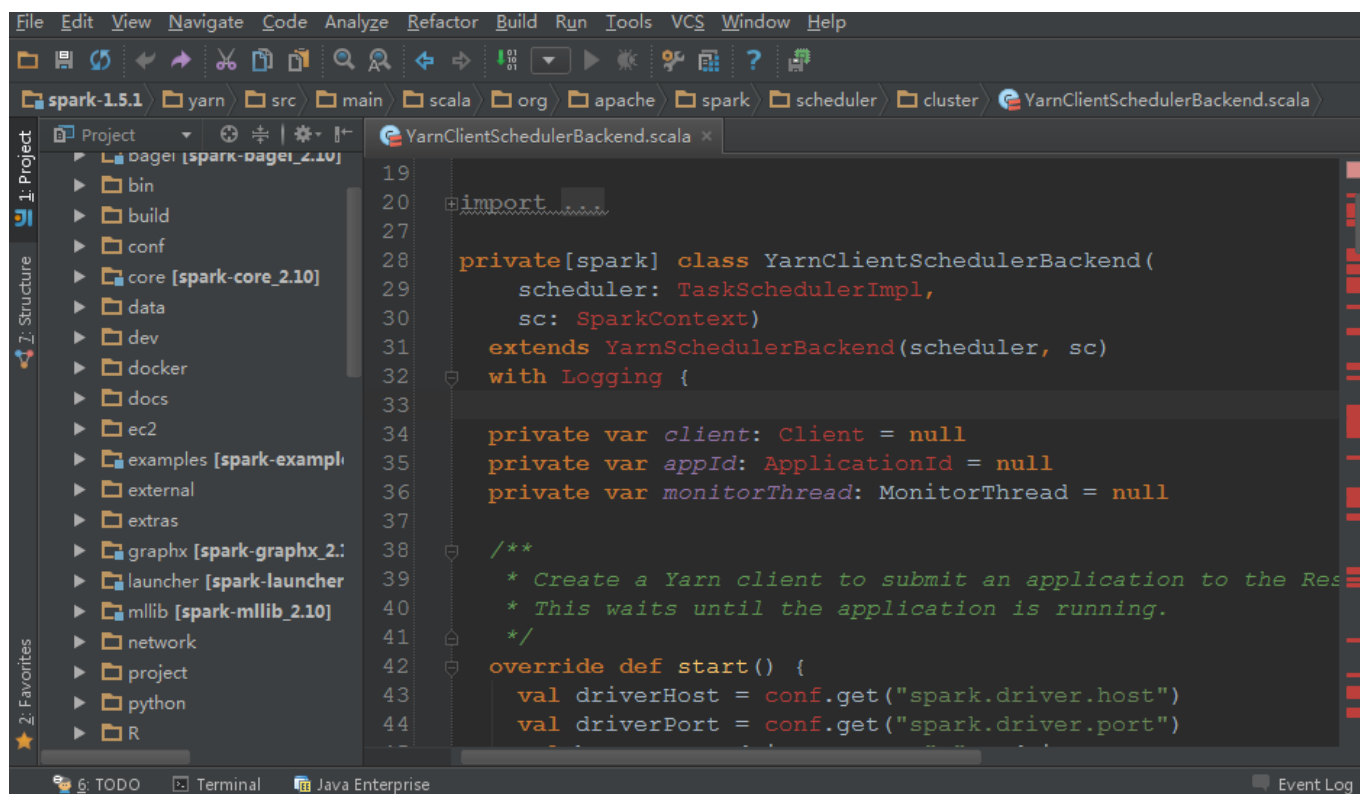
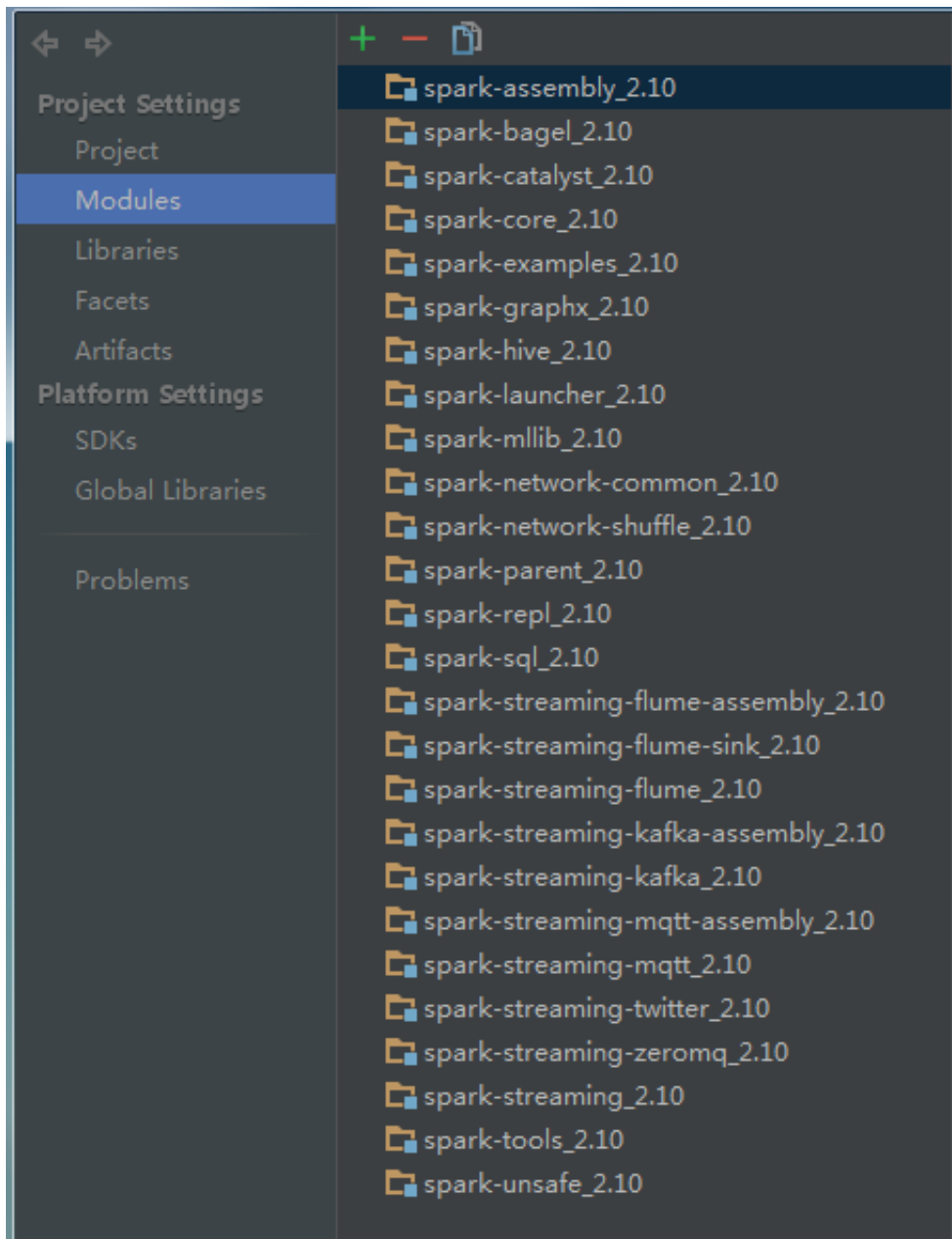


怎么在Idea IDE里面打开Spark源码而不报错

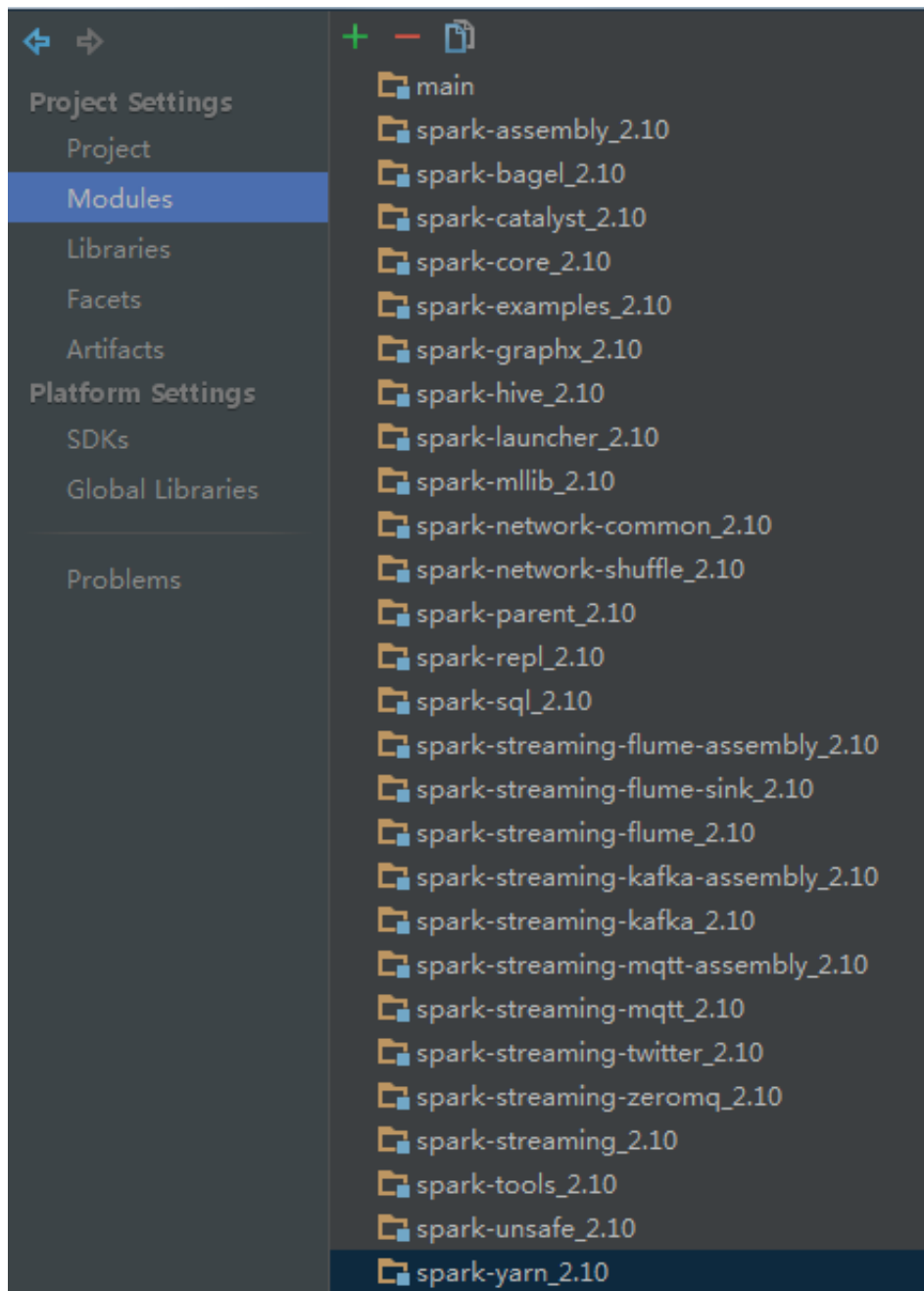
我们在学习或者使用Spark的时候都会选择下载Spark的源码包来加强Spark的学习。但是在导入Spark代码的时候，我们会发现yarn模块的相关代码总是有相关类依赖找不到的错误（如下图），而且搜索（快捷键Ctrl+N）里面的类时会搜索不到！这给我们带来了很不便。



本文就是来解决这个问题的。我使用的是Idea IDE工具阅读代码的，操作系统是window，其他的系统和IDE我没试过。我们先点击这个工程的Project Structure菜单，这时候会弹出一个对话框，部分截图如下：



仔细的用户肯定会发现里面列出来的模块(Module)居然没有yarn！就是这个原因导致yarn模块相关的代码老是报错！我们只需要将yarn模块加入到这里即可。步骤依次选择 Add->Import Module->选择pom.xml，然后一步一步点击确定，这时候会在对话框里面多了spark-yarn_2.10模块，如下：



然后点击Maven Projects里面的Reimport All Maven Projects，等yarn模块里面的所有依赖全部下载完的时候，我们就可以看到这个模块里面的代码终于不再报错了！！

```
import org.apache.hadoop.yarn.api.records.{ApplicationId, YarnApplicationState}

import org.apache.spark.{SparkException, Logging, SparkContext}
import org.apache.spark.deploy.yarn.{Client, ClientArguments, YarnSparkHadoopUtil}
import org.apache.spark.scheduler.TaskSchedulerImpl

private[spark] class YarnClientSchedulerBackend(
  scheduler: TaskSchedulerImpl,
  sc: SparkContext)
  extends YarnSchedulerBackend(scheduler, sc)
  with Logging {

  private var client: Client = null
  private var appId: ApplicationId = null
  private var monitorThread: MonitorThread = null
```

甚至你都可以在这里面编译和调试Spark代码了，终于不用担心出错了。。现在赶紧去学习Spark的代码吧。。

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】（）](#)