

LAB 3: AUTOSCALING

AIM: To create an autoscaling group.

THEORY:

- 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 instances.
- If you specify scaling policies, then Amazon EC2 Auto Scaling can launch or terminate instances as demand on your application increases or decreases.

PROCEDURE:

1. Create a launch template.

The screenshot shows the 'Create Auto Scaling group' wizard in the AWS Management Console, specifically Step 1: Choose launch template. The breadcrumb navigation at the top reads 'EC2 > Auto Scaling groups > Create Auto Scaling group'. On the left, a sidebar lists the steps: Step 1: Choose launch template (active), Step 2: Choose instance launch options, Step 3 - optional: Configure advanced options, Step 4 - optional: Configure group size and scaling, Step 5 - optional: Add notifications, Step 6 - optional: Add tags, and Step 7: Review. The main content area is titled 'Choose launch template' with an 'Info' icon. Below the title, it says 'Specify a launch template that contains settings common to all EC2 instances that are launched by this Auto Scaling group.' There are two main sections: 'Name' and 'Launch template'. The 'Name' section has a text input field with 'demoASG' and a note: 'Must be unique to this account in the current Region and no more than 255 characters.' The 'Launch template' section has a blue information box stating: '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.' Below this, there is a 'Launch template' section with a dropdown menu labeled 'Select a launch template' and a 'Create a launch template' link. At the bottom right, there are 'Cancel' and 'Next' buttons.

EC2 > Launch templates > 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

demoTemplate

Must be unique to this account. Max 128 chars. No spaces or special characters like %, ", *, "g".

Template version description

asg-template

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.

Application and OS Images (Amazon Machine Image) - required

info

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below

Summary

Software Image (AMI)

Amazon Linux 2023 AMI 2023.4.2...read more

ami-05295b6dc790593e

Virtual server type (instance type)

t2.micro

Firewall (security group)

launch-wizard-2

Storage (volumes)

1 volume(s) - 8 GB

Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 750 hours of public IPv4 address usage per month, 30 GiB of EBS storage, 2 million I/Os, 1 GiB of snapshots, and 100 GB of bandwidth to the internet.

Cancel

Create launch template

EC2 > Launch templates > 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

demoTemplate

Must be unique to this account. Max 128 chars. No spaces or special characters like %, ", *, "g".

Template version description

asg-template

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.

Application and OS Images (Amazon Machine Image) - required

info

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below

Summary

Software Image (AMI)

Amazon Linux 2023 AMI 2023.4.2...read more

ami-05295b6dc790593e

Virtual server type (instance type)

t2.micro

Firewall (security group)

launch-wizard-2

Storage (volumes)

1 volume(s) - 8 GB

Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 750 hours of public IPv4 address usage per month, 30 GiB of EBS storage, 2 million I/Os, 1 GiB of snapshots, and 100 GB of bandwidth to the internet.

Cancel

Create launch template

Application and OS Images (Amazon Machine Image) - required

info

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below

Search our full catalog including 1000s of application and OS images

RecentsQuick Start

Amazon Linux

macOS

Ubuntu

Windows

Red Hat

SUSE L

Browse more AMIs

Including AMIs from AWS Marketplace and the Community

Amazon Machine Image (AMI)

Amazon Linux 2023 AMI

ami-05295b6dc790593e (64-bit x86), self-preferred / ami-0c265d90f3bedf8e (64-bit ARM), self

Virtualization: hvmENA enabled: trueRoot device type: ebs

Free tier eligible

Description

Amazon Linux 2023 AMI 2023.4.20240319.1 x86_64 HVM kernel-6.1

Architecture

64-bit (x86)

Boot mode

uefi-preferred

AMI ID

ami-05295b6dc790593e

Verified provider

Instance type

infoGet advice

Advanced

Instance type

Summary

Software Image (AMI)

Amazon Linux 2023 AMI 2023.4.2...read more

ami-05295b6dc790593e

Virtual server type (instance type)

t2.micro

Firewall (security group)

launch-wizard-2

Storage (volumes)

1 volume(s) - 8 GB

Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 750 hours of public IPv4 address usage per month, 30 GiB of EBS storage, 2 million I/Os, 1 GiB of snapshots, and 100 GB of bandwidth to the internet.

Cancel

Create launch template

A2305221030

10

2. Proceed with launching an autoscaling group.

Choose instance launch options

Step 3 - optional
Configure advanced options

Step 4 - optional
Configure group size and scaling

Step 5 - optional
Add notifications

Step 6 - optional
Add tags

Step 7
Review

Name

Auto Scaling group name
Enter a name to identify the group.

Must be unique to this account in the current Region and no more than 255 characters.

Launch template info

For accounts created after May 31, 2023, the EC2 console only supports creating Auto Scaling groups with launch templates. Creating Auto Scaling groups with launch configurations is not recommended but still available via the CLI and API until December 31, 2023.

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

[Create a launch template](#)

Version

[Create a launch template version](#)

Description
asg-template

AMI ID
ami-05295b6e6c790593e

Key pair name
pringle

Launch template
[demoTemplate](#)
lt-07c1ade3e2dc2cbb

Security groups
-

Security group IDs
[sg-03ae4b3e0f5717e7c](#)

Instance type
t2.micro

Request Spot Instances
No

Additional details

Storage (volumes)
-

Date created
Sun Mar 31 2024 01:39:38
GMT+05:30 (India Standard Time)

Cancel

Next

Step 2
Choose instance launch options

Step 3 - optional
[Configure advanced options](#)

Step 4 - optional
[Configure group size and scaling](#)

Step 5 - optional
[Add notifications](#)

Step 6 - optional
[Add tags](#)

Step 7
[Review](#)

Instance type requirements info [Override launch template](#)

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
[demoTemplate](#)
lt-07c1ade3e2dc2cbb

Version
Default

Description
asg-template

Instance type
t2.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.

172.31.0.0/16 Default
[Create a VPC](#)

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

ap-south-1a | subnet-0d2a8a51b672e15d9
172.31.32.0/20 Default

ap-south-1b | subnet-08aa87792b7effa26
172.31.0.0/20 Default

ap-south-1c | subnet-0b678bf93da49f82d
172.31.16.0/20 Default

[Create a subnet](#)

Your requested instance type (t2.micro) is not available in 1 Availability Zone. You may need to change the instance type or choose other Availability Zones for better resiliency. [Learn more](#)

Cancel

Skip to review

Previous

Next

A2305221030

11

EC2 > Auto Scaling groups > Create Auto Scaling group

Step 1
[Choose launch template](#)

Step 2
[Choose instance launch options](#)

Step 3 - optional
Configure advanced options

Step 4 - optional
[Configure group size and scaling](#)

Step 5 - optional
[Add notifications](#)

Step 6 - optional
[Add tags](#)

Step 7
[Review](#)

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.

EC2 health checks [Info](#)
☒ **Always enabled**

Additional health check types - optional [Info](#)
☐ **Turn on Elastic Load Balancing health checks**
 Elastic Load Balancing monitors whether instances are available to handle requests. When it reports an unhealthy instance, EC2 Auto Scaling can replace it on its next periodic check.

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

Additional settings

Monitoring [Info](#)
☐ **Enable group metrics collection within CloudWatch**

Default instance warmup [Info](#)
 The amount of time that CloudWatch metrics for new instances do not contribute to the group's aggregated instance metrics, as their usage data is not available yet.

3. Configure the required no of instances.

EC2 > Auto Scaling groups > Create Auto Scaling group

Step 1
[Choose launch template](#)

Step 2
[Choose instance launch options](#)

Step 3 - optional
[Configure advanced options](#)

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.

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 [Info](#)
 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.

Instance maintenance policy [Info](#)
 Control your Auto Scaling group's availability during instance replacement events. This includes health checks, instance refreshes, maximum instance lifetime features and events that happen automatically to keep your group balanced, called rebalancing events.

EC2 > Auto Scaling groups

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

[Refresh](#) [Launch config](#)

<input type="checkbox"/>	Name	Launch template/configuration	Instances	Status	Desired capacity	Min	Max	Availability Zones
<input type="checkbox"/>	demoASG	demoTemplate Version Default	0	Updating capacity...	2	1	3	ap-south-1c, ap-south-1b, ap-south-1a

4. We have successfully created the autoscaling groups

demoASG

Details

Activity

Automatic scaling

Instance management

Monitoring

Instance refresh

Activity notifications (0)

Filter notifications

Send to

On instance action

No notifications are currently specified

Create notification

Activity history (5)

Filter activity history

Activity history

Instances (2)

Find Instance by attribute or tag (case-sensitive)

Running

Connect

Instance state

Actions

Launch instances

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 ...	Elastic IP	IPv6 IPs	Monitoring	Security grou
	i-0f6b6f95d79a7dfcb	Running	t2.micro	Initializing	View alarms +	ap-south-1b	ec2-13-201-99-135.ap-...	13.201.99.135	-	-	disabled	launch-wizard
	i-0ea5229f2162ecba7	Running	t2.micro	Initializing	View alarms +	ap-south-1a	ec2-15-206-122-105.ap-...	15.206.122.105	-	-	disabled	launch-wizard

5. Testing the autoscaling groups.

Hello World from ip-172-31-3-106.ap-south-1.compute.internal