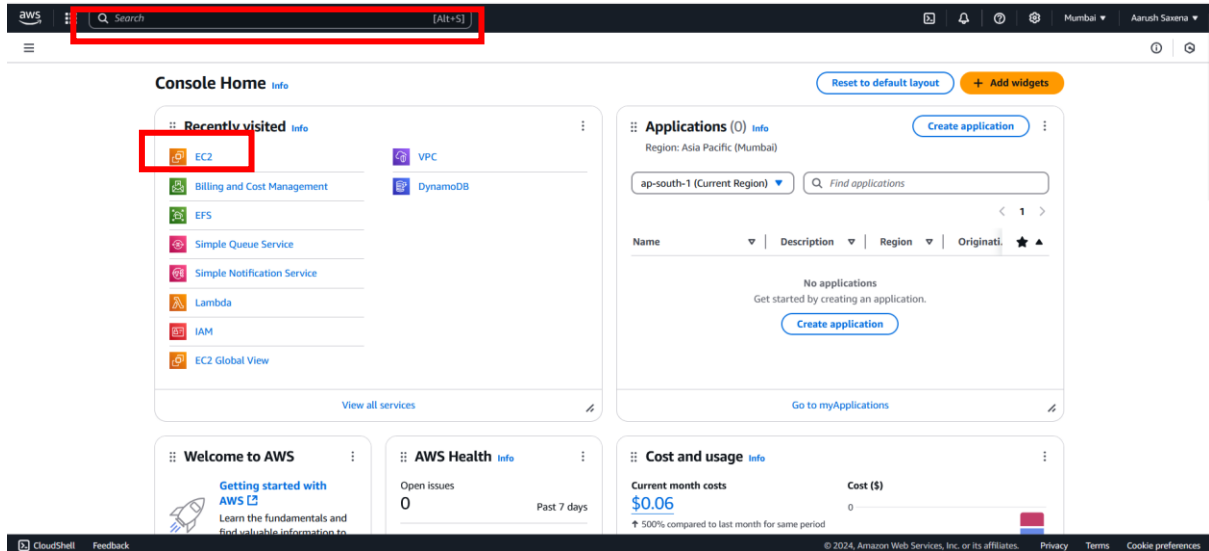
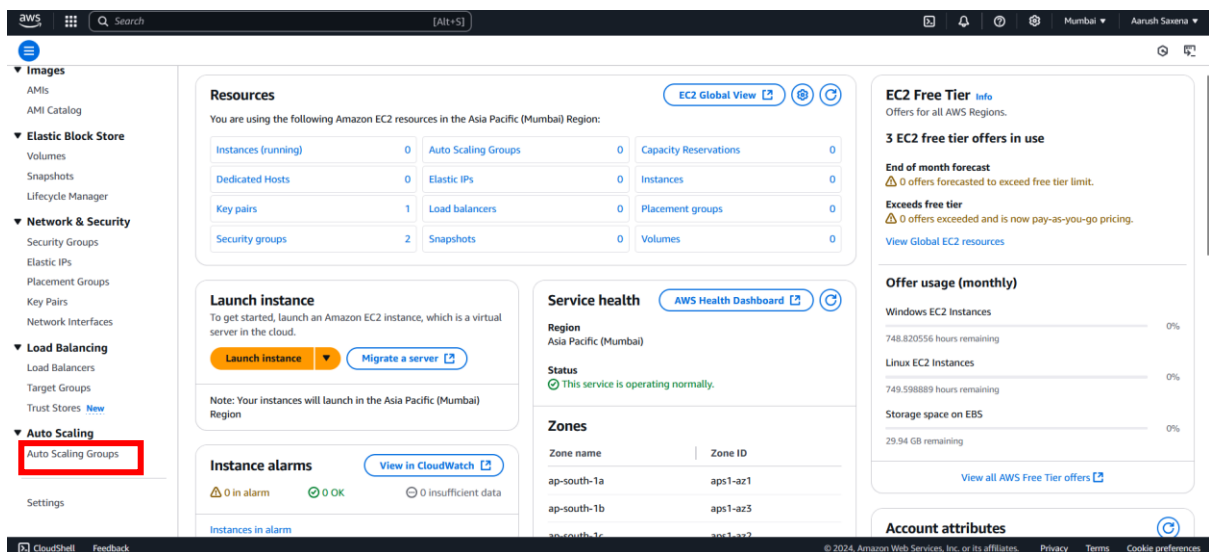


# Auto Scaling(Basic )

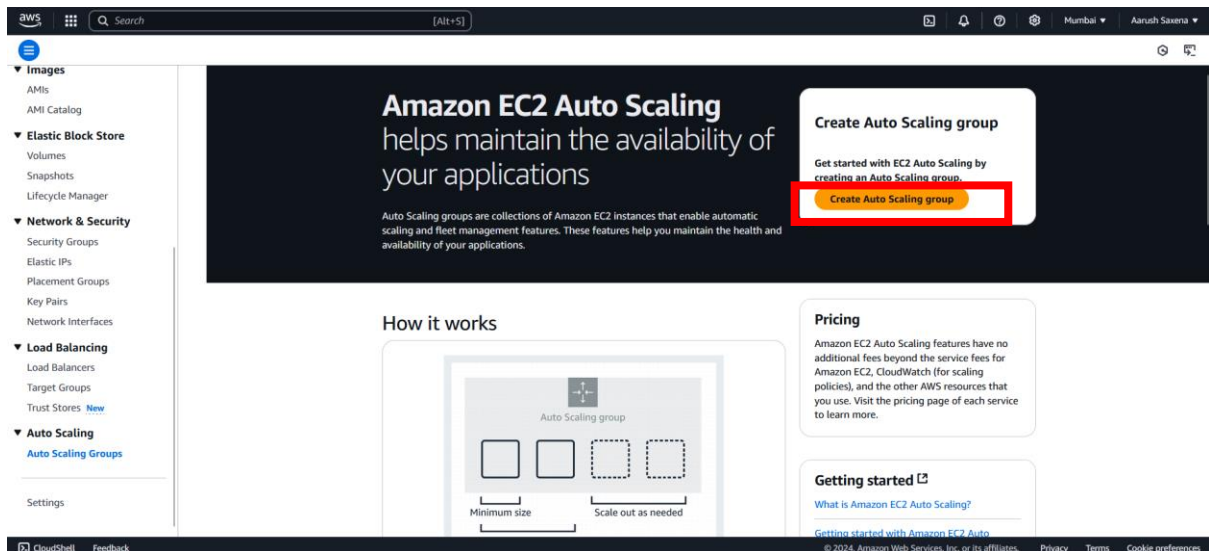
Step1: Log-in to your account from dashboard or management console go for EC2 from search bar.



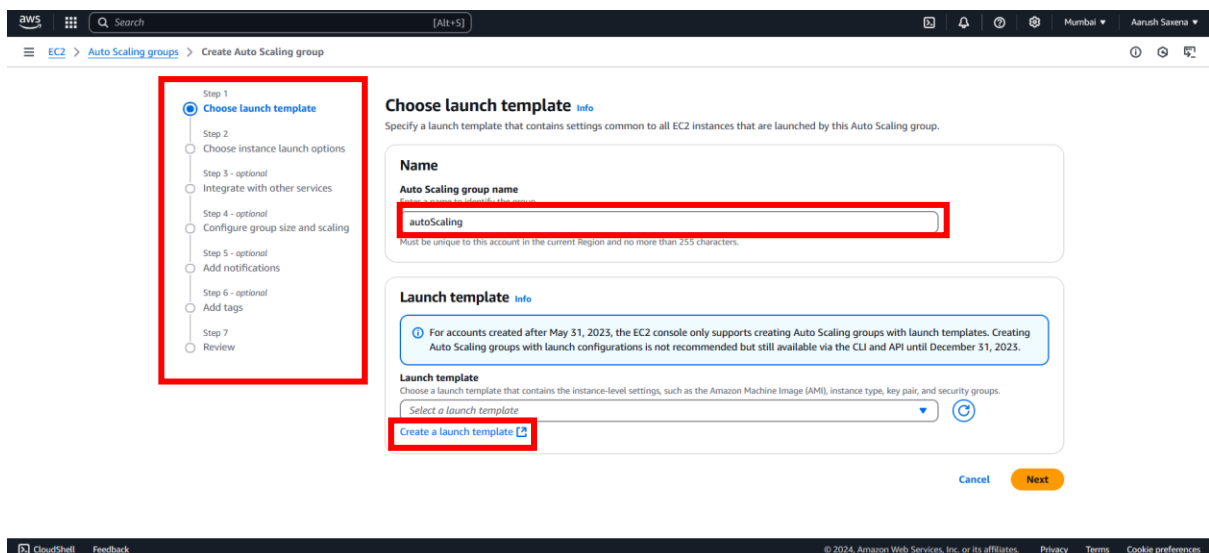
Step2: Scroll down from ec2 dashboard and there you can see auto scaling option click on that.



After clicking on it window will appear as shown in image given below there you can see create auto scaling group click on it.



After clicking on it 7 steps will show from them first step is to name your group and launch template on which you want to apply this service. As we have not created templates yet we are going to click on create a launch template.



After clicking on it a window will appear as shown in image given below. We can create templates from the EC2 dashboard from the option of Launch Templates and window will show same result. I am choosing Aws Linux choose according to your requirements. I am choosing t2.micro

aws

Q Search

[Alt+S]

EC2 > Launch templates > Create launch template

Dashboard

EC2 Global View

Events

Instances

Instances

Instance Types

Launch Templates

Spot Requests

Savings Plans

Reserved Instances

Dedicated Hosts

Capacity Reservations

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

TemplateForAutoScaling

Must be unique to this account. Max 128 chars. No spaces or special characters like "&", ":", "(", ")", "/\*", "/\*", "/\*".

Template version description

version1

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

Key [Info](#)

Value [Info](#)

Q name

Q autoScalingTemplate

Remove tag

Add new tag

You can add up to 49 more tags.

► Source template

Launch template contents

Consider the default and optional tags before. Launch a field block will result in the field set being included in the launch template.

▼ Summary

Software Image (AMI)

Amazon Linux 2023 AMI 2023.6.2...[read more](#)

ami-0f0d05997bdaff7aac

Virtual server type (instance type)

-

Firewall (security group)

-

Storage (volumes)

1 volume(s) - 8 GiB

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 GB of snapshots, and 100 GB of bandwidth to the internet.

✕

Cancel

Create launch template

Keep other settings to default or give it as per requirements. Then click on Launch Template.

**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:  [Create new key pair](#)

**Network settings** info

Subnet:  [Create new subnet](#)

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

**Firewall (security groups)** info

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.

☒ Select existing security group ☐ Create security group

Security groups:  [Compare security group rules](#)

[X](#)

**Storage (volumes)** info

1 volume(s) - 8 GiB

**Summary**

**Software Image (AMI)**  
Amazon Linux 2023 AMI 2023.6.2...[read more](#)  
ami-0f05997b4dff7aac

**Virtual server type (instance type)**  
t2.micro

**Firewall (security group)**  
launch-wizard-1

**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 IP-v4 address usage per month, 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](#)

**Launch Templates (1/1)** info

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

Launch Template ID	Launch Template Name	Default Version	Latest Version	Create Time	Created By	Managed	Operator
<input checked="" type="checkbox"/> lt-0104dbe0f64a73e26	TemplateForAutoScaling	1	1	2024-12-23T17:11:57.000Z	arnawsiam:992382687678root	false	-

**TemplateForAutoScaling (lt-0104dbe0f64a73e26)**

**Launch template details**

Launch template ID:  Launch template name:  Default version:  Owner:

[Actions](#) [Delete template](#)

**Launch template version details**

Version	Description	Date created	Created by
version1			

Now you can see your template has been launched as shown in image above.

Step3: Go to autoscaling group and provide auto scaling template and click on next

**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**  
version1

**AMI ID**  
ami-0f05997b4dff7aac

**Key pair name**  
keyPair

**Launch template**  
TemplateForAutoScaling  
lt-0104dbe0f64a73e26

**Security groups**  
-

**Security group IDs**  
sg-0ee470d20eb13722b

**Instance type**  
t2.micro

**Request Spot Instances**  
No

**Additional details**

**Storage (volumes)**  
-

**Date created**  
Mon Dec 23 2024 22:41:57 GMT+0530 (India Standard Time)

[Cancel](#) [Next](#)

In next step choose availability zone and keep other settings to default or as per requirement and click on next.

EC2 > Auto Scaling groups > Create Auto Scaling group

Create a VPC

Availability Zones and subnets

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

Select Availability Zones and subnets

ap-south-1a | subnet-05d19f63af3d53842

172.31.32.0/20 Default

ap-south-1b | subnet-0a067d3638e89fe45

172.31.0.0/20 Default

ap-south-1c | subnet-0b845d8dd14dc62f0

172.31.16.0/20 Default

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.

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

CloudShell Feedback

© 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

- 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

### Integrate with other services - optional

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

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.

#### 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.

[Create new VPC Lattice service](#)

#### Application Recovery Controller (ARC) zonal shift - new

During an Availability Zone impairment, target instance launches towards other healthy Availability Zones.

- ☐ Enable zonal shift  
New instance launches will be retargeted towards healthy Availability Zones until the zonal shift is canceled.

#### 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

[Always enabled](#)

##### Additional health check types - optional

- ☐ 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.
- ☐ Turn on VPC Lattice health checks  
VPC Lattice can monitor whether instances are available to handle requests. If it considers a target as failed a health check, EC2 Auto Scaling replaces it after its next periodic check.
- ☐ Turn on Amazon EBS health checks  
EBS monitors whether an instance's root volume or attached volume stalls. When it reports an unhealthy volume, EC2 Auto Scaling can replace the instance on its next periodic health check.

##### Health check grace period

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.

300 seconds

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

Keep it default and click on next. I have give desired capacity as 1. Which means one server has to be run at any condition.

- 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. vCPU and Memory(GiB) are only supported for mixed instance groups configured with a set of instance attributes.

Units (number of instances)

#### 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

##### Choose whether to use a target tracking policy [info](#)

You can set up other metrics-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.

#### Choose a replacement behavior depending on your availability requirements

##### ☒ Launch before terminate

###### ☒ No policy

For rebalancing events, new instances will launch before terminating others. For all other events, instances terminate and launch at the same time.

##### ☐ Prioritize availability

☐ Launch before terminating  
Launch new instances and wait for them to be ready before terminating others. This allows you to go above your desired capacity by a given percentage and may temporarily increase costs.

##### ☐ Control costs

###### ☐ Terminate and launch

Terminate and launch instances at the same time. This allows you to go below your desired capacity by a given percentage and may temporarily reduce availability.

##### ☐ Flexible

###### ☐ Custom behavior

Set custom values for the minimum and maximum amount of available capacity. This gives you greater flexibility in setting how far below and over your desired capacity EC2 Auto Scaling goes when replacing instances.

### Additional capacity settings

#### Capacity Reservation preference [info](#)

Select whether you want Auto Scaling to launch instances into an existing Capacity Reservation or Capacity Reservation resource group.

##### ☒ Default

Auto Scaling uses the Capacity Reservation preference from your launch template.

##### ☐ None

Instances will not be launched into a Capacity Reservation.

##### ☐ Capacity Reservations only

Instances will only be launched into a Capacity Reservation. If capacity isn't available, the instances fail to launch.

##### ☐ Capacity Reservations first

Instances will attempt to launch into a Capacity Reservation first. If capacity isn't available, instances will run in On-Demand capacity.

### Additional settings

#### Instance scale-in protection

If protect from scale-in is enabled, newly launched instances will be protected from scale-in by default.

☐ Enable instance scale-in protection

#### 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 reliable yet.

☐ Enable default instance warmup

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

- 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

## 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.

[Add notification](#)

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

Click on next again. Give tags Optional but I recommend to give it for easiness.

AWS Search [Alt+S]

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

### Add tags - optional [Info](#)

Add tags to help you search, filter, and track your Auto Scaling group across AWS. You can also choose to automatically add these tags to instances when they are launched.

**Tags (1)**

Key	Value - optional	Tag new instances
name	autoScaling	<input checked="" type="checkbox"/>

[Add tag](#) [Remove](#)

49 remaining

[Cancel](#) [Previous](#) [Next](#)

In last step click on create auto scaling group.

AWS Search [Alt+S]

EC2 > Auto Scaling groups > Create Auto Scaling group

### Additional settings

Instance scale-in protection Disabled	Monitoring Disabled	Default instance warmup Disabled
--	------------------------	-------------------------------------

### Capacity Reservation preference

Preference Default	Capacity Reservation IDs -	Resource Groups -
-----------------------	-------------------------------	----------------------

**Step 5: Add notifications** [Edit](#)

Notifications  
No notifications

**Step 6: Add tags** [Edit](#)

### Tags (1)

Key	Value	Tag new instances
name	autoScaling	Yes

[Preview code](#) [Cancel](#) [Previous](#) [Create Auto Scaling group](#)

AWS Search [Alt+S]

EC2 > Auto Scaling groups

Dashboard < EC2 Global View Events

**Instances**  
Instances  
Instance Types  
Launch Templates  
Spot Requests  
Savings Plans  
Reserved Instances  
Dedicated Hosts  
Capacity Reservations

**Images**  
AMIs  
AMI Catalog

**Elastic Block Store**  
Volumes  
Snapshots  
Lifecycle Manager

**Network & Security**  
Security Groups  
Elastic IPs

**Auto Scaling groups (1/1) Info** [Launch configurations](#) [Launch templates](#) [Actions](#) [Create Auto Scaling group](#)

Search your Auto Scaling groups

<input checked="" type="checkbox"/>	Name	Launch template/configuration	Instances	Status	Desired capacity	Min	Max	Availability...
<input checked="" type="checkbox"/>	AutoScaling	TemplateForAutoScaling   Version Default	0	Updating capacity	1	1	1	ap-south-1c...

### Auto Scaling group: AutoScaling

Details Integrations - new Automatic scaling **Instance management** Instance refresh Activity Monitoring

### Instances (1)

Filter instances

<input type="checkbox"/>	Instance ID	Lifecycle	Instance type	Weighted ca...	Launch temp...	Availability ...	Health status	Protected from
<input type="checkbox"/>	i-0cf5b36700735198b	InService	t2.micro	-	TemplateForAutoScal	ap-south-1a	Healthy	

### Lifecycle hooks (0) Info

Filter lifecycle hooks

[Actions](#) [Create lifecycle hook](#)

