

## PROBLEM STATEMENT

Import Acadgild Spark VM

BIG DATA  
HADOOP  
&  
SPARK  
TRAINING

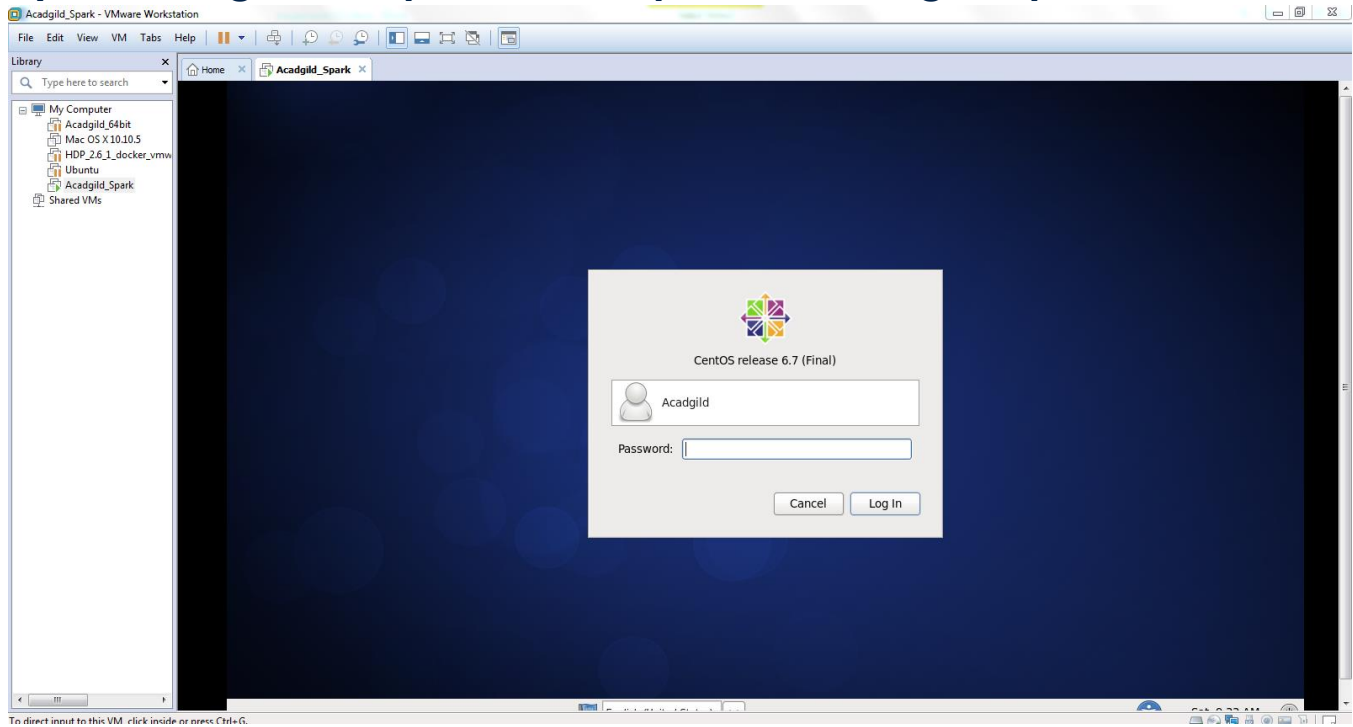
ACADGILD

ASSIGNMENT  
12.1

BY :-

SAHIL  
KHURANA

## By following the steps I have imported Acadgild Spark VM



### Starting namenode:

```
cd /usr/local/hadoop-2.6.0/sbin
```

**Command: ./hadoop-daemon.sh start namenodeemp\_details.txt**

```

Applications Places System
acadgild@localhost:/usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ cd /usr/local/hadoop-2.6.0/sbin
[acadgild@localhost sbin]$ ./hadoop-daemon.sh start namenode
starting namenode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadgild-namenode-localhost.localdomain.out
[acadgild@localhost sbin]$ jps
4231 NameNode
4300 Jps
[acadgild@localhost sbin]$
[acadgild@localhost sbin]$

```

### Starting datanode:

**Command: ./hadoop-daemon.sh start datanode**

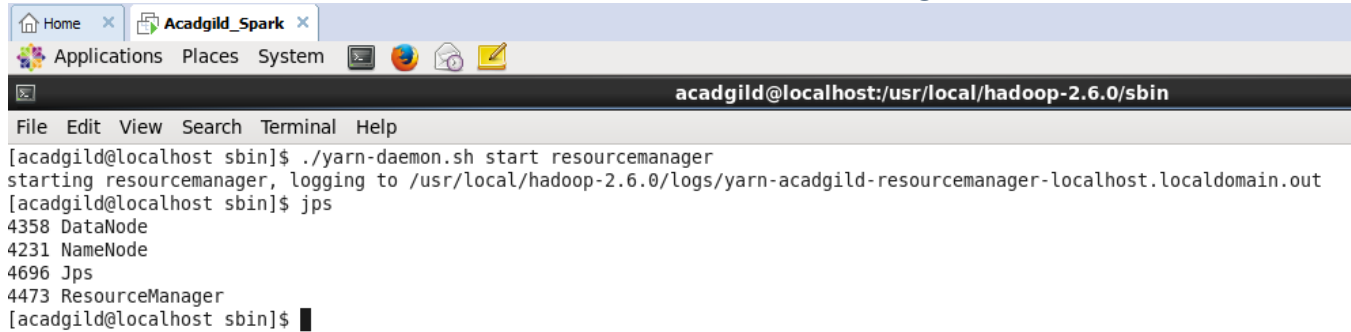
```

Home Acadgild_Spark
Applications Places System
acadgild@localhost:/usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ ./hadoop-daemon.sh start datanode
starting datanode, logging to /usr/local/hadoop-2.6.0/logs/hadoop-acadgild-datanode-localhost.localdomain.out
[acadgild@localhost sbin]$ jps
4433 Jps
4358 DataNode
4231 NameNode
[acadgild@localhost sbin]$

```

## Starting resourcemanager:

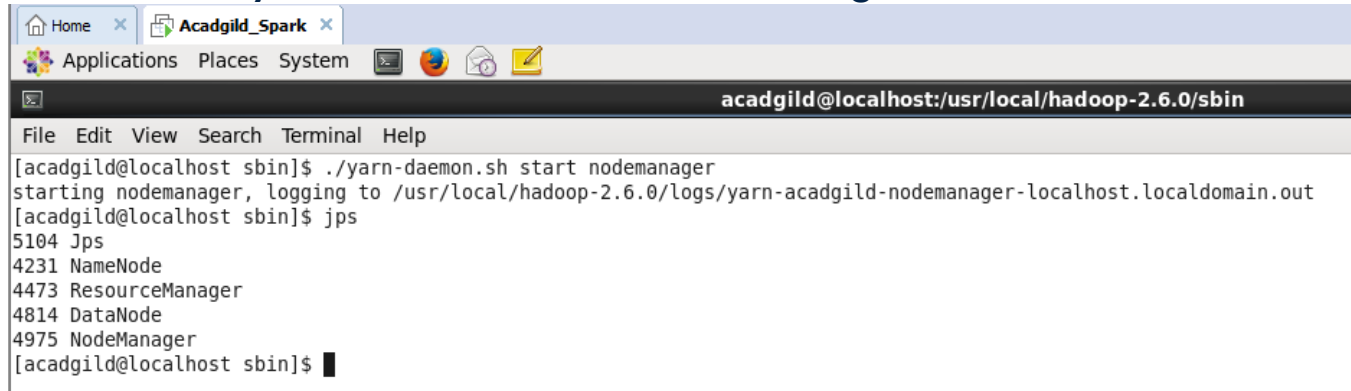
**Command: `./yarn-daemon.sh start resourcemanager`**



```
acadgild@localhost: /usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ ./yarn-daemon.sh start resourcemanager
starting resourcemanager, logging to /usr/local/hadoop-2.6.0/logs/yarn-acadgild-resourcemanager-localhost.localdomain.out
[acadgild@localhost sbin]$ jps
4358 DataNode
4231 NameNode
4696 Jps
4473 ResourceManager
[acadgild@localhost sbin]$
```

## Starting nodemanager:

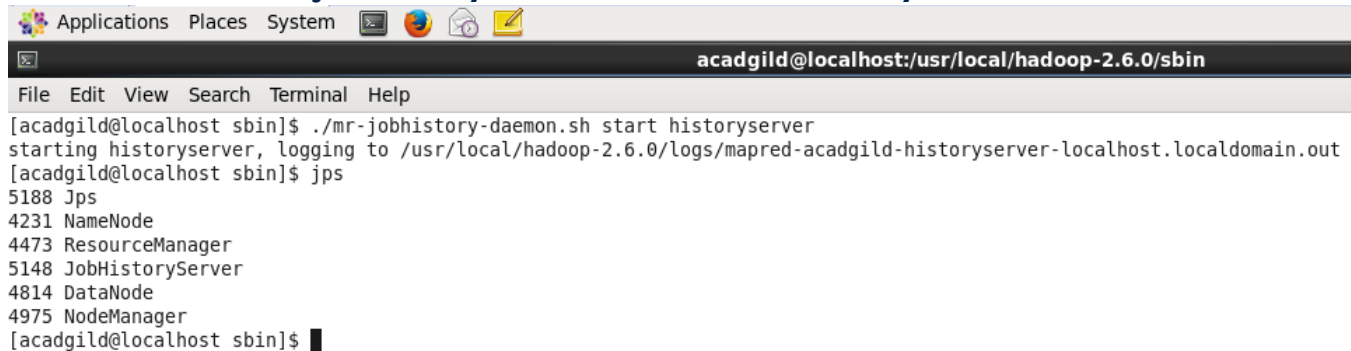
**Command: `./yarn-daemon.sh start nodemanager`**



```
acadgild@localhost: /usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ ./yarn-daemon.sh start nodemanager
starting nodemanager, logging to /usr/local/hadoop-2.6.0/logs/yarn-acadgild-nodemanager-localhost.localdomain.out
[acadgild@localhost sbin]$ jps
5104 Jps
4231 NameNode
4473 ResourceManager
4814 DataNode
4975 NodeManager
[acadgild@localhost sbin]$
```

## Starting historyserver:

**Command: `./mr-jobhistory-daemon.sh start historyserver`**



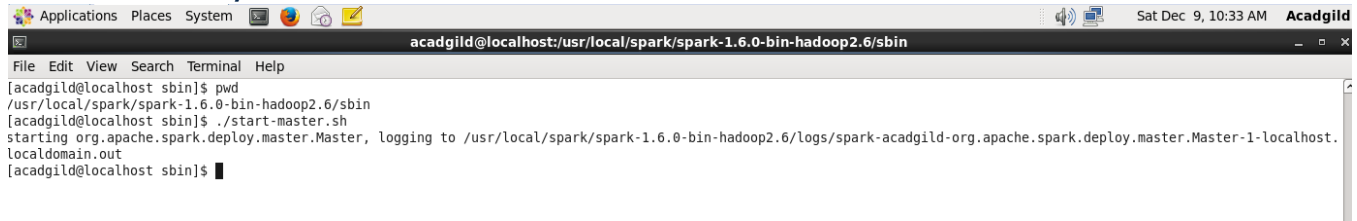
```
acadgild@localhost: /usr/local/hadoop-2.6.0/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ ./mr-jobhistory-daemon.sh start historyserver
starting historyserver, logging to /usr/local/hadoop-2.6.0/logs/mapred-acadgild-historyserver-localhost.localdomain.out
[acadgild@localhost sbin]$ jps
5188 Jps
4231 NameNode
4473 ResourceManager
5148 JobHistoryServer
4814 DataNode
4975 NodeManager
[acadgild@localhost sbin]$
```

## Starting Spark daemons Master, Workers

**Move into the sbin folder of spark installed directory  
(Here spark is installed in /usr/local/spark/ directory.)**

## Starting Master:

**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]$ pwd
/usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
[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]$
```

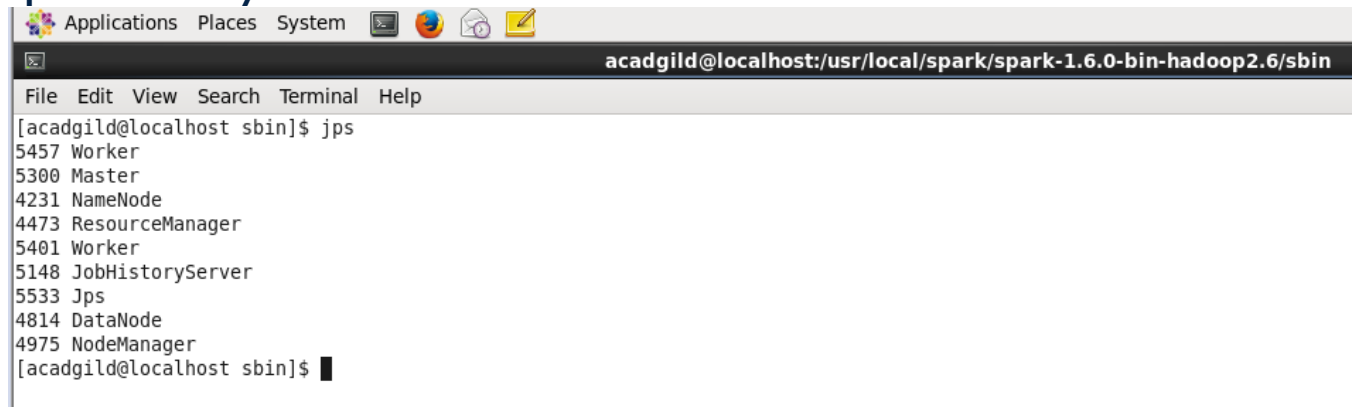
## Starting Slaves:

**Command: `./start-slaves.sh`**



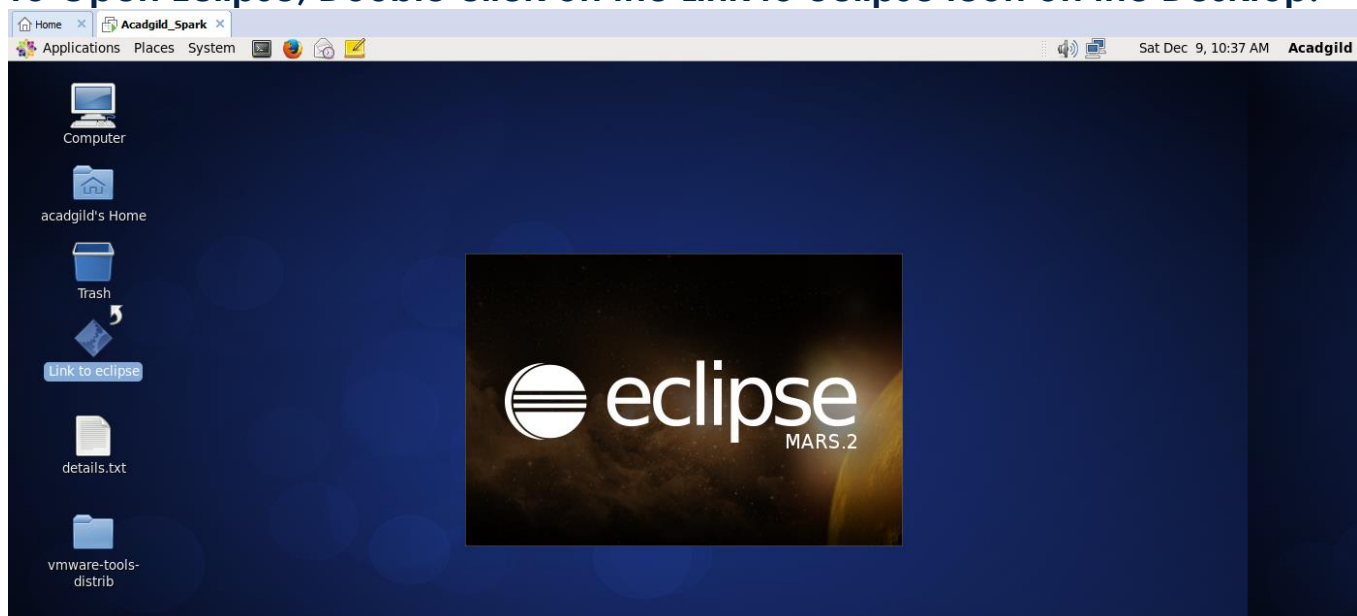
```
Acadgild_Spark
Applications Places System
acadgild@localhost:usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ ./start-slaves.sh
localhost: starting org.apache.spark.deploy.worker.Worker, logging to /usr/local/spark/spark-1.6.0-bin-hadoop2.6/logs/spark-acadgild-org.apache.spark.deploy.worker.Worker-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

**Check the health of the daemons, if all the daemons will be running then spark is ready.**



```
Applications Places System
acadgild@localhost:usr/local/spark/spark-1.6.0-bin-hadoop2.6/sbin
File Edit View Search Terminal Help
[acadgild@localhost sbin]$ jps
5457 Worker
5300 Master
4231 NameNode
4473 ResourceManager
5401 Worker
5148 JobHistoryServer
5533 Jps
4814 DataNode
4975 NodeManager
[acadgild@localhost sbin]$
```

**To Open Eclipse, Double click on the Link to eclipse icon on the Desktop.**



## The screenshot shows the Eclipse IDE's 'Workspace Launcher' dialog box. The dialog has a title bar with a gear icon and the text 'Workspace Launcher'. Inside, the section 'Select a workspace' is followed by the text: 'Eclipse stores your projects in a folder called a workspace. Choose a workspace folder to use for this session.' Below this is a text field labeled 'Workspace:' containing the path '/home/acadgild/workspace'. To the right of the text field is a 'Browse...' button. At the bottom, there is a checkbox labeled 'Use this as the default and do not ask again' which is currently unchecked. The bottom right corner of the dialog contains 'Cancel' and 'OK' buttons. The background of the image shows a Linux desktop environment with a dark blue theme, a top panel with application icons, and a sidebar with desktop icons like 'Computer', 'Home', 'Trash', and 'Link to eclipse'.

```
acagild@localhost:~/Desktop
File Edit View Search Terminal Help
Welcome to

    /\_/\
   /__\/  version 1.6.0

Using Scala version 2.10.5 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_65)
Type in expressions to have them evaluated.
Type :help for more information.
17/12/09 11:01:23 WARN Utils: Your hostname, localhost.localdomain resolves to a loopback address: 127.0.0.1; using 192.168.19.141 instead (on interface eth1)
17/12/09 11:01:23 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
17/12/09 11:01:23 WARN SparkConf:
SPARK_WORKER_INSTANCES was detected (set to '2').
This is deprecated in Spark 1.0+.

Please instead use:
- ./spark-submit with --num-executors to specify the number of executors
- Or set SPARK_EXECUTOR_INSTANCES
- spark.executor.instances to configure the number of instances in the spark config.

Spark context available as sc.
17/12/09 11:01:33 WARN Connection: BoneCP specified but not present in CLASSPATH (or one of dependencies)
17/12/09 11:01:34 WARN Connection: BoneCP specified but not present in CLASSPATH (or one of dependencies)
17/12/09 11:01:42 WARN ObjectStore: Version information not found in metastore. hive.metastore.schema.validation is not enabled so recording the schema version 1.2.0
17/12/09 11:01:42 WARN ObjectStore: Failed to get database default, returning NoSuchElementException
17/12/09 11:01:46 WARN Connection: BoneCP specified but not present in CLASSPATH (or one of dependencies)
17/12/09 11:01:46 WARN Connection: BoneCP specified but not present in CLASSPATH (or one of dependencies)
17/12/09 11:01:51 WARN ObjectStore: Version information not found in metastore. hive.metastore.schema.validation is not enabled so recording the schema version 1.2.0
17/12/09 11:01:52 WARN ObjectStore: Failed to get database default, returning NoSuchElementException
SQL context available as sqlContext.
```