计算机与信息工程学院实验报告

姓名： 苗智军 学号： 2012080014 专业： 数据科学与大数据技术 年级： 2020

课程： 大数据平台技术 主讲教师： 刘春 辅导教师： 刘春

实验时间： 2022 年 3月 2 日 中 午 10 时至 12:30时，实验地点 计算机大楼201

实验题目： 开发与运行环境安装与介绍

实验目的： 本次实验主要为hadoop相关环境的安装，如虚拟机系统Linux，必须的jdk，和IDE idea以及依赖管理工具maven

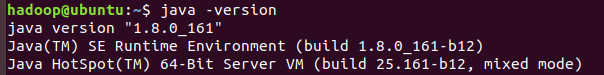
的安装 。

实验环境（硬件和软件）：

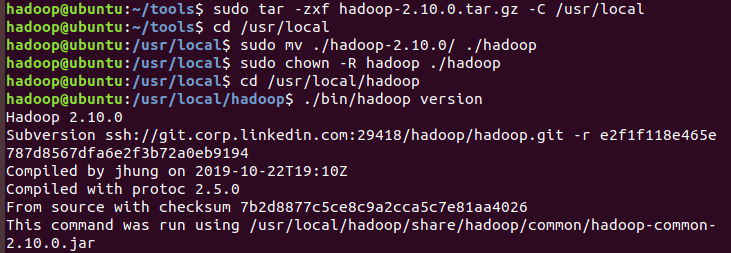
Linux版本：



Jdk版本：



Hadoop版本：



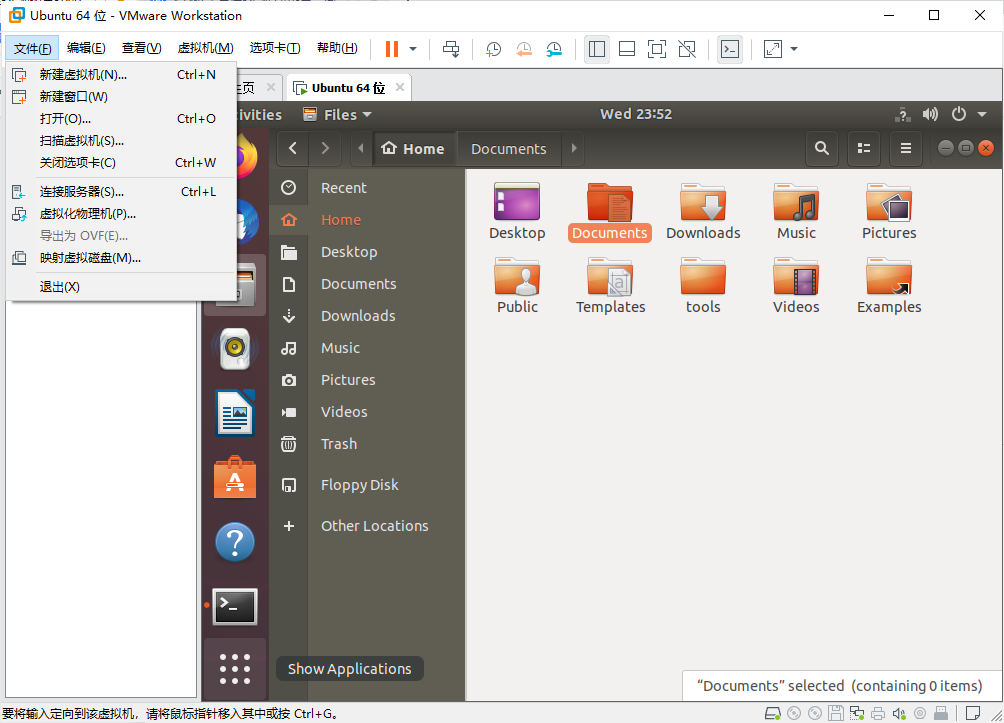
### 2.1虚拟机的创建

#### 2.1.1 虚拟化软件的安装



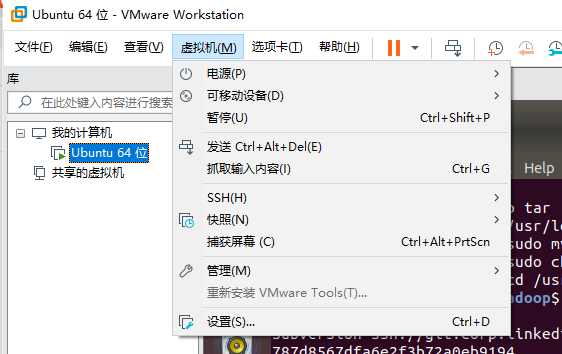
下载群文件里的VMware 解压安装

#### 2.1.2 虚拟机的创建



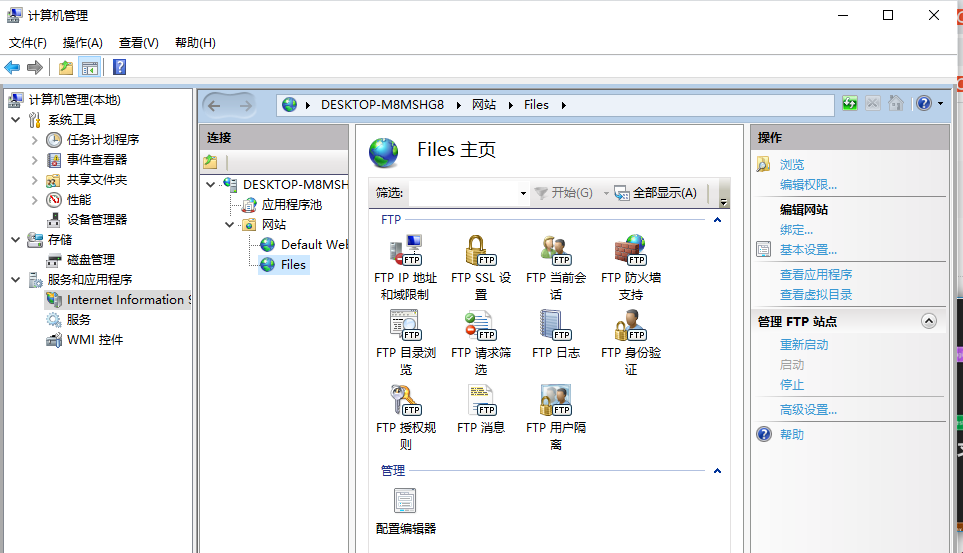
点击文件中的新建虚拟机，在调整虚拟机性能后选择群文件里的Ubuntu镜像文件，进去虚拟机后激活系统。

#### 2.1.3 Vmware tools的安装

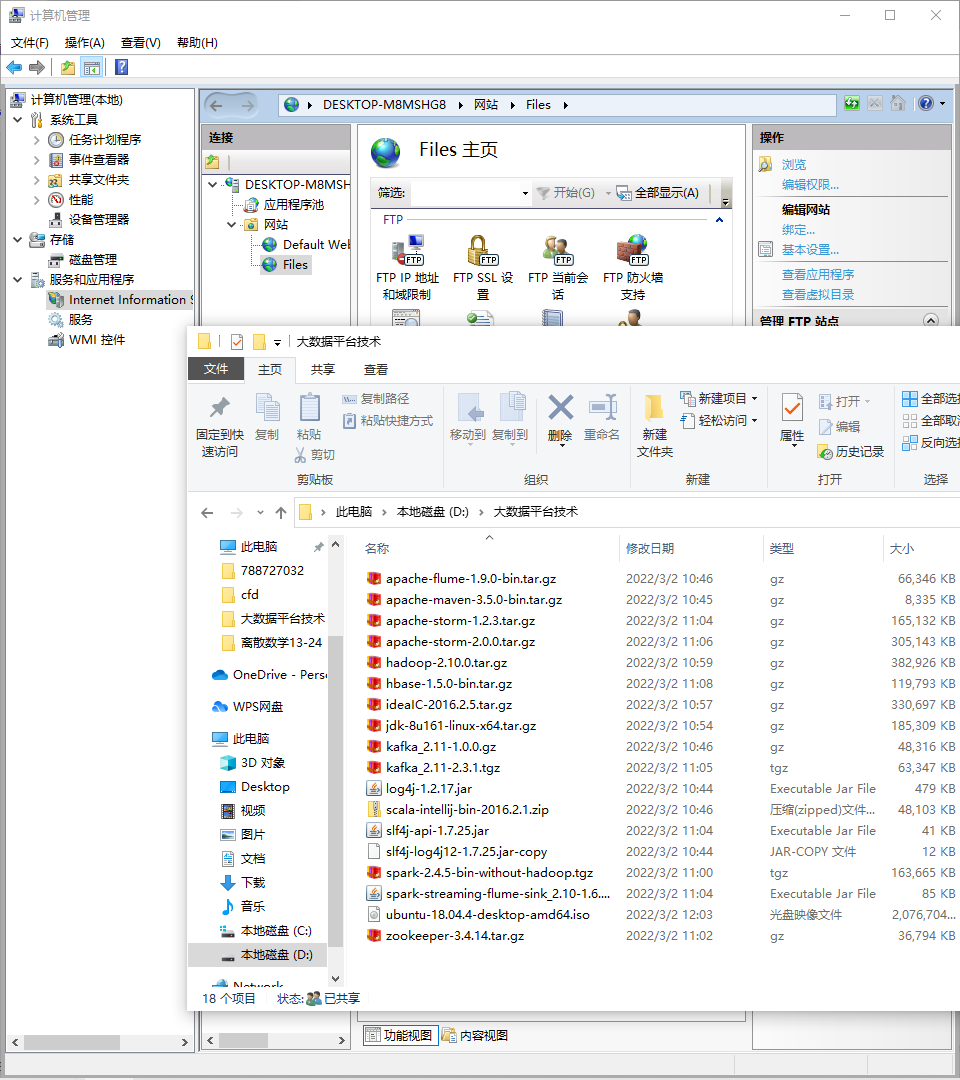


因为我的VMware tools是灰色的，于是再传输文件时我就用了其他方法。

在宿主机windows系统中开启IIS（ftp）设置ip地址和端口后，在设置用户权限和要分享的文件夹即可。

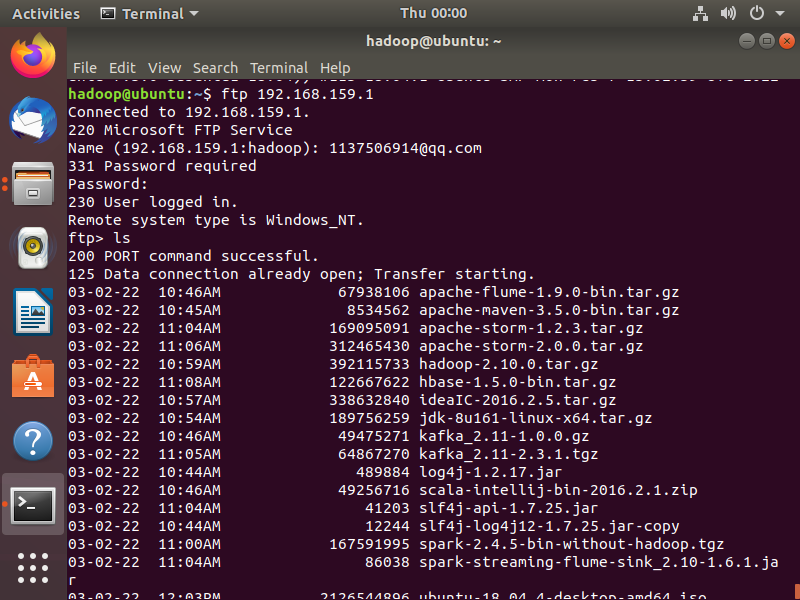


其中共享的文件夹为本课程所需内容：



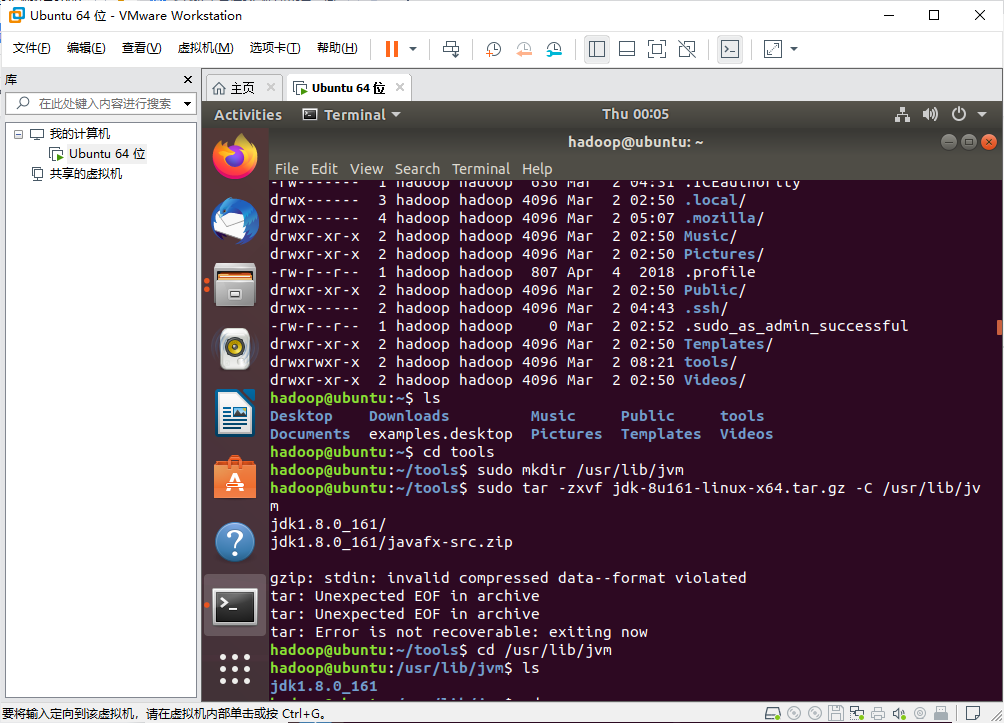
之后在linux中使用apt-get install一个lftp

使用lftp连接IIS，



连接后要先调整传输方式为二进制传输（binary）

IIs默认传输方式为ASCII传输，传输过程中由于编码问题，压缩文件会造成损坏，无法正常解压和使用



之后使用get命令下载服务器文件

### 2.2Linux的常用命令

在lftp的苦苦挣扎中一直在用，，，，

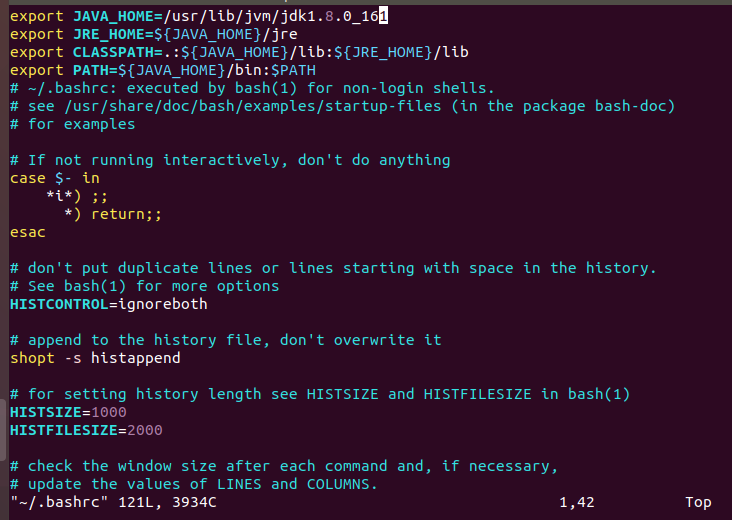
### 2.3JDK的安装

使用 tar指令解压



使用vim配置环境变量

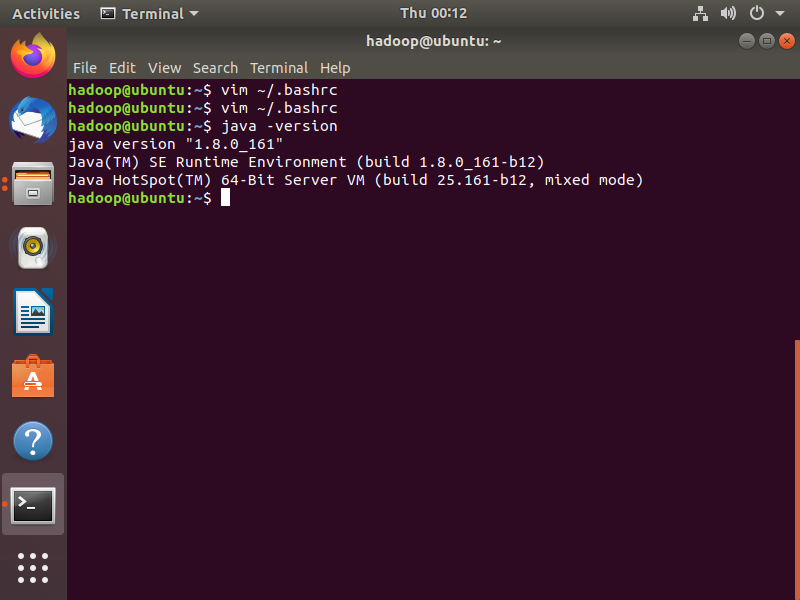




之后再次执行使其生效



测试

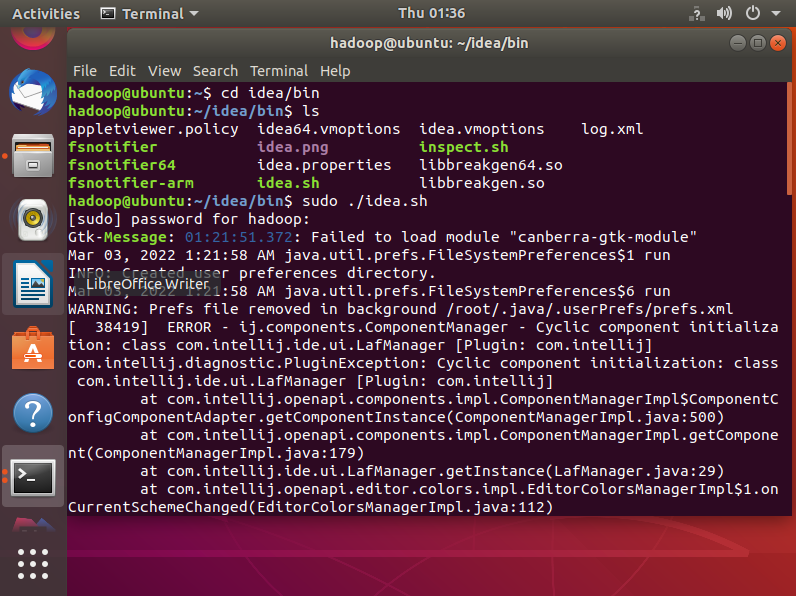


### 2.4 IDEA+Maven的安装

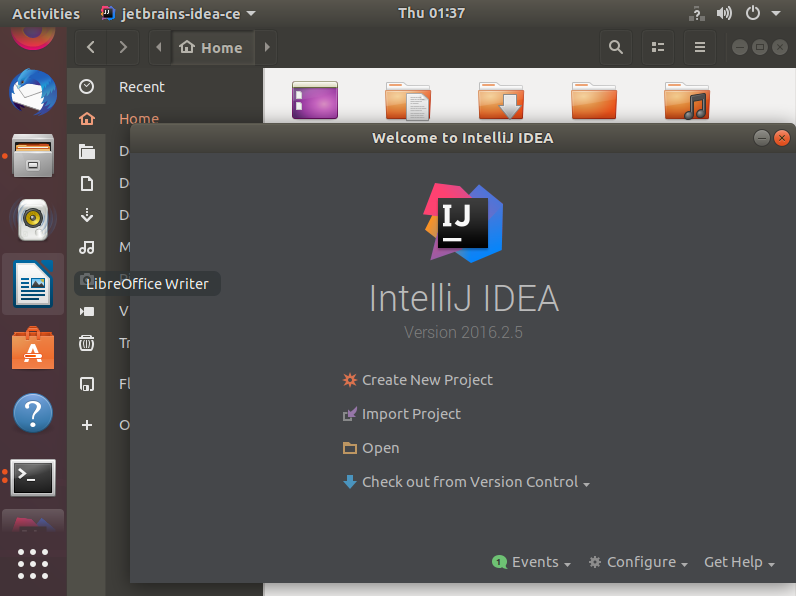
#### 2.4.1 IDEA的安装

使用tar命令解压文件

之后使用sudo权限 打开bin里面的idea.sh

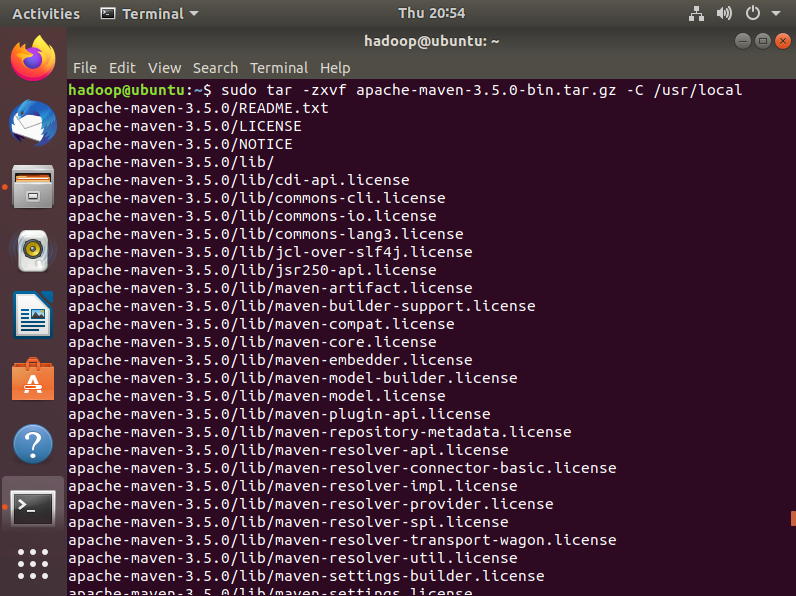


跟随步骤安装即可

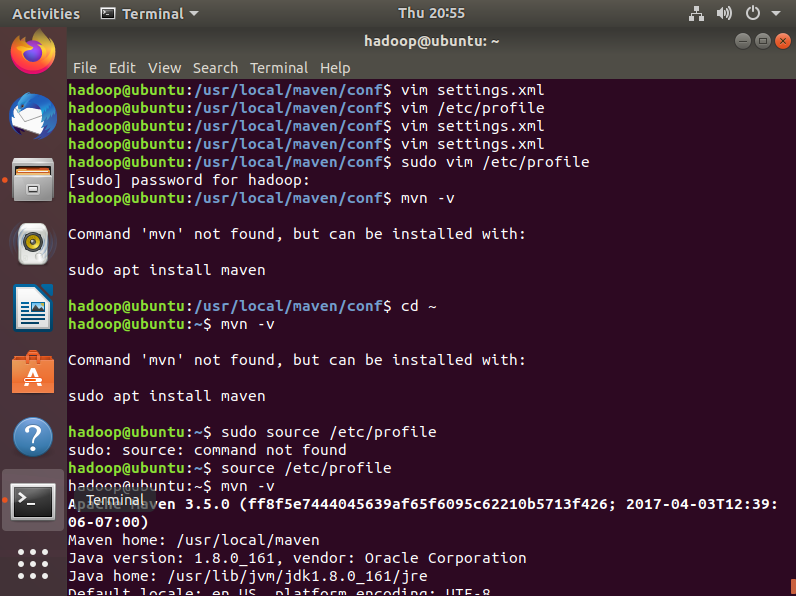


#### 2.4.2 Maven的安装

解压到/usr/local



在maven的settings.xml里修改仓库地址



可以试试改变maven中心仓库的镜像地址，提高下载jar包的速度，同样在setting.xml里可以改仓库的地址，手动设置jar包的安装位置。

之后在 /etc/profile里配置环境变量

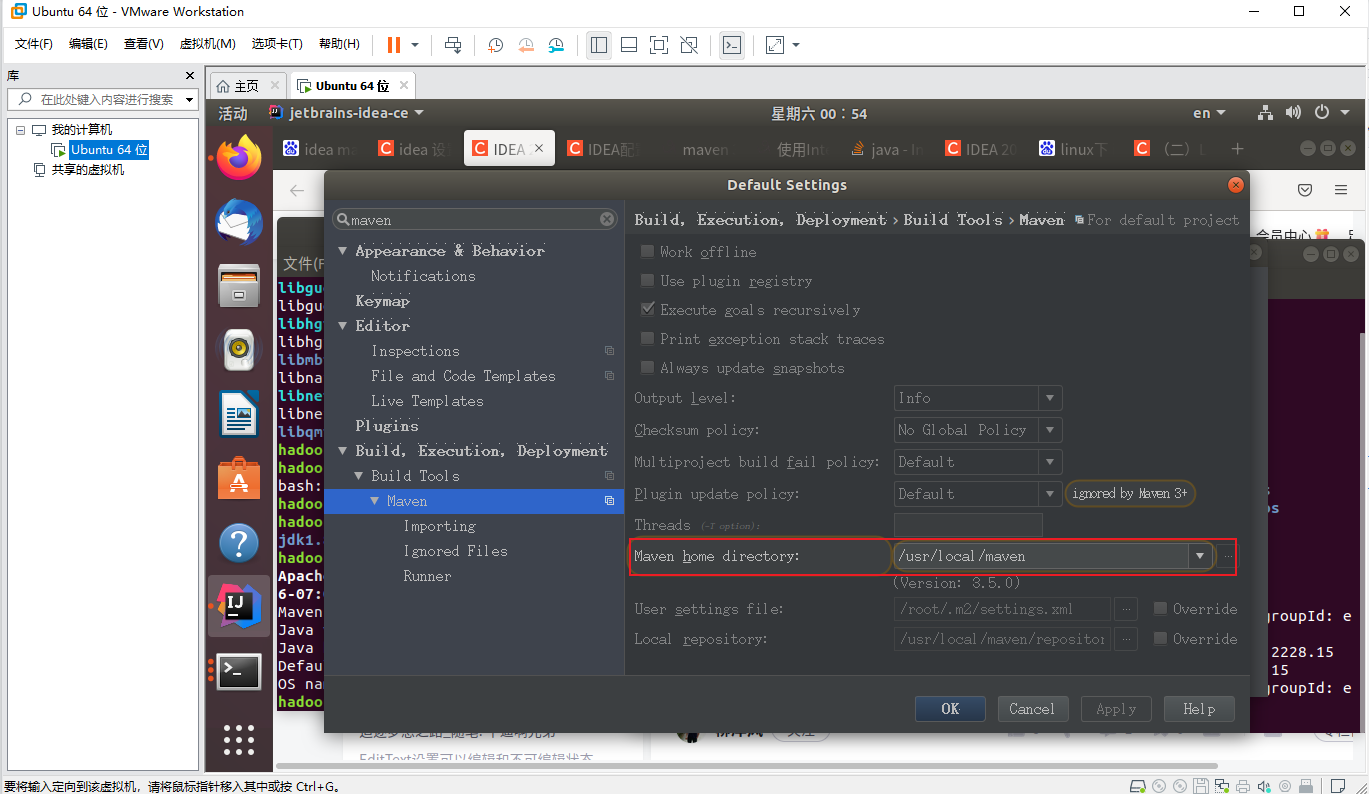
Source 重新使用/etc/profile文件 应用新配置的环境变量

mvn -v测试

显示出maven和java版本代表安装完毕

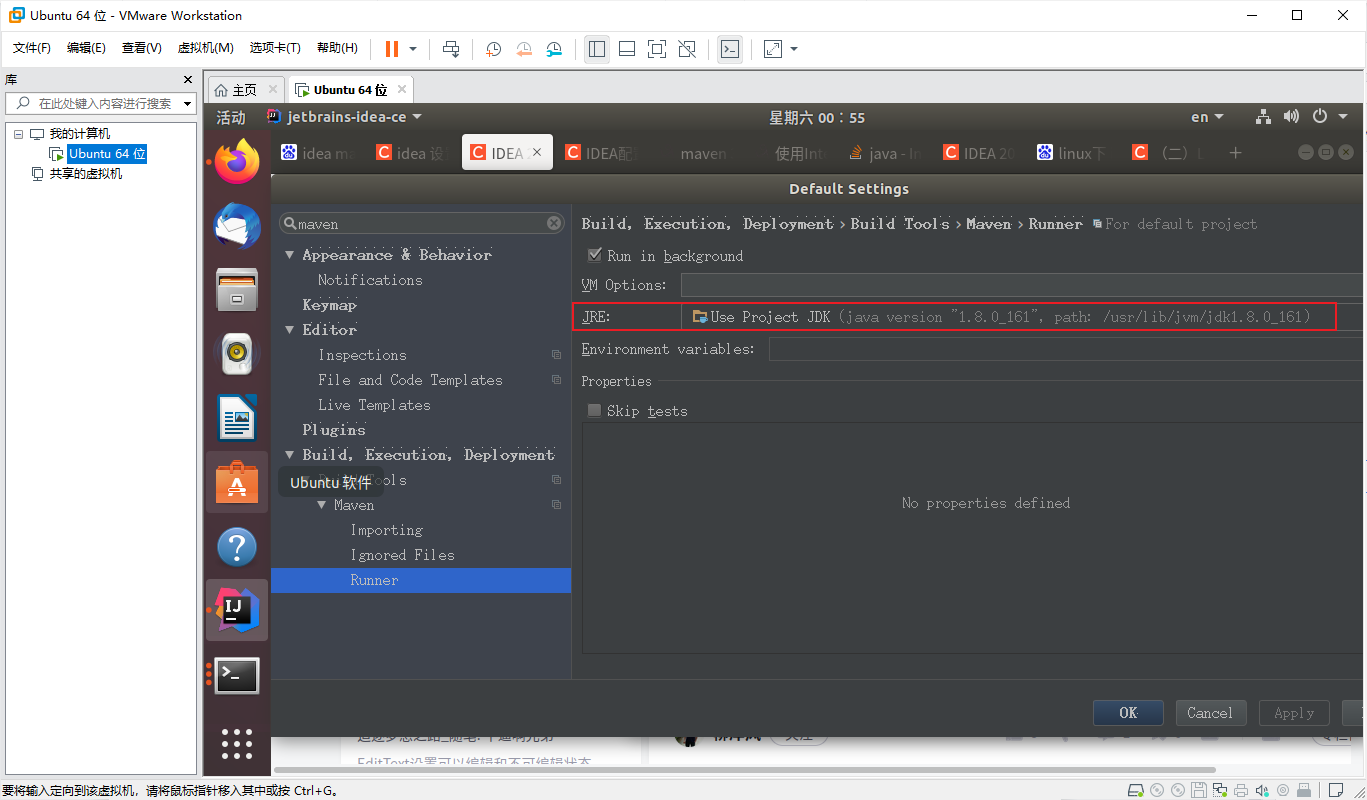
#### 2.4.3 在IDEA项目中配置JDK和Maven

打开idea的maven选项



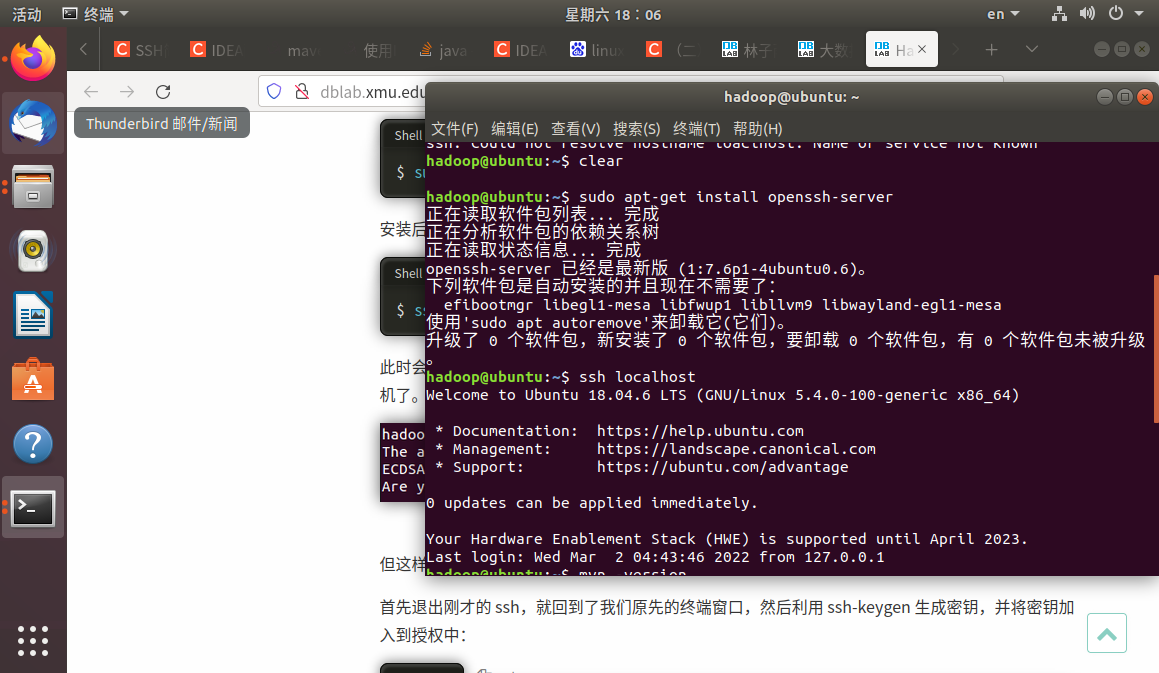
将maven地址改为安装包地址

将maven运行时的jre改为我们下载的版本（idea集成了maven，用的jre为线上版本）



### 2.5 Hadoop的安装

#### 2.5.1 SSH的安装



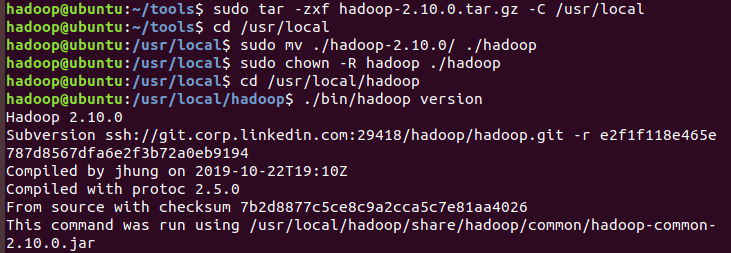
apt-get Install下载openssh-server

ssh localhost尝试连接本地服务器

连接成功

#### 2.5.2 Hadoop的安装

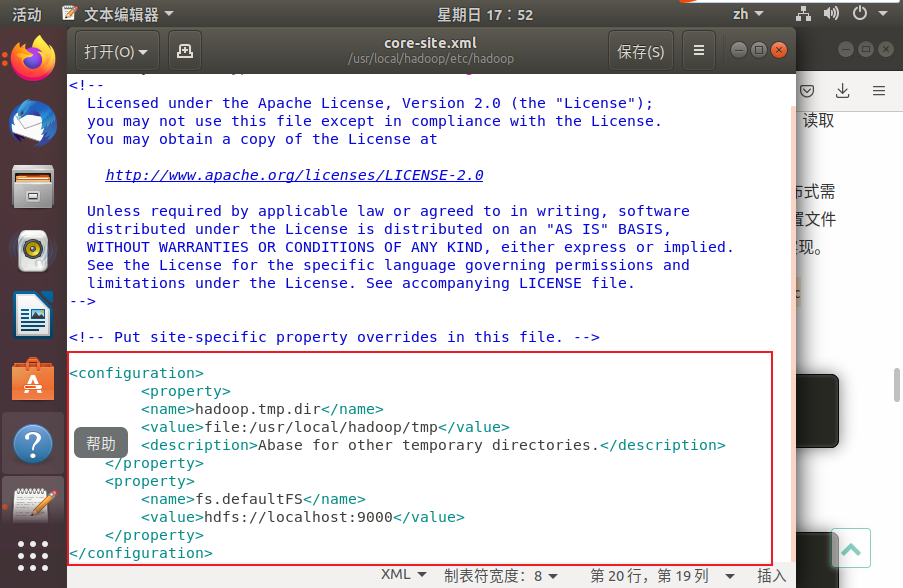
解压Hadoop压缩文件

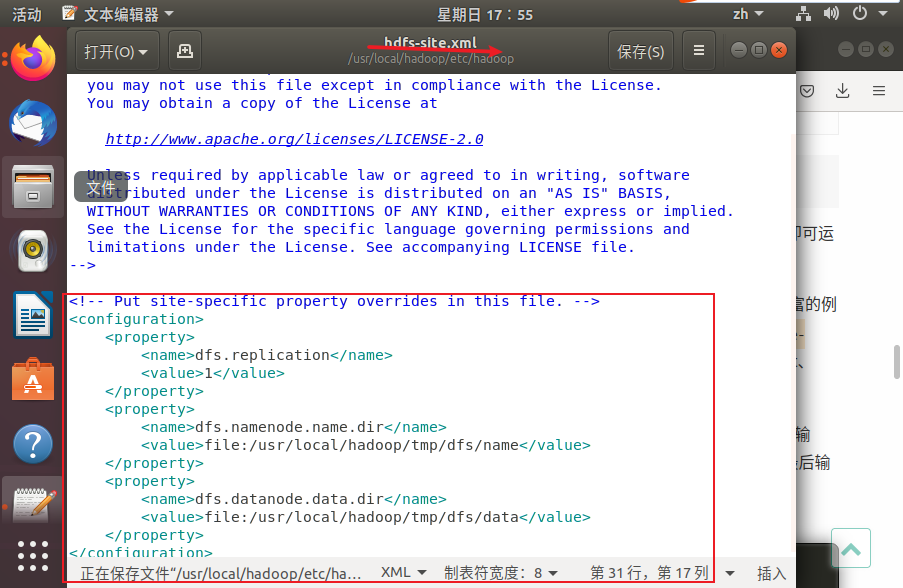


#### 2.5.3 伪分布式环境配置

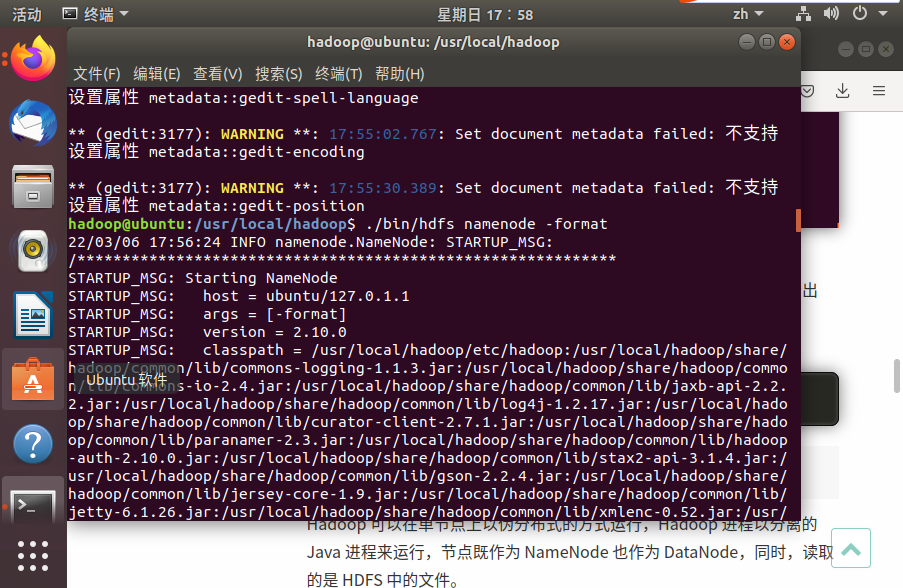
找到hadoop文件里etc/hadoop里的core-site.xml和hdfs-site.xml

使用gedit编辑：配置xml文件

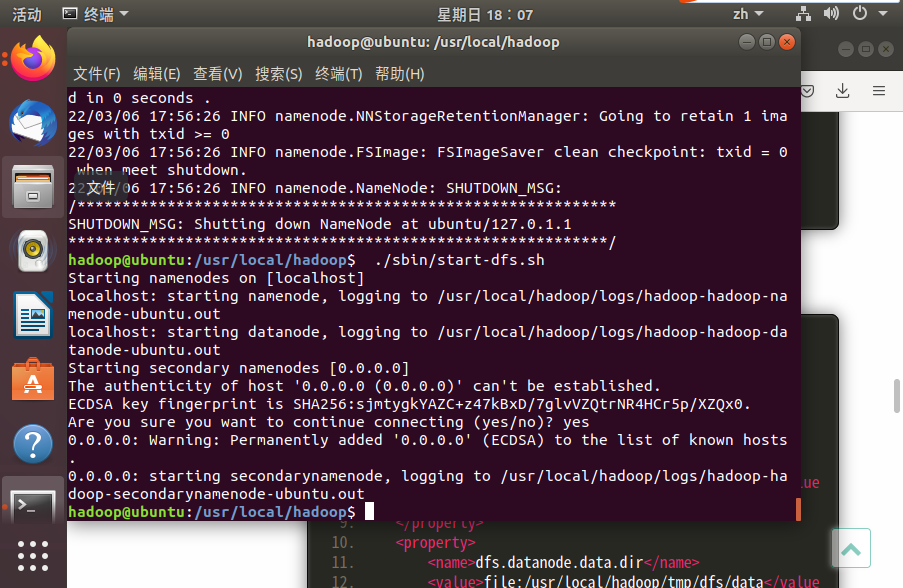




对hdfs进行格式化：

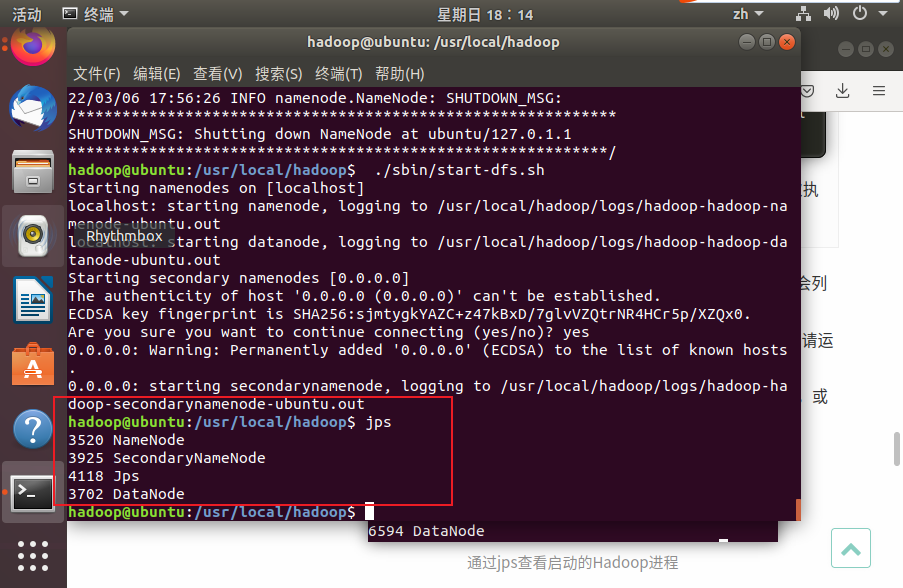


开启 NameNode 和 DataNode 守护进程。



用jps命令查看进程

“NameNode”、”DataNode” 和 “SecondaryNameNode”已启动



浏览器输入 http://localhost:50070/

查看hadoop版本

