**HDFS Quotas:**

You can set two types of quotas in HDFS

a. Space Quota: The amount of space used by given directory

b. Name Quota: The number of file and directory names used.

Notes:

Quotas for space and names are independent of each other

File and directory creation fails if creation would cause the quota to be exceeded.

Block allocations fail if the quota would not allow a full block to be written.

Each replica counts against quota. For eg. if user is writing 3GB file with replication factor of 3, then 9GB will be consumed from his quota.

Largest quota is Long.Max\_Value

HDFS Quota Operations:

a. Set a Name quota:

ACL: Only admin can perform this operation.

Command: Hadoop admin can use following command to set name quota.

hadoop dfsadmin -setQuota number\_of\_files path

eg.

hadoop dfsadmin -setQuota 100 /grid/landing

Explanation: It sets hadoop quota to 100, which means user can create 100 files including directories under /grid/landing path.

b. Clear a Name quota:

ACL: Only admin can perform this operation.

Command: Hadoop admin can use following command to clear name quota.

hadoop dfsadmin -clearQuota path

eg.

hadoop dfsadmin -clearQuota /grid/landing

c. Set Space quota:

ACL: Only admin can perform this operation.

Command: To set space quota, hadoop admin can use following command,

hadoop dfsadmin -setSpaceQuota size path

eg.

hadoop dfsadmin -setSpaceQuota 15G /grid/landing

Explanation: It means user can write upto 5GB ( 5 \* 3 = 15) of data under /grid/landing path , assuming the replication factor of 3. Here user cannot write data less than block size. why? Because HDFS assumes an entire block will be filled, when its allocated. eg. say, if some path /projects/ingestion has quota of 50 MB, and if someone is writing a file of 10MB under this path, it'll fail, because of quota violation. Here HDFS thinks user is writing 384MB (128 \* 3) data, instead of (10 \* 3 = 30 MB).

d. **Clear Space quota:**

ACL: Only admin can perform this operation.

Command: To clear space quota, admin can use following command,

hadoop dfsadmin -clearSpaceQuota path

eg.

hadoop dfsadmin -clearSpaceQuota /grid/landing

e. Get quota allocation of Path:

ACL: Anyone can check quota allocation of path

Command: Hadoop admin can use following command to check quota allocation of a path,

hadoop fs -count -q path

eg.

hadoop fs -count -q /grid

Explanation: Above command will give output of following format,

hadoop fs -count -q /grid

9223372036854775807 9223372036854775333 none inf 141 333 655855032 /grid

where ,

column1 --> Namespace quota, which means total 9223372036854775807 files can be created

column2 --> Available Namespace quota, user can add 9223372036854775333 files .

column3 --> Space quota

column4 --> Available space quota

column5 --> Number of directories

column6 --> Number of Files

column7 --> Size of content available

column8 --> Path

**Using Quotas to control Amount of Data Written in HDFS**

# su – hdfs

# hdfs dfs –ls /

# hdfs dfs -mkdir /mydir

#hdfs dfsadmin –setQuota 2 /mydir

#hdfs dfs –count –q /mydir

#hdfs dfs -touchz /mydir/test1.txt

#hdfs dfs -touchz /test1/test2.txt

#hdfs dfs –ls /mydir

**Let’s Create a new user :-**

#hdfs dfs –mkdir /mydir2

#hdfs dfsadmin –setSpaceQuota 384m /mydir2

#hdfs dfs –count –q /mydir2

#hdfs dfs –put test1.txt /mydir2

#hdfs dfsadmin –clrQuota /mydir2

#hdfs dfsadmin –clrSpaceQuota /mydir2