Installing Spark On Ubuntu/:-

1: Download Spark from the apache spark site:-

<https://spark.apache.org/downloads.html>

2: Move spark-hadoop.tar to new folder where you want to install spark.

3: tar – xvf spark-hadoop.tar

4: cd spark folder , cd conf

5: mv spark-env.sh.template spark-env.sh

6: mv slaves.template slaves

7. vim ~/.bashrc

8: set following path

export SPARK\_HOME = /home/sears/sears/spark-bin-hadoop-2.2.3/

export PATH=$SPARK\_HOME/bin:$PATH

9:

export JAVA\_HOME = /usr/lib/jvm/jdk.1.8.0open

export PATH=$JAVA\_HOME/bin:$PATH

10: Go to the /home/sears/sears/spark-bin-hadoop-2.2.3/sbin

Do following ssh connection now

1. sudo apt-get install ssh
2. ssh [user@127.0.0.1](mailto:user@127.0.0.1) –p 22
3. ssh [user@127.0.0.1](mailto:user@127.0.0.1)
4. sudo nano /etc/ssh/ssh\_config
5. sudo service ssh force-reload
6. ssh user@127.0.0.1 –p 2222

Change password of the sudo root if you don’t know

sudo passwd root

cent-os

generate ssh key

11. Type ./start-all.sh

12 Now Type spark-shell

Now enjoy the SPARK connection

Creating user

sudo addgroup hadoop

sudo adduser –ingroup hadoop hduser

sudo adduser hduser sudo

Creating ssh connection:-

Sudo apt-get install ssh

Su hduser

Ssh-keygen –t rsa –P “”

enter

cat $HOME/.ssh/id\_rsa.pub >> $HOME/.ssh/authorized\_keys

chmod 777 $HOME/.ssh/authorized\_keys

Installing multimode Hadoop and Spark

sudo apt-get update

sudo apt-get install default-jdk

java –version

cd /usr/local/

Wget [**http://mirrors.estointernet.in/apache/hadoop/common/hadoop-2.6.5/hadoop-2.6.5.tar.gz**](http://mirrors.estointernet.in/apache/hadoop/common/hadoop-2.6.5/hadoop-2.6.5.tar.gz)

Tar –xvf Hadoop-2.6.5.tar.gz

Remove /usr/local/Hadoop

And mv hadoop-2.6.5 hadoop

1: open vi ~/.bashrc :- vonfiguration file for all linux softwares mysql,hadooop,spark,hive,casssendra

Add following for Hadoop

1. #Hadoop variables
2. export JAVA\_HOME=/usr/lib/jvm/jdk/jdk1.7.0\_71
3. export HADOOP\_INSTALL=/usr/hadoop
4. export PATH=$PATH:$HADOOP\_INSTALL/bin
5. export PATH=$PATH:$HADOOP\_INSTALL/sbin
6. export HADOOP\_MAPRED\_HOME=$HADOOP\_INSTALL
7. export HADOOP\_COMMON\_HOME=$HADOOP\_INSTALL
8. export HADOOP\_HDFS\_HOME=$HADOOP\_INSTALL
9. export YARN\_HOME=$HADOOP\_INSTALL

2.Open core-site.xml

1. **<configuration>**
2. **<property>**
3. **<name>**fs.default.name**</name>**
4. **<value>**hdfs://hadoop-master:9000**</value>**
5. **</property>**
6. **<property>**
7. **<name>**dfs.permissions**</name>**
8. **<value>**false**</value>**
9. **</property>**
10. **</configuration>**

3.Open mapred-site.xml

1. **<configuration>**
2. **<property>**
3. **<name>**mapred.framework.name**</name>**
4. **<value>**yarn**</value>**
5. **</property>**
6. **</configuration>**

5. Open /etc/local/Hadoop/bin/Hadoop\_env.sh

JAVA\_HOME=/usr/local/jvm/openjdk

Javac

* + path

JAVA\_HOME =Readlink –f path

6:Open hdfs-site.xml

1. **<configuration>**
2. **<property>**
3. **<name>**dfs.data.dir**</name>**
4. **<value>**usr/hadoop/dfs/name/data**</value>**
5. **<final>**true**</final>**
6. **</property>**
7. **<property>**
8. **<name>**dfs.name.dir**</name>**
9. **<value>**usr/hadoop/dfs/name**</value>**
10. **<final>**true**</final>**
11. **</property>**
12. **<property>**
13. **<name>**dfs.replication**</name>**
14. **<value>**1**</value>**
15. **</property>**
16. **</configuration>**

Open yarn-site

Add following

<?xml version="1.0">

<configuration>

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>

<value>org.apache.hadoop.mapred.ShuffleHandler</value>

</property>

</configuration>

Very important

Now type

Hadoop namenode -format

$HADOOP\_INSTALL/sbin/stop-all.sh

$HADOOP\_INSTALL/abin/start-all.sh

If connection refused error occurred:-

Do telnet localhost 9000

Then if it is refused then type

ifconfig

telnet 192.168.0.103(ip) 9000

if it is working add this ip adrees at localhost place

./hadoop-daemon.sh start namenode

./hadoop-daemon.sh start datanode

**localhost**:**50070/dfshealth.html**

Enjoy Hadoop Singlenode connection

All the best

Installing firefox :-

sudo apt-get install firefox

If ip address has changed:-

sudo ssh-keygen – f “/root/.ssh/known\_hosts” -R 192.168.0.103

sudo ssh-keygen – f “/root/.ssh/known\_hosts” -R 172.29.7.165

/usr/local/hadoop/etc/hadoop/

sudo –u hdheena –i

hadoop namenode –format

/usr/local/hadoop/sbin/start-all,sh

hadoop fs –ls /user/heena

Id having error to put file

Use hadoop namenode – format :--don’t format the namenode unless and until it is exetremly important

ConnectionRefused Error:-

hadoop/sbin/start-dfs.sh

Hive Installation:-

wget <http://www-us.apache.org/dist/hive/hive-2.1.0/apache-hive-2.1.0-bin.tar.gz>

sudo tar xvzf apache-hive-2.1.0-bin.tar.gz -C /usr/local

**vi ~/.bashrc**

export HIVE\_HOME=/usr/local/apache-hive-2.1.0-bin

export HIVE\_CONF\_DIR=/usr/local/apache-hive-2.1.0-bin/conf

export PATH=$HIVE\_HOME/bin:$PATH

export CLASSPATH=$CLASSPATH:/usr/local/hadoop/lib/\*:.

export CLASSPATH=$CLASSPATH:/usr/local/apache-hive-2.1.0-bin/lib/\*:export DERBY\_HOME=/usr/local/db-derby-10.13.1.1-bin

export PATH=$PATH:$DERBY\_HOME/bin

export CLASSPATH=$CLASSPATH:$DERBY\_HOME/lib/derby.jar:$DERBY\_HOME/lib/derbytools.jar

echo $HADOOP\_HOME

/usr/local/hadoop

hduser@laptop:~$ hdfs dfs -mkdir /user/hive/warehouse

hduser@laptop:~$ hdfs dfs -chmod g+w /tmp

hduser@laptop:~$ hdfs dfs -chmod g+w /user/hive/warehouse

cd $HIVE\_HOME/conf

sudo cp hive-env.sh.template hive-env.sh

vi  **hive-env.sh**

export HADOOP\_HOME=/usr/local/hadoop

$ cd /tmp

$ wget http://archive.apache.org/dist/db/derby/db-derby-10.13.1.1/db-derby-10.13.1.1-bin.tar.gz

$ sudo tar xvzf db-derby-10.13.1.1-bin.tar.gz -C /usr/local

sudo mkdir $DERBY\_HOME/data

cd $HIVE\_HOME/conf

sudo cp hive-default.xml.template hive-site.xml

vi hive-site.xml

<property>

<name>javax.jdo.option.ConnectionURL</name>

<value>jdbc:derby:;databaseName=metastore\_db;create=true</value>

<description>

JDBC connect string for a JDBC metastore.

To use SSL to encrypt/authenticate the connection, provide database-specific SSL flag in the connection URL.

For example, jdbc:postgresql://myhost/db?ssl=true for postgres database.

</description>

</property>

**Create new vi jpox.properties**

**Add following**

javax.jdo.PersistenceManagerFactoryClass =

org.jpox.PersistenceManagerFactoryImpl

org.jpox.autoCreateSchema = false

org.jpox.validateTables = false

org.jpox.validateColumns = false

org.jpox.validateConstraints = false

org.jpox.storeManagerType = rdbms

org.jpox.autoCreateSchema = true

org.jpox.autoStartMechanismMode = checked

org.jpox.transactionIsolation = read\_committed

javax.jdo.option.DetachAllOnCommit = true

javax.jdo.option.NontransactionalRead = true

javax.jdo.option.ConnectionDriverName = org.apache.derby.jdbc.ClientDriver

javax.jdo.option.ConnectionURL = jdbc:derby://hadoop1:1527/metastore\_db;create = true

javax.jdo.option.ConnectionUserName = APP

javax.jdo.option.ConnectionPassword = mine

sudo chown -R hduser:hadoop apache-hive-2.1.0-bin

schematool -dbType derby -initSchema

echo $HIVE\_HOME

$HIVE\_HOME/bin/hive

**hive>Link for reference** [**https://www.bogotobogo.com/Hadoop/BigData\_hadoop\_Hive\_Install\_On\_Ubuntu\_16\_04.php**](https://www.bogotobogo.com/Hadoop/BigData_hadoop_Hive_Install_On_Ubuntu_16_04.php)

Heena

Installing HBASE ,Cassendra and MySQL

How to find localhost in vmware

vi /etc/resolv.conf

hostname –I

Download Putty :-

<https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>

download msi version

in virtual machine type ifconfig

* + 192.168.240.130

HostName :- 192.168.240.130

Port 22

Login id :- username

Password :- respective user password

Putty connection for vmware

https://www.cyberciti.biz/faq/centos-ssh/

Installing java

1. Download java rpm package from https://www.oracle.com/technetwork/java/javase/downloads/index.html (**jdk-8-linux-x64.rpm**)
2. Install from the rpm. (**rpm -Uvh jdk-8-linux-x64.rpm**)
3. Open **/etc/profile**, and set the java paths, save it.
4. Check the java installation path, and java version, with the commands: **which java**, **java -version**