**HDFS - Task 2**

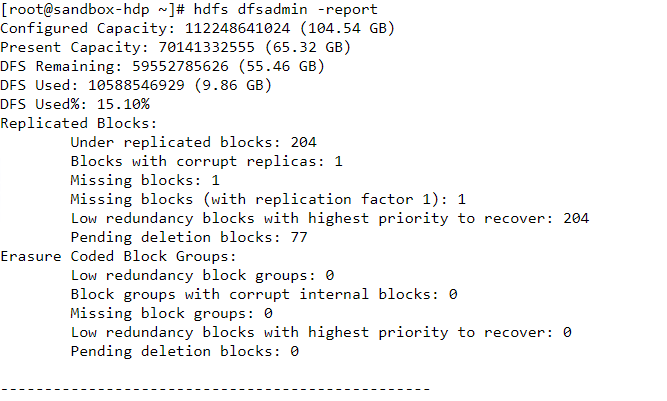
**HDFS:**

* Hadoop distributed file system
* written in java
* store files across machine in a cluster

**HDFS current State Report:**

The below command provide the report of the current state of hdfs.

hdfs dfsadmin –report



1. **Under replicate Blocks:**

The blocks under replication is 204.

1. **Configured Capacity**

It is the total capacity available to HDFS for storage is 104.54 GB. It is calculated as follows:

Configured Capacity = Total Disk Space - Reserved Space

1. **Present Capacity**

The total amount actual space available to store files is 65.32 GB

1. **DFS Remaining**

It is the amount of storage space still available to the HDFS to store more files is 9.86 GB.

1. **DFS Used**

The total space used by DFS is 55.46 GB.

Present Capacity = DFS Used + DFS Remaining

1. **DFS used Percentage**

The total space used by DFS in percentage = 15.10%.

1. **Blocks with corrupt replicas**

No. of blocks which has not valid replica of hdfs data block = 204

1. **Corrupted blocks**

No. invalid blocks = 1

1. **Missing blocks**

Block missing in hdfs = 1

1. **Pending Deleted blocks**

Blocks which are pending for deletion = 77

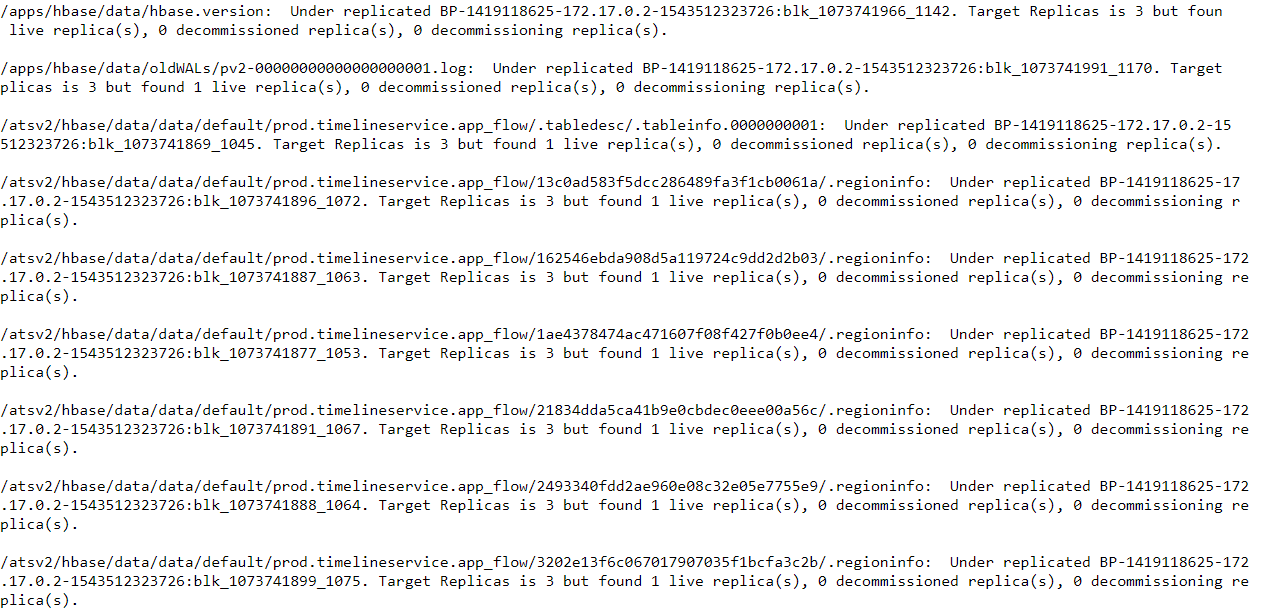
**HDFS File system check**

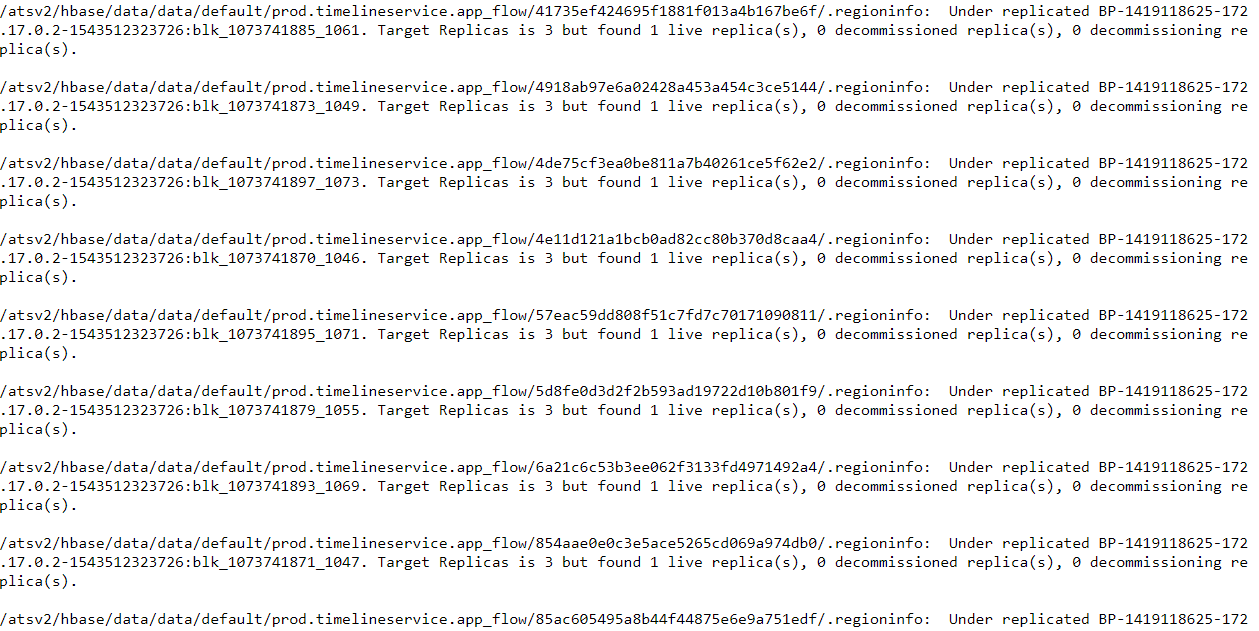
The below command can use to determine health of file system of hadoop:

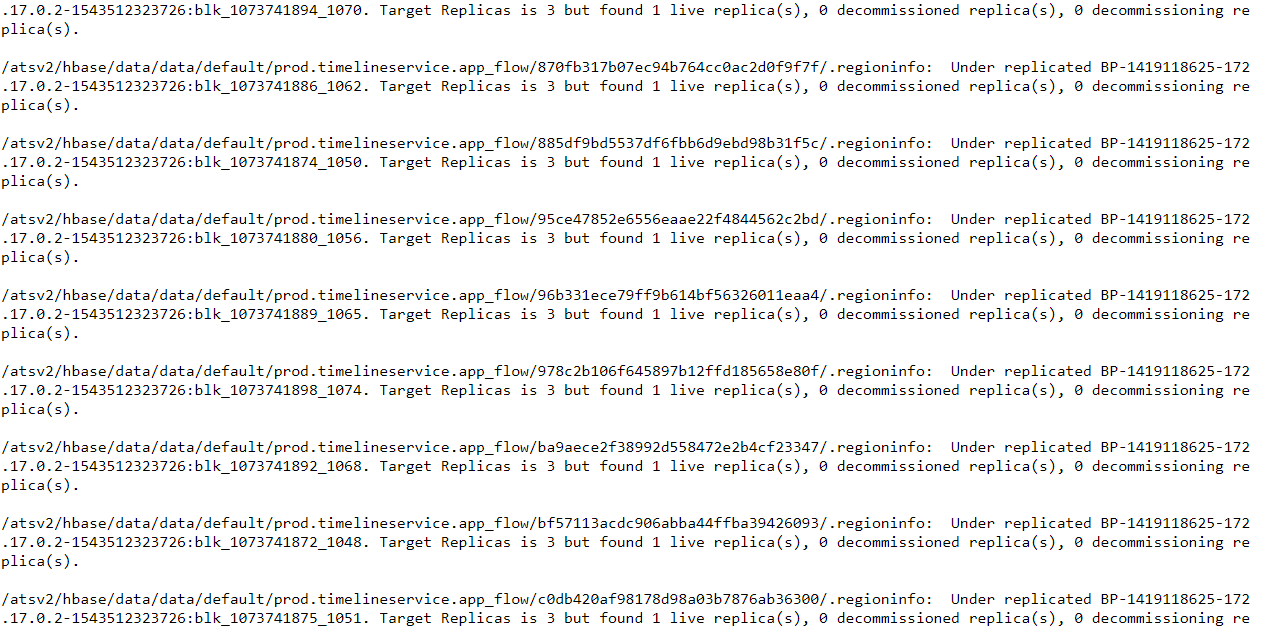
hdfs fsck /

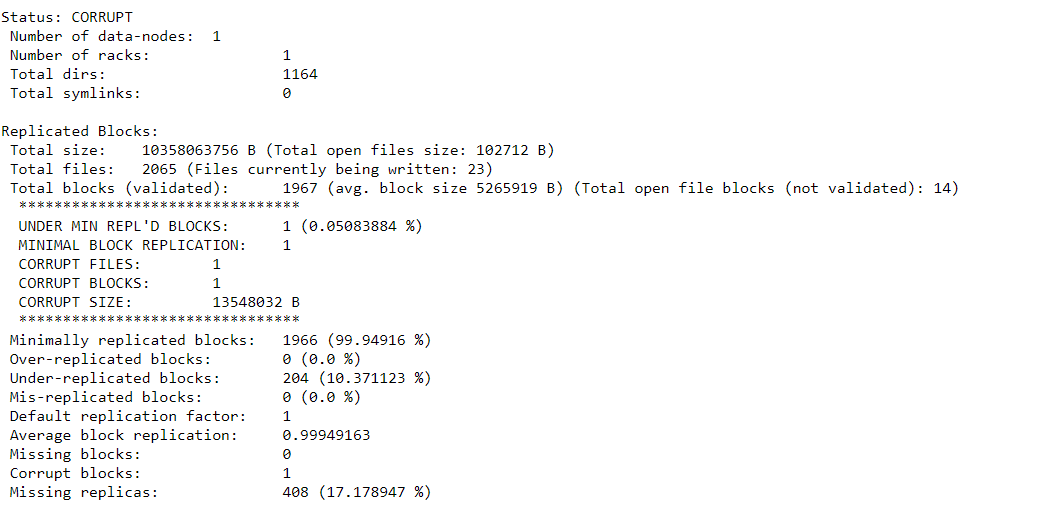
Note: we can use sudo to provide admin permission.

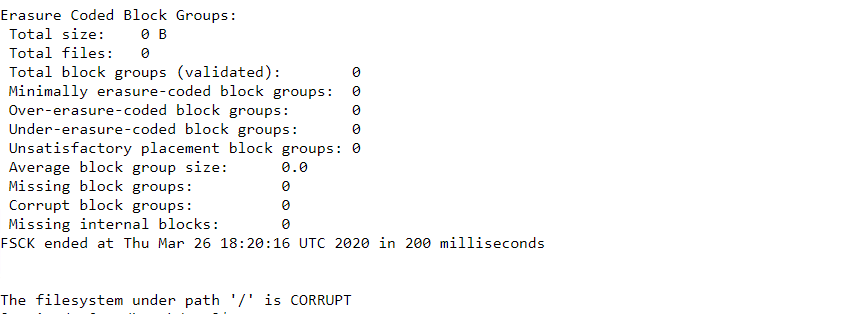










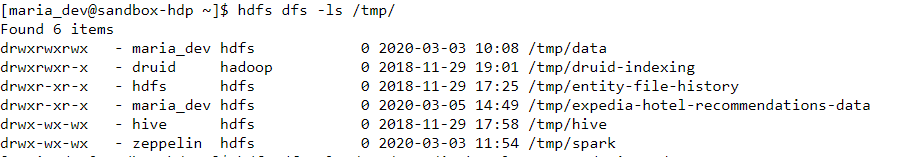


1. **Status:** Corrupted means corrupted files are found in hdfs.
2. **No. of Data Nodes:** The no. of data in which Hadoop can store data or connect to Hadoop = 1
3. **Total Size:** Total size of replicated blocks is 10358063756 B
4. **Total files:** Total replicate files present is 2065
5. **Total Blocks:** Total valid blocks present is 1967
6. **Corrupted Files:** Total number of invalid files found = 1
7. **Corrupted blocks:** Total no. of invalid blocks corrupted in hdfs = 1
8. **Corrupted size:** Total blocks invalid blocks size = 13548032 B

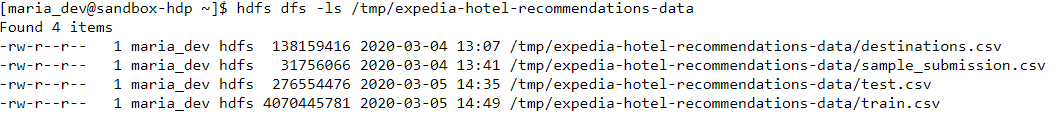
**Move file in hdfs:**

The below command can use to list files.

Hdfs dfs –ls /tmp/



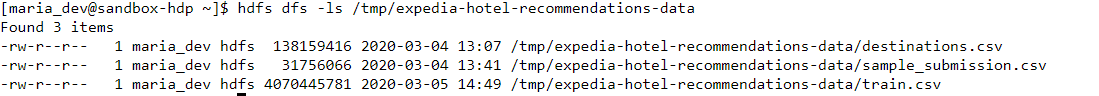
In the above image, I want to move the test.csv file



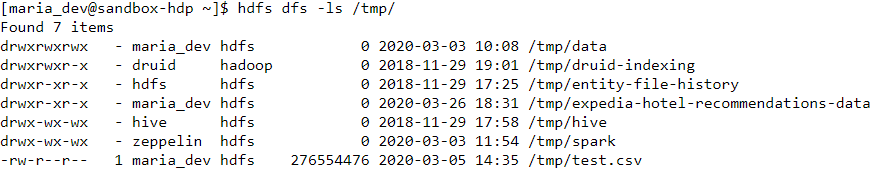
In the above image, I found test.csv file at location /tmp/expedia-hotel-recommendation-data/



In the above image, I have moved test.csv to location /tmp/ with name test.csv



In above image, now the file test.csv not present in /tmp/expedia-hotel-recommendations-data



In the above image, we can see the file test.csv is successfully moved to /tmp location