# 编译安装spark-2.2.0

虚拟机环境：centos6.5

**说明：曾经尝试阿里云ecs服务器上编译，由于各种原因：如内存啊，网络啊，jar包下载不下来等，放弃了阿里云上直接编译，而采用虚拟机编译好，上传到阿里云上的方式部署。**

软件准备：spark-2.2.0.tgz源码

jdk-8u144-linux-x64.tar.gz JDK1.8以上

apache-maven-3.3.9-bin.tar.gz

scala-2.11.8.tgz

hadoop-2.6.0-cdh5.7.0.tar.gz

喜好: app目录：存放各种软件

data目录：存放测试数据

source目录：存放下载源码的软件

software目录：存放tar.gz这种软件包

lib目录:存放自己的jar包

maven\_repo ： 存放maven的依赖

mkdir -p /{app,data,source,software,lib,maven\_repo}

 在 / 目录下创建6个新目录，它们分别为：app,data,source,software,lib,maven\_repo

* **linux查看版本当前操作系统发行信息： cat /etc/issue 或 cat /etc/centos-release**

CentOS release 6.5 (Final)

* **Java查看版本 ：**

[root@mini1 local]# java -version

java version "1.8.0\_141"

Java(TM) SE Runtime Environment (build 1.8.0\_141-b15)

Java HotSpot(TM) 64-Bit Server VM (build 25.141-b15, mixed mode)

[root@mini1 local]#

* **Maven查看版本 ：**

[root@mini1 local]# mvn -v

Apache Maven 3.3.9 (bb52d8502b132ec0a5a3f4c09453c07478323dc5; 2015-11-11T00:41:47+08:00)

Maven home: /app/apache-maven-3.3.9

Java version: 1.8.0\_141, vendor: Oracle Corporation

Java home: /app/jdk1.8.0\_141/jre

Default locale: zh\_CN, platform encoding: UTF-8

OS name: "linux", version: "2.6.32-431.el6.x86\_64", arch: "amd64", family: "unix"

[root@mini1 local]#

* **Scala查看版本 ：**

[root@mini1 local]# scala

Welcome to Scala 2.11.8 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0\_141).

Type in expressions for evaluation. Or try :help.

scala>

* **Hadoop查看版本 ：**

## [【Linux】CentOS7下安装JDK详细过程](https://www.cnblogs.com/sxdcgaq8080/p/7492426.html)

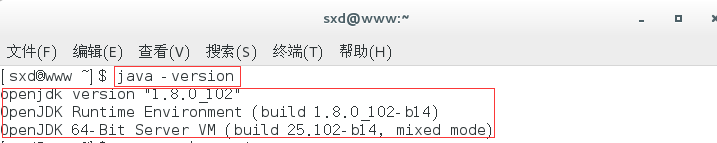
Linux上一般会安装Open JDK,关于OpenJDK和JDK的区别：http://www.cnblogs.com/sxdcgaq8080/p/7487369.html

下面开始安装步骤：

### 第一步　卸载系统自带的OpenJDK以及相关的java文件

①在命令窗口键入：

java -version



可以看到系统自带的OpenJDK版本信息。

②在命令窗口键入：

rpm -qa | grep java

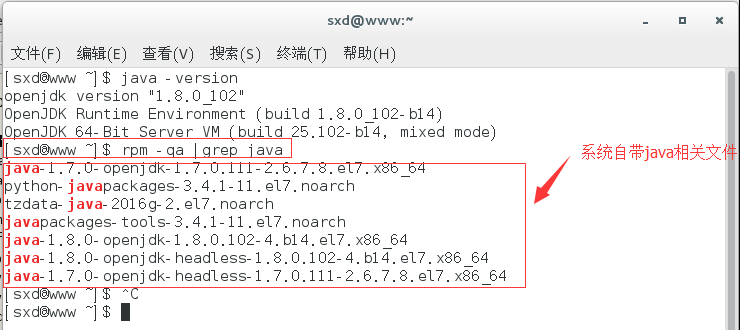
命令说明：

rpm 　　管理套件

-qa 　　使用询问模式，查询所有套件

grep　　查找文件里符合条件的字符串

java 　　查找包含java字符串的文件



以上文件中：

下面这几个可以删除

java-1.7.0-openjdk-1.7.0.111-2.6.7.8.el7.x86\_64

java-1.8.0-openjdk-1.8.0.102-4.b14.el7.x86\_64

java-1.8.0-openjdk-headless-1.8.0.102-4.b14.el7.x86\_64

java-1.7.0-openjdk-headless-1.7.0.111-2.6.7.8.el7.x86\_64

noarch文件可以不用删除

python-javapackages-3.4.1-11.el7.noarch

tzdata-java-2016g-2.el7.noarch

javapackages-tools-3.4.1-11.el7.noarch

③在命令窗口键入：

rpm -e --nodeps java-1.7.0-openjdk-1.7.0.111-2.6.7.8.el7.x86\_64

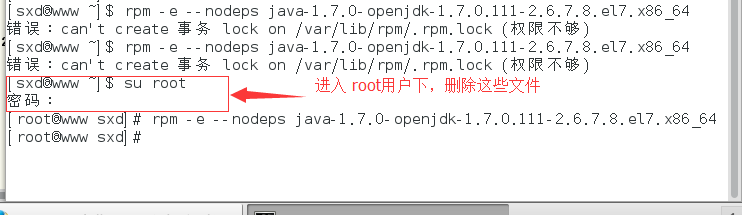
但是会出错，因为在普通用户sxd用户下，并没有操作这几个文件的权限。

解决：

　　在命令窗口键入：

su root

进入root用户，可以有权限操作这几个文件。



完整的删除文件的命令，在命令窗口键入：

rpm -e --nodeps java-1.7.0-openjdk-1.7.0.111-2.6.7.8.el7.x86\_64

rpm -e --nodeps java-1.8.0-openjdk-1.8.0.102-4.b14.el7.x86\_64

rpm -e --nodeps java-1.8.0-openjdk-headless-1.8.0.102-4.b14.el7.x86\_64

rpm -e --nodeps java-1.7.0-openjdk-headless-1.7.0.111-2.6.7.8.el7.x86\_64

命令介绍：

rpm 　　　　管理套件

-e　　　　　删除指定的套件

--nodeps　　不验证套件档的相互关联性

④检查是否已经删除成功

在命令窗口键入：

java -version

IMG_259

代表已经删除成功了。

⑤如果还没有删除，则用yum -y remove去删除他们

### **第二步　下载最新稳定JDK**

**【注意】:JDK安装在哪个用户下，就是给哪个用户使用**

①下载地址为

当前最新版本下载地址：

http://www.oracle.com/technetwork/java/javase/downloads/index.html

历史版本下载地址：

　http://www.oracle.com/technetwork/java/javase/archive-139210.html

**wget --no-cookies --no-check-certificate --header "Cookie: gpw\_e24=http%3A%2F%2Fwww.oracle.com%2F; oraclelicense=accept-securebackup-cookie" "http://download.oracle.com/otn-pub/java/jdk/8u141-b15/336fa29ff2bb4ef291e347e091f7f4a7/jdk-8u141-linux-x64.tar.gz"**

**tar xzf jdk-8u141-linux-x64.tar.gz**

### **第三步　配置JDK环境变量**

①编辑全局变量

在命令行键入：

vim /etc/profile

 进入文本编辑状态下，在文本的最后一行粘贴如下：  
export JAVA\_HOME=/usr/java/jdk1.8.0\_144

### **第四步　让刚刚设置的环境变量生效并检查是否安装成功**

①让刚刚设置的环境变量生效

 键入： source /etc/profile或 . /etc/profile

 ②检查是否配置成功

键入：java -version

## Maven安装

Maven3.3.9安装包下载地址：[https://mirrors.tuna.tsinghua.edu.cn/apache//maven/maven-3/3.3.9/binaries/](https://mirrors.tuna.tsinghua.edu.cn/apache//maven/maven-3/3.3.9/binaries/" \t "https://www.imooc.com/article/_blank)

<https://archive.apache.org/dist/maven/maven-3/3.3.9/binaries/apache-maven-3.3.9-bin.tar.gz>

检查版本：maven -v

## scala-2.11.8.tgz安装

参考教程：<http://www.runoob.com/scala/scala-tutorial.html>

下载Linux版本：<https://downloads.lightbend.com/scala/2.11.8/scala-2.11.8.tgz>

检查版本：scala

## hadoop-2.6.0-cdh5.7.0.tar.gz 安装

教程：<https://www.yiibai.com/hbase/hbase_installation.html#>

wget <http://archive.cloudera.com/cdh5/cdh/5/hadoop-2.6.0-cdh5.7.0.tar.gz>

在外网是可以访问[http://Master:50070](https://link.jianshu.com?t=http://Master:50070" \t "https://www.jianshu.com/p/_blank)

运行yarn测试用例：

**hadoop jar $HADOOP\_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0-cdh5.7.0.jar pi 2 3**

## Spark 安装

1. 下载源码：

**wget <https://archive.apache.org/dist/spark/spark-2.2.0/spark-2.2.0.tgz>**