

Datanode Decommissioning in HDFS

Source:-<https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/HdfsDataNodeAdminGuide.html>

To decommission a datanode we need to add the following properties in the hdfs-site.xml in both the namenodes..

cat /etc/hadoop/conf/hdfs-site.xml

```
<property>
  <name>dfs.namenode.replication.work.multiplier.per.iteration</name>
  <value>100</value>
</property>

<property>
  <name>dfs.balancer.max-size-to-move</name>
  <value>26843545600</value>
</property>

<property>
  <name>dfs.hosts.exclude</name>
  <value>path_to_/dfs.exclude</value>
</property>
</configuration>
```

Namenodes needs to be **restarted** once the above configurations are added..

we need to add one more file as mentioned on the above path `path_to_/dfs.exclude` in both the namenodes

In the above file we need to mention the datanodes that needs to be decommissioned..

```
10.0.9.****
10.0.8.****
```

we need to execute the following command to start the decommissioning

```
hdfs dfsadmin -refreshNodes
```

Monitor the namenode UI for any missing blocks report or we can use the following command to check on the missing blocks and datanode decommissioning

```
hdfs dfsadmin -report
```