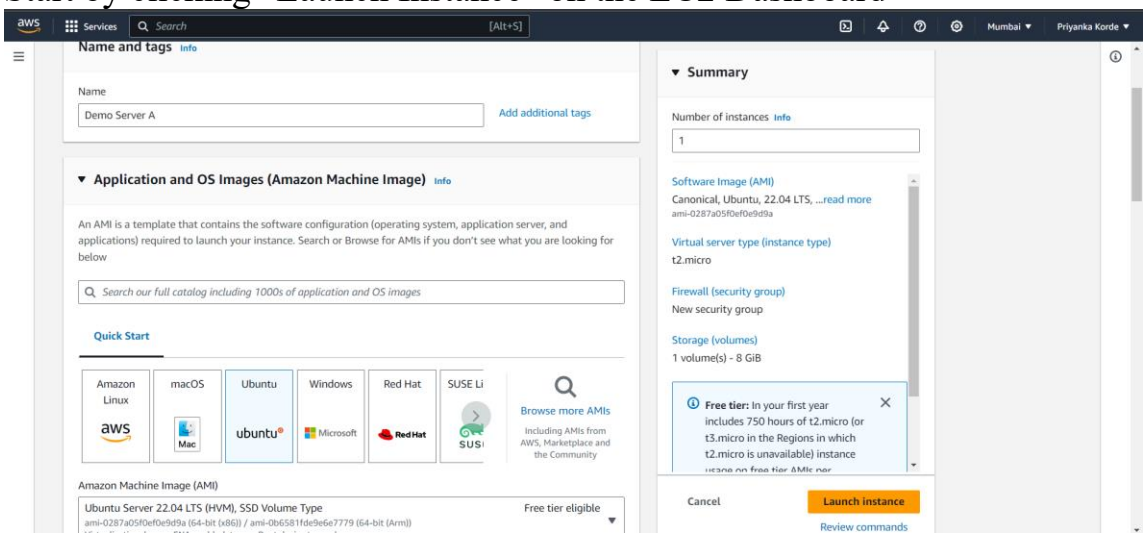
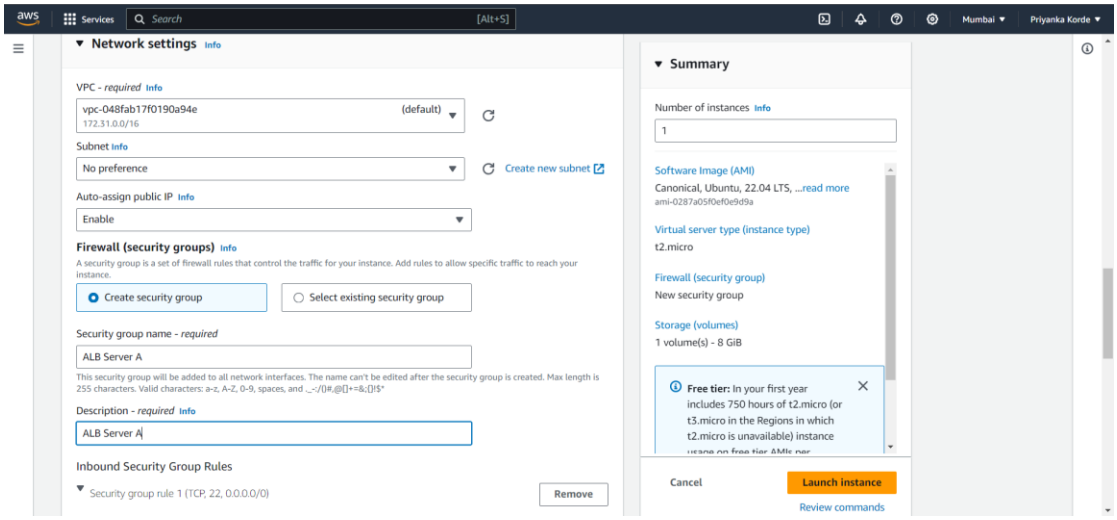
Priyanka Korde

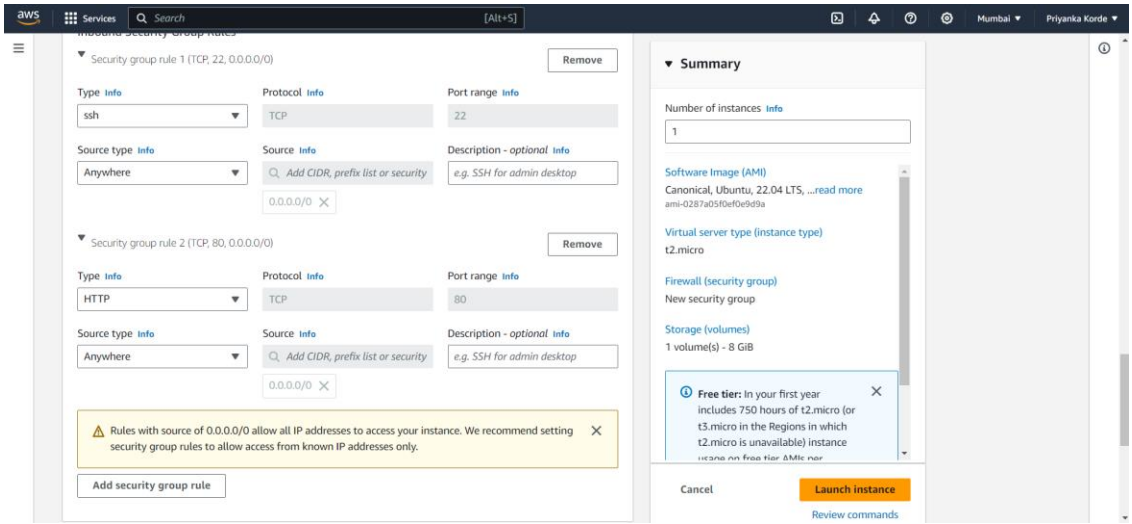
AWS Athena S3 EC2 ALB analysis for Accessing Logs
Analyzing ALB Access Logs with Amazon Athena

Step 1: Launch EC2 instance

Start by clicking “Launch Instance” on the EC2 Dashboard



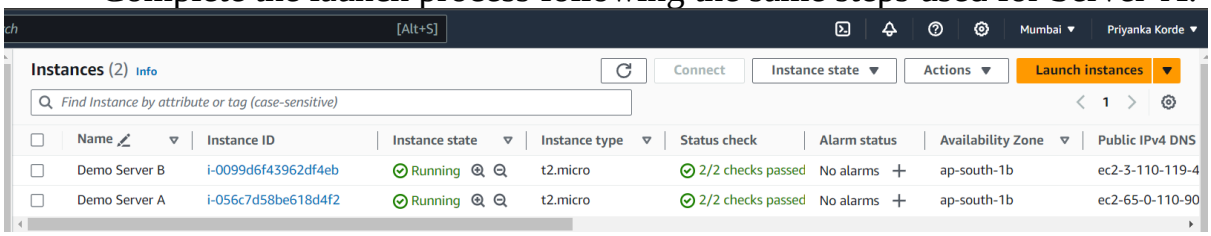




Finally, select **"launch instance"** to create the server.

To launch another instance identical to Server-A and name it Server-B:

- Create a new EC2 instance using the same configuration as Server-A.
- During the setup, assign the name "Server-B" to the instance.
- Complete the launch process following the same steps used for Server-A.



Connect the both instance and run commands:

sudo apt-get update
sudo apt-get upgrade
sudo apt-get install apache2
cd var/www/html
sudo rm index.html
sudo vi html.html

Put simple html code inside index file

Server A: <html> <h1>Hello, This is Priyanka Korde from Server A !</h1> </html>

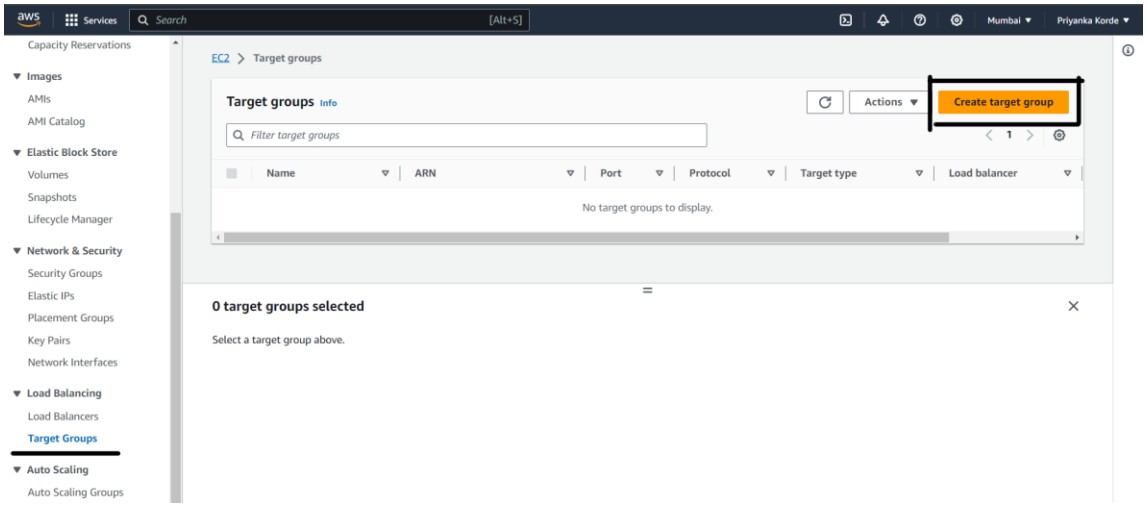
Server B: <html> <h1>Hello, This is Priyanka Korde from Server B !</h1> </html>

Open your browser and paste the public IPv4 address of the Server-A

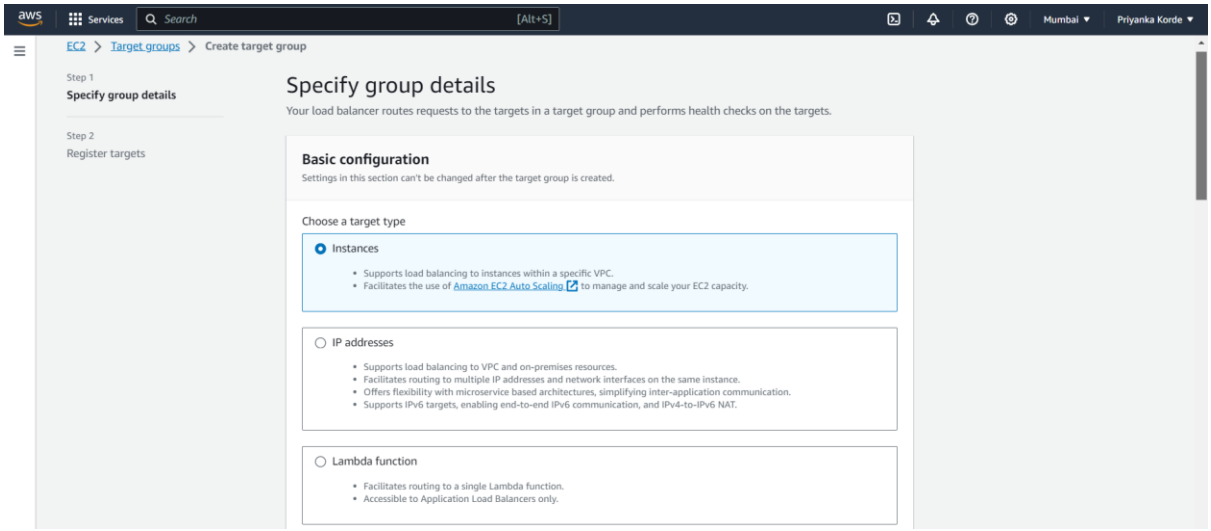
Also verify that Server B is also functioning correctly.

Step 2: **Create Target Group**

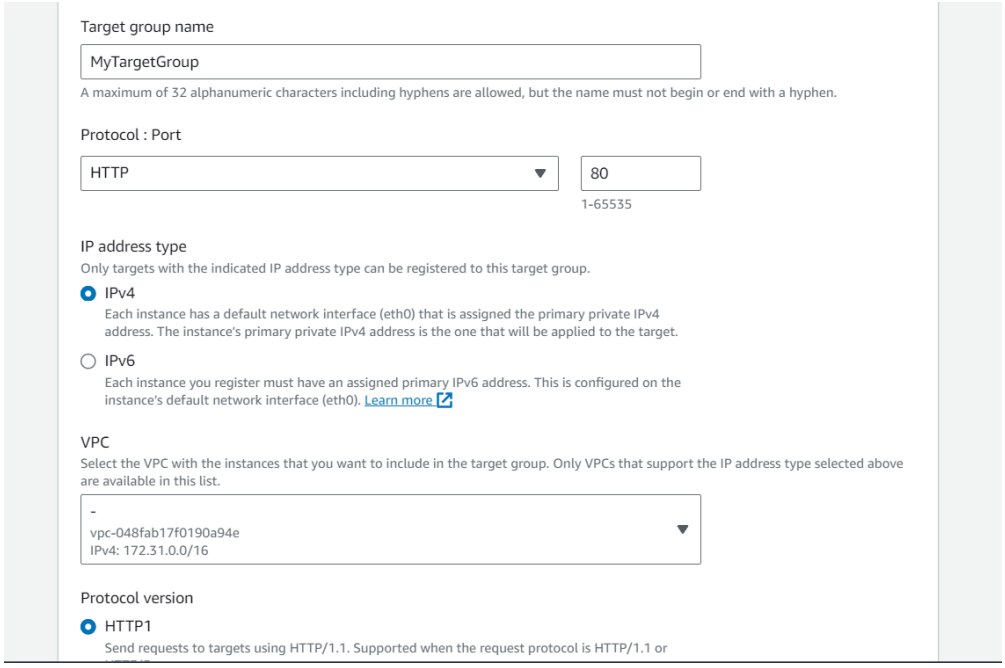
In the EC2 Dashboard, locate the "Load Balancing" section in the left-hand navigation pane and select "Target Groups."



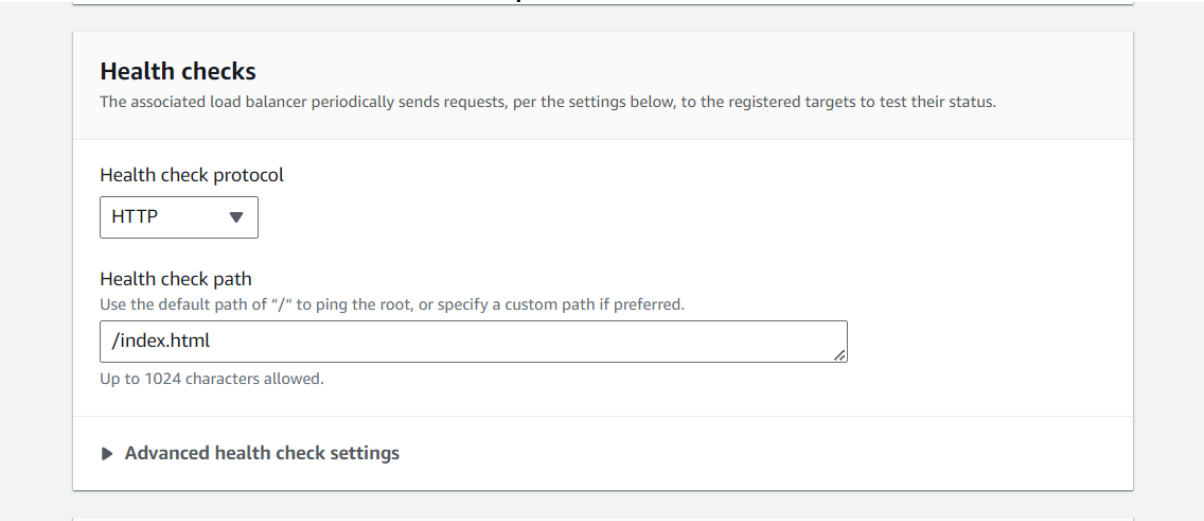
Click the "Create target group" button to start the target group creation process , ensure that you have selected the instances as **target type**



- **Name**: Enter a descriptive name for your target group.
- **Protocol**: Choose the protocol (HTTP, HTTPS, etc.) based on your application's needs.
- **Port**: Specify the port that your instances are listening on.
- **VPC**: Choose the Virtual Private Cloud (VPC) where your instances are located.



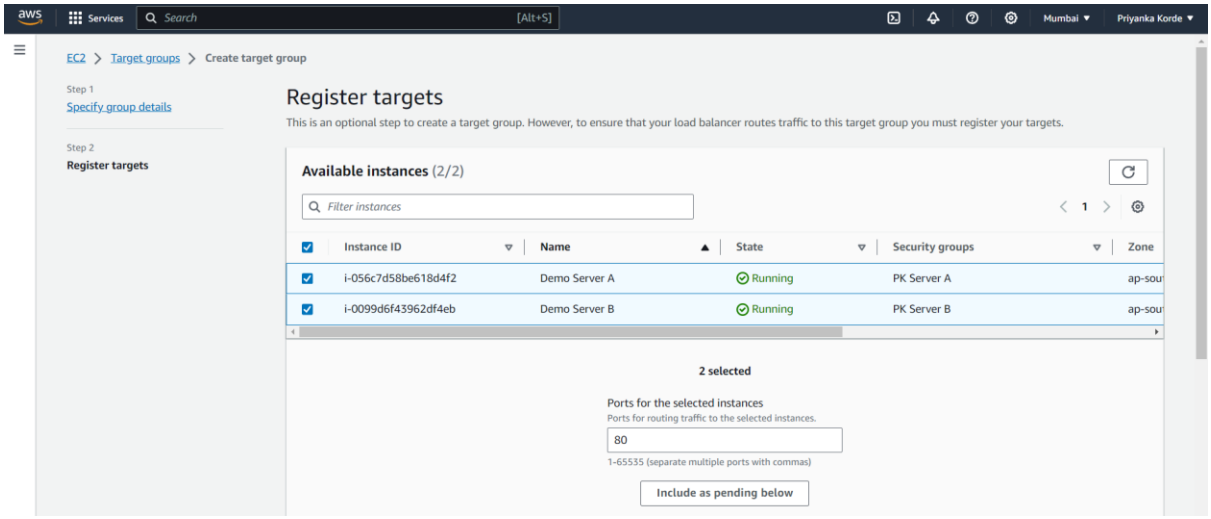
- **Health Checks**: Define health check settings, including the path as **index.html** health check protocol and then click on next



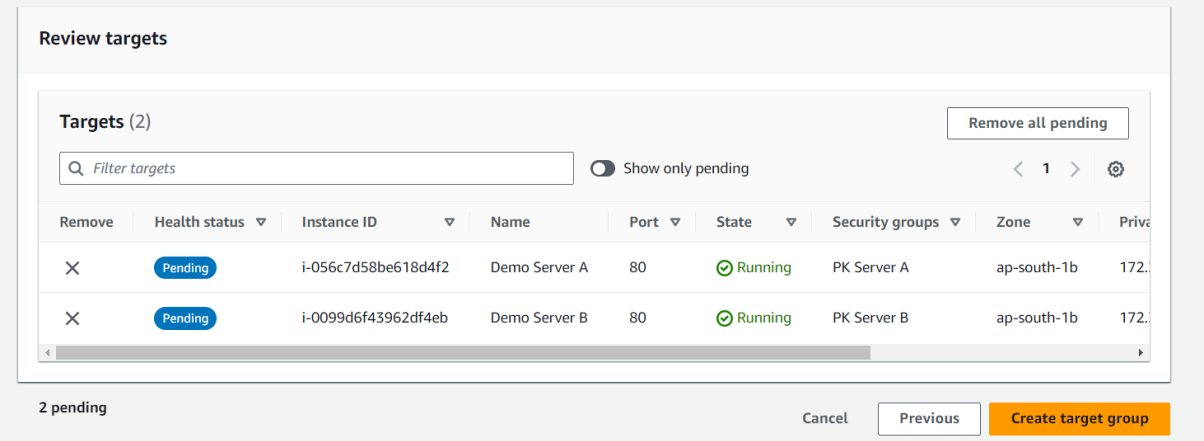
Register Targets

Under the "Targets" section, you'll need to register instances that should be part of this target group.

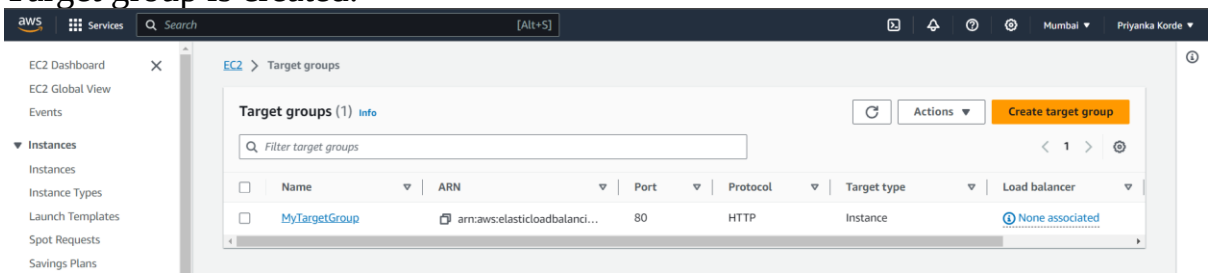
Select the instances by their instance ID or IP address and click on **include as pending below**



Click on create target group

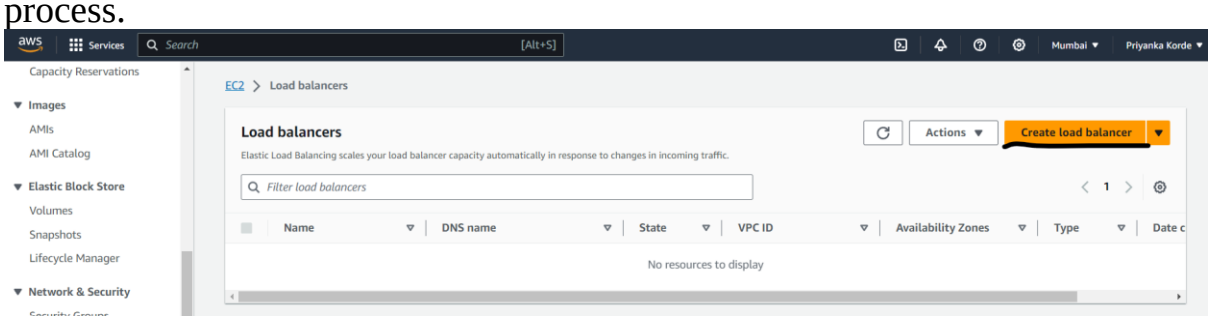


Target group is created.

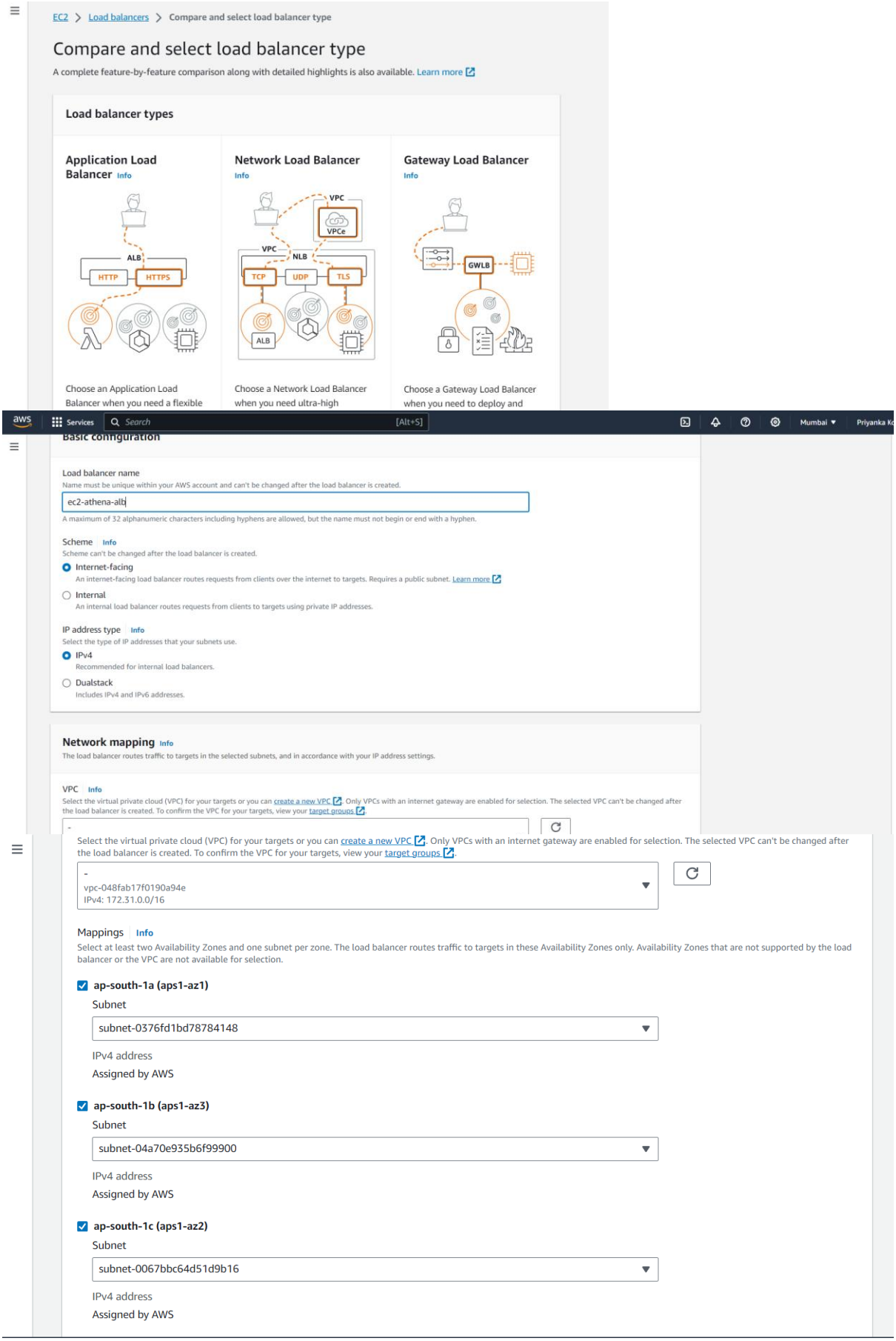


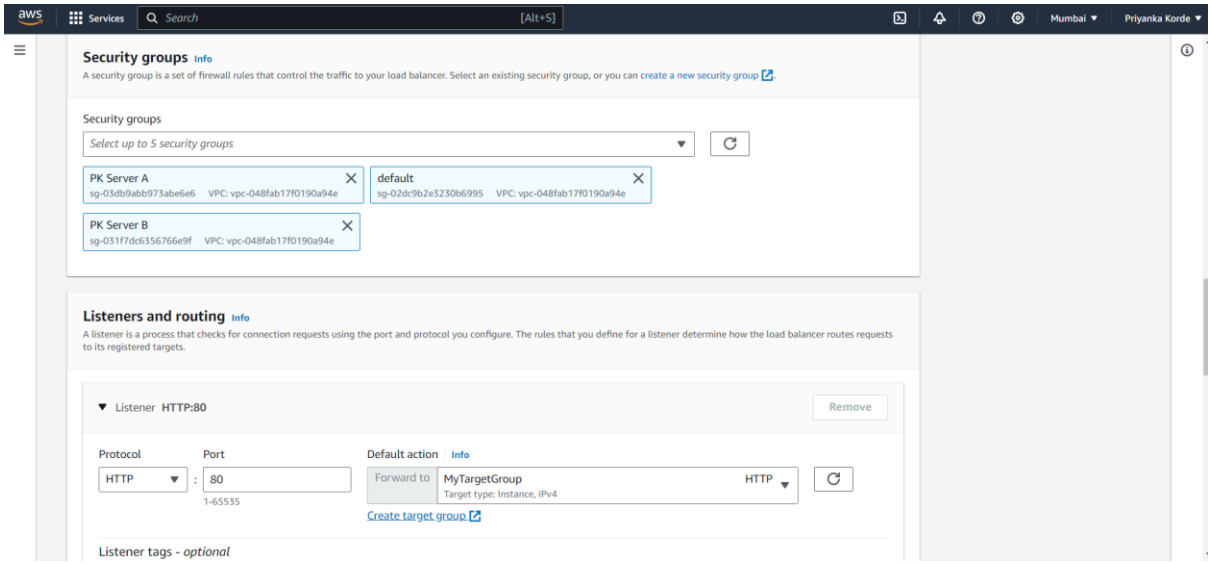
Step 3 : Create a Load Balancer

Click the “Create Load Balancer” button to start the load balancer creation process.

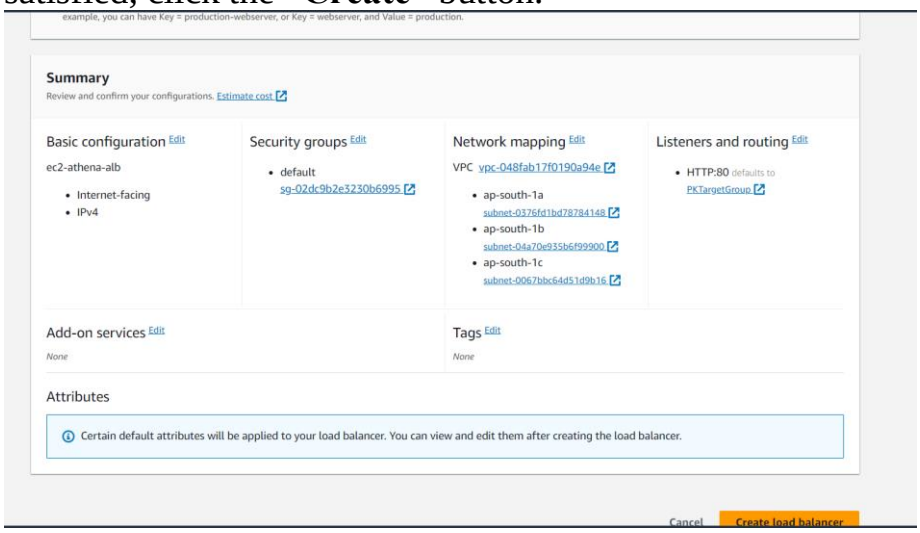


Select the type of load balancer you want to create:
Choose the appropriate type for your use case and click “Create.”


EC2 > Load balancers > Compare and select load balancer type
Compare and select load balancer type
A complete feature-by-feature comparison along with detailed highlights is also available. Learn more
Load balancer types
Application Load Balancer Info
Network Load Balancer Info
Gateway Load Balancer Info
ALB
HTTP
HTTPS
VPC
VPCe
NLB
TCP
UDP
TLS
GWLB
Choose an Application Load Balancer when you need a flexible
Choose a Network Load Balancer when you need ultra-high
Choose a Gateway Load Balancer when you need to deploy and
Basic configuration
Load balancer name
Name must be unique within your AWS account and can't be changed after the load balancer is created.
ec2-athena-alb
A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.
Scheme Info
Scheme can't be changed after the load balancer is created.
Internet-facing
An internet-facing load balancer routes requests from clients over the internet to targets. Requires a public subnet. Learn more
Internal
An internal load balancer routes requests from clients to targets using private IP addresses.
IP address type Info
Select the type of IP addresses that your subnets use.
IPv4
Recommended for internal load balancers.
Dualstack
Includes IPv4 and IPv6 addresses.
Network mapping Info
The load balancer routes traffic to targets in the selected subnets, and in accordance with your IP address settings.
VPC Info
Select the virtual private cloud (VPC) for your targets or you can create a new VPC. Only VPCs with an internet gateway are enabled for selection. The selected VPC can't be changed after the load balancer is created. To confirm the VPC for your targets, view your target groups.
-
vpc-048fab17f0190a94e
IPv4: 172.31.0.0/16
Mappings Info
Select at least two Availability Zones and one subnet per zone. The load balancer routes traffic to targets in these Availability Zones only. Availability Zones that are not supported by the load balancer or the VPC are not available for selection.
ap-south-1a (aps1-az1)
Subnet
subnet-0376fd1bd78784148
IPv4 address
Assigned by AWS
ap-south-1b (aps1-az3)
Subnet
subnet-04a70e935b6f99900
IPv4 address
Assigned by AWS
ap-south-1c (aps1-az2)
Subnet
subnet-0067bbc64d51d9b16
IPv4 address
Assigned by AWS

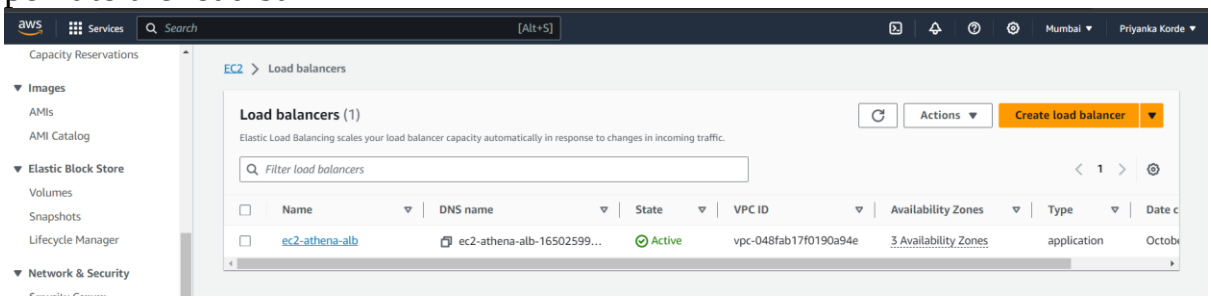


Review the configuration details to ensure they are correct. Once you're satisfied, click the **"Create"** button.



You should see a confirmation message indicating that your load balancer has been created successfully.

Your load balancer is now ready to distribute traffic across the registered instances in the target group, providing high availability and scalability for your application. Make sure to update your DNS or application configurations to point to the load balancer's DNS name or IP address.



Upon refreshing the page, you will receive a response from Server A and B, demonstrating that the load balancer successfully directs traffic to both EC2

instances, ensuring redundancy and optimal load distribution.

Hello, This is Priyanka Korde from Server A !

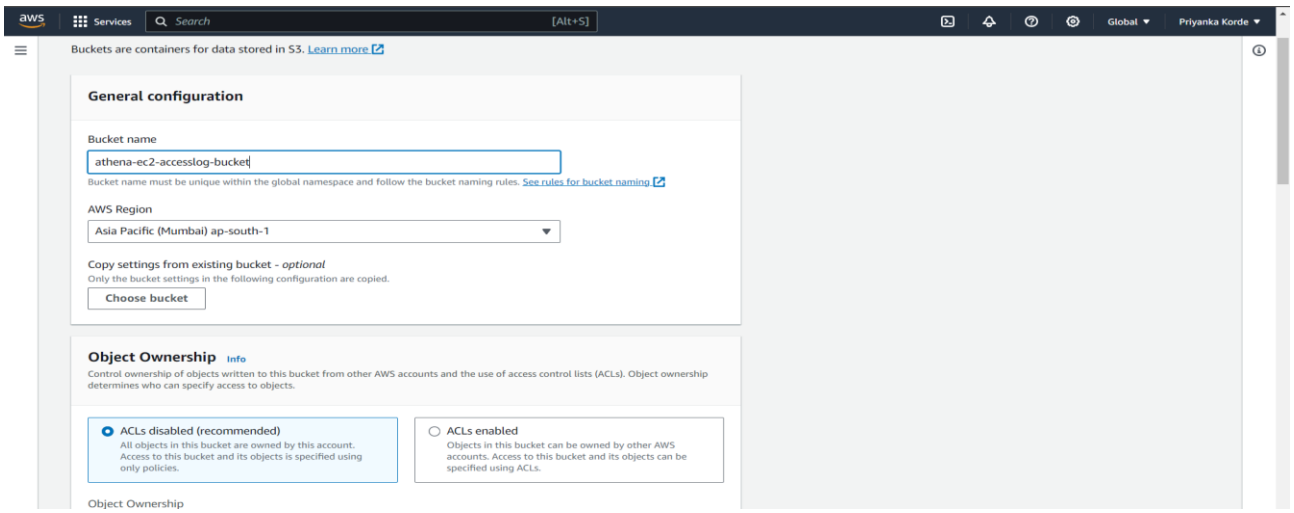
Hello, This is Priyanka Korde from Server B !

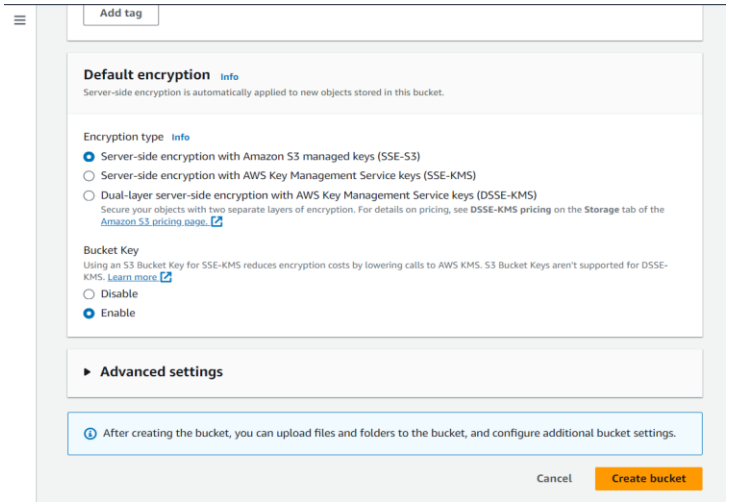
Step 3 : Configure Access Logs

To enabling access logs for your Application Load Balancer (ALB), you'll also need to create an Amazon S3 bucket and set up a policy to allow the ALB to write logs to the bucket.

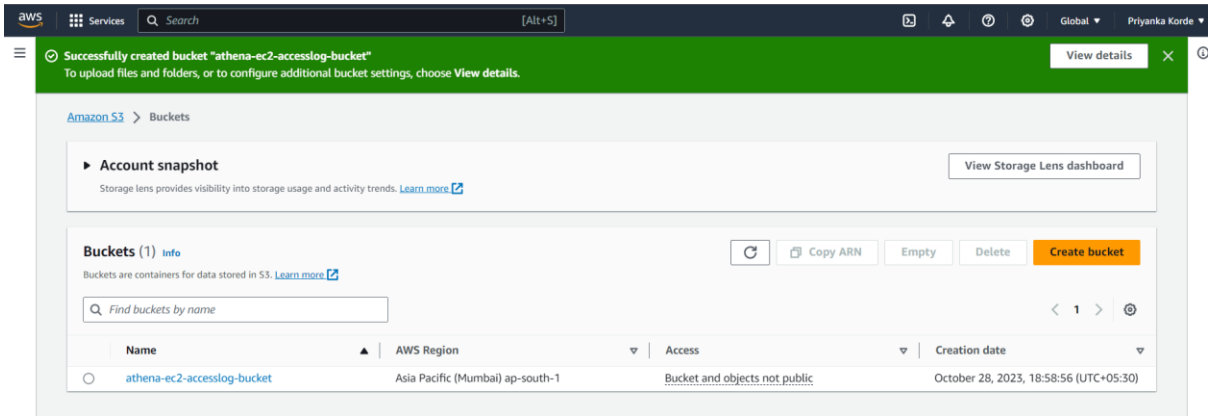
Navigate to the S3 service

Create the bucket





Review your bucket settings, and if everything looks correct, click the “Create bucket” button.



The bucket creation process has been completed successfully.
Once the bucket is created create 3 folders /prefix /Awslogs/ (youraccountid) for
logs.

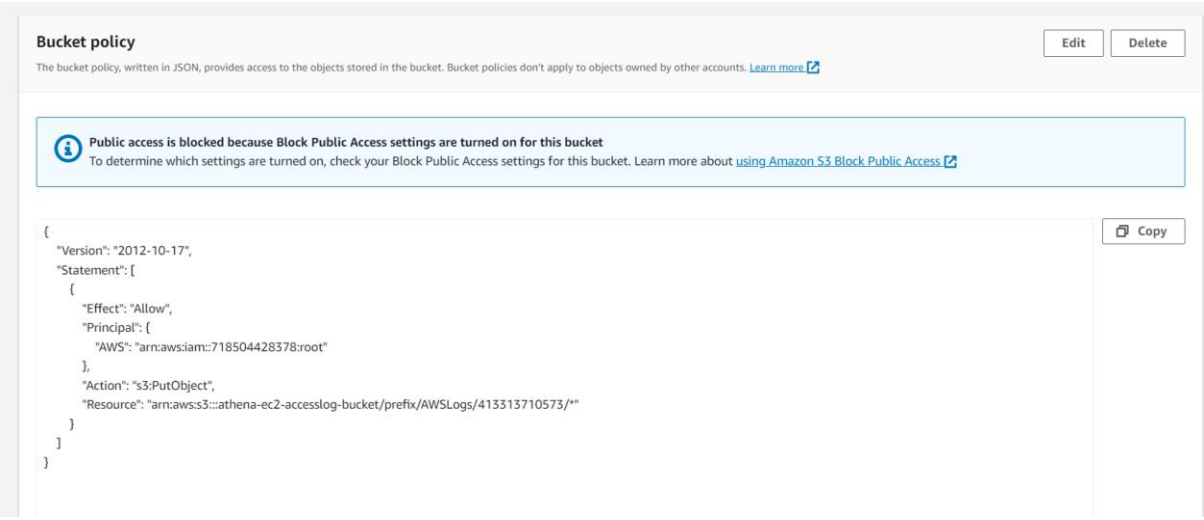
Step 4: Set Up a Bucket Policy

After creating the S3 bucket, you’ll need to set up a bucket policy to allow the ALB to write logs to it:
In the bucket policy editor, you’ll need to define a policy that grants the necessary permissions to the ALB.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::account-id:root"
      },
      "Action": "s3:PutObject",
      "Resource": "my-s3-arn"
```

```
    }
  ]
}
```

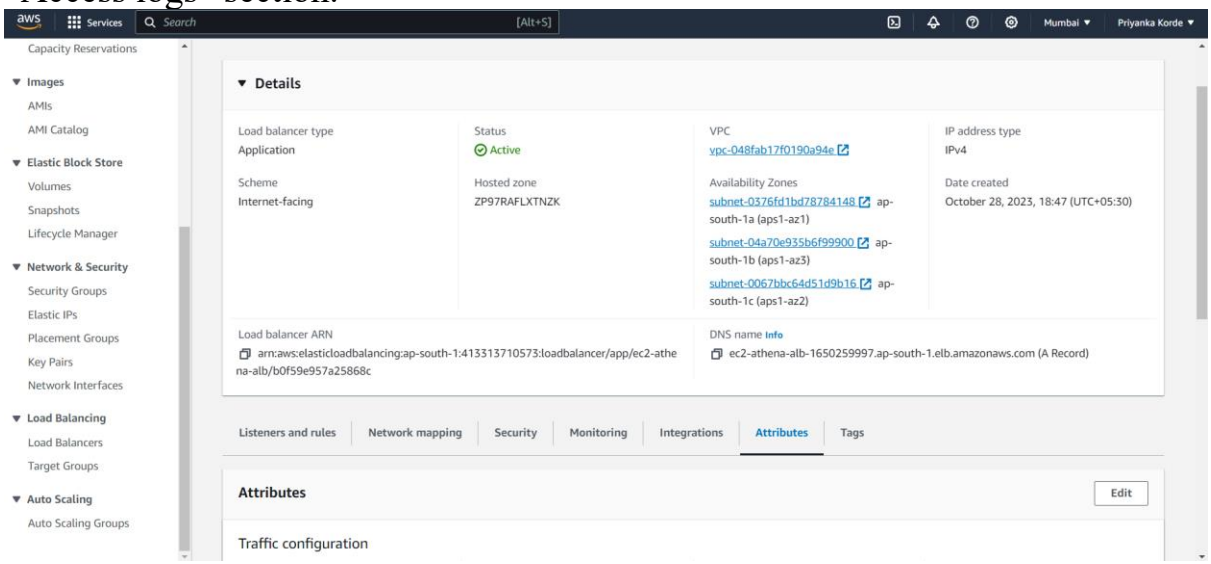
- Replace account-id with the ID of the AWS account for Elastic Load Balancing for your Region: Asia Pacific (Mumbai) – 718504428378
- Replace *my-s3-arn* with the ARN of the location for your access logs. arn:aws:s3:::*bucket-name*/*prefix*/AWSLogs/*aws-account-id*/*



Step 5: Enabling logs for an Application Load Balancer

Click on the name of your Application Load Balancer (ALB) for which you want to enable access logs.

Inside the ALB details page, navigate to the "Attributes" tab and find the "Access logs" section.

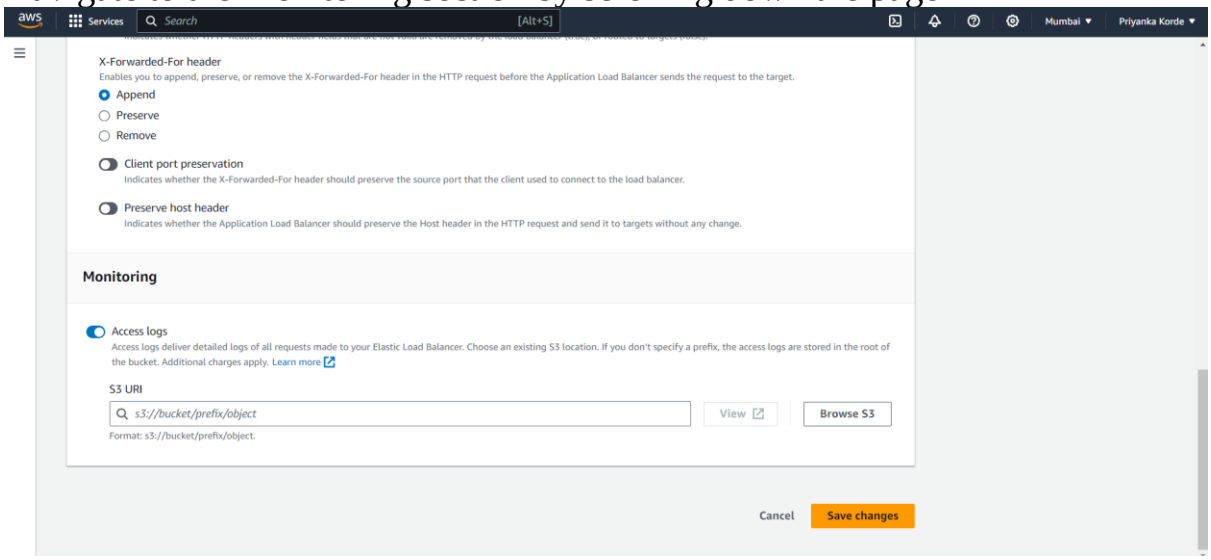


Click on "Edit"

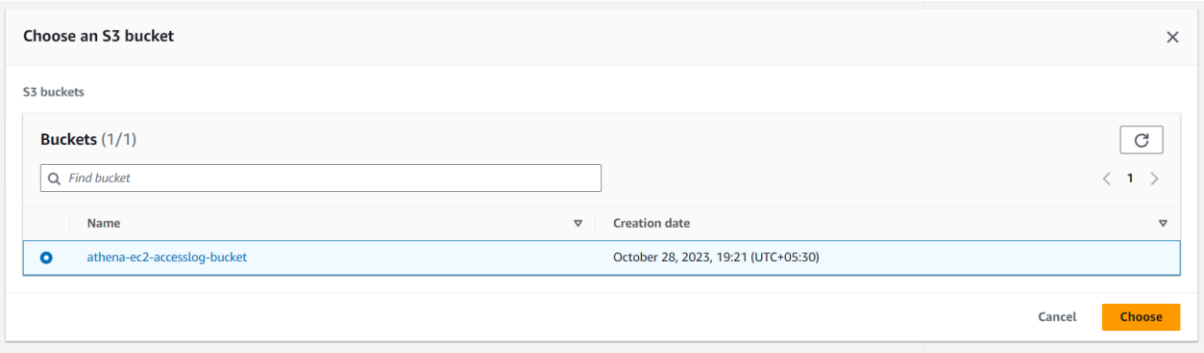
Click the "Edit" button to configure access logs.

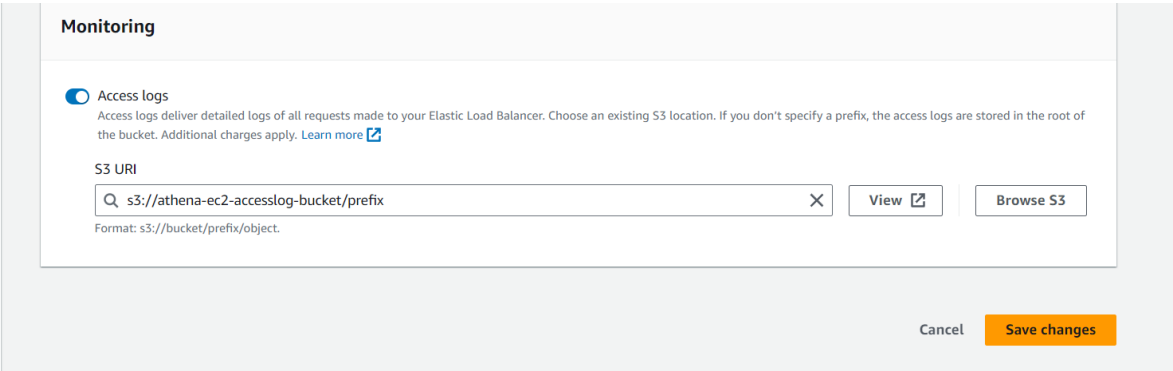
Enable Access Logs

In the “Access logs” section, do the following:
Navigate to the Monitoring section by scrolling down the page

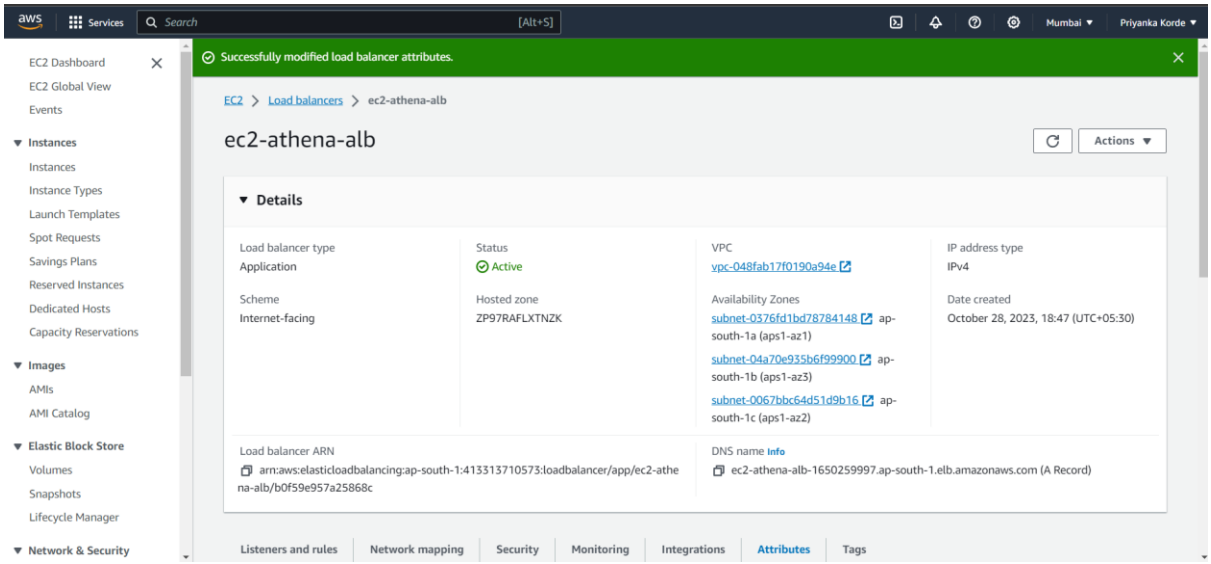


Specify the S3 bucket where you want to store the access logs. You can choose an existing bucket or create a new one.



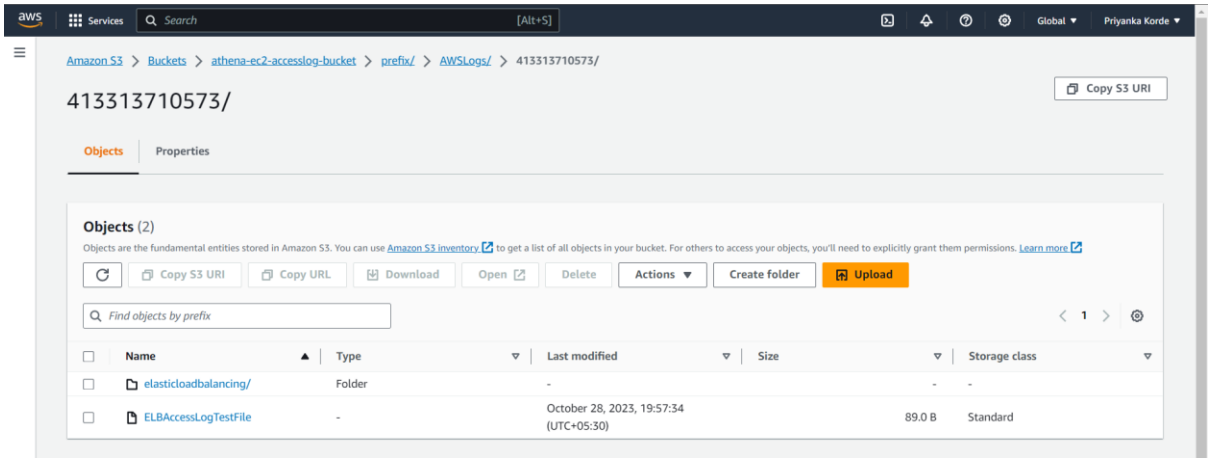


Click the “Save” button to save your access log configuration.

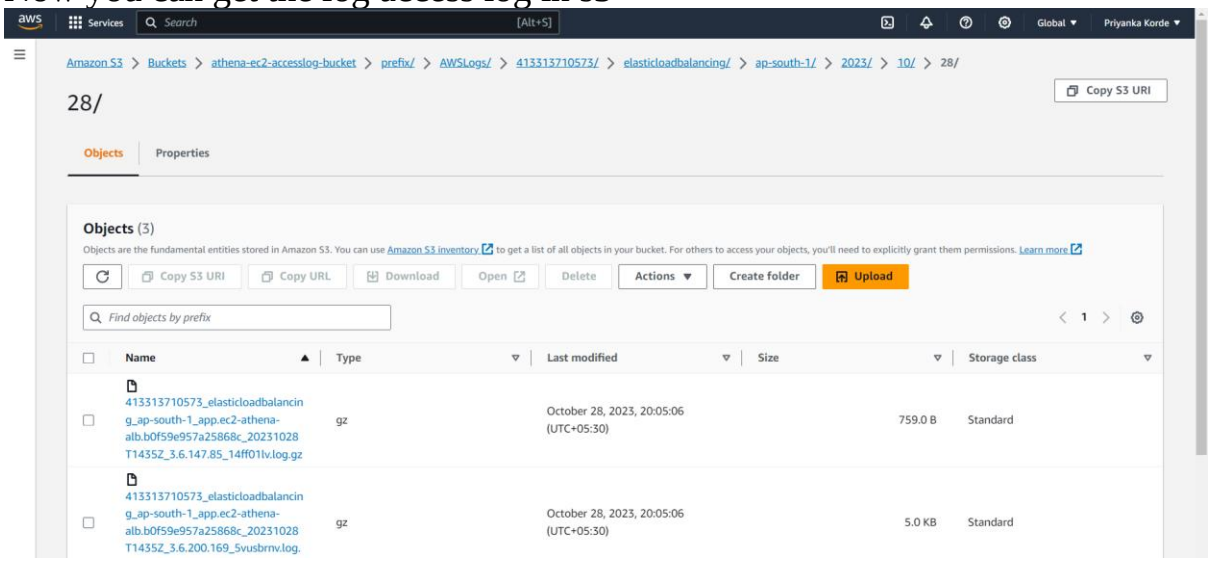


From this point forward, your ALB will start generating access logs and storing them in the specified S3 bucket. You can use these logs for monitoring and analysis of traffic to your load balancer.

Now check in the s3 bucket you will receive the **[ELBAccessLogTestFile]**

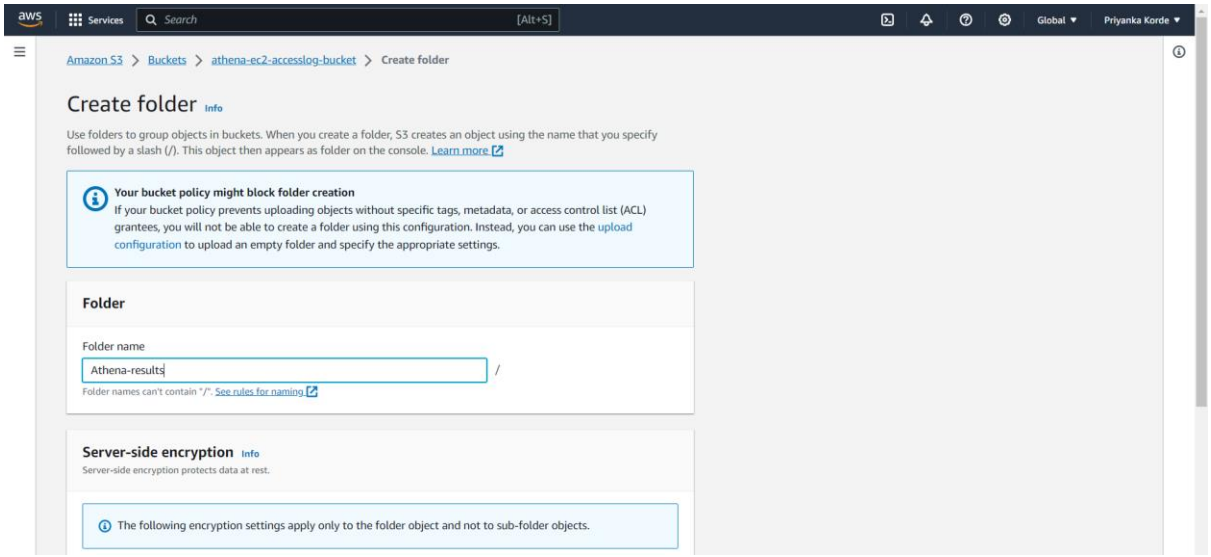


Now you can get the log access log in s3



Step 6: Querying AWS Application Load Balancer (ALB) access logs using Amazon Athena:

- Navigate to the Athena service by clicking on “Services” and then selecting “Athena” under the Analytics section.
- Make sure your ALB access logs are being delivered to the S3 bucket, as previously configured.
- Create a folder in the same s3 bucket for the Athena query results.

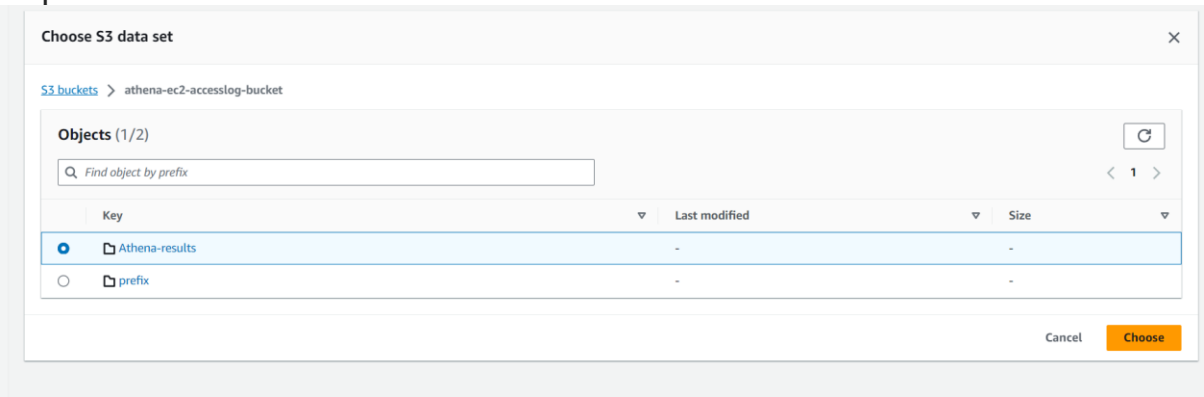


In the Athena console, click on “Query Editor” on the left-hand side.

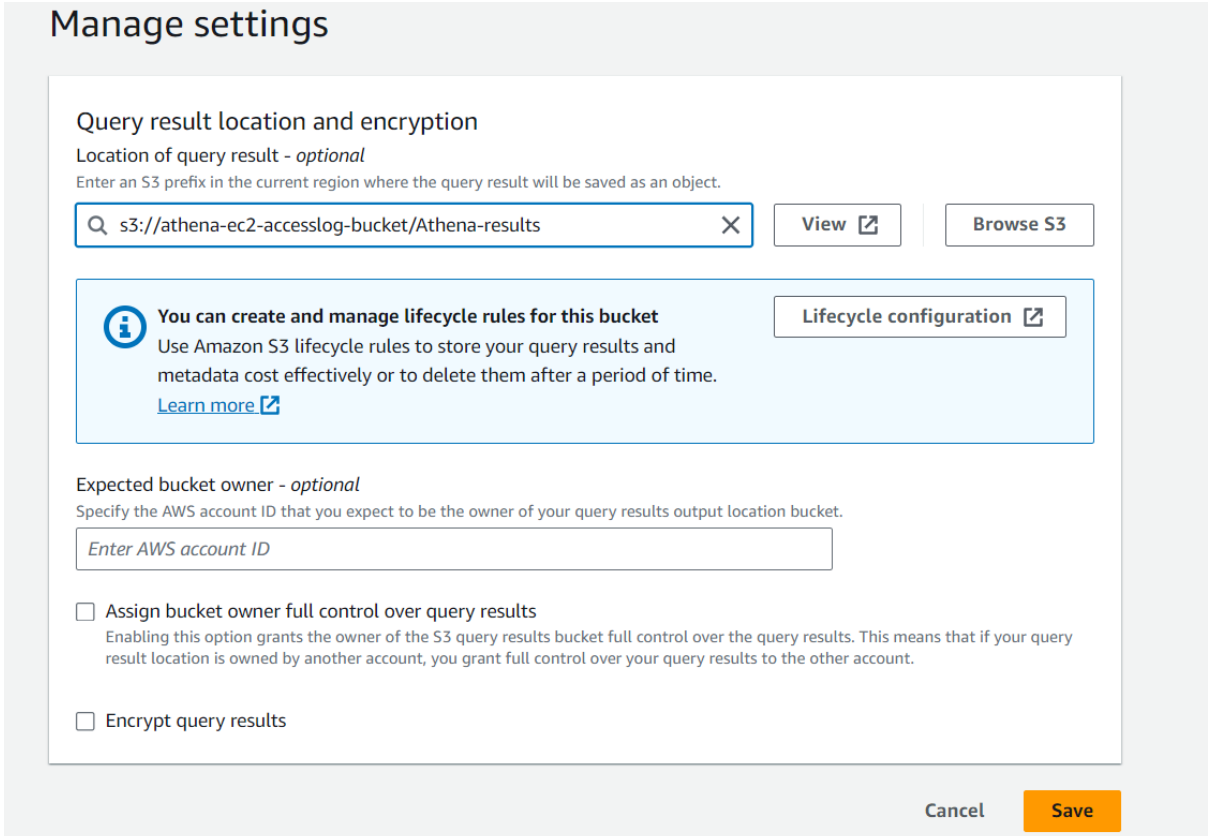
Click on settings and manage

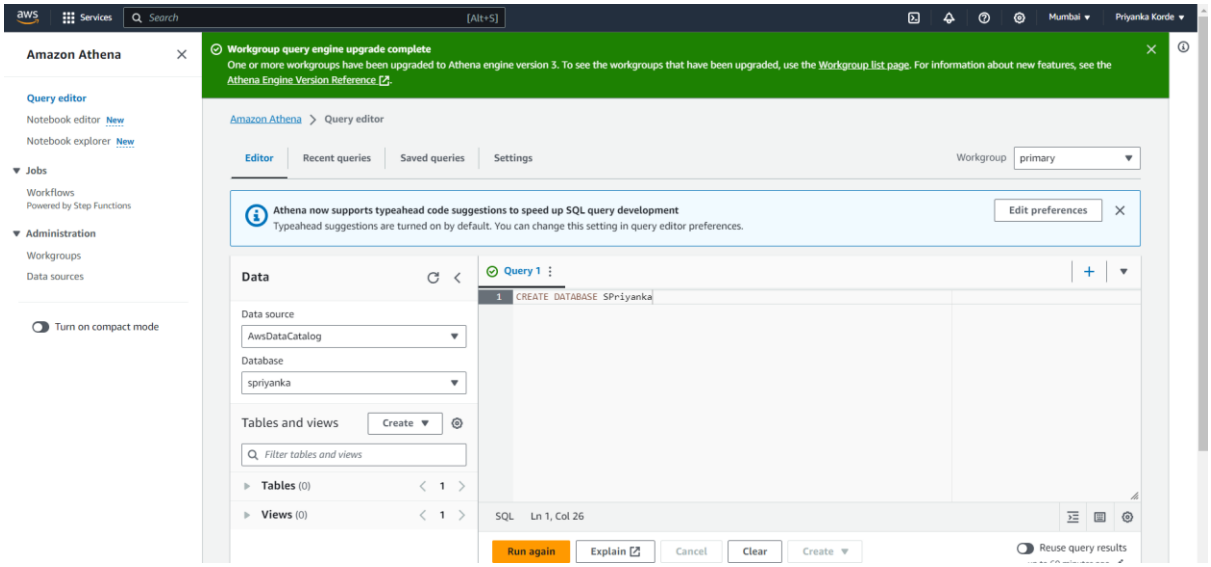
“Access the settings menu and then choose ‘Manage.’ From there, select the S3 bucket you’ve designated for storing the ALB logs.”

click on **browse** select the bucket and folder and which we have created earlier steps



Click on Choose and **Save**



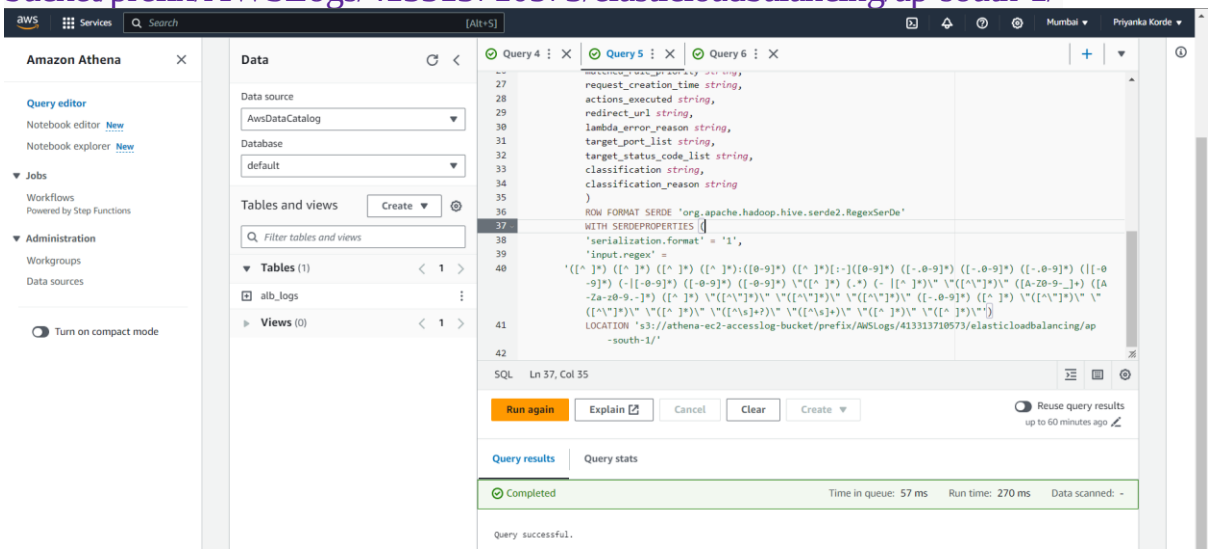


Step 8: Create a Table

After creating the database, you need to create a table that defines the structure of your ALB access logs. You can do this with a CREATE TABLE statement.

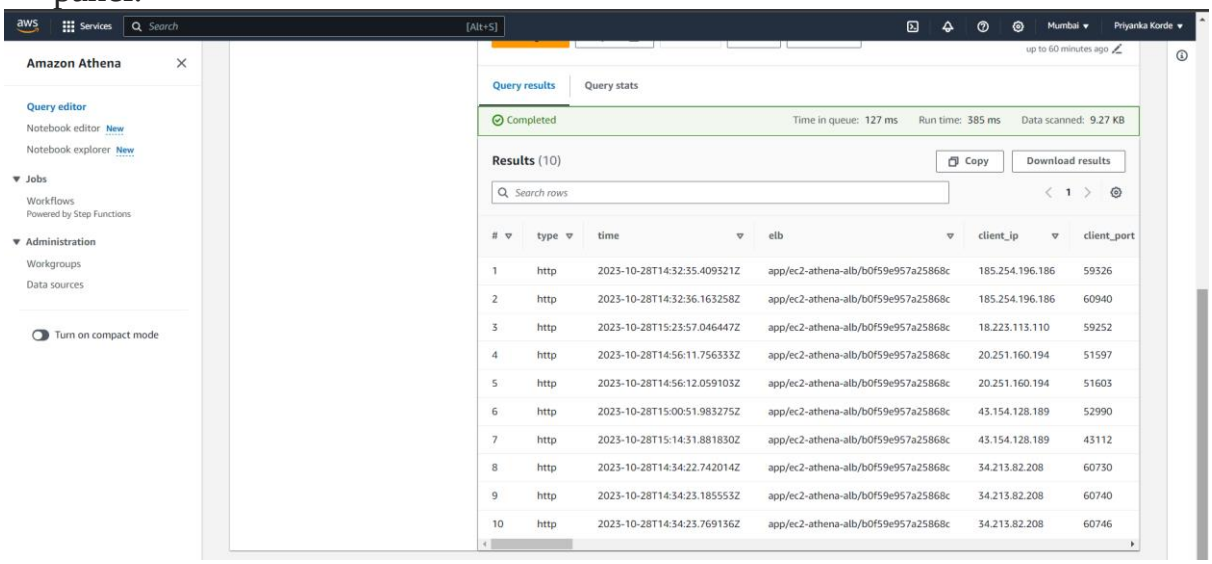
```
CREATE EXTERNAL TABLE IF NOT EXISTS alb_logs (
        type string,
        time string,
        elb string,
        client_ip string,
        client_port int,
        target_ip string,
        target_port int,
        request_processing_time double,
        target_processing_time double,
        response_processing_time double,
        elb_status_code int,
        target_status_code string,
        received_bytes bigint,
        sent_bytes bigint,
        request_verb string,
        request_url string,
        request_proto string,
        user_agent string,
        ssl_cipher string,
        ssl_protocol string,
        target_group_arn string,
        trace_id string,
        domain_name string,
        chosen_cert_arn string,
        matched_rule_priority string,
```

```
        request_creation_time string,
        actions_executed string,
        redirect_url string,
        lambda_error_reason string,
        target_port_list string,
        target_status_code_list string,
        classification string,
        classification_reason string
        )
        ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.RegexSerDe'
        WITH SERDEPROPERTIES (
        'serialization.format' = '1',
        'input.regex' =
    '([^ ]*) ([^ ]*) ([^ ]*) ([^ ]*):([0-9]*) ([^ ]*)[:-]([0-9]*) ([-.0-9]*) ([-.0-9]*) ([-.0-9]*) (|[-0-9]*) (-|[-0-9]*) ([-0-9]*) ([-0-9]*) \"([^ ]*) (.*) (- |[^ ]*)\" \"([^\"]*)\" ([A-Z0-9-_]+) ([A-Za-z0-9.-]*) ([^ ]*) \"([^\"]*)\" \"([^\"]*)\" \"([^\"]*)\" ([-.0-9]*) ([^ ]*) \"([^\"]*)\" \"([^\"]*)\" \"([^ ]*)\" \"([^\s]+?)\" \"([^\s]+)\" \"([^ ]*)\" \"([^ ]*)\"')
        LOCATION 's3://athena-ec2-accesslog-bucket/prefix/AWSLogs/413313710573/elasticloadbalancing/ap-south-1/'
```



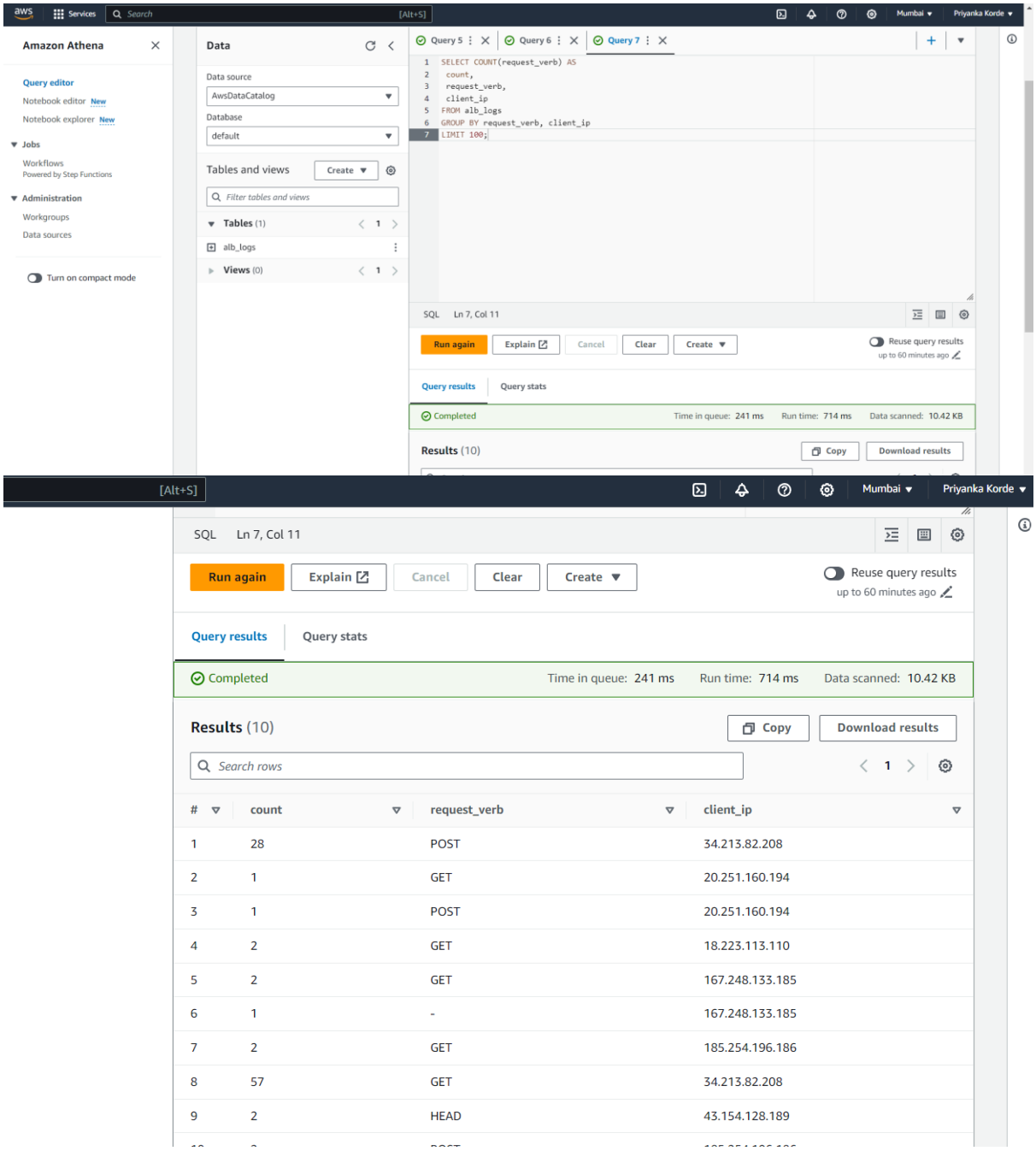
To preview a table in Amazon Athena:

1. Click on the desired table name in the database list.
2. Open the table's menu by clicking the three dots next to the name.
3. Choose "Preview Table" to view a sample of the table's data in a new tab or panel.



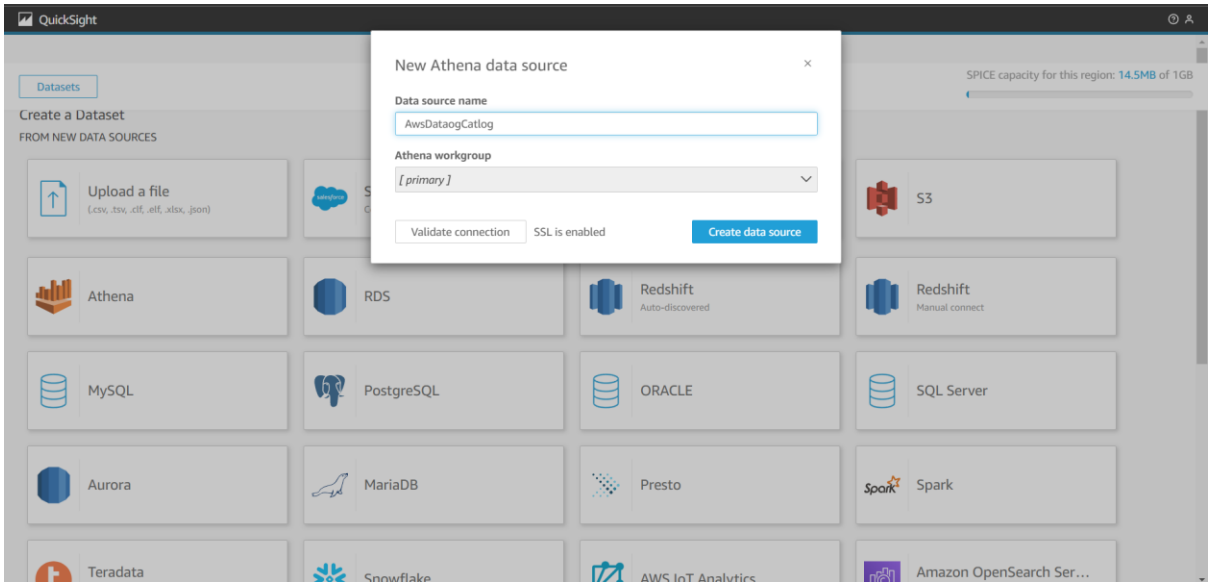
query that performs a count of HTTP GET requests grouped by the client IP address:

```
SELECT COUNT(request_verb) AS
 count,
 request_verb,
 client_ip
FROM alb_logs
GROUP BY request_verb, client_ip
LIMIT 100;
```

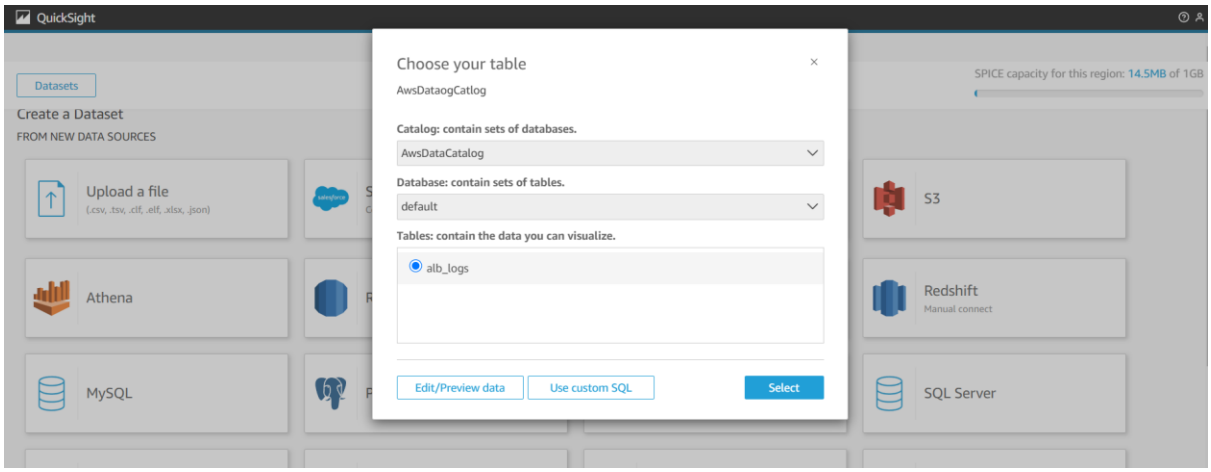


Step 8: Quicksite Analysis

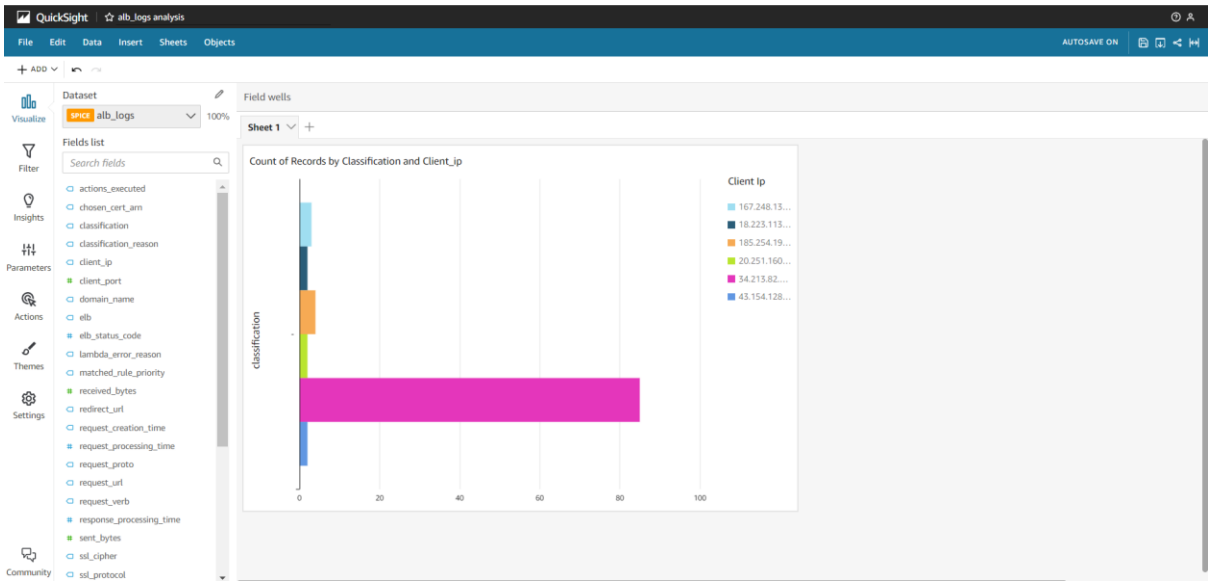
Navigate to New Analysis > New dataset > Athena

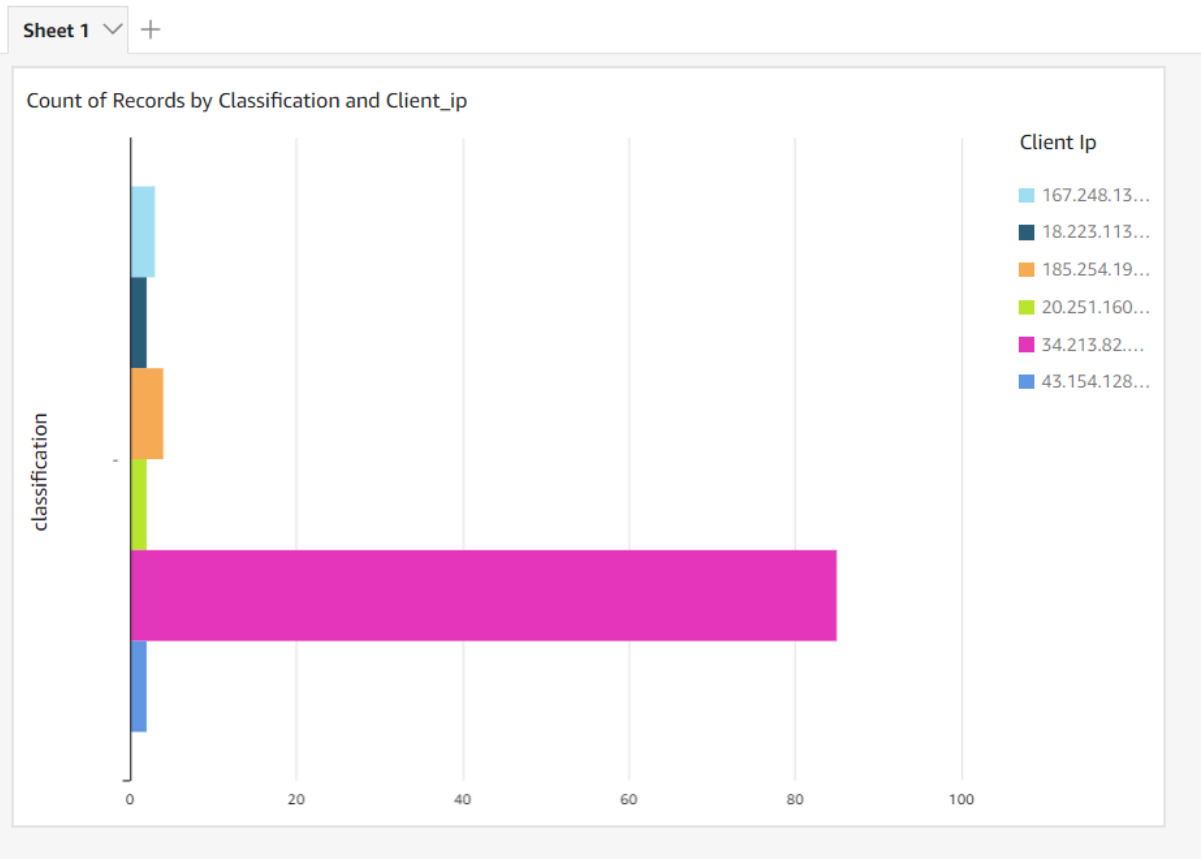


Click Create data source



Select the Y-axis dimension and Group/Color dimension





In This Way you can use the Amazon Quick sight Analysis for Analysing the overall tasks.