**家乐福销售需求计划项目运维文档**

# 检查运行状态

前往Airflow安装目录，确保下列两个pid格式文件存在。

|  |  |
| --- | --- |
| airflow-webserver.pid | Airflow UI 管理界面的进程pid |
| airflow-scheduler.pid | Airflow scheduler 调度器的进程pid |

使用ps命令查看运行中的进程，确保上述文件中的pid存在与查询结果。如pid在结果中不存在，需要删除pid文件并重启对应的airflow服务。

检查目录中的err文件。正常运行状态下err文件应该为空。如不为空需打开文件，查看当前问题。

# 检查运行结果

需求计划项目每次运行结果都会写入表vartefact.forecast\_script\_runs 中。表包含如下信息

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 序号 | 字段英文名称 | 字段中文名称 | 类型 | 长度 | 允许空 | 缺省值 | 说明 |
| 1 | insert\_time | 写入时间 | TIMESTAMP |  |  |  | 写入时间 |
| 2 | run\_date | 运行日期 | STRING |  |  |  | 运行输入日期 |
| 3 | run\_status | 运行状态 | STRING |  |  |  | 运行结果 |
| 4 | script\_name | 脚本名称 | STRING |  |  |  | 运行模块名称 |
| 5 | script\_type | 脚本类型 | STRING |  |  |  | 运行类型，如模拟运算 |
| 6 | script\_parameter | 脚本参数 | STRING |  |  |  | 运行输入参数 |
| 7 | output | 输出 | STRING |  |  |  | 模块运行输出，如订单数量 |
| 8 | info | 信息 | STRING |  |  |  | 模块运行相关信息，如运行时长 |
| 9 | error | 错误 | STRING |  |  |  | 模块运行产生的错误 |

需求计划项目日志在如下文件夹中

|  |
| --- |
| /data/jupyter/Carrefour-China-Supply-Chain-Forecast/logs/forecast\_orderflow |

需求计划项目会在运行时生成下列日志文件

|  |  |
| --- | --- |
| 日志文件名 | 说明 |
| run\_parameter.log | 运行输入参数 |
| run\_dm\_order.log | DM 订单日志 |
| run\_onstock\_store\_order.log | 存库商品门店订单生成日志 |
| run\_xdock\_order.log | 越库商品门店订单生成日志 |
| run\_dc\_order.log | 大仓订单生成日志 |

Airflow也会在统一文件夹生成运行日志。运行日志中的重要内容均已同时写入上述日志。通常不需要查看Airflow运行日志

# 订单逻辑调试

订单逻辑调试可以重新计算给定单品的订货，并将计算过程和结果输出至给定的表。订单逻辑调试生成的订单信息不会写入订单表，只错调试使用。

使用如下模板编辑订单逻辑调整运行语句，

spark-submit --class 运行主程序 --master yarn --driver-memory 6G --num-executors 14 jar/carrefour\_forecast\_2.11-0.1.jar 订单日 其它参数 debug\_table=数据库名.表名 >> 日志文件

* 运行主程序

此选项有三个可选值，对应不同类型的订单计算

|  |  |
| --- | --- |
| carrefour.forecast.process.OnStockForecastProcess | 计算存库单品门店订单 |
| carrefour.forecast.process.XDockingForecastProcess | 计算越库单品订单 |
| carrefour.forecast.process.DcForecastProcess | 计算寸库单品大仓订单 |

* 订单日

订单日为yyyMMdd 格式订单日期，如20190902

* 其它参数

订单逻辑调试可使用其它参数来限制运行范围。用户可选择一个或多个下列参数。

|  |  |  |
| --- | --- | --- |
| 其它参数 | 参考格式 | 用途 |
| item\_id | item\_id=1380542 | 限制程序只运行此item\_id对应的单品 |
| sub\_id | sub\_id=1380542 | 限制程序只运行此sub\_id对应的单品 |
| store\_code | store\_code=101 | 限制程序只运行此store\_code对应的门店 |

在linux环境中进入下列文件夹，运行编辑好的语句

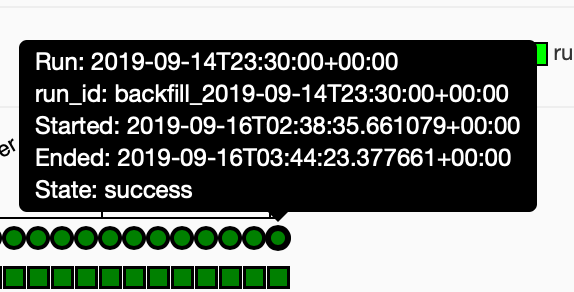
|  |  |
| --- | --- |
| /data/jupyter/Carrefour-China-Supply-Chain-Forecast/workflow\_integration | 需求计划程序根文件夹 |

# 重新生成

打开浏览器，进入Airflow UI 管理界面。点击对应DAG进入DAG界面。进入后默认显示Tree View。如未显示点击Tree View。

图片包含 屏幕截图

描述已自动生成

讲鼠标移至圆形图标，即可查看DAG的每次运行。选择需要的运行日期。如需重新生成2019-09-15日的订单，需选择Run：2019-09-13T23:30:00+00:00的运行。

点击圆形图标来选择此次运行，在弹出的窗口中选择Clear。此操作会将运行状态重设为未执行。

图片包含 屏幕截图

描述已自动生成

在新窗口中选择OK！，完成重设。

Airflow只会自动启动最新的运行。对于过去日期的运行，还需要在linux环境中执行以下指令。

airflow backfill {DAG名} -s {开始日期，包括} -e {结束日期，不包括}

如需重新执行Run：2019-09-13T23:30:00+00:00的运行，命令为

airflow backfill forecast\_orderflow -s 20190913 -e 20190914