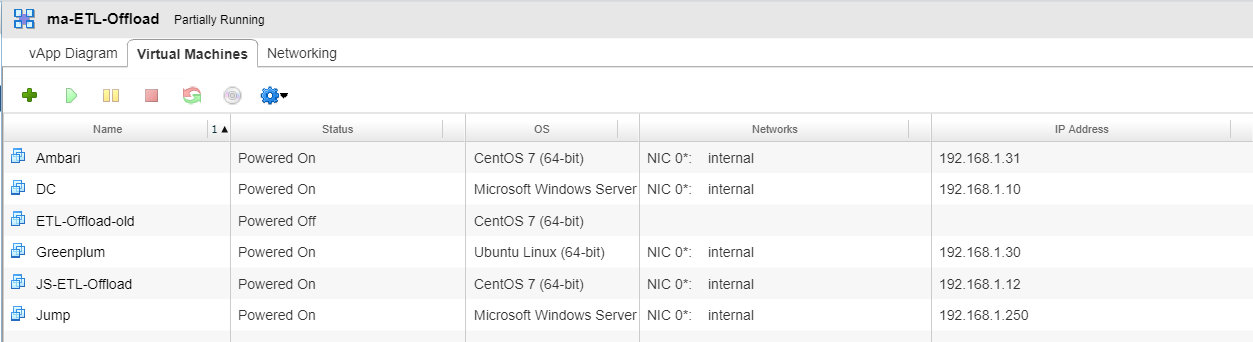
Step Accessing/starting Hadoop cluster

1. Right click on the JS-ETL-Offload VM and click on pop out console. Alternatively, you can right click on the Jump console and open the windows interface.



* 1. If the Jump console has been selected, unlock the screen by using the password Dees12345!.
  2. Click on Putty and select the ETL offload VM.

1. Enter your user name: root and password:Dees12345 to gain access to the root folder on the Virtual Machine.

Open the terminal on the ETL offload by using the search or clicking the icon on the sidebar.

4. If this is your first time accessing the server, start the Hadoop cluster. Run the following line of code in the terminal:

./start-all.sh

5. You may type the following command to make sure NameNode, DataNode, ResourceManager, and NodeManager are started and running:

jps

Lab steps

1. STEP:

Check your Hadoop environment:

1. Open the terminal.

2. Connect to the Hadoop server using SSH. When prompted for the password, enter the password for hdp user provided to you:

ssh hdp@192.168.2.20

3. Execute the command:

printenv | grep '^H'

4. Are there any Hadoop variables defined? Which ones?

5. List the folders in the home directory of the user, to have an idea of the location of different programs to be used in this lab. You should see a LAB11 directory beside other directories.

6. Execute the command:

ls

7. Execute the command:

hadoop fs -ls

8. What do you see?

9. This should be different.

2. STEP:

Get Help:

1. Execute the command: hadoop -help 2>&1 | tee Hadoop.hlp

2. Execute the command:

hadoop fs -help | tee HDFS.hlp

3. Now, list the files by executing the commands:

clear && more Hadoop.hlp clear && more HDFS.hlp

4. fs is a general file system shell which is not specific to hdfs and can be used to transfer files from other file systems into hdfs. Now, create an alias for using hadoop fs: alias hf="hadoop fs"

3. STEP:

Load data into HDFS:

In this step, we will be loading an input data file from a local directory into HDFS that we will be using in later activities:

1. This command will copy the file from local system to HDFS system:

hf -copyFromLocal ~/LAB11/shakespearehamlet.txt /warehouse/shakespearehamlet.txt

2. Display the contents of the “warehouse” directory in HDFS:

hf -ls /user/hdp/warehouse hf -ls /warehouse

The files shakespearehamlet.txt and shakespearehamletshort.txt are now in HDFS