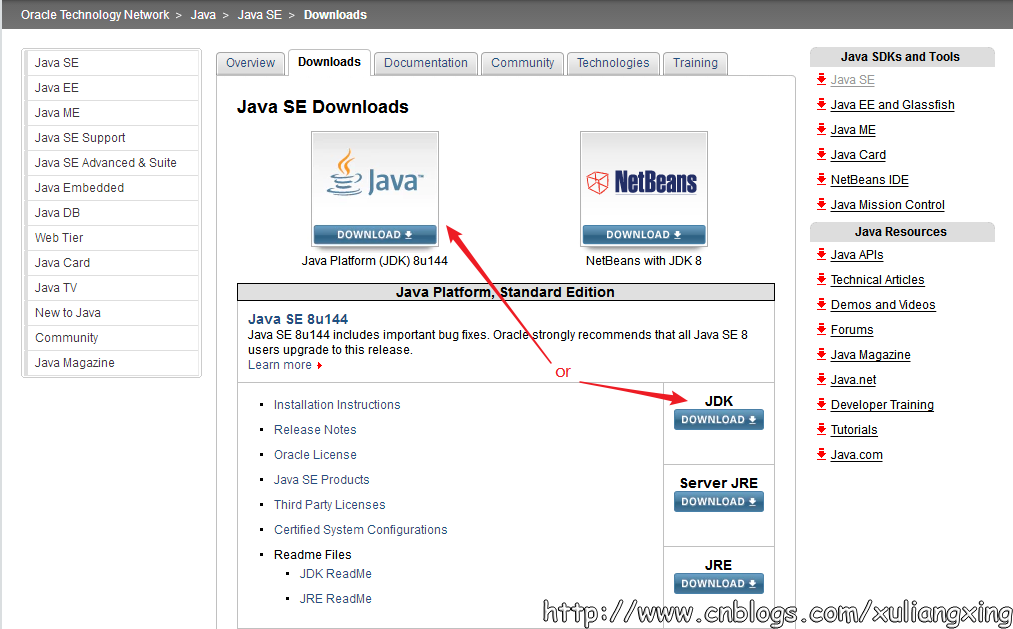
windows下安装spark

# 一、安装JDK

## 1.下载JDK

　首先需要安装JDK，并且将环境变量配置好，如果已经安装了的老司机可以忽略。JDK（全称是JavaTM Platform Standard Edition Development Kit）的安装，去Oracle官网下载，下载地址是[Java SE Downloads](http://www.oracle.com/technetwork/java/javase/downloads/index.html) 。



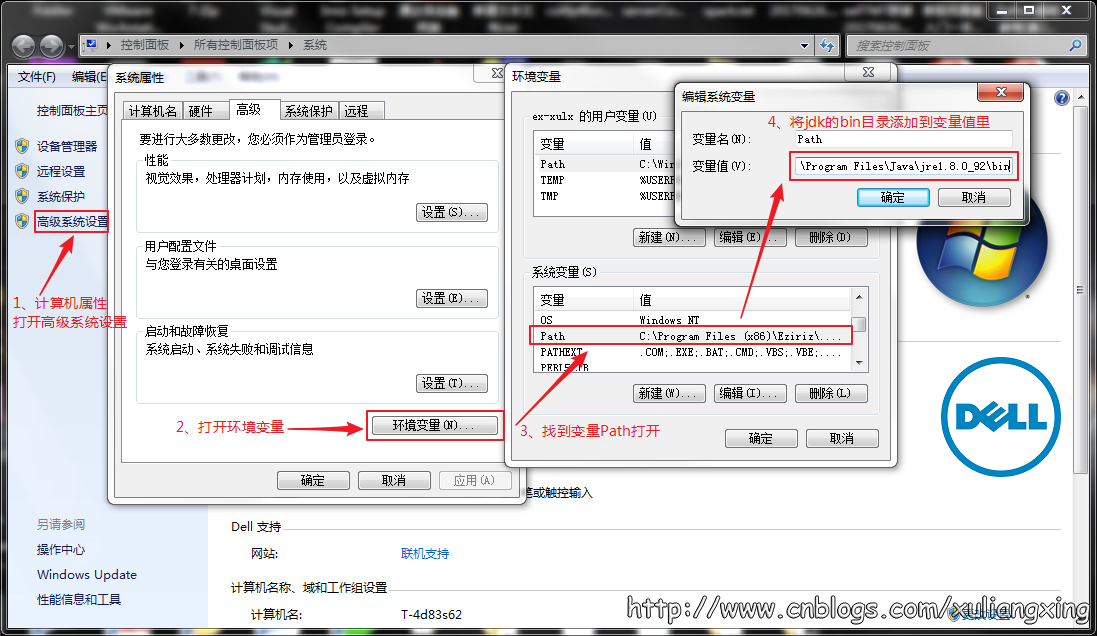
　　上图中两个用红色标记的地方都是可以点击的，点击进去之后可以看到这个最新版本的一些更为详细的信息，如下图所示：



　　下载完之后，我们安装就可以直接JDK，JDK在windows下的安装非常简单，按照正常的软件安装思路去双击下载得到的exe文件，然后设定你自己的安装目录（这个安装目录在设置环境变量的时候需要用到）即可。

## 2 JDK环境变量设置

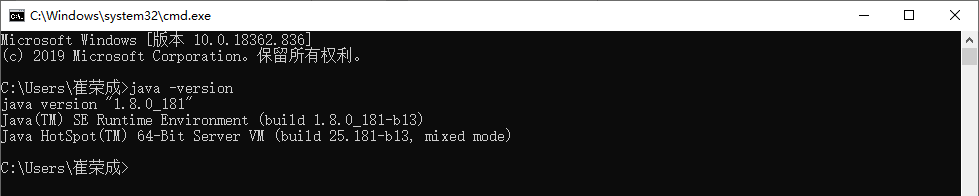
　　接下来设置相应的环境变量，设置方法为：在桌面右击【计算机】－－【属性】－－【高级系统设置】，然后在系统属性里选择【高级】－－【环境变量】，然后在系统变量中找到**“Path”**变量，并选择**“编辑”**按钮后出来一个对话框，可以在里面添加上一步中所安装的JDK目录下的bin文件夹路径名。如图所示：



　　这样设置好后，便可以在任意目录下打开的cmd命令行窗口下运行下面命令。查看是否设置成功。

java -version

　　观察是否能够输出相关java的版本信息，如果能够输出，说明JDK安装这一步便全部结束了。如图所示：

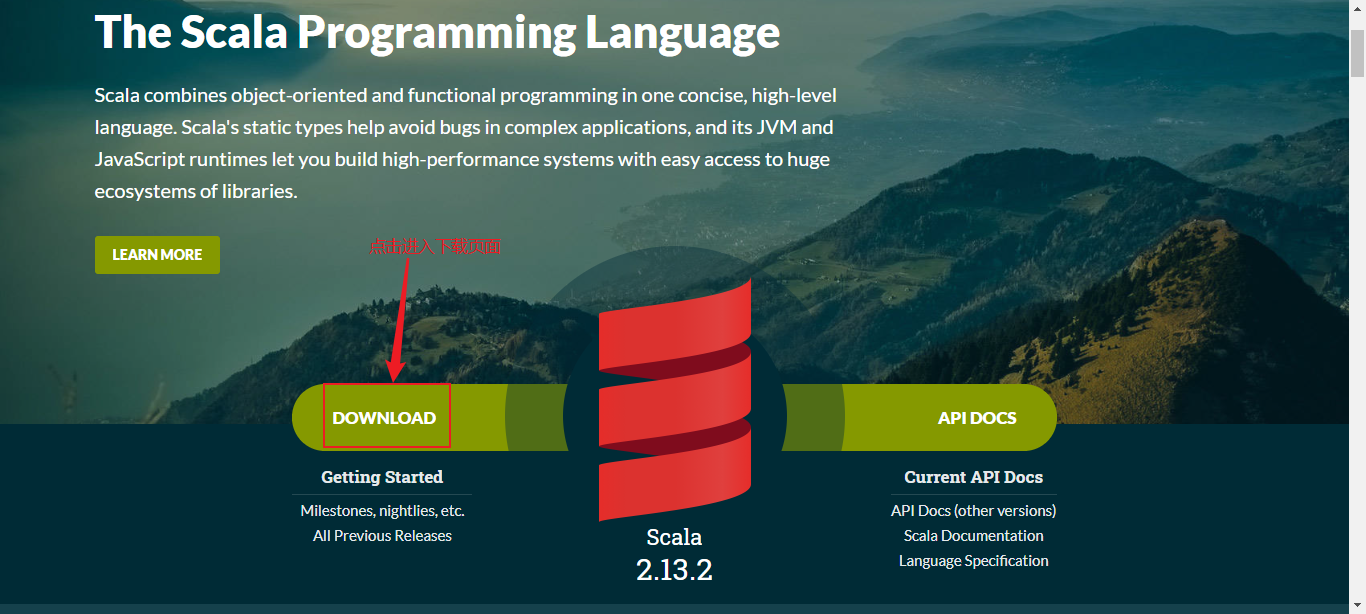


# 二、安装Hadoop

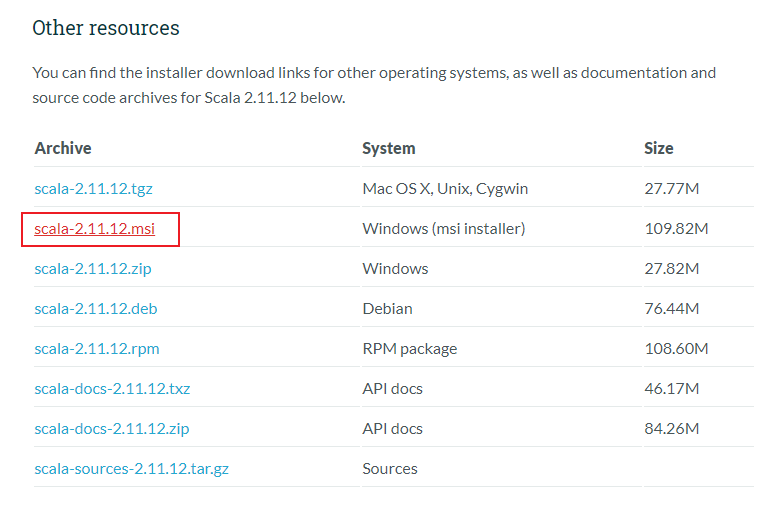
参考之前Hadoop安装说明

# 三、安装Scala

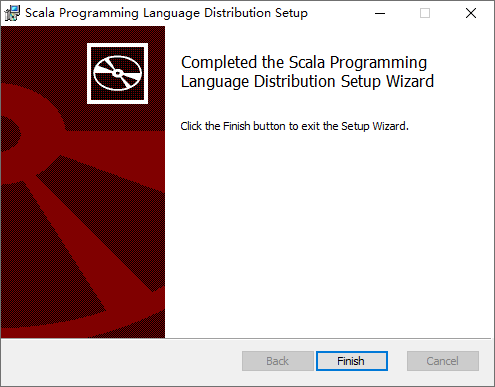
我们从官网：<http://www.scala-lang.org/> 下载Scala，最新的版本为2.13.2，如图所示



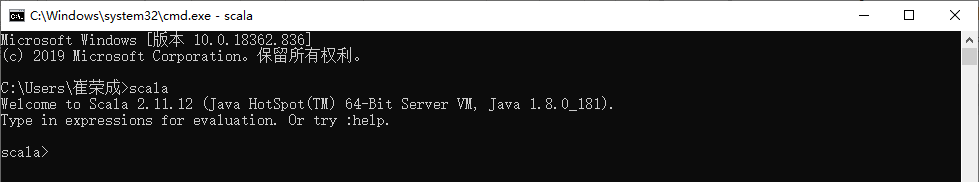
因为我们是在Windows环境下，这也是本文的目的，我们选择对应的Windows版本下载，如图所示：



下载得到Scala的msi文件后，可以双击执行安装。



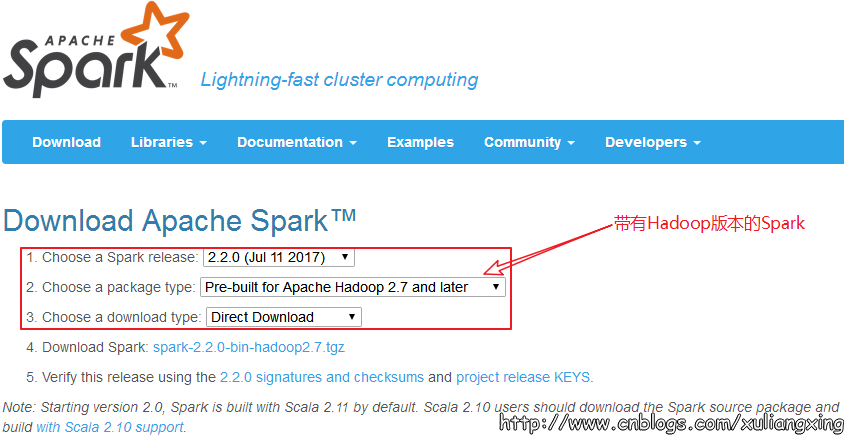
安装成功后，默认会将Scala的bin目录添加到**PATH**系统变量中去（如果没有，和上面JDK安装步骤中类似，将Scala安装目录下的bin目录路径，添加到系统变量**PATH**中），为了验证是否安装成功，开启一个新的cmd窗口，输入scala然后回车，如果能够正常进入到Scala的交互命令环境则表明安装成功。如下图所示：



备注：如果不能显示版本信息，并且未能进入Scala的交互命令行，通常有两种可能性：   
1、Path系统变量中未能正确添加Scala安装目录下的bin文件夹路径名，按照JDK安装中介绍的方法添加即可。   
2、Scala未能够正确安装，重复上面的步骤即可。

# 三、安装Spark

我们到Spark官网进行下载：<http://spark.apache.org/> ，我们选择带有Hadoop版本的Spark，如图所示：

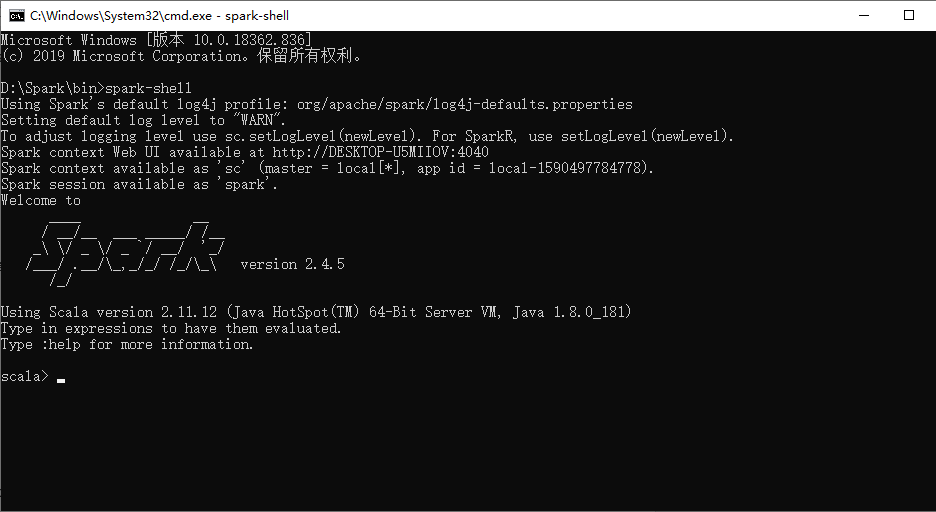


　　下载后得到了大约200M的文件： spark-2.2.0-bin-hadoop2.7

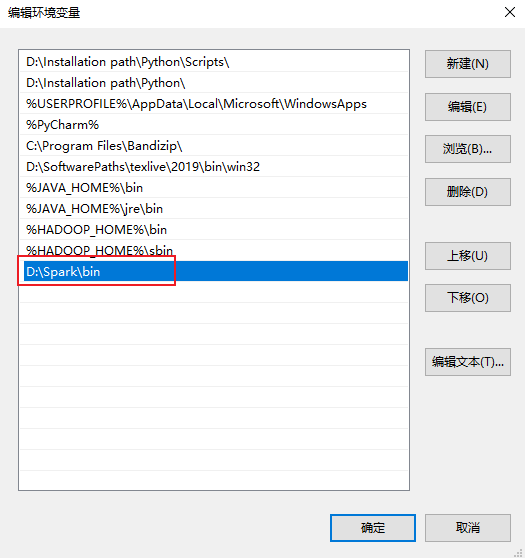
　　这里使用的是Pre-built的版本，意思就是已经编译了好了，下载来直接用就好，Spark也有源码可以下载，但是得自己去手动编译之后才能使用。下载完成后将文件进行解压（可能需要解压两次），最好解压到一个盘的根目录下，并重命名为Spark，简单不易出错。并且需要注意的是，在Spark的文件目录路径名中，不要出现空格，类似于“Program Files”这样的文件夹名是不被允许的。我们在D盘新建一个Spark文件夹存放，如图所示：

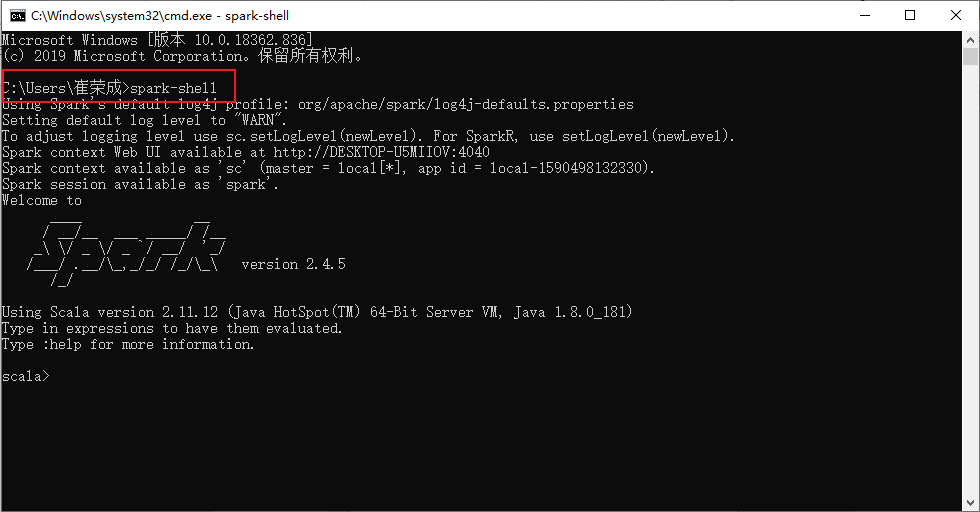


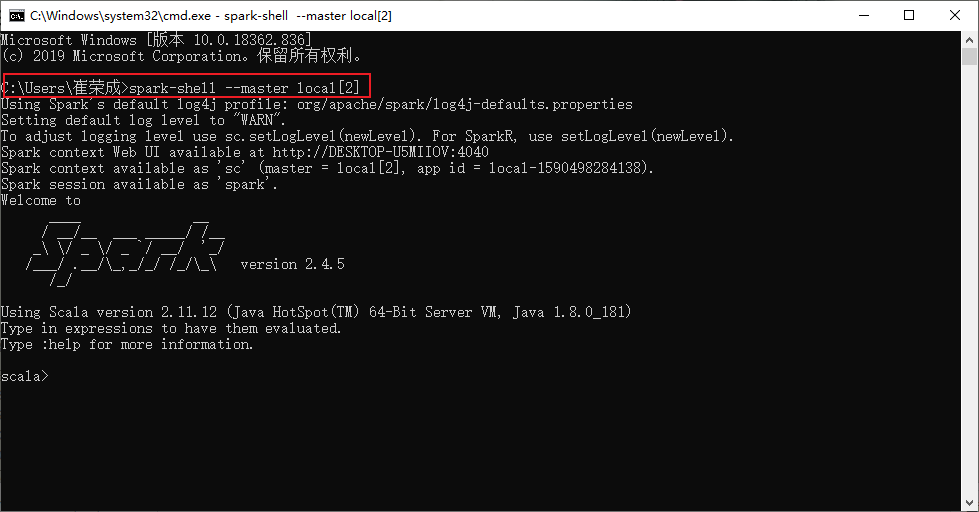
解压后基本上就差不多可以到cmd命令行下运行了。



但这个时候每次运行spark-shell（spark的命令行交互窗口）的时候，都需要先cd到Spark的安装目录下，比较麻烦，因此可以将Spark的bin目录添加到系统变量**PATH**中。例如我这里的Spark的bin目录路径为D:\Spark\bin，那么就把这个路径名添加到系统变量的**PATH**中即可，方法和JDK安装过程中的环境变量设置一致，设置完系统变量后，在任意目录下的cmd命令行中，直接执行spark-shell命令，即可开启Spark的交互式命令行模式。



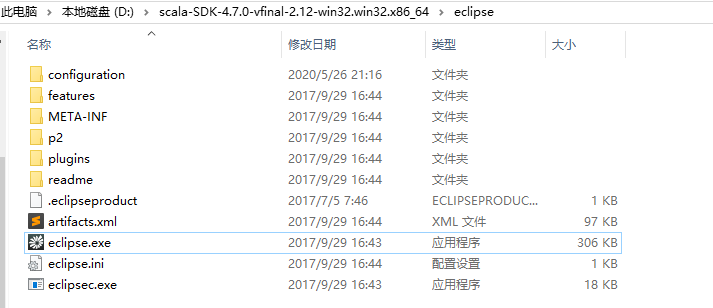




# 四、配置环境变量

# 四、Eclipse for Scala 开发spark程序

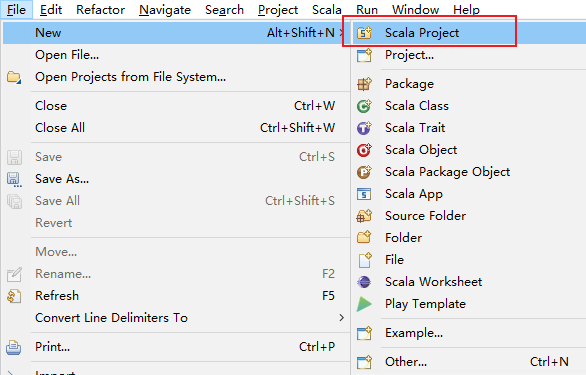
1、解压解压缩scala-SDK-4.7.0-vfinal-2.12-win32.win32.x86\_64



2、运行eclipse.exe



3.新建Scala project



4、建立Scala object

