**西南石油大学实验报告**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **课 程** | 大数据处理技术 | | **实验项目** | | | HDFS安装、性能调优调试 | | | | | **成绩** | | |  |
| **专业班级** | 网络工程1602 | | **组别** | | 无 | | **学号** | 201631063211 | | **指导教师** | | 赖俊良 | | |
| **姓 名** | 古皓 | **同组人姓名** | | 无 | | | | | **实验日期** | | | | 2019/6/17 | |

1. **实验目的**

掌握HDFS安装、性能调优调试。

1. **实验内容**

在集群节点中输入HDFS命令行完成目录管理和文件上传/下载等操作

1. **实验步骤**

1. 查看HDFS根目录；

2. 在HDFS根目录创建文件夹test；

3. 从Linux中上传一个文件至HDFS的/test目录；

4. 将该文件从HDFS下载至/Home目录中。

在Eclipse环境中创建一个Maven项目，并使用HDFS的Java API操作HDFS

1. 下载Eclipse，并创建Maven项目，菜单项File/New/Other…选择Maven Project；

2. 编辑pom.xml文件，引用必要的Hadoop和HDFS组件，注意将文件中JDK的路径和版本改为自己计算机上的实际路径和版本，编辑完成后在项目上点右键，执行Maven/Update Project和Run As/Maven Install。

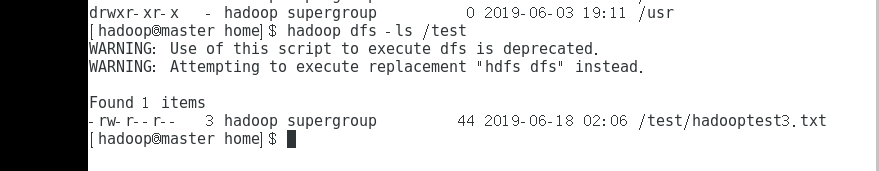
3. 使用Java API编写代码并调试，将Windows中的一个文件上传至HDFS指定目录；常见异常参见《HDFS Java API常见异常解决方案》；

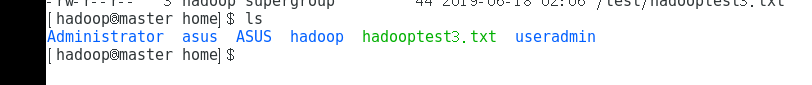
4. 在集群中使用HDFS命令查看是否上传成功。

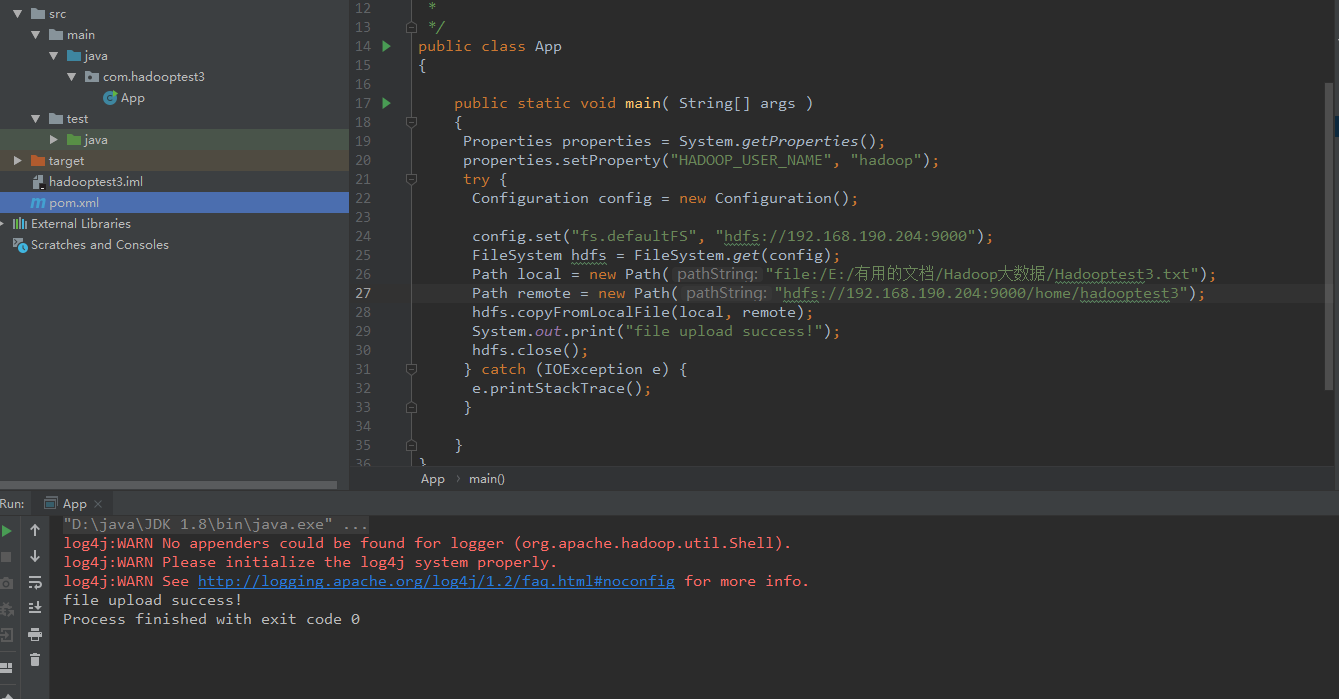
**四、实验结果**

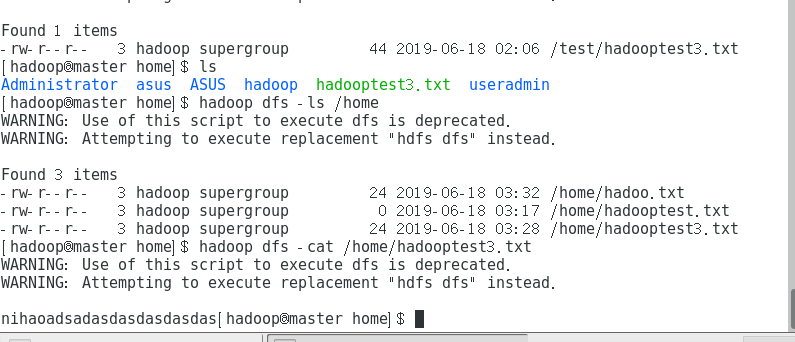
试验运行过程及结果

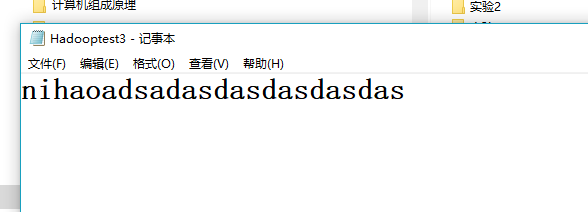
根目录，创建了test目录

从Linux上传的hadooptest3.txt文件

下载到Linux /home目录里

本地上传文件Hadooptest3.txt到hdfs /home目录下

Hadooptest3.txt是上传的文件

内容是

**五、实验体会**

此次试验内容不难，遇到的问题主要还是文件权限的问题，通过百度查找解决了问题成功将代码跑出来了，通过自己去查资料解决问题也是一种提升自己知识认知的一个过程，对实验有了更广泛深入的理解，有了更全面的认识，总之这次实验受益良多获益匪浅！