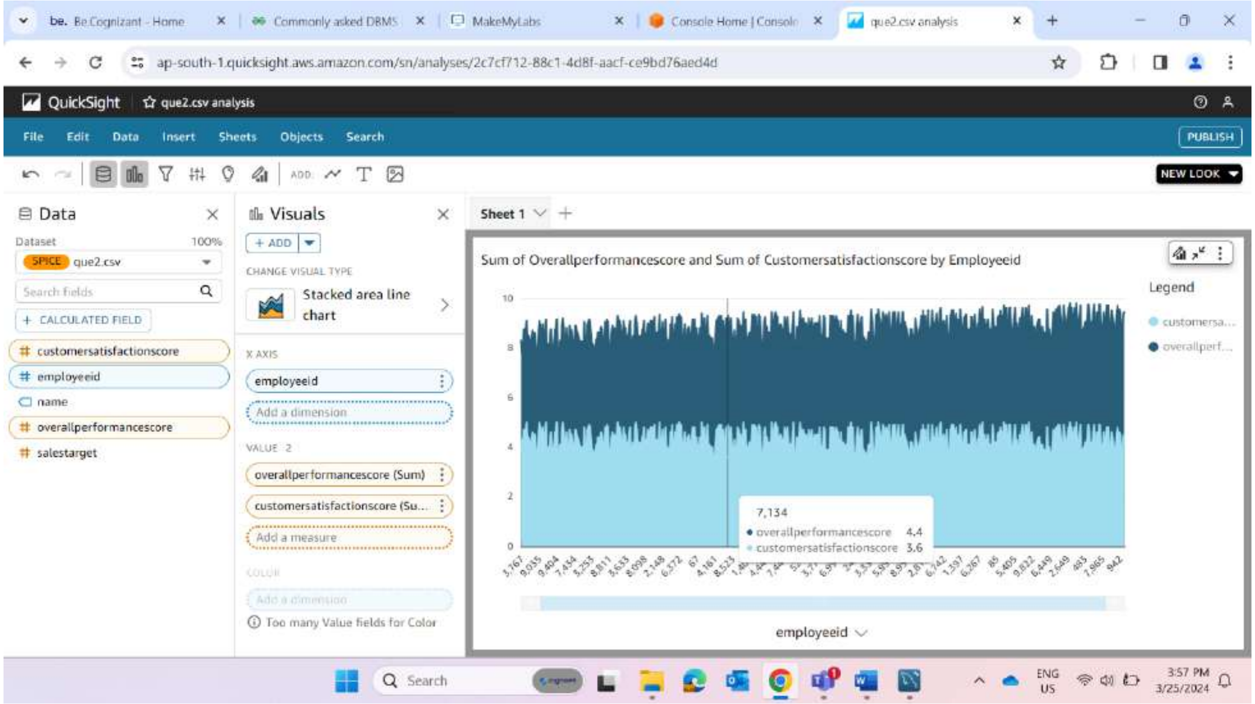
Result 1 (1922)

employeeid	name	salestarget	customersatisfactionsc...
5	Cynthia Smith	52881	4.8
12	Jessica Williams	122845	3.7
28	Matthew Moore	82076	4.8
30	David Wise	116166	3.7
36	Laura Long	136440	4.4

1 to 100 of 1922 rows

Elapsed time: 27 ms Total rows: 1922

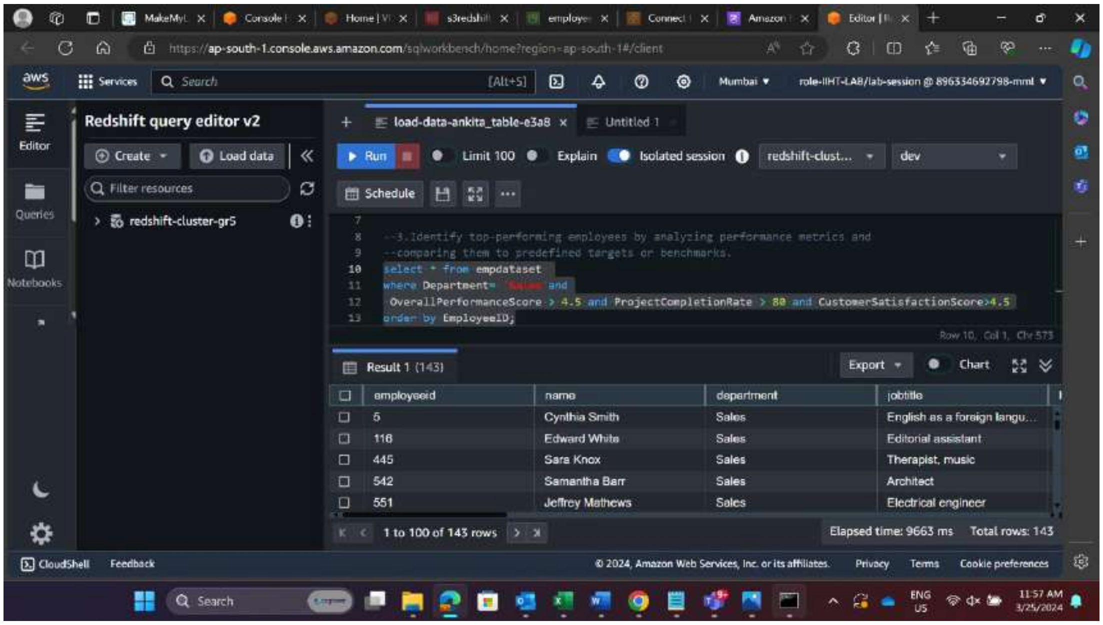
b. Visualization



Sum of Overallperformancescore and Sum of Customersatisfactionscore by Employeeid

2. Identify top-performing employees by analyzing performance metrics and comparing them to predefined targets or benchmarks.

a. Query Execution



```
--3.Identify top-performing employees by analyzing performance metrics and
--comparing them to predefined targets or benchmarks.
select * from empdataset
where Department= 'Sales' and
 OverallPerformanceScore > 4.5 and ProjectCompletionRate > 80 and CustomerSatisfactionScore>4.5
order by EmployeeID;
```

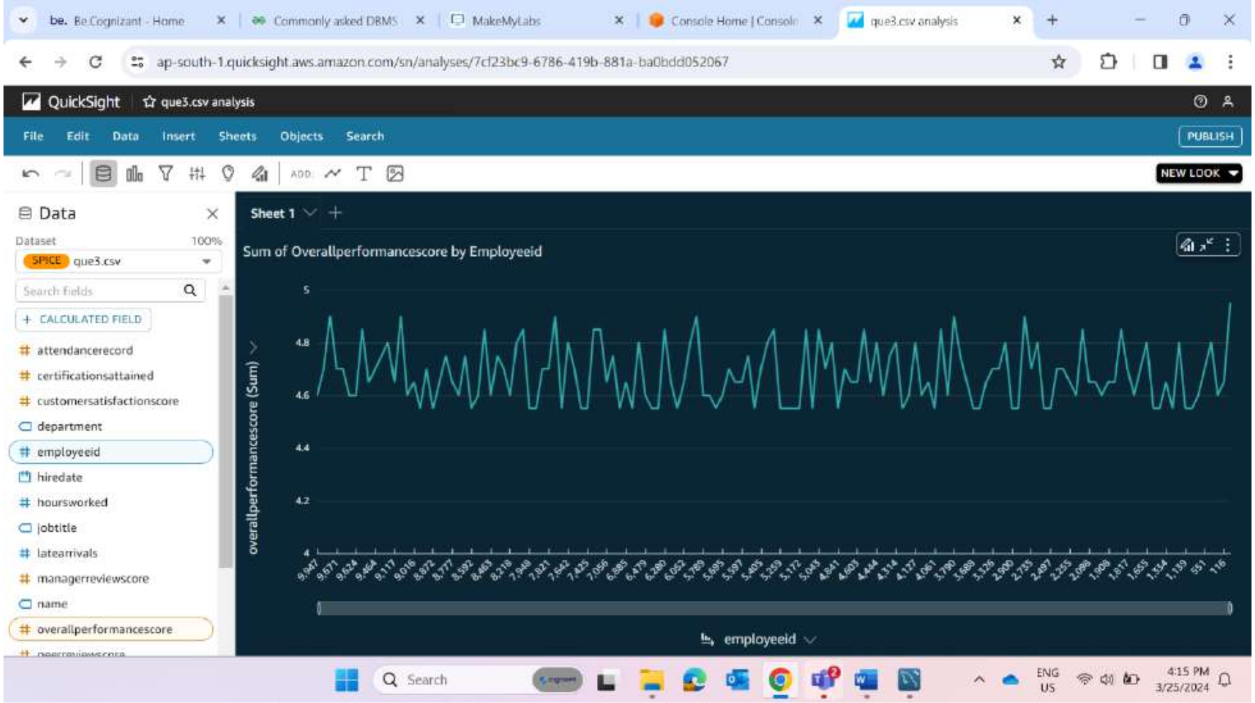
Row 10, Col 1, Chr 573

employeeid	name	department	jobtitle
5	Cynthia Smith	Sales	English as a foreign langu...
116	Edward White	Sales	Editorial assistant
445	Sara Knox	Sales	Therapist, music
542	Samantha Barr	Sales	Architect
551	Jeffrey Mathews	Sales	Electrical engineer

1 to 100 of 143 rows

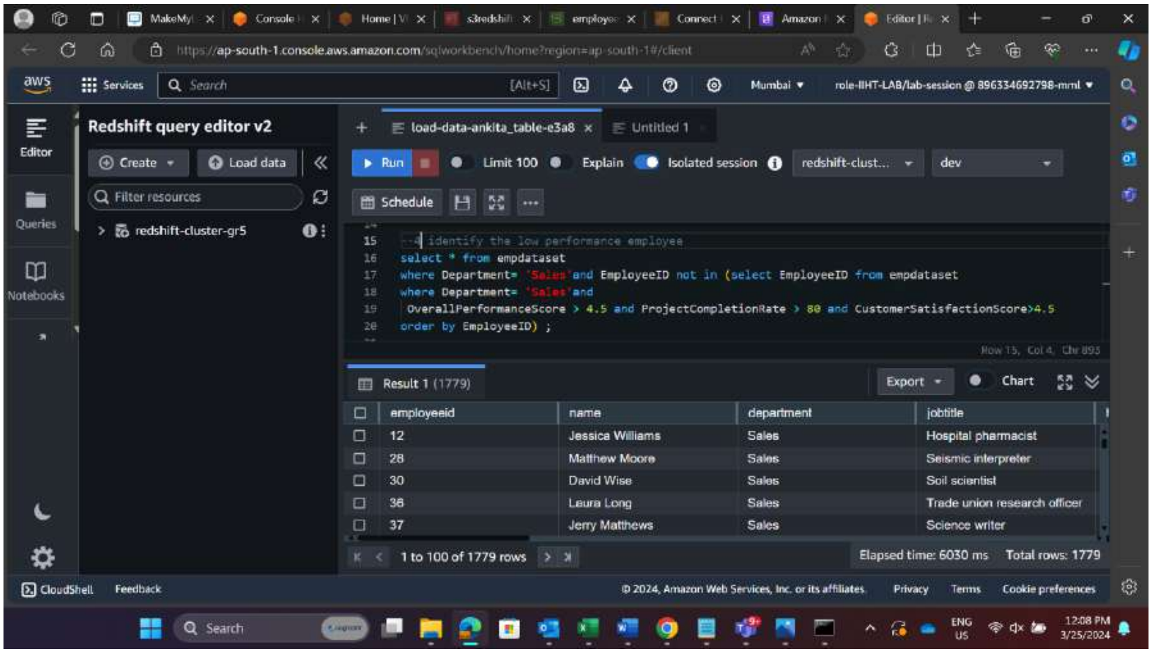
Elapsed time: 9663 ms Total rows: 143

b. Visualization



3. Identify the employees having low performance in the sales department.

a. Query Execution



Redshift query editor v2

```
--4 identify the low performance employee
select * from empdataset
where Department= 'Sales'and EmployeeID not in (select EmployeeID from empdataset
where Department= 'Sales'and
 OverallPerformanceScore > 4.5 and ProjectCompletionRate > 80 and CustomerSatisfactionScore>4.5
order by EmployeeID) ;
```

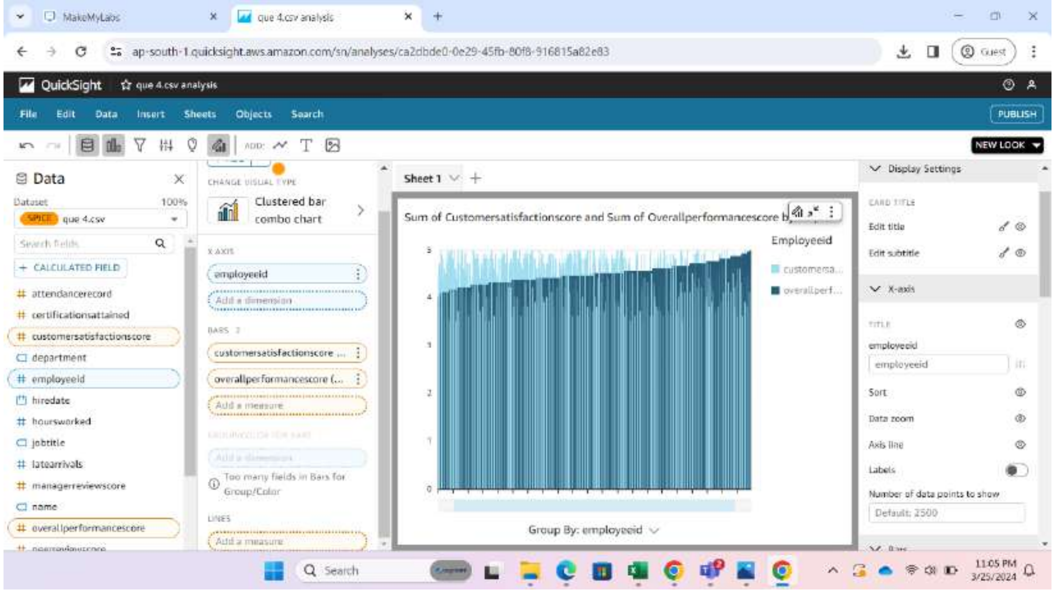
Result 1 (1779)

employeeid	name	department	jobtitle
12	Jessica Williams	Sales	Hospital pharmacist
28	Matthew Moore	Sales	Seismic interpreter
30	David Wise	Sales	Soil scientist
36	Laura Long	Sales	Trade union research officer
37	Jerry Matthews	Sales	Science writer

1 to 100 of 1779 rows

Elapsed time: 6030 ms   Total rows: 1779

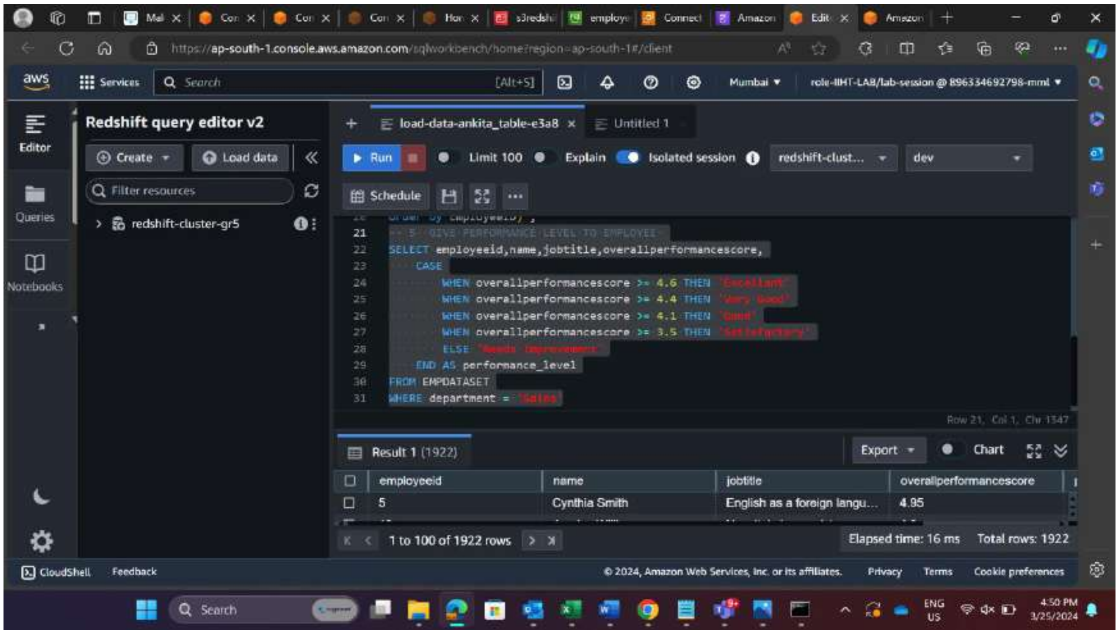
b. Visualization



Sum of Customersatisfactionscore and Sum of Overallperformancescore by Employeeid

4. Give a performance value to all the employees of the sales department based on some specified criteria.

a. Query Execution



```
-- 5. GIVE PERFORMANCE LEVEL TO EMPLOYEE
SELECT employeeid,name,jobtitle,overallperformancescore,
    CASE
        WHEN overallperformancescore >= 4.6 THEN 'Excellent'
        WHEN overallperformancescore >= 4.4 THEN 'Very Good'
        WHEN overallperformancescore >= 4.1 THEN 'Good'
        WHEN overallperformancescore >= 3.5 THEN 'Satisfactory'
        ELSE 'Needs Improvement'
    END AS performance_level
FROM EMPDATASET
WHERE department = 'Sales'
```

Result 1 (1922)

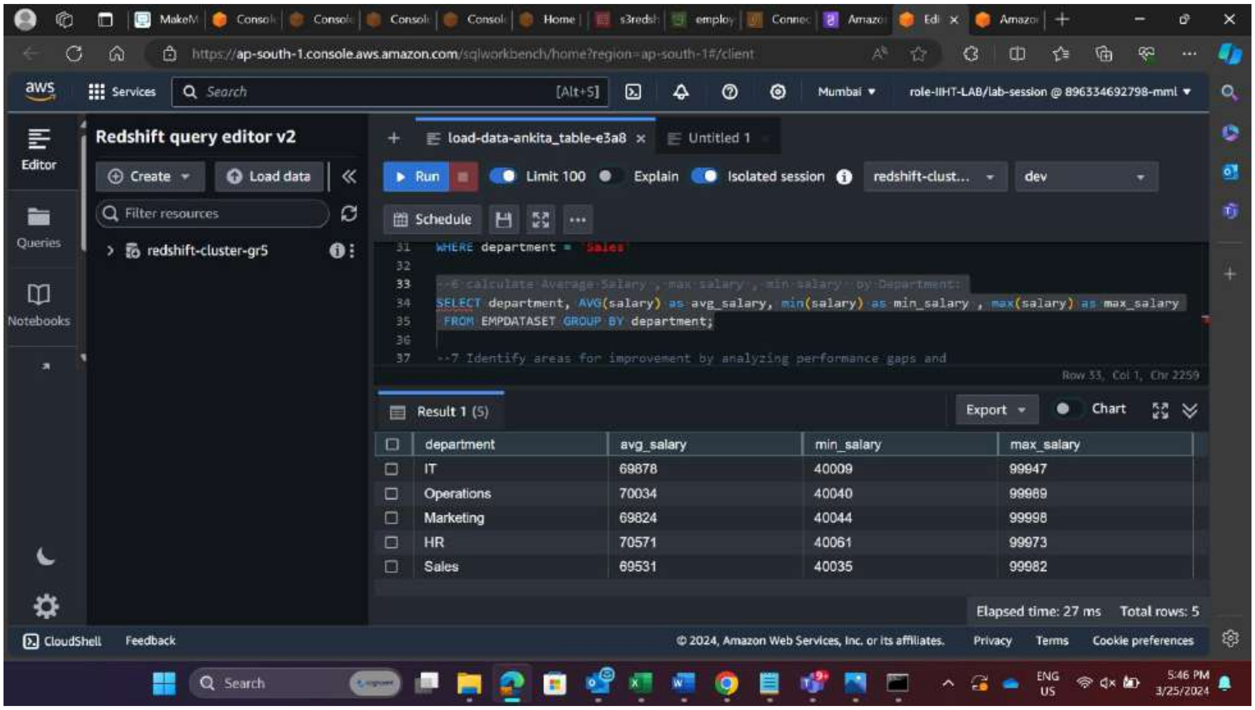
	employeeid	name	jobtitle	overallperformancescore
□	5	Cynthia Smith	English as a foreign langu...	4.95

1 to 100 of 1922 rows

Elapsed time: 16 ms Total rows: 1922

5. Calculate the average, minimum and maximum salary of each department in the employee table.

a. Query Execution



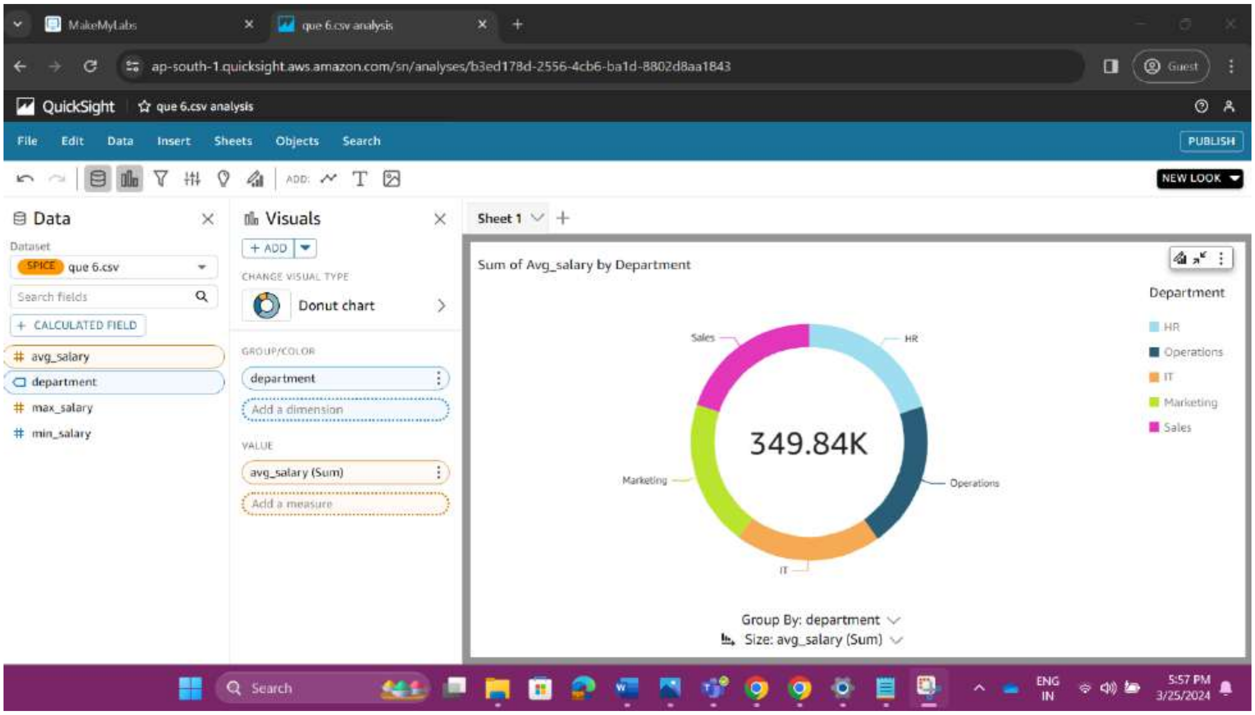
```
WHERE department = 'Sales'

--6 calculate Average Salary , max salary , min salary  by Department:
SELECT department, AVG(salary) as avg_salary, min(salary) as min_salary , max(salary) as max_salary
 FROM EMPDATASET GROUP BY department;

--7 Identify areas for improvement by analyzing performance gaps and
```

Result 1 (5)

department	avg_salary	min_salary	max_salary
IT	69878	40009	99947
Operations	70034	40040	99989
Marketing	69824	40044	99998
HR	70571	40061	99973
Sales	69531	40035	99982

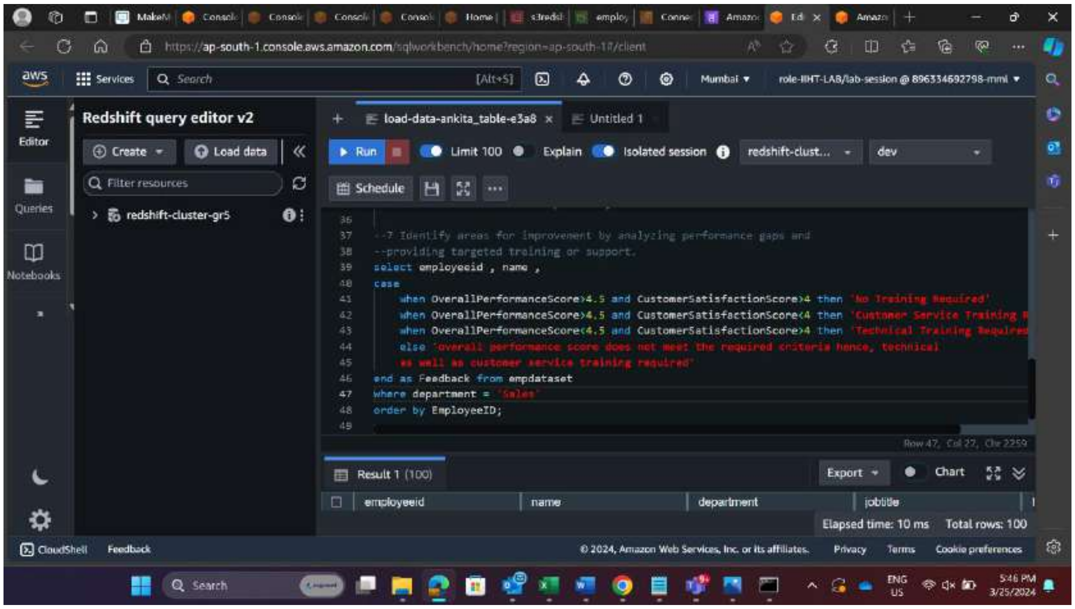
Elapsed time: 27 ms Total rows: 5

b. Visualization



6. Identify areas for improvement by analyzing performance gaps and providing targeted training or support.

a. Query Execution



```
--7 Identify areas for improvement by analyzing performance gaps and
--providing targeted training or support.
select employeeid , name ,
case
    when OverallPerformanceScore>4.5 and CustomerSatisfactionScore>4 then 'No Training Required'
    when OverallPerformanceScore>4.5 and CustomerSatisfactionScore<4 then 'Customer Service Training R
    when OverallPerformanceScore<4.5 and CustomerSatisfactionScore>4 then 'Technical Training Required
    else 'overall performance score does not meet the required criteria hence, technical
    as well as customer service training required'
end as Feedback from empdataset
where department = 'Sales'
order by EmployeeID;
```

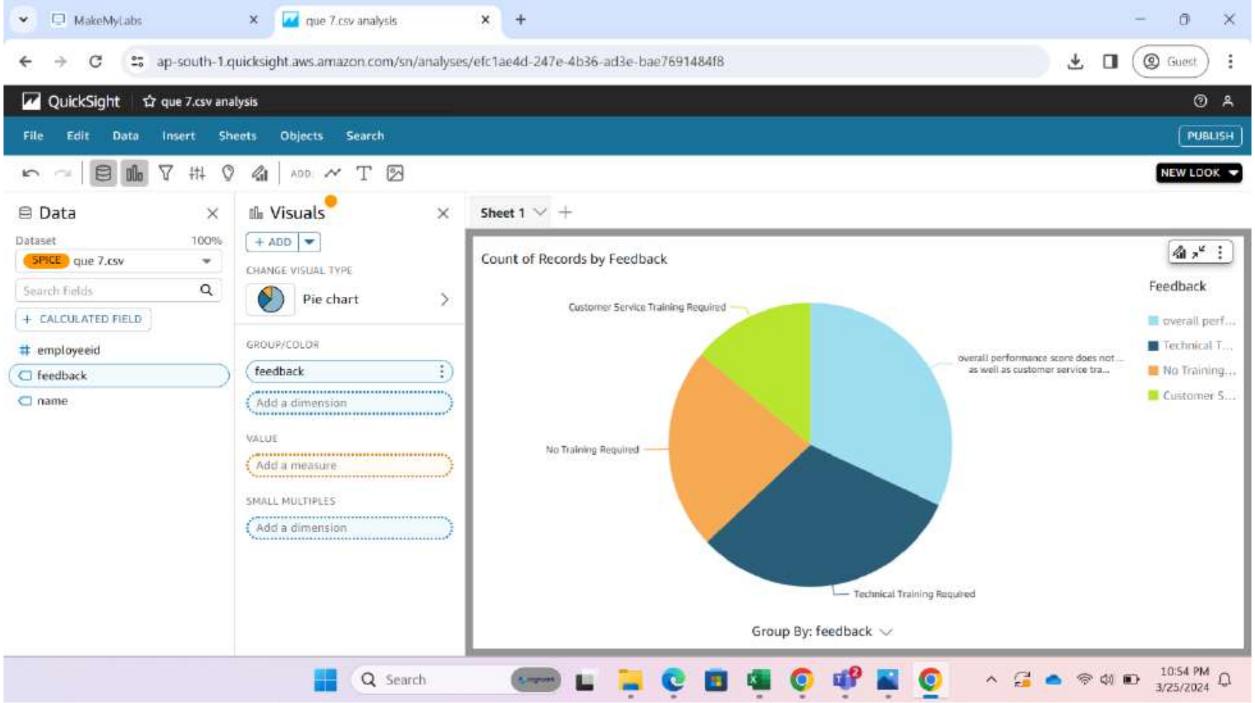
Row 47, Col 22, Chr 2259

Result 1 (100)

employeeid	name	department	jobtitle

Elapsed time: 10 ms   Total rows: 100

b. Visualization



Dataset: SPICE que 7.csv (100%)

CHANGE VISUAL TYPE: Pie chart

GROUP/COLOR: feedback

VALUE: Add a measure

SMALL MULTIPLES: Add a dimension

CONCLUSION

In conclusion, the project exemplifies the seamless integration and powerful capabilities of AWS services in constructing a robust data infrastructure. By harnessing Amazon S3 as a scalable storage solution, we ensured the efficient and secure handling of our datasets. The coupling of S3 with Amazon Redshift allowed for the establishment of a high-performance data warehousing environment, facilitating quick access and analysis of our structured data. Moreover, the utilization of Amazon EC2 instances enabled us to execute complex SQL operations, ensuring the manipulation and transformation of data according to our analytical needs.

Furthermore, the deployment of AWS Quicksight for visualization added a layer of insight to our data exploration process. With Quicksight's intuitive interface and comprehensive visualization options, we were able to craft informative reports and dashboards that provided stakeholders with clear and actionable insights. This not only enhanced our understanding of the underlying data but also empowered decision-makers to make informed choices based on the analyzed information.

In summary, the successful execution of this project underscores the value proposition of AWS as a leading cloud provider for data analytics initiatives. The combination of S3, Redshift, EC2, and Quicksight offered a cohesive and scalable solution that streamlined the entire data processing pipeline. As organizations continue to grapple with increasing volumes of data, leveraging AWS services provides a reliable and efficient means to extract actionable insights and drive business outcomes. Moving forward, the lessons learned from this project will serve as a foundation for future endeavors in harnessing the full potential of cloud-based data analytics solutions.