

## Big Data And Hadoop

### Session 12 – Assignment3

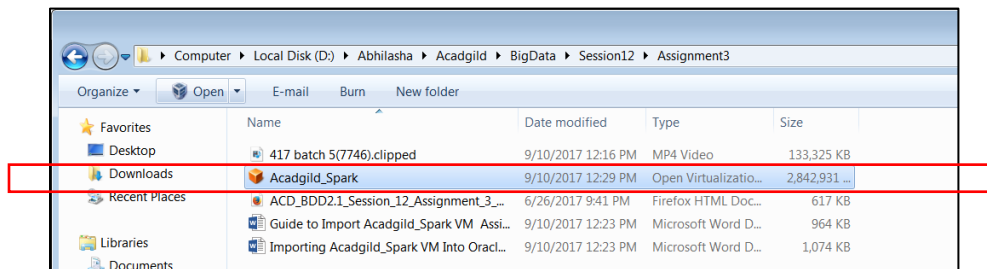
#### Problem Statement:

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

#### Solution:

##### Part A: Import acadgild vm

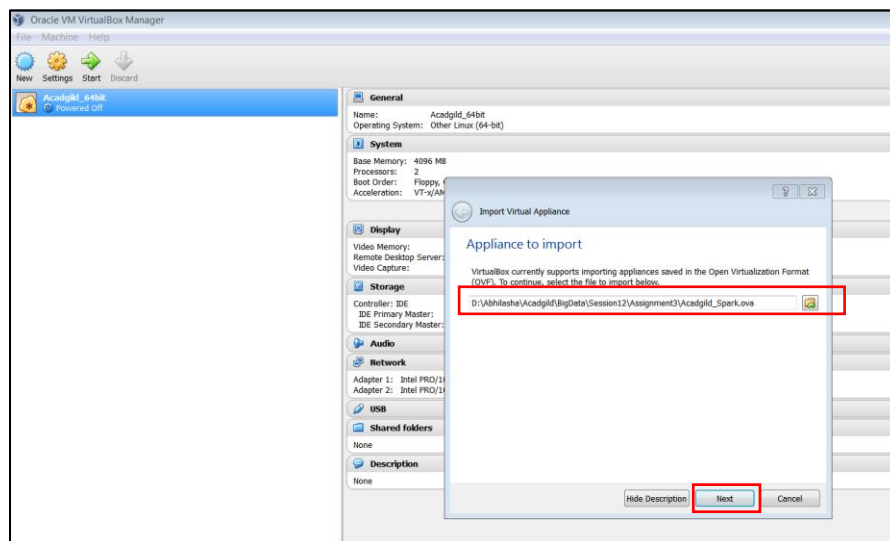
Step 1: Downloaded the acadgild vm's ova file and placed it locally, as seen in the screenshot.



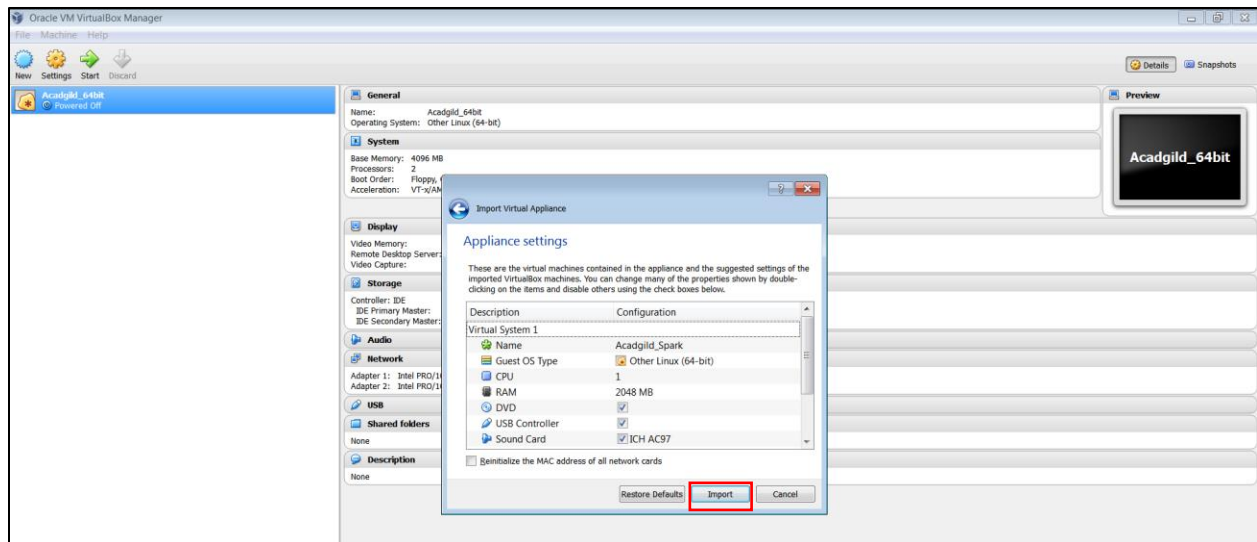
Step2: Started virtual box. Virtual box is already installed. Hence, didn't install it again.

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

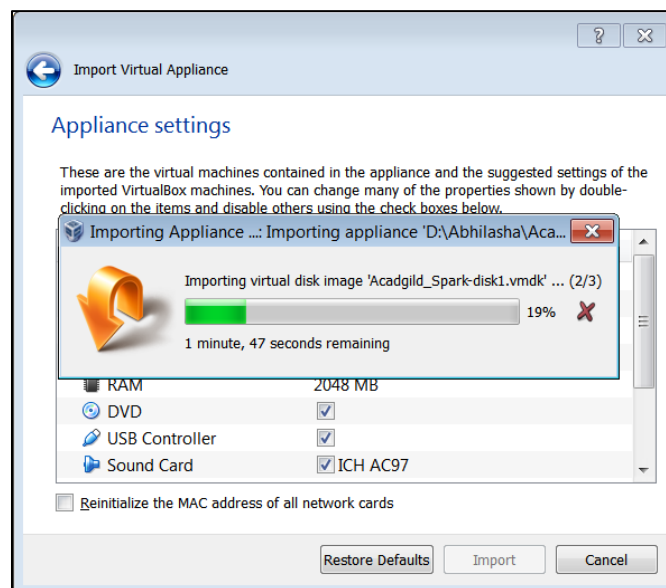
Step4: The following window appeared and we browsed the location of the ova file, as shown below. Then clicked on **next button** at the end of that window.



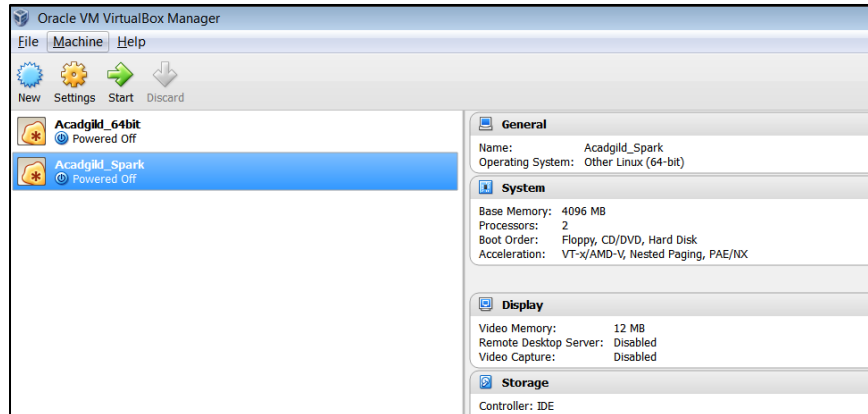
Step5: After browsing the file, clicked on **Import button** shown below.



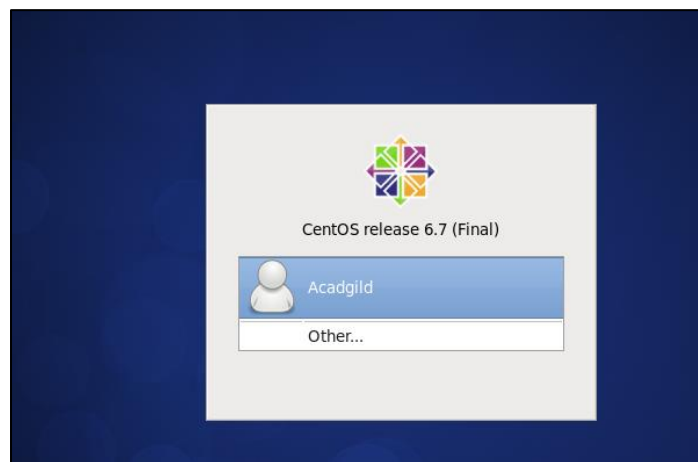
Step6: Import started.



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`

```
acadgild@localhost:/usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ pwd
/usr/local/hadoop-2.6.0/sbin
[acadgild@localhost sbin]$
```

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

```
acadmild@localhost:/usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadmild@localhost sbin]$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
17/09/10 21:51:48 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadmild-namenode-localhost.localdomain.out
localhost: starting datanode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadmild-datanode-localhost.localdomain.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadmild-secondarynamenode-localhost.localdomain.out
17/09/10 21:52:06 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop-2.6.0/logs/yarn-acadmild-resourcemanager-localhost.localdomain.out
localhost: starting nodemanager, logging to /usr/local/hadoop-2.6.0/logs/yarn-acadmild-nodemanager-localhost.localdomain.out
[acadmild@localhost sbin]$
```

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

```
acadmild@localhost:/usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadmild@localhost sbin]$ jps
4161 ResourceManager
3719 NameNode
3848 DataNode
4265 NodeManager
4586 Jps
4012 SecondaryNameNode
[acadmild@localhost sbin]$
```

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

```
acadmild@localhost:/usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
File Edit View Search Terminal Help
[acadmild@localhost sbin]$ stop-all.sh
This script is Deprecated. Instead use stop-dfs.sh and stop-yarn.sh
17/09/10 21:57:18 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Stopping namenodes on [localhost]
localhost: stopping namenode
localhost: stopping datanode
Stopping secondary namenodes [0.0.0.0]
0.0.0.0: stopping secondarynamenode
17/09/10 21:57:39 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
stopping yarn daemons
stopping resourcemanager
localhost: stopping nodemanager
no proxyserver to stop
```

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

```
acadmild@localhost:/usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
File Edit View Search Terminal Help
[acadmild@localhost sbin]$ pwd
/usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
[acadmild@localhost sbin]$
```

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

```
acadgild@localhost:usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ ./start-m
start-master.sh start-mesos-dispatcher.sh start-mesos-shuffle-service.sh
[acadgild@localhost sbin]$ ./start-master.sh
starting org.apache.spark.deploy.master.Master, logging to /usr/local/spark/spark-1.6.0-bin-hadoop2.6/logs/spark-acadgild-org
.apache.spark.deploy.master.Master-1-localhost.localdomain.out
[acadgild@localhost sbin]$
```

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

```
acadgild@localhost:usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ ./start-slaves.sh
localhost: starting org.apache.spark.deploy.worker.Worker, logging to /usr/local/spark/spark-1.6.0-bin-hadoop2.6/logs/spark-a
cadgild-org.apache.spark.deploy.worker.Worker-1-localhost.localdomain.out
localhost: starting org.apache.spark.deploy.worker.Worker, logging to /usr/local/spark/spark-1.6.0-bin-hadoop2.6/logs/spark-a
cadgild-org.apache.spark.deploy.worker.Worker-2-localhost.localdomain.out
[acadgild@localhost sbin]$
```

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

```
acadgild@localhost:usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ jps
4630 Master
5352 Jps
4731 Worker
4780 Worker
[acadgild@localhost sbin]$
```