大数据课程学习笔记

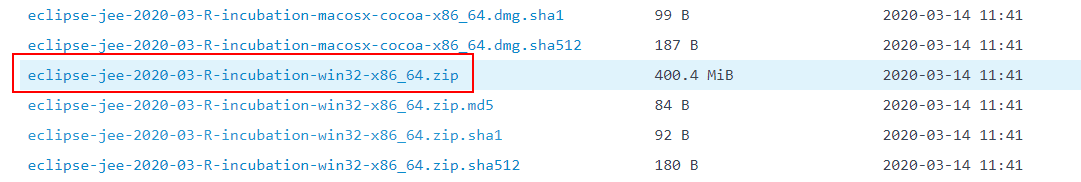
## 利用 Eclipse 可视化 HDFS 上的文件信息

下载eclipse

官网下载链接：<https://www.eclipse.org/downloads/>

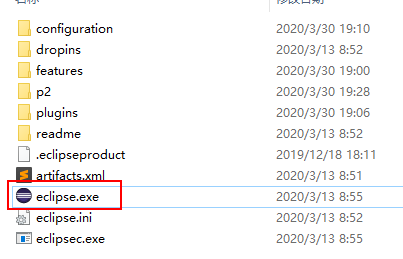
官网下载较慢，可以使用镜像下载：

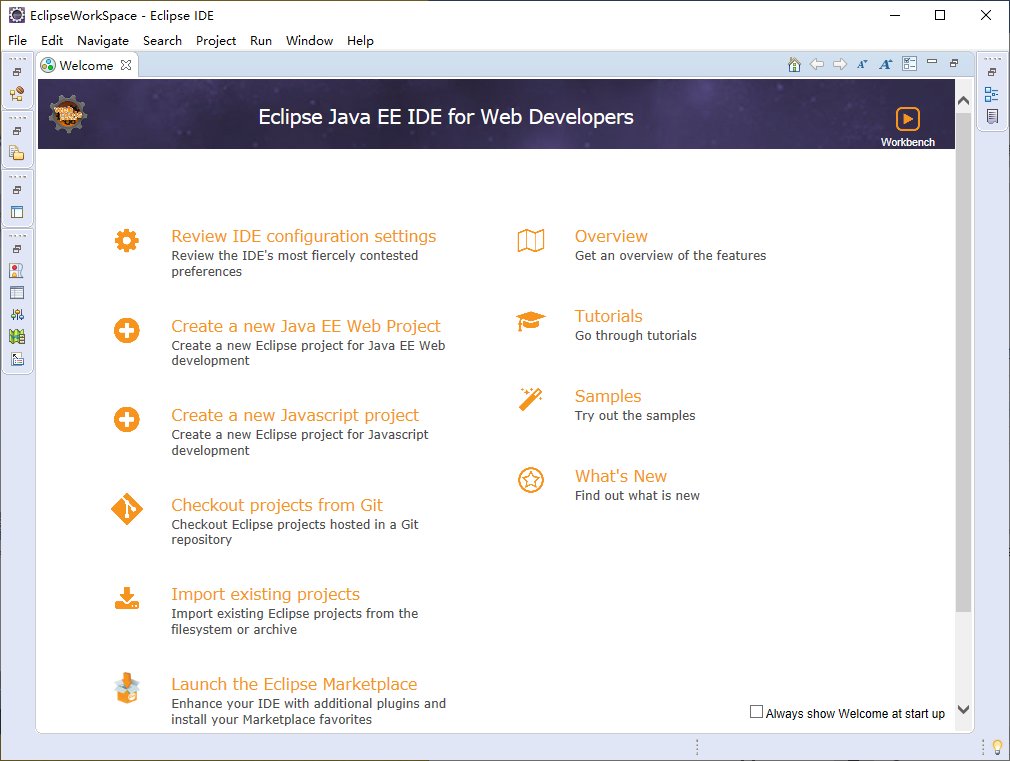
<https://mirrors.tuna.tsinghua.edu.cn/eclipse/technology/epp/downloads/release/>



选择eclipse-jee-xxx-win32-xxx版本

下载解压，然后运行exe



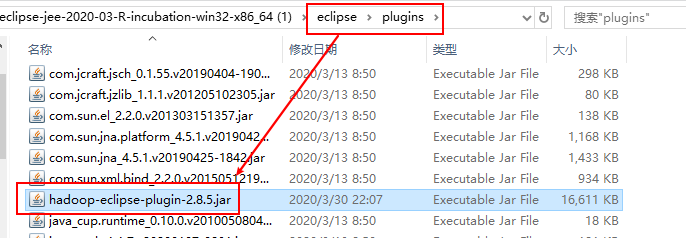


### 编译hadoop-eclipse插件

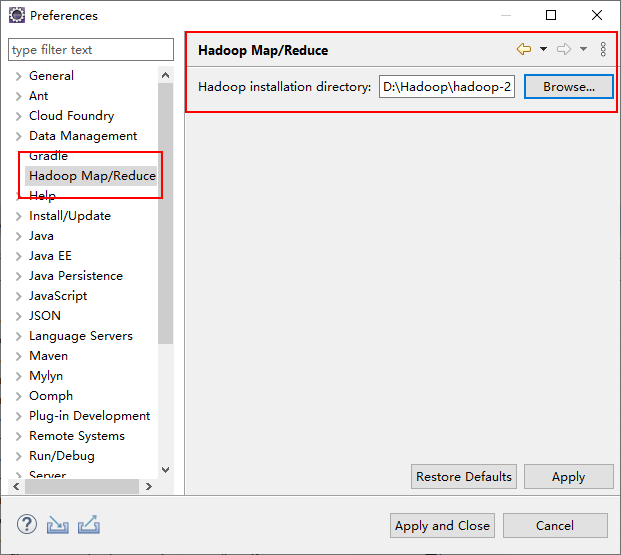
可以下载ant自行编译对应环境的插件，本次用的编译好的插件。

下载链接：<https://github.com/DoubleBirdsU/Hadoop-eclipse-plugin>

将 hadoop-eclipse-plugin-2.8.5.jar 插件包放到 eclipse/plugins 文件夹下

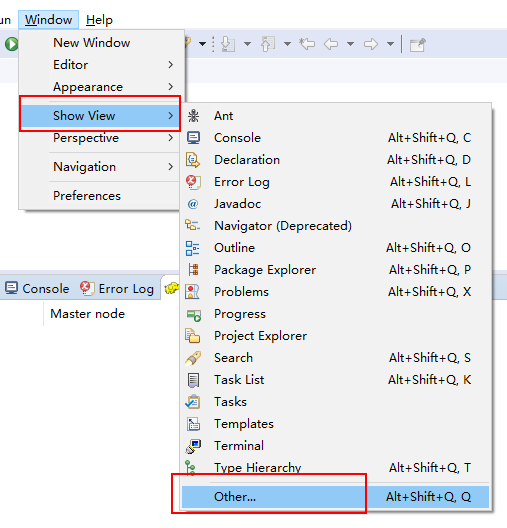


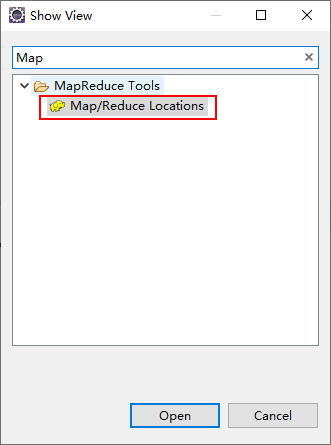
启动 eclipse ，点击 Window -> Prefrences ，在弹出的框选中Hadoop Map/Reduce ，然后在右边配置hadoop解压的路径

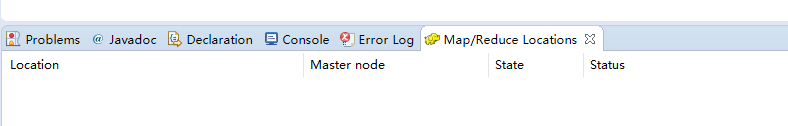


### 切换"Map/Reduce"工作目录

选择"Window"菜单下选择"Show view"，选择others，弹出一个窗体，搜索Map，从中选择"Map/Reduce"选项即可进行切换



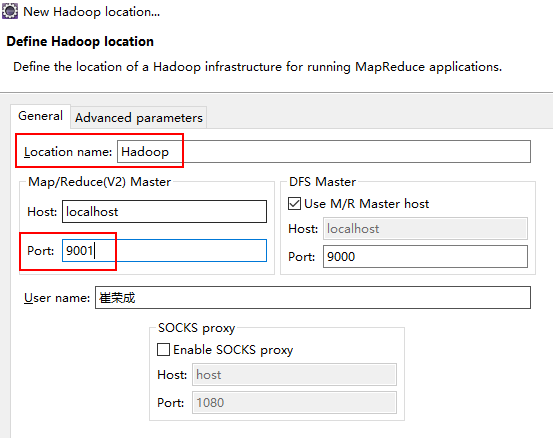




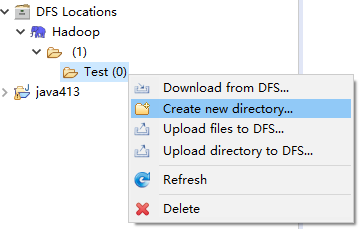
### 建立与Hadoop集群的连接

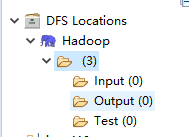
在Eclipse软件下面的"Map/Reduce Locations"进行右击，弹出一个选项，选择"New Hadoop Location"，然后弹出一个窗体。





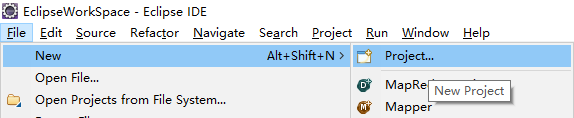
连接成功后可以看到HDFS结构，因为第一次连接所以没有东西，可以右击闯进文件夹。



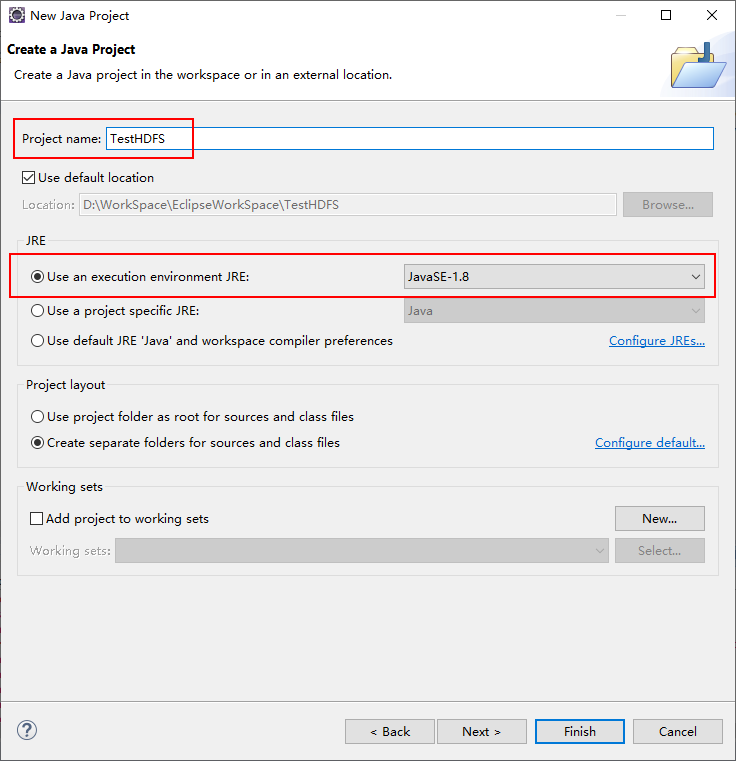


## Eclipse编程进行hdfs常用操作

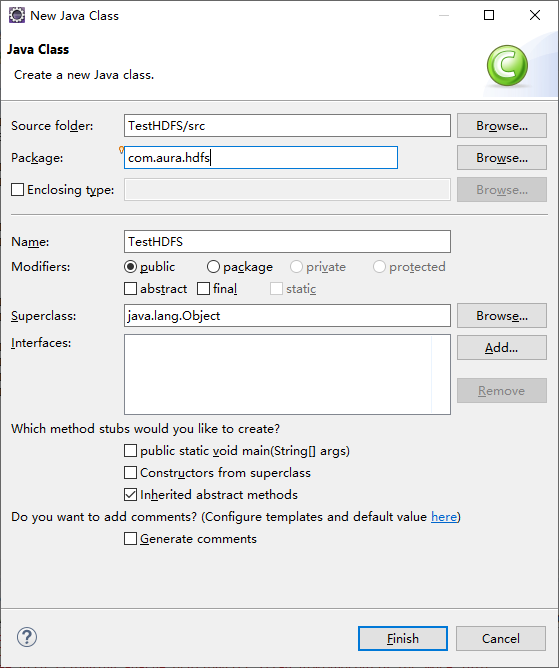
### 创建TestHDFS程序



选择Java程序

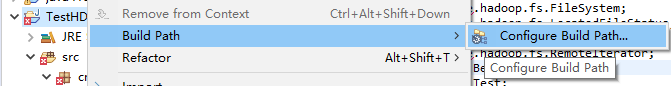


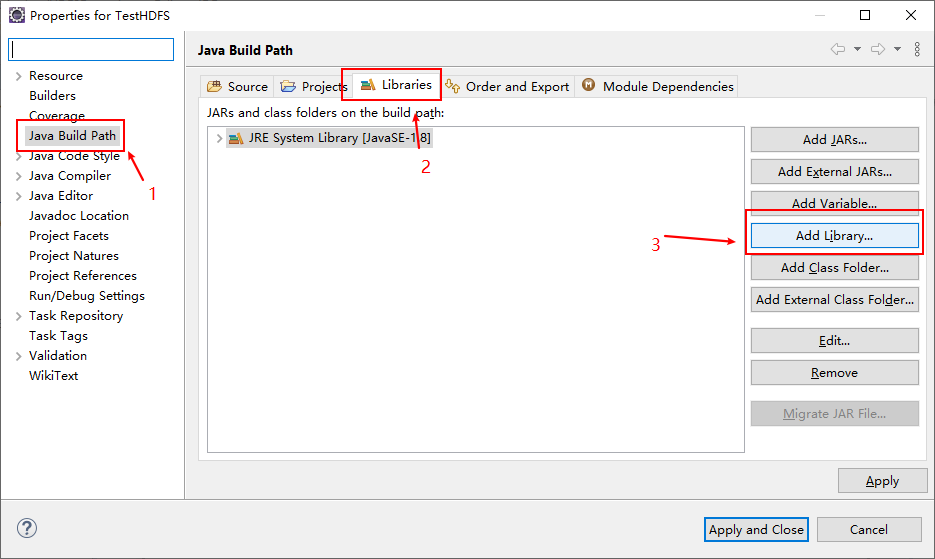
### 创建TestHDFS类

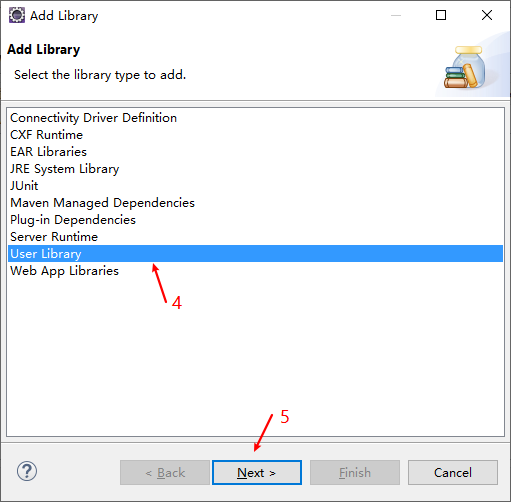


### 添加依赖库

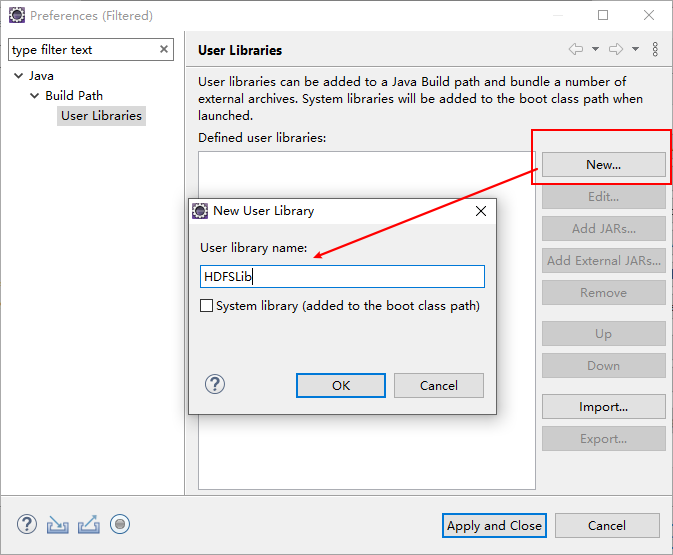
右击项目，选择build path，进行配置

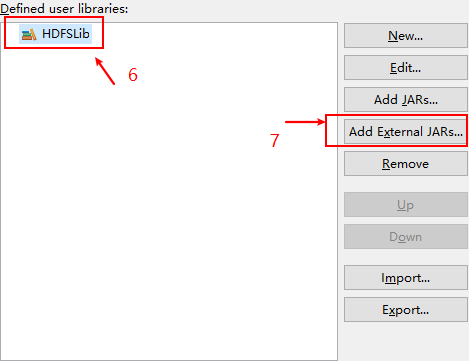




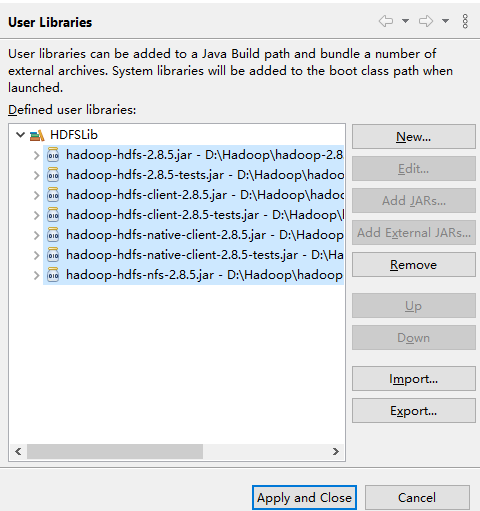


接下来选中User Libraries -> New ，输入HDFS库的名字





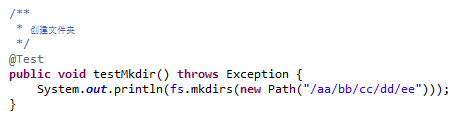
然后把hadoop目录下/share/hadoop/hdfs的所有jar包放进去

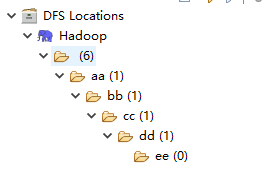


### 创建fs对象

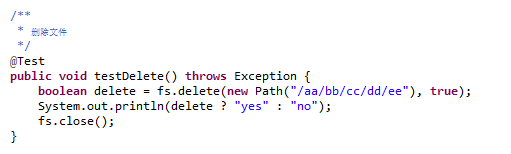


### 创建文件夹



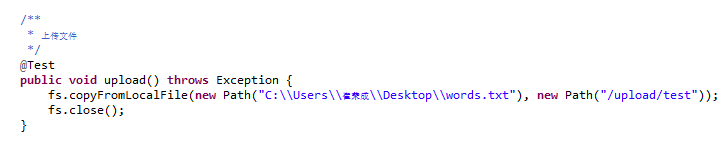


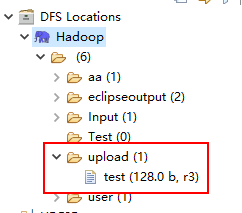
### 删除文件



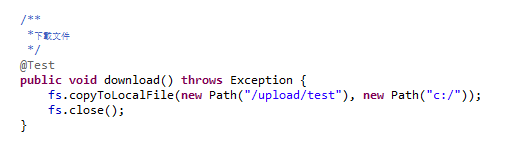


### 上传文件





### 下载文件



### 查看文件及文件夹信息

