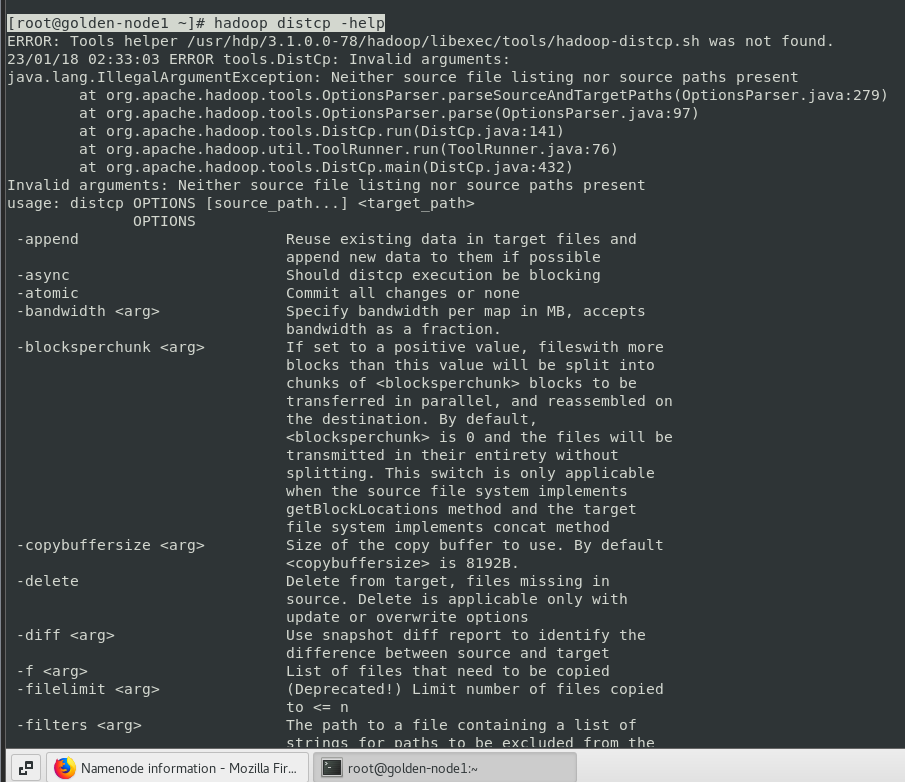
Lab 11: Using DistCp

Objective: To use the DistCp command and options to copy HDFS files and directories

[root@golden-node1 ~]# hadoop distcp –help

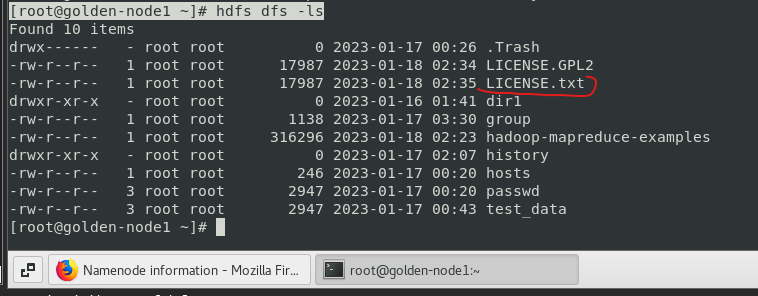


Configuring the DistCp Environment

* Configure a DistCp test environment on your cluster

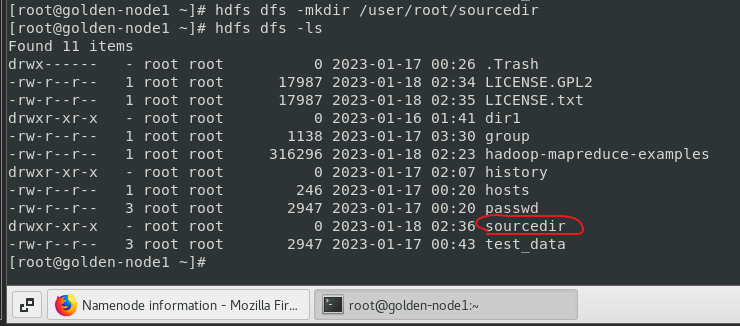
[root@golden-node1 ~]# hdfs dfs -put /usr/share/licenses/systemd-219/LICENSE.GPL2 /user/root/LICENSE.txt

[root@golden-node1 ~]# hdfs dfs –ls



[root@golden-node1 ~]# hdfs dfs -mkdir /user/root/sourcedir

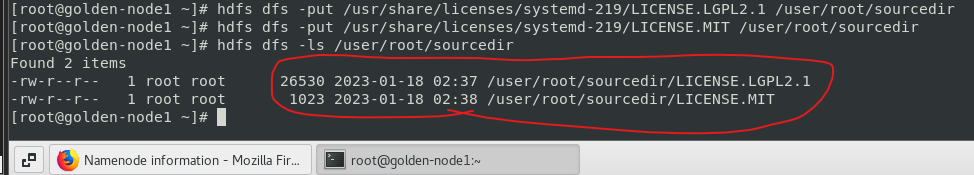
[root@golden-node1 ~]# hdfs dfs –ls



[root@golden-node1 ~]# hdfs dfs -put /usr/share/licenses/systemd-219/LICENSE.LGPL2.1 /user/root/sourcedir

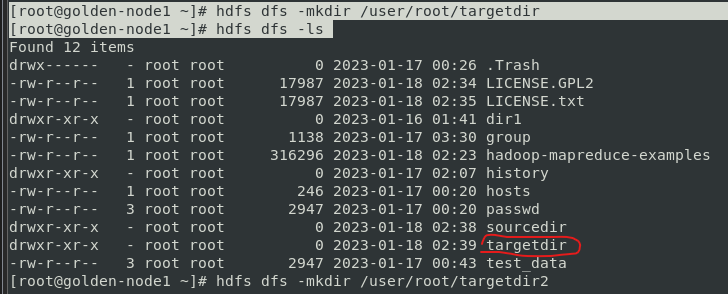
[root@golden-node1 ~]# hdfs dfs -put /usr/share/licenses/systemd-219/LICENSE.MIT /user/root/sourcedir

[root@golden-node1 ~]# hdfs dfs -ls /user/root/sourcedir



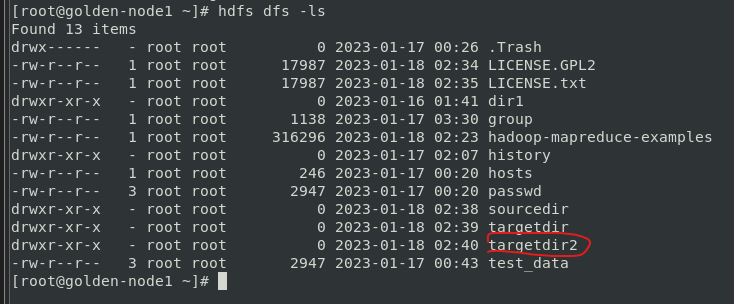
[root@golden-node1 ~]# hdfs dfs -mkdir /user/root/targetdir

[root@golden-node1 ~]# hdfs dfs –ls



[root@golden-node1 ~]# hdfs dfs -mkdir /user/root/targetdir2

[root@golden-node1 ~]# hdfs dfs -ls



Copying Files and Directories

* Copy files and directories to another HDFS directory using DistCp and DistCp options