**1.**       **Local debug**

《大数据平台核心技术》实践2：编写MR完成Group By+Join操作，大家可以先自己本地验证通过之后，再在学堂在线上提交代码：

附件解压之后，进入对应目录

**1.1**    **编译并打包**

$javac -cp .:odps/lib/\* src/example/Join.java -d .

$jar -cf xdj\_example.jar example/

**1.2**    **local执行**

open\_bigdata这个project中两张源表orders和customers都有一些测试数据。

本地执行时，会从odps上的源表copy 100行数据到本地（放到了本地warehouse目录下），然后执行各位的逻辑，同样在本地生成了对应的输出表（join\_example\_out）

注意：文件是windows格式，可能需要将其转换为unix格式

$./odps/bin/odps

Type `help` for more information.

odps:open\_bigdata> jar -classpath .:./\* -local example.Join

Usage: Join <orders><customers><join\_example\_out>

odps:open\_bigdata> jar -classpath .:./\* -local example.Join  orders customers join\_example\_out;

#执行信息

OK

**1.3**    **查看源表内容和结果表内容，来验证MapRedcue逻辑是否符合预期**

每个表有一个对应的目录

$ls warehouse/open\_bigdata/\_\_tables\_\_/

customers  join\_example\_out  orders

每个表目录有两个文件：schema文件、和数据文件

$ls warehouse/open\_bigdata/\_\_tables\_\_/join\_example\_out/

R\_000000  \_\_schema\_\_

**2.**       **同时邀请各位同学体验ODPS各个功能**

**2.1**    **注册并邮件申请授权**

注册链接：[www.aliyun.com/](http://www.aliyun.com/)

我们有一个免费的Project（open\_bigdata）可以share给大家来用。注册成功之后，邮件我（邮件内容包括姓名、阿里云上的注册账号），我会给大家授权，授权成功会给大家邮件。

注册登录之后，会进入”管理控制台”，在”AccessKeys”中”创建Access Key”。

**2.2 修改odps.conf并执行SQL或MapReduce**

修改附件中的odps/conf/odps.conf中的access\_id和access\_key，就可以开始使用odps了，比如查看open\_bigdata中有哪些表：

$./odps/bin/odps

Type `help` for more information.

odps:open\_bigdata>sql

odps:sql:open\_bigdata> show tables;

[ALIYUN$bigdata\_course@aliyun.com:customers](mailto:ALIYUN$bigdata_course@aliyun.com:customers)

[ALIYUN$bigdata\_course@aliyun.com:gby\_out](mailto:ALIYUN$bigdata_course@aliyun.com:gby_out)

[ALIYUN$bigdata\_course@aliyun.com:join\_example\_out](mailto:ALIYUN$bigdata_course@aliyun.com:join_example_out)

[ALIYUN$bigdata\_course@aliyun.com:join\_out](mailto:ALIYUN$bigdata_course@aliyun.com:join_out)

[ALIYUN$bigdata\_course@aliyun.com:orders](mailto:ALIYUN$bigdata_course@aliyun.com:orders)

OK

跑一个join语句：

odps:sql:open\_bigdata> select \* from orders o join customers c on o.customer\_id = c.customer\_id limit 1;

+------------+-------------+-------------+------------+-------------+---------------+---------+------+---------+

| order\_id   | customer\_id | employee\_id | order\_date | customer\_id | customer\_name | address | city | country |

+------------+-------------+-------------+------------+-------------+---------------+---------+------+---------+

| 10254      | CHOPS       | 5           | 1996-07-11 00:00:00.000 | CHOPS       | Yang Wang     | Hauptstr. 29 | Bern | Switzerland |

+------------+-------------+-------------+------------+-------------+---------------+---------+------+---------+

这样大家，可以按文档说明，使用一下ODPS的SQL和MapReduce功能了，文档见：

<http://help.aliyun.com/product/8314999_odps.html?spm=5176.750001.2.16.arsusn>

**NOTE：附件解压之后能够拿到课上Join示例实现Join.java**

谢德军