

HDFS基本操作

hdfs是文件系统，所以和linux管理文件方式区别不大。例如linux中文件管理使用 `ls /` 查看根目录内容，在hdfs中用 `hdfs dfs -ls /` 查看，区别就在于前面加了 `hdfs dfs -`

上传文件命令 put

```
hdfs dfs -put <localsrc> ... <dst[hdfs]>
```

Copy single src, or multiple srcs from local file system to the destination file system. Also reads input from stdin and writes to destination file system.

```
hdfs dfs -put localfile /user/hadoop/hadoopfile
hdfs dfs -put localfile1 localfile2 /user/hadoop/hadoopdir
hdfs dfs -put localfile hdfs://nn.example.com/hadoop/hadoopfile
hdfs dfs -put - hdfs://nn.example.com/hadoop/hadoopfile Reads the input from stdin.
eg:hdfs dfs -put test2.txt /usr/input/
```

查看目录 ls

Usage: `hdfs dfs -ls [-R]`

```
eg:hdfs dfs -ls /user/hadoop/file1
```

查看文本 text

3、cat

Usage: `hdfs dfs -text URI [URI ...]`

eg

```
hdfs dfs -text /usr/input/test2.txt
hdfs dfs -cat hdfs://nn1.example.com/file1 hdfs://nn2.example.com/file2
hdfs dfs -cat file:///file3 /user/hadoop/file4
```

删除文件 rm

Usage: `hdfs dfs -rm [-f] [-r] [-R] [-skipTrash] URI [URI ...]`

Options:

- f选项不会显示信息或者在文件不存在导致报错的情况下修改退出状态。
- R选项递归的删除子文件和目录里的内容。
- r选项等同于-R。
- skipTrash选项会略过回收站删除指定文件。当删除超过quota配额的目录文件时这个选项会很有用。

```
eg:hdfs dfs -rm /usr/input/test2.txt
hdfs dfs -rm hdfs://nn.example.com/file /user/hadoop/emptydir
```

下载文件 get

Usage: `hdfs dfs -get [-ignorecrc] [-crc]`

拷贝文件到本地文件系统。如果在CRC校验失败的文件可以加-ignorecrc选项。使用-crc选项，文件和CRC都可以被拷贝。

```
eg: hdfs dfs -get /user/hadoop/file localfile  
hdfs dfs -get hdfs://nn.example.com/user/hadoop/file localfile
```

创建目录 mkdir

Usage: hdfs dfs -mkdir [-p]

```
eg: hdfs dfs -mkdir /user/hadoop/dir1 /user/hadoop/dir2  
hdfs dfs -mkdir hdfs://nn1.example.com/user/hadoop/dir >  
hdfs://nn2.example.com/user/hadoop/dir
```

创建文件 touchz

Usage: hdfs dfs -touchz URI [URI ...]

创建一个空文件。

```
eg:hdfs dfs -touchz pathname  
hdfs dfs -touchz /usr/input/a.txt /usr/input/b.txt
```