

HADOOP INSTALL

COMMANDS

HADOOP VERSION – 2.9.0

JAVA – 8 (1.8.0)

WORKS ON ANY UBUNTU VERSION NO HDUSER REQUIRED

Installing and moving (make sure the java version is same as given below)

```
sudo apt install openjdk-8-jdk -y
```

```
java -version
```

```
sudo apt install ssh -y
```

```
ssh-keygen -t rsa -P ""
```

```
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
```

```
chmod 600 ~/.ssh/authorized_keys
```

```
ssh localhost
```

```
tar -xzf hadoop-2.9.0.tar.gz
```

```
sudo mv hadoop-2.9.0 /usr/local/hadoop
```

Set Environment Variables

```
nano ~/.bashrc ### go to end and paste it
```

```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
```

```
export HADOOP_HOME=/usr/local/hadoop
```

```
export PATH=$PATH:$HADOOP_HOME/bin
```

```
export PATH=$PATH:$HADOOP_HOME/sbin
```

```
export HADOOP_MAPRED_HOME=$HADOOP_HOME
```

```
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib"
export HADOOP_CONF_DIR=/usr/local/hadoop/etc/hadoop

source ~/.bashrc
```

Step 3: Configure Hadoop

```
cd $HADOOP_HOME/etc/hadoop
```

```
nano core-site.xml
```

```
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://localhost:9000</value>
  </property>
</configuration>
```

```
nano hdfs-site.xml
```

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

```
<name>dfs.namenode.name.dir</name>

<value>file:///usr/local/hadoop/hdfs/namenode</value>

</property>

<property>

<name>dfs.datanode.data.dir</name>

<value>file:///usr/local/hadoop/hdfs/datanode</value>

</property>

<property>

<name>dfs.namenode.rpc-address</name>

<value>localhost:9000</value>

</property>

<property>

<name>dfs.namenode.http-address</name>

<value>localhost:9870</value>

</property>

</configuration>
```

cp mapred-site.xml.template mapred-site.xml

nano mapred-site.xml

```
<configuration>

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

</configuration>
```

nano yarn-site.xml

```
<configuration>
  <property>
    <name>yarn.resourcemanager.address</name>
    <value>localhost:8032</value>
  </property>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
</configuration>
```

nano hadoop-env.sh

```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
```

hdfs namenode -format

start-dfs.sh

start-yarn.sh

go to <http://localhost:9870> for Hadoop dashboard and <http://localhost:8088> for manager