

# Getting Started with AWS - Launch Configuration for an Auto Scaling Group in AWS

## Launch Configurations

A launch configuration is an instance configuration template that an Auto Scaling group uses to launch EC2 instances. When you create a launch configuration, you specify information for the instances. Include the ID of the Amazon Machine Image (AMI), the instance type, a key pair, one or more security groups, and a block device mapping. If you've launched an EC2 instance before, you specified the same information in order to launch the instance.

### Step 1: To create the launch configuration for an Auto Scaling group (console).

1. Open the Amazon EC2 console at <https://console.aws.amazon.com/ec2/>.
2. In the navigation pane, choose **Launch Configurations**.
3. Select or choose **"Create launch configuration."**
4. Select an AMI (operating system) Linux **"Next"**
5. Select an instance types optimized to fit different use cases **"general purpose t2.micro"** Next

**Provide the name for your launch configuration and you can add or skip the script for HTTPD service installation.**

The screenshot shows the 'Create Launch Configuration' page in the AWS Management Console. The page is divided into six steps: 1. Choose AMI, 2. Choose Instance Type, 3. Configure details, 4. Add Storage, 5. Configure Security Group, and 6. Review. Step 3 is currently active. The 'Name' field is set to 'Launch Configuration 1'. The 'Purchasing option' is 'Request Spot Instances'. The 'IAM role' is 'None'. The 'Monitoring' section has 'Enable CloudWatch detailed monitoring' checked. The 'Advanced Details' section is expanded, showing 'Kernel ID' as 'Use default', 'RAM Disk ID' as 'Use default', and 'User data' as 'As text'. The 'User data' field contains a script for installing and configuring HTTPD. A red box highlights the script content, and a red box highlights the 'Next: Add Storage' button.

```
#!/bin/bash
sudo su
yum update -y
yum install httpd -y
service httpd start
chkconfig httpd on
echo "<html><h1>This is served from <i>${HOSTNAME}</i></h1>"
```

6. Add the storage click "Next"
7. Create or assign a security group "Next"
8. Create and launch the **configuration**
9. Select the security **Key** pair

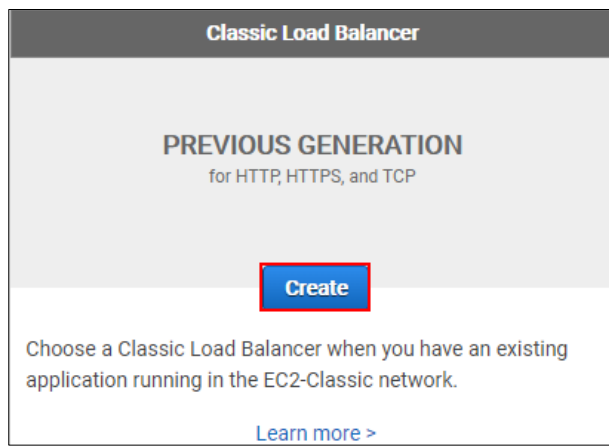
Launch configuration creation status

✓ Successfully created launch configuration: Launch Configuration 1  
[View creation log](#)

## Add a Load Balancer (Console)

### Step 2: Attaching a Load Balancer to Your Auto Scaling Group

1. On the navigation pane, under **Load Balancers**, choose **Create Load Balancer**
2. Select **Classic Load Balancers**



## Define the name of Load Balancer and select the port “80”

1. Define Load Balancer 2. Assign Security Groups 3. Configure Security Settings 4. Configure Health Check 5. Add EC2 Instances 6. Add Tags 7. Review

### Step 1: Define Load Balancer

#### Basic Configuration

This wizard will walk you through setting up a new load balancer. Begin by giving your new load balancer a unique name so that you can identify it from other load balancers you might create. You will also need to configure ports and protocols for your load balancer. Traffic from your clients can be routed from any load balancer port to any port on your EC2 instances. By default, we've configured your load balancer with a standard web server on port 80.

Load Balancer name:

Create LB inside:  (where's that?)

Create an internal load balancer: ☐

Enable advanced VPC configuration: ☐

Listener Configuration:

Load Balancer Protocol	Load Balancer Port	Instance Protocol	Instance Port
HTTP	80	HTTP	80

[Add](#)

[Cancel](#) [Next: Assign Security Groups](#)

## Select the Security group

### Step 2: Assign Security Groups

You have selected the option of having your Elastic Load Balancer inside of a VPC, which allows you to assign security groups to your load balancer. Please select the security groups to assign to this load balancer. This can be changed at any time.

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

Filter: VPC security groups

Security Group ID	Name	Description	Actions
sg-d34c81ac	default	default VPC security group	<a href="#">Copy to new</a>
sg-0b522a419b6fb8f8d	SG-1	launch-wizard-1 created 2019-09-27T09:40:57.325+05:30	<a href="#">Copy to new</a>

[Cancel](#) [Previous](#) [Next: Configure Security Settings](#)

## Configure security settings click “next”

## Configure Health Check

### Step 4: Configure Health Check

Your load balancer will automatically perform health checks on your EC2 instances and only route traffic to instances that pass the health check. If an instance fails the health check, it is automatically removed from the load balancer. Customize the health check to meet your specific needs.

Ping Protocol:

Ping Port:

Ping Path:

#### Advanced Details

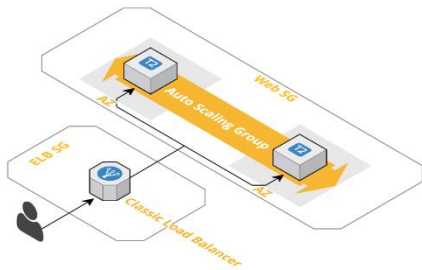
Response Timeout:  seconds

Interval:  seconds

Unhealthy threshold:

Healthy threshold:

[Cancel](#) [Previous](#) [Next: Add EC2 Instances](#)



On the confirmation page, we choose to go and create your Auto Scaling groups.  
Choose **Create Auto Scaling group**

**Save up to 90% on Compute**  
Optimize compute costs by creating your Auto Scaling group with a launch template to combine EC2 On-Demand, Spot, and RIs. [Learn more.](#)

### Welcome to Auto Scaling

You can use Auto Scaling to manage Amazon EC2 capacity automatically, maintain the right number of instances for your application, operate a healthy group of instances, and scale it according to your needs. [Learn more](#)

[Create Auto Scaling group](#)

Note: To create your Auto Scaling groups in a different region, select your region from the navigation bar.

### Benefits of Auto Scaling

**Automated Provisioning**

Keep your Auto Scaling group healthy and balanced, whether you need one instance or 1,000. [Learn more](#)

**Adjustable Capacity**

Maintain a fixed group size or adjust dynamically based on Amazon CloudWatch metrics. [Learn more](#)

**Launch Template Support**

Provision instances easily using EC2 Launch Templates. [Learn more](#)

### Additional Information

- [Getting Started Guide](#)
- [Documentation](#)
- [All EC2 Resources](#)
- [Forums](#)
- [Pricing](#)
- [Contact Us](#)

### Create Auto Scaling Group

Complete this wizard to create your Auto Scaling group. First, choose either a launch configuration or a launch template to specify the parameters that your Auto Scaling group uses to launch instances.

**Launch Configuration**

You can continue to use your launch configurations if they support the Amazon EC2 features you need. [Learn more](#)

[Create a new launch configuration](#)

**Launch Template** New

Launch templates give you the option of launching one type of instance, or a combination of instance types and purchase options. Launch templates include the latest Amazon EC2 features and can be updated and versioned. [Learn more](#)

[Create new launch template](#)

Name	AMI ID	Instance Type	Spot Price	Security Groups
Launch Configuration 1	ami-05c859630889c79c8	t2.micro		sg-0b522a41968fb8fd

**“Group name”**

For **Subnet**, choose a subnet for the VPC.

### Note

You can choose the Availability Zone for your instance by choosing its corresponding default subnet.

1. Configure Auto Scaling group details 2. Configure scaling policies 3. Configure Notifications 4. Configure Tags 5. Review

### Create Auto Scaling Group

Group Name: **CASG1**

Launch Configuration: CLC1

Group size: Start with **1** instances. **We will keep the Instance 1 only**

Network: **vpc-92a5e5e8 (172.31.0.0/16) (default)**

Subnet: **subnet-6770d459 (172.31.48.0/20) (Default in us-east-1a)**

You can select all the subnet which are available within the region

Advanced Details

Load Balancing: ☒ Receive traffic from one or more load balancers

Classic Load Balancers: **ELB1**

Target Groups: **ELB1**

Health Check Type: ☒ ELB ☐ EC2

Health Check Grace Period: **300** seconds

Monitoring: ☒ Enable CloudWatch detailed monitoring

Instance Protection: ☐ Enable

Service-Linked Role: **AWSAutoScalingRoleForECS**

Cancel Next: Configure scaling policies

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

1. Configure Auto Scaling group details 2. Configure scaling policies 3. Configure Notifications 4. Configure Tags 5. Review

### Create Auto Scaling Group

You can optionally add scaling policies if you want to adjust the size (number of instances) of your group automatically. A scaling policy is a set of instructions for making such adjustments in response to an Amazon CloudWatch alarm that you assign to it. In each policy, you can choose to add or remove a specific number of instances or a percentage of the existing group size, or you can set the group to an exact size. When the alarm triggers, it will execute the policy and adjust the size of your group accordingly. [Learn more about scaling policies.](#)

**Keep this group at its initial size**

Use scaling policies to adjust the capacity of this group

You can setup the policies to this group and setup the later

Cancel Previous **Review** Next: Configure Notifications

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

1. Configure Auto Scaling group details 2. Configure scaling policies 3. Configure Notifications 4. Configure Tags 5. Review

### Create Auto Scaling Group

Configure your Auto Scaling group to send notifications to a specified endpoint, such as an email address, whenever a specified event takes place, including: successful launch of an instance, failed instance launch, instance termination, and failed instance termination.

If you created a new topic, check your email for a confirmation message and click the included link to confirm your subscription. Notifications can only be sent to confirmed addresses.

Add notification

Cancel Previous **Review** Next: Configure Tags

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

1. Configure Auto Scaling group details 2. Configure scaling policies 3. Configure Notifications 4. Configure Tags 5. Review

### Create Auto Scaling Group

Please review your Auto Scaling group details. You can go back to edit changes for each section. Click **Create Auto Scaling group** to complete the creation of an Auto Scaling group.

Auto Scaling Group Details

Group name	ASG 1
Group size	1
Minimum Group Size	1
Maximum Group Size	1
Subnet(s)	subnet-a4f6e1c2, subnet-a27af6a, subnet-42e44619
Health Check Grace Period	300
Detailed Monitoring	No
Instance Protection	None
Service-Linked Role	AWSAutoScalingRoleForECS

Scaling Policies

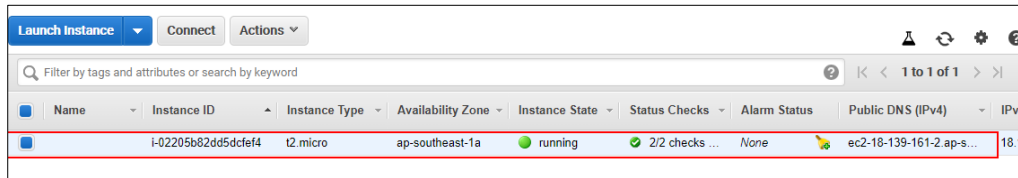
Notifications

Tags


Cancel Previous **Create Auto Scaling group**

Done

### Step 3: Please check your EC2 Instance is ready from your auto scaling lunch configuration.

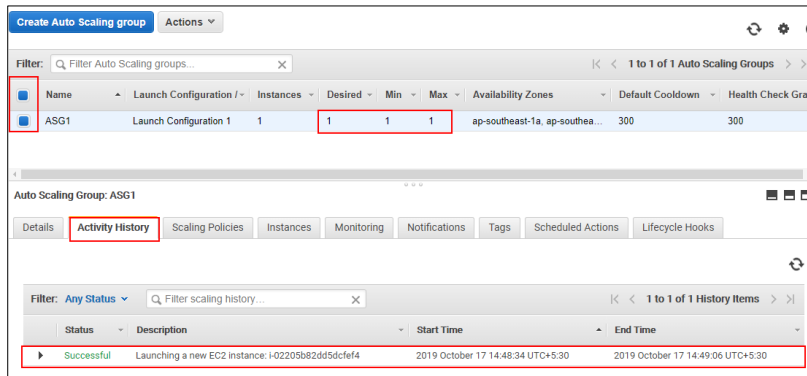


Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4
	i-02205b82dd5dcfe4	t2.micro	ap-southeast-1a	running	2/2 checks ...	None	ec2-18-139-161-2.ap-s...	18.1...



This is served from *ip-172-31-21-215*

### Step 4: Check the status of autoscaling group with one instance.

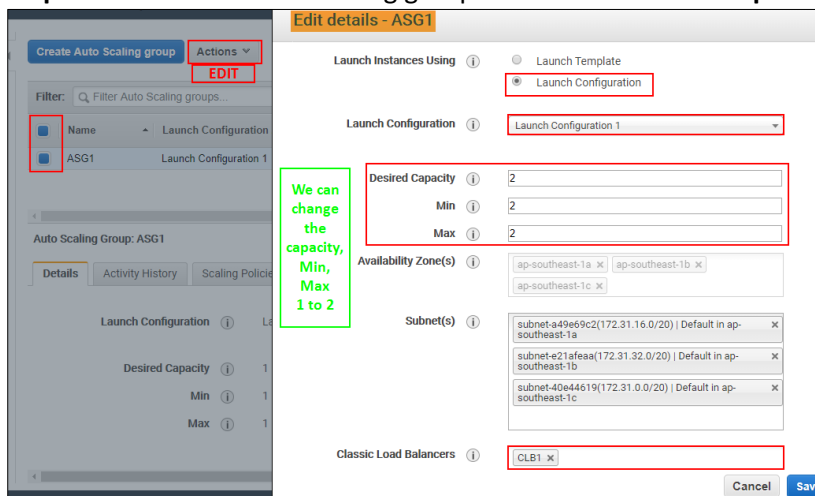


Name	Launch Configuration	Instances	Desired	Min	Max	Availability Zones	Default Cooldown	Health Check Grace
ASG1	Launch Configuration 1	1	1	1	1	ap-southeast-1a, ap-southea...	300	300

Status	Description	Start Time	End Time
Successful	Launching a new EC2 instance: i-02205b82dd5dcfe4	2019 October 17 14:48:34 UTC+5:30	2019 October 17 14:49:06 UTC+5:30

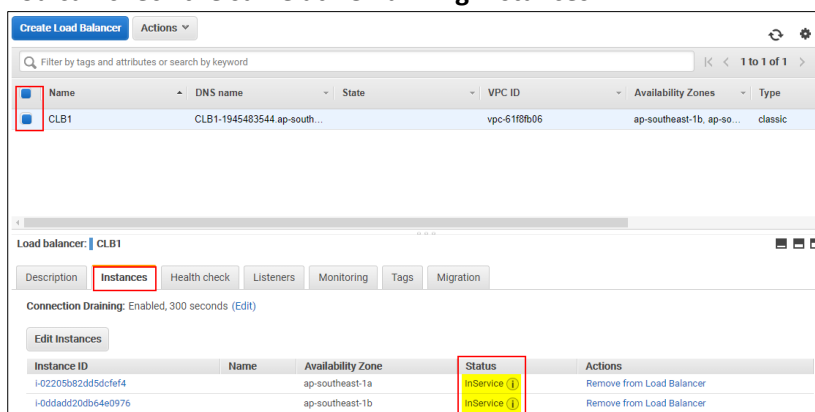
### Steps 5: We can edit auto scaling group and make it Desired capacity, Min, Max 2



We can change the capacity, Min, Max 1 to 2

Desired Capacity: 2  
Min: 2  
Max: 2

### Steps 6: Now we can check in Load Balancers if there is “two Instance are there” You can check the same at EC Running Instances.



Name	DNS name	State	VPC ID	Availability Zones	Type
CLB1	CLB1-1945483544.ap-south...		vpc-61f8b06	ap-southeast-1b, ap-so...	classic

Instance ID	Name	Availability Zone	Status	Actions
i-02205b82dd5dcfe4		ap-southeast-1a	InService	Remove from Load Balancer
i-0ddadd20db64e0976		ap-southeast-1b	InService	Remove from Load Balancer

**Steps 7:** Same we can check the Instances health status.

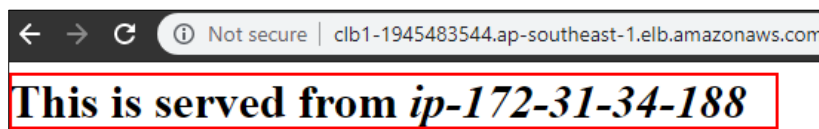
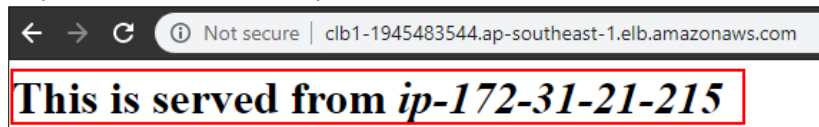
Name	Launch Configuration / Template	Instances	Desired	Min	Max	Availability Zones	Default Cooldown	Health Check Grace Period
ASG1	Launch Configuration 1	2	2	2	2	ap-southeast-1a, ap-southeast-1b	300	300

Instance ID	Lifecycle	Launch Configuration / Template	Availability Zone	Health Status	Protected from
i-02205b82d5d5cfe4	InService	Launch Configuration 1	ap-southeast-1a	Healthy	
i-0ddadd20db64e0976	InService	Launch Configuration 1	ap-southeast-1b	Healthy	

**Steps 8:** Now you can access the Instance using DNS service.

<http://clb1-1945483544.ap-southeast-1.elb.amazonaws.com/>



Done

### A) Configure simple scaling policy

Go to Auto scaling group configure the settings as per the table below.

1. Scaling Policies tab selected

2. Add policy button clicked

3. Policy Name: ScaleOut

4. Create a simple scaling policy button selected

**Creating simple scaling Out policy for average above the 60% (Add the new Instance)**

1. Policy Name: ScaleOut

2. Create new alarm button

3. Execute policy when: No alarm selected

4. Take the action: Add

5. And then wait: 120 seconds before allowing another scaling activity

6. Create a simple scaling policy button

7. Create button

For how to setup the alarm please see the...

## Alarm Settings: For creating a simple scaling policy

### Create Alarm

You can use CloudWatch alarms to be notified automatically whenever metric data reaches a level you define.  
To edit an alarm, first choose whom to notify and then define when the notification should be sent.

☐ Send a notification to: No SNS topics found...

Whenever: Average of CPU Utilization

Is: >= 60 Percent

For at least: 1 consecutive period(s) of 1 Minute

Name of alarm: ScaleOut-Alarm

Cancel Create Alarm

CPU Utilization Percent

CASG1

## Creating simple scaling In policy for average below the 60% (Remove the new Instance automatically)

### Create Scaling policy

Name: ScaleIn

Execute policy when: ScaleIn Create new alarm

breaches the alarm threshold: CPUUtilization <= 59 for 60 seconds  
for the metric dimensions: AutoScalingGroupName = CASG1

Take the action: Remove 1 instances

And then wait: 120 seconds before allowing another scaling activity

Create a target tracking scaling policy  
Create a scaling policy with steps

Cancel Create

For how to setup the alarm please see the next screen...

## Alarm Settings: For creating a simple scaling policy

### Create Alarm

You can use CloudWatch alarms to be notified automatically whenever metric data reaches a level you define.  
To edit an alarm, first choose whom to notify and then define when the notification should be sent.

☐ Send a notification to: No SNS topics found...

Whenever: Average of CPU Utilization

Is: <= 59 Percent

For at least: 1 consecutive period(s) of 1 Minute

Name of alarm: ScaleIn

Cancel Create Alarm

CPU Utilization Percent

CASG1

The screenshot shows the 'Scaling Policies' tab in the AWS Management Console. There are two policies listed:

- ScaleIn:**
  - Policy type: Simple scaling
  - Execute policy when: ScaleIn (breaches the alarm threshold: CPUUtilization <= 59 for 60 seconds for the metric dimensions AutoScalingGroupName = CASG1)
  - Take the action: Remove 1 instances
  - And then wait: 120 seconds before allowing another scaling activity
- ScaleOut:**
  - Policy type: Simple scaling
  - Execute policy when: No alarm selected
  - Take the action: Add 1 instances
  - And then wait: 120 seconds before allowing another scaling activity

1. On the **Configure scaling policies** page, do the following:
  - a. Select **Use scaling policies to adjust the capacity of this group**.
  - b. Specify the minimum and maximum size for your Auto Scaling group using the row that begins with **Scale between**. For example, if your group is already at its maximum size, you need to specify a new maximum in order to scale out.

Scale between  and  instances. These will be the minimum and maximum size of your group.

- c. Specify your scale-out policy under **Increase Group Size**. You can optionally specify a name for the policy, then choose **Add new alarm**.
  - d. On the **Create Alarm** page, choose **create topic**. For **Send a notification to**, type a name for the SNS topic. For **With these recipients**, type one or more email addresses to receive notification. You can replace the default name for your alarm with a custom name. Next, specify the metric and the criteria for the policy. For example, you can leave the default settings for **Whenever** (Average of CPU Utilization). For **Is**, choose **>=** and type **80** percent. For **For at least**, type **1** consecutive period of **5** Minutes. Choose **Create Alarm**.

The 'Create Alarm' dialog box is shown with the following configuration:

- Send a notification to:** AddCapacityNotification
- With these recipients:** mymail@example.com
- Whenever:** Average of CPU Utilization
- Is:** >= 80 Percent
- For at least:** 1 consecutive period(s) of 5 Minutes
- Name of alarm:** AddCapacityAlarm

A line graph on the right shows 'CPU Utilization Percent' over time, with a single data point at 12/10 04:00.

- e. For **Take the action**, choose **Add**, type **30** in the next field, and then choose **percent** of **group**. By default, the lower bound for this step adjustment is the alarm threshold and the upper bound is null (positive infinity).



To add another step adjustment, choose **Add step**. To set a minimum number of instances to scale, update the number field in **Add instances in increments of at least 1 instance(s)**.

(Optional) We recommend that you use the default to create scaling policies with steps. To create simple scaling policies, choose **Create a simple scaling policy**. For more information, see [Simple and Step Scaling Policies for Amazon EC2 Auto Scaling](#).

### Increase Group Size

**Name:**

**Execute policy when:** [AddCapacityAlarm](#) [Edit](#) [Remove](#)  
breaches the alarm threshold: CPUUtilization >= 80 for 300 seconds  
for the metric dimensions AutoScalingGroupName = my-asg

**Take the action:**    when  <= CPUUtilization < +infinity  
[Add step](#) ⓘ  
Add instances in increments of at least  instance(s)

**Instances need:**  seconds to warm up after each step

[Create a simple scaling policy](#) ⓘ

- f. Specify an instance warm-up value for **Instances need**, which allows you to control the amount of time until a newly launched instance can contribute to the CloudWatch metrics.
- g. Specify your scale-in policy under **Decrease Group Size**. You can optionally specify a name for the policy, then choose **Add new alarm**.
- h. On the **Create Alarm** page, you can select the same notification that you created for the scale-out policy or create a new one for the scale-in policy. You can replace the default name for your alarm with a custom name. Keep the default settings for **Whenever** (Average of CPU Utilization). For **Is**, choose <= and type 40 percent. For **For at least**, type 1 consecutive period of 5 Minutes. Choose **Create Alarm**.
- i. For **Take the action**, choose **Remove**, type 2 in the next field, and then choose **instances**. By default, the upper bound for this step adjustment is the alarm threshold and the lower bound is null (negative infinity). To add another step adjustment, choose **Add step**.

(Optional) We recommend that you use the default to create scaling policies with steps. To create simple scaling policies, choose **Create a simple scaling policy**. For more information, see [Scaling Policy Types](#).

### Decrease Group Size

**Name:**

**Execute policy when:** [DecreaseCapacityAlarm](#) [Edit](#) [Remove](#)  
breaches the alarm threshold: CPUUtilization <= 40 for 300 seconds  
for the metric dimensions AutoScalingGroupName = my-asg

**Take the action:**    when  >= CPUUtilization > -infinity  
[Add step](#) ⓘ  
[Create a simple scaling policy](#) ⓘ

- j. Choose **Review**.
- k. On the **Review** page, choose **Create Auto Scaling group**.
2. Use the following steps to verify the scaling policies for your Auto Scaling group.

- a. The **Auto Scaling Group creation status** page confirms that your Auto Scaling group was successfully created. Choose **View your Auto Scaling Groups**.
- b. On the **Auto Scaling Groups** page, select the Auto Scaling group that you just created.
- c. On the **Activity History** tab, the **Status** column shows whether your Auto Scaling group has successfully launched instances.
- d. On the **Instances** tab, the **Lifecycle** column contains the state of your instances. It takes a short time for an instance to launch. After the instance starts, its lifecycle state changes to InService.

The **Health Status** column shows the result of the EC2 instance health check on your instance.

- e. On the **Scaling Policies** tab, you can see the policies that you created for the Auto Scaling group.