

Prácticas BigData

1. YARN EN ENTORNOS PSEUDO-DISTRIBUIDO

- Copiamos el fichero "mapred-site.xml.template" a "mapred-site.xml"
- Ponemos la siguiente propiedad.

```
<name>mapreduce.framework.name
```

• Y en el fichero yarn-site.xml ponemos las siguientes:

- Arrancamos HDFS si no lo está
- Arrancamos los servicios YARN

start-yarn.sh

- Comprobamos con jps que tenemos los dos procesos de YARN
- Arrancamos HDFS si no lo está
- Arrancamos los servicios YARN
- Start-yarn.sh
- Arrancamos el servicio que permite guardar el histórico de los Jobs lanzados
- En el caso de tener hadoop versión 2 ejecutamos el comando

/mr-jobhistory-daemon.sh starthistoryserver

• En el caso de tener haddop 3 ejecutamos el comando



mapred --daemon start historyserver

- Comprobamos con "jsp" que tenemos los procesos lanzados
- Arrancamos HDFS si no lo está
- Arrancamos los servicios YARN

Start-yarn.sh

• Comprobamos con jps que está funcionando

ips

1137 Jps

22962 DataNode

23187 SecondaryNameNode

419 ResourceManager

22837 NameNode

541 NodeManager

- Arrancamos el servicio que permite guardar el histórico de los Jobs lanzados
 - o En versión 2 de Hadoop

mr-jobhistory-daemon.sh start historyserver

starting historyserver, logging to /opt/hadoop/logs/mapred-hadoop-historyserver-nodo1.out

- o En Versión 3 de Hadoop
- Comprobamos con jps que está funcionando

jps

22962 DataNode

23187 SecondaryNameNode

419 ResourceManager

1475 JobHistoryServer

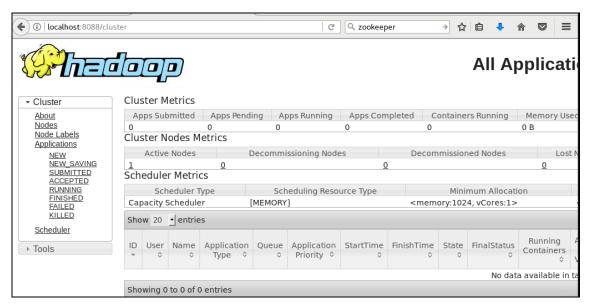
22837 NameNode

1515 Jps

541 NodeManager

 Accedemos a la página WEB de Administración para comprobar que está funcionando.





• Comprobamos en Nodes que tenemos el nodo activo

