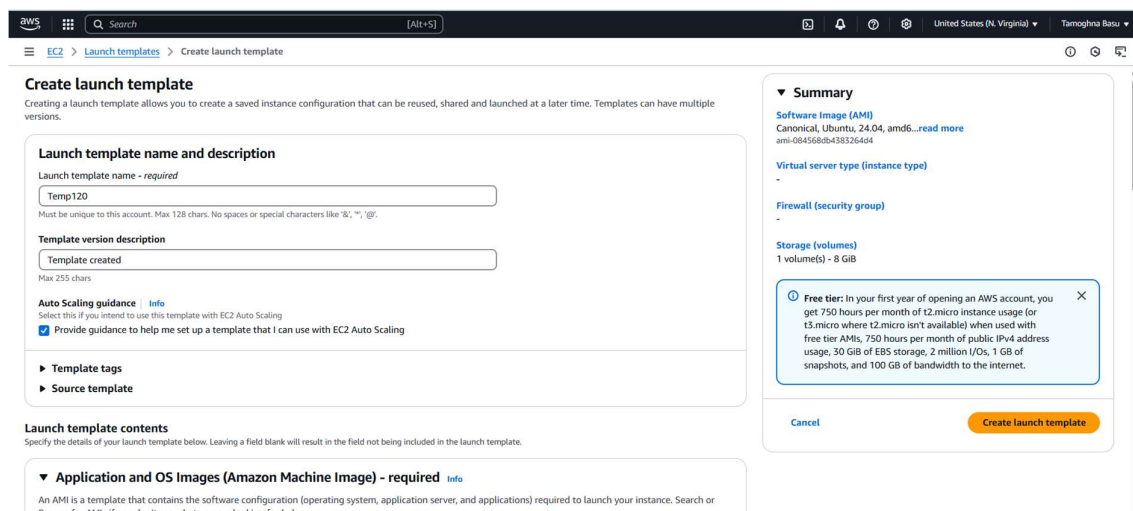


# Assignment-11

Problem Definition: Build scaling plans in AWS that balance the load on different EC2 instance.

**Step-1:** Signing in to the AWS Management console, go to launch template. A template name should be given along with a description and the auto-scaling option is to be checked. Then ubuntu is selected from quickstart.



The screenshot shows the 'Create launch template' page in the AWS Management Console. The page is titled 'Create launch template' and includes a sub-header 'Creating a launch template allows you to create a saved instance configuration that can be reused, shared and launched at a later time. Templates can have multiple versions.'

The main form is divided into several sections:

- Launch template name and description:** A text input field contains 'Temp120'. Below it, a note states: 'Must be unique to this account. Max 128 chars. No spaces or special characters like '&', "'", '@'.'
- Template version description:** A text input field contains 'Template created'. Below it, a note states: 'Max 255 chars'.
- Auto Scaling guidance:** A checkbox labeled 'Provide guidance to help me set up a template that I can use with EC2 Auto Scaling' is checked.
- Template tags:** A section with a plus icon and the text 'Source template'.
- Launch template contents:** A section with a plus icon and the text 'Application and OS Images (Amazon Machine Image) - required'. Below it, a note states: 'An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or'.

On the right side, there is a 'Summary' panel with the following details:

- Software Image (AMI):** Canonical, Ubuntu, 24.04, amd64...read more  
ami-084568db433316404
- Virtual server type (instance type):** -
- Firewall (security group):** -
- Storage (volumes):** 1 volume(s) - 8 GiB

At the bottom right, there is a 'Free tier' notification: 'Free tier: In your first year of opening an AWS account, you get 750 hours per month of t2.micro instance usage (or t3.micro where t2.micro isn't available) when used with free tier AMIs, 750 hours per month of public IPv4 address usage, 30 GiB of EBS storage, 2 million I/Os, 1 GiB of snapshots, and 100 GiB of bandwidth to the internet.'

At the bottom, there are two buttons: 'Cancel' and 'Create launch template'.

**Step-2:** Instance type is selected as t2.micro.

Search your full catalog including 1000s of application and OS images

Recent Quick Start

Amazon Linux macOS Ubuntu Windows Red Hat SUSE Linux Debian

Amazon Machine Image (AMI)

Ubuntu Server 24.04 LTS (HVM), SSD Volume Type  
ami-084568db438326404 (64-bit x86) / ami-044709339f08521a (64-bit ARM)  
Virtualization: hvm, ENX: enabled, true Root device type: ebs

Description

Ubuntu Server 24.04 LTS (HVM), EBS General Purpose (SSD) Volume Type. Support available from Canonical (<http://www.ubuntu.com/cloud/services>).

Canonical, Ubuntu, 24.04, amd64 noble image

Architecture AMI ID Publish Date Username Verified provider

64-bit (x86) ami-084568db438326404 2025-03-05 ubuntu

Free tier eligible

▼ Instance type Info | Get advice Advanced

Instance type

t2.micro  
Family: t2 1 vCPU 1 GiB Memory Current generation: true On-Demand Windows base pricing: 0.0162 USD per Hour Free tier eligible  
On-Demand Ubuntu Pro base pricing: 0.0134 USD per Hour On-Demand SUSE base pricing: 0.0116 USD per Hour  
On-Demand RHEL base pricing: 0.026 USD per Hour On-Demand Linux base pricing: 0.0116 USD per Hour

Compare Instance types

▼ Summary

Software image (AMI)  
Canonical, Ubuntu, 24.04, amd64...read more  
ami-084568db438326404

Virtual server type (instance type)  
t2.micro

Firewall (security group)  
-

Storage (volumes)  
1 volume(s) - 8 GiB

Free tier: In your first year of opening an AWS account, you get 750 hours per month of t2.micro instance usage (or t3.micro where t2.micro isn't available) when used with free tier AMIs, 750 hours per month of public IPv4 address usage, 30 GiB of EBS storage, 2 million I/Os, 1 GiB of snapshots, and 100 GiB of bandwidth to the internet.

Cancel Create launch template

Step-3: A key-pair is selected along with the security group and the user data is given. Create launch template is selected to create a new template.

Additional costs apply for AMIs with pre-installed software

▼ Key pair (login) Info

You can use a key pair to securely connect to your instance. Ensure that you have access to the selected key pair before you launch the instance.

Key pair name  
key3 Create new key pair

▼ Network settings Info

Subnet  
Don't include in launch template Create new subnet

When you specify a subnet, a network interface is automatically added to your template.

Firewall (security groups)  
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.

Select existing security group Create security group

Security groups  
Select security groups

default: sg-0d46f1de5afa5934c VPC: vpc-0d57a2959495f22a Compare security group rules

▼ Storage (volumes) Info

▼ Summary

Software image (AMI)  
Canonical, Ubuntu, 24.04, amd64...read more  
ami-084568db438326404

Virtual server type (instance type)  
t2.micro

Firewall (security group)  
default

Storage (volumes)  
1 volume(s) - 8 GiB

Free tier: In your first year of opening an AWS account, you get 750 hours per month of t2.micro instance usage (or t3.micro where t2.micro isn't available) when used with free tier AMIs, 750 hours per month of public IPv4 address usage, 30 GiB of EBS storage, 2 million I/Os, 1 GiB of snapshots, and 100 GiB of bandwidth to the internet.

Cancel Create launch template

**Summary**

**Software image (AMI)**  
Canonical, Ubuntu, 24.04, amd64...  
ami-08456db4383264d4

**Virtual server type (instance type)**  
t2.micro

**Firewall (security group)**  
default

**Storage (volumes)**  
1 volume(s) - 8 GiB

**Free tier:** In your first year of opening an AWS account, you get 750 hours per month of t2.micro instance usage (or t3.micro where t2.micro isn't available) when used with free tier AMIs. 750 hours per month of public IPv4 address usage, 30 GiB of EBS storage, 2 million I/Os, 1 GB of snapshots, and 100 GB of bandwidth to the internet.

**Success**  
Successfully created Temp120lt-07426884fc689de9f.

**Next Steps**

**Launch an instance**  
With On-Demand Instances, you pay for compute capacity by the second (for Linux, with a minimum of 60 seconds) or by the hour (for all other operating systems) with no long-term commitments or upfront payments. Launch an On-Demand instance from your launch template.  
[Launch instance from this template](#)

**Create an Auto Scaling group from your template**  
Amazon EC2 Auto Scaling helps you maintain application availability and allows you to scale your Amazon EC2 capacity up or down automatically according to conditions you define. You can use Auto Scaling to help ensure that you are running your desired number of Amazon EC2 instances during demand spikes to maintain performance and decrease capacity during lulls to reduce costs.  
[Create Auto Scaling group](#)

**Create Spot Fleet**  
A Spot Instance is an unused EC2 instance that is available for less than the On-Demand price. Because Spot Instances enable you to request unused EC2 instances at steep discounts, you can lower your Amazon EC2 costs significantly. The hourly price for a Spot Instance (of each instance type in each Availability Zone) is set by Amazon EC2, and adjusted gradually based on the long-term supply of and demand for Spot Instances. Spot instances are well-suited for data-analysis, batch jobs, background processing, and optional tasks.  
[Create Spot Fleet](#)

[View launch templates](#)

Step-4:Auto-scaling group is reached ,a name is given,template is selected,zones are selected,along with Balanced best effort.

Step 1: Choose launch template

**Choose launch template**

Specify a launch template that contains settings common to all EC2 instances that are launched by this Auto Scaling group.

**Name**

Auto Scaling group name

Enter a name to identify the group.

Auto

Must be unique to this account in the current Region and no more than 255 characters.

**Launch template**

Choose a launch template that contains the instance-level settings, such as the Amazon Machine Image (AMI), instance type, key pair, and security groups.

Temp120

Create a launch template

**Version**

Default (1)

Create a launch template version

**Description**

Template created

**Launch template**

Temp120

lt-07426884f3c689de9

**Instance type**

t2.micro

**Security groups**

-

**AMI ID**

ami-084568db4381264d4

**Request Spot Instances**

No

Step 2: Add Tags

**Network**

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-0d57ad950495f022a

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-0f064e0b82b84

172.31.0.0/20 Default

us-east-1b | subnet-0a9ca49076c52510d

172.31.80.0/20 Default

us-east-1c | subnet-009aba00b760ee48

172.31.16.0/20 Default

Create a subnet

**Availability Zone distribution - new**

Auto Scaling automatically balances instances across Availability Zones. If launch failures occur in a zone, select a strategy.

☒ **Balanced best effort**

If launches fail in one Availability Zone, Auto Scaling will attempt to launch in another healthy Availability Zone.

☐ **Balanced only**

If launches fail in one Availability Zone, Auto Scaling will continue to attempt to launch in the unhealthy Availability Zone to preserve balanced distribution.

Cancel Skip to review Previous Next

**Step-5:**On the next page,attach to a new load balancer is selected along with Internet-facing as the load balancer scheme.Additional health check types is enabled.

Step 1

Choose launch template

Step 2

Choose instance launch options

Step 3 - optional

Integrate with other services

Step 4 - optional

Configure group size and scaling

Step 5 - optional

Add notifications

Step 6 - optional

Add tags

Step 7

Review

### Integrate with other services - optional

Use a load balancer to distribute network traffic across multiple servers. Enable service-to-service communications with VPC Lattice. Shift resources away from impaired Availability Zones with zonal shift. You can also customize health check replacements and monitoring.

#### Load balancing

Use the options below to attach your Auto Scaling group to an existing load balancer, or to a new load balancer that you define.

☐ No load balancer  
Traffic to your Auto Scaling group will not be fronted by a load balancer.

☐ Attach to an existing load balancer  
Choose from your existing load balancers.

☒ Attach to a new load balancer  
Quickly create a basic load balancer to attach to your Auto Scaling group.

#### Attach to a new load balancer

Define a new load balancer to create for attachment to this Auto Scaling group.

**Load balancer type**  
Choose from the load balancer types offered below. Type selection cannot be changed after the load balancer is created. If you need a different type of load balancer than those offered here, visit the [Load Balancing console](#).

☒ Application Load Balancer  
HTTP, HTTPS

☐ Network Load Balancer  
TCP, UDP, TLS

**Load balancer name**  
Name cannot be changed after the load balancer is created.

Auto-1

**Load balancer scheme**  
Scheme cannot be changed after the load balancer is created.

☐ Internal

☒ Internet-facing

**Network mapping**  
Your new load balancer will be created using the same VPC and Availability Zone selections as your Auto Scaling group. You can select different subnets and add subnets from additional Availability Zones.

VPC

Availability Zones and subnets

You must select a single subnet for each Availability Zone enrolled. Only public subnets are available for selection to support DNS resolution.

☒ us-east-1a

subnet-0afba49076e52510d

☒ us-east-1b

subnet-0f8ef064a8b82bd4

☒ us-east-1c

subnet-0c7ab8a00b760ee48

☐ us-east-1f

Select a subnet

☐ us-east-1e

Select a subnet

☐ us-east-1d

Select a subnet

**Listeners and routing**  
If you require secure listeners, or multiple listeners, you can configure them from the [Load Balancing console](#) after your load balancer is created.

**Protocol** **Port** **Default routing (forwarded)**

HTTP

4000

Create a target group

**New target group name**  
An instance target group with default settings will be created.

Auto-1

**Tips - optional**  
Consider adding tags to your load balancer. Tags enable you to categorize your AWS resources so you can more easily manage them.

[Add tag](#)

50 remaining

#### VPC Lattice integration options

To improve networking capabilities and scalability, integrate your Auto Scaling group with VPC Lattice. VPC Lattice facilitates communications between AWS services and helps you connect and manage your applications across compute services in AWS.

☒ No VPC Lattice service  
VPC Lattice will not manage your Auto Scaling group's network access and connectivity with other services.

☐ Attach to VPC Lattice service  
Routing requests associated with specified VPC Lattice target groups will be routed to your Auto Scaling group.

[Create new VPC Lattice service](#)

VPC Lattice integration options

To improve networking capabilities and scalability, integrate your Auto Scaling group with VPC Lattice. VPC Lattice facilitates communications between AWS services and helps you connect and manage your applications across compute services in AWS.

☒ No VPC Lattice service  
VPC Lattice will not manage your Auto Scaling group's network access and connectivity with other services.

☐ Attach to VPC Lattice service  
Routing requests associated with specified VPC Lattice target groups will be routed to your Auto Scaling group.

[Create new VPC Lattice service](#)

#### Application Recovery Controller (ARC) zonal shift - new

During an Availability Zone impairment, target instance launches towards other healthy Availability Zones.

☐ Enable zonal shift  
New instance launches will be integrated towards healthy Availability Zones until the zonal shift is canceled.

#### 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**

☒ Turn on Elastic Load Balancing health checks  
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.

☐ Turn on Amazon EBS health checks  
EBS monitors whether an instance's root volume or attached volume status. When it reports an unhealthy volume, EC2 Auto Scaling can replace the instance on its next periodic health check.

**Health check grace period**  
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.

300 seconds

Cancel

[Skip to review](#)

[Previous](#)

[Next](#)

**Step-6:**In configure group and scaling,desired capacity is selected as 2,along with min value as 2 and max as 3.Target tracking scaling policy is selected.Instance warmup is set for 200 sec.

The screenshot shows the 'Configure group size and scaling' step in the AWS Management Console. On the left, a navigation pane lists steps 1 through 7, with 'Configure group size and scaling' selected. The main content area is titled 'Configure group size and scaling - optional' and includes a description: 'Define your group's desired capacity and scaling limits. You can optionally add automatic scaling to adjust the size of your group.'

**Group size**

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.

2

**Scaling**

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**

2

**Max desired capacity**

3

Equal or less than desired capacity

Equal or greater than desired capacity

**Automatic scaling - optional**

Choose whether to use a target tracking policy.

You can set up other metric-based scaling policies and scheduled scaling after creating your Auto Scaling group.

☐ No scaling policies

Your Auto Scaling group will remain at its initial size and will not dynamically resize to meet demand.

☒ Target tracking scaling policy

Choose a CloudWatch metric and target value and let the scaling policy adjust the desired capacity in proportion to the metric's value.

**Scaling policy name**

Target Tracking Policy

The screenshot shows the 'Instance maintenance policy' and 'Additional capacity settings' sections of the AWS Management Console. The 'Target tracking scaling policy' is selected, and the 'Metric type' is set to 'Average CPU utilization'. The 'Target value' is set to 50. The 'Instance warmup' is set to 200 seconds. The 'Disable scale in to create only a scale-out policy' checkbox is unchecked.

☐ No scaling policies

Your Auto Scaling group will remain at its initial size and will not dynamically resize to meet demand.

☒ Target tracking scaling policy

Choose a CloudWatch metric and target value and let the scaling policy adjust the desired capacity in proportion to the metric's value.

**Scaling policy name**

Target Tracking Policy

**Metric type**

Monitored metric that determines if resource utilization is too low or high. If using EC2 metrics, consider enabling detailed monitoring for better scaling performance.

Average CPU utilization

**Target value**

50

**Instance warmup**

200 seconds

☐ Disable scale in to create only a scale-out policy

**Instance maintenance policy**

Control your Auto Scaling group's availability during instance replacement events. This includes health checks, instance refreshes, maximum instance lifetime features and events that happen automatically to keep your group balanced, called rebalancing events.

**Choose a replacement behavior depending on your availability requirements**

**Mixed behavior**

☒ No policy

For rebalancing events, new instances will launch before terminating others. For all other events, instances terminate and launch at the same time.

**Priority availability**

☐ Launch before terminating

Launch new instances and wait for them to be ready before terminating others. This allows you to go above your desired capacity by a given percentage and may temporarily increase costs.

**Control costs**

☐ Terminate and launch

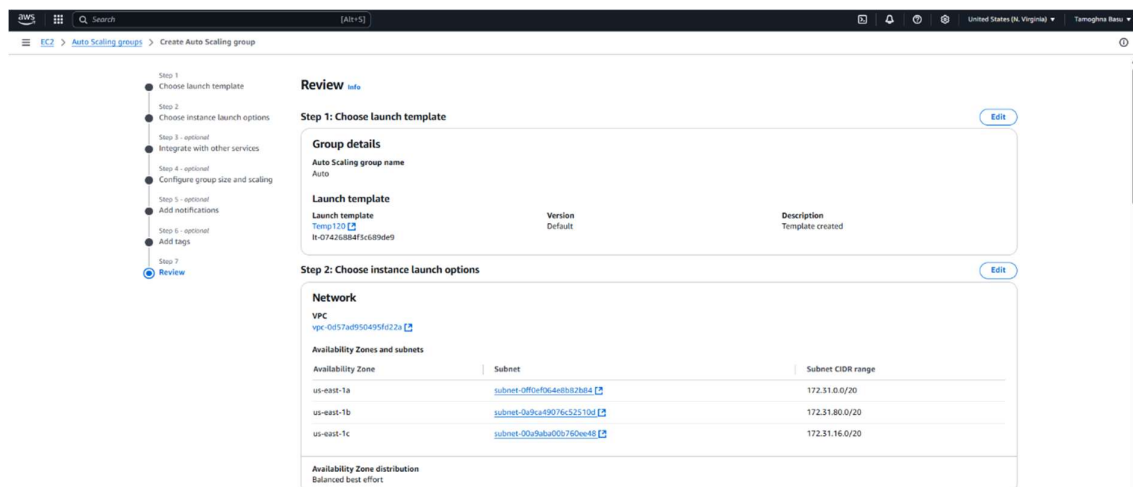
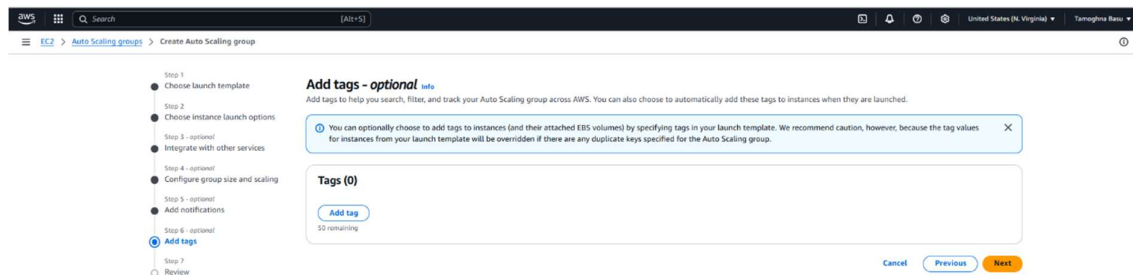
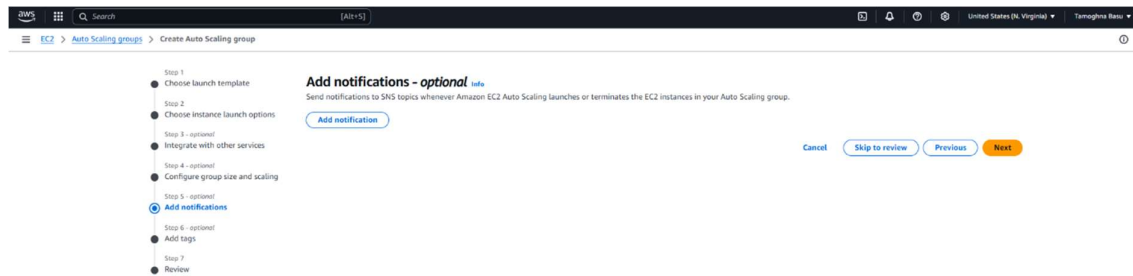
Terminate and launch instances at the same time. This allows you to go below your desired capacity by a given percentage and may temporarily reduce availability.

**Flexibility**

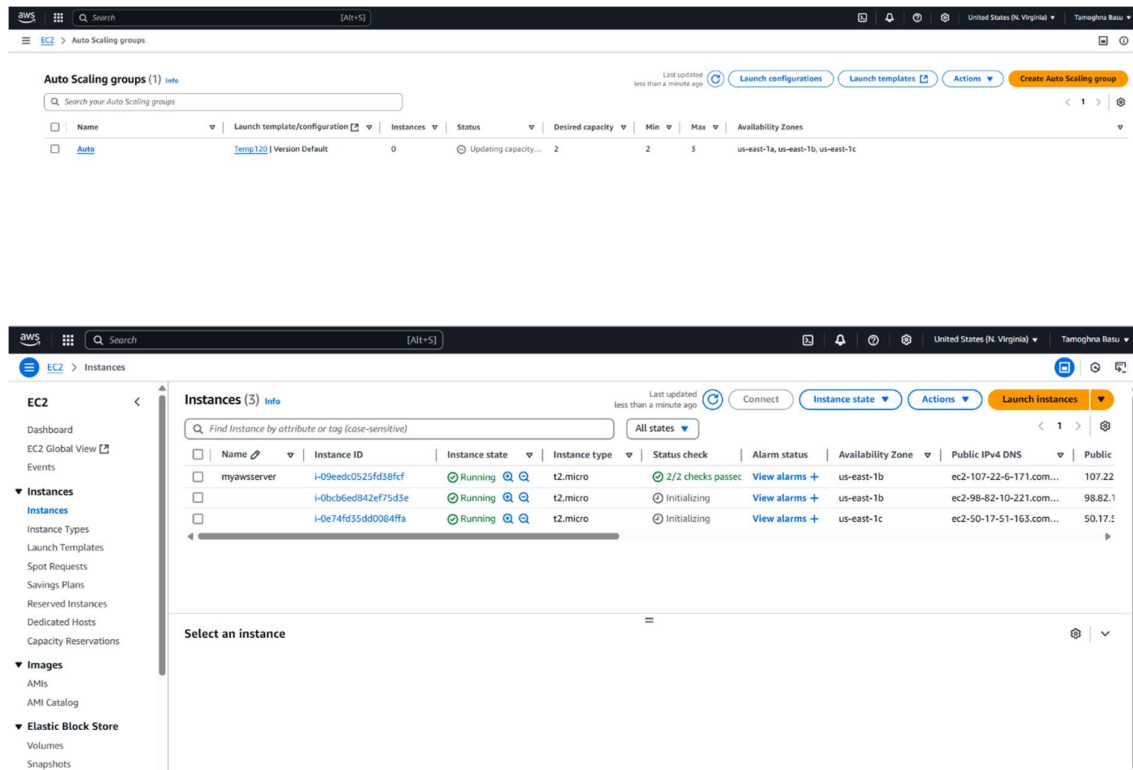
☐ Custom behavior

Set custom values for the minimum and maximum amount of available capacity. This gives you greater flexibility in setting how far below and over your desired capacity EC2 Auto Scaling goes when replacing instances.

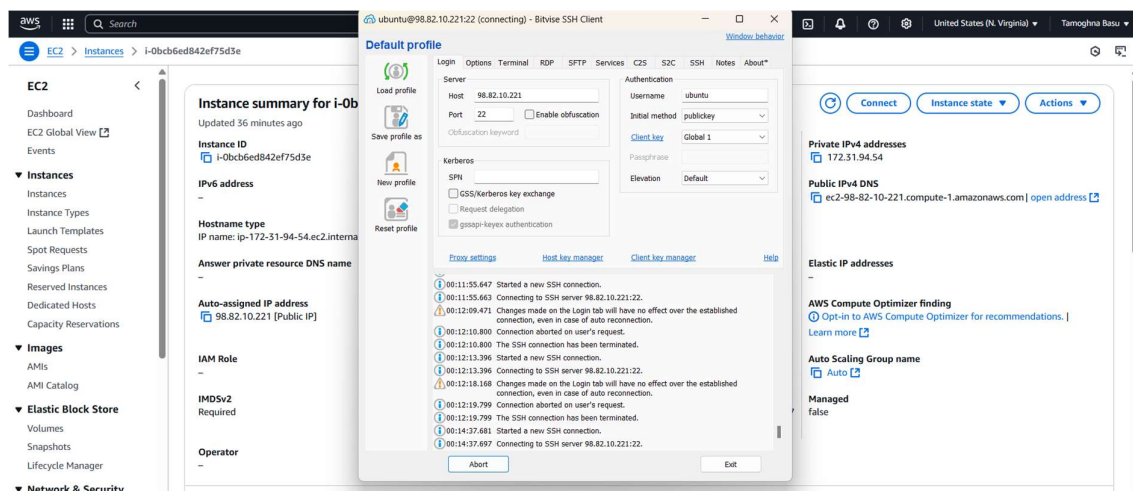
**Additional capacity settings**



Step-7: Without making further modifications, simply clicking on next ,finally the auto scaling group has been created.



Step-8:Bitwise is opened,properly logged in with the key-pair and in the terminal following commands are written.



//Sudo nano ri.sh



Inside the file we have to write

```
#!/bin/bash
```

```
while(true)
```

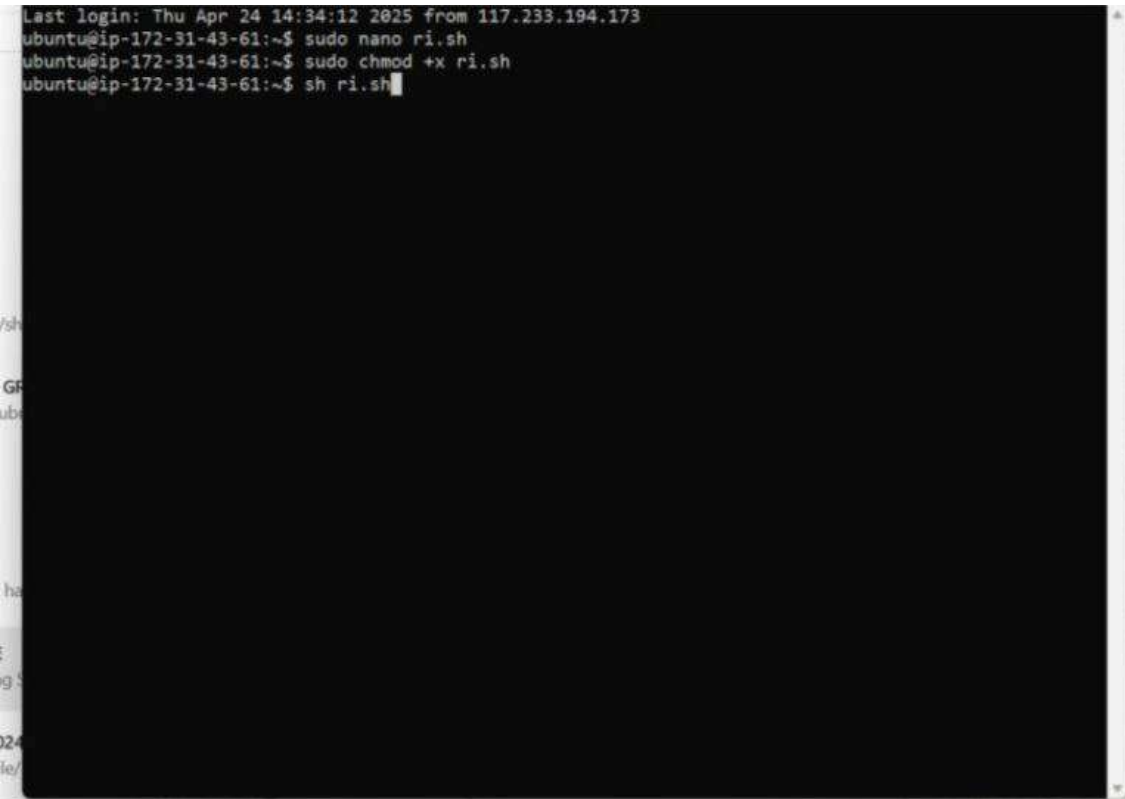
```
do
```

```
    echo "Inside loop"
```

```
done
```

```
//sudo chmod +x ri.sh
```

```
//sh ri.sh
```

A terminal window screenshot showing the execution of the script. The terminal output is as follows:

```
Last login: Thu Apr 24 14:34:12 2025 from 117.233.194.173
ubuntu@ip-172-31-43-61:~$ sudo nano ri.sh
ubuntu@ip-172-31-43-61:~$ sudo chmod +x ri.sh
ubuntu@ip-172-31-43-61:~$ sh ri.sh
```

The terminal window has a black background and a white border. The prompt is 'ubuntu@ip-172-31-43-61:~\$'. The command 'sh ri.sh' is being executed, and the terminal is currently blank, indicating the script is running in a loop.

```
GNU nano 7.2      ri.sh *
#!/bin/bash
while(true)
do
    echo "Inside loop"
done
```

024  
gle/^G Help  
^X Exit

^O Write Out  
^R Read File

^W Where Is  
^\_ Replace

[ Cancelled ]  
^K Cut  
^U Paste

^T Execute  
^J Justify

^C Location  
^\_ Go To Line

M-U Undo  
M-E Redo



Step-8:CPU utilization is checked on the monitoring section after the bash file is executed and the third instance is created as a result of it.

EC2

Dashboard

EC2 Global View

Events

Instances

Instance Types

Launch Templates

Spot Requests

Savings Plans

Reserved Instances

Dedicated Hosts

Capacity Reservations

Images

AMI

AMI Catalog

Elastic Block Store

Volumes

EBS Snapshots

Lifecycle Manager

Network & Security

Security Groups

Elastic IP

Placement Groups

Key Pairs

Network Interfaces

Load Balancing

Instances (2/3) help

Find instance by attribute or tag (none specified)

All status

Connect

Instance state

Actions

Launch instance

2 instances selected

Alert recommendations

Investigate with AI

15 30 120 360 5d Custom

Local timezone

Explore related

CPU utilization (%)

Percent

10.0

0.00

0

10.00

10.00

10.00

20.00

Network in (bytes)

Bytes

17.00

0.000

0

10.00

10.00

10.00

20.00

Network out (bytes)

Bytes

0.00

0.00

0

10.00

10.00

10.00

20.00

Network packets in (count)

Count

10.00

0.00

0

10.00

10.00

10.00

20.00

Network packets out (count)

Count

1.00

0.00

0

10.00

10.00

10.00

20.00

Metadata no token (count)

No unit

1

0.00

0

10.00

10.00

10.00

20.00

CPU credit usage (count)

Count

0.00

0.00

0

10.00

10.00

10.00

20.00

CPU credit balance (count)

Count

20.00

10.00

0

10.00

10.00

10.00

20.00