

华东师范大学数据科学与工程学院实验报告

课程名称： 分布式模型与编程	年级： 2016 级	上机实践成绩：
指导教师： 徐辰	姓名： 张宏伟	
上机实践名称： MapReduce 编程	学号： 10165101180	上机实践日期：
上机实践编号： 05	组号：	上机实践时间：

一、实验目的

使用 MapReduce 模型编程

二、实验任务

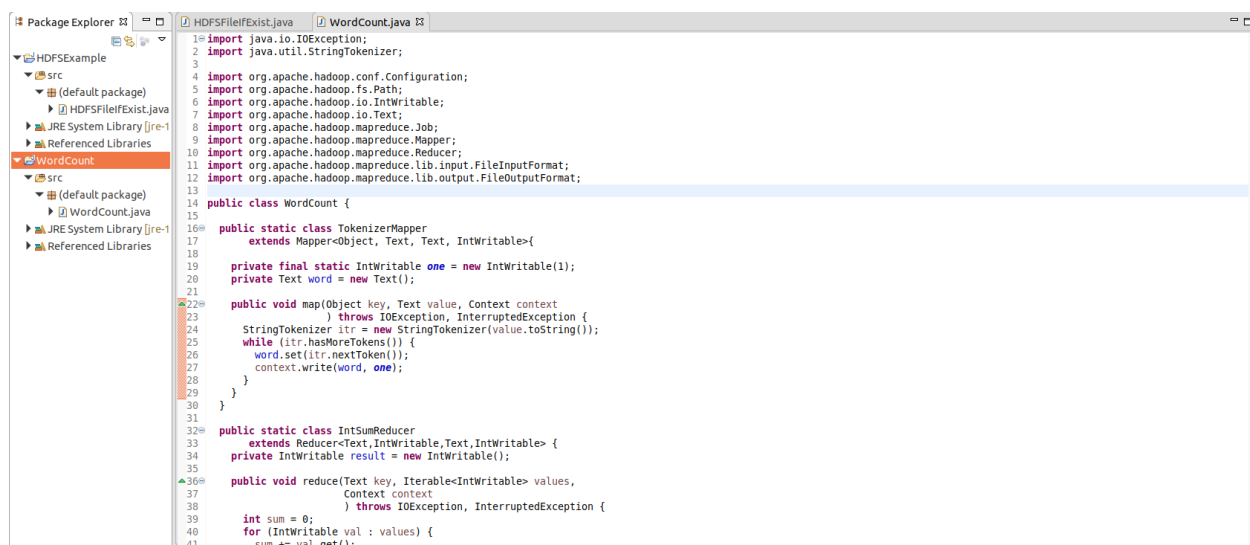
编写 wordCount 程序，统计文件中各单词出现次数

三、使用环境

Hadoop2.7.3+Eclipse+Jre1.8

四、实验过程

1. 编写 WordCount 程序并打包成 WordCount.jar



2. 编写统计文本 wordFile1.txt 和 wordFile2.txt

WordFile1.txt 内容：

```
I love Spark
I love Hadoop
```

WordFile2.txt 内容：

```
Hadoop is good
Spark is fast
```

3.伪分布式

将文本传输到 hdfs 中

```
hadoop@PC-honwee:/usr/local/hadoop-2.7.3_pseudo$ ./bin/hdfs dfs -ls /user/hadoop/input
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (f
ile:/usr/local/hadoop-2.7.3_pseudo/share/hadoop/common/lib/hadoop-auth-2.7.3.jar) to method sun.secu
rity.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authenticat
ion.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operation
s
WARNING: All illegal access operations will be denied in a future release
Found 2 items
-rw-r--r-- 1 hadoop supergroup 27 2018-10-22 19:42 /user/hadoop/input/wordFile1.txt
-rw-r--r-- 1 hadoop supergroup 29 2018-10-22 19:42 /user/hadoop/input/wordFile2.txt
```

执行命令(部分截图):

```
hadoop@PC-honwee:/usr/local/hadoop-2.7.3_pseudo$ ./bin/hadoop jar myapp/WordCount.jar input output
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (f
ile:/usr/local/hadoop-2.7.3_pseudo/share/hadoop/common/lib/hadoop-auth-2.7.3.jar) to method sun.secu
rity.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authenticat
ion.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operation
s
WARNING: All illegal access operations will be denied in a future release
```

运行结果

```
hadoop@PC-honwee:/usr/local/hadoop-2.7.3_pseudo$ ./bin/hdfs dfs -cat output/*
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.
util.KerberosUtil (file:/usr/local/hadoop-2.7.3_pseudo/share/hadoop/common/lib/h
adoop-auth-2.7.3.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.
security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflect
ive access operations
WARNING: All illegal access operations will be denied in a future release
Hadoop 2
I 2
Spark 2
fast 1
good 1
is 2
love 2
```

4.集群下

创建个人文件

```
hadoop24@ubuntu16g-1:/usr/local/hadoop$ ./bin/hadoop fs -mkdir /user/zhw
18/10/22 18:38:14 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platfor
m... using builtin-java classes where applicable
```

传输 wordFile1.txt 和 wordFile2.txt 到 hdfs 的/user/zhw/input 下。

```
hadoop24@ubuntu16g-1:/usr/local/hadoop$ ./bin/hadoop fs -ls /user/zhw/input
18/10/22 18:51:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platfor
m... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 3 hadoop24 supergroup 27 2018-10-22 18:49 /user/zhw/input/wordfile1.txt
-rw-r--r-- 3 hadoop24 supergroup 29 2018-10-22 18:49 /user/zhw/input/wordfile2.txt
```

第一次跑的时候用 jre10 版本编译，报错，后改用 jre1.8 正常

```
hadoop24@ubuntu16g-1:/usr/local/hadoop$ ./bin/hadoop jar ~/MapReduce.jar /user/zhw/input /user/zhw/output
Exception in thread "main" java.lang.UnsupportedClassVersionError: WordCount has been compiled by a more recent version of the Java Runtime (class file version 54.0), this version of the Java Runtime only recognizes class file versions up to 52.0
```

注意运行程序 WordCount.jar 是放在本地而不是 hdfs 上的，若放在 hdfs 上运行会报 “not a valid jar”，最后运行结果：

```
hadoop24@ubuntu16g-1:/usr/local/hadoop$ ./bin/hadoop dfs -cat /user/zhw/output/*
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

18/10/22 19:15:50 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Hadoop 2
I 2
Spark 2
fast 1
good 1
is 2
love 2
```

5.尝试在 Docker 中也完成分布式的执行，结果相同

```
root@master:/usr/local/hadoop# hdfs dfs -cat ./output/*
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.7.7.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
Hadoop 2
I 2
Spark 2
fast 1
good 1
is 2
love 2
root@master:/usr/local/hadoop#
```

五、总结

1. 跑的 jar 程序放在本地而不是 hdfs
2. MapReduce 将编程高度抽象到 Map 函数和 Reduce 函数，虽然隐藏了很多细节，但也因此使得表达能力较差，需要编写的代码量较大。