Case study – ELB,ASG AND ROUTE 53

Problem Statement: You work for XYZ Corporation that uses on premise solutions and a limited number of systems. With the increase in requests in their application, the load also increases. So, to handle the load the corporation has to buy more systems almost on a regular basis. Realizing the need to cut down the expenses on systems, they decided to move their infrastructure to AWS

Tasks To Be Performed:

1. Manage the scaling requirements of the company by:

a. Deploying multiple compute resources on the cloud as soon as the load increases and the CPU utilization exceeds 80%

b. Removing the resources when the CPU utilization goes under 60%

2. Create a load balancer to distribute the load between compute resources.

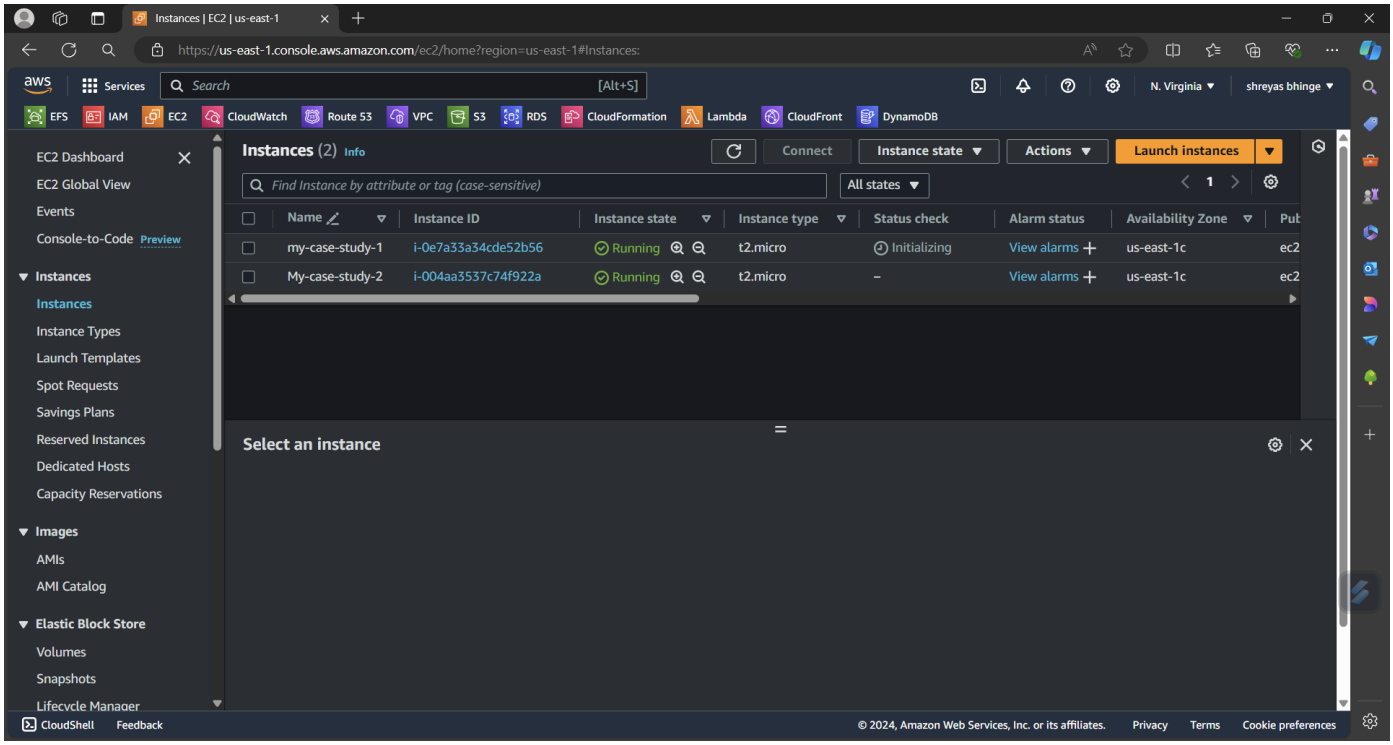
3. Route the traffic to the company

Tasks To Be Performed:

1. Manage the scaling requirements of the company by:

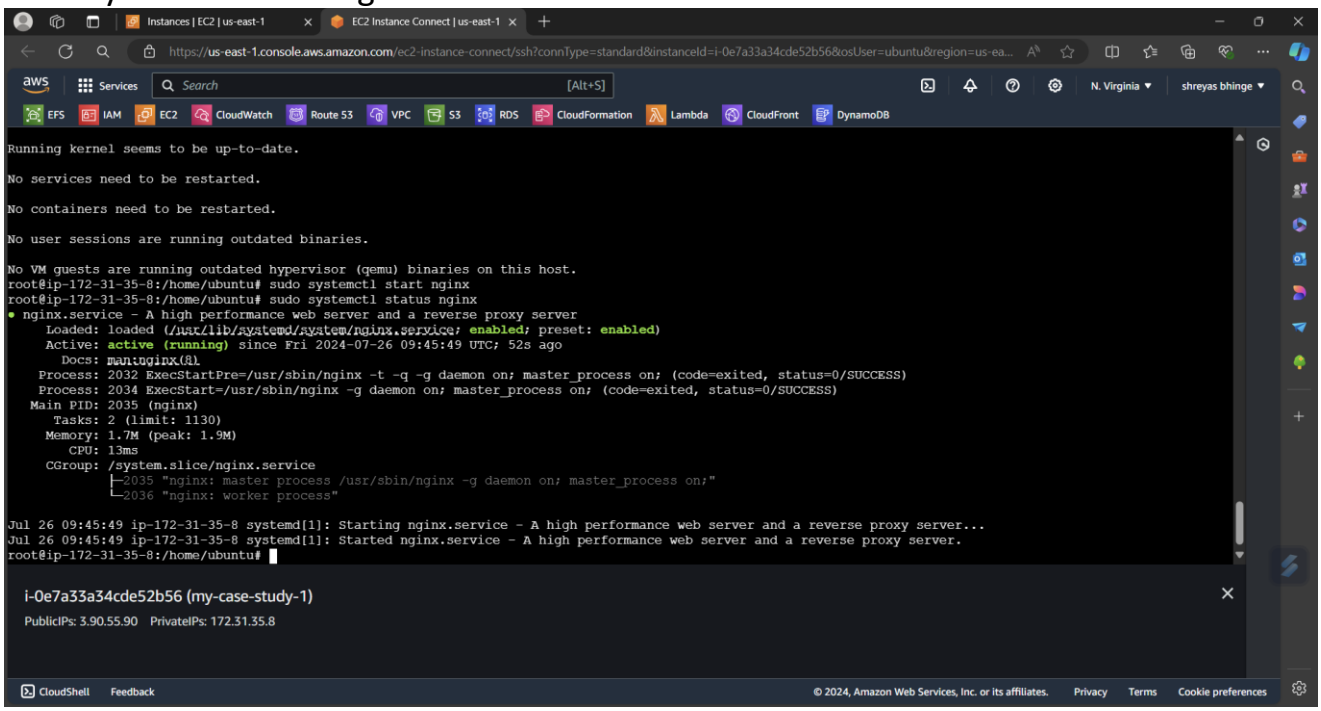
a. Deploying multiple compute resources on the cloud as soon as the load increases and the CPU utilization exceeds 80%

Create 2 EC2 Instance

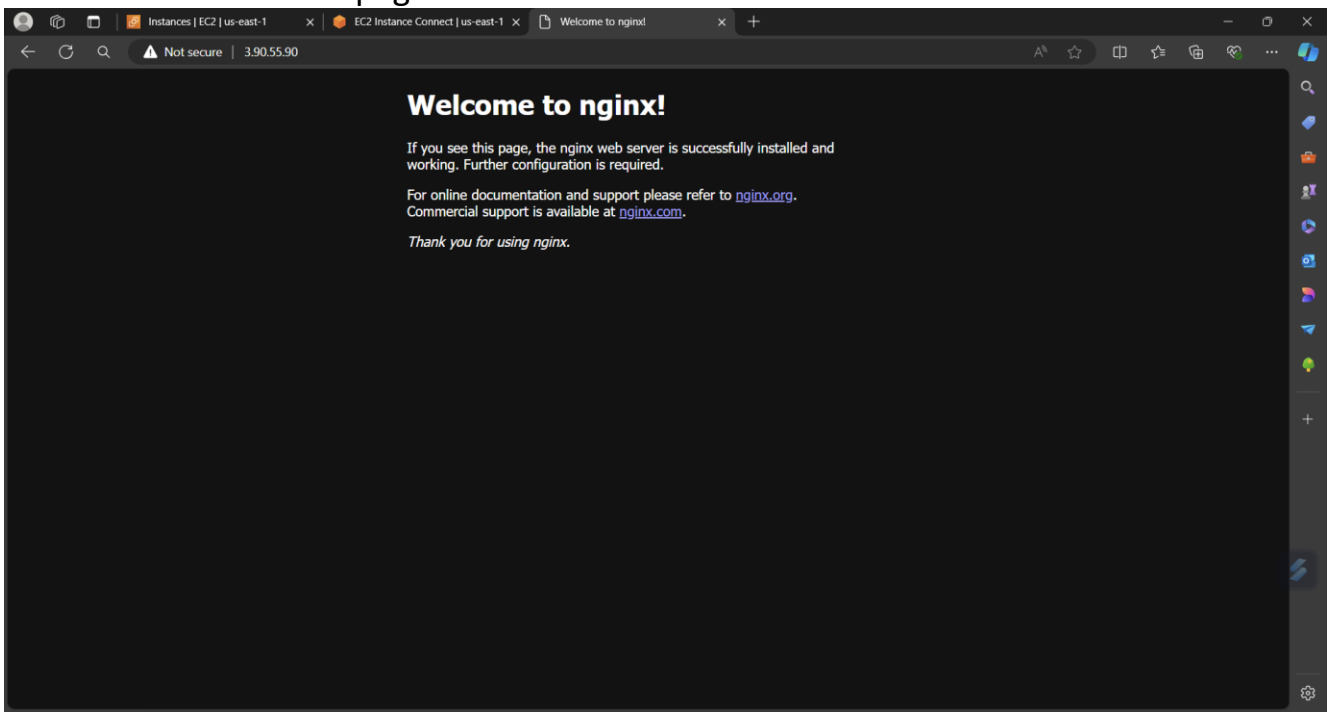


- In first instance
  I Installed nginx

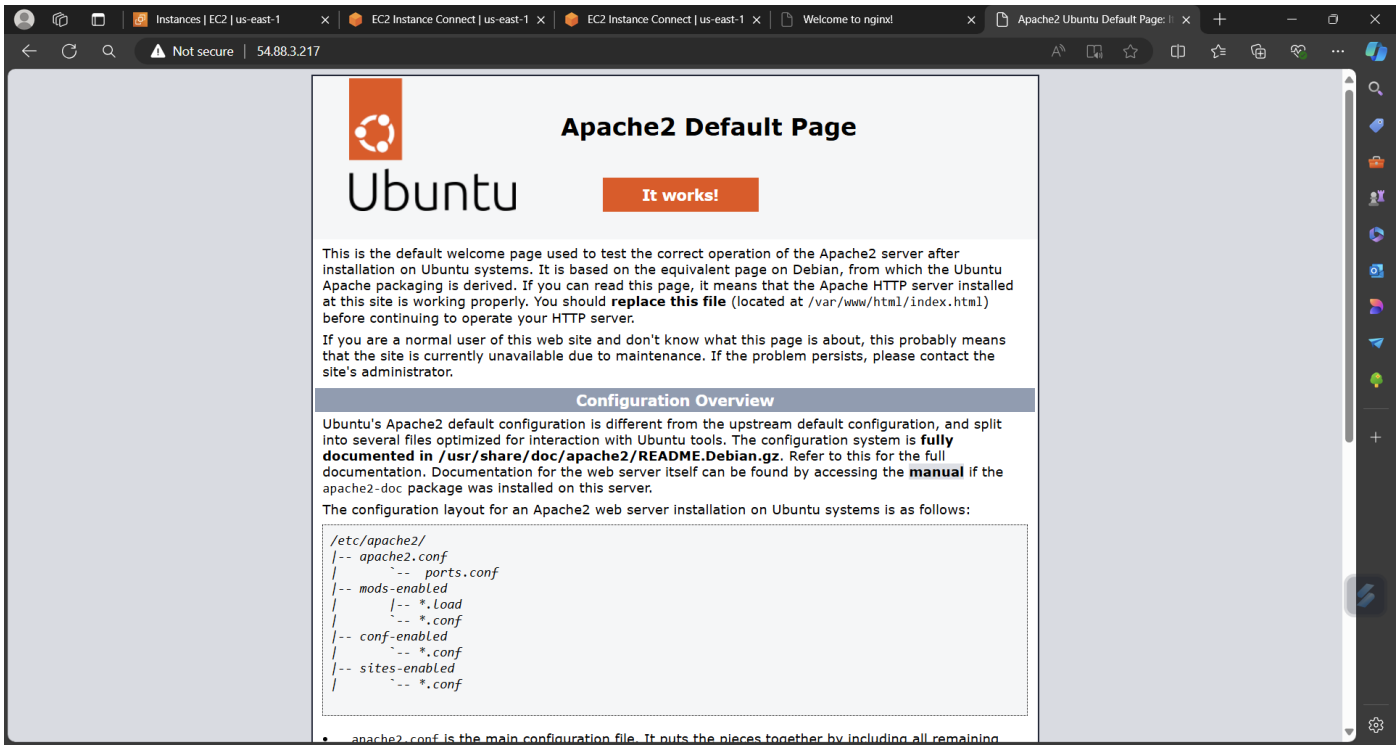
By using below command
Sudo su sudo apt-get update
Sudo apt install ngnix
Sudo systemctl start nginx
Sudo systemctl status nginx



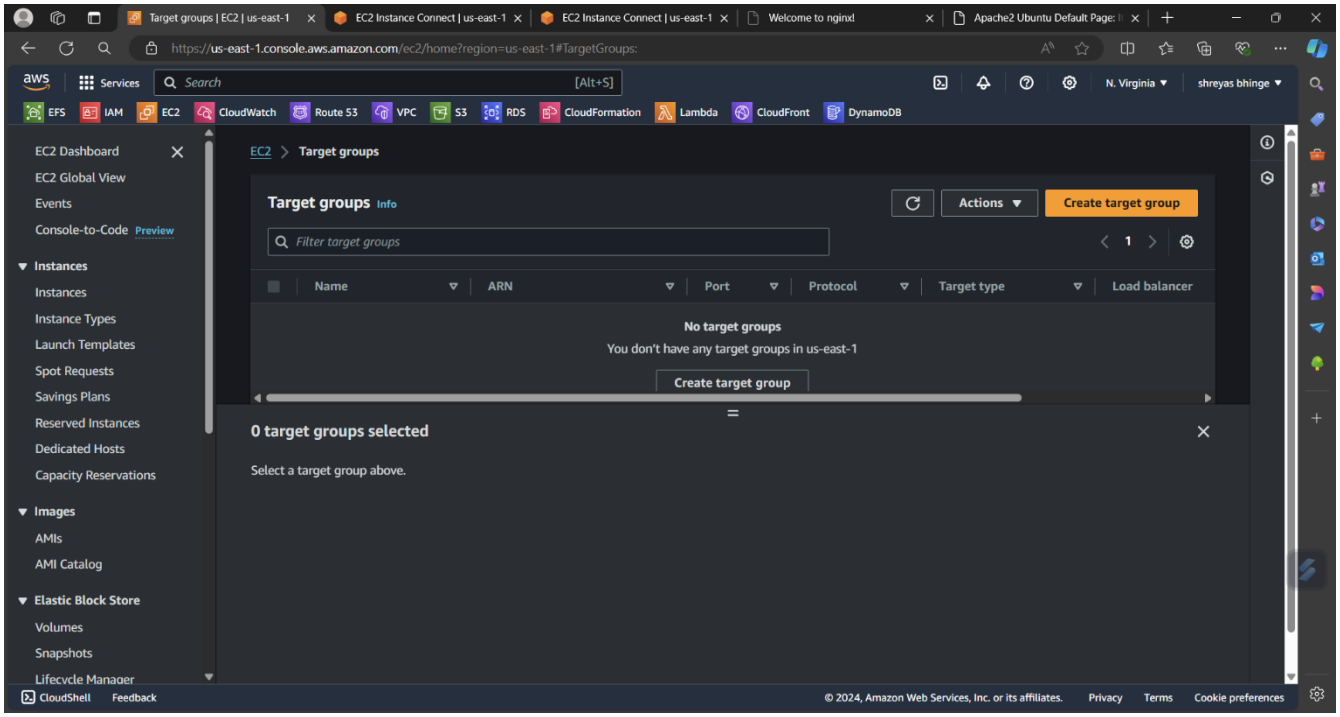
- You will find this page

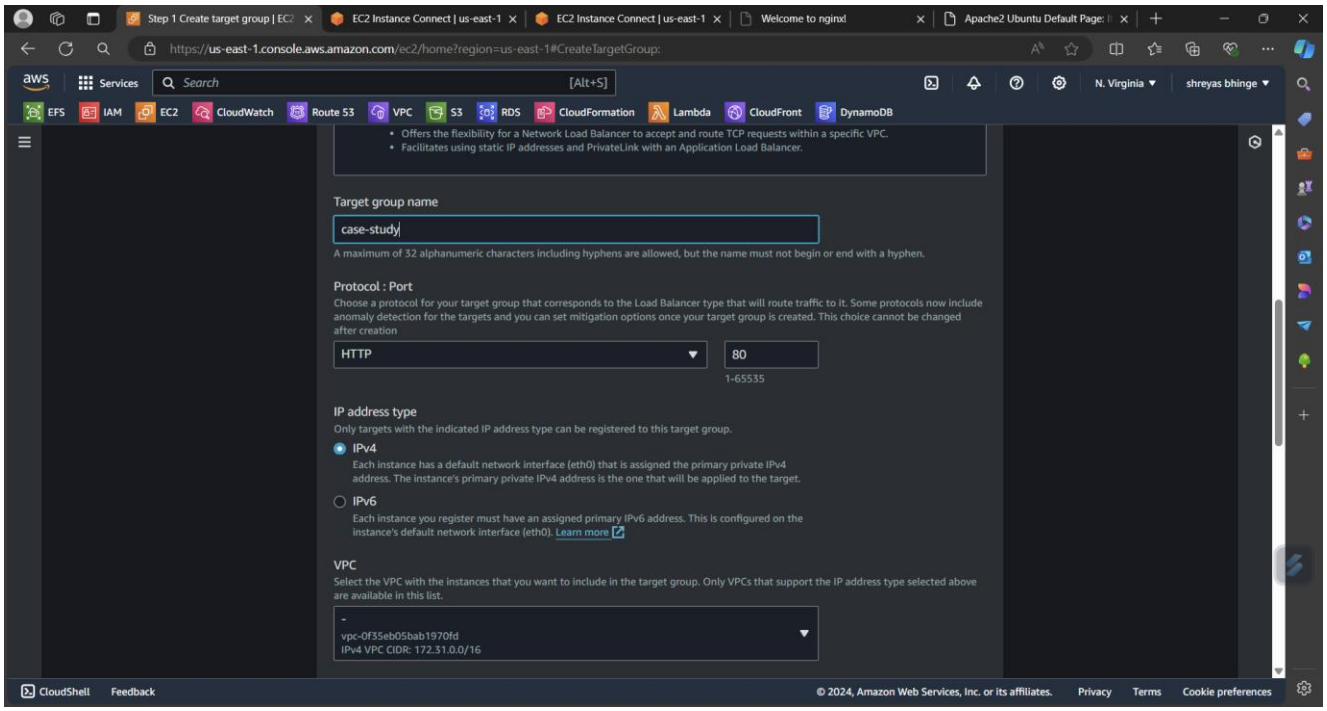


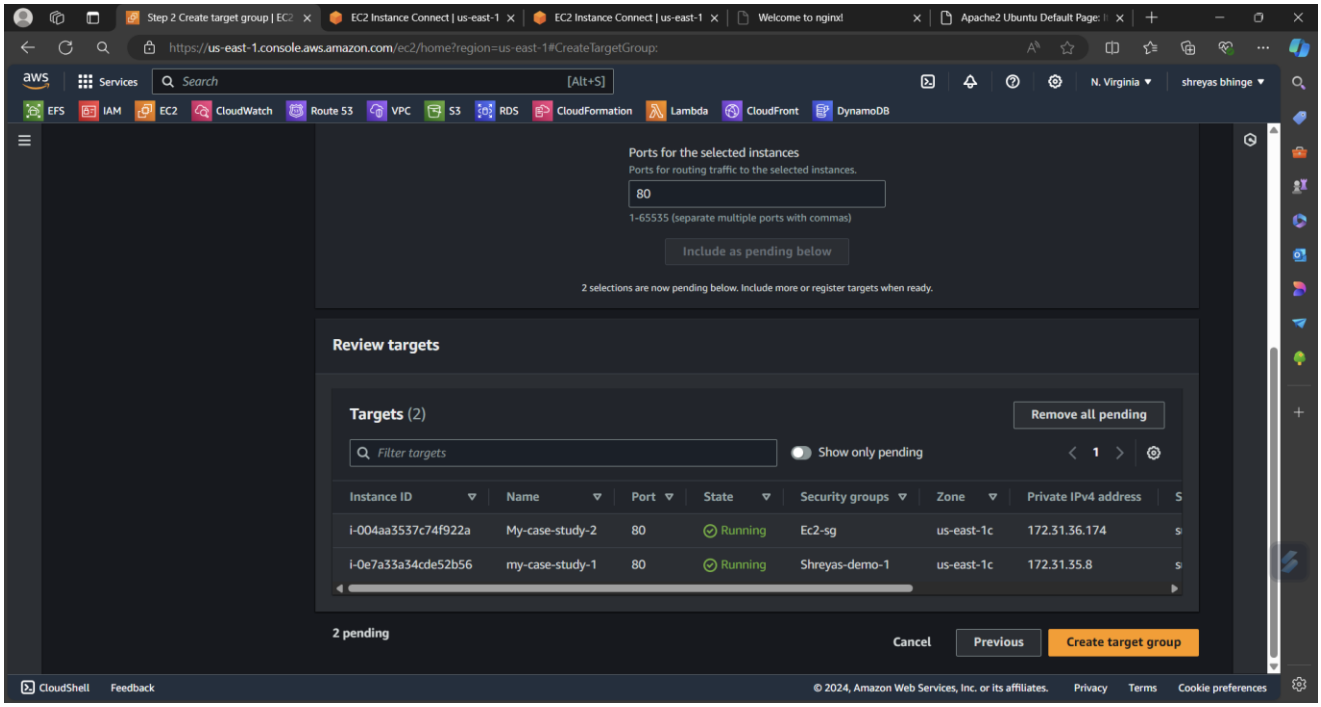
In the same way create 2nd instance
Sudo apt-get update
Sudo apt install apache2
Sudo systemctl start apache2



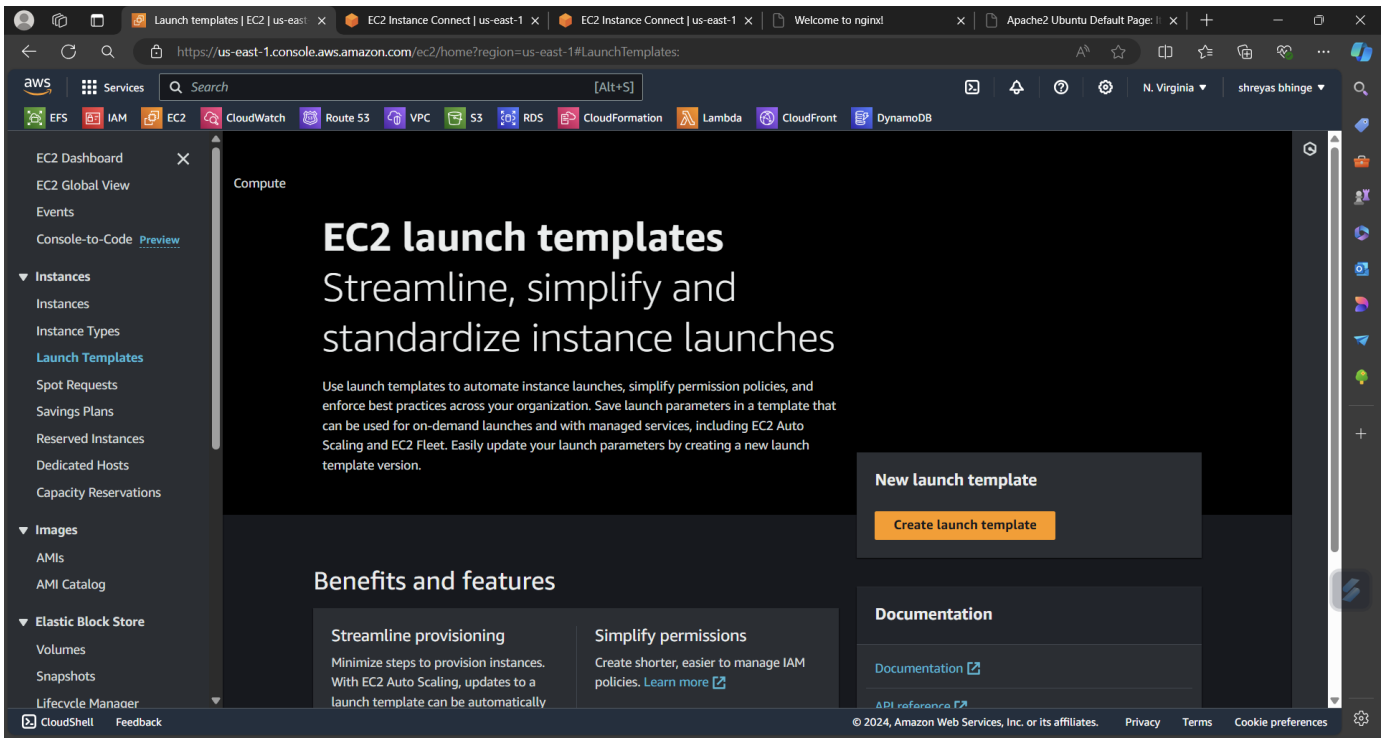
- Now Create target group

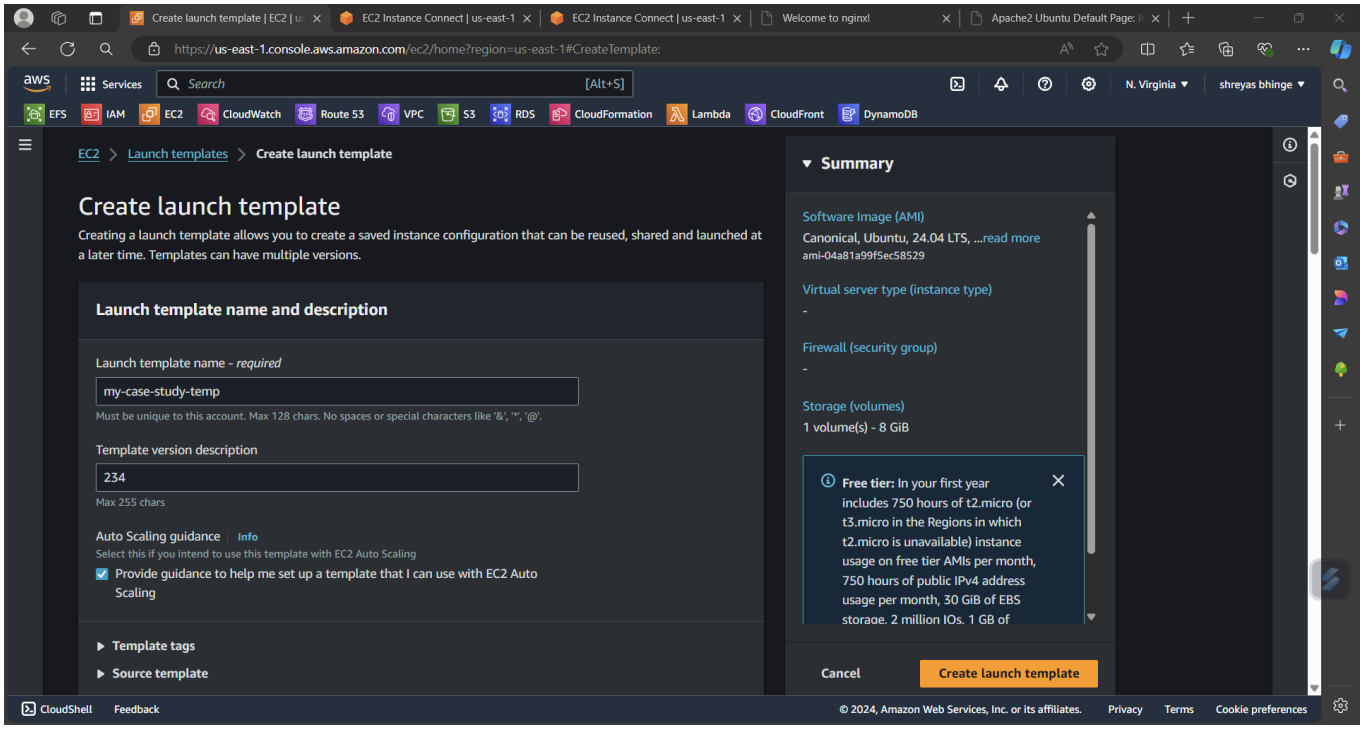




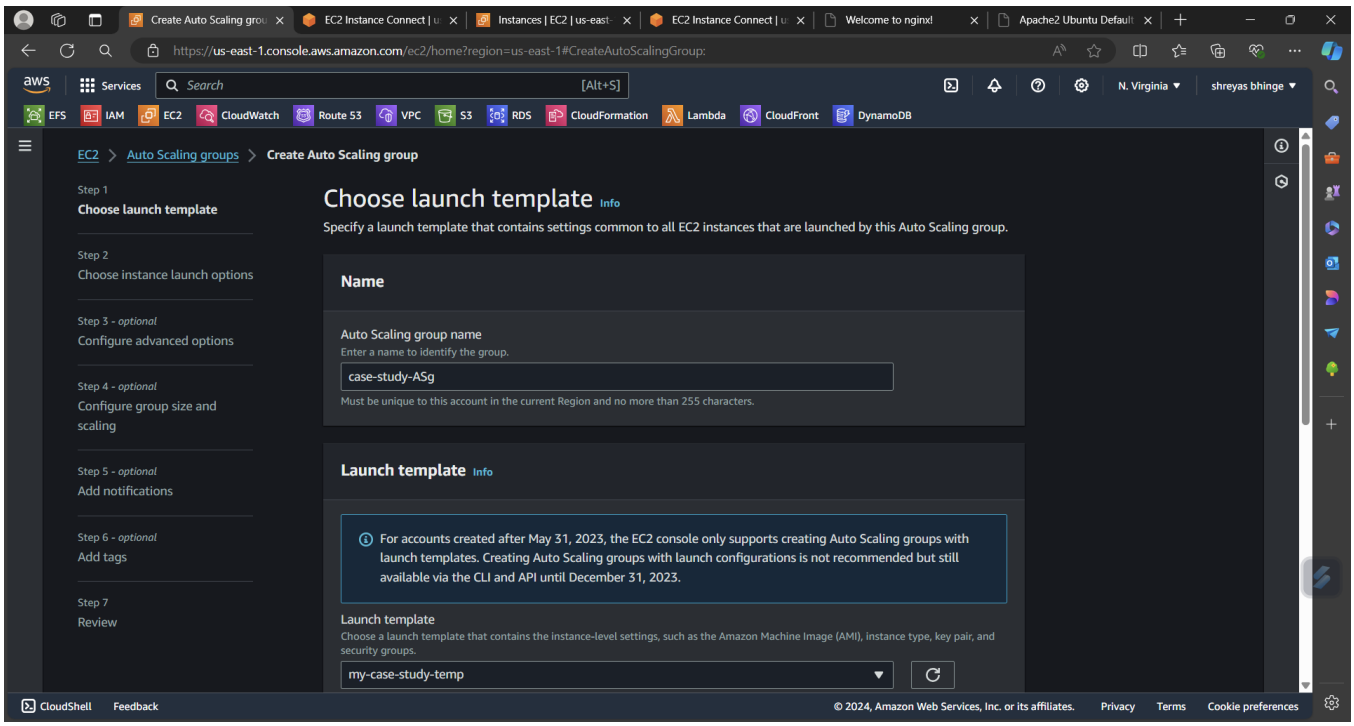


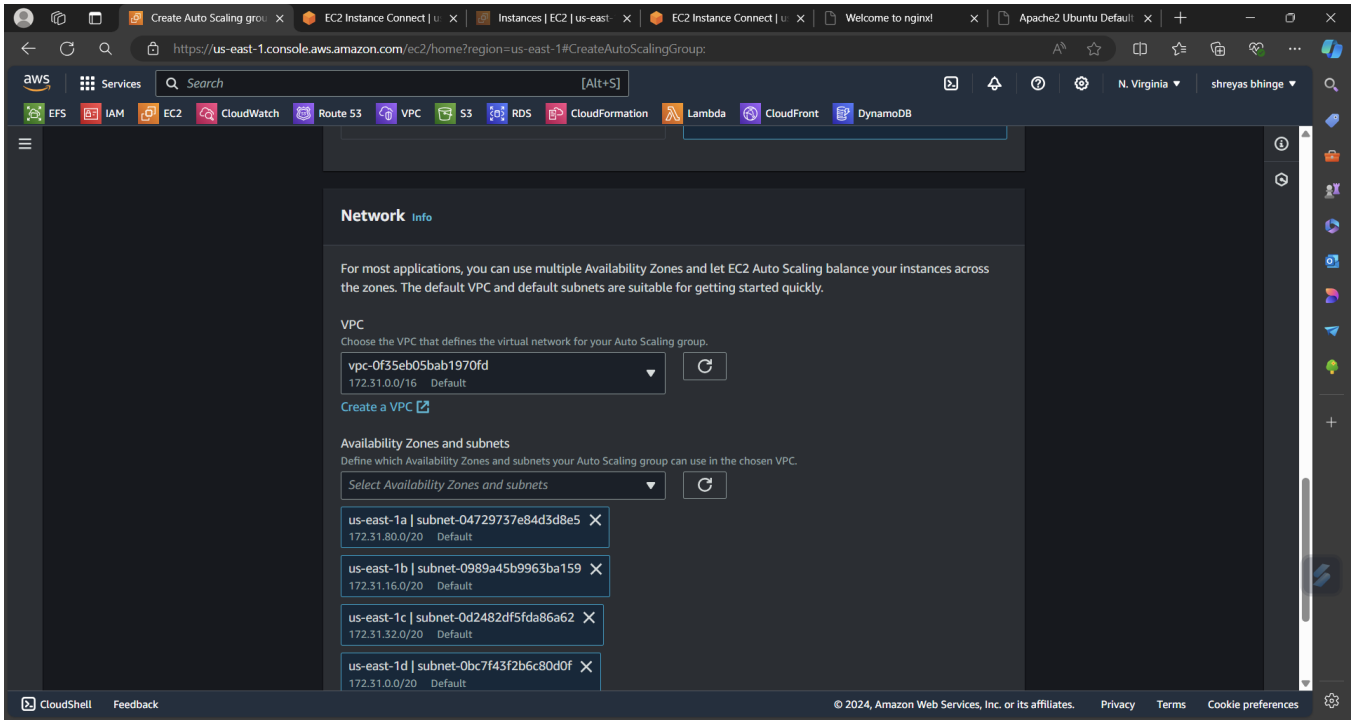
- Now launch template

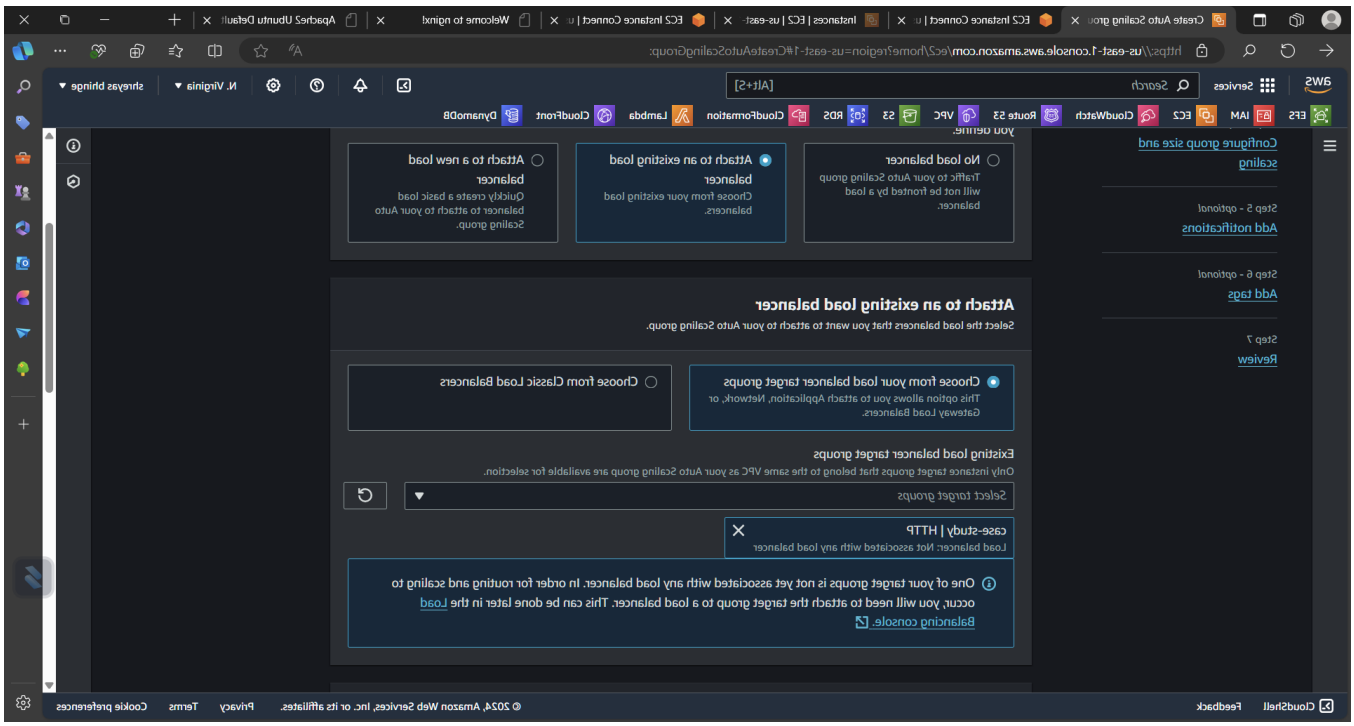


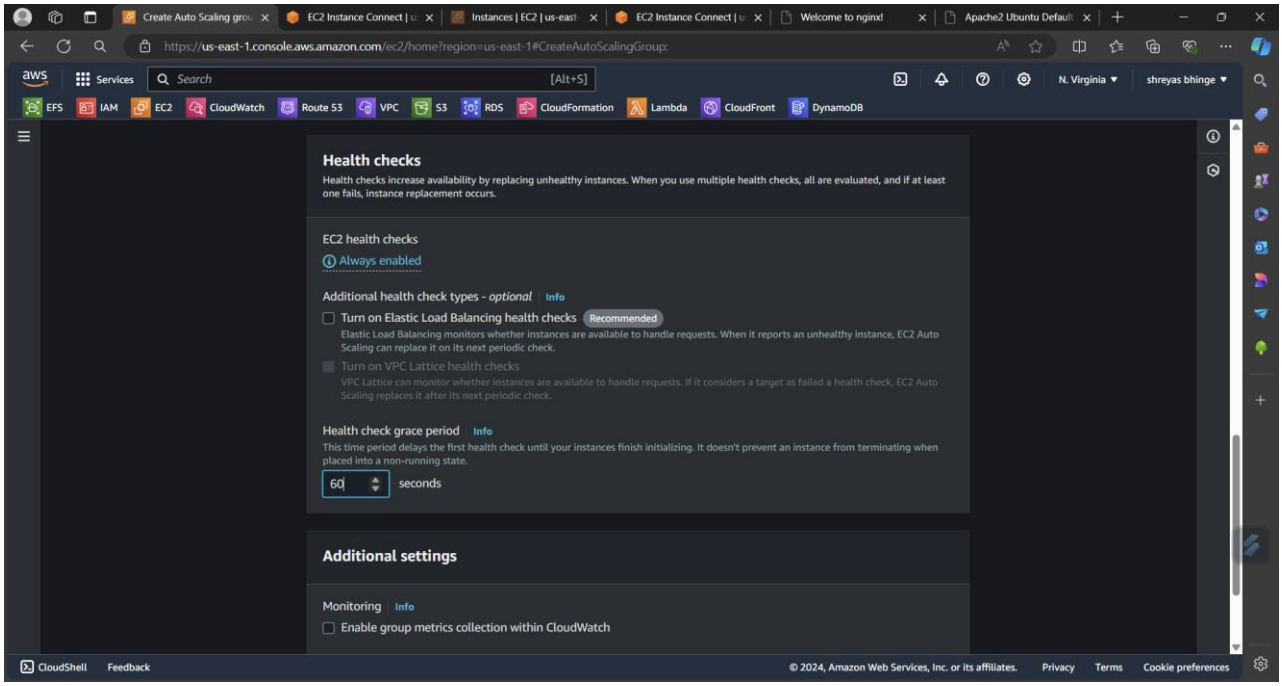


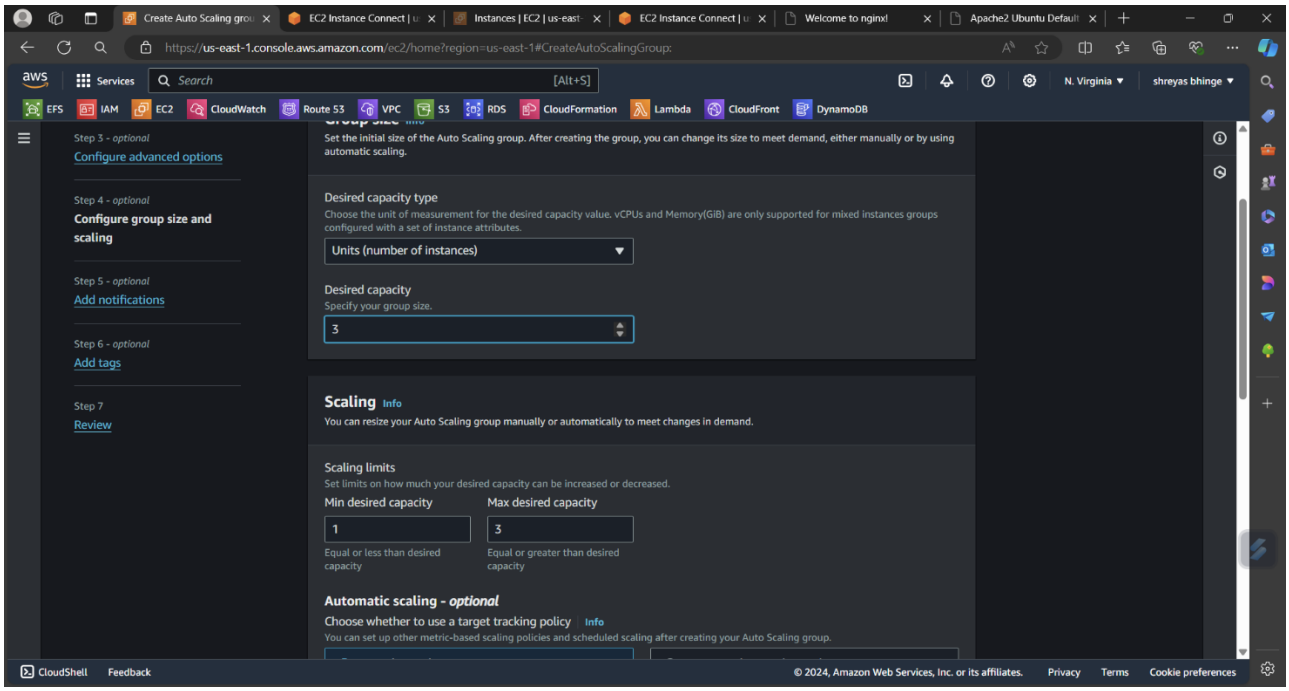
- Now launch Auto scaling

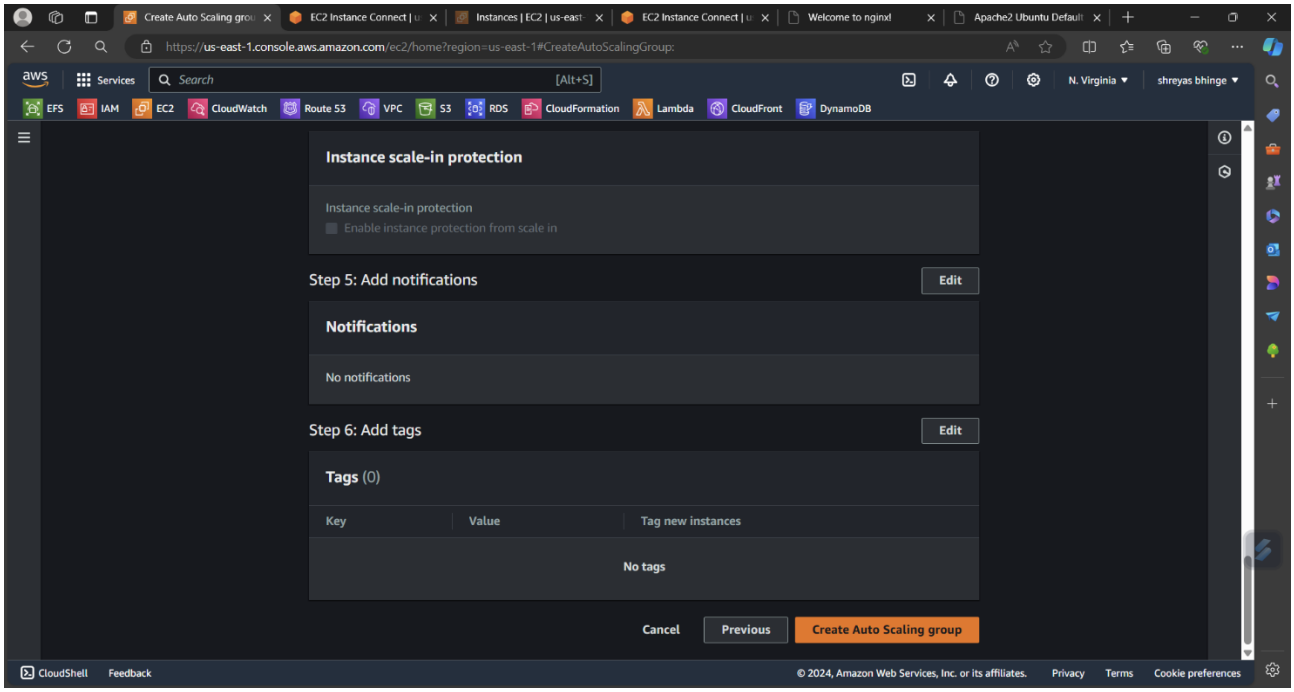


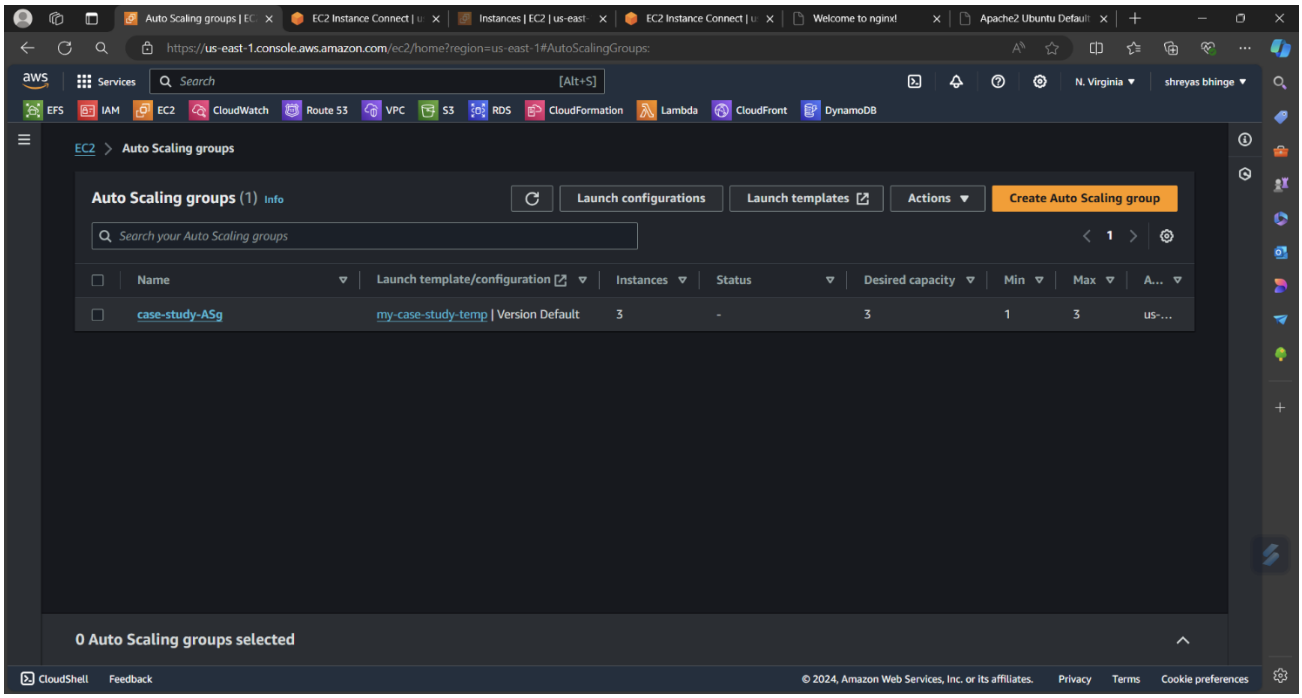

Create Auto Scaling grou
EC2 Instance Connect | u
Instances | EC2 | us-east-
EC2 Instance Connect | u
Welcome to nginx!
Apache2 Ubuntu Default
https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateAutoScalingGroup:
aws
Services
Search
[Alt+S]
N. Virginia
shreyas bhinge
EFS
IAM
EC2
CloudWatch
Route 53
VPC
S3
RDS
CloudFormation
Lambda
CloudFront
DynamoDB
Network Info
For most applications, you can use multiple Availability Zones and let EC2 Auto Scaling balance your instances across the zones. The default VPC and default subnets are suitable for getting started quickly.
VPC
Choose the VPC that defines the virtual network for your Auto Scaling group.
vpc-0f35eb05bab1970fd
172.31.0.0/16 Default
Create a VPC
Availability Zones and subnets
Define which Availability Zones and subnets your Auto Scaling group can use in the chosen VPC.
Select Availability Zones and subnets
us-east-1a | subnet-04729737e84d3d8e5
172.31.80.0/20 Default
us-east-1b | subnet-0989a45b9963ba159
172.31.16.0/20 Default
us-east-1c | subnet-0d2482df5fda86a62
172.31.32.0/20 Default
us-east-1d | subnet-0bc7f43f2b6c80d0f
172.31.0.0/20 Default
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

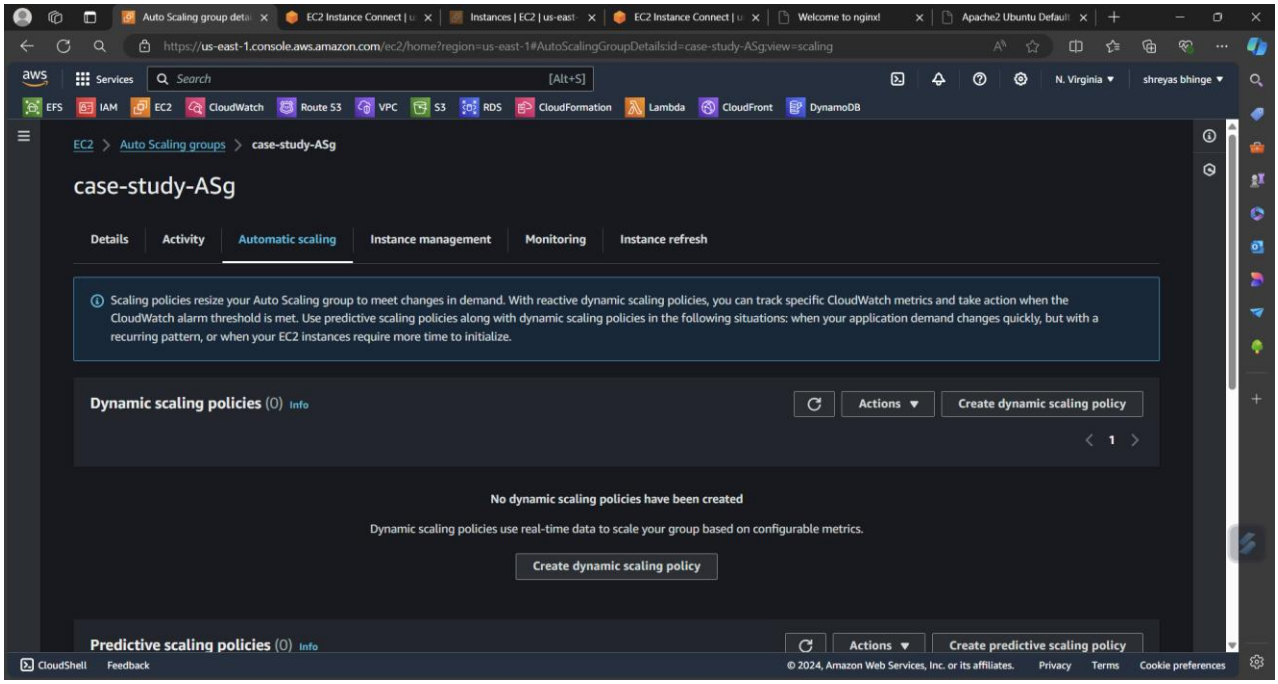


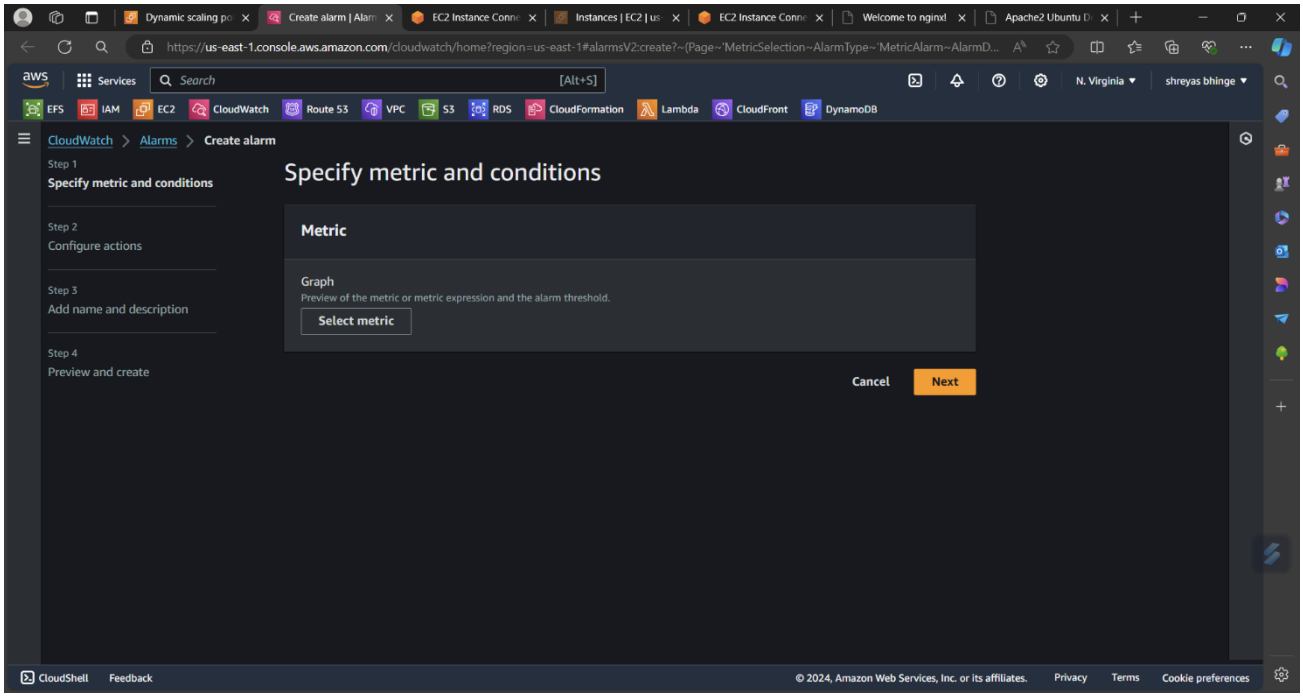

Create Auto Scaling grou
EC2 Instance Connect | u
Instances | EC2 | us-east
EC2 Instance Connect | u
Welcome to nginx!
Apache2 Ubuntu Default
https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateAutoScalingGroup:
Services
Search
[Alt+S]
N. Virginia
shreyas bhinge
EFS
IAM
EC2
CloudWatch
Route 53
VPC
S3
RDS
CloudFormation
Lambda
CloudFront
DynamoDB
Health checks
Health checks increase availability by replacing unhealthy instances. When you use multiple health checks, all are evaluated, and if at least one fails, instance replacement occurs.
EC2 health checks
Always enabled
Additional health check types - optional
Info
Turn on Elastic Load Balancing health checks
Recommended
Elastic Load Balancing monitors whether instances are available to handle requests. When it reports an unhealthy instance, EC2 Auto Scaling can replace it on its next periodic check.
Turn on VPC Lattice health checks
VPC Lattice can monitor whether instances are available to handle requests. If it considers a target as failed a health check, EC2 Auto Scaling replaces it after its next periodic check.
Health check grace period
Info
This time period delays the first health check until your instances finish initializing. It doesn't prevent an instance from terminating when placed into a non-running state.
60
seconds
Additional settings
Monitoring
Info
Enable group metrics collection within CloudWatch
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

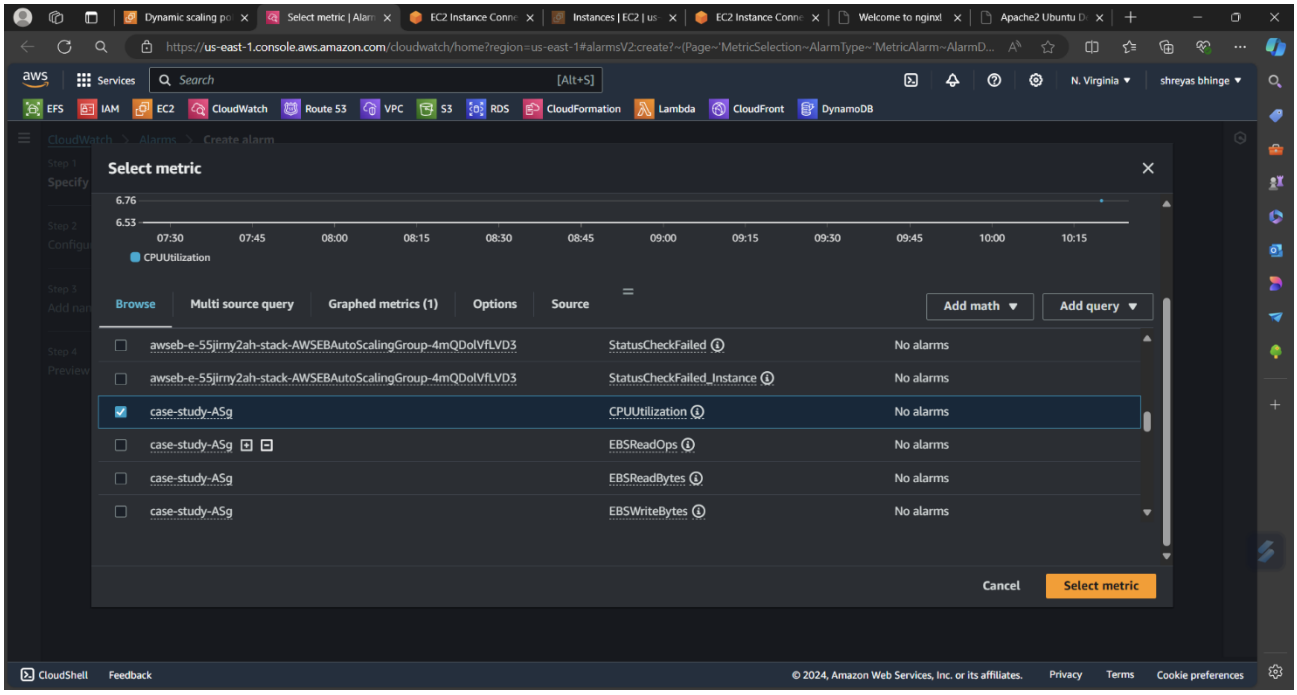

Create Auto Scaling grou
EC2 Instance Connect | u
Instances | EC2 | us-east
EC2 Instance Connect | u
Welcome to nginx!
Apache2 Ubuntu Default
https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateAutoScalingGroup:
Services
Search
[Alt+S]
N. Virginia
shreyas bhinge
EFS
IAM
EC2
CloudWatch
Route 53
VPC
S3
RDS
CloudFormation
Lambda
CloudFront
DynamoDB
Step 3 - optional
Configure advanced options
Step 4 - optional
Configure group size and scaling
Step 5 - optional
Add notifications
Step 6 - optional
Add tags
Step 7
Review
Set the initial size of the Auto Scaling group. After creating the group, you can change its size to meet demand, either manually or by using automatic scaling.
Desired capacity type
Choose the unit of measurement for the desired capacity value. vCPUs and Memory(GiB) are only supported for mixed instances groups configured with a set of instance attributes.
Units (number of instances)
Desired capacity
Specify your group size.
3
Scaling
Info
You can resize your Auto Scaling group manually or automatically to meet changes in demand.
Scaling limits
Set limits on how much your desired capacity can be increased or decreased.
Min desired capacity
1
Equal or less than desired capacity
Max desired capacity
3
Equal or greater than desired capacity
Automatic scaling - optional
Choose whether to use a target tracking policy
Info
You can set up other metric-based scaling policies and scheduled scaling after creating your Auto Scaling group.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

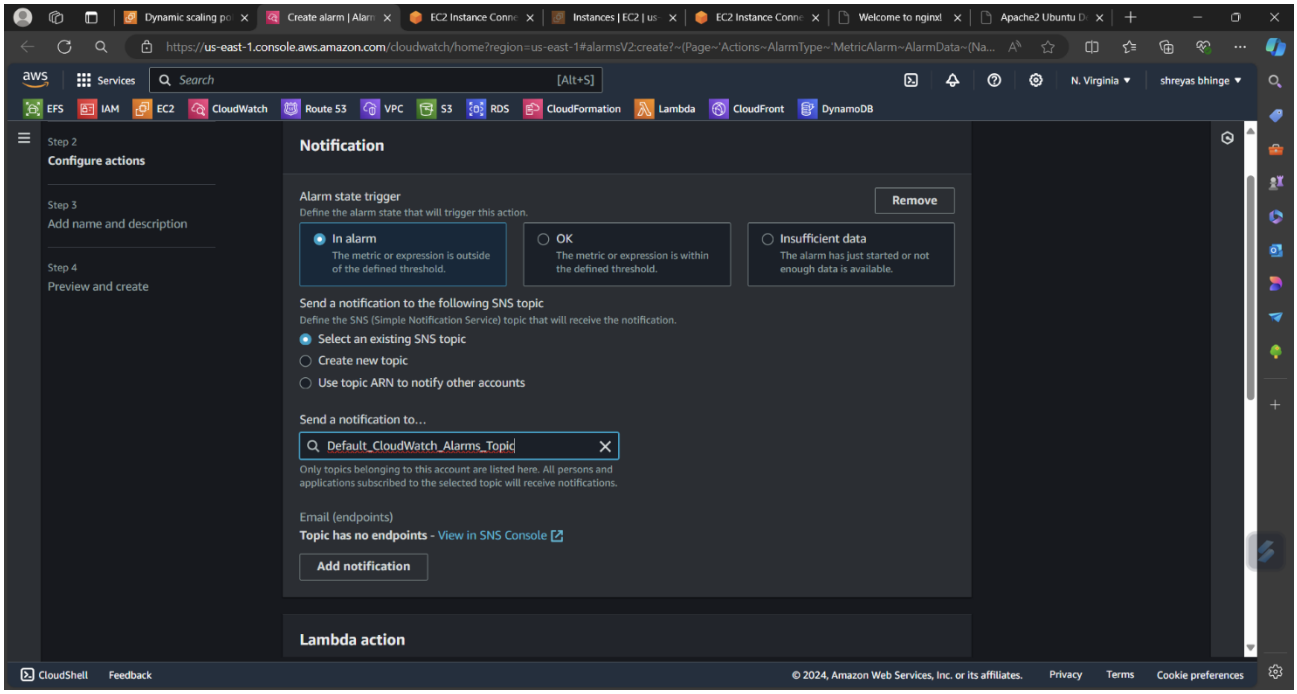

Instance scale-in protection
Instance scale-in protection
Enable instance protection from scale in
Step 5: Add notifications
Edit
Notifications
No notifications
Step 6: Add tags
Edit
Tags (0)
Key
Value
Tag new instances
No tags
Cancel
Previous
Create Auto Scaling group

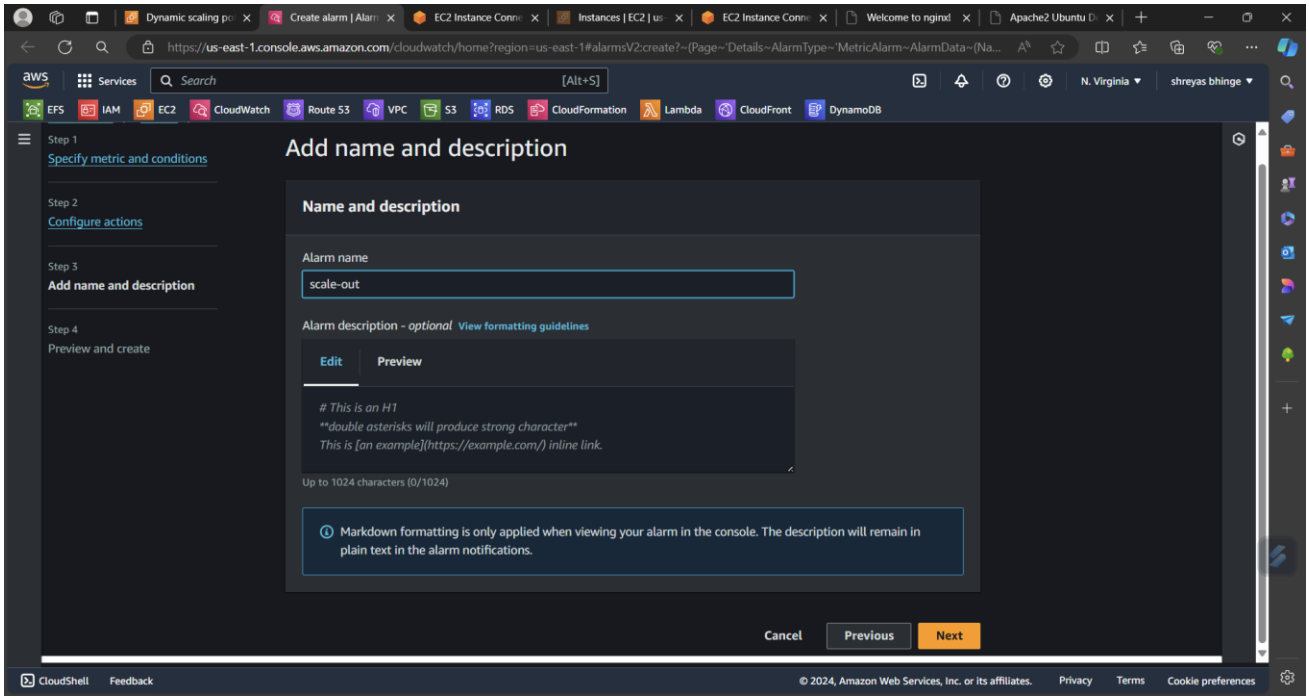

EC2 > Auto Scaling groups
Auto Scaling groups (1) Info
Launch configurations
Launch templates
Actions
Create Auto Scaling group
Search your Auto Scaling groups
Name
Launch template/configuration
Instances
Status
Desired capacity
Min
Max
case-study-ASg
my-case-study-temp | Version Default
3
-
3
1
3
0 Auto Scaling groups selected

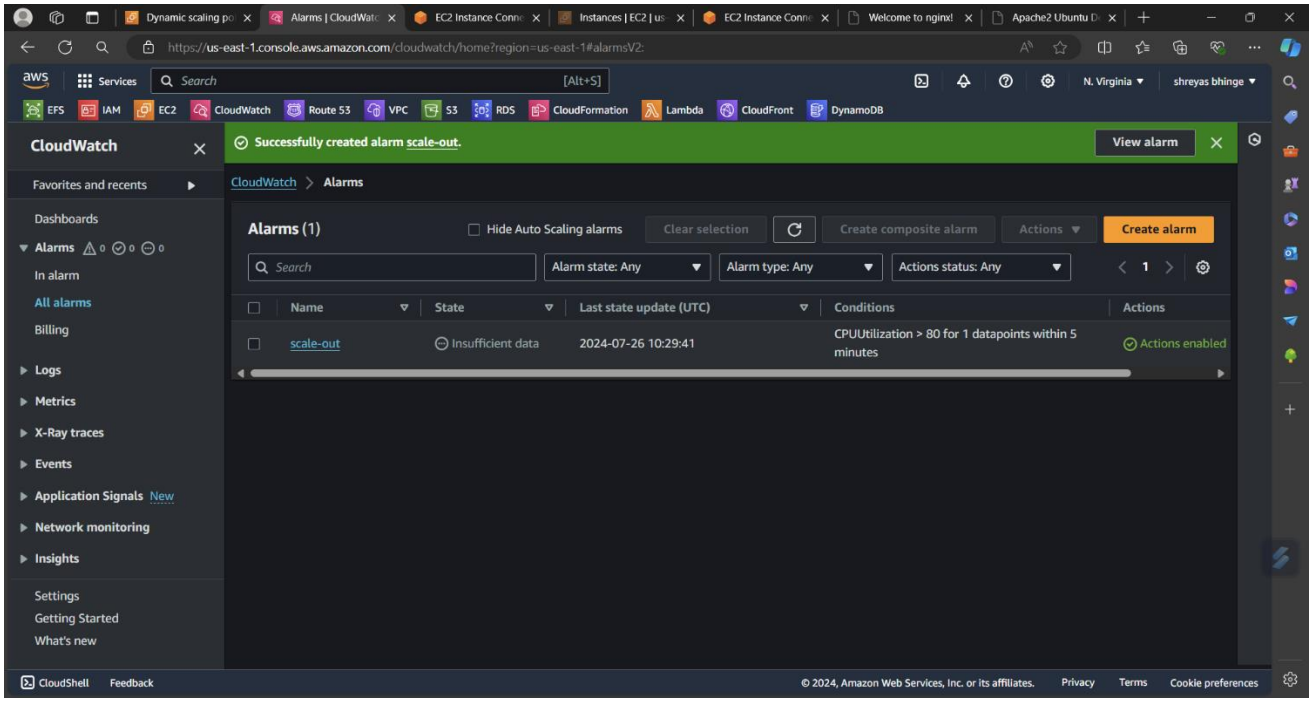

Auto Scaling group detai
EC2 Instance Connect | u
Instances | EC2 | us-east
EC2 Instance Connect | u
Welcome to nginx!
Apache2 Ubuntu Default
https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#AutoScalingGroupDetails:id=case-study-ASg;view=scaling
aws
Services
Search
[Alt+S]
N. Virginia
shreyas bhinge
EFS
IAM
EC2
CloudWatch
Route 53
VPC
S3
RDS
CloudFormation
Lambda
CloudFront
DynamoDB
EC2 > Auto Scaling groups > case-study-ASg
case-study-ASg
Details
Activity
Automatic scaling
Instance management
Monitoring
Instance refresh
Scaling policies resize your Auto Scaling group to meet changes in demand. With reactive dynamic scaling policies, you can track specific CloudWatch metrics and take action when the CloudWatch alarm threshold is met. Use predictive scaling policies along with dynamic scaling policies in the following situations: when your application demand changes quickly, but with a recurring pattern, or when your EC2 instances require more time to initialize.
Dynamic scaling policies (0) Info
Actions
Create dynamic scaling policy
1
No dynamic scaling policies have been created
Dynamic scaling policies use real-time data to scale your group based on configurable metrics.
Create dynamic scaling policy
Predictive scaling policies (0) Info
Actions
Create predictive scaling policy
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Dynamic scaling po
Create alarm | Alarm
EC2 Instance Conne
Instances | EC2 | us
EC2 Instance Conne
Welcome to nginx!
Apache2 Ubuntu D
https://us-east-1.console.aws.amazon.com/cloudwatch/home?region=us-east-1#alarmsV2:create?~(Page~'MetricSelection~AlarmType~'MetricAlarm~AlarmD...
aws
Services
Search
[Alt+S]
N. Virginia
shreyas bhinge
EFS
IAM
EC2
CloudWatch
Route 53
VPC
S3
RDS
CloudFormation
Lambda
CloudFront
DynamoDB
CloudWatch > Alarms > Create alarm
Step 1
Specify metric and conditions
Step 2
Configure actions
Step 3
Add name and description
Step 4
Preview and create
Specify metric and conditions
Metric
Graph
Preview of the metric or metric expression and the alarm threshold.
Select metric
Cancel
Next
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Select metric
6.76
6.53
07:30
07:45
08:00
08:15
08:30
08:45
09:00
09:15
09:30
09:45
10:00
10:15
CPUUtilization
Browse
Multi source query
Graphed metrics (1)
Options
Source
Add math
Add query
awseb-e-55jirny2ah-stack-AWSEBAutoScalingGroup-4mQDolVfLVD3
StatusCheckFailed
No alarms
awseb-e-55jirny2ah-stack-AWSEBAutoScalingGroup-4mQDolVfLVD3
StatusCheckFailed_Instance
No alarms
case-study-ASg
CPUUtilization
No alarms
case-study-ASg
EBSReadOps
No alarms
case-study-ASg
EBSReadBytes
No alarms
case-study-ASg
EBSWriteBytes
No alarms
Cancel
Select metric
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

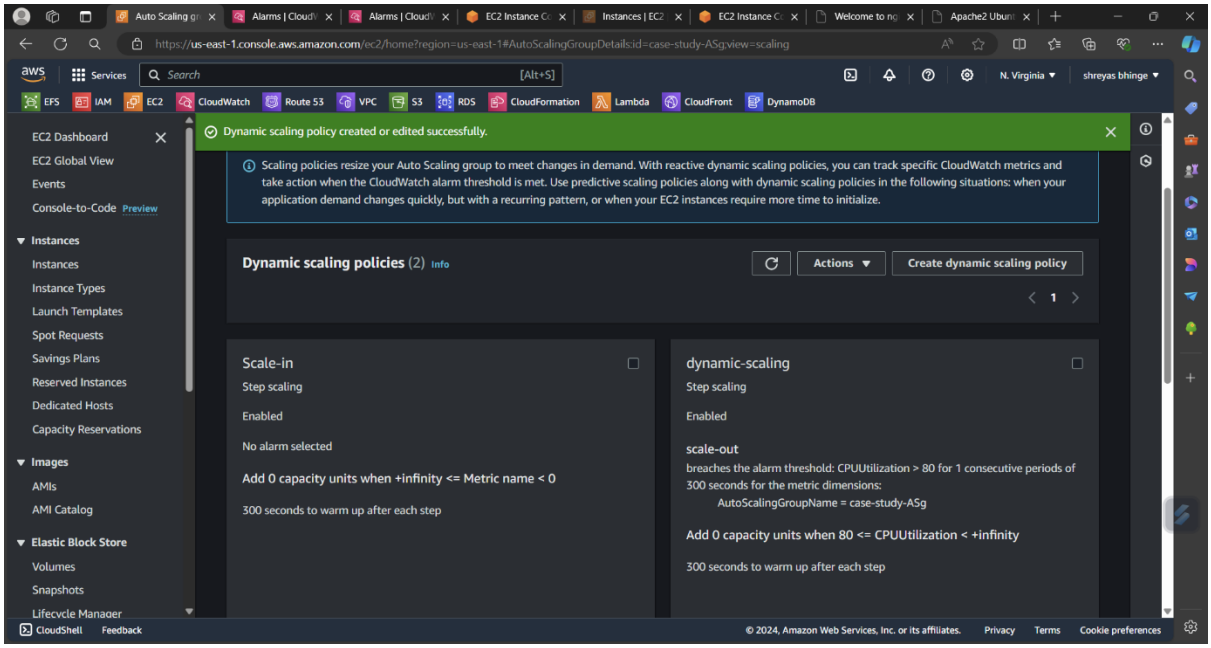

Step 2
Configure actions
Step 3
Add name and description
Step 4
Preview and create
Notification
Alarm state trigger
Define the alarm state that will trigger this action.
Remove
In alarm
The metric or expression is outside of the defined threshold.
OK
The metric or expression is within the defined threshold.
Insufficient data
The alarm has just started or not enough data is available.
Send a notification to the following SNS topic
Define the SNS (Simple Notification Service) topic that will receive the notification.
Select an existing SNS topic
Create new topic
Use topic ARN to notify other accounts
Send a notification to...
Default_CloudWatch_Alarms_Topic
Only topics belonging to this account are listed here. All persons and applications subscribed to the selected topic will receive notifications.
Email (endpoints)
Topic has no endpoints - View in SNS Console
Add notification
Lambda action
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences





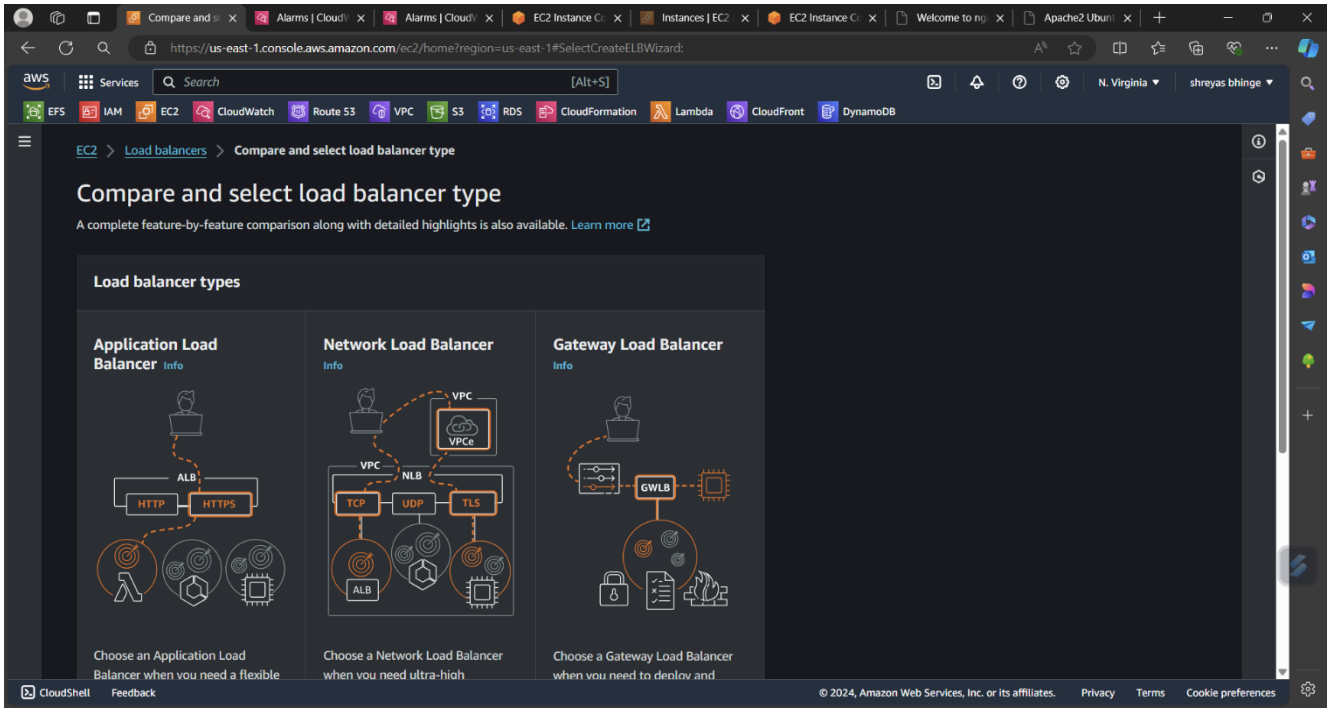
b. Removing the resources when the CPU utilization goes under 60%

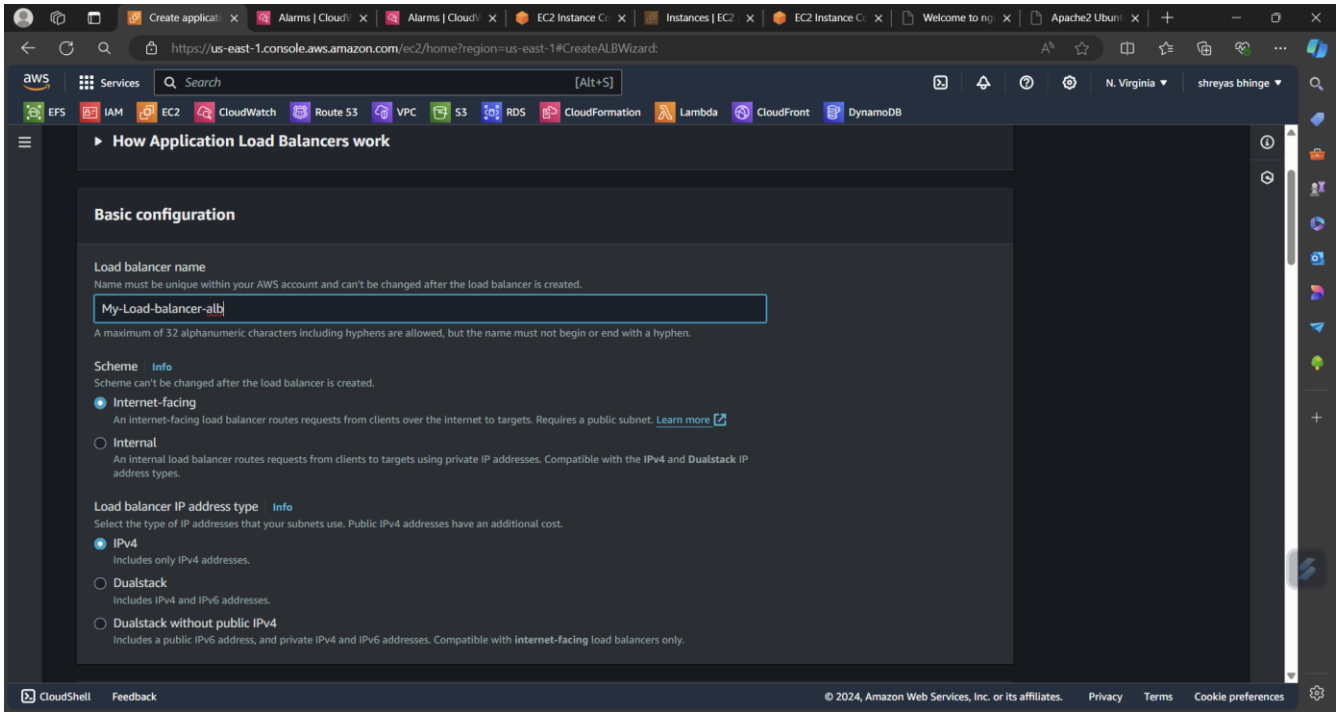
Follow the same steps from dynamic scaling for scale in

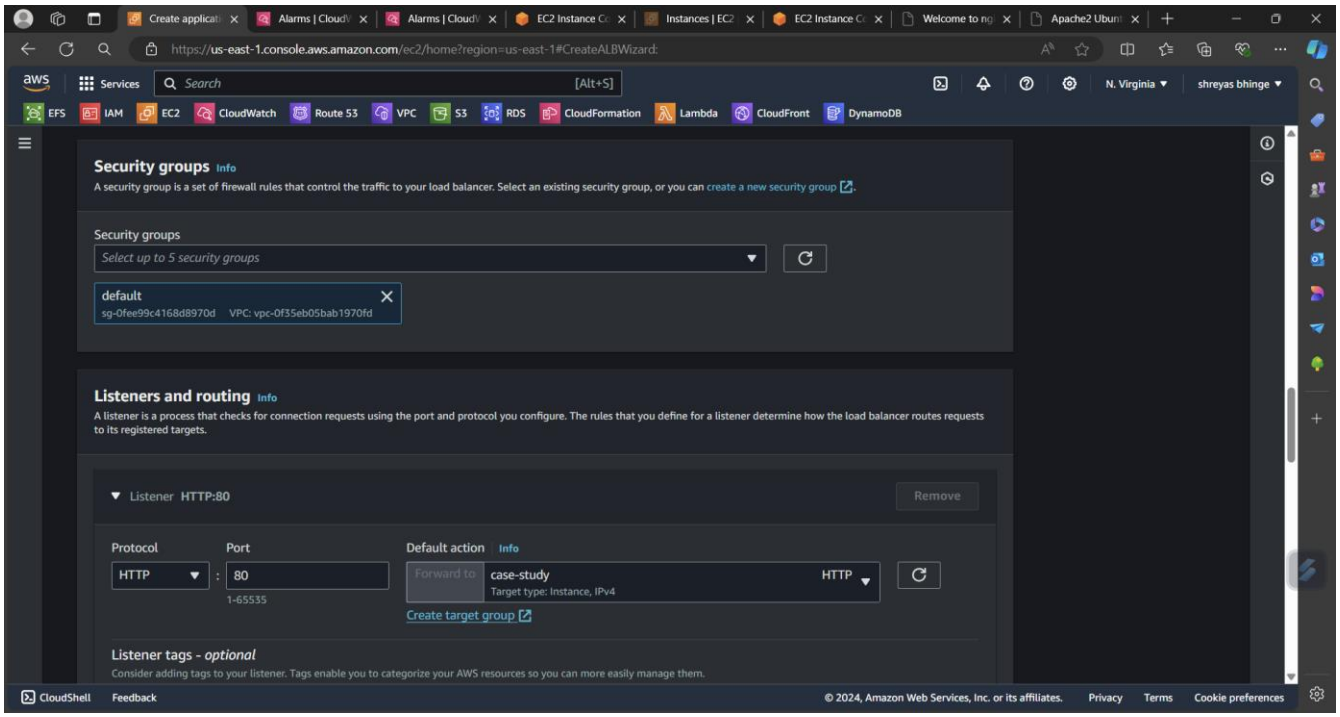


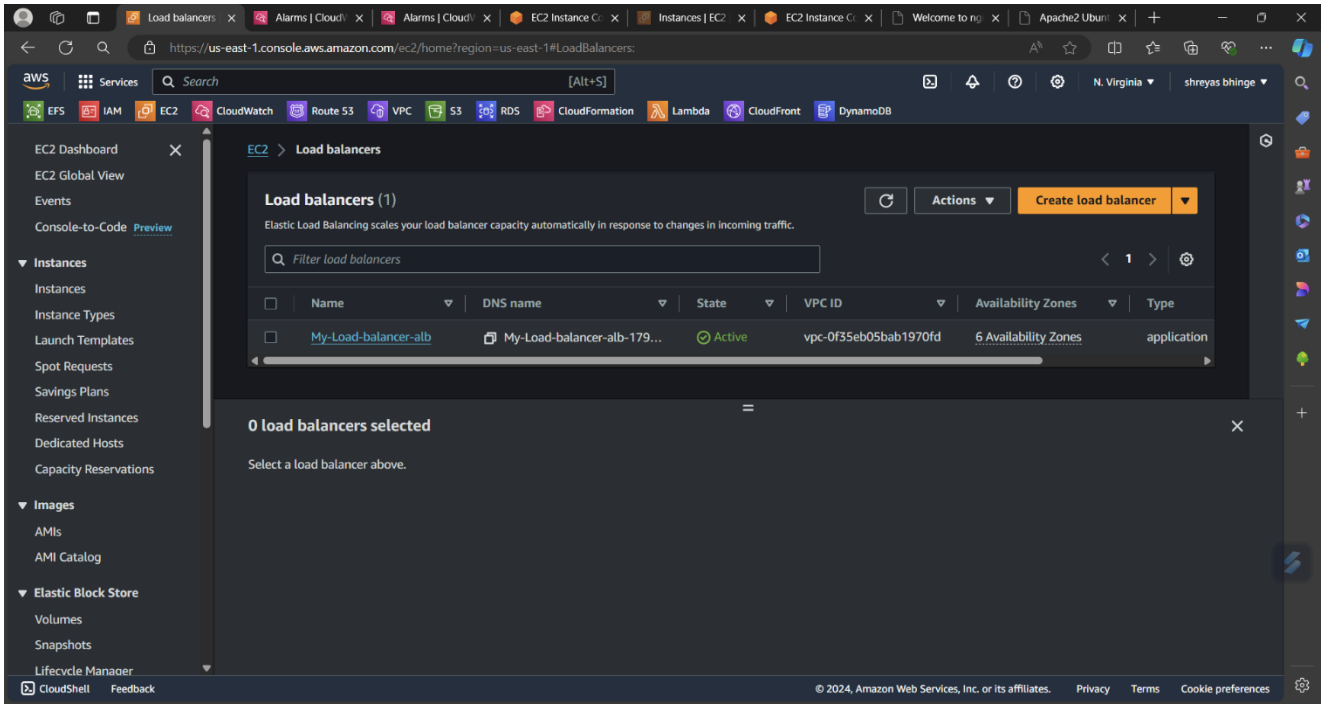
2.Create a load balancer to distribute the load between the compute resource

- Create a load balancer (Application L.B)

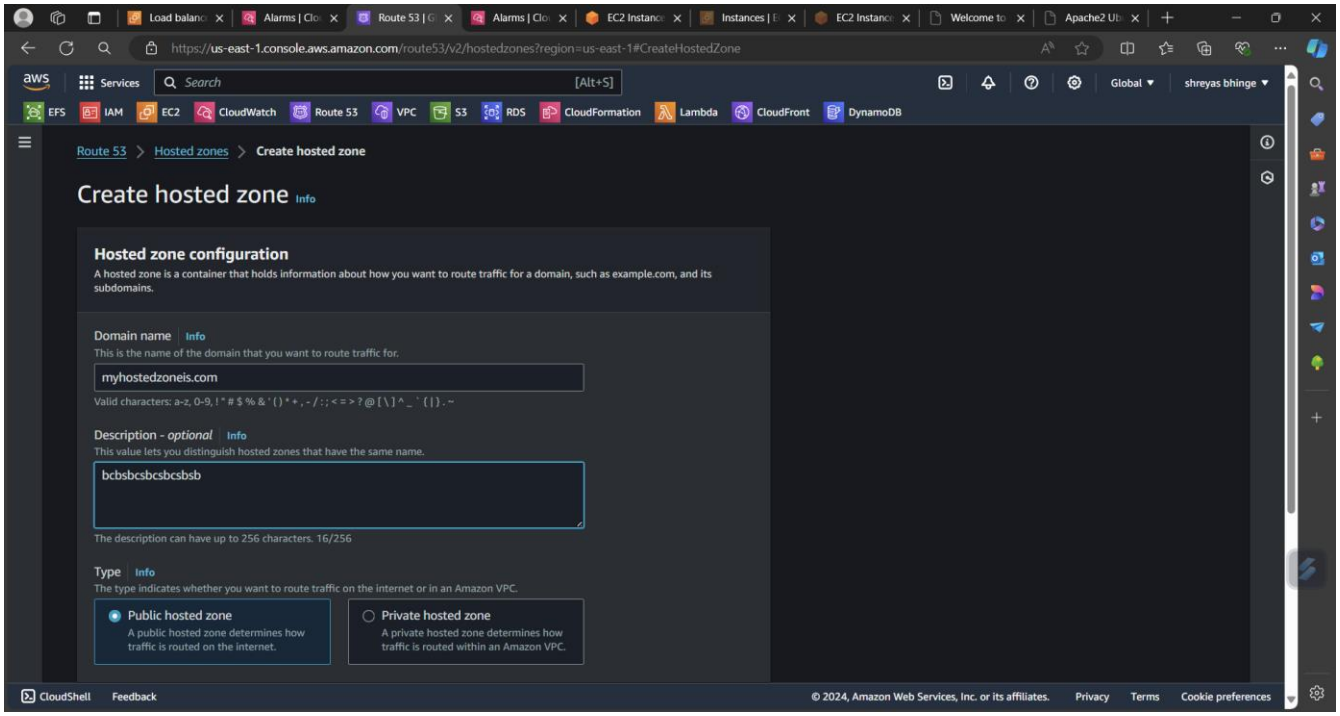


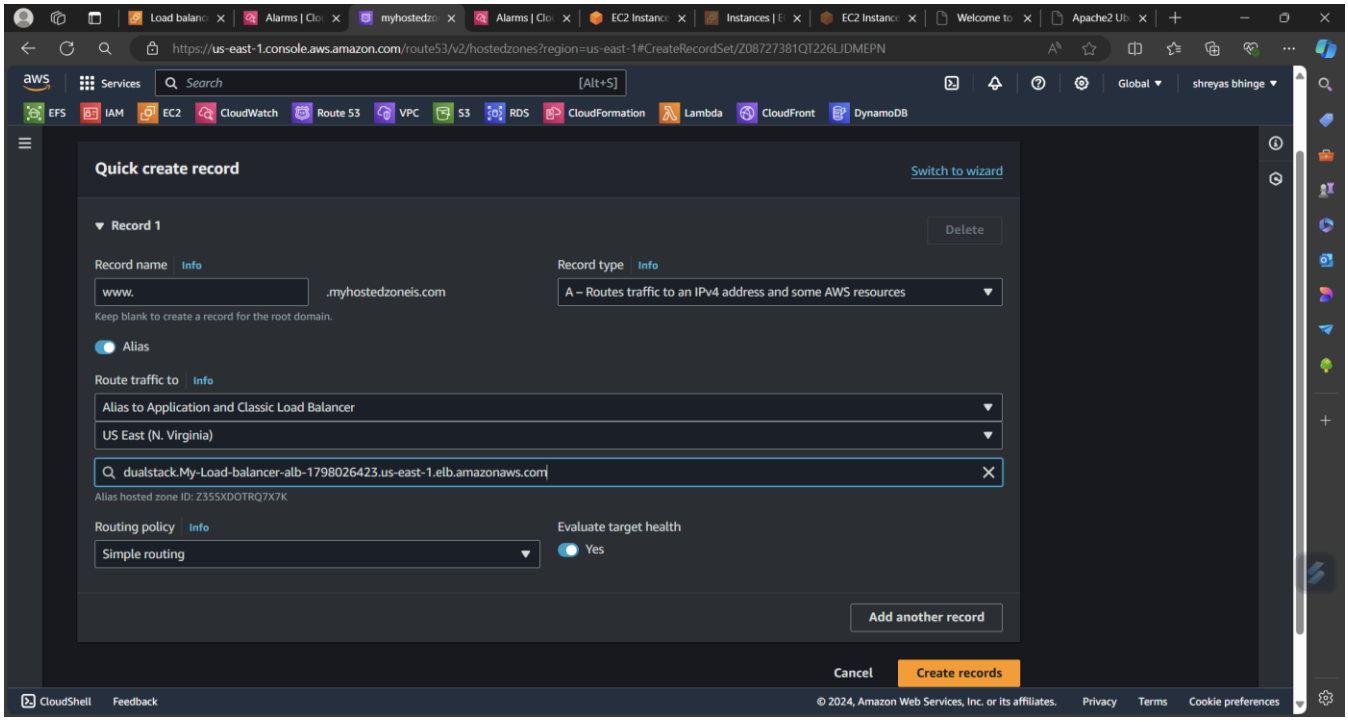

https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateALBWizard:
Search
[Alt+S]
N. Virginia
shreyas bhinge
How Application Load Balancers work
Basic configuration
Load balancer name
Name must be unique within your AWS account and can't be changed after the load balancer is created.
My-Load-balancer-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. Compatible with the IPv4 and Dualstack IP address types.
Load balancer IP address type Info
Select the type of IP addresses that your subnets use. Public IPv4 addresses have an additional cost.
IPv4
Includes only IPv4 addresses.
Dualstack
Includes IPv4 and IPv6 addresses.
Dualstack without public IPv4
Includes a public IPv6 address, and private IPv4 and IPv6 addresses. Compatible with internet-facing load balancers only.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

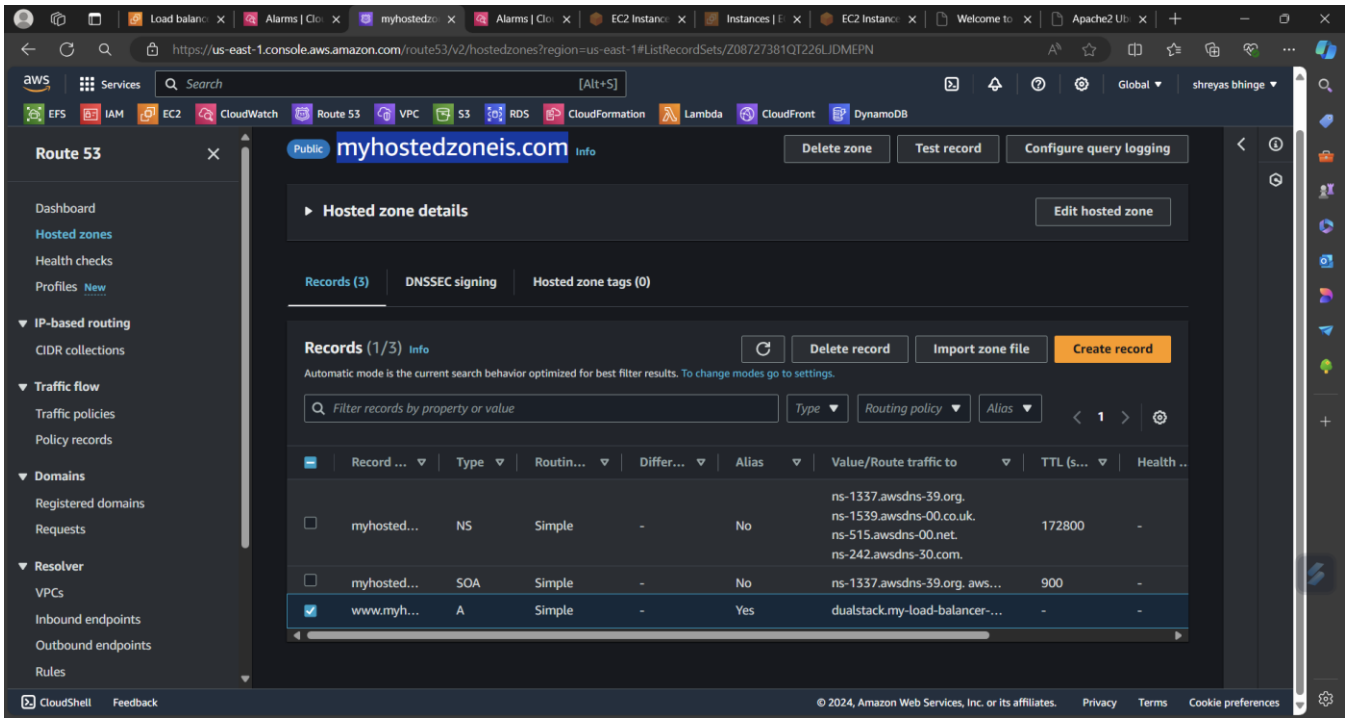

https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateALBWizard:
Search
[Alt+S]
N. Virginia
shreyas bhinge
Security groups Info
A security group is a set of firewall rules that control the traffic to your load balancer. Select an existing security group, or you can create a new security group.
Security groups
Select up to 5 security groups
default
sg-0fee99c4168d8970d VPC: vpc-0f35eb05bab1970fd
Listeners and routing Info
A listener is a process that checks for connection requests using the port and protocol you configure. The rules that you define for a listener determine how the load balancer routes requests to its registered targets.
Listener HTTP:80
Remove
Protocol
HTTP
Port
80
1-65535
Default action Info
Forward to
case-study
HTTP
Target type: Instance, IPv4
Create target group
Listener tags - optional
Consider adding tags to your listener. Tags enable you to categorize your AWS resources so you can more easily manage them.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences



3. Route traffic to the company
Now go to route 53
Select hosted zone







- Hit the Domain after few hours