File Transfer – Linux/Mac

Uploading data file from the local machine to the EMR instance

1. Open the terminal on your machine.
2. Login to your EMR instance using SSH.
3. Open another terminal tab and browse to the directory where the data file is present.

```
composer117@Composer117:~/File_transfer_demo$ ls
test.txt
composer117@Composer117:~/File_transfer_demo$
```

In our case, we want to upload this **test.txt** file to the EMR instance.

4. Use the following command to upload the data into your EMR instance.

```
scp -i <path to pem file> <path to the data file> hadoop@<public
DNS>:<destination path in the EMR instance>
```

We have used the following command to upload the test.txt file.

```
scp -i ~/RHEL.pem test.txt hadoop@ec2-34-207-142-34.compute-1.amazonaws.com:~
```

```
composer117@Composer117:~/File_transfer_demo$ scp -i ~/RHEL.pem test.txt hadoop@ec2-34-207-142-34.compute-1.amazonaws.co
m:~
test.txt                                                      100%   15     0.0KB/s   00:00
composer117@Composer117:~/File_transfer_demo$
```

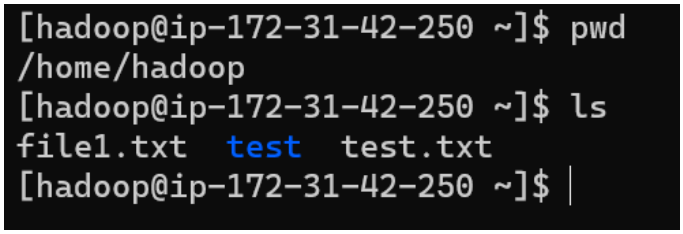
Note: We are running this command from the directory where the data file is present.

5. Verify the data file in the EMR instance.
   Go to the directory where the file was uploaded. (**~ or /home/hadoop** in our case)
   Run the `ls` command to verify that the file is present.

```
[hadoop@ip-172-31-42-250 ~]$ pwd
/home/hadoop
[hadoop@ip-172-31-42-250 ~]$ ls
file1.txt  test  test.txt
[hadoop@ip-172-31-42-250 ~]$
```

Downloading data file from the EMR to the local machine

1. Open the terminal on your machine.
2. Login to your EMR instance using the SSH command and browse to the directory where the data file is present.
   In our case, the data is named **test.txt** and is present in the '**/home/hadoop directory**'.



3. Open another terminal tab on your computer. We use the following command to download the data from the EMR instance to our local machine.

```
scp -i <path to the pem file> hadoop@<public DNS>:<path of the file to be downloaded> <destination path in the local machine>
```

We have used the following command:

```
scp -i ~/RHEL.pem hadoop@ec2-34-207-142-34.compute-1.amazonaws.com:~/test.txt ~/File_transfer_demo/download_test/
```

In our case, we have downloaded a file named **test.txt** from **/home/hadoop** in the EMR instance to **~/File_transfer_demo/download_test/**.

```
composer117@Composer117:~/File_transfer_demo$ scp -i ~/RHEL.pem hadoop@ec2-34-207-142-34.compute-1.amazonaws.com:~/test.
txt ~/File_transfer_demo/download_test/
test.txt                                                                    100%   15     0.0KB/s   00:00
composer117@Composer117:~/File_transfer_demo$ ls download_test/
test.txt
composer117@Composer117:~/File_transfer_demo$
```

4. You can verify the downloaded file in our system once the transfer is complete.