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：** 吴双

上机实践名称： HDFS 编程实践　　**学号：** 10164102141　　**上机实践日期：**

上机实践编号： #4　　　　　　　**组号：**　　　　　　　　**上机实践时间：**

一、 实验目的

熟悉 HDFS 编程过程

二、 实验任务

测试 HDFS 中某文件的操作与鉴别，并在本地及集群环境下运行。

三、 使用环境

本地：
　　Ubunut LTS18.04
　　Hadoop v2.7.3
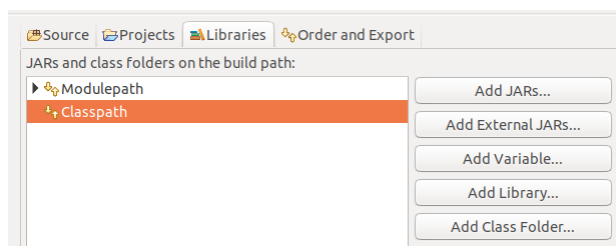　　Eclipse
　　jre10.0.2
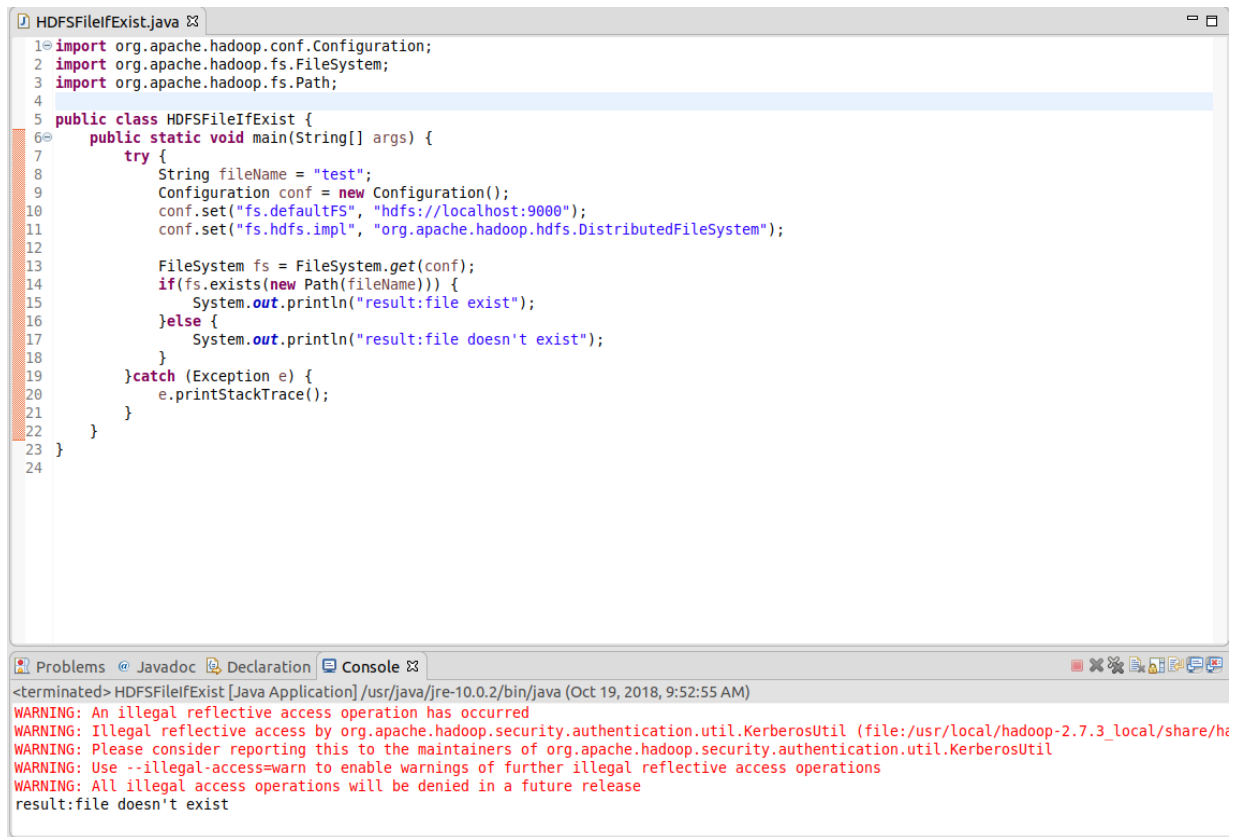Docker 集群：
　　jre10.0.2
班级集群：
　　jre1.8.0_181

四、 实验过程

1.创建 javaProject，导入 HDFS 和 hadoop 相关 jar，注意需要将 jar 导入 Classpath。之前导入 Modulepath，一直报类名重复的错误。

2.输入代码，启动 hdfs，在 Eclipse 中运行：run as JAVA Application，结果如下图

3.将代码环境设置中的 localhost 改成 master，export 为 jar 包，传输到 Docker，在
Docker 集群中运行，结果如下图

4.将代码环境设置中的 localhost 改成 10.11.6.91，编译，export 为 jar 包，scp 到班级集
群自己用户 hadoop23 下，ssh 并运行，结果如下图

五、 总结
导入 jar 的时候要注意，在本地 Elicpse 中运行 hdfs 代码时需先启动本地的 hdfs。