25 instructions to Install Hadoop 2.6 on Ubuntu Server 14.04- LTS.

1) (Execute all the commands as root user) #apt-get update
2) # apt-get install default-jdk
3) # java –version (Type this command)
4) # ssh-keygen -t rsa -P ''
   (Press enter after this don't put the file name)
5) # cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
6) # wget http://mirrors.sonic.net/apache/hadoop/common/hadoop-2.6.0/hadoop-2.6.0.tar.gz
7) # tar xvzf hadoop-2.6.0.tar.gz
8) # mv hadoop-2.6.0 /usr/local/hadoop
9) # update-alternatives --config java

   (You will get below output)

   root@hadoop:~# update-alternatives --config java

   There is only one alternative in link group java (providing /usr/bin/java): **/usr/lib/jvm/java-7-openjdk-amd64**/jre/bin/java Nothing to configure.

10) # nano ~/.bashrc

   (Copy paste below things in the file and replace red line with green line if they are not same)

```
#HADOOP VARIABLES START
export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64
export HADOOP_INSTALL=/usr/local/hadoop
export PATH=$PATH:$HADOOP_INSTALL/bin
export PATH=$PATH:$HADOOP_INSTALL/sbin
export HADOOP_MAPRED_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_HOME=$HADOOP_INSTALL
export HADOOP_HDFS_HOME=$HADOOP_INSTALL
export YARN_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_INSTALL/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_INSTALL/lib"
#HADOOP VARIABLES END
```

   Now we need to edit some file. All the files are in below directory -

11) # cd /usr/local/hadoop/etc/hadoop
12) # ls

   You will see many files here ---

13) # nano hadoop-env.sh

Copy paste below line at the end of the above file. (Replace red line with green line if they are not same, just like we did above)

export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64

14) # nano core-site.xml

Open the file and enter the following in between the <configuration></configuration> tag

```
<property>
 <name>hadoop.tmp.dir</name>
 <value>/app/hadoop/tmp</value>
 <description>A base for other temporary directories.</description>
</property>
<property>
 <name>fs.default.name</name>
 <value>hdfs://localhost:54310</value>
 <description>The name of the default file system.  A URI whose
 scheme and authority determine the FileSystem implementation.  The
 uri's scheme determines the config property (fs.SCHEME.impl) naming
 the FileSystem implementation class.  The uri's authority is used to
 determine the host, port, etc. for a filesystem.</description>
</property>
```

15) Copy **mapred-site.xml.template** file to **mapred-site.xml**

```
# cp /usr/local/hadoop/etc/hadoop/mapred-site.xml.template
/usr/local/hadoop/etc/hadoop/mapred-site.xml
```

16) # nano mapred-site.xml

Open the file and enter the following in between the <configuration></configuration> tag

```
<property>
 <name>mapred.job.tracker</name>
 <value>localhost:54311</value>
 <description>The host and port that the MapReduce job tracker runs
 at.  If "local", then jobs are run in-process as a single map
 and reduce task.
 </description>
</property>
```

(I hope you are still this **/usr/local/hadoop/etc/Hadoop** directory)

17) # mkdir -p /usr/local/hadoop_store/hdfs/namenode
18) # mkdir -p /usr/local/hadoop_store/hdfs/datanode

19) # nano hdfs-site.xml

Open the file and enter the following in between the <configuration></configuration> tag

```
<property>
 <name>dfs.replication</name>
 <value>1</value>
 <description>Default block replication.
 The actual number of replications can be specified when the file is created.
 The default is used if replication is not specified in create time.
 </description>
</property>
<property>
 <name>dfs.namenode.name.dir</name>
 <value>file:/usr/local/hadoop_store/hdfs/namenode</value>
</property>
<property>
 <name>dfs.datanode.data.dir</name>
 <value>file:/usr/local/hadoop_store/hdfs/datanode</value>
</property>
```

20) # hdfs namenode -format

(Do this only once, if this command is executed again after Hadoop has been used, it'll destroy all the data on the Hadoop file system. Sometimes this command do not work in 1st attempt. So change the user and again change it back to 'root'.)

Below commands are to test the successful installation of Hadoop.

21) # start-dfs.sh (Type YES both the time when asked)
22) # start-yarn.sh
23) # jps
24) # netstat -plten | grep java
25) # stop-all.sh

Reference Links ---

1) https://www.digitalocean.com/community/tutorials/how-to-install-hadoop-on-ubuntu-13-10

2) http://www.bogotobogo.com/Hadoop/BigData_hadoop_Install_on_ubuntu_single_node_cluster.php