编译hadoop2.7.3源码并导入eclipse

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 编译hadoop2.7.3源码简单步骤如下：  1：先配置好maven，包括环境变量以及eclipse的maven路径  2：下载hadoop2.7.3-src源文件  3：protoc 2.5.0，他是google的一个数据缓冲高效可扩展的服务包  下载protoc2.5，只有linux的tar.gz和win32的，我的是win10的64位，也可以使用win32的  地址：<https://github.com/google/protobuf/releases?after=v2.6.1>    4：将下载下来的protoc2.5.0里面的exe文件拷贝到windows的system32文件里面    5：在cmd里面进行输入测试，发现有版本信息就ok了    6：进入下载下来的hadoop源文件的hadoop-maven-plugins目录中，比如hadoop2.7.3    执行mvn install命令  （执行mvn install ，这个过程比较久，需要下载很多的jar包），最好将maven的仓库镜像设置成国内的阿里云  这样速度快很多    （出现如上画面，有BULID SUCCESS就说明编译好了）  在cd到hadoop-2.7.3-src 目录下  7：在hadoop2.7.3-src目录下执行    设置成eclipe的工程结构项目    （如果看到如上的BUILD SUCCESS，那么我们就创建成功了）  8：将编译的项目导入到eclipse中  9：一般设置好了以后，导入eclipse中会有三类错误   |  |  | | --- | --- | | **如果出现访问限制：由于对必需的库 C:/Program Files/Java/jre6/lib/rt.jar 具有一定限制，因此无法访问类型 AudioPlayer**  **等等之类的** |  | | **AvroRecord不是一个类型** | 执行这个命令，这个avro-tools-1.7.7.jar下载地址  <http://mvnrepository.com/artifact/org.apache.avro/avro-tools>  下载的这个jar包我是放到 | | **EchoRequestProto无法解析为类型** | 执行上面的这个命令就行了 | | **hadoop-streaming里面的build path有问题**， | 右键hadoop-streaming项目-properties->左边Java Build Path->Source->选定错误项remove掉引用 |   设置好了以后，就可以查看hadoop2.7.3的源码了 |
|  |