For Linux/Mac OS users to connect to the EMR Instance

Prerequisite:

- Make sure you have set up MyIP in your instance's inbound security group.

For Linux/Mac systems, **you do not need to convert your .pem file to a .ppk file**.

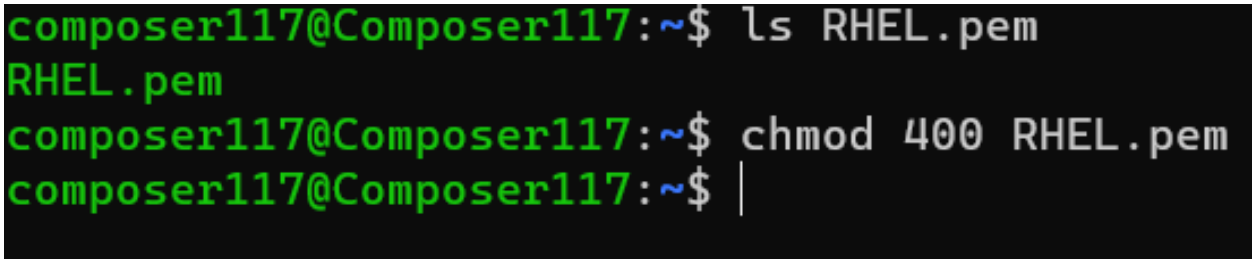
1. Open **'Terminal'** on your system and go to the location where you downloaded the .pem file.
   - Let's say that your .pem file was downloaded in the 'Downloads' folder.
   - You need to first change your current working directory to the 'Downloads' directory.
   - To do that, use the following '**cd**' command: `cd Downloads/`

2. Next, run the '`ls`' command, which lists all the files in a given Linux directory. Verify that your .pem file exists in the given directory.

3. Change the permissions of the .pem file to 400, which gives the user only read permission and removes all other permissions.

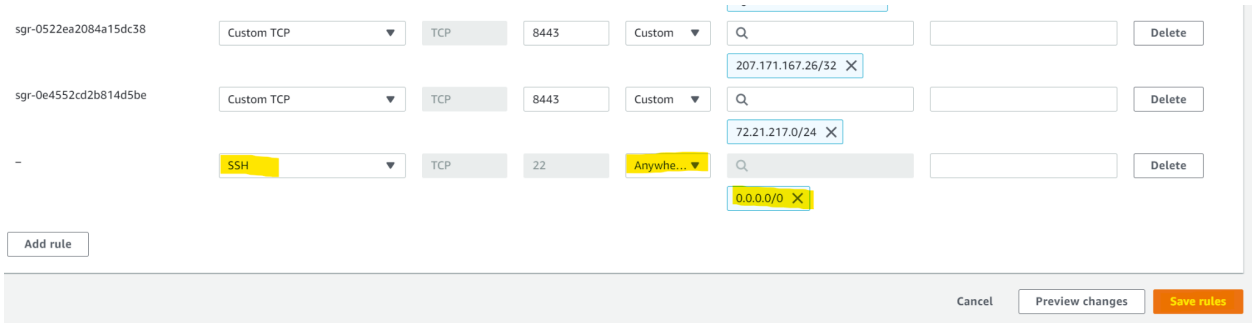
   The command is as shown below. (RHEL.pem is the filename in our case).

   ```
   chmod 400 RHEL.pem
   ```

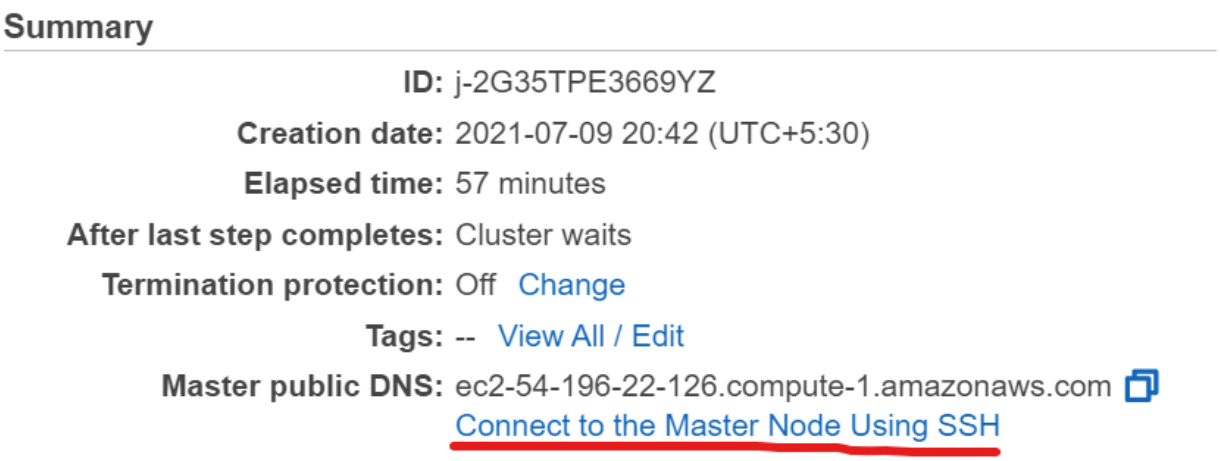


Note:

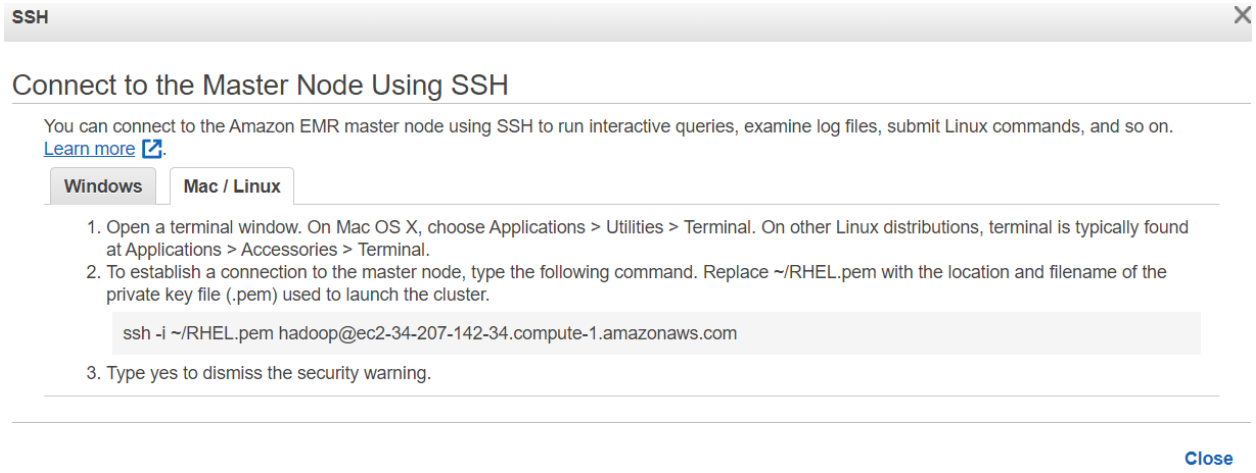
- Make sure you have set the Port 22(SSH) as open as mentioned in the EMR Setup document before using Putty to login to your EMR cluster otherwise you won't be able to SSH to your EMR cluster



4. Now, go back to your EMR instance page and click on the '**Connect to the Master Node Using SSH'** link.



You should now be able to see the following screen, where you should then go to the Mac/Linux tab.



5. Copy the command shown at the end of step 2.

   The command will look like this:

```
ssh -i pemfile.pem hadoop@public_dns_name
```

   - Before running this command, ensure that you are present in the directory in which your .pem file is present.
   - This can be checked using the '**pwd**' command, which writes the full path of the current working directory. The command is: `pwd`

6. Then, run the following command to establish the connection:

```
ssh -i ~/RHEL.pem hadoop@ec2-34-207-142-34.compute-1.amazonaws.com
```

   Next, enter **'yes'**, when prompted by the terminal.

```
composer117@Composer117:~$ ssh -i ~/RHEL.pem hadoop@ec2-34-207-142-34.compute-1.amazonaws.com
The authenticity of host 'ec2-34-207-142-34.compute-1.amazonaws.com (34.207.142.34)' can't be established.
ECDSA key fingerprint is SHA256:UbZwrW1QeQR4s+9Thsi6T9I3oTrZVj0Yqn/VeyadI18.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'ec2-34-207-142-34.compute-1.amazonaws.com,34.207.142.34' (ECDSA) to the list of known hosts.
Last login: Sat Jul 17 14:44:19 2021

       __|  __|_  )
       _|  (     /   Amazon Linux 2 AMI
      ___|\___|___|

https://aws.amazon.com/amazon-linux-2/
101 package(s) needed for security, out of 171 available
Run "sudo yum update" to apply all updates.

EEEEEEEEEEEEEEEEEEE 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
EEEEEEEEEEEEEEEEEEE MMMMMMM             MMMMMMM RRRRRRR      RRRRRR
```

7. After this, run the following command to check whether S3 is accessible from your instance.

```
aws s3 ls
```

8. If the above command is not working, then you will need to follow the S3 access document given on the platform. Otherwise, you have successfully logged in to your EMR cluster.

```
[hadoop@ip-172-31-46-44 ~]$ aws s3 ls
2020-06-28 04:56:43 atmetloutput5files
2020-12-14 17:59:20 aws-emr-resources-367134191692-us-east-1
2020-12-14 17:47:40 aws-logs-367134191692-us-east-1
2021-03-07 14:56:29 demobucket00123
2020-09-11 06:49:25 rawattestbucket123
2020-06-24 15:33:12 shrianshs3testbucket
2020-12-04 11:36:10 testbucketupgrad
[hadoop@ip-172-31-46-44 ~]$
```

9. Before logging off, make sure that you have run the following command.

```
sudo yum update
```

```
[hadoop@ip-172-31-32-207 ~]$ sudo yum update
Loaded plugins: extras_suggestions, langpacks, priorities, update-motd
amzn2-core                                                                | 3.7 kB  00:00:00
10 packages excluded due to repository priority protections
Resolving Dependencies
--> Running transaction check
---> Package amazon-linux-extras.noarch 0:1.6.10-1.amzn2 will be updated
---> Package amazon-linux-extras.noarch 0:2.0.0-1.amzn2 will be an update
---> Package amazon-linux-extras-yum-plugin.noarch 0:1.6.10-1.amzn2 will be updated
---> Package amazon-linux-extras-yum-plugin.noarch 0:2.0.0-1.amzn2 will be an update
---> Package amazon-ssm-agent.x86_64 0:2.3.714.0-1.amzn2 will be updated
---> Package amazon-ssm-agent.x86_64 0:3.0.1124.0-1.amzn2 will be an update
---> Package avahi-libs.x86_64 0:0.6.31-19.amzn2.0.1 will be updated
---> Package avahi-libs.x86_64 0:0.6.31-20.amzn2 will be an update
---> Package aws-cfn-bootstrap.noarch 0:1.4-31.amzn2 will be updated
---> Package aws-cfn-bootstrap.noarch 0:2.0-6.amzn2 will be an update
--> Processing Dependency: python3-pystache for package: aws-cfn-bootstrap-2.0-6.amzn2.noarch
--> Processing Dependency: python3-daemon for package: aws-cfn-bootstrap-2.0-6.amzn2.noarch
---> Package awscli.noarch 0:1.16.300-1.amzn2.0.1 will be updated
---> Package awscli.noarch 0:1.18.147-1.amzn2.0.1 will be an update
---> Package bash.x86_64 0:4.2.46-33.amzn2 will be updated
---> Package bash.x86_64 0:4.2.46-34.amzn2 will be an update
```

When you get a prompt to verify if the update can be done, type in '**y',** and then press Enter. This will successfully update your EMR instance.

```
Transaction Summary
================================================
====Install   10 Packages (+6 Dependent packages)
Upgrade  161 Packages

Total download size: 308 M
Is this ok [y/d/N]: y
```

Wait for a few minutes until the update has been completed. You will eventually be able to see the following screen.

```
Replaced:
  ec2-net-utils.noarch 0:1.3-1.amzn2
86_64 1:2.02-35.amzn2.0.4
  python3-tools.x86_64 0:3.7.6-1.amzn2.0.1

Complete!
[hadoop@ip-172-31-32-207 ~]$
```

You have now successfully logged into your EMR instance.