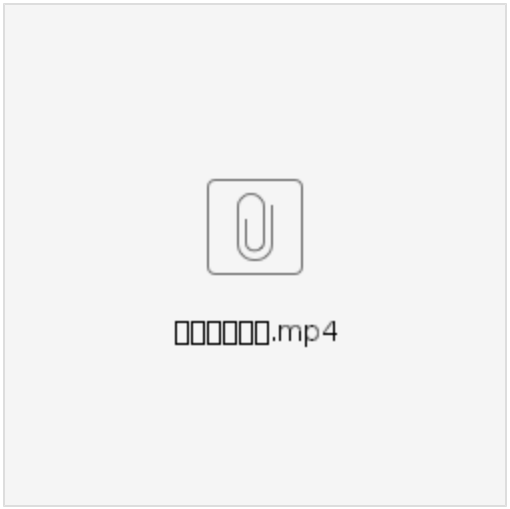
数据调度介绍

了解数据调度

数据调度是我公司自主研发的一款数据处理工具，提供可视化的数据调度作业界面，支持作业的试运行和断点调试等，提供丰富的组件。

- 了解数据调度
  - 1. 概述
  - 2. 界面概览
  - 3. 创建数据调度
    - 3.1 创建管道
    - 3.2 管道流程设计
    - 3.3 保存&发布
  - 4. 查看数据调度日志

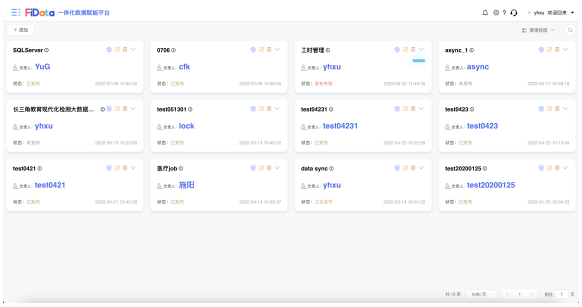


1. 概述

数据调度是我公司自主研发的一款数据处理工具，提供可视化的数据调度作业界面，支持作业的试运行和断点调试等，提供丰富的组件。

2. 界面概览

入口



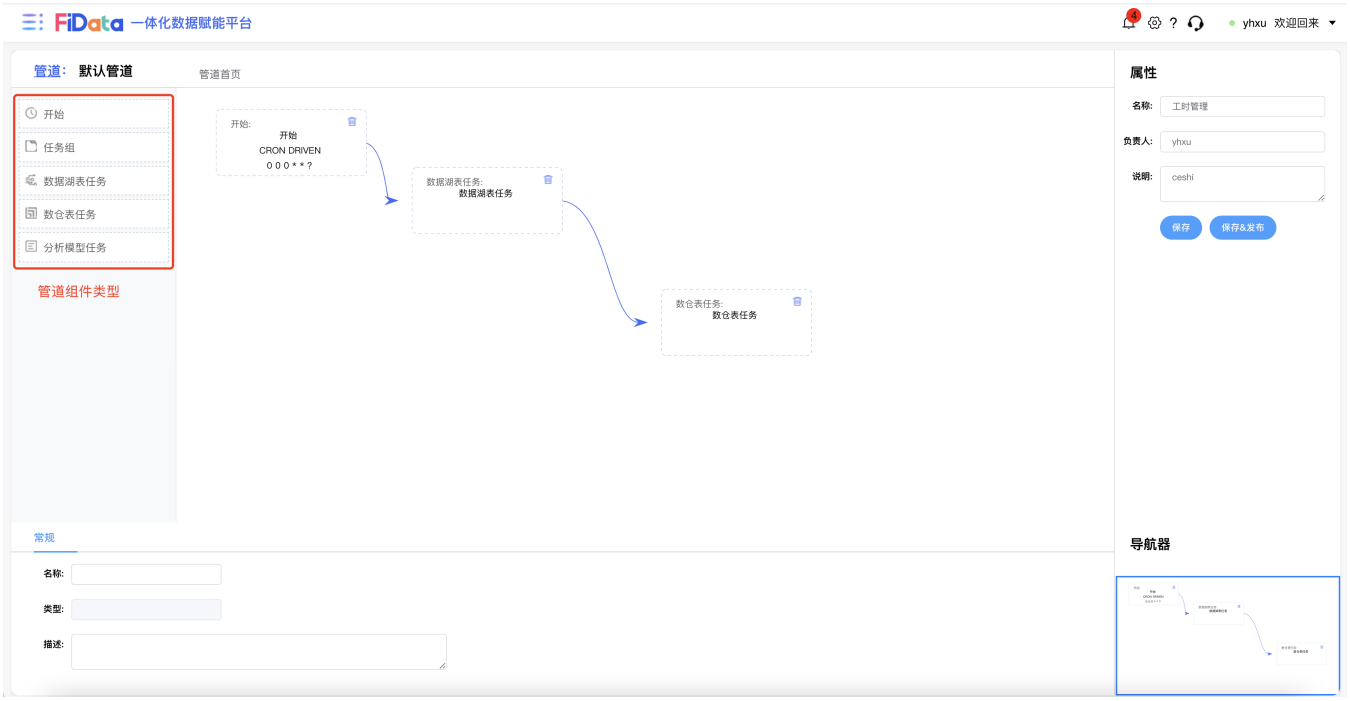
调度设计页面

3. 创建数据调度

3.1 创建管道

添加管道： 添写基本信息后确定，会在页面新增一条数据。

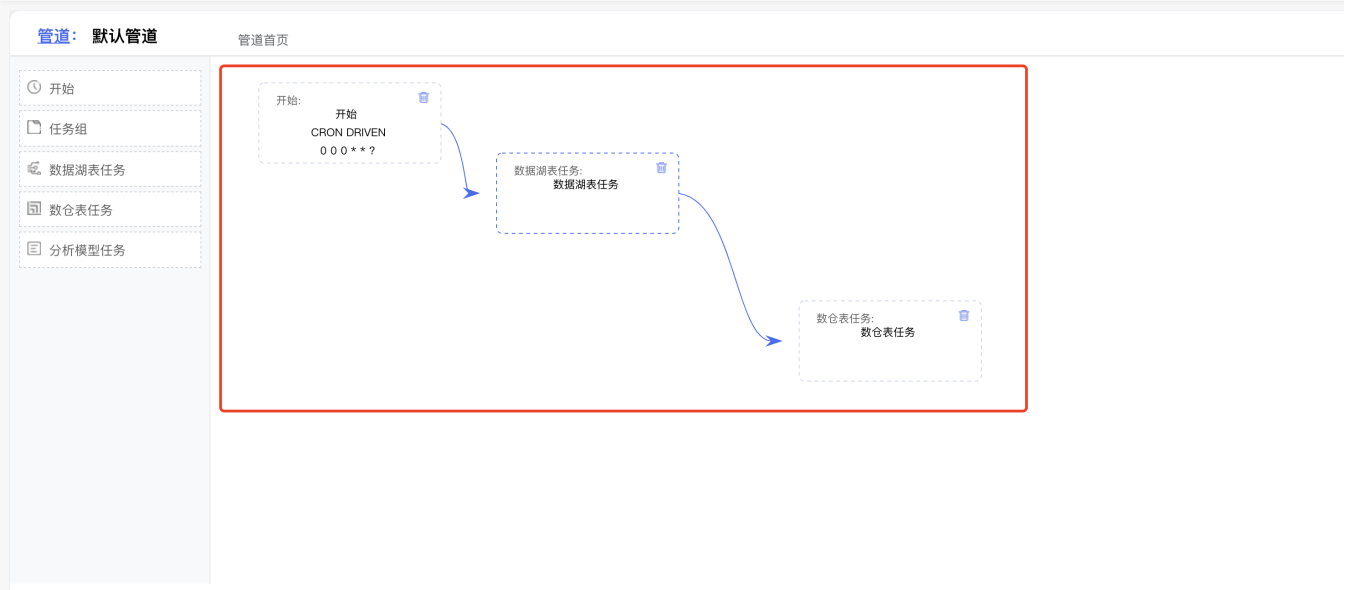
3.2 管道流程设计



点击管道进入管道设计界面，左侧是可拖拽的任务组件，分别是：开始、任务组、数据湖表任务、数仓表任务、分析模型任务。

1. 开始：流程起点，用来设置整个流程的调度时间、频率，调度类型有两种：TIMER DRIVEN ： TIMER DRIVEN 输入框输入的为秒（只能输入数字）或 CRON DRIVEN ： 自定义设置。
2. 任务组：可以理解为任务的集合，可将多个任务（数据湖表任务、数仓表任务、分析模型任务）放在一个任务组里。
3. 数据湖表任务：数据湖表对应数据接入中的物理表。

4. 数仓表任务：数仓表对应数据建模中的物理表。
5. 分析模型任务：分析模型对应数仓建模中分析指标模块中创建的表。



在左侧选中需要用到的组件拖拽到右侧的设计面板中。

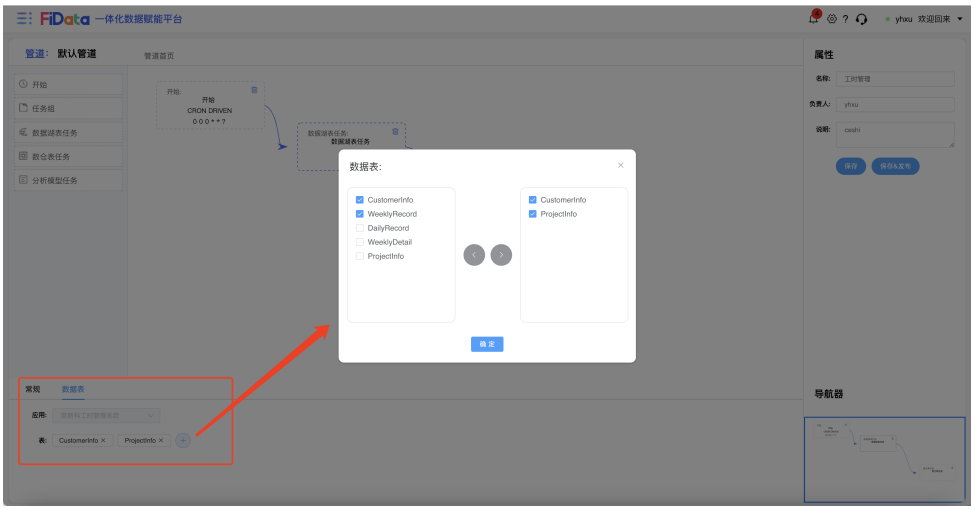
连接管道，利用连接点连接管道活动，连接的顺序代表数据同步的顺序。

任务组件：

1.开始：

选择 CRON DRIVEN 点击自定义设计时间调度，弹出弹窗可自定义设计调度。

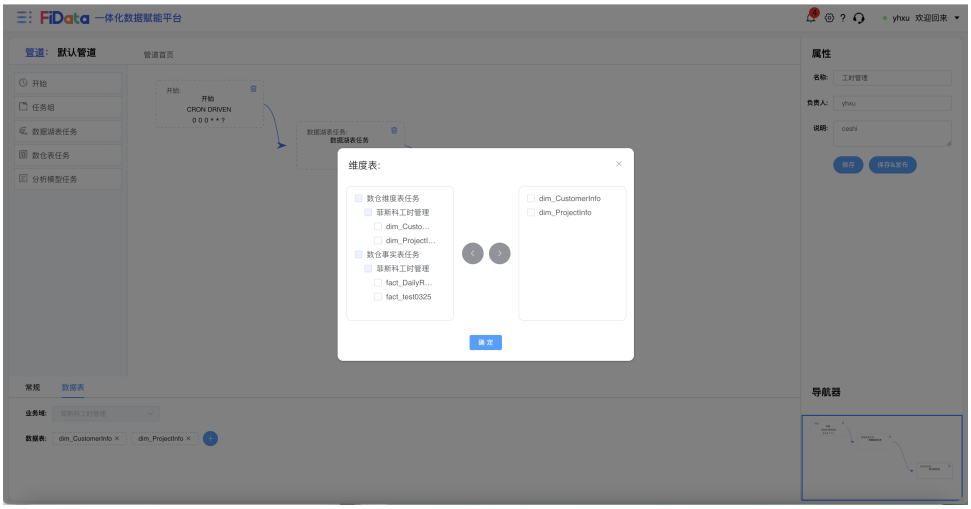
2.数据湖表任务



选中数据湖表任务，在左下角切换到数据表选项点击“+”图标弹出需要同步数据的物理表供用户选择。

3.数仓表任务：

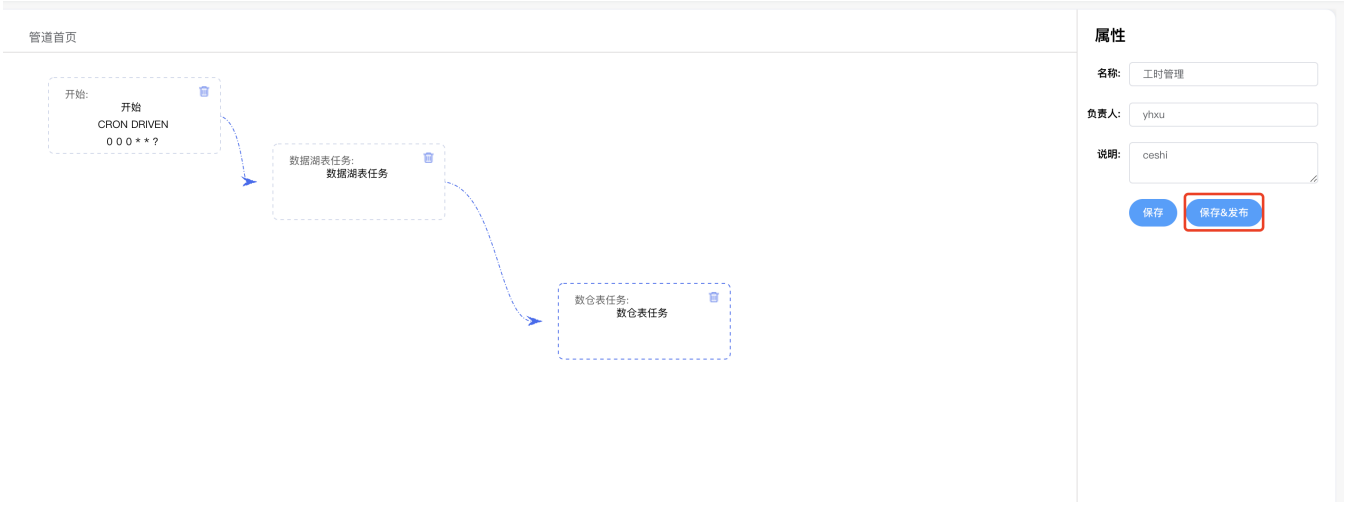
⚠️数仓表任务的操作流程和数据湖表任务的流程相同，唯一不同点在于数仓表任务把物理表分成了维度表和事实表两类；

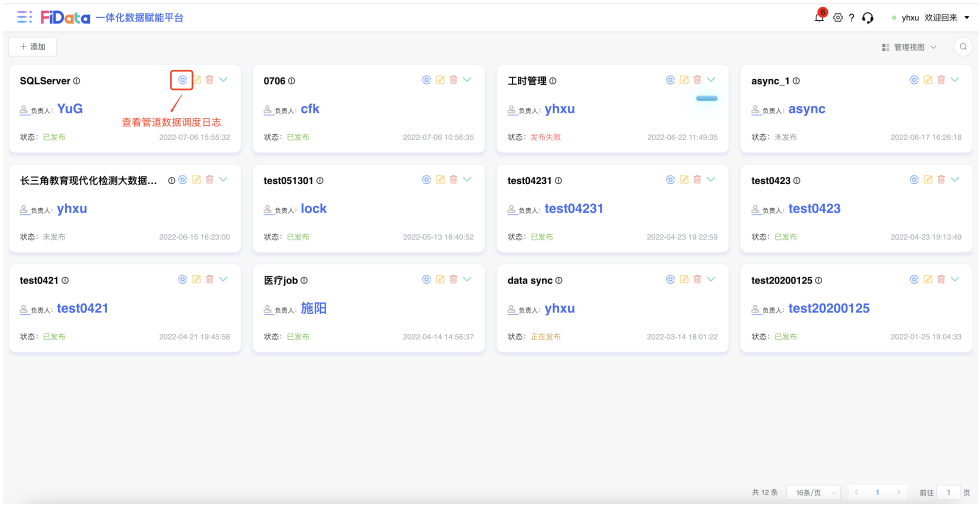


4.分析模型任务

同2、3 。

3.3 保存&发布



数据同步流程设计好了之后点击右侧的“保存&发布”按钮既可，稍后会根据设计的流程后台自动生产数据同步的调度任务，根据设置的调度频率自行调用。如果是点击的“保存”按钮，那么只会保存设计的流程，不会在系统后台生成调度任务。

4. 查看数据调度日志