App Deployment Inside VPC On Private Server in Production

About the Project:

This example demonstrates how to create a VPC that you use for servers in a production environment.

To improve resiliency, you deploy the servers in two AZs by using an auto scaling group and an application load balancer. For additional security, you deploy the servers in private subnets. The server receives requests through the load balancer. The servers can connect to the internet gateway by using a NAT gateway. To improve resiliency, you deploy the NAT gateway in both availability zones.

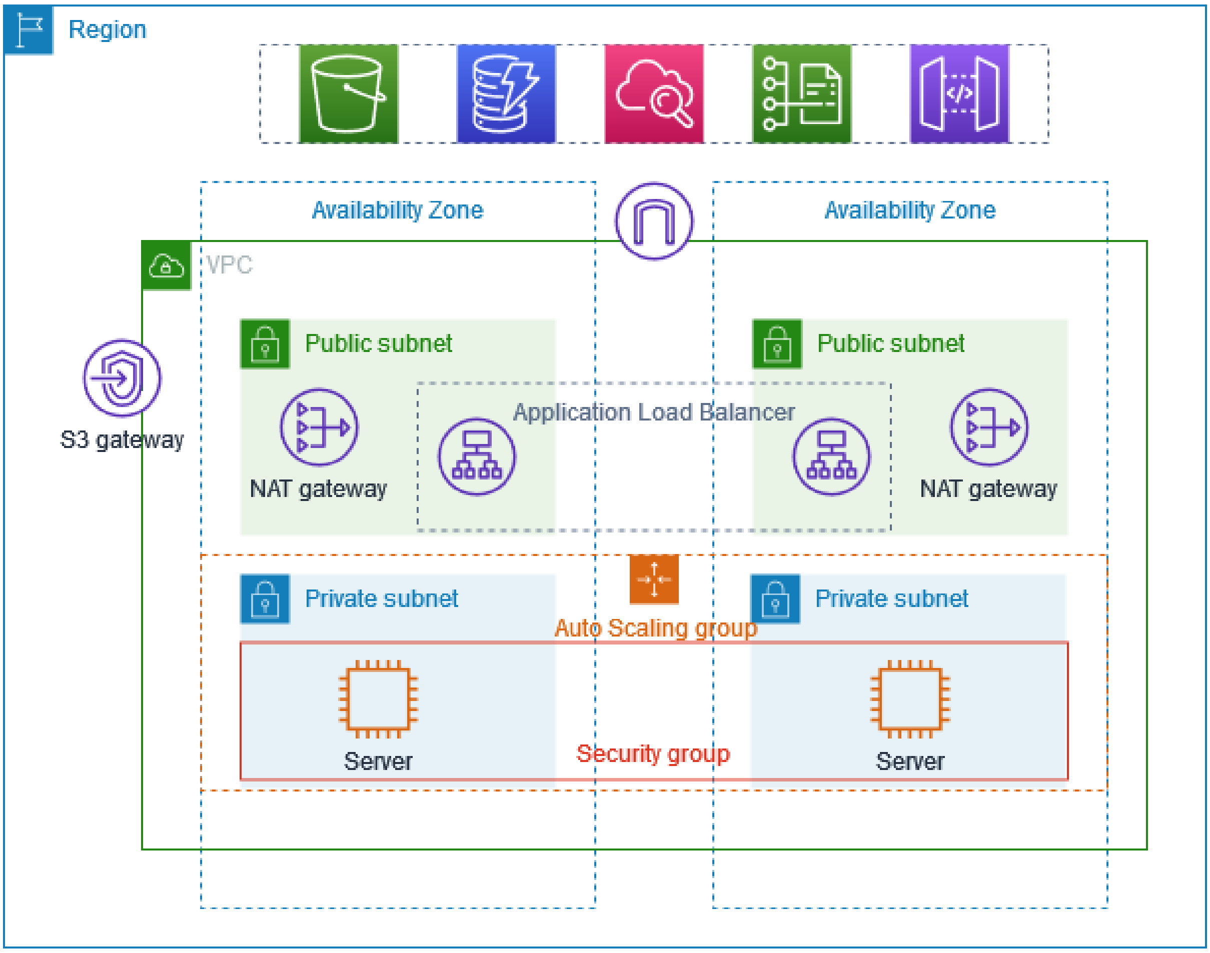
Project Overview:

The VPC has public subnets and private subnets in two availability zones.

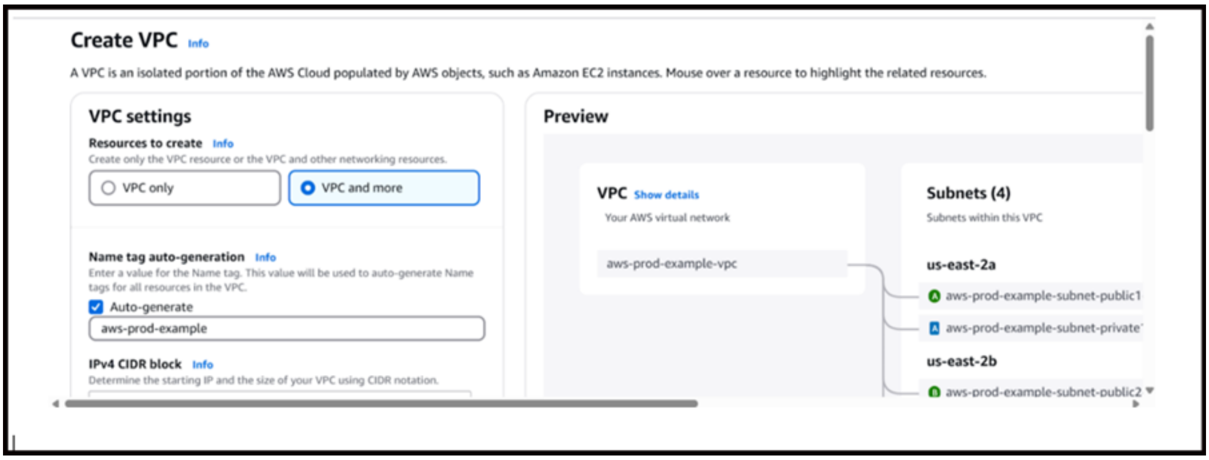
Each public subnet contains a NAT gateway and a load balancer.

The servers run in the private subnets are launched and terminated by using an auto-scaling group, and receives traffic from the load balancer.

The server can connect to the internet by using NAT gateway.



1. Go to VPC service and click on create VPC.
2. Follow the configuration and create a VPC as per your need.



Create VPC Info

A VPC is an isolated portion of the AWS Cloud populated by AWS objects, such as Amazon EC2 instances. Mouse over a resource to highlight the related resources.

VPC settings

Resources to create Info

Create only the VPC resource or the VPC and other networking resources.

○ VPC only ● VPC and more

Name tag auto-generation Info

Enter a value for the Name tag. This value will be used to auto-generate Name tags for all resources in the VPC.

☑ Auto-generate

aws-prod-example

IPv4 CIDR block Info

Determine the starting IP and the size of your VPC using CIDR notation.

Preview

VPC Show details

Your AWS virtual network

aws-prod-example-vpc

Subnets (4)

Subnets within this VPC

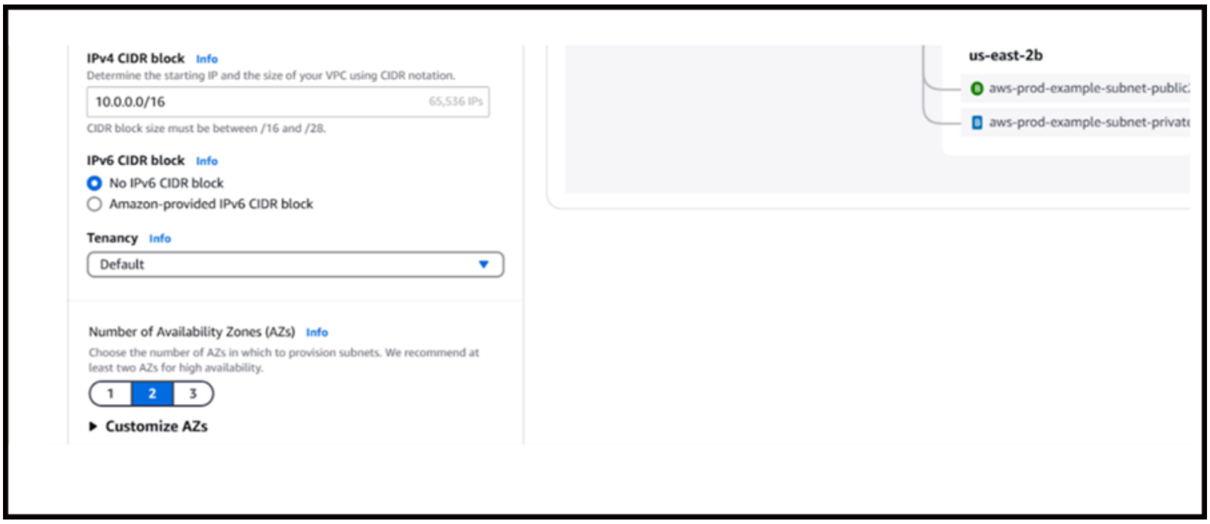
us-east-2a

aws-prod-example-subnet-public1

aws-prod-example-subnet-private

us-east-2b

aws-prod-example-subnet-public2



IPv4 CIDR block Info

Determine the starting IP and the size of your VPC using CIDR notation.

10.0.0.0/16 — 65,536 IPs

CIDR block size must be between /16 and /28.

IPv6 CIDR block Info

● No IPv6 CIDR block

○ Amazon-provided IPv6 CIDR block

Tenancy Info

Default ▼

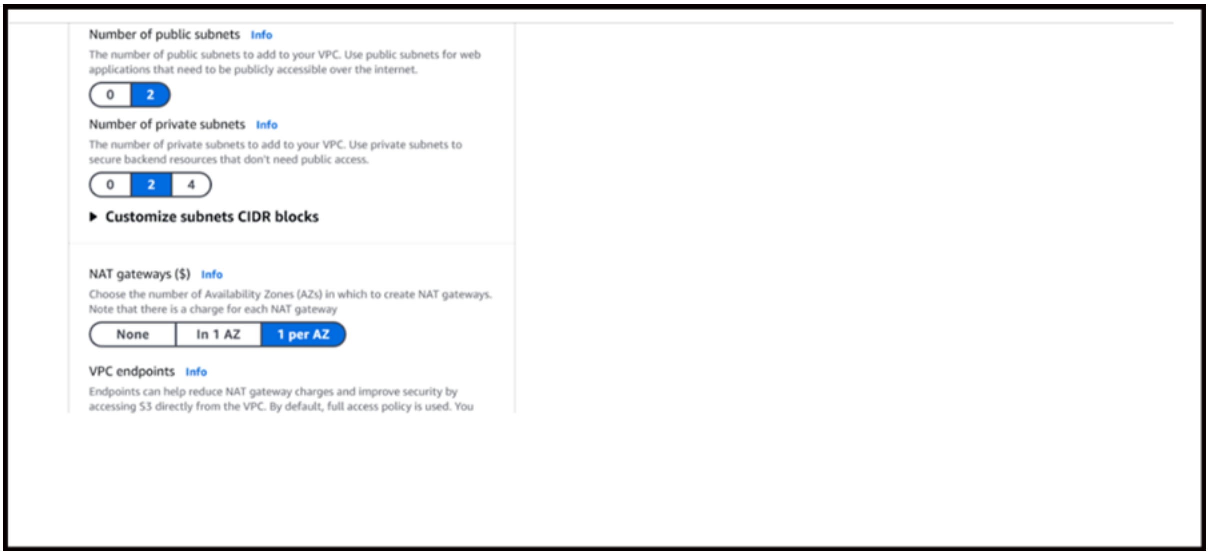
Number of Availability Zones (AZs) Info

Choose the number of AZs in which to provision subnets. We recommend at least two AZs for high availability.

1 | [✓ 2] | 3

▸ **Customize AZs**

us-east-2b
aws-prod-example-subnet-public
aws-prod-example-subnet-privat



Number of public subnets Info

The number of public subnets to add to your VPC. Use public subnets for web applications that need to be publicly accessible over the internet.

0 | [✓ 2]

Number of private subnets Info

The number of private subnets to add to your VPC. Use private subnets to secure backend resources that don't need public access.

0 | [✓ 2] | 4

▸ **Customize subnets CIDR blocks**

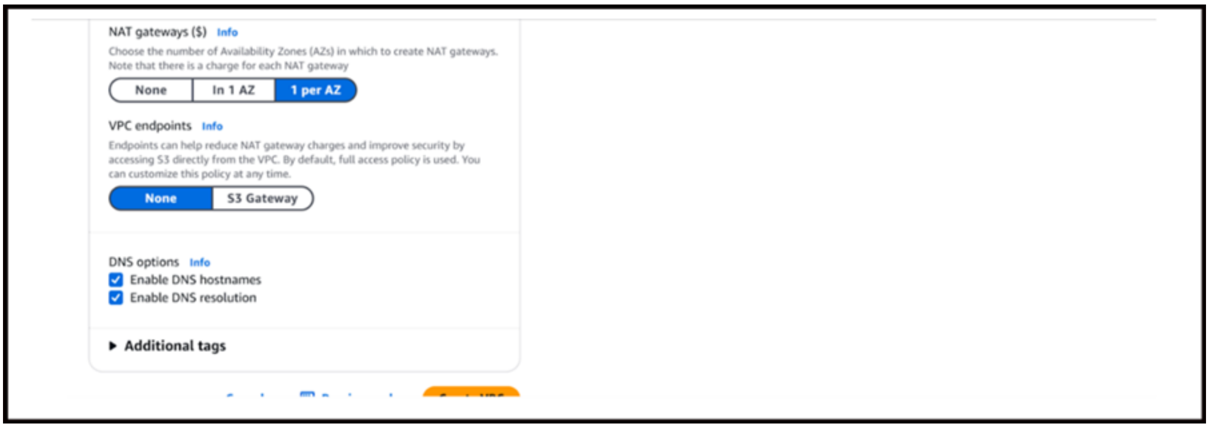
NAT gateways ($) Info

Choose the number of Availability Zones (AZs) in which to create NAT gateways. Note that there is a charge for each NAT gateway

None | In 1 AZ | [✓ 1 per AZ]

VPC endpoints Info

Endpoints can help reduce NAT gateway charges and improve security by accessing S3 directly from the VPC. By default, full access policy is used. You



NAT gateways ($) Info

Choose the number of Availability Zones (AZs) in which to create NAT gateways. Note that there is a charge for each NAT gateway

None | In 1 AZ | [✓ 1 per AZ]

VPC endpoints Info

Endpoints can help reduce NAT gateway charges and improve security by accessing S3 directly from the VPC. By default, full access policy is used. You can customize this policy at any time.

[✓ None] | S3 Gateway

DNS options Info

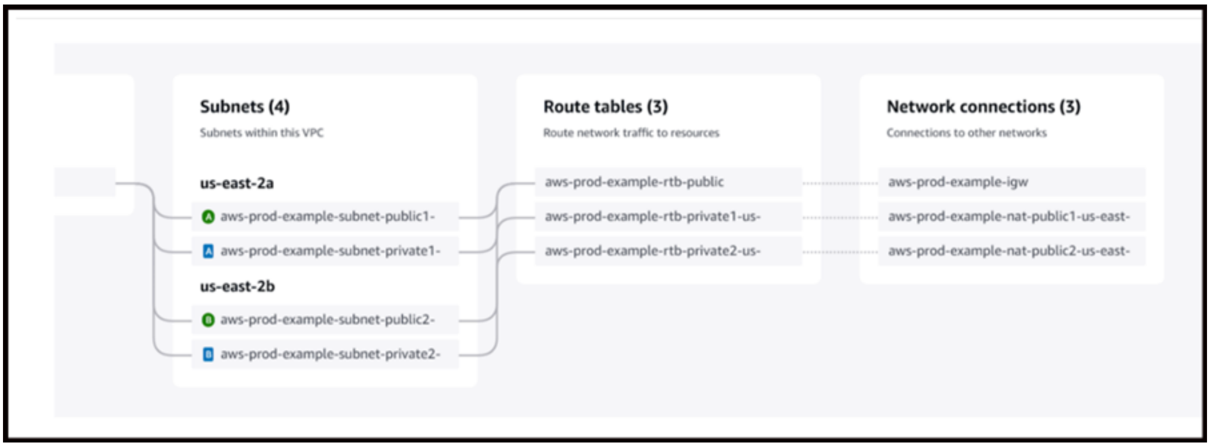
☑ Enable DNS hostnames

☑ Enable DNS resolution

▸ **Additional tags**

Now click on create VPC.

This would be the architecture of our VPC



Subnets (4)

Subnets within this VPC

us-east-2a

aws-prod-example-subnet-public1-

aws-prod-example-subnet-private1-

us-east-2b

aws-prod-example-subnet-public2-

aws-prod-example-subnet-private2-

Route tables (3)

Route network traffic to resources

aws-prod-example-rtb-public

aws-prod-example-rtb-private1-us-

aws-prod-example-rtb-private2-us-

Network connections (3)

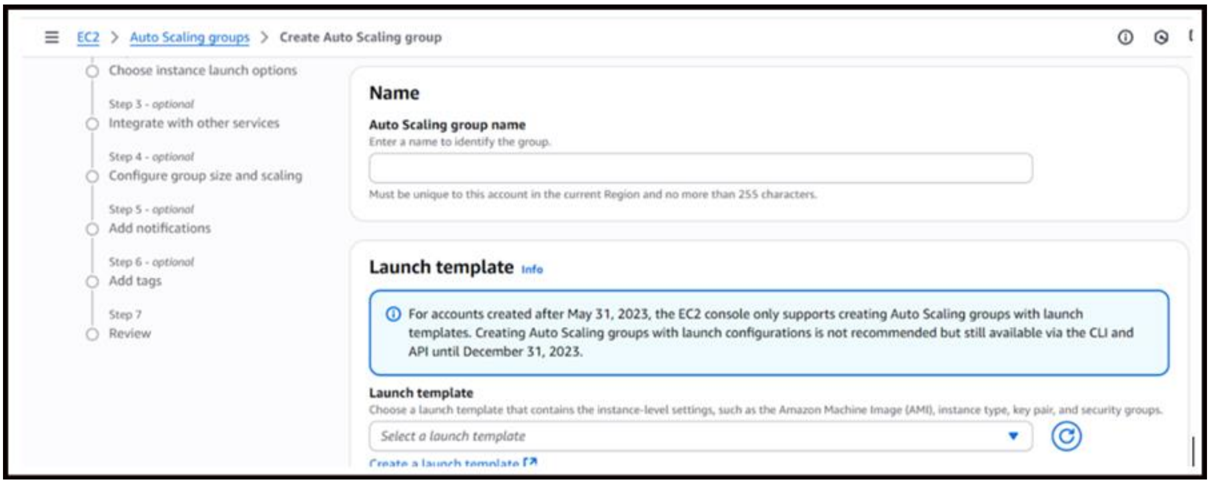
Connections to other networks

aws-prod-example-igw

aws-prod-example-nat-public1-us-east-

aws-prod-example-nat-public2-us-east-

3. Now create two EC2 instances with auto scaling and a load balancer. Go to the auto scaling group and create an auto-scaling group.



EC2 > Auto Scaling groups > Create Auto Scaling group

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

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.

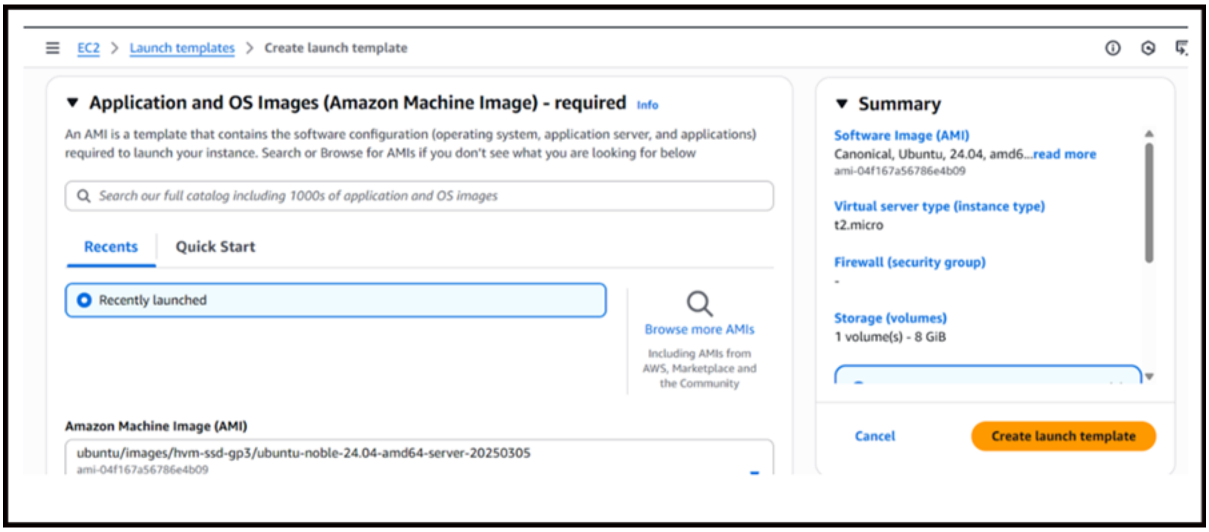
Launch template

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

Select a launch template

Create a launch template

4. Click on create a launch templates and select following configurations:



EC2 > Launch templates > Create 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

Recents | Quick Start

● Recently launched

Browse more AMIs

Including AMIs from AWS, Marketplace and the Community

Amazon Machine Image (AMI)

ubuntu/images/hvm-ssd-gp3/ubuntu-noble-24.04-amd64-server-20250305
ami-04f167a56786e4b09

Summary

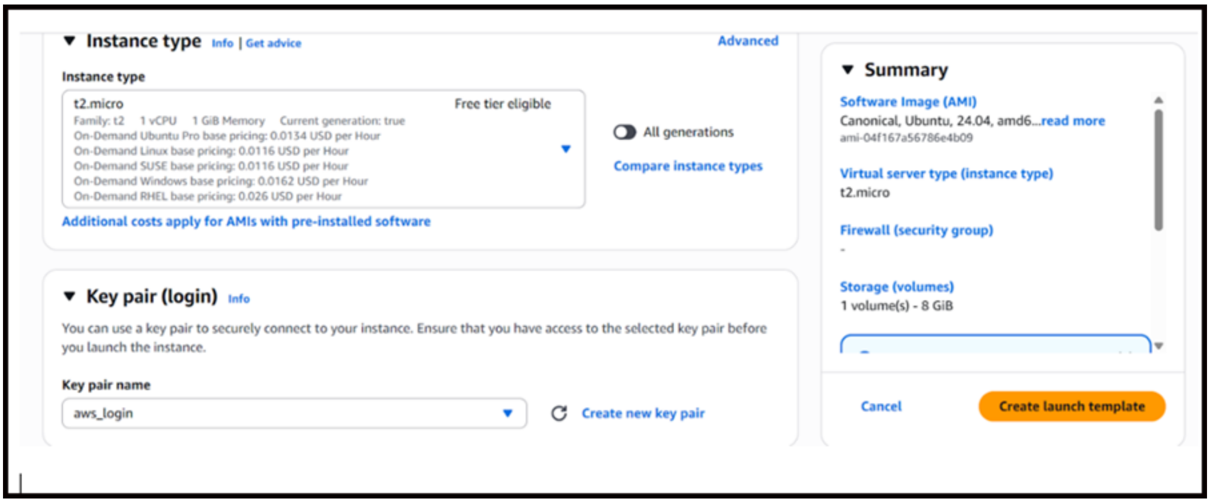
Software Image (AMI)
Canonical, Ubuntu, 24.04, amd6...read more
ami-04f167a56786e4b09

Virtual server type (instance type)
t2.micro

Firewall (security group)
-

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

Cancel | Create launch template



▼ Instance type Info | Get advice

Advanced

Instance type

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

All generations

Compare instance types

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: aws_login

Create new key pair

▼ Summary

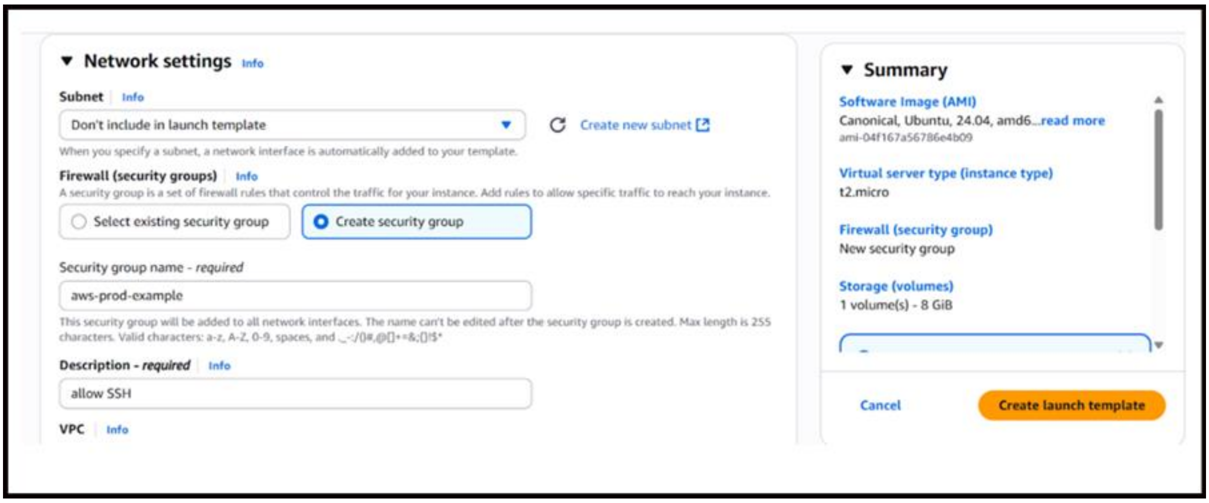
Software Image (AMI)
Canonical, Ubuntu, 24.04, amd6...read more
ami-04f167a56786e4b09

Virtual server type (instance type)
t2.micro

Firewall (security group)
-

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

Cancel | Create launch template



▼ Network settings Info

Subnet Info: 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) 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.

○ Select existing security group ● Create security group

Security group name - *required*: aws-prod-example

This security group will be added to all network interfaces. The name can't be edited after the security group is created. Max length is 255 characters. Valid characters: a-z, A-Z, 0-9, spaces, and ._-:/()#,@[]+=&;{}!$*

Description - *required* Info: allow SSH

VPC Info

▼ Summary

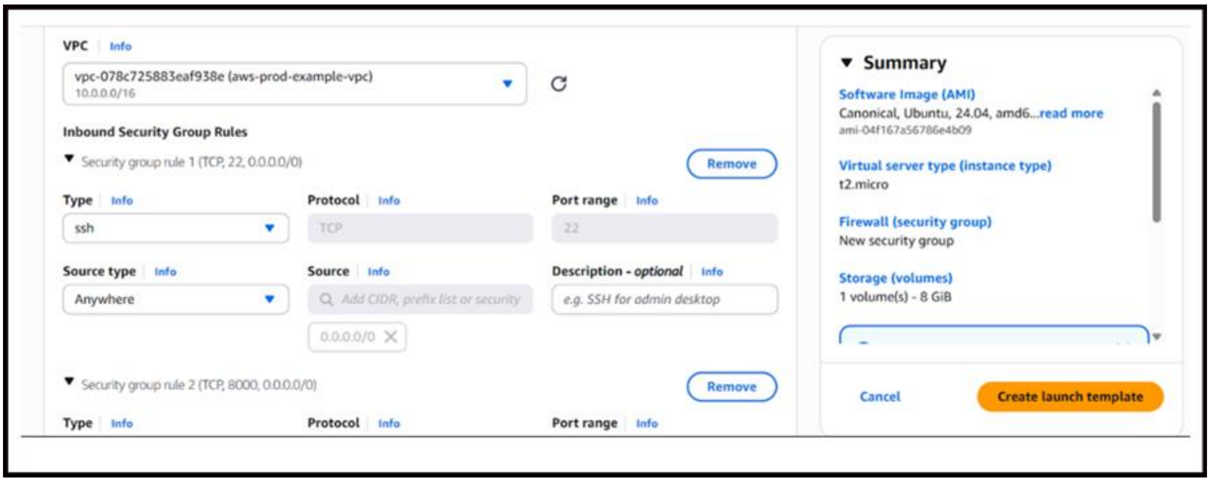
Software Image (AMI)
Canonical, Ubuntu, 24.04, amd6...read more
ami-04f167a56786e4b09

Virtual server type (instance type)
t2.micro

Firewall (security group)
New security group

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

Cancel | Create launch template



VPC Info: vpc-078c725883eaf938e (aws-prod-example-vpc)
10.0.0.0/16

Inbound Security Group Rules

▼ Security group rule 1 (TCP, 22, 0.0.0.0/0) — Remove

Type Info: ssh
Protocol Info: TCP
Port range Info: 22
Source type Info: Anywhere
Source Info: Add CIDR, prefix list or security — 0.0.0.0/0
Description - *optional* Info: e.g. SSH for admin desktop

▼ Security group rule 2 (TCP, 8000, 0.0.0.0/0) — Remove

Type Info | Protocol Info | Port range Info

▼ Summary

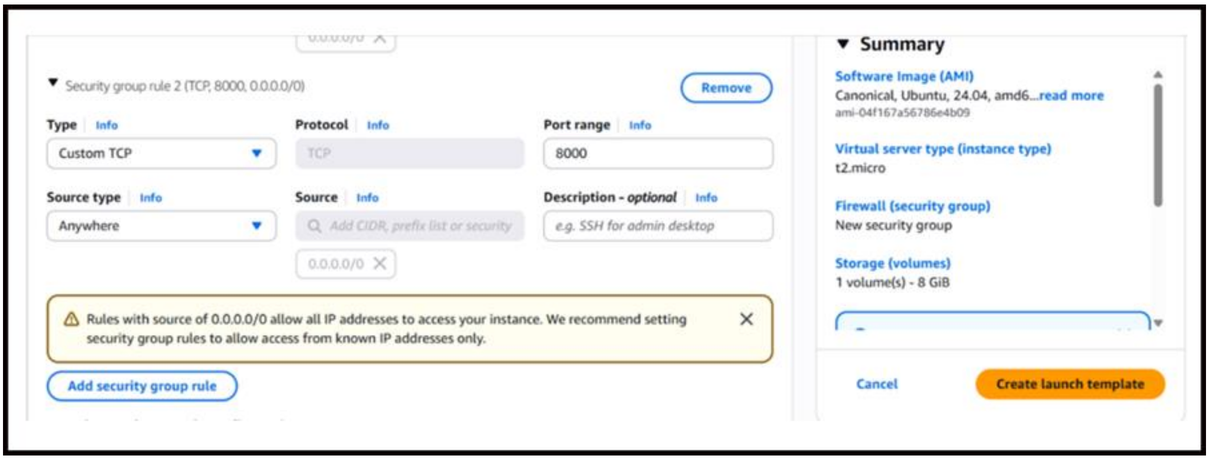
Software Image (AMI)
Canonical, Ubuntu, 24.04, amd6...read more
ami-04f167a56786e4b09

Virtual server type (instance type)
t2.micro

Firewall (security group)
New security group

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

Cancel | Create launch template



Security group rule 2 (TCP, 8000, 0.0.0.0/0)

Type: Custom TCP

Protocol: TCP

Port range: 8000

Source type: Anywhere

Source: 0.0.0.0/0

Description - *optional*: e.g. SSH for admin desktop

Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

Summary

Software Image (AMI)
Canonical, Ubuntu, 24.04, amd6...read more
ami-04f167a56786e4b09

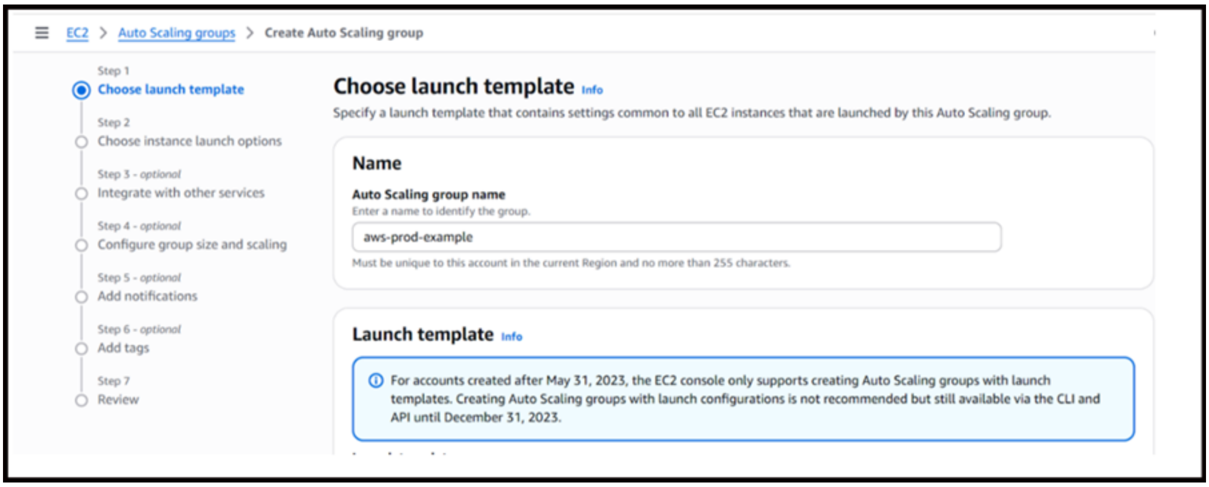
Virtual server type (instance type)
t2.micro

Firewall (security group)
New security group

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

And click on the launch template.

5. Do refresh the page and go back to the auto-scaling group and fill the necessary thing.



EC2 > Auto Scaling groups > Create Auto Scaling group

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

Choose launch template

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

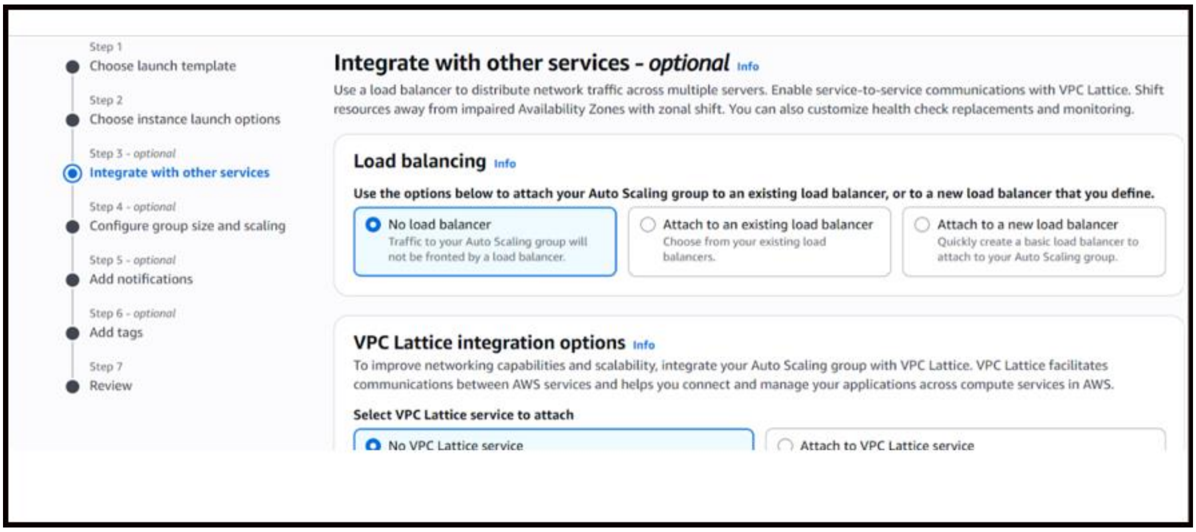
Name

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

Launch template

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.

and choose the template you created and click on the next button.



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.

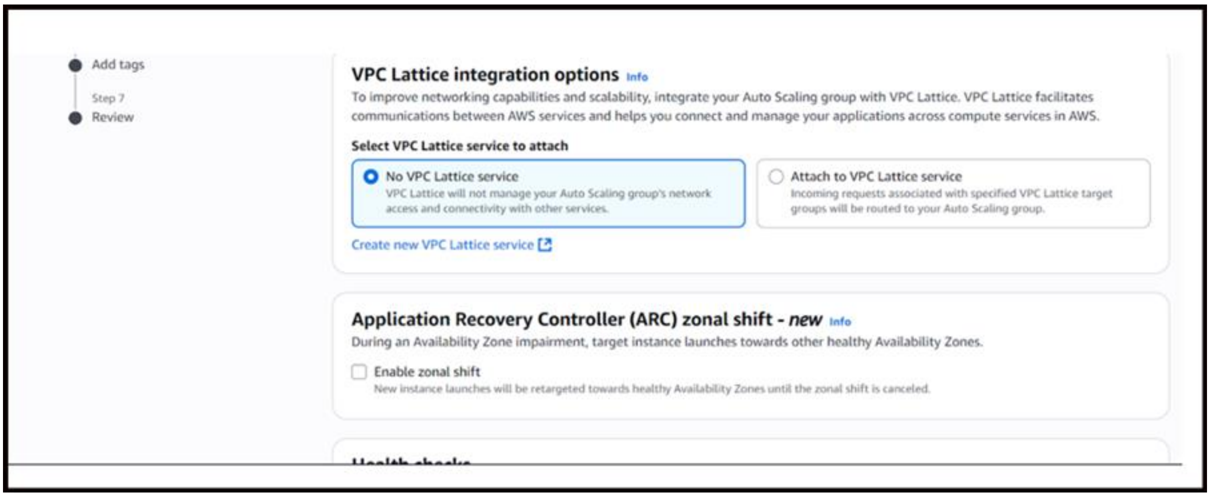
● No load balancer — Traffic to your Auto Scaling group will not be fronted by a load balancer.
○ Attach to an existing load balancer — Choose from your existing load balancers.
○ Attach to a new load balancer — Quickly create a basic load balancer to attach to your Auto Scaling group.

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.

Select VPC Lattice service to attach

● No VPC Lattice service
○ Attach to VPC Lattice service



● Add tags

Step 7

● Review

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 network access and connectivity with other services.

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

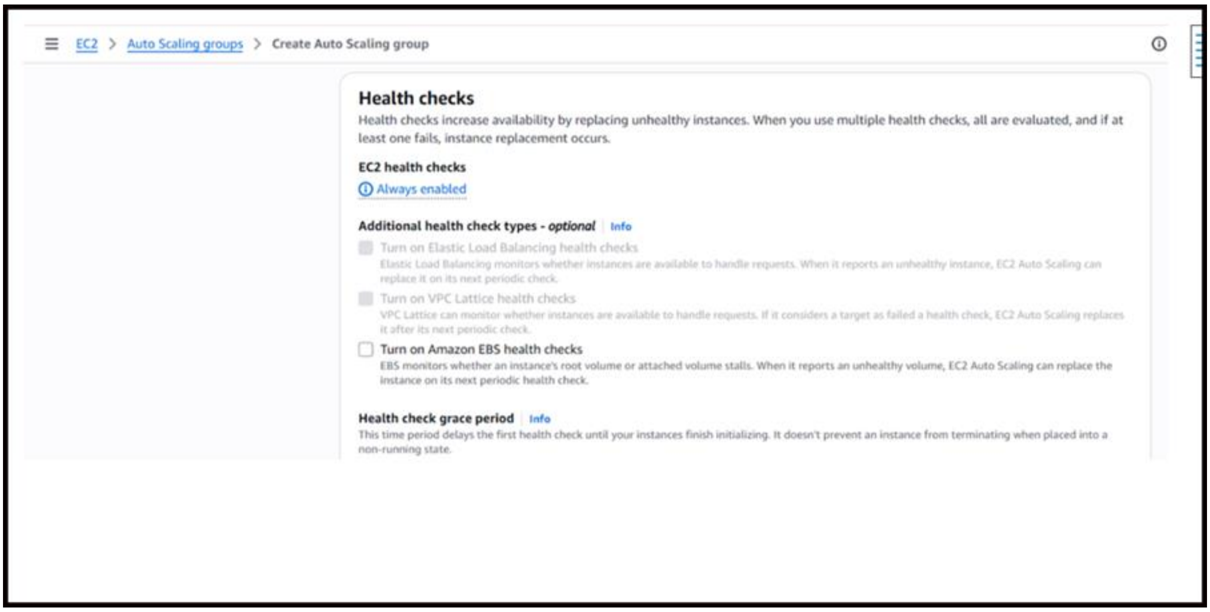
Create new VPC Lattice service

Application Recovery Controller (ARC) zonal shift - *new* Info

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

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

Health checks



EC2 > Auto Scaling groups > Create Auto Scaling group

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

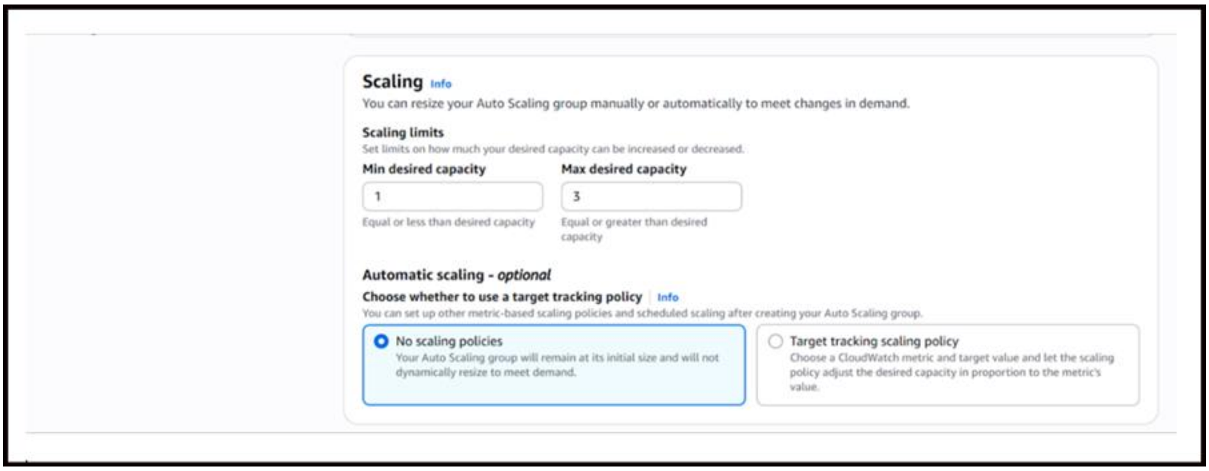
□ Turn on Elastic Load Balancing health checks
Elastic Load Balancing monitors whether instances are available to handle requests. When it reports an unhealthy instance, EC2 Auto Scaling can replace it on its next periodic check.

□ 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 stalls. When it reports an unhealthy volume, EC2 Auto Scaling can replace the instance on its next periodic health 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.



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

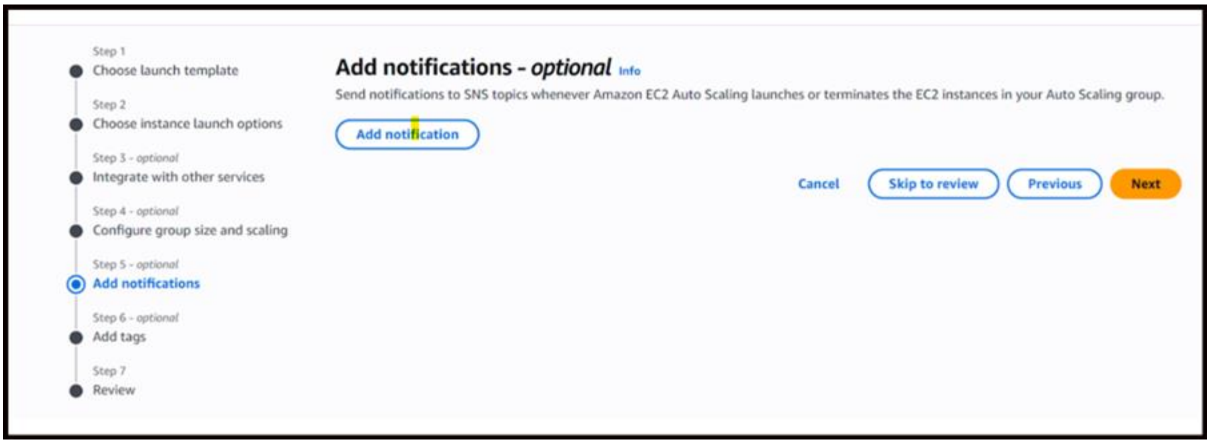
Automatic scaling - *optional*

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.

● No scaling policies
Your Auto Scaling group will remain at its initial size and will not dynamically resize to meet demand.

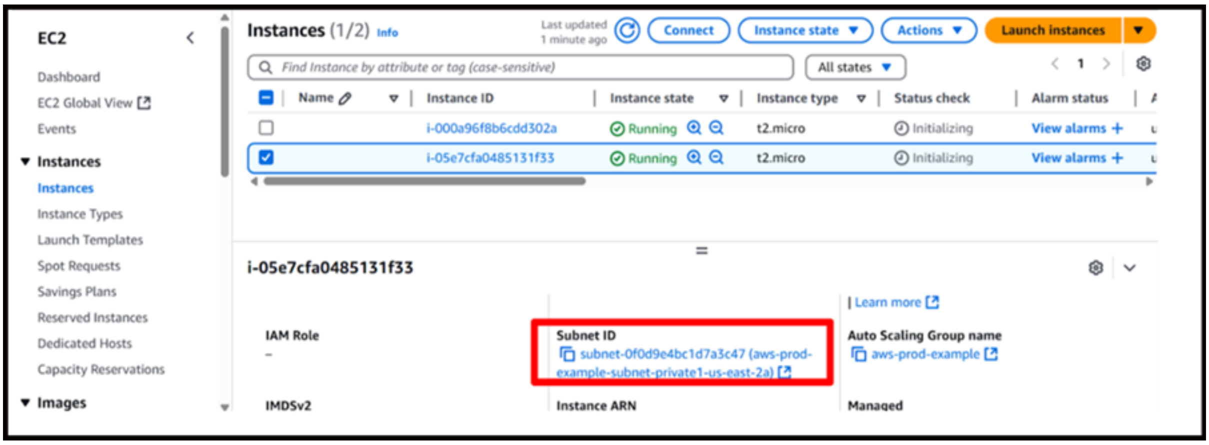
○ Target tracking scaling policy
Choose a CloudWatch metric and target value and let the scaling policy adjust the desired capacity in proportion to the metric's value.



Add notifications - *optional* Info

Send notifications to SNS topics whenever Amazon EC2 Auto Scaling launches or terminates the EC2 instances in your Auto Scaling group.

click on the next button after doing all above configurations. Verify all your configurations and click on create auto-scaling group.



Instances (1/2) Info

Find Instance by attribute or tag (case-sensitive)

	Name	Instance ID	Instance state	Instance type	Status check	Alarm status
□		i-000a96f8b6cdd302a	Running	t2.micro	Initializing	View alarms +
☑		i-05e7cfa0485131f33	Running	t2.micro	Initializing	View alarms +

i-05e7cfa0485131f33

| Learn more

IAM Role
–

Subnet ID
subnet-0f0d9e4bc1d7a3c47 (aws-prod-example-subnet-private1-us-east-2a)

Auto Scaling Group name
aws-prod-example

IMDSv2

Instance ARN

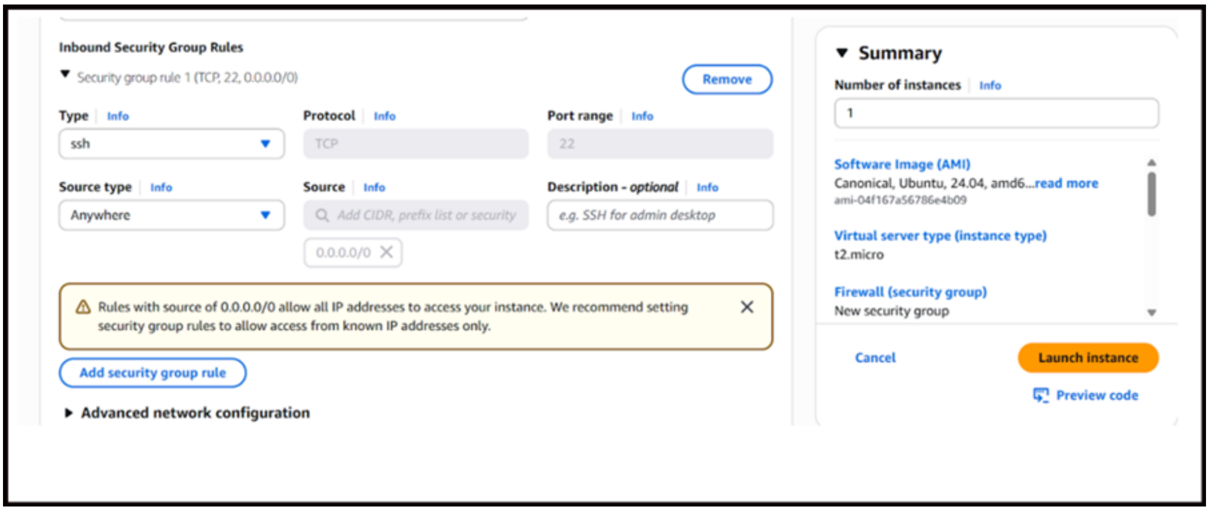
Managed

Two instances got created in us-east-2a and us-east-2b

These instances don't have an IP address in public because we created a private subnets server. So we use a jump server approach to connect to them.

Jump Server - In AWS, a jump server (also called a bastion host) is a dedicated server that acts as a secure gateway, providing access to private instances within a VPC (Virtual Private Cloud) from outside the VPC. It's a common security practice that reduces the risk of unauthorized access to internal resources by funneling all external SSH traffic through a single, secure point.

For this go to launch instance and create a new instance as bastion host or jump server with created VPC and public subnets and give access to SSH login.



Inbound Security Group Rules

Security group rule 1 (TCP, 22, 0.0.0.0/0)

Type: ssh

Protocol: TCP

Port range: 22

Source type: Anywhere

Source: 0.0.0.0/0

Description - optional: e.g. SSH for admin desktop

Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

Summary

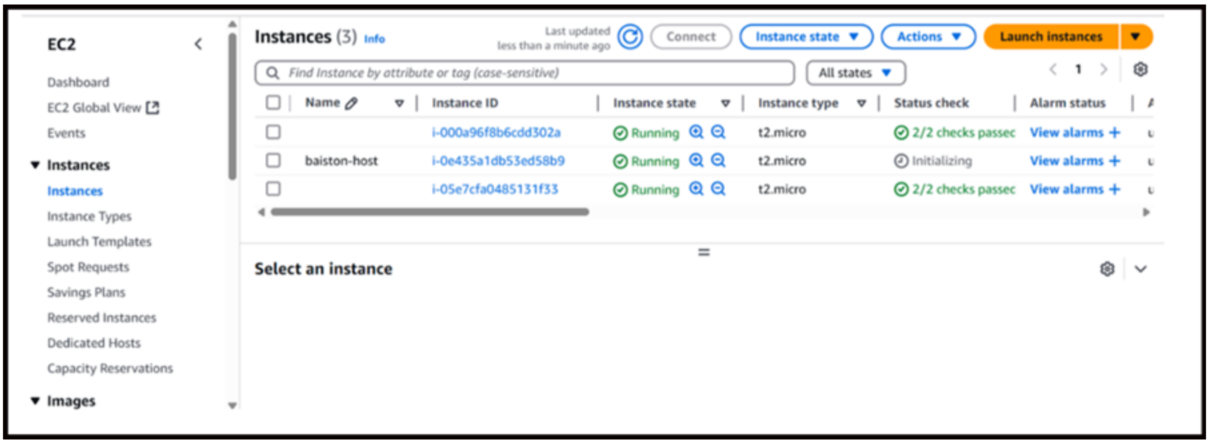
Number of instances: 1

Software Image (AMI): Canonical, Ubuntu, 24.04, amd6...read more ami-04f167a56786e4b09

Virtual server type (instance type): t2.micro

Firewall (security group): New security group

And launch the instance. Now you will have 3 instances.



Instances (3)

Last updated less than a minute ago

Name	Instance ID	Instance state	Instance type	Status check	Alarm status
	i-000a96f8b6cdd302a	Running	t2.micro	2/2 checks passec	View alarms +
baiston-host	i-0e435a1db53ed58b9	Running	t2.micro	Initializing	View alarms +
	i-05e7cfa0485131f33	Running	t2.micro	2/2 checks passec	View alarms +

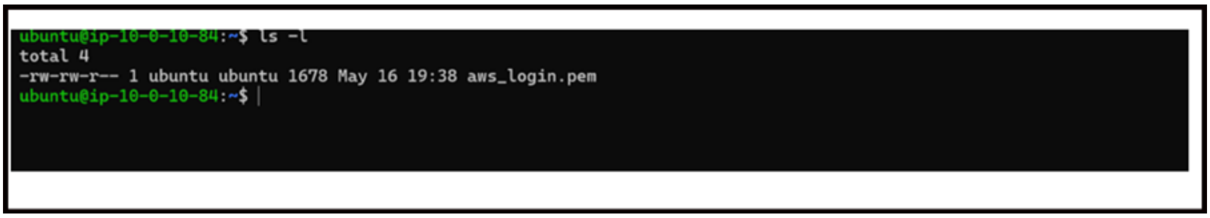
Select an instance

2 created with an auto-scaling group and one as a jump server.

To connect with private subnets, your jump server should connect to the SSH because we want to connect 2 instances created by auto-scaling.

So we have to copy the pem file of 2 instances from our local server to the jump server.SCP Command - scp -i "aws_login.pem" aws_login.pem ubuntu@ec2-18-116-80-56.us-east-2.compute.amazonaws.com:/home/ubuntu

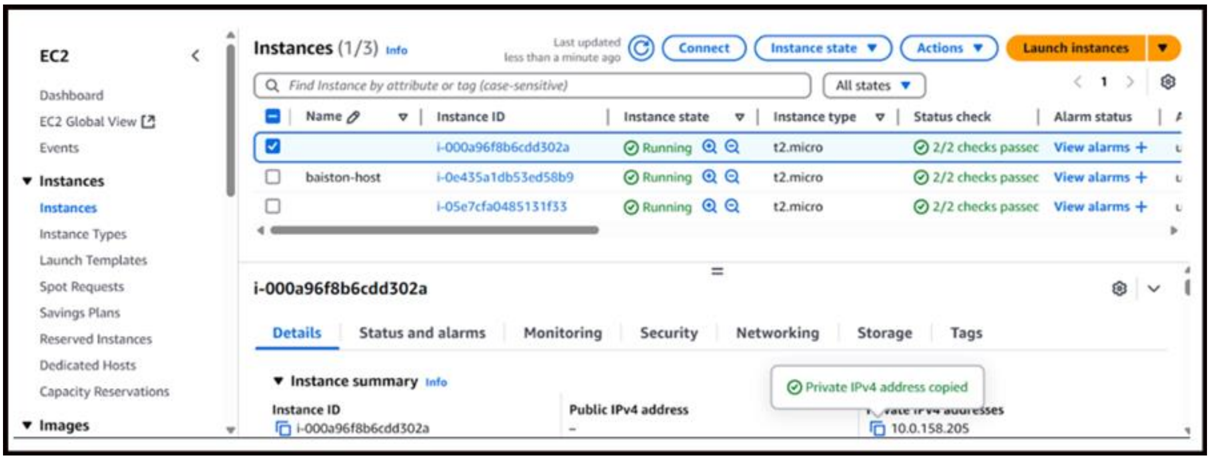
After copying file login to the jump server by using SSH key.



```
ubuntu@ip-10-0-10-84:~$ ls -l
total 4
-rw-rw-r-- 1 ubuntu ubuntu 1678 May 16 19:38 aws_login.pem
ubuntu@ip-10-0-10-84:~$ |
```

This pem file should be here in your jump server.

Now go to any one instance and take the private IP address and launch an application.



	Name	Instance ID	Instance state	Instance type	Status check	Alarm status
☑		i-000a96f8b6cdd302a	Running	t2.micro	2/2 checks passec	View alarms +
□	baiston-host	i-0e435a1db53ed58b9	Running	t2.micro	2/2 checks passec	View alarms +
□		i-05e7cfa0485131f33	Running	t2.micro	2/2 checks passec	View alarms +

i-000a96f8b6cdd302a

▼ Instance summary Info

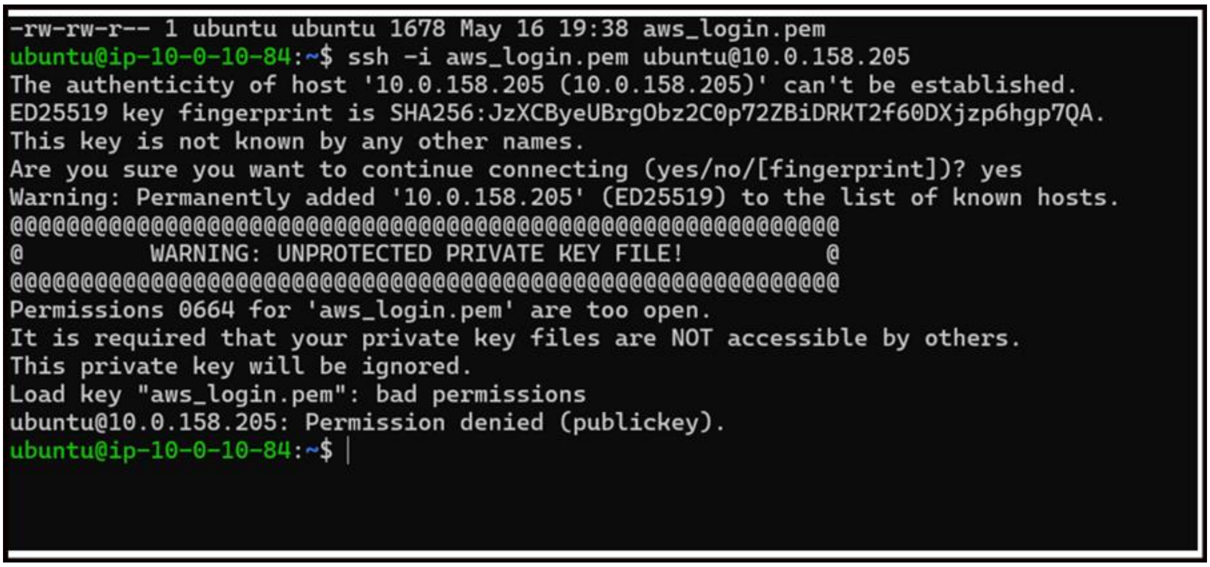
Instance ID: i-000a96f8b6cdd302a

Public IPv4 address: –

Private IPv4 addresses: 10.0.158.205

Private IPv4 address copied

And do ssh in instance.

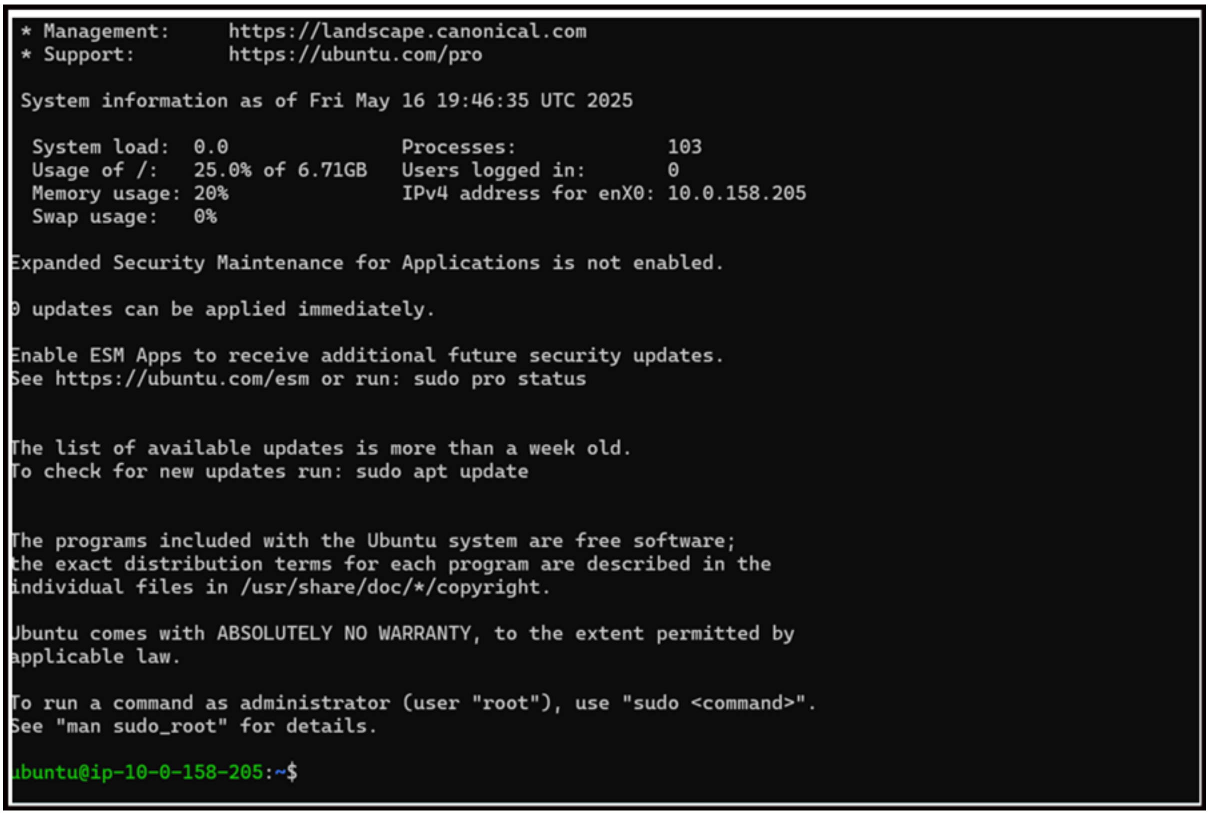


```
-rw-rw-r-- 1 ubuntu ubuntu 1678 May 16 19:38 aws_login.pem
ubuntu@ip-10-0-10-84:~$ ssh -i aws_login.pem ubuntu@10.0.158.205
The authenticity of host '10.0.158.205 (10.0.158.205)' can't be established.
ED25519 key fingerprint is SHA256:JzXCByeUBrgObz2C0p72ZBiDRKT2f60DXjzp6hgp7QA.
This key is not known by any other names.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added '10.0.158.205' (ED25519) to the list of known hosts.
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@         WARNING: UNPROTECTED PRIVATE KEY FILE!          @
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
Permissions 0664 for 'aws_login.pem' are too open.
It is required that your private key files are NOT accessible by others.
This private key will be ignored.
Load key "aws_login.pem": bad permissions
ubuntu@10.0.158.205: Permission denied (publickey).
ubuntu@ip-10-0-10-84:~$
```

Change the permission of pem to login to the server.

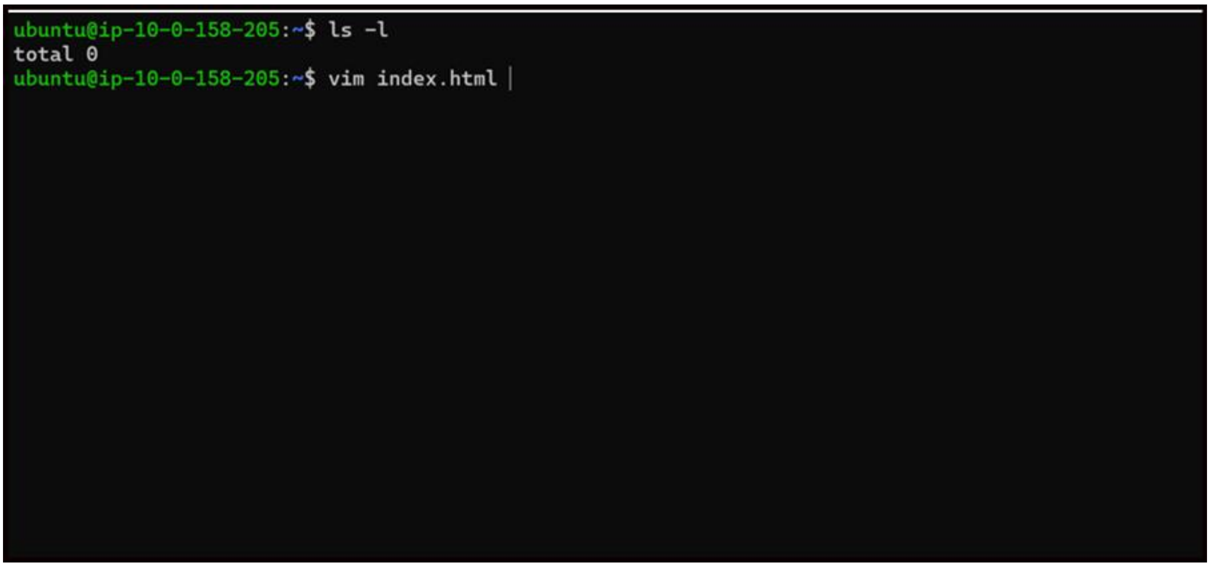
chmod 600 aws_login.pem

and do ssh -i aws_login.pem ubuntu@10.0.158.205 and get connected to the server.

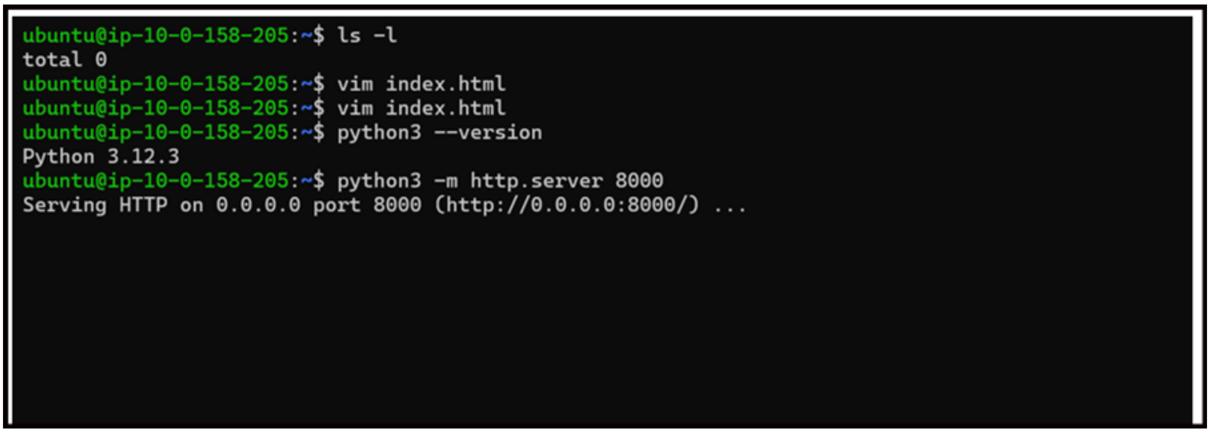


```
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

 System information as of Fri May 16 19:46:35 UTC 2025

  System load:  0.0               Processes:             103
  Usage of /:   25.0% of 6.71GB   Users logged in:       0
  Memory usage: 20%               IPv4 address for enX0: 10.0.158.205
  Swap usage:   0%

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status


The list of available updates is more than a week old.
To check for new updates run: sudo apt update


The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

ubuntu@ip-10-0-158-205:~$
```

create an index.html and write a basic html inside the index.html file.

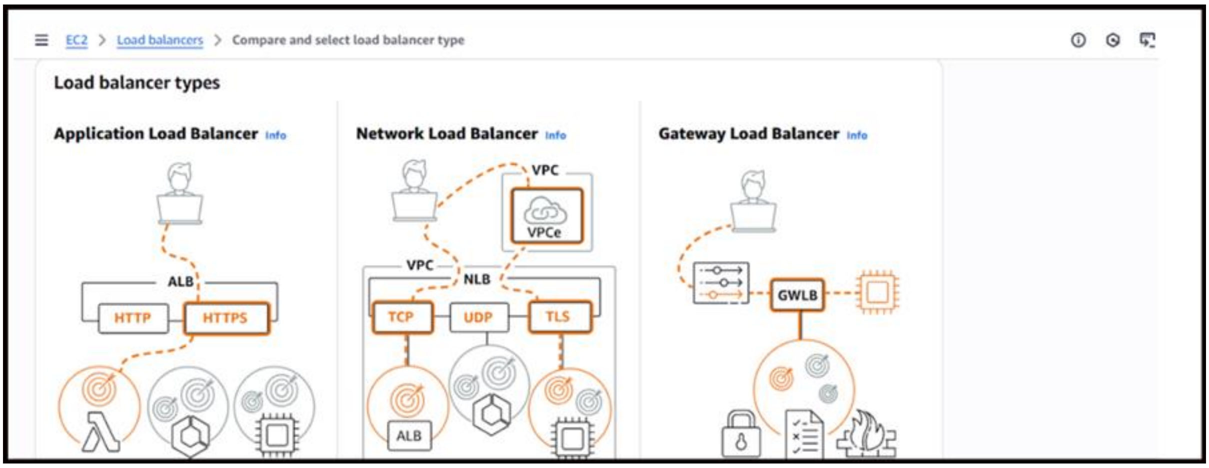


```
ubuntu@ip-10-0-158-205:~$ ls -l
total 0
ubuntu@ip-10-0-158-205:~$ vim index.html
```



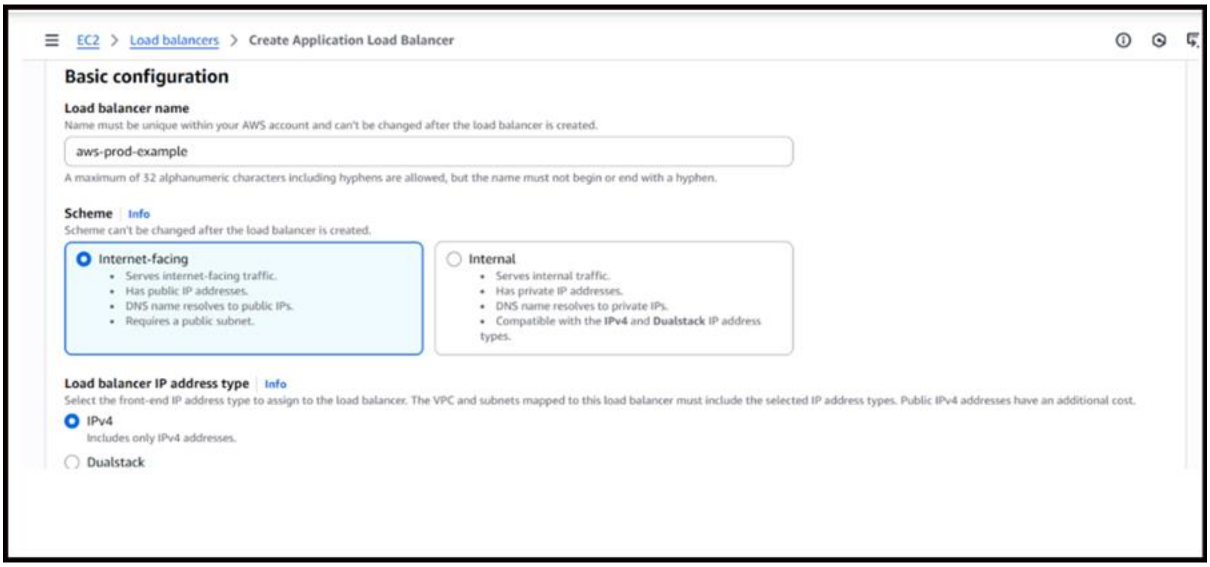
```
ubuntu@ip-10-0-158-205:~$ ls -l
total 0
ubuntu@ip-10-0-158-205:~$ vim index.html
ubuntu@ip-10-0-158-205:~$ vim index.html
ubuntu@ip-10-0-158-205:~$ python3 --version
Python 3.12.3
ubuntu@ip-10-0-158-205:~$ python3 -m http.server 8000
Serving HTTP on 0.0.0.0 port 8000 (http://0.0.0.0:8000/) ...
```

Now go to the load balancers section and create a default application load balancer which moves all your traffic to application (python).



EC2 > Load balancers > Compare and select load balancer type
Load balancer types
Application Load Balancer Info
ALB
HTTP
HTTPS
Network Load Balancer Info
VPC
VPCe
VPC
NLB
TCP
UDP
TLS
ALB
Gateway Load Balancer Info
GWLB

Application load balancer - The Application Load Balancer distributes incoming HTTP and HTTPS traffic across multiple targets such as Amazon EC2 instances, microservices, and containers, based on request attributes. When the load balancer receives a connection request, it evaluates the listener rules in priority order to determine which rule to apply, and if applicable, it selects a target from the target group for the rule action.



Basic configuration

Load balancer name

Name must be unique within your AWS account and can't be changed after the load balancer is created.

aws-prod-example

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
- Serves internet-facing traffic.
- Has public IP addresses.
- DNS name resolves to public IPs.
- Requires a public subnet.

○ Internal
- Serves internal traffic.
- Has private IP addresses.
- DNS name resolves to private IPs.
- Compatible with the **IPv4** and **Dualstack** IP address types.

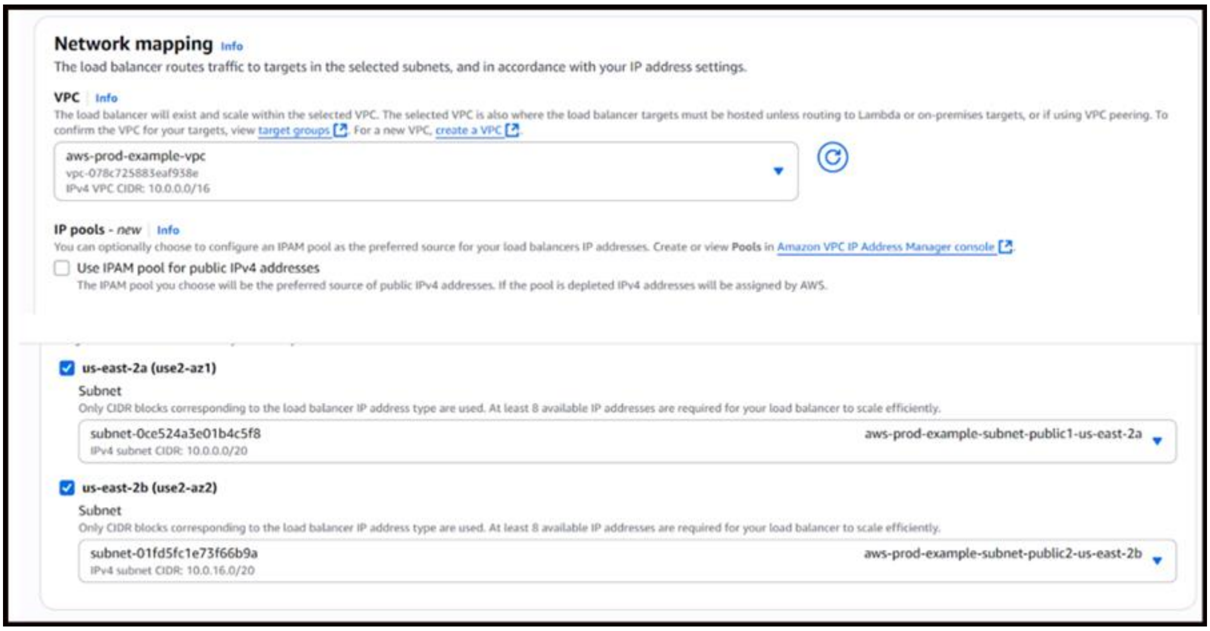
Load balancer IP address type | Info

Select the front-end IP address type to assign to the load balancer. The VPC and subnets mapped to this load balancer must include the selected IP address types. Public IPv4 addresses have an additional cost.

● IPv4

Includes only IPv4 addresses.

○ Dualstack



Network mapping Info

The load balancer routes traffic to targets in the selected subnets, and in accordance with your IP address settings.

VPC | Info

The load balancer will exist and scale within the selected VPC. The selected VPC is also where the load balancer targets must be hosted unless routing to Lambda or on-premises targets, or if using VPC peering. To confirm the VPC for your targets, view target groups. For a new VPC, create a VPC.

aws-prod-example-vpc
vpc-078c725883eaf938e
IPv4 VPC CIDR: 10.0.0.0/16

IP pools - *new* | Info

You can optionally choose to configure an IPAM pool as the preferred source for your load balancers IP addresses. Create or view **Pools** in Amazon VPC IP Address Manager console.

□ Use IPAM pool for public IPv4 addresses

The IPAM pool you choose will be the preferred source of public IPv4 addresses. If the pool is depleted IPv4 addresses will be assigned by AWS.

☑ us-east-2a (use2-az1)

Subnet

Only CIDR blocks corresponding to the load balancer IP address type are used. At least 8 available IP addresses are required for your load balancer to scale efficiently.

subnet-0ce524a3e01b4c5f8 aws-prod-example-subnet-public1-us-east-2a
IPv4 subnet CIDR: 10.0.0.0/20

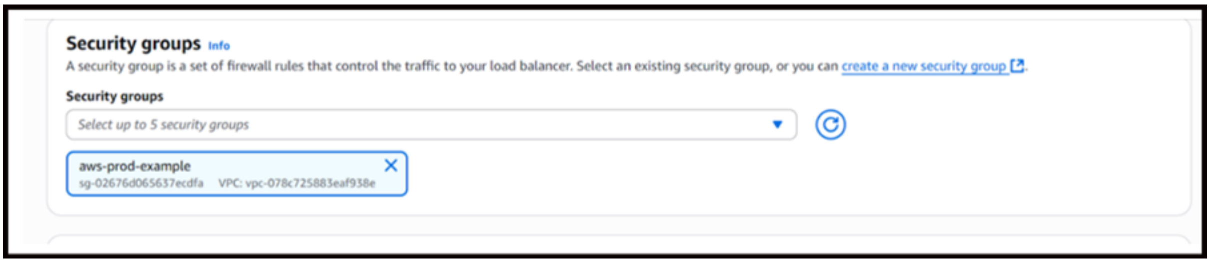
☑ us-east-2b (use2-az2)

Subnet

Only CIDR blocks corresponding to the load balancer IP address type are used. At least 8 available IP addresses are required for your load balancer to scale efficiently.

subnet-01fd5fc1e73f66b9a aws-prod-example-subnet-public2-us-east-2b
IPv4 subnet CIDR: 10.0.16.0/20

Give the public subnet in subnet section because in our diagram load balancer attached to the public subnet.



Security groups Info

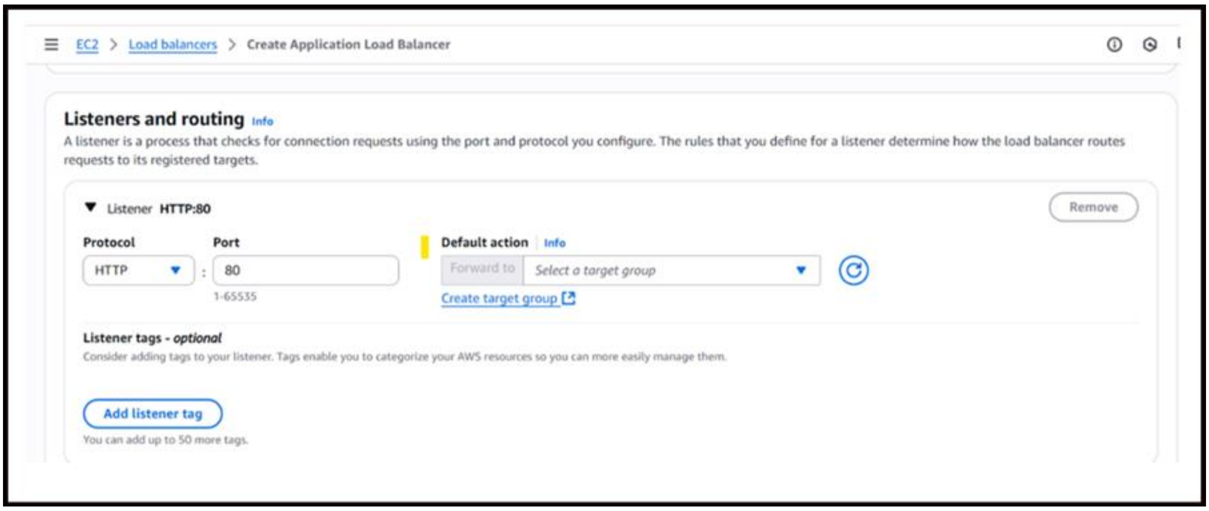
A security group is a set of firewall rules that control the traffic to your load balancer. Select an existing security group, or you can create a new security group.

Security groups

Select up to 5 security groups

aws-prod-example ✕
sg-02676d065637ecdfa VPC: vpc-078c725883eaf938e

create a target group for listening to port 8000 and after configuring all the things click on next.



EC2 > Load balancers > Create Application Load Balancer

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: *Select a target group* ▼

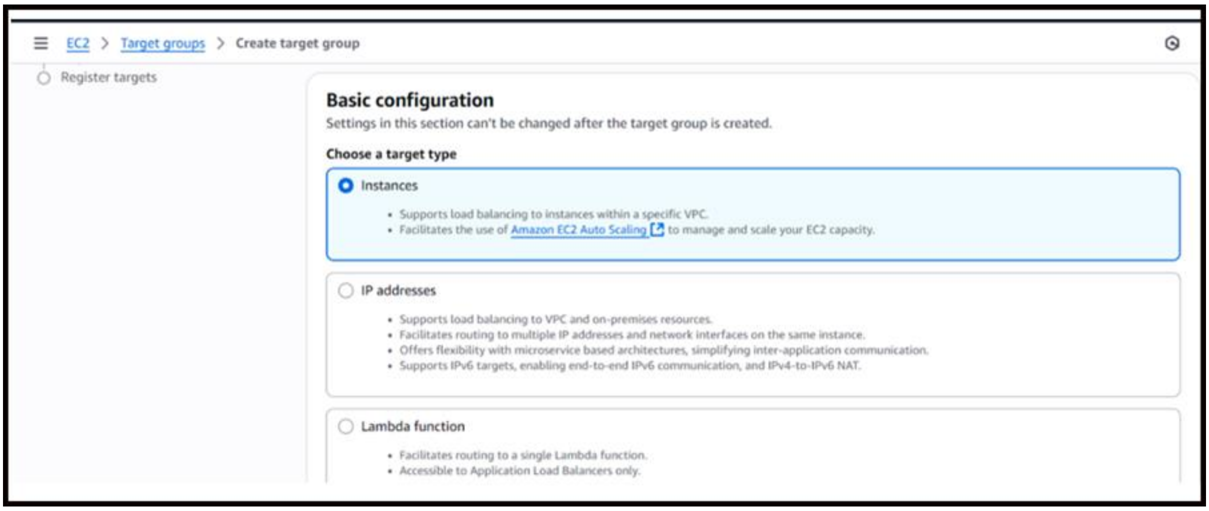
Create target group

Listener tags - *optional*

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

Add listener tag

You can add up to 50 more tags.



EC2 > Target groups > Create target group

○ Register targets

Basic configuration

Settings in this section can't be changed after the target group is created.

Choose a target type

● Instances

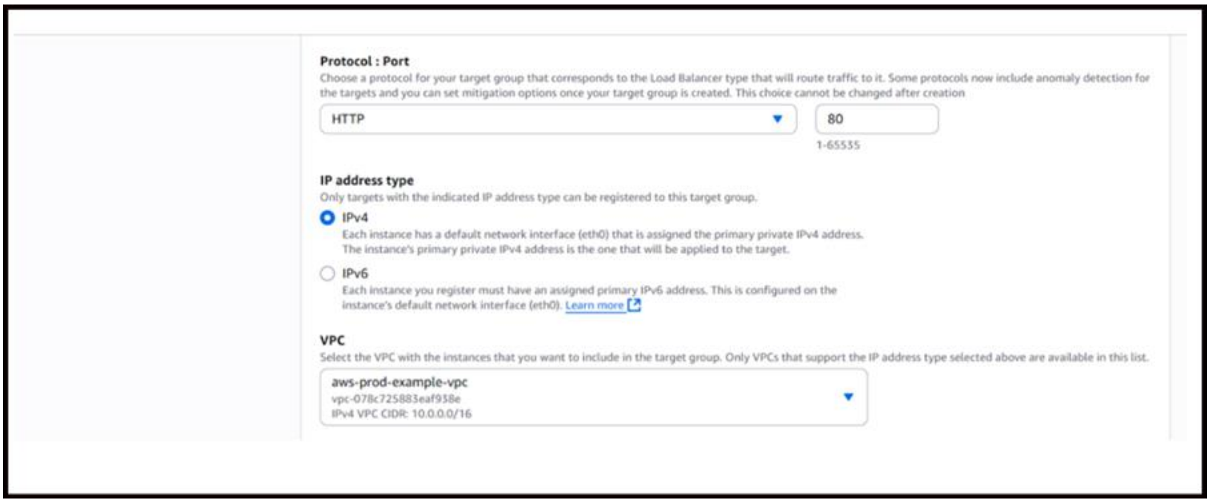
- Supports load balancing to instances within a specific VPC.
- Facilitates the use of Amazon EC2 Auto Scaling to manage and scale your EC2 capacity.

○ IP addresses

- Supports load balancing to VPC and on-premises resources.
- Facilitates routing to multiple IP addresses and network interfaces on the same instance.
- Offers flexibility with microservice based architectures, simplifying inter-application communication.
- Supports IPv6 targets, enabling end-to-end IPv6 communication, and IPv4-to-IPv6 NAT.

○ Lambda function

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



Protocol : Port

Choose a protocol for your target group that corresponds to the Load Balancer type that will route traffic to it. Some protocols now include anomaly detection for the targets and you can set mitigation options once your target group is created. This choice cannot be changed after creation

HTTP ▼ : 80

1-65535

IP address type

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

● IPv4

Each instance has a default network interface (eth0) that is assigned the primary private IPv4 address. The instance's primary private IPv4 address is the one that will be applied to the target.

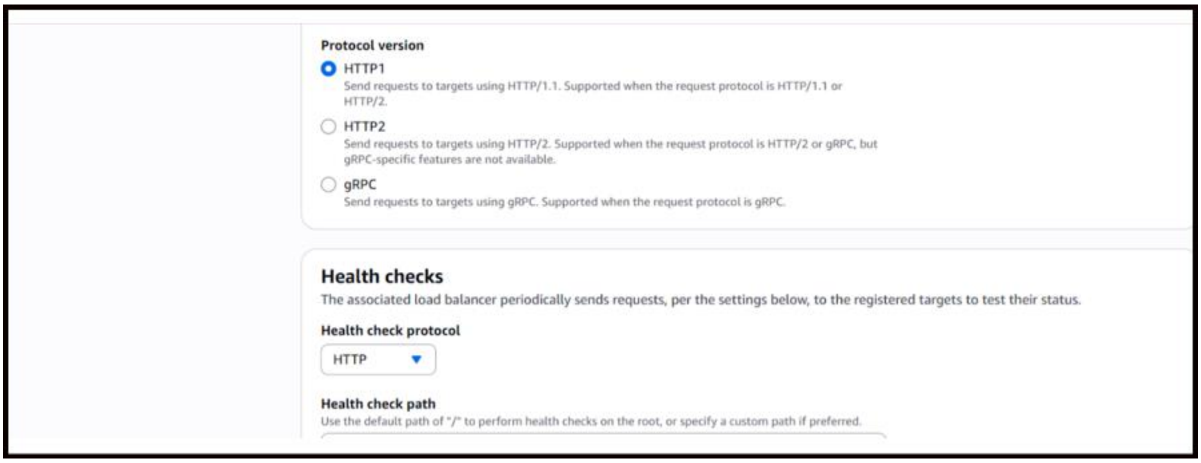
○ IPv6

Each instance you register must have an assigned primary IPv6 address. This is configured on the instance's default network interface (eth0). Learn more

VPC

Select the VPC with the instances that you want to include in the target group. Only VPCs that support the IP address type selected above are available in this list.

aws-prod-example-vpc
vpc-078c725883eaf938e
IPv4 VPC CIDR: 10.0.0.0/16 ▼



Protocol version

● HTTP1
Send requests to targets using HTTP/1.1. Supported when the request protocol is HTTP/1.1 or HTTP/2.

○ HTTP2
Send requests to targets using HTTP/2. Supported when the request protocol is HTTP/2 or gRPC, but gRPC-specific features are not available.

○ gRPC
Send requests to targets using gRPC. Supported when the request protocol is gRPC.

Health checks

The associated load balancer periodically sends requests, per the settings below, to the registered targets to test their status.

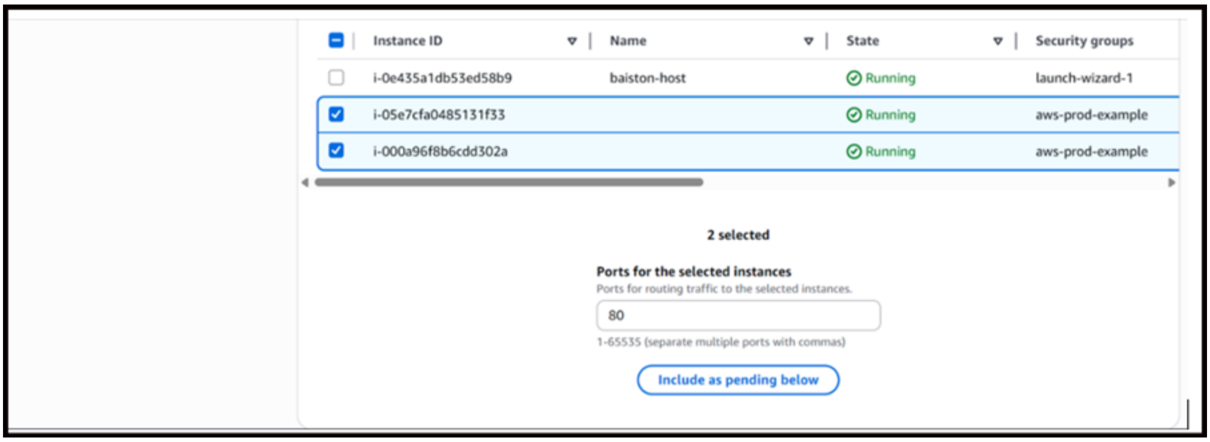
Health check protocol

HTTP

Health check path

Use the default path of "/" to perform health checks on the root, or specify a custom path if preferred.

It will open a instance section and choose both the instance because one has an application and other not we have to check the incoming traffic of private subnets in one server and other one throw an error because it doesn't have the application.



	Instance ID	Name	State	Security groups
□	i-0e435a1db53ed58b9	baiston-host	✓ Running	launch-wizard-1
☑	i-05e7cfa0485131f33		✓ Running	aws-prod-example
☑	i-000a96f8b6cdd302a		✓ Running	aws-prod-example

2 selected

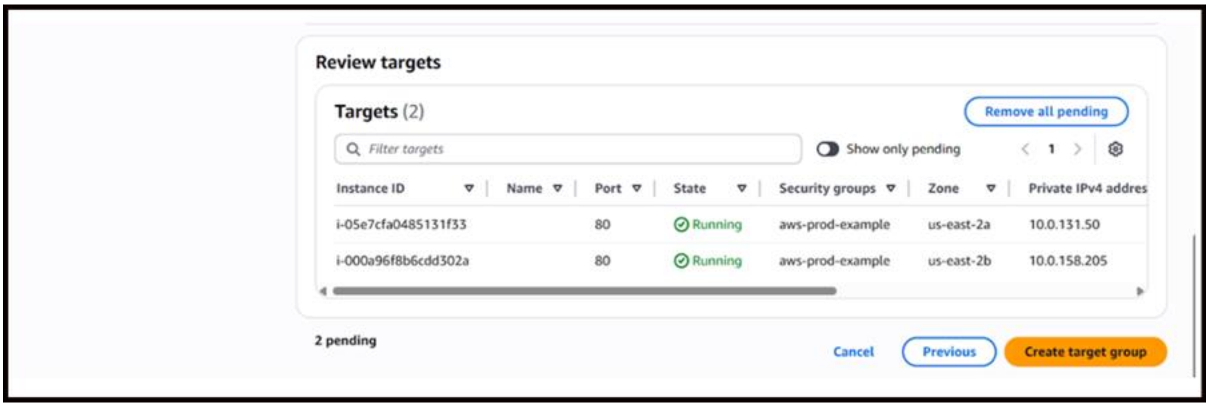
Ports for the selected instances

Ports for routing traffic to the selected instances.

80

1-65535 (separate multiple ports with commas)

Include as pending below, click on that and it shows the target instance where the flow of incoming traffic should go and create target groups.



Review targets

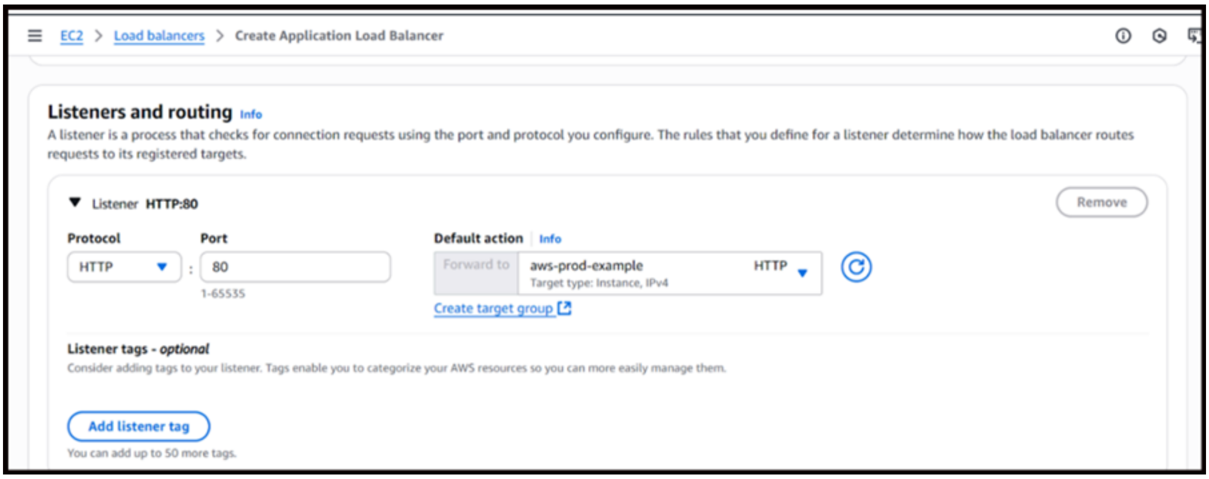
Targets (2)

Instance ID	Name	Port	State	Security groups	Zone	Private IPv4 addres
i-05e7cfa0485131f33		80	✓ Running	aws-prod-example	us-east-2a	10.0.131.50
i-000a96f8b6cdd302a		80	✓ Running	aws-prod-example	us-east-2b	10.0.158.205

2 pending

Run your application on port instead of 8000. Relaunch your application in port 80 likewise you can configure 8000 also.

Now go back to the load balancer and do configure for 80 port and target group as created.



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**

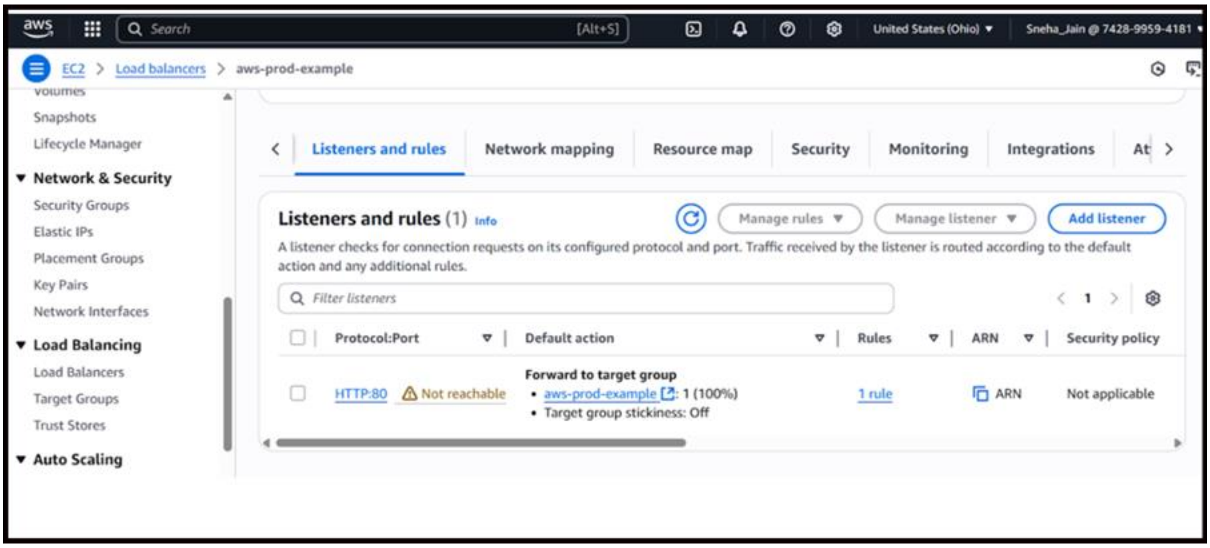
Protocol	Port	Default action Info
HTTP	80 1-65535	Forward to aws-prod-example HTTP Target type: Instance, IPv4 Create target group

Listener tags - *optional*

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

You can add up to 50 more tags.

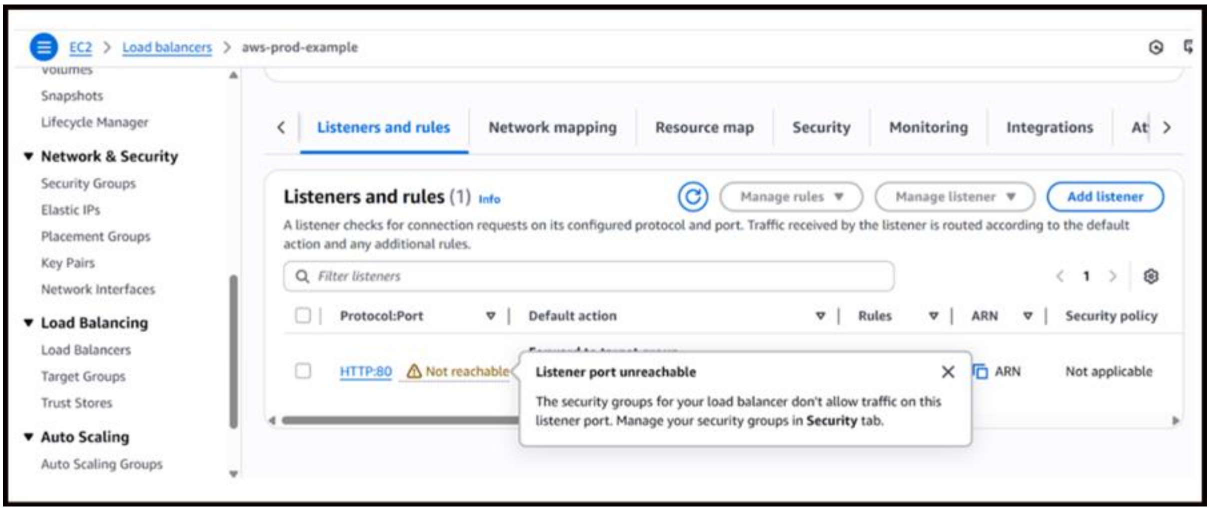
When you try to access the application at port, you won’t allow it because the load balancer is allowing port 80 access.



Listeners and rules (1) Info

A listener checks for connection requests on its configured protocol and port. Traffic received by the listener is routed according to the default action and any additional rules.

□	Protocol:Port	Default action	Rules	ARN	Security policy
□	HTTP:80 ⚠ Not reachable	**Forward to target group** • aws-prod-example: 1 (100%) • Target group stickiness: Off	1 rule	ARN	Not applicable



Listeners and rules (1) Info

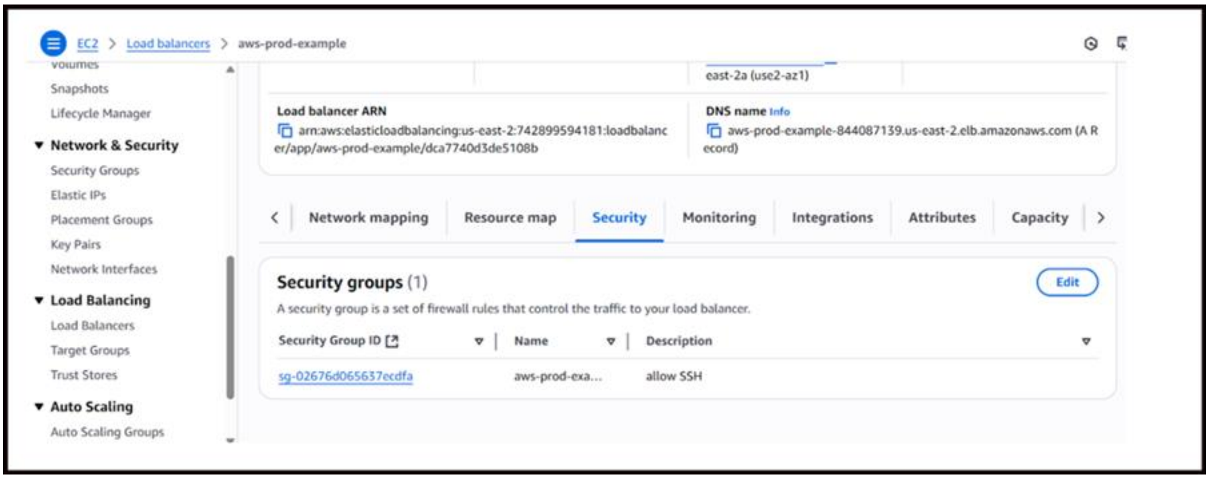
A listener checks for connection requests on its configured protocol and port. Traffic received by the listener is routed according to the default action and any additional rules.

□	Protocol:Port	Default action	Rules	ARN	Security policy
□	HTTP:80 ⚠ Not reachable			ARN	Not applicable

Listener port unreachable

The security groups for your load balancer don't allow traffic on this listener port. Manage your security groups in **Security** tab.

Go to the security group and allow the traffic for port 80.



east-2a (use2-az1)

Load balancer ARN
arn:aws:elasticloadbalancing:us-east-2:742899594181:loadbalancer/app/aws-prod-example/dca7740d3de5108b

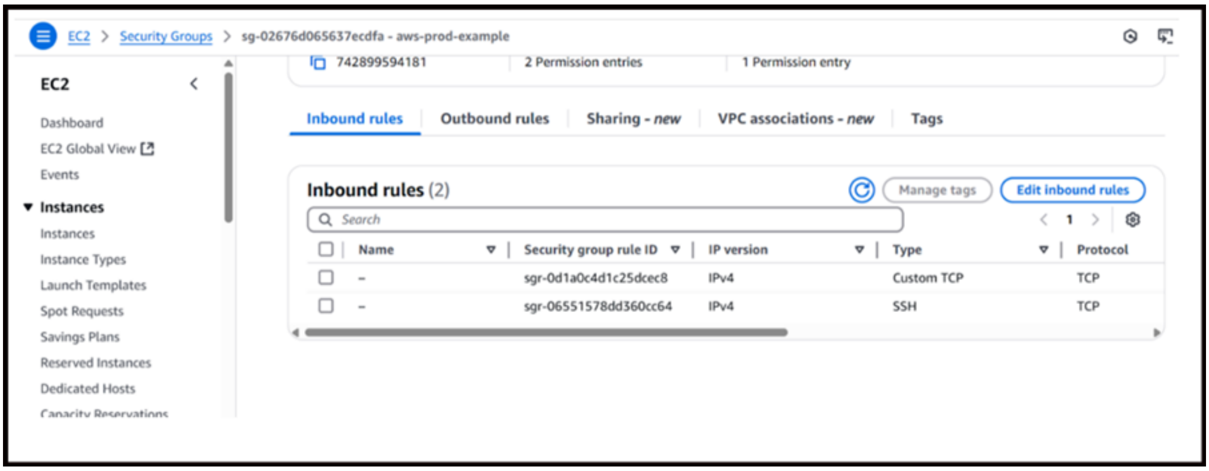
DNS name Info
aws-prod-example-844087139.us-east-2.elb.amazonaws.com (A Record)

Security groups (1)

A security group is a set of firewall rules that control the traffic to your load balancer.

Security Group ID	Name	Description
sg-02676d065637ecdfa	aws-prod-exa...	allow SSH

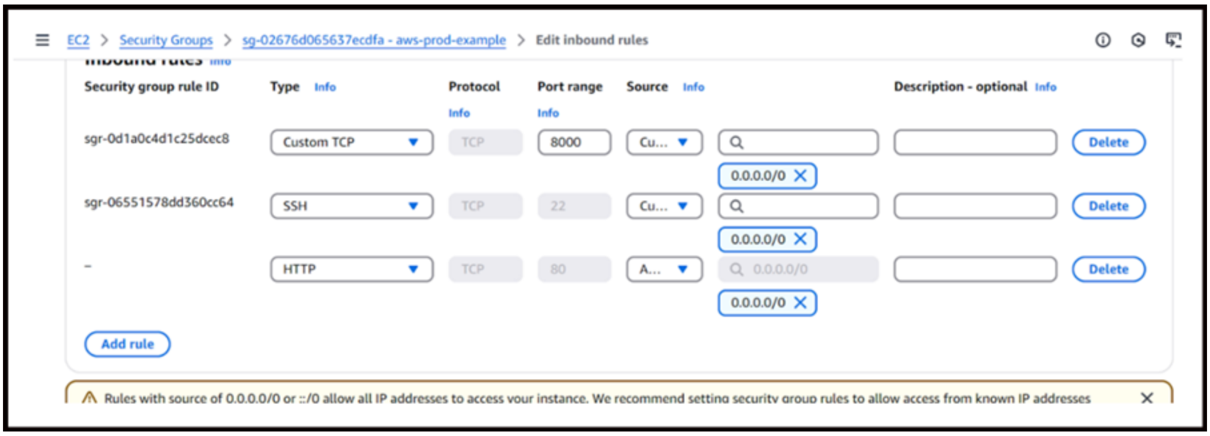
Go to inbound traffic rules and add a port 80 configuration.



742899594181 | 2 Permission entries | 1 Permission entry

Inbound rules (2)

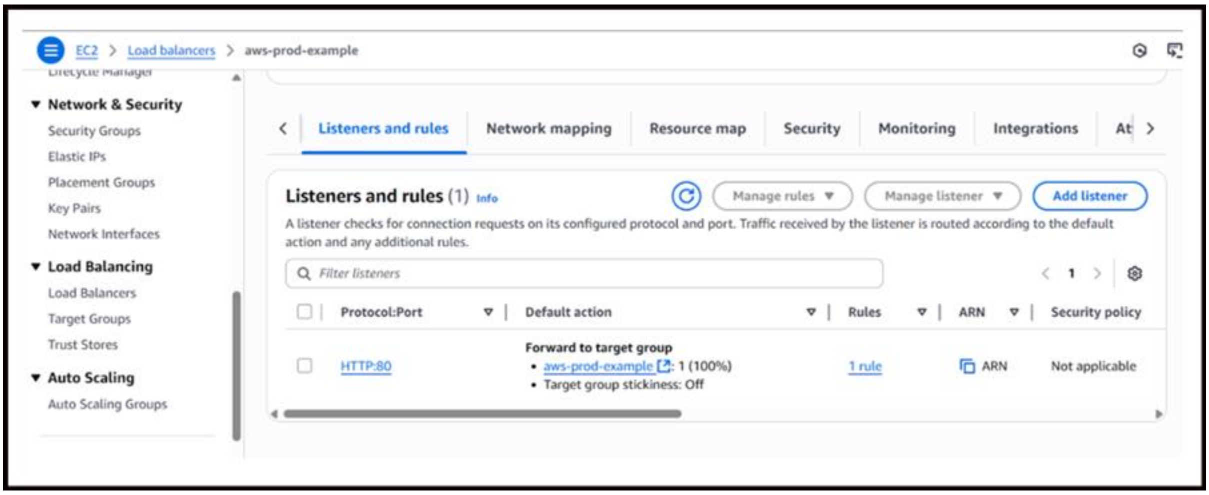
□	Name	Security group rule ID	IP version	Type	Protocol
□	–	sgr-0d1a0c4d1c25dcec8	IPv4	Custom TCP	TCP
□	–	sgr-06551578dd360cc64	IPv4	SSH	TCP


EC2 > Security Groups > sg-02676d065637ecdfa - aws-prod-example > Edit inbound rules
Inbound rules Info
Delete
Add rule

Security group rule ID	Type Info	Protocol Info	Port range Info	Source Info		Description - optional Info
sgr-0d1a0c4d1c25dcec8	Custom TCP	TCP	8000	Cu...	0.0.0.0/0	
sgr-06551578dd360cc64	SSH	TCP	22	Cu...	0.0.0.0/0	
-	HTTP	TCP	80	A...	0.0.0.0/0	

⚠ Rules with source of 0.0.0.0/0 or ::/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses

save this rule. and after adding this 80 port in group the error in load balancer disappeared.

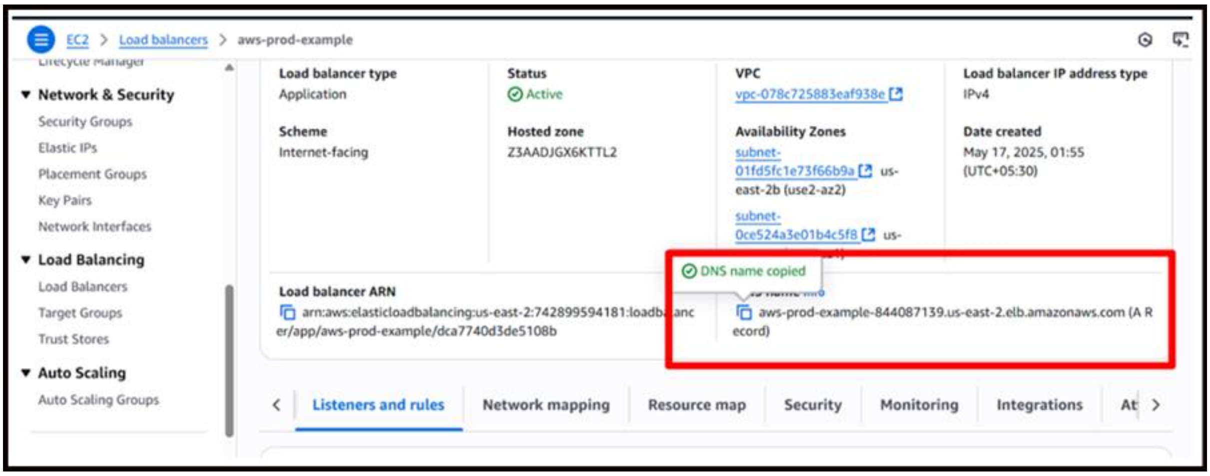


Listeners and rules (1) Info

A listener checks for connection requests on its configured protocol and port. Traffic received by the listener is routed according to the default action and any additional rules.

Protocol:Port	Default action	Rules	ARN	Security policy
HTTP:80	Forward to target group • aws-prod-example: 1 (100%) • Target group stickiness: Off	1 rule	ARN	Not applicable

Now access the DNS for the server.



Load balancer type
Application

Status
Active

VPC
vpc-078c725883eaf938e

Load balancer IP address type
IPv4

Scheme
Internet-facing

Hosted zone
Z3AADJGX6KTTL2

Availability Zones
subnet-01fd5fc1e73f66b9a us-east-2b (use2-az2)
subnet-0ce524a3e01b4c5f8 us-

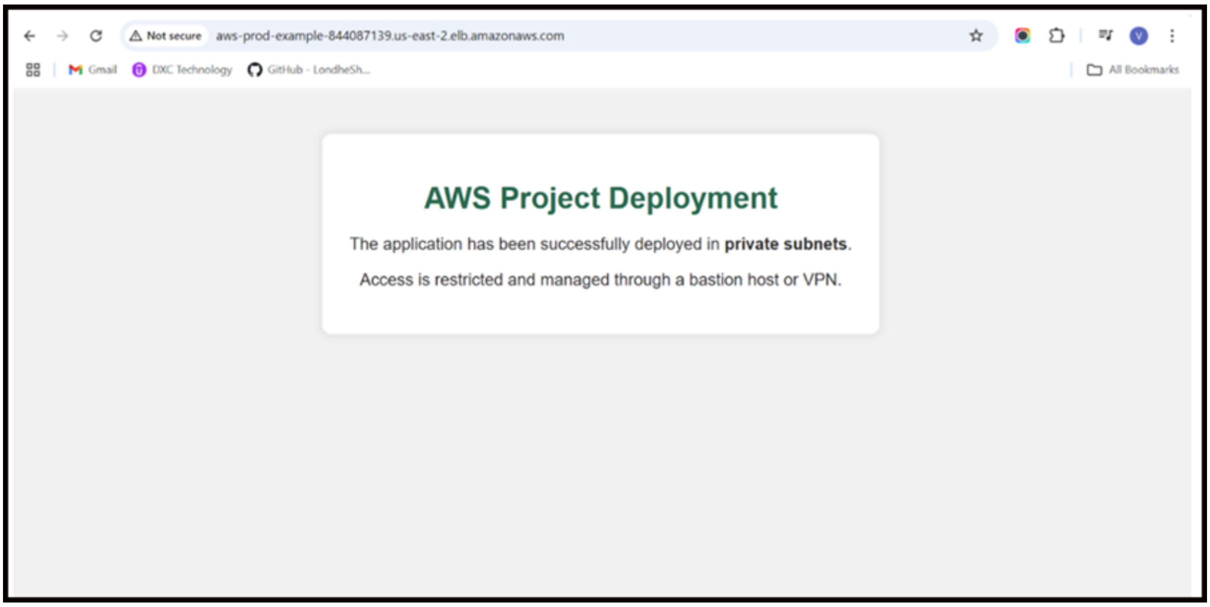
Date created
May 17, 2025, 01:55 (UTC+05:30)

Load balancer ARN
arn:aws:elasticloadbalancing:us-east-2:742899594181:loadbalancer/app/aws-prod-example/dca7740d3de5108b

DNS name copied

aws-prod-example-844087139.us-east-2.elb.amazonaws.com (A Record)

Your application in the private subnet server has deployed now.



AWS Project Deployment

The application has been successfully deployed in **private subnets**.

Access is restricted and managed through a bastion host or VPN.

Enjoy:)