通过Java程序向HDFS里面写入文件

1.以用户hdfs身份创建需写入文件的目录

passwd hdfs //设置HDFS用户密码

chsh -s /bin/bash hdfs //更换shell

su hdfs //以HDFS身份登录

yum install java-devel //安装jdk

hadoop fs -mkdir /output //创建写入文件的目录/output

hadoop fs -chmod 777 /output //修改权限

exit

2. 创建目录 /root/java 存放编写及编译的代码

mkdir java //在root下创建目录/java

cd java //进入/java目录

vi HDFSCreateJavaFile.java //创建java文件

编辑文件名为 HDFSCreateJavaFile.java 内容如下

import org.apache.hadoop.conf.Configuration;

import org.apache.hadoop.fs.FSDataOutputStream;

import org.apache.hadoop.fs.FileSystem;

import org.apache.hadoop.fs.Path;

public class HDFSCreateJavaFile {

public static void main(String[] args){

try{

Configuration conf = new Configuration();

conf.set("fs.defaultFS","hdfs://namenode01.hjw.com:8020");

conf.set("fs.hdfs.impl","org.apache.hadoop.hdfs.DistributedFileSystem");

FileSystem fs = FileSystem.get(conf);

byte[] buff = "Bye Bye= =".getBytes();

String fileName = "/output/test.txt";

FSDataOutputStream os =fs.create(new Path(fileName));

os.write(buff,0,buff.length);

System.out.println("Create:"+fileName);

os.close();

fs.close();

} catch (Exception e){

e.printStackTrace();

}

}

}

3.编译java源文件

javac -Djava.ext.dirs=/opt/cloudera/parcels/CDH/jars HDFSCreateJavaFile.java

4. 创建manifest.txt，该文件指定主程序入口

echo Main-class: HDFSCreateJavaFile > manifest.txt

5.生成可执行jar文件

jar cvfm HDFSCreateJavaFile.jar manifest.txt HDFSCreateJavaFile.class

6.测试执行文件

hadoop jar HDFSCreateJavaFile.jar

7.查看程序建立的文件

hadoop fs -cat /output/test.txt

运行结果如下：

