****

**大数据技术原理及应用**

实验报告



学 院： 计算机科学与技术学院

专 业： 电力信息技术

报告名称： HDFS Shell实践

学生姓名： 崔荣成 学号： 19108002

指导老师： 周 平

时 间： 2020年4月27日

目 录

[HDFS Shell 练习 2](#_Toc43970220)

[(1)-help 2](#_Toc43970221)

[(2)-ls 2](#_Toc43970222)

[(3)-mkdir 3](#_Toc43970223)

[(4)-copyFromLocal 3](#_Toc43970224)

[(5)-copyToLocal 4](#_Toc43970225)

[(6)-appendToFile 4](#_Toc43970226)

[(7)-cat 5](#_Toc43970227)

[(8)-checksum 5](#_Toc43970228)

[(9)-count 5](#_Toc43970229)

[(10)-cp 6](#_Toc43970230)

[(11)-find 6](#_Toc43970231)

[(12)-get 7](#_Toc43970232)

[(13)-put 7](#_Toc43970233)

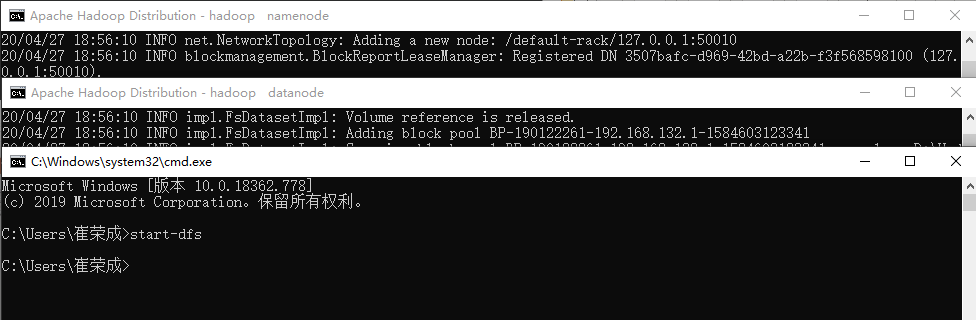
[(14)-rm 8](#_Toc43970234)

[(15)-rmdir 8](#_Toc43970235)

[(16)-text 9](#_Toc43970236)

HDFS Shell 练习

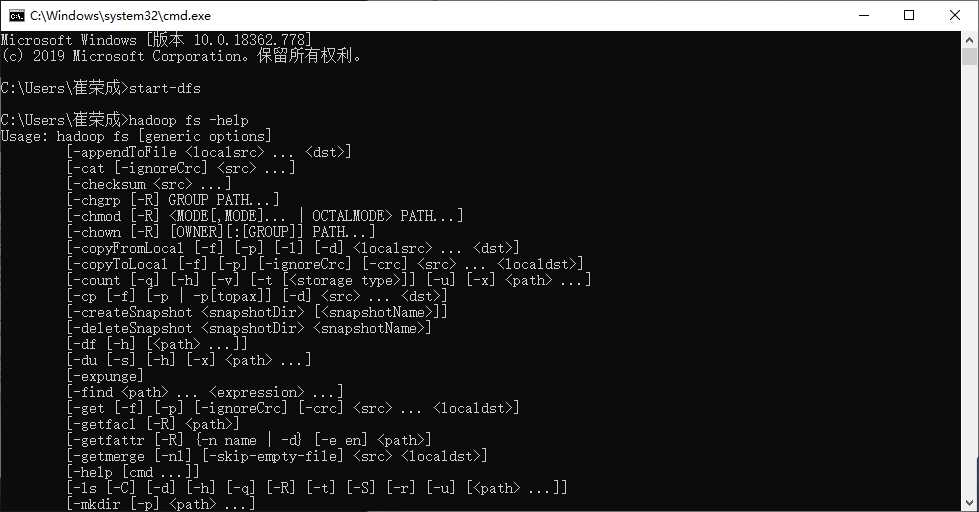
开始使用命令前，必须启动Hadoop



# (1)-help

作用：帮助文档

用法：hadoop fs -help



# (2)-ls

作用：查看文件，与linux下ls命令基本类似。

用法:hadoop fs -ls [-d] [-h] [-R] <args>

选项：

-d:只展示查询展示目录；

-h:显示为人眼更易识别的单位（原来是字节）。

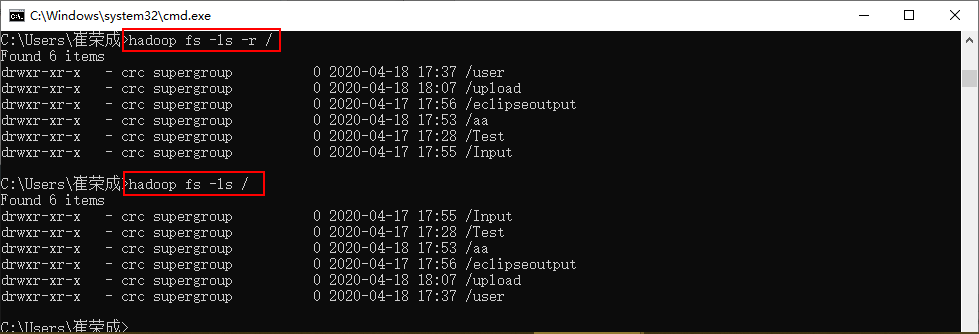
-R:递归展示，显示所有的文件夹及文件

示例：

hadoop fs -ls -d /

hadoop fs -ls -h /

hadoop fs -ls -R /



# (3)-mkdir

作用：创建文件夹。

用法：hadoop fs -mkdir [-p] <paths>

选项：

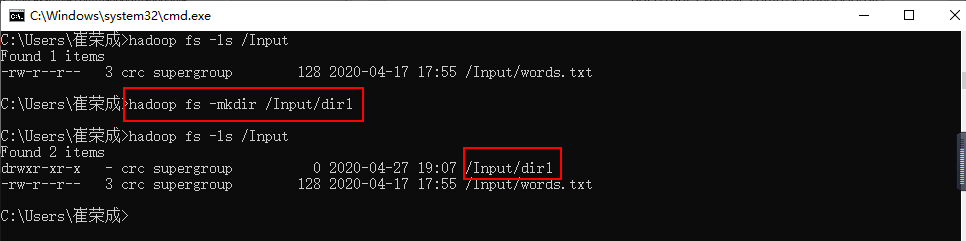
-p:创建父目录。类似于Unix的mkdir -p命令。

示例：

hadoop fs -mkdir /user/hadoop/dir1 /user/hadoop/dir2

hadoop fs -mkdir hdfs://nn1.example.com/user/hadoop/dir

hdfs://nn2.example.com/user/hadoop/dir



# (4)-copyFromLocal

用法：hadoop fs -copyFromLocal <localsrc> URI

作用：类似于put命令，和put不同的是，拷贝的源地址必须是本地文件地址。

-f 参数 当拷贝的目标文件存在时，进行覆盖。

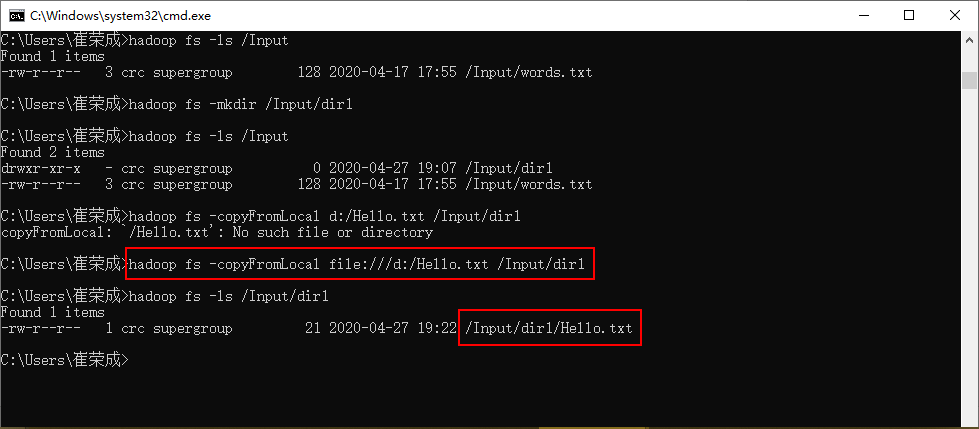
示例：

hadoop fs -copyFromLocal testFlatMap.txt /1.txt

这个时候加上-f参数。即可覆盖。

hadoop fs -copyFromLocal -f testFlatMap.txt /1.txt

**hadoop fs -copyFromLocal d:/Hello2.java /mydir/input/Hello2.java**



# (5)-copyToLocal

用法： hadoop fs -copyToLocal [-ignorecrc] [-crc] URI <localdst>

作用：类似于get指令。和get不同的是，拷贝的目的地址必须是本地文件地址。

# (6)-appendToFile

用法： hadoop fs -appendToFile <localsrc> ... <dst>

作用：是将一个或者多个文件添加到HDFS系统中。

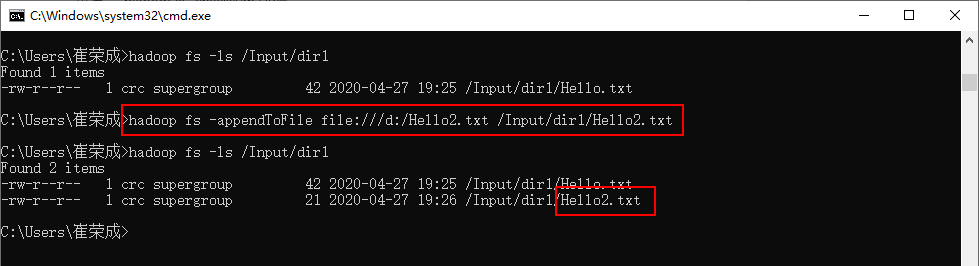
示例:

hadoop fs -appendToFile localfile /user/hadoop/hadoopfile

hadoop fs -appendToFile localfile1 localfile2 /user/hadoop/hadoopfile

hadoop fs -appendToFile localfile hdfs://nn.example.com/hadoop/hadoopfile

**hadoop fs -appendToFile d:/Hello2.java /mydir/input/Hello2.java**



# (7)-cat

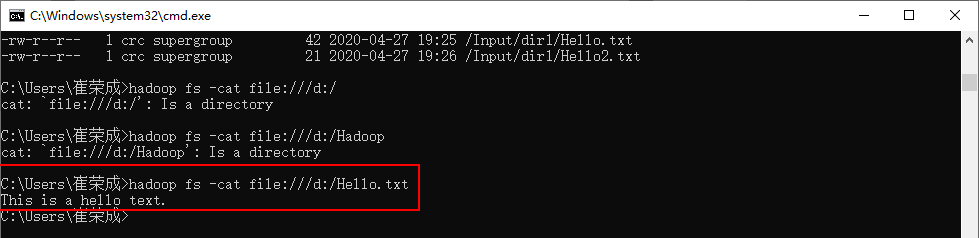
用法：hadoop fs -cat URI [URI ...]

作用：查看文件内容（可以查看本地和HDFS上的内容）。

示例：

hadoop fs -cat hdfs://nn1.example.com/file1 hdfs://nn2.example.com/file2

hadoop fs -cat file:///file3 /user/hadoop/file4



# (8)-checksum

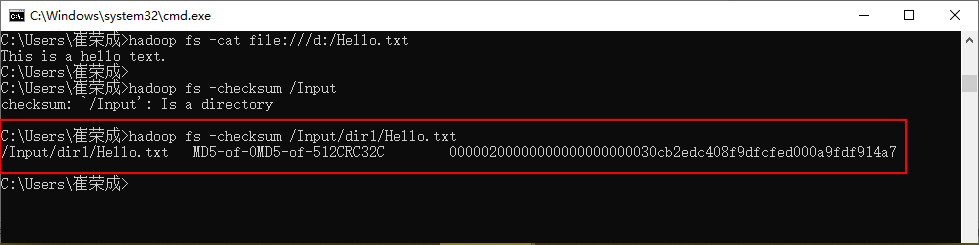
用法： hadoop fs -checksum URI

作用：查看校验码信息。（例子显示了MD5）

示例：

hadoop fs -checksum hdfs://nn1.example.com/file1

hadoop fs -checksum [file:///etc/hosts](file:///C:\etc\hosts)



# (9)-count

作用：计算paths下的目录数，文件数和字节数。

用法： hadoop fs -count [-q] [-h] [-v] <paths>

hadoop fs -count hdfs://nn1.example.com/file1 hdfs://nn2.example.com/file2

hadoop fs -count -q hdfs://nn1.example.com/file1

hadoop fs -count -q -h hdfs://nn1.example.com/file1

hdfs dfs -count -q -h -v hdfs://nn1.example.com/file1



# (10)-cp

用法：hadoop fs -cp [-f] [-p | -p[topax]] URI [URI ...] <dest>

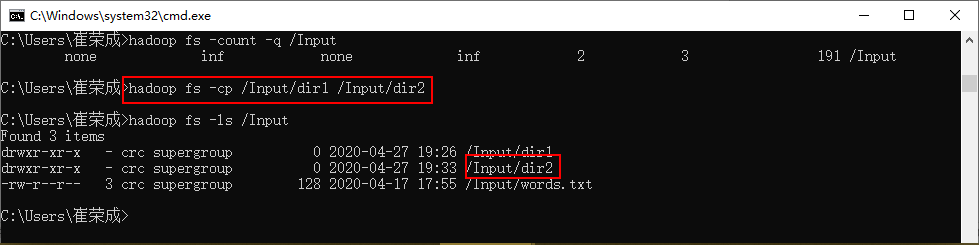
作用：拷贝，HDFS文件系统中进行的拷贝操作。

-f 参数选项：当文件存在时，进行覆盖。

-p 参数选项：将权限、所属组、时间戳、ACL以及XAttr等也进行拷贝。

示例：

hadoop fs -cp -p /tmp/fan /tmp/fan1



# (11)-find

作用：查找满足表达式的文件和文件夹。没有配置path的话，默认的就是全部目录/；如果表达式没有配置，则默认为-print。

用法: hadoop fs -find <path> ... <expression> ...

-name pattern 所要查找文件的文件名。

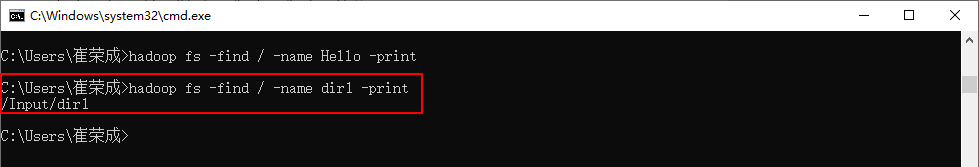
-iname pattern 所要查找的文件名，不区分大小写。

-print 打印。

-print0 打印在一行，如下图所示。

示例：

**hadoop fs -find / -name test -print**



# (12)-get

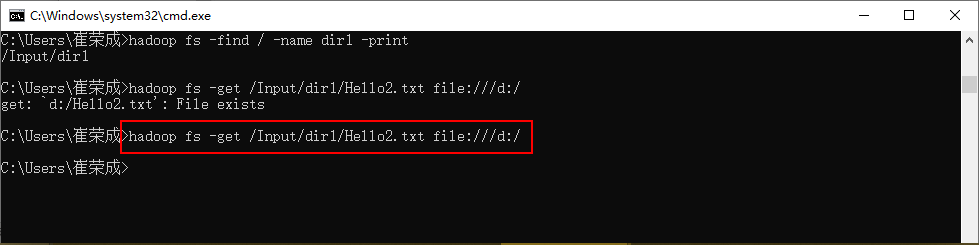
**作用**：从HDFS上拷贝文件到本地。

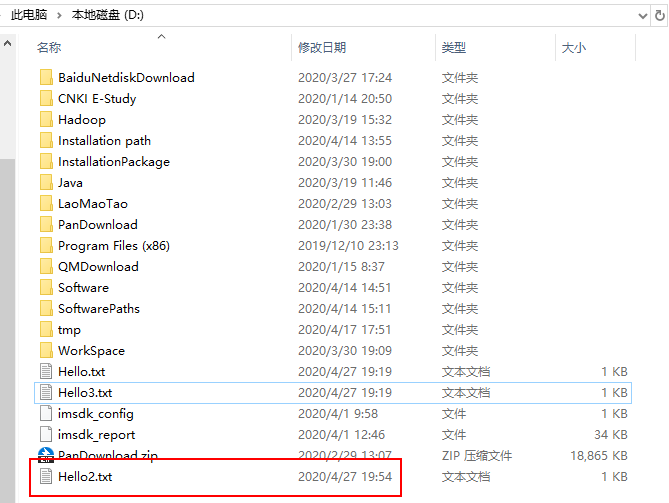
用法：hadoop fs -get [-ignorecrc] [-crc] <src> <localdst>

示例：

hadoop fs -get /user/hadoop/file localfile

hadoop fs -get hdfs://nn.example.com/user/hadoop/file localfile





# (13)-put

用法： hadoop fs -put <localsrc> ... <dst>

作用：将本地的文件上传（复制）到HDFS是dst目录下。

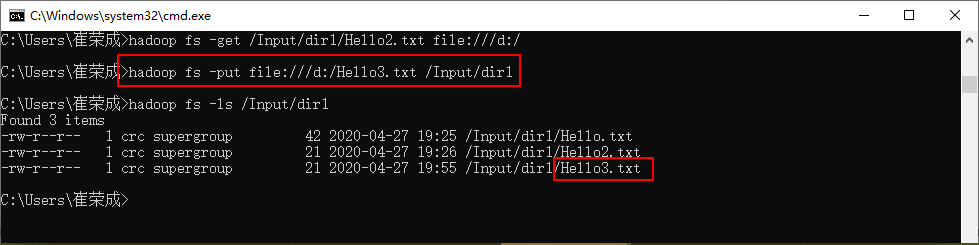
示例：

hadoop fs -put localfile /user/hadoop/hadoopfile

hadoop fs -put localfile1 localfile2 /user/hadoop/hadoopdir

hadoop fs -put localfile hdfs://nn.example.com/hadoop/hadoopfile

hadoop fs -put - hdfs://nn.example.com/hadoop/hadoopfile Reads the input from stdin.



# (14)-rm

用法：hadoop fs -rm [-f] [-r |-R] [-skipTrash] URI [URI ...]

作用：删除文件。

选项：

The -f option will not display a diagnostic message or modify the exit status to reflect an error if the file does not exist.

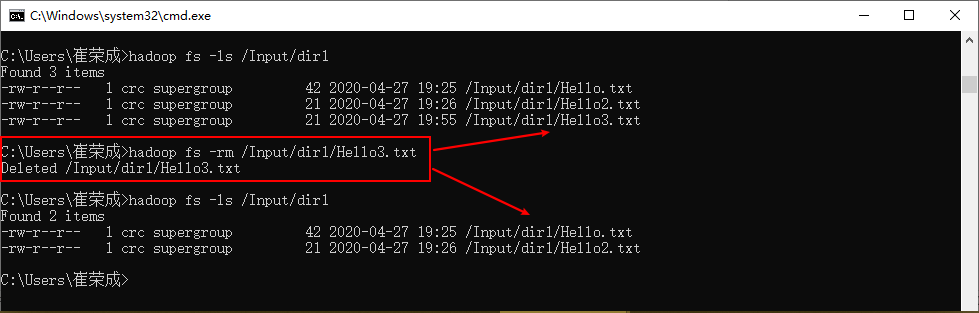
The -R option deletes the directory and any content under it recursively.

The -r option is equivalent to -R.

The -skipTrash option will bypass trash, if enabled, and delete the specified file(s) immediately. This can be useful when it is necessary to delete files from an over-quota directory.

示例：

hadoop fs -rm hdfs://nn.example.com/file /user/hadoop/emptydir



# (15)-rmdir

用法：hadoop fs -rmdir [--ignore-fail-on-non-empty] URI [URI ...]

作用：删除空目录。

选项：

—ignore-fail-on-non-empty：使用它的时候，忽略因文件夹非空删除失败的信息。

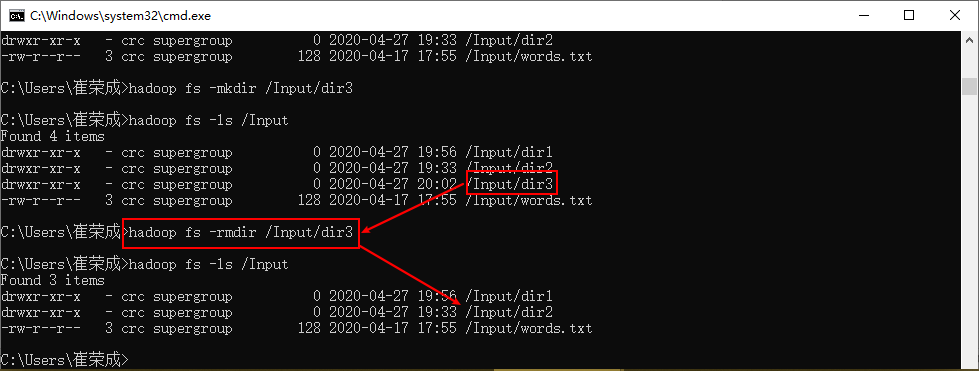
hadoop fs -setfacl -b /file

hadoop fs -setfacl -k /dir

hadoop fs -setfacl --set user::rw-,user:hadoop:rw-,group::r--,other::r-- /file

hadoop fs -setfacl -R -m user:hadoop:r-x /dir

hadoop fs -setfacl -m default:user:hadoop:r-x /dir



# (16)-text

用法：hadoop fs -text <src>

作用：将HDFS中文件以文本形式输出（包括zip包，jar包等形式）

示例：hadoop fs -text /wc.jar

