

Variable Costs

- Description
- Tools: Talend
 - From source to ODS
 - From DM to Operational Dashboard
 - Access rights
- Source
- Destination
 - DataOcean
 - Product
 - Format
 - Sizing
- Loading
 - 1.1 Full load
 - 1.2. Reloading data
 - 1.3 Plan to schedule
 - 1.4 Timing
- Criticality
- Logging

Description

The data source getting from prj-data-dq-selfservice, which get from DataOcean_industrialreporting_workspace.VariableCosts tables in order to get the information of VariableCosts.

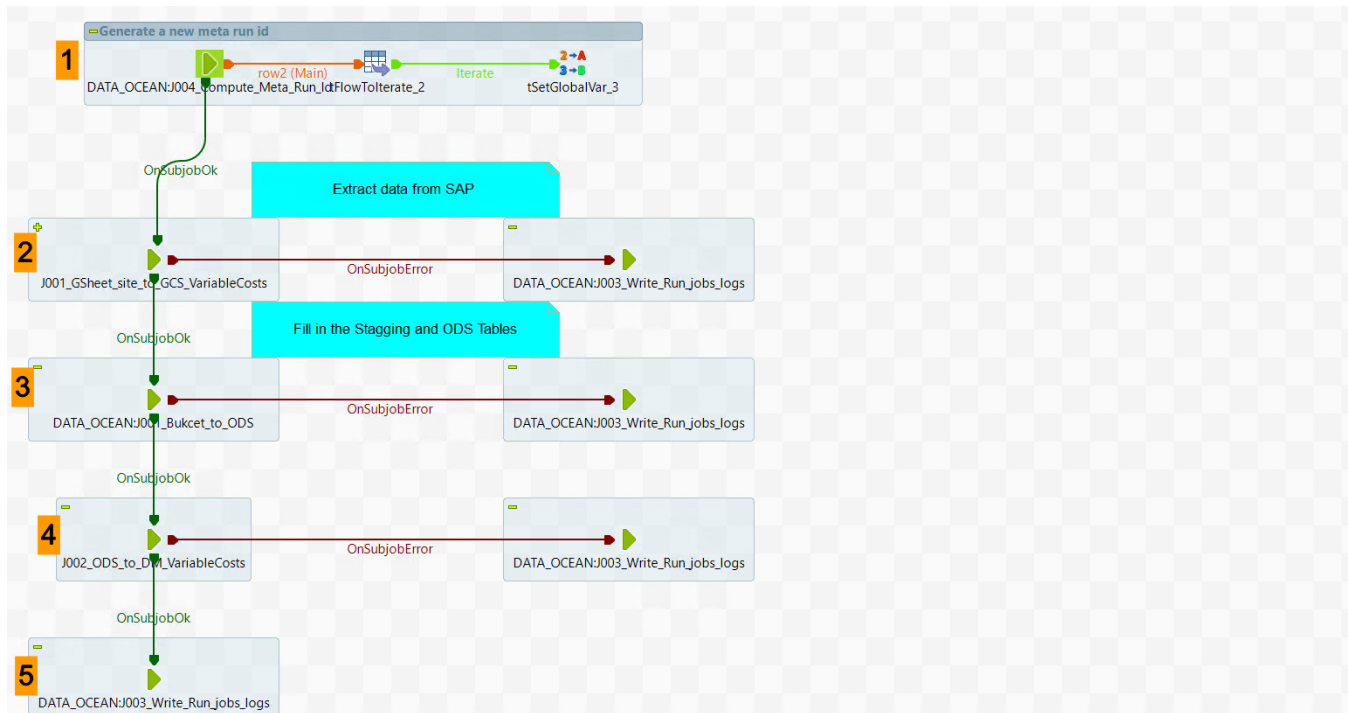
Tools: Talend

Talend project IND_DASHBOARD

- F001_GSheet_site_to_BQ_VariableCosts

From source to ODS

Job: F001_GSheet_site_to_BQ_VariableCosts (VariableCosts)



1. Define variable and Generate meta_run_id
2. Call reference job to extract from prj-data-dq-selfservice-test.DataOcean_industrialreporting_workspace.VariableCosts and save the output file to bucket cs-ew1-prj-data-industrial-dash-dev-staging/vc
3. Load from Bucket to STG and ODS

4. Move data from ODS to DM
5. Update log

From DM to Operational Dashboard

Dataflow

prj-data-dq-selfservice-test.DataOcean_industrialreporting_workspace.VariableCosts prj-data-industrial-dash-dev.STG.
STG_FIL_0000_0000_F001_F_M_vc prj-data-industrial-dash-dev.ODS.ODS_FIL_0000_F001_F_M_vc prj-data-industrial-dash-dev.DM.
FACT_VariableCosts prj-data-industrial-dash-dev.DPL.V_FACT_VariableCosts

Access rights

It is required to access prj-data-dq-selfservice-test project

Source

BigQuery

Project = prj-data-dq-selfservice-dev

BQ Dataset = DataOcean_industrialreporting_workspace

BQ Table = VariableCosts

Destination

DataOcean

- Bucket = cs-ew1-prj-data-industrial-dash-dev-staging/
 - vc
 - FIL_IND_0000_0000_F001_20241126224947_0000_F_M_vc.csv
- STG Table names =
 - STG_FIL_0000_0000_F001_F_M_vc
- ODS Table names =
 - ODS_FIL_0000_F001_F_M_vc
- DM Table names =
 - FACT_variableCosts
- DPL View names =
 - V_FACT_variableCosts

Product

- GCP = prj-data-industrial-dash-[env]

Format

columnar format

Sizing

- TG_FIL_0000_0000_F001_F_M_vc around 7,392 records

Loading

1.1 Full load

- **VariableCosts** Source to ODS : TASK_DailyExecution_VariableCosts

1.2. Reloading data

VariableCosts

- Just run the TASK_DailyExecution_VariableCosts.

1.3 Plan to schedule

It is scheduled by plans below on WS_IND_DASHBOARD

- TASK_DailyExecution_VariableCosts - Weekly every 30 min starting at 07am CET ~ 6pm CET

1.4 Timing

- Case 5 minutes from source to ODS (full)

Criticality

- Low?

Logging

Check the loading records

```
select job.job_name, job.meta_start_date, job.meta_execution_id, logs.meta_run_id, logs.meta_source_system, logs.meta_step, logs.meta_status, logs.
meta_num_lines, logs.meta_error_lines from STG.log_tables logs join STG.run_jobs job on logs.meta_run_id = job.meta_run_id
where logs.meta_run_id in (SELECT meta_run_id FROM STG.run_jobs order by meta_start_date desc limit 1000)
and job_name in ('F001_GSheet_site_to_BQ_VariableCosts')
and meta_start_date > DATE_SUB(CURRENT_TIMESTAMP(), INTERVAL 32 DAY)
order by job.meta_start_date desc
```