**HDFS 原理**

1. ****实验目标****

熟悉HDFS的原理及其相关常用命令操作。

1. ****实验内容****

* 了解HDFS架构
* 了解HDFS 读操作和写操作
* 熟悉HDFS 中的常用命令

1. ****实验步骤及结果截图****

* 在 Hadoop 集群中编译并运行《权威指南》中的例 3.2，读取 HDFS 文件内容。

1. 创建代码目录

（1）配置本机主机名为 hadoop，sudo 时需要输入 shiyanlou 用户的密码。将 hadoop 添加到最后一行的末尾。

sudo vim /etc/hosts # 将hadoop添加到最后一行的末尾

ping hadoop



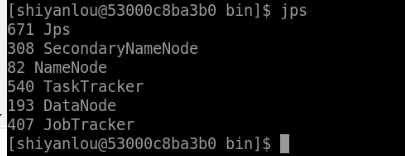
（2）启动 Hadoop

使用代码：

cd /app/hadoop-1.1.2/bin

./start-all.sh

jps # 查看启动的进程，确保 NameNode 和 DataNode 都有启动



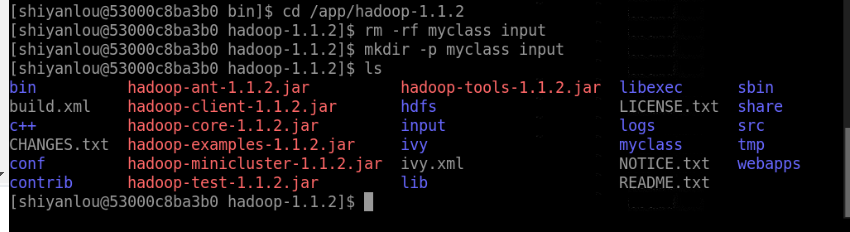
1. 在 /app/hadoop-1.1.2 目录下使用如下命令建立 myclass 和 input 目录

使用代码：

cd /app/hadoop-1.1.2

rm -rf myclass input

mkdir -p myclass input



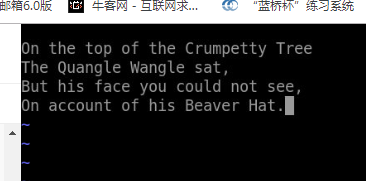
1. 建立例子文件上传到 HDFS 中

（1）进入 /app/hadoop-1.1.2/input 目录，在该目录中创建 quangle.txt 文件。

cd /app/hadoop-1.1.2/input

touch quangle.txt

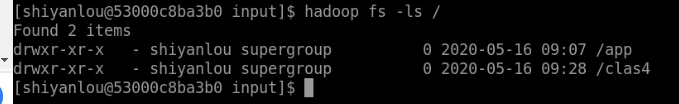
vi quangle.txt



1. 使用如下命令在 hdfs 中建立目录 /class4

hadoop fs -mkdir /class4

hadoop fs -ls /

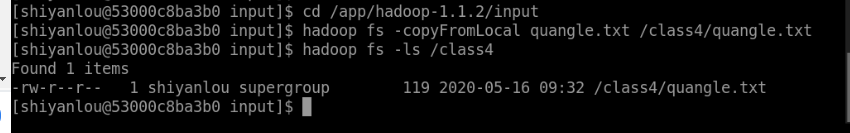


1. 把例子文件上传到 hdfs 的 /class4 文件夹中。

cd /app/hadoop-1.1.2/input

hadoop fs -copyFromLocal quangle.txt /class4/quangle.txt

hadoop fs -ls /class4



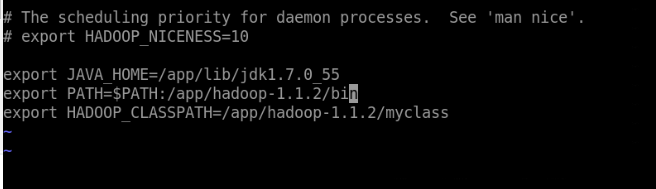
1. 配置本地环境
2. 对 /app/hadoop-1.1.2/conf 目录中的 hadoop-env.sh 进行配置

cd /app/hadoop-1.1.2/conf

sudo vi hadoop-env.sh

1. 加入 HADOOP\_CLASPATH 变量值，值为 /app/hadoop-1.1.2/myclass，设置完毕后编译该配置文件，使配置生效

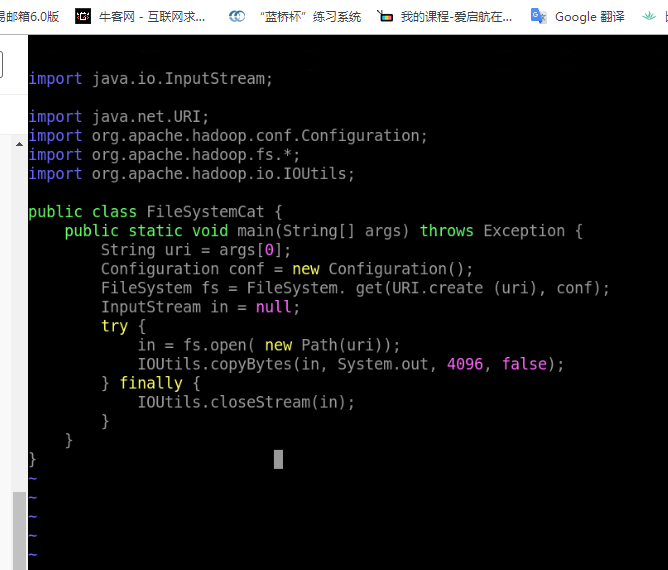
export HADOOP\_CLASSPATH=/app/hadoop-1.1.2/myclass



1. 编写代码
2. 进入 /app/hadoop-1.1.2/myclass 目录，在该目录中建立 FileSystemCat.java 代码文件

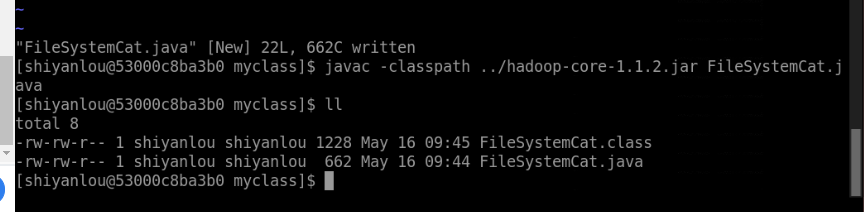
cd /app/hadoop-1.1.2/myclass/

vi FileSystemCat.java



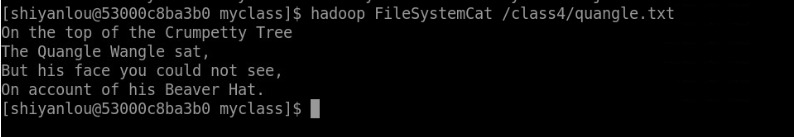
1. 编译代码
2. 在 /app/hadoop-1.1.2/myclass 目录中，使用如下命令编译代码：

javac -classpath ../hadoop-core-1.1.2.jar FileSystemCat.java



1. 使用编译代码读取 HDFS 文件
2. 使用如下命令读取 HDFS 中 /class4/quangle.txt内容：

hadoop FileSystemCat /class4/quangle.txt

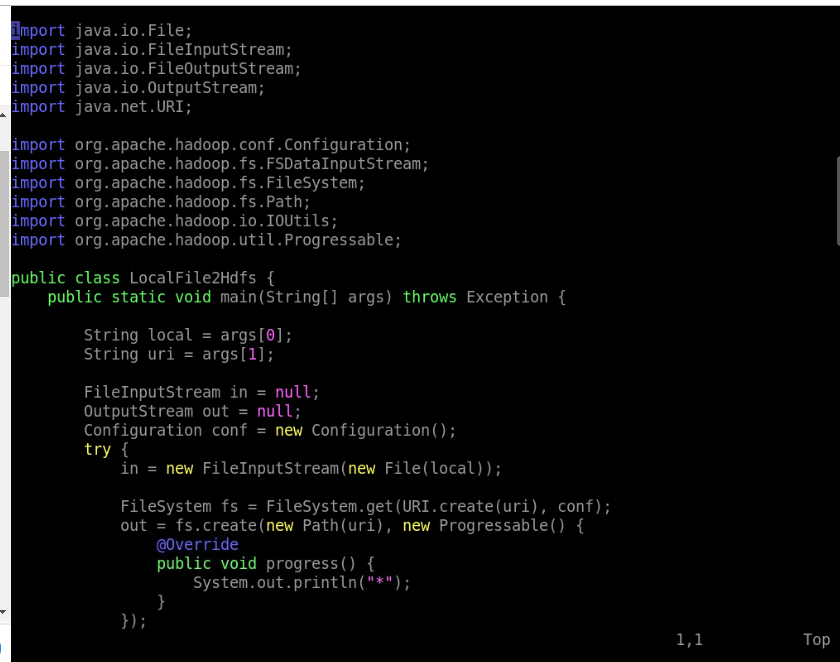


* 在本地文件系统生成一个大约 100 字节的文本文件，写一段程序读入这个文件并将其第 101-120 字节的内容写入 HDFS 成为一个新文件

1. 编写代码
2. 进入 /app/hadoop-1.1.2/myclass 目录，在该目录中立 LocalFile2Hdfs.java 代码文件，命令如下：

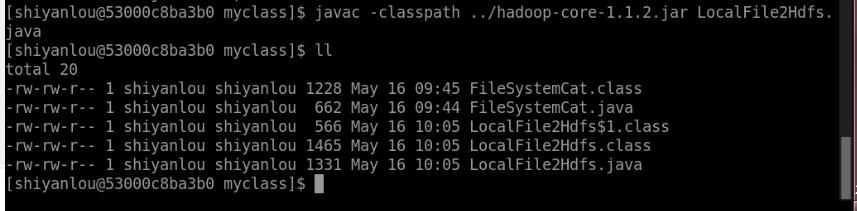
cd /app/hadoop-1.1.2/myclass/

vi LocalFile2Hdfs.java



1. 编译代码
2. 在 /app/hadoop-1.1.2/myclass 目录中，使用如下命令编译代码：

javac -classpath ../hadoop-core-1.1.2.jar LocalFile2Hdfs.java



3. 建立测试文件

（1）进入 /app/hadoop-1.1.2/input 目录，在该目录中建立 local2hdfs.txt 文件

cd /app/hadoop-1.1.2/input/

vi local2hdfs.txt

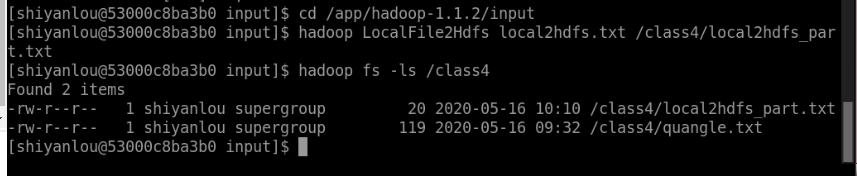
4.使用编译代码上传文件内容到 HDFS

（1）使用如下命令读取 local2hdfs 第 101-120 字节的内容写入 HDFS 成为一个新文件：

cd /app/hadoop-1.1.2/input

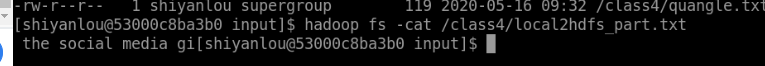
hadoop LocalFile2Hdfs local2hdfs.txt /class4/local2hdfs\_part.txt

hadoop fs -ls /class4



1. 验证是否成功
2. 使用如下命令读取 local2hdfs\_part.txt 内容：

hadoop fs -cat /class4/local2hdfs\_part.txt

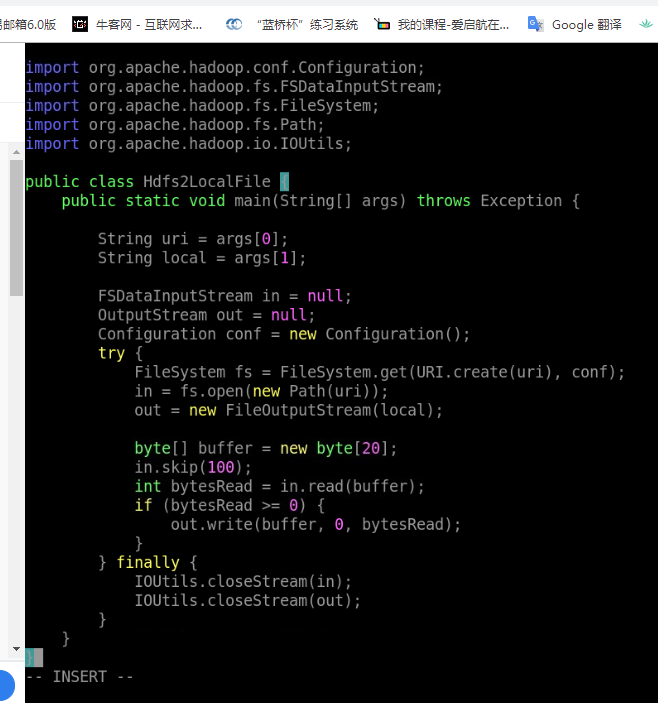


* 实验案例 3 的反向操作，在 HDFS 中生成一个大约 100 字节的文本文件，写一段程序读入这个文件，并将其第 101-120 字节的内容写入本地文件系统成为一个新文件。

1. 编写代码
2. 进入 /app/hadoop-1.1.2/myclass 目录，在该目录中建立 Hdfs2LocalFile.java 代码文件，命令如下：

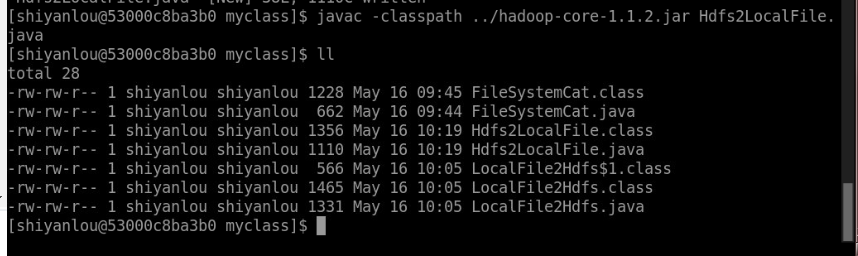
cd /app/hadoop-1.1.2/myclass/

vi Hdfs2LocalFile.java



1. 编译代码
2. 在 /app/hadoop-1.1.2/myclass 目录中，使用如下命令编译代码：

javac -classpath ../hadoop-core-1.1.2.jar Hdfs2LocalFile.java



1. 建立测试文件
2. 进入 /app/hadoop-1.1.2/input 目录，在该目录中建立 hdfs2local.txt 文件

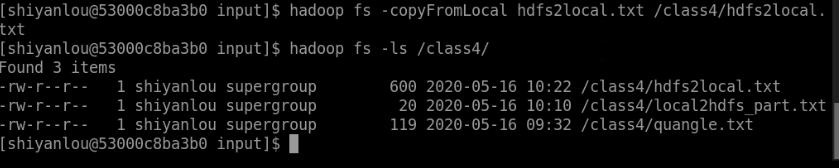
cd /app/hadoop-1.1.2/input/

vi hdfs2local.txt

1. 在 /app/hadoop-1.1.2/input 目录下把该文件上传到 hdfs 的 /class4/ 文件夹中：

hadoop fs -copyFromLocal hdfs2local.txt /class4/hdfs2local.txt

hadoop fs -ls /class4/



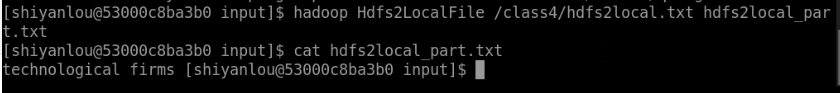
1. 使用编译代码把文件内容从 HDFS 输出到文件系统中
2. 使用如下命令读取 hdfs2local.txt 第 101-120 字节的内容写入本地文件系统成为一个新文件：

hadoop Hdfs2LocalFile /class4/hdfs2local.txt hdfs2local\_part.txt

1. 验证是否成功

（1）使用如下命令读取 hdfs2local\_part.txt 内容：

cat hdfs2local\_part.txt



1. ****实验分析****

本实验介绍了 HDFS 的原理及一些基础的操作，三个实验案例分别运用了三段代码进行读取HDFS文件内容，写入新文件，写入文件系统等操作，得到了相对应的结果。

1. ****心得体会****

经过本次实验，我充分的了解到了HDFS的工作原理，了解了它的架构；同时运用到了HDFS的读和写操作来运用代码熟悉了相关的文件操作；运用了很多HDFS相关的常用命令，大大激发了对于大数据这门课程的学习兴趣，提高了学习热情