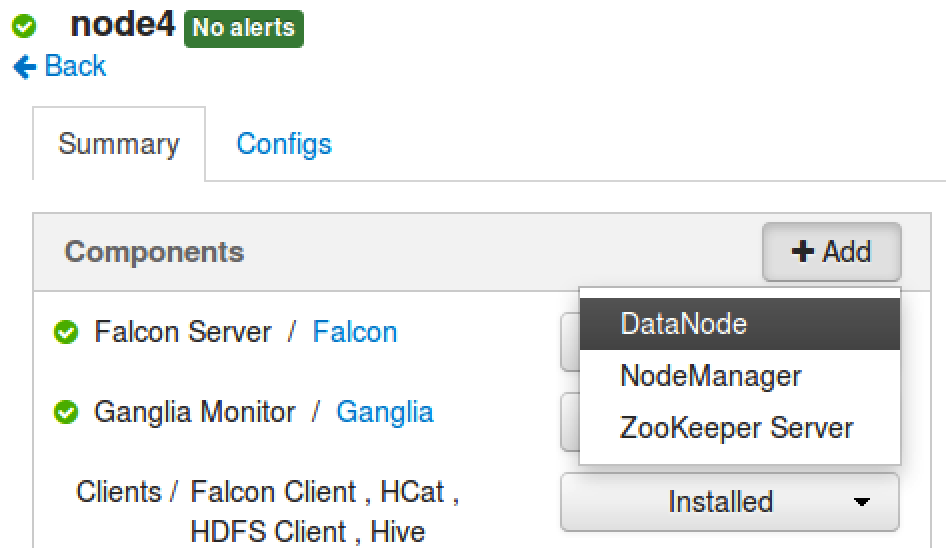
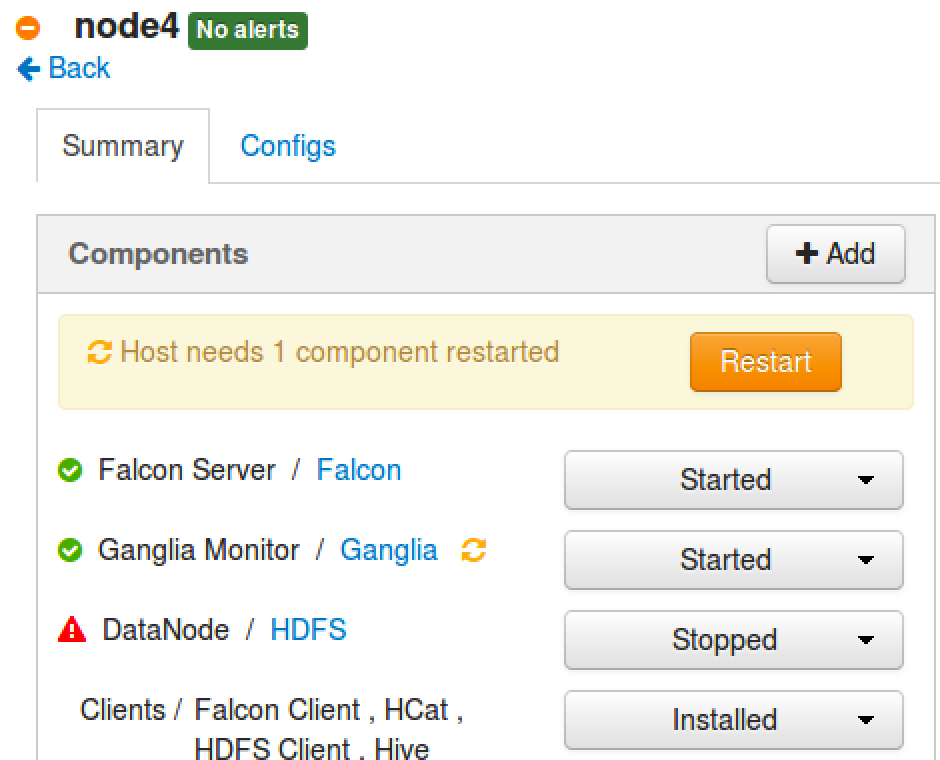
Lab: Commissioning & Decommissioning Services

|  |  |
| --- | --- |
| **Objective:** | Commission a new DataNode to an existing cluster, and also decommission a node. |
| **Successful Outcome:** | A hadoop-slave is successfully added to your cluster as a DataNode, and hadoop-master is successfully decommissioned from the cluster. |
| **Before You Begin:** | Open the Ambari UI (port 8080) |

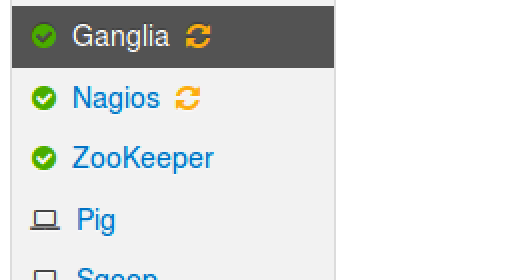
1. Copy a significant big directory to HDFS
   1. Copy /root/repo directory to HDFS.
2. Commission a New DataNode.
   1. Go to the Hosts page of Ambari.
   2. Click on the host for hadoop-slaveX.
   3. Click on the + Add button and select DataNode:



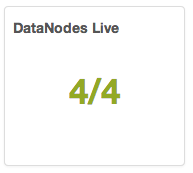
* 1. Click ‘Confirm Add’ when the confirmation dialog appears.
  2. Wait for the DataNode component to be installed. When the install is complete, DataNode should appear in the list of Components on hadoop-slaveX:



1. Start the DataNode
   1. Click on the Action menu next to DataNode and choose Start.
   2. Wait for the DataNode service to start.
2. Click on Restart button to start any dependent services. Also restart Ganglia and Nagios services if “restart” icon is there on Ambari Dashboard.



1. Verify the Commissioned Node
   1. Once the DataNode is started successfully on hadoop-slaveX, go back to the Ambari Dashboard. You should see 4 live DataNodes:



1. Run the Balancer
   1. Go to the NameNode UI. Notice it also shows 4 live DataNodes.
   2. Click on Live Nodes link.
   3. Notice that the new DataNode does not have any data blocks currently.
   4. Run following command to run the balancer process on your cluster:

# su -l hdfs -c "hdfs balancer -threshold 1"

* 1. Wait a couple minutes for the balancer to even out the block storage. You will see at the command prompt as blocks get moved from one node to another:

INFO balancer.Balancer: 0 over-utilized: []

INFO balancer.Balancer: 1 underutilized: [BalancerDatanode[10.222.133.205:50010, utilization=0.40288804420577484]]

INFO balancer.Balancer: Need to move 173.74 MB to make the cluster balanced.

INFO balancer.Balancer: Decided to move 89.95 MB bytes from 10.170.202.246:50010 to 10.222.133.205:50010

INFO balancer.Balancer: Decided to move 151.79 MB bytes from 10.174.49.252:50010 to 10.222.133.205:50010

INFO balancer.Balancer: Will move 241.74 MB in this iterationINFO balancer.Balancer: Moving block 1073742573 from 10.174.49.252:50010 to 10.222.133.205:50010 through 10.174.50.60:50010 is succeeded.

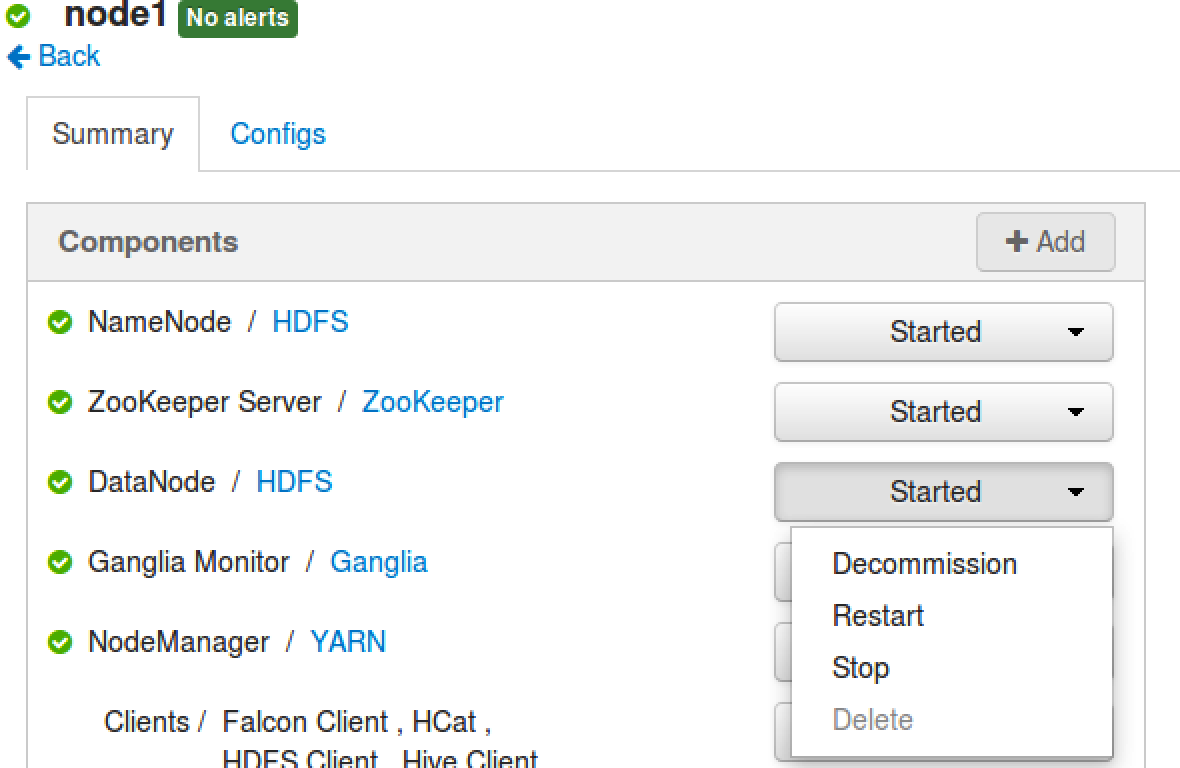
INFO balancer.Balancer: Moving block 1073742572 from 10.174.49.252:50010 to 10.222.133.205:50010 through 10.174.50.60:50010 is succeeded.

...

* 1. Refresh the Live Nodes page of the NameNode UI. Your hadoop-slaveX DataNode should now have blocks on it, and the number of blocks will gradually increase as the balancer app continue to even out the block storage on your cluster.

**NOTE**: The balancer app will run for a long time. Just leave the process open in your terminal window. If you need to perform any future tasks on hadoop-master, just open a new terminal window.

1. Decommission a DataNode
   1. From the Hosts page of Ambari, click on the host for hadoop-master.
   2. In the Components section, click on the Action menu next to DataNode and select Decommission:



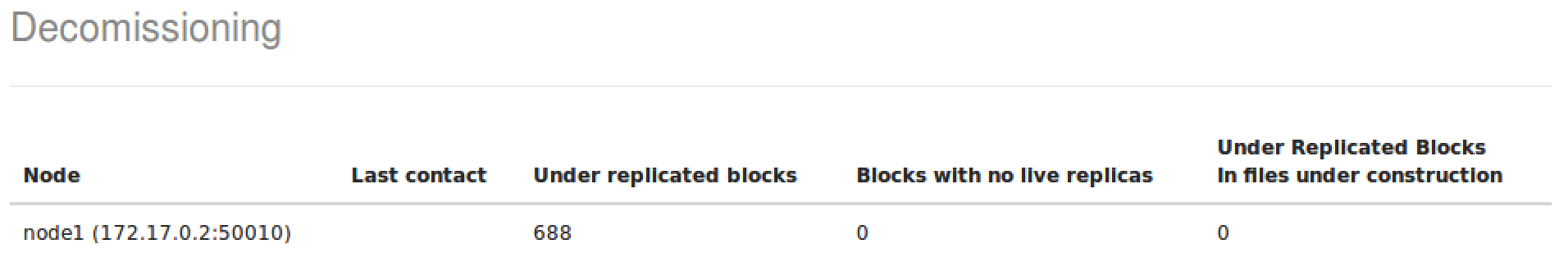
* 1. Click OK in the confirmation dialog, and wait for the decommissioning task to complete.

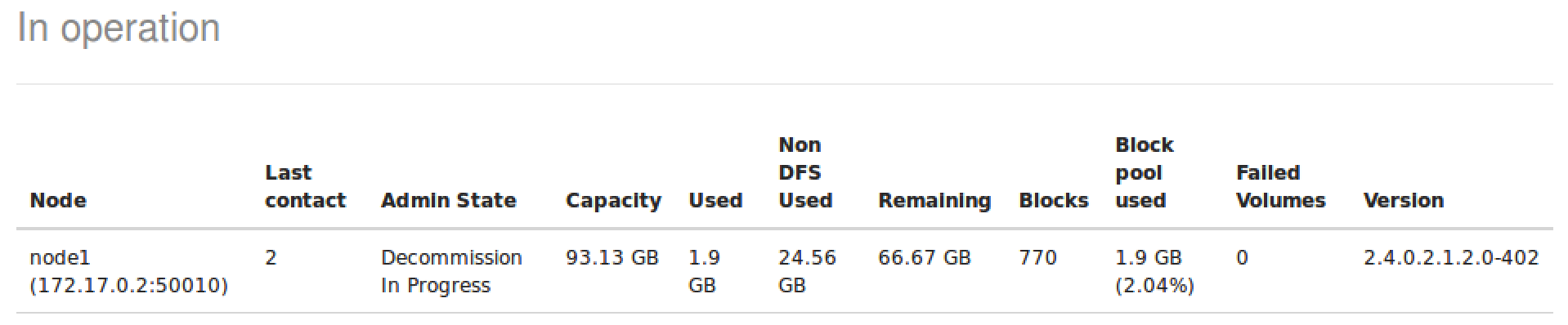
**NOTE**: There is a minimal chance that the decommissioning task may fail due to a known bug in Hadoop 2.0 where the node contains a block that belongs to a file with a replication factor larger than the rest of the cluster size. The work-around is to locate and delete any files that have a replication factor larger than 3. View <https://issues.apache.org/jira/browse/HDFS-5662> for more details.

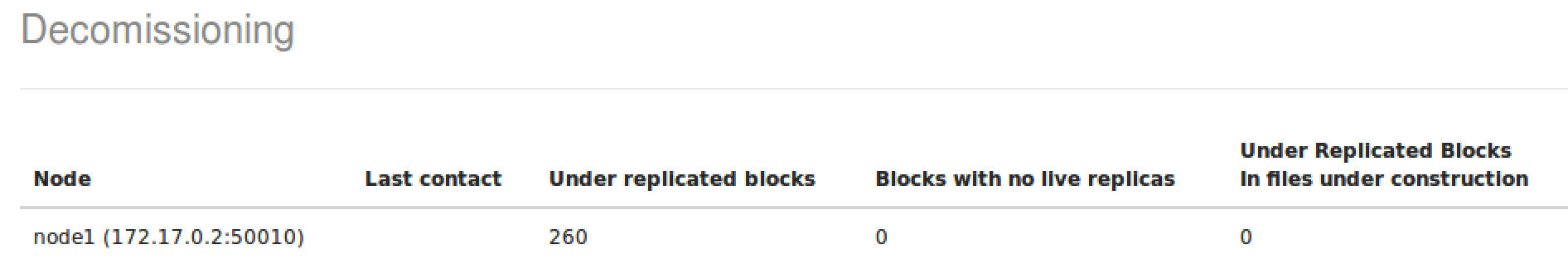
1. View the NameNode UI
   1. Go to the Namenode UI. Notice that Decommissioning Nodes is 1 and Live Nodes is still X:

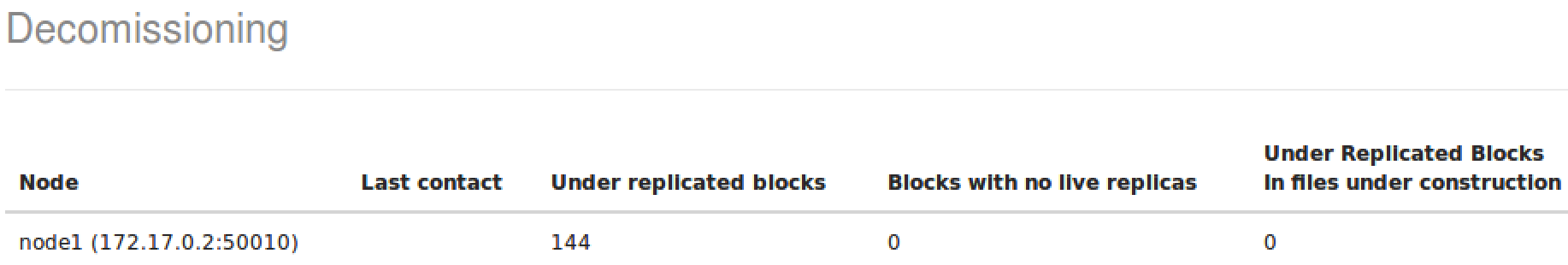


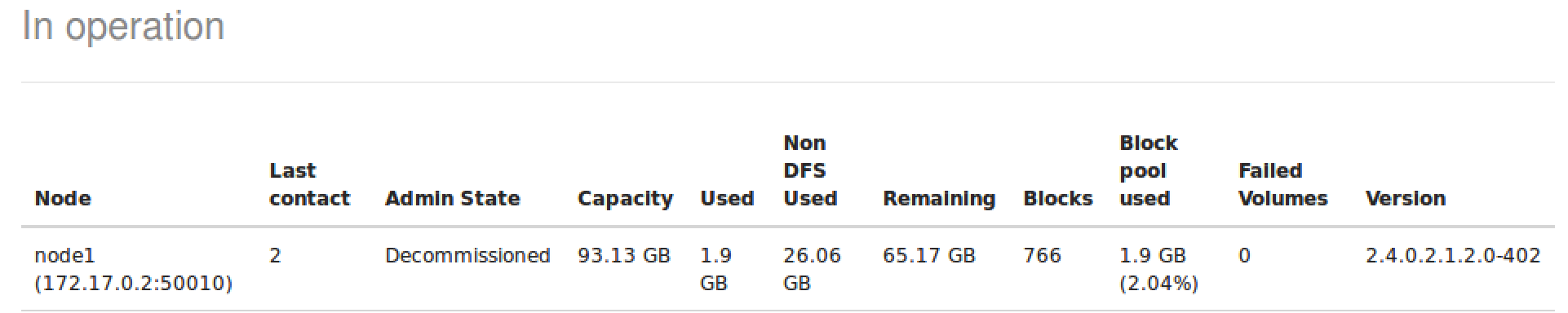
* 1. Click on Decommissioning Nodes and it will show that hadoop-master is undergoing the decommission process.



* 1. You will also see that blocks are gradually being copied from hadoop-master to the other nodes. The Admin State of hadoop-master is going to be either Decommission in Progress or Decommissioned. Refresh the page until the status is Decommissioned.

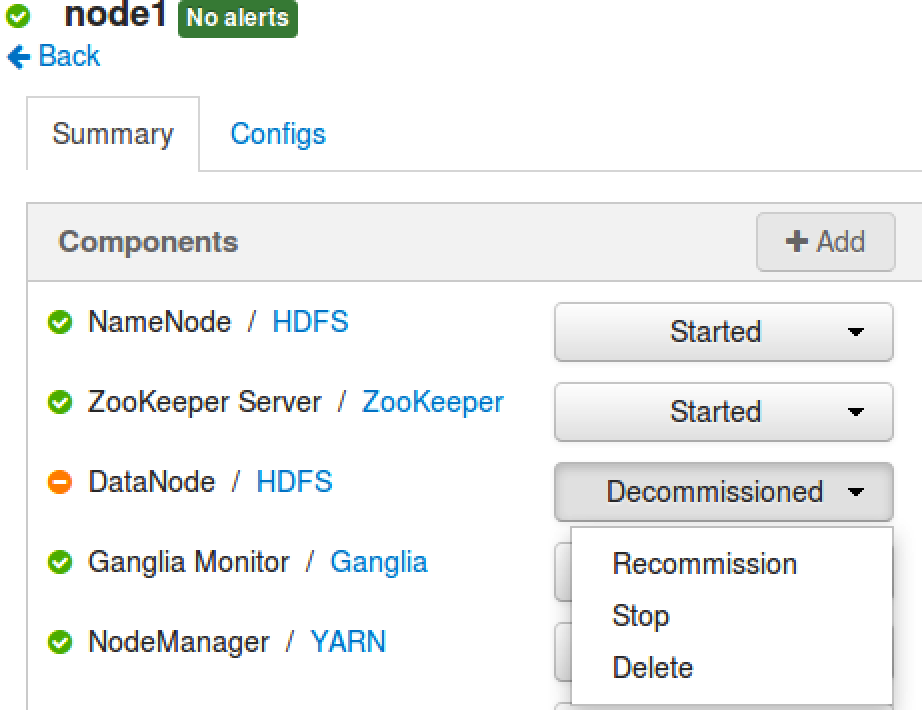




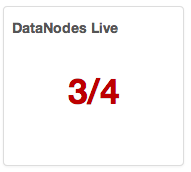


* 1. Go back to the NameNode UI page. Notice you have 4 Live Nodes, and 1 of them is Decommissioned:

1. Stop the DataNode
   1. Once the decommissioning process is complete, go back to the Hosts page for hadoop-master in Ambari.
   2. In the Action menu next to DataNode, select Stop:



* 1. From the Ambari Dashboard page, you should see 3/4 live DataNodes:



**NOTE**: You can decommission/commission a nodemanager also in the same way.

**RESULT**: You have now seen how to commission a new DataNode, and also how to run the balancer tool to balance the blocks across a cluster once new DataNodes are commissioned. You also have decommissioned one of the DataNodes from your cluster.