

# 1、任务调度配置

(MySQL 5.7.13-Ubuntu0.16.04.2-log) dw\_master/rt/ETL\_CONTROL/ETL\_CONTROL\_TASK

The screenshot shows the MySQL Workbench interface with the following details:

- Database:** ETL\_CONTROL
- Tables:** BACK\_FLOW/BACK\_FLOW, nono\_new\_canal/db\_no..., DB\_NONO\_FINANCE/db\_no..., ETL\_CONTROL/ETL\_CONT..., nonobank\_app/hor/maiya/..., nonobank\_app/hor/maiya/...
- Structure:** Structure tab selected.
- Task Configuration:** A large table view showing task details. Key columns include:
  - TASK\_ID:** Task ID (e.g., 110000001, 110000002, ...)
  - TASK\_NAME:** Task Name (e.g., usr\_member.borrows\_dunning\_contact, db\_nono.borrows, ...)
  - RUN\_FLAG:** Run Flag (e.g., 1 0 0 0 1/1 ? \*)
  - REFRESH\_FREQ:** Refresh Frequency (e.g., 每日0点执行)
  - REFRESH\_FREQ\_DESC:** Refresh Frequency Description (e.g., 每日0点执行)
  - TASK\_PRD:** Task Product (e.g., 5, 9, 5, 5, 5, ...)
  - LAST\_START\_TIME:** Last Start Time (e.g., 2018-10-10 00:52:02, 2018-10-10 00:00:02, ...)
  - LAST\_END\_TIME:** Last End Time (e.g., 2018-10-09 01:01, 2018-10-10 00:31, ...)
- Table Information:** Shows table metadata including creation date (2018/7/25), last update (2018/10/10), engine (InnoDB), rows (1,002), size (176.0 KB), and encoding (utf8).
- Procs & Funcs:** SP\_GET\_ETL\_JOB\_INFO, SP\_TASK\_INSERT\_LOCK.

Rows 1 - 100 of 1,002 from table; 1 row selected

## 2、正在执行任务列表

(MySQL 5.7.13-Ubuntu 0.16.04.2-0-log) dw\_master/t/ETL\_CONTROL/ETL\_CONTROL\_RESOURCE\_POOL

ETL\_CONTROL  
Select Database

Structure Content Relations Triggers Table Info Query

BACK\_FLOW/BACK\_FLOW nono\_new\_cana/db\_no... DB\_NOVO\_FINANCE/db\_no... ETL\_CONTROL/ETL\_CONT... nonobank\_app/hor/malya/... nonobank\_app/hor/malya/... +

Q Filter

Search: TASK\_ID LIKE

TABLES

- ETL\_CONTROL\_JDBC
- ETL\_CONTROL\_JOB
- ETL\_CONTROL\_JOB\_LOG
- ETL\_CONTROL\_JOB\_LOG\_DETAIL
- ETL\_CONTROL\_LOG
- ETL\_CONTROL\_RESOURCE\_POOL\_0929
- ETL\_CONTROL\_TASK
- ETL\_CONTROL\_TASK\_H2M
- ETL\_CONTROL\_TASK\_LOG
- ETL\_CONTROL\_TASK\_LOG\_DETAIL
- ETL\_CONTROL\_TASK\_M2H
- ETL\_CONTROL\_TASK\_M2H\_COLUMN\_MAP
- ETL\_CONTROL\_TASK\_M2H\_DATA\_TYPE\_MAP
- ETL\_CONTROL\_TASK\_M2M
- ETL\_CONTROL\_TASK\_M2M\_COLUMN\_MAP
- ETL\_CONTROL\_TASK\_PARENT
- ETL\_CONTROL\_TASK\_SCRIPT
- ETL\_CONTROL\_TASK\_SCRIPT\_LOG
- ETL\_CONTROL\_TASK\_SPARK

PROCS & FUNCS

- SP\_GET\_ETL\_JOB\_INFO
- SP\_TASK\_INSERT\_LOCK

TABLE INFORMATION

- created: 2018/3/7
- updated: 2018/10/10
- engine: InnoDB
- rows: 15
- size: 16.0 KB
- encoding: utf8

+ × × × × | || + - × × × 15 rows in table

### 3、任务执行状态

(MySQL 5.7.13-Ubuntu0.16.04.2-log) dw\_master/rt/ETL\_CONTROL/ETL\_CONTROL\_TASK\_LOG

TABLES	TASK_ID	JOB_ID	RUN_ID	TASK_NAME	TASK_BEGIN_TIME	TASK_END_TIME	JOB_NAME	STATUS
ETL_CONTROL	110000117	3	271515	nonobank_app.account_relation	2018-10-10 01:04:02	NULL	BI-APP03	2
ETL_CONTROL_JOB	110000111	2	234386	db_nono.borrows_guarantee	2018-10-10 01:04:02	NULL	BI-APP02	2
ETL_CONTROL_JOB_LOG	110000005	3	271514	usr_member.user_student_info	2018-10-10 01:03:03	NULL	BI-APP03	2
ETL_CONTROL_JOB_LOG_DETAIL	110000228	3	271513	db_nono.coupon_us	2018-10-10 01:01:02	NULL	BI-APP03	2
ETL_CONTROL_LOG	110000162	3	271513	nonobank_app.user_interest	2018-10-10 01:01:02	NULL	BI-APP03	2
ETL_CONTROL_RESOURCE_POOL	110000161	2	234385	nonobank_app.interest_info	2018-10-10 01:00:02	NULL	BI-APP02	2
ETL_CONTROL_RESOURCE_POOL_0929	110000694	1	234430	usr_member.user_relation_person_info	2018-10-10 00:59:02	NULL	BI-APP01	2
ETL_CONTROL_TASK	110000692	2	234384	db_user_log.user_register_detail_log	2018-10-10 00:59:01	NULL	BI-APP02	2
ETL_CONTROL_TASK_H2M	110000685	1	234429	usr_member.user_pengyuan_edu_record	2018-10-10 00:58:01	NULL	BI-APP01	2
ETL_CONTROL_TASK_LOG	110000004	2	234383	usr_member.user_info	2018-10-10 00:58:01	NULL	BI-APP02	2
ETL_CONTROL_TASK_LOG_DETAIL	110000481	3	271512	db_nono.inv_trd_order	2018-10-10 00:57:02	NULL	BI-APP03	2
ETL_CONTROL_TASK_M2H	110000107	2	234382	db_nono.activity_prolocutor	2018-10-10 00:55:02	2018-10-10 00:59:19	BI-APP02	1
ETL_CONTROL_TASK_M2H_COLUMN_MAP	110000098	3	271511	db_nono.user_hkconpon	2018-10-10 00:55:02	2018-10-10 01:03:01	BI-APP03	1
ETL_CONTROL_TASK_M2H_DATA_TYPE_MAP	110000680	1	234428	usr_member.user_mobile_online	2018-10-10 00:54:03	2018-10-10 00:58:34	BI-APP01	1
ETL_CONTROL_TASK_M2M	110000003	2	234381	db_nono.borrows_prepare	2018-10-10 00:54:03	NULL	BI-APP02	2
ETL_CONTROL_TASK_M2M_COLUMN_MAP	110000437	2	234380	usr_member.user_face_beyond_result	2018-10-10 00:53:02	2018-10-10 00:57:16	BI-APP02	1
ETL_CONTROL_TASK_PARENT	110000001	1	234427	usr_member.borrows_dunning_contact	2018-10-10 00:52:02	NULL	BI-APP01	2
ETL_CONTROL_TASK_SCRIPT	110000650	2	234379	db_nono.user_csmfinfo	2018-10-10 00:51:02	2018-10-10 00:53:36	BI-APP02	1
ETL_CONTROL_TASK_SCRIPT_LOG	110001353	3	271510	usr_member.user_address_info	2018-10-10 00:51:01	2018-10-10 01:03:25	BI-APP03	1
ETL_CONTROL_TASK_SPARK	110000647	3	271509	BIA_FACT_CHANNEL_GROUP	2018-10-10 00:49:01	2018-10-10 00:50:44	BI-APP03	1
PROCS & FUNCS	110000687	2	234378	db_nono.user_credit_investigation	2018-10-10 00:47:01	2018-10-10 00:58:44	BI-APP02	1
SP_GET_ETL_JOB_INFO	110000647	2	234378	CHANNEL_MANAGER.ssb1.market_settle	2018-10-10 00:47:01	2018-10-10 00:50:57	BI-APP02	1
SP_TASK_INSERT_LOCK	110000526	2	234377	db_nono.products_config	2018-10-10 00:44:02	2018-10-10 00:47:01	BI-APP02	1
SP_TASK_INSERT_LOCK	110000224	2	234376	platform_honor.HOR_THIRD_CHANNEL	2018-10-10 00:43:02	2018-10-10 00:46:06	BI-APP02	1
SP_TASK_INSERT_LOCK	110001346	3	271508	usr_member.user_loan_info_new	2018-10-10 00:08:02	2018-10-10 00:13:42	BI-APP03	1
SP_TASK_INSERT_LOCK	110001333	3	271508	market.mkt_continue_bid_order	2018-10-10 00:08:02	2018-10-10 00:12:27	BI-APP03	1
SP_TASK_INSERT_LOCK	110001332	2	234375	market.mkt_red_packet	2018-10-10 00:08:02	2018-10-10 00:12:13	BI-APP02	1
SP_TASK_INSERT_LOCK	110000426	1	234426	db_nono.debt_package	2018-10-10 00:07:03	2018-10-10 00:11:00	BI-APP01	1
SP_TASK_INSERT_LOCK	110001331	3	271507	db_nono.inv_va_reinvest_set_log	2018-10-10 00:07:02	2018-10-10 00:11:02	BI-APP03	1
SP_TASK_INSERT_LOCK	110001265	3	271506	CHANNEL_MANAGER.reg_resource_mark	2018-10-10 00:06:02	2018-10-10 00:10:37	BI-APP03	1
SP_TASK_INSERT_LOCK	11000266	1	234425	db_nono.mobile_site	2018-10-10 00:06:02	2018-10-10 00:10:56	BI-APP01	1
SP_TASK_INSERT_LOCK	110001239	3	271506	db_nono.borrows_credit_promote_amount	2018-10-10 00:06:02	2018-10-10 00:48:16	BI-APP03	1
SP_TASK_INSERT_LOCK	11000382	1	234425	db_nono.borrows_guarantor	2018-10-10 00:06:02	2018-10-10 00:10:36	BI-APP01	1

## 4、任务执行日志

(MySQL 5.7.13-Ubuntu0.16.04.2-log) dw\_master/rt/ETL\_CONTROL/ETL\_CONTROL\_TASK\_LOG\_DETAIL

Select Database: BACK\_FLOW/BACK\_FLOW

Structure Content Relations Triggers Table Info Query

ETL\_CONTROL

DB\_NONO\_FINANCE/db\_no...

ETL\_CONTROL/ETL\_CONT...

nonobank\_app/hor/maiya...

nonobank\_app/hor/maiya...

Search: ID LIKE Filter

Tables:

ID	TASK_ID	JOB_ID	RUN_ID	EVENT_TIME	EVENT_TYPE	EVENT_DESC
2236974	110000004	2	234383	2018-10-10 01:08:10	INFO	任务类型是F,跳过CUR MERGER。
2236973	110000004	2	234383	2018-10-10 01:08:10	INFO	CUR MERGER 开始。
2236975	110000004	2	234383	2018-10-10 01:08:10	INFO	MERGE HST 开始。
2236970	110000111	2	234386	2018-10-10 01:05:32	INFO	判断目标字段是否有变更。
2236971	110000111	2	234386	2018-10-10 01:05:32	INFO	开始生成sqoop脚本
2236972	110000111	2	234386	2018-10-10 01:05:32	INFO	SQOOP TMP表数据同步开始
2236969	110000117	3	271515	2018-10-10 01:04:06	INFO	SQOOP TMP表数据同步开始
2236968	110000117	3	271515	2018-10-10 01:04:05	INFO	开始生成sqoop脚本
2236967	110000117	3	271515	2018-10-10 01:04:04	INFO	判断目标字段是否有变更。
2236966	110000111	2	234386	2018-10-10 01:04:03	INFO	判断目标表是否存在。
2236965	110000117	3	271515	2018-10-10 01:04:03	INFO	判断目标表是否存在。
2236963	110000111	2	234386	2018-10-10 01:04:03	INFO	比较和同步元数据。
2236964	110000117	3	271515	2018-10-10 01:04:03	INFO	比较和同步元数据。
2236962	110000117	3	271515	2018-10-10 01:04:02	INFO	判断源表是否存在。
2236961	110000111	2	234386	2018-10-10 01:04:02	INFO	判断源表是否存在。
2236957	110000111	2	234386	2018-10-10 01:04:02	INFO	开始: 处理任务
2236958	110000111	2	234386	2018-10-10 01:04:02	INFO	开始: 执行优先级为 1 的子任务。
2236960	110000117	3	271515	2018-10-10 01:04:02	INFO	开始: 执行优先级为 1 的子任务。
2236959	110000117	3	271515	2018-10-10 01:04:03	INFO	开始: 处理任务
2236956	110000228	3	271513	2018-10-10 01:04:01	INFO	MERGE HST 开始。
2236955	110000162	3	271513	2018-10-10 01:03:57	INFO	MERGE HST 开始。
2236954	110000005	3	271514	2018-10-10 01:03:56	INFO	SQOOP TMP表数据同步开始
2236953	110000005	3	271514	2018-10-10 01:03:55	INFO	开始生成sqoop脚本
2236952	110000005	3	271514	2018-10-10 01:03:55	INFO	判断目标字段是否有变更。
2236950	110000209	2	234371	2018-10-10 01:03:25	INFO	完成: 优先级为 1 的子任务。
2236949	110001353	3	271510	2018-10-10 01:03:25	INFO	成功! 结束处理
2236951	110000209	2	234371	2018-10-10 01:03:25	INFO	成功! 结束处理
2236947	110000209	2	234371	2018-10-10 01:03:24	INFO	开始刷新IMPALA元数据。
2236948	110001353	3	271513	2018-10-10 01:03:24	INFO	完成: 优先级为 1 的子任务。
2236946	110001353	3	271513	2018-10-10 01:03:24	INFO	开始刷新IMPALA元数据。
2236945	110000162	3	271513	2018-10-10 01:03:13	INFO	CUR MERGER 开始。
2236944	110000228	3	271513	2018-10-10 01:03:05	INFO	CUR MERGER 开始。
2236942	110000005	3	271514	2018-10-10 01:03:04	INFO	比较和同步元数据。

Table Information:

- created: 2018/8/10
- updated: 2018/10/10
- engine: InnoDB
- rows: ~1,905,622
- size: 184.7 MB
- encoding: utf8
- auto\_increment: 2,236,976

Rows 1 - 100 of ~1,905,622 from table

## 5、任务执行详细日志

```
(MySQL 5.7.13-Ubuntu0.16.04.2-log) dw_master(rt)/ETL_CONTROL/ETL_CONTROL_LOG
ETL_CONTROL Structure Content Relations Triggers Table Info Query
BACK_FLOW/BACK_FLOW nono_new_canal/db_nono... DB_NONO_FINANCE/db_no... ETL_CONTROL/ETL_CONT... nonobank_app/hor/maliya... nonobank_app/hor/maliya...
Field: "LOG_FIELD" - LONGTEXT(1431655765) ut8
Edit All Fields in pop-up Sheet
2018/10/10 00:55:05 - 040102_SET_J_SQOOP_SQL - Loading transformation from repository [040102_SET_J_SQOOP_SQL] in directory [/ETL_SCHEDULING/03_TASK_EXECUTOR/0302_M2H_EXECUTOR]
2018/10/10 00:55:05 - 040102_SET_J_SQOOP_SQL - Dispatching started for transformation [040102_SET_J_SQOOP_SQL]
2018/10/10 00:55:05 - 040102_SET_J_SQOOP_SQL - Disposing reading query, closing connection.
2018/10/10 00:55:05 - 040102_SET_J_SQOOP_SQL - Setting environment variables...
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - Set variable V_T_Sqoop to value [
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - sqoop import)
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - D-mapsreduce.output.fileoutputformat.compress=false )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - D-mapsreduce.output.format=csv )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - D-mapsreduce.job.name=Staging-TMP_ACTIVITY_PROLOCUTOR )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --connect jdbc:mysql://172.16.2.109:3313/db_nono?allowPublicKeyRetrieval=true --username risk --password "Ee9kzcorDZnjln4" )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --query "SELECT * FROM db_nono.activity_prolocutor WHERE update_time>='2018-10-08 00:00:00' AND CONDITIONS"
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --map-column-types id=integer "
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --boundary-query "select MIN(id),MAX(id) from db_nono.activity_prolocutor WHERE (update_time>='2018-10-08 00:00:00' )"
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - -m 11
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - -hive-drop-import-delims )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --fields-terminated-by '\001' --lines-terminated-by '\n' --null-string '\N' --null-non-string '\n'
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --hive-import )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --hive-database tmp )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --hive-overwrite )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --hive-table TMP_ACTIVITY_PROLOCUTOR )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --delete-target-dir '
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --target-dir '/user/hive/warehouse/tmp.db/TMP_ACTIVITY_PROLOCUTOR' )
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - --default-character-set=utf-8 ;
2018/10/10 00:55:05 - 表插入 创建临时表,O - Finished reading query, closing connection.
2018/10/10 00:55:05 - 表插入 T_Sqoop_O - Finished processing (I=1, O=0, R=0, W=1, U=0, E=0)
2018/10/10 00:55:05 - 表插入 创建临时表,O - Finished reading query, closing connection.
2018/10/10 00:55:05 - 表插入 T_Sqoop_O - Finished processing (I=1, O=0, R=0, W=1, U=0, E=0)
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - Finished after 1 rows
2018/10/10 00:55:05 - 读设置 T_Sqoop_O - Finished processing (I=0, O=0, R=1, W=1, U=0, E=0)
2018/10/10 00:55:05 - 执行SQL文本[字段映射转换] - O - Finished reading query, closing connection.
2018/10/10 00:55:05 - 执行SQL文本[字段映射转换] - O - Finished reading query, closing connection (I=0, O=1, R=1, W=1, U=0, E=0)
2018/10/10 00:55:05 - 0401_SET_J_SQOOP_SQL - Starting entry [Success] [result=true]
2018/10/10 00:55:05 - 0401_SET_J_SQOOP_SQL - Finished job entry [Success] [result=true]
2018/10/10 00:55:06 - 0401_SET_J_SQOOP_SQL - Finished job entry [040102_SET_J_SQOOP_SQL] [result=true]
2018/10/10 00:55:06 - 0401_SET_J_SQOOP_SQL - Finished job entry [TASK_ETL_TYPE] [result=true]
2018/10/10 00:55:06 - 0401_SET_J_SQOOP_SQL - Finished reading entry [LOG_EXECUTE_DETAIL] [执行sqoop脚本]
2018/10/10 00:55:06 - 04_SYNC_TABLE_DATA - Starting entry [0402_EXEC_SQOOP_SCRIPT]
2018/10/10 00:55:06 - 0402_EXEC_SQOOP_SCRIPT - Starting entry [Shell]
2018/10/10 00:55:06 - Shell - Running on platform: Linux
2018/10/10 00:55:06 - Shell - Executing command: /tmp/sqoop-12d52329-cba4-11e8-8852-000c2b666363/bin/hadoop fs -put /tmp/sqoop-12d52329-cba4-11e8-8852-000c2b666363/* /user/hive/warehouse/tmp.db/TMP_ACTIVITY_PROLOCUTOR
2018/10/10 00:55:06 - Shell - (stdout) Please set ACCUMULO_HOME to the root of your Accumulo installation.
2018/10/10 00:55:07 - Shell - (stdout) 2018-10-10 00:55:07.966 INFO [main] sqoop.Sqoop: Sqoop version: 1.4.6-cdh5.8.0
2018/10/10 00:55:07 - Shell - (stdout) 2018-10-10 00:55:07.982 WARN [main] tool.BaseSqoopTool: [BaseSqoopTool]java:applyCredentialsOptions(1042) - Setting your password on the command-line is insecure. Consider using -P instead.
2018/10/10 00:55:07 - Shell - (stdout) 2018-10-10 00:55:07.983 WARN [main] tool.BaseSqoopTool: [BaseSqoopTool]java:validateHiveOptions(1443) - It seems that you're doing hive import directly into default
2018/10/10 00:55:07 - Shell - (stderr) 2018-10-10 00:55:07.983 WARN [main] tool.BaseSqoopTool: [BaseSqoopTool]java:validateHiveOptions(1443) - hive warehouse directtrue which is not supported. Sqoop is
Open Save... Import Export Cancel OK
```

# 6、Yarn 资源调度

The screenshot shows the YARN scheduler interface. At the top, there's a tree view of application logs:

- root.ad-hoc
- root.default
- root.etl-kpl
- root.etl-log
- root.online
- root.report
- root.spark-streaming
- root.etl-dw

Below the tree view is a table titled "Show 20 entries" displaying the following information for four running applications:

ID	User	Name	Application Type	Queue	Fair Share	StartTime	FinishTime	State
application_1539047048418_2314	root	Sqoop_TMP_BORROWS_PREPARE	MAPREDUCE	root.etl-dw	161320	Wed Oct 10 00:54:47 +0800 2018	N/A	RUNNING
application_1539047048418_2030	root	Sqoop_TMP_INVT_BORROWS_ACCEPT_UNPAID	MAPREDUCE	root.etl-dw	161320	Wed Oct 10 00:00:32 +0800 2018	N/A	RUNNING
application_1539047048418_2364	hive	KETTLE_MERGER_ODS_T_USER_REGISTER_DETAIL_LOG_HIST	MAPREDUCE	root.etl-dw	161320	Wed Oct 10 01:09:14 +0800 2018	N/A	RUNNING
application_1539047048418_2363	hive	KETTLE_MERGER_ODS_T_USER_INFO_HIST	MAPREDUCE	root.etl-dw	161320	Wed Oct 10 01:08:20 +0800	N/A	RUNNING

## 7、调度监控

