

Maintaining High Availability with Auto Scaling

edureka!

edureka!

© Brain4ce Education Solutions Pvt. Ltd.

Overview

Auto Scaling allows you to scale your Amazon EC2 capacity up or down automatically according to conditions you define. With Auto Scaling, you can ensure that the number of Amazon EC2 instances you're using increases seamlessly during demand spikes to maintain performance and decreases automatically during demand lulls to minimize costs. Auto Scaling is particularly well suited for applications that experience hourly, daily, or weekly variability in usage.

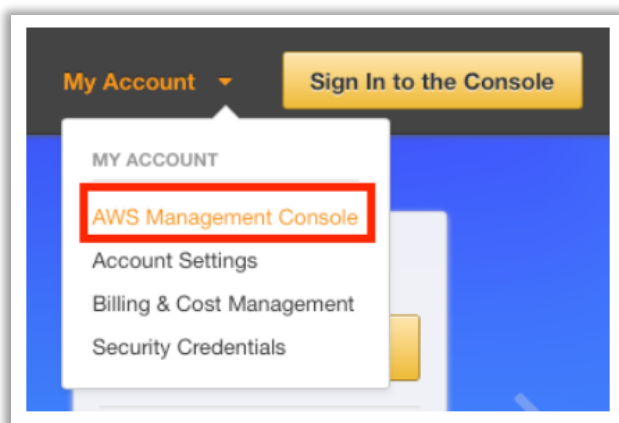
The following step- by-step instructions help you create a template that defines your EC2 instances and create an Auto Scaling group to maintain the healthy number of instances at all times.

Create a Launch Configuration

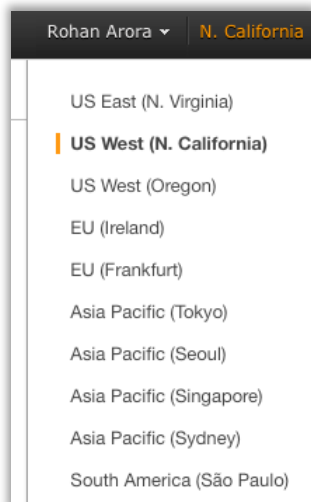
A launch configuration specifies the type of EC2 instance that Auto Scaling creates for you. You create the launch configuration by including information such as the Amazon Machine Image (AMI) ID to use for launching the EC2 instance, the instance type, key pairs, security groups, and block device mappings, among other configuration settings.

To create a launch configuration

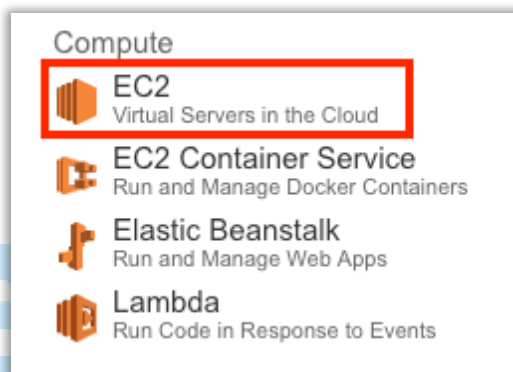
1. Login to AWS Management Console.



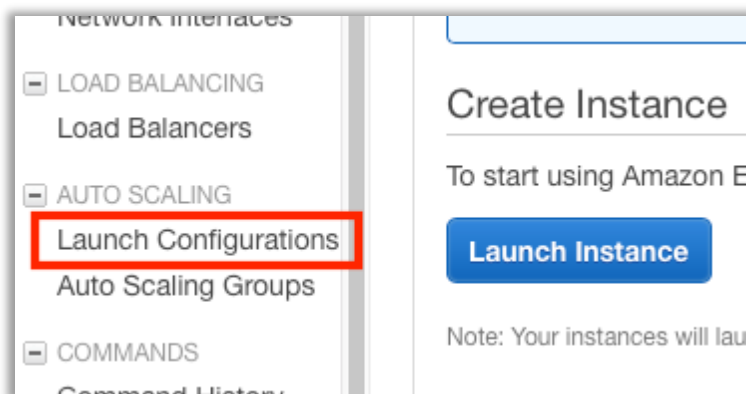
2. Select your preferred Region. The Auto Scaling resources that you create are tied to the region you specify and are not replicated across regions.



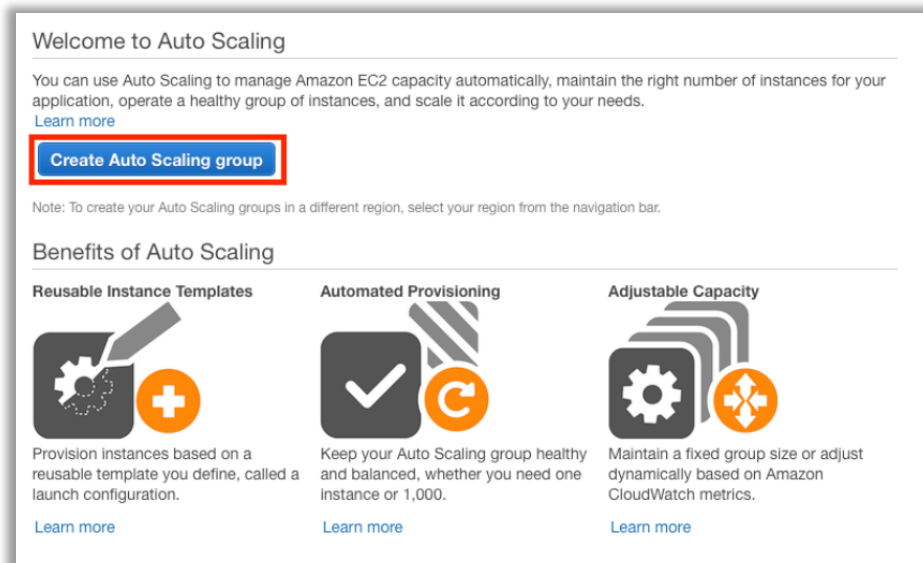
3. Click [EC2](#) under Compute section. This will take you to EC2 dashboard.



4. On the navigation pane, under [Auto Scaling](#), choose [Launch Configurations](#).



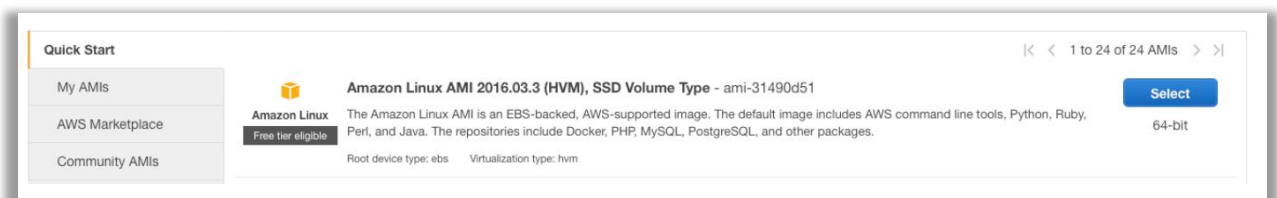
5. On the [Welcome to Auto Scaling](#) page, choose [Create Auto Scaling group](#).



6. On the [Create Auto Scaling Group](#) page, choose [Create launch configuration](#).



7. On the [Choose AMI](#) page, there is a list of basic configurations, called Amazon Machine Images (AMIs), that serve as templates for your instance. Select the 64-bit Amazon Linux AMI.



8. On the [Choose Instance Type](#) page, select a hardware configuration for your instance. We recommend that you keep the default, a *t2.micro* instance. Choose [Next: Configure details](#).

Currently selected: t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)

	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)
<input type="checkbox"/>	General purpose	t2.nano	1	0.5	EBS only
<input checked="" type="checkbox"/>	General purpose	t2.micro Free tier eligible	1	1	EBS only
<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only

9. On the Configure Details page, type a name for your launch configuration in the **Name** section.

Create Launch Configuration

Name

Purchasing option ☐ Request Spot Instances

IAM role

Monitoring ☐ Enable CloudWatch detailed monitoring
[Learn more](#)

10. Keep other settings to default and click [Skip to review](#).
11. On the [Review](#) page, choose [Edit security groups](#). Follow the instructions to choose an existing security group, and then choose [Review](#).

Assign a security group: ☒ Create a new security group
☐ Select an existing security group

Security group name:

Description:

Type	Protocol	Port Range	Source
SSH	TCP	22	Anywhere 0.0.0.0/0

[Add Rule](#)

12. On the [Review](#) page, choose [Create launch configuration](#).

AMI Details [Edit AMI](#)

Amazon Linux AMI 2016.03.3 (HVM), SSD Volume Type - ami-31490d51
The Amazon Linux AMI is an EBS-backed, AWS-supported image. The default image includes AWS command line tools, Python, Ruby, Perl, and Java. The repositories include Docker, PHP, MySQL, PostgreSQL, and other packages.
Root device type: ebs Virtualization Type: hvm

Free tier eligible

Instance Type [Edit instance type](#)

Instance Type	ECUs	vCPUs	Memory GiB	Instance Storage (GiB)	EBS-Optimized Available	Network Performance
t2.micro	Variable	1	1	EBS only	-	Low to Moderate

Launch configuration details [Edit details](#)

Name: Test-LC
Purchasing option: On demand

[Cancel](#) [Previous](#) [Create launch configuration](#)

13. On the [Select an existing key pair or create a new key pair](#) page, select one of the listed options.

Select an existing key pair or create a new key pair [×](#)

A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair

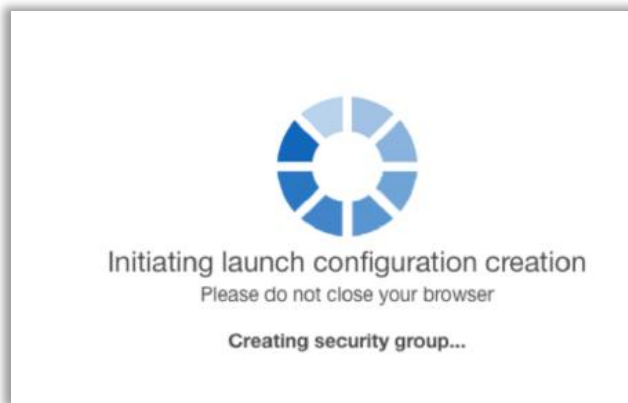
Select a key pair

linuxec2-kp

☒ I acknowledge that I have access to the selected private key file (linuxec2-kp.pem), and that without this file, I won't be able to log into my instance.

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

14. Choose [Create launch configuration](#). This initiates the process to create this Launch Configuration.

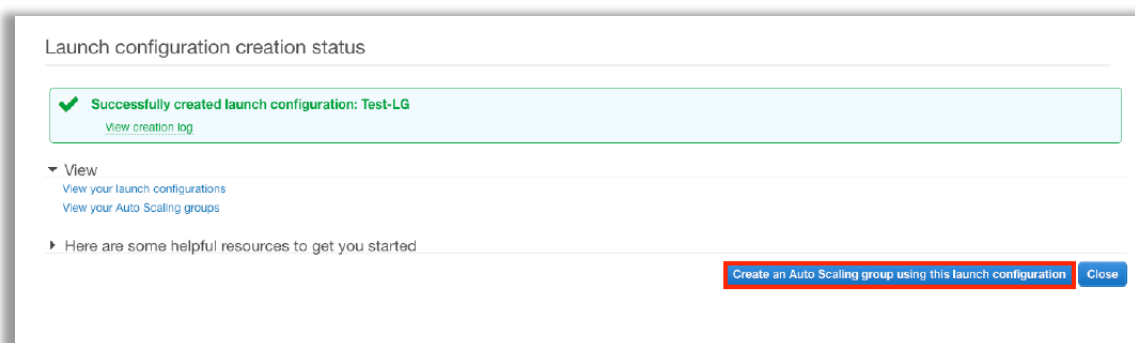


Create an Auto Scaling Group

An Auto Scaling group is a collection of EC2 instances, and the core of the Auto Scaling service. You create an Auto Scaling group by specifying the launch configuration you want to use for launching the instances and the number of instances your group must maintain at all times. You also specify the Availability Zone in which you want the instances to be launched.

To create an Auto Scaling group

15. Click **Create an Auto Scaling group using this launch configuration**.



16. For **Group name**, type a name for your Auto Scaling group.

17. Keep **Group size** set to the default value of 1 instance for this tutorial.

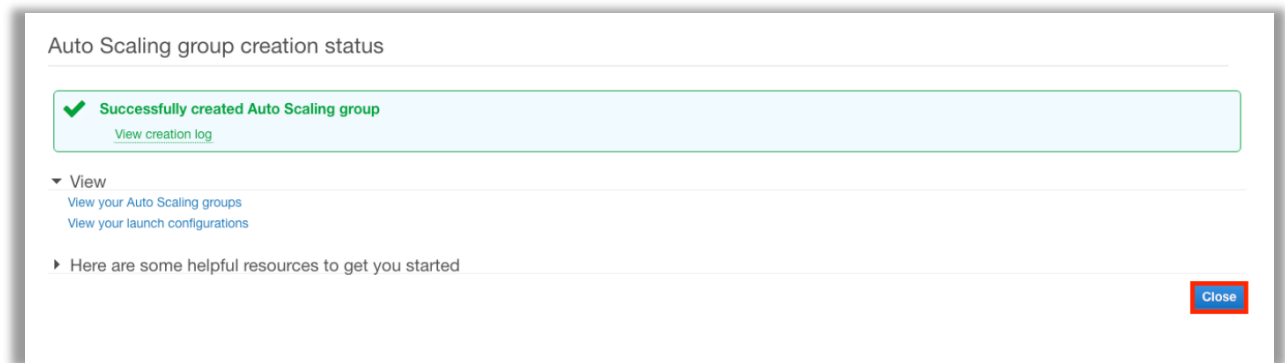
18. Choose VPC in **Network** and its associated subnets in **Subnet**. In this example, we select our default VPC and one of its subnets.

19. Choose **Next: Configure scaling policies**.

20. On the **Configure scaling policies** page, select **Keep this group at its initial size** and choose **Review**.

21. On the **Review** page, choose **Create Auto Scaling group**.

22. On the **Auto Scaling group creation status** page, choose **Close**.



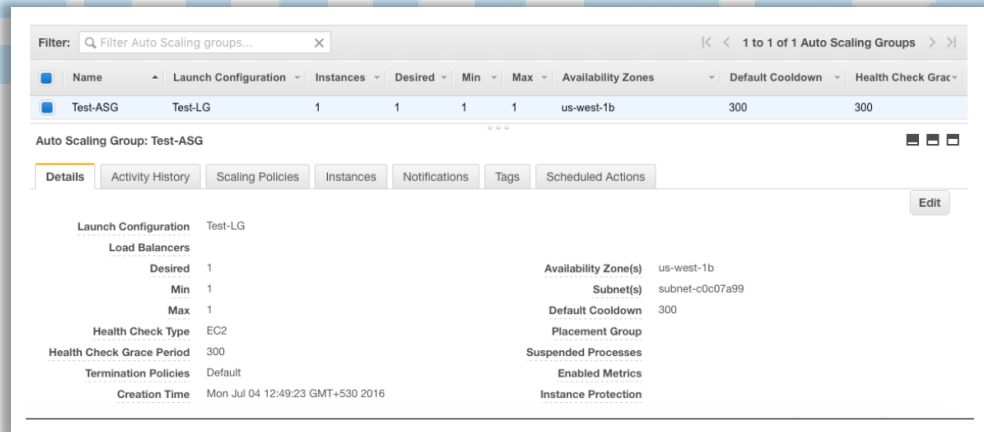
Verify your Auto Scaling Group

Now that you have created your Auto Scaling group, you are ready to verify that the group has launched an EC2 instance.

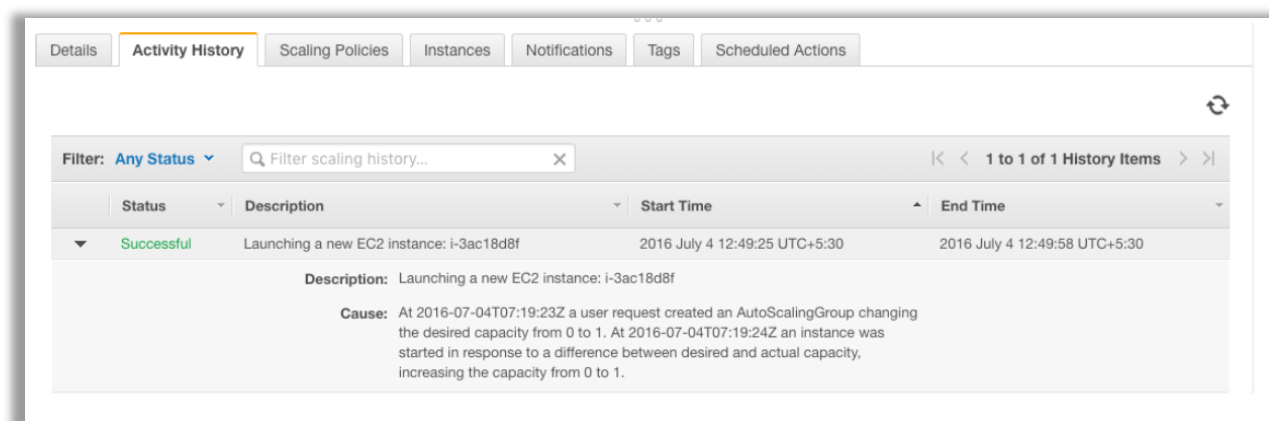
To verify that your Auto Scaling group has launched an EC2 instance

23. On the Auto Scaling Groups page, select the Auto Scaling group that you just created.

24. The **Details** tab provides information about the Auto Scaling group.



25. On the **Activity History** tab, the **Status** column shows the current status of your instance. While your instance is launching, the status column shows in progress. The status changes too Successful after the instance is launched. You can also use the refresh button to see the current status of your instance.



26. On the **Instances** tab, the **Lifecycle column** shows the state of your instance. You can see that your Auto Scaling group has launched your EC2 instance, and that it is in the *InService* lifecycle state. The **Health Status** column shows the result of the EC2 instance health check on your instance.

