

Monitoring and Alerting with CloudWatch

1.1 Creating an AMI for auto scaling

- create an AMI from the existing Web Server 1 Instance
- In the left navigation pane, locate the Instances section, and choose Instances.
- From the **Actions** dropdown list, choose **Image and templates** > **Create image**, and then configure the following.

Create image info

An image (also referred to as an AMI) defines the programs and settings that are applied when you launch an EC2 instance. You can create an image from the configuration of an existing instance.

Instance ID
i-00f322f216b4b64e4 (Web Server 1)

Image name
web server AMI
Maximum 127 characters. Can't be modified after creation.

Image description - optional
Lab AMI for Web Server
Maximum 255 characters

☒ **Reboot instance**
When selected, Amazon EC2 reboots the instance so that data is at rest when snapshots of the attached volumes are taken. This ensures data consistency.

Instance volumes

Storage type	Device	Snapshot	Size	Volume type	IOPS	Throughput	Delete on termination	Encrypted
EBS	/dev/xv...	Create new snapshot from v...	8	EBS General Purpose SSD - ...	100		<input checked="" type="checkbox"/> Enable	<input type="checkbox"/> Enable

[Add volume](#)

During the image creation process, Amazon EC2 creates a snapshot of each of the above volumes.

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.

☒ **Tag image and snapshots together**
Tag the image and the snapshots with the same tag.

☐ **Tag image and snapshots separately**
Tag the image and the snapshots with different tags.

No tags associated with the resource.

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

[Cancel](#) [Create image](#)

2 Creating a load Balancer

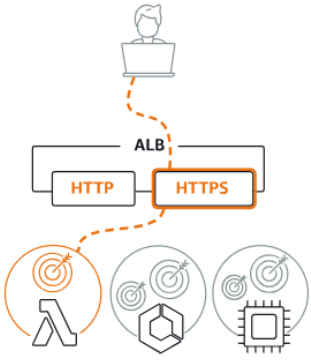
- In the left navigation pane, locate the Load Balancing section, and choose Load Balancers
- In the Load balancer types section, for Application Load Balancer, choose Create.
- Select Application load balancer

Compare and select load balancer type

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

Load balancer types

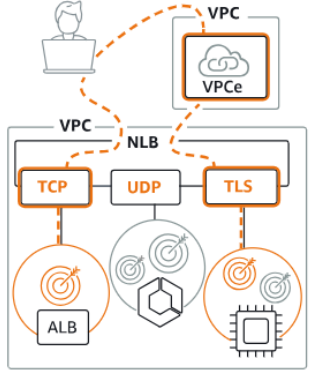
Application Load Balancer [Info](#)



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 containers.

[Create](#)


Network Load Balancer [Info](#)



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 second securely while maintaining ultra-low latencies.

[Create](#)

Gateway Load Balancer [Info](#)



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](#)

► **Classic Load Balancer - previous generation**

[Close](#)

- Map network to the VPC you create with private and public subnets and select two public subnets
- Select at least two Availability Zones and a subnet for each zone. A load balancer node will be placed in each selected zone and will automatically scale in response to traffic. The load balancer routes traffic to targets in the selected Availability Zones only.

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](#).

Lab VPC
vpc-0e4cca92a56b2d3d5
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.

Availability Zones and subnets [Info](#)

Select at least two Availability Zones and a subnet for each zone. A load balancer node will be placed in each selected zone and will automatically scale in response to traffic. The load balancer routes traffic to targets in the selected Availability Zones only.

☒ **us-west-2a (usw2-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-0e44b3809b0e64d02
IPv4 subnet CIDR: 10.0.0.0/24
Public Subnet 1

☒ **us-west-2b (usw2-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-01e2bcae9a475fe1d
IPv4 subnet CIDR: 10.0.2.0/24
Public Subnet 2

- Select security group which you have to create and enable port 80 for http.

- At Listeners and routing section and click on create target group to create a target group.

Listeners and routing [Info](#)

A listener is a process that checks for connection requests using the port and protocol you configure. The rules that you define for a listener determine how the load balancer routes requests to its registered targets.

▼ Listener HTTP:80

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.

[Add listener](#)

Create target group

Step 2
Register targets

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 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.

☒ 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

lab-target-group

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

TCP 80 1-65535

VPC

Select the VPC with the Application Load Balancer that you want to include in the target group.

1 us VPC

- Add instance to the target

Filter instances

< 1 > ⚙

<input checked="" type="checkbox"/>	Instance ID	Name	State	Security groups	Zone
<input checked="" type="checkbox"/>	i-00f322f216b4b64e4	Web Server 1	Running	Web Security Group	us-west-2b

1 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](#)

Review targets

Targets (0)

[Remove all pending](#)

☐ Show only pending < 1 > ⚙

Instance ID	Name	Port	State	Security groups	Zone	Private IPv4 address	Subnet ID	Launch time
-------------	------	------	-------	-----------------	------	----------------------	-----------	-------------

- Add the target group in Listing and Routing

LabELB

Successfully created load balancer: LabELB
It might take a few minutes for your load balancer to fully set up and route traffic. Targets will also take a few minutes to complete the registration process and pass initial health checks.

Application Load Balancers now support public IPv4 IP Address Management (IPAM)
You can get started with this feature by configuring IP pools in the Network mapping section. [Edit IP pools](#)

LabELB [Actions](#)

▼ Details

Load balancer type Application	Status Provisioning	VPC vpc-0e4cca92a56b2d3d5	Load balancer IP address type IPv4
Scheme Internet-facing	Hosted zone Z1H1FL5HABSF5	Availability Zones subnet-0e44b3809b0e64d02 us-west-2a (usw2-az1) subnet-01e2bcae9a475fe1d us-west-2b (usw2-az2)	Date created April 28, 2025, 21:29 (UTC+05:30)
Load balancer ARN arn:aws:elasticloadbalancing:us-west-2:662892777925:loadbalancer/app/LabELB/607ca1df0087208e		DNS name info LabELB-344476822.us-west-2.elb.amazonaws.com (A Record)	

- We have successfully created load Balancer.

3. Creating a launch template

- create a launch template for your Auto Scaling group. A launch template is a template that an Auto Scaling group uses to launch EC2 instances. When you create a launch template, you specify information for the instances, such as the AMI, instance type, key pair, security group, and disks
- In the left navigation pane, locate the **Instances** section, and choose **Launch Templates** and create launch template

Create launch template
Creating a launch template allows you to create a saved instance configuration that can be reused, shared and launched at a later time. Templates can have multiple versions.

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 Scaling

► **Template tags**
► **Source template**

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

▼ **Summary**

Software Image (AMI)
Lab AMI for Web Server
ami-0c1105ef7bb23591f

Virtual server type (instance type)
t3.micro

Firewall (security group)
Web 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](#)

- For Auto Scaling guidance, choose Provide guidance to help me set up a template that I can use with EC2 Auto Scaling.
- In the **Application and OS Images (Amazon Machine Image) - required** section, choose the **My AMIs** tab. Notice that **Web Server AMI** and choose that AMI.
- In the Key pair (login) section, confirm that you create a key pair to log to instances in autoscaling group.

Recents
My AMIs
Quick Start

☒ Owned by me
☐ Shared with me

Browse more AMIs
Including AMIs from AWS, Marketplace and the Community

Amazon Machine Image (AMI)

web server AMI
ami-0c1105ef7bb23591f
2025-04-28T15:39:58.000Z Virtualization: hvm ENA enabled: true Root device type: ebs

Description
Lab AMI for Web Server

Architecture

AMI ID

x86_64

ami-0c1105ef7bb23591f

▼ Instance type Info | Get advice Advanced

Instance type

t3.micro
Family: t3 2 vCPU 1 GiB Memory Current generation: true On-Demand SUSE base pricing: 0.0104 USD per Hour
On-Demand Ubuntu Pro base pricing: 0.0139 USD per Hour On-Demand Windows base pricing: 0.0196 USD per Hour
On-Demand RHEL base pricing: 0.0392 USD per Hour On-Demand Linux base pricing: 0.0104 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

vockey

Create new key pair

- Click on launch template

EC2 > Launch templates > Create launch template

Success
Successfully created lab-app-launch-template(lt-026e052075913c2f6).

▶ 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 commitment 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 a 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 use Spot Instances (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 used for processing, and optional tasks.
Create Spot Fleet

4. Creating an Auto Scaling group

- Choose lab-app-launch-template, and then from Actions dropdown list, choose Create Auto Scaling group

Step 1
Choose launch template or configuration
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 or configuration [Info](#)

Specify a launch template that contains settings common to all EC2 instances that are launched by this Auto Scaling group. If you currently use launch configurations, you might consider migrating to launch templates.

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](#) [Switch to launch configuration](#)

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
 [Create a launch template version](#)

Description
A web server for the load test app

AMI ID
ami-0c1105ef7bb23591f

Launch template
[lab-app-launch-template](#)
lt-026e052075913c2f6

Security groups
-

Instance type
t3.micro

Request Spot Instances
No

- On the Choose instance launch options page, in the Network section, configure the VPC and SUBNETS

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

[Override launch template](#)

Instance type requirements [Info](#)

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	Version	Description
lab-app-launch-template lt-026e052075913c2f6	Default	A web server for the load test app

Instance type
t3.micro

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.

10.0.0.0/16
[Create a VPC](#)

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

[×](#)
10.0.3.0/24

[×](#)
10.0.2.0/24

[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.

- In the Integrate with other services – optional page choose existing load balancer which we have already created before.

Step 1: Choose launch template or configuration
 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.

Select target groups

lab-target-group | HTTP
Application Load Balancer: LabELB

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](#)

- On the Configure group size and scaling policies – optional page
- In the **Group size – optional** section, enter the following values: **Desired capacity:** 2, **Minimum capacity:** 2 **Maximum capacity:** 4
- In the Scaling policies – optional section, configure the following options:
- Choose **Target tracking scaling policy**.
- For **Metric type**, choose **Average CPU utilization**.
- Change the **Target value** to 50
- This change tells Auto Scaling to maintain an average CPU utilization across all instances of 50 percent. Auto Scaling automatically adds or removes capacity as required to keep the metric at or close to the specified target value. It adjusts to fluctuations in the metric due to a fluctuating load pattern.

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 **Max desired capacity**
Equal or less than desired capacity Equal or greater than desired capacity

Automatic scaling - optional
Choose whether to use a target tracking policy [Info](#)
You can set up other metric-based scaling policies and scheduled scaling after creating your Auto Scaling group.

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

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

Scaling policy name

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.

Target value

Instance warmup [Info](#)
 seconds

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

EC2 > Auto Scaling groups > Create Auto Scaling group

Step 1 Choose launch template or configuration
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

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.

▼ Notification 1 [Remove](#)

SNS Topic
Choose an SNS topic to use to send notifications

[Create a topic](#)

Event types
Notify subscribers whenever instances

☒ Launch
☒ Terminate
☒ Fail to launch
☒ Fail to terminate

[Add notification](#)

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

- Create a topic

New Feature
Amazon SNS now supports High Throughput FIFO topics. [Learn more](#)

Create topic

Details

Type [Info](#)

Topic type cannot be modified after topic is created

☐ FIFO (first-in, first-out)

- Strictly-preserved message ordering
- Exactly-once message delivery
- Subscription protocols: SQS

☒ Standard

- Best-effort message ordering
- At-least once message delivery
- Subscription protocols: SQS, Lambda, Data Firehose, HTTP, SMS, email, mobile application endpoints

Name

MyTopic

Maximum 256 characters. Can include alphanumeric characters, hyphens (-) and underscores (_).

Display name - optional [Info](#)

To use this topic with SMS subscriptions, enter a display name. Only the first 10 characters are displayed in an SMS message.

My Topic

Maximum 100 characters.

- Review everything and choose create auto scaling group.

Lab Auto Scaling Group, 1 Scaling policy created successfully

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

Last updated
less than a minute ago

[Launch configurations](#)

[Launch templates](#)

[Actions](#)

Search your Auto Scaling groups

<input type="checkbox"/>	Name	Launch template/configuration	Instances	Status	Desired capacity	Min	Max	Availability Zones
<input type="checkbox"/>	Lab Auto Scaling Group	lab-app-launch-template Version Default	0	Updating capacity...	2	2	4	us-west-2b

- In the left navigation pane, locate the Instances section, and choose Instances.
- You should see two new instances named **Lab Instance**. These instances were launched by auto scaling.

Instances (1/3) [Info](#)

Last updated
1 minute ago

[Connect](#)

[Instance state](#)

[Actions](#)

[Launch instances](#)

Find Instance by attribute or tag (case-sensitive)

All states

<input checked="" type="checkbox"/>	Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IP
<input checked="" type="checkbox"/>	Web Server 1	i-00f322f216b4b64e4	Running	t3.micro	3/3 checks passed	View alarms	us-west-2b	-	34.22.1.1
<input type="checkbox"/>	Lab Instance	i-0292e65c2a73825d4	Running	t3.micro	Initializing	View alarms	us-west-2b	-	-
<input type="checkbox"/>	Lab Instance	i-07994dc6e699aec01	Running	t3.micro	Initializing	View alarms	us-west-2b	-	-

i-00f322f216b4b64e4 (Web Server 1)

- Verifying that load balancing is working

- in the Load Balancing section, choose Target Groups.
- Choose lab-target-group which we have created.
- In the Registered targets section, two Lab Instance targets should be listed for this target group.
- Wait until the Health status of both instances changes to healthy.

The screenshot shows the AWS Management Console interface for the 'lab-target-group' in the 'Load Balancing' section. The left sidebar contains navigation links for various AWS services. The main content area displays the details of the target group, including its ARN, target type (Instance), IP address type (IPv4), protocol (HTTP), and port (80). Below this, a summary bar shows 2 total targets, all of which are 'Healthy'. The 'Registered targets' table lists two targets, both 'Lab Instance' types, with 'Healthy' status. The table columns include Instance ID, Name, Port, Zone, Health status, Health status details, Admin..., Override..., Launch..., and Anomaly detection...

Instance ID	Name	Port	Zone	Health status	Health status details	Admin...	Override...	Launch...	Anomaly detection...
i-0292e65c2a75825d4	Lab Instance	80	us-west-2b (us...)	Healthy	-	No override	No overrid...	April 28, 2...	Normal
i-07994dc6e699a0c01	Lab Instance	80	us-west-2b (us...)	Healthy	-	No override	No overrid...	April 28, 2...	Normal

5. Verifying that load balancing is working

- You should see two new instances named Lab Instance. These instances were launched by auto scaling.

The screenshot shows the AWS Management Console 'Instances' page. It displays a table of EC2 instances. The first instance is 'Lab Instance' with ID 'i-095d6da31d8981032', status 'Running', and type 't3.micro'. The second instance is 'Web Server 1' with ID 'i-052cec3b78b16ba18', status 'Running', and type 't3.micro'. The third instance is 'Lab Instance' with ID 'i-04358426af5092c4f', status 'Running', and type 't3.micro'. The table columns include Name, Instance ID, Instance state, Instance type, Status check, Alarm status, Availability Zone, and Public IPv4.

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4
Lab Instance	i-095d6da31d8981032	Running	t3.micro	Initializing	View alarms +	us-west-2a	-
Web Server 1	i-052cec3b78b16ba18	Running	t3.micro	3/3 checks passed	View alarms +	us-west-2b	-
Lab Instance	i-04358426af5092c4f	Running	t3.micro	Initializing	View alarms +	us-west-2b	-

- Open a new web browser tab, paste the DNS name that you copied before, and press Enter.

← → ↻ ⚠ Not secure labelb-173541881.us-west-2.elb.amazonaws.com

aws Load Test

Meta-Data	Value
InstanceId	i-095d6da31d8981032
Availability Zone	us-west-2a

Current CPU Load: 0%

- If we refresh website we can see change in availability zone

→ ↻ ⚠ Not secure labelb-173541881.us-west-2.elb.amazonaws.com

aws Load Test

Meta-Data	Value
InstanceId	i-04358426af5092c4f
Availability Zone	us-west-2b

Current CPU Load: 0%

6: Testing auto scaling

- In the AWS Management Console, in the search bar, enter and choose CloudWatch
- In the left navigation pane, in the **Alarms** section, choose **All alarms**.
- Two alarms are displayed. The Auto Scaling group automatically created these two alarms. These alarms automatically keep the average CPU load close to 50 percent while also staying within the limitation of having 2–4 instances.
- Choose the alarm that has **AlarmHigh** in its name. This alarm should have a **State** of OK. The *OK* state indicates that the alarm has not been initiated. It is the alarm for **CPU Utilization > 50**, which adds instances when the average CPU utilization is high. The chart should show very low levels of CPU at the moment.
- You should see the **AlarmHigh** chart indicating an increasing CPU percentage. Once it crosses the 50 percent line for more than 3 minutes, it initiates auto scaling to add additional instances.

CloudWatch > Alarms

loudWatch

avorites and recents

dashboards

Operations [Preview](#)

Alarms 1 alarm

logs [New](#)

Metrics

Cloud-Ray traces [New](#)

Events

Application Signals

Network Monitoring

Insights

Alarms (2)

☐ Hide Auto Scaling alarms

Clear selection [Create composite alarm](#) Actions [Create alarm](#)

Search Alarm state: Any Alarm type: Any Actions status: Any < 1 >

<input type="checkbox"/>	Name	State	Last state update (UTC)	Conditions	Actions
<input type="checkbox"/>	TargetTracking-Lab Auto Scaling Group-AlarmLow-4b51397d-dab4-4201-a4e1-be7499455969	In alarm	2025-05-05 11:10:43	CPUUtilization < 35 for 15 datapoints within 15 minutes	Actions enabled
<input type="checkbox"/>	TargetTracking-Lab Auto Scaling Group-AlarmHigh-aefc72b1-dc59-487a-9865-d93b1e3f67a3	OK	2025-05-05 10:57:57	CPUUtilization > 50 for 3 datapoints within 3 minutes	Actions enabled

- More than two instances named **Lab Instance** should now be running. Auto scaling created the new instances in response to the alarm.

Instances (5) [Info](#)

Last updated less than a minute ago [Connect](#) [Instance state](#) [Actions](#) [Launch instances](#)

Find Instance by attribute or tag (case-sensitive) All states < 1 >

<input type="checkbox"/>	Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IP
<input type="checkbox"/>	Lab Instance	i-095d6da31d8981032	Running	t3.micro	3/3 checks passed	View alarms +	us-west-2a	-
<input type="checkbox"/>	Web Server 1	i-052cec3b78b16ba18	Running	t3.micro	3/3 checks passed	View alarms +	us-west-2b	-
<input type="checkbox"/>	Lab Instance	i-09e257af2c8396962	Running	t3.micro	Initializing	View alarms +	us-west-2b	-
<input type="checkbox"/>	Lab Instance	i-04358426af5092c4f	Running	t3.micro	3/3 checks passed	View alarms +	us-west-2b	-
<input type="checkbox"/>	Lab Instance	i-09e61a2d00f49c32b	Running	t3.micro	Initializing	View alarms +	us-west-2a	-

Select an instance