**大数据管理实验报告 （ 2 ）**

**班级： 学号： 姓名： 成绩：**

|  |  |
| --- | --- |
| **实验名称**：熟悉常用的HDFS操作 | **实验地点**：信息楼-318 |
| **所使用的工具软件及环境：**VirtualBox6、Ubuntu Linux18、Hadoop-3.3.1、Eclipse | |
| **一、实验目的：**  本实验的目的是理解HDFS在Hadoop体系结构中的角色、熟练使用HDFS操作常用的Shell命令、、熟悉HDFS操作常用的Java API、编写Java程序访问HDFS文件系统。 | |
| **二、实验内容：**  按照https://dblab.xmu.edu.cn/blog/2460/上的步骤进行下面操作：  1、利用Shell命令与HDFS进行交互  2、利用Web界面管理HDFS  3、利用Java API与HDFS进行交互 | |
| **三、实验过程及结果：**  1、给出运行./bin/hadoop fs命令的结果截图  2、在HDFS中创建用户目录：“/user/hadoop”，在该用户目录下再创建一个input目录，创建好后列出“/user/hadoop”目录下的内容（给出命令和结果截图）  3、在HDFS的根目录下创建一个名称为input的目录，列出根目录下的内容（给出命令和结果截图）  4在本地Linux文件系统的“/home/hadoop/”目录下创建一个文件myLocalFile.txt，编辑该文件，里面随意输入一些单词（要求：其中一个单词是你的学号+Hadoop）  （4-1）把本地文件系统的“/home/hadoop/myLocalFile.txt”上传到HDFS中的当前用户目录的input目录下，并使用ls命令查看文件已经成功上传到HDFS中（给出命令和结果截图）。  （4-2）查看HDFS中的myLocalFile.txt这个文件的内容（给出命令和结果截图）  5、把HDFS中的myLocalFile.txt文件下载到本地文件系统中的“/home/hadoop/下载/”，列出“/home/hadoop/下载/”下的文件，并查看myLocalFile.txt。（给出命令和结果截图）  6、把HDFS的“/user/hadoop/input/myLocalFile.txt”文件，拷贝到HDFS的另外一个目录“/input”中。（给出命令和拷贝成功结果截图）  7、打开Linux自带的Firefox浏览器，输入http://localhost:9870，给出HDFS的web管理界面截图。  8、在Ubuntu中安装好Eclipse，给出启动好Eclipse的界面。  9、用Eclipse工具编写Java程序。要执行的任务是：假设在目录“hdfs://localhost:9000/user/hadoop”下面有几个文件，分别是file1.txt、file2.txt、file3.txt、file4.abc和file5.abc，这里需要从该目录中过滤出所有后缀名不为“.abc”的文件，对过滤之后的文件进行读取，并将这些文件的内容合并到文件“hdfs://localhost:9000/user/hadoop/merge.txt”中。  （9-1）给出eclipse中显示工程和类以及程序代码的截图。  （9-2）给出运行成功后，底部的“Console”面板中显示运行结果信息的截图。  （9-3）给出运行成功后的到HDFS中查看生成的merge.txt文件的结果截图。  （9-4) 把该Java应用程序生成JAR包，部署到Hadoop平台上运行，给出查看JAR文件的命令和结果的截图  （9-5）给出直接运行jar文件的命令和查看生成的merge.txt文件的命令和结果截图。  **四、实验过程中遇到的问题及解决方法：** | |

任课教师签名： 2022 年 5 月 日