Big Data And Hadoop

Session 12 – Assignment3

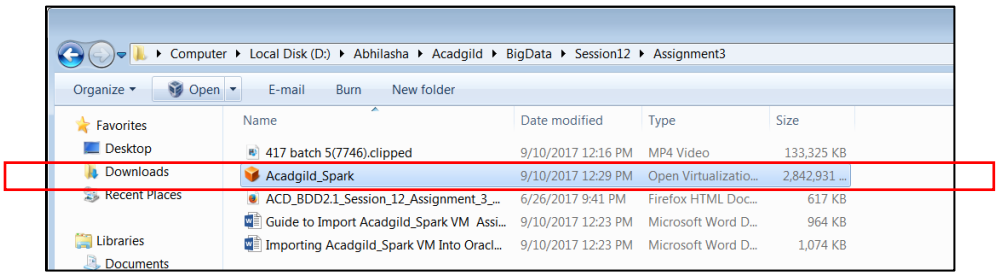
Problem Statement:

Import AcadGild VM into Virtual box and start all the Spark daemons.

Solution:

Part A: Import acadgild vm

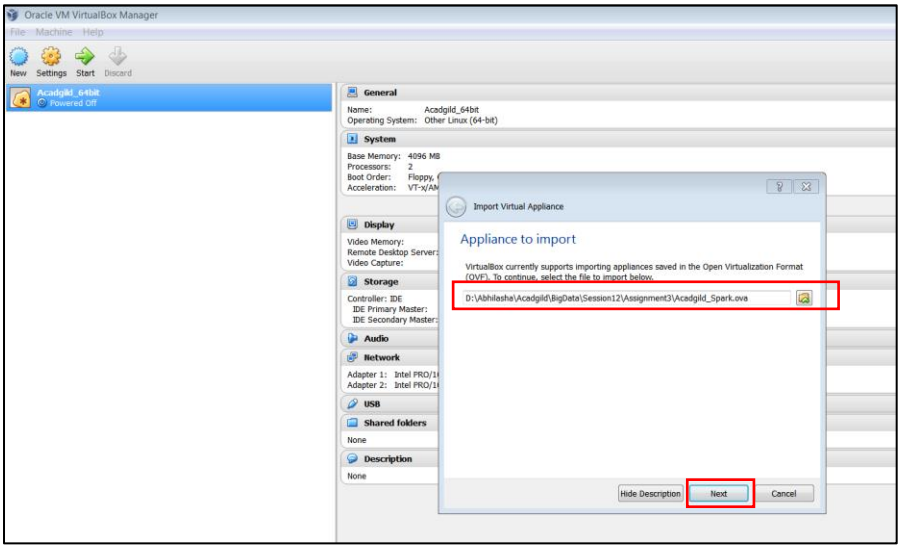
Step 1: Downloaded the acadgild vm's ova file and placed it locally, as seen in the screenshot.



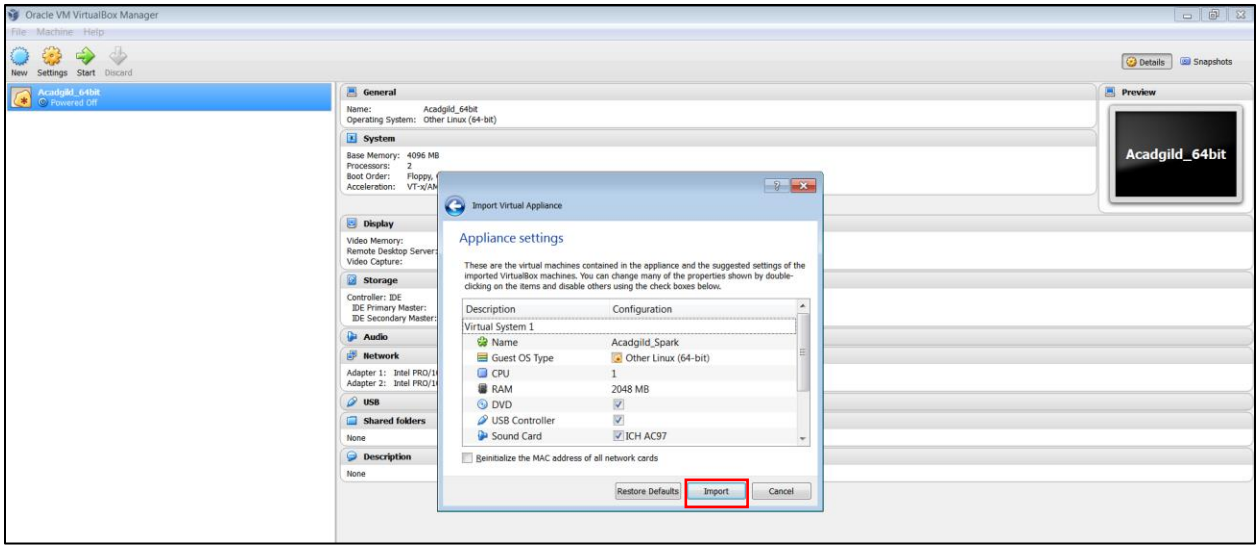
Step2: Started virtual box. Virtual box is already installed. Hence, didn't install it again.

Step3: Then went to **File -> Import appliance**

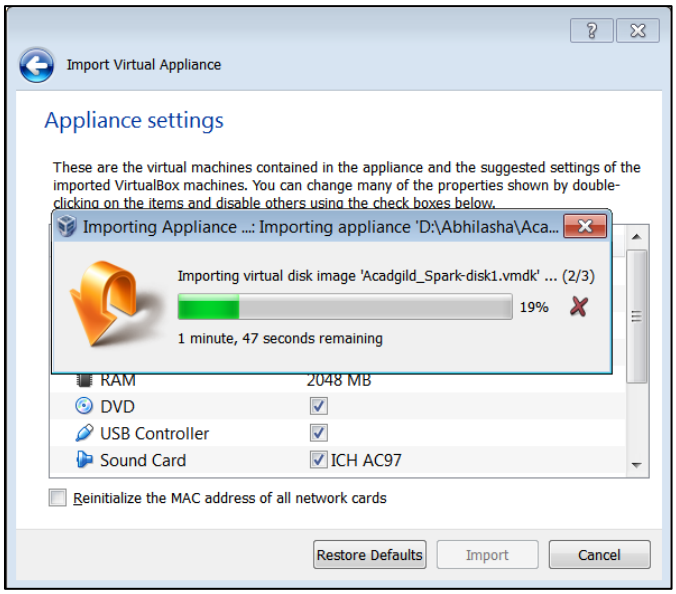
Step4: The following window appeared and we browsed the location of the ova file, as shown below. Then clicked on **next button** at the end of that window.



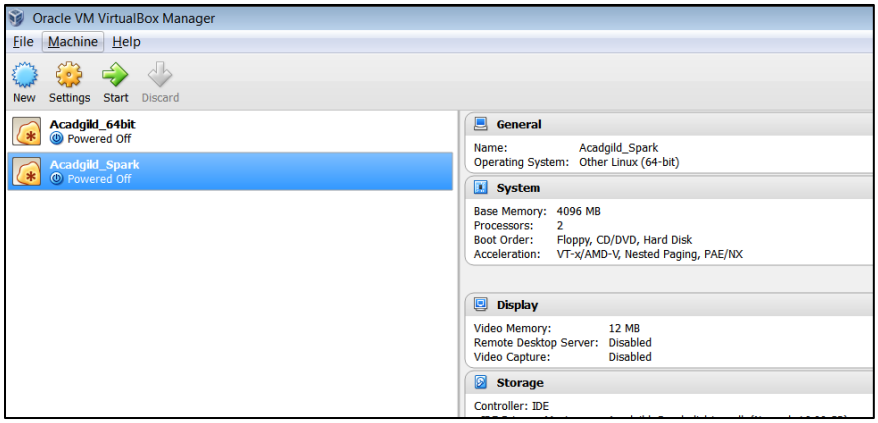
Step5: After browsing the file, clicked on **Import button** shown below.



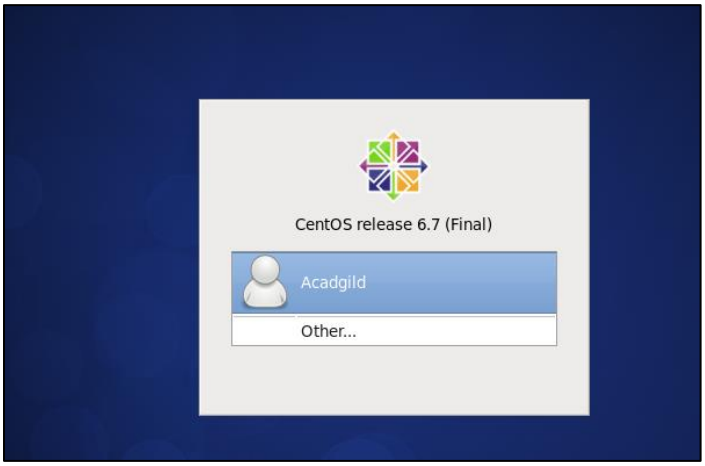
Step6: Import started.



Step7: After the import is complete, the vm appears in the window in the left side. We then start the vm.

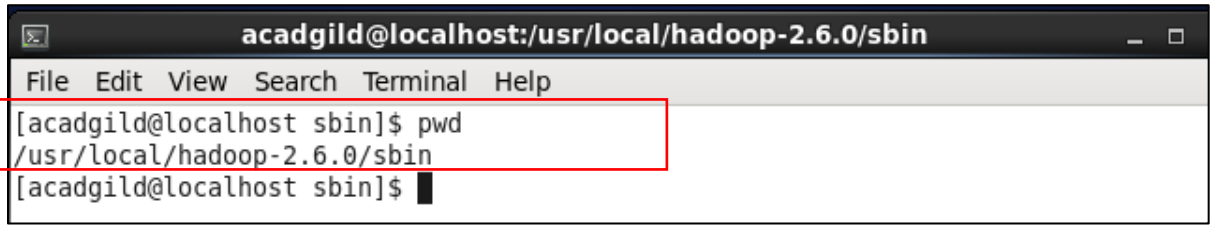


Step8: We now log into the vm as follows



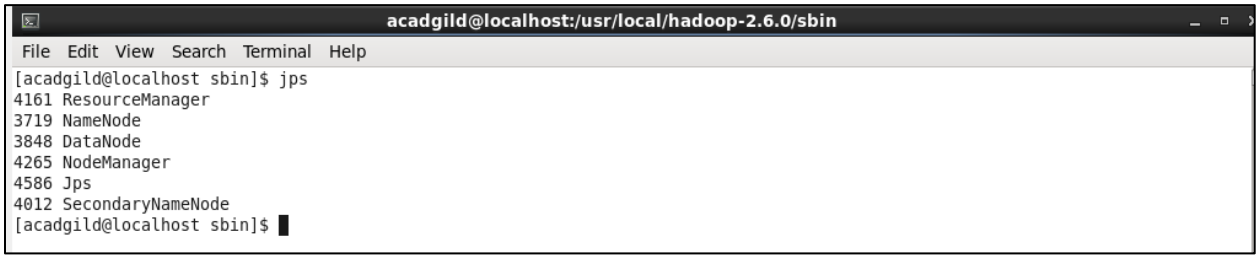
Part B: Start all the daemons

Step1: We first go to the bin directory of Hadoop. Its location is **/usr/local/Hadoop – 2.6.0/sbin**

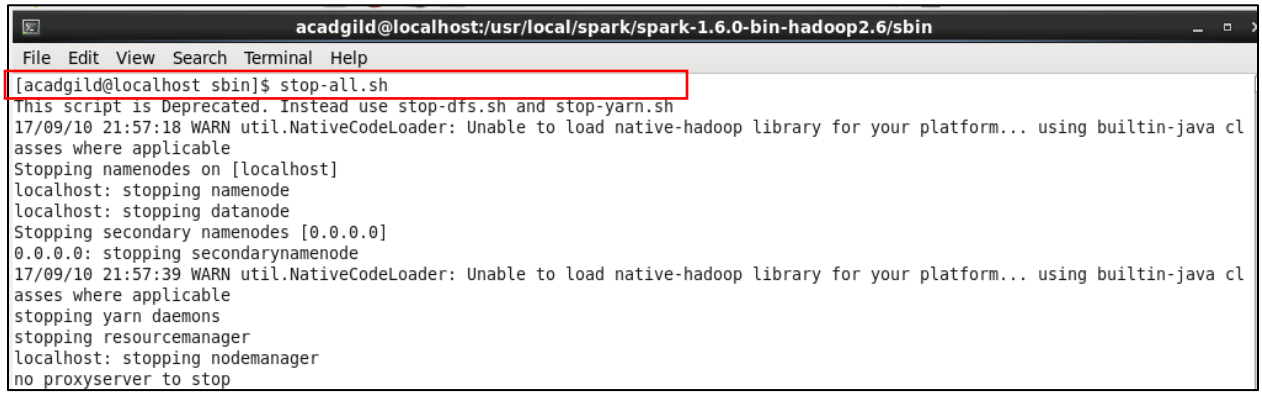


Step2: Next step, we start all the daemons of Hadoop using the command **start-all.sh** as follows

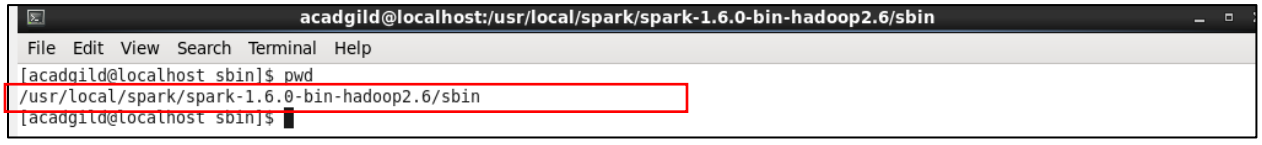
Step3: To check if all the daemons are up, we use **jps** command as follows



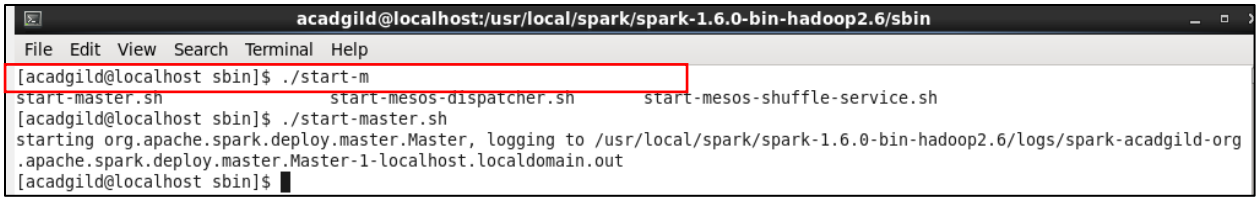
Step4: We use **stop-all.sh** command to stop these daemons



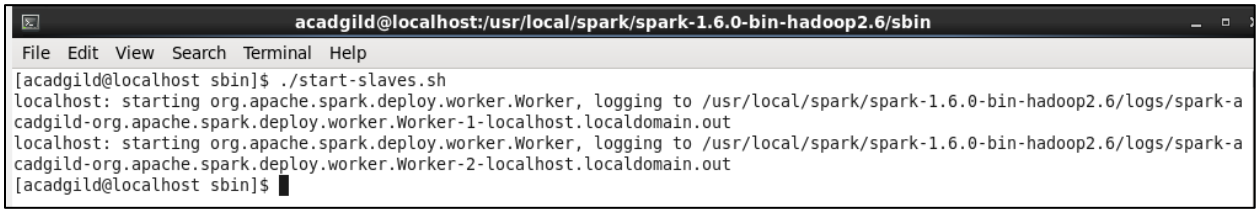
Step5: Next, we go to the bin directory of spark. Its location is **/usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin**



Step6: To start master daemon in spark, we use the command **./start-master.sh**



Step7: To start slave daemon, we use the command **./start-slaves.sh**



Step8: We then verify if all the daemons are up using **jps** command as follows: