**HDFS REPLICATION AND BLOCK SIZE**

Replication factor dictates how many copies of a block should be kept in your cluster. The replication factor is 3 by default and hence any file you create in HDFS will have a replication factor of 3 and each block from the file will be copied to 3 different nodes in your cluster.

1. **Configure replication factor**

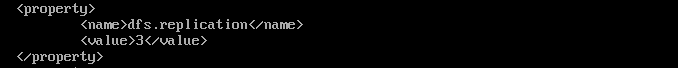
Edit **$HADOOP\_HOME/etc/hadoop/hdfs-site.xml** file. Here the replication is 3.

***<property>***

***<name>dfs.replication<name>***

***<value>3<value>***

***<property>***



We will put a file to HDFS and check the replication.

(You can access the following link to download the example files in **Data** folder.

**https://drive.google.com/open?id=1WYIOJfupaq2X411KWEROJF1anAD5mfYQ**

)



Check the replication after finishing copying by using the following command.

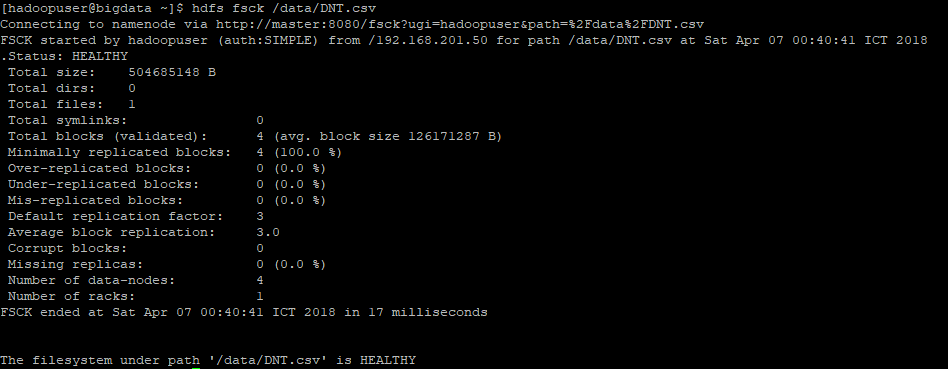
***hdfs dfs –ls /data***



As the result, the second column of the output shows the replication factor of the file.

Here is the command to check the health of file.

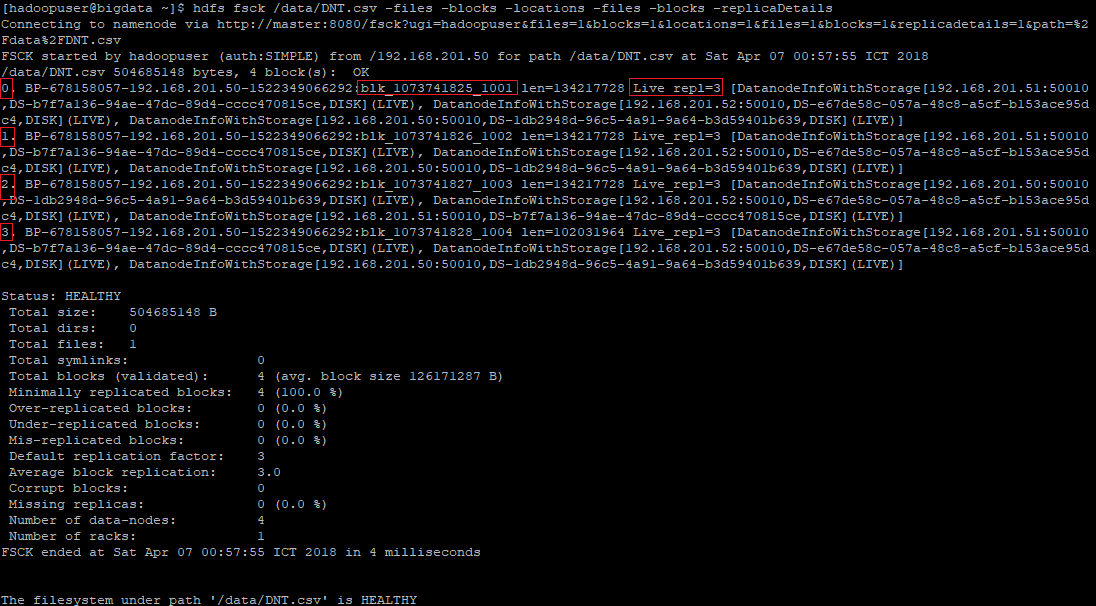
***hdfs fsck /data/DNT.csv***



The default Block Size on Hadoop 1 - 64MB, on Hadoop 2 – 128 MB. Here I am using Hadoop 2.x, so my Block Size is 128 MB. In the above example, I copied a ~ 482 MB to HDFS, and it will be splitted to 4 blocks.

The following command will print out locations for every block of the file.

***hdfs fsck /data/DNT.csv -files -blocks -locations -files -blocks -replicaDetails***

****

* BP-678158057-192.168.201.50-15223490: Block Pool ID

Read about Block Pools here: <https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/Federation.html>

* blk\_1073741825\_1001: Block ID
* len=134217728: Length of the block. Number of bytes in the block
* Live repl=3: There are 3 replicas of this block
* DatanodeInfoWithStorage[192.168.201.51:50010,DS-b7f7a136….,DISK]:
  + 192.168.201.51 => IP address of the Data Node holding this block
  + 50010: => Data streaming port
  + DS-b7f7a136…: => Storage ID
  + DISK: => Storage type. Storage type can be: RAM\_DISK, SSD, DISK and ARCHIVE

We can use the -setrep command to change the replication factor for files that already exist in HDFS. -R flag would recursively change the replication factor on all the files under the specified folder.

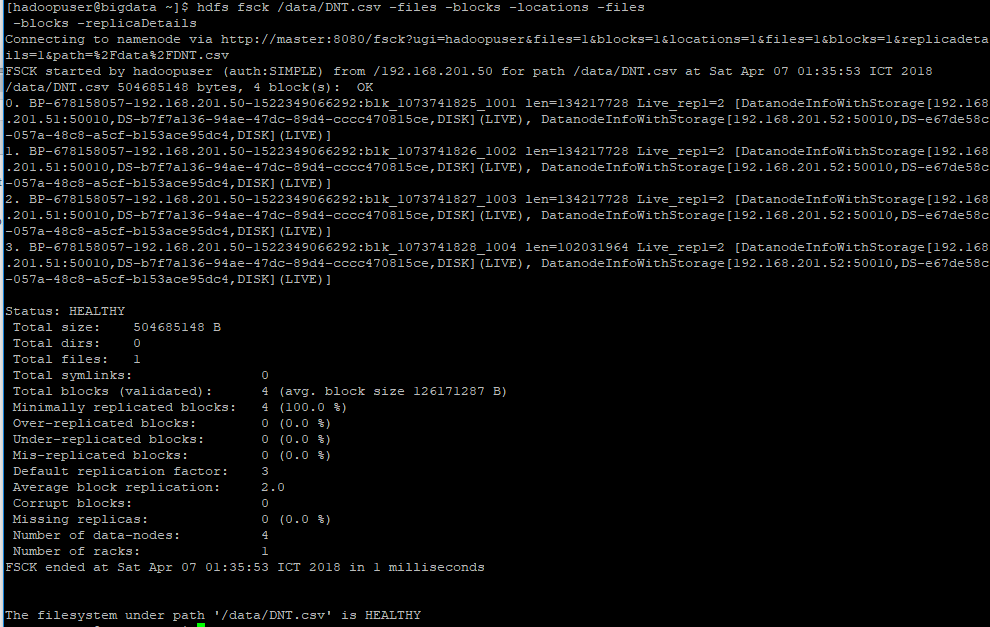
For example,

***hdfs dfs -setrep 2 /data/DNT.csv***



Check the result.



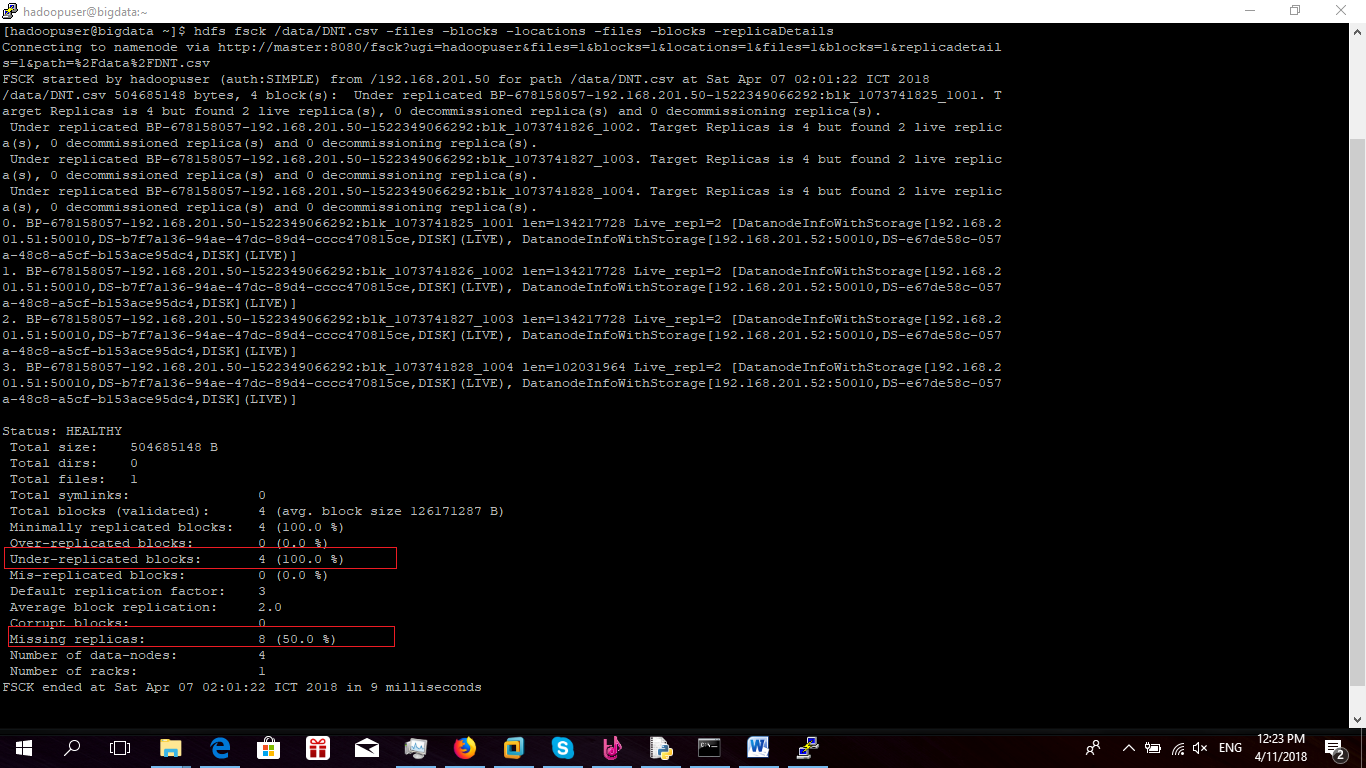


Now we try to set the replication factor is greater than number of nodes. Here I set the replication 4 for DNT.csv file.

***hdfs dfs -setrep 4 /data/DNT.csv***

We check the result and see that

* Under-replicated blocks: 4
* Missing replicas: 8



So what are **Under-replicated blocks** and **Missing replicas**?

* **Under-replicated blocks**: These are blocks that do not meet their target replication for the file they belong to. HDFS will automatically create new replicas of under-replicated blocks until they meet the target replication. You can get information about the blocks being replicated (or waiting to be replicated) using hdfs dfsadmin –metasave.
* **Missing replicas**: These are blocks that do not satisfy the block replica placement policy. For example, for a replication level of three in a multirack cluster, if all three replicas of a block are on the same rack, then the block is misreplicated because the replicas should be spread across at least two racks for resilience. HDFS will automatically re-replicate misreplicated blocks so that they satisfy the rack placement policy.

Next we will fix this by commissioning node3 to Hadoop cluster system.

* **On master**, comment node3 in **$HADOOP\_HOME/etc/hadoop/exclude**



* Refresh nodes

*hdfs dfsadmin -refreshNodes*

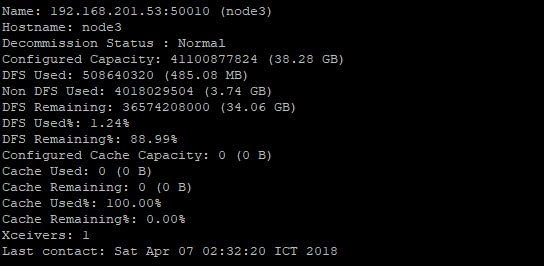
*yarn rmadmin –refreshNodes*





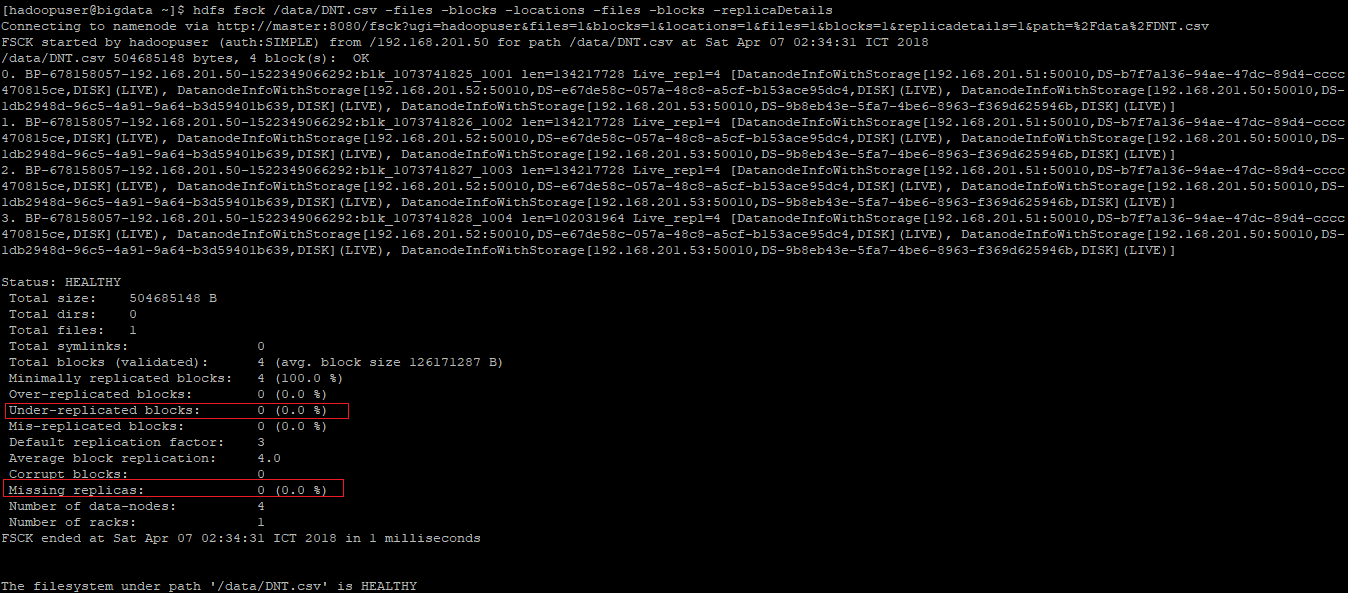
* Check health of nodes: We will see Decommission Status is normal.

*hdfs dfsadmin –report*



The HDFS try to recover the under replicated blocks automatically. Check the result again.

***hdfs fsck /data/DNT.csv -files -blocks -locations -files -blocks -replicaDetails***



* Under-replicated blocks: 0
* Missing replicas: 0

1. **Configure blocksize**

The default Block Size on Hadoop 1 - 64MB, on Hadoop 2 – 128 MB.

As the above example, we see that length of block (len=134217728) is 128 MB (128 \* 1024 \* 1024)



We can change the block size by setting the following property in **hdfs-site.xml**. Here I set it 3 MB (3 \* 1024 \* 1024)

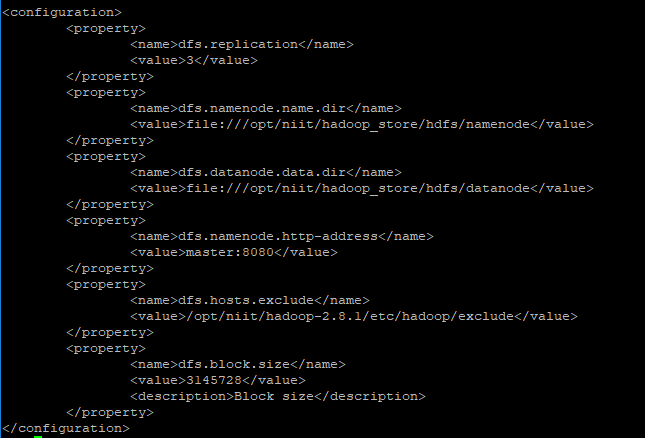
*<property>*

*<name>dfs.block.size<name>*

*<value>3145728<value>*

*<description>Block size<description>*

*</property>*



We will copy another file (39.1 MB) to HDFS.





And check the result.

***hdfs fsck /data/NYSE.csv -files -blocks -locations -files -blocks -replicaDetails***

