TCDM Práctica 1

José Antonio

- Tarea 1: Añadir al cluster un servidor de Backup y un TimeLineServer

1º) Servidor de backup

Ejecución del backupnode:

```
[jue oct 16 19:41:12 (0)] hdadmin@backupnode:~$ hdfs namenode -backup
WARNING: /opt/bd/hadoop/logs does not exist. Creating.
NOTE: Picked up JDK_JAVA_OPTIONS: --add-opens=java.base/java.lang=ALL-UNNAMED --add-opens=java.base/java.lang.invoke=ALL-UNNAMED --add-opens=java.base/java.
lang.reflect=ALL-UNNAMED --add-opens=java.base/java.net=ALL-UNNAMED --add-opens=java.base/sun.net.www.protocol.http=ALL-UNNAMED --add-opens=java.base/sun.ne
t.www.protocol.https=ALL-UNNAMED --add-opens=java.base/sun.util.calendar=ALL-UNNAMED --add-opens=jdk.zipfs/jdk.nio.zipfs=ALL-UNNAMED
2025-10-16 19:41:54,201 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = backupnode/172.19.0.7
STARTUP_MSG:   args = [-backup]
STARTUP_MSG:   version = 3.4.1
```

Mensajes que genera dicho servicio, destacando la creación del checkpoint:

```
2025-10-22 12:27:04,841 INFO namenode.FSImageTransactionalStorageInspector: No version file in /var/data/hdfs/backup/
dfs/name
2025-10-22 12:27:04,871 INFO namenode.TransferFsImage: Image Transfer timeout configured to 60000 milliseconds
2025-10-22 12:27:04,872 INFO namenode.TransferFsImage: Sending fileName: /var/data/hdfs/backup/dfs/name/current/fsima
ge_0000000000000000295, fileSize: 2939. Sent total: 2939 bytes. Size of last segment intended to send: -1 bytes.
2025-10-22 12:27:04,889 INFO namenode.TransferFsImage: Uploaded image with txid 295 to namenode at http://namenode:98
70 in 0.032 seconds
2025-10-22 12:27:04,895 INFO namenode.FSImage: Going to finish converging with remaining 1 txns from in-progress stre
am /var/data/hdfs/backup/dfs/name/current/edits_inprogress_0000000000000000296
2025-10-22 12:27:04,896 INFO namenode.RedundantEditLogInputStream: Fast-forwarding stream '/var/data/hdfs/backup/dfs/
name/current/edits_inprogress_0000000000000000296' to transaction ID 296
2025-10-22 12:27:04,899 INFO namenode.FSImage: Successfully synced BackupNode with NameNode at txnid 296
2025-10-22 12:27:04,900 INFO namenode.Checkpointer: Checkpoint completed in 0 seconds. New Image Size: 2939
```

Ficheros después del backup (directorio de backup y namenode):

```
jose@DESKTOP-VUD8TI9:~/master/TCDM$ docker exec -ti backupnode ls -l /var/data/hdfs/backup/dfs/name/current
total 4116
-rw-r--r-- 1 hdadmin hadoop     216 Oct 16 19:42 VERSION
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct 16 19:42 edits_0000000000000000001-0000000000000000001
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct 16 19:42 edits_0000000000000000002-0000000000000000470
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct 16 19:42 edits_0000000000000000471-0000000000000000471
-rw-r--r-- 1 hdadmin hadoop      42 Oct 16 19:42 edits_0000000000000000472-0000000000000000473
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct 16 19:42 edits_inprogress_0000000000000000474
-rw-r--r-- 1 hdadmin hadoop    5118 Oct 16 19:42 fsimage_0000000000000000473
-rw-r--r-- 1 hdadmin hadoop      62 Oct 16 19:42 fsimage_0000000000000000473.md5
jose@DESKTOP-VUD8TI9:~/master/TCDM$ docker exec -ti namenode ls -l /var/data/hdfs/namenode/current
total 4132
-rw-r--r-- 1 hdadmin hadoop     216 Oct 16 16:04 VERSION
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct  2 18:49 edits_0000000000000000001-0000000000000000001
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct  9 19:50 edits_0000000000000000002-0000000000000000470
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct 15 15:19 edits_0000000000000000471-0000000000000000471
-rw-r--r-- 1 hdadmin hadoop      42 Oct 16 19:42 edits_0000000000000000472-0000000000000000473
-rw-r--r-- 1 hdadmin hadoop 1048576 Oct 16 19:42 edits_inprogress_0000000000000000474
-rw-r--r-- 1 hdadmin hadoop    5084 Oct 16 16:04 fsimage_0000000000000000471
-rw-r--r-- 1 hdadmin hadoop      62 Oct 16 16:04 fsimage_0000000000000000471.md5
-rw-r--r-- 1 hdadmin hadoop    5118 Oct 16 19:42 fsimage_0000000000000000473
-rw-r--r-- 1 hdadmin hadoop      62 Oct 16 19:42 fsimage_0000000000000000473.md5
-rw-r--r-- 1 hdadmin hadoop       4 Oct 16 19:42 seen_txid
```

Ficheros del directorio de metadatos del NameNode dentro de current antes de backup:

```
[Thu Oct 09 19:57:16 (0)] root@namenode:/# ls /var/data/hdfs/namenode/curren
t
VERSION
edits_0000000000000000001-0000000000000000001
edits_inprogress_0000000000000000002
fsimage_0000000000000000000
fsimage_0000000000000000000.md5
fsimage_0000000000000000001
fsimage_0000000000000000001.md5
seen_txid
[Thu Oct 09 20:10:36 (0)] root@namenode:/# 
```

Interfaz web del nodo de backup, nos aseguramos que está corriendo:

Hadoop	Overview	Datanodes	Datanode Volume Failures	Snapshot	Startup Progress	Utilities ▾

Upgrade in progress. Not yet finalized.   ✕

Overview 'backupnode:50100' (⏻standby)

Started:	Thu Oct 16 20:18:35 +0200 2025
Version:	3.4.1, r4d7825309348956336b8f06a08322b78422849b1
Compiled:	Wed Oct 09 16:57:00 +0200 2024 by mthakur from branch-3.4.1
Cluster ID:	CID-1e7e20aa-d3fd-4875-bacd-e16a34f89cd0
Block Pool ID:	BP-1824411926-172.19.0.2-1759423600963

2º) Timeline Server

Mostrando la información de la tarea ejecutada QuasiMonteCarlo:

hadoop — All Applications

Show 20 v entries   Search:

ID	User	Name	Application Type	Application Tags	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Progress	Tracking UI
application_1761134638821_0001	hdadmin	QuasiMonteCarlo	MAPREDUCE		root.default	0	Wed Oct 22 14:12:09 +0200 2025	N/A	N/A	ACCEPTED	UNDEFINED		Unassigned

Application History
- About
- Applications
  - FINISHED
  - FAILED
  - KILLED
- Tools

Showing 1 to 1 of 1 entries   First Previous 1 Next Last

---

hadoop — Application application_1761134638821_0001

Application History
- About
- Applications
  - FINISHED
  - FAILED
  - KILLED
- Tools

Application Overview

User:	hdadmin
Name:	QuasiMonteCarlo
Application Type:	MAPREDUCE
Application Tags:	
Application Priority:	0 (Higher Integer value indicates higher priority)
YarnApplicationState:	ACCEPTED: waiting for AM container to be allocated, launched and register with RM.
Queue:	root.default
FinalStatus Reported by AM:	Application has not completed yet.
Started:	Wed Oct 22 14:12:09 +0200 2025
Launched:	N/A
Finished:	N/A
Elapsed:	48sec
Tracking URL:	Unassigned
Diagnostics:	
Unmanaged Application:	false
Application Node Label expression:	<Not set>
AM container Node Label expression:	<DEFAULT_PARTITION>

Show 20 v entries   Search:

Attempt ID	Started	Node	Logs
No data available in table			

Showing 0 to 0 of 0 entries   First Previous Next Last

- Tarea 2: Añadir un nuevo DataNode/NodeManager

1º) Creación de ficheros de nodos incluidos y excluidos

Una vez incluidos los datanodes al HDFS y Yarn mediante los ficheros include podemos ver que se han añadido exitosamente mediante los ficheros log

```
hdadmin@namenode:~$ grep -Ei "include|exclude" archivo.txt $HADOOP_HOME/logs/hadoop-hdadmin-resourcema
nager-namenode.log
grep: archivo.txt: No such file or directory
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-10-14 12:02:56,252 INFO org.apach
e.hadoop.util.HostsFileReader: Refreshing hosts (include/exclude) list
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-10-14 12:11:58,167 INFO org.apach
e.hadoop.util.HostsFileReader: Refreshing hosts (include/exclude) list
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-10-22 10:53:02,949 INFO org.apach
e.hadoop.util.HostsFileReader: Refreshing hosts (include/exclude) list
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-10-22 14:03:55,452 INFO org.apach
e.hadoop.util.HostsFileReader: Refreshing hosts (include/exclude) list
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-12-11 18:00:22,063 INFO org.apach
e.hadoop.util.HostsFileReader: Refreshing hosts (include/exclude) list
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-12-11 18:53:46,506 INFO org.apach
e.hadoop.util.HostsFileReader: Refreshing hosts (include/exclude) list
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-12-11 18:53:46,506 INFO org.apach
e.hadoop.util.HostsFileReader: Adding a node "datanode1" to the list of included hosts from /opt/bd/ha
doop/etc/hadoop/yarn.include
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-12-11 18:53:46,507 INFO org.apach
e.hadoop.util.HostsFileReader: Adding a node "datanode2" to the list of included hosts from /opt/bd/ha
doop/etc/hadoop/yarn.include
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-12-11 18:53:46,507 INFO org.apach
e.hadoop.util.HostsFileReader: Adding a node "datanode3" to the list of included hosts from /opt/bd/ha
doop/etc/hadoop/yarn.include
/opt/bd/hadoop/logs/hadoop-hdadmin-resourcemanager-namenode.log:2025-12-11 18:53:46,507 INFO org.apach
e.hadoop.util.HostsFileReader: Adding a node "datanode4" to the list of included hosts from /opt/bd/ha
doop/etc/hadoop/yarn.include
hdadmin@namenode:~$ grep -Ei "include|exclude" archivo.txt $HADOOP_HOME/logs/hadoop-hdadmin-namenode-n
amenode.log
grep: archivo.txt: No such file or directory
/opt/bd/hadoop/logs/hadoop-hdadmin-namenode-namenode.log:2025-12-11 18:53:31,866 INFO org.apache.hadoo
p.util.HostsFileReader: Adding a node "datanode1" to the list of included hosts from /opt/bd/hadoop/et
c/hadoop/dfs.include
/opt/bd/hadoop/logs/hadoop-hdadmin-namenode-namenode.log:2025-12-11 18:53:31,866 INFO org.apache.hadoo
p.util.HostsFileReader: Adding a node "datanode2" to the list of included hosts from /opt/bd/hadoop/et
c/hadoop/dfs.include
/opt/bd/hadoop/logs/hadoop-hdadmin-namenode-namenode.log:2025-12-11 18:53:31,866 INFO org.apache.hadoo
p.util.HostsFileReader: Adding a node "datanode3" to the list of included hosts from /opt/bd/hadoop/et
c/hadoop/dfs.include
/opt/bd/hadoop/logs/hadoop-hdadmin-namenode-namenode.log:2025-12-11 18:53:31,866 INFO org.apache.hadoo
p.util.HostsFileReader: Adding a node "datanode4" to the list of included hosts from /opt/bd/hadoop/et
c/hadoop/dfs.include
hdadmin@namenode:~$
```

2º) Añadir un nuevo datanode/nodemanager

Conjunto de pasos para añadir un nuevo datanode/nodemanager llamado datanode5
Paso 1) y 2) Lo añadimos primero a YARN

```
hdadmin@namenode:~$ vim $HADOOP_HOME/etc/hadoop/yarn.include
hdadmin@namenode:~$ vim $HADOOP_HOME/etc/hadoop/yarn.include
hdadmin@namenode:~$ yarn rmadmin -refreshNodes
2025-12-11 19:39:31,210 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager
at resourcemanager/172.18.0.2:8033
hdadmin@namenode:~$ cat $HADOOP_HOME/etc/hadoop/yarn.include
datanode1
datanode2
datanode3
datanode4
datanode5
hdadmin@namenode:~$
```

Paso 3) Creamos el contenedor datanode5 con docker

Paso 4) Comprobamos que efectivamente se añade al YARN pero no al HDFS


```
-------------------------------------------------
Live datanodes (4):

Name: 172.18.0.3:9866 (datanode1.hadoop-cluster)
Hostname: datanode1
Decommission Status : Normal
Configured Capacity: 1081101176832 (1006.85 GB)
DFS Used: 338309120 (322.64 MB)
Non DFS Used: 29109428224 (27.11 GB)
DFS Remaining: 996661084160 (928.21 GB)
DFS Used%: 0.03%
DFS Remaining%: 92.19%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 0
Last contact: Thu Dec 11 19:43:06 CET 2025
Last Block Report: Thu Dec 11 18:53:36 CET 2025
Num of Blocks: 34


Name: 172.18.0.4:9866 (datanode2.hadoop-cluster)
Hostname: datanode2
Decommission Status : Normal
Configured Capacity: 1081101176832 (1006.85 GB)
DFS Used: 338481152 (322.80 MB)
Non DFS Used: 29109256192 (27.11 GB)
DFS Remaining: 996661084160 (928.21 GB)
DFS Used%: 0.03%
DFS Remaining%: 92.19%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 0
Last contact: Thu Dec 11 19:43:06 CET 2025
Last Block Report: Thu Dec 11 19:15:46 CET 2025
Num of Blocks: 31


Name: 172.18.0.5:9866 (datanode3.hadoop-cluster)
Hostname: datanode3
Decommission Status : Normal
Configured Capacity: 1081101176832 (1006.85 GB)
DFS Used: 70508544 (67.24 MB)
Non DFS Used: 29377228800 (27.36 GB)
DFS Remaining: 996661084160 (928.21 GB)
DFS Used%: 0.01%
DFS Remaining%: 92.19%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 0
Last contact: Thu Dec 11 19:43:06 CET 2025
Last Block Report: Thu Dec 11 18:53:36 CET 2025
Num of Blocks: 30


Name: 172.18.0.6:9866 (datanode4.hadoop-cluster)
Hostname: datanode4
Decommission Status : Normal
Configured Capacity: 1081101176832 (1006.85 GB)
DFS Used: 271376384 (258.80 MB)
Non DFS Used: 29176360960 (27.17 GB)
DFS Remaining: 996661084160 (928.21 GB)
DFS Used%: 0.03%
DFS Remaining%: 92.19%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 0
Last contact: Thu Dec 11 19:43:06 CET 2025
Last Block Report: Thu Dec 11 18:53:36 CET 2025
Num of Blocks: 36


hdadmin@namenode:~$ yarn node -list
2025-12-11 19:43:36,111 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at resourcemanager/172.18.0.2:8032
2025-12-11 19:43:36,279 INFO client.AHSProxy: Connecting to Application History server at timelineserver/172.18.0.7:10200
Total Nodes:5
         Node-Id             Node-State Node-Http-Address       Number-of-Running-Containers
    datanode1:33565             RUNNING     datanode1:8042                                  0
    datanode4:37361             RUNNING     datanode4:8042                                  0
    datanode5:36029             RUNNING     datanode5:8042                                  0
    datanode3:38285             RUNNING     datanode3:8042                                  0
    datanode2:44921             RUNNING     datanode2:8042                                  0
hdadmin@namenode:~$
```

Último nodo del hdfs es datanode4, pero en yarn si que está

```
Name: 172.18.0.6:9866 (datanode4.hadoop-cluster)
Hostname: datanode4
Decommission Status : Normal
Configured Capacity: 1081101176832 (1006.85 GB)
DFS Used: 271376384 (258.80 MB)
Non DFS Used: 29176360960 (27.17 GB)
DFS Remaining: 996661084160 (928.21 GB)
DFS Used%: 0.03%
DFS Remaining%: 92.19%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 0
Last contact: Thu Dec 11 19:43:06 CET 2025
Last Block Report: Thu Dec 11 18:53:36 CET 2025
Num of Blocks: 36


hdadmin@namenode:~$ yarn node -list
2025-12-11 19:43:36,111 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager
at resourcemanager/172.18.0.2:8032
2025-12-11 19:43:36,279 INFO client.AHSProxy: Connecting to Application History server at timelineserv
er/172.18.0.7:10200
Total Nodes:5
         Node-Id             Node-State Node-Http-Address     Number-of-Running-Containers
  datanode1:33565              RUNNING    datanode1:8042                               0
  datanode4:37361              RUNNING    datanode4:8042                               0
  datanode5:36029              RUNNING    datanode5:8042                               0
  datanode3:38285              RUNNING    datanode3:8042                               0
  datanode2:44921              RUNNING    datanode2:8042                               0
hdadmin@namenode:~$
```

Pasos 5) 6) 7) Lo añadimos al HDFS

```
hdadmin@namenode:~$ vim $HADOOP_HOME/etc/hadoop/dfs.include
hdadmin@namenode:~$
hdadmin@namenode:~$
hdadmin@namenode:~$ cat $HADOOP_HOME/etc/hadoop/dfs.include
datanode1
datanode2
datanode3
datanode4
datanode5
hdadmin@namenode:~$ hdfs dfsadmin -refreshNodes
Refresh nodes successful
hdadmin@namenode:~$
```

Comprobamos que ahora sí está en el hdfs

```
Name: 172.18.0.9:9866 (datanode5.hadoop-cluster)
Hostname: datanode5
Decommission Status : Normal
Configured Capacity: 1081101176832 (1006.85 GB)
DFS Used: 314905 (307.52 KB)
Non DFS Used: 29448348135 (27.43 GB)
DFS Remaining: 996660158464 (928.21 GB)
DFS Used%: 0.00%
DFS Remaining%: 92.19%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 0
Last contact: Thu Dec 11 19:53:06 CET 2025
Last Block Report: Thu Dec 11 19:52:33 CET 2025
Num of Blocks: 2
```

Salida del balanceador de carga, se han movido 0 bytes.

```
hdadmin@namenode:~$ hdfs balancer
2025-12-11 19:55:41,811 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-12-11 19:55:41,910 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-12-11 19:55:41,910 INFO impl.MetricsSystemImpl: Balancer metrics system started
2025-12-11 19:55:41,951 INFO balancer.Balancer: namenodes  = [hdfs://namenode:9000]
2025-12-11 19:55:41,953 INFO balancer.Balancer: parameters = Balancer.BalancerParameters [BalancingPol
icy.Node, threshold = 10.0, max idle iteration = 5, #excluded nodes = 0, #included nodes = 0, #source
nodes = 0, #blockpools = 0, run during upgrade = false, sort top nodes = false, hot block time interva
l = 0]
2025-12-11 19:55:41,953 INFO balancer.Balancer: included nodes = []
2025-12-11 19:55:41,953 INFO balancer.Balancer: excluded nodes = []
2025-12-11 19:55:41,953 INFO balancer.Balancer: source nodes = []
Time Stamp               Iteration# Bytes Already Moved  Bytes Left To Move  Bytes Being Moved  NameN
ode
2025-12-11 19:55:41,957 INFO balancer.NameNodeConnector: getBlocks calls for hdfs://namenode:9000 will
 be rate-limited to 20 per second
2025-12-11 19:55:43,304 INFO balancer.Balancer: dfs.namenode.get-blocks.max-qps = 20 (default=20)
2025-12-11 19:55:43,304 INFO balancer.Balancer: dfs.balancer.movedWinWidth = 5400000 (default=5400000)
2025-12-11 19:55:43,304 INFO balancer.Balancer: dfs.balancer.moverThreads = 1000 (default=1000)
2025-12-11 19:55:43,304 INFO balancer.Balancer: dfs.balancer.dispatcherThreads = 200 (default=200)
2025-12-11 19:55:43,304 INFO balancer.Balancer: dfs.balancer.getBlocks.size = 2147483648 (default=2147
483648)
2025-12-11 19:55:43,305 INFO balancer.Balancer: dfs.balancer.getBlocks.min-block-size = 10485760 (defa
ult=10485760)
2025-12-11 19:55:43,305 INFO balancer.Balancer: dfs.datanode.balance.max.concurrent.moves = 100 (defau
lt=100)
2025-12-11 19:55:43,305 INFO balancer.Balancer: dfs.datanode.balance.bandwidthPerSec = 104857600 (defa
ult=104857600)
2025-12-11 19:55:43,309 INFO balancer.Balancer: dfs.balancer.max-size-to-move = 10737418240 (default=1
0737418240)
2025-12-11 19:55:43,310 INFO balancer.Balancer: dfs.blocksize = 67108864 (default=134217728)
2025-12-11 19:55:43,333 INFO net.NetworkTopology: Adding a new node: /default-rack/172.18.0.3:9866
2025-12-11 19:55:43,333 INFO net.NetworkTopology: Adding a new node: /default-rack/172.18.0.4:9866
2025-12-11 19:55:43,333 INFO net.NetworkTopology: Adding a new node: /default-rack/172.18.0.9:9866
2025-12-11 19:55:43,333 INFO net.NetworkTopology: Adding a new node: /default-rack/172.18.0.5:9866
2025-12-11 19:55:43,333 INFO net.NetworkTopology: Adding a new node: /default-rack/172.18.0.6:9866
2025-12-11 19:55:43,334 INFO balancer.Balancer: 0 over-utilized: []
2025-12-11 19:55:43,334 INFO balancer.Balancer: 0 underutilized: []
Dec 11, 2025 7:55:43 PM           0                    0 B                 0 B                0 B
         0  hdfs://namenode:9000
The cluster is balanced. Exiting...
Dec 11, 2025 7:55:43 PM  Balancing took 1.939 seconds
hdadmin@namenode:~$
```

- Tarea 3: Retirar un DataNode/NodeManager

Interfaz web del YARN

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Node
4	0	1	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum Cluster Application Priority	Scheduler Busy %
Capacity Scheduler	[memory-mb (unit=Mi), vcores]	<memory:128, vCores:1>	<memory:2560, vCores:1>	0	0

Show 20 entries

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Allocation Tags	Mem Used	M A
	/default-rack	DECOMMISSIONED	datanode4:37361	N/A	Thu Dec 11 20:16:31 +0100 2025		0		0 B	0 B

Showing 1 to 1 of 1 entries

Interfaz web de HDFS

DataNode State [All]   Show [25] entries   Search: [   ]

Node	Http Address	Last contact	Last Block Report	Used	Non DFS Used	Capacity	Blocks	Block pool used	Block pool usage StdDev	Version
✔/default-rack/datanode2:9866 (172.18.0.4:9866)	http://datanode2:9864	1s	56m	323.33 MB	27.36 GB	1006.85 GB	38	323.33 MB (0.03%)	0%	3.4.1
✔/default-rack/datanode5:9866 (172.18.0.9:9866)	http://datanode5:9864	0s	23m	126.48 MB	27.56 GB	1006.85 GB	14	126.48 MB (0.01%)	0%	3.4.1
✔/default-rack/datanode3:9866 (172.18.0.5:9866)	http://datanode3:9864	0s	26m	198.51 MB	27.48 GB	1006.85 GB	37	198.51 MB (0.02%)	0%	3.4.1
✔/default-rack/datanode1:9866 (172.18.0.3:9866)	http://datanode1:9864	1s	76m	322.96 MB	27.36 GB	1006.85 GB	42	322.96 MB (0.03%)	0%	3.4.1
⊘/default-rack/datanode4:9866 (172.18.0.6:9866)	http://datanode4:9864	0s	76m	258.8 MB	27.43 GB	1006.85 GB	36	258.8 MB (0.03%)	0%	3.4.1

Showing 1 to 5 of 5 entries   Previous [1] Next

- Tarea 4: Retirar un DataNode/NodeManager

Topología antes de añadir racks

```
hdadmin@namenode:~$
hdadmin@namenode:~$ hdfs dfsadmin -printTopology
Rack: /default-rack
   172.18.0.9:9866 (datanode5.hadoop-cluster) In Service
   172.18.0.4:9866 (datanode2.hadoop-cluster) In Service
   172.18.0.6:9866 (datanode4.hadoop-cluster) Decommissioned
   172.18.0.5:9866 (datanode3.hadoop-cluster) In Service
   172.18.0.3:9866 (datanode1.hadoop-cluster) In Service

hdadmin@namenode:~$ hdfs --daemon stop namenode
```

Topología de la red después de añadir rack awareness

```
hdadmin@namenode:~$ hdfs dfsadmin -printTopology
Rack: /default-rack
   172.18.0.6:9866 (datanode4.hadoop-cluster) Decommissioned

Rack: /rack1
   172.18.0.4:9866 (datanode2.hadoop-cluster) In Service
   172.18.0.3:9866 (datanode1.hadoop-cluster) In Service

Rack: /rack2
   172.18.0.9:9866 (datanode5.hadoop-cluster) In Service
   172.18.0.5:9866 (datanode3.hadoop-cluster) In Service

hdadmin@namenode:~$ |
```