**仅限阅读 请勿传播**

当您阅读本文档时，即表示您同意不传播本方案的所有内容

**ETL\_运维操作手册**

**文件信息**

|  |  |  |  |
| --- | --- | --- | --- |
| 文件状态： | [ ] 草稿文件  [ √ ] 正式文件  [ ] 更改正式文件 | 文件名称： | ETL\_运维操作手册 |
| 作者： | 蔡伟 |
| 审核： | 胡东云 |
| 完成日期： | 20240825 |
| 客户名称： | 庞贝捷管理（上海）有限公司 | | |
| 项目名称： | 庞贝捷管理（上海）有限公司\_PPG的数仓\_20230919002 | | |
| 项目编号： | 20230919002 | | |
| 适用阶段： | 系统建设阶段 | | |
| 版权所有： | 帆软软件有限公司 | | |

**版本修订：**

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **负责人** | **版本** | **说明** |
| 2024-08-25 |  | 1.0 |  |
|  |  |  |  |
|  |  |  |  |

**确认记录：**

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **负责人** | **版本** | **说明** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

目录

[一、 前言 1](#_Toc175505936)

[1. 本文说明 1](#_Toc175505937)

[2. 适用对象 1](#_Toc175505938)

[二、 ETL流程设计 1](#_Toc175505939)

[1. 总作业 1](#_Toc175505940)

[1.1 示意图 1](#_Toc175505941)

[1.2 描述 1](#_Toc175505942)

[2. Ods层作业main\_ods 2](#_Toc175505943)

[1.3 示意图 2](#_Toc175505944)

[1.4 描述 4](#_Toc175505945)

[3. Dw层作业main\_dw 4](#_Toc175505946)

[1.5 示意图 4](#_Toc175505947)

[1.6 描述 7](#_Toc175505948)

[4. Dm层main\_dm 7](#_Toc175505949)

[1.7 示意图 7](#_Toc175505950)

[1.8 描述 8](#_Toc175505951)

[5. 执行异常处理 8](#_Toc175505952)

[1.9 示意图 9](#_Toc175505953)

[1.10 描述 9](#_Toc175505954)

[三、 ETL执行计划配置 9](#_Toc175505955)

[1. 定时任务 9](#_Toc175505956)

[2. 日志输出目录 10](#_Toc175505957)

[四、 ETL异常处理 10](#_Toc175505958)

[1.ETL是否执行成功 10](#_Toc175505959)

[3. 执行失败时：查询执行日志 11](#_Toc175505960)

# 前言

## 本文说明

本文适用于ETL的维护工作.

## 适用对象

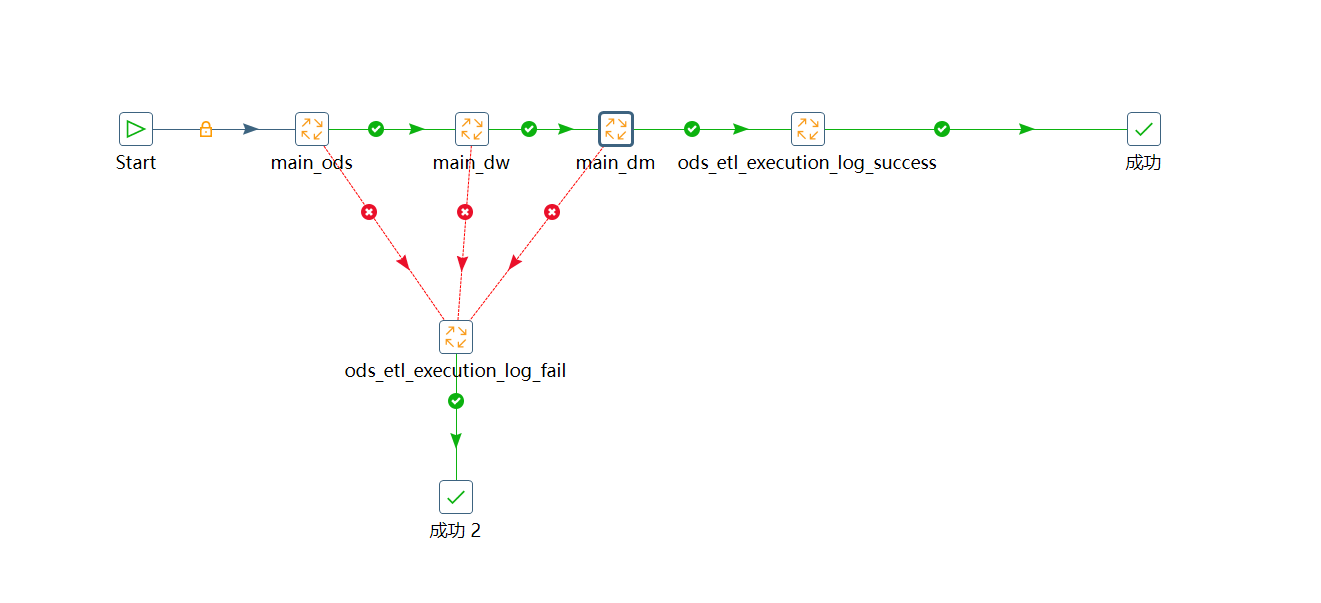
本文供ETL相关的运维人员及技术人员参考使用。

。

# ETL流程设计

## 总作业

1. 示意图

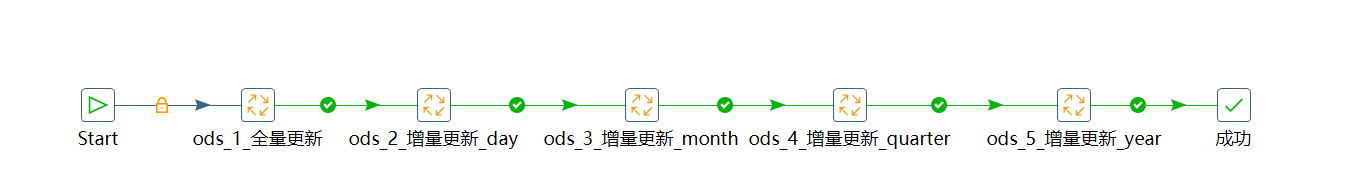


1. 描述

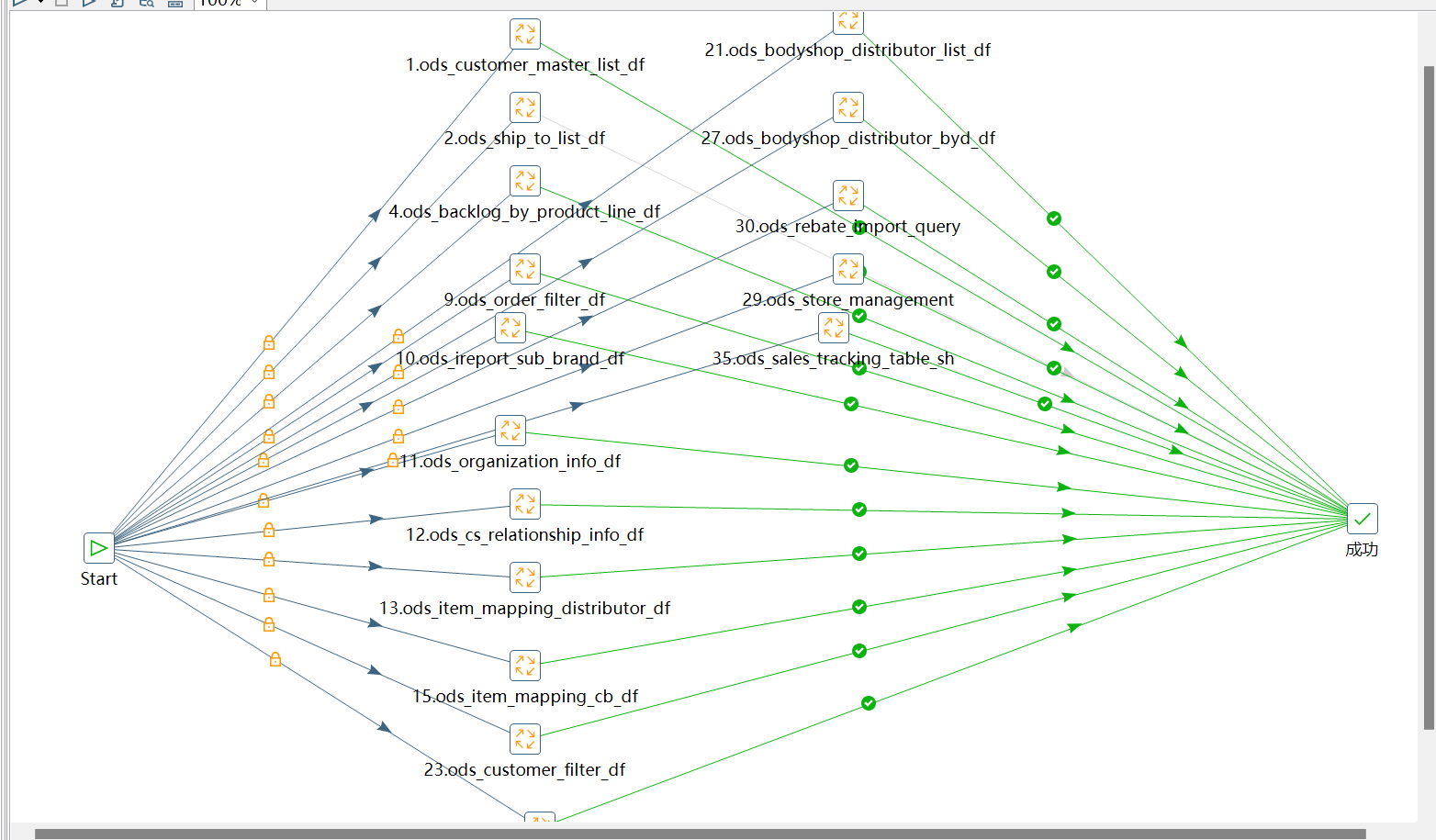
整体分为四个模块，main\_ods、main\_dw、main\_dm、执行异常处理（成功ods\_etl\_execution\_log\_success/失败ods\_etl\_execution\_log\_fail）。

## Ods层作业main\_ods

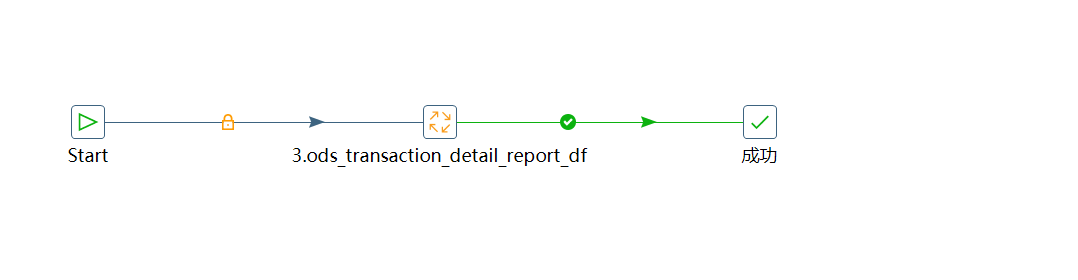
1. 示意图



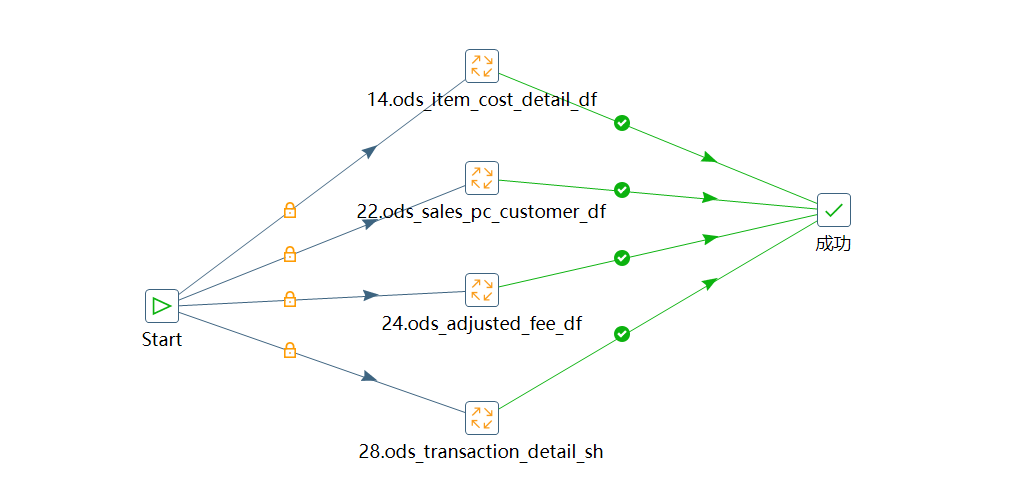
ods\_1\_全量更新



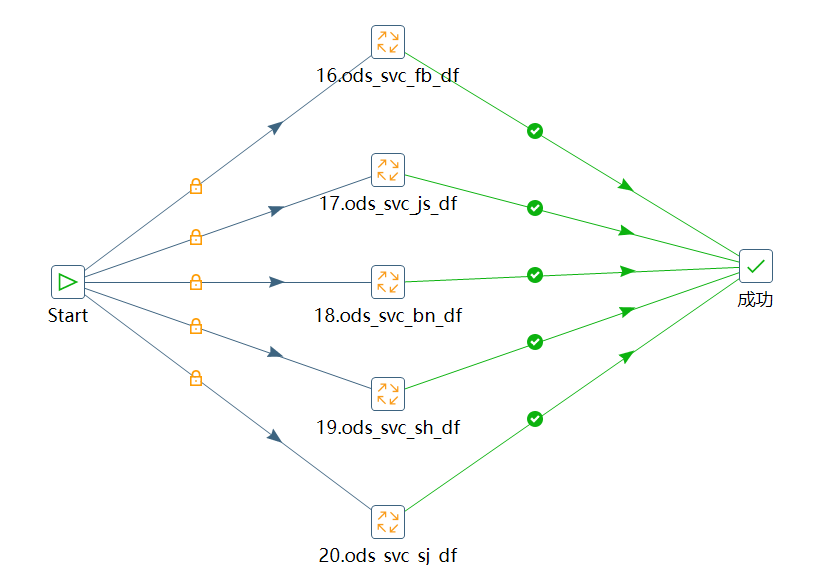
ods\_2\_增量更新\_day



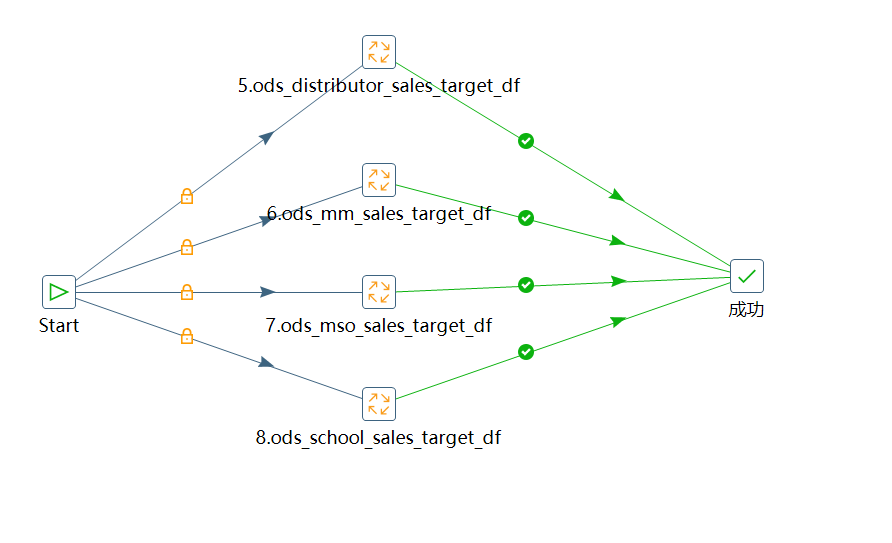
ods\_3\_增量更新\_month



ods\_4\_增量更新\_quarter



ods\_5\_增量更新\_year



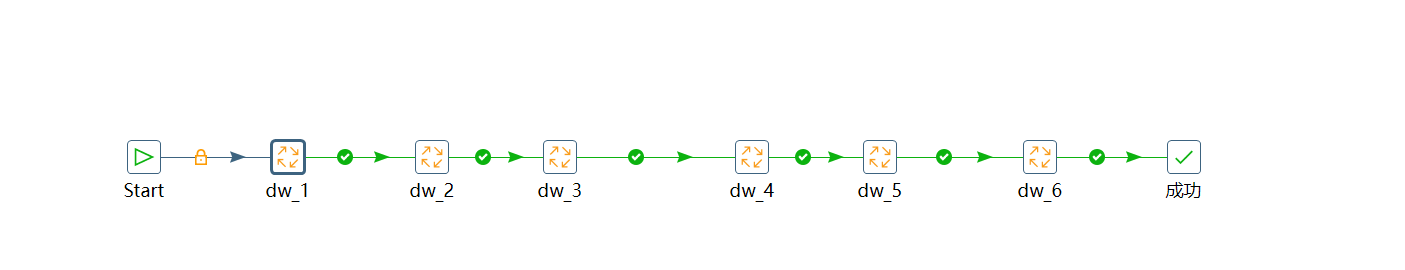
1. 描述

整体分为五个模块，ods\_1\_全量更新、ods\_2\_增量更新\_day、ods\_3\_增量更新\_month、

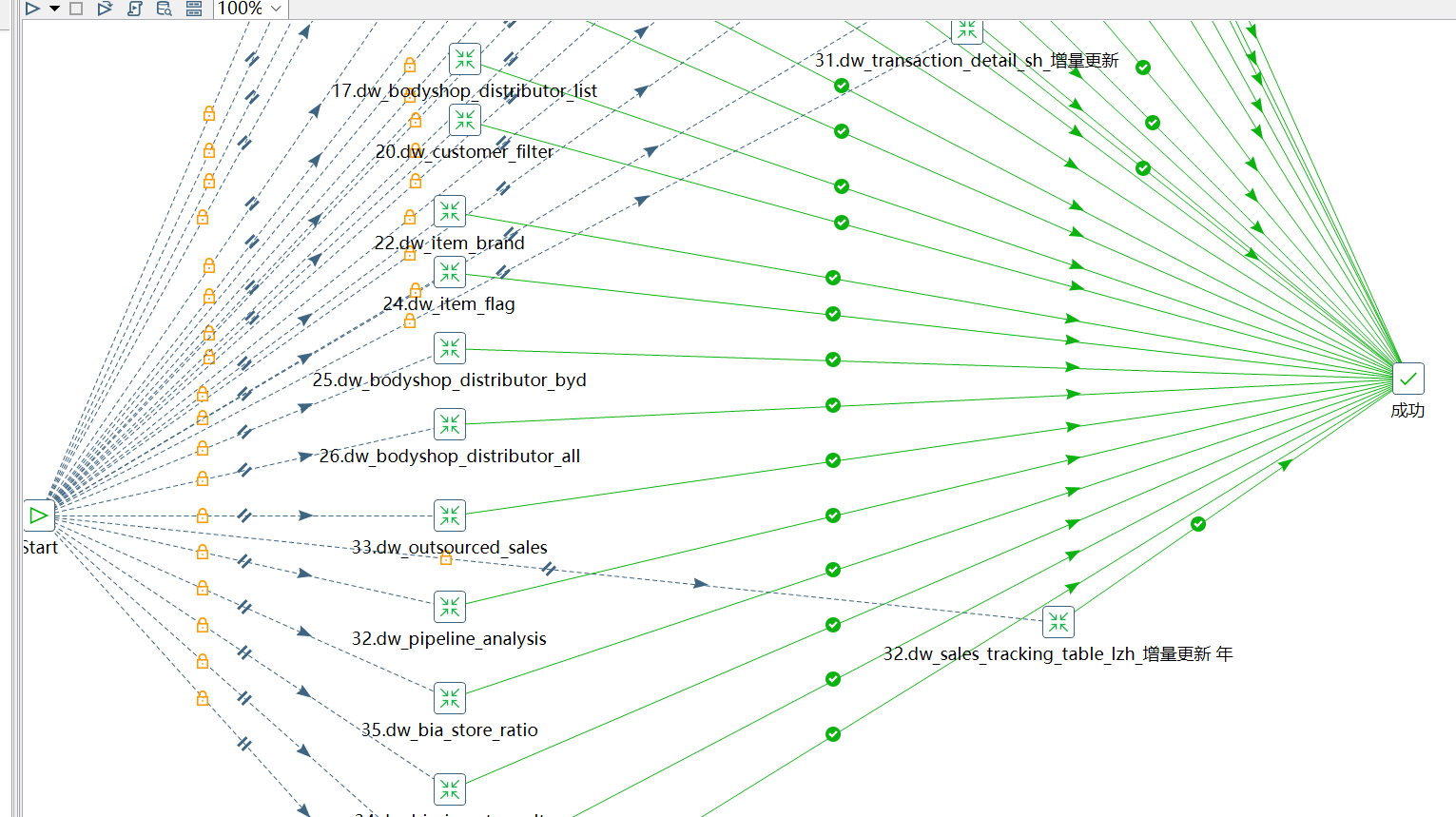
ods\_4\_增量更新\_quarter、ods\_5\_增量更新\_year

## Dw层作业main\_dw

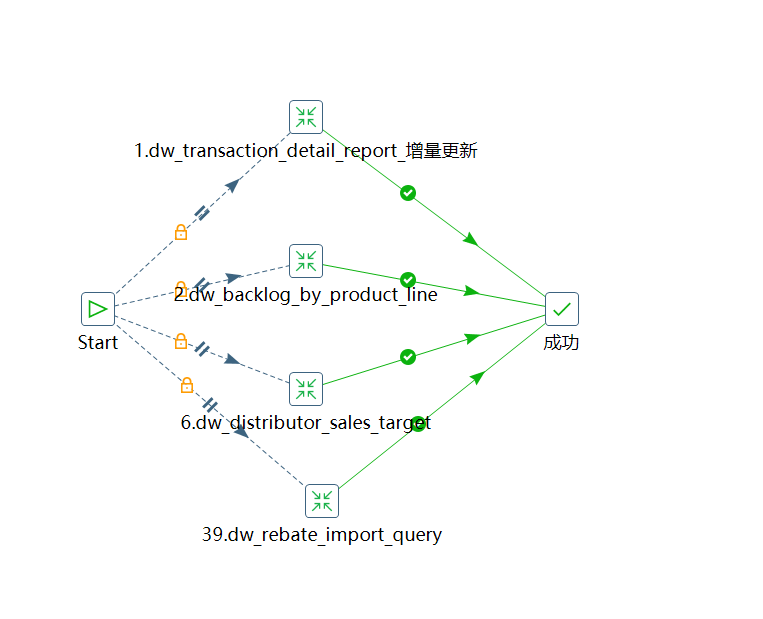
1. 示意图



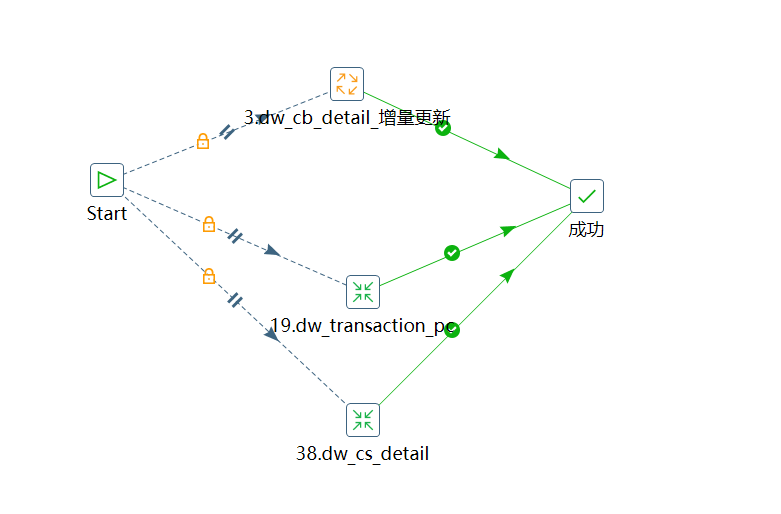
dw\_1



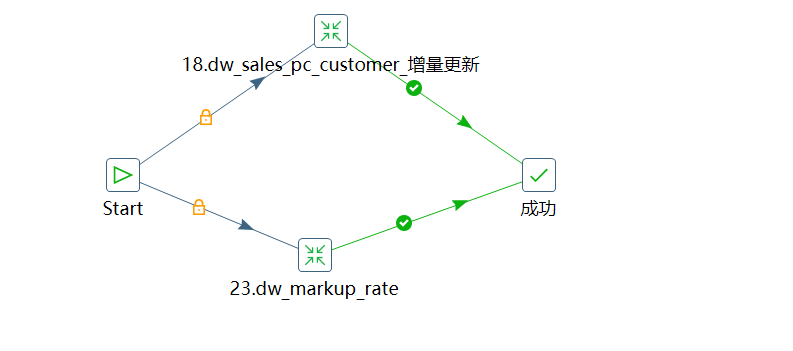
dw\_2



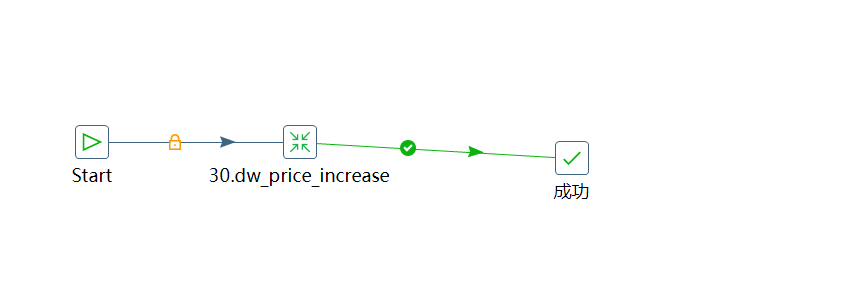
dw\_3



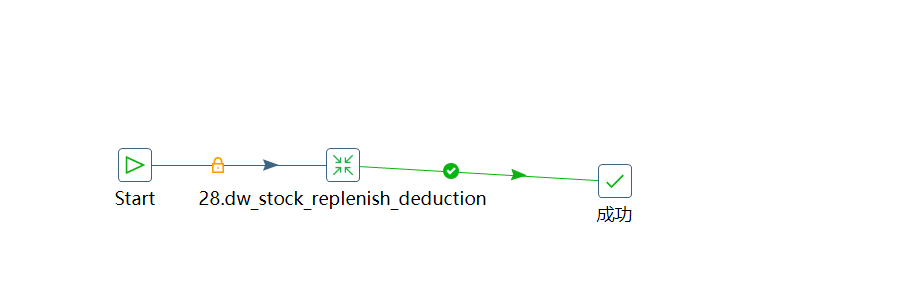
dw\_4



dw\_5



dw\_6

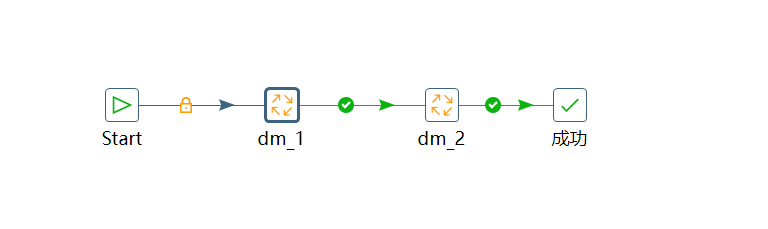


1. 描述

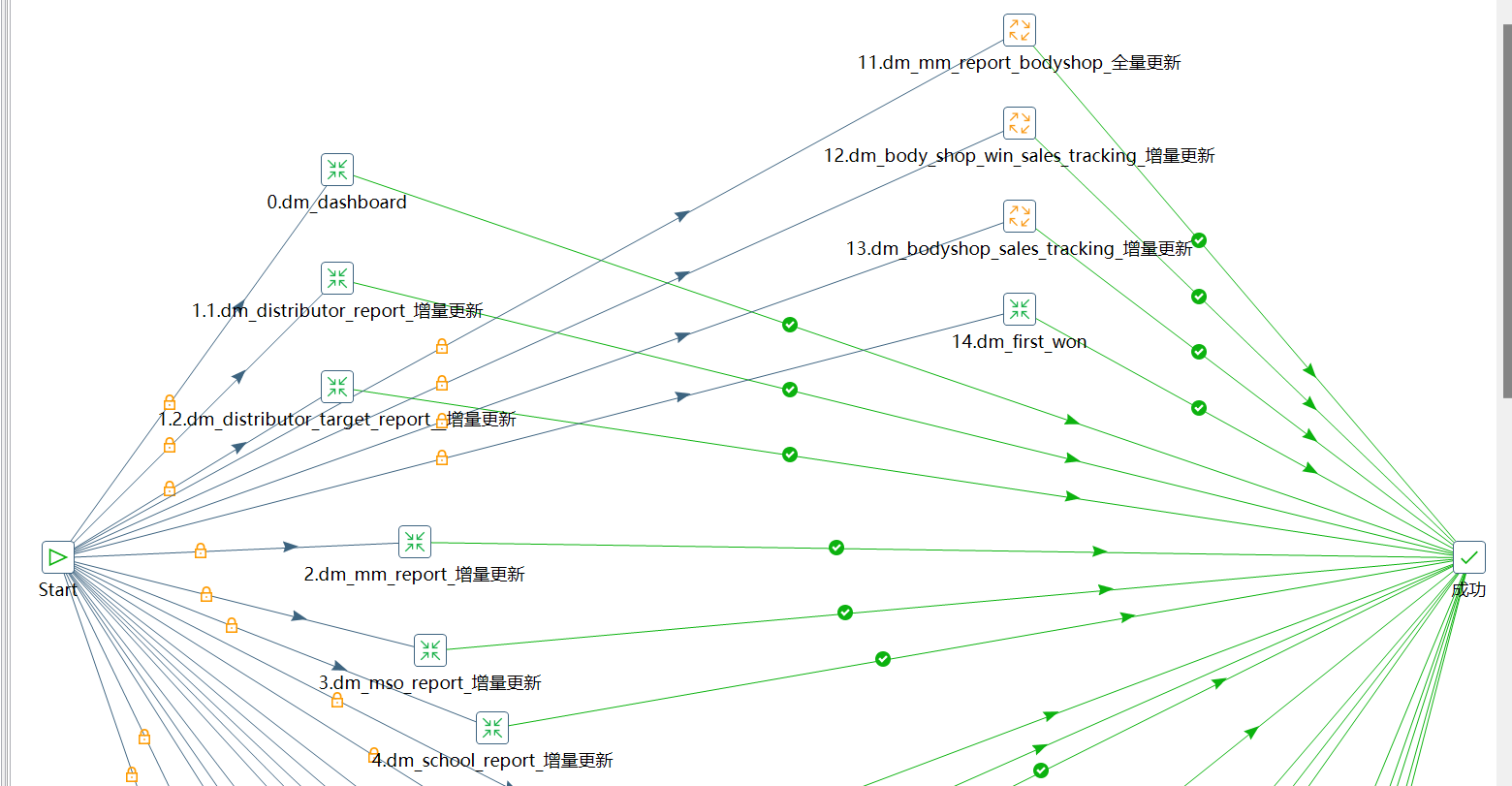
整体分为六个模块，dw\_1，dw\_2，dw\_3，dw\_4，dw\_5，dw\_6。

## Dm层main\_dm

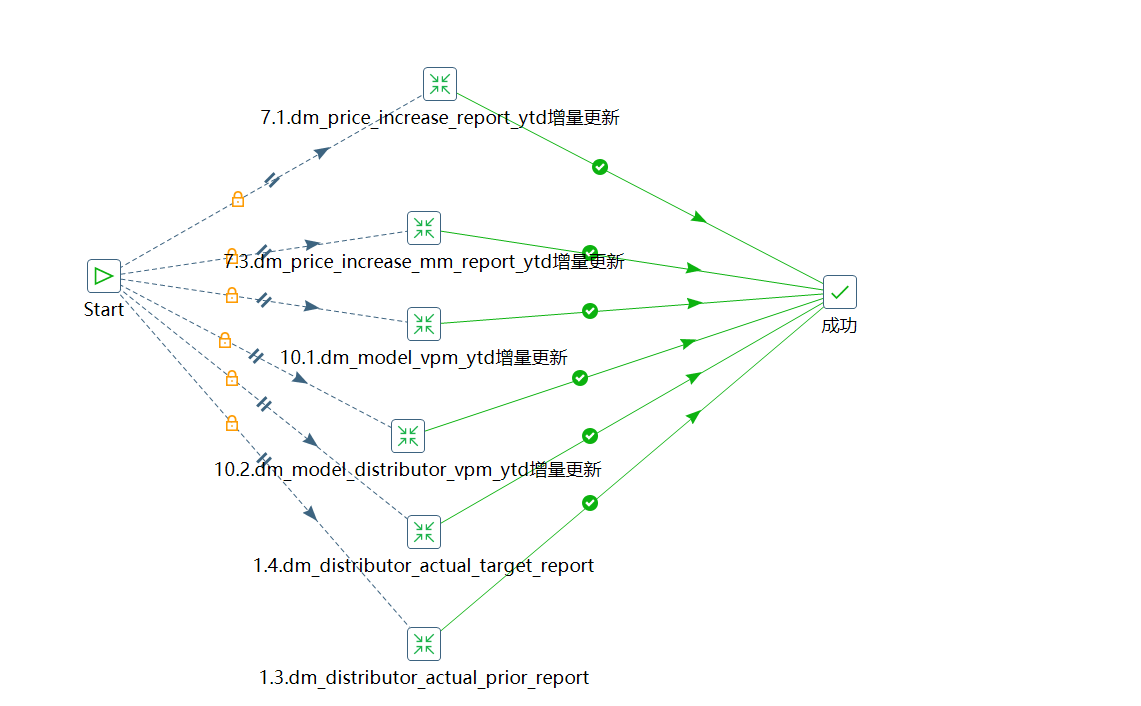
1. 示意图



dm\_1



dm\_2

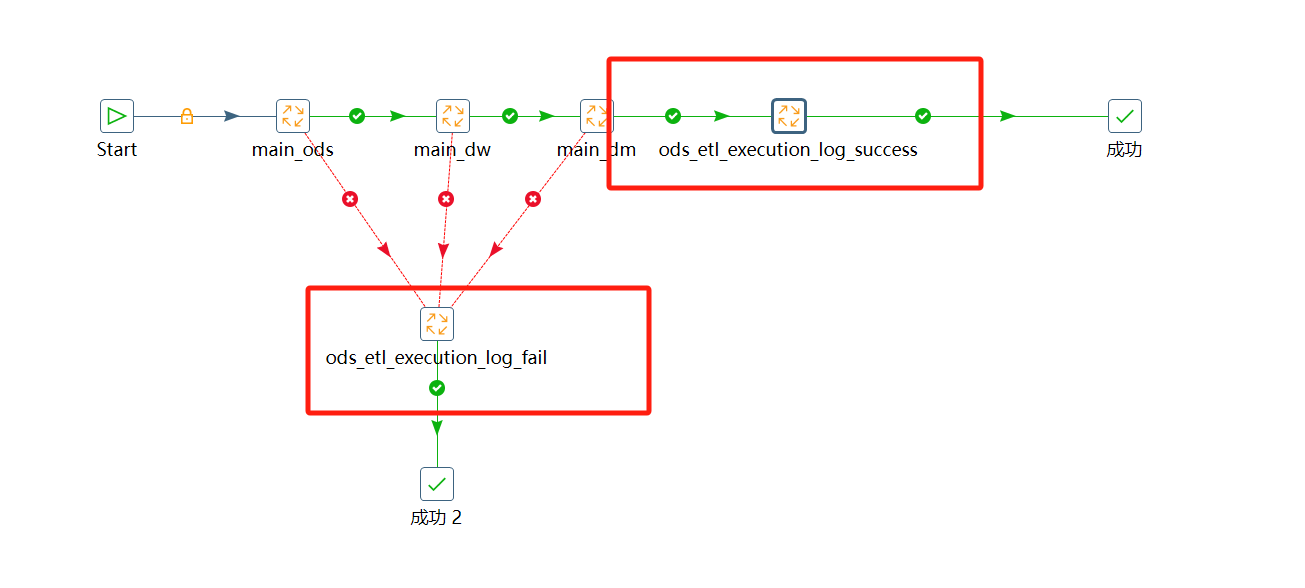


1. 描述

整体分为两个个模块，dm\_1,dm\_2.

## 执行异常处理

1. 示意图



1. 描述

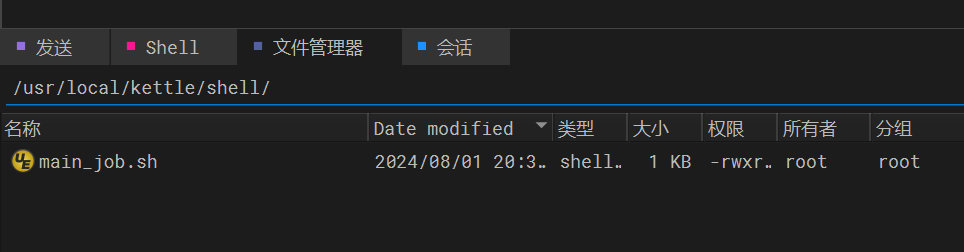
整体分为两个个模块，ods\_etl\_execution\_log\_success，ods\_etl\_execution\_log\_fail

# ETL执行计划配置

## 定时任务

Linux定时任务每天上午10点执行/usr/local/kettle/shell/main\_job.sh。

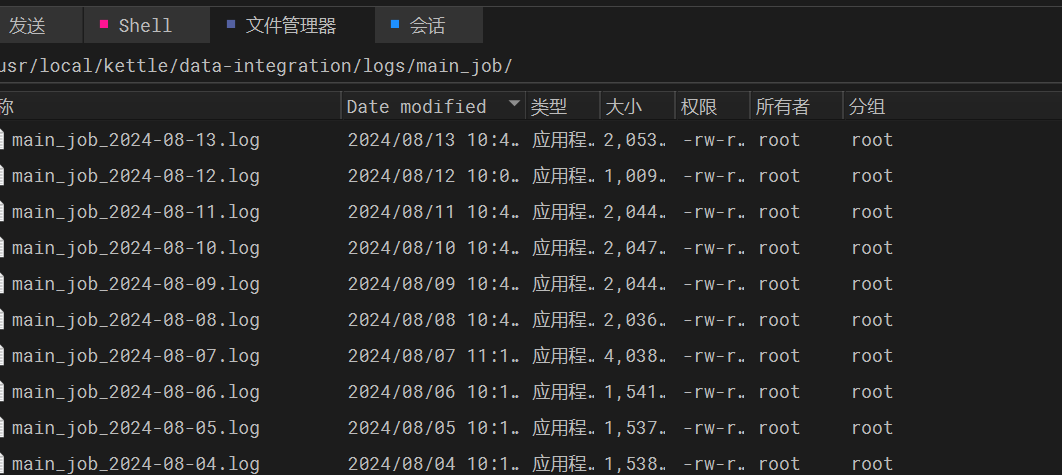
main\_job.sh该脚本执行main\_job作业，当前执行时长大概1小时。



## 日志输出目录

目录：/usr/local/kettle/data-integration/logs/main\_job/

每天生产一份执行日志。



# ETL异常处理

## 1.ETL是否执行成功

查询执行日志表 fine\_ods. ods\_etl\_execution\_log，该表只保存一条记录。

execution\_datetime :执行时间

execution\_result：执行结果 （success成功、fail失败）

执行时间为当天，执行结果为success，表示ETL正常执行，其他情况均为执行失败。

## 执行失败时：查询执行日志

目录：/usr/local/kettle/data-integration/logs/main\_job/

每天生产一份执行日志，分析日志查询具体问题，可联系运维处理。

