

1.1 Talend 总体介绍

1.2 Talend Studio安装

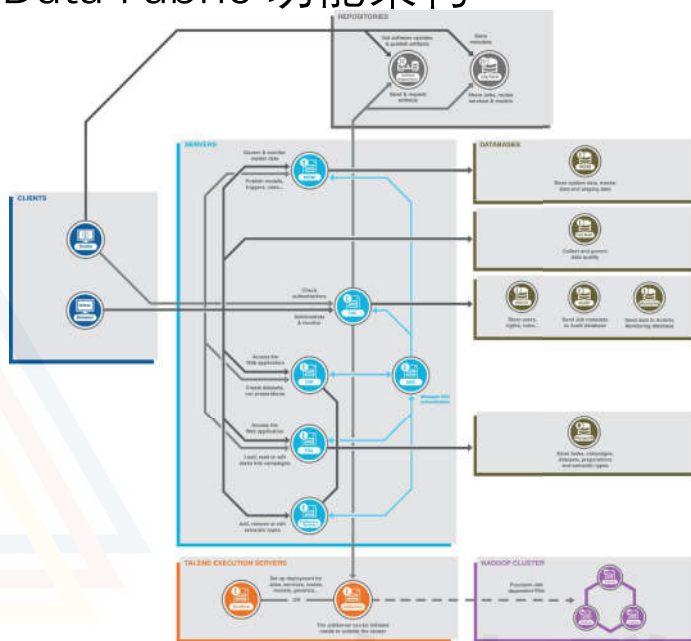
1.3 Talend Studio 介绍

1.4 Talend开发模式介绍

- 创建个标准job
- 文件处理

1.1 Talend 总体介绍

Talend Data Fabric 功能架构



3GOLDEN

金电联行


1.2 Talend Studio安装

- 客户端安装文件下载

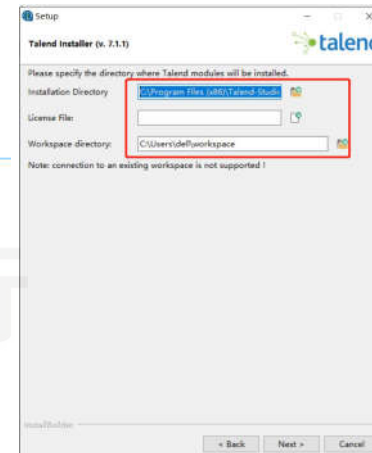


- Talend Studio 安装

打开客户端安装文件
TalendStudio-20181026_1147-
V7.1.1-windows-installer

 TalendStudio-20181026_1147-V7.1.1-windows-installer.exe

填写安装路径和workspace路径
以及license路径，然后一直默认
下一步就可以

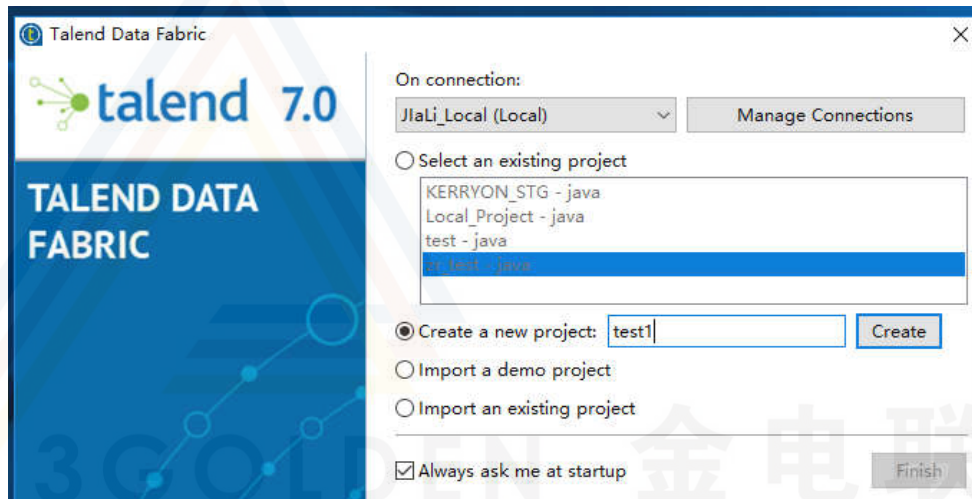


1.3 Talend Studio 介绍

1、双击图标，打开Talend Studio



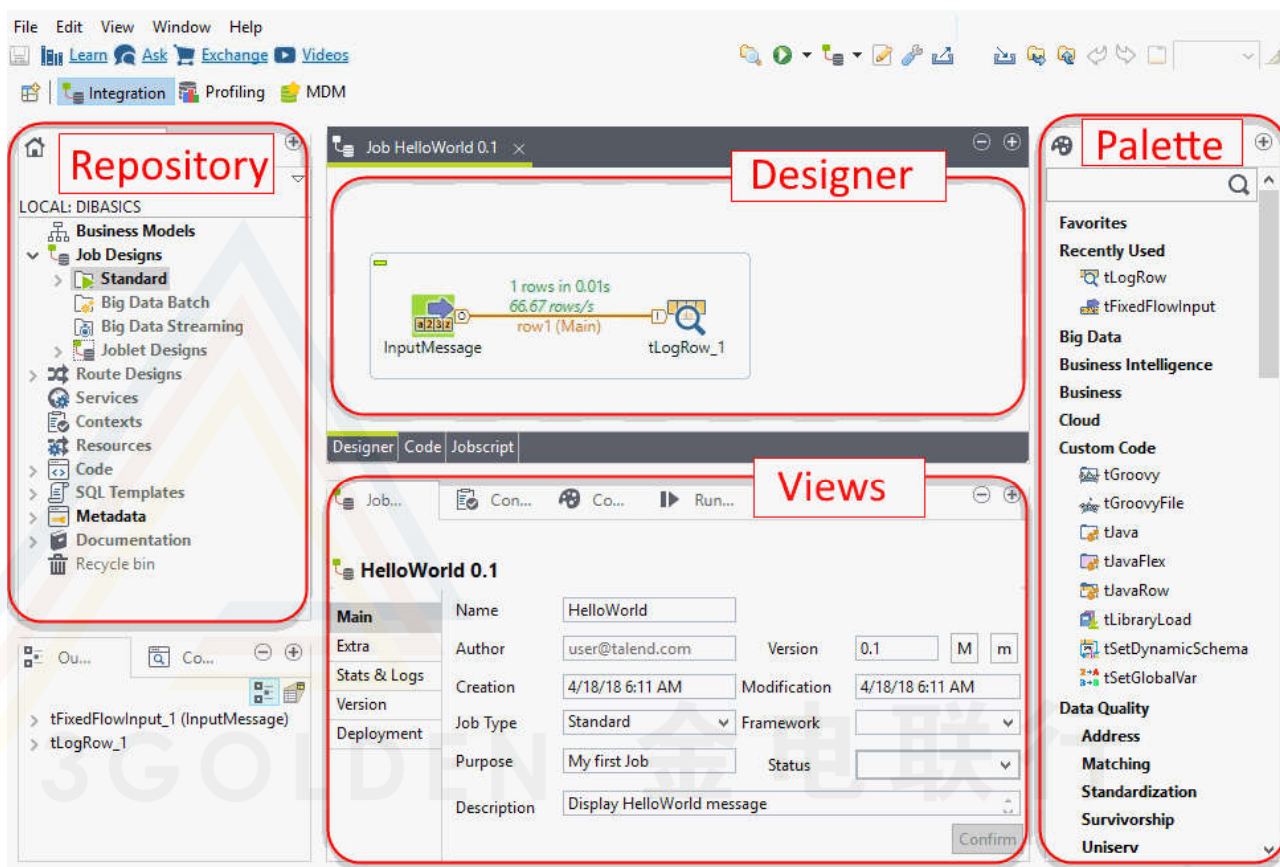
2、创建一个Project，点击finish



3、选择Integration

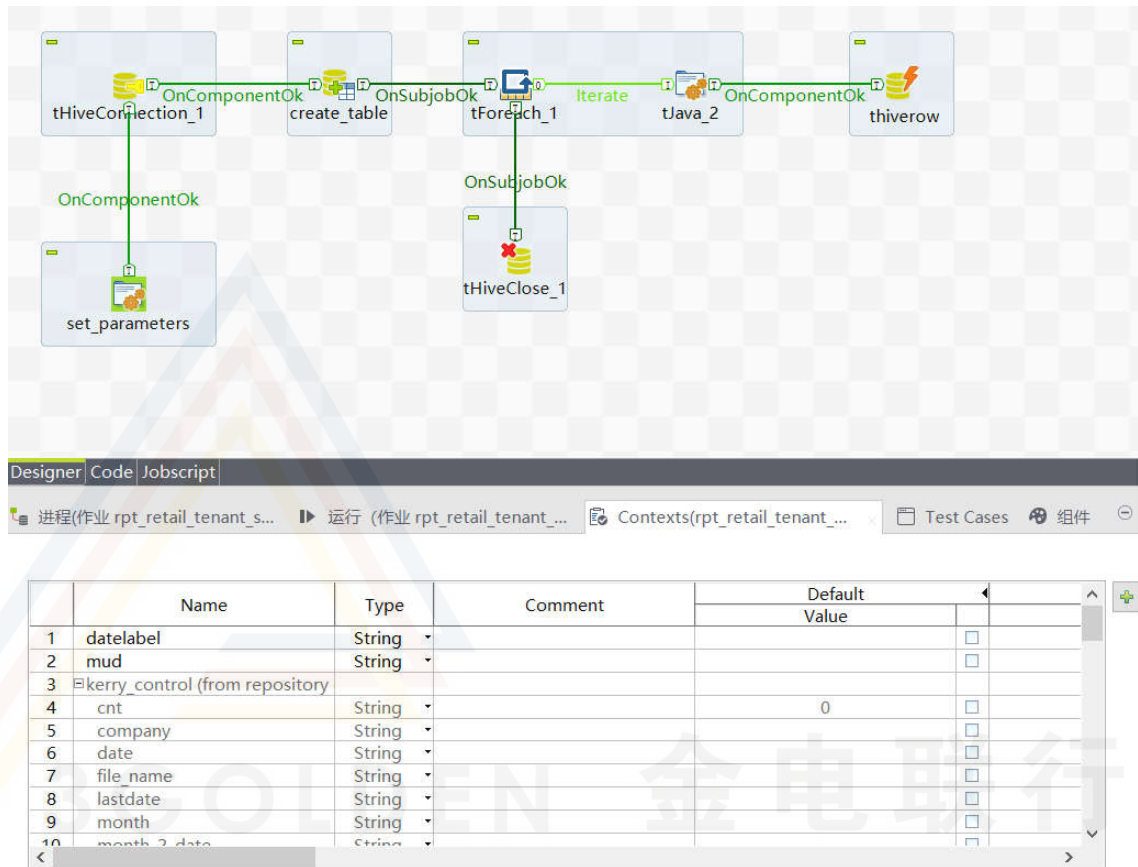


4、开发设计job的操作界面

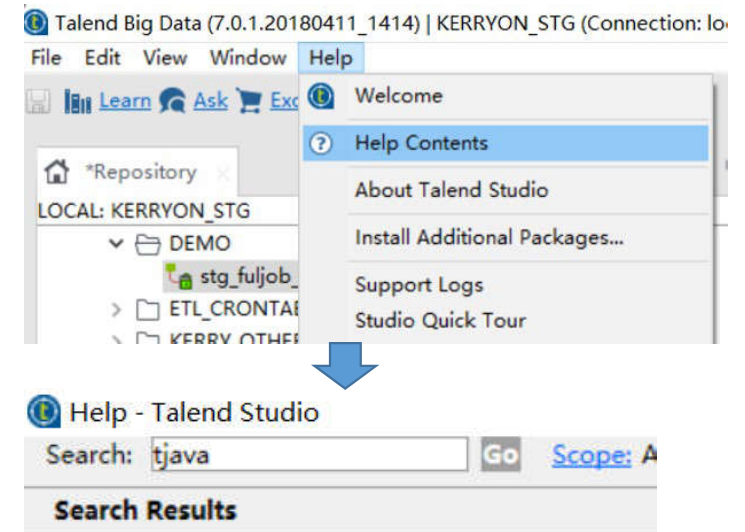


Repository:存储库
包含Job Designs, Contexts,
Metadata等
Palette: 组件界面
提供可添加的组件
Designer: 设计者界面
Views: 视图界面
包含job 视图, 上下文视图, 组件视
图, 运行视图等

1.4 Talend开发模式



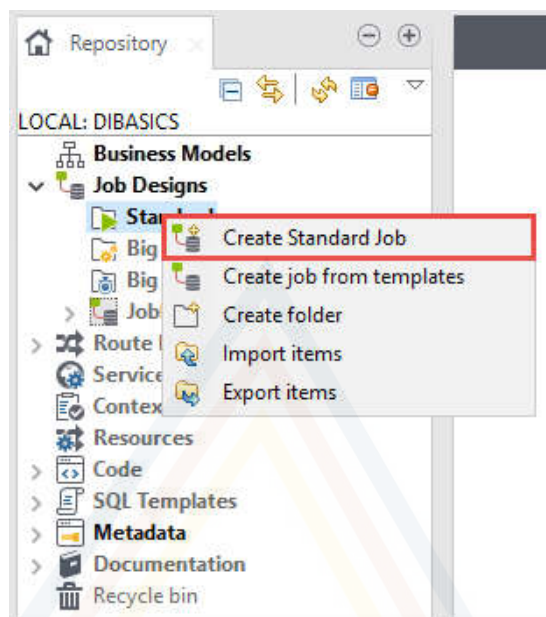
Tips:想了解组件的功能点Help>Help Contents, 然后在搜索框中输入你想了解的组件名称



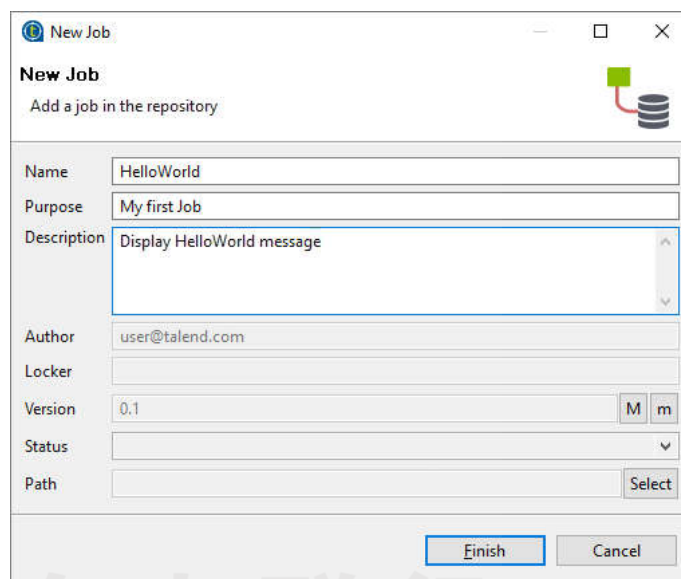
想知道更多的开发案例可以上网搜索help.talend.com

创建一个标准job (STANDARD JOB)

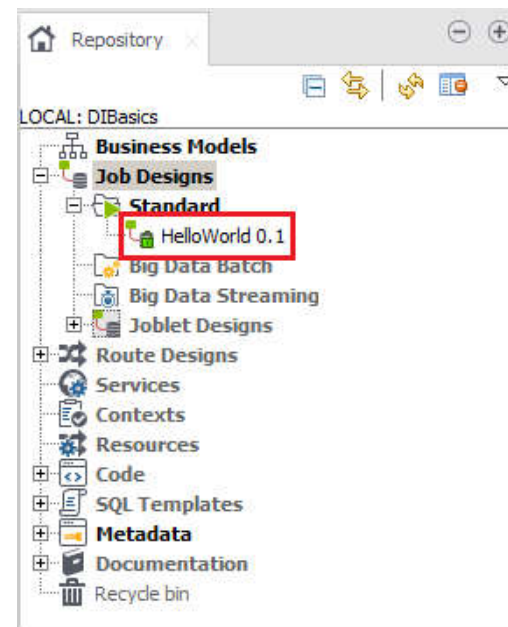
第一步：Repository > Job Designs,选择Create Standard Job



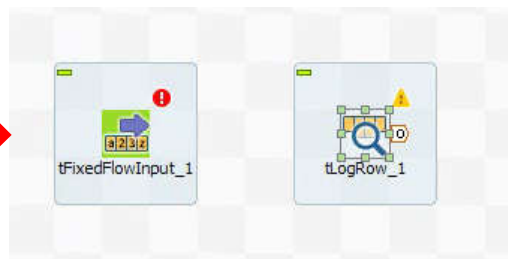
第二步：填写Name, Purpose, Description, 然后Finish



第三步：在Standard下面就可以看到新建的job，然后双击打开



用到两个组件
如右边



InputMessage(tFixedFlowInput_1)

Column	Value
Message	"Hello World"



配置好组件
(Compent)
后, 点击运行
(run)

作业 HelloWorld

基本运行

开始调试

高级设置

内存运行

运行

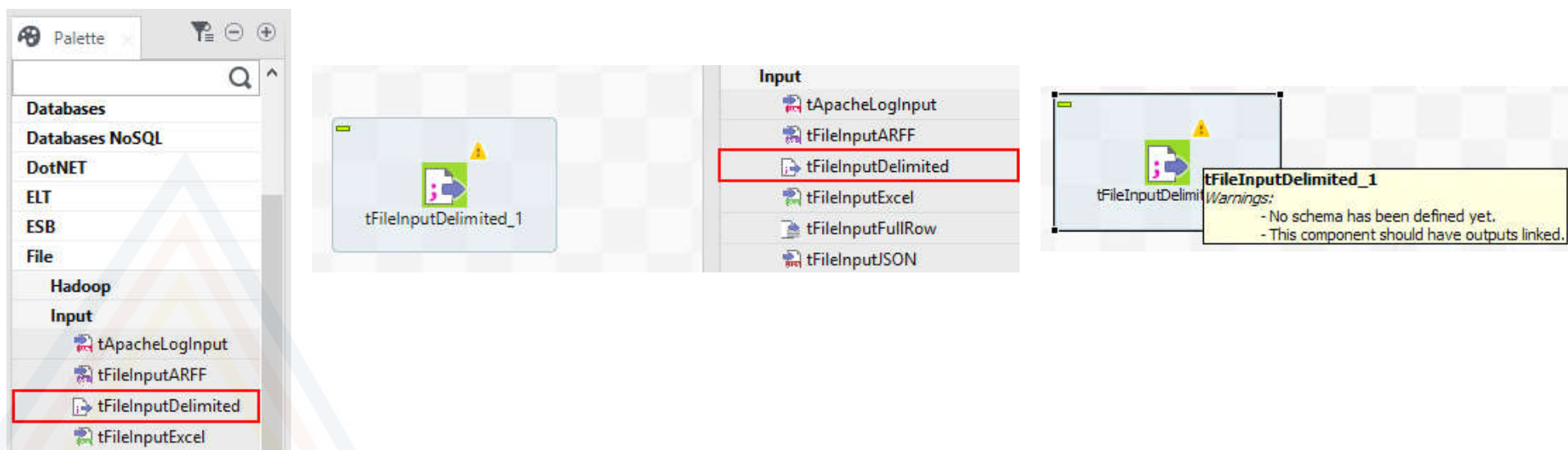
终止

清理

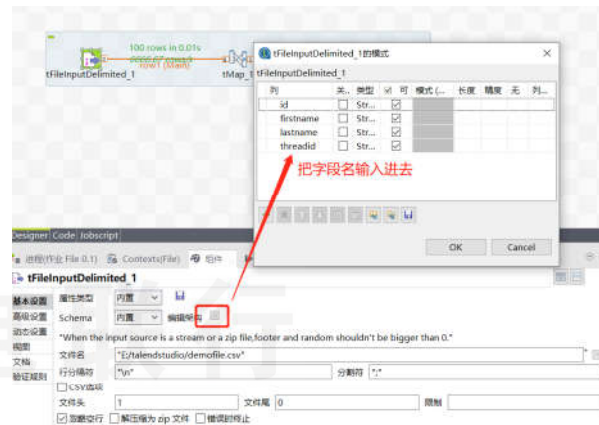
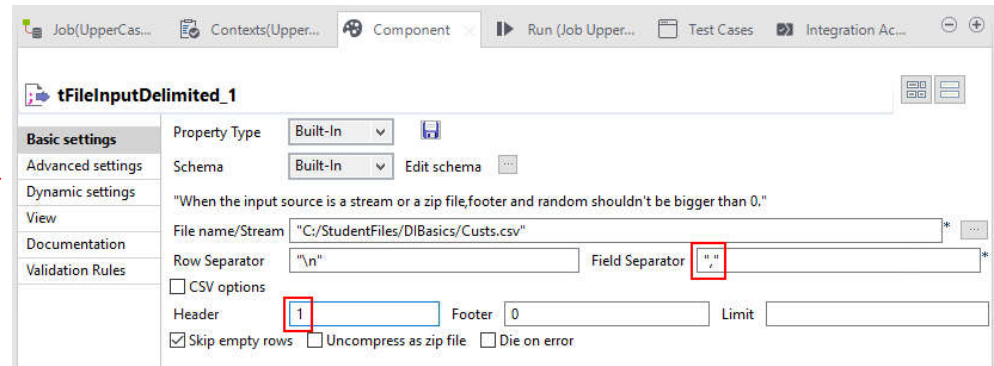
在 22:53 11/02/2020 开始作业 HelloWorld.
[statistics] connecting to socket on port 3974
[statistics] connected
Hello Word!
[statistics] disconnected
在 22:53 11/02/2020 结束作业 HelloWorld. [退出代码 = 0]

文件处理

- 从右边Palette拖动tFileInputDelimited组件到设计者界面



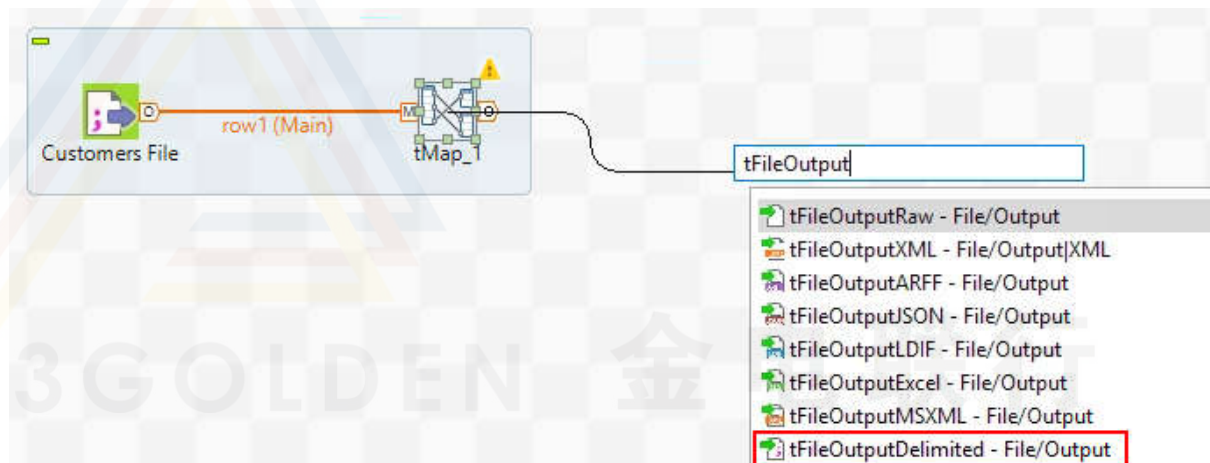
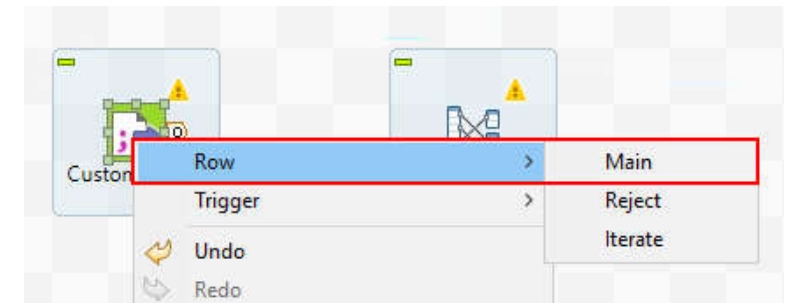
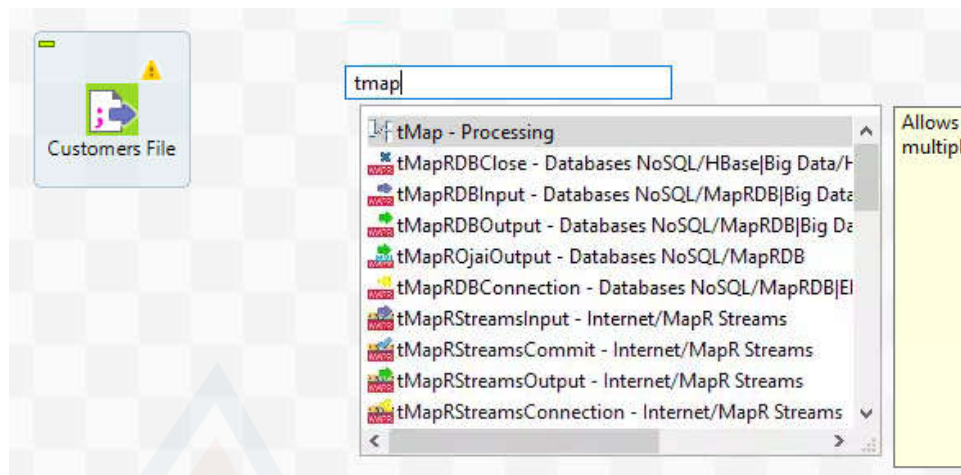
- 配置好组件tFileInptDelimited, 填写文件路径, 配置分隔符为并去掉第一行
- 配置好schema



3GOLDEN

金电

- 加入tmap组件和输出组件tFileOutputDelimited



- 双击tmap组件，做好映射，左右两边的字段类型必须一样



Talend Data Fabric - tMap - tMap_1

Find: ↓ ↑ ↺ ↻

Var + × ↑ ↓ ↺ ↻ 📄 自动映射!

row1

Column
id
firstname
lastname
threadid

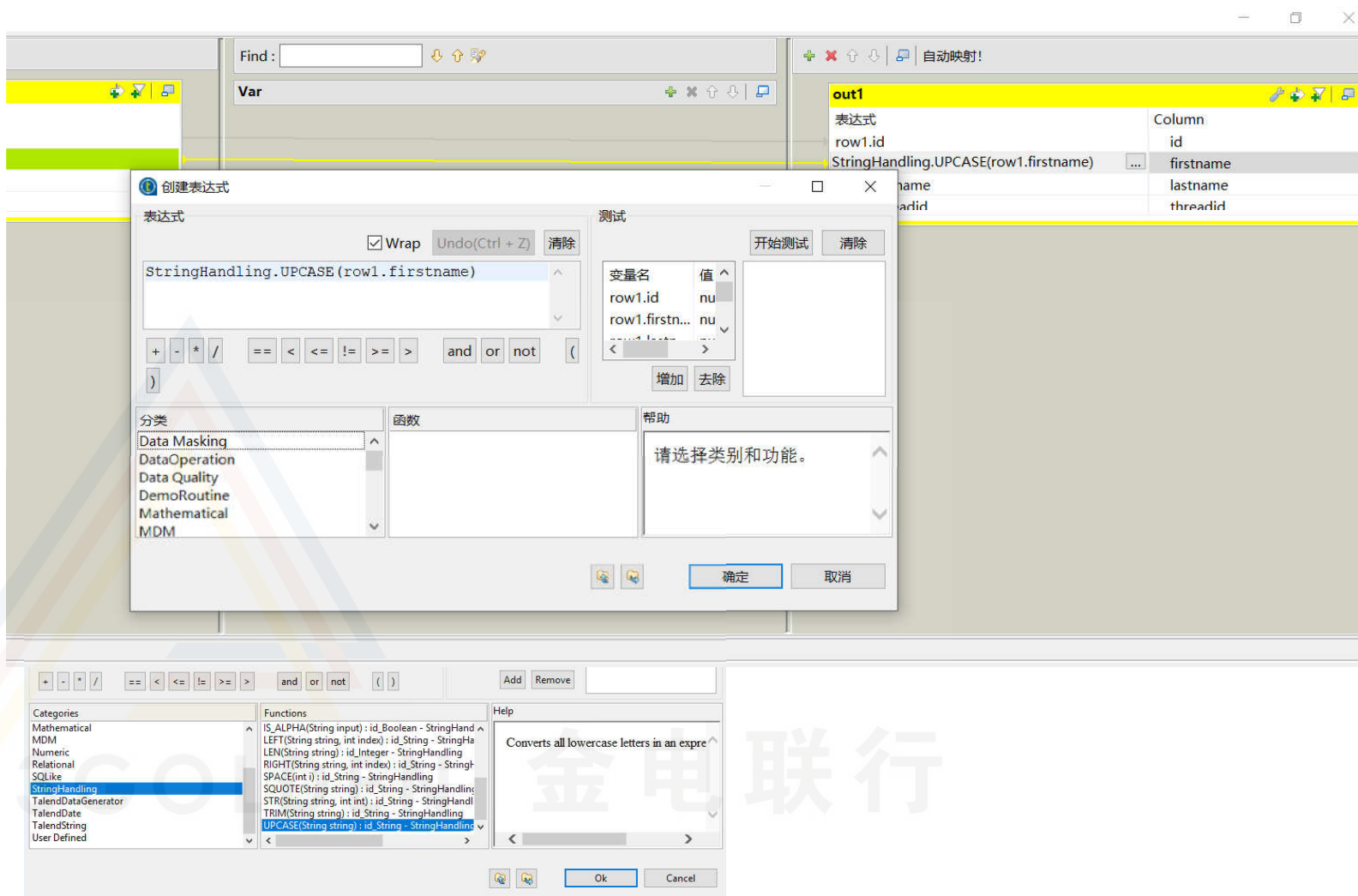
图解编辑器 表达式编辑器

row1

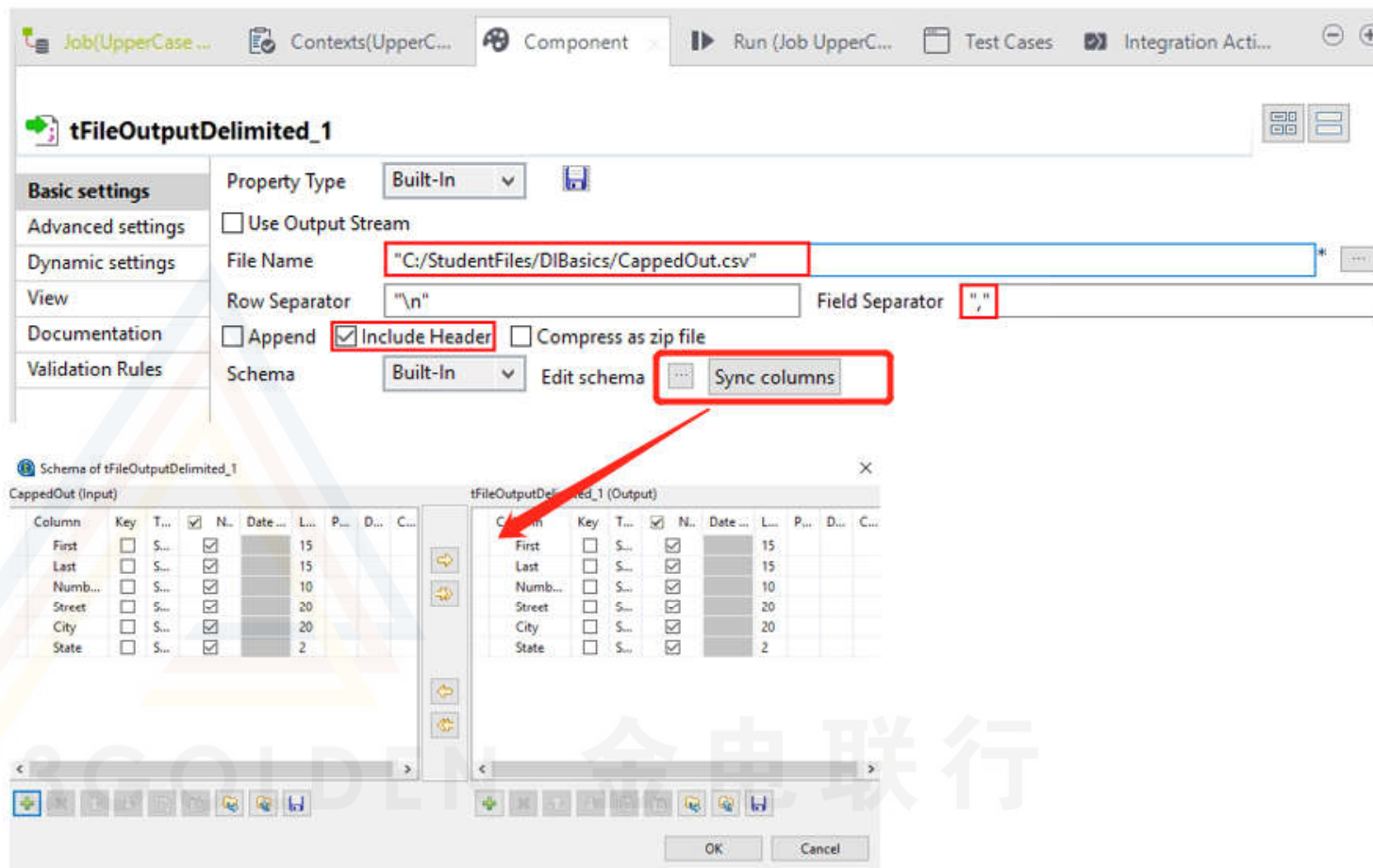
列	关..	类型	✓	可	模式 (只限...	长度	精度	无	列注释
id	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					
firstname	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					
lastname	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					
threadid	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					

列	关..	类型	✓	可	模式 (只限...	长度	精度	无	列注释
---	-----	----	---	---	-----------	----	----	---	-----

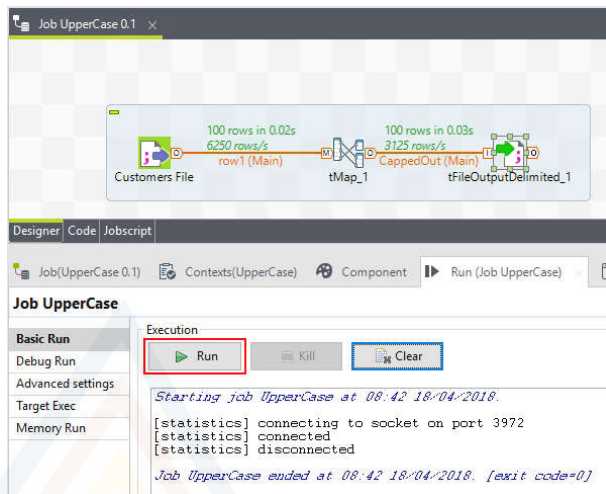
- 在Express Budiler做个简单的处理，把字段firstname内容全换成大写



- 配置tFileOutputDelimited，路径，分隔符，输出加上表头，并同步列（Sync columns）



- 然后点击下运行（Run）



输出结果

id;firstname;lastname;threadid	
1;GROVER;Harris;main	
2;HERBERT;Willis;main	
3;RONALD;Higgins;main	
4;HARRY;Higgins;main	
5;BENJAMIN;Higgins;main	
6;JAMES;Willis;main	
7;THEODORE;Thomson;main	
8;LYNDON;Thomson;main	
9;ULYSSES;Smith;main	
10;LYNDON;Thomson;main	
11;BILL;Harris;main	
12;BILL;Willis;main	
13;JIMMY;Thomson;main	
14;JOHN;Thomson;main	
15;BILL;Higgins;main	
16;JOHN;Smith;main	
17;WILLIAM;Harris;main	
18;HARRY;Willis;main	
19;GEORGE;Smith;main	

作业

- 1、创建一个简单的标准JOB
- 2、创建一个文件处理JOB

