

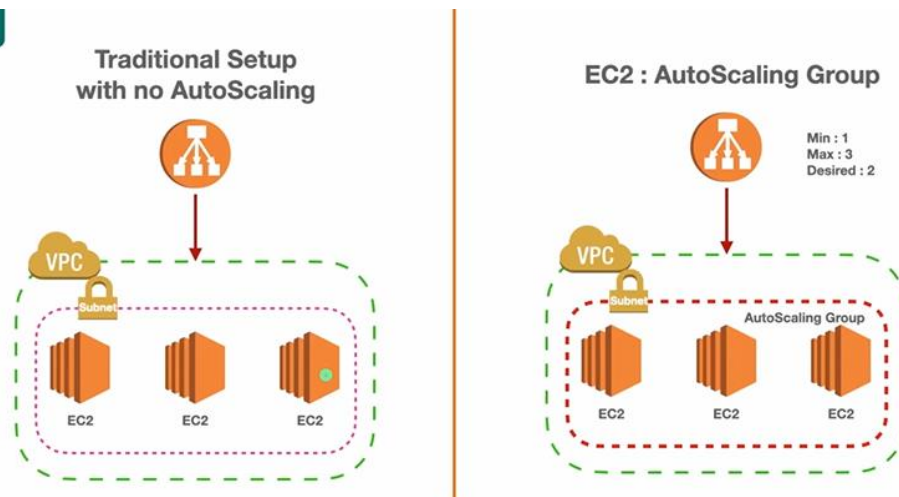
EC2 AUTO SCALING

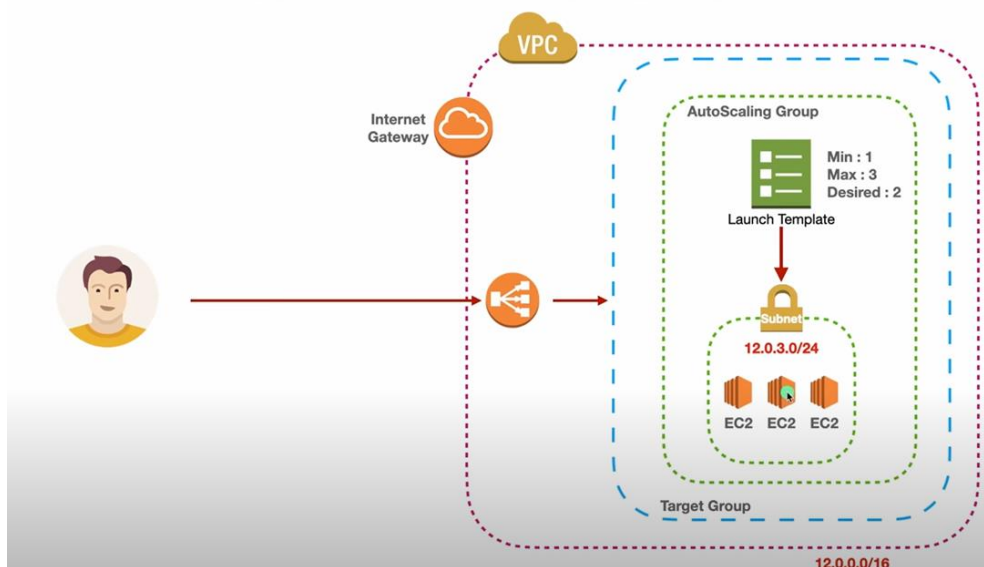
INTRODUCTION:

Implemented EC2 Auto Scaling to dynamically scale instances as per application needs.

What is Amazon EC2 Auto Scaling?

Amazon EC2 Auto Scaling helps you ensure that you have the correct number of Amazon EC2 instances available to handle the load for your application. You create collections of EC2 instances, called *Auto Scaling groups*. You can specify the minimum number of instances in each Auto Scaling group, and Amazon EC2 Auto Scaling ensures that your group never goes below this size. You can specify the maximum number of instances in each Auto Scaling group, and Amazon EC2 Auto Scaling ensures that your group never goes above this size. If you specify the desired capacity, either when you create the group or at any time thereafter, Amazon EC2 Auto Scaling ensures that your group has this many instance. If you specify scaling policies, then Amazon EC2 Auto Scaling can launch or terminate instances as demand on your application increases or decreases.

0

ARCHITECTURE:**SERVICES USED:**

- **Application Load Balancer** - An Application Load Balancer functions at the application layer, the seventh layer of the Open Systems Interconnection (OSI) model. After the load balancer receives a request, it evaluates the listener rules in priority order to determine which rule to apply, and then selects a target from the target group for the rule action. You can configure listener rules to route requests to different target groups based on the content of the application traffic. Routing is performed independently for each target group, even when a target is registered with multiple target groups.
- **VPC (Virtual Private Cloud)** - which is a virtual network infrastructure provided by AWS. It enables you to create a logically isolated section of the AWS cloud where you can launch resources such as EC2 instances, RDS databases, and more.
- **Internet Gateway (IGW)** - is a horizontally scaled, redundant, and highly available component in AWS that allows communication between resources in a VPC and the internet. An Internet Gateway is a virtual router that connects a VPC to the internet. It provides a target for traffic destined for the public internet from instances in the VPC and a source for traffic originating from the internet and intended for instances in the VPC.

IMPLEMENTATION:**STEP1) Create VPC**

By default, when you create a new AWS account, a default VPC (Virtual Private Cloud) is created for you in each AWS region. The default VPC comes preconfigured with several default settings, including an Internet Gateway and a default subnet in each Availability Zone within the region. This means that you can launch your resources in the default VPC without having to worry about configuring networking settings.

-Here, we will create custom VPC

- Click on “Create VPC” ->select VPC only -> IPV4 CIDR block (12.0.0.0/16)
- name- “test-vpc”

Create VPC [Info](#)

A VPC is an isolated portion of the AWS Cloud populated by AWS objects, such as Amazon EC2 instances.

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 - optional [Info](#)

Creates a tag with a key of 'Name' and a value that you specify.

test-vpc

IPv4 CIDR block [Info](#)

☒ IPv4 CIDR manual input
 ☐ IPAM-allocated IPv4 CIDR block

IPv4 CIDR

12.0.0.0/16

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

IPv6 CIDR block [Info](#)

☒ No IPv6 CIDR block
 ☐ IPAM-allocated IPv6 CIDR block
 ☐ Amazon-provided IPv6 CIDR block
 ☐ IPv6 CIDR owned by me

Tenancy [Info](#)

Default

Tags

A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

Key

Value - optional

You can add 49 more tags

You successfully created vpc-0f90fe1b94d4f174e / test-vpc

Your VPCs (2) [Info](#)

Last updated less than a minute ago

| <input type="checkbox"/> | Name | VPC ID | State | Block Public... | IPv4 CIDR | IPv6 CIDR | DHCP option se |
|--------------------------|----------|---------------------------------------|-----------|---------------------------|---------------|-----------|-------------------------------|
| <input type="checkbox"/> | - | vpc-04deda9b41c70d7be | Available | <input type="radio"/> Off | 172.31.0.0/16 | - | dopt-00a4eef5 |
| <input type="checkbox"/> | test-vpc | vpc-0f90fe1b94d4f174e | Available | <input type="radio"/> Off | 12.0.0.0/16 | - | dopt-00a4eef5 |

- test-vpc created

STEP2) Create an Internet gateway to your Custom VPC

- Name- “**igw-test**” ->click on “Create Internet gateway”

Create internet gateway [Info](#)

An internet gateway is a virtual router that connects a VPC to the internet. To create a new internet gateway specify the name for the gateway below.

Internet gateway settings

Name tag
Creates a tag with a key of 'Name' and a value that you specify.

igw-test

Tags - optional
A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

Key **Value - optional**

igw-test igw-test

[Add new tag](#) [Remove](#)

You can add 49 more tags.

[Cancel](#) [Create internet gateway](#)

igw-0611373a4d0feed56 / igw-test

Details [Info](#)

Internet gateway ID: igw-0611373a4d0feed56 State: Detached VPC ID: - Owner: 897722695580

Tags

[Attach to VPC](#) [Detach from VPC](#) [Manage tags](#) [Delete](#)

- igw-test created but it is not attached to custom VPC
- Attach IGW to VPC by clicking on “Attach to VPC” and select the VPC to attach.

Attach to VPC (igw-0611373a4d0feed56) [Info](#)

VPC
Attach an internet gateway to a VPC to enable the VPC to communicate with the internet. Specify the VPC to attach below.

Available VPCs
Attach the internet gateway to this VPC.

vpc-0f90fe1b94d4f174e

[AWS Command Line Interface command](#)

[Cancel](#) [Attach internet gateway](#)

Internet gateway igw-0611373a4d0feed56 successfully attached to vpc-0f90fe1b94d4f174e

igw-0611373a4d0feed56 / igw-test

Details [Info](#)

Internet gateway ID: igw-0611373a4d0feed56 State: Attached VPC ID: vpc-0f90fe1b94d4f174e | test-vpc Owner: 897722695580

- “igw-test” is now attached to “test-vpc”.

STEP3) Create public subnets

- Click on Create subnet, select your custom VPC, Give name to subnet, select Availability zone according your convince and assign IPV4 CDIR block to this subnet.
- Create 2 public subnets (“**test-public-subnet-1a**” and “**test-public-subnet-1b**”) in different availability zones (“**us-east-1a**” and “**us-east-1b**”) and IPV4 CDIR blocks (**12.0.1.0/24** and **12.0.3.0/24**)

Create subnet Info

VPC
VPC ID
Create subnets in this VPC.
vpc-0f90fe1b94d4f174e (test-vpc) ▼

Associated VPC CIDRs
IPv4 CIDRs
12.0.0.0/16

Subnet settings
Specify the CIDR blocks and Availability Zone for the subnet.

Subnet 1 of 2
Subnet name
Create a tag with a key of 'Name' and a value that you specify.
test-public-subnet-1a
The name can be up to 256 characters long.
Availability Zone Info
Choose the zone in which your subnet will reside, or let Amazon choose one for you.
US East (N. Virginia) / us-east-1a ▼
IPv4 VPC CIDR block Info
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.
12.0.0.0/16 ▼
IPv4 subnet CIDR block
12.0.1.0/24 256 IPs
◀ ▶ ▲ ▼
Tags - optional

| Key | Value - optional |
|--------|-------------------------|
| Q Name | Q test-public-subnet-1a |

Add new tag Remove

You can add 49 more tags.

Subnet 2 of 2
Subnet name
Create a tag with a key of 'Name' and a value that you specify.
test-public-subnet-1b
The name can be up to 256 characters long.
Availability Zone Info
Choose the zone in which your subnet will reside, or let Amazon choose one for you.
US East (N. Virginia) / us-east-1b ▼
IPv4 VPC CIDR block Info
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.
12.0.0.0/16 ▼
IPv4 subnet CIDR block
12.0.3.0/24 256 IPs
◀ ▶ ▲ ▼
Tags - optional

| Key | Value - optional |
|--------|-------------------------|
| Q Name | Q test-public-subnet-1b |

Add new tag Remove

You can add 49 more tags.

Add new subnet

Cancel Create subnet

You have successfully created 2 subnets: subnet-057805f0964b94d89, subnet-0dfe16d8e0d66e191

Subnets (2) [Info](#)

Find resources by attribute or tag

Subnet ID: subnet-057805f0964b94d89 Subnet ID: subnet-0dfe16d8e0d66e191 [Clear filters](#)

| <input type="checkbox"/> | Name | Subnet ID | State | VPC | Block Public... | IPv4 CIDR | IPv6 CIDR | IPv6 C |
|--------------------------|-----------------------|--------------------------|-----------|----------------------------------|-----------------|-------------|-----------|--------|
| <input type="checkbox"/> | test-public-subnet-1b | subnet-0dfe16d8e0d66e191 | Available | vpc-0f90fe1b94d4f174e test-... | Off | 12.0.3.0/24 | - | - |
| <input type="checkbox"/> | test-public-subnet-1a | subnet-057805f0964b94d89 | Available | vpc-0f90fe1b94d4f174e test-... | Off | 12.0.1.0/24 | - | - |

- subnets created for custom vpc

STEP4) Create Route table (“ret-test-public”) for public subnet.

VPC > Route tables > Create route table

Create route table [Info](#)

A route table specifies how packets are forwarded between the subnets within your VPC, the internet, and your VPN connection.

Route table settings

Name - optional
Create a tag with a key of 'Name' and a value that you specify.
ret-test-public

VPC
The VPC to use for this route table.
vpc-0f90fe1b94d4f174e (test-vpc)

Tags
A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

Key
Name

Value - optional
ret-test-public

[Add new tag](#)
You can add 49 more tags.

[Cancel](#) [Create route table](#)

- Associate “ret-test-public” with public subnet-> Click on “subnet association” ->“Edit subnet association”-> click on public subnet.

VPC > Route tables > rtb-081f4ed067fd95279 > Edit subnet associations

Edit subnet associations

Change which subnets are associated with this route table.

Available subnets (2/2)

Filter subnet associations

| <input type="checkbox"/> | Name | Subnet ID | IPv4 CIDR | IPv6 CIDR | Route table ID |
|-------------------------------------|-----------------------|--------------------------|-------------|-----------|------------------------------|
| <input checked="" type="checkbox"/> | test-public-subnet-1b | subnet-0dfe16d8e0d66e191 | 12.0.3.0/24 | - | Main (rtb-0f246645d029256ac) |
| <input checked="" type="checkbox"/> | test-public-subnet-1a | subnet-057805f0964b94d89 | 12.0.1.0/24 | - | Main (rtb-0f246645d029256ac) |

Selected subnets

subnet-0dfe16d8e0d66e191 / test-public-subnet-1b subnet-057805f0964b94d89 / test-public-subnet-1a

[Cancel](#) [Save associations](#)

You have successfully updated subnet associations for rtb-081f4ed067fd95279 / ret-test-public.

rtb-081f4ed067fd95279 / ret-test-public [Actions](#)

Details [Info](#)

Route table ID
rtb-081f4ed067fd95279

VPC
vpc-0f90fe1b94d4f174e | test-vpc

Main
☒ No

Owner ID
897722695580

Explicit subnet associations
2 subnets

Edge associations
-

Subnet associations

Find subnet association

| Name | Subnet ID | IPv4 CIDR | IPv6 CIDR |
|-----------------------|--------------------------|-------------|-----------|
| test-public-subnet-1b | subnet-0dfe16d8e0d66e191 | 12.0.3.0/24 | - |
| test-public-subnet-1a | subnet-057805f0964b94d89 | 12.0.1.0/24 | - |

- For directing traffic within the VPC public route is associated with public subnet and public subnet needs internet access.

- Edit route by associating it with internal gateway.

VPC > Route tables > rtb-081f4ed067fd95279 > Edit routes

Edit routes

| Destination | Target | Status | Propagated |
|-------------|-----------------------|--------|------------|
| 12.0.0.0/16 | local | Active | No |
| 0.0.0.0/0 | Internet Gateway | - | No |
| | igw-0611373a4d0feed5d | | |

[Add route](#) [Cancel](#) [Preview](#) [Save changes](#)

STEP 5) Create Target group

- Target group is responsible for pointing the EC2 instances.
- Target group name ("tg-ec2-apache2")->Protocol (HTTP)

EC2 > Target groups > Create target group

Specify group details

Your load balancer routes requests to the targets in a target group and performs health checks on the 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 microservices 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.
- ☐ **Application Load Balancer**
 - Offers the flexibility for a Network Load Balancer to accept and route TCP requests within a specific VPC.
 - Facilitates using static IP addresses and PrivateLink with an Application Load Balancer.

Target group name

Target group name

tg-ec2-apache2

A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.

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.

test-vpc
vpc-0f90fe1094d6f174e
IPv4 VPC CIDR: 12.0.0.0/16

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.

/

Up to 1024 characters allowed.

► **Advanced health check settings**

Attributes

○ Certain default attributes will be applied to your target group. You can view and edit them after creating the target group.

► **Tags - optional**

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

Cancel Next

- Blank target group is created as instances will be created via Autoscale policy.

EC2 > Target groups > Create target group

Step 1: Specify group details
Step 2: Register targets

Register targets

This is an optional step to create a target group. However, to ensure that your load balancer routes traffic to this target group you must register your targets.

Available instances (0)

Filter instances

| Instance ID | Name | State | Security groups | Zone | Private IPv4 address | Subnet |
|--------------|------|-------|-----------------|------|----------------------|--------|
| No instances | | | | | | |

0 selected

Ports for the selected instances

Ports for routing traffic to the selected instances.

80
1-65535 (separate multiple ports with comma)

Include as pending below

Review targets

Targets (0)

Remove all pending

Successfully created the target group: tg-ec2-apache2. Anomaly detection is automatically applied to all registered targets. Results can be viewed in the Targets tab.

tg-ec2-apache2

Details

Target type: Instance
Protocol: Port: HTTP-80
Protocol version: HTTP1
VPC: vpc-0f90fe1b94d4f174e

IP address type: IPv4
Load balancer: [None associated](#)

0 Total targets
0 Healthy
0 Unhealthy
0 Anomalous

0 Unused
0 Initial
0 Draining

Registered targets (0)

Target groups route requests to individual registered targets using the protocol and port number specified. Health checks are performed on all registered targets according to the target group's health check settings. Anomaly detection is automatically applied to HTTP/HTTPS target groups with at least 3 healthy targets.

Filter targets

Instance ID | Name | Port | Zone | Health status | Health status details | Administrative... | Override details | Launch time

- Target group created.
- As we can see there is no load balancer associated to the target group.

STEP 6) Create load balancer

- Select and create Application load balancer ("*alb-ec2-instance-with-asg*")->internet facing

Compare and select load balancer type

A complete feature-by-feature comparison along with detailed highlights is also available. [Learn more](#)

Load balancer types

Application Load Balancer

Choose an Application Load Balancer when you need a flexible feature set for your applications with HTTP and HTTPS traffic. Operating at the request level, Application Load Balancers provide advanced routing and visibility features targeted at application architectures, including microservices and

Network Load Balancer

Choose a Network Load Balancer when you need ultra-high performance, TLS offloading at scale, centralized certificate deployment, support for UDP, and static IP addresses for your applications. Operating at the connection level, Network Load Balancers are capable of handling millions of requests per

Gateway Load Balancer

Choose a Gateway Load Balancer when you need to deploy and manage a fleet of third-party virtual appliances that support GENEVE. These appliances enable you to improve security, compliance, and policy controls.

[Create](#)

Create Application Load Balancer [Info](#)

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 evaluating 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.

► 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.

alb-ec2-instance-with-asg

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 is publicly resolvable.
- Requires a public subnet.

☐ Internal

- Serves internal traffic.
- Has private IP addresses.
- DNS name is publicly resolvable.
- 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

Includes IPv4 and IPv6 addresses.

☐ Dualstack without public IPv4

Includes a public IPv6 address, and private IPv4 and IPv6 addresses. Compatible with **Internet-facing** load balancers only.

- Create a new security group ("**alb-sg-for-http-request**") in addition to the default security group.

Create security group [Info](#)

A security group acts as a virtual firewall for your instance to control inbound and outbound traffic. To create a new security group, complete the fields below.

Basic details

Security group name [Info](#)

alb-sg-for-http-request

Name cannot be edited after creation.

Description [Info](#)

allow http request

VPC [Info](#)

vpc-0f90fe1b94d4f174e (test-vpc)

Inbound rules [Info](#)

Type [Info](#)

HTTP

Protocol [Info](#)

TCP

Port range [Info](#)

80

Source [Info](#)

Anywhere...

0.0.0.0/0

Description - optional [Info](#)

Add rule

Delete

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 only.

Security group (sg-0539ee19e9a4aaed6 | alb-sg-for-http-request) was created successfully

► Details

Security Groups (3) [Info](#)

Find resources by attribute or tag

| <input type="checkbox"/> | Name | Security group ID | Security group name | VPC ID | Description | Owner |
|--------------------------|------|----------------------|-------------------------|-----------------------|----------------------------|--------------|
| <input type="checkbox"/> | - | sg-075c96376642d5236 | default | vpc-04deda9b41c70d7be | default VPC security group | 897722695580 |
| <input type="checkbox"/> | - | sg-018131d29b9fccc3d | default | vpc-0f90fe1b94d4f174e | default VPC security group | 897722695580 |
| <input type="checkbox"/> | - | sg-0539ee19e9a4aaed6 | alb-sg-for-http-request | vpc-0f90fe1b94d4f174e | allow http request | 897722695580 |

- Select the security groups

EC2 > Load balancers > Create Application Load Balancer

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

Availability Zones

☒ **us-east-1a (use1-az2)**

Subnet

subnet-057805f0964b94d89
IPv4 subnet CIDR: 12.0.1.0/24

test-public-subnet-1a

IPv4 address
Assigned by AWS

☒ **us-east-1b (use1-az4)**

Subnet

subnet-0dfe16d8e0d66e191
IPv4 subnet CIDR: 12.0.3.0/24

test-public-subnet-1b

IPv4 address
Assigned by AWS

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

Q

☒ default
sg-018131d29b9fccc3d VPC: vpc-0f90fe1b94d4f174e

☒ alb-sg-for-http-request
sg-0539ee19e9a4aaed6 VPC: vpc-0f90fe1b94d4f174e

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

- Load balancer needs to point at the target group (“tg-ec2-apache2”)
- A **listener** checks for connection requests from clients, using the protocol(HTTP) and port (80) that you configure.

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 : HTTP

Port : 80
1-65535

Default action [info](#)

Forward to tg-ec2-apache2
Target type: Instance, IPv4

HTTP

[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.

Summary:

Summary

Review and confirm your configurations. [Estimate cost](#)

Basic configuration [Edit](#)

alb-ec2-instance-with-asg

- Internet-facing
- IPv4

Security groups [Edit](#)

- default
sg-018131d29b9fccc3d
- alb-sg-for-http-request
sg-0539ee19e9a4aaed6

Network mapping [Edit](#)

VPC vpc-0f90fe1b94d4f174e

test-vpc

- us-east-1a
subnet-057805f0964b94d89
- us-east-1b
subnet-0dfe16d8e0d66e191

test-public-subnet-1a
test-public-subnet-1b

Listeners and routing [Edit](#)

- HTTP:80 defaults to tg-ec2-apache2

Service integrations [Edit](#)

Amazon CloudFront + AWS Web Application Firewall (WAF): None

AWS WAF: None

AWS Global Accelerator: None

Tags [Edit](#)

None

Load balancer created.

Network settings:

▼ 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

lt-sg-ec2-instance-apache2

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 and http request

VPC

Info

vpc-0f90fe1b94d4f174e (test-vpc)

12.0.0.0/16

▼

↻

Inbound Security Group Rules

▼ Security group rule 1 (TCP, 80, 0.0.0.0/0)

Remove

Type

Info

HTTP

▼

Protocol

Info

TCP

Port range

Info

80

Source type

Info

Anywhere

▼

Source

Info

Q Add CIDR, prefix list or security group

0.0.0.0/0

×

Description - optional

Info

e.g. SSH for admin desktop

▼ Security group rule 2 (TCP, 22, 0.0.0.0/0)

Remove

Type

Info

ssh

▼

Protocol

Info

TCP

Port range

Info

22

Source type

Info

Anywhere

▼

Source

Info

Q Add CIDR, prefix list or security group

0.0.0.0/0

×

Description - optional

Info

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.

×

Add security group rule

▼ Advanced network configuration

Network interface 1

Remove

Device index

Info

0

Network interface

Info

New interface

▼

Existing network interfaces are not recommended when creating a template for auto-scaling.

Description

Info

Subnet

Info

Don't include in launch template

Not applicable for EC2 Auto Scaling

Security groups

Info

New security group

Auto-assign public IP

Info

Enable

▼

Primary IP

Info

Not applicable for EC2 Auto Scaling

Secondary IP

Info

Don't include in launch template

▼

Not applicable for EC2 Auto Scaling

IPv6 IPs

Info

Don't include in launch template

▼

Not applicable for EC2 Auto Scaling

IPv4 Prefixes

Info

Don't include in launch template

▼

The selected instance type does not support IPv4 prefixes.

IPv6 Prefixes

Info

Don't include in launch template

▼

The selected instance type does not support IPv6 prefixes.

Assign Primary IPv6 IP

Info

Don't include in launch template

▼

Delete on termination

Info

Don't include in launch template

▼

Interface type

Info

Don't include in launch template

▼

Network card index

Info

Don't include in launch template

▼

The selected instance type does not support multiple network cards.

ENA Express

Info

Don't include in launch template

▼

The selected instance type does not support ENA Express.

ENA Express UDP

Info

Don't include in launch template

▼

The selected instance type does not support ENA Express.

Idle connection tracking timeout

Info

Enable

Idle connection tracking timeout is only supported on Nitro instances.

Add network interface

▼ Storage (volumes)

Info

EBS Volumes

Hide details

In advanced details under user data, paste the code to install apache server in ec2.

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

```
yes | sudo apt update
```

```
yes | sudo apt install apache2
```

```
echo "<h1>Server Details</h1><p><strong>Hostname:</strong> $(hostname)
```

```
</p><p><strong>IP Address:</strong> $(hostname -I | cut -d " " -f1)</p>" >  
/var/www/html/index.html
```

```
sudo systemctl restart apache2
```

User data - optional [Info](#)

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

[Choose file](#)

```
#!/bin/bash  
yes | sudo apt update  
yes | sudo apt install apache2  
echo "<h1>Server Details</h1><p><strong>Hostname:</strong> $(hostname)</p><p><strong>IP  
Address:</strong> $(hostname -I | cut -d " " -f1)</p>" > /var/www/html/index.html  
sudo systemctl restart apache2]
```

☐ User data has already been base64 encoded

- Launch template create

| Launch Templates (1) Info | | | | | | | | |
|--|----------------------|--------------------------|-----------------|----------------|--------------------------|--------------------------------|---------|----------|
| <input type="text" value="Search"/> Actions Create launch template | | | | | | | | |
| <input type="checkbox"/> | Launch Template ID | Launch Template Name | Default Version | Latest Version | Create Time | Created By | Managed | Operator |
| <input type="checkbox"/> | lt-07d485156ecc7632c | lt-ec2-instances-apache2 | 1 | 1 | 2025-01-21T01:39:22.000Z | arn:aws:iam::897722695580:root | false | - |

- Get back to the auto scaling group page after the launch template is created
- Auto Scaling group name ("**asg-ec2-instance-test-demo**")->select the launch template created.

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 Info

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.

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](#)

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

Description
-

Launch template
lt-ec2-instances-apache2
lt-07d485156ecc7632c

Instance type
t2.micro

AMI ID
ami-0c55b1405c2b1c32c

Security groups

Request Spot Instances

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.

12.0.0.0/16

[Create a VPC](#)

Availability Zones and subnets

Define which Availability Zones and subnets your Auto Scaling group can use in the chosen VPC.

12.0.1.0/24

12.0.3.0/24

[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 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 Info

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 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.

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.

Application Load Balancer: alb-ec2-instance-with-asg

Enable health check

Health checks

Health checks increase availability by replacing unhealthy instances. When you use multiple health checks, all are evaluated, and if at least one fails, instance replacement occurs.

EC2 health checks

[Always enabled](#)

Additional health check types - optional [Info](#)

☒ **Turn on Elastic Load Balancing health checks** Recommended

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

EC2 Auto Scaling will start to detect and act on health checks performed by Elastic Load Balancing. To avoid unexpected terminations, first verify the settings of these health checks in the [Load Balancer console](#).

☐ 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.

seconds

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

- Configure capacity:
->Desired-2, Min-1, Max-3

Configure group size and scaling - optional [Info](#)

Define your group's desired capacity and scaling limits. You can optionally add automatic scaling to adjust the size of your group.

Group size [Info](#)

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.

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 Equal or less than desired capacity

Max 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 and let the scaling policy adjust the desired capacity in proportion to the metric's value.

- Autoscaling group created

EC2 > Load balancers > alb-ec2-instance-with-asg

Dashboard
EC2 Global View
Events

Instances

Instances
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances
Dedicated Hosts
Capacity Reservations

alb-ec2-instance-with-asg

Details

| | | | |
|---|--------------------------------------|--|--|
| Load balancer type Application | Status Active | VPC vpc-0f90fe1b94d4f174e | Load balancer IP address type IPv4 |
| Scheme Internet-facing | Hosted zone Z355XDOTRQ7X7K | Availability Zones subnet-0dfe16d8e0d6e191 us-east-1b (use1-az4) subnet-057805f0964b94d89 us-east-1a (use1-az2) | Date created January 21, 2025, 06:48 (UTC+05:30) |
| Load balancer ARN arn:aws:elasticloadbalancing:us-east-1:897722695580:loadbalancer/app/alb-ec2-instance-with-asg/0631282a1626b064 | | DNS name info alb-ec2-instance-with-asg-442633224.us-east-1.elb.amazonaws.com (A Record) | |

EC2 > Auto Scaling groups

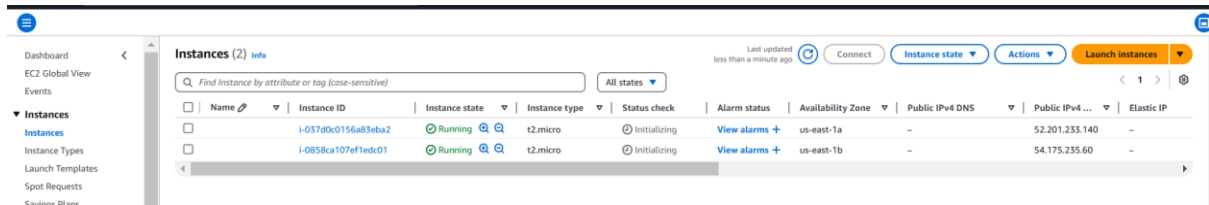
Auto Scaling groups (1) [Info](#)

[Launch configurations](#) [Launch templates](#) [Actions](#) [Create Auto Scaling group](#)

Search your Auto Scaling groups

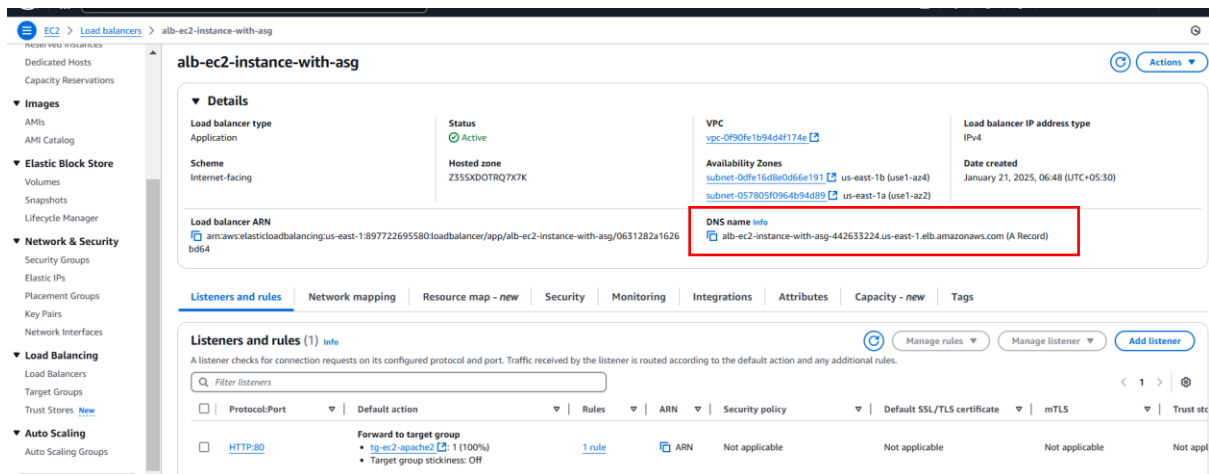
| Name | Launch template/configuration | Instances | Status | Desired capacity | Min | Max | Availability Zones |
|--|--|-----------|----------------------|------------------|-----|-----|------------------------|
| asg-ec2-instance-test-demo | lt-ec2-instances-apache2 Version Default | 0 | Updating capacity... | 2 | 1 | 3 | us-east-1a, us-east-1b |

- Autoscaling group will automatically start creating EC2 instances based on the desired capacity(here, 2).
- Check desired capacity ec2 instance provisioned.



| Name | Instance ID | Instance state | Instance type | Status check | Alarm status | Availability Zone | Public IPv4 DNS | Public IPv4 ... | Elastic IP |
|------|---------------------|----------------|---------------|--------------|---------------|-------------------|-----------------|-----------------|------------|
| | i-037d0c0156a83eba2 | Running | t2.micro | Initializing | View alarms + | us-east-1a | - | 52.201.233.140 | - |
| | i-0858ca107ef1edc01 | Running | t2.micro | Initializing | View alarms + | us-east-1b | - | 54.175.235.60 | - |

- To check, go to load balancer created and copy the dns name and paste it in new tab



alb-ec2-instance-with-asg

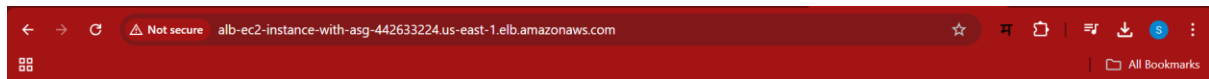
Details

- Load balancer type: Application
- Status: Active
- VPC: vpc-0f90fe1b94d4f174e
- Availability Zones:
 - subnet-0dfe16d8e0d66e191 (us-east-1b (use1-az4))
 - subnet-057805f0964b94d89 (us-east-1a (use1-az2))
- Load balancer IP address type: IPv4
- Date created: January 21, 2025, 06:48 (UTC+05:30)
- Load balancer ARN: arn:aws:elasticloadbalancing:us-east-1:897722695580:loadbalancer/app/alb-ec2-instance-with-asg/0631282a1626bd64
- DNS name info:** alb-ec2-instance-with-asg-442633224.us-east-1.elb.amazonaws.com (A Record)

Listeners and rules (1)

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 | Default SSL/TLS certificate | mTLS | Trust store |
|---------------|---|--------|-----|-----------------|-----------------------------|----------------|----------------|
| HTTP:80 | Forward to target group <ul style="list-style-type: none"> tg-ec2-apache2 (100%) Target group stickiness: Off | 1 rule | ARN | Not applicable | Not applicable | Not applicable | Not applicable |



Server Details

Hostname: ip-12-0-3-162

IP Address: 12.0.3.162

- EC2

Instances

i-037d0c0156a83eba2

Dashboard

EC2 Global View

Events

Instances

Instance Types

Launch Templates

Spot Requests

Savings Plans

Reserved Instances

Dedicated Hosts

Capacity Reservations

Images

AMIs

AMI Catalog

Elastic Block Store

Volumes

Snapshots

Instance summary for i-037d0c0156a83eba2

Updated less than a minute ago

Instance ID

i-037d0c0156a83eba2

IPv6 address

-

Hostname type

IP name: ip-12-0-1-212.ec2.internal

Answer private resource DNS name

-

Auto-assigned IP address

52.01.233.140 [Public IP]

IAM Role

-

IMDSv2

Optional

EC2 recommends setting IMDSv2 to required | Learn more

Public IPv4 address

52.201.233.140 | open address

Instance state

Running

Private IP DNS name (IPv4 only)

ip-12-0-1-212.ec2.internal

Instance type

t2.micro

VPC ID

vpc-0f90fe1b94d4f174e (test-vpc)

Subnet ID

subnet-057805f0964b94d89 (test-public-subnet-1a)

Instance ARN

arn:aws:ec2:us-east-1:897722695580:instance/i-037d0c0156a83eba2

Connect

Instance state

Actions

Private IPv4 addresses

12.0.1.212

Public IPv4 DNS

-

Elastic IP addresses

-

AWS Compute Optimizer finding

Opt-in to AWS Compute Optimizer for recommendations. | Learn more

Auto Scaling Group name

asg-ec2-instance-test-demo

Managed

false



Hostname: ip-12-0-1-212

Hostname: ip-12-0-1-212

IP Address: 12.0.1.212

- The ec2 instance private ip and the server details private ip given is same for both ec2 servers

Check how autoscaling works:

- In Autoscaling both servers are shown in healthy state

asg-ec2-instance-test-demo Capacity overview

amawsautoscaling:us-east-1:897722695580:autoScalingGroup:50c2ef18-d0b5-490e-be04-751b150e84b3:autoScalingGroupName/asg-ec2-instance-test-demo

| Desired capacity | Scaling limits (Min - Max) | Desired capacity type | Status |
|------------------|----------------------------|-----------------------------|--------|
| 2 | 1 - 5 | Units (number of instances) | - |

Date created: Tue Jan 21 2025 07:18:49 GMT+0530 (India Standard Time)

Details | Integrations - new | Automatic scaling | **Instance management** | Instance refresh | Activity | Monitoring

Instances (2)

| Instance ID | Lifecycle | Instance type | Weighted capacity | Launch template... | Availability Zone | Health status | Protected from |
|---------------------|-----------|---------------|-------------------|--------------------------|-------------------|---------------|----------------|
| i-037d0c0156a83eba2 | InService | t2.micro | - | lt-ec2-instances-apache2 | us-east-1a | Healthy | - |
| i-0858ca107ef1edc01 | InService | t2.micro | - | lt-ec2-instances-apache2 | us-east-1b | Healthy | - |

Lifecycle hooks (0) Info

- Now, manually terminate one instance,

Instances (1/2) Info

| Name | Instance ID | Instance state | Instance type | Status check | Alarm status | Availability Zone | Public IPv4 DNS | Public IPv4 ... | Elastic IP |
|---------------------|---------------------|----------------|---------------|-------------------|---------------|-------------------|-----------------|-----------------|------------|
| i-037d0c0156a83eba2 | i-037d0c0156a83eba2 | Running | t2.micro | 2/2 checks passed | View alarms + | us-east-1a | 52.201.233.140 | - | - |
| i-0858ca107ef1edc01 | i-0858ca107ef1edc01 | Running | t2.micro | 2/2 checks passed | View alarms + | us-east-1b | 54.175.235.60 | - | - |

Terminate (delete) instance?

On an EBS-backed instance, the default action is for the root EBS volume to be deleted when the instance is terminated. Storage on any local drives will be lost.

Are you sure you want to terminate these instances?

Instance ID: i-0858ca107ef1edc01 | Termination protection: Disabled

To confirm that you want to delete the instances, choose the terminate button below. Instances with termination protection enabled will not be terminated. Terminating the instance cannot be undone.

Cancel | **Terminate (delete)**

- It has started provisioning new instance

Successfully initiated termination (deletion) of i-0858ca107ef1edc01

Instances (1/3) Info

| Name | Instance ID | Instance state | Instance type | Status check | Alarm status | Availability Zone | Public IPv4 DNS | Public IPv4 ... | Elastic IP |
|---------------------|---------------------|----------------|---------------|-------------------|---------------|-------------------|-----------------|-----------------|------------|
| i-037d0c0156a83eba2 | i-037d0c0156a83eba2 | Running | t2.micro | 2/2 checks passed | View alarms + | us-east-1a | 52.201.233.140 | - | - |
| i-0858ca107ef1edc01 | i-0858ca107ef1edc01 | Terminated | t2.micro | - | View alarms + | us-east-1b | - | - | - |
| i-022da2c5052cb363c | i-022da2c5052cb363c | Running | t2.micro | Initializing | View alarms + | us-east-1b | 54.227.67.184 | - | - |

i-0858ca107ef1edc01

Instance summary

Instance ID: i-0858ca107ef1edc01 | Public IPv4 address: - | Private IPv4 addresses: -

IPv6 address: - | Instance state: Terminated | Public IPv4 DNS: -

Hostname type: IP name: ip-12-0-5-162.ec2.internal | Private IP DNS name (IPv4 only): ip-12-0-5-162.ec2.internal

- We can see the terminated instance showing unhealthy state

EC2 > Auto Scaling groups > asg-ec2-instance-test-demo

asg-ec2-instance-test-demo

asg-ec2-instance-test-demo Capacity overview [Edit](#)

arn:aws:autoscaling:us-east-1:897722695580:autoScalingGroup:50c2ef18-d0b5-490e-be04-751b150e84b3:autoScalingGroupName/asg-ec2-instance-test-demo

| | | | |
|-----------------------|-------------------------------------|--|-------------|
| Desired capacity 2 | Scaling limits (Min - Max) 1 - 3 | Desired capacity type Units (number of instances) | Status - |
|-----------------------|-------------------------------------|--|-------------|

Date created
Tue Jan 21 2025 07:18:49 GMT+0530 (India Standard Time)

Details | Integrations - new | Automatic scaling | **Instance management** | Instance refresh | Activity | Monitoring

Instances (3)

[Filter instances](#)

| <input type="checkbox"/> | Instance ID | Lifecycle | Instance type | Weighted capacity | Launch template... | Availability Zone | Health status | Protected from |
|--------------------------|-------------------------------------|-------------|---------------|-------------------|--|-------------------|---------------|----------------|
| <input type="checkbox"/> | i-022da2c5052cb363c | InService | t2.micro | - | lt-ec2-instances-apache2 | us-east-1b | Healthy | |
| <input type="checkbox"/> | i-037d0c0156a83eba2 | InService | t2.micro | - | lt-ec2-instances-apache2 | us-east-1a | Healthy | |
| <input type="checkbox"/> | i-0858ca107ef1edc01 | Terminating | t2.micro | - | lt-ec2-instances-apache2 | us-east-1b | Unhealthy | |

Lifecycle hooks (0) [Info](#) [Actions](#) [Create lifecycle hook](#)

- Autoscaling group has automatically created EC2 instances based on the desired capacity (here, 2) even after the one instance is terminated.