

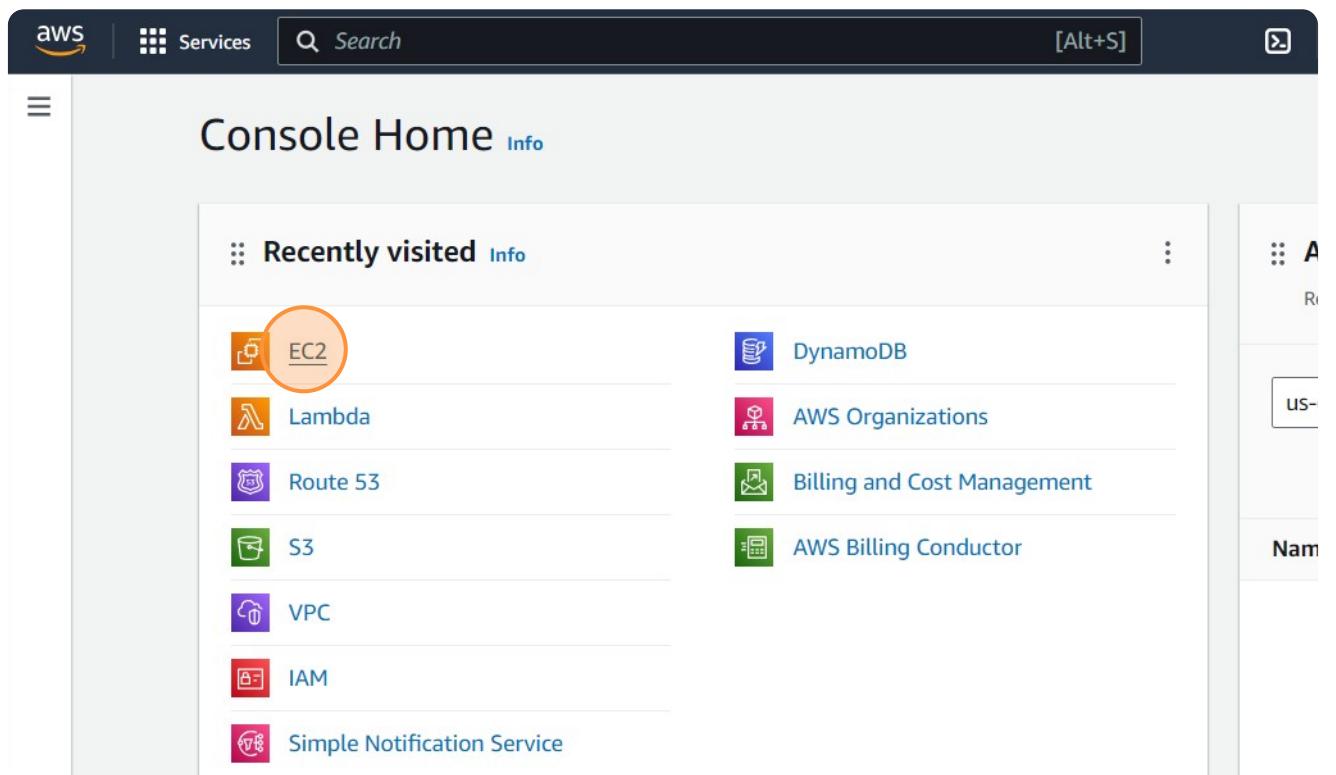
# EC2 with ELB and ASG

Learning how to create a scalable and highly available web application environment using Amazon EC2 instances, ELB, and ASG.

## Steps to launch EC2 instance

- 1 Navigate to AWS Management Console

- 2 Navigate to EC2 services



### 3 Click "Launch instance"

The screenshot shows the AWS EC2 console. On the left, there's a sidebar with links like Instances, Instance Types, Launch Templates, Spot Requests, Savings Plans, Reserved Instances, Dedicated Hosts, Capacity Reservations, and Images. Under Images, there are AMIs and AMI Catalog. Under Elastic Block Store, there are Buttons and Volumes. At the bottom of the sidebar are CloudShell and Feedback buttons. The main area has four boxes at the top: Load balancers (0), Placement groups (0), Snapshots (1), and Volumes (1). Below these is a large box titled "Launch instance" with the sub-instruction: "To get started, launch an Amazon EC2 instance, which is a virtual server in the cloud." It contains two buttons: "Launch instance" (highlighted with a red circle) and "Migrate a server". A note below says: "Note: Your instances will launch in the US East (N. Virginia) Region". To the right, there's a "Service health" section with "AWS Health Data" and a "Regions" section showing "Region: US East (N. Virginia)". Below that is a "Zones" section with "Zone name: us-east-1a".

### 4 Provide a name for EC2 instance.

#### 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

e.g. My Web Server

Add additional tags

##### ▼ 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

- 5 Select an AMI. Here, I selected "Amazon Linux 2 AMI (HVM) - Kernel 5.10, SSD Volume Type".

Amazon Machine Image (AMI)

Amazon Linux 2023 AMI  
Free tier eligible  
ami-0e731c8a588258d0d (64-bit (x86), uefi-preferred) / ami-0bbebc09f0a12d4d9 (64-bit (Arm), uefi)  
Virtualization: hvm ENA enabled: true Root device type: ebs

Amazon Linux 2023 AMI  
Free tier eligible  
ami-0e731c8a588258d0d (64-bit (x86), uefi-preferred) / ami-0bbebc09f0a12d4d9 (64-bit (Arm), uefi)  
Virtualization: hvm ENA enabled: true Root device type: ebs

**Amazon Linux 2 AMI (HVM) - Kernel 5.10, SSD Volume Type** ✓  
Free tier eligible  
ami-0cf10cdf9fcfd62d37 (64-bit (x86)) / ami-023c9d94d7c3b72 (64-bit (Arm))  
Virtualization: hvm ENA enabled: true Root device type: ebs

Deep Learning OSS Nvidia Driver AMI GPU PyTorch 2.0.1 (Amazon Linux 2) 20240206  
ami-079ce88a030285488 (64-bit (x86))  
Virtualization: hvm ENA enabled: true Root device type: ebs

Deep Learning OSS Nvidia Driver AMI GPU TensorFlow 2.15 (Amazon Linux 2) 20240213  
ami-0f1dfb67b5f95748 (64-bit (x86))  
Virtualization: hvm ENA enabled: true Root device type: ebs

Deep Learning OSS Nvidia Driver AMI GPU PyTorch 1.13.1 (Amazon Linux 2) 20240123  
ami-0dc430554d1998447 (64-bit (x86))  
Virtualization: hvm ENA enabled: true Root device type: ebs

Amazon Linux 2 LTS with SQL Server 2019 Standard  
ami-04142897ceadbbb7d (64-bit (x86))  
Virtualization: hvm ENA enabled: true Root device type: ebs

**Summary**

Number of instances: 1

Software Image (AMI): Amazon Linux 2023 AMI

Virtual server type (inst): t2.micro

Firewall (security group): New security group

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

**Free tier:** In you includes 750 hours of t3.micro in the Free Tier. t2.micro is unavailable.

- 6 Select an existing key pair or create a new one. I selected an existing key pair named "advance".

Key pair name - required

Select ✓ Create new key pair

advance (highlighted)

No preference (Default subnet in any availability zone)

Auto-assign public IP | Info

Enable

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-3' with the following rules:

7

Configure security group. We can add rules to allow specific traffic to reach our instance according to the requirements of the application. Here, we need to run a simple web application. So, I selected an existing security group that provides access to HTTP protocol.

Network [Info](#)  
vpc-0289f69e3626447a7

Subnet [Info](#)  
No preference (Default subnet in any availability zone)

Auto-assign public IP [Info](#)  
Enable

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

Common security groups [Info](#)

Select security groups	SG ID
launch-wizard-1 VPC: vpc-0289f69e3626447a7	sg-09ef620202b81d9ca
launch-wizard-4 VPC: vpc-0289f69e3626447a7	sg-02fc4a99ec9fd3a96
default	sg-07fb5906e636efbe1
launch-wizard-2	sg-07e1a5c1552b90b78

Compare security group rules

Number of instances [Info](#)  
1

Software Image (AMI)  
Amazon Linux 2 Kernel 5.10 AMI...[read more](#)  
ami-0cf10cdf9fd62d37

Virtual server type (instance type)  
t2.micro

Firewall (security group)

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

**Free tier:** In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance

Cancel [Launch instance](#) Review commands

8

Under Advanced details, I will set user data to following:

Upload a file with your user data or enter it in the field.

```
#!/bin/bash
yum update -y
yum install -y httpd
systemctl start httpd
systemctl enable httpd
echo "<h1> Hello World from $(hostname -f)</h1>" > /var/www/html/index.html
```

User data has already been base64 encoded

9

Launch more than 1 instance so that the load balancer can analyze incoming requests and divert them to the relevant servers.

The screenshot shows the AWS Lambda function configuration interface. On the left, there's a code editor containing a base64-encoded Lambda function. On the right, the configuration panel is displayed. The 'Number of instances' dropdown is set to 1, and this value is circled in orange. Below it, the 'Software Image (AMI)' is set to Amazon Linux 2 Kernel 5.10 AMI, and the 'Virtual server type (instance type)' is t2.micro. A tooltip for the free tier is visible, stating: 'Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance'. At the bottom right of the configuration panel, there are 'Cancel', 'Launch instance', and 'Review commands' buttons.

10

I created 3 instances. Then, click "Launch instance".

This screenshot shows the same AWS Lambda configuration interface as the previous one, but with a key difference: the 'Number of instances' dropdown is now set to 3, indicated by an orange circle. The rest of the configuration remains the same: Software Image (AMI) is Amazon Linux 2 Kernel 5.10 AMI, and Virtual server type is t2.micro. The 'Launch instance' button at the bottom right is also circled in orange. The rest of the interface, including the code editor and the free tier tooltip, appears identical to the previous screenshot.

## Steps to configure load balancer

### 11 Click "Load Balancers"

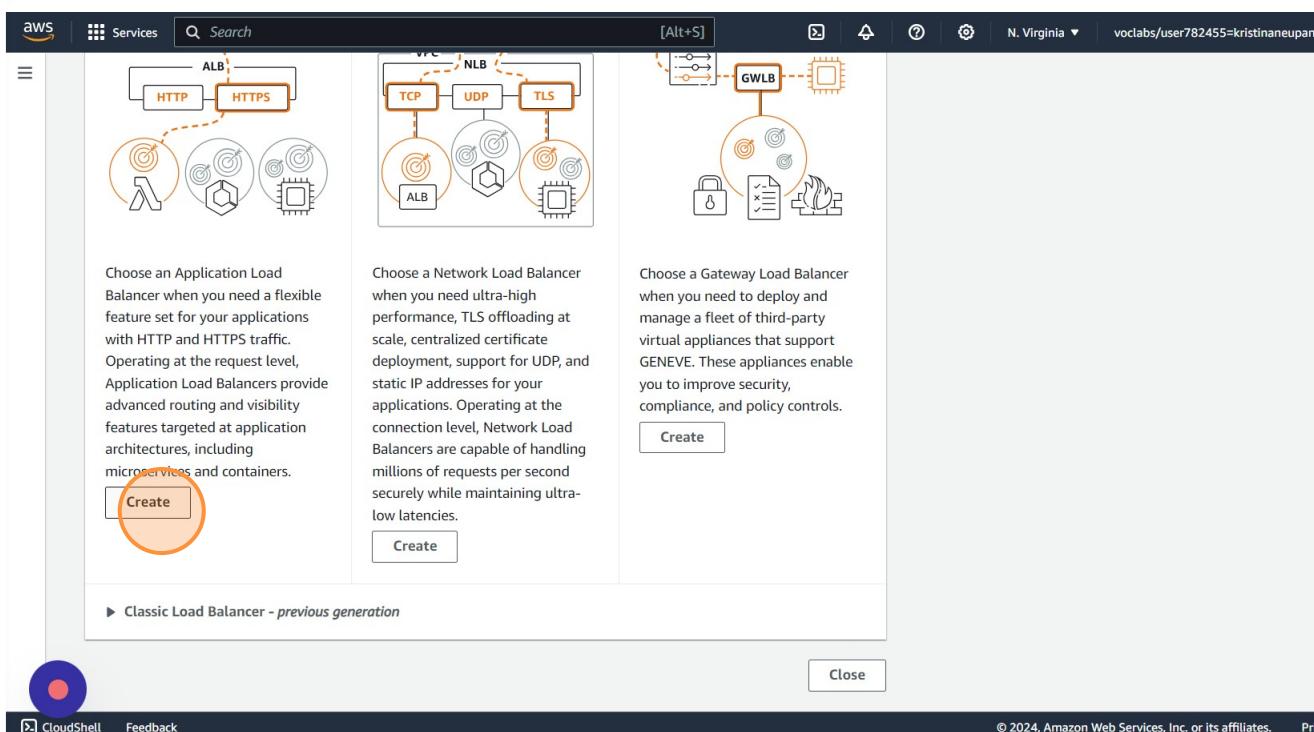
The screenshot shows the AWS CloudShell interface. On the left, there's a sidebar with various service links: AMI Catalog, Elastic Block Store (Volumes, Snapshots, Lifecycle Manager), Network & Security (Security Groups, Elastic IPs, Placement Groups, Key Pairs, Network Interfaces), Load Balancing (Load Balancers, Target Groups, Trust Stores), and Auto Scaling (Auto Scaling Groups). The 'Load Balancers' link under 'Load Balancing' is circled in red. At the bottom of the sidebar, there are 'CloudShell' and 'Feedback' buttons. The main area is a table titled 'Select an instance' showing EC2 instances with columns for Name, Instance ID, Instance state, and Instance type. One instance named 'test' is terminated.

### 12 Click "Create load balancer"

The screenshot shows the EC2 Load Balancers page. At the top, there's a search bar, a user icon, and a 'N. Virginia' dropdown. Below the header, there's a breadcrumb trail: EC2 > Load balancers. The main title is 'Load balancers'. A sub-header says 'Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic.' There's a 'Create load balancer' button at the top right, which is circled in red. Below the button is a 'Actions' dropdown and a refresh icon. A filter bar with a search icon and a placeholder 'Filter load balancers' is present. The main table has columns for Name, DNS name, State, VPC ID, Availability Zones, and Type. A message at the bottom states 'No load balancers' and 'You don't have any load balancers in us-east-1'. At the very bottom, it says '0 load balancers selected' and 'Select a load balancer above.'

13

Click "Create" on Application load balancer.



14

Provide name for load balancer.

## ▶ 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.

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



Search type

Select the type of IP addresses that your subnets use.

15

Select multiple availability zones so that the load balancer will route traffic to targets in these availability zones only. I selected 3 of them.

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-0289f69e3626447a7  
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.

us-east-1a (use1-az2)  
Subnet  
subnet-012f57e25edc2d99c

us-east-1b (use1-az4)  
Subnet  
subnet-01104b0fef4a971d9

us-east-1c (use1-az6)  
Subnet  
subnet-01104b0fef4a971d9

us-east-1d (use1-az1)

CloudShell Feedback © 2024, Amazon Web Services, Inc. or its affiliates. [Pri](#)

16

Select security group. I selected the one that was associated with EC2 instance.

## 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



launch-wizard-1

sg-09ef620202b81d9ca VPC: vpc-0289f69e3626447a7

launch-wizard-4

sg-02fc4a99ec9fd3a96 VPC: vpc-0289f69e3626447a7

launch-wizard-2

sg-07e1a5c1552b90b78 VPC: vpc-0289f69e3626447a7

default

sg-07fb5906e636efbe1 VPC: vpc-0289f69e3626447a7

▼ Listener HTTP:80

17

Select existing target group or create a new. I will be creating a new target group. To create new target group click on "Create target group".

**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 traffic to its registered targets.

▼ Listener **HTTP:80**

Protocol	Port	Default action <a href="#">Info</a>
HTTP	80 1-65535	Forward to <a href="#">Select a target group</a> <a href="#">Create target group</a>

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

[Add listener tag](#)  
You can add up to 50 more tags.

[Add listener](#)

► **Load balancer tags - optional**

18

Provide a name for target group.

• Facilitates routing to a single Lambda function.  
• Accessible to Application Load Balancers only.

Application Load Balancer  
• Offers the flexibility for a Network Load Balancer to accept and route TCP requests.  
• Facilitates using static IP addresses and PrivateLink with an Application Load Balancer.

**Target group name**  
A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must start with a letter.

**Protocol : Port**  
Choose a protocol for your target group that corresponds to the Load Balancer type that will be used. You can also choose a port for your target group and enable anomaly detection for the targets and you can set mitigation options once your target group is created.

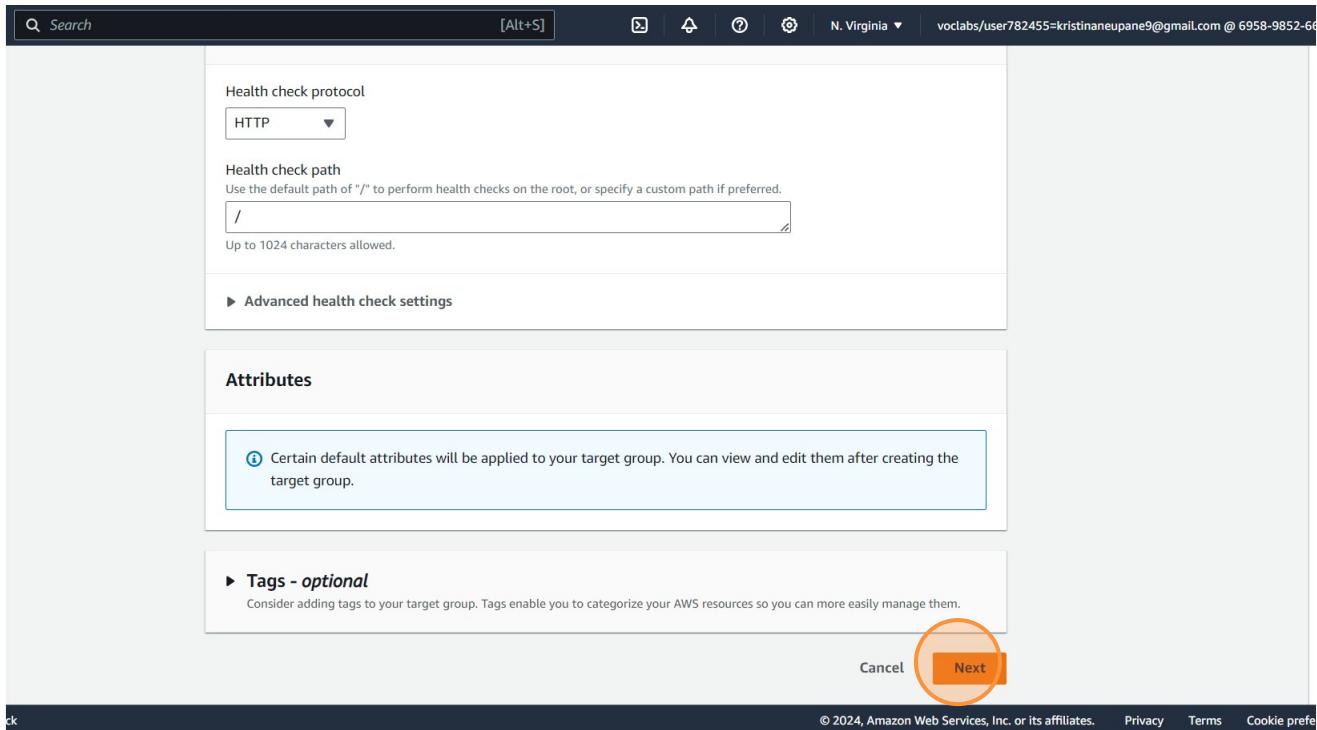
HTTP	80 1-65535
------	---------------

### IP address type

Only targets with the indicated IP address type can be registered to this target group.

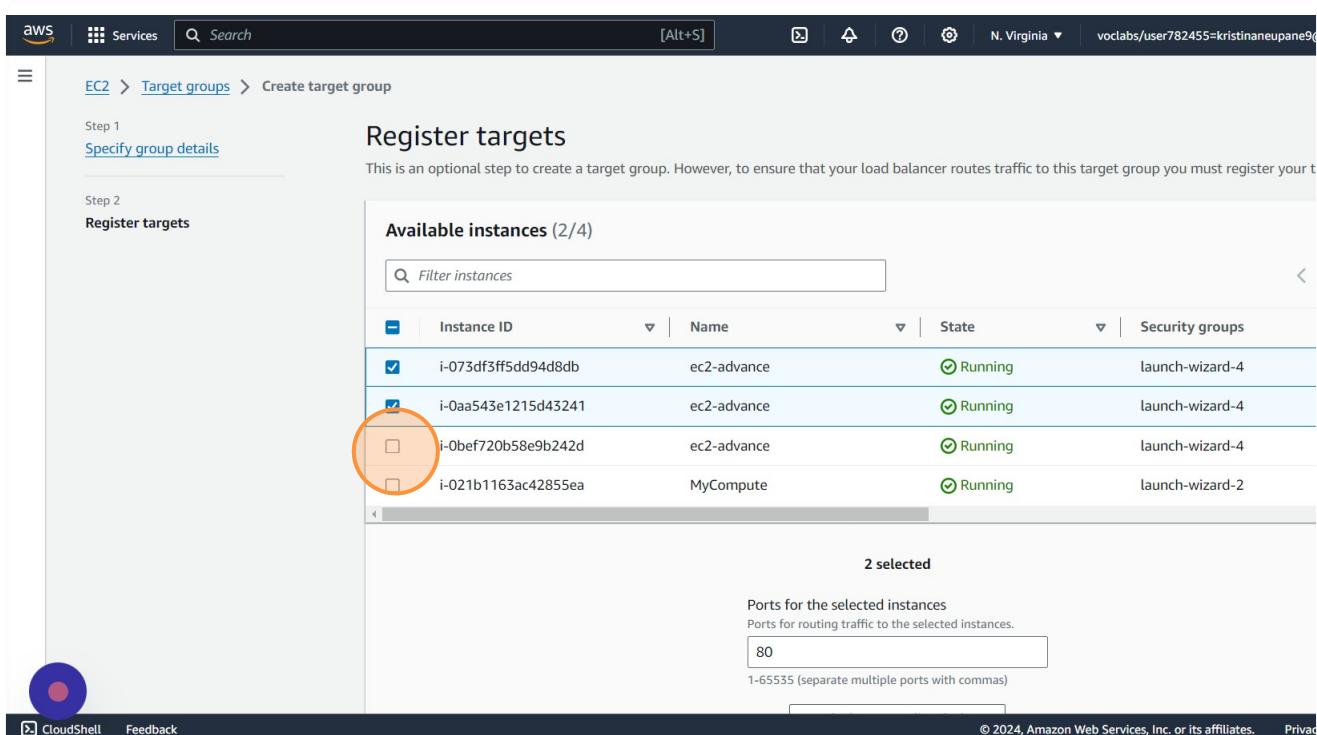
19

Click "Next"



20

Select the instances that need to be included in this target group.



**21** Click "Include as pending below"

The screenshot shows the AWS Lambda Targets configuration page. At the top, there is a table listing three instances:

Instance ID	Name	State	Security groups
i-0bef720b58e9b242d	ec2-advance	Running	launch-wizard-4
i-021b1163ac42855ea	MyCompute	Running	launch-wizard-2

Below the table, it says "3 selected". Under "Ports for the selected instances", there is a text input field containing "80". A button labeled "Include as pending below" is highlighted with an orange circle.

**Review targets**

**Targets (0)**

Filter targets Show only pending < 1 Remove all

Instance ID	Name	Port	State	Security groups	Zone	Private IPv4 address	Subnet ID
No instances added yet							

**22** Click "Create target group"

The screenshot shows the AWS Lambda Targets configuration page. At the top, it says "3 selections are now pending below. Include more or register targets when ready." Below this, it says "Review targets".

**Targets (3)**

Filter targets Show only pending < 1 > Remove all pending

Instance ID	Name	Port	State	Security groups	Zone	Private IPv4 address	Subnet ID
i-073df3ff5dd94d8db	ec2-advance	80	Running	launch-wizard-4	us-east-1c	172.31.34.135	subnet-04
i-0aa543e1215d43241	ec2-advance	80	Running	launch-wizard-4	us-east-1c	172.31.35.137	subnet-04
i-0bef720b58e9b242d	ec2-advance	80	Running	launch-wizard-4	us-east-1c	172.31.43.203	subnet-04

3 pending

Create target group

Cancel Previous

© 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

23

Now, go back to the page where target group need to be selected. After clicking on refresh button we can see the newly created target group.

The screenshot shows the AWS CloudFront Listener configuration page. A newly created target group, "my-target-group", is highlighted with an orange circle. The "Default action" dropdown is set to "Forward to Select a target group". The "Protocol" is set to "HTTP" and the "Port" is "80". The "Listener tags - optional" section shows an "Add listener tag" button. The "Load balancer tags - optional" section is collapsed. The bottom right corner of the page includes links for CloudShell, Feedback, and cookie preferences.

24

Select target group.

The screenshot shows the "Default action" configuration for a CloudFront listener. The "Forward to" dropdown is open, showing the "Select a target group" option, which is also highlighted with an orange circle. Below the dropdown, a search bar contains the text "my-target-group". The "Create target" button is visible to the left of the search bar. The "HTTP" protocol is selected at the bottom right. A note at the bottom states: "Tags enable you to categorize your AWS resources so you can more easily manage them."

25

View the summary and click "Create load balancer".

The screenshot shows the AWS Load Balancers console. A new load balancer named 'my-load-balancer' is being created. The configuration includes:

- Internet-facing**: IPv4
- VPC**: [vpc-0289f69e3626447a7](#)
- HTTP:80** default target group: [my-target-group](#)
- Service integrations**: AWS WAF: None, AWS Global Accelerator: None
- Tags**: None
- Attributes**: A note states: "Certain default attributes will be applied to your load balancer. You can view and edit them after creating the load balancer."
- Creation workflow and status**: A step titled "Server-side tasks and status" is shown, with a note: "After completing and submitting the above steps, all server-side tasks and their statuses become available for monitoring." The "Create load balancer" button is highlighted with a red circle.

26

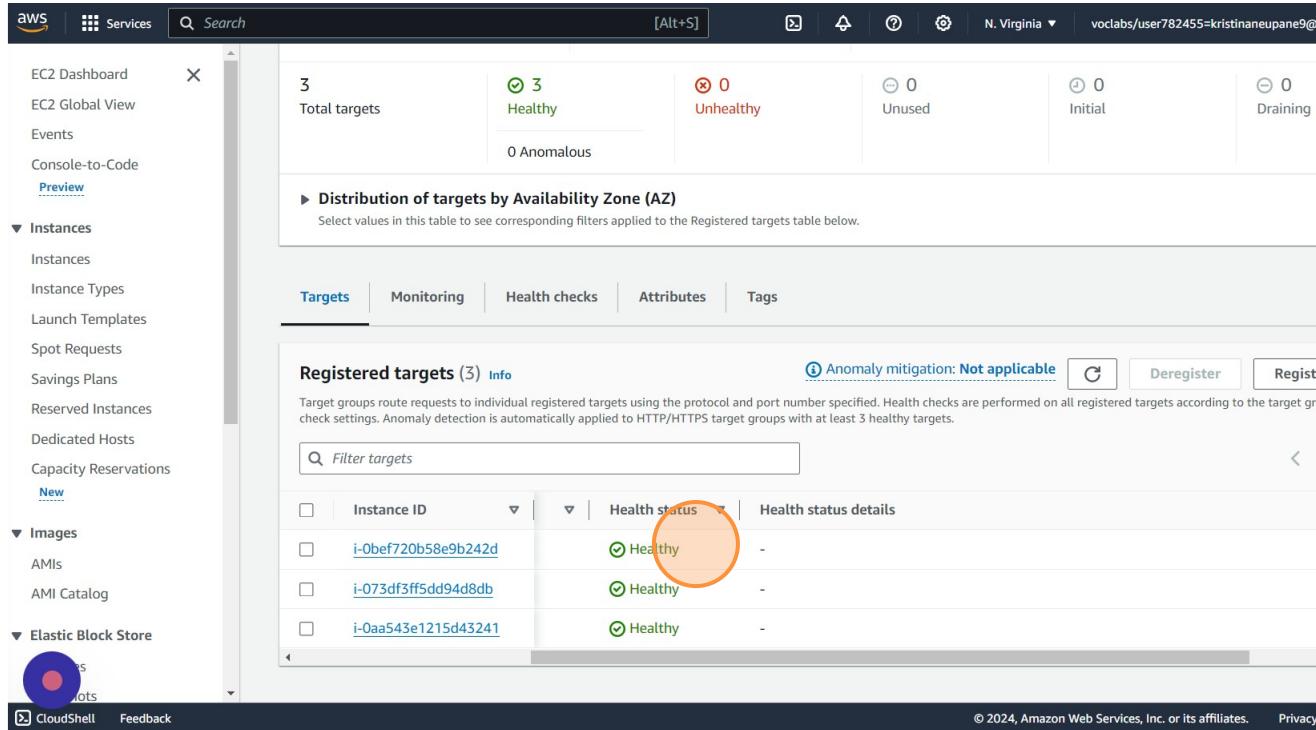
Finally, load balancer is successfully created.

The screenshot shows the AWS EC2 Dashboard under the 'Load balancers' section. A message at the top says: "Successfully created load balancer: my-load-balancer. It might take a few minutes for your load balancer to fully set up and route traffic. Targets will also take a few minutes to complete the registration process and pass initial health checks." Below this, the 'my-load-balancer' details are listed:

Details	Value
Load balancer type	Application
Status	Provisioning (highlighted with a red circle)
Scheme	Internet-facing
VPC	<a href="#">vpc-0289f69e3626447a7</a>
IP address type	IPv4
Availability Zones	subnet-048e167e8141311cb us-east-1c (use1-az6), subnet-012f57e25edc2d99c us-east-1a (use1-az2), subnet-01104b0fe4a971d9 us-east-1b (use1-az4)
Date created	February 17, 2024, 08:44 (UTC+05:45)
Load balancer ARN	<a href="#">arn:aws:elasticloadbalancing:us-east-1:695898526652:loadbalancer/app/my-load-balancer/262f32259da63e5e</a>
DNS name	<a href="#">my-load-balancer-203143361.us-east-1.elb.amazonaws.com (A Record)</a>

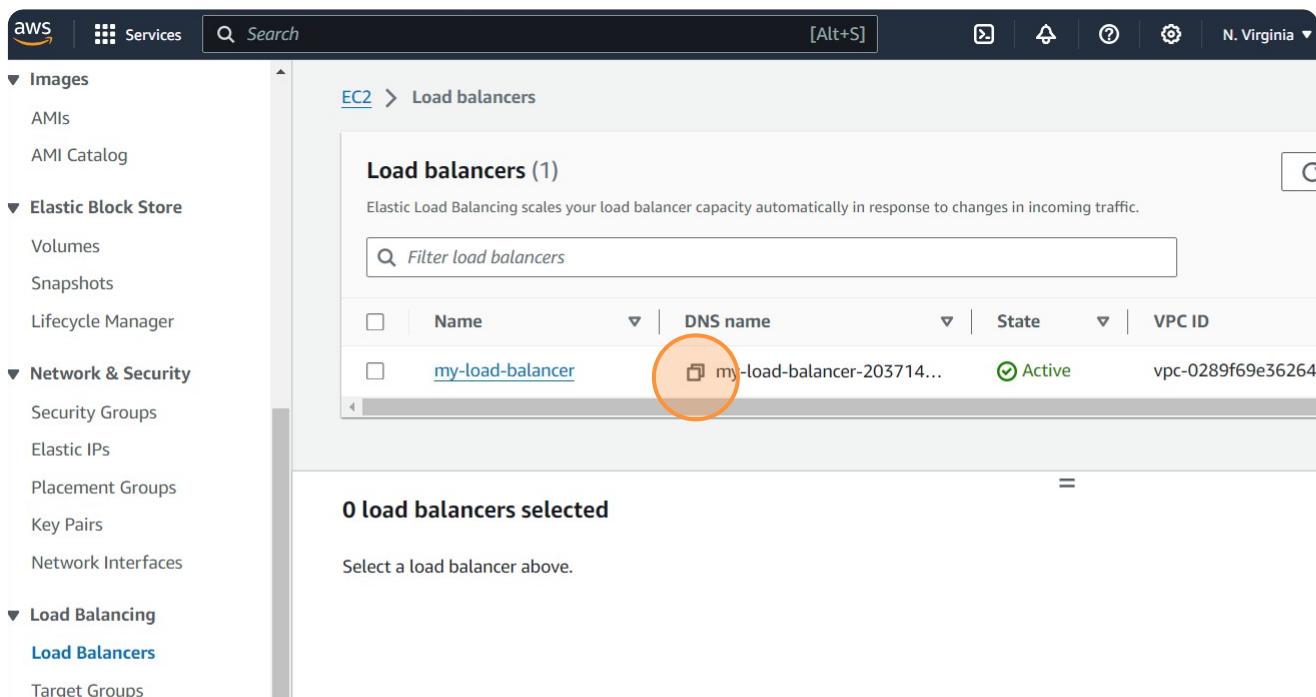
## Steps to test load balancer

- 27 We can see all 3 instances are healthy. Now, we should be able to access the web application.



The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with links like EC2 Dashboard, EC2 Global View, Events, Console-to-Code, Instances, Images, and Elastic Block Store. The main area displays a summary of targets: 3 Total targets, 3 Healthy, 0 Unhealthy, 0 Unused, 0 Initial, and 0 Draining. Below this is a table titled "Registered targets (3) Info". The table has columns for Instance ID, Health status, and Health status details. Three rows are listed, each with an orange circle around the "Healthy" status indicator. The rows are: i-0bef720b58e9b242d, i-073df3ff5dd94d8db, and i-oaa543e1215d43241. At the bottom right of the table, there are buttons for Anomaly mitigation, Deregister, and Register.

- 28 Copy DNS name of load balancer and paste it in new tab and click "Enter".



The screenshot shows the AWS EC2 Load Balancers page. The sidebar includes links for Images, Elastic Block Store, Network & Security, and Load Balancing. Under Load Balancing, "Load Balancers" is selected. The main content area shows a table titled "Load balancers (1)". The table has columns for Name, DNS name, State, and VPC ID. One row is present: my-load-balancer, my-load-balancer-203714..., Active, vpc-0289f69e36264. An orange circle highlights the "DNS name" column for the first row. Below the table, it says "0 load balancers selected" and "Select a load balancer above."

29

We can see a simple web application that specifies the IP address of the instance to which load balancer forwarded request.

**Hello World from ip-172-31-34-135.ec2.internal**



30

After refreshing the page, we can see another IP address.

**Hello World from ip-172-31-43-203.ec2.internal**





In this way, an AWS load balancer analyzes incoming requests and diverts them to the relevant servers.

## Set Up Auto Scaling Group (ASG)

31

Click "Auto Scaling Groups"

The screenshot shows the AWS Lambda console interface. On the left, there's a sidebar with 'Lifecycle Manager' at the top, followed by sections for 'Network & Security' (Security Groups, Elastic IPs, Placement Groups, Key Pairs, Network Interfaces), 'Load Balancing' (Load Balancers, Target Groups, Trust Stores), and 'Auto Scaling' (Auto Scaling Groups). The 'Auto Scaling Groups' link is circled in orange. At the bottom of the sidebar are 'CloudShell' and 'Feedback' buttons. The main area displays two EC2 instances: 'ec2-advance-2' and 'ec2-advance-3', both marked as 'Running'. Below this is a 'Select an instance' section.

32

Click "Create Auto Scaling group"

The screenshot shows the AWS Auto Scaling Groups page. At the top, there's a green success message: 'Created successfully'. Below it, there's a table with one row, 'Scaling groups'. At the bottom of the page, there are tabs for 'Scaling groups (1) Info', 'Launch configurations', 'Launch templates', 'Actions', and a prominent 'Create Auto Scaling group' button, which is also circled in orange. The URL in the address bar is 'vocabs/user782455=kristinaneupane9@gmail.com @ 6958-9852-6'.

The screenshot shows the AWS Auto Scaling console. At the top, there's a navigation bar with tabs for 'Launch template/configuration', 'Instances', 'Status', 'Desired capacity', and 'Min'. Below the navigation bar, a table row displays the launch template 'my-template | Version Default' with a value of '4', a status of 'Updating capacity...', a 'Desired capacity' of '1', and a 'Min' value of '1'. A large, semi-transparent callout box labeled 'ups selected' covers the middle portion of the screen. In the bottom right corner of this callout box, there are two small icons: a gear and a close button.

33 Provide name for ASG.

This screenshot shows the 'Choose launch template' step in the AWS Auto Scaling group creation wizard. On the left, there's a sidebar with several optional configuration sections: 'Configure launch template', 'Configure instance launch options', 'Configure advanced options', 'Configure group size and scaling', 'Configure notifications', and 'Add tags'. The main area is titled 'Choose launch template' with an 'Info' link. It instructs the user to 'Specify a launch template that contains settings common to all EC2 instances that are lau...'. Below this is a 'Name' input field where the user can enter the Auto Scaling group name. A note says 'Enter a name to identify the group.' A blue input field is shown with a red circle highlighting its center. Another note below it states 'Must be unique to this account in the current Region and no more than 255 characters.' At the bottom, there's a 'Launch template' section with an 'Info' link and a note: 'For accounts created after May 31, 2023, the EC2 console only supports creating launch templates. Creating Auto Scaling groups with launch configurations is no longer available via the CLI and API until December 31, 2023.'

34 Select an existing launch template or create new one. For creating new, click "Create a launch template".

This screenshot shows the 'Configure group size and scaling' step in the AWS Auto Scaling group creation wizard. On the left, there are three optional configuration sections: 'Step 5 - optional' (which includes 'Add notifications'), 'Step 6 - optional' (which includes 'Add tags'), and 'Configure group size and scaling'. The 'Configure group size and scaling' section has a note: 'Must be unique to this account in the current Region and no more than 255 characters.' On the right, there's a 'Launch template' section with an 'Info' link and a note: 'For accounts created after May 31, 2023, the EC2 console only supports creating launch templates. Creating Auto Scaling groups with launch configurations is no longer available via the CLI and API until December 31, 2023.'

Step 7  
Review

Launch templates. Creating Auto Scaling groups with launch configurations available via the CLI and API until December 31, 2023.

Launch template  
Choose a launch template that contains the instance-level settings, such as the Amazon Machine Image, security groups, and more.

Select a launch template

Create a launch template 

Feedback

### 35 Provide name for template.

#### Launch template name and description

Launch template name - *required*

my-template2

Must be unique to this account. Max 128 chars. No spaces or special characters like '&', '\*', '@'.

Template version description

A prod webserver for MyApp

Max 255 chars

Auto Scaling guidance | [Info](#)

Select this if you intend to use this template with EC2 Auto Scaling

Provide guidance to help me set up a template that I can use with EC2 Auto Scaling

► [Template tags](#)

► [Source template](#)

### 36 Select an AMI.

#### Launch template contents

Specify the details of your launch template below. Leaving a field blank will result in the field not being included in the launch template.

##### ▼ Application and OS Images (Amazon Machine Image) - *required* [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](#)

Software Image (AMI)

Virtual server type (inst)

Firewall (security group)

Storage (volumes)

Recents My AMIs Quick Start

Amazon Linux macOS Ubuntu Windows Red Hat SUSE Li [Browse more AMIs](#)

Including AMIs from AWS, Marketplace and the Community

Cancel

CloudShell Feedback © 2024, Amazon Web Services

37 Select instance type.

▼ Instance type [Info](#) | [Get advice](#) [Advanced](#)

Instance type

Don't include in launch template

Get advice on instance type selection...

Don't include in launch template

t2.nano

Family: t2 1 vCPU 0.5 GiB Memory Current generation: true  
On-Demand Linux base pricing: 0.0058 USD per Hour  
On-Demand SUSE base pricing: 0.0058 USD per Hour  
On-Demand Windows base pricing: 0.0081 USD per Hour

t2.micro

Family: t2 1 vCPU 1 GiB Memory Current generation: true  
On-Demand Windows base pricing: 0.0162 USD per Hour  
On-Demand SUSE base pricing: 0.0116 USD per Hour  
On-Demand RHEL base pricing: 0.0716 USD per Hour  
On-Demand Linux base pricing: 0.0116 USD per Hour

All generations

Compare instance types

CloudShell Feedback

38 Select existing key pair or create a new one.

On-Demand RHEL base pricing: 0.016 USD per Hour  
On-Demand Linux base pricing: 0.0116 USD per Hour

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.

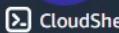
Key pair name

Don't include in launch template

Create new key pair

▼ Network settings [Info](#)

Subnet [Info](#)



CloudShell

Feedback

39

Select existing security group or create a new one.

Select existing security group

Create security group

Security groups [Info](#)

Select security groups



Specify a custom value...

launch-wizard-1  
VPC: vpc-0289f69e3626447a7

sg-09ef620202b81d9ca

launch-wizard-4  
VPC: vpc-0289f69e3626447a7

sg-02fc4a99ec9fd3a96

default  
VPC: vpc-0289f69e3626447a7

sg-07fb5906e636efbe1

launch-wizard-2  
VPC: vpc-0289f69e3626447a7

sg-07e1a5c1552b90b78

default  
VPC: vpc-019a041aa289e56d9

sg-0cf9fee934bb80f57

Compare security group rules

Hide details



CloudShell

Feedback

40

Click "Create launch template"

Search [Alt+S] ✖

Image (volumes) [Info](#)

Volumes Hide details

1 (AMI Root) (8 GiB, EBS, General purpose SSD (gp3))  
Volumes are not included in the template unless modified

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage ✖

New volume

Summary

Software Image (AMI)  
Amazon Linux 2023 AMI 2023.3.2... [read more](#)

ami-0e731ca588258bd0d

Virtual server type (instance type)  
t2.micro

Firewall (security group)  
launch-wizard-4

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

Free tier: In your first year ✖

**Source tags** [Info](#)

No tags are currently included in this template. Add a resource tag to include it in the launch template.

**New tag** Add up to 50 more tags.

**Selected details** [Info](#)

Includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 30 GiB of EBS storage, 2 million IOs, 1 GB of snapshots, and 100 GB of bandwidth to the internet.

[Cancel](#) [Create launch template](#) 

© 2024, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

**41** Now, go back to page where launch template need to be selected. Click on refresh button to see newly created template.

**Name**

Auto Scaling group name  
Enter a name to identify the group.  
 Must be unique to this account in the current Region and no more than 255 characters.

**Launch template** [Info](#)

For accounts created after May 31, 2023, the EC2 console only supports creating Auto Scaling groups with launch templates. Creating Auto Scaling groups with launch configurations is not recommended but still available via the CLI and API until December 31, 2023.

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.  
   
[Create a launch template](#) 

[Cancel](#) [Next](#)

© 2024, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

**42** Select a launch template.

**Name**

Auto Scaling group name  
Enter a name to identify the group.  
 Must be unique to this account in the current Region and no more than 255 characters.

**Launch template** [Info](#)

For accounts created after May 31, 2023, the EC2 console only supports creating Auto Scaling groups with launch templates. Creating Auto Scaling groups with launch configurations is not recommended but still available via the CLI and API until December 31, 2023.



Step 7  
Review

Search launch templates

my-template2  
my-template

Select a launch template

Create a launch template [Create a launch template](#)

Cancel [Next](#)

© 2024, Amazon Web Services, Inc. or its affiliates.

Feedback

#### 43 Click "Next"

Search [Alt+S]

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.

my-template2

Create a launch template [Create a launch template](#)

Version

Default (1) [Create a launch template version](#)

Description	Launch template	Instance type
my-template2	my-template2 <a href="#">Edit</a> lt-0e2fd94f8ff06c6cb	t2.micro

AMI ID	Security groups	Request Spot Instances
ami-0e731c8a588258d0d	-	No

Key pair name	Security group IDs
advance	sg-02fc4a99ec9fd3a96 <a href="#">Edit</a>

Additional details

Storage (volumes)	Date created
-	Sat Feb 17 2024 16:17:43 GMT+0545 (Nepal Time)

Cancel [Next](#)

© 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

#### 44 You can select multiple availability zones and let ASG balance the instances across the zones.

Step 5 - optional

Add notifications

Step 6 - optional

Add tags

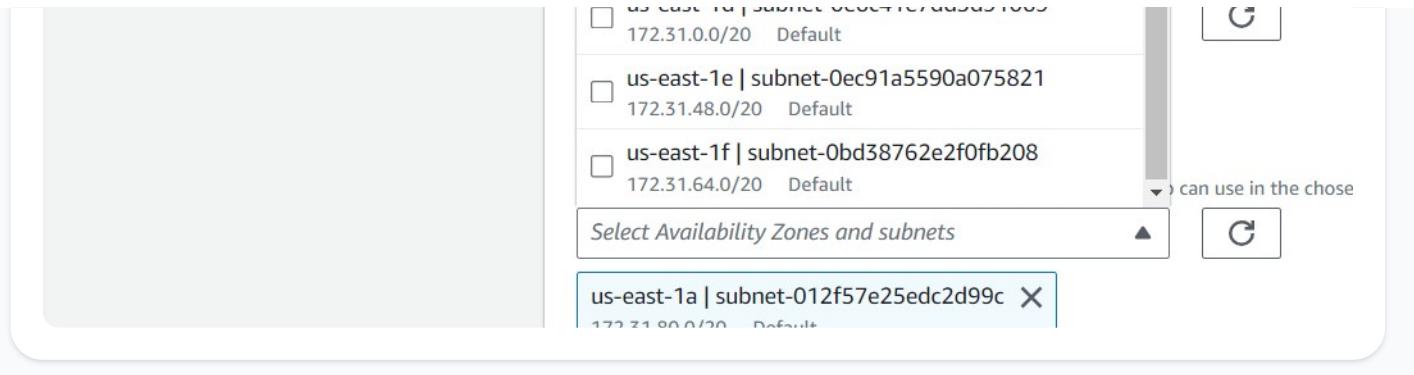
Step 7

Review

Instance type  
t2.micro

**Network Info**

Availability Zone	Subnet	IP Range	Default
us-east-1a	subnet-012f57e25edc2d99c	172.31.80.0/20	Default
us-east-1b	subnet-01104b0fef4a971d9	172.31.16.0/20	Default
us-east-1c	subnet-048e167e8141311cb	172.31.32.0/20	Default
us-east-1d	subnet-0ef6c41e7dd3d91069		



45 Click "Next"

Search [Alt+S] N. Virginia v vocabs/user782455=kristinaneupane9@gmail.com @ 6958-9852

**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-0289f69e3626447a7 ▾ C  
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 ▾ C  
us-east-1a | subnet-012f57e25edc2d99c X  
172.31.80.0/20 Default  
us-east-1b | subnet-01104b0fef4a971d9 X  
172.31.16.0/20 Default  
us-east-1c | subnet-048e167e8141311cb X  
172.31.32.0/20 Default  
Create a subnet

Cancel Skip to review Previous Next

© 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie pr

46 There are various options related to load balancing. I attached ASG to an existing load balancer that we created before.

Launch template

Instance launch options

Advanced options

Group size and

Applications

**Configure advanced options - optional** Info

Integrate your Auto Scaling group with other services to distribute network traffic across multiple servers using a load balancer or to establish service-to-service communications using VPC Lattice. You can also set options that give you more control over health check replacements and monitoring.

**Load balancing** Info

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.

ional

### VPC Lattice integration options Info

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.

Select VPC Lattice service to attach

No VPC Lattice service

VPC Lattice will not manage your Auto Scaling group's

Attach to VPC Lattice service

Incoming requests associated with specified VPC Lattice

Feedback

© 2024, Amazon Web Services, Inc. c

47

I selected an existing target group "my-target-group".

### Attach to an existing load balancer

Select the load balancers that you want to attach to your Auto Scaling group.

Choose from your load balancer target groups

This option allows you to attach Application, Network, or Gateway Load Balancers.

Choose from Classic Load Balancers

#### Existing load balancer target groups

Only instance target groups that belong to the same VPC as your Auto Scaling group are available for selection.

Select target groups



my-target-group | HTTP

Application Load Balancer: my-load-balancer



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.

Select VPC Lattice service to attach

© 2024, Amazon

48

Click "Next"

Search [Alt+S] N. Virginia ▾ [vocabs/user782455=krishnanepane9@gmail.com @ 6958-9852](#)

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.

300 seconds

**Additional settings**

Monitoring Info

Enable group metrics collection within CloudWatch

**Default instance warmup** [Info](#)  
The amount of time that CloudWatch metrics for new instances do not contribute to the group's aggregated instance metrics, as their usage data is not reliable yet.

Enable default instance warmup

Cancel [Skip to review](#) Previous **Next**

© 2024, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

## 49 Set group size by specifying desired capacity, min and max desired capacity.

Step 5 - optional  
[Add notifications](#)

Desired capacity  
Specify your group size.  
1

Step 6 - optional  
[Add tags](#)

Step 7  
[Review](#)

**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.

<b>Min desired capacity</b> 1	<b>Max desired capacity</b> 4
Equal or less than desired capacity	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.

**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 a scaling policy adjust the desired capacity in proj the metric's value.

CloudShell Feedback © 2024, Amazon Web Services, Inc. or its affiliates.

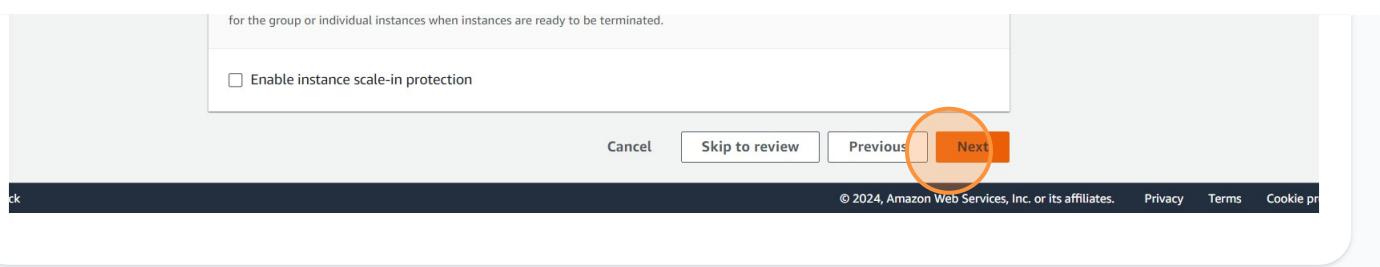
## 50 Click "Next"

An instance maintenance policy determines how much availability your application has when EC2 Auto Scaling replaces instances. It also establishes guardrails that limit the amount of capacity that can be added or removed when replacing instances.

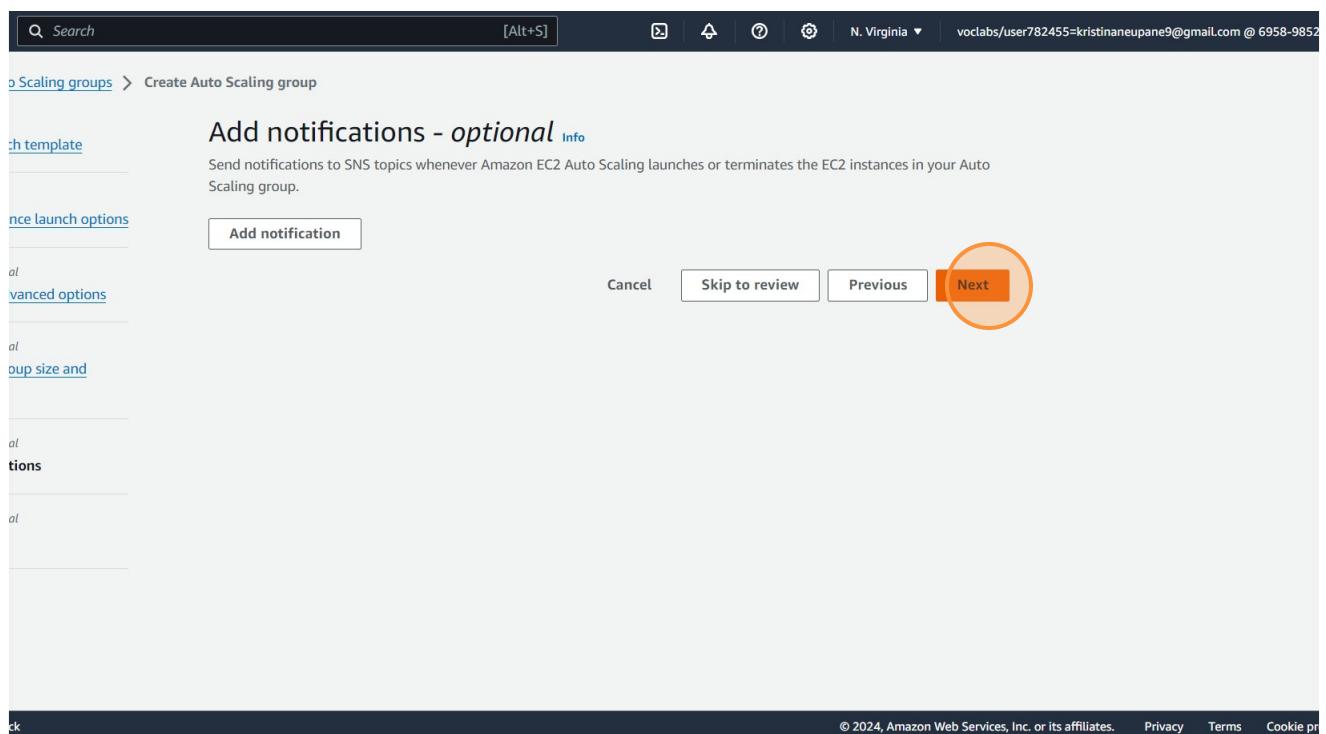
Choose a replacement behavior depending on your availability requirements

<b>Mixed behavior</b> <input checked="" type="radio"/> <b>No policy</b> For rebalancing events, new instances will launch before terminating others. For all other events, instances terminate and launch at the same time.	<b>Prioritize availability</b> <input type="radio"/> <b>Launch before terminating</b> 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.	<b>Control costs</b> <input type="radio"/> <b>Terminate and launch</b> 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.	<b>Flexible</b> <input type="radio"/> <b>Custom behavior</b> 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.
---	---	--	---

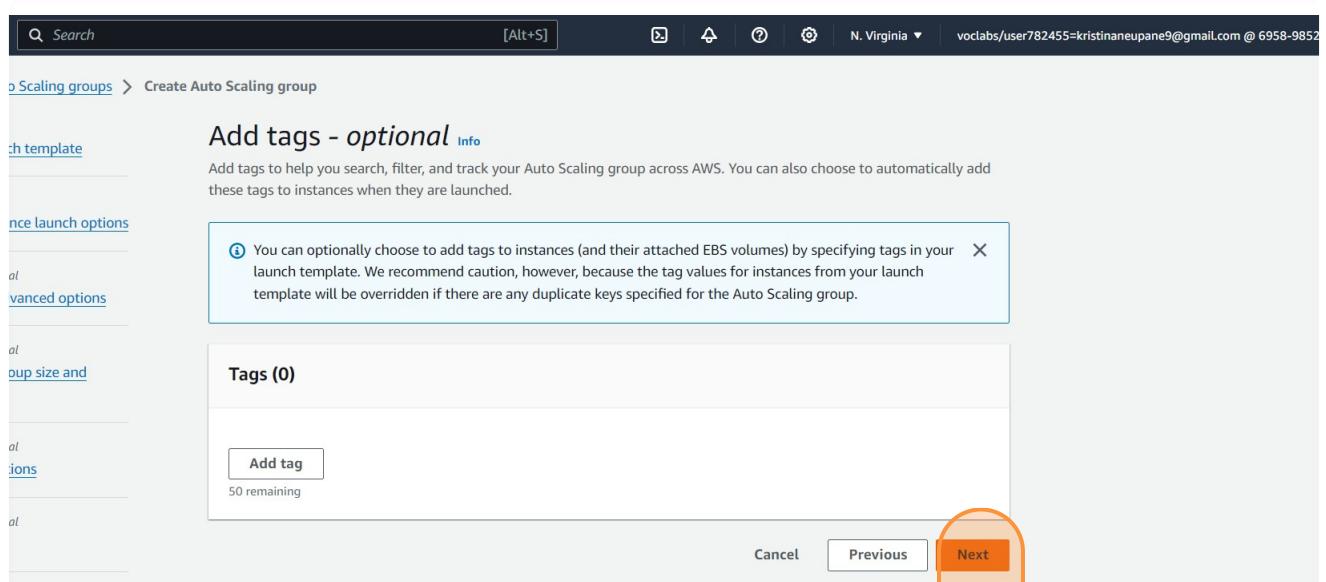
**Instance scale-in protection**  
Scale-in protection prevents newly launched instances from being terminated by scaling activities. Make sure to remove scale-in protection



## 51 Click "Next"



## 52 Click "Next"



**53**

Click "Create Auto Scaling group"

The screenshot shows the final step of creating an Auto Scaling group. It includes sections for Instance scale-in protection, Notifications, and Tags. At the bottom, there are 'Cancel', 'Previous', and 'Create Auto Scaling group' buttons. The 'Create Auto Scaling group' button is highlighted with a red circle.

Instance scale-in protection

Instance scale-in protection

Enable instance protection from scale in

Step 5: Add notifications

Notifications

No notifications

Step 6: Add tags

Tags (0)

Key	Value	Tag new instances
No tags		

Cancel Previous Create Auto Scaling group

**54**

We can see ASG named "my-asg" is created.

The screenshot shows the 'Auto Scaling groups' page with one group listed. The group is named 'my-asg' and is associated with 'my-template2 | Version Default'. The 'Name' column has a red circle around it.

Auto Scaling groups (1) info

Name	Launch template/configuration	Instances	Status	Desired capacity	Min	Max
my-asg	my-template2   Version Default	0	Updating capacity...	1	1	4

0 Auto Scaling groups selected

CloudShell Feedback © 2024, Amazon Web Services, Inc. or its affiliates. Pri

## Testing ASG

- 55 Copy DNS name of load balancer.

EC2 > Load balancers

### Load balancers (1)

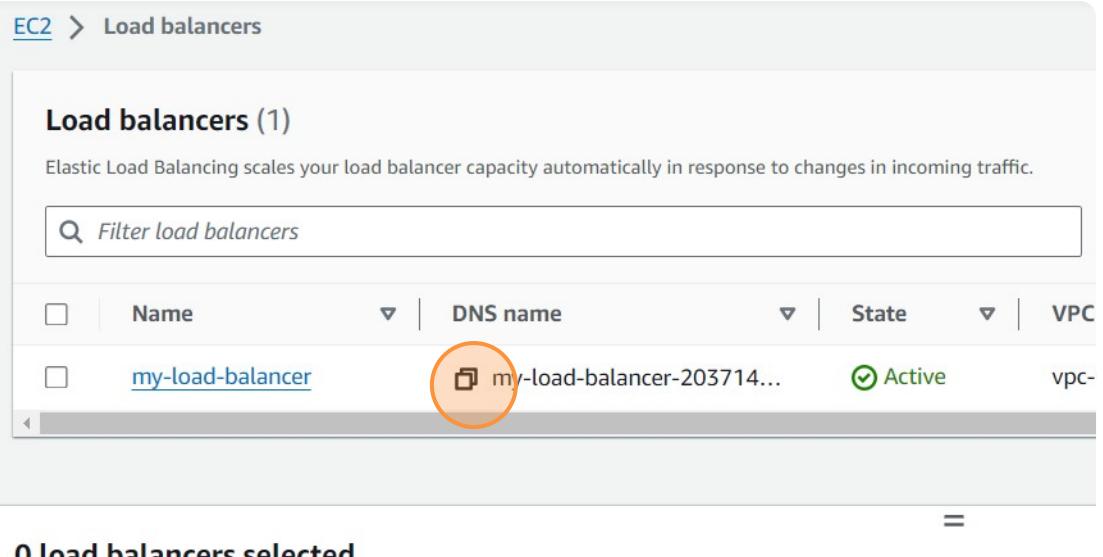
Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic.

Filter load balancers

Name	DNS name	State	VPC
my-load-balancer	my-load-balancer-203714...	Active	vpc-

0 load balancers selected

Select a load balancer above.



- 56 In a new tab, navigate to

<http://my-load-balancer-2037143361.us-east-1.elb.amazonaws.com/>

- 57 Here, we can see web application is running.

**Hello World from ip-172-31-34-135.ec2.internal**

- 58** Now lets try what happens if we stop all the instances that are associated with load balancer.

The screenshot shows the AWS CloudWatch Metrics Insights interface. At the top, there's a search bar and a toolbar with various icons. Below that is a table titled "Instances (4/5) Info" with columns for Name, Instance ID, Instance state, and Instance type. Five instances are listed, all currently running. To the right of the table is a context menu with options: Stop instance (highlighted with an orange circle), Start instance, Reboot instance, Hibernate instance, Terminate instance, Alarm status, and Availability Zone. Below the table, a message states: "Instances: i-05846983a782418e2 (ec2-advance), i-09a92f1458f7a6d6b (ec2-advance), i-01ed61f04006e4aae, i-0f7631b46dd630b4b (ec2-advance)". At the bottom, there's a "Monitoring" section with a chart for CPU utilization (%) and other network metrics like Network in (bytes), Network out (bytes), and Network packets in (cou...).

- 59** After refreshing on the page we can see "502 Bad Gateway".

**502 Bad Gateway**

60

Refresh instances.

Instances (4/5) [Info](#)

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone
MyCompute	i-021b1163ac42855ea	Running	t2.micro	2/2 checks passed	<a href="#">View alarms</a> +	us-east-1b
<input checked="" type="checkbox"/> ec2-advance	i-09a92f1458f7a6d6b	Stopped	t2.micro	2/2 checks passed	<a href="#">View alarms</a> +	us-east-1b
<input checked="" type="checkbox"/> ec2-advance	i-01ed61f04006e4aae	Stopped	t2.micro	2/2 checks passed	<a href="#">View alarms</a> +	us-east-1b
<input checked="" type="checkbox"/> ec2-advance	i-0f7631b46dd630b4b	Stopped	t2.micro	2/2 checks passed	<a href="#">View alarms</a> +	us-east-1a
<input checked="" type="checkbox"/> ec2-advance	i-05846983a782418e2	Stopped	t2.micro	2/2 checks passed	<a href="#">View alarms</a> +	us-east-1c

Instances: i-05846983a782418e2 (ec2-advance), i-09a92f1458f7a6d6b (ec2-advance), i-01ed61f04006e4aae, i-0f7631b46dd630b4b (ec2-advance)

Monitoring

CPU utilization (%) Network in (bytes) Network out (bytes) Network packets in (bytes)

Percent Bytes Bytes Count

15.9 7.41k 6.72k 64.4

61

We can see new EC2 instance is created automatically to handle the request.

Instances (4/6) [Info](#)

Name	Instance ID	Instance state	Instance type	Status check	Alarm status
MyCompute	i-021b1163ac42855ea	Running	t2.micro	2/2 checks passed	<a href="#">View alarms</a> +
<input checked="" type="checkbox"/> ec2-advance	i-09a92f1458f7a6d6b	Stopped	t2.micro	-	<a href="#">View alarms</a> +
<input checked="" type="checkbox"/> ec2-advance	i-01ed61f04006e4aae	Stopped	t2.micro	-	<a href="#">View alarms</a> +
<input checked="" type="checkbox"/> ec2-advance	i-0f7631b46dd630b4b	Stopped	t2.micro	-	<a href="#">View alarms</a> +

The screenshot shows the AWS CloudWatch Metrics interface for monitoring EC2 instances. At the top, there are two rows of instance details:

	i-07115a32e51be3428	Running	t2.micro	Initializing	View alarms +
<input checked="" type="checkbox"/>	ec2-advance	Stopped	t2.micro	-	View alarms +

Below this, a message states: "Instances: i-05846983a782418e2 (ec2-advance), i-09a92f1458f7a6d6b (ec2-advance), i-01ed61f04006e4aae, i-0f7631b46dd630b4b (ec2-advance)".

The main area is titled "Monitoring" and displays four metrics over the last hour:

- CPU utilization (%): Percent, 15.9
- Network in (bytes): Bytes, 7.41k
- Network out (bytes): Bytes, 6.72k
- Network packets in (count): Count, 64.4

At the bottom right, it says "© 2024, Amazon Web Services, Inc. or its affiliates. Privacy"

62 Again refresh on the page. Now, we can see web application is running.

## Hello World from ip-172-31-89-248.ec2.internal

In this way, Auto Scaling Group continues to maintain a fixed number of instances even if an instance becomes unhealthy.