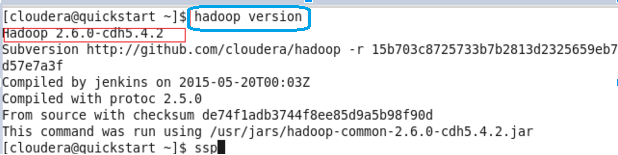
Hand on Lab: Understanding HDFS

Aim: The main aim of this lab is to understand the basic Hadoop commands, using which you can perform various operations on HDFS.

Note: Please type the commands manually instead of copy paste the commands for good practice.

Command to find the Version of Hadoop:

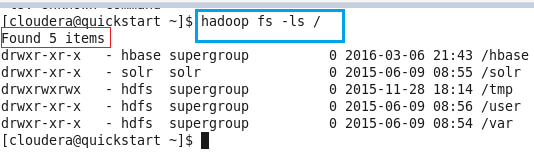
Command: hadoop version



LS command:

Displays List of Files and Directories in HDFS file Path

Command: hadoop fs -ls /

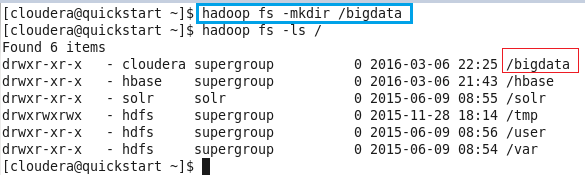


MKDIR command:

It creates the directory in HDFS

Syntax: hadoop fs -mkdir /directory\_name

Command: hadoop fs -mkdir /bigdata



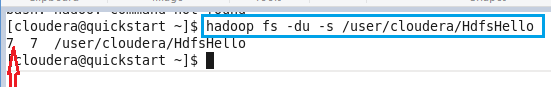
DU command:

Displays the summary of file lengths.

Syntax: hadoop fs -du -s /path/to/file\_in\_hdfs

Command: hadoop fs -du -s /user/cloudera/HdfsHello

Note: Here HdfsHello is a file that exists in HDFS in the directory cloudera



TOUCHZ command:

Create a file in HDFS with file size 0 bytes

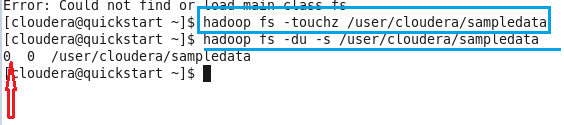
Syntax: hadoop fs -touchz /directory/filename

Command: hadoop fs -touchz /user/cloudera/sampledata

Note: Here we are trying to create a file named “sampledata” in the directory ‘cloudera’ of hdfs with file size 0 bytes.

Use the du command to check the file size

Command: hadoop fs -du -s /user/cloudera/sampledata



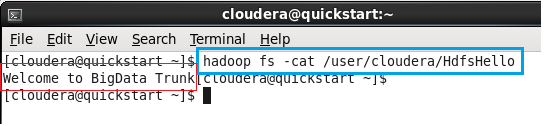
CAT Command:

Copies source paths to stdout.

Syntax: hadoop fs -cat /path/to/file\_in\_hdfs

Command: hadoop fs -cat /user/cloudera/HdfsHello

Note: Here HdfsHello is a file that exists in HDFS in the directory /user/cloudera



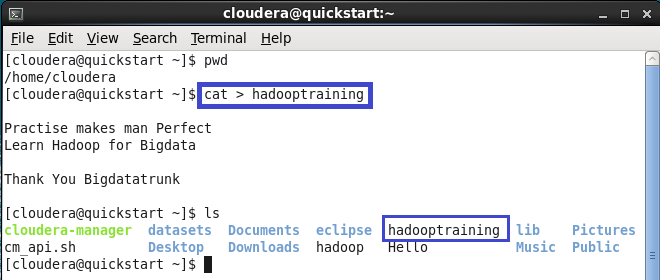
Cat > Command:

This command is used to create file in specified directory or in default home directory.

Syntax: cat >  <filename>  (or)  cat > <filepath with new filename>

Command: cat > hadooptraining

Note: After typing text in file press “ctrl + d” for saving and exiting from file.



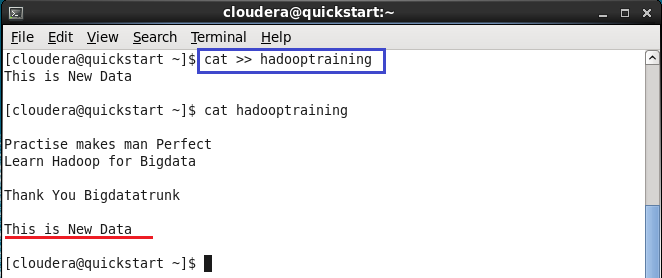
Cat >> Command:

This command is used to append the text from existing file text.

Syntax: cat >> <filename>

Command: cat >> hadooptrianing

Note: After typing text in file press “ctrl + d” for saving and exiting from file.



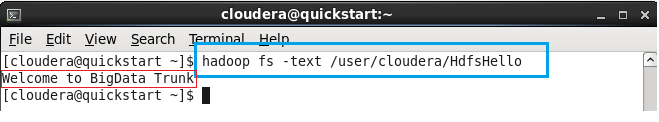
TEXT command:

Takes a source file and outputs the file in text format.( Same as Cat command )

Syntax: hadoop fs -text /path/to/file\_in\_hdfs

Command: hadoop fs -text /user/cloudera/HdfsHello

Note: Here HdfsHello is a file that exists in HDFS in the directory /user/cloudera



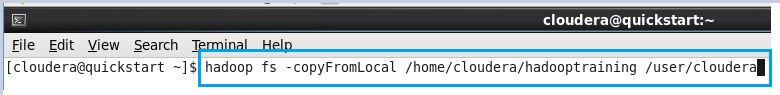
copyFromLocal command:

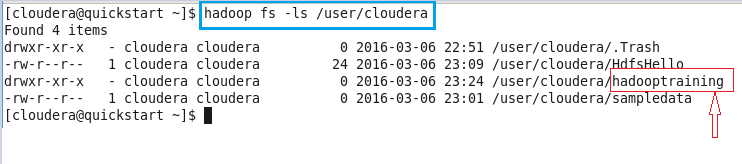
Copy the file from Local file system to HDFS.

Syntax: hadoop fs -copyFromLocal <localsrc> URI

Command: hadoop fs -copyFromLocal /home/cloudera/hadooptraining /user/cloudera

Note: Here hadooptraining is the file present in the local directory - /home/cloudera





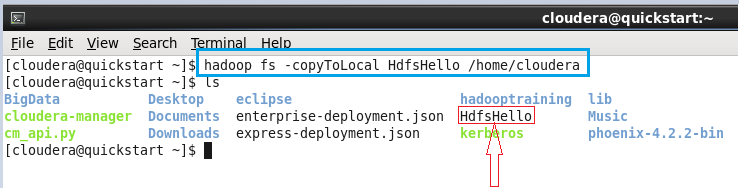
copyToLocal command:

Copy the file from HDFS to Local File System.

Syntax: hadoop fs -copyToLocal URI <localdst>

Command: hadoop fs -copyToLocal HdfsHello /home/cloudera

Note: Here HdfsHello is a file present in default hdfs directory ie “/user/cloudera” of HDFS



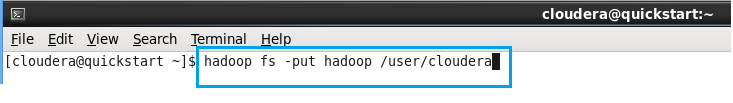
put command:

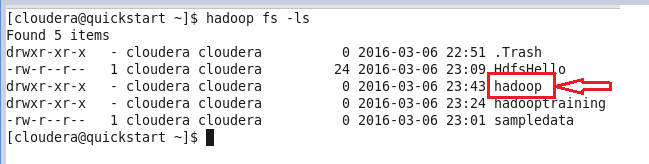
Copy single source, or multiple srcs from local file system to the destination file system.

Syntax: hadoop fs -put <localsrc> ... <dst>

Command: hadoop fs -put hadoop /user/cloudera

Note: copyFromLocal is similar to put command, except that the source is restricted to a local file reference





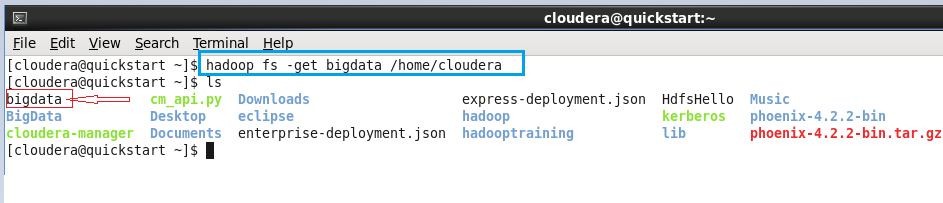
GET command:

Copy files from hdfs to the local file system.

Syntax: hadoop fs -get [-ignorecrc] [-crc] <src> <localdst>

Command: hadoop fs -get bigdata /home/cloudera

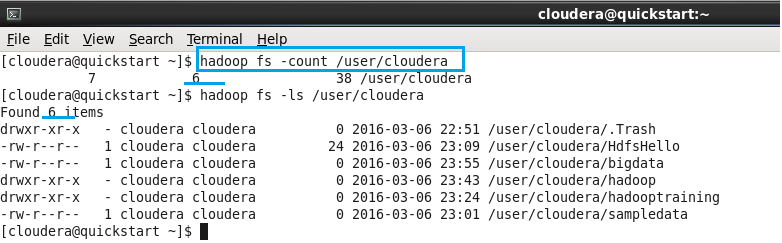
Note: copyToLocal is similar to get command, except that the destination is restricted to a local file reference.



COUNT command:

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

Command: hadoop fs -count /user/cloudera

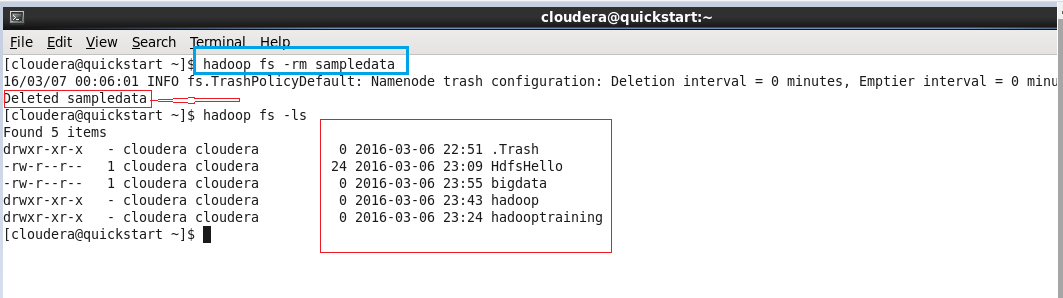


RM command:

Remove the file from HDFS.

Syntax: hadoop fs -rm /path/to/file\_in\_hdfs

Command: hadoop fs -rm sampledata

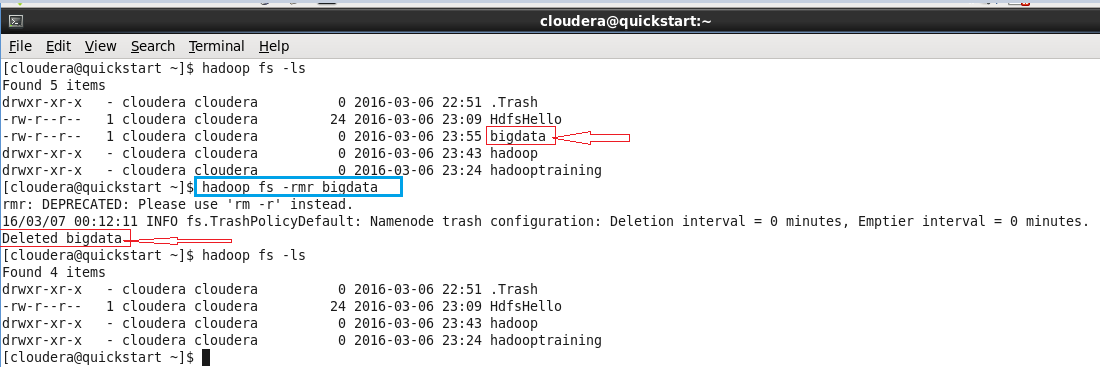


RMR command:

Remove the directory to HDFS

Syntax: hadoop fs -rmr /path/to/directory\_in\_hdfs

Command: hadoop fs -rmr bigdata



|  |  |  |
| --- | --- | --- |
| Commands | How To Use | Explanation |
| hadoop version | hadoop version | Displays the hadoop version information |
| hadoop fs -ls / | hadoop fs -ls / | Displays List of Files and Directories in HDFS file Path |
| hadoop fs -mkdir | hadoop fs -mkdir /directory\_name  Ex: hadoop fs -mkdir /bigdata | It creates the directory in HDFS |
| hadoop fs -du -s | hadoop fs -du -s /path/to/file\_in\_hdfs  Ex:  hadoop fs -du -s /user/cloudera/HdfsHello | Displays the summary of file lengths. |
| hadoop fs -touchz | hadoop fs -touchz /directory/filename  Ex:  hadoop fs -touchz /user/cloudera/sampledata | Create a file in HDFS with file size 0 bytes |
| hadoop fs -cat | hadoop fs -cat /path/to/file\_in\_hdfs  Ex:  hadoop fs -cat /user/cloudera/HdfsHello | Copies source paths to stdout. |
| hadoop fs -cat > | cat >  <filename>  (or)  cat > <filepath with new filename>  Ex: cat > hadooptraining | This command is used to create file in specified directory or in default home directory. |
| hadoop fs - cat >> | hadoop fs - cat >> <filename>  Ex: cat >> hadooptrianing | To append the text from existing file text. |
| hadoop fs -text | hadoop fs -text /path/to/file\_in\_hdfs  Ex:  hadoop fs -text /user/cloudera/HdfsHello | Takes a source file and outputs the file in text format. |
| hadoop fs -copyFromLocal | hadoop fs -copyFromLocal <localsrc> URI  Ex: hadoop fs -copyFromLocal /home/cloudera/hadooptraining /user/cloudera | Copy the file from Local file system to HDFS. |
| hadoop fs -copyToLocal | hadoop fs -copyToLocal URI <localdst>  Ex: hadoop fs -copyToLocal HdfsHello /home/cloudera | Copy the file from HDFS to Local File System |
| hadoop fs -put | hadoop fs -put <localsrc> ... <dst>  Ex: hadoop fs -put hadoop /user/cloudera | Copy single source, or multiple srcs from local file system to the destination file system |
| hadoop fs -get | hadoop fs -get [-ignorecrc] [-crc] <src> <localdst>  Ex: hadoop fs -get bigdata /home/cloudera | Copy files from hdfs to the local file system |
| hadoop fs -count | hadoop fs -count /filepath  Ex: hadoop fs count /user/cloudera | Count the number of directories, files and bytes under the paths that match the specified file |
| hadoop fs -rm | hadoop fs -rm /path/to/file\_in\_hdfs  Ex: hadoop fs -rm sampledata | Remove the file from HDFS |
| hadoop fs -rmr | hadoop fs -rmr /path/to/directory\_in\_hdfs  Ex: hadoop fs -rmr bigdata | Remove the directory from HDFS |