

HADOOP - 2.6.0 INSTALLATION GUIDE

By. Mr.Gopal Krishna

Sr. Hadoop Architect

STEP 1: First we need to update the “**System Software Repositories**” using below command [**sudo apt-get update**]

```
gopalkrishna@ubuntu: ~  
gopalkrishna@ubuntu:~$  
gopalkrishna@ubuntu:~$ sudo apt-get update  
Hit http://us.archive.ubuntu.com trusty/universe amd64 Packages  
Hit http://us.archive.ubuntu.com trusty/multiverse amd64 Packages  
Hit http://us.archive.ubuntu.com trusty/main i386 Packages  
Hit http://us.archive.ubuntu.com trusty/restricted i386 Packages  
Hit http://us.archive.ubuntu.com trusty/universe i386 Packages  
Hit http://us.archive.ubuntu.com trusty/multiverse i386 Packages  
Hit http://us.archive.ubuntu.com trusty/main Translation-en  
Hit http://us.archive.ubuntu.com trusty/multiverse Translation-en  
Hit http://us.archive.ubuntu.com trusty/restricted Translation-en  
Hit http://us.archive.ubuntu.com trusty/universe Translation-en  
Ign http://us.archive.ubuntu.com trusty/main Translation-en_US  
Ign http://us.archive.ubuntu.com trusty/multiverse Translation-en_US  
Ign http://us.archive.ubuntu.com trusty/restricted Translation-en_US  
Ign http://us.archive.ubuntu.com trusty/universe Translation-en_US  
Fetched 6,394 kB in 30s (209 kB/s)  
Reading package lists... Done  
gopalkrishna@ubuntu:~$
```

STEP 2: Install **java-1.7** version using below command. [This is a mandate step for Hadoop Installation]

sudo apt-get install openjdk-7-jdk

Before Java Installation:

```
gopalkrishna@ubuntu:~$ java -version
The program 'java' can be found in the following packages:
 * default-jre
 * gcj-4.8-jre-headless
 * openjdk-7-jre-headless
 * gcj-4.6-jre-headless
 * openjdk-6-jre-headless
Try: sudo apt-get install <selected package>
gopalkrishna@ubuntu:~$
```

For Java Installation

```
gopalkrishna@ubuntu: ~
gopalkrishna@ubuntu:~$
gopalkrishna@ubuntu:~$ sudo apt-get install openjdk-7-jdk
```

```
Adding debian:Verisign_Class_3_Public_Primary_Certification_Authority.pem
Adding debian:Equifax_Secure_Global_eBusiness_CA.pem
Adding debian:Verisign_Class_3_Public_Primary_Certification_Authority_-_G3.pem
Adding debian:AffirmTrust_Premium.pem
Adding debian:DigiCert_Global_Root_G2.pem
Adding debian:EBG_Elektronik_Sertifika_Hizmet_Saglayicisi.pem
Adding debian:E-Tugra_Certification_Authority.pem
Adding debian:D-TRUST_Root_Class_3_CA_2_EV_2009.pem
Adding debian:Verisign_Class_3_Public_Primary_Certification_Authority_2.pem
done.
Setting up libatk-wrapper-java-jni:amd64 (0.30.4-4) ...
Setting up libgnome2-bin (2.32.1-4ubuntu1) ...
Setting up libgnome2-0:amd64 (2.32.1-4ubuntu1) ...
Processing triggers for libc-bin (2.19-0ubuntu6.7) ...
Processing triggers for ca-certificates (20141019ubuntu0.14.04.1) ...
Updating certificates in /etc/ssl/certs... 0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
done.
gopalkrishna@ubuntu:~$
```

After Java Installation:

```
gopalkrishna@ubuntu:~$ java -version
java version "1.7.0_121"
OpenJDK Runtime Environment (IcedTea 2.6.8) (7u121-2.6.8-1ubuntu0.14.04.3)
OpenJDK 64-Bit Server VM (build 24.121-b00, mixed mode)
gopalkrishna@ubuntu:~$
```

STEP 3: Install ssh (secure shell) using below command

sudo apt-get install ssh

Before ssh installation

```
gopalkrishna@ubuntu: ~  
gopalkrishna@ubuntu:~$  
gopalkrishna@ubuntu:~$  
gopalkrishna@ubuntu:~$ ssh localhost  
ssh: connect to host localhost port 22: Connection refused  
gopalkrishna@ubuntu:~$
```

For ssh installation

```
gopalkrishna@ubuntu:~$  
gopalkrishna@ubuntu:~$ sudo apt-get install ssh
```

```
Adding debian:AffirmTrust_Premium.pem  
Adding debian:DigiCert_Global_Root_G2.pem  
Adding debian:EBG_Elektronik_Sertifika_Hizmet_Saglayicisi.pem  
Adding debian:E-Tugra_Certification_Authority.pem  
Adding debian:D-TRUST_Root_Class_3_CA_2_EV_2009.pem  
Adding debian:Verisign_Class_3_Public_Primary_Certification_Authority_2.pem  
done.  
Setting up libatk-wrapper-java-jni:amd64 (0.30.4-4) ...  
Setting up libgnome2-bin (2.32.1-4ubuntu1) ...  
Setting up libgnome2-0:amd64 (2.32.1-4ubuntu1) ...  
Processing triggers for libc-bin (2.19-0ubuntu6.7) ...  
Processing triggers for ca-certificates (20141019ubuntu0.14.04.1) ...  
Updating certificates in /etc/ssl/certs... 0 added, 0 removed; done.  
Running hooks in /etc/ca-certificates/update.d...  
done.  
done.  
gopalkrishna@ubuntu:~$
```

After ssh installation

```
gopalkrishna@ubuntu:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is bb:a7:0c:37:23:64:5f:ac:2c:59:32:56:69:30:03:d8.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
gopalkrishna@localhost's password:
Welcome to Ubuntu 14.04.4 LTS (GNU/Linux 4.2.0-27-generic x86_64)

 * Documentation:  https://help.ubuntu.com/

274 packages can be updated.
0 updates are security updates.

New release '16.04.1 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
```

Secure Shell (SSH) Communication
prompting for **PASSWORD**

```
Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

gopalkrishna@ubuntu:~$
```

To Achieve the Password Less SSH Communication, enter the below commands at any terminal:

```
gopalkrishna@ubuntu:~$ ssh localhost
```

```
gopalkrishna@ubuntu:~$ ssh-keygen -t dsa -P "" -f ~/.ssh/id_dsa
```

```
gopalkrishna@ubuntu:~$ cat ~/.ssh/id_dsa.pub >> ~/.ssh/authorized_keys
```

```
gopalkrishna@ubuntu:~$ ssh localhost
Welcome to Ubuntu 14.04.4 LTS (GNU/Linux 4.2.0-27-generic x86_64)

 * Documentation:  https://help.ubuntu.com/

273 packages can be updated.
168 updates are security updates.

New release '16.04.1 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Thu Feb 16 05:07:50 2017 from localhost
gopalkrishna@ubuntu:~$
```

Secure Shell (SSH) Communication
WITHOUT prompting for **PASSWORD**

STEP 4: To Install eclipse IDE, use the below command

sudo apt-get install eclipse

Before eclipse install

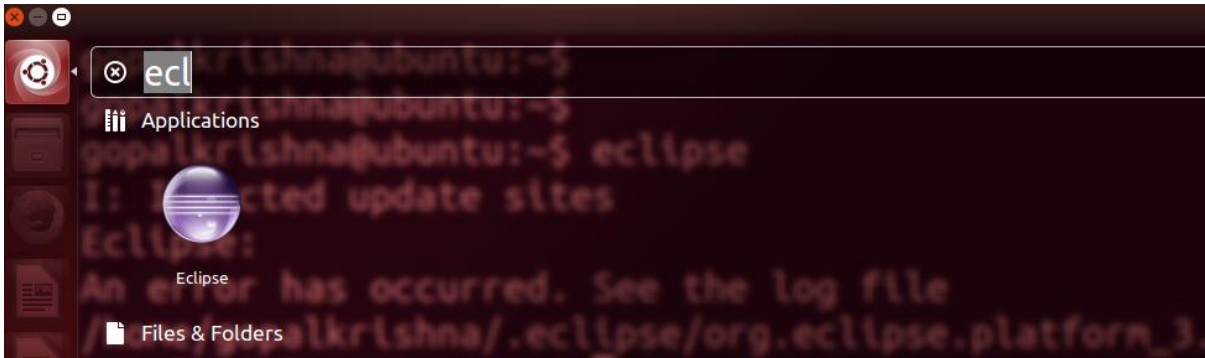
```
gopalkrishna@ubuntu:~$  
gopalkrishna@ubuntu:~$ eclipse  
The program 'eclipse' is currently not installed. You can install it by typing:  
sudo apt-get install eclipse-platform  
gopalkrishna@ubuntu:~$
```

For Eclipse Installation

```
gopalkrishna@ubuntu:~$  
gopalkrishna@ubuntu:~$ sudo apt-get install eclipse
```

```
Setting up libkxml2-java (2.3.0+ds1-2) ...  
Processing triggers for ureadahead (0.100.0-16) ...  
Setting up jarwrapper (0.45ubuntu1) ...  
Setting up sat4j (2.3.2-1) ...  
Setting up libgeronimo-jpa-2.0-spec-java (1.1-2) ...  
Setting up libosgi-compendium-java (4.3.0-1) ...  
Setting up libgeronimo-osgi-support-java (1.0-2) ...  
Setting up libfelix-gogo-runtime-java (0.10.0-2) ...  
Setting up libfelix-gogo-command-java (0.12.0-2) ...  
Setting up libfelix-gogo-shell-java (0.10.0-2) ...  
Setting up eclipse-rcp (3.8.1-5.1) ...  
Setting up eclipse-platform (3.8.1-5.1) ...  
Setting up eclipse-jdt (3.8.1-5.1) ...  
Setting up eclipse-pde (3.8.1-5.1) ...  
Setting up eclipse (3.8.1-5.1) ...  
Setting up libfelix-shell-java (1.4.2-3) ...  
Setting up libfelix-utils-java (1.1.0-3) ...  
Processing triggers for libc-bin (2.19-0ubuntu6.7) ...  
gopalkrishna@ubuntu:~$
```

After Eclipse Installation



STEP 5: Download Hadoop-2.6.0 version tar ball from Apache Mirrors

<http://apache.mirrors.tds.net/hadoop/core/>

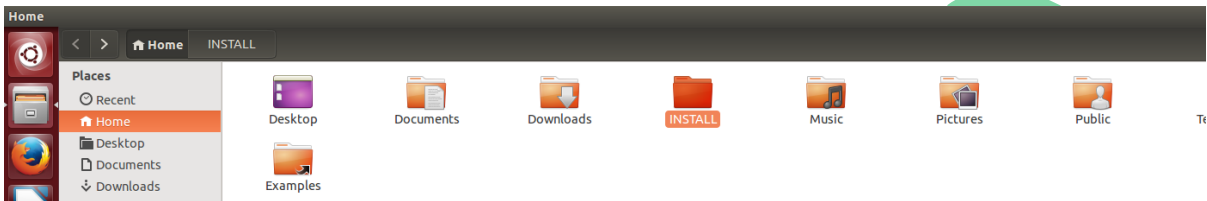


STEP 6: Create 'INSTALL' directory in '/home/gopalkrishna' folder

```

gopalkrishna@ubuntu:~$ mkdir INSTALL
gopalkrishna@ubuntu:~$
gopalkrishna@ubuntu:~$ cd INSTALL/
gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$ ll
total 8
drwxrwxr-x  2 gopalkrishna gopalkrishna 4096 Mar 10 01:40 ./
drwxr-xr-x 20 gopalkrishna gopalkrishna 4096 Mar 10 01:40 ../
gopalkrishna@ubuntu:~/INSTALL$

```



STEP 7: Copy the Downloaded 'hadoop-2.6.0.tar.gz' file to **INSTALL** directory

```

gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$ cp /home/gopalkrishna/Downloads/hadoop-2.6.0.tar.gz .
gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$ ll
total 8
drwxrwxr-x  2 gopalkrishna gopalkrishna 4096 Mar 10 01:41 ./
drwxr-xr-x 20 gopalkrishna gopalkrishna 4096 Mar 10 01:40 ../
-rw-r--r--  1 gopalkrishna gopalkrishna   0 Mar 10 01:41 hadoop-2.6.0.tar.gz
gopalkrishna@ubuntu:~/INSTALL$

```

STEP 8: Extract the copied tar ball [**hadoop-2.6.0.tar.gz**] using below command:

```

gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$ ll hadoop-2.6.0.tar.gz
-rw-r--r--  1 gopalkrishna gopalkrishna 0 Mar 10 01:41 hadoop-2.6.0.tar.gz
gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$ tar -xzf hadoop-2.6.0.tar.gz

```

After UNTARING: hadoop-2.6.0 directory is extracted

```
gopalkrishna@ubuntu:~/INSTALL
gopalkrishna@ubuntu:~/INSTALL$
gopalkrishna@ubuntu:~/INSTALL$ LL
LL: command not found
gopalkrishna@ubuntu:~/INSTALL$ ll
total 190696
drwxrwxr-x 3 gopalkrishna gopalkrishna 4096 Mar 10 01:45 ./
drwxr-xr-x 20 gopalkrishna gopalkrishna 4096 Mar 10 01:40 ../
drwxr-xr-x 9 gopalkrishna gopalkrishna 4096 Nov 13 2014 hadoop-2.6.0/
-rw-r--r-- 1 gopalkrishna gopalkrishna 195257604 Mar 10 01:44 hadoop-2.6.0.tar.gz
gopalkrishna@ubuntu:~/INSTALL$
```

STEP 9: Go to the below path to update configuration files:

[/home/gopalkrishna/INSTALL/hadoop-2.6.0/etc/hadoop](#)

Below are the total configuration files in 'hadoop' directory

```
gopalkrishna@ubuntu:~/INSTALL/hadoop-2.6.0$ cd etc/hadoop/
gopalkrishna@ubuntu:~/INSTALL/hadoop-2.6.0/etc/hadoop$
gopalkrishna@ubuntu:~/INSTALL/hadoop-2.6.0/etc/hadoop$ pwd
/home/gopalkrishna/INSTALL/hadoop-2.6.0/etc/hadoop
gopalkrishna@ubuntu:~/INSTALL/hadoop-2.6.0/etc/hadoop$ ls
capacity-scheduler.xml      httpfs-env.sh              mapred-env.sh
configuration.xml           httpfs-log4j.properties   mapred-queues.xml.template
container-executor.cfg     httpfs-signature.secret   mapred-site.xml
core-site.xml               httpfs-site.xml           mapred-site.xml.template
hadoop-env.cmd             kms-acls.xml              slaves
hadoop-env.sh              kms-env.sh                ssl-client.xml.example
hadoop-metrics2.properties kms-log4j.properties     ssl-server.xml.example
hadoop-metrics.properties kms-site.xml               yarn-env.cmd
hadoop-policy.xml          log4j.properties         yarn-env.sh
hdfs-site.xml              mapred-env.cmd            yarn-site.xml
gopalkrishna@ubuntu:~/INSTALL/hadoop-2.6.0/etc/hadoop$
```

STEP 10: We have to edit the below 8 configuration files as part of HADOOP Installation:

1. core-site.xml
2. mapred-site.xml
3. mapred-env.sh
4. yarn-site.xml
5. hdfs-site.xml
6. hadoop-env.sh
7. yarn-env.sh
8. slaves

STEP 11: Open **'core-site.xml'** file, add the below 2 properties in between **<configuration>** tags

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl" />
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://localhost:8020</value>
  </property>
  <property>
    <name>hadoop.tmp.dir</name>
    <value>file:/home/gopalkrishna/data/tmp</value>
  </property>
</configuration>
```

Name Node Configuration Property...
RPC port is: 8020

To Store the data in Local File System

```
<property>
  <name>fs.defaultFS</name>
  <value>hdfs://localhost:8020</value>
</property>
<property>
  <name>hadoop.tmp.dir</name>
  <value>file:/home/gopalkrishna/data/tmp</value>
</property>
```

STEP 12: Open **'hadoop-env.sh'** file and update **JAVA_HOME** path

```
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.

# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME. All others are
# optional. When running a distributed configuration it is best
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

# The java implementation to use.
export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64
export JAVA_HOME=${JAVA_HOME}

# The jsvc implementation to use. Jsvc is required to run secure datanodes
# that bind to privileged ports to provide authentication of data transfer
# protocol. Jsvc is not required if SASL is configured for authentication of
# data transfer protocol using non-privileged ports.
#export JSVC_HOME=${JSVC_HOME}
```

STEP 13: Open **mapred-env.sh** and update **JAVA_HOME** in that

```
# http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.

# export JAVA_HOME=/home/y/libexec/jdk1.6.0/
export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64

export HADOOP_JOB_HISTORYSERVER_HEAPSIZE=1000

export HADOOP_MAPRED_ROOT_LOGGER=INFO,RFA

#export HADOOP_JOB_HISTORYSERVER_OPTS=
#export HADOOP_MAPRED_LOG_DIR="" # Where log files are stored. $HADOOP_MAPRED_HOME/logs by default.
#export HADOOP_JHS_LOGGER=INFO,RFA # Hadoop JobSummary logger.
#export HADOOP_MAPRED_PID_DIR= # The pid files are stored. /tmp by default.
#export HADOOP_MAPRED_IDENT_STRING= #A string representing this instance of hadoop. $USER by default
#export HADOOP_MAPRED_NICENESS= #The scheduling priority for daemons. Defaults to 0.
```

Updating **JAVA_HOME** path in **mapred_env.sh** file

STEP 14: Open **hdfs-site.xml** file and add the below properties:

```
<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>

<property>
  <name>dfs.namenode.name.dir</name>
  <value>file:/home/gopalkrishna/data/dfs/name</value>
</property>

<property>
  <name>dfs.datanode.data.dir</name>
  <value>file:/home/gopalkrishna/data/dfs/data</value>
</property>
```

```
mapred-env.sh x *hdfs-site.xml x
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>file:/home/gopalkrishna/data/dfs/name</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>file:/home/gopalkrishna/data/dfs/data</value>
  </property>
</configuration>
```

First Property is to configuration "REPLICATION" factor...In Single Node Cluster it's always one...default is: 3

Second Property is to specify the path to store Metadata in Name Node

Third property is to store the specify the path to store the Actual Data on Data Node

STEP 15: Open **mapred-site.xml** and update the framework architecture details as 'yarn'

```
<property>
  <name>mapreduce.framework.name</name>
  <value>yarn</value>
</property>
```

```
mapred-env.sh x hdfs-site.xml x core-site.xml x *mapred-site.xml x
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

This Property is to Specify MR2 architecture is "yarn"

STEP 16: Open **yarn-env.sh** and update **JAVA_HOME** path in that

```
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.

# User for YARN daemons
export HADOOP_YARN_USER=${HADOOP_YARN_USER:-yarn}

# resolve links - $0 may be a softlink
export YARN_CONF_DIR="${YARN_CONF_DIR:-$HADOOP_YARN_HOME/conf}"

# some Java parameters
# export JAVA_HOME=/home/y/libexec/jdk1.6.0/
export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64

if [ "$JAVA_HOME" != "" ]; then
#echo "run java in $JAVA_HOME"
JAVA_HOME=$JAVA_HOME
fi

if [ "$JAVA_HOME" = "" ]; then
echo "Error: JAVA_HOME is not set."
exit 1
```

Updating JAVA_HOME in yarn-env.sh

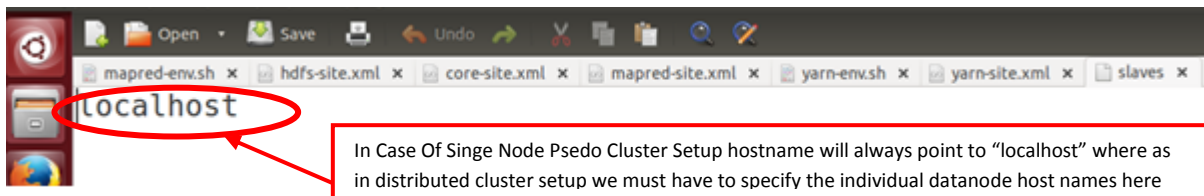
STEP 17: Open **yarn-site.xml** and add the below properties to configure **“Resource Manager”** & **“Node Manager”** details:

```
<property>
  <name>yarn.resourcemanager.hostname</name>
  <value>localhost</value>
</property>
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
<property>
  <name>yarn.nodemanager.aux-services.mapreduce_shuffle.class</name>
  <value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>
```

```
<?xml version="1.0"?>
<configuration>
<!-- Site specific YARN configuration properties -->
<property>
  <name>yarn.resourcemanager.hostname</name>
  <value>localhost</value>
</property>
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
<property>
  <name>yarn.nodemanager.aux-services.mapreduce_shuffle.class</name>
  <value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>
</configuration>
```

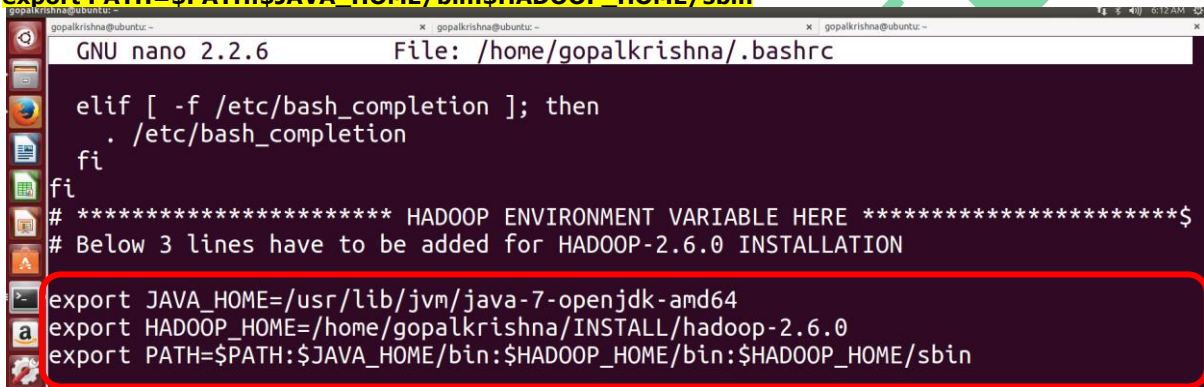
These 3 properties are to specify **“ResourceManager”** & **“NodeManager”** Configurations.

STEP 18: Open **slaves** file and check whether the hostname is **localhost** or not

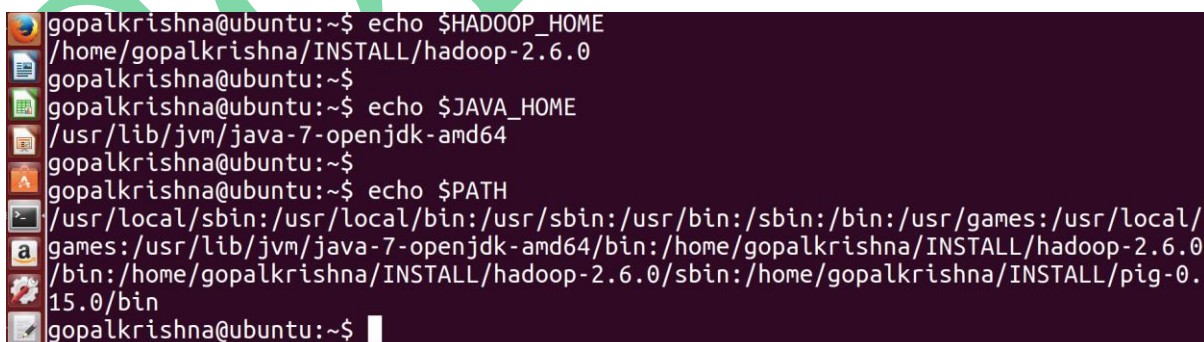


STEP 19: Open **bashrc** file (**~/.bashrc**) and update **JAVA_HOME, HADOOP_HOME & PATH** variables:

```
export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64
export HADOOP_HOME=/home/gopalkrishna/INSTALL/hadoop-2.6.0
export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/sbin
```



STEP 20: To check the bashrc changes, open a new terminal and type the below command: (**echo \$HADOOP_HOME**)



STEP 21: Before starting **Name Node**, we **must** have to **format** the name node using below command: (**hadoop namenode -format**)

```
gopalkrishna@ubuntu:~$  
gopalkrishna@ubuntu:~$ hadoop namenode -format  
DEPRECATED: Use of this script to execute hdfs command is deprecated.  
Instead use the hdfs command for it.  
  
17/03/10 02:13:25 INFO namenode.NameNode: STARTUP_MSG:  
/*****  
STARTUP_MSG: Starting NameNode  
STARTUP_MSG: host = ubuntu/127.0.1.1  
STARTUP_MSG: args = [-format]  
STARTUP_MSG: version = 2.6.0  
STARTUP_MSG: classpath = /home/gopalkrishna/INSTALL/hadoop-2.6.0/etc/hadoop:/home/  
gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/commons-collections-3.2.1.  
jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/httpcore-4.2.5.j  
ar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/commons-el-1.0.ja  
r:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/jasper-runtime-5.5  
.23.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/jersey-serve  
r-1.9.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/jackson-ja  
xrs-1.9.13.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/jerse  
y-json-1.9.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/activ  
ation-1.1.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/jettis  
on-1.1.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/slf4j-api  
-1.7.5.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/hadoop-au
```

Before starting name node, we must have to format the name node, otherwise daemons wont start even

```
a/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/protobuf-java-2.5.0.jar:/home/gopalkr  
ishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/xmlenc-0.52.jar:/home/gopalkrishn  
a/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/httpclient-4.2.5.jar:/home/gopalkrish  
na/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/commons-cli-1.2.jar:/home/gopalkrish  
na/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/apacheds-kerberos-codec-2.0.0-M15.ja  
r:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/commons-compress-1  
.4.1.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/apacheds-i1  
8n-2.0.0-M15.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/com  
mons-digester-1.8.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/li  
b/jetty-6.1.26.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/c  
ommons-io-2.4.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/ja  
xb-api-2.2.2.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/lib/cur  
ator-framework-2.6.0.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common  
lib/jsr305-1.3.9.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/common/li  
b/commons-httpclient-3.1.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/co  
mmon/lib/jetty-util-6.1.26.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/hadoop/  
common/lib/curator-recipes-2.6.0.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/share/h  
adoop/common/lib/hadoop-annotations-2.6.0.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.  
0/share/hadoop/common/lib/jets3t-0.9.0.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/s  
hare/hadoop/common/lib/jersey-core-1.9.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0/s  
hare/hadoop/common/lib/commons-codec-1.4.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0  
/share/hadoop/common/lib/commons-net-3.1.jar:/home/gopalkrishna/INSTALL/hadoop-2.6.0  
/share/hadoop/common/lib/jackson-xc-1.9.13.jar:/home/gopalkrishna/INSTALL/hadoop-2.6
```

```

STARTUP_MSG:   build = https://git-wip-us.apache.org/repos/asf/hadoop.git -r e349649
9ecb8d220fba99dc5ed4c99c8f9e33bb1; compiled by 'jenkins' on 2014-11-13T21:10Z
STARTUP_MSG:   java = 1.7.0_121
*****/
17/03/10 02:13:26 INFO namenode.NameNode: registered UNIX signal handlers for [TERM,
HUP, INT]
17/03/10 02:13:26 INFO namenode.NameNode: createNameNode [-format]
Formatting using clusterid: CID-ffadd70a-09c8-4309-afa2-e685db7daf89
17/03/10 02:13:32 INFO namenode.FSNamesystem: No KeyProvider found.
17/03/10 02:13:32 INFO namenode.FSNamesystem: fsLock is fair:true
17/03/10 02:13:32 INFO blockmanagement.DatanodeManager: dfs.block.invalidate.limit=1
000
17/03/10 02:13:32 INFO blockmanagement.DatanodeManager: dfs.namenode.datanode.regist
ration.ip-hostname-check=true
17/03/10 02:13:32 INFO blockmanagement.BlockManager: dfs.namenode.startup.delay.bloc
k.deletion.sec is set to 000:00:00:00.000
17/03/10 02:13:32 INFO blockmanagement.BlockManager: The block deletion will start a
round 2017 Mar 10 02:13:32
17/03/10 02:13:32 INFO util.GSet: Computing capacity for map BlocksMap
17/03/10 02:13:32 INFO util.GSet: VM type          = 64-bit
17/03/10 02:13:32 INFO util.GSet: 2.0% max memory 966.7 MB = 19.3 MB
17/03/10 02:13:32 INFO util.GSet: capacity        = 2^21 = 2097152 entries
0.9990000128746033
17/03/10 02:13:33 INFO namenode.FSNamesystem: dfs.namenode.safemode.min.datanodes =
0
17/03/10 02:13:33 INFO namenode.FSNamesystem: dfs.namenode.safemode.extension      =
30000
17/03/10 02:13:33 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
17/03/10 02:13:33 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total hea
p and retry cache entry expiry time is 600000 millis
17/03/10 02:13:33 INFO util.GSet: Computing capacity for map NameNodeRetryCache
17/03/10 02:13:33 INFO util.GSet: VM type          = 64-bit
17/03/10 02:13:33 INFO util.GSet: 0.029999999329447746% max memory 966.7 MB = 297.0
KB
17/03/10 02:13:33 INFO util.GSet: capacity        = 2^15 = 32768 entries
17/03/10 02:13:34 INFO namenode.NNConf: ACLs enabled? false
17/03/10 02:13:34 INFO namenode.NNConf: XAttrs enabled? true
17/03/10 02:13:34 INFO namenode.NNConf: Maximum size of an xattr: 16384
17/03/10 02:13:34 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1331610154-12
7.0.1.1-1489140814171
17/03/10 02:13:34 INFO common.Storage: Storage directory /home/gopalkrishna/data/dfs
/name has been successfully formatted.
17/03/10 02:13:35 INFO namenode.NNStorageRetentionManager: Going to retain 1 images
with txid >= 0
17/03/10 02:13:35 INFO util.ExitUtil: Exiting with status 0
17/03/10 02:13:35 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1
*****/
gopalkrishna@ubuntu:~$ █
17/02/16 04:41:06 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
17/02/16 04:41:06 INFO util.ExitUtil: Exiting with status 0
17/02/16 04:41:06 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at ubuntu/127.0.1.1
*****/
gopalkrishna@ubuntu:~$ █

```

STEP 22: To start all the daemons of hadoop in 2.X.X use “**start-all.sh**” command

```

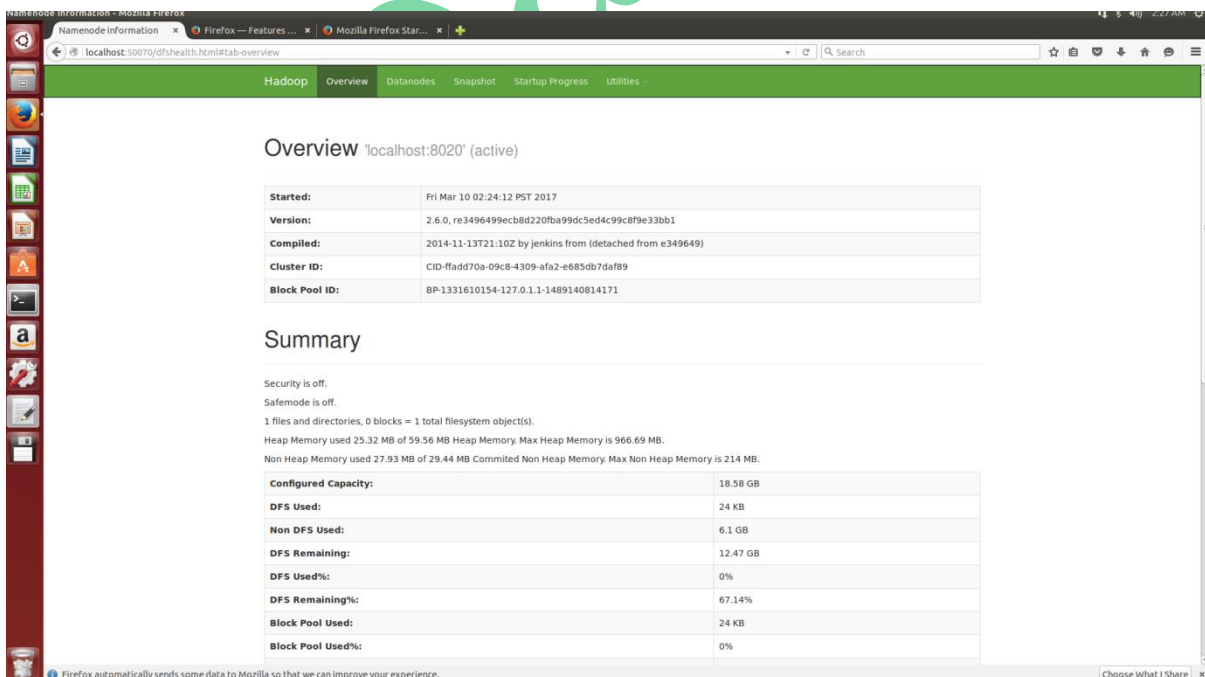
gopalkrishna@ubuntu:~$ jps
13104 Jps
gopalkrishna@ubuntu:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [localhost]
localhost: starting namenode, logging to /home/gopalkrishna/INSTALL/hadoop-2.6.0/lo
gs/hadoop-gopalkrishna-namenode-ubuntu.out
localhost: starting datanode, logging to /home/gopalkrishna/INSTALL/hadoop-2.6.0/lo
gs/hadoop-gopalkrishna-datanode-ubuntu.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /home/gopalkrishna/INSTALL/hadoop-2
.6.0/logs/hadoop-gopalkrishna-secondarynamenode-ubuntu.out
starting yarn daemons
starting resourcemanager, logging to /home/gopalkrishna/INSTALL/hadoop-2.6.0/logs/y
arn-gopalkrishna-resourcemanager-ubuntu.out
localhost: starting nodemanager, logging to /home/gopalkrishna/INSTALL/hadoop-2.6.0
/logs/yarn-gopalkrishna-nodemanager-ubuntu.out
gopalkrishna@ubuntu:~$
gopalkrishna@ubuntu:~$ jps
2929 NameNode
3302 SecondaryNameNode
3674 Jps
3499 ResourceManager
3076 DataNode
3632 NodeManager
gopalkrishna@ubuntu:~$

```

HDFS Storage Daemons: Name Node
: Data Node
: Secondary Name Node

Map Reduce Processing Daemons: Resource Manager
: Node Manager

STEP 23: To Access the Name Node information in GUI:
<http://localhost:50070>



DataNodes usages% (Min/Median/Max/stdDev):	0.03% / 0.03% / 0.03% / 0.00%
Live Nodes	1 (Decommissioned: 0)
Dead Nodes	0 (Decommissioned: 0)
Decommissioning Nodes	0
Number of Under-Replicated Blocks	0
Number of Blocks Pending Deletion	0
Block Deletion Start Time	3/10/2017, 6:16:55 AM

NameNode Journal Status

Current transaction ID: 545

Journal Manager	State
FileJournalManager(root=/home/gopalkrishna/data/dfs/name)	EditLogOutputStream(/home/gopalkrishna/data/dfs/name/current/edits_inprogress_00000000000000000545)

NameNode Storage

Storage Directory	Type	State
/home/gopalkrishna/data/dfs/name	IMAGE_AND_EDITS	Active

Hadoop, 2014.

To Access Legacy UI page, click here

Legacy UI

NameNode 'localhost:8020' (active)

Started:	Fri Mar 10 06:16:55 PST 2017
Version:	2.6.0, e3496499ecb8d220fba99dc5ed4c99c8f9e33bb1
Compiled:	2014-11-13T21:10Z by jenkins from (detached from e349649)
Cluster ID:	CID-ffadd70a-09c8-4309-afa2-e685db7daf89
Block Pool ID:	BP-1331610154-127.0.1.1-1489140814171

[Browse the filesystem](#)

[NameNode Logs](#)

Cluster Summary

Security is OFF
 49 files and directories, 20 blocks = 69 total.
 Heap Memory used 32.39 MB is 54% of Committed Heap Memory 59.56 MB. Max Heap Memory is 966.69 MB.
 Non Heap Memory used 31.41 MB is 95% of Committed Non Heap Memory 32.88 MB. Max Non Heap Memory is 214 MB.

Configured Capacity	:	18.58 GB
DFS Used	:	6.29 MB
Non DFS Used	:	6.57 GB
DFS Remaining	:	12.00 GB
DFS Used%	:	0.03%
DFS Remaining%	:	64.57%
Block Pool Used	:	6.29 MB
Block Pool Used%	:	0.03%
DataNodes usages	:	Min % Median % Max % stdev %
	:	0.03% 0.03% 0.03% 0.00%
Live Nodes	:	1 (Decommissioned: 0)
Dead Nodes	:	0 (Decommissioned: 0)

STEP 24: To Start Job History Server in 2.x.x.

mr-jobhistory-daemon.sh start historyserver

```

gopal.krishna@ubuntu:~$ mr-jobhistory-daemon.sh start historyserver
starting historyserver, logging to /home/gopal.krishna/INSTALL/hadoop-2.6.0/logs/mapred-gopal.krishna-hi
storyserver-ubuntu.out
gopal.krishna@ubuntu:~$ jps
6145 JobHistoryServer
5297 NodeManager
5177 ResourceManager
4828 DataNode
4983 SecondaryNameNode
6197 Jps
4712 NameNode
gopal.krishna@ubuntu:~$ █

```

STEP 25: To Access Resource Manager:

<http://localhost:8088>

The screenshot shows the Hadoop All Applications web interface. The main content area displays 'Cluster Metrics' and 'Scheduler Metrics'. Below these, there is a table for applications with columns for ID, User, Name, Application Type, Queue, Start Time, Finish Time, State, Final Status, Progress, Tracking UI, and Blacklisted Nodes. The table is currently empty, showing 'Showing 0 to 0 of 0 entries'.

STEP 26: To Access Job History Server:

<http://localhost:19888>

The screenshot shows the Hadoop JobHistory web interface. The main content area displays 'Retired Jobs' with a table containing columns for Submit Time, Start Time, Finish Time, Job ID, Name, User, Queue, State, Maps Total, Maps Completed, Reduces Total, and Reduces Completed. The table is currently empty, showing 'Showing 0 to 0 of 0 entries'.

STEP 27: To stop all the daemons of hadoop in 2.X.X use "stop-all.sh" command

```
gopalkrishna@ubuntu:~$ stop-all.sh
This script is Deprecated. Instead use stop-dfs.sh and stop-yarn.sh
Stopping namenodes on [localhost]
localhost: stopping namenode
localhost: stopping datanode
Stopping secondary namenodes [0.0.0.0]
0.0.0.0: stopping secondarynamenode
stopping yarn daemons
stopping resourcemanager
localhost: stopping nodemanager
localhost: nodemanager did not stop gracefully after 5 seconds: killing with kill -9
no proxyserver to stop
gopalkrishna@ubuntu:~$
```

STEP 28: To Stop Job History Server in 2.x.x.

```
gopalkrishna@ubuntu:~$ mr-jobhistory-daemon.sh stop historyserver
stopping historyserver
gopalkrishna@ubuntu:~$ jps
13073 Jps
gopalkrishna@ubuntu:~$
```