

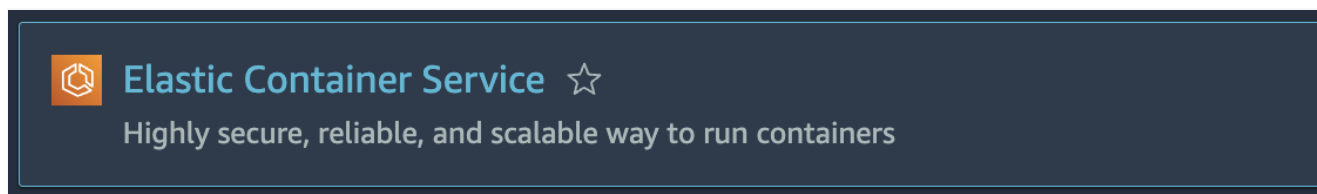
## Introduction

An Amazon Elastic Container Service (ECS) cluster is a logical grouping of container instances where you can define task definitions to execute tasks. A task is the instantiation of a task definition on a container instance within an ECS cluster. You can customize several features of a cluster such as EC2 instance type, storage allocation, and networking specifics.

In this lab step, you will create an ECS cluster using a pre-configured Amazon EC2 Auto Scaling group (ASG), and a pre-configured Amazon Virtual Private Cloud (VPC) network.

## Instructions

1. In the search bar at the top of the AWS Management Console, enter *ECS*, and under **Services**, click the **Elastic Container Service** result:



2. To navigate to the clusters page, in the left-hand menu, click **Clusters**:



You will see an empty clusters list, no ECS clusters have been created yet in this lab.

3. To begin creating a new ECS cluster, in the top-right, click **Create cluster**:

An orange rectangular button with the text "Create cluster" in white.

A form page titled **Create cluster** will load.

4. To name your cluster, in the **Cluster configuration** section, in the **Cluster name** textbox, enter *ecslab-cluster*:

A light blue rectangular input field with the text "Cluster name" in black at the top.

Cluster name

ecslab-cluster

*Warning:* Ensure that your cluster name matches **ecslab-cluster** (all lowercase) exactly. The pre-configured Amazon EC2 Auto Scaling group has been configured to launch container instances that will join an ECS cluster with this name. If the name does not match, later lab steps will not work correctly.

5. In the rest of the form, configure the following:

- **Networking:**
  - **VPC:** Select the VPC named **cloudacademylabs**
  - **Subnets:** Select both available subnets

**VPC**  
Use a VPC with public and private subnets. By default, VPCs are created for your AWS account. To create a new VPC, go to the [VPC Console](#).

vpc-094ed95e04c5fadca  
cloudacademylabs ▼

**Subnets**  
Select the subnets where your tasks run. We recommend that you use three subnets for production.

Choose subnets ▼

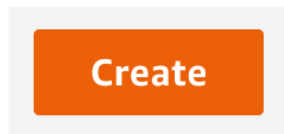
subnet-03af16e824e766c5f ✕  
public | us-west-2a

subnet-0aa5c21aca69c7d94 ✕  
public | us-west-2b

**Note:** Ensure all non-greyed-out options in the **Infrastructure** section are unchecked, you will add infrastructure to the cluster later in this step.

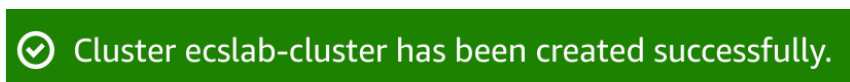
You have specified that the cluster will be deployed into a VPC with two public subnets across two availability zones (AZs). Using two AZs enables deployments in the cluster to be highly available (HA).

6. To finish creating your ECS cluster, at the bottom of the page, click **Create**:



You will be returned to the **All Clusters** page, and you will see a blue notification that your cluster is being created.

The cluster creation process can take up to three minutes to complete. Once complete, you will see a green notification:



*Note:* If you don't see your cluster listed after the success notification, click the refresh icon:



The ECS cluster creation process is a fairly complex operation that utilizes the following Amazon services:

- Elastic Compute Cloud (EC2): The cluster is powered by EC2 instances and networking features
- Identity and Access Management (IAM): The EC2 instances require roles with defined policies to securely interact with other services
- AWS CloudFormation: CloudFormation is the mechanism that deploys and manages your cluster as configured

7. To view details of your cluster, in the clusters list, click **ecslab-cluster**:

[ecslab-cluster](#)

A page showing a **Cluster overview** will load.

8. To see infrastructure information about this cluster, in the row of tabs under the **Cluster overview**, click **Infrastructure**:



9. To create an Amazon ECS capacity provider for an Amazon EC2 Autoscaling Group, on the right-hand side of the **Capacity providers** section, click **Create**:

An orange rectangular button with the word "Create" in white text.

The **Create capacity providers** form will load.

10. To configure a capacity provider, enter and select the following:

- **Basic details:**
  - **Capacity provider name:** Enter *lab-capacity-provider*
- **Auto Scaling group:**
  - Select the existing ASG whose name begins with **cloudacademylabs-EcsAutoScalingGroup-**
- **Scaling policies** (Expand this)
  - Uncheck **Turn on managed scaling**

Cluster name

ecslab-cluster

Capacity provider name

lab-capacity-provider

Up to 255 characters are allowed, including letters (upper and lowercase), numbers, underscores, and hyphens. The name cannot be prefixed with "aws", "ecs", or "fargate"

## Auto Scaling group [Info](#)

### Use an existing Auto Scaling group

You must create an Auto Scaling group before creating a capacity provider.

cloudacademylabs-EcsAutoScalingGroup-7671LJDVQJBU

### ▼ Scaling policies - *optional* [Info](#)

Use the Service metrics and CloudWatch metrics to scale in your service (run fewer tasks) to reduce costs.

☐ Turn on managed scaling

Use managed scaling to have Amazon ECS manage the scaling of your service.

You have specified that the cluster will use Amazon EC2 instances in an existing Auto Scaling group to run containers on.

The Auto Scaling group has been pre-configured to launch instances that:

- Use the latest ECS-optimized Amazon Linux 2 image for 64-bit x86 for the operating system
- Will attempt to join an Amazon ECS cluster named **ecslab-cluster**
- Have an IAM instance profile attached with an IAM role appropriate for ECS container instances
- Use a pre-configured security group that allows inbound traffic on port 80 and port 8081
- Are tagged with the name **ecs-lab-instance**

The most notable aspects of the ASG are the operating system image and the user data. The image is provided by AWS and has the [Amazon ECS container agent](#) pre-installed, as well as other optimizations for running Docker containers. The user data is used to configure the ECS

container agent's configuration file (`/etc/ecs/ecs.config`). To see other available options, visit the [Bootstrapping container instances with Amazon EC2 user data page](#) of the Amazon ECS Developer Guide.

Apart from those elements, the Auto Scaling group is standard and configured similarly to an ASG that could be used to manage autoscaling for an application running on Amazon EC2 instances without using ECS.

11. To finish creating the capacity provider, at the bottom of the page, click **Create**:

A rectangular orange button with the word "Create" in white text.

You will be returned to the **Infrastructure** tab of your cluster's overview page. You will see a blue notification at the top of the page informing you that the capacity provider is being created. Momentarily, you will see a green notification that it is successful:

A green rectangular notification box with a white checkmark icon on the left and the text "Capacity Provider esclab-capacity-provider has been deployed successfully." in white.

12. Scroll down and observe the **Container instances** table:

Container instances (1) <a href="#">Info</a>				
<input type="text" value="Filter container instances by property or value"/>				
<input type="checkbox"/>	Container instance ▾	Status ▾	Type ▾	Instance ID <a href="#">↗</a> ▾
<input type="checkbox"/>	a093085635a34bb...	✔ Active	EC2	i-0011c8cd2a6a...

You will see one container instance listed.

*Note:* If you do not see a container instance, wait thirty seconds and click the refresh icon.

This Amazon EC2 instance was launched by the pre-configured Amazon EC2 Auto Scaling group that was created during lab setup. Now that you have created a cluster named **ecslab-cluster**, the instance has joined your Amazon Elastic Container Service cluster.

## Summary

In this lab step, you created a new Amazon ECS cluster, and you configured it to use a pre-existing VPC and ASG.

### VALIDATION CHECKS

#### 1 Checks

Check again 



#### Created an ECS Cluster

Check if at least one Amazon ECS cluster has been created

Amazon Elastic Container Service (ECS)