

Talend data integration对接oracle提取过滤数据导入数据库新表

发表者：admin 分类：数据库 2022-08-26 16:39:40 阅读[303]

Talend data integration对接oracle提取过滤数据导入数据库新表

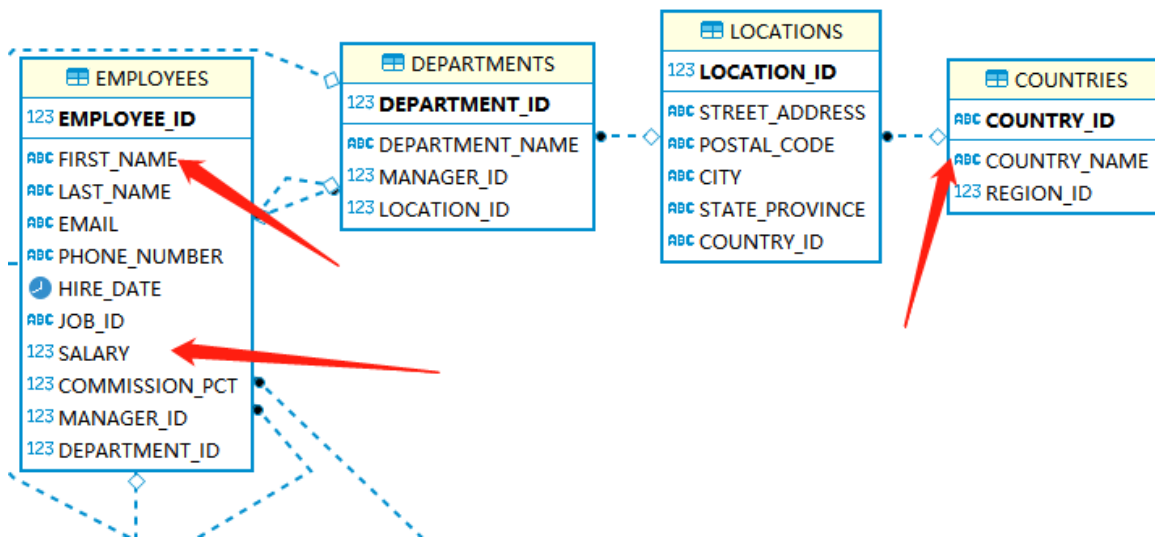
1,环境介绍。

Oracle 19C IP：192.168.1.252 账号:hr 密码: 123456789,使用oracle 19c的hr账号schema自带数据

数据集成工作: talend data intagratiion 8.0

2，利用 talend 数据集成工作，生成一个新表 emp_con，包含 name,salary,country_name，
将hr账号下的数据提取出来，再过滤后，导入新表emp_con。

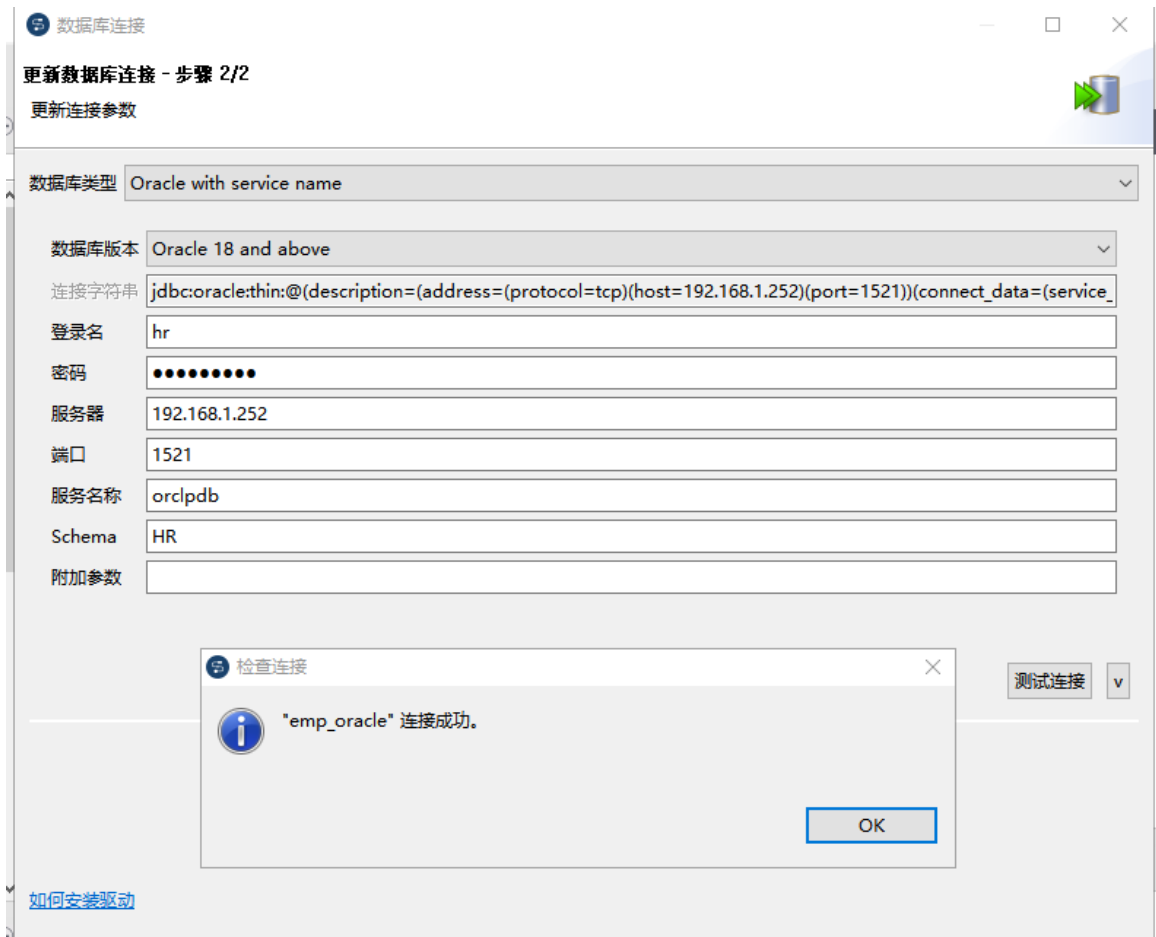
3，HR模式下的ER图如下：



4，打开 talend 软件，新建一个emp_con工程，再新建作业文件夹，然后新建作业名为：EMP_SAL_Country

5，点击 元数据，数据库连接 点右键 新建连接 emp_oracle，选择 数据库类型 oracle with service name 后，填写相关账号与密码，ip等信息。

如下图



6，要提取 country_name, First_name, Salary, 需要关联操作四个表，以下是相关运行SQL。

```
SELECT e.FIRST_NAME ,e.SALARY ,e.DEPARTMENT_ID FROM HR.EMPLOYEES e ;
```

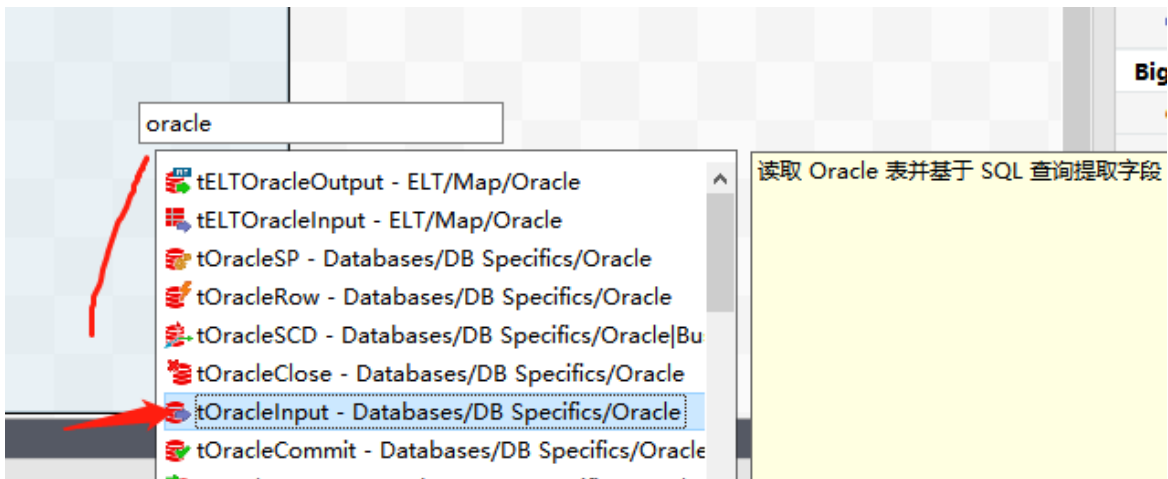
```
SELECT d.DEPARTMENT_ID ,d.LOCATION_ID FROM HR.DEPARTMENTS d ;
```

```
SELECT l.LOCATION_ID ,l.COUNTRY_ID FROM HR.LOCATIONS l ;
```

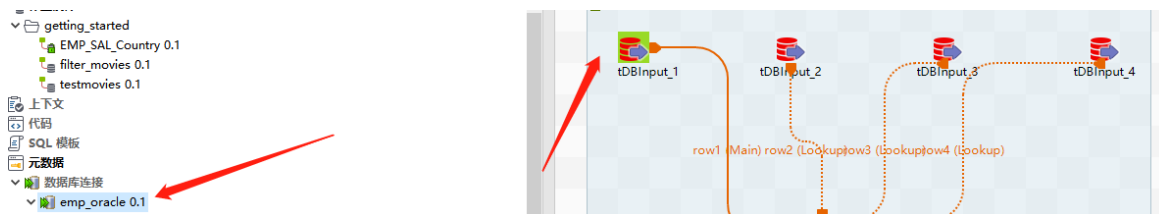
```
SELECT c.COUNTRY_ID ,c.COUNTRY_NAME FROM HR.COUNTRIES c ;
```

7，根据以上SQL，双击 EMP_SAL_country 作业，给作业添加组件，在空白作业处，点击后，输入 oracle，再找到 tOracleInput。

注意，这里添加4个一样的。

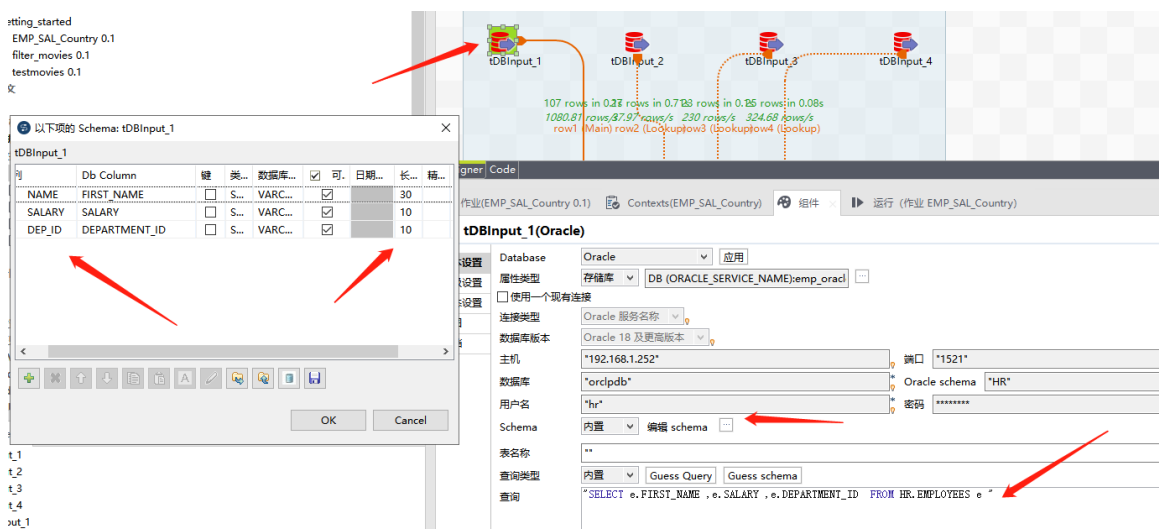


8, 将左边 元数据中的 数据库连接 --> emp_oracle , 拖到 四个组件中。

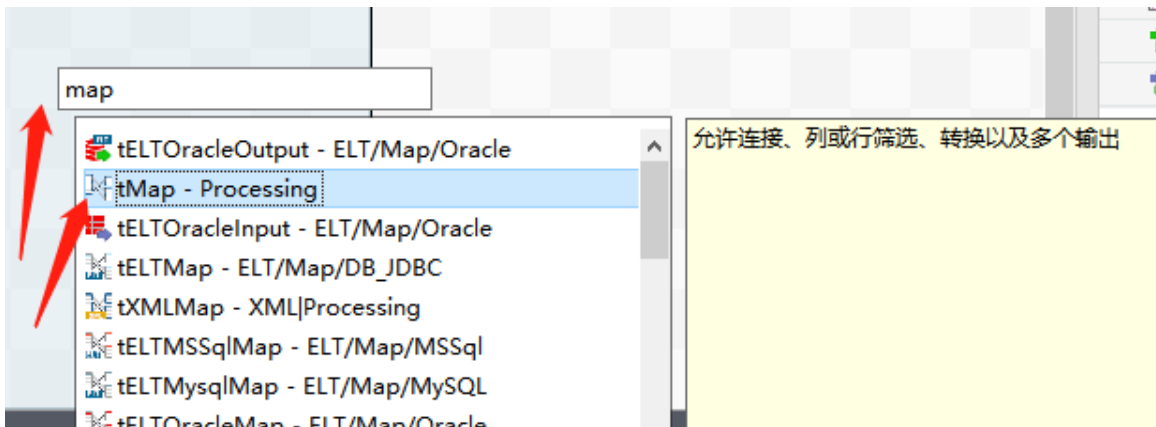


9, 双击 tDBInput_1 组件, 点击 编写schema, 填写相关 输入的 字段名称, 数据类型, 长度(可以不填写, 看情况填写)等信息, 再后填写查询SQL。

4个组件都要填写。其它3个类似。



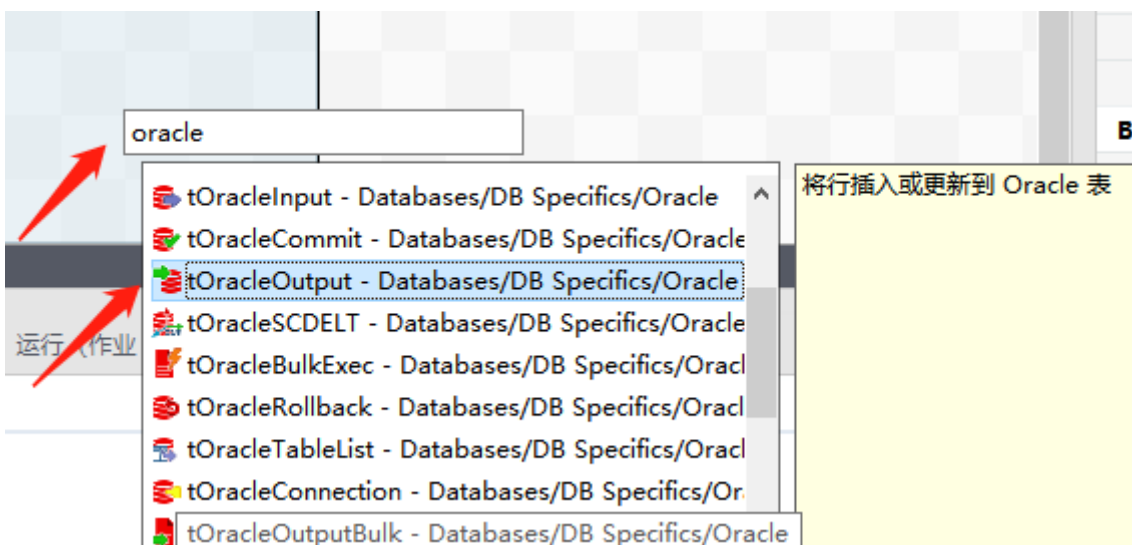
10, 在4个组件下方 新建 tmap组件, 一样的方法, 空白地方 输入 map 会自己提示, 这里选择 tMap - Processing.

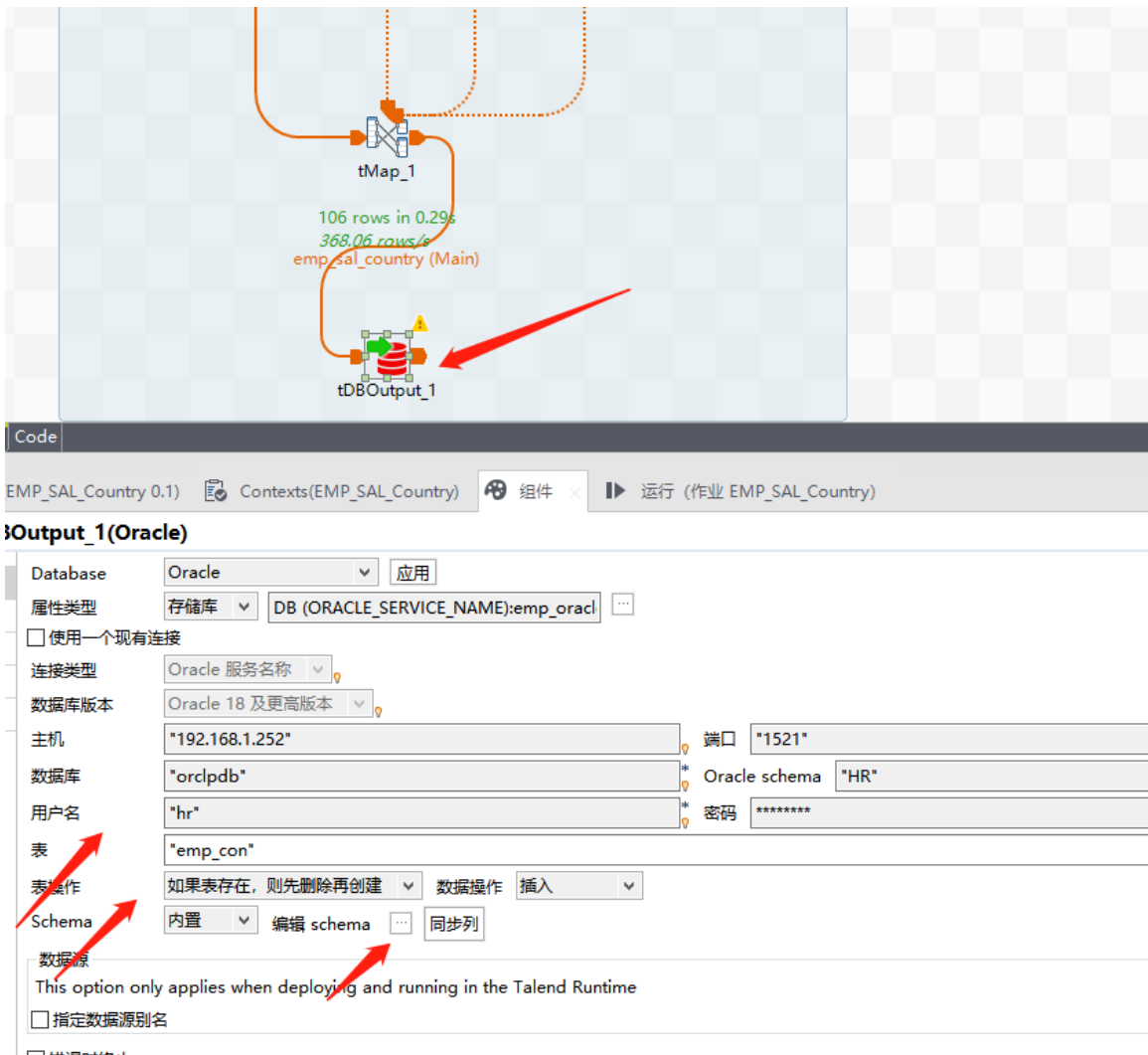


11，对4个 tDBInput组件 点 右键，从上下文菜单中选择 Row (行) > Main (主)，然后单击 tMap 组件，tDBInput 现在即已连接到 tMap。

12，添加 数据库输出 组件，添加两个 tOracleOutput 组件。这些组件用于将已处理的信息写入指定的数据库表。

点击 tMap 点 右键 选择 Row (行) > new output (主) 连线到 tDBOutput输出集。如有提示，填写 新输出连接 的名称 emp_sal_country。





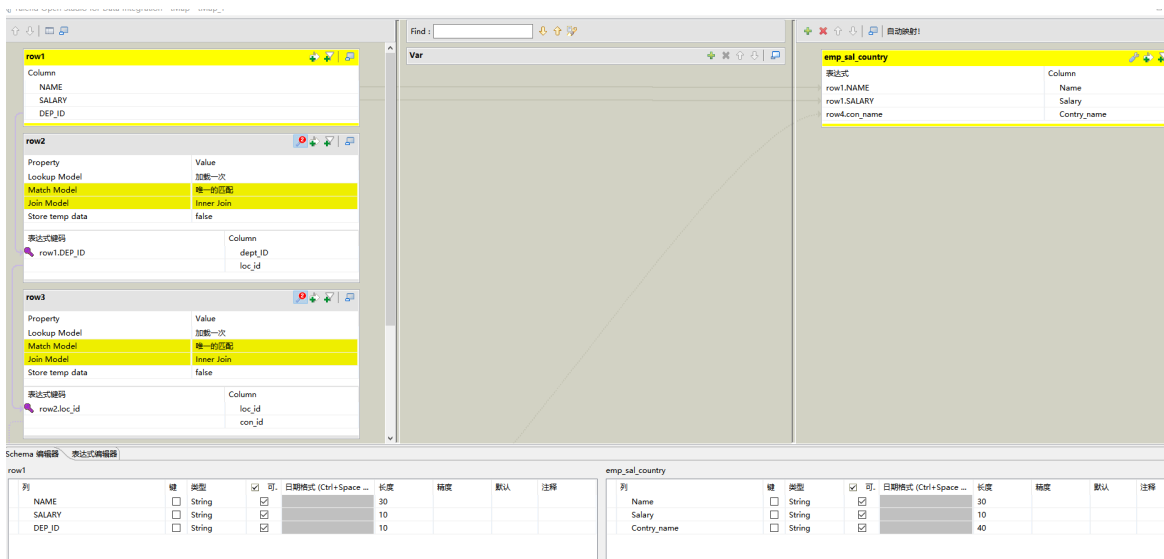
13, 双击 tMap 组件，打开 Map 编辑器。将 row1 中的 dep_id 拖到row2 中的dep_id，创建两个输入数据集之间的连接。

单击 tMap settings (tMap 设置) 按钮，然后单击 Join Model (连接模型) 的 Value (值) 字段，再单击出现的 [...] 按钮打开 Options (选项) 对话框。

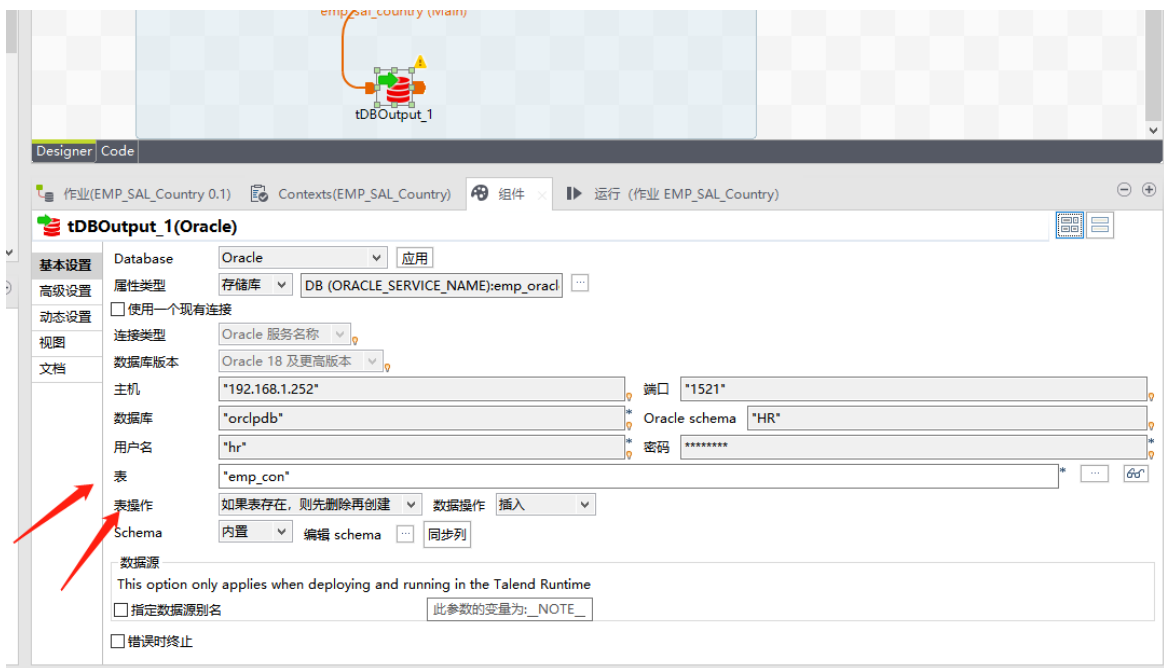
在对话框中，选择 Inner Join (内连接)，然后单击 OK (确定) 将联接定义为内连接。

14, 其它数据集类似操作。最后将 row1的 NAME，SALARY，COU_NAME 三列拖到 右边 emp_sal_country数据集中。最后点 确定。

最后如下图：



15, 最后 将左边 元数据中的 数据库连接 --> emp_oracle 拖到 tDBOutput 组件中。填写下图中的输出 表名 : emp_con



19,最后点击 运行 , 提示运行成功, 查看 emp_con表, 数据已经 生成。

可以将 作业 构建 成软件包, 支持win,linux等系统。



无报错，就是执行完成。