Lab 3. Installing, configuring and using YARN.

Exercise 1. Install YARN.

Step 1. Install the packages.

Bring back node4 (The one that we disabled the I/O device on in the last lab.

```
ssh node4 'hostname –i >> /etc/hadoop/conf/dfs.hosts
su –c 'hdfs dfsadmin –refreshNodes' hdfs
ssh node4 'service hadoop-hdfs-datanode restart'
hdfs dfs –df –h /
```

Now let's install the node manager on all four nodes.

```
allnodes 'yum install –y hadoop-yarn-nodemanager'
```

We'll install the resource manager on node 3.

Connect to node3 and then run the following:

```
yum install –y hadoop-yarn-resourcemanager
rpm –ql hadoop-yarn-{node,resource}manager
```

Look at the changes made to the system account database.

```
grep yarn /etc/{passwd,group}
```

We should see both a user and a group added for the YARN daemons.

Additionally, the yarn user should now be part of the hadoop group.

Now make sure that the new services are configured to start automatically on node 3.

```
chkconfig –list | grep yarn
```

Step 2. Configure YARN.

We start by removing all of the properties in the /etc/hadoop/conf/yarn-site.xml file except for the yarn.application.classpath property.

Now add the following properties to the file.

```
<configuration>
  <property>

  <name>yarn.resourcemanager.hostname</name>

  <value>node3</value>
 </property>
 <property>

<name>yarn.nodemanager.local-dirs</name>

<value>file:///yarn/local</value>
 </property>
 <property>

<name>yarn.nodemanager.log-dirs</name>

<value>file:///yarn/logs</value>
 </property>
 </configuration>
```

Next we create the directories where YARN will store local data.

Use either allnodes or pssh to run the following commands:

```
mkdir –pm 755 /yarn/{local,logs}
chown –R yarn: /yarn
```

Now start the resource manager and verify that it is running.

```
service hadoop-yarn-resourcemanager start
ps –u yarn eww
lsof –Pi –au yarn
```

Start the nodemanager and verify that both daemons are running

```
service hadoop-yarn-nodemanager start
ps –u yarn w
```

Verify that the resource manager shows the single node manager joined and ready to go,.

```
yarn node -list
```