HDFS的Shell操作

基本语法

- 1. hadoop fs 具体命令【推荐】
- 2. hdfs dfs 具体命令

常用命令大全

• 启动Hadoop集群

```
// 这些命令也都可以按照自己的方式组成脚本哦 start-dfs.sh start-yarn.sh
```

• -help: 输出这个命令参数

```
hadoop fs -help rm
```

• -ls: 显示目录信息

```
hadoop fs -ls /
hadoop fs -ls -R / 递归查看
```

```
[zhutian@hadoop102 shell]$ hadoop fs -ls -R /
drwxrwxrwx - zhutian supergroup 0 2020-05-13 15:12 /input
-rwxrwxrwx 3 zhutian supergroup 38 2020-05-13 15:12 /input/word_data.txt
drwxr-xr-x - zhutian supergroup 0 2020-05-13 15:17 /output
-rw-r--r-- 3 zhutian supergroup 0 2020-05-13 15:17 /output/_SUCCESS
-rw-r--r-- 3 zhutian supergroup 25 2020-05-13 15:17 /output/part-r-00000
```

• -mkdir: 在HDFS上创建目录

```
hadoop fs -mkdir -p /input/word_data
```

• -moveFromLocal: 从本地剪切粘贴到HDFS

```
touch new_data.txt 创建文件
hadoop fs -moveFromLocal ./new_data.txt /input/
```

• -appendToFile: 追加一个文件到已经存在的文件末尾

```
echo "hello xiaofei" >> xiaofei.txt
hadoop fs -appendToFile ./xiaofei.txt /input/new_data.txt
```

```
[zhutian@hadoop102 module]$ echo "hello xiaofei" >> xiaofei.txt
[zhutian@hadoop102 module]$ hadoop fs -appendToFile ./xiaofei.txt /input/new_data.txt
[zhutian@hadoop102 module]$ hadoop fs -cat /input/new_data.txt
hello xiaofei
```

• -cat: 显示文件内容

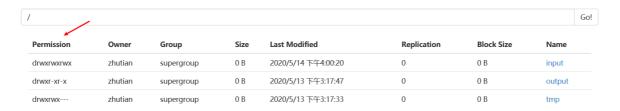
```
hadoop fs -cat /input/new_data.txt
```

看上图

• -chgrp 、-chmod、-chown: Linux文件系统中的用法一样,修改文件所属权限

```
hadoop fs -chmod 777 /input
hadoop fs -chown zhutian:zhutian /input
```

Browse Directory



• -copyFromLocal: 从本地文件系统中拷贝文件到HDFS路径去

```
hadoop fs -copyFromLocal ./xiaofei.txt /input
```

```
[zhutian@hadoop102 module]$ hadoop fs -copyFromLocal ./xiaofei.txt /input
[zhutian@hadoop102 module]$ hadoop fs -ls -R /
drwxrwxrwx - zhutian supergroup
                                              0 2020-05-14 16:06 /input
rw-r--r-- 3 zhutian supergroup
                                             14 2020-05-14 16:02 /input/new_data.txt
drwxr-xr-x
             - zhutian supergroup
                                              0 2020-05-14 15:56 /input/word_data
                                             38 2020-05-13 15:12 /input/word_data.txt
rwxrwxrwx 3 zhutian supergroup
                                             14 2020-05-14 16:06 /input/xiaofei.txt 0 2020-05-13 15:17 /output 0 2020-05-13 15:17 /output/_SUCCESS
rw-r--r-- 3 zhutian supergroup
drwxr-xr-x
             - zhutian supergroup
-rw-r--r--
             3 zhutian supergroup
```

• -copyToLocal: 从HDFS拷贝到本地

```
hadoop fs -copyToLocal /input/xiaofei.txt ./
```

• -cp: 从HDFS的一个路径拷贝到HDFS的另一个路径

```
hadoop fs -cp /input/xiaofei.txt /output/xiaofei
```

```
[zhutian@hadoop102 module]$ hadoop fs -cp /input/xiaofei.txt /output/xiaofei
[zhutian@hadoop102 module]$ hadoop fs -ls -R /
drwxrwxrwx - zhutian supergroup
                                            0 2020-05-14 16:06 /input
                                            14 2020-05-14 16:02 /input/new_data.txt 0 2020-05-14 15:56 /input/word_data
             3 zhutian supergroup
-rw-r--r--
drwxr-xr-x - zhutian supergroup
-rwxrwxrwx 3 zhutian supergroup
                                            38 2020-05-13 15:12 /input/word_data.txt
rw-r--r-- 3 zhutian supergroup
                                            14 2020-05-14 16:06 /input/xiaofei.txt
                                            0 2020-05-14 16:08 /output
drwxr-xr-x - zhutian supergroup
rw-r--r-- 3 zhutian supergroup
                                            0 2020-05-13 15:17 /output/_SUCCESS
                                            25 2020-05-13 15:17 /output/part-r-00000
rw-r--r-- 3 zhutian supergroup
rw-r--r-- 3 zhutian supergroup
                                           14 2020-05-14 16:08 /output/xiaofei
```

• -mv: 在HDFS目录中移动文件

```
hadoop fs -mv /input/new_data.txt /output/new_data.txt
```

```
[zhutian@hadoop102 module]$ hadoop fs -mv /input/new_data.txt /output/new_data.txt
[zhutian@hadoop102 module]$ hadoop fs -ls -R /
                                               0 2020-05-14 16:09 /input
drwxrwxrwx

    zhutian supergroup

drwxr-xr-x - zhutian supergroup
                                              0 2020-05-14 15:56 /input/word_data
-rwxrwxrwx 3 zhutian supergroup
                                              38 2020-05-13 15:12 /input/word_data.txt
-rw-r--r-- 3 zhutian supergroup
                                             14 2020-05-14 16:06 /input/xiaofei.txt
drwxr-xr-x
                                              0 2020-05-14 16:09 /output
0 2020-05-13 15:17 /output/_SUCCESS
14 2020-05-14 16:02 /output/new_data.txt
              - zhutian supergroup
- rw- r- - r- -
- rw- r- - r- -
              3 zhutian supergroup
              3 zhutian supergroup
-rw-r--r--
                                               25 2020-05-13 15:17 /output/part-r-00000
             3 zhutian supergroup
```

• -get: 等同于copyToLocal, 就是从HDFS下载文件到本地

```
hadoop fs -get /input/xiaofei.txt ./
```

```
[zhutian@hadoop102 module]$ hadoop fs -get /input/xiaofei.txt ./
get: `./xiaofei.txt': File exists
[zhutian@hadoop102 module]$
```

• -getmerge: 合并下载多个文件, 比如HDFS的目录 /aaa/下有多个文件:log.1, log.2,log.3,...

```
hadoop fs -getmerge /input/* ./together.txt
```

```
[zhutian@hadoop102 module]$ echo "hello zhutian" >> zhutian.txt
[zhutian@hadoop102 module]$ hadoop fs -copyFromLocal ./zhutian.txt /input
[zhutian@hadoop102 module]$ hadoop fs -getmerge /input/* ./together.txt
[zhutian@hadoop102 module]$ ls
flink-1.7.2 hadoop-2.7.2 jdk1.8.0_144 scala-2.11.8 together.txt xiaofei.txt zhutian.txt
[zhutian@hadoop102 module]$ cat together.txt
hadoop
hadoop
spark
flink
flink
flink
flink
hello xiaofei
hello zhutian
```

• -put: 等同于copyFromLocal

```
hadoop fs -put ./together.txt /input
```

• -tail: 显示一个文件的末尾

hadoop fs -tail /input/xiaofei.txt

[zhutian@hadoop102 module]\$ hadoop fs -tail /input/xiaofei.txt
hello xiaofei
[zhutian@hadoop102 module]\$ |

• -rm: 删除文件或文件夹

hadoop fs -rm /output/new_data.txt

[zhutian@hadoop102 module]\$ hadoop fs -rm /output/new_data.txt 20/05/14 16:33:10 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion = 0 minutes, Emptier interval = 0 minutes. Deleted /output/new data.txt

• -rmdir: 删除空目录

hadoop fs -mkdir /test

• -du统计文件夹的大小信息

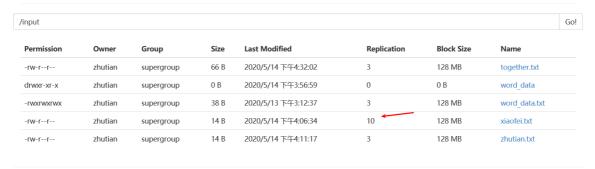
hadoop fs -du -h /input/xiaofei.txt

[zhutian@hadoop102 module]\$ hadoop fs -du -s -h /input/xiaofei.txt
14 /input/xiaofei.txt
[zhutian@hadoop102 module]\$

• -setrep: 设置HDFS中文件的副本数量

hadoop fs -setrep 10 /input/xiaofei.txt

Browse Directory



Hadoop, 2015.

这里设置的副本数只是记录在NameNode的元数据中,是否真的会有这么多副本,还得看DataNode的数量。因为目前只有3台设备,最多也就3个副本,只有节点数的增加到10台时,副本数才能达到10。