命令基本格式:

hadoop fs -**cmd** < args >

1.ls

hadoop fs -ls /

列出hdfs文件系统根目录下的目录和文件

hadoop fs -ls -R /

列出hdfs文件系统所有的目录和文件

2.put

hadoop fs -**put** < **local** file > < hdfs file >

hdfs file的父目录一定要存在，否则命令不会执行

hadoop fs -**put** < **local** file **or** dir >...< hdfs dir >

hdfs dir 一定要存在，否则命令不会执行

hadoop fs -**put** - < hdsf file>

从键盘读取输入到hdfs file中，按Ctrl+D结束输入，hdfs file不能存在，否则命令不会执行

2.1.moveFromLocal

hadoop fs -moveFromLocal < local src > ... < hdfs dst >

与put相类似，命令执行后源文件 local src 被删除，也可以从从键盘读取输入到hdfs file中

2.2.copyFromLocal

hadoop fs -copyFromLocal < local src > ... < hdfs dst >

与put相类似，也可以从从键盘读取输入到hdfs file中

3.get

hadoop fs -**get** < hdfs file > < **local** file **or** dir>

local file不能和 hdfs file名字不能相同，否则会提示文件已存在，没有重名的文件会复制到本地

hadoop fs -**get** < hdfs file **or** dir > ... < **local** dir >

拷贝多个文件或目录到本地时，本地要为文件夹路径  
**注意：**如果用户不是root， local 路径要为用户文件夹下的路径，否则会出现权限问题，

3.1.moveToLocal

当前版本中还未实现此命令

3.2.copyToLocal

hadoop fs -copyToLocal < local src > ... < hdfs dst >

与get相类似

4.rm

hadoop fs -**rm** < hdfs **file** > ...

hadoop fs -**rm** -r < hdfs **dir**>...

每次可以删除多个文件或目录

5.mkdir

hadoop fs -mkdir < hdfs path>

只能一级一级的建目录，父目录不存在的话使用这个命令会报错

hadoop fs -mkdir -p < hdfs path>

所创建的目录如果父目录不存在就创建该父目录

6.getmerge

hadoop fs -getmerge < hdfs **dir** > < **local** **file** >

将hdfs指定目录下所有文件排序后合并到local指定的**文件**中，文件不存在时会自动创建，文件存在时会覆盖里面的内容

hadoop fs -getmerge -**nl** < hdfs **dir** > < **local** **file** >

加上nl后，合并到local file中的hdfs文件之间会空出一行

7.cp

hadoop fs -**cp** < hdfs **file** > < hdfs **file** >

目标文件不能存在，否则命令不能执行，相当于给文件重命名并保存，源文件还存在  
hadoop fs -cp < hdfs file or dir >... < hdfs dir >

目标文件夹要存在，否则命令不能执行

8.mv

hadoop fs -mv < hdfs file > < hdfs file >

目标文件不能存在，否则命令不能执行，相当于给文件重命名并保存，源文件不存在

hadoop fs -mv < hdfs file or dir >... < hdfs dir >

源路径有多个时，目标路径必须为目录，且必须存在。  
**注意：**跨文件系统的移动（local到hdfs或者反过来）都是不允许的

9.count

hadoop fs -count < hdfs path >

统计hdfs对应路径下的目录个数，文件个数，文件总计大小  
显示为目录个数，文件个数，文件总计大小，输入路径

10.du

hadoop fs -du < hdsf path>

显示hdfs对应路径下每个文件夹和文件的大小

hadoop fs -du -s < hdsf path>

显示hdfs对应路径下所有文件和的大小

hadoop fs -du - h < hdsf path>

显示hdfs对应路径下每个文件夹和文件的大小,文件的大小用方便阅读的形式表示，例如用64M代替67108864

11.text

hadoop fs -text < hdsf file>

将文本文件或某些格式的非文本文件通过文本格式输出

12.setrep

hadoop fs -setrep -R 3 < hdfs path >

改变一个文件在hdfs中的副本个数，上述命令中数字3为所设置的副本个数，-R选项可以对一个人目录下的所有目录+文件递归执行改变副本个数的操作

13.stat

hdoop fs -stat [format] < hdfs path >

返回对应路径的状态信息  
[format]可选参数有：%b（文件大小），%o（Block大小），%n（文件名），%r（副本个数），%y（最后一次修改日期和时间）  
可以这样书写*hadoop fs -stat %b%o%n < hdfs path >*，不过不建议，这样每个字符输出的结果不是太容易分清楚

14.tail

hadoop fs -tail < hdfs file >

在标准输出中显示文件末尾的1KB数据

15.archive

hadoop archive -archiveName name.har -p < hdfs parent dir > < src >\* < hdfs dst >

命令中参数name：压缩文件名，自己任意取；< hdfs parent dir > ：压缩文件所在的父目录；< src >*：要压缩的文件名；< hdfs dst >：压缩文件存放路径  
\*示例：hadoop archive -archiveName hadoop.har -p /user 1.txt 2.txt /des*  
示例中将hdfs中/user目录下的文件1.txt，2.txt压缩成一个名叫hadoop.har的文件存放在hdfs中/des目录下，如果1.txt，2.txt不写就是将/user目录下所有的目录和文件压缩成一个名叫hadoop.har的文件存放在hdfs中/des目录下  
显示har的内容可以用如下命令：

hadoop fs -**ls** /**des**/hadoop.jar

显示har压缩的是那些文件可以用如下命令

hadoop fs -ls -R har:///des/hadoop.har

**注意：**har文件不能进行二次压缩。如果想给.har加文件，只能找到原来的文件，重新创建一个。har文件中原来文件的数据并没有变化，har文件真正的作用是减少NameNode和DataNode过多的空间浪费。

16.balancer

hdfs **balancer**

如果管理员发现某些DataNode保存数据过多，某些DataNode保存数据相对较少，可以使用上述命令手动启动内部的均衡过程

17.dfsadmin

hdfs dfsadmin -help

管理员可以通过dfsadmin管理HDFS，用法可以通过上述命令查看  
hdfs dfsadmin -report

显示文件系统的基本数据

hdfs dfsadmin -safemode < enter | **leave** | **get** | **wait** >

enter：进入安全模式；leave：离开安全模式；get：获知是否开启安全模式；  
wait：等待离开安全模式

18.distcp

用来在两个HDFS之间拷贝数据