Workflow for Programming Assignment -2

28 Steps

Created by

Rama Chetan Atmudi

Creation Date

November 3, 2022

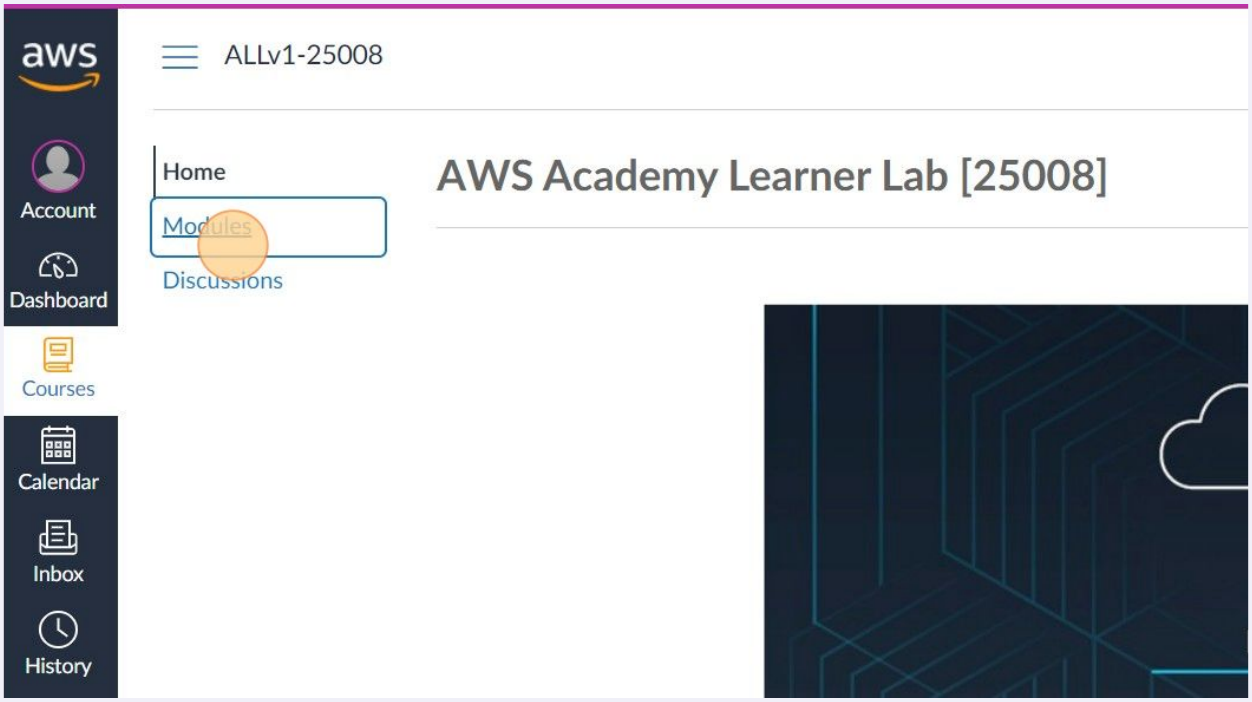
Last Updated

November 3, 2022

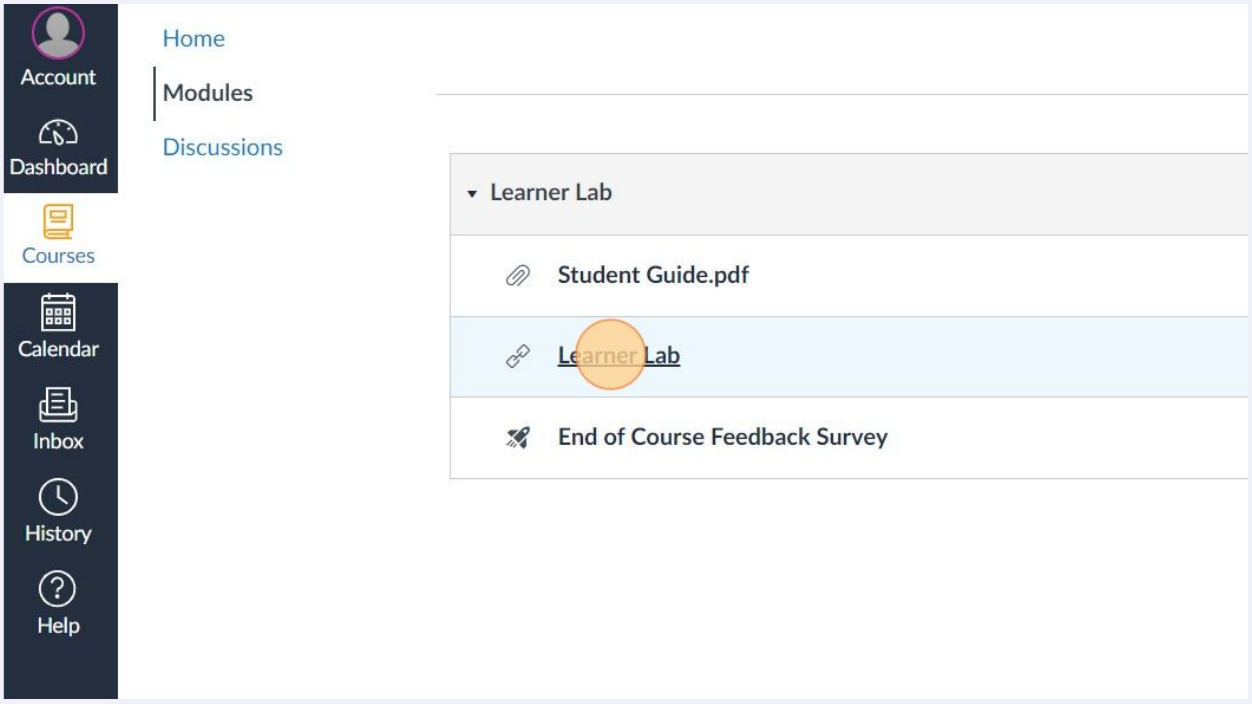
Programming Assignment -2 Workflow

1 Navigate to https://awsacademy.instructure.com/courses/25008

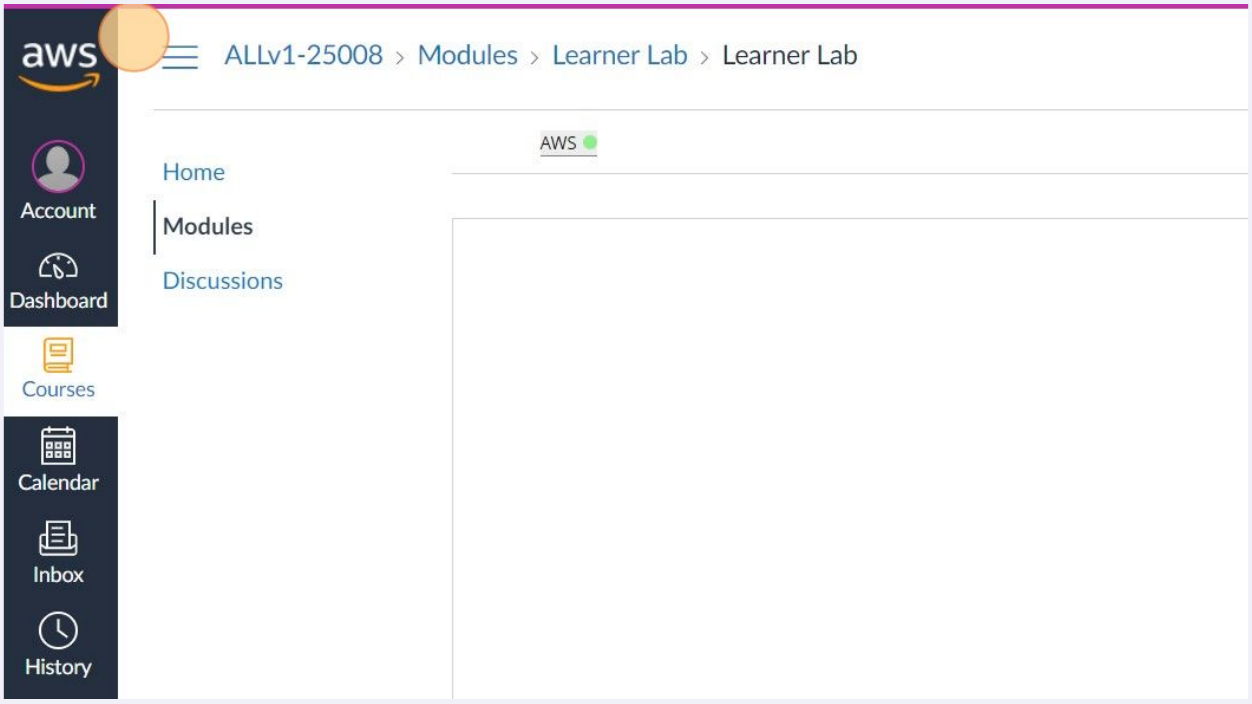
2 Click "Modules"



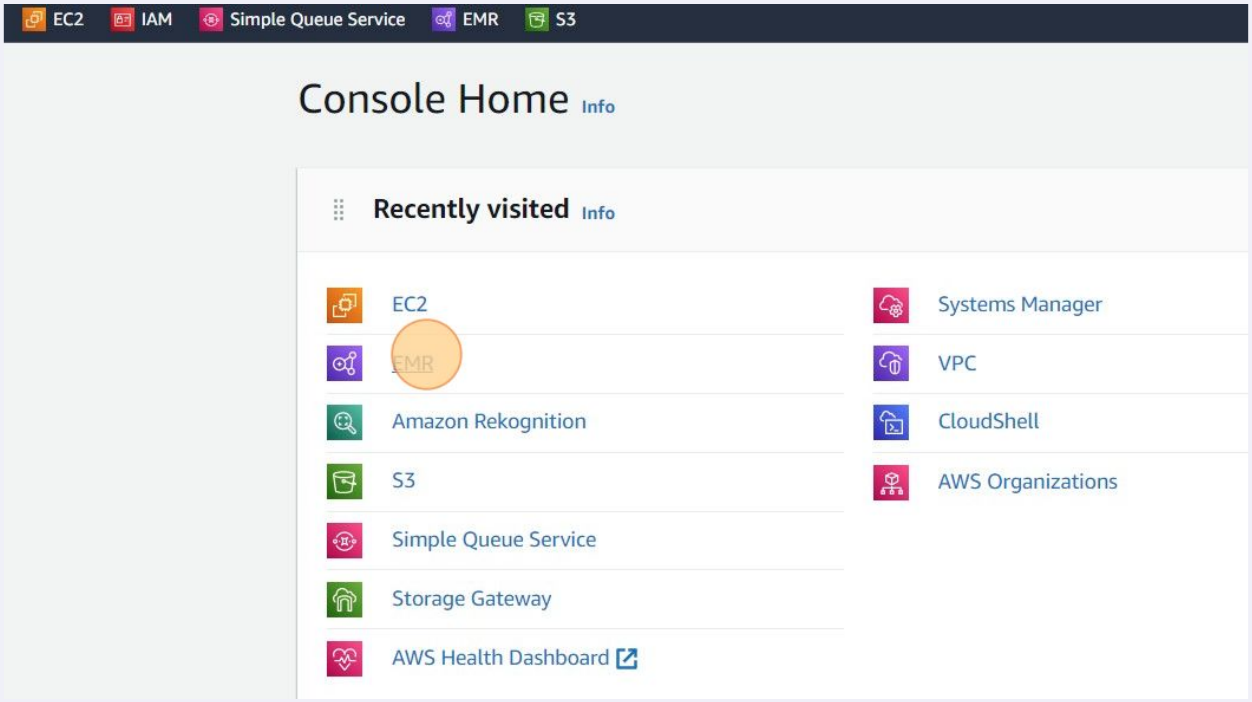
3 Click "Learner Lab"



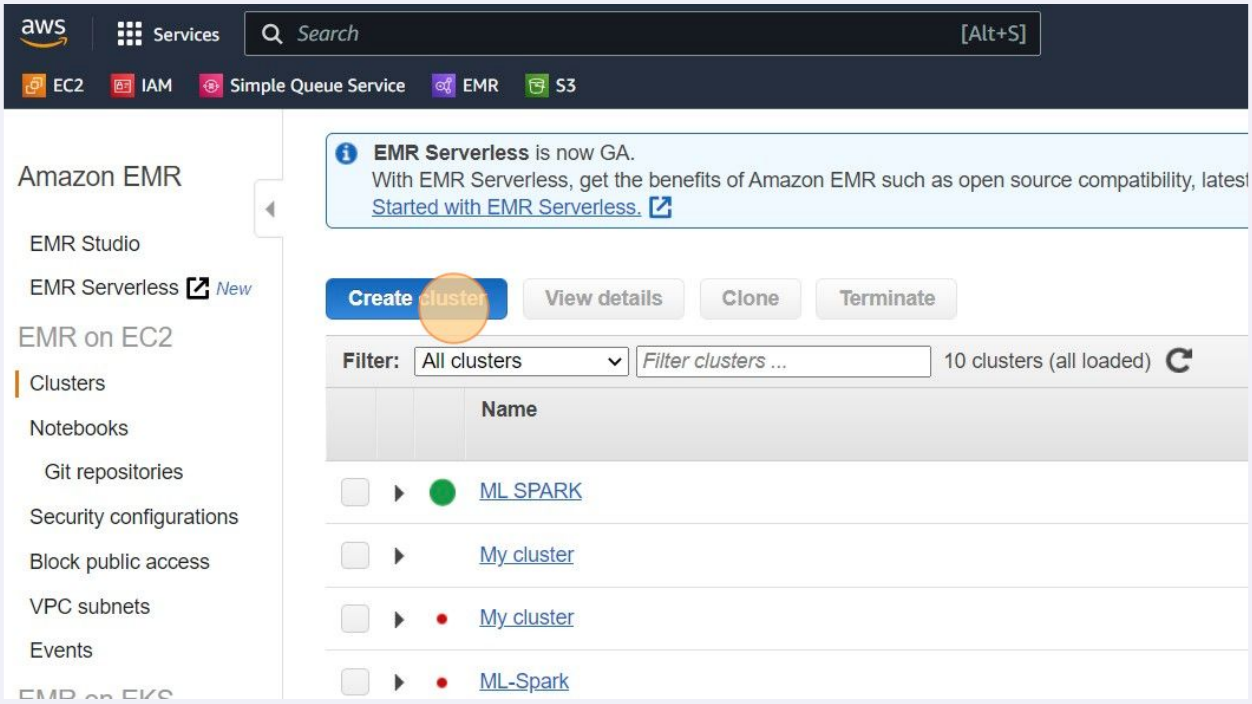
4 Start your lab session and Click "AWS"



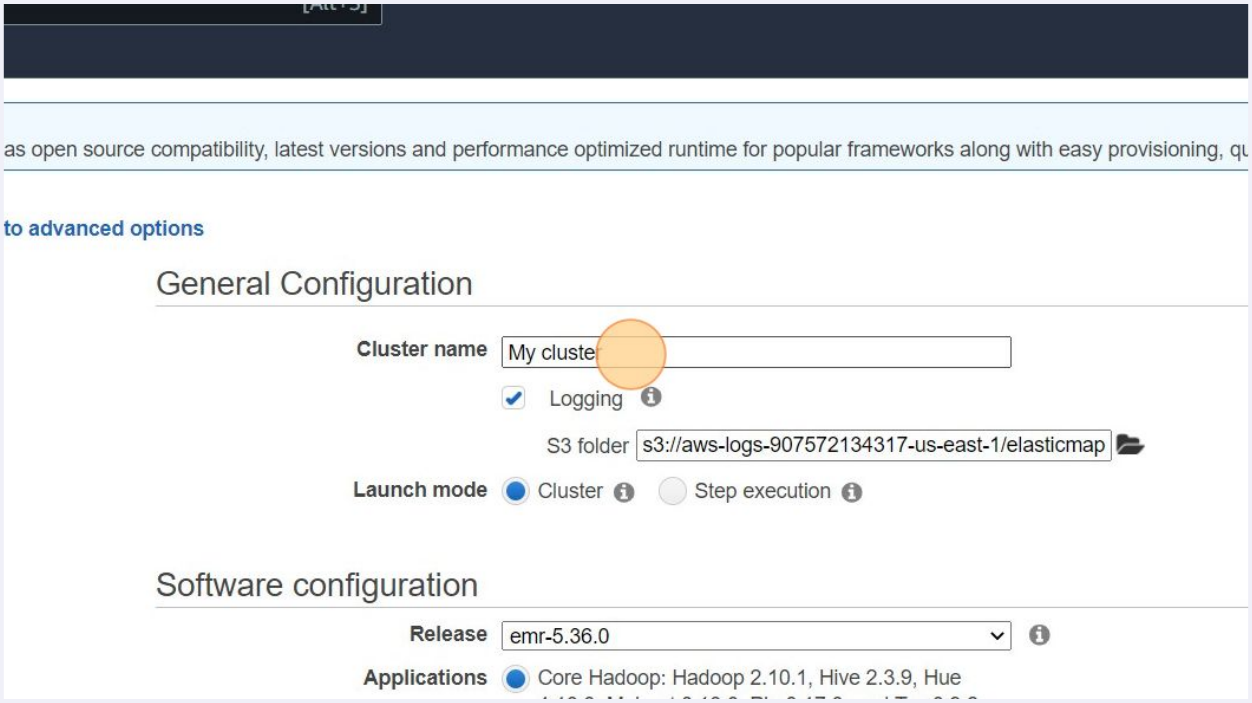
5 Search for "EMR" in the Services section of your AWS Management Console and then click "EMR." In the screenshot shown below I accessed EMR from "Recently Visited" tab.



6 Click "Create cluster"

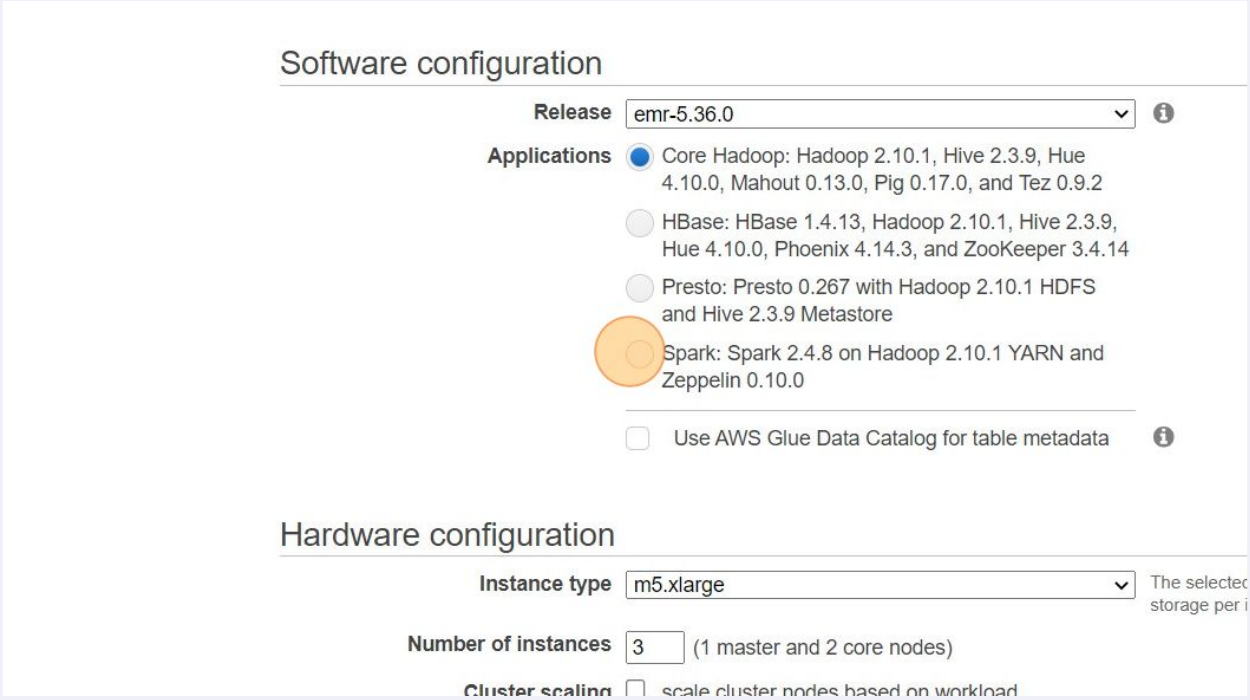


7 Name your Cluster

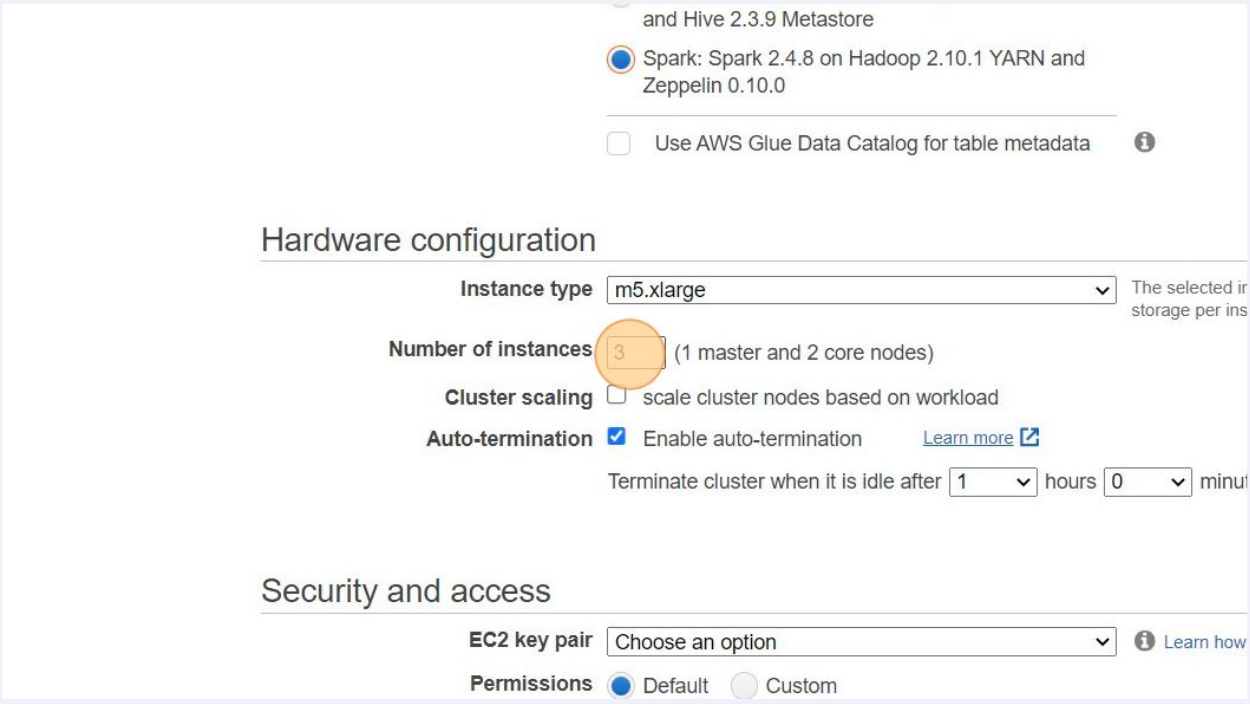


8 I named my Cluster as "ML Spark"

9 Click on "Spark" in Applications as we are dealing with Spark in this assignment.

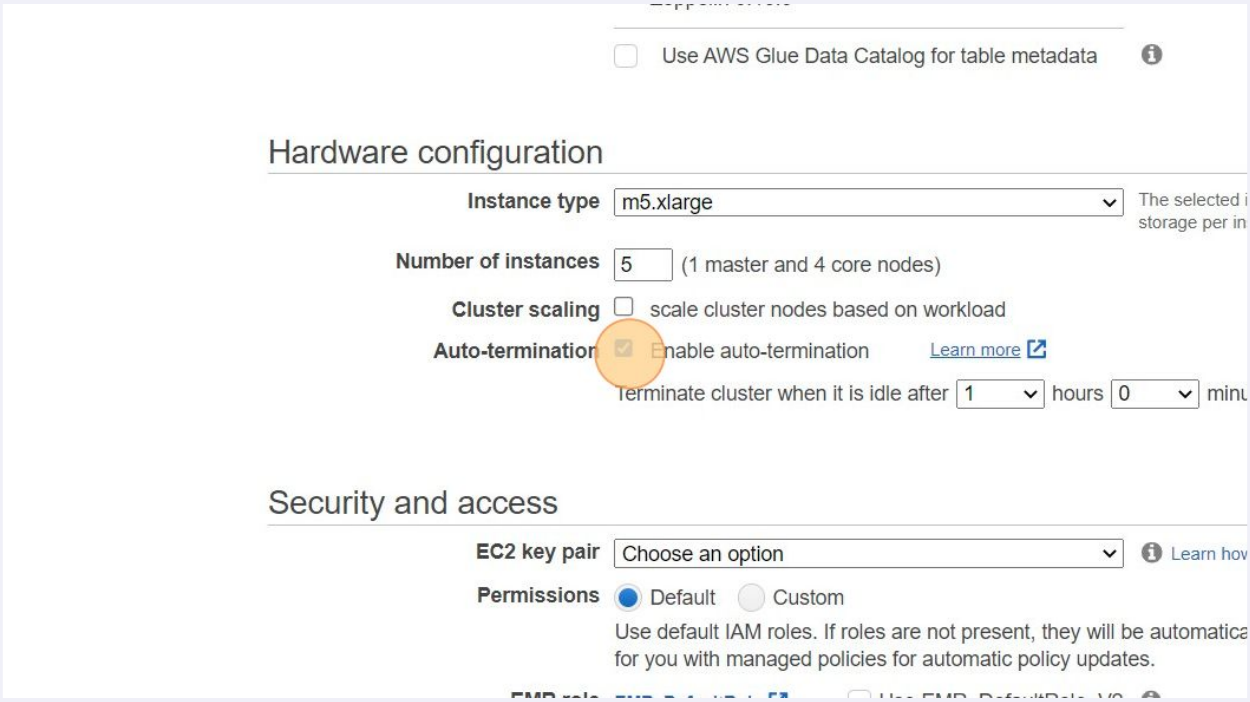


10 We need to run Spark in 4 EC2 Instances so, create 5 Instances 1 master and 4 core nodes

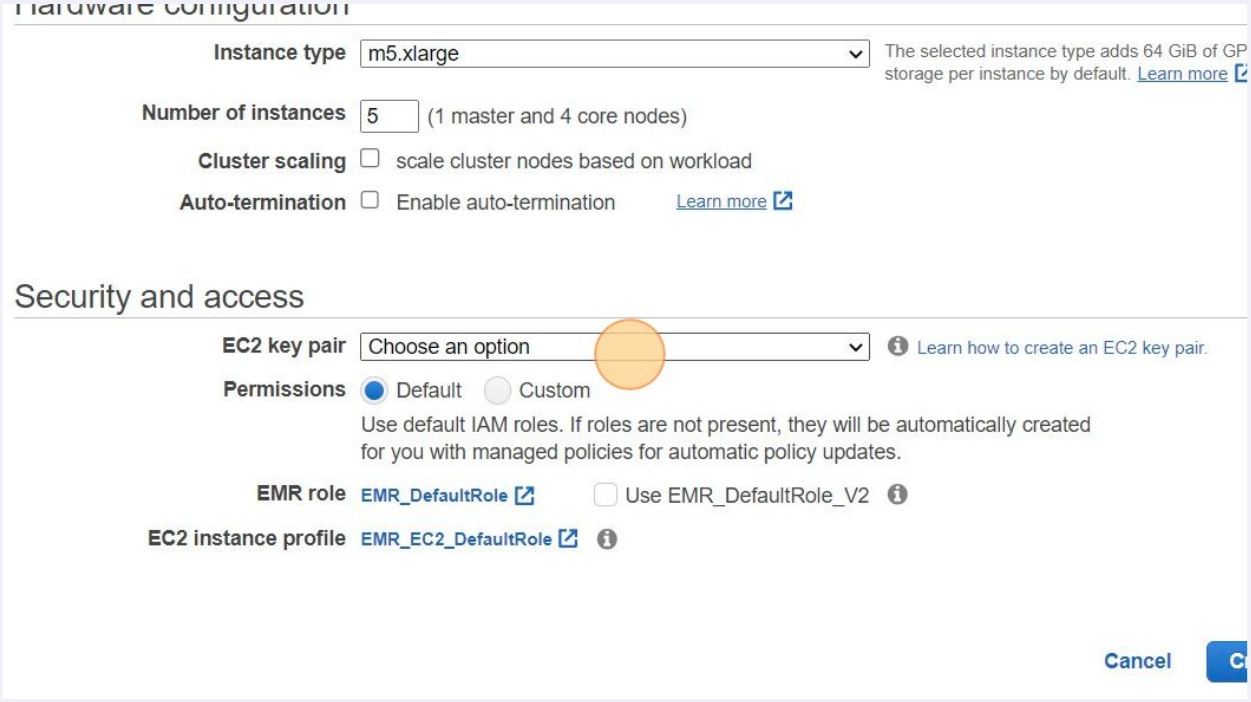


11 Enter "5" in Number of Instances

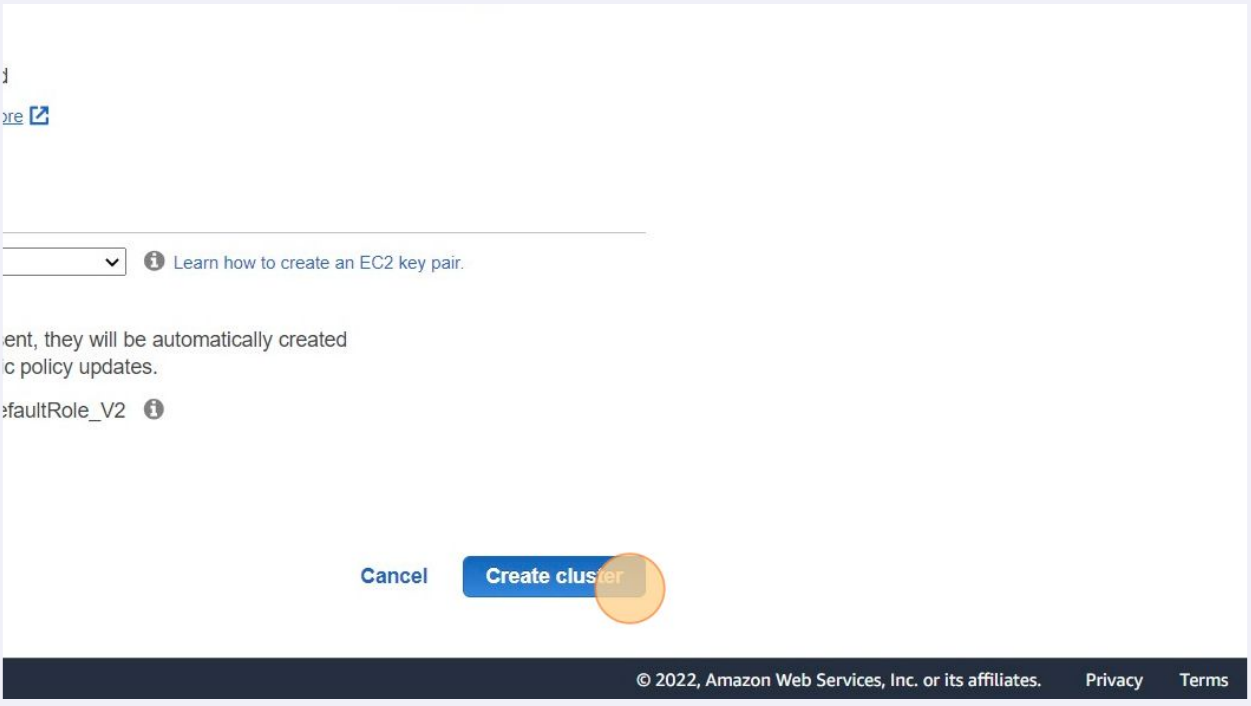
12 Disable "Auto-Termination"



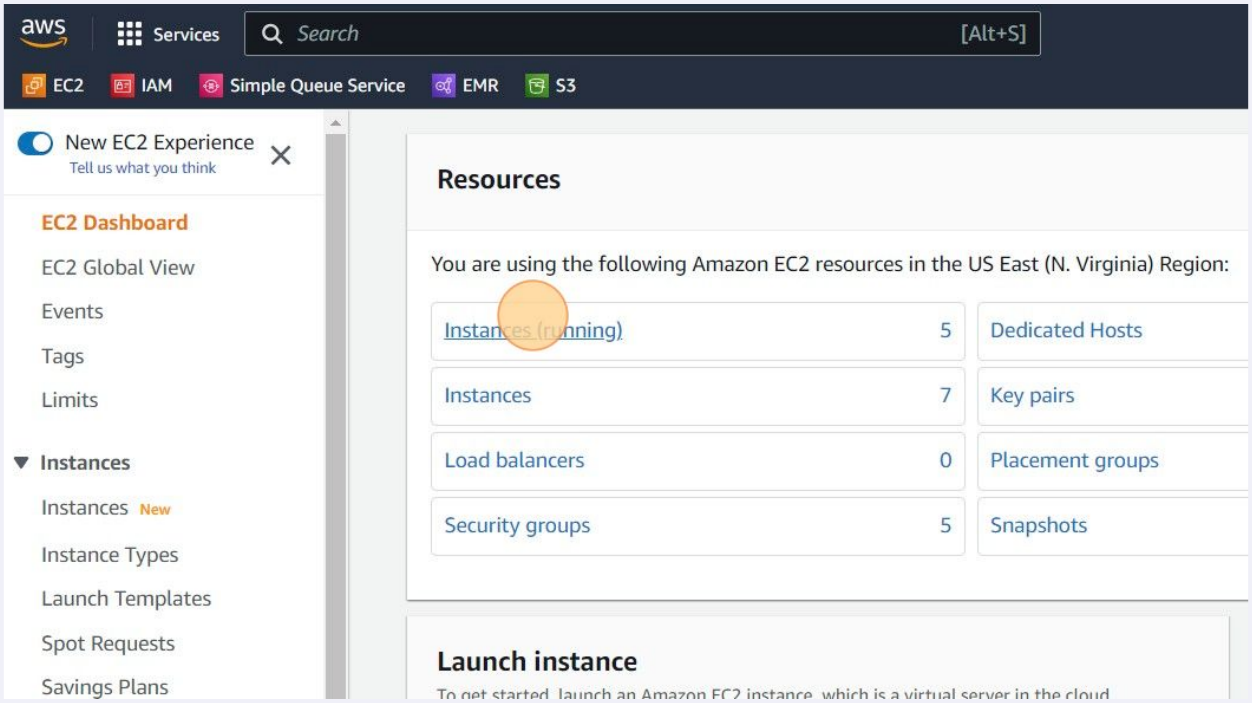
13 Create a New Key Pair for this cluster to access it.



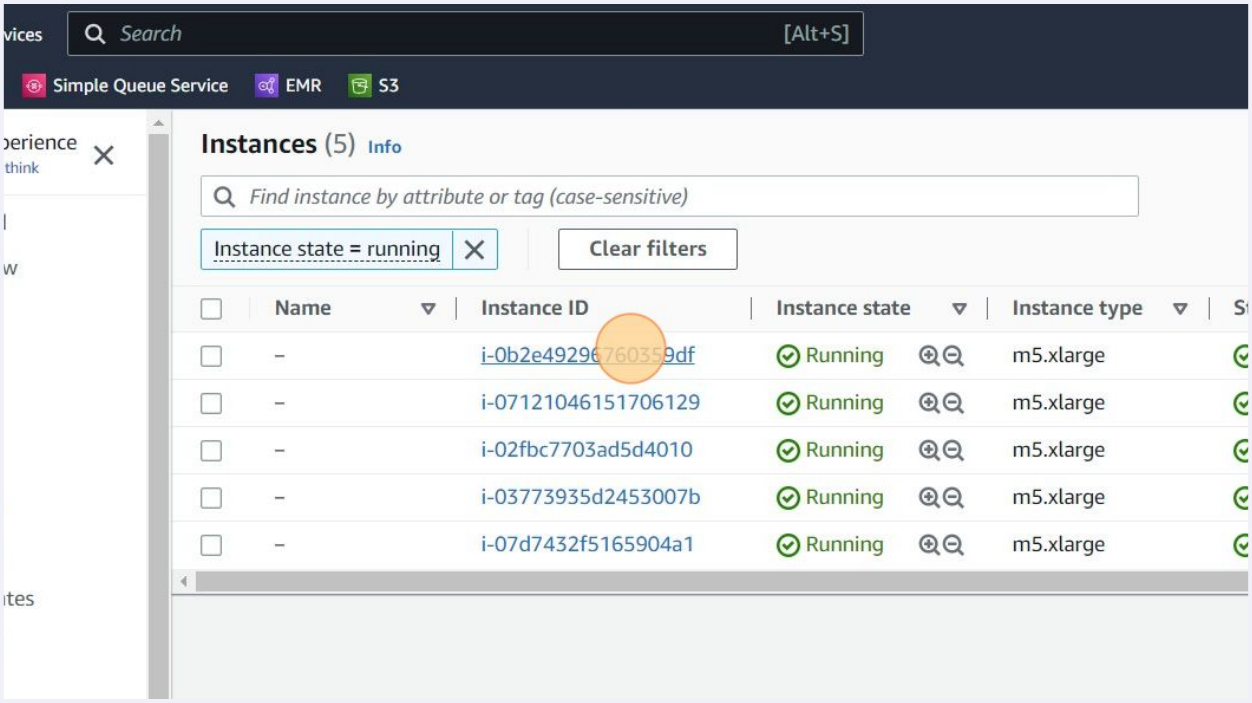
14 After you setup everything you are ready to "Create cluster"



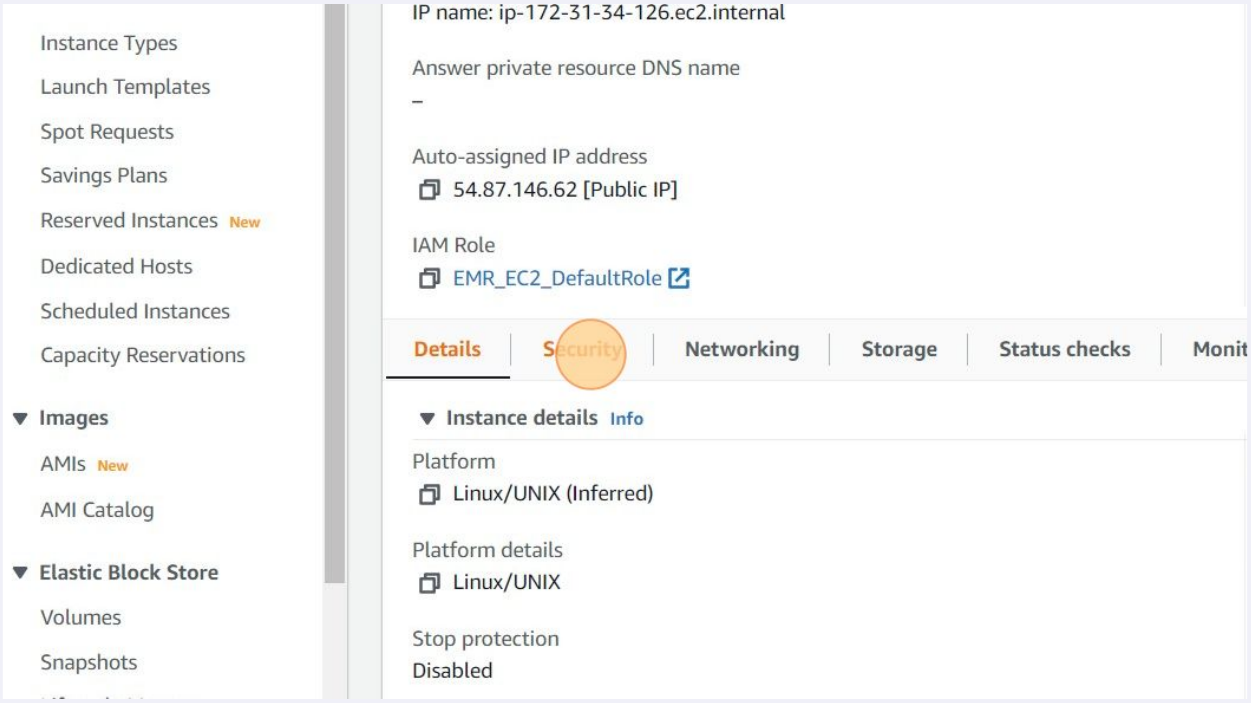
15 After you have configured your cluster, navigate to the EC2 Dashboard and you will notice 5 new instances created.



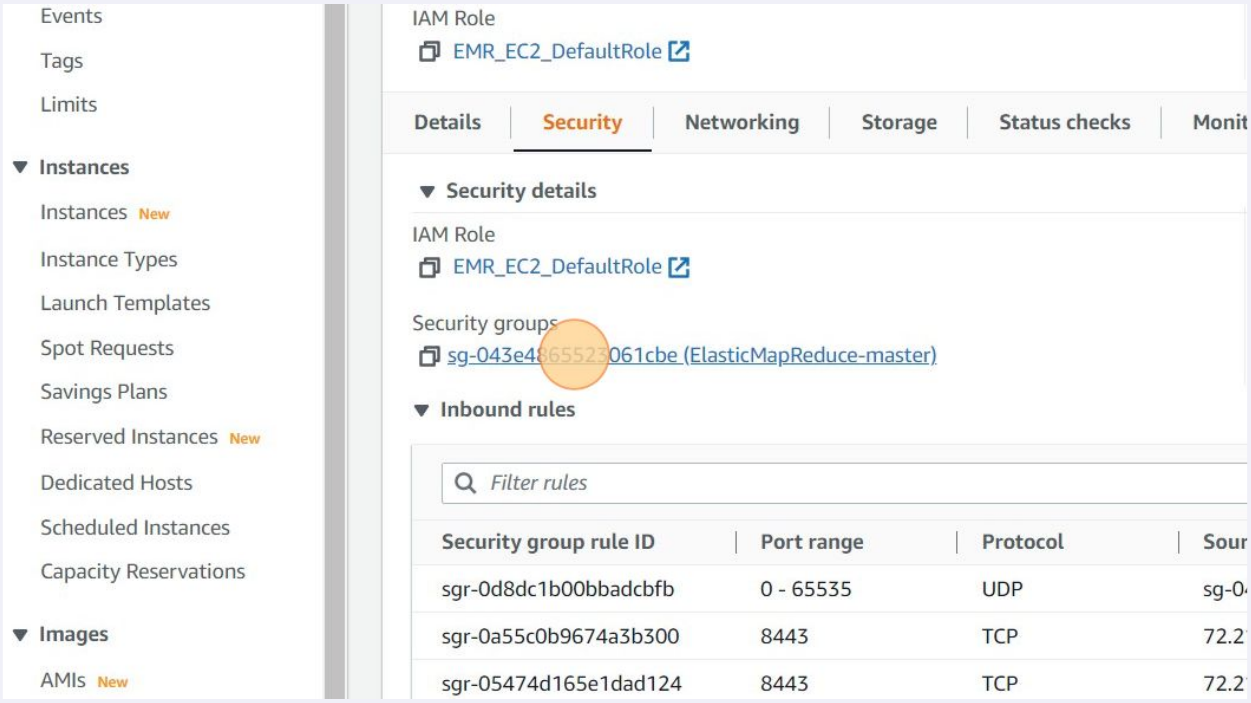
16 Select the first instance you created.



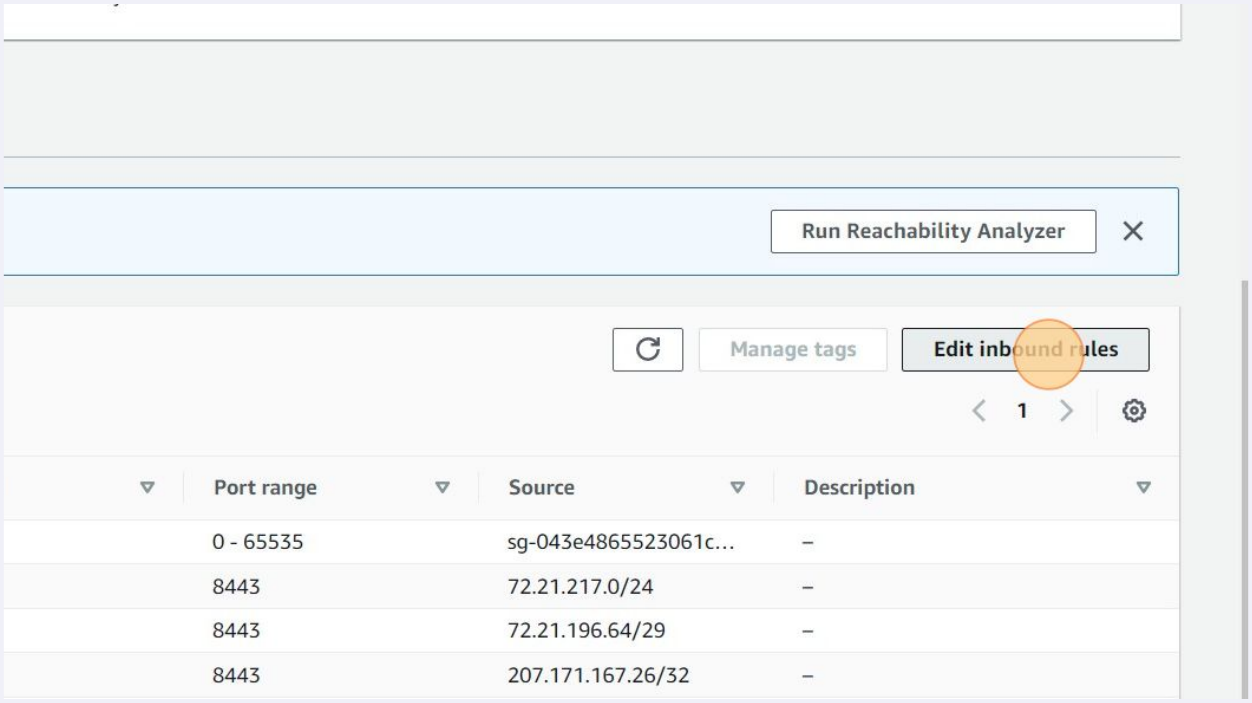
17 Click "Security"



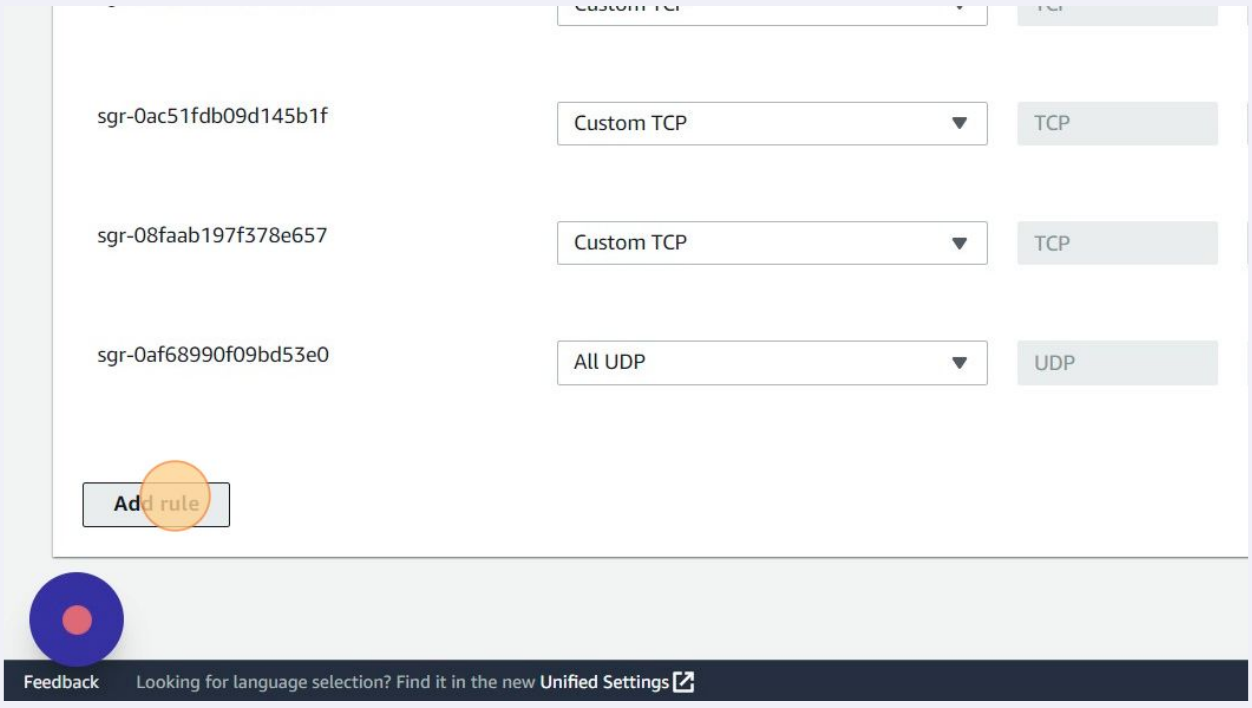
18 Click the link below Security Groups as shown in the screenshot.



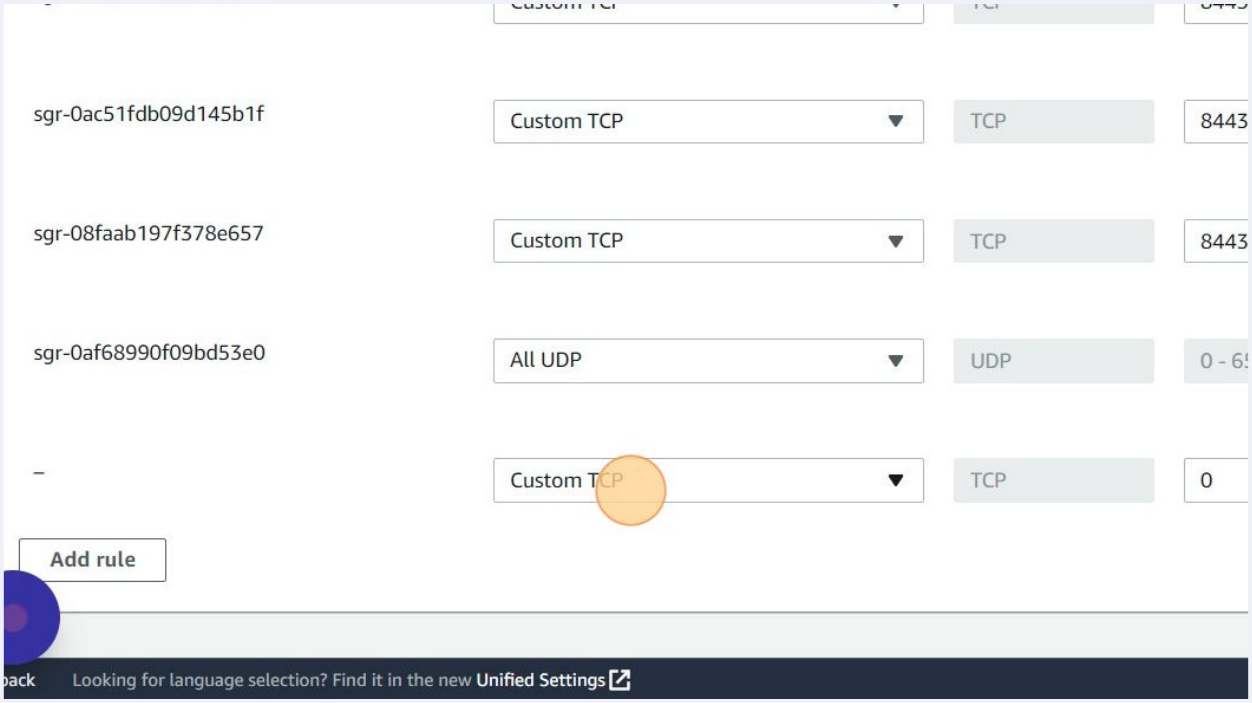
19 Click "Edit inbound rules"



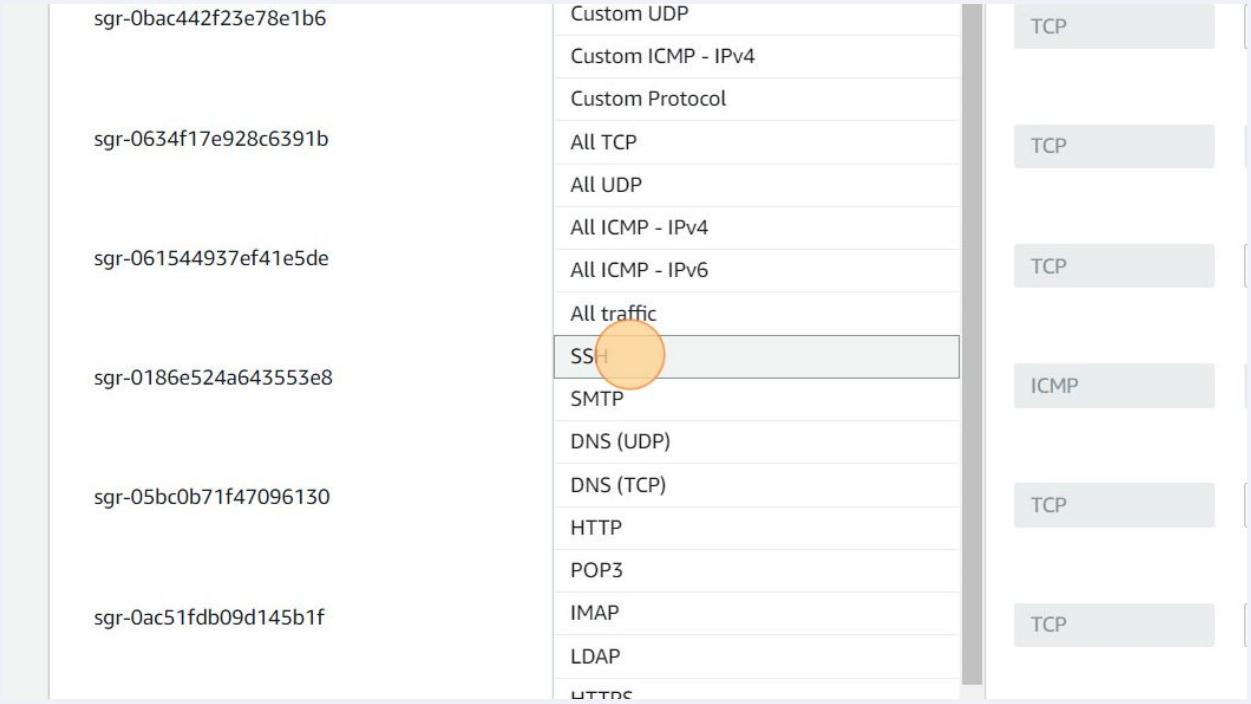
20 Click "Add rule"



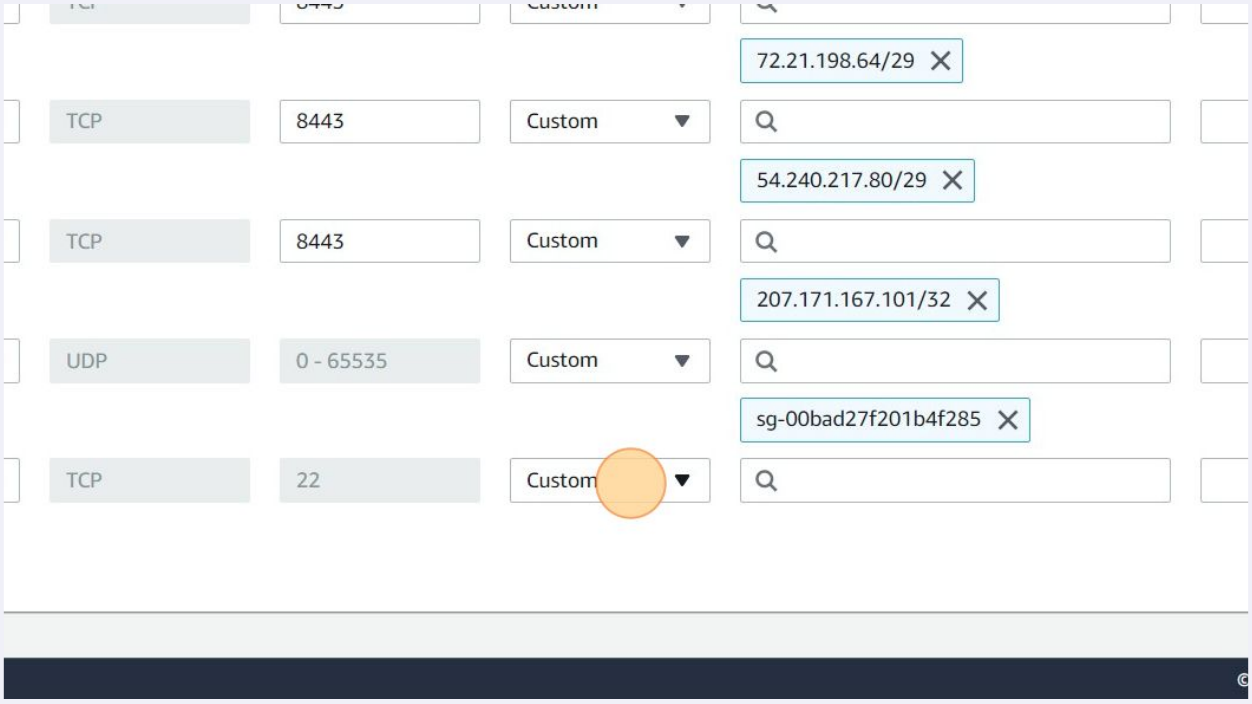
21 Click "Custom TCP"



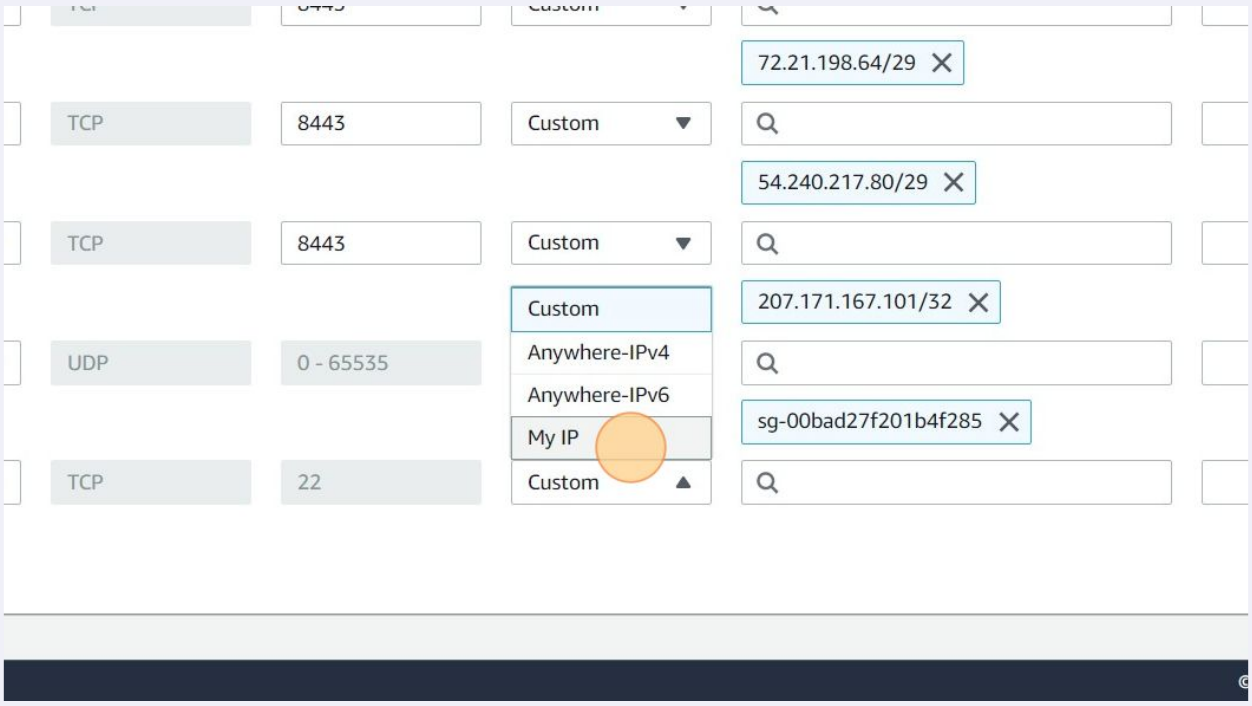
22 Search for "SSH"



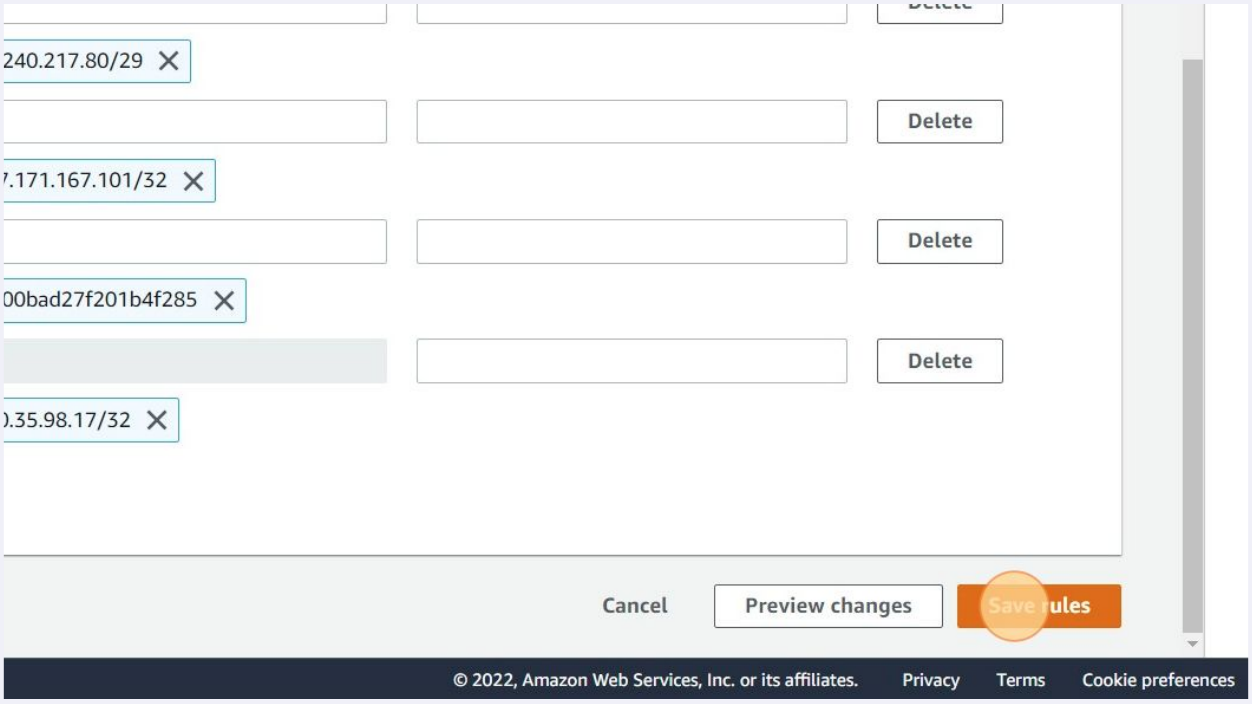
23 Click "Custom"



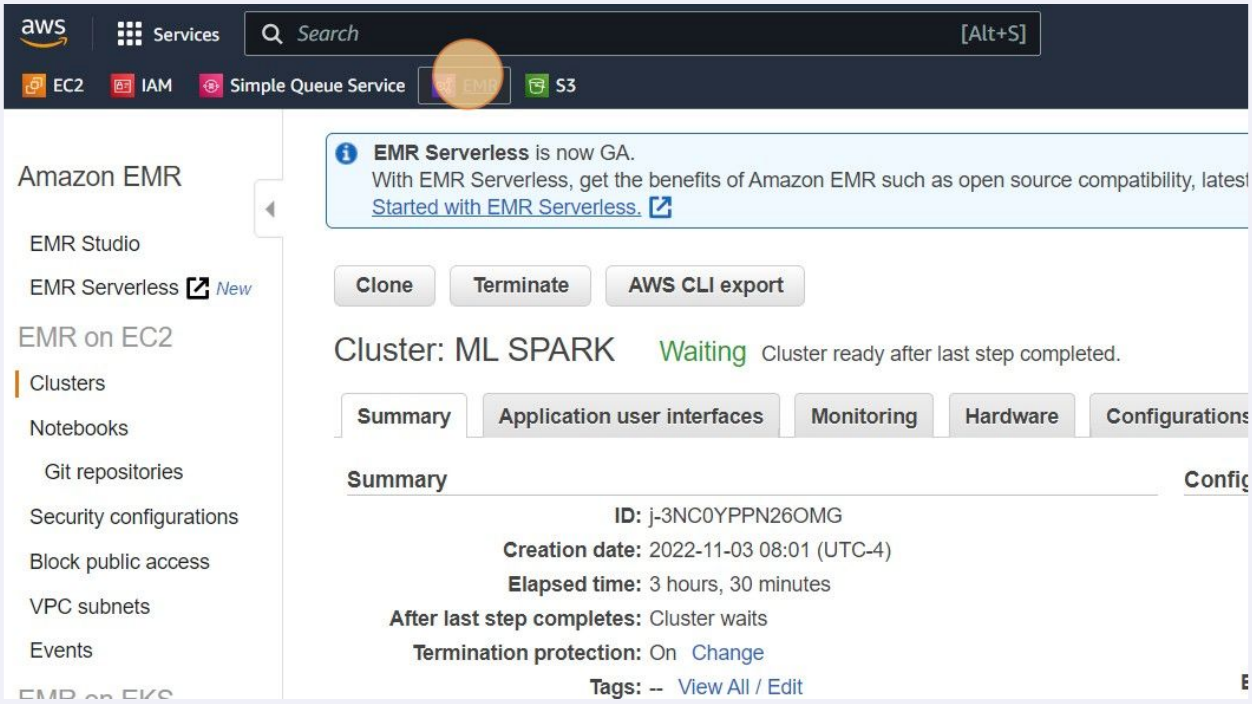
24 Change it to "My IP"



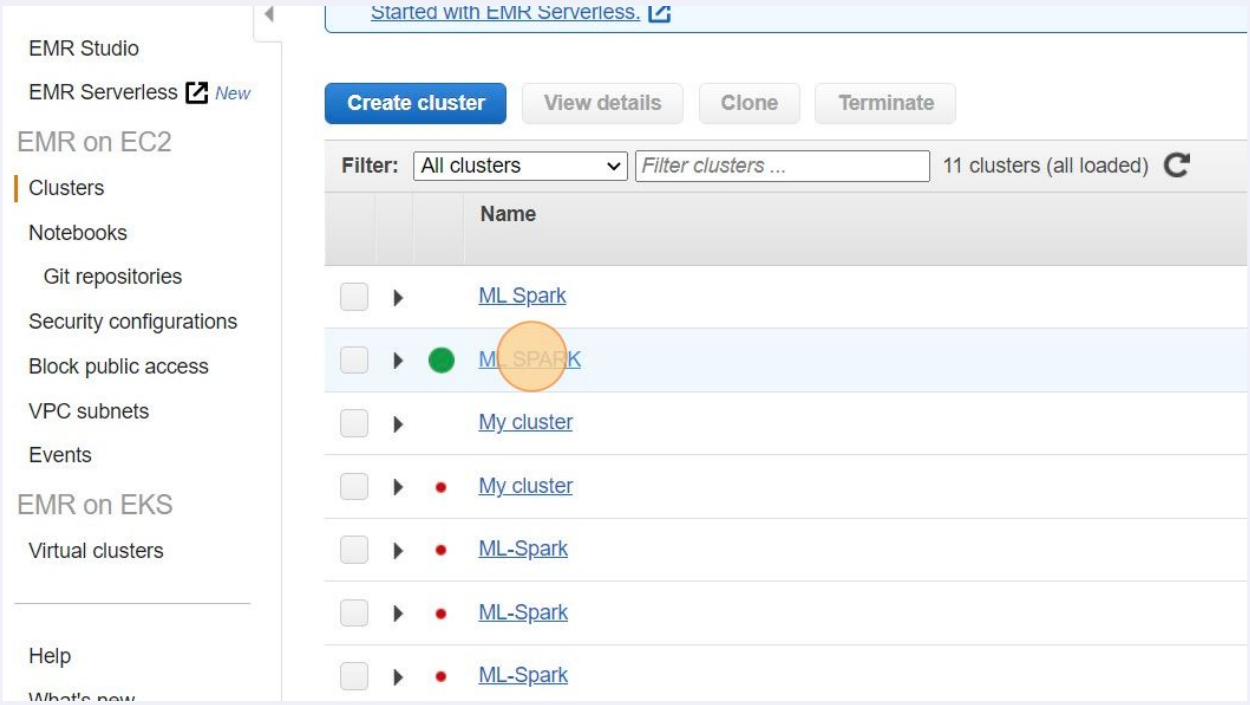
25 Click "Save rules"



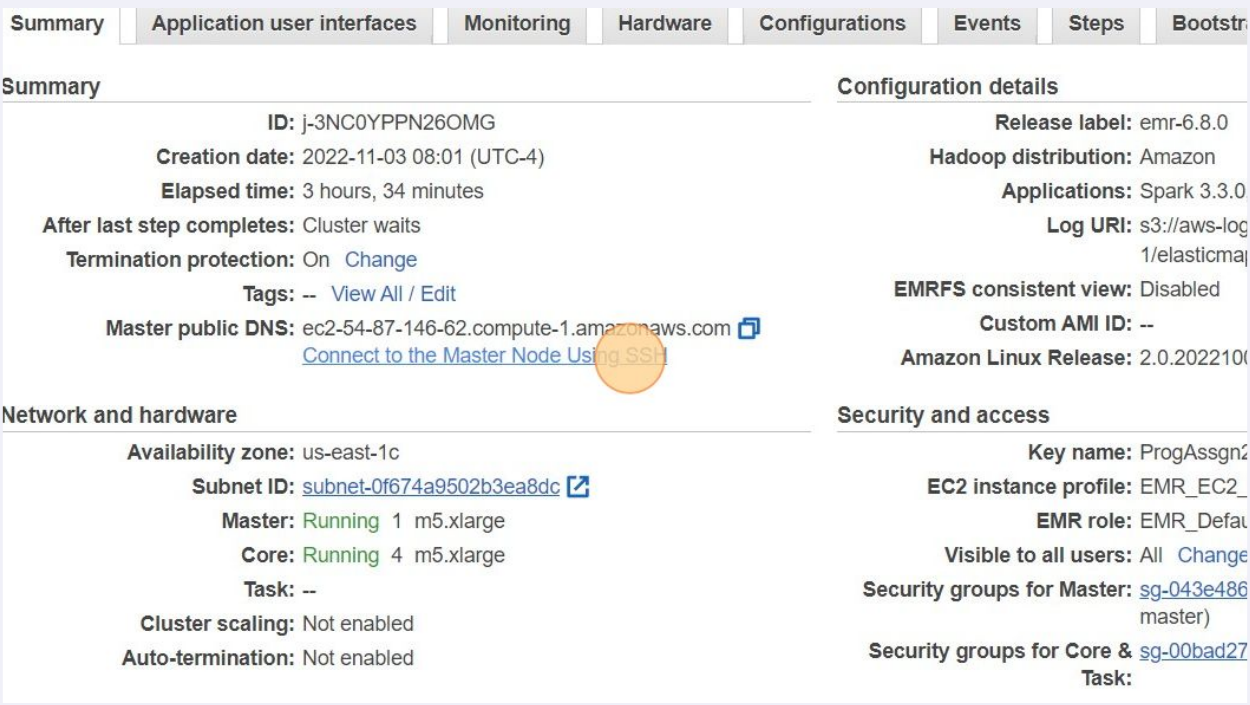
26 To connect to your Cluster go back to "EMR Dashboard"



27 Click on your Cluster

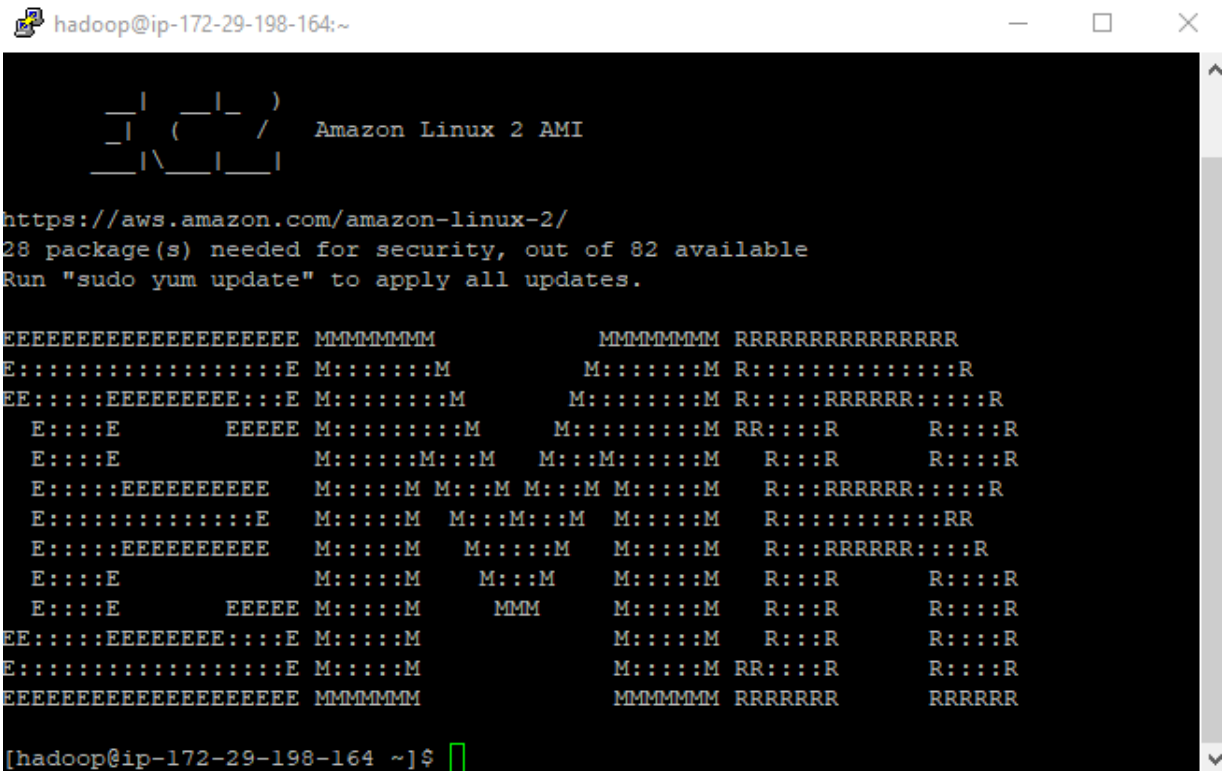


28 Click "Connect to the Master Node Using SSH"



You can access your Cluster after following the procedures provided there in accordance with your OS.

It should look something like this.



End