**Importance of Namenode in Hadoop cluster**

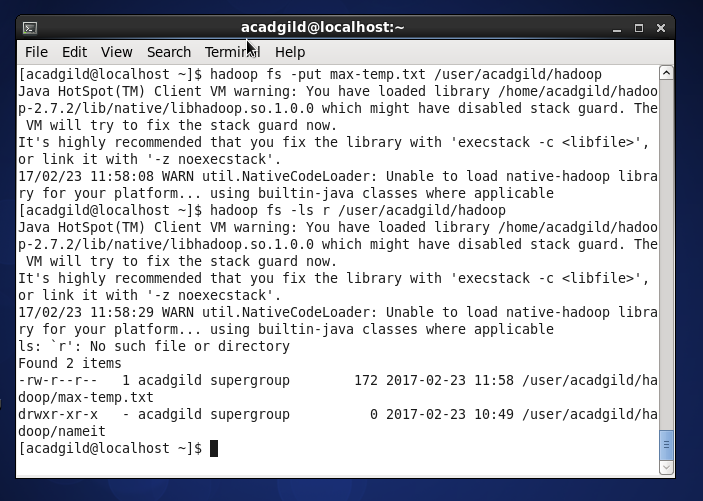
* NameNode is the centerpiece of HDFS.
* NameNode is also known as the Master
* NameNode only stores the metadata of HDFS – the directory tree of all files in the file system, and tracks the files across the cluster.
* NameNode does not store the actual data or the dataset. The data itself is actually stored in the DataNodes.
* NameNode knows the list of the blocks and its location for any given file in HDFS. With this information NameNode knows how to construct the file from blocks.
* NameNode is so critical to HDFS and when the NameNode is down, HDFS/Hadoop cluster is inaccessible and considered down.
* NameNode is a single point of failure in Hadoop cluster.
* NameNode is usually configured with a lot of memory (RAM). Because the block locations are help in main memory.

**Beginners commands for HDFS**

get - get a file present in hadoop system to local, it has checksum. hadoop fs -put /home/acadgild/max-temp.txt /user/acadgild/hadoop/max-temp.txt



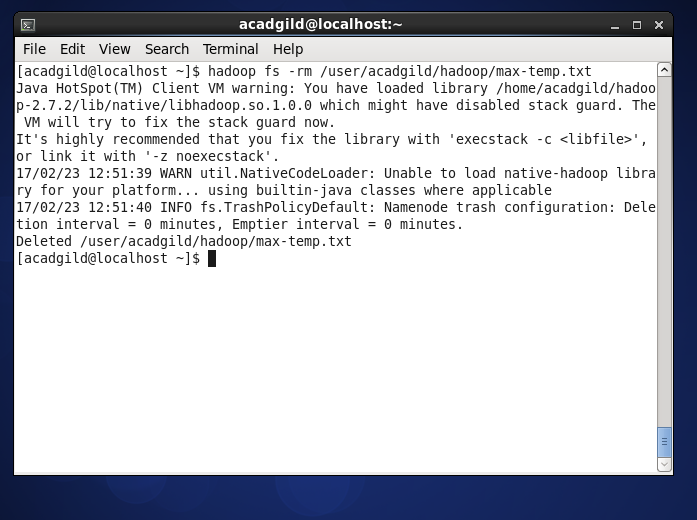
put- put file from hadoop to local, has checksum hadoop fs -put /user/acadgild/hadoop/max-temp.txt /home/acadgild



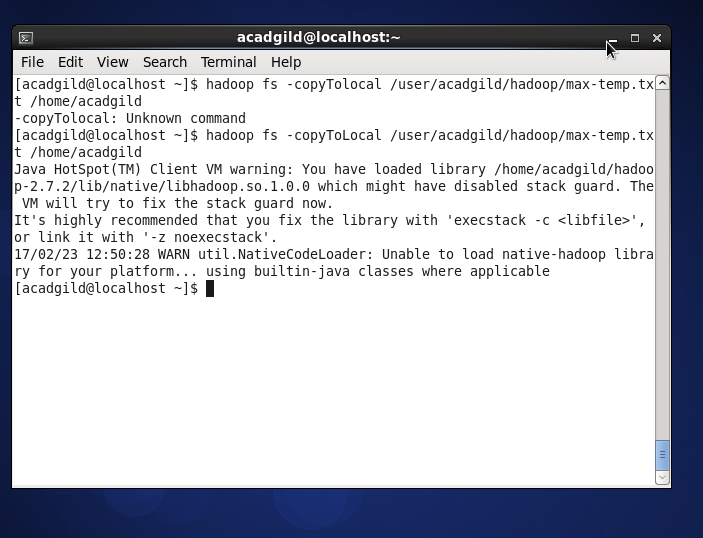
mkdir- make new directory hadoop fs -mkdir shrey



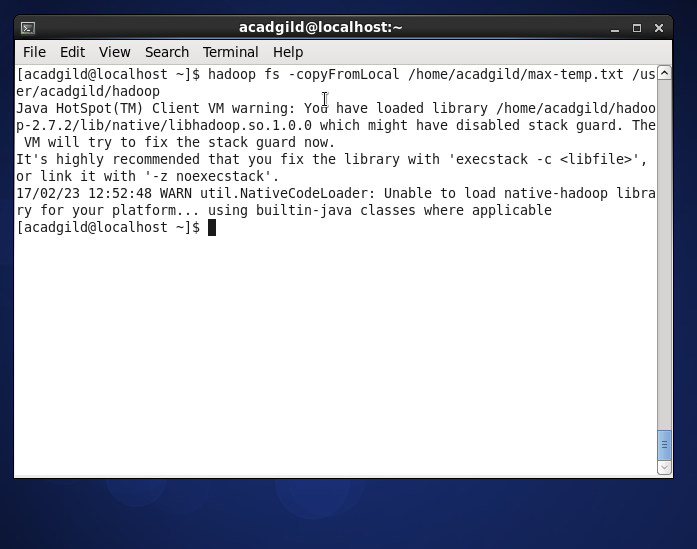
rm- remove file from directory hadoop fs -rm /user/acadgild/hadoop/max-temp.txt



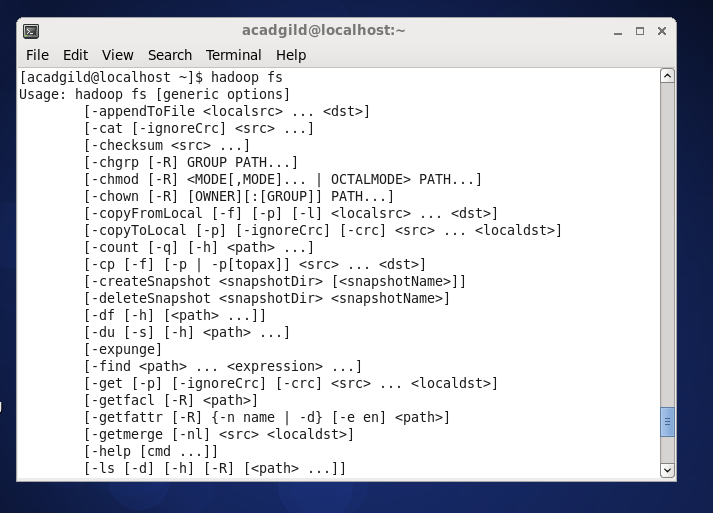
copyToLocal copyToLocal will copy file from hadoop fs to local hadoop fs -copyToLocal /user/acadgild/hadoop/max-temp.txt /home/acadgild



6.copyFromLocal-will copy file from local to hadoop fs hadoop fs -copyFromLocal /home/acadgild /user/acadgild/hadoop/max-temp.txt

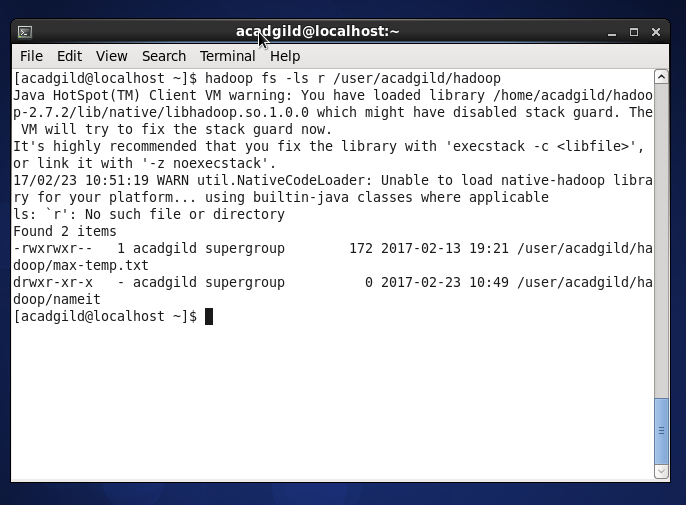


7.Option - hadoop fs will give every possible option of the command

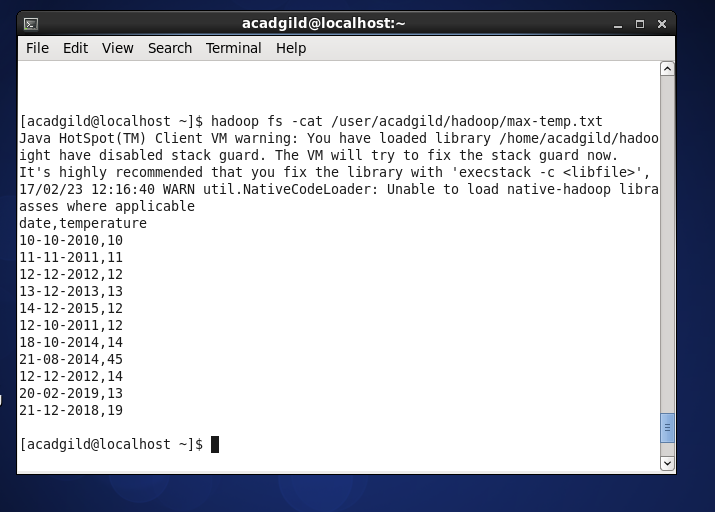


8.Help will give neccessary help when required hadoop fs -help

9.list list of all the files present hadoop fs -ls /user/acadgild/hadoop

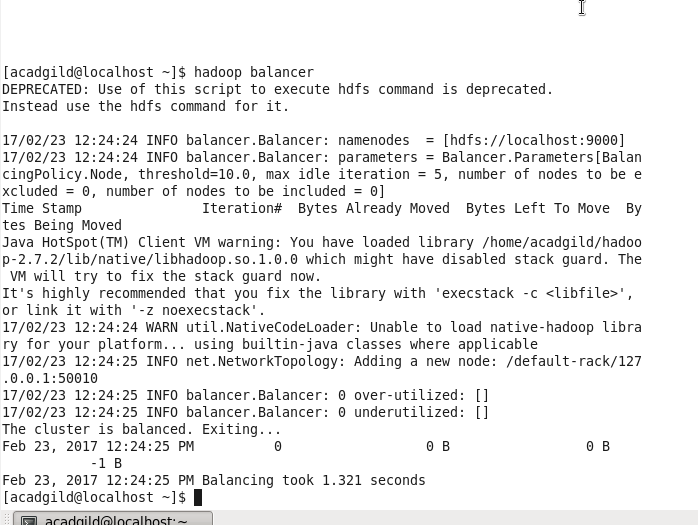


10.cat display the containt present in the file hadoop fs -cat /user/acadgild/hadoop/max-temp.txt



fsck- The command ‘fsck’ is used for checking the consistency of a file system hadoop fsck /user/acadgild/hadoop/max-temp.txt

12.Balancer- The command ‘balancer’ will check for work load on nodes in cluster and balance it. Syntax: hadoop balancer



13.Check directory space- The command will show the file size occupied by file inside cluster. Syntax: hadoop dfs -du -s -h /user/acadgild/hadoop/max-temp.txt

