

作业流如何识别“外部ETL”跑完状态？

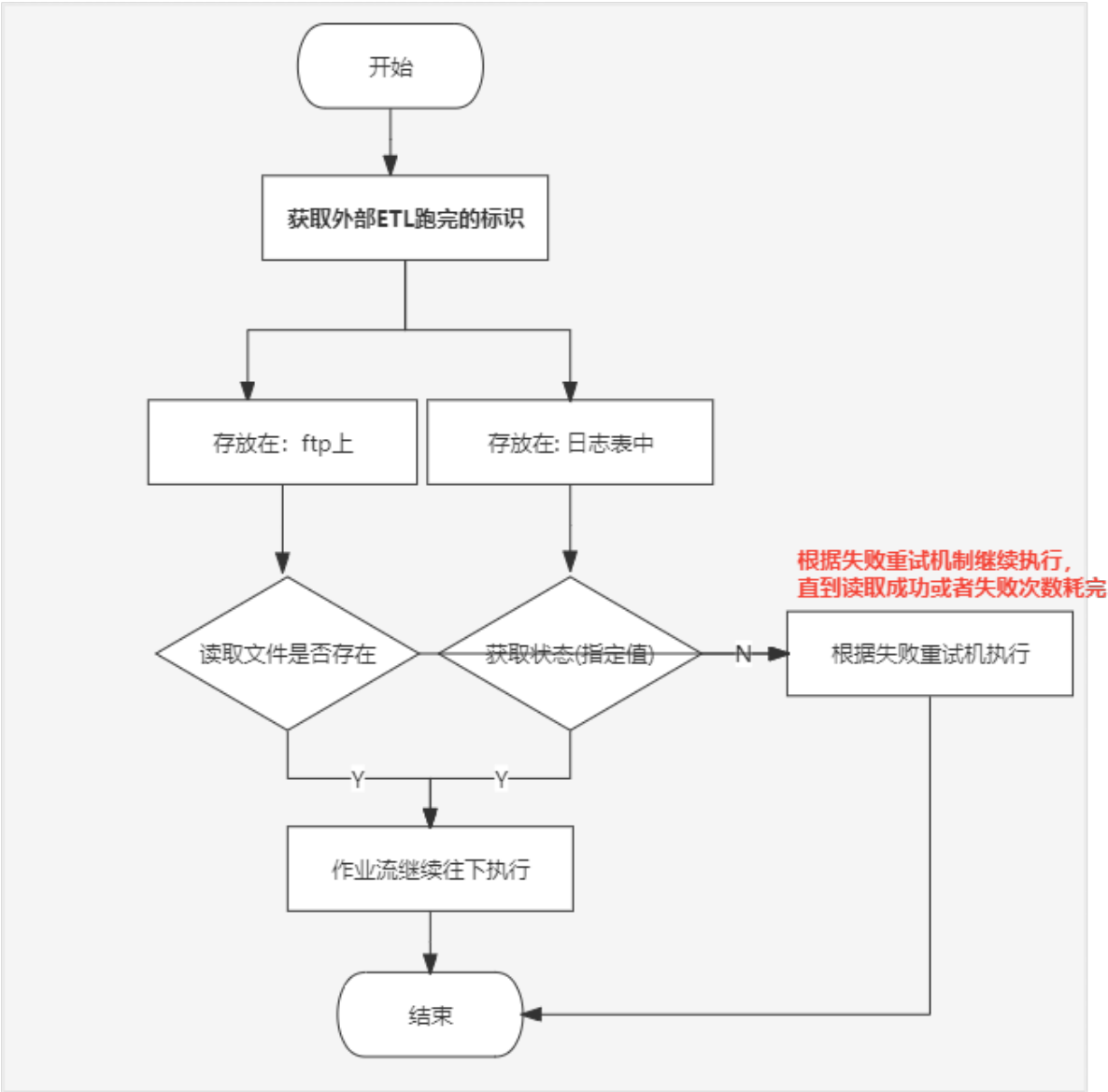
- 1、概述
 - 2、场景举例：通过“ftp方式”检测文件是否存在
 - 3、场景举例：获取每天ETL执行状态

1、概述

在V10.5beta版本中，识别“外部ETL”跑完状态有以下方法：

- 1) “源库脚本”节点：需要写SQL语句并且该节点无法与其他节点关联，需要单独建一个ETL资源，不但操作麻烦而且增加了维护成本。

而V10.5.8，支持”标识“存放在ftp上，也支持”标识“存放在具体的表中，如下图，扩展了识别方法，操作也更为直观简单，节省了实施成本，确保了数据正确。



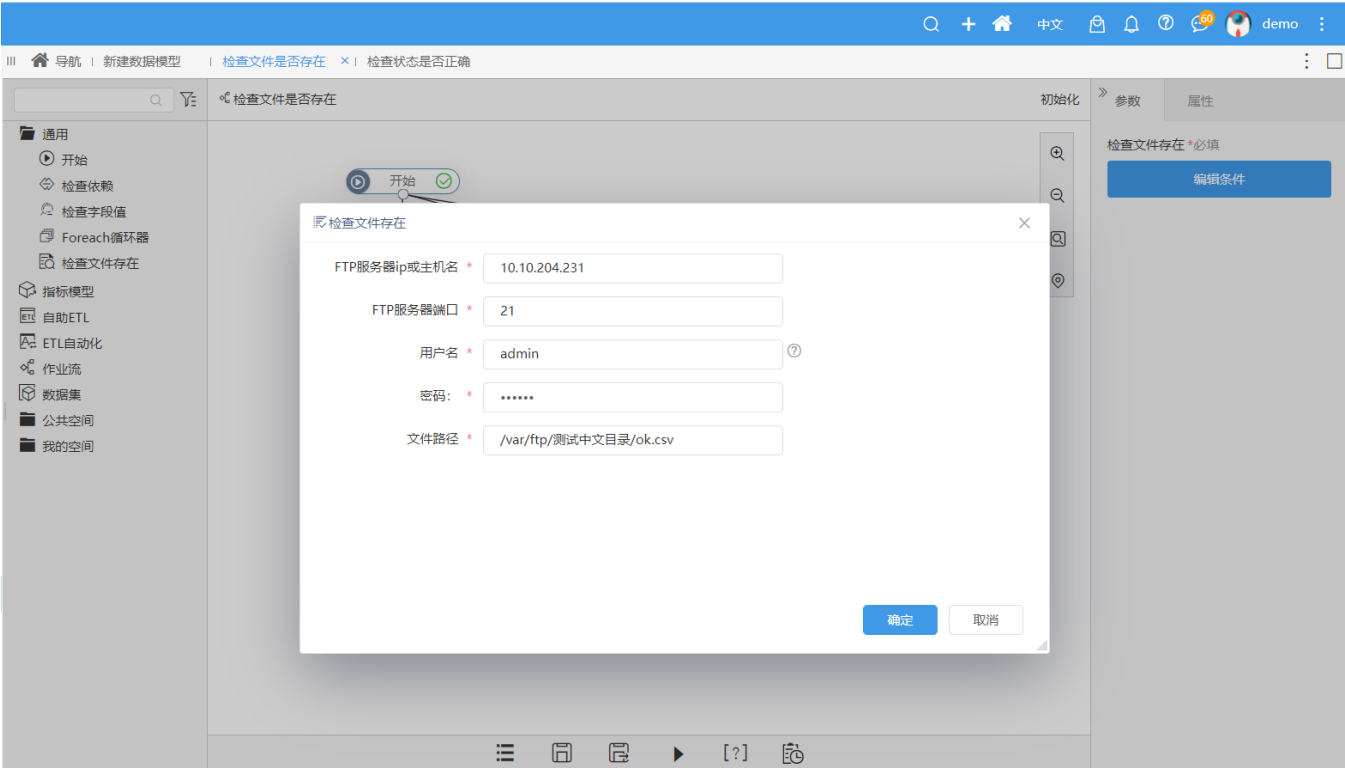
2、场景举例：通过“ftp方式”检测文件是否存在

某银行项目，客户有自己的ETL工具，会设置时间频率跑定时任务，跑完会把定时任务文件输出到ftp下，并且会生成一个ok.csv文件；如果存在了ok文件，那么则认为，这个外部任务已跑成功，数据已写入完成。

此时smartbi需要去检测是否存在ok文件，如果存在，才可以继续往下执行作业流。

2.1 功能介绍

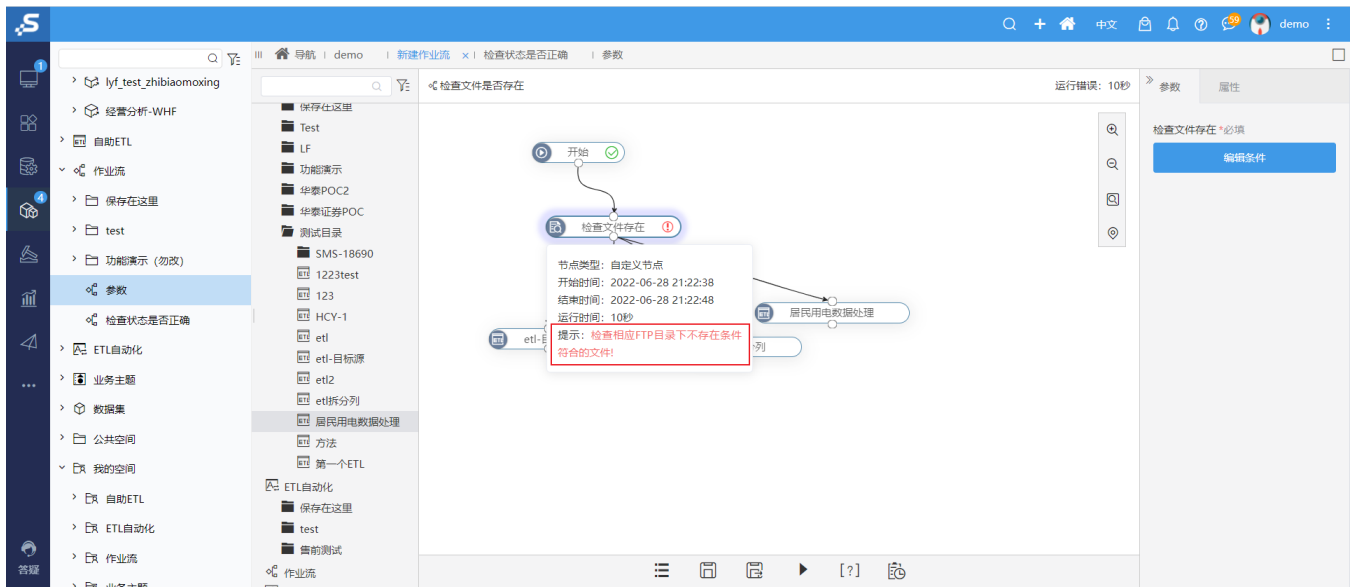
step1：在作业流拖入”检查文件存在”节点并且进入编辑状态，配置ftp地址、登录账号以及输入”文件路径“，如下图：



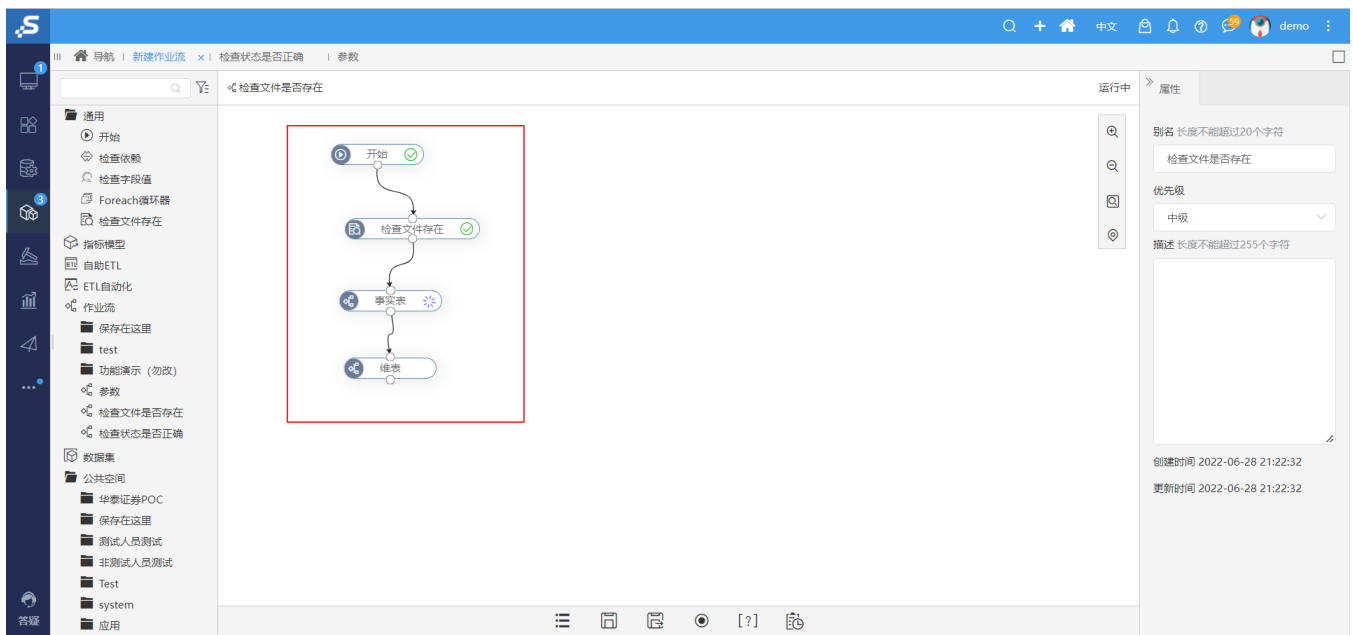
以下是“检查文件存在”节点的配置界面字段说明：

序号	选项	备注
1	ftp服务器IP或主机名	ftp地址
2	用户名/密码	默认是匿名用户：anonymous，密码可随便填写
3	文件路径	ftp下的绝对路径；不是读取文件数据，而是判断文件存在不存在

step2：运行节点，如果ok.csv文件不存会报错并不往下执行，如下图：



相反如果ok.csv文件存在则继续往下执行，直到所有的节点都执行完：如下图



3、场景举例： 获取每天ETL执行状态

某些银行项目有自己的ETL工具,但也使用Smartbi的ETL，需要先跑完自己的ETL，才开始跑Smartbi 的ETL；

那Smartbi 怎么通过动态的参数是获取到这个”状态“呢？

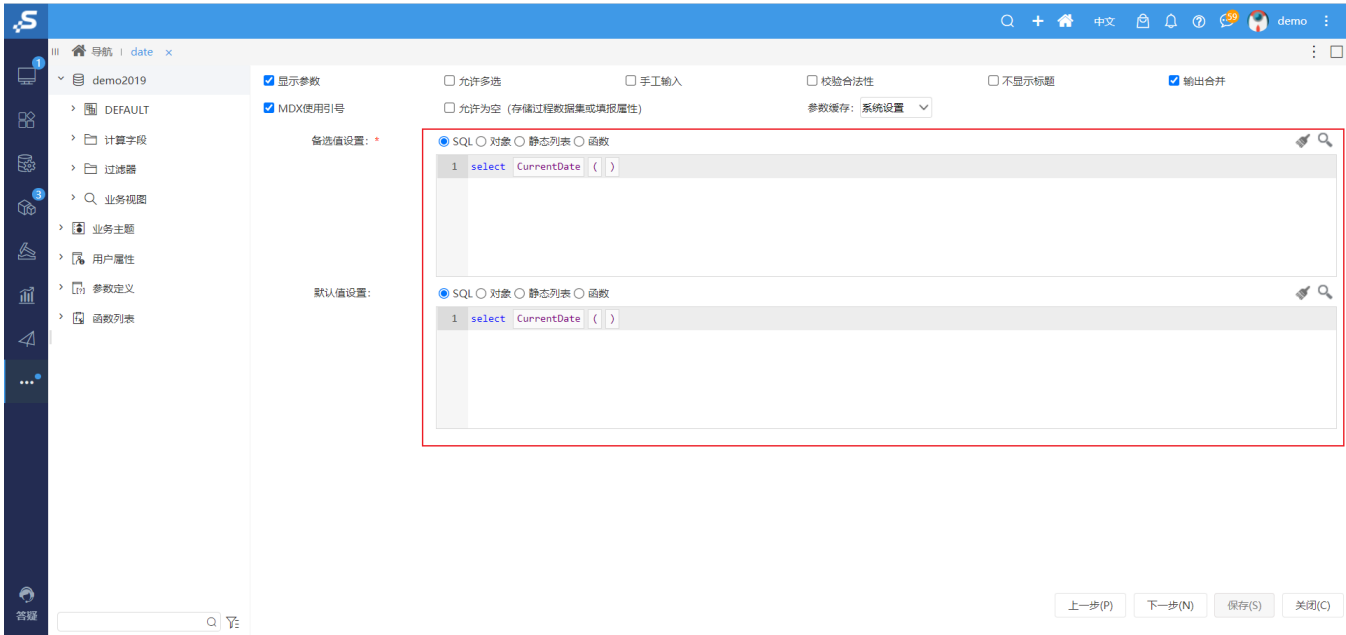
3.1、功能介绍

某项目，客户有自己的ETL工具，会设置时间频率跑定时任务，跑完会每在指定表生成一条记录，记录的状态1为成功，其他值则代表失败，如下表格：

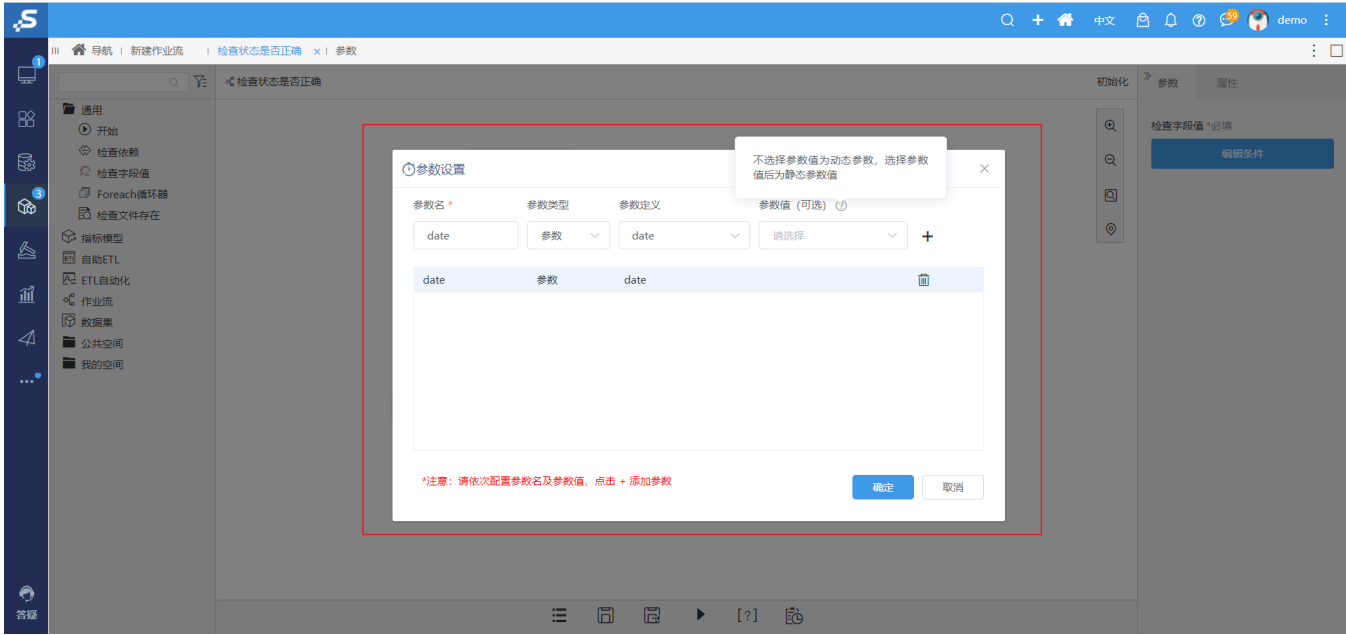
日期	状态	执行时间
2022-02-01	0	2022-02-01 01:13
2022-02-01	1	2022-02-01 02:10

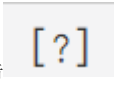
方案：Smartbi 通过作业流的参数支持动态参数(公共参数)获取当前日期，根据当前日期，再结合“检查字段值”节点的SQL语句进行判断；


step1: 首先在Smatbi “公共参数”新增一个获取动态日期的参数“date”



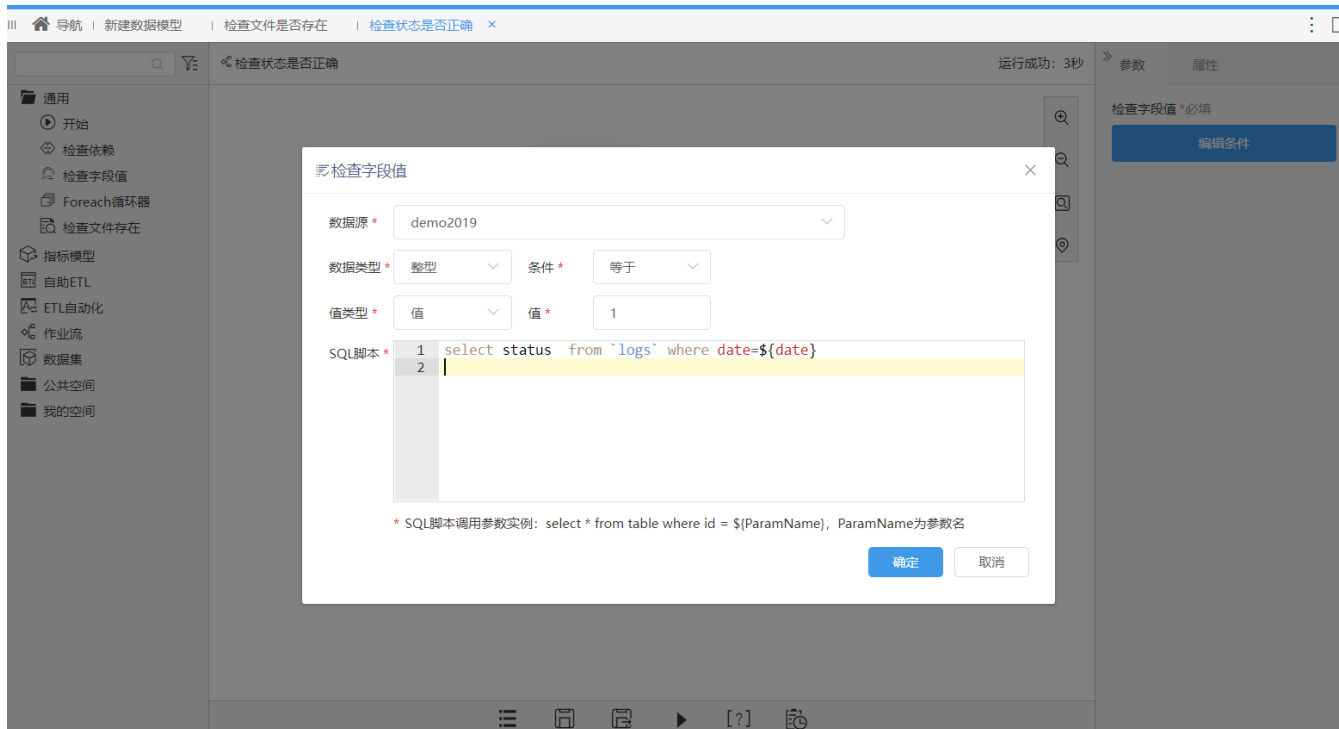
setp2: 再到作业流中参数界面新增参数“date”，并且与公共参数进行关联，如下图：



序号	功能选项	说明
1	参数入口	点击  图标
1	参数名	必填
2	参数类型	分类“值”、“参数” 值：枚举型，即具体某一个值，可以是整形、字符串、日期等等 参数：通过公共参数定义的参数来获取值
3	参数定义	根据“参数类型”显示，如果是参数类型是“值”，则显示数据类型 如果参数类型是“参数”，则可选择具体的参数

4	参数值	 如果想要实现动态获取参数值，参数值不需要填写； 如果想实现静态值，则直接输入具体的值
---	-----	--

step3: **参数+“检查字段值”节点的sql脚本**，把新建的参数date值动态带入到SQL语句中，实现根据动态条件查询状态，如下图：



如果值为“1”的话,则继续执行,如果不为1,则不往下执行:

