|  |
| --- |
|  |
| #Install Java 8 |
|  | sudo add-apt-repository ppa:webupd8team/java |
|  | sudo apt-get update |
|  | sudo apt-get install oracle-java8-installer -y |
|  |  |
|  | #Create a Hadoop user for accessing HDFS and MapReduce |
|  | sudo addgroup hadoop |
|  | sudo adduser --ingroup hadoop hduser |
|  | sudo adduser hduser sudo |
|  | sudo su hduser |
|  |  |
|  | #Install SSH |
|  | sudo apt-get install openssh-server -y |
|  |  |
|  | #Configure SSH |
|  | echo -e 'y\n'| ssh-keygen -t rsa -P "" -f $HOME/.ssh/id\_rsa |
|  | cat $HOME/.ssh/id\_rsa.pub >> $HOME/.ssh/authorized\_keys |
|  |  |
|  | #Disable IPV6 |
|  | sudo su -c 'cat >>/etc/sysctl.conf <<EOL |
|  | net.ipv6.conf.all.disable\_ipv6 = 1 |
|  | net.ipv6.conf.default.disable\_ipv6 = 1 |
|  | net.ipv6.conf.lo.disable\_ipv6 = 1 |
|  | EOL' |
|  |  |
|  |  |
|  | #Download Hadoop |
|  | wget http://apache.mirrors.tds.net/hadoop/common/stable/hadoop-2.7.2.tar.gz |
|  |  |
|  | #Extract and Install Hadoop tar ball |
|  | tar -xzvf hadoop-2.7.2.tar.gz |
|  | sudo mv hadoop-2.7.2 /usr/local/hadoop |
|  | sudo chown hduser:hadoop -R /usr/local/hadoop |
|  |  |
|  | sudo mkdir -p /usr/local/hadoop\_tmp/hdfs/namenode |
|  | sudo mkdir -p /usr/local/hadoop\_tmp/hdfs/datanode |
|  | sudo chown hduser:hadoop -R /usr/local/hadoop\_tmp |
|  |  |
|  | #Update Linux configuration files |
|  | cat >>$HOME/.bashrc <<EOL |
|  | # -- HADOOP ENVIRONMENT VARIABLES START -- # |
|  | export JAVA\_HOME=/usr/lib/jvm/java-8-oracle |
|  | export HADOOP\_HOME=/usr/local/hadoop |
|  | export PATH=\$PATH:\$HADOOP\_HOME/bin |
|  | export PATH=\$PATH:\$HADOOP\_HOME/sbin |
|  | export PATH=\$PATH:/usr/local/hadoop/bin/ |
|  | 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" |
|  | # -- HADOOP ENVIRONMENT VARIABLES END -- # |
|  | EOL |
|  |  |
|  | exec bash |
|  |  |
|  | #Update hadoop-env.sh |
|  | sudo su -c 'echo export JAVA\_HOME=/usr/lib/jvm/java-8-oracle >> /usr/local/hadoop/etc/hadoop/hadoop-env.sh' |
|  | sudo su -c 'echo export HADOOP\_LOG\_DIR=/var/log/hadoop/ >> /usr/local/hadoop/etc/hadoop/hadoop-env.sh' |
|  | sudo mkdir /var/log/hadoop/ |
|  | sudo chown hduser:hadoop -R /var/log/hadoop |
|  |  |
|  | #Update core-site.xml |
|  | sudo sed -i '/<configuration>/,/<\/configuration>/d' /usr/local/hadoop/etc/hadoop/core-site.xml |
|  | sudo su -c 'cat >> /usr/local/hadoop/etc/hadoop/core-site.xml <<EOL |
|  | <configuration> |
|  | <property> |
|  | <name>fs.default.name</name> |
|  | <value>hdfs://localhost:8000</value> |
|  | </property> |
|  | </configuration> |
|  | EOL' |
|  |  |
|  | #Update hdfs-site.xml |
|  | sudo sed -i '/<configuration>/,/<\/configuration>/d' /usr/local/hadoop/etc/hadoop/hdfs-site.xml |
|  | sudo su -c 'cat >>/usr/local/hadoop/etc/hadoop/hdfs-site.xml <<EOL |
|  | <configuration> |
|  | <property> |
|  | <name>dfs.replication</name> |
|  | <value>1</value> |
|  | </property> |
|  | <property> |
|  | <name>dfs.namenode.name.dir</name> |
|  | <value>file:/usr/local/hadoop\_tmp/hdfs/namenode</value> |
|  | </property> |
|  | <property> |
|  | <name>dfs.datanode.data.dir</name> |
|  | <value>file:/usr/local/hadoop\_tmp/hdfs/datanode</value> |
|  | </property> |
|  | </configuration> |
|  | EOL' |
|  |  |
|  | #Update yarn-site.xml |
|  | sudo sed -i '/<configuration>/,/<\/configuration>/d' /usr/local/hadoop/etc/hadoop/yarn-site.xml |
|  | sudo su -c 'cat >>/usr/local/hadoop/etc/hadoop/yarn-site.xml <<EOL |
|  | echo "write to yarn-site.xml" |
|  | <configuration> |
|  | <property> |
|  | <name>yarn.nodemanager.aux-services</name> |
|  | <value>mapreduce\_shuffle</value> |
|  | </property> |
|  | </configuration> |
|  | EOL' |
|  |  |
|  | #Update mapred-site.xml |
|  | sudo sed -i '/<configuration>/,/<\/configuration>/d' /usr/local/hadoop/etc/hadoop/mapred-site.xml |
|  | sudo su -c 'cat >>/usr/local/hadoop/etc/hadoop/mapred-site.xml <<EOL |
|  | <configuration> |
|  | <property> |
|  | <name>mapreduce.framework.name</name> |
|  | <value>yarn</value> |
|  | </property> |
|  | </configuration> |
|  | EOL' |
|  |  |
|  | #Format Namenode |
|  | hdfs namenode -format |
|  | start-all.sh |
|  | or start-dfs.sh |
|  | start-yarn.sh |
|  |  |
|  | jps |
|  |  |
|  | ps -aux | grep jvm |
|  |  |
|  | hadoop jar /usr/local/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar pi 5 10 |
|  | nano inputdata # write anything in it for processing. |
|  | hadoop fs -put inputdata . |
|  | hadoop fs -ls -R |
|  | df -h inputdata |
|  | hadoop fs -df -h inputdata |
|  | hadoop jar /usr/local/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar wordcount inputdata outputresults |
|  | hadoop fs -ls |
|  | hadoop fs -mkdir yourname |
|  | hadoop fs -ls |
|  |  |