## 实验3：HDFS操作及编程

### 一、实验目的

* 继续练习Linux命令；
* 熟练使用HDFS操作常用的Shell命令；
* 熟悉HDFS操作常用的Java API。

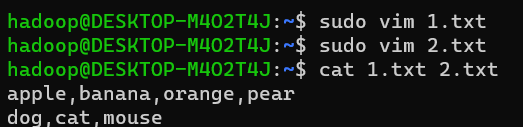
### 二、实验平台

* 操作系统：Linux（建议Ubuntu16.04）；
* Hadoop版本：2.7.1；
* JDK版本：1.7或以上版本；
* Java IDE：Eclipse。

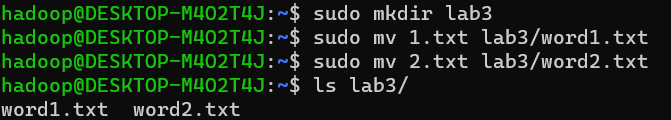
### 三、实验步骤

**（一）HDFS Shell 命令综合使用**

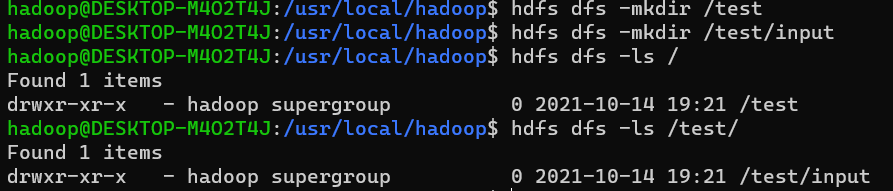
1、在当前用户目录下新建文本文件1.txt及2.txt，文件中分别录入若干单词

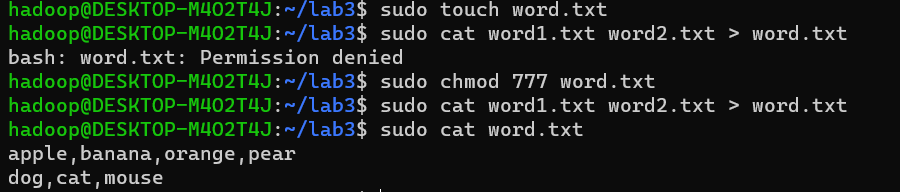


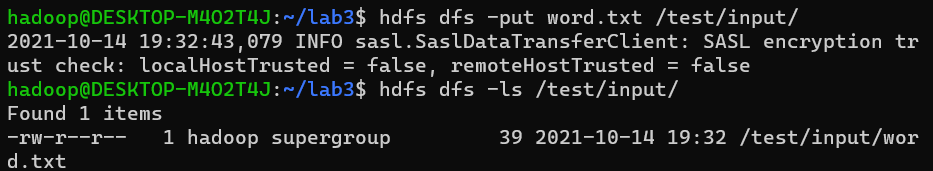
2、在当前用户目录下新建lab2目录，将1.txt及2.txt移动到lab2的test目录下并改名为word1.txt及word2.txt



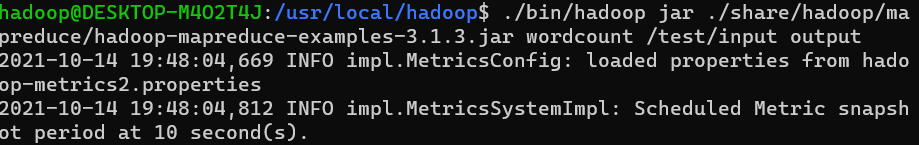
3、在HDFS的根目录下新建test/input目录，将本地文件系统中的word1.txt和word2.txt合并成word.txt后上传到HDFS的input目录中，查看HDFS中test目录下所有文件的读写权限、大小、创建时间、路径等，如果某输出文件是目录，则递归输出该目录下所有文件相关信息

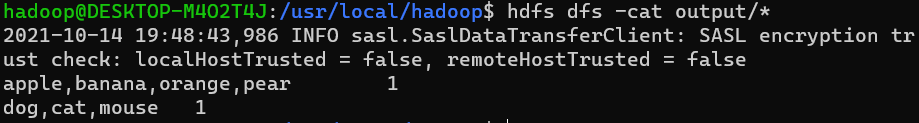




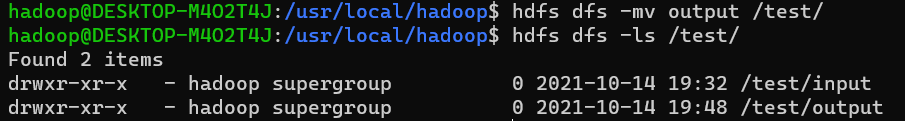


4、调用wordcount示例程序统计word.txt中的单词出现次数，将结果写入 /test/out中，查看统计结果

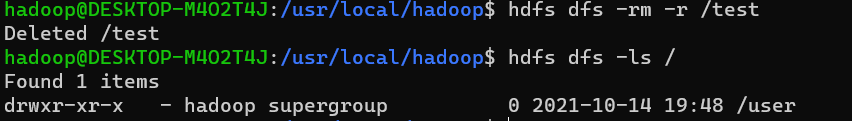




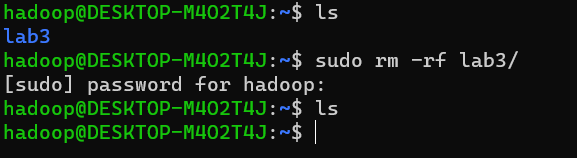
5、将统计结果文件移动到HDFS的/test目录下，并下载到本地文件系统的lab2的test目录下，然后删除统计结果文件



6、删除HDFS的/test目录及包含的所有文件



7、删除本地文件系统的test目录及包含的所有文件



**（二）在eclipse中导入与HDFS相关的jar包，编程检测HDFS的/test/input目录下是否存在word.txt文件，并将该程序发布成jar包（若在自己电脑上完成，需先安装Hadoop及eclipse）**

1. **将本地文件系统中的word1.txt和word2.txt上传到HDFS的/test/input目录下，编程实现将word1.txt和word2.txt合并成merge.txt放到/test/input目录下。**
2. **（选做）编程实现将HDFS的/test/input目录下的所有.txt文件合并成merge2.txt放到/test/input目录下。**

### 实验报告

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验报告题目： | | | | | |
| 学号（班级）： |  | 姓名： |  | 日期： |  |
| 实验环境： | | | | | |
| 实验内容与完成情况： | | | | | |
| 出现的问题： | | | | | |
| 解决方案（列出遇到的问题和解决办法，列出没有解决的问题）： | | | | | |