

## 14-mapreduce-java代码- ( 本地和集群 )

-----成都尚学堂-mr-zeng-----

### Eclipse安装mapreduce插件

把我们的"hadoop-eclipse-plugin-2.7.0.jar"放到Eclipse的目录的"plugins"中，然后重新Eclipse即可生效。

- a) 提供可视化操作hdfs
- b) 方便编写mapreduce程序

...

### MapReduce程序的基本结构

#### 新建MapReduce项目

右键--> new--> other--> map / reduce project

**程序执行入口类**：整个Mapreduce程序执行的入口，这里配置任务的执行信息，任务信息，map和reduce的类，输出数据的类型，程序输入数据（要计算的数  
据），程序输出数据（计算的结果数据）。。。

右键--> new--> other--> mapreduce driver

**Map执行入口类**：这里是map阶段的执行代码

右键--> new--> other--> mapper

**Reduce执行入口类**：这里是reduce阶段的执行代码

右键--> new--> other--> reducer

### Api位置

hadoop1位于mapred包

hadoop2位于mapreduce包

### 相关类 (

Configuration：配置信息类，hadoop环境配置

如：reduce任务数，mapreduce运行方式（local，yarn）

Job：任务类，任务环境配置

如：任务名，任务各入口类，任务数据类型等。

FileInputFormat：map输入文件的类，配置输入信息

如：输入路径

FileOutputFormat：reduce输出文件的类，配置输出信息

如：输出路径

Mapper：用于map执行类基础的基类，提供模版功能，和运行入口map方法

Reduce：用于reduce执行类基础的基类，提供模版功能，和运行入口reduce方法

### 例子

实现单词统计的mapreduce代码（local和yarn分别执行）