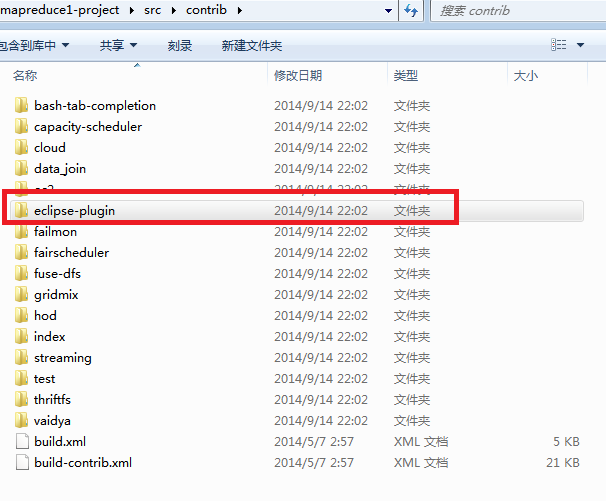
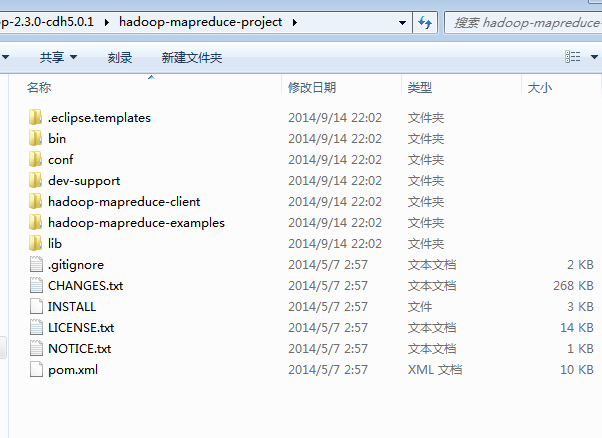
[cloudera CDH（5）开发方式及CDH eclipse插件编译总结](http://blog.csdn.net/qiezikuaichuan/article/details/46829679)

**问题导读**

**1.CDH有几种开发方式？**  
**2.如何在CDH源码中，如何查找eclipse插件源码？**  
  
http://www.aboutyun.com/static/image/hrline/4.gif  
  
  
  
  
我们知道hadoop的开发方式有两种  
**hadoop插件开发方式**  
**1.插件开发**  
对于插件开发，有些小麻烦。使用插件最好自己编译，因为eclipse插件能否连接成功，不止跟你的配置和集群有关系，还和你的eclipse版本有关系。这样造成很多新手链接不成功.  
  
**2.无插件开发**  
无插件开发，对于新手则需要熟悉Linux，因为你上传的文件，删除文件，没有使用插件方便。插件可以帮助你删除文件、上传文件后，不用通过Linux即可看到。插件开发的优点则是你只需要添加hadoop的相关包即可，不会因为插件问题在程序的运行过程中产生错误，导致找不到错误原因。  
  
更详细可以查看  
[**hadoop开发方式总结及操作指导**](http://www.aboutyun.com/forum.php?mod=viewthread&tid=6950&highlight=hadoop%2B%BF%AA%B7%A2%B7%BD%CA%BD)  
[**从零教你在Linux环境下（ubuntu 12.04）如何编译hadoop2.4**](http://www.aboutyun.com/forum.php?mod=viewthread&tid=8130&highlight=%B4%D3%C1%E3%2B%2B%2Bhadoop%2B%2B%2B%2B%B1%E0%D2%EB)  
  
  
**CDH开发方式：**  
同样也有两种，并且它们的开发有所区别，但也是大同小异。  
  
**1.插件开发**  
插件开发，有很多同学找不到eclipse插件源码，自己想编译无门，所以这里给大家介绍一个方法，该如何找到eclipse插件源码，那就是全局搜索。因为不同的版本，插件位置有所不同，这是个万能方法。  
那么CDH5插件在什么位置：  
  
{hadoop-home}\hadoop-mapreduce1-project\src\contrib  
我这里是  
hadoop-2.3.0-cdh5.0.1\hadoop-mapreduce1-project\src\contrib  
  
   
  
  
下面还有关于eclipse文件夹，供大家参考，它的用途以后在研究。  
   
  
  
**2.无插件开发**  
无插件开发也是通过加入一些jar包，然后打包到集群运行  
  
http://www.aboutyun.com/static/image/hrline/2.gif  
  
开发方式我们知道了，那么剩下的就是编译了。该如何编译，下面（centos 6.2为例）内容可以参考：

由于与eclipse端不兼容，各种神奇，把cdh4降成cdh3u5，然后又降成cdh3u4，总之很烦，后来看到了篇文章关于如何自己编译eclipse plugin的jar包，然后自己修改修改，在这里跟大家分享分享

首先，无论你的服务器上Hadoop 使用的是什么版本，你都需要下载对应的这个版本的源代码包进行编译. 你是Apache Hadoop 0.20.2 就去Apache 下0.20.2 的，你是Apache 0.21 版本的就去下0.21 版本的. 你是Cloudera CHDu.x 版本的，就去Cloudera 网站上下它的CDHu.x  版本对应的tar 包. 确保你要连接的服务器版本和你打算编译的eclipse-plugin版本是一致的.

在自己编译eclipse-plugin之前，你需要apache-ant, apache-maven, apache-ivy如果你打算编译整个包括hadoop的服务器版本并生成二进制包，你还需要apache-forrest 用来输出文档. ant,maven,forrest 的安装跟java 的安装没有区别，解压，然后添加对应的bin目录到path 变量. ivy 的安装就是添加一个apache-ivy-xxx.jar 到你的ant  安装目录下的lib 目录. 这些前提工具就算安装好了.

然后从命令行进入你解压的$hadoop\_home （cdh4的源码包中的目录为/root/hadoop-2.0.0-cdh4.1.1/src/hadoop-mapreduce-project）源代码目录, 执行ant compile-core , 这个会编译基础的hadoop-core 包. 注意如果你是在linux 下这个命令不会报错，但是如果你是windows 会报can’t run  program mvn , 在对应的build.xml 文件的42 行，你需要进入$Hadoop\_home/build.xml 文件的42行，修改

1. <exec  executable="mvn"  xxxxxxxxxxxxxxxxxxxxx

改成

1. <exec  executable="mvn.bat"  xxxxxxxxxxxxxxxxxx

如果你打算编译整个包的话对应的1230  行里面的forrest 在windows 下同样会出现这个错误.

然后你需要将$Hadoop\_home 下面的hadoop-core-xxx.jar 文件拷贝到$Hadoop\_home/build 目录下. 你是什么版本的就拷贝对应版本的hadoop-core-xxx.jar 文件. 比如我编译的是hadoop-core-0.20.2-cdh3u4.jar 文件.

进入$Hadoop\_home/src/contrib/eclipse-plugin 目录, 执行

1. ant –Declipse.home=/path/to/eclipse  -Dversion=0.20.2-cdh3u4 jar

注意，如果你没有执行前面一步的copy 工作，你会报错找不到build/hadoop-core-xxx.jar 文件. 你也可以修改对应eclipse-plugin 目录下的build.xml 文件的69行左右对应的两个copy 任务.

1. <copy file="${hadoop.root}/build/hadoop-core-${version}.jar" tofile="${build.dir}/lib/hadoop-core.jar" verbose="true"/>
2. <copy file="${hadoop.root}/build/ivy/lib/Hadoop/common/commons-cli-${commons-cli.version}.jar"  todir="${build.dir}/lib" verbose="true"/>

确保copy 的这些文件能够在对应的目录找到. 如果你是0.21 版本的apache-hadoop ，由于0.21 版本完全maven 化了. 你可能需要添加hadoop-core-xxx , hadoop-mapreduce-xxx.jar hadoop-hdfs-xxx.jar , 我的是CDHu4 的，你在执行完这个之后在$hadoop\_home/build/contrib/eclipse-plugin 下面会有一个hadoop-eclipse-plugin-0.20.2-cdh3u4.jar 的文件. 不过这个文件是不能用的，会提示你差一些jar 文件. 你需要添加 jackson-core-asl-xxx.jar , jackson-mapper-asl-xxx.jar , commons-httpclient-xxx.jar 文件. 具体的version 你需要到$hadoop\_home/lib 目录下自己找. apache-hadoop 0.20.205 版本的好像用的是jackson-core-1.0.1版本的. CDHu4 版本用的好像是jackson-core-1.5.2 版本的.

1. <copy file="${hadoop.root}/lib/jackson-core-asl-1.5.2.jar" todir="${build.dir}/lib" verbose="true"/>
2. <copy file="${hadoop.root}/lib/jackson-mapper-asl-1.5.2.jar" todir="${build.dir}/lib" verbose="true"/>
3. <copy file="${hadoop.root}/lib/commons-httpclient-3.1.jar" todir="${build.dir}/lib" verbose="true"/>

编译好之后放在eclipse 的dropin 目录下就可以了.

注意连接的时候MapReduce-Master port  别和DFS port 填反了，不然会报fail to open /xxxx 的错误. 默认安装的时候应该是mapreduce 是8021 , dfs  是8020 , 查看你对应服务器上的端口配置

如果出现这样的问题

（1）/src/contrib/eclipse-plugin/build.xml:62: warning: 'includeantruntime' was not set, defaulting to build.sysclasspath=last; set to false for repeatable builds，则是ant版本有点高，1.8以上需要设置一下includeantruntime

具体设置

<javac.....includeantruntime=false>在其中加上这一项即可

(2)/root/hadoop-2.0.0-cdh4.1.1/src/hadoop-mapreduce-project/src/contrib/eclipse-plugin/bu exist出现这个问题，需要到这个目录下

1. SRC\_BASE\_DIR/hadoop-common/hadoop-mapreduce-project

执行命令mvn -DskipTests install即可

(3).编译心得，开始的时候不会整，按照网上介绍的开始弄，总是会有这样那样的奇葩的问题，最近总结了一些，跟大家分享下，每次遇到错误的时候，第一反应就是百度，谷歌，其实我们自己可以根据错误提示，像编译cdh4的时候遇到这个错误

1. /root/hadoop-2.0.0-cdh4.1.1/src/hadoop-mapreduce-project/src/contrib/eclipse-plugin/build.xml:72: /root/hadoop-2.0.0-cdh4.1.1/src/hadoop-mapreduce-project/build/ivy/lib/Hadoop/common does not exist.
2. vim /root/hadoop-2.0.0-cdh4.1.1/src/hadoop-mapreduce-project/src/contrib/eclipse-plugin/build.xml

看看它的第72行是什么，虽然里面语言看不太懂，但也大致可以理解，

1. <copy tofile="${build.dir}/lib/hadoop-common.jar">
2. <fileset dir="${hadoop.root}/build/ivy/lib/Hadoop/common" includes="hadoop-common-\*.jar" excludes="hadoop-common-\*tests.jar"/>
3. </copy>

可以看出是将${hadoop.root}/build/ivy/lib/Hadoop/common目录下的hadoop-common-\*.jar包（除了hadoop-common-\*tests.jar）包拷贝到

${build.dir}/lib/hadoop-common.jar

这就能大致理解了，去找${hadoop.root}/build/ivy/lib/Hadoop/common目录，发现没有这个目录，于是创建这个目录

然后问题又来了，紧接着报了这个错/root/hadoop-2.0.0-cdh4.1.1/src/hadoop-mapreduce-project/src/contrib/eclipse-plugin/build.xml:72: Cannot perform operation from directory to file

当时就恍然大悟，刚创建的那个目录里面没有hadoop-commons这个包，然后find 命令，找到这个包hadoop-common-2.0.0-cdh4.1.1.jar，放进去

然后就成功了，其实都很简单，只是我们不注意，这些错误，如果直接百度，谷歌的话，还真是收获很少