

电 子 科 技 大 学

实 验 报 告

学生姓名：Lolipop 学号：2018091202000 指导教师：xx

实验地点：信软学院楼西 400 实验时间：2020.11.18

一、实验名称：Eclipse 编译 MapReduce 程序

二、实验学时：4 学时

三、实验目的：

1. 掌握用 Eclipse 和 Hadoop 开发应用基础；
2. 熟悉 Eclipse 和 hadoop 编程基本操作。

四、实验原理：

Eclipse 是著名的跨平台开源集成开发环境（IDE）。最初主要用来 Java 语言开发，目前亦有人通过插件使其作为 C++、Python、PHP 等其他语言的开发工具。

Eclipse 的本身只是一个框架平台，但是众多插件的支持，使得 Eclipse 拥有较佳的灵活性，所以许多软件开发商以 Eclipse 为框架开发自己的 IDE。

Eclipse 的插件机制是轻型软件组件化架构。在富客户机平台上，Eclipse 使用插件来提供所有的附加功能，例如支持 Java 以外的其他语言。已有的分离的插件已经能够支持 C/C++（CDT）、PHP、Perl、Ruby，Python、telnet 和数据库开发。插件架构能够支持将任意的扩展加入到现有环境中，例如配置管理，而决不仅仅限于支持各种编程语言。

利用 hadoop-eclipse-plugin 插件可以实现在 Eclipse 上进行 Hadoop 开发。

五、实验内容：

使用 Eclipse 配置 Hadoop 开发环境，并在伪分布式环境中执行代码。在此基础上熟悉 Eclipse 和 hadoop 编程基本操作。

六、实验器材（设备、元器件）：

1. Ubuntu 20.04
2. jdk 1.8
3. hadoop 2.10.1
4. Eclipse Luna 4.4

七、实验步骤：

1. 安装 Eclipse
2. 安装并配置 hadoop-eclipse-plugin
3. 查看 HDFS 中的文件内容
4. 在 Eclipse 中创建 MapReduce 项目
5. 通过 Eclipse 运行 MapReduce 程序

八、实验结果与分析（含重要数据结果分析或核心代码流程分析）

1. 安装 Eclipse

- a) 解压缩 Eclipse 文件。
- b) 创建 Eclipse 快捷图标。编辑文件如图 1-1 所示。

```
cd /usr/share/applications  
sudo gedit eclipse.desktop
```

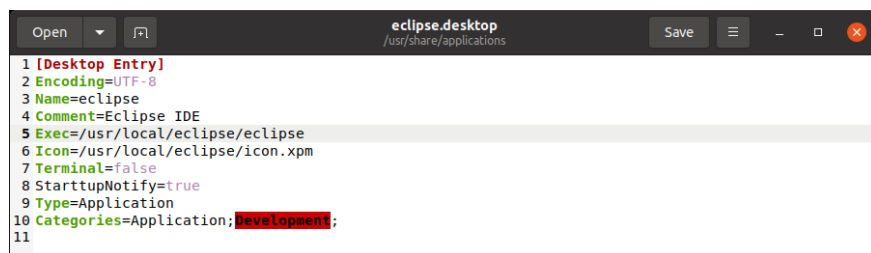


图 1-1 编辑 eclipse.desktop

- c) 结果如图 1-2 所示。

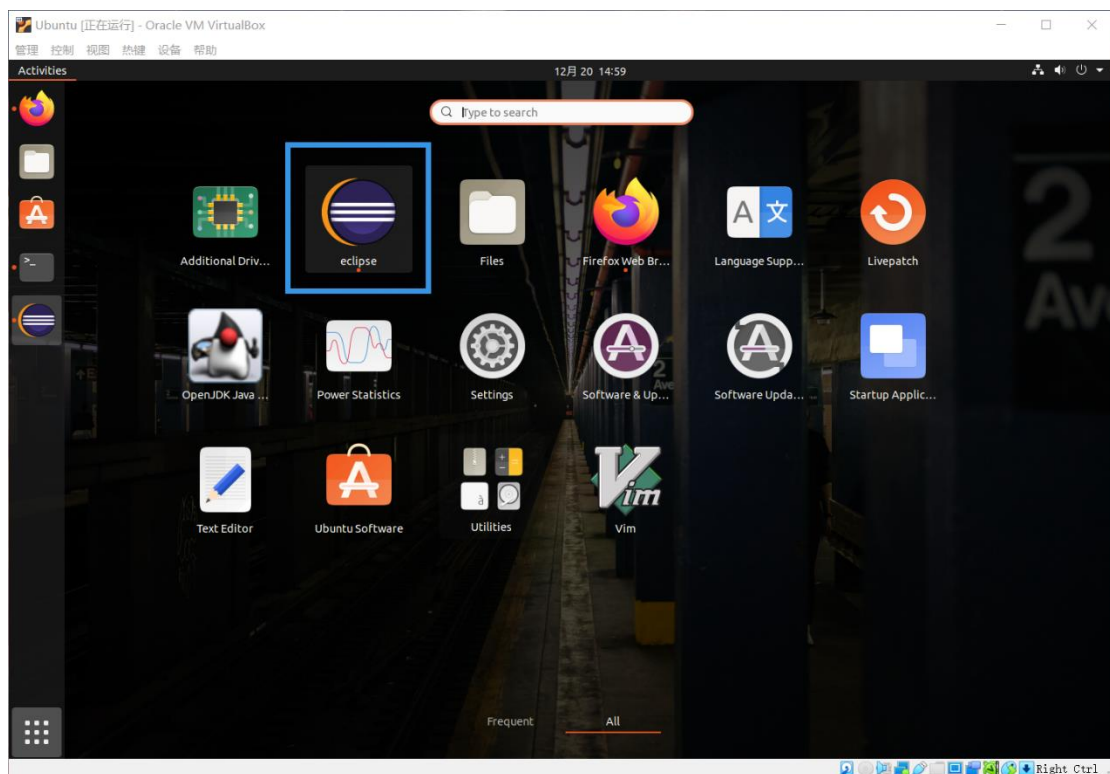


图 1-2 Eclipse 快捷方式

2. 安装并配置 hadoop-eclipse-plugin

- a) 拷贝 hadoop-eclipse-plugin-2.7.3.jar 到 Eclipse/plugins 目录下。如图 2-1 所示。

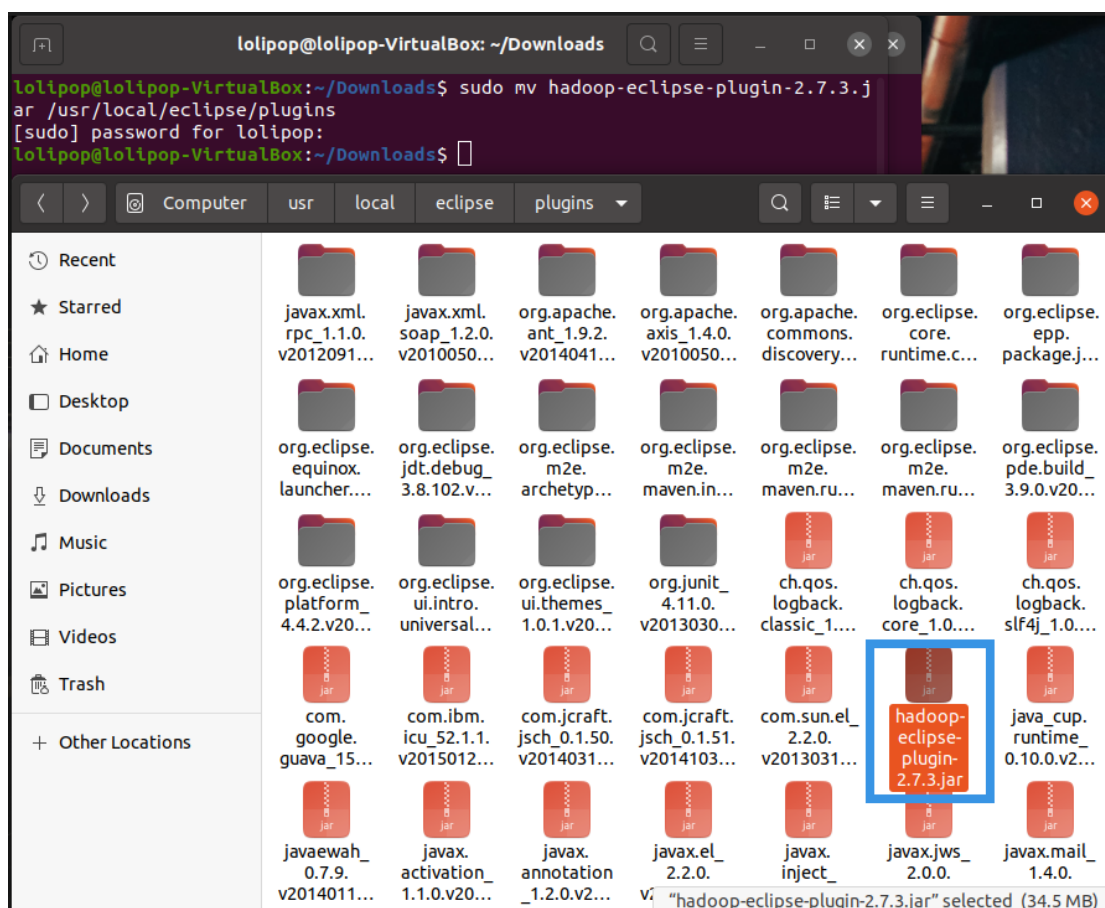


图 2-1 安装插件

- b) 启动 Eclipse 后就可以在左侧的 Project Explorer 中看到 DFS Locations。
- c) 选择 Window 菜单下的 Preference。此时会弹出一个窗体，窗体的左侧会多出 Hadoop Map/Reduce 选项，点击此选项，选择 Hadoop 的安装目录。
- d) 切换 Map/Reduce 工作目录，选择 Window 菜单下选择 Open Perspective -> Other，弹出一个窗体，从中选择 Map/Reduce 选项即可进

行切换。

- e) 建立与 Hadoop 集群的连接, 点击 Eclipse 软件右下角的 Map/Reduce Locations 面板, 在面板中单击右键, 选择 New Hadoop Location。
- f) 在弹出来的 General 选项面板中进行 Master 的设置, 设置要要 Hadoop 的配置一致, 如设置了 fs.defaultFS 为 hdfs://localhost:9000, 则 DFS Master 的 Port 也应改为 9000。如图 2-2 所示。

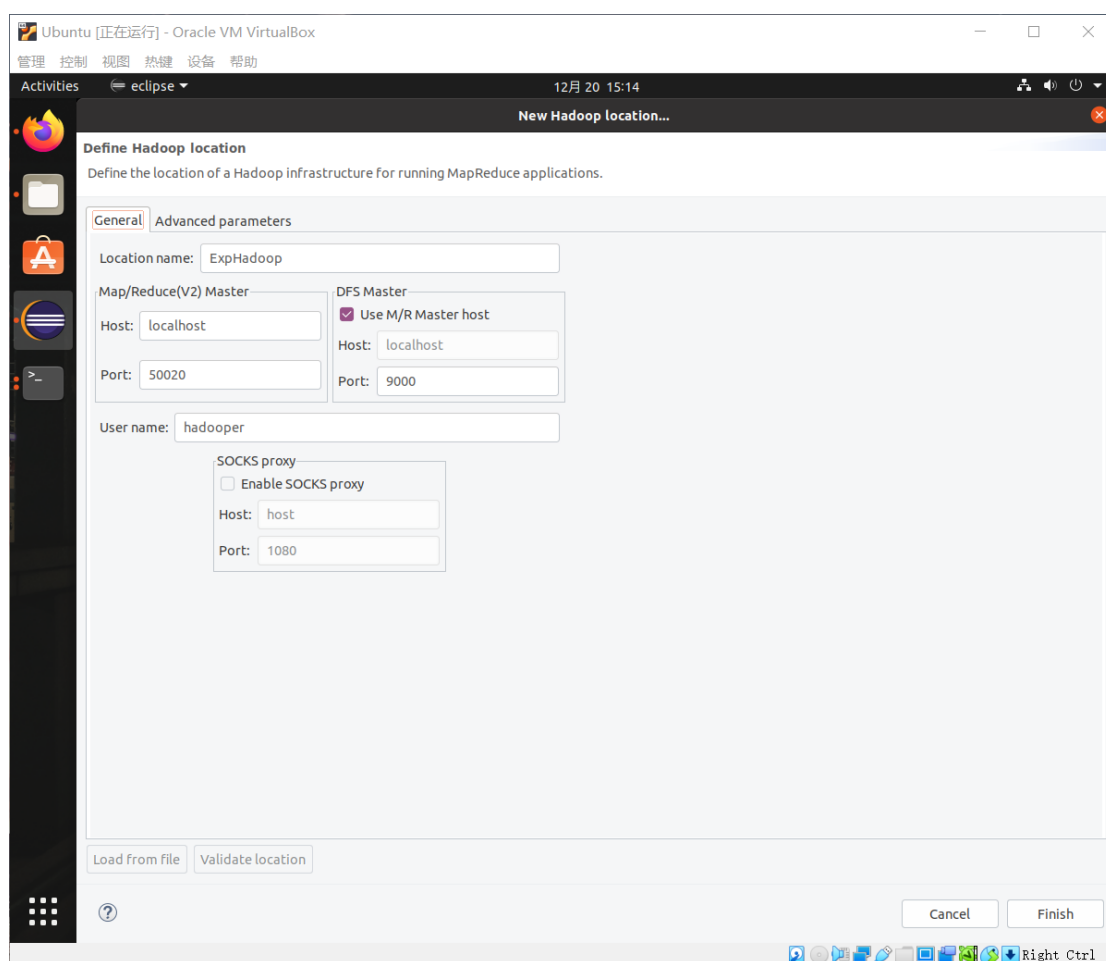


图 2-2 配置 Hadoop Location

g) 接着再切换到 Advanced parameters 选项面板, 配置 `hadoop.tmp.dir` 项。

如图 2-3 所示。

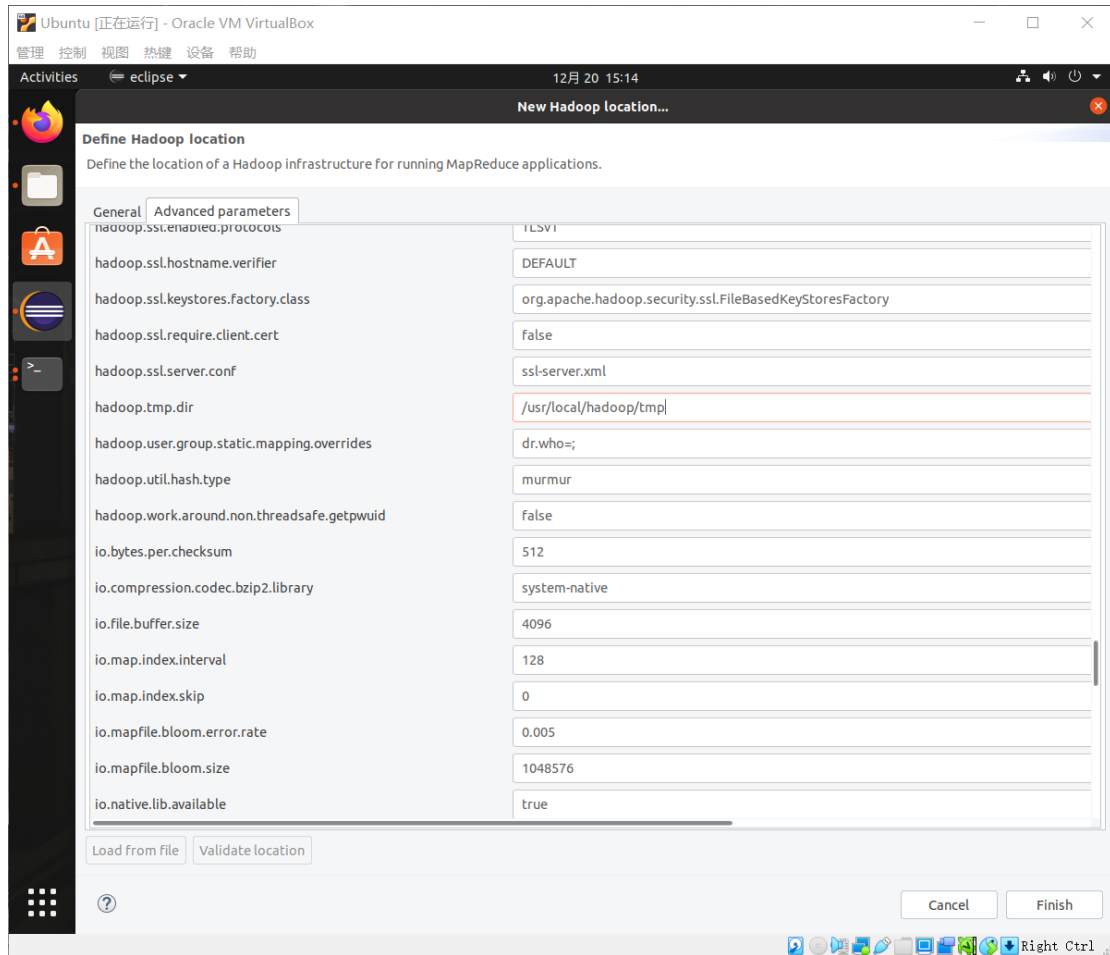


图 2-3 详细配置 Hadoop Location

h) 配置完成后点击 finish, Map/Reduce Location 就创建好了。

3. 查看 HDFS 中的文件内容

配置成功后，点击左侧 Project Explorer 中的 MapReduce Location 就能直接查看 HDFS 中的文件内容了，而无需再通过繁琐的 `hdfs dfs -ls` 命令。如图 3 所示。

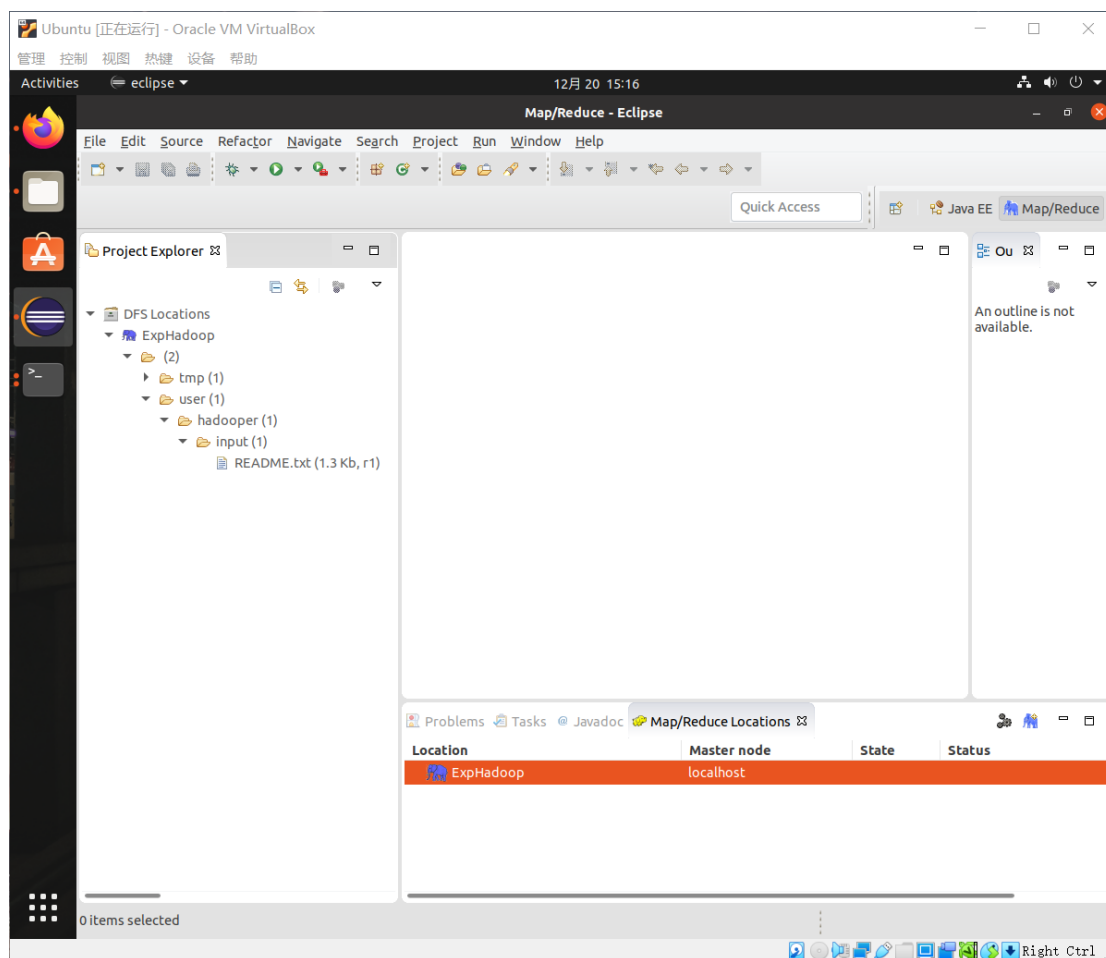


图 3 查看 HDFS 中的文件内容

4. 在 Eclipse 中创建 MapReduce 项目

- a) 点击 File 菜单，选择 New -> Project。
- b) 选择 Map/Reduce Project，点击 Next。
- c) 填写 Project name 为 WordCount 即可，点击 Finish 就创建好了项目。

如图 4-1 所示。

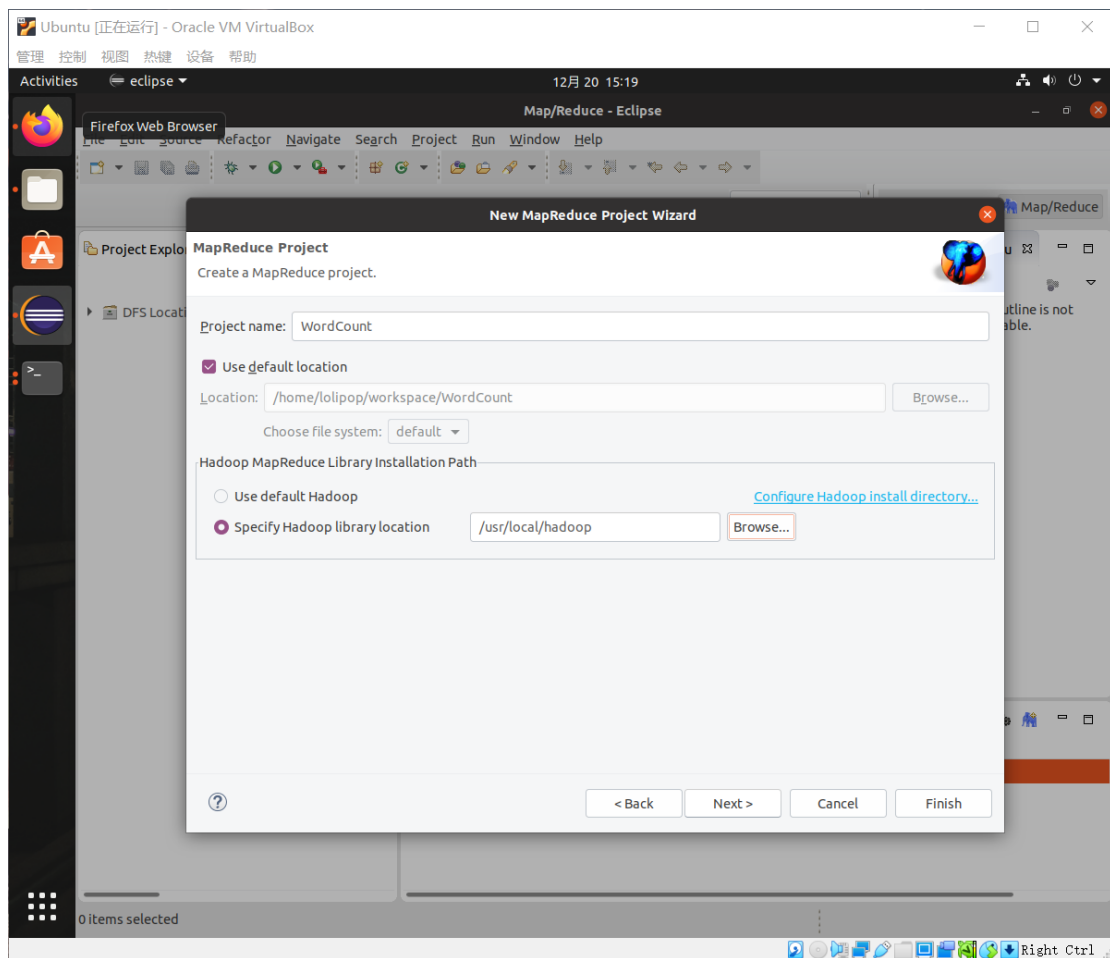


图 4-1 创建新的项目

- d) 接着右键点击刚创建的 WordCount 项目，选择 New -> Class。在 Package 处填写 org.apache.hadoop.examples；在 Name 处填写 WordCount。如图 4-2 所示。

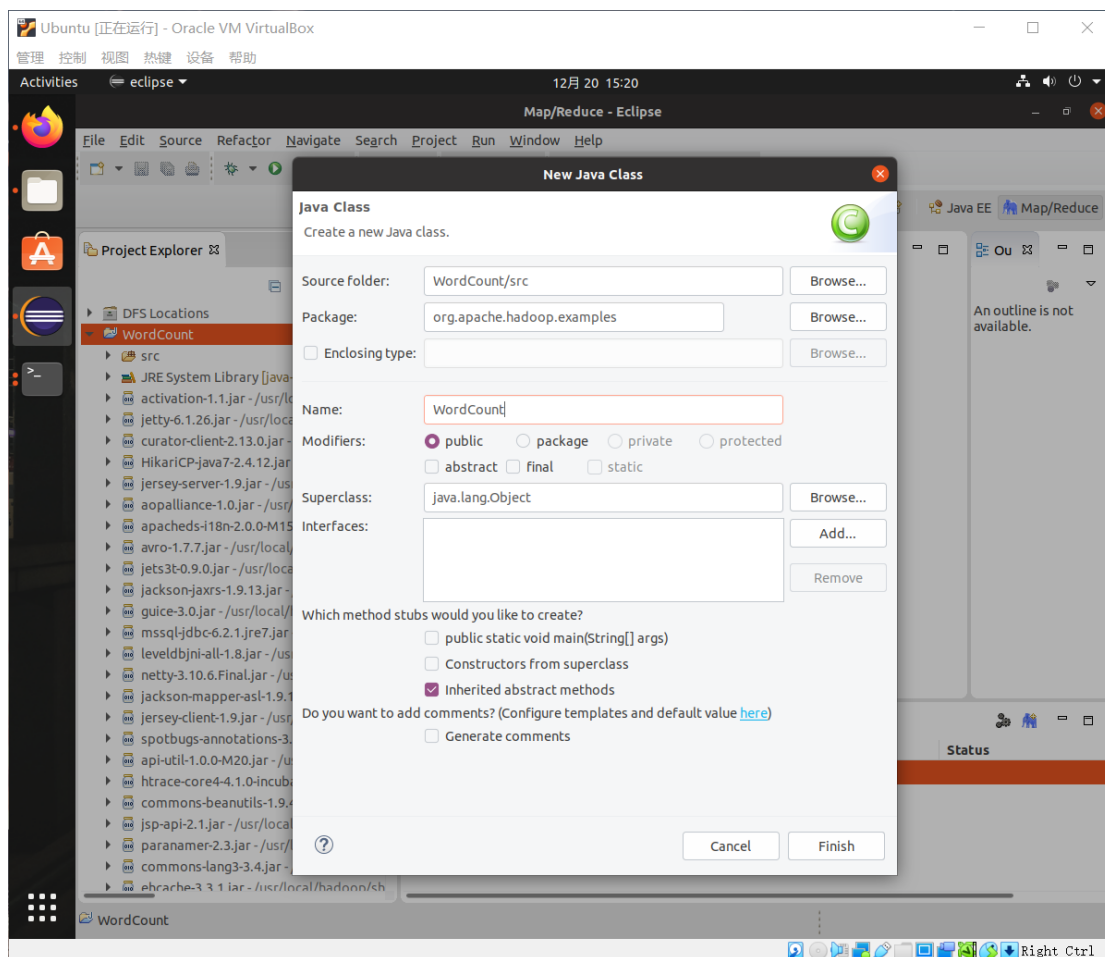


图 4-2 创建 WordCount 测试类

e) 仿照实验指导书给的示例程序编写 Java 文件，如图 4-3 所示。

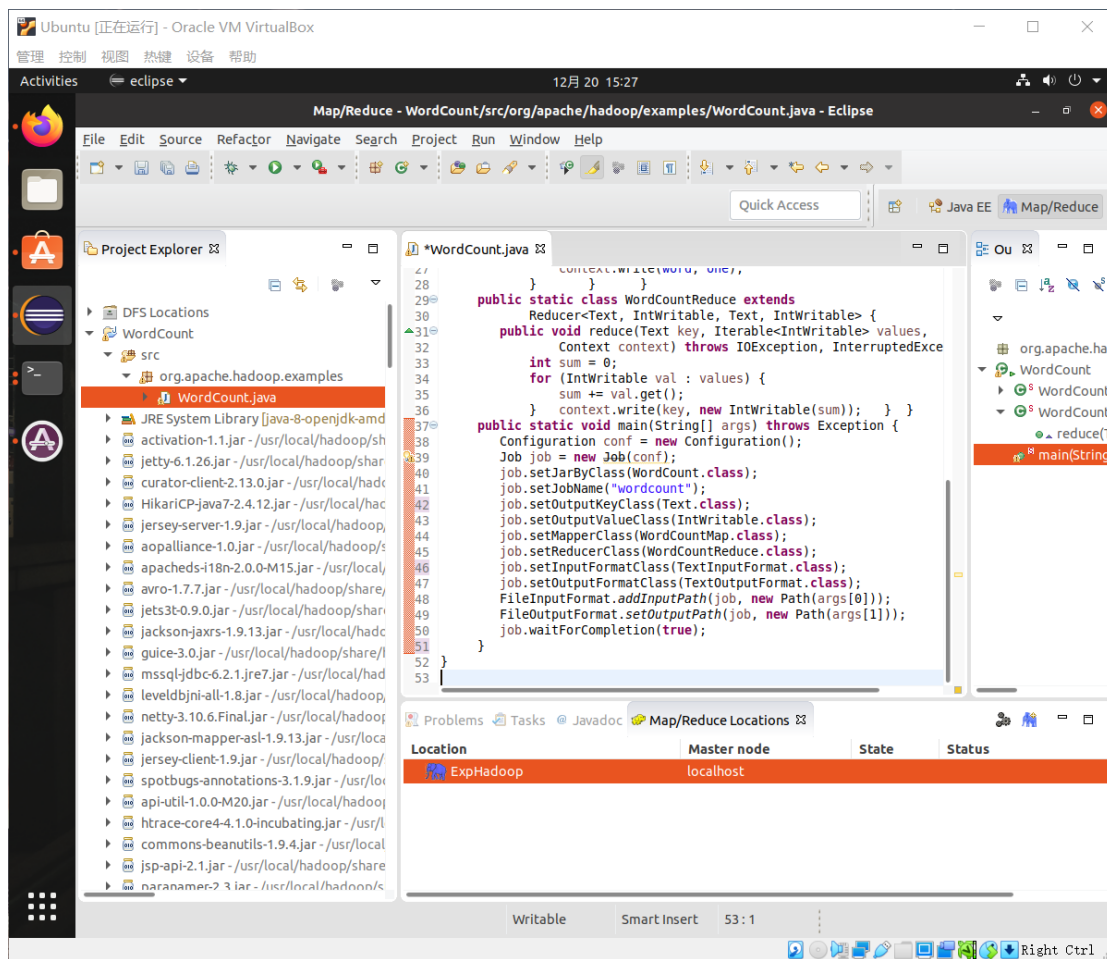


图 4-3 编写 WordCount 代码

5. 通过 Eclipse 运行 MapReduce 程序

- a) 将/usr/local/hadoop/etc/hadoop 中的 log4j.properties 复制到 WordCount 项目下的 src 文件夹中。如图 5-1 所示。

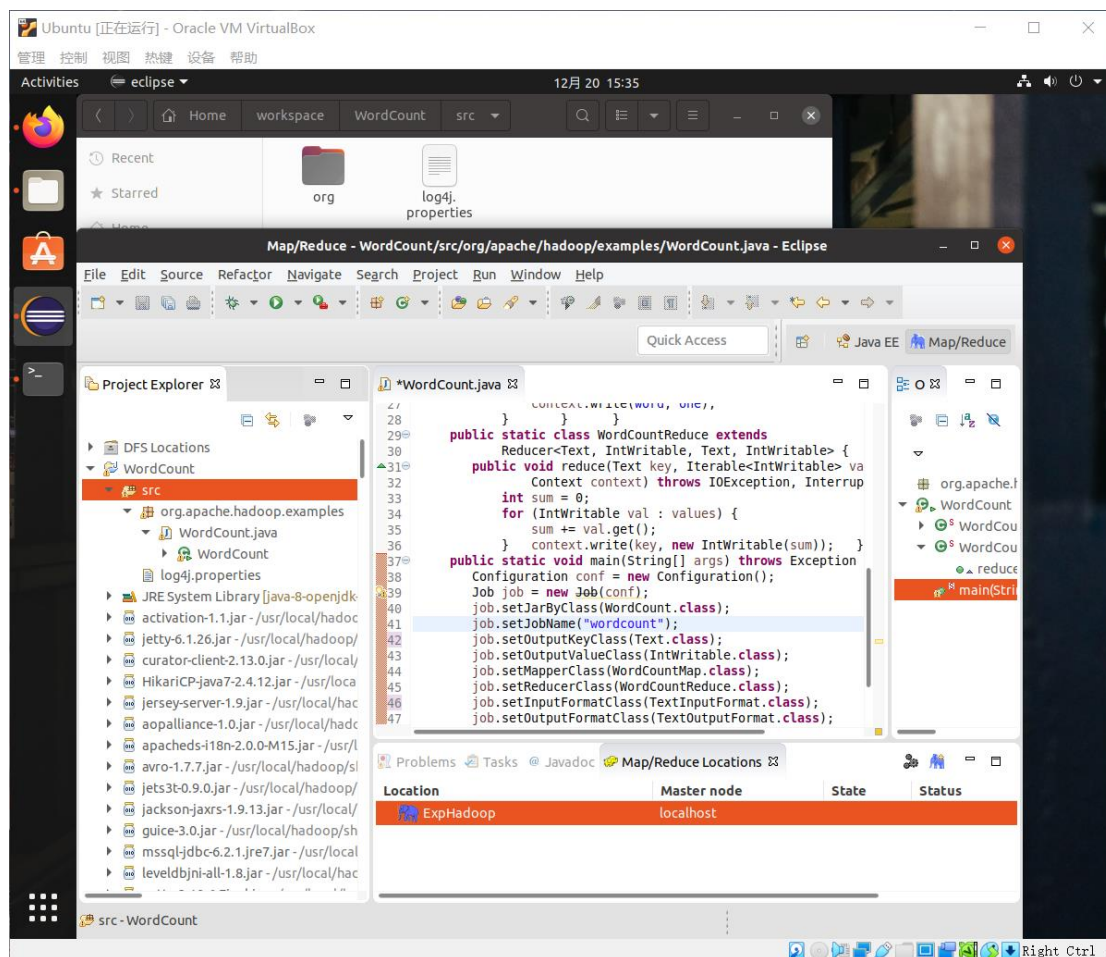


图 5-1 复制依赖文件 log4j.properties

b) 选择 Run As -> Run Configurations, 在此处可以设置运行时的相关参数。

切换到“Arguments”栏, 在 Program arguments 处填写参数信息。如图

5-2 所示

```
hdfs://localhost:9000/user/hadooper/input  
hdfs://localhost:9000/user/hadooper/output
```

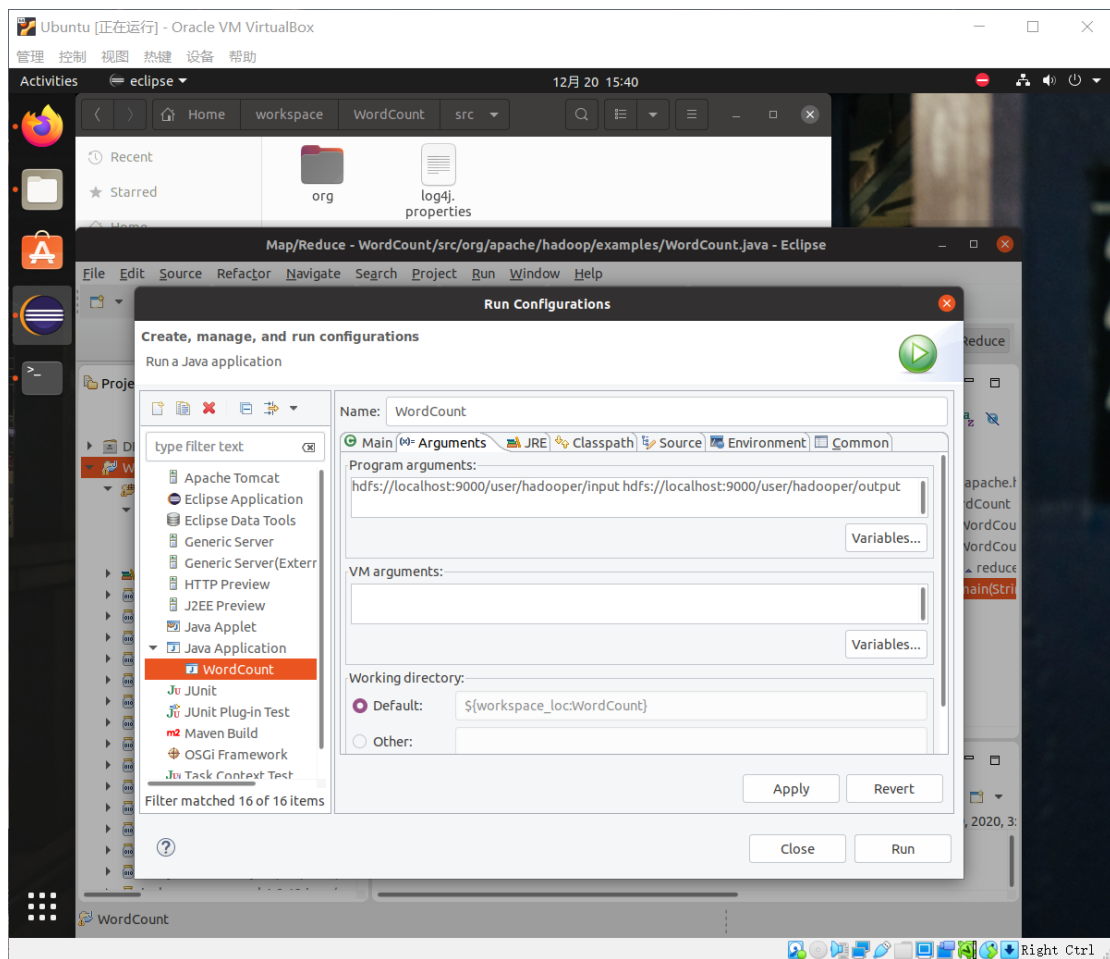


图 5-2 配置运行参数

- c) 将提供的 The_Notebooks_of_Leonardo_Da_Vinci.txt 和 Ulysses.txt 文件传输到伪分布式文件系统中的/user/hadooper/input 目录下，如图 5-3 所示。

```
hadooper@lolipop-VirtualBox:/usr/local/hadoop$ bin/hdfs dfs -put input/*.txt input
hadooper@lolipop-VirtualBox:/usr/local/hadoop$ bin/hdfs dfs -ls input
Found 2 items
-rw-r--r-- 1 hadoop supergroup 2792076 2020-12-20 22:27 input/The_Notebooks_of_Leonardo_Da_Vinci.txt
-rw-r--r-- 1 hadoop supergroup 2834481 2020-12-20 22:27 input/Ulysses.txt
hadooper@lolipop-VirtualBox:/usr/local/hadoop$
```

图 5-3 传输测试文件

- d) 执行 WordCount 程序，在伪分布式文件系统中的/user/hadooper/output 目录下生成两个结果文件。执行的结果如图 5-4 所示。

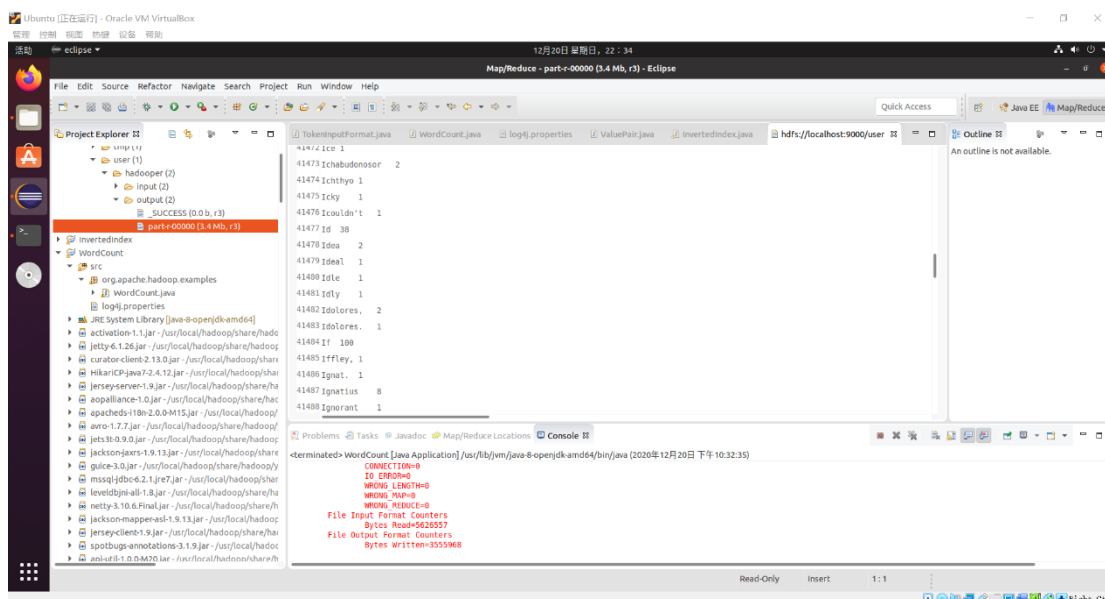


图 5-4 程序执行结果

- e) 输出伪分布式文件系统中的部分结果如下所示。

A.	33
A.,	5
A.;	1
A.D.)	1

A.D.,	1
A.E.I.O.U.	1
A.M.	1
ADONAI	2
ALBERTA	1
ALEXANDER	2
ALF	1
ALL	4
ALONG	1
AM.	1
AN	3
AND	10
.....	
soft	45
soft!	1
soft,	1
soft.	1
softcreakfooted,	1
soften	1
softened	1
softens	1
softens.)	1
softer	4
softer,	1
softer.	1
softlier,	1
softly	13
softly,	7
softly.	5

softly.) 3

softlyfeatured 1

softness 1

softnosed 1

softy 1

softy. 1

soi-disant 2

soil 5

soil, 3

soil. 2

soiled 9

soiled, 1

soiling 1

soirée, 1

sojourn 1

solace 1

solar 5

sold 15

soldier 8

soldier. 1

soldiers 13

.....

within 36

within, 2

within. 2

without 131

without, 1

without. 1

withouten 1

withsay. 2

withstand, 1

withstood1

witness 7

witness), 1

witness, 1

witness. 4

witnessbox 2

witnessbox, 1

witnessed2

witnesses 2

witnesses' 1

witnesses. 1

witnessing 2

wits 5

.....

九、总结及心得体会：

// removed

十、对本实验过程及方法、手段的改进建议：

// removed

报告评分：

指导教师签字：