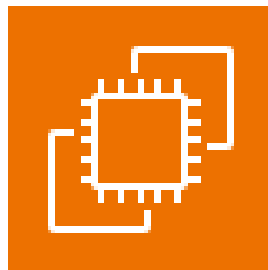
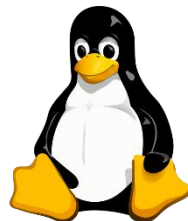




AWS Solution Architect Training with AWS Cloud Practitioner Global Certification Training

Trainer: Aravindraaj.G- Nminds Academy

Configure Network Load Balancer with 3 Linux Web Servers in AWS



Deepakraj R

B.E -Computer Science Engineering

Mepco Schlenk Engineering College

Sivakasi



Objective

An AWS Network Load Balancer (NLB) is a highly scalable and high-performance Layer 4 (TCP/UDP) load balancing service provided by Amazon Web Services (AWS). It is designed to handle millions of requests per second while maintaining ultra-low latencies, making it ideal for applications that require high availability and fault tolerance at the transport layer.

Key Features of AWS Network Load Balancer:

1. **Layer 4 Load Balancing:** NLB operates at the Transport Layer (Layer 4) of the OSI model, meaning it forwards network packets based on IP protocol (TCP, UDP, or TLS), and it can efficiently balance traffic for any application protocol that operates at this layer.
2. **High Availability and Scalability:** NLB automatically scales to handle increasing traffic volumes. It distributes incoming traffic across multiple targets (EC2 instances, containers, IP addresses) within one or more availability zones.
3. **Static IP Support:** NLB provides a single static IP address for each availability zone, which simplifies DNS management. You can also associate an Elastic IP (EIP) with the NLB to maintain a static, publicly reachable IP address.
4. **TLS Termination:** NLB supports TLS termination (encrypted traffic), allowing secure connections from clients to be offloaded at the load balancer. This reduces the burden on backend servers.
5. **Health Checks:** NLB performs health checks on targets to ensure traffic is only forwarded to healthy instances. If a target becomes unhealthy, traffic is redirected to healthy targets.
6. **High Throughput:** Designed to handle millions of requests per second, NLB is optimized for performance and can be used for real-time applications, such as gaming, IoT, and financial applications that require very low latencies.
7. **Connection-based Load Balancing:** NLB uses the client's IP address and port to route traffic, maintaining the source IP, which can be beneficial in some use cases like logging and monitoring.
8. **Cross-Zone Load Balancing:** By default, cross-zone load balancing is enabled, which means traffic is distributed evenly across all available targets in all availability zones. You can disable it if you want to control traffic distribution more finely.



NLB Architecture:

1. **Listeners:** A listener is a process that checks for connection requests. When a request arrives at the NLB, it listens on a specific port (e.g., port 80 for HTTP, port 443 for HTTPS) and forwards the request to an appropriate target group.
2. **Target Groups:** A target group is a set of EC2 instances, containers, or IP addresses that will receive the traffic forwarded by the NLB. You can configure health checks for each target group to monitor the health of your resources.
3. **Target Registration:** Targets (e.g., EC2 instances) need to be registered with the NLB. These targets can be in any available zone in the region, and the NLB will distribute traffic based on their health and availability.

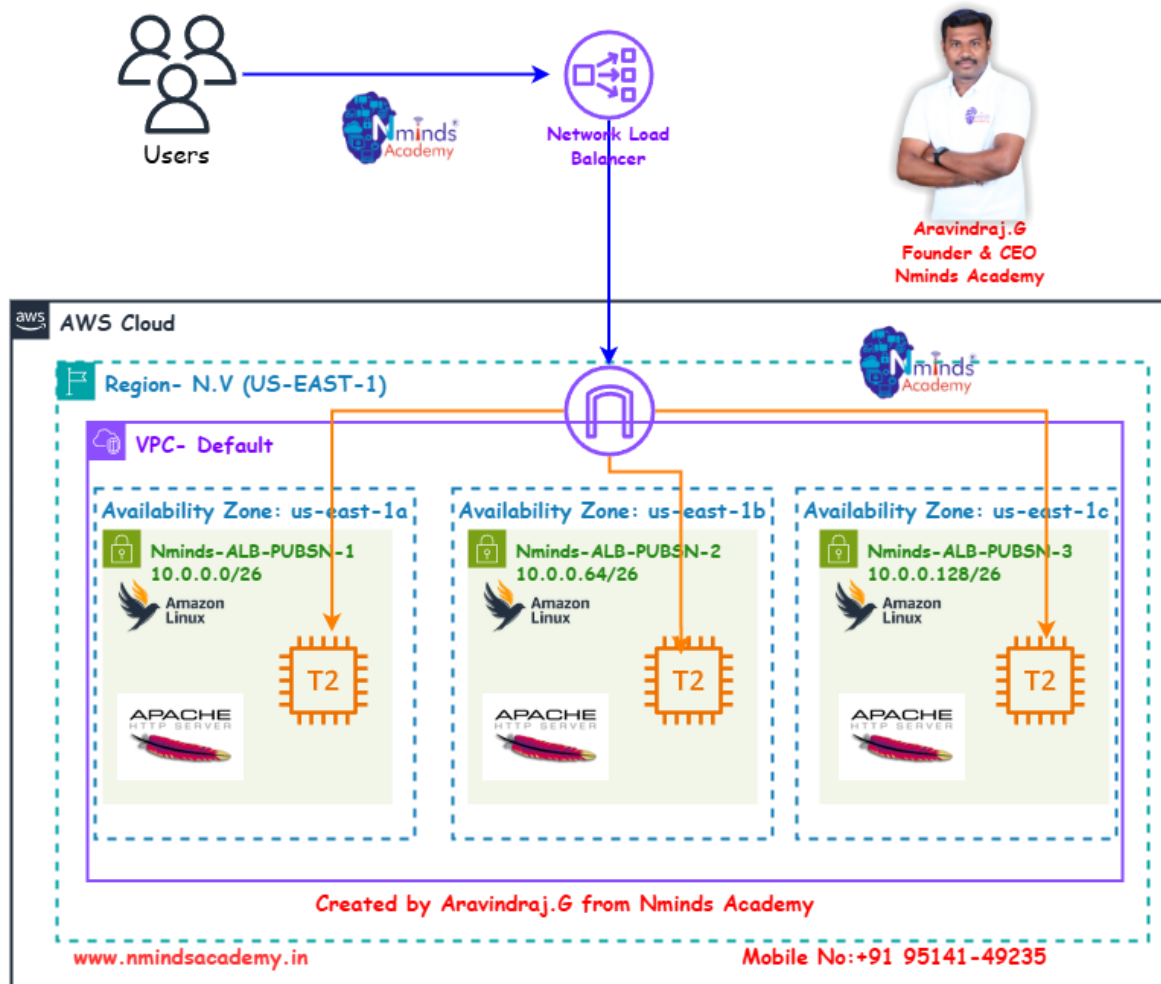
Benefits:

- **Fast performance:** Designed for low latency and high throughput.
- **Highly scalable:** Can handle millions of requests per second.
- **Fault tolerance:** Automatically redirects traffic to healthy targets.
- **Support for IP address-based routing:** Retains the original IP of the client request.



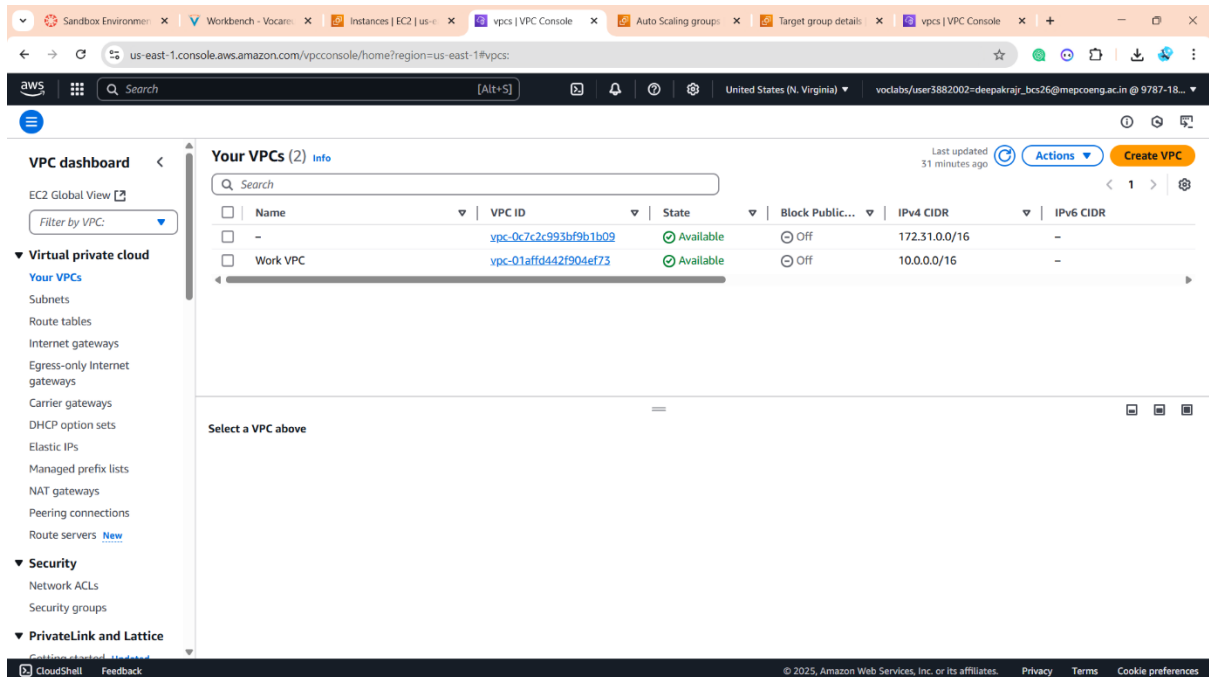
Topology

Configure High Availability with 3 Webservers using Network Load Balancer in AWS



Execution Tasks:

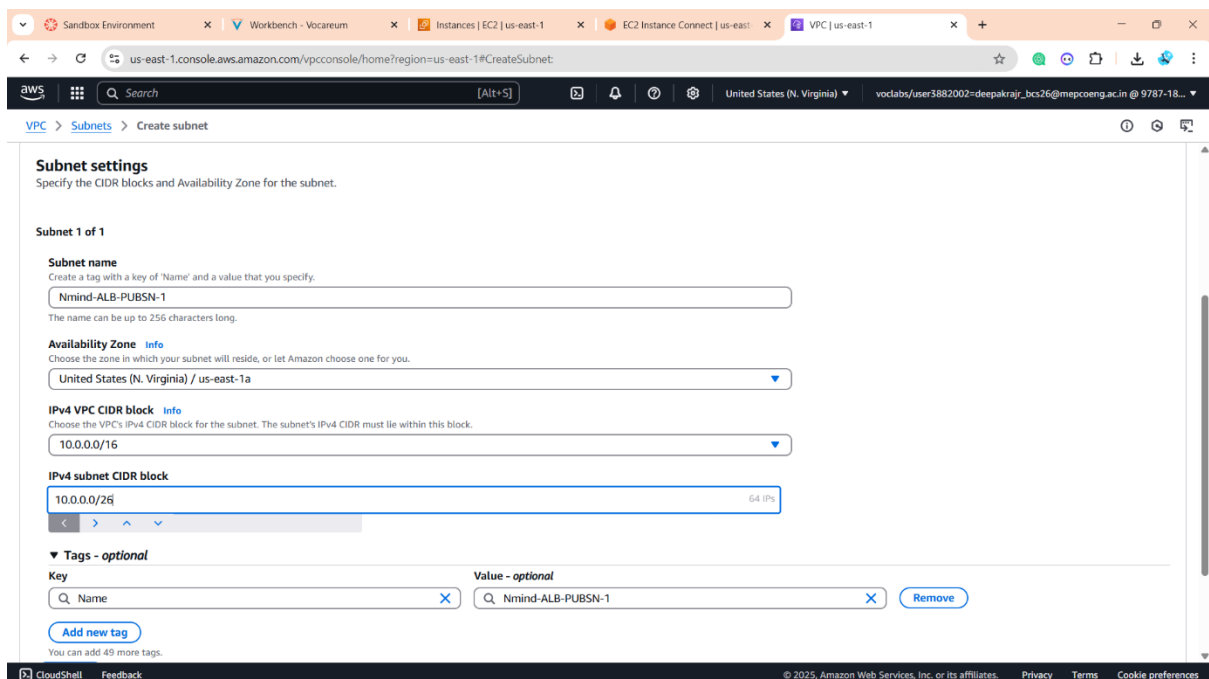
Step1: Set Up the VPC and Subnets



The screenshot shows the AWS VPC console dashboard. The left sidebar contains navigation links for VPC dashboard, EC2 Global View, and various VPC resources. The main area displays 'Your VPCs (2)' with a table listing two VPCs:

Name	VPC ID	State	Block Public...	IPv4 CIDR	IPv6 CIDR
-	vpc-0c7c2c993bf9b1b09	Available	Off	172.31.0.0/16	-
Work VPC	vpc-01affd442f904ef73	Available	Off	10.0.0.0/16	-

Step2: Go to Subnets, create the public subnets within the default VPC



The screenshot shows the 'Create subnet' page in the AWS VPC console. The page is titled 'Subnet settings' and includes the following configuration details:

- Subnet name:** Nminds-ALB-PUBSN-1
- Availability Zone:** United States (N. Virginia) / us-east-1a
- IPv4 VPC CIDR block:** 10.0.0.0/16
- IPv4 subnet CIDR block:** 10.0.0.0/26 (64 IPs)
- Tags - optional:** Key: Name, Value: Nminds-ALB-PUBSN-1



VPC dashboard

You have successfully created 1 subnet: subnet-04aab8f5cb08bf2d3

Subnets (9)

Name	Subnet ID	State	VPC	Block Public...	IPv4 CIDR
-	subnet-016d49178fd5afbf	Available	vpc-0c7c2c993bf9b1b09	Off	172.31.48.0/
-	subnet-0eaf91e3010a1f01c	Available	vpc-0c7c2c993bf9b1b09	Off	172.31.80.0/
-	subnet-06aa8c1526ef0fbc6	Available	vpc-0c7c2c993bf9b1b09	Off	172.31.32.0/
-	subnet-045383b44fca5431a	Available	vpc-0c7c2c993bf9b1b09	Off	172.31.64.0/
Nimind-ALB-PUBSN-2	subnet-0088e5eb1cdc379d7	Available	vpc-01affd442f904ef73 Work ...	Off	10.0.64.0/26
Nimind-ALB-PUBSN-1	subnet-04aab8f5cb08bf2d3	Available	vpc-01affd442f904ef73 Work ...	Off	10.0.0.0/26
Nimind-ALB-PUBSN-3	subnet-027659d6be7958d7c	Available	vpc-01affd442f904ef73 Work ...	Off	10.0.128.0/2

Step3: enable Auto-assign public IPv4 address.

Edit subnet settings

Subnet
Subnet ID: subnet-0088e5eb1cdc379d7
Name: Nimind-ALB-PUBSN-2

Auto-assign IP settings
Enable AWS to automatically assign a public IPv4 or IPv6 address to a new primary network interface for an instance in this subnet.

☒ Enable auto-assign public IPv4 address

☐ Enable auto-assign customer-owned IPv4 address

Resource-based name (RBN) settings
Specify the hostname type for EC2 instances in this subnet and optional RBN DNS query settings.

☐ Enable resource name DNS A record on launch

☐ Enable resource name DNS AAAA record on launch

Hostname type
☐ Resource name
☒ IP name

DNS64 settings
Enable DNS64 to allow IPv6-only services in Amazon VPC to communicate with IPv4-only services and networks.

Step4: Go to Route Tables, select the main route table associated with the default VPC, and ensure a route exists for 0.0.0.0/0 with a target of the Internet Gateway



us-east-1.console.aws.amazon.com/vpcconsole/home?region=us-east-1#EditRoutes:RouteTableId=rtb-0827bfe9282b8dcfe

VPC > Route tables > rtb-0827bfe9282b8dcfe > Edit routes

Edit routes

Destination	Target	Status	Propagated
10.0.0.0/16	local	Active	No
0.0.0.0/0	Internet Gateway	-	No
	igw-0bdf4ebc906d0e142		

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

Step5: Navigate to Launch Templates and Click Create Launch Template

us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateTemplate:

EC2 > Launch templates > Create launch template

Instance type

Instance type

t2.micro

Family: t2 1 vCPU 1 GiB Memory Current generation: true

On-Demand Windows base pricing: 0.0162 USD per Hour

On-Demand Ubuntu Pro base pricing: 0.0134 USD per Hour

On-Demand SUSE base pricing: 0.0116 USD per Hour

On-Demand Linux base pricing: 0.0116 USD per Hour

Free tier eligible

[Compare instance types](#)

Additional costs apply for AMIs with pre-installed software

Key pair (login)

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

WebServer-Key-Pair

[Create new key pair](#)

Network settings

Subnet

Don't include in launch template

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

[Create new subnet](#)

Firewall (security groups)

A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.

Summary

Software Image (AMI)

Amazon Linux 2023 AMI 2023.7.2...read more

ami-00a929b66ed6e0de6

Virtual server type (instance type)

t2.micro

Firewall (security group)

-

Storage (volumes)

1 volume(s) - 8 GiB

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

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

Step6: Configure Template



Inbound Security Group Rules

▼ Security group rule 1 (TCP; 22; 42.104.211.153/32) [Remove](#)

Type: [Info](#) ssh Protocol: [Info](#) TCP Port range: [Info](#) 22

Source type: [Info](#) My IP Name: [Info](#) 42.104.211.153/32 Description - optional: [Info](#) e.g. SSH for admin desktop

▼ Security group rule 2 (TCP; 80; 0.0.0.0/0) [Remove](#)

Type: [Info](#) HTTP Protocol: [Info](#) TCP Port range: [Info](#) 80

Source type: [Info](#) Custom Source: [Info](#) 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**

▼ **Summary**

Software Image (AMI)
Amazon Linux 2023 AMI 2023.7.2...[read more](#)
ami-00a929b66ed6e0de6

Virtual server type (instance type)
t2.micro

Firewall (security group)
New security group

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

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

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

✔ **Success**
Successfully created Nmind-WebServer-Template(tt-09fb281e2423ceb).

► **Actions log**

Next Steps

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

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

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

[View launch templates](#)

Step7: Create a Network Load Balancer and select Network Load Balancer.



Sandbox Environ... Workbench - Vocare... Create network load... RouteTables | VPC C... Auto Scaling group... Target group detail... vpcs | VPC Console

us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateNLBWizard:

Network Load Balancer now supports UDP for Dualstack
Set your IP address type as dualstack and enable prefix for IPv6 source NAT. Then configure UDP-based listeners to route to IPv6 targets.

► How Network 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.
Network-Load-Balancer
A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.

Scheme
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.

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.

☒ **IPv4**
Includes only IPv4 addresses.

☐ **Dualstack**
Includes IPv4 and IPv6 addresses.

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Sandbox Eni... Workbench... Load balanc... Target group... Auto Scaling... RouteTables... Auto Scaling... Target group... vpcs | VPC C...

us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#LoadBalancers:

EC2 > Load balancers

Load balancers (2)

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

Filter load balancers

<input type="checkbox"/>	Name	DNS name	State	VPC ID	Availability Zones	Type	Date create
<input type="checkbox"/>	Network-Load-Balancer	Network-Load-Balancer-d9...	Active	vpc-01affd442f904ef73	3 Availability Zones	network	April 17, 20...
<input type="checkbox"/>	Application-Load-Balan...	Application-Load-Balancer-...	Active	vpc-01affd442f904ef73	3 Availability Zones	application	April 17, 20...

0 load balancers selected

Select a load balancer above.

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Step8: Create an Auto Scaling Group (ASG)



Screenshot of the AWS Management Console showing the 'Create Auto Scaling group' wizard, Step 2: Choose instance launch options.

The page title is 'Create Auto Scaling group'. The breadcrumb navigation shows 'EC2 > Auto Scaling groups > Create Auto Scaling group'.

Step 2: Choose instance launch options

Choose the VPC network environment that your instances are launched into, and customize the instance types and purchase options.

Instance type requirements

You can keep the same instance attributes or instance type from your launch template, or you can choose to override the launch template by specifying different instance attributes or manually adding instance types.

Launch template
Nmind-WebServer-Template
lt-09fb281e1e2423ceb

Version
Default

Description
Nmind-WebServer-Template

Instance type
t2.micro

Network

For most applications, you can use multiple Availability Zones and let EC2 Auto Scaling balance your instances across the zones. The default VPC and default subnets are suitable for getting started quickly.

VPC
Choose the VPC that defines the virtual network for your Auto Scaling group.
vpc-0c7c2c993bf9b1b09
172.31.0.0/16 Default

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

Step9: Configure ASG

Screenshot of the AWS Management Console showing the 'Create Auto Scaling group' wizard, Step 3: Integrate with other services.

The page title is 'Create Auto Scaling group'. The breadcrumb navigation shows 'EC2 > Auto Scaling groups > Create Auto Scaling group'.

Step 3: Integrate with other services

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.

Select target groups

Nmind-TG-Proj4 | TCP
Network Load Balancer: Network-Load-Balancer

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
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.



us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#CreateAutoScalingGroup:

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

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.

3

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
6
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.

CloudShell Feedback

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#AutoScalingGroups:

EC2 > Auto Scaling groups

WebServer-ASG-PROJ4 created successfully

Auto Scaling groups (2) [Info](#) [Launch configurations](#) [Launch templates](#) [Actions](#) [Create Auto Scaling group](#)

Search your Auto Scaling groups

<input type="checkbox"/>	Name	Launch template/configuration Info	Instances	Status	Desired capacity	Min	Max	Availability Zones
<input type="checkbox"/>	WebServer-ASG	Nminds-WebServer-Template Version Del	4	-	3	1	6	us-east-1a, us-east-1b, us-east-1c
<input type="checkbox"/>	WebServer-ASG-PROJ4	Nminds-WebServer-Template Version Del	3	-	3	1	6	us-east-1a, us-east-1b, us-east-1c

0 Auto Scaling groups selected

CloudShell Feedback

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Step10: Test the Configuration



us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#Instances:

EC2 > Instances

Instances (1/12) Info

Find Instance by attribute or tag (case-sensitive)

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IP address
	i-0b18883fae5c83e24	Running	t2.micro	Initializing	View alarms +	us-east-1a	ec2-3
	i-082ef995c4edb5441	Running	t2.micro	Initializing	View alarms +	us-east-1b	ec2-5
	i-0229e37005ee979c9	Terminated	t2.micro	The instance ID	View alarms +	us-east-1b	-
	i-076f5adac618f5353	Running	t2.micro	Initializing	View alarms +	us-east-1b	ec2-3
	i-0213ff3ce7bae4caa	Terminated	t2.micro	-	View alarms +	us-east-1c	-
	i-0b36d751e754aa910	Terminated	t2.micro	The instance ID	View alarms +	us-east-1c	-

i-00b230736ac3c2e7b

Details | Status and alarms | Monitoring | Security | Networking | Storage | Tags

▼ Instance summary Info

Instance ID i-00b230736ac3c2e7b	Public IPv4 address -	Private IPv4 addresses -
IPv6 address -	Instance state Terminated	Public IPv4 DNS -
Hostname type -		
Answer private resource DNS name	Instance type	Elastic IP addresses

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

