# 使用IDEA运行Spark程序

梁世威

2014.09.29

## 1. 安装IDEA

从IDEA官网下载Community版本，解压到/usr/local/idea目录下。

tar –xzf ideaIC-13.1.4b.tar.gz

mv idea-IC-135.1230/ /usr/local/idea

为了方便使用其bin目录下的命令，把它配置在“/etc/profile”目录下。

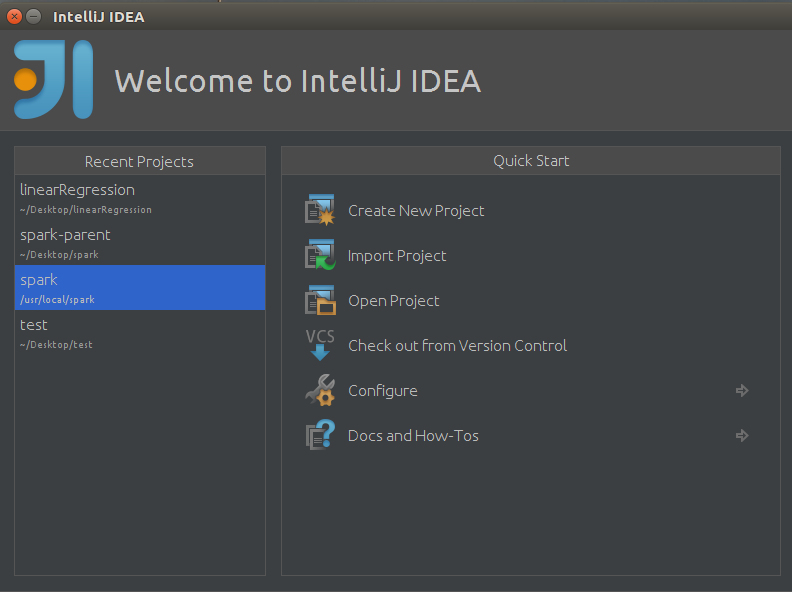
export PATH=$PATH:/usr/local/idea/ idea-IC-135.1230/bin

## 2. 安装Scala插件

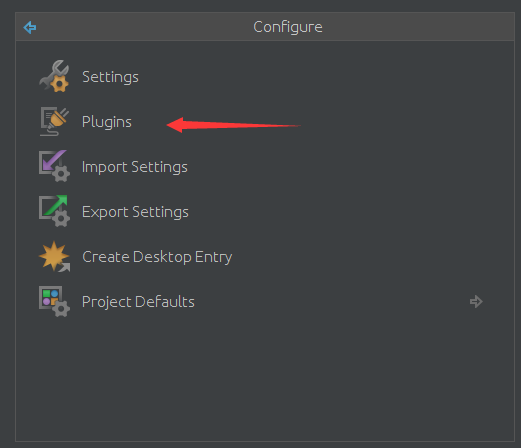
1. 在终端输入idea.sh进入如下界面：



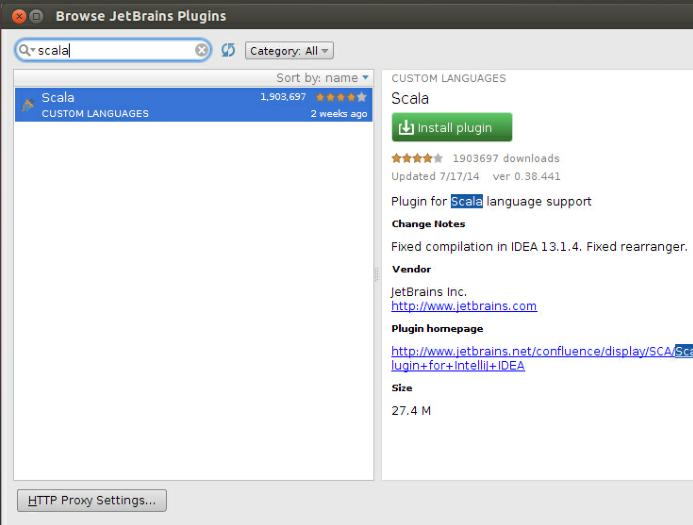
2. 点击”Configure”进入IDEA的配置界面：



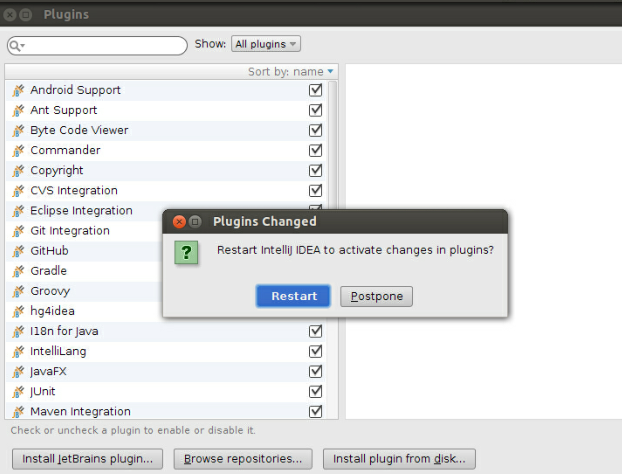
3. 选择“Plugins”,进入插件安装界面：



4. 点击左下角的“Install JetBrains plugin”选项进入如下界面在左上方的输入框输入“scala”来查找scala插件，点击右侧的“Install plugin”开始下载插件：

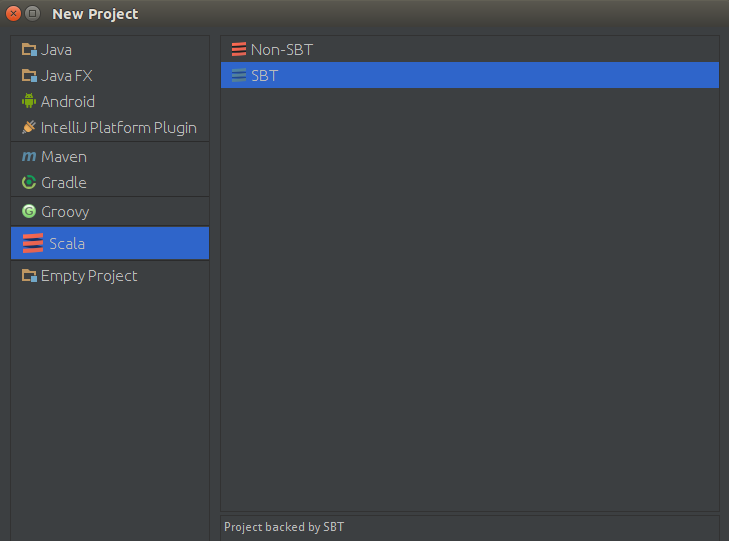


5. 安装完后重启IDEA。

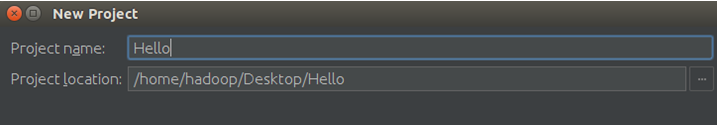


## 3. 使用scala SBT工程

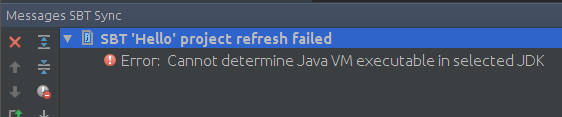
1. 重启之后进入如下界面，选择“Create New Project”,选择左侧列表中的“Scala”选项，为了方便以后的开发，选择右侧的“SBT”选项：

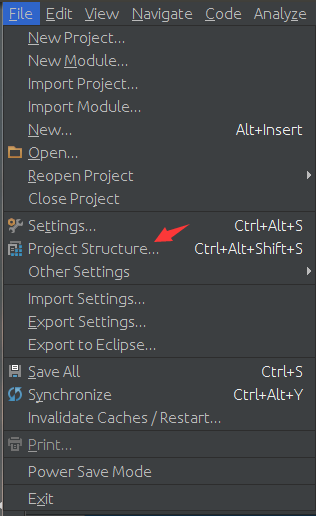


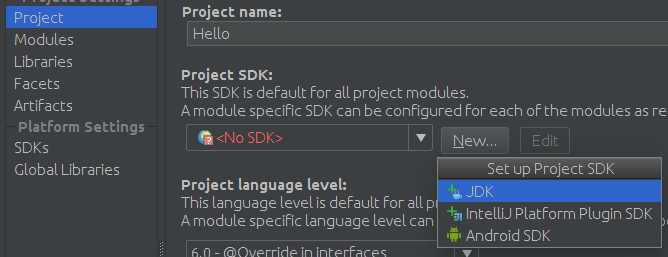
2. 设置scala工程的名称和目录，点击Finish：

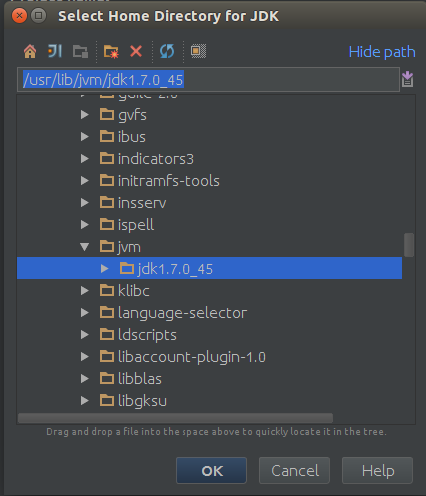


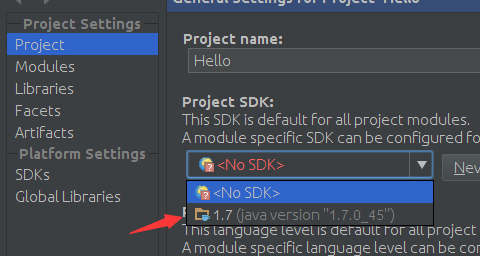
3. 此时发现没有设置Java的JDK路径，点击进入Project Structure目录，点击New，选择JDK路径后，关闭后重新打开项目：

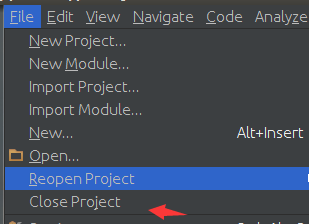




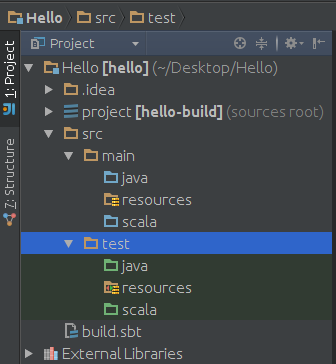




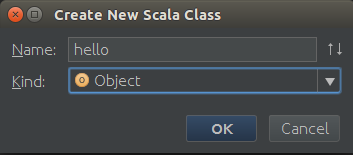




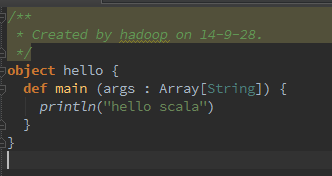
4. 此时IDEA会智能的帮助我们构建SBT工具，IDEA自动完成SBT工具的安装需要一段时间（1-2分钟），SBT会自动帮我们建好一些目录：



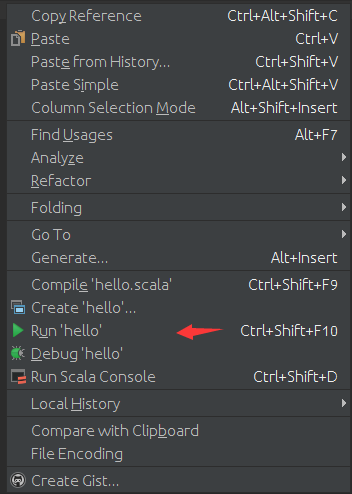
5. 点击src下的main下的scala在弹出的“New”下选择“Scala Class，输入文件名称并选择类型:



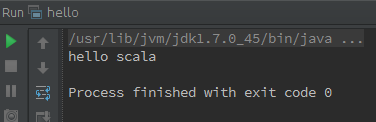
6. 写入代码：



7. 右击代码区点击运行，第一次运行scala程序会有点缓慢：

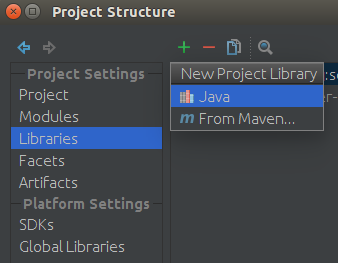


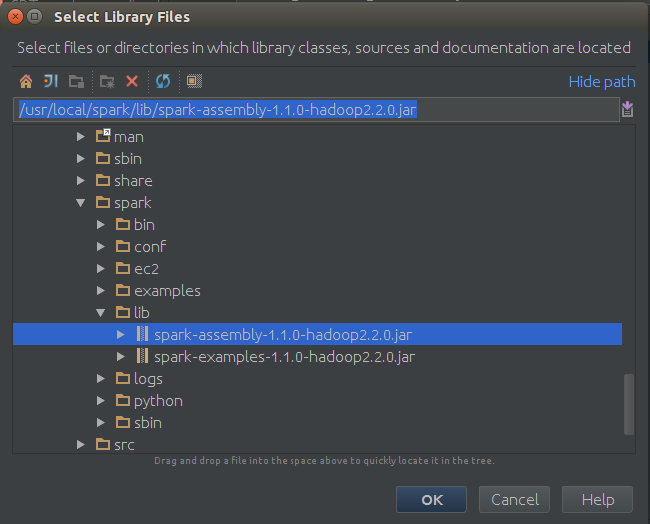
8. 运行的结果如下：

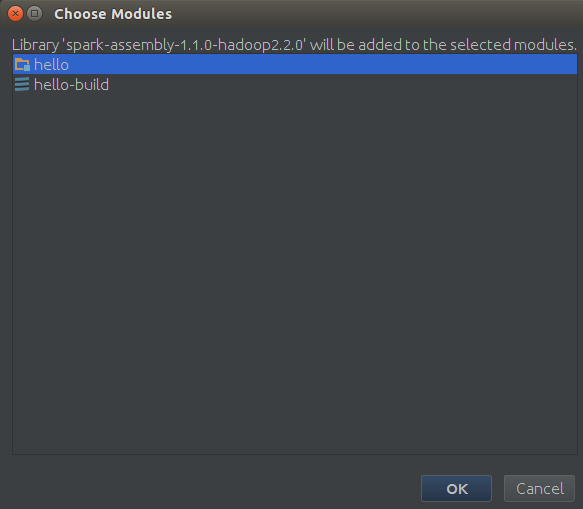


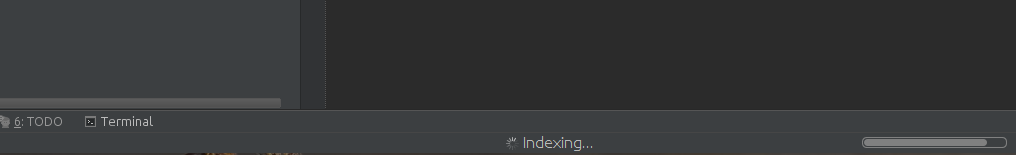
## 4. 搭建Spark运行环境

1.导入Spark的包，依次点击File -> Project Structure -> Libraries ->点击“+”号 ->选择Java -> 选择spark lib目录下的assembly包

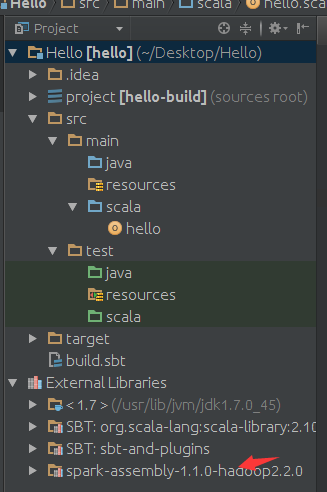






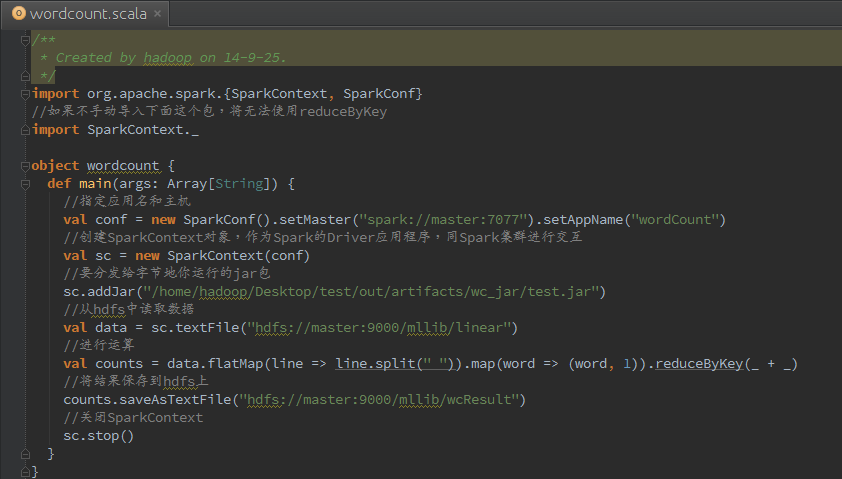


IDEA工作完后发现Spark的jar包导入到我们的工程中

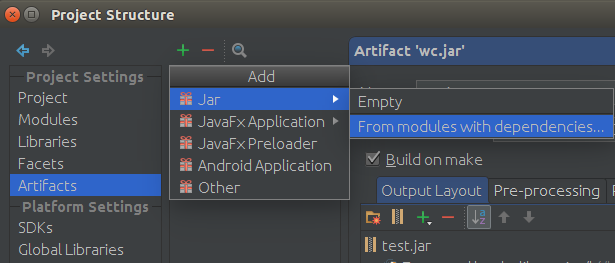


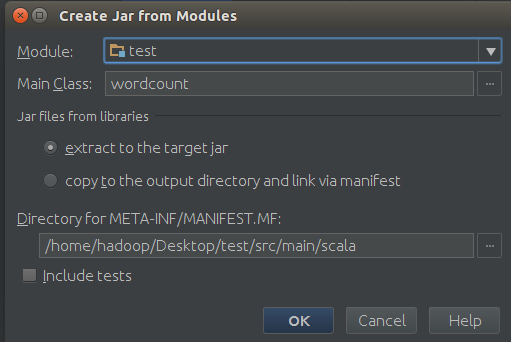
## 5. 运行Spark程序

1.我们这里采用standalone模式运行wordcount作代码演示。数据从hdfs读取，代码如下。

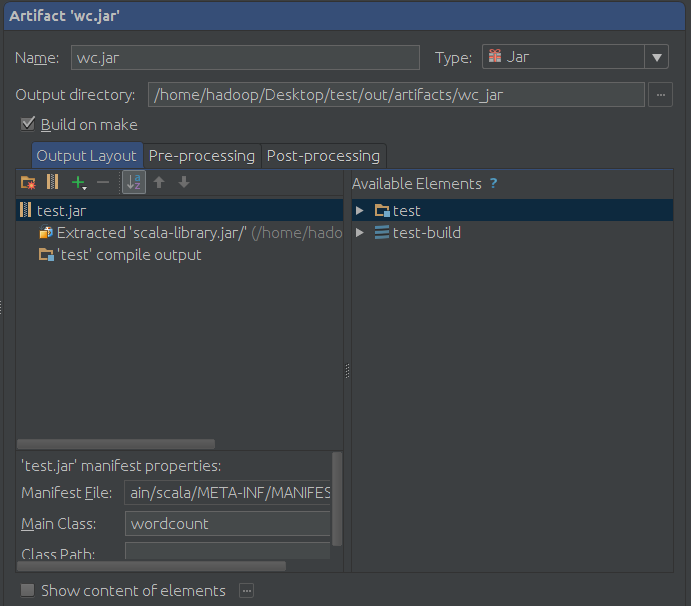


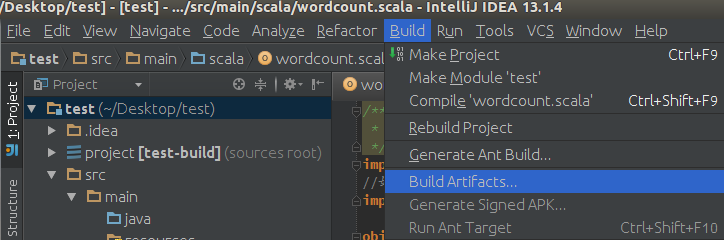
2. 下面描述如何打jar包：

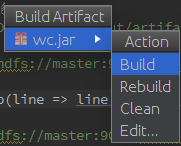




配置路径和包名







之后便可在集群上运行spark程序。

注：每次修改代码后都要rebuild jar包。