**Assignment – 11**

**Problem Statement:**

**Build scaling plans in AWS that balance the load on different EC2 instances.**

1. Go to EC2 then in Instances click on Launch Templates.

48 A screenshot of a computer

Description automatically generated

1. Click on Create launch template.

A screenshot of a computer

Description automatically generated

1. Give Launch template name and then click on checkbox for Auto Scaling guidance.

A screenshot of a computer

Description automatically generated

1. In Quick Start select Ubuntu.

A screenshot of a computer

Description automatically generated

1. Select Instance type – either t2.micro or t3.micro. Create new key pair.

A screenshot of a computer

Description automatically generated

1. Select existing security group.

A screenshot of a computer

Description automatically generated

1. Now go to Advanced details and then in User data write given statements and then Create launch template.

#!/bin/bash

apt -get update

apt -get install -y nginx

systemctl start nginx

systemctl enable nginx

apt -get install -y git

curl -SL https://deb.nodesource.com/setup\_16.x|sudo -E bash -

apt -get install -y nodejs

git clone https://github.com/UnderDevelopment10/new-repo1.git

cd new-repo1

npm install

node index.js

A screenshot of a computer

Description automatically generated

1. After creating Launch template, click on Auto Scaling Groups in left pane.

Click on Create Auto Scaling group.

A screenshot of a computer

Description automatically generated

1. Give Auto scaling group name and then select your recently launched template. Then click on Next.

A screenshot of a launch template

Description automatically generated

1. Select availability zones and subnets then click on Next.

A screenshot of a computer

Description automatically generated

1. Click on Attach to a new load balancer.

A screenshot of a computer screen

Description automatically generated

1. Select Application Load Balancer as load balancer type and Internet-facing as Load balancer scheme.

A screenshot of a computer

Description automatically generated

1. Select port no. 4000 for HTTP and select New target group name.

A screenshot of a computer

Description automatically generated

1. Now turn on Elastic load balancing health check and in health check grace period give 240 seconds. Then click on Next.

A screenshot of a computer

Description automatically generated

1. Select 2 in Desired capacity in group size and in Scaling 2 as Min desired capacity and 3 as Max desired capacity.

A screenshot of a computer

Description automatically generated

1. Next in Automatic scaling select Target tracking scaling policy and give 240 in Instance warmup.

A screenshot of a computer

Description automatically generated

1. Then click on Next, then again next.

A screenshot of a computer error

Description automatically generated A screenshot of a computer

Description automatically generated

A screenshot of a computer

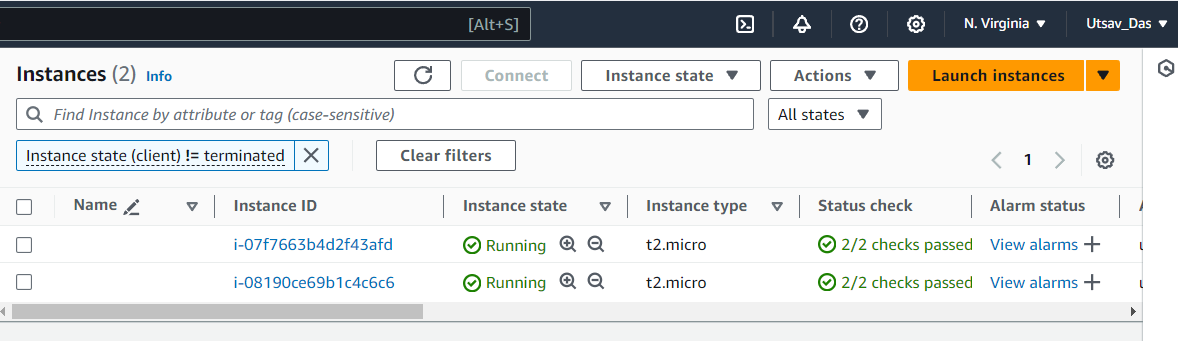
Description automatically generated

1. Now click on Create Auto Scaling group.

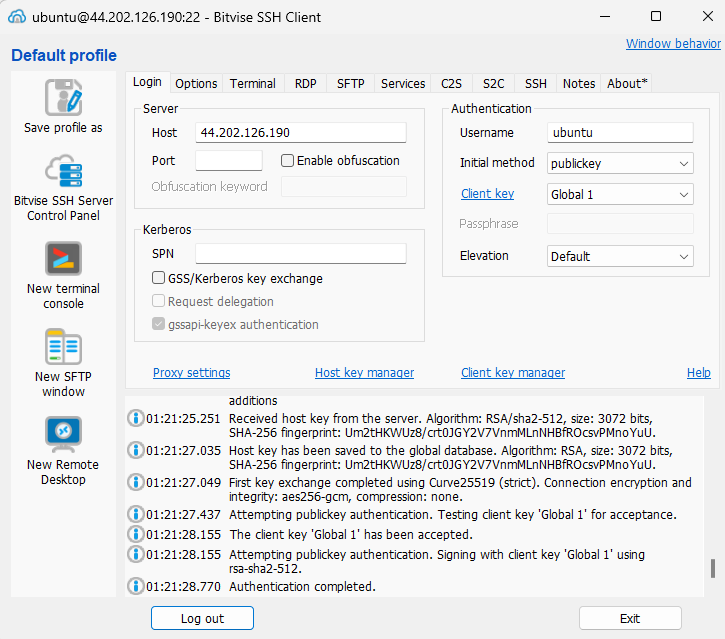
A screenshot of a computer

Description automatically generated

1. Come back to EC2 dashboard and go to Instance and here you can see 2 instances running as minimum capacity 2 is chosen.



1. Click on any instance and then copy public IPv4 address.
2. Open Bitvise SSH client and then paste that public IPv4 address and then in client key manager import the generated key and then Login.



1. Open New terminal console and then write command 'sudo nano infi.sh'. A new .sh file will be created. Now write this code to run an infinite loop.

!/bin/bash

while(true)

do

echo"Inside Loop"

done

A black screen with white text

Description automatically generated

1. Then do Ctrl+X then Y then click Enter. And write command 'sudo chmod 777 infy.sh' to provide all permissions to file and then to run give 'sh infy.sh'.

A group of white text on a black background

Description automatically generated

1. Here we can see infinite loop running.

A black screen with white text

Description automatically generated

1. Go back to instances and then select both instances and then in bottom pane, go to CPU utilization and select Enlarge.

A graph with numbers and lines

Description automatically generated

1. The graph displays CPU utilization for both instances. When CPU utilization for one instance is very high, then another instance will be created as we have set maximum capacity to 3.
2. Finally in Instances, we can see another new instance is created.

A screenshot of a computer

Description automatically generated

A graph on a white background

Description automatically generated