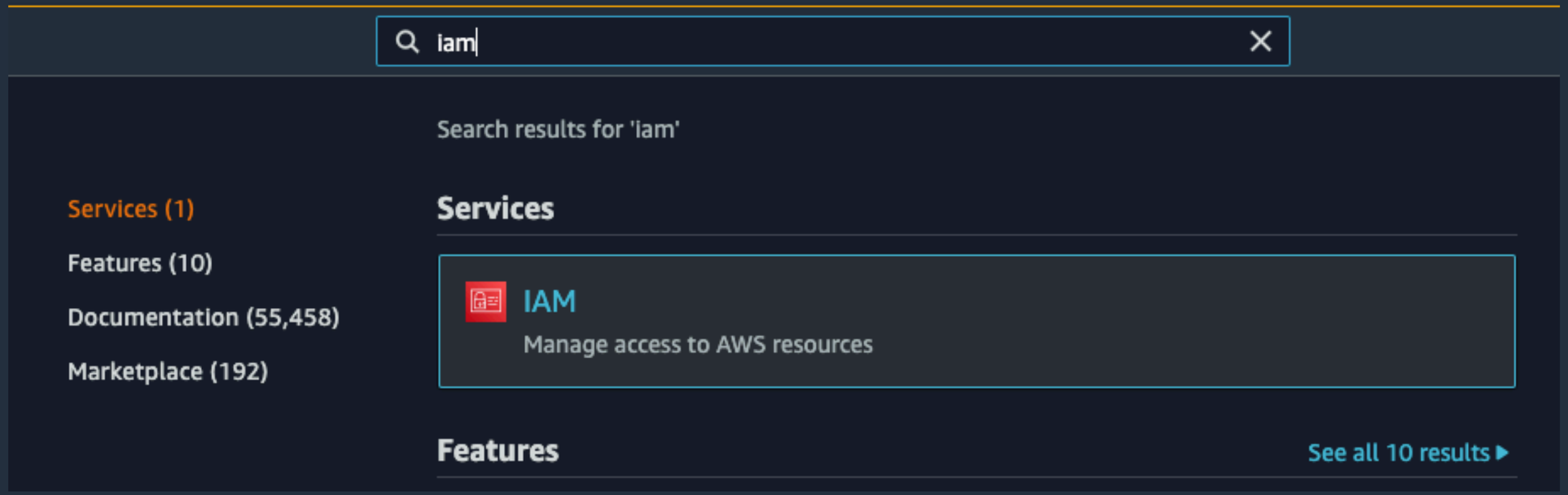
# Taming EMR on AWS

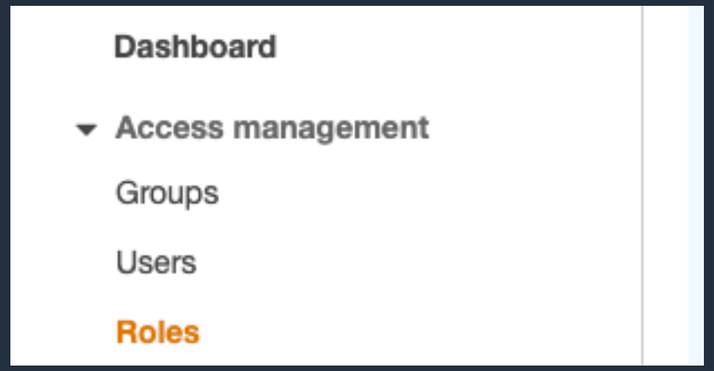
## Guidance Pack

# How to launch an EMR cluster

First we need to create a Role, open the IAM console

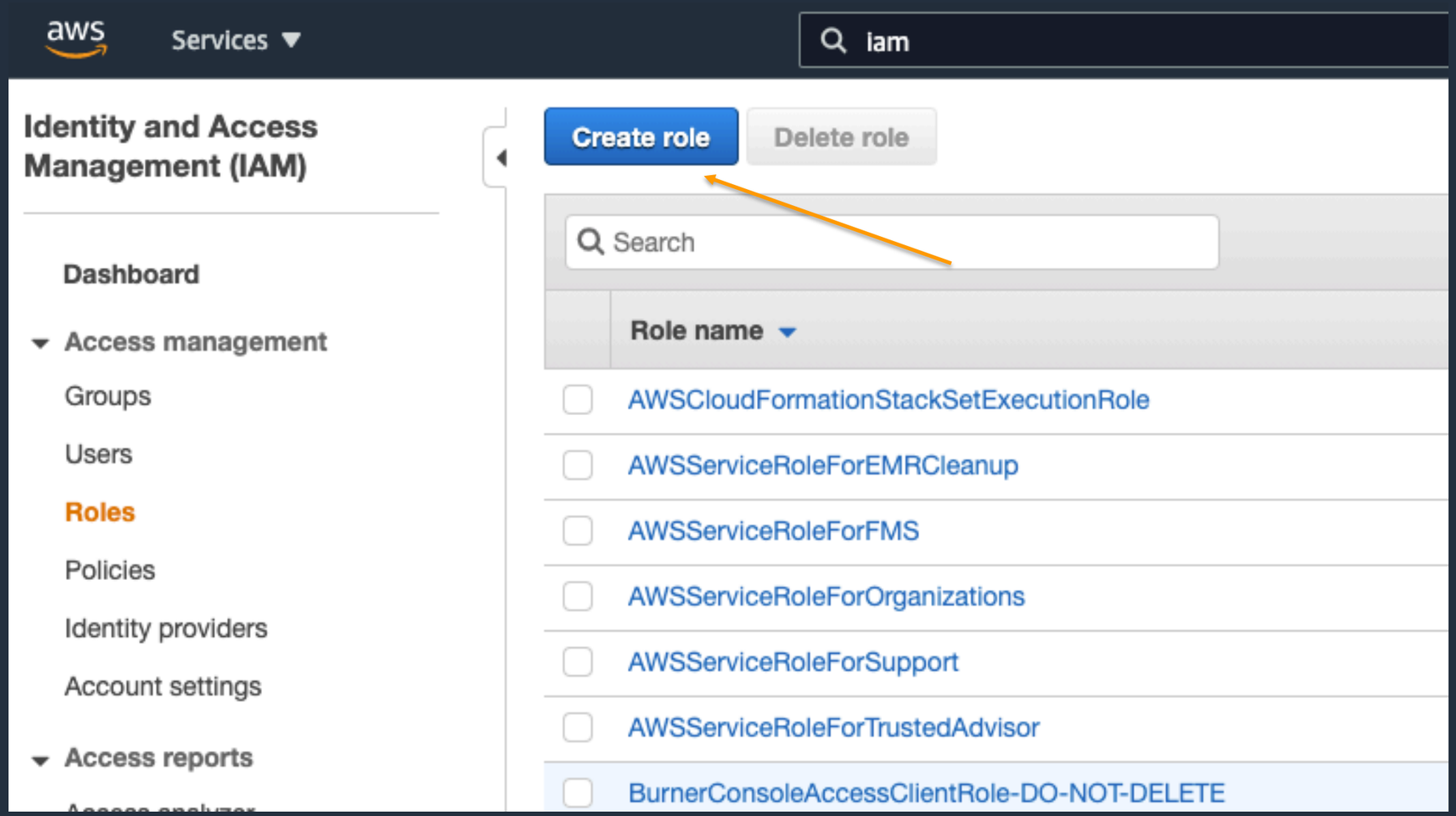



And select the Roles



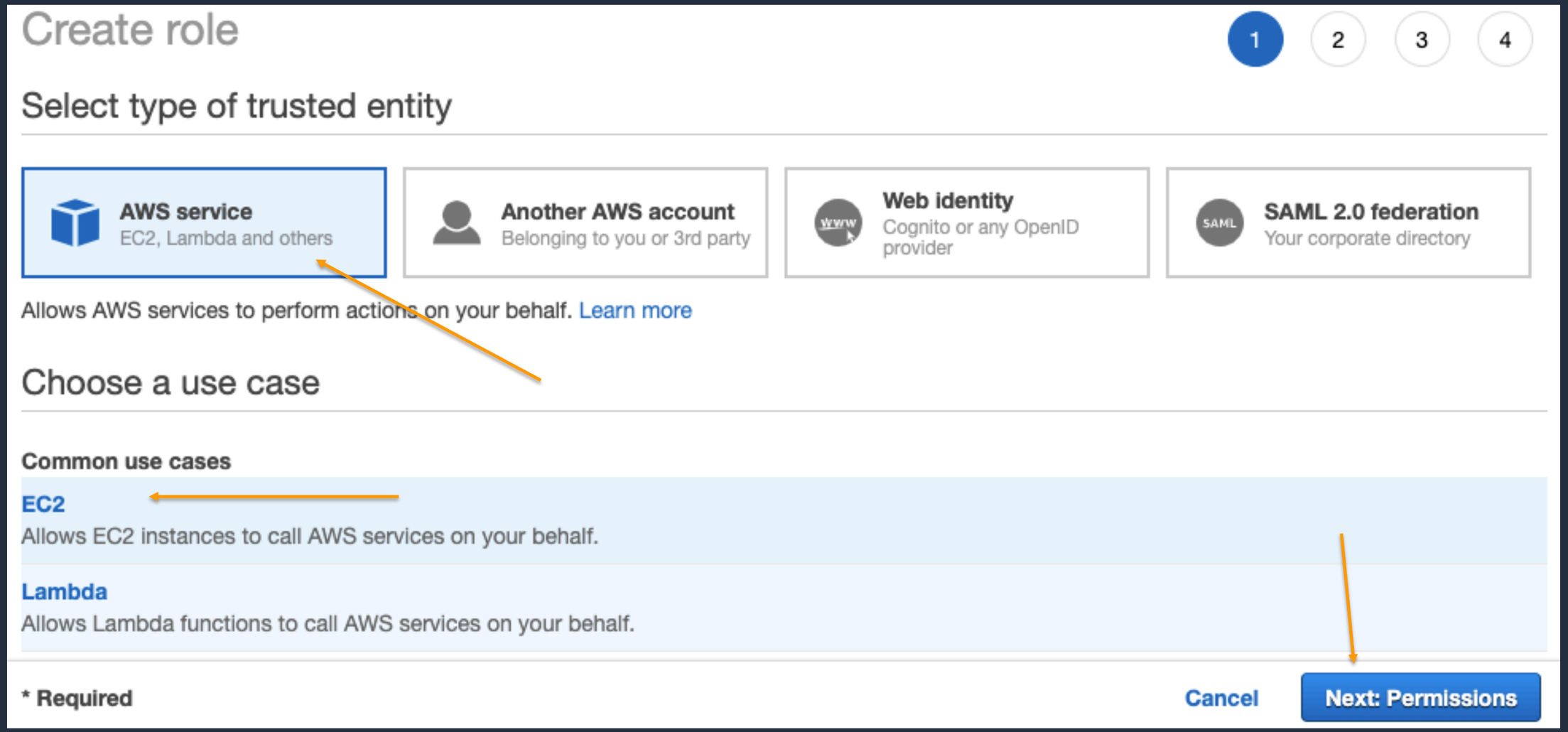
# How to launch an EMR cluster

Click in Create role



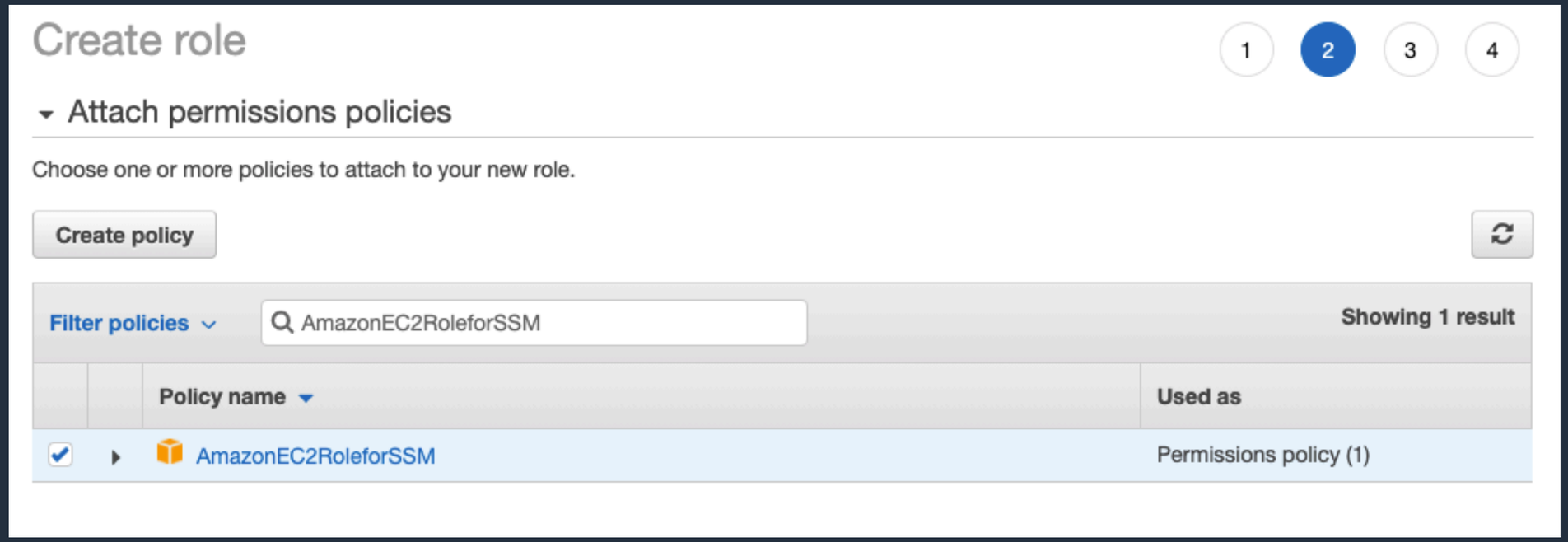
# How to launch an EMR cluster

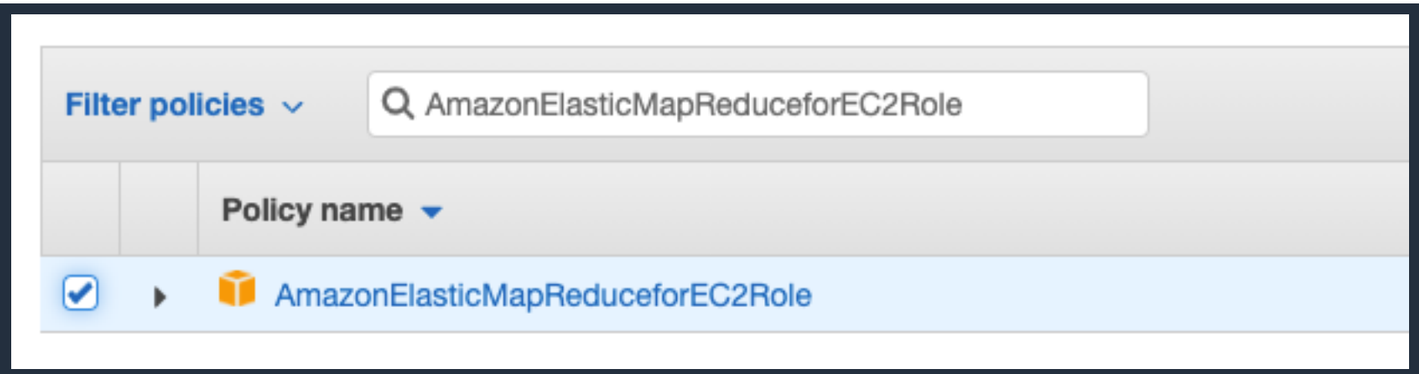
Select AW Service as trusted entity, EC2 as common use cases and click in Next



# How to launch an EMR cluster

Search and select AmazonEC2RoleforSSM, after that search and select AmazonElasticMapReduceforEC2Role. Click in Next



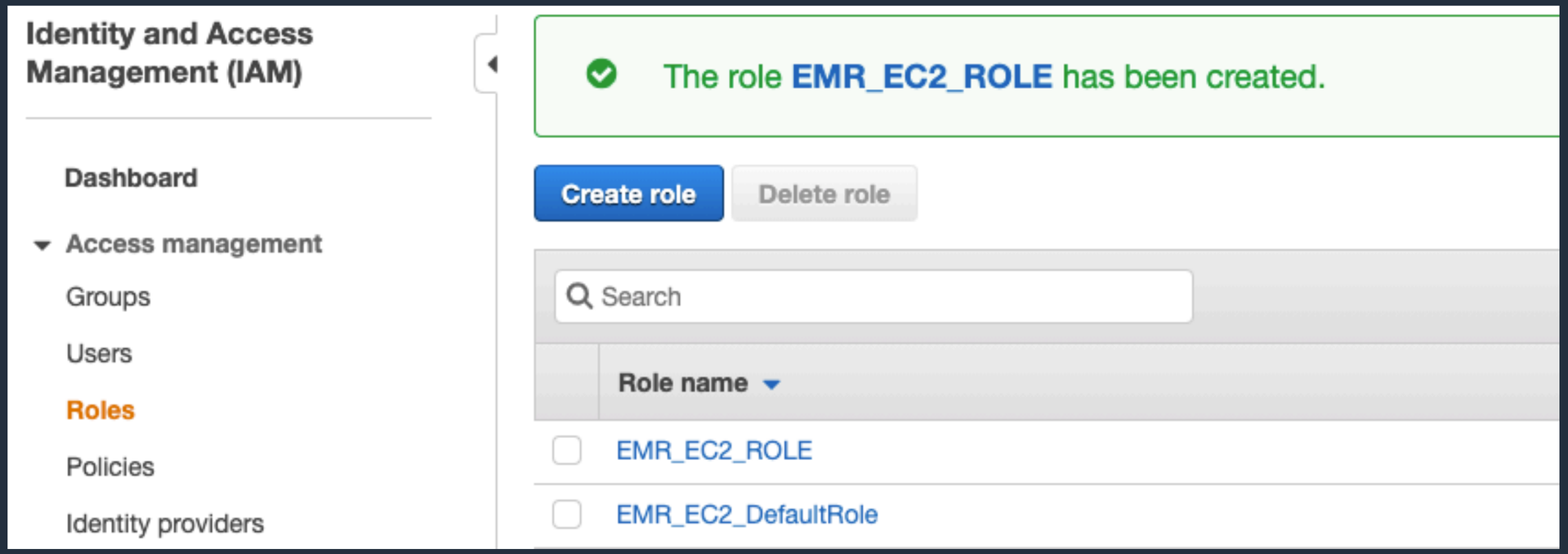
# How to launch an EMR cluster

In the Tags session click in Next, In the Review add EMR_EC2_ROLE as Role name.

Check if both policies(AmazonEC2RoleforSSM, AmazonElasticMapReduceforEC2Role) are selected and if Trusted entities is correct, after that click in Create Role

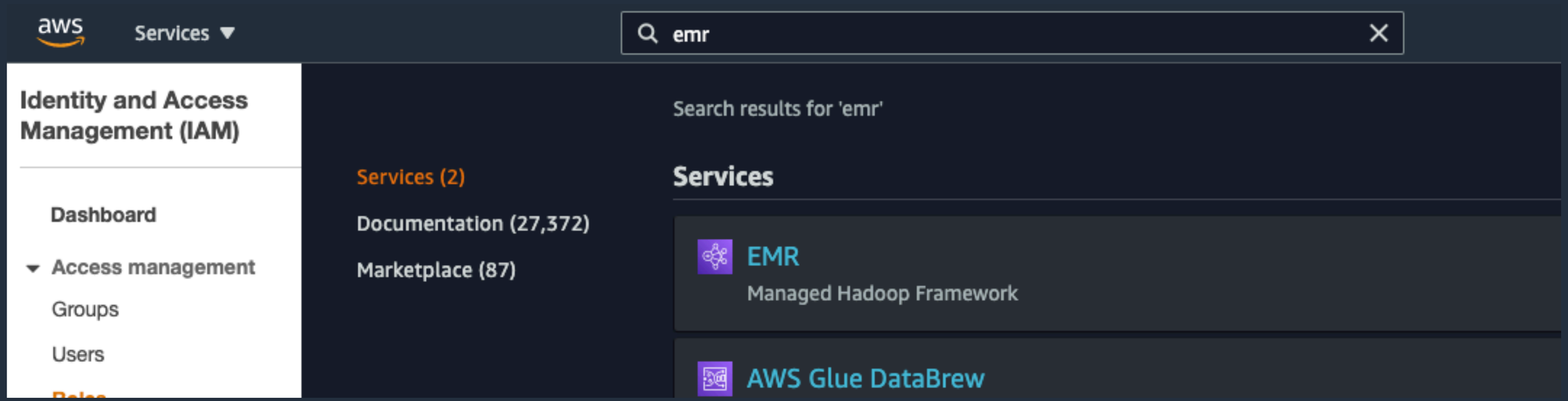
# How to launch an EMR cluster

Check if you can see the role EMR_EC2_ROLE in the Role name List



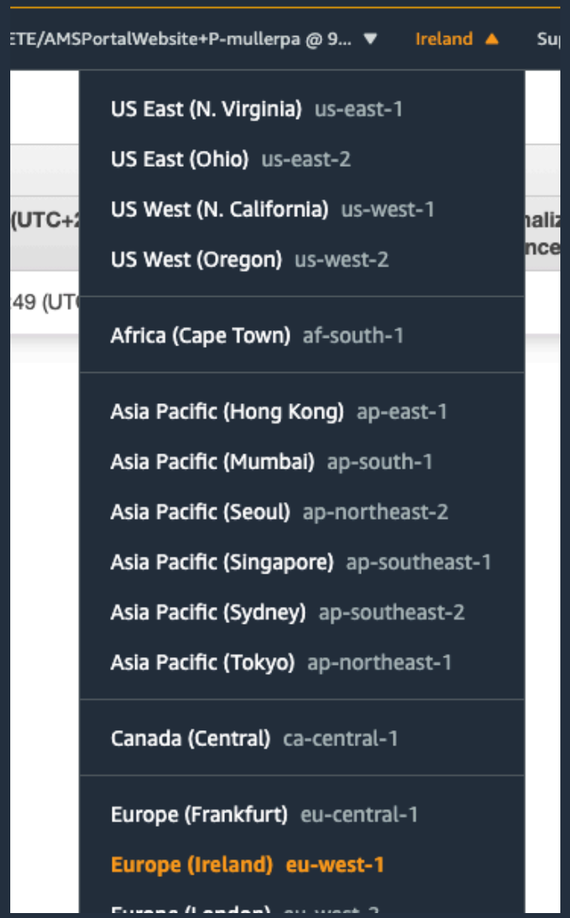
# How to launch an EMR cluster

On the AWS console, select search and select EMR from the list of services



# How to launch an EMR cluster

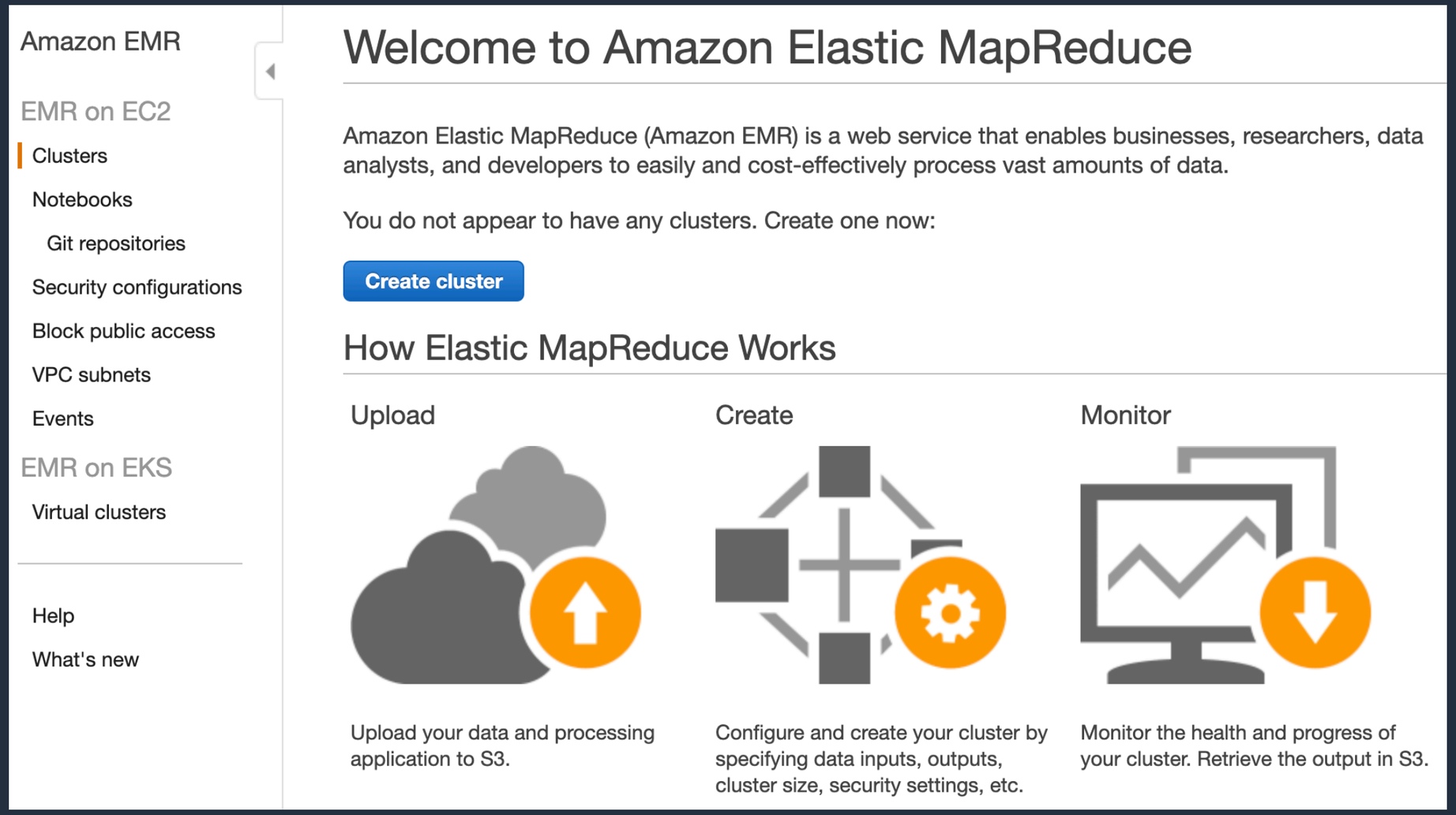
## Select the Ireland Region



# How to launch an EMR cluster

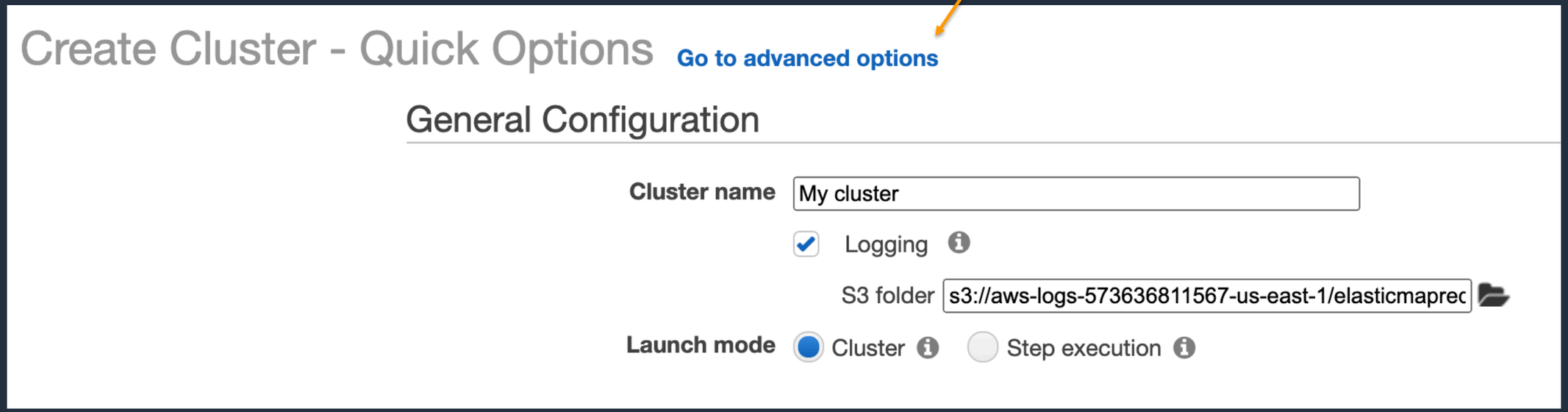
Click on

"Create cluster"



# How to launch an EMR cluster

On the create cluster page, you'll see an option to "Go to advanced options". Click on that

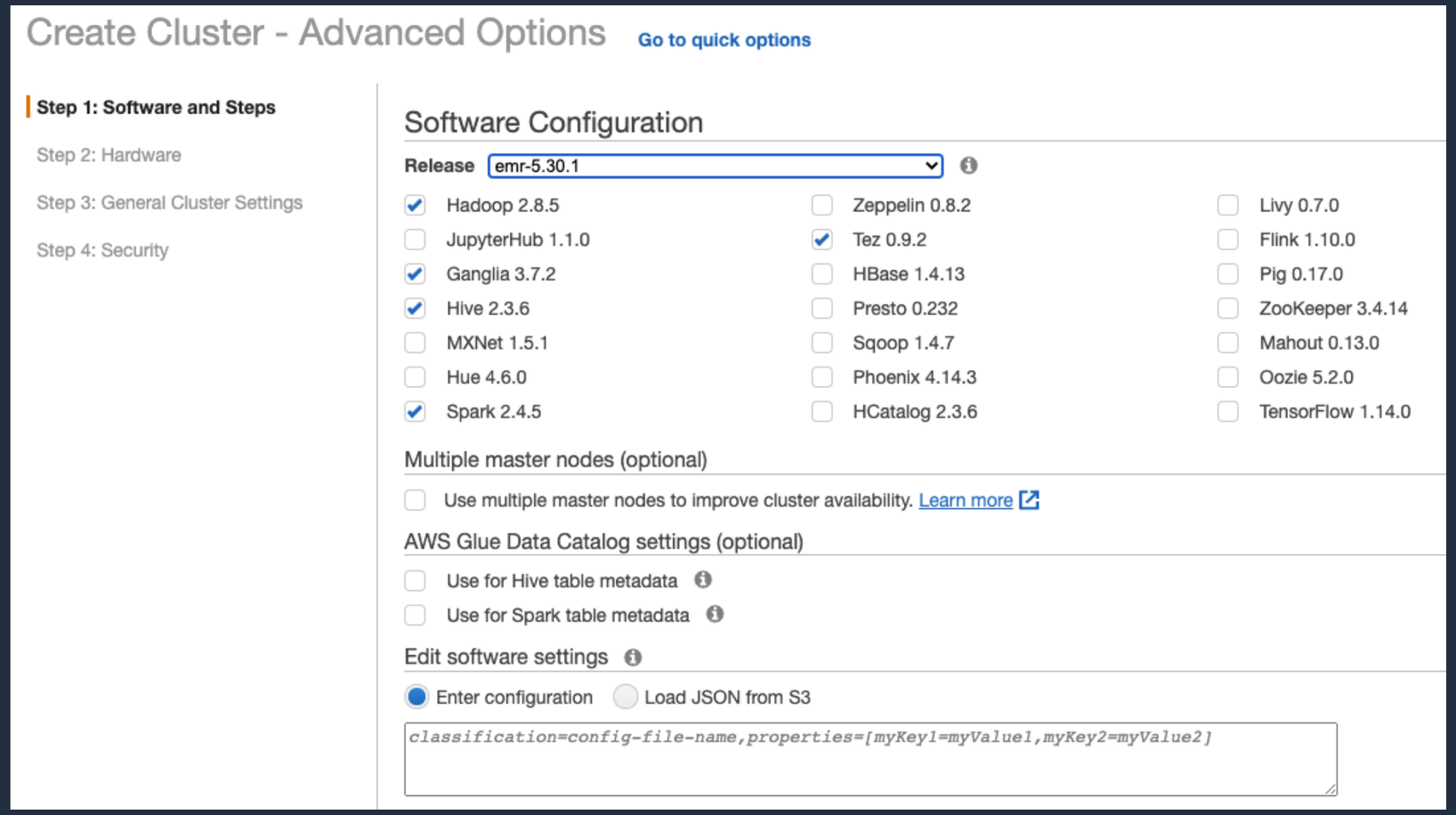


# How to launch an EMR cluster

During the Advanced Options, you will be able to select the applications you will be using during this course. Please select EMR release 5.32.0 and applications:

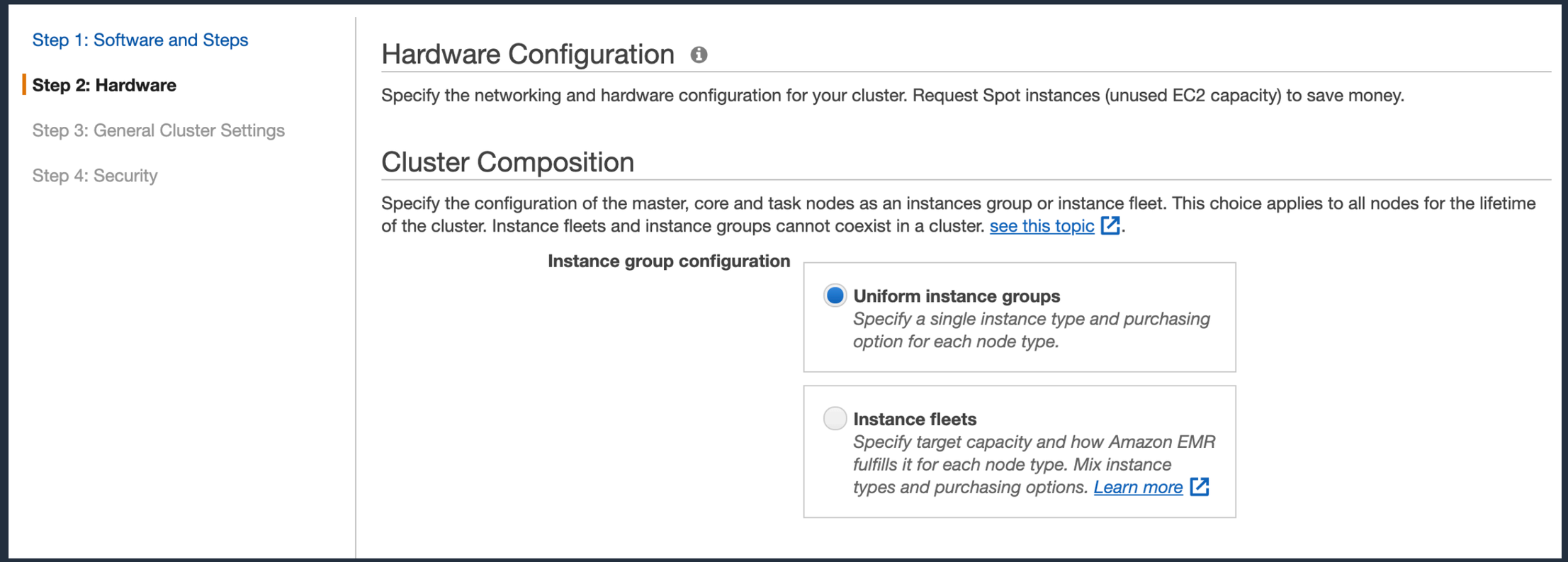
- Hadoop
- Ganglia
- Hive
- Spark
- Tez

Then click “Next”



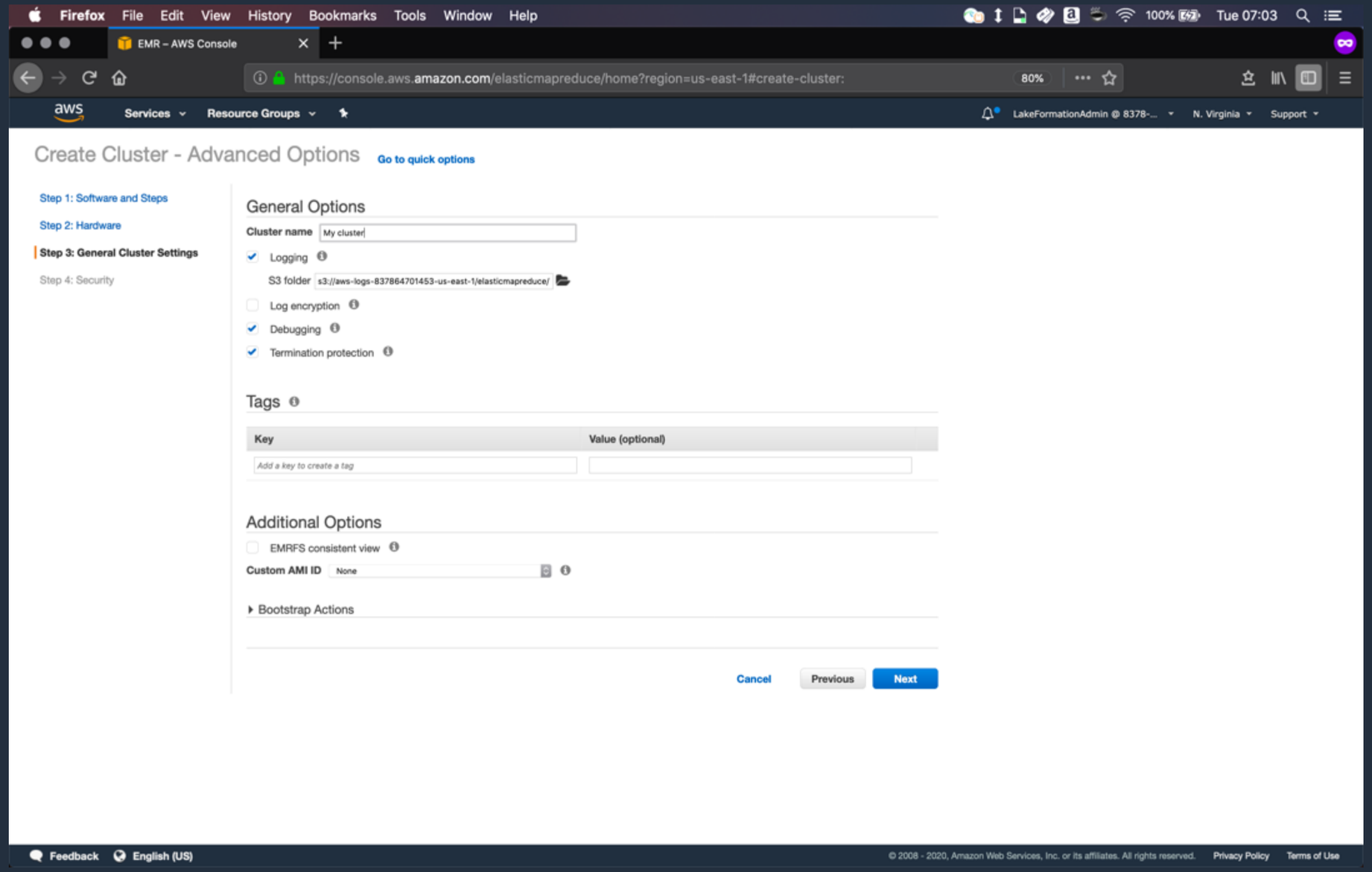
# How to launch an EMR cluster

In the Hardware Configuration just click in Next



# How to launch an EMR cluster

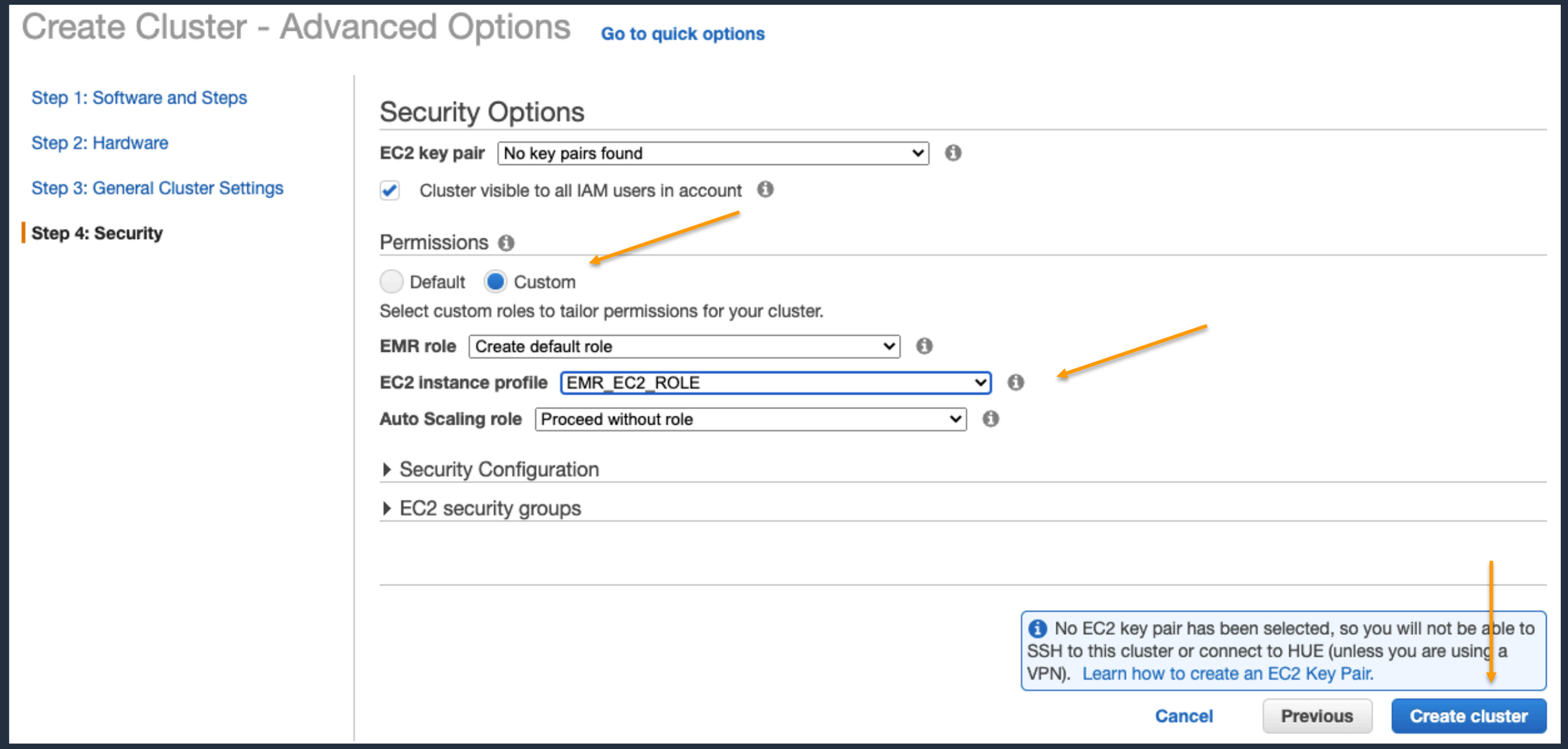
On the “General Cluster Settings” page, you can name your cluster and leave the rest of The config as it is. Click “Next”.



# How to launch an EMR cluster

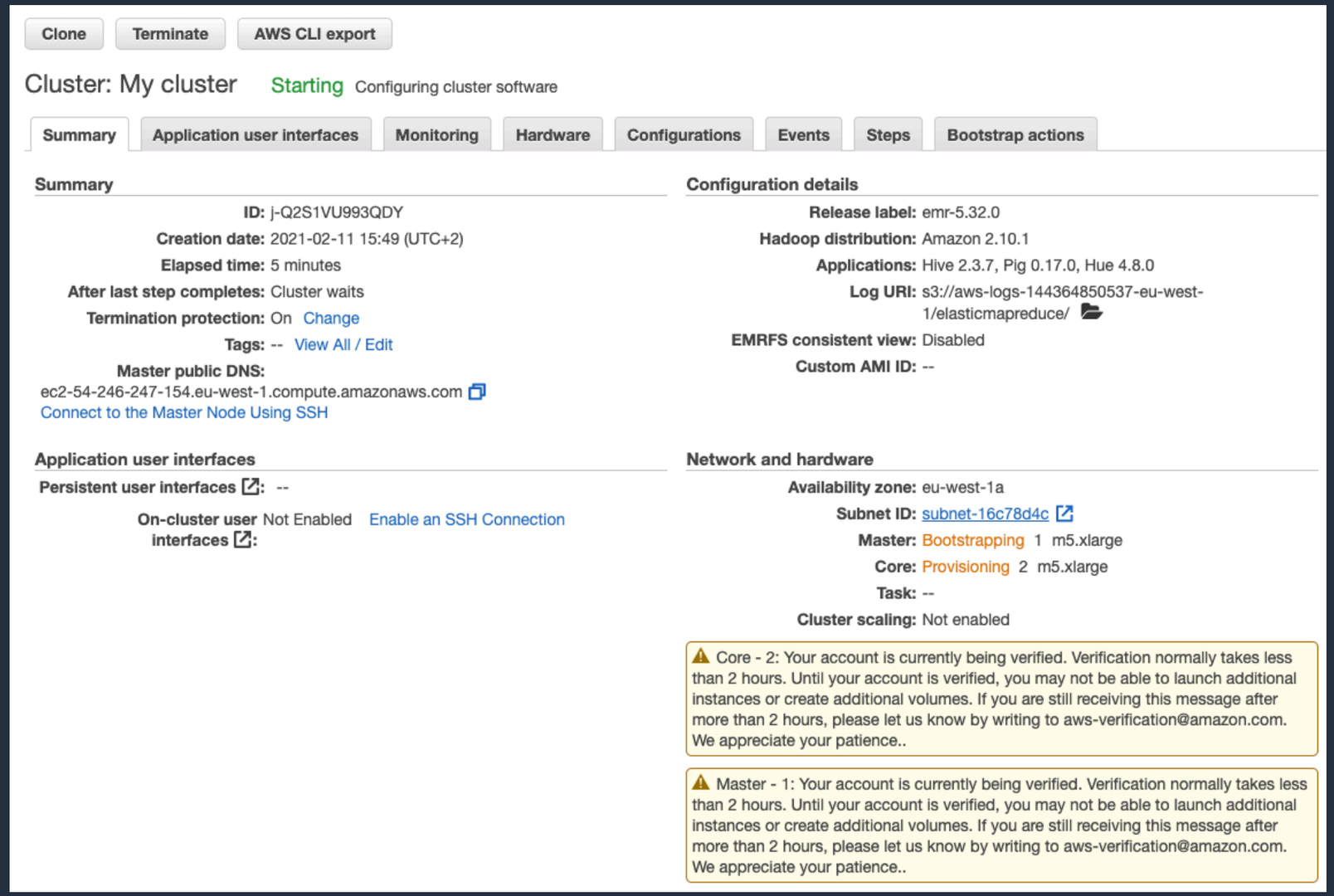
In the Security Tab, in the Permissions session, select Custom

And in the EC2 Instance profile select EMR_EC2_ROLE and click in Create Cluster



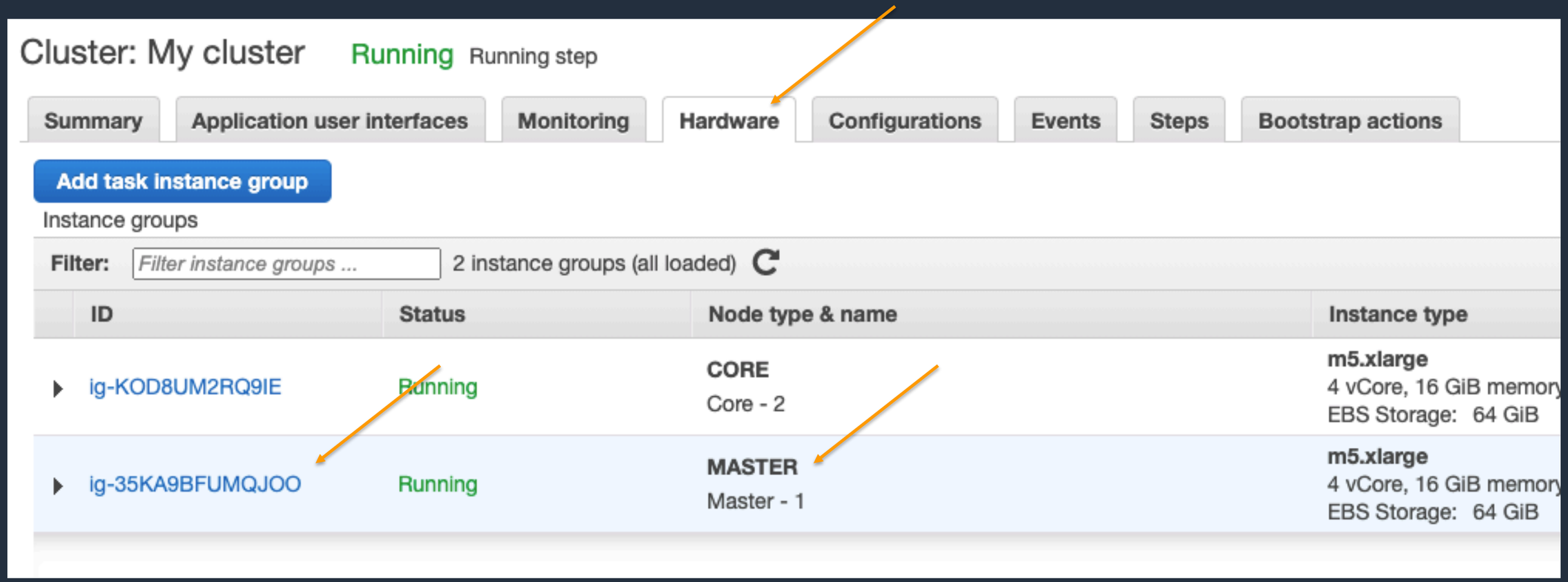
# How to launch an EMR cluster

Your cluster will now be
starting, and resources
will be provisioning.
Please allow 5-10min for your cluster to provision successfully. The status will move from starting,
to waiting.



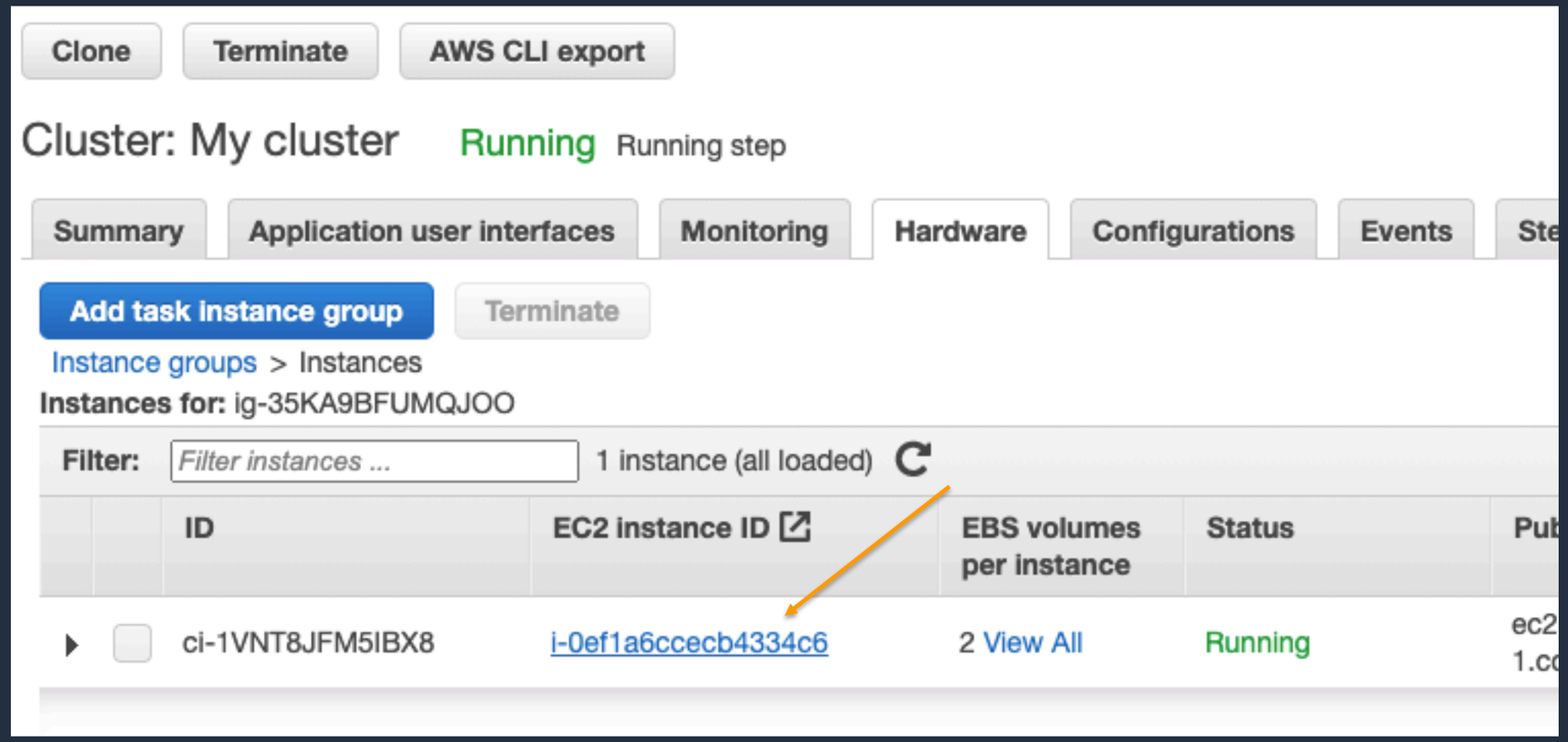
# How to launch an EMR cluster

To connect to your cluster select the Hardware Tab, and click in the ID for the Master Node



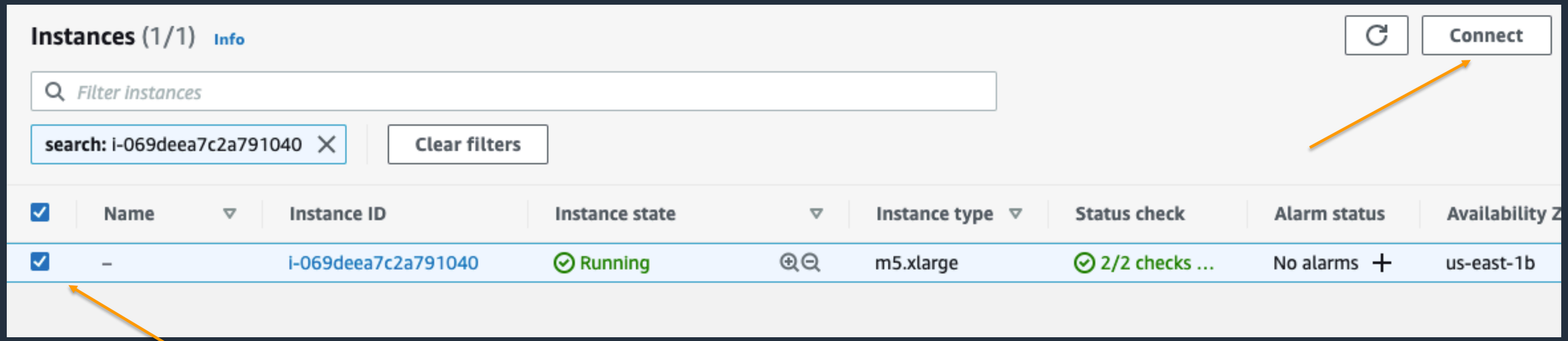
# How to launch an EMR cluster

Click in the EC2 Instance ID Link



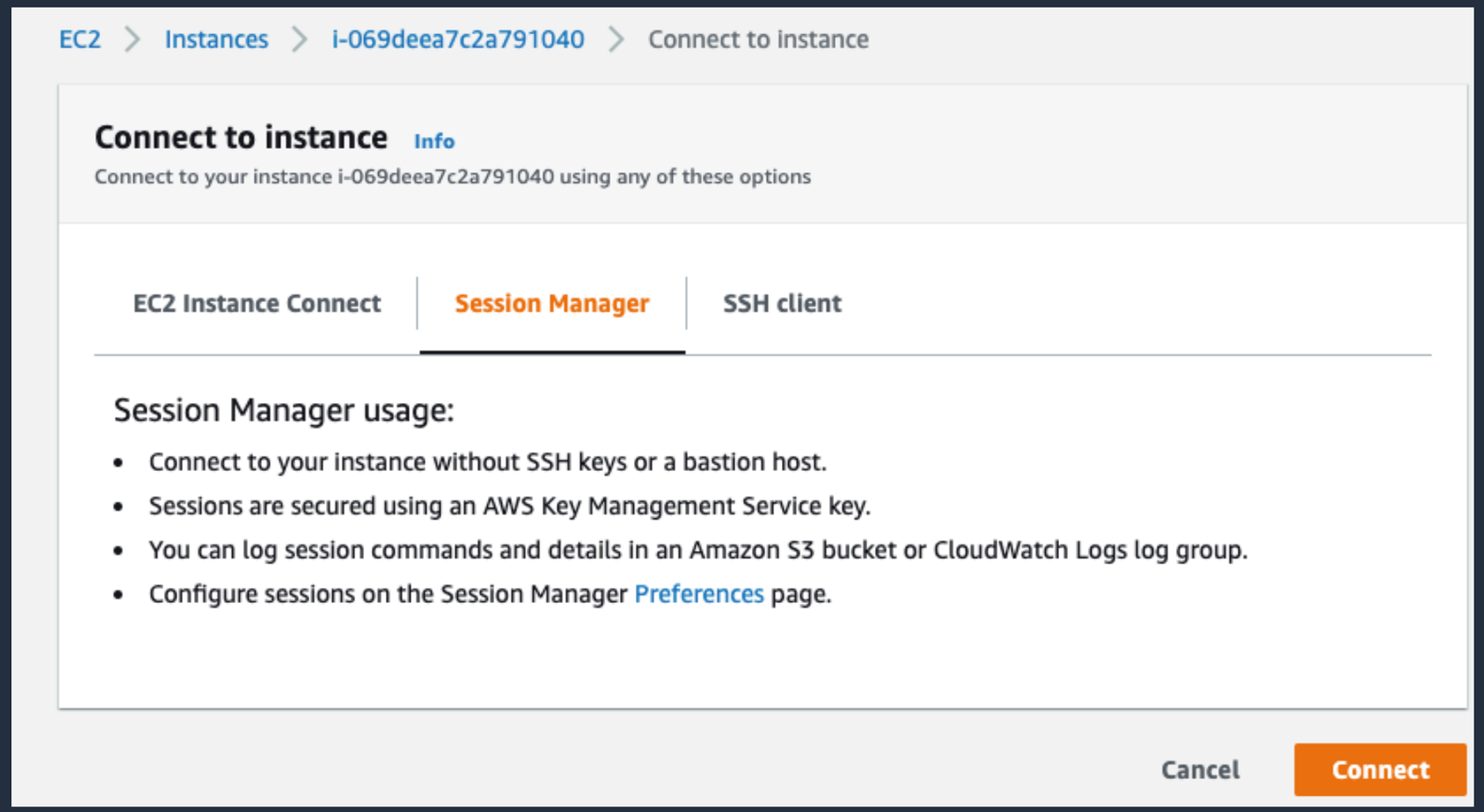
# How to launch an EMR cluster

In the EC2 Instance Page, select the instance ID and Click in Connect



# How to launch an EMR cluster

Select the Session Manager option and click in Connect



# How to launch an EMR cluster

In the terminal change run the below commands

sudo su - hadoop

You need to work with the hadoop user

```
sh-4.2$ whoami
ssm-user
sh-4.2$ sudo su - hadoop
Last login: Thu Feb 11 10:05:37 UTC 2021 on pts/0

EEEEEEEEEEEEEEEEEEEE MMMMMMMM           MMMMMMMM RRRRRRRRRRRRRRR
E::::::::::::::::::E M:::::::M         M:::::::M R::::::::::::::R
EE:::::EEEEEEEEE:::E M::::::::M       M::::::::M R:::::RRRRRR:::::R
  E::::E       EEEEE M:::::::::M     M:::::::::M RR::::R      R::::R
  E::::E             M::::::M:::M   M:::M::::::M   R:::R      R::::R
  E:::::EEEEEEEEEE   M:::::M M:::M M:::M M:::::M   R:::RRRRRR:::::R
  E::::::::::::::E   M:::::M  M:::M:::M  M:::::M   R::::::::::::RR
  E:::::EEEEEEEEEE   M:::::M   M:::::M   M:::::M   R:::RRRRRR::::R
  E::::E             M:::::M    M:::M    M:::::M   R:::R      R::::R
  E::::E       EEEEE M:::::M     MMM     M:::::M   R:::R      R::::R
EE:::::EEEEEEEE::::E M:::::M             M:::::M   R:::R      R::::R
E::::::::::::::::::E M:::::M             M:::::M RR::::R      R::::R
EEEEEEEEEEEEEEEEEEEE MMMMMMM             MMMMMMM RRRRRRR      RRRRRR

[hadoop@ip-172-31-33-10 ~]$ whoami
hadoop
[hadoop@ip-172-31-33-10 ~]$
```