**Recovering-Accidentally-Deleted-file-with-skip-trash-in-HDFS**

we are going to learn how to recover the accidentally deleted files (with skip-trash) command in hdfs.

Please open raw document for full article...,

---Follow this 9 step procedure for recovering accidental permanent deletion of file from HDFS----

please proceed with caution in production environments .

**Important:**

**\*\*\*\*\*please stop the namenode immidiatly after the file deletion other wise its hard to recover as namenode already send out block deletion**

**request to datanode , which will delete physical blocks upon request.\*\*\*\*\***

**\*\*\*\*\* always take backuo of "FSimage" and "edits" file before hand.\*\*\*\*\***

step-1 : lets create a samle file in HDFS

first switch to hdfs user or to the user which has hdfs write access.

eg: su - hdfs

create or put a sample file in hdfs using

hdfs dfs -put /tmp/file1 /temp

check weather the file created exists in hdfs or not using command

hdfs dfs -ls /temp/file1

step-2: Delete the file and also make sure that the file doesnt go to trash.

hdfs dfs -rmr -skipTrash /temp/file1

once its accidentally removed then if u want to recover it

step-3: stop the HDFS service via "Cloudera manager" or "Ambari"

step-4: open the Namenode Metadata folder--- if u dont know the folder u can find the path under the property (dfs.namenode.name.dir)

In my case : /data/nn/

Inside that directory there will be a folder called "current"( which holds current FSIMAGE) open that folder and look for "edits\_inprogress".

"edits\_inprogress"-- file is the most recent file where the latest transactions are pushed to.

step-5: Convert the binary edits file to xml format. to convert it use the following command

hdfs oev -i edits\_inprogress\_\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* -o edits\_inprogress\_\*\*\*\*\*\*\*\*\*\*\*\*.xml

step-6: open the file and look for the trasaction which recordered delete operation of the file "/temp/file1" , the transactions will look like below:::

<RECORD>

<OPCODE>OP\_DELETE</OPCODE>

<DATA>

<TXID>\*\*\*\*\*</TXID>

<LENGTH>0</LENGTH>

<PATH>/temp/file1</PATH>

<TIMESTAMP>\*\*\*\*\*\*\*\*\*\*\*\*</TIMESTAMP>

<RPC\_CLIENTID>\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*</RPC\_CLIENTID>

<RPC\_CALLID>1</RPC\_CALLID>

</DATA>

</RECORD>

Remove the complete entry and save the xml file .

step-7: now convert the xml file into binary format using,

go to the current edits or fsimage directory

cd /data/nn/current

and then execute the below command with appropriate file names

hdfs oev -i edits\_inprogess\_\*\*\*\*\*\*\*\*\*\*\*\*\*.xml -o edits\_inprogrss\_\*\*\*\*\*\*\*\*\*\*\*\*\* -p binary

step-8: if the transaction entry in the edits\_inprogress is the last one , then you could just try to startup the namenode, which brings you back the deleted file.

if the transaction entry is not the last one and in any other position then you must run recovery command as below

hadoop namenode -recover

step-9: now restart the HDFS service via "cloudera manager" or "ambari" and check for the deleted file.

this way u can recover the file deleted accidentally from hdfs using -skipTrash command.