|  |
| --- |
| **Open Terminal**  **# Update the source list** devesh@laptop:~$ **sudo apt-get update** |
| # The OpenJDK project is the default version of Java  **Install JDK**  devesh@laptop:~$ **sudo apt-get install default-jdk`**  devesh@laptop:~$ **java -version** java version "1.7.0\_95" |
| **Adding a dedicated Hadoop user**  devesh@laptop:~$ **sudo addgroup hadoop** Adding group `hadoop' (GID 1002) ... Done.  devesh@laptop:~$ **sudo adduser --ingroup hadoop hduser (Press Enter)**  Is the information correct? [Y/n] Y  **Enter new password for new hduser**  devesh@laptop:~$ **sudo adduser hduser sudo** Adding user `hduser' to group `sudo' ... Adding user hduser to group sudo Done.  **Logout from current user and log in to hduser** |
| **Installing SSH**  devesh@laptop:~$ **su hduser**  hduser@laptop:~$ **sudo apt-get install ssh** |
| **Create and Setup SSH Certificates**  hduser@laptop:~$ **cd /home/hduser/**  hduser@laptop:~$ **ssh-keygen -t rsa -P ""** Generating public/private rsa key pair.  Enter file in which to save the key (/home/hduser/.ssh/id\_rsa):  **(\*\*press enter)** Created directory '/home/hduser/.ssh'. Your identification has been saved in /home/hduser/.ssh/id\_rsa. Your public key has been saved in /home/hduser/.ssh/id\_rsa.pub. The key fingerprint is: 50:6b:f3:fc:0f:32:bf:30:79:c2:41:71:26:cc:7d:e3 hduser@laptop The key's randomart image is: +--[ RSA 2048]----+ | .oo.o | | . .o=. o | | . + . o . | | o = E | | S + | | . + | | O + | | O o | | o.. | +-----------------+  hduser@laptop:$ **cat /home/hduser/.ssh/id\_rsa.pub >> /home/hduser/.ssh/authorized\_keys** |
| hduser@laptop:$ **ssh localhost**  hduser@laptop:$ **exit**  hduser@laptop:$ **cd /home/hduser** |
| **Download Hadoop**  hduser@laptop:~$  **wget http://mirrors.sonic.net/apache/hadoop/common/hadoop-2.7.2/hadoop-2.7.2.tar.gz**  Or from <http://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-2.7.2/hadoop-2.7.2.tar.gz>  hduser@laptop:~$ **tar xvzf hadoop-2.8.3.tar.gz**  hduser@laptop:~$ **cd hadoop-2.8.3**  hduser@laptop:~$ **sudo mkdir /usr/local/hadoop**  hduser@laptop:~/hadoop-2.7.2$ **sudo mv \* /usr/local/hadoop** hduser@laptop:~/hadoop-2.7.2$ **sudo chown -R hduser:hadoop /usr/local/hadoop**  hduser@laptop:~/hadoop-2.7.2$ **sudo apt-get install vim**  **Check Java** hduser@laptop:~$ **readlink -f /usr/bin/javac** /usr/lib/jvm/java-8-openjdk-amd64/bin/javac \*\*\* for 64 bit  /usr/lib/jvm/java-8-openjdk-i386/bin/javac \*\*\*for 32 bit |
| **Set-up the Configuration Files**  hduser@laptop:~$ **vim ~/.bashrc** **#HADOOP VARIABLES START \*\*\*\*Append this at end of file\*\*\*\*\*\*\*\*\* export JAVA\_HOME=/usr/lib/jvm/java-8-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**  Save the file and Exit ( **esc:wq!** to save the file in vim) hduser@laptop:~$ **source ~/.bashrc** |
| hduser@laptop:~$ **vim /usr/local/hadoop/etc/hadoop/hadoop-env.sh**  (Change existing JAVA\_HOME to **/usr/lib/jvm/java-7-openjdk-amd64 if machine is 64 Bit** or  **If machine is 32 bit change it to /usr/lib/jvm/java-7-openjdk-i386**)  **export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64 Save and Exit** |
| hduser@laptop:~$ **vim /usr/local/hadoop/etc/hadoop/yarn-site.xml**  **Open the file and enter the following in between <configuration></configuration> tag and save :**  **<property>**  **<name>yarn.nodemanager.aux-services</name>**  **<value>mapreduce\_shuffle</value>**  **</property>** |
| Create Temporary Directory to store App-data  hduser@laptop:~$ **sudo mkdir -p /app/hadoop/tmp**  hduser@laptop:~$ **sudo chown hduser:hadoop /app/hadoop/tmp**  hduser@laptop:~$ **vim /usr/local/hadoop/etc/hadoop/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></property>** |
| [hduser@laptop](mailto:hduser@laptop):~$ **cp /usr/local/hadoop/etc/hadoop/mapred-site.xml.template /usr/local/hadoop/etc/hadoop/mapred-site.xml**  hduser@laptop:~$ **vim /usr/local/hadoop/etc/hadoop/mapred-site.xml**  **Add these properties in <configuration></configuration> tag:**  **<property>  <name>mapred.job.tracker</name>  <value>localhost:54311</value>  </property>**  **<property>**  **<name>mapreduce.framework.name</name>**  **<value>yarn</value>**  **</property>** Save and Exit |
| hduser@laptop:~$ **sudo mkdir -p /usr/local/hadoop\_store/hdfs/namenode**  hduser@laptop:~$ **sudo mkdir -p /usr/local/hadoop\_store/hdfs/datanode** hduser@laptop:~$ **sudo chown -R hduser:hadoop /usr/local/hadoop\_store**  hduser@laptop:~$ **vim /usr/local/hadoop/etc/hadoop/hdfs-site.xm**l **Open the file and add the following properties between the <configuration></configuration> tag:**  **<property>  <name>dfs.replication</name> <value>1</value>  </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>** |
| hduser@laptop:~$ **hadoop namenode -format**  **Starting Hadoop** hduser@laptop$: **/usr/local/hadoop/sbin/start-all.sh**  **We can check if it's really up and running:**  hduser@laptop$: **jps** 9026 NodeManager 7348 NameNode 9766 SecondaryNameNode 8887 ResourceManager 7507 DataNode  **Hadoop UI visible in browser at localhost:50070**  **MR Job progress and history UI visible at** **localhost:8088**  **To Stop Hadoop**  hduser@laptop$: **/usr/local/hadoop/sbin/stop-all.sh**  **\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Done\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*** |