HDFS Commands

1. To display the version of hadoop use

**Hadoop version**

vagrant@master:~$ hadoop version

Hadoop 2.8.0

Subversion https://git-wip-us.apache.org/repos/asf/hadoop.git -r 91f2b7a13d1e97be65db92ddabc627cc29ac0009

Compiled by jdu on 2017-03-17T04:12Z

Compiled with protoc 2.5.0

From source with checksum 60125541c2b3e266cbf3becc5bda666

This command was run using /home/vagrant/bigdata/hadoop/share/hadoop/common/hadoop-common-2.8.0.jar

vagrant@master:~$

1. List the contents of the root directory in HDFS

**hdfs dfs –ls**

vagrant@master:~$ hdfs dfs -ls

Found 1 items

drwxr-xr-x - vagrant supergroup 0 2019-05-15 04:30 wordcount

1. Report the amount of space used and available on currently mounted file system

**hdfs dfs -df hdfs:/**

vagrant@master:~$ hdfs dfs -df hdfs:/

Filesystem Size Used Available Use%

hdfs://localhost:9000 42241163264 241664 34749714432 0%

1. Count the number of directories, files and bytes under the paths that match the specified file pattern

**hdfs dfs -count hdfs:/**

vagrant@master:~$ hdfs dfs -count hdfs:/

24 6 171538 hdfs:///

1. Run a cluster balancing utility

**hadoop balancer**

DEPRECATED: Use of this script to execute hdfs command is deprecated.

Instead use the hdfs command for it.

19/07/05 10:37:33 INFO balancer.Balancer: namenodes = [hdfs://localhost:9000]

19/07/05 10:37:33 INFO balancer.Balancer: parameters = Balancer.BalancerParameters [BalancingPolicy.Node, threshold = 10.0, max idle iteration = 5, #excluded nodes = 0, #included nodes = 0, #source nodes = 0, #blockpools = 0, run during upgrade = false]

19/07/05 10:37:33 INFO balancer.Balancer: included nodes = []

19/07/05 10:37:33 INFO balancer.Balancer: excluded nodes = []

19/07/05 10:37:33 INFO balancer.Balancer: source nodes = []

Time Stamp Iteration# Bytes Already Moved Bytes Left To Move Bytes Being Moved

19/07/05 10:37:35 INFO balancer.Balancer: dfs.balancer.movedWinWidth = 5400000 (default=5400000)

19/07/05 10:37:35 INFO balancer.Balancer: dfs.balancer.moverThreads = 1000 (default=1000)

19/07/05 10:37:35 INFO balancer.Balancer: dfs.balancer.dispatcherThreads = 200 (default=200)

19/07/05 10:37:35 INFO balancer.Balancer: dfs.datanode.balance.max.concurrent.moves = 50 (default=50)

19/07/05 10:37:35 INFO balancer.Balancer: dfs.balancer.getBlocks.size = 2147483648 (default=2147483648)

19/07/05 10:37:35 INFO balancer.Balancer: dfs.balancer.getBlocks.min-block-size = 10485760 (default=10485760)

19/07/05 10:37:35 INFO balancer.Balancer: dfs.balancer.max-size-to-move = 10737418240 (default=10737418240)

19/07/05 10:37:35 INFO balancer.Balancer: dfs.blocksize = 134217728 (default=134217728)

19/07/05 10:37:35 INFO net.NetworkTopology: Adding a new node: /default-rack/127.0.0.1:50010

19/07/05 10:37:35 INFO balancer.Balancer: 0 over-utilized: []

19/07/05 10:37:35 INFO balancer.Balancer: 0 underutilized: []

The cluster is balanced. Exiting...

Jul 5, 2019 10:37:35 AM 0 0 B 0 B 0 B

Jul 5, 2019 10:37:35 AM Balancing took 2.826 seconds

1. Creating a new directory

**hdfs dfs -mkdir /abc**

1. To view the created directory

**vagrant@master:~$ hdfs dfs -ls abc**

Found 1 items

-rw-r--r-- 1 vagrant supergroup 66 2019-07-06 10:05 abc/Test.txt

1. To remove an already existing file use the command

**hdfs fs –rm - r abc/Test.txt**

Deleted abc/Test.txt

1. To see the content of the file

**vagrant@master:~$ hadoop fs -cat Test1.txt**

Hi all

Welcome to the session on HDFS

Hope you are enjoying it

1. To copy from the remote file of the local system to the hadoop distributed file system

Copy the file ABC in the folder abc/ABC to the test folder in hadoop with a file name ABC!

**vagrant@master:~$ hdfs dfs -copyFromLocal /home/vagrant/abc/ABC/ABC.txt /user/vagrant/test/ABC1.txt**

1. To see how much space is occupied in HDFS.

**vagrant@master:~$ hdfs dfs -du -s -s abc/ABC**

26 abc/ABC

1. To delete a specific file in a specific folder

**vagrant@master:~$ hadoop fs -rm -r /user/vagrant/test/ABC1.txt**

Deleted /user/vagrant/test/ABC1.txt

1. To empty a trash the following command is used

**vagrant@master:~$ hadoop fs –expunge**

1. To remove files within a directory

**vagrant@master:~$ hadoop fs -rm /user/vagrant/abc/ABC/\***

Deleted /user/vagrant/abc/ABC/ABC.txt