

Load balancer and Auto Scaling Group

Load Balancer-

Load balancing plays an important security role as computing moves evermore to the cloud. Load Balancer improves application performance by increasing response time and reducing network latency. Also, it distributes the load evenly between servers to improve application performance.

There are four types of Load Balancer :

- a. Application Load Balancer
- b. Network Load Balancer
- c. Classic Load Balancer
- d. Gateway Load Balancer


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

Scheme [Info](#)

Scheme can't be changed after the load balancer is created.

 Internet-facing

2. after this create security group

Basic details

Security group name [Info](#)

Name cannot be edited after creation.

Description [Info](#)

VPC [Info](#)

3. Attach security group which you have create one then create target group. And the it will show 2 servers running. These 2 ervers which I want to behind load balancer.

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 (2)			
<input type="text" value="Filter instances"/>			
<input type="checkbox"/>	Instance ID	Name	State
<input type="checkbox"/>	i-0a0c255fda0936fca	server-1	Running
<input type="checkbox"/>	i-0f733ca8876b7d727	server-2	Running

4. This is how you create load balancer and distribute the traffics.

Load balancers (1)			
Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic			
<input type="text" value="Filter load balancers"/>			
<input type="checkbox"/>	Name	DNS name	State
<input type="checkbox"/>	ALB-loadbalancer	ALB-loadbalancer-744493...	Active

5. In this scenario I have created 2 servers manually myself if traffic is increase I want to have more servers, how can I do that. In this situation we need to **Auto Scaling group**.

Auto Scaling Group

Auto scaling is autometically scale up and scale down of your number of instances. Using AWS Auto Scaling, it's easy to setup application scaling for multiple resources across multiple services in minutes.

6. Give here name to create autoscaling 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

7. Templates which define what should be the settings for new instances created by auto scaling.

[EC2](#) > [Instances](#) > [i-Oa0c255fda0936fca](#) > Create image

Create image [Info](#)

An image (also referred to as an AMI) defines the programs and settings that

Instance ID


 [i-Oa0c255fda0936fca](#) (server-1)

Image name

Maximum 127 characters. Can't be modified after creation.

Image description - *optional*

Maximum 255 characters

Launch template name and description

Launch template name - *required*

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

Template version description

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

8. select security group which you have created

☒ Select existing security group

☐ Create security group

Security groups [Info](#)

Select security groups ▼

ALB_securitygroup sg-040453414855a60dd ✕
VPC: vpc-0a23cefc657885b5

► Advanced network configuration

Configure advanced options - *optional* [Info](#)

Choose a load balancer to distribute incoming traffic for your application across instances to make it more reliable and easily scalable. 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.

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.

Select how many instances you want

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

1

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.

Select target tracking policy Which means you can create more servers or launch servers depending upon specific criteria CPU utilization network in, network out or application load balancer.

Scaling policy name

Target Tracking Policy

Metric type [Info](#)

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

Average CPU utilization ▼

Target value

80

Instance warmup [Info](#)

300 seconds

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

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.

Add notification

Cancel Skip to review Previous **Next**

Then auto scaling group will be created.

EC2 > Auto Scaling groups

Auto Scaling groups (1) [Info](#) [Refresh](#) [Launch configurations](#) [Launch templates](#) [Actions](#) [Create Auto Scaling group](#)

Q Search your Auto Scaling groups < 1 >

<input type="checkbox"/>	Name	Launch template/configuration	Instances	Status	Desired capacity
<input type="checkbox"/>	myASG	mytemplate Version Default	1	-	1