

Hadoop 分布式文件系统 HDFS

基本命令

侯在钱

目 录

| | |
|--------------------------------------|---|
| 1. 查看目录列表信息 | 2 |
| 2. 递归列出目录列表 | 2 |
| 3. 创建文件夹 | 2 |
| 4. 上传文件 | 2 |
| 5. 下载文件 | 2 |
| 6. 查看文件内容 | 3 |
| 7. 删除文件/空白文件夹 | 3 |
| 8. 递归删除目录及目录下的所有文件..... | 3 |
| 9. 修改文件所属组 | 3 |
| 10. 修改文件权限 | 3 |
| 11. 修改文件所有者，文件所属组，其他用户的读、写、执行权限..... | 3 |
| 12. 复制文件到 HDFS | 3 |
| 13. 复制文件到本地 | 3 |
| 14. 统计文件及文件夹数目..... | 4 |
| 15. HDFS 文件系统间的文件复制 | 4 |
| 16. 统计目录下的文件及大小..... | 4 |
| 17. 汇总目录下的文件总大小..... | 4 |
| 18. 清空回收站 | 4 |
| 19. 查看 ACL（访问权限组拥有者） | 4 |
| 20. 合并下载文件到本地 | 4 |
| 21. 剪切文件到 HDFS | 4 |
| 22. 剪切文件到本地 | 4 |
| 23. 移动 HDFS 文件 | 5 |
| 24. 修改副本数 | 5 |
| 25. 显示文件统计信息 | 5 |
| 26. 查看文件尾部内容 | 5 |

Hadoop 1.x 使用 “hadoop fs ” 格式命令操作 HDFS，Hadoop 2.x 改用 “hdfs dfs” 格式命令操作 HDFS。其他参数一样。

1. 查看目录列表信息

查看某个目录下有哪些文件及子目录

hdfs dfs -ls <args>

例如查看根目录下有哪些文件及子目录：

hdfs dfs -ls /

2. 递归列出目录列表

hdfs dfs -lsr <args>

例如查看根目录下的文件：

hdfs dfs -lsr /

3. 创建文件夹

hdfs dfs -mkdir [-p] <paths>

例如创建目录/test：

hdfs dfs -mkdir /test

例如创建目录/test/hzq：

hdfs dfs -mkdir /test/hzq

4. 上传文件

从本地上传文件到 HDFS 文件系统中。

hdfs dfs -put <localsrc> ... <dst>

例如，把本地当前目录下的 **test.txt** 文件上传到 Hadoop 文件系统中的/test/目录下：

./hdfs dfs -put test.txt /test

5. 下载文件

从 HDFS 文件系统中下载文件到本地。

命令语法：

hdfs dfs -get [-ignorecrc] [-crc] <src> <localdst>

例如，把 Hadoop 文件系统中的/test/test.txt 文件下载到本地/tmp/目录下：

./hdfs dfs -get /test/test.txt /tmp/

6. 查看文件内容

```
hdfs dfs -cat URI [URI ...]
```

例如查看 HDFS 中的/test/hello.txt 这个文件：

```
hdfs dfs -cat /test/hello.txt
```

7. 删除文件/空白文件夹

删除 HDFS 中文件。

```
hdfs dfs -rm [-skipTrash] URI [URI ...]
```

例如，删除 HDFS 中/test/test.txt 文件：

```
hdfs dfs -rm /test/test.txt
```

8. 递归删除目录及目录下的所有文件

```
hdfs dfs -rmr [-skipTrash] URI [URI ...]
```

例如，删除 HDFS 中/test/目录：

```
hdfs dfs -rmr /test
```

9. 修改文件所属组

```
hdfs dfs -chgrp [-R] GROUP URI [URI ...]
```

10. 修改文件权限

```
hdfs dfs -chmod [-R] <MODE[,MODE]... | OCTALMODE> URI  
[URI ...]
```

11. 修改文件所有者，文件所属组，其他用户的读、写、执行权限

```
hdfs dfs -chown [-R] [OWNER][:[GROUP]] URI [URI ]
```

12. 复制文件到 HDFS

```
hdfs dfs -copyFromLocal <localsrc> URI
```

13. 复制文件到本地

```
hdfs dfs -copyToLocal [-ignorecrc] [-crc] URI <localdst>
```

14. 统计文件及文件夹数目

```
hdfs dfs -count [-q] <paths>
```

15. HDFS 文件系统间的文件复制

```
hdfs dfs -cp [-f] URI [URI ...] <dest>
```

16. 统计目录下的文件及大小

```
hdfs dfs -du [-s] [-h] URI [URI ...]
```

17. 汇总目录下的文件总大小

```
hdfs dfs -dus <args>
```

18. 清空回收站

文件被删除时，它首先会移到临时目录.Trash/中，当超过延迟时间之后，文件才会被永久删除

```
hdfs dfs -expunge
```

19. 查看 ACL（访问权限组拥有者）

```
hdfs dfs -getfacl [-R] <path>
```

20. 合并下载文件到本地

```
hdfs dfs -getmerge <src> <localdst> [addnl]
```

21. 剪切文件到 HDFS

```
hdfs dfs -moveFromLocal <localsrc> <dst>
```

22. 剪切文件到本地

```
hdfs dfs -moveToLocal [-crc] <src> <dst>
```

例如：

```
hdfs dfs -moveToLocal /test/test.txt /tmp
```

执行是提示：“moveToLocal: Option '-moveToLocal' is not implemented yet.”，意思是当前版本中未实现此命令。

23. 移动 HDFS 文件

在 HDFS 中移动文件位置。

```
hdfs dfs -mv URI [URI ...] <dest>
```

24. 修改副本数

```
hdfs dfs -setrep [-R] [-w] <numReplicas> <path>
```

25. 显示文件统计信息

```
hdfs dfs -stat URI [,URI ...]
```

26. 查看文件尾部内容

```
hdfs dfs -tail [-f] URI
```