Prácticas BigData

1. Comando YARN

- Vamos a ver como trabajar con el comando YARN

- Mantenemos dos terminales abiertos

- En el primero lanzamos un proyecto mapreduce.

  - Si es en versiones anteriores a la 2.9 tenemos que lanzarlo de esta manera. Es debido a que en la 2.9 se ha modificado la propiedad queue.name por queuename

```
hadoop jar /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-
examples-2.9.0.jar wordcount -Dmapred.job.queuename=warehouse
/practicas/cite75_99.txt /salida20
```

  - En la 2.9 lo hacemos así

```
hadoop jar /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-
examples-2.9.0.jar wordcount -Dmapred.job.queue.name=warehouse
/practicas/cite75_99.txt /salida20En el segundo vemos la aplicación con el
comando yarn
```

- Vamos a comprobar que se ha lanzado

```
yarn application -list
```

18/01/07 20:57:28 INFO client.RMProxy: Connecting to ResourceManager at nodo1/192.168.56.101:8032

Total number of applications (application-types: [], states: [SUBMITTED, ACCEPTED, RUNNING] and tags: []):1

Application-Id	Application-Name	Application-Type	User	Queue	State	Final-State	Progress	Tracking-URL
application_1515354119376_0002	word count	MAPREDUCE	hadoop	default	ACCEPTED	UNDEFINED	0%	N/A

- Comprobamos en que nodos se está ejecutando

```
yarn node -list
```

18/01/07 20:57:38 INFO client.RMProxy: Connecting to ResourceManager at nodo1/192.168.56.101:8032

Total Nodes:6

Node-Id	Node-State	Node-Http-Address	Number-of-Running-Containers
nodo7:34788	RUNNING	nodo7:8042	0
nodo3:40979	RUNNING	nodo3:8042	0
nodo4:36792	RUNNING	nodo4:8042	0
nodo6:33922	RUNNING	nodo6:8042	0
nodo2:36376	RUNNING	nodo2:8042	3
nodo5:37517	RUNNING	nodo5:8042	0

- Comprobamos el estado completo de la aplicación

---

yarn application -status application_1515354119376_0002

18/01/07 20:57:50 INFO client.RMProxy: Connecting to ResourceManager at nodo1/192.168.56.101:8032

Application Report :

       Application-Id : application_1515354119376_0002

       Application-Name : word count

       Application-Type : MAPREDUCE

       User : hadoop

       Queue : default

       Application Priority : 0

       Start-Time : 1515355041442

       Finish-Time : 0

       Progress : 5%

       State : RUNNING

       Final-State : UNDEFINED

       Tracking-URL : http://nodo2:34944

       RPC Port : 41274

       AM Host : nodo2

       Aggregate Resource Allocation : 91429 MB-seconds, 59 vcore-seconds

       Aggregate Resource Preempted : 0 MB-seconds, 0 vcore-seconds

       Log Aggregation Status : DISABLED

       Diagnostics :

       Unmanaged Application : false

       Application Node Label Expression : <Not set>

       AM container Node Label Expression : <DEFAULT_PARTITION>

       TimeoutType : LIFETIME   ExpiryTime : UNLIMITED   RemainingTime : -1seconds

---

- Comprobamos los Application Attempts:

---

yarn applicationattempt -list application_1515354119376_0002

18/01/07 20:58:50 INFO client.RMProxy: Connecting to ResourceManager at nodo1/192.168.56.101:8032

Total number of application attempts :1

    ApplicationAttempt-Id   State   AM-Container-Id   Tracking-URL

appattempt_1515354119376_0002_000001   RUNNING   container_1515354119376_0002_01_000001   http://nodo1:8088/proxy/application_1515354119376_0002/

- Vemos los contenedores asociados

---

yarn container -list appattempt_1515354119376_0002_000001

```
18/01/07  20:59:13  INFO  client.RMProxy:  Connecting  to  ResourceManager  at
nodo1/192.168.56.101:8032

Total number of containers :2

            Container-Id      Start Time              Finish Time              State
Host      Node Http Address                LOG-URL

container_1515354119376_0002_01_000004   dom ene 07 20:58:51 +0100 2018
N/A             RUNNING          nodo2:36376      http://nodo2:8042
        http://nodo2:8042/node/containerlogs/container_1515354119376_0002_01_000
004/hadoop

container_1515354119376_0002_01_000001   dom ene 07 20:57:21 +0100 2018
N/A             RUNNING          nodo2:36376      http://nodo2:8042
        http://nodo2:8042/node/containerlogs/container_1515354119376_0002_01_000
001/hadoop
```

- Eliminamos la aplicación

```
yarn application -kill application_1515354119376_0002
```

- Comprobamos que se ha eliminado