**

**HUNAN UNIVERSITY**

《数据仓库技术与应用》

实验报告三

|  |  |
| --- | --- |
| **报告名称：** | 大数据实验三 |
| **学生姓名：** | 杨超然 |
| **学生学号：** | 202106060220 |
| **专业班级：** | 电商2102班 |
| **学 院：** | 工商管理学院 |
| **指导老师：** | 李信儒 |
| **日 期：** | 2023.3.6 |

目录

[一、实验过程 3](#_Toc69208567)

[（一）利用Shell命令与HDFS进行交互 3](#_Toc1399732542)

[1． 目录操作 4](#_Toc1768964156)

[2． 文件操作 5](#_Toc1216960824)

[（二） 利用Web界面管理HDFS 6](#_Toc98451431)

[（三） 利用Java API与HDFS进行交互 7](#_Toc1110792627)

[1．在Ubuntu中安装Eclipse 7](#_Toc1016338618)

[2．在Eclipse中创建项目 8](#_Toc518224488)

[3． 为项目添加需要用到的JAR包 9](#_Toc1752781231)

[4． 编写Java应用程序代码 11](#_Toc1960963518)

[5． 编译运行程序 14](#_Toc482316517)

[6． 应用程序的部署 15](#_Toc1690417441)

[（四） 相关代码文件 18](#_Toc1784764724)

[1． 写入文件 18](#_Toc489134972)

[2． 判断文件是否存在 18](#_Toc324073688)

[3． 读取文件 19](#_Toc687945424)

[二、易错点总结 20](#_Toc246785720)

[（一）mkdir:Permission denied 20](#_Toc936673683)

[（二） Error occurred during initialization of boot layer 20](#_Toc1619457671)

[（三） Provider class not in moudle 21](#_Toc1017334419)

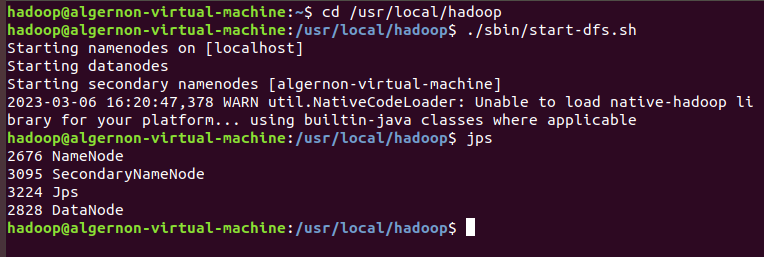
[三、 学习小记 22](#_Toc74782719)

[（一） vim与gedit的不同 22](#_Toc595224738)

# 一、实验过程

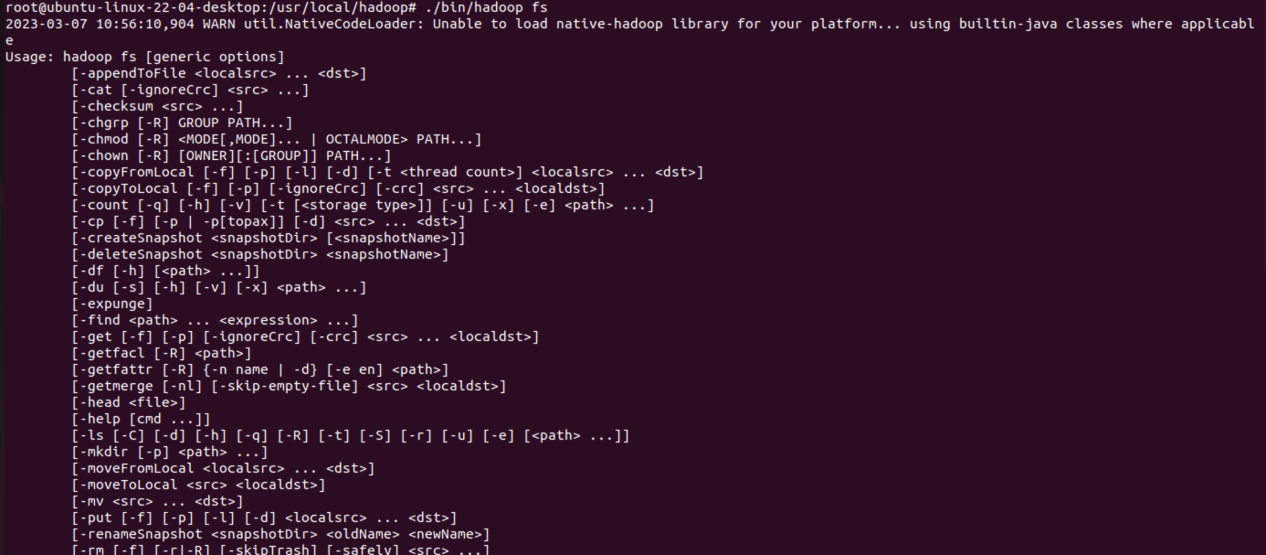
## （一）利用Shell命令与HDFS进行交互

在进行HDFS编程实践前，首先启动Hadoop

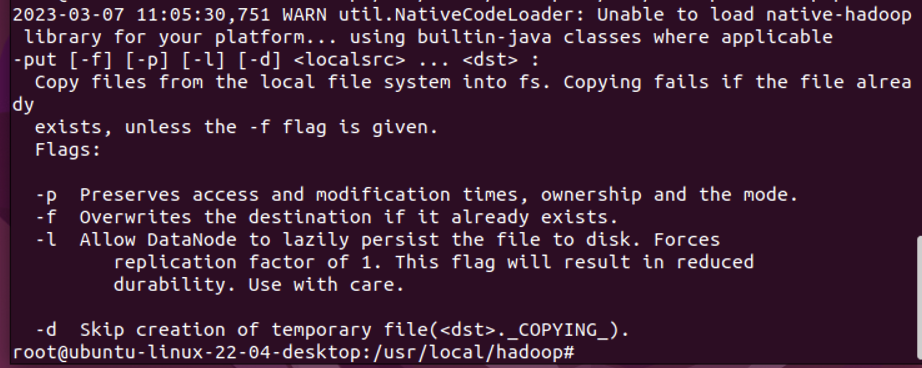


我们可以输入如下指令，看看fs总共支持哪些命令

截屏2023-03-07 19.01.51

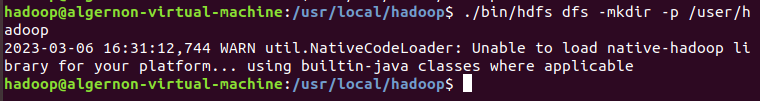


输入如下命令，可以查看具体某个命令的作用，下图演示的是查看put命令如何使用

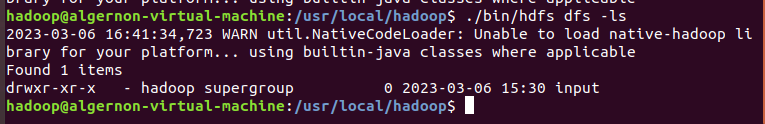
截屏2023-03-07 19.05.25

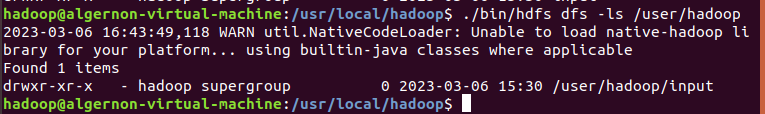
### 目录操作

第一次使用HDFS时，需要首先在HDFS中创建用户目录，该实验以hadoop用户登陆Linux系统，因此首先需要在HDFS中为hadoop创建一个用户目录

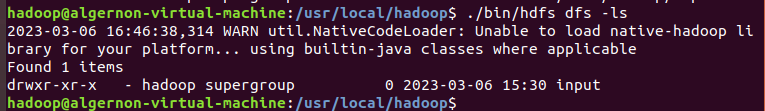


可用如下两条等价命令显示HDFS中与用户hadoop对应的用户目录下的内容

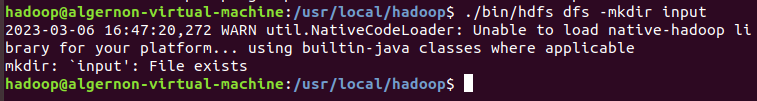




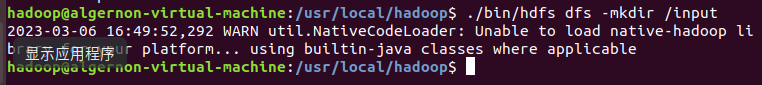
查看HDFS上的所有目录



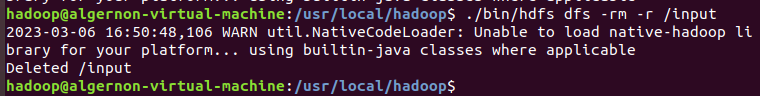
创建input目录



在HDFS根目录下创建一个名为input的目录

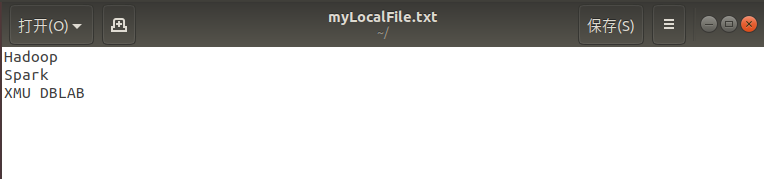


使用rm命令删除一个目录

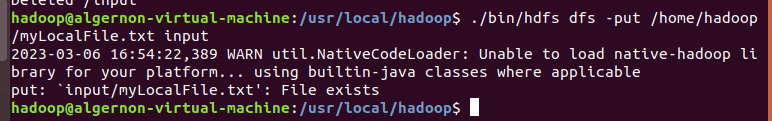


1. 文件操作

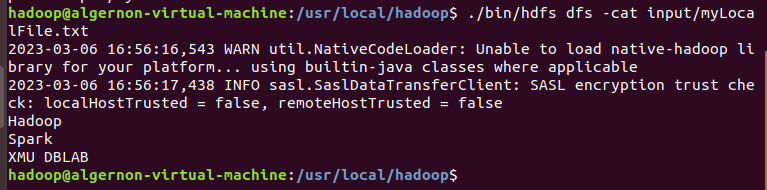
在/home/hadoop/目录下创建文件myLocalFile.txt,并随意输入一些单词



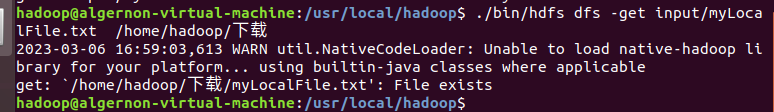
将本地文件系统的“/home/hadoop/myLocalFile.txt”上传到HDFS中的当前用户目录的input目录下，也就是上传到HDFS的“/user/hadoop/input/”目录下



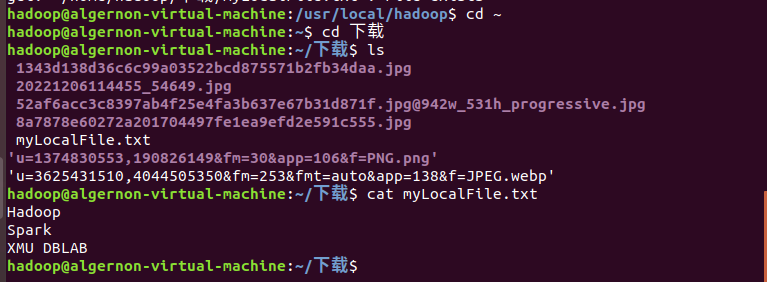
查看HDFS中myLocalFile.txt文件的内容



将HDFS中的myLocalFile.txt文件下载到本地文件系统中的“/home/hadoop/下载/”这个目录下

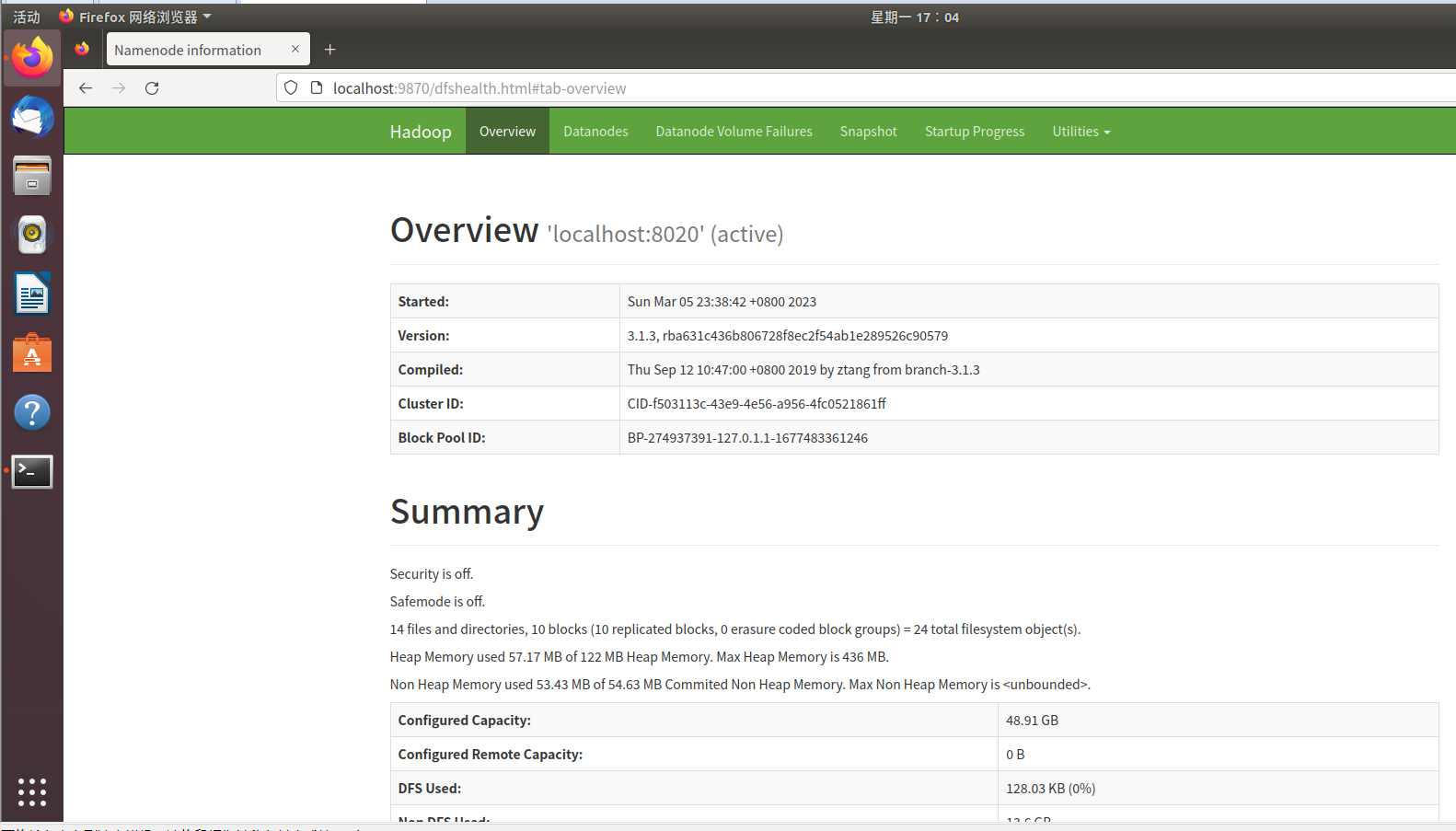


到本地文件系统来查看myLocalFile.txt



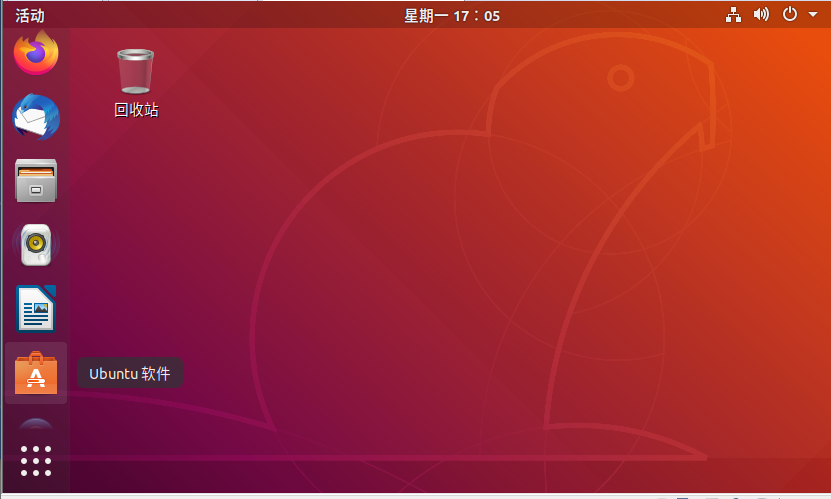
## 利用Web界面管理HDFS

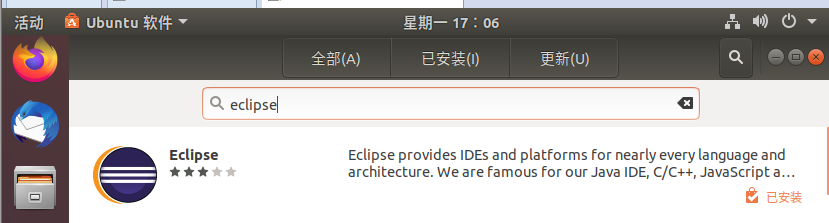
利用Firefox浏览器即可查看HDFS的web管理界面

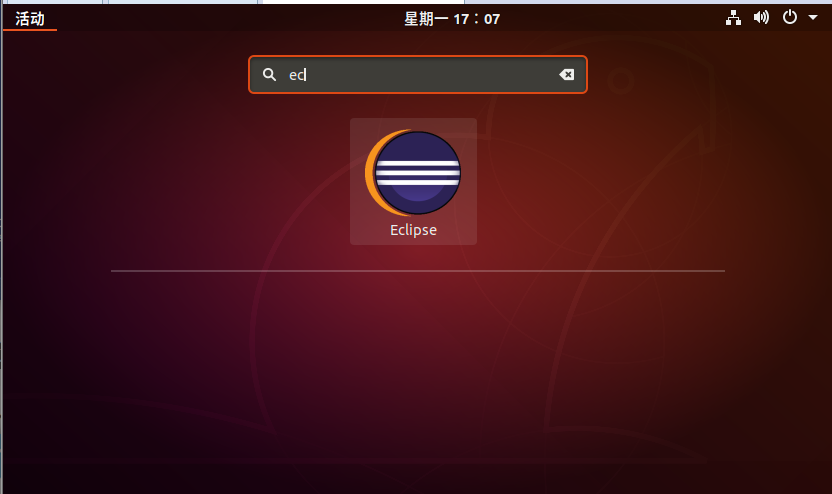


## 利用Java API与HDFS进行交互

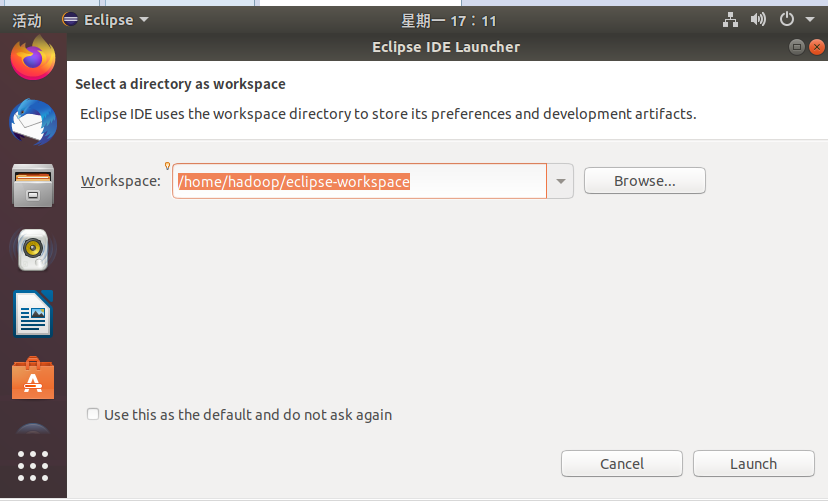
### 1．在Ubuntu中安装Eclipse

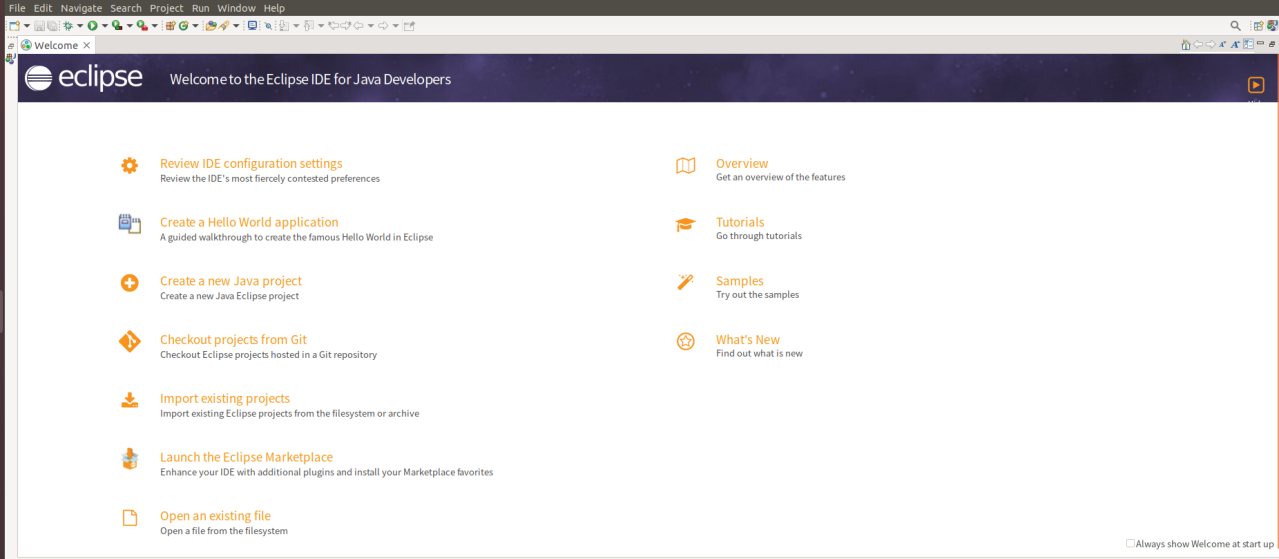


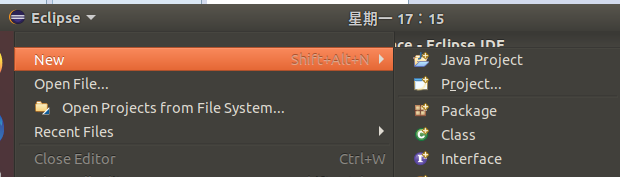


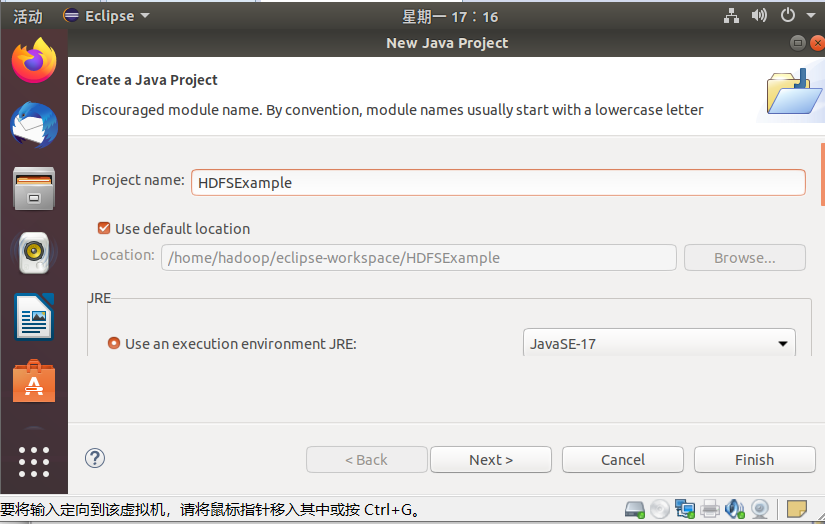


### 2．在Eclipse中创建项目



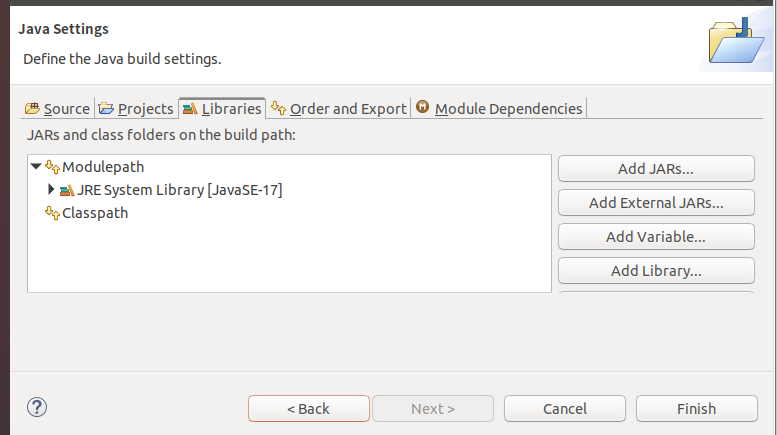


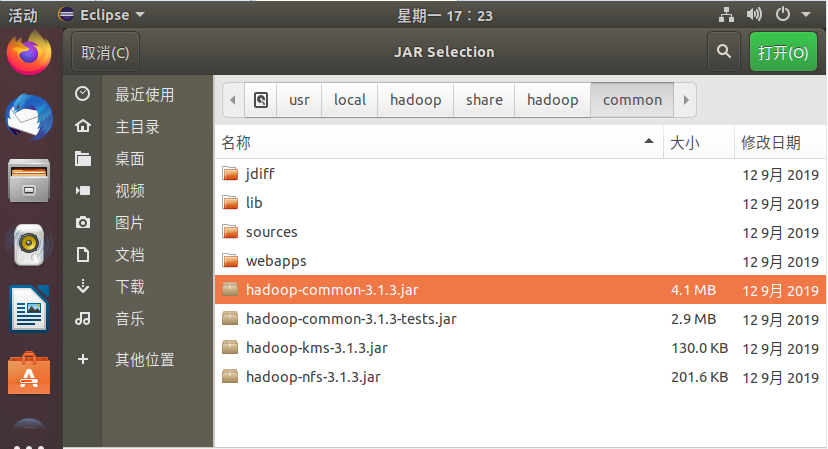


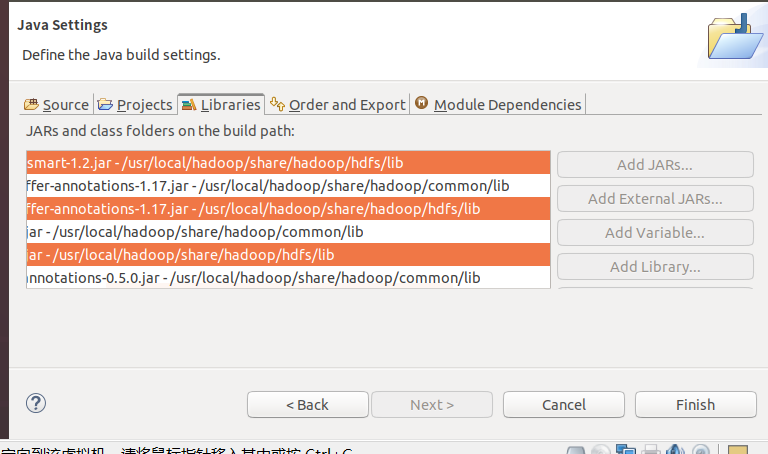


### 为项目添加需要用到的JAR包

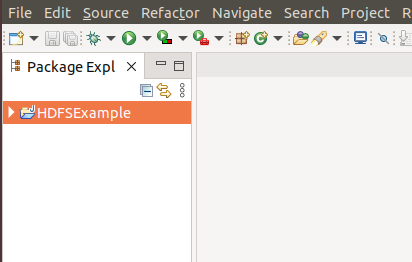
为了编写一个能够与HDFS交互的Java应用程序，一般需要向Java工程中添加以下JAR包：  
（1）"/usr/local/hadoop/share/hadoop/common”目录下的hadoop-common-2.7.1.jar和haoop-nfs-2.7.1.jar；  
（2）/usr/local/hadoop/share/hadoop/common/lib”目录下的所有JAR包；  
（3）“/usr/local/hadoop/share/hadoop/hdfs”目录下的haoop-hdfs-2.7.1.jar和haoop-hdfs-nfs-2.7.1.jar；  
（4）“/usr/local/hadoop/share/hadoop/hdfs/lib”目录下的所有JAR包。

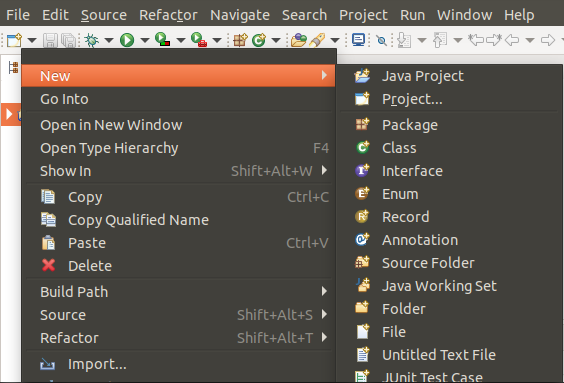


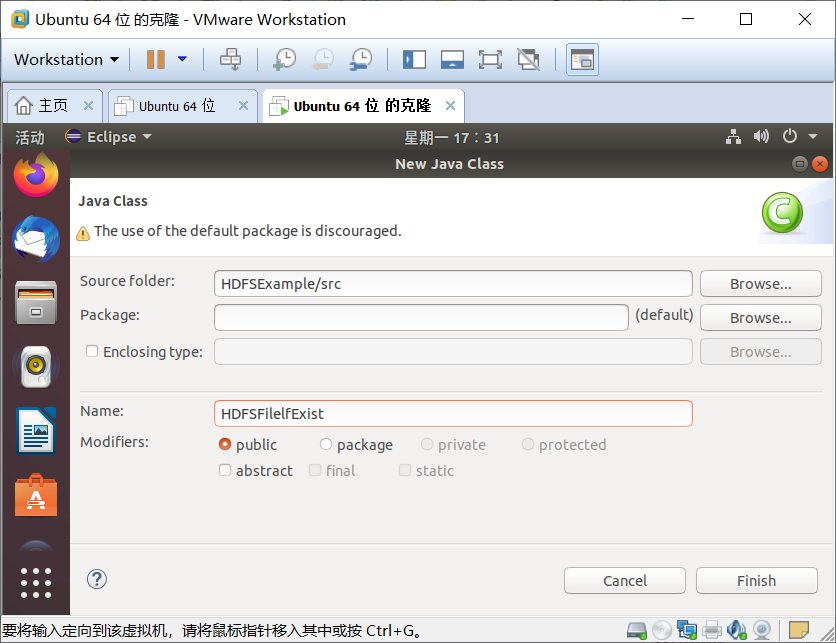


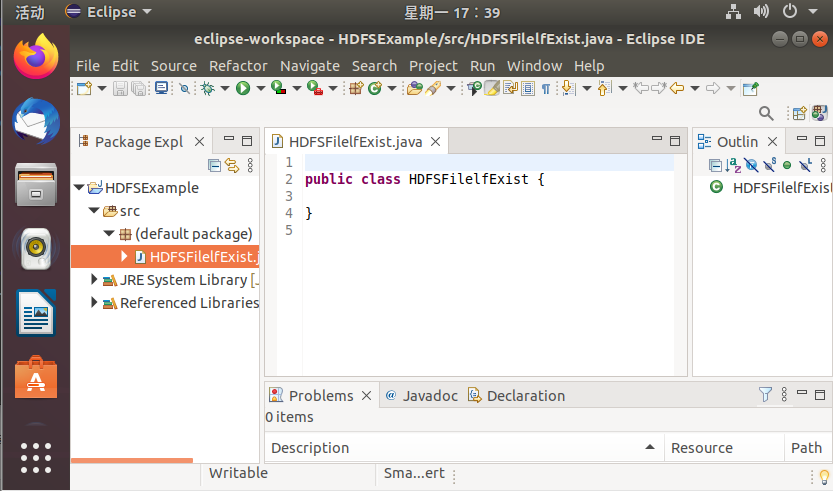


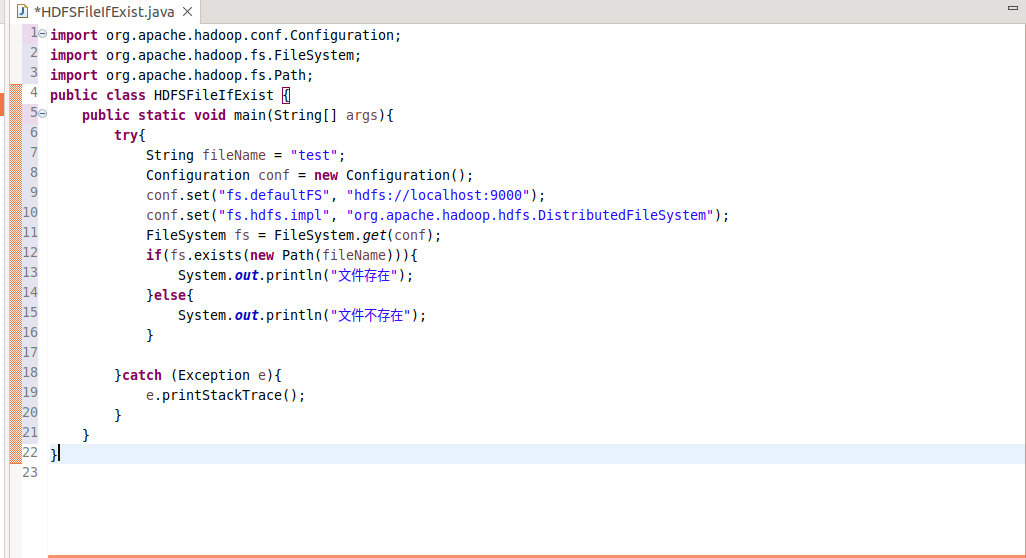
### 编写Java应用程序代码



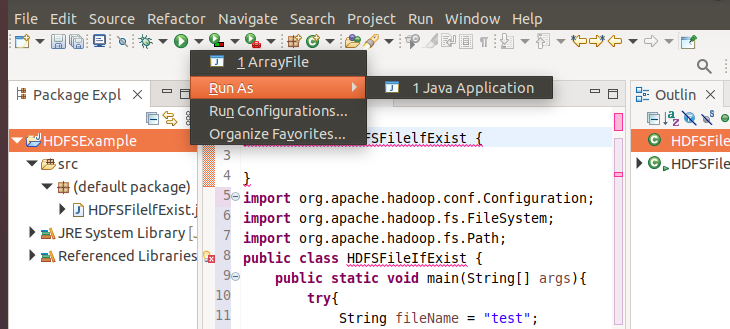


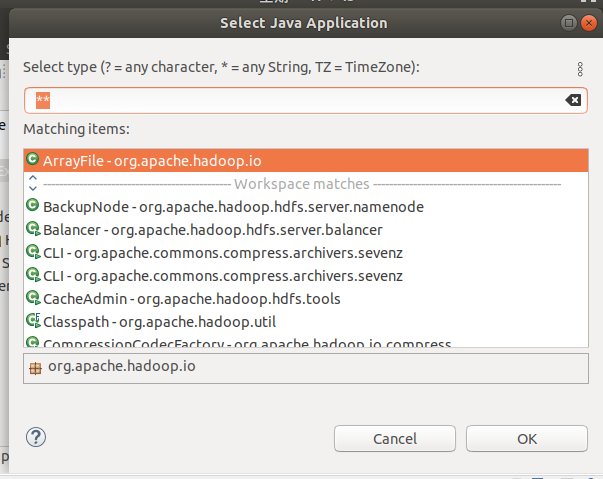


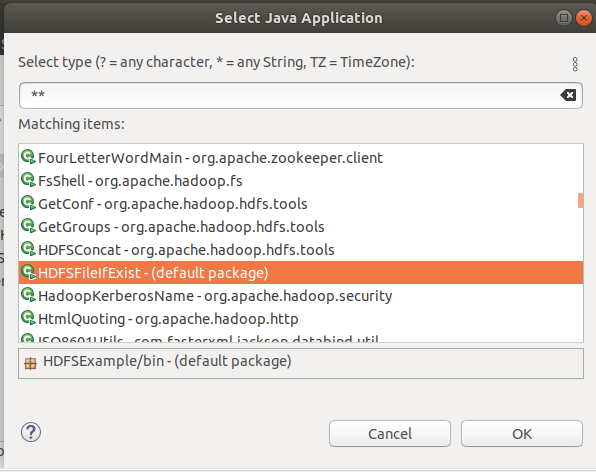


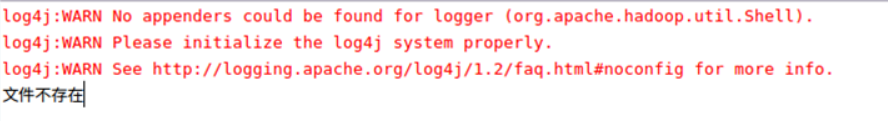


### 编译运行程序

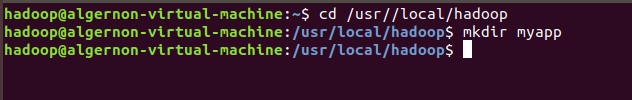


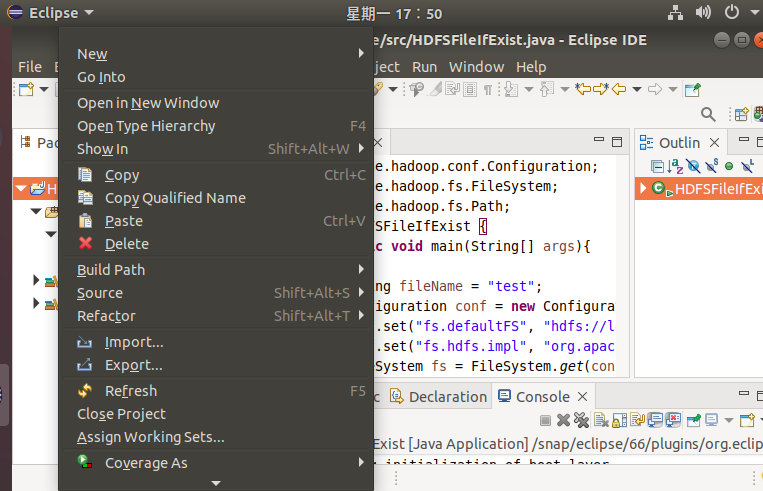


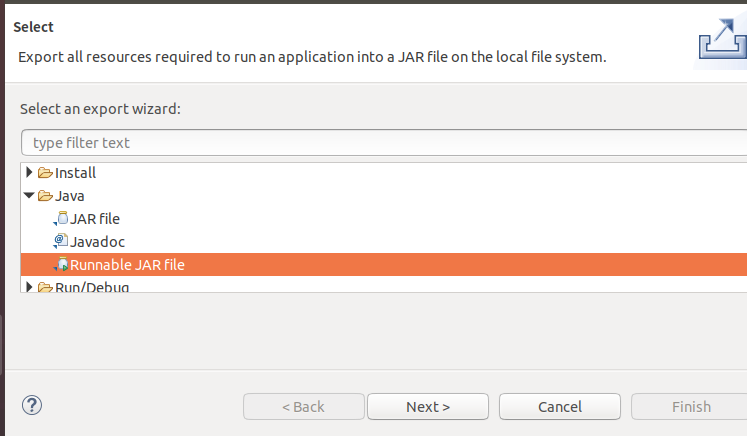


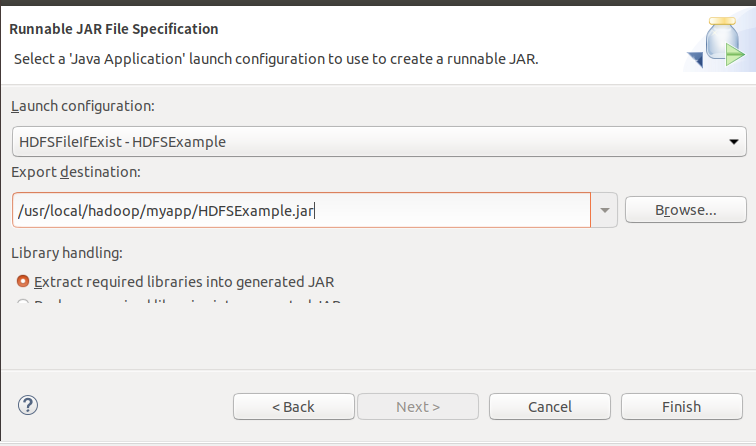


### 应用程序的部署

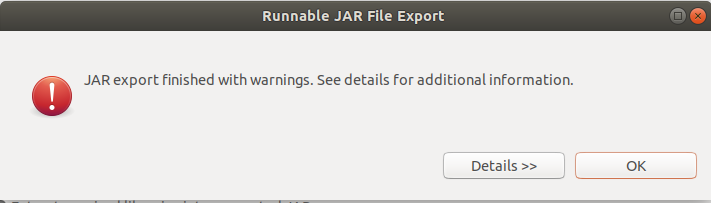






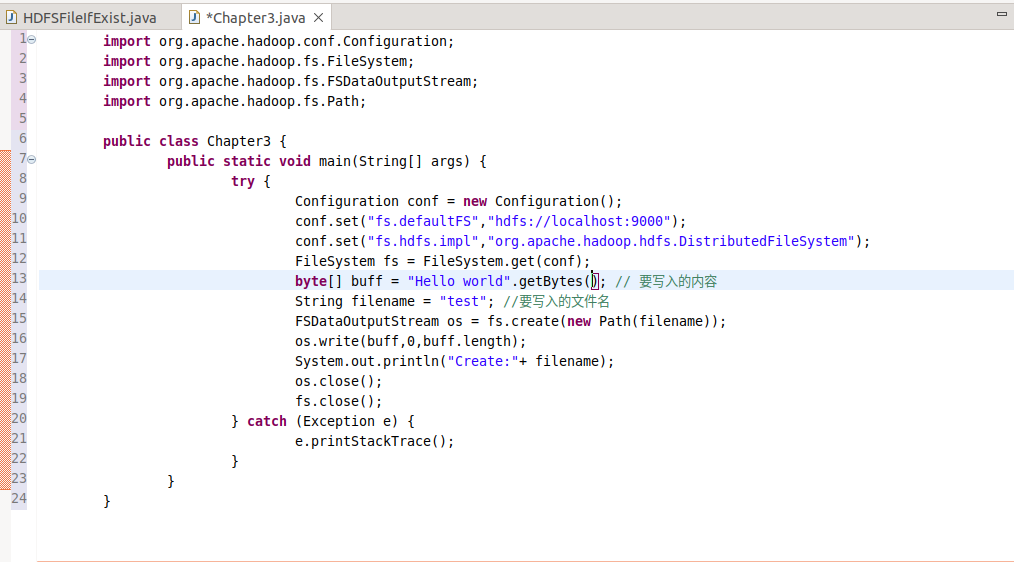




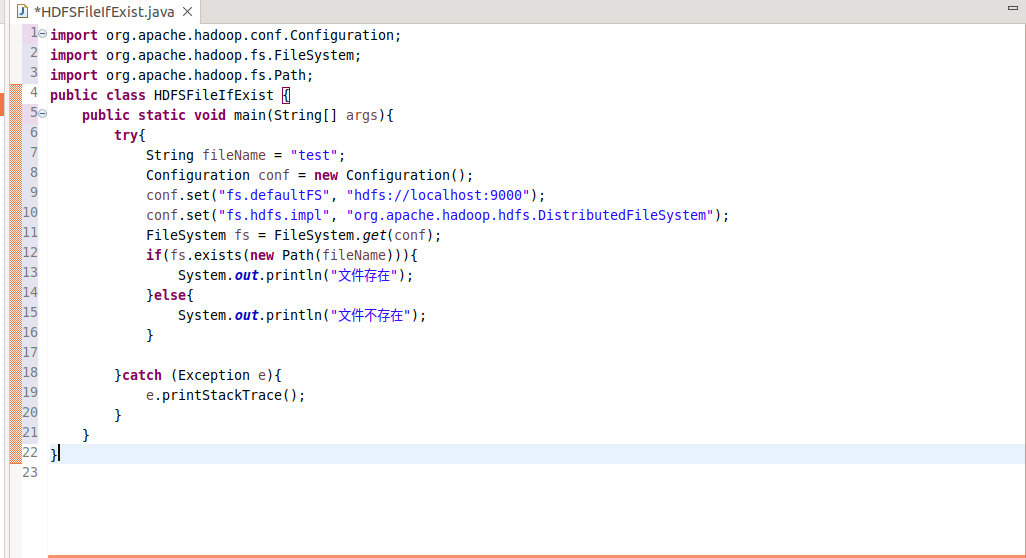


## 相关代码文件

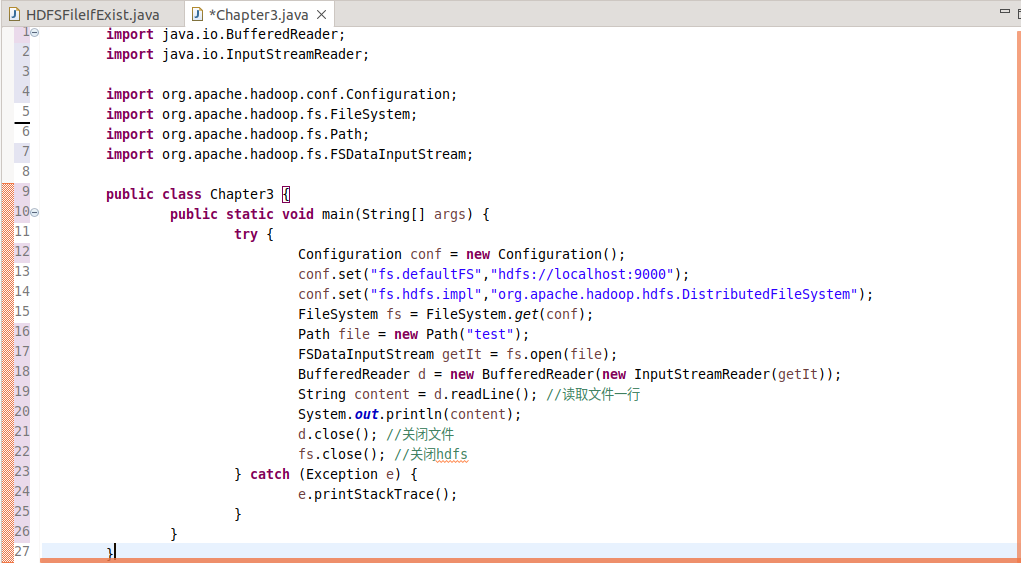
### 写入文件



### 判断文件是否存在



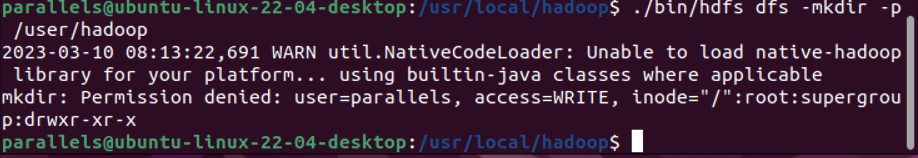
### 读取文件



# 二、易错点总结

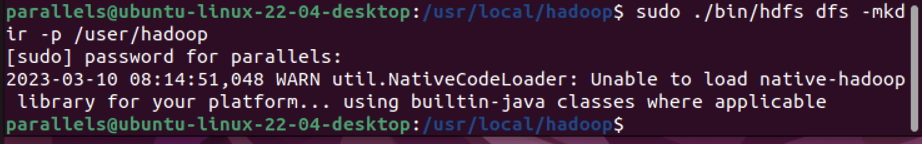
## （一）**mkdir:Permission denied**

在操作 linux 时,有时会遇到 ****mkdir(): Permission denied**** 错误 , 这是操作权限不够导致的



**解决方案：**

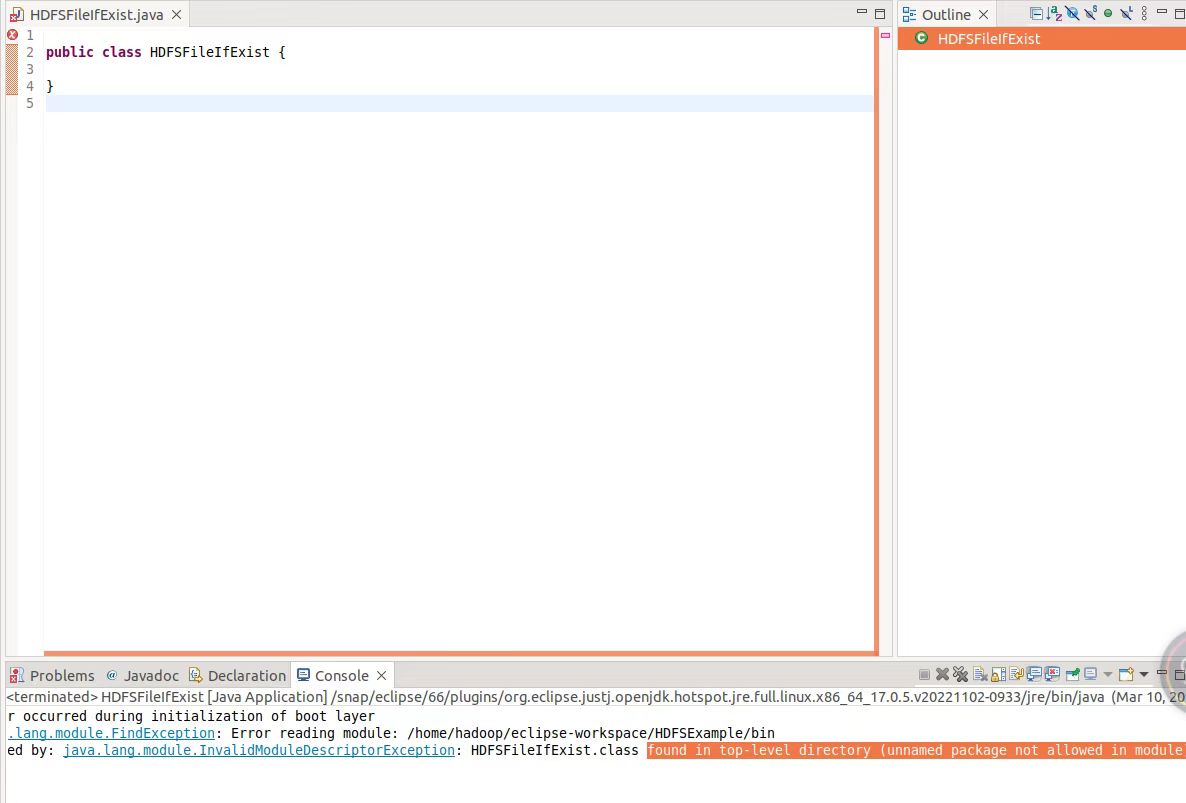
1. 使用sudo命令，以root身份来创建文件夹



1. 修改文件的操作权限，通过命令 sudo chmod -R 777 实现

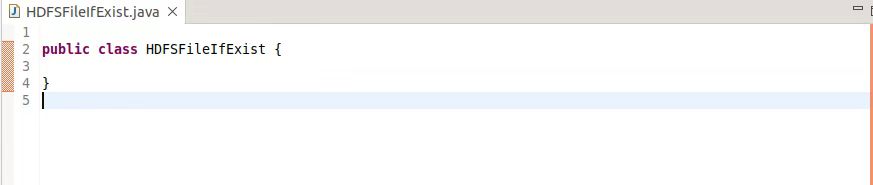
## **Error occurred during initialization of boot layer**

在使用eclipse创建源代码文件HDFSFileIfExist.java时，在文件刚刚建好便出现了醒目的红叉，并提示：error occurred during initialization of boot layer; HDFSFileIfExist.class found in top level directory(unnamed package not allowed in module)。



**解决方案：**

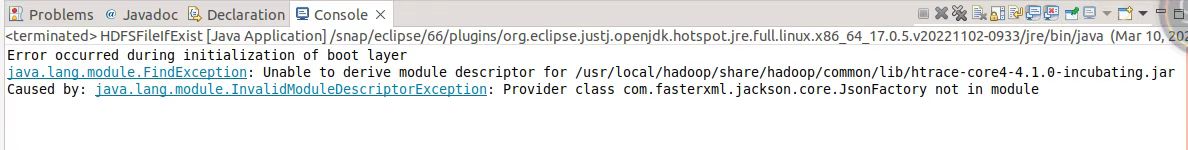
原因是由于JDK9及以上版本引入了模块，（而在JDK8中就没有这样的问题）所以在default [package](https://so.csdn.net/so/search?q=package&spm=1001.2101.3001.7020" \t "/Users/yangchaoran/Documents\\x/_blank)建立一个单独运行的类，就无法通过编辑。需要先删除module-info.java，删除之后，再运行就可以了。



删除module.info.java后，文件不再出现报错提醒。

## **Provider class not in moudle**

在编译运行程序的过程中遇到如下报错：



**解决方案：**

该报错产生的原因是导入jar包时出现重复导入相同jar包的情况，将重复的jar包remove即可解决

# 学习小记

## vim与gedit的不同

在实验二三的操作过程种，都有涉及到用vim和gedit两种方式修改代码的方法，那么两者有何差别呢？

vim：不能随意定位光标，编辑是需要按i键，退出编辑esc键，保存需要先按esc键，在打出:wq（w是保存，q是退出）。

复制：Ctrl+shift+C

粘贴是Ctrl+shift+V

gedit：可以用法与window文本编辑一样。

复制：Ctrl+C

粘贴：Ctrl+V

保存：Ctrl+S