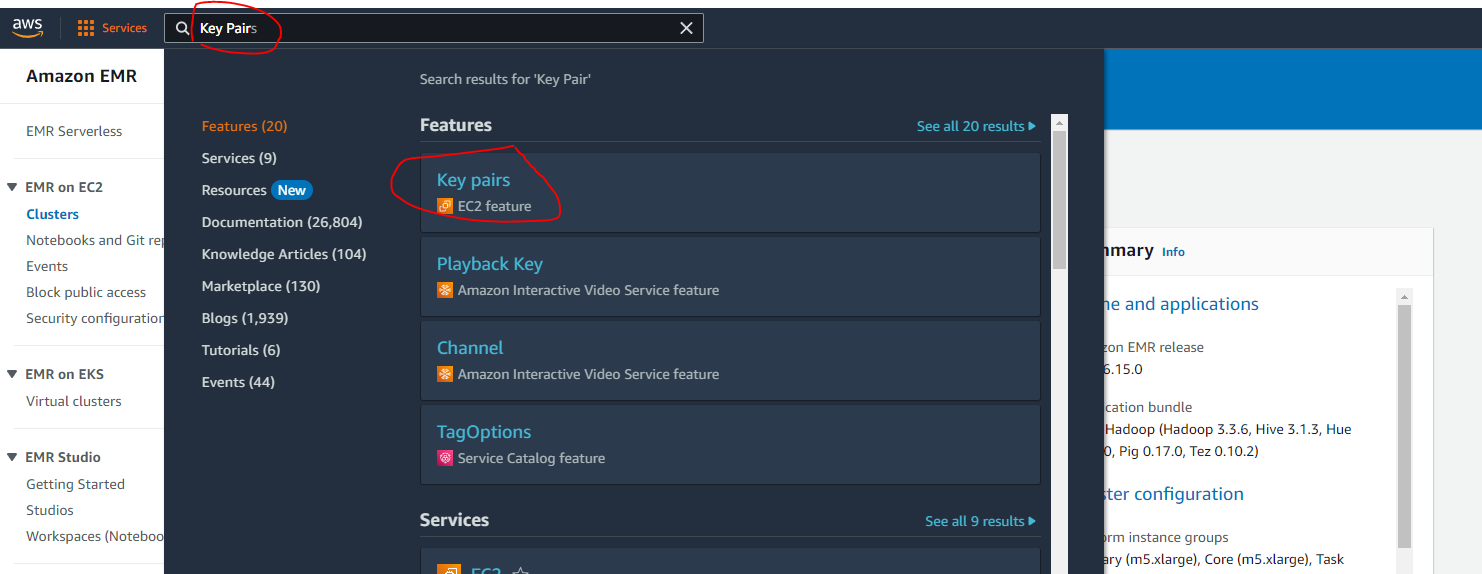
# AWS PREREQUISITE

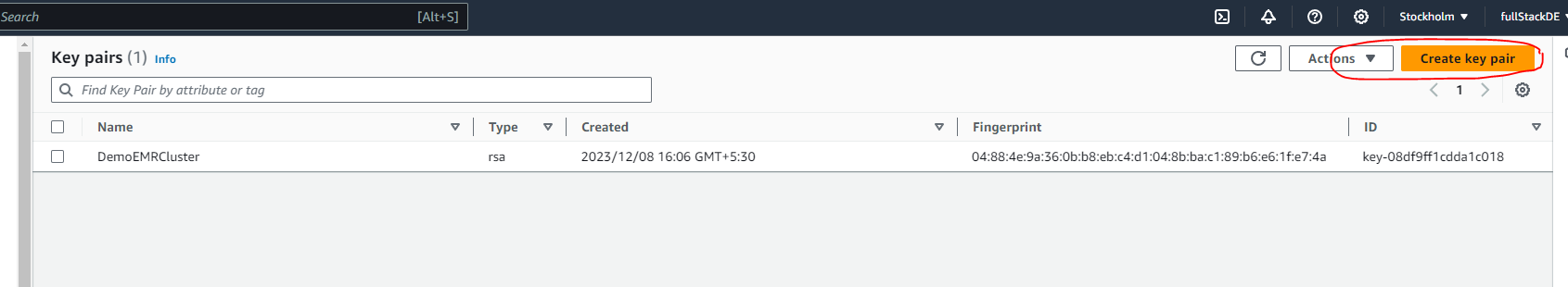
This has to be done once

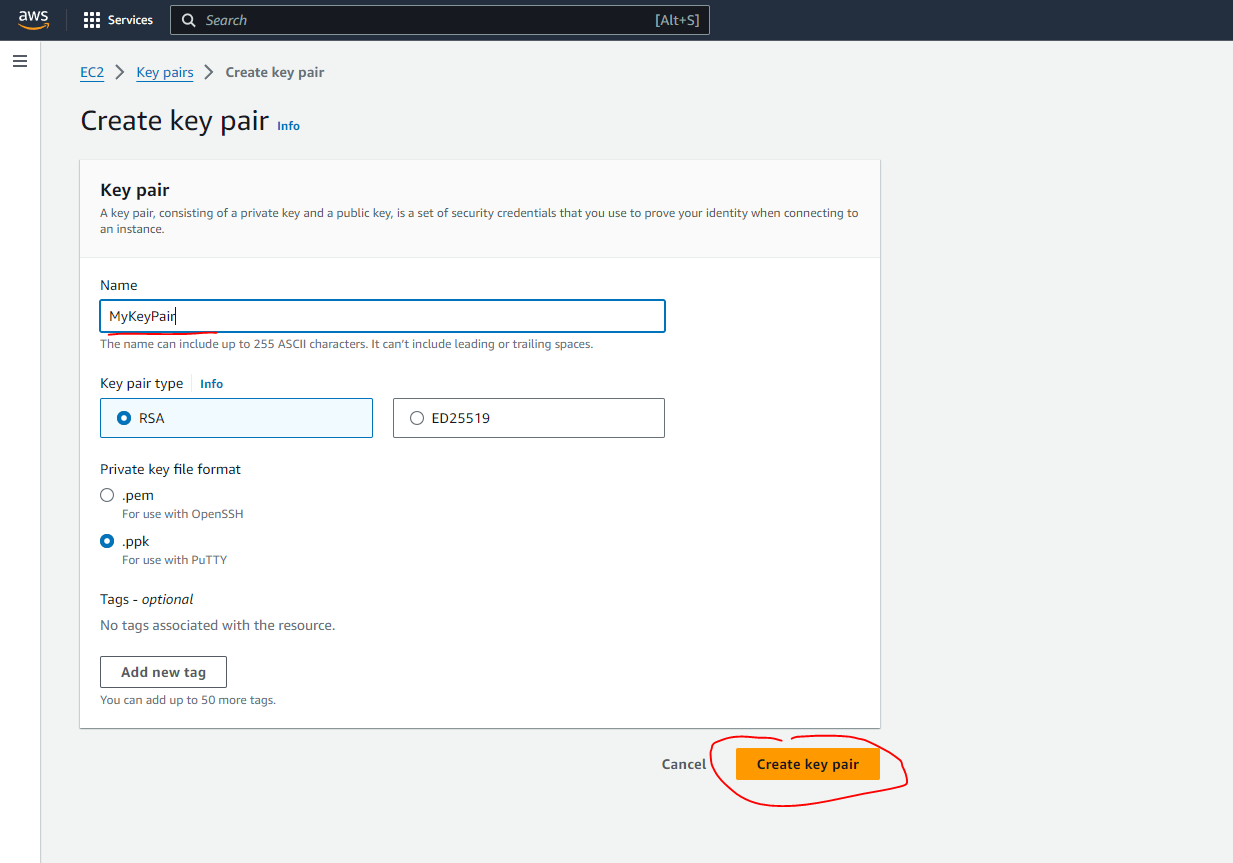
### **Create Key Value Pair**

1. Search “Key Pairs” on aw
2. s console Home – Click on it.

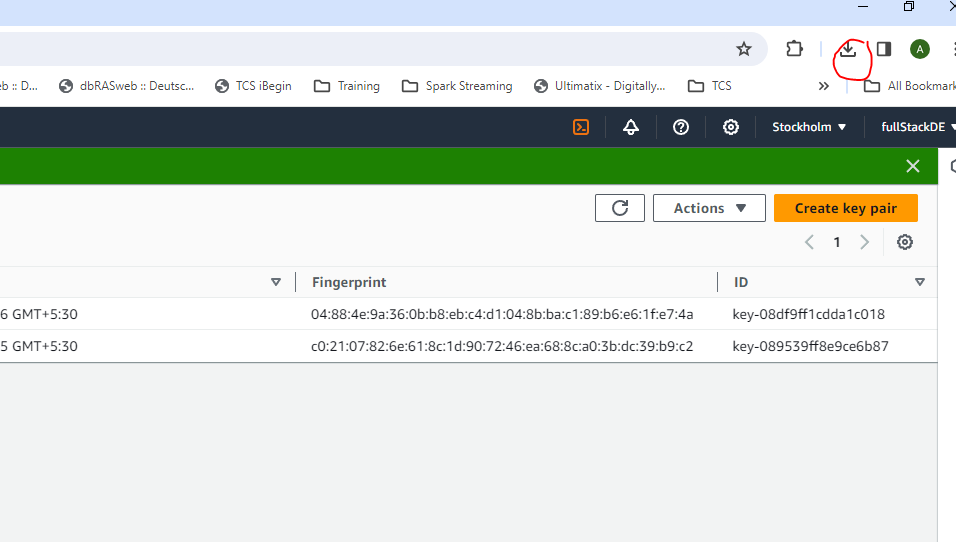


1. Create key pair



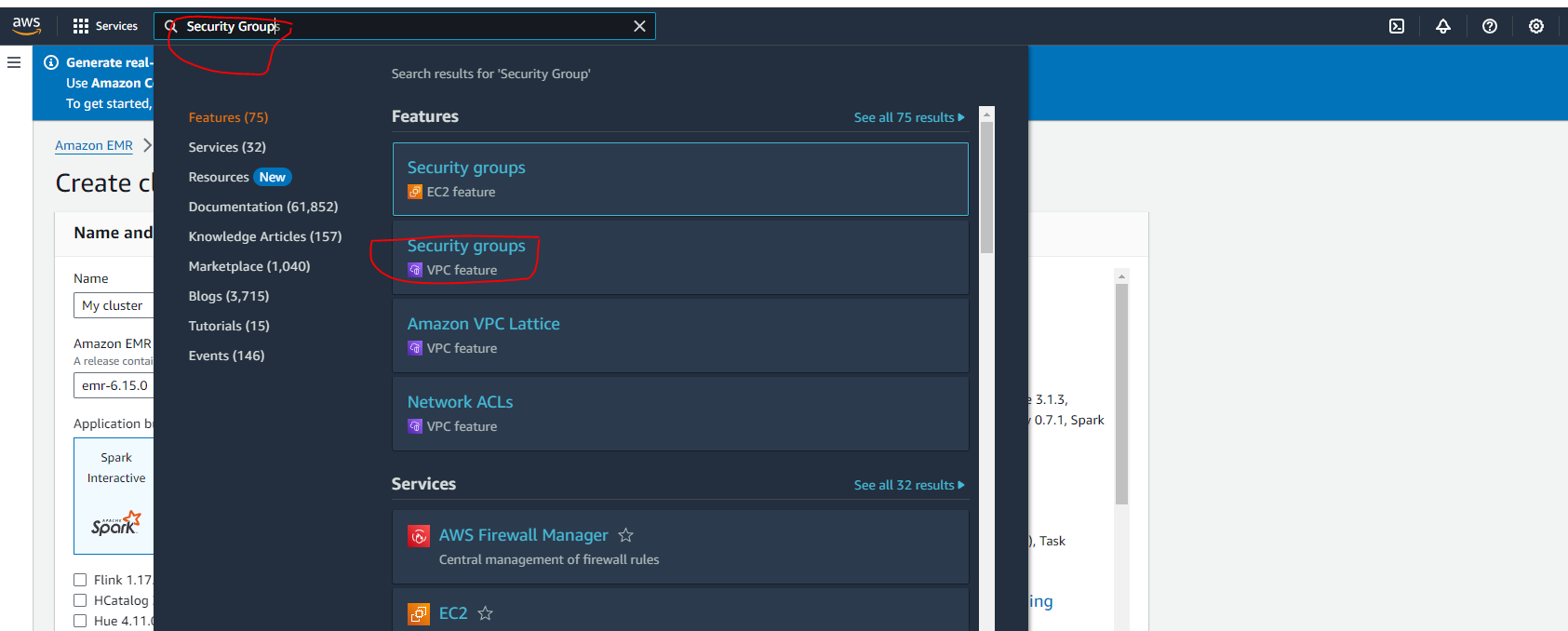


1. MyKeyPair will be downloaded in ppk format in your download folder

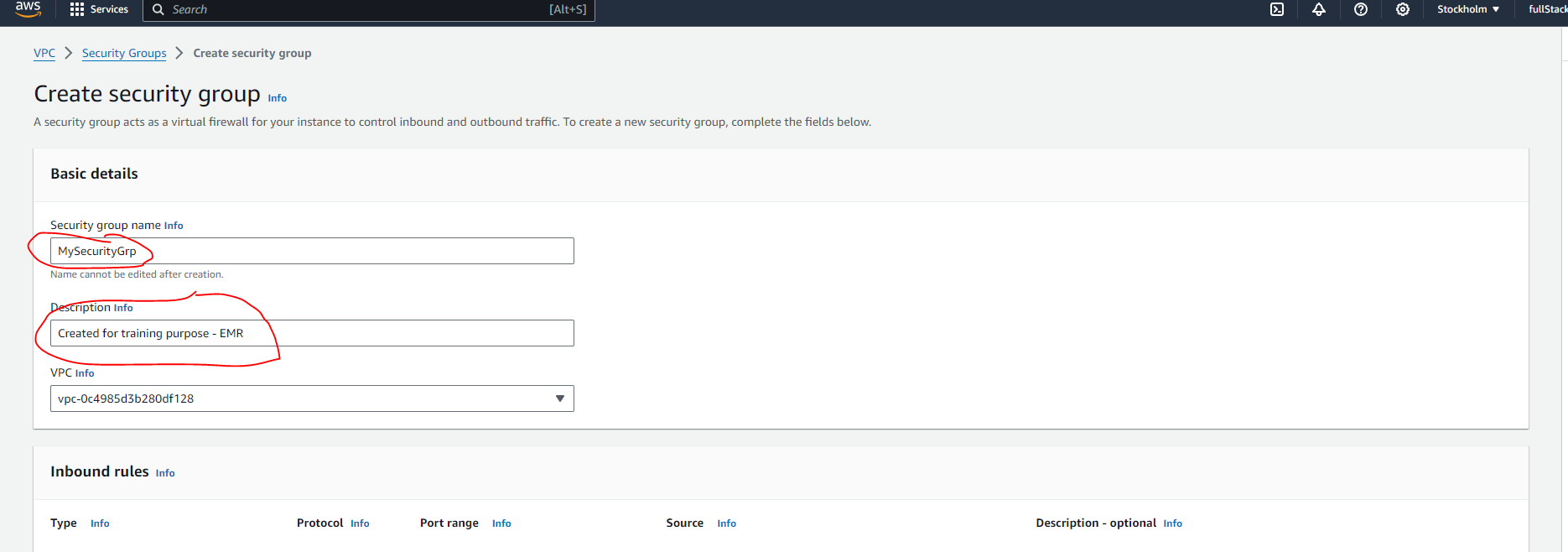


### Create Security Group

1. Search Security Group

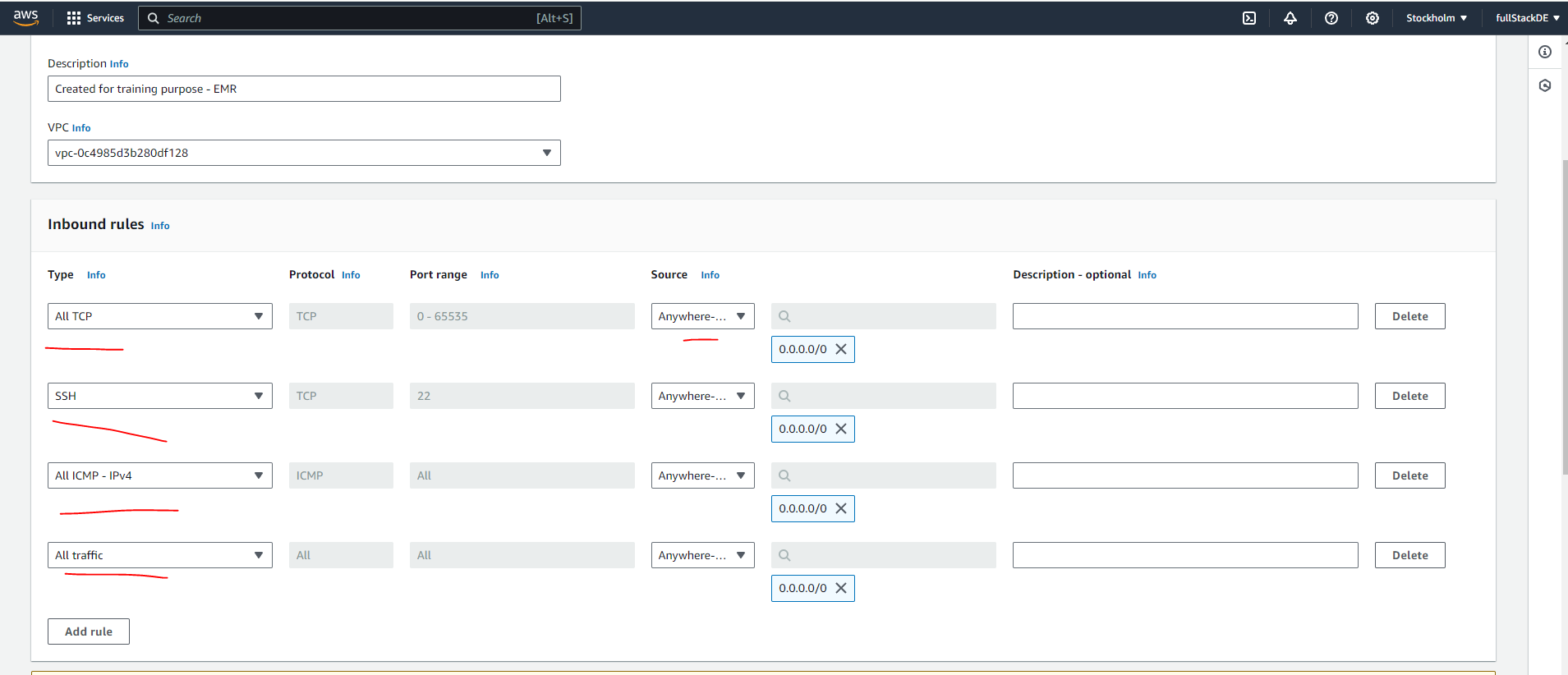


1. Create Security Group

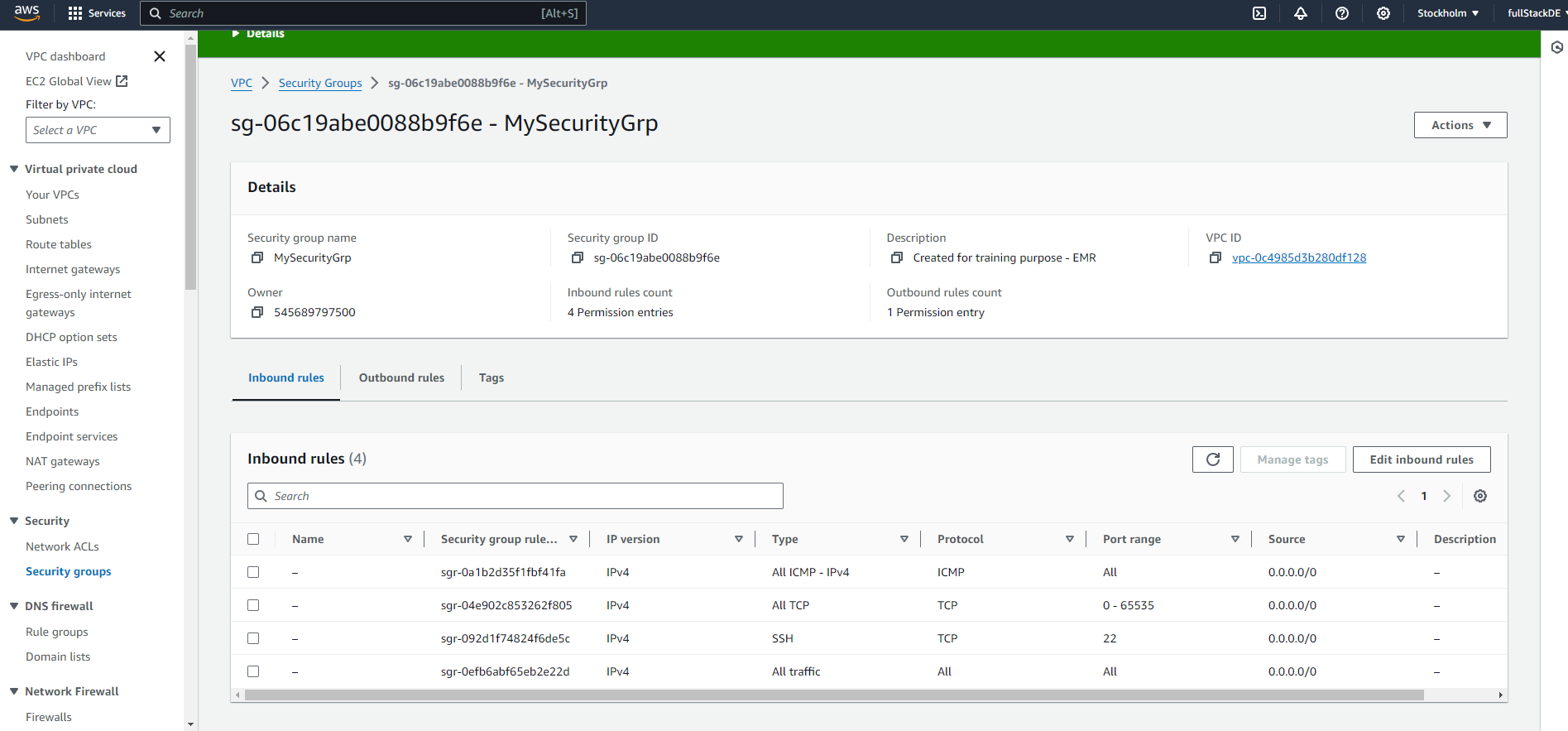


1. Add Inbound Rules

Under Source Tab – “***Anywhere IPv4***”

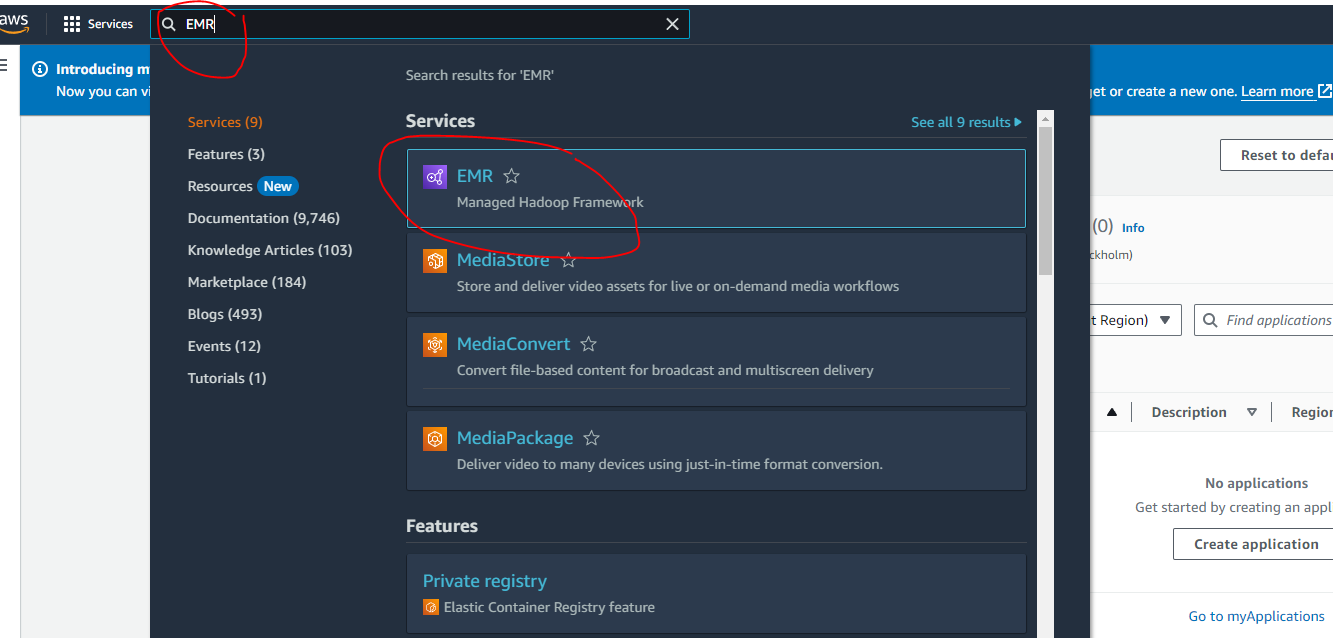


1. Click on Create Security Group and MySecurityGrp is created

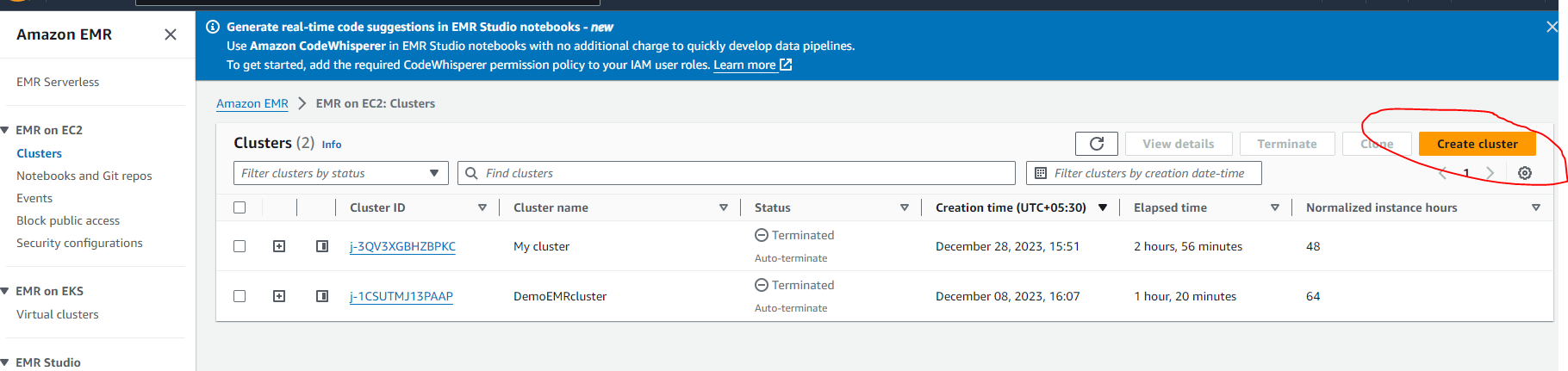


# Launching An EMR cluster on AWS for Hadoop and Hive Services

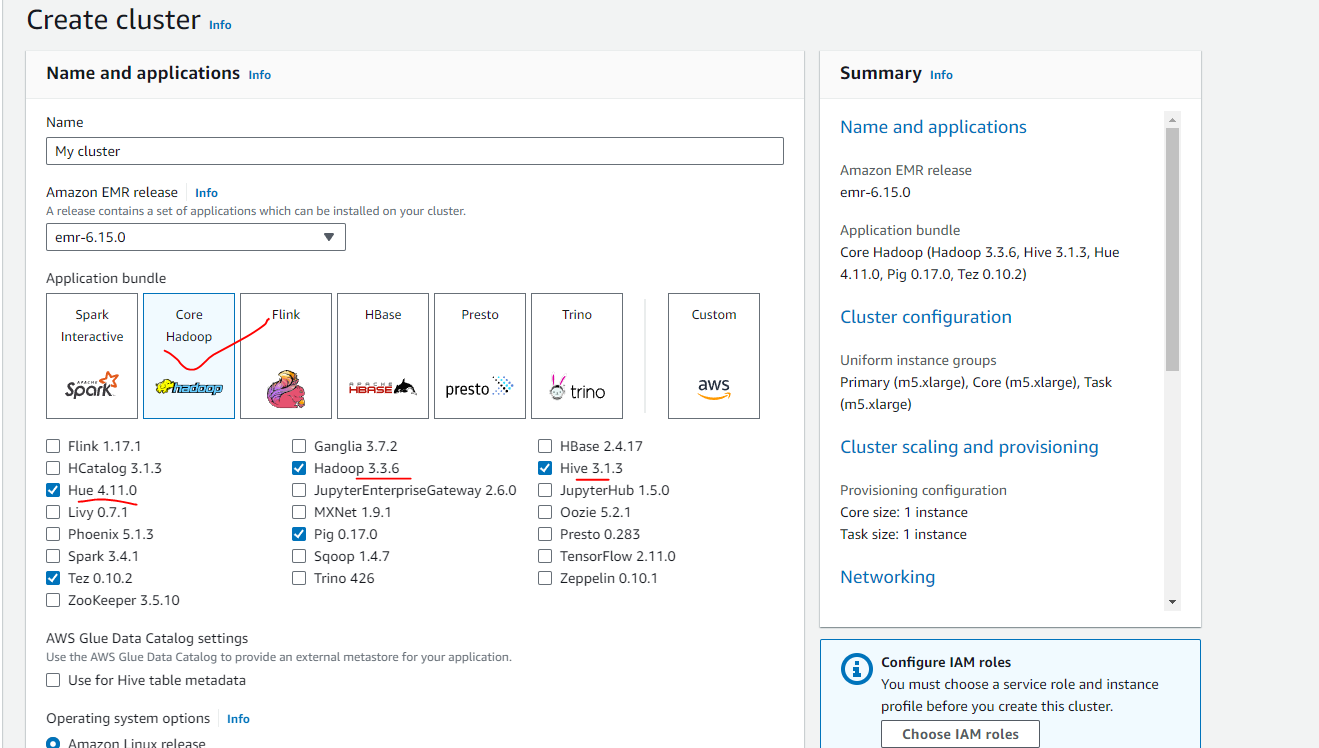
1. Search EMR on aws Console Home – Click on EMR



1. Create Cluster – give a name which we can remember – for ex – “My Cluster”



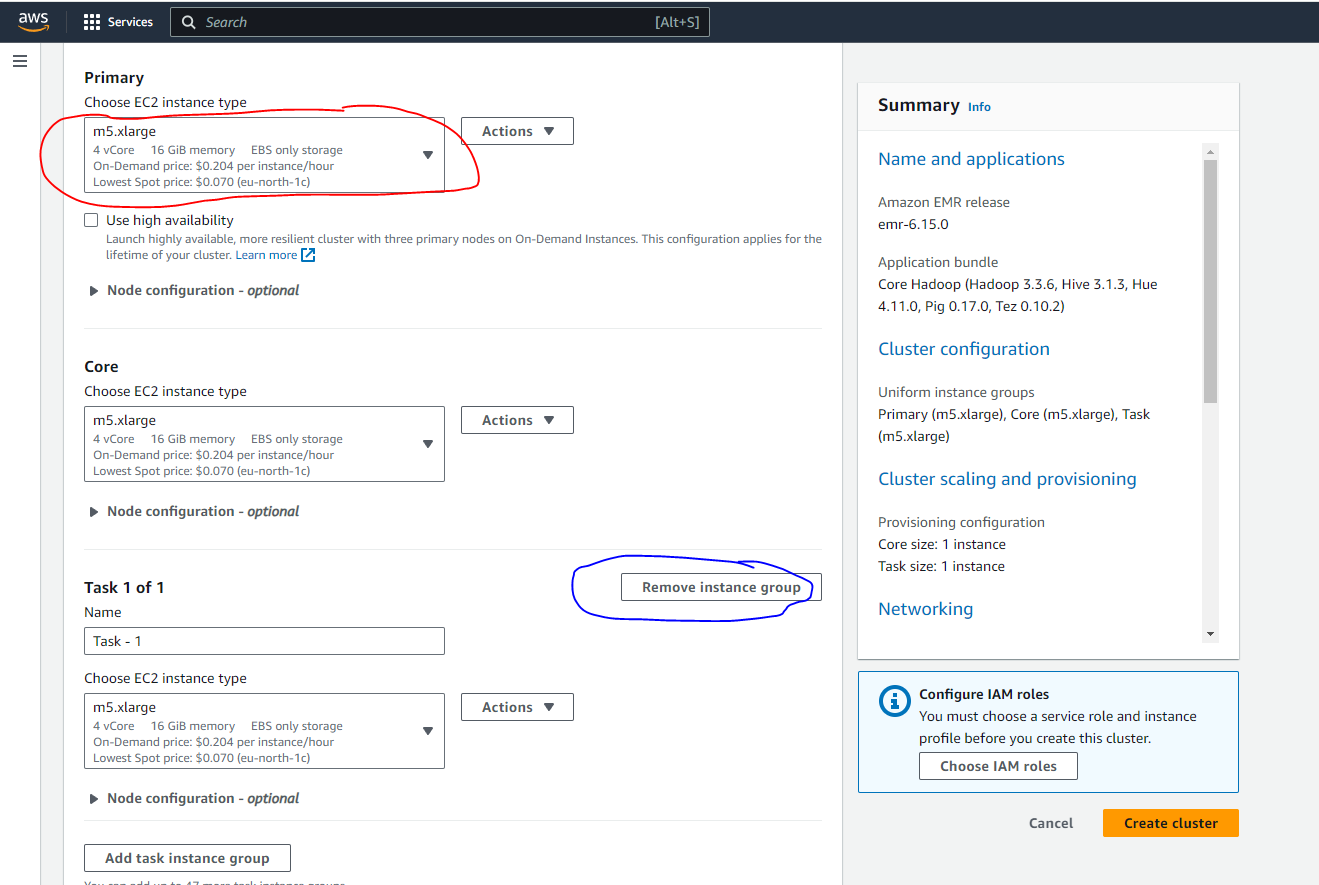
1. Select Core Hadoop which will select the core Hadoop applications – like Hadoop, Hive , Hue and others.



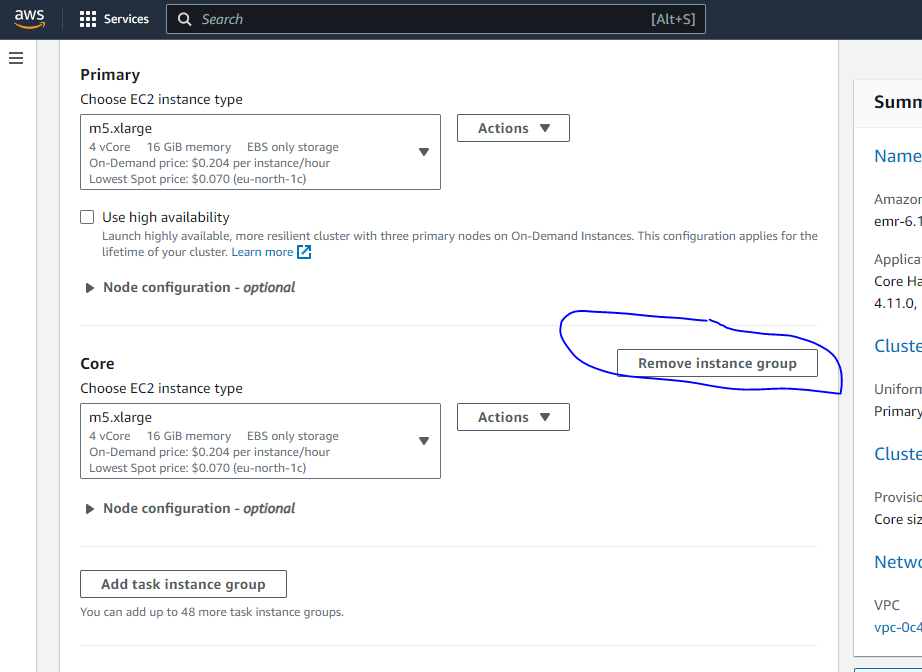
1. By Default following EC2 Instance is selected- Its not free hence always terminated after the practice.

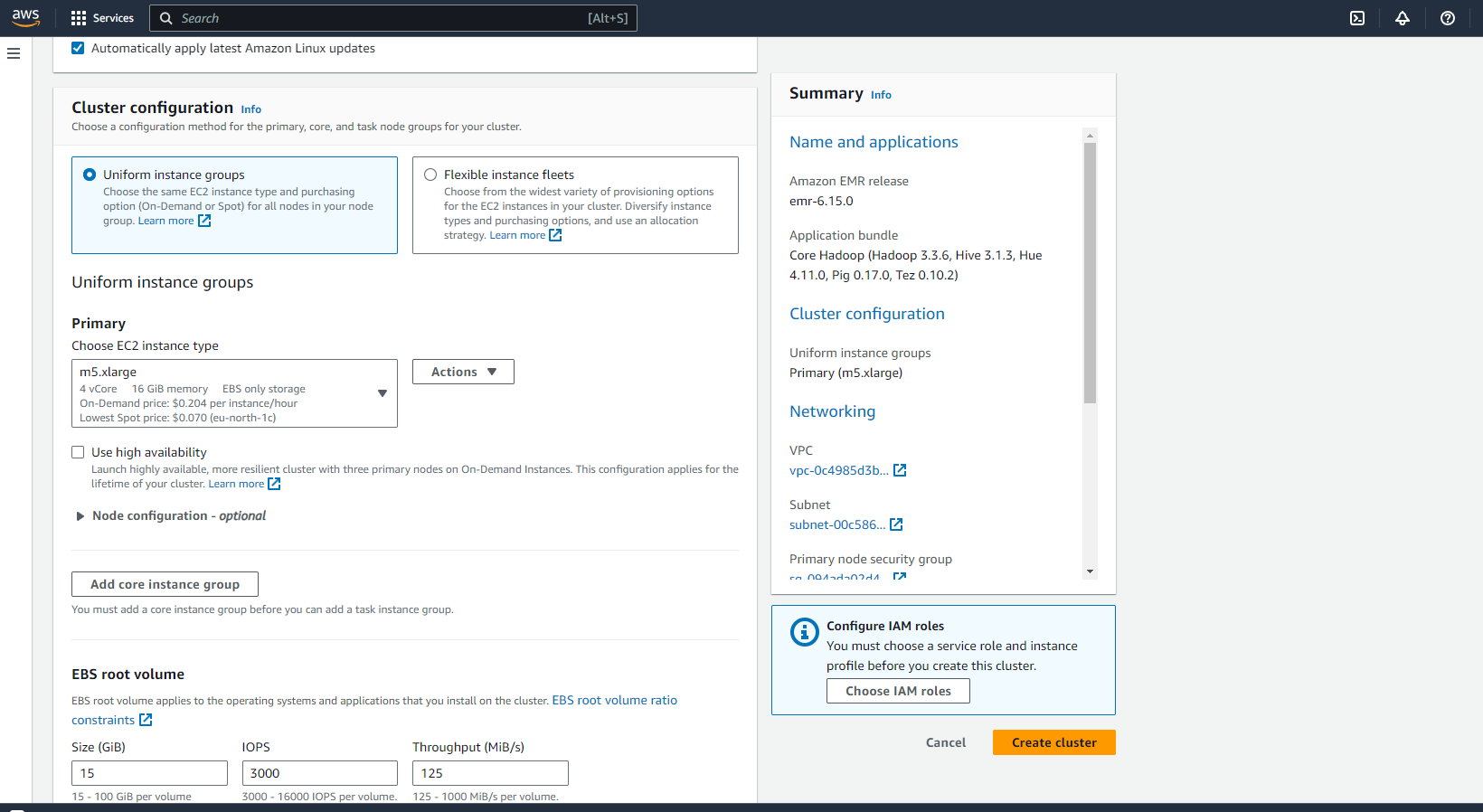
Click on the Remove Instance Group for **Task 1 of 1.**

We are removing it as it will add to extra cost and in general practice not require to launch 2 node cluster.

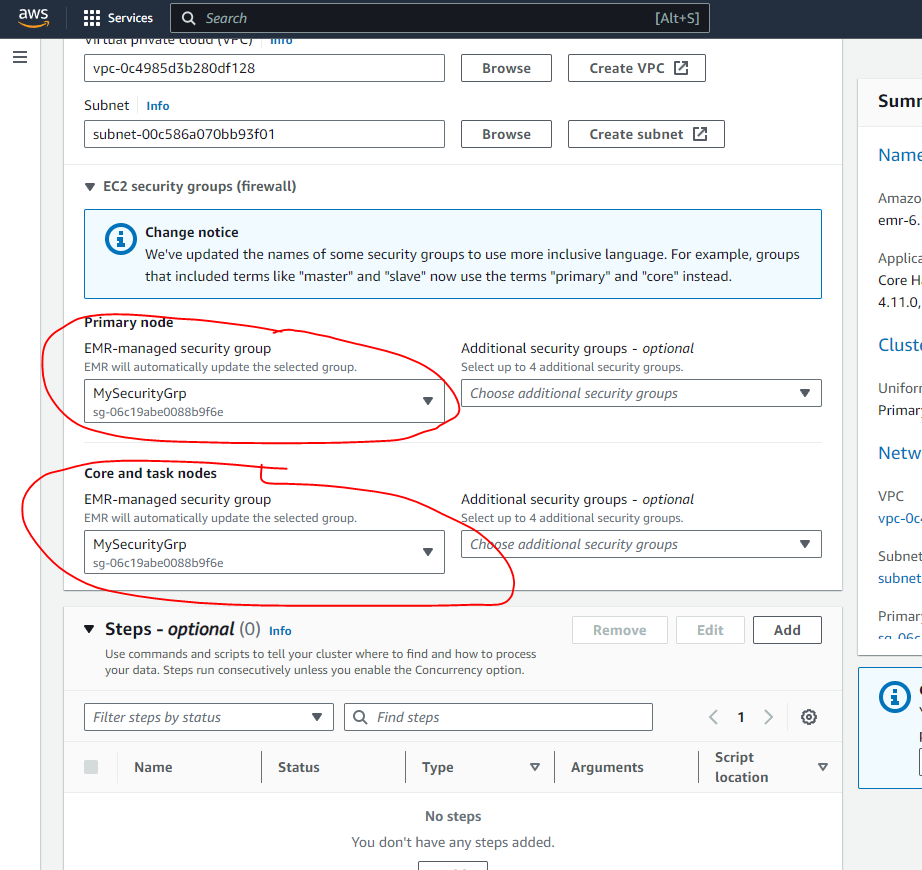


**Remove the following Core instance group as well.**





Select “**MySecurityGrp” (**one which is created earlier**)** for Primary node and Core and task nodes.



1. Select EC2 key pair for SSH to the cluster

