

Hadoop Shell命令

目录

1 FS Shell	3
1.1 cat	3
1.2 chgrp	3
1.3 chmod	3
1.4 chown	4
1.5 copyFromLocal.....	4
1.6 copyToLocal.....	4
1.7 cp	4
1.8 du.....	4
1.9 dus	5
1.10 expunge	5
1.11 get	5
1.12 getmerge	5
1.13 ls	6
1.14 lsr.....	6
1.15 mkdir	6
1.16 movefromLocal	7
1.17 mv	7
1.18 put	7
1.19 rm	7
1.20 rmr	8

1.21	setrep	8
1.22	stat	8
1.23	tail	9
1.24	test	9
1.25	text	9
1.26	touchz	9

1. FS Shell

调用文件系统(FS)Shell命令应使用 `bin/hadoop fs <args>`的形式。所有的FS shell命令使用URI路径作为参数。URI格式是`scheme://authority/path`。对HDFS文件系统, `scheme`是`hdfs`, 对本地文件系统, `scheme`是`file`。其中`scheme`和`authority`参数都是可选的, 如果未加指定, 就会使用配置中指定的默认`scheme`。一个HDFS文件或目录比如`/parent/child`可以表示成`hdfs://namenode:namenodeport/parent/child`, 或者更简单的`/parent/child` (假设你配置文件中的默认值是`namenode:namenodeport`)。大多数FS Shell命令的行为和对应的Unix Shell命令类似, 不同之处会在下面介绍各命令使用详情时指出。出错信息会输出到`stderr`, 其他信息输出到`stdout`。

1.1. cat

使用方法: `hadoop fs -cat URI [URI ...]`

将路径指定文件的内容输出到`stdout`。

示例:

- `hadoop fs -cat hdfs://host1:port1/file1 hdfs://host2:port2/file2`
- `hadoop fs -cat file:///file3 /user/hadoop/file4`

返回值:

成功返回0, 失败返回-1。

1.2. chgrp

使用方法: `hadoop fs -chgrp [-R] GROUP URI [URI ...]` Change group association of files. With `-R`, make the change recursively through the directory structure. The user must be the owner of files, or else a super-user. Additional information is in the [Permissions User Guide](#). -->

改变文件所属的组。使用`-R`将使改变在目录结构下递归进行。命令的使用者必须是文件的所有者或者超级用户。更多的信息请参见[HDFS权限用户指南](#)。

1.3. chmod

使用方法: `hadoop fs -chmod [-R] <MODE[,MODE]... | OCTALMODE> URI [URI ...]`

改变文件的权限。使用-R将使改变在目录结构下递归进行。命令的使用者必须是文件的所有者或者超级用户。更多的信息请参见[HDFS权限用户指南](#)。

1.4. chown

使用方法: `hadoop fs -chown [-R] [OWNER][:[GROUP]] URI [URI]`

改变文件的拥有者。使用-R将使改变在目录结构下递归进行。命令的使用者必须是超级用户。更多的信息请参见[HDFS权限用户指南](#)。

1.5. copyFromLocal

使用方法: `hadoop fs -copyFromLocal <localsrc> URI`

除了限定源路径是一个本地文件外，和[put](#)命令相似。

1.6. copyToLocal

使用方法: `hadoop fs -copyToLocal [-ignorecrc] [-crc] URI <localdst>`

除了限定目标路径是一个本地文件外，和[get](#)命令类似。

1.7. cp

使用方法: `hadoop fs -cp URI [URI ...] <dest>`

将文件从源路径复制到目标路径。这个命令允许有多个源路径，此时目标路径必须是一个目录。

示例:

- `hadoop fs -cp /user/hadoop/file1 /user/hadoop/file2`
- `hadoop fs -cp /user/hadoop/file1 /user/hadoop/file2 /user/hadoop/dir`

返回值:

成功返回0，失败返回-1。

1.8. du

使用方法: `hadoop fs -du URI [URI ...]`

显示目录中所有文件的大小，或者当只指定一个文件时，显示此文件的大小。

示例：

```
hadoop fs -du /user/hadoop/dir1 /user/hadoop/file1
```

```
hdfs://host:port/user/hadoop/dir1
```

返回值：

成功返回0，失败返回-1。

1.9. dus

使用方法：hadoop fs -dus <args>

显示文件的大小。

1.10. expunge

使用方法：hadoop fs -expunge

清空回收站。请参考[HDFS设计](#)文档以获取更多关于回收站特性的信息。

1.11. get

使用方法：hadoop fs -get [-ignorecrc] [-crc] <src> <localdst>

复制文件到本地文件系统。可用-ignorecrc选项复制CRC校验失败的文件。使用-crc选项复制文件以及CRC信息。

示例：

- `hadoop fs -get /user/hadoop/file localfile`
- `hadoop fs -get hdfs://host:port/user/hadoop/file localfile`

返回值：

成功返回0，失败返回-1。

1.12. getmerge

使用方法：hadoop fs -getmerge <src> <localdst> [addnl]

接受一个源目录和一个目标文件作为输入，并且将源目录中所有的文件连接成本地目

标文件。addnl是可选的，用于指定在每个文件结尾添加一个换行符。

1.13. ls

使用方法: `hadoop fs -ls <args>`

如果是文件，则按照如下格式返回文件信息：

文件名 <副本数> 文件大小 修改日期 修改时间 权限 用户ID 组ID

如果是目录，则返回它直接子文件的一个列表，就像在Unix中一样。目录返回列表的信息如下：

目录名 <dir> 修改日期 修改时间 权限 用户ID 组ID

示例：

```
hadoop fs -ls /user/hadoop/file1 /user/hadoop/file2
```

```
hdfs://host:port/user/hadoop/dir1 /nonexistentfile
```

返回值：

成功返回0，失败返回-1。

1.14. lsr

使用方法: `hadoop fs -lsr <args>`

ls命令的递归版本。类似于Unix中的ls -R。

1.15. mkdir

使用方法: `hadoop fs -mkdir <paths>`

接受路径指定的uri作为参数，创建这些目录。其行为类似于Unix的mkdir -p，它会创建路径中的各级父目录。

示例：

- `hadoop fs -mkdir /user/hadoop/dir1 /user/hadoop/dir2`
- `hadoop fs -mkdir hdfs://host1:port1/user/hadoop/dir`
`hdfs://host2:port2/user/hadoop/dir`

返回值：

成功返回0，失败返回-1。

1.16. movefromLocal

使用方法: `dfs -moveFromLocal <src> <dst>`

输出一个” not implemented “信息。

1.17. mv

使用方法: `hadoop fs -mv URI [URI ...] <dest>`

将文件从源路径移动到目标路径。这个命令允许有多个源路径，此时目标路径必须是一个目录。不允许在不同的文件系统间移动文件。

示例:

- `hadoop fs -mv /user/hadoop/file1 /user/hadoop/file2`
- `hadoop fs -mv hdfs://host:port/file1 hdfs://host:port/file2`
`hdfs://host:port/file3 hdfs://host:port/dir1`

返回值:

成功返回0，失败返回-1。

1.18. put

使用方法: `hadoop fs -put <localsrc> ... <dst>`

从本地文件系统中复制单个或多个源路径到目标文件系统。也支持从标准输入中读取输入写入目标文件系统。

- `hadoop fs -put localfile /user/hadoop/hadoopfile`
- `hadoop fs -put localfile1 localfile2 /user/hadoop/hadoopdir`
- `hadoop fs -put localfile hdfs://host:port/hadoop/hadoopfile`
- `hadoop fs -put - hdfs://host:port/hadoop/hadoopfile`
从标准输入中读取输入。

返回值:

成功返回0，失败返回-1。

1.19. rm

使用方法: `hadoop fs -rm URI [URI ...]`

删除指定的文件。只删除非空目录和文件。请参考`rmr`命令了解递归删除。

示例:

- `hadoop fs -rm hdfs://host:port/file /user/hadoop/emptydir`

返回值:

成功返回0, 失败返回-1。

1.20. `rmr`

使用方法: `hadoop fs -rmr URI [URI ...]`

`delete`的递归版本。

示例:

- `hadoop fs -rmr /user/hadoop/dir`
- `hadoop fs -rmr hdfs://host:port/user/hadoop/dir`

返回值:

成功返回0, 失败返回-1。

1.21. `setrep`

使用方法: `hadoop fs -setrep [-R] <path>`

改变一个文件的副本系数。`-R`选项用于递归改变目录下所有文件的副本系数。

示例:

- `hadoop fs -setrep -w 3 -R /user/hadoop/dir1`

返回值:

成功返回0, 失败返回-1。

1.22. `stat`

使用方法: `hadoop fs -stat URI [URI ...]`

返回指定路径的统计信息。

示例:

- `hadoop fs -stat path`

返回值:

成功返回0, 失败返回-1。

1.23. tail

使用方法: `hadoop fs -tail [-f] URI`

将文件尾部1K字节的内容输出到stdout。支持-f选项, 行为和Unix中一致。

示例:

- `hadoop fs -tail pathname`

返回值:

成功返回0, 失败返回-1。

1.24. test

使用方法: `hadoop fs -test -[ezd] URI`

选项:

- e 检查文件是否存在。如果存在则返回0。
- z 检查文件是否是0字节。如果是则返回0。
- d 如果路径是个目录, 则返回1, 否则返回0。

示例:

- `hadoop fs -test -e filename`

1.25. text

使用方法: `hadoop fs -text <src>`

将源文件输出为文本格式。允许的格式是zip和TextRecordInputStream。

1.26. touchz

使用方法: `hadoop fs -touchz URI [URI ...]`

创建一个0字节的空文件。

示例:

- `hadoop -touchz pathname`

返回值:

成功返回0, 失败返回-1。