

Customer Complaint

- [Description](#)
- [Tools: Talend](#)
 - [From source to ODS](#)
 - [From DM to Operational Dashboard](#)
- [Access rights](#)
- [Source](#)
 - [Format](#)
- [Destination](#)
 - [DataOcean](#)
 - [Product](#)
 - [Format](#)
 - [Sizing](#)
 - [Assessment](#)
- [Loading](#)
 - [1.1 Incremental Load](#)
 - [1.3. Reloading data](#)
 - [1.4 Plan to schedule](#)
 - [1.5 Timing](#)
- [Criticality](#)
- [Logging](#)

Description

The data source getting from data ocean marketing, which get from salesforce (SFC) on the Case tables in order to get the number of complaints comparing with BW query QVBW_QRY_MVSDSO57_0002 QVBW - Deliveries Detials SO&PO (core query), which get the number of delivery items by filtering Solvay gbu only

Dataflow

prj-data-dm-marketing-dev.ODS.ODS_BWH_0000_F001_F_D_qvbw_qry_mvdsso57_0002 prj-data-dm-marketing-dev.ODS.V_ODS_BWH_0000_F001_F_D_qvbw_qry_mvdsso57_0002 prj-data-dm-marketing-dev.ODS.VB_BWH_no_of_delivery_orders prj-data-dm-marketing-dev.DS_prj_industrial_dash.V_BWH_no_of_delivery_orders => prj-data-industrial-dash-dev.DataOcean.V_BWH_no_of_delivery_orders prj-data-industrial-dash-dev.DPL.V_OP_QUALITY_CustomerComplaints_ListOfDeliveries

prj-data-dm-marketing-dev.ODS.ODS_SFC_0000_F001_F_M_case prj-data-dm-marketing-dev.ODS.V_ODS_SFC_0000_F001_F_M_case prj-data-dm-marketing-dev.ODS.VB_sfc_case prj-data-dm-marketing-dev.DS_prj_industrial_dash.V_sfc_case ==> prj-data-industrial-dash-dev.DataOcean.V_sfc_case DPL.V_OP_QUALITY_CustomerComplaints_ListOfComplaints

From these 2 sources, they will combine to DM.FACT_customer_complaints on GCP project prj-data-industrial-dash-dev by using Talend job

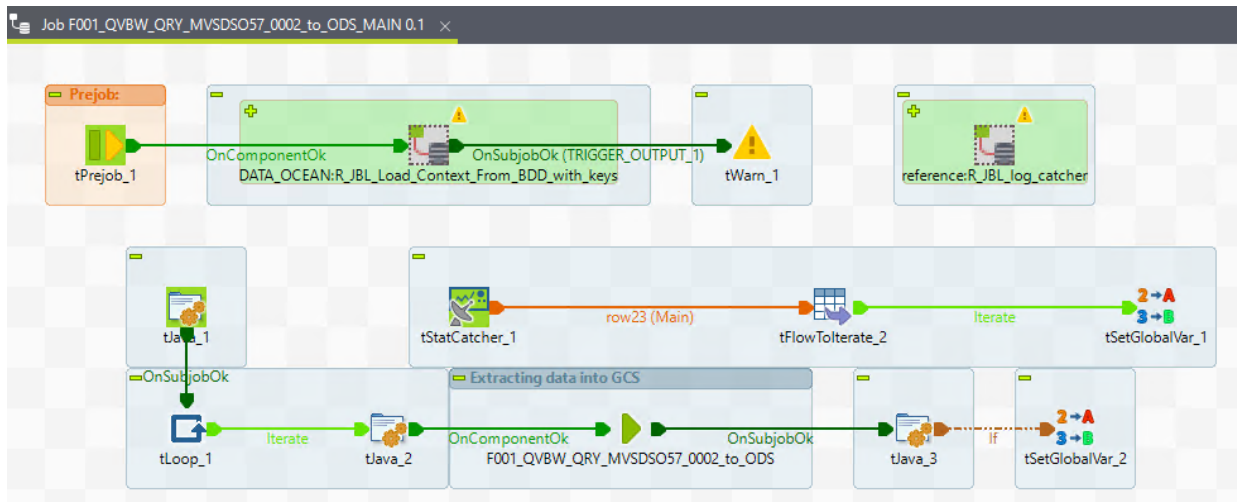
FlowName	Tasks	Description	Source_table/query	Target_Table	Complex transformation, if Yes,then logic
F001_QVBW_QRY_MVSDSO57_0002_to_ODS_MAIN	F001_QVBW_QRY_MVSDSO57_0002_to_ODS	Delivery query to get number of delivery	Xtract job TALEND_DEV_QVBW_QRY_MVSDSO57_0002	prj-data-dm-marketing-dev. ODS_BWH_0000_F001_F_D_qvbw_qry_mvdsso57_0002	
F020_SFC_case_customer_complaint_to_ODS	J020_SFC_case_customer_complaint_to_ODS	number of complaints	Salesforce on case table	prj-data-dm-marketing-dev. ODS_SFC_0000_F001_F_M_case	
F010_FACT_customer_complaints	J010_FACT_customer_complaints	Fact table of customer complaint	ODS_BWH_0000_F001_F_D_qvbw_qry_mvdsso57_0002 ODS_SFC_0000_F001_F_M_case	prj-data-industrial-dash-dev. FACT_customer_complaints	count distinct of delivery number and case id. Also calculate rolling 12 months by this script

Tools: Talend

From source to ODS

Talend project DATA_OCEAN_DOMAIN_MARKETING

F001_QVBW_QRY_MVSDSO57_0002_to_ODS_MAIN (number of delivery item)



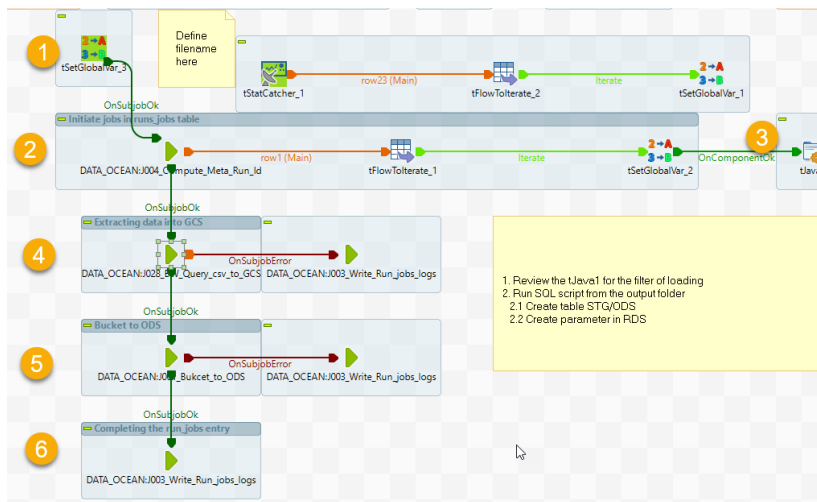
tJava will check the loading that it will be

- Load last 3 month : I_VAR_XTRACT_PARA_TALEND_QVBW_QRY_MVSDSO57_0002 = "last3month" and today is NOT the day in this parameter I_VAR_FULL_LOAD_WEEKDAY_QVBW_QRY_MVSDSO57_0002 (1 = Monday , 7 = Sunday)
- Full load 2 years : I_VAR_XTRACT_PARA_TALEND_QVBW_QRY_MVSDSO57_0002 = "last3month" and today is the day in this parameter I_VAR_FULL_LOAD_WEEKDAY_QVBW_QRY_MVSDSO57_0002
 - Loading 2 years will have one additional parameter to control the selection number of month to extract from BW query
 - I_VAR_FULL_LOAD_NB_MTH_QVBW_QRY_MVSDSO57_0002 = 3 meaning, it will extract 4 months (n+1). until it is current month. The parameter should not enter over 3 to avoid memory issue in Xtract.
- Reload : I_VAR_XTRACT_PARA_TALEND_QVBW_QRY_MVSDSO57_0002 = "&YYYYMM_Start=202306&YYYYMM_End=202309" (in case reload 202306 to 202309). It should not over 4 months to avoid Xtract memory issue.

It will loop to call job F001_QVBW_QRY_MVSDSO57_0002_to_ODS until current month. It will split loading max 6 months to avoid out of memory.

F001_QVBW_QRY_MVSDSO57_0002_to_ODS (number of delivery item)

Xtract job = TALEND_DEV_QVBW_QRY_MVSDSO57_0002, which filter on GBU = SD,SI,CH,CT,GY,PE and input period on actual goods issue date with format YYYYMM to YYYYMM

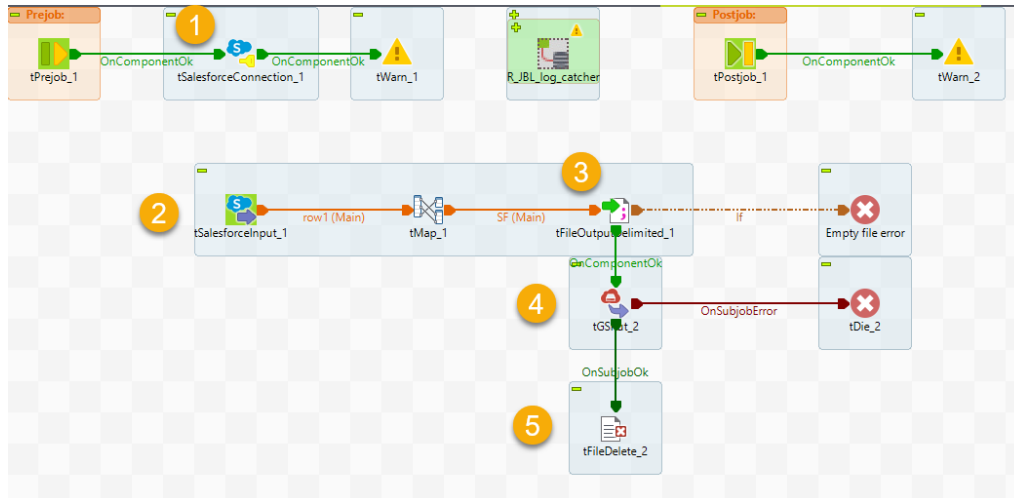


1. Define variable
2. Generate meta_run_id
3. Just print variable to check
4. Call reference job to load from BW and save the output file to bucket cs-ew1-prj-data-dm-marketing-dev-staging /TALEND_DEV_QVBW_QRY_MVSDSO57_0002
5. Load from Bucket to STG and ODS
6. Update log

⚠ Please note that most of the data in this query get from attribute of master data c_itm_gs which can be change automatically in BW when master data updated but it will not change in GCP. Therefore, the important field c_itm_gs__c_ingbuf2(intra gbu flag) can be changed over time and it may not aligned with BW. This is why we need to do the full reload 2 years weekly.

F020_SFC_case_customer_complaint_to_ODS (number of complaints)

Detail job J020_SFC_case_customer_complaint_to_ODS

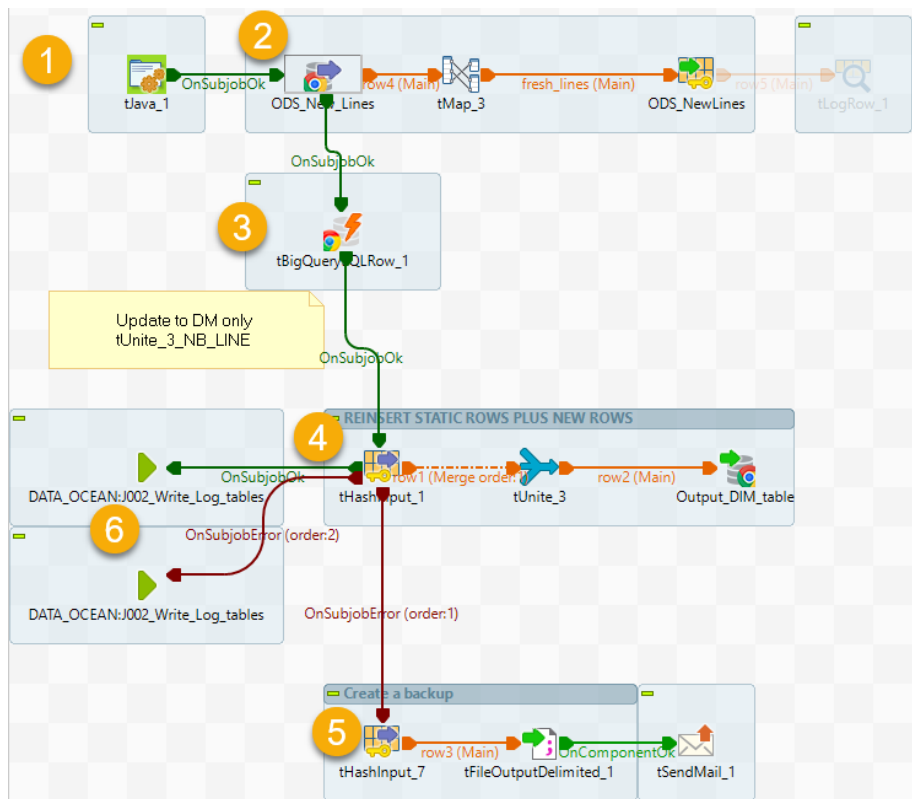


1. Connect to Salesforce (SFC)
2. Query SFC on the case table (`select only required fields FROM Case WHERE CreatedDate >= 2020-01-01T00:00:00Z and PO2_CASE_Organization__c like '%ECO%'`)
3. Write output to file in local PC
4. Put the file to bucket cs-ew1-prj-data-dm-marketing-dev-staging/Case/
5. Delete the output file in local PC
6. Write log on the main flow job

From DM to Operational Dashboard

Talend Project IND_DASHBOARD

Talend job F010_FACT_customer_complaints (J010_FACT_customer_complaints)



1. Check execution time format
2. Select data from DataOcean.V_sfc_case sfc and DataOcean.V_BWH_no_of_delivery_orders ([script](#))

- a. Create a template internal table (period x gbu x plant x serivity)
 - i. calendar to have period in format YYYY-MM back to n month, which control by variable I_VAR_IND_DASH_customer_complaint_t_nb_month_to_DM
 - ii. gbu_mapping to have text and key, which get from variable I_VAR_IND_DASH_customer_complaint_gbu_mapping = internal table for gbu, add more in the variable if we have new gbu
 - iii. severity_mapping to have list of severity value = not specified, Low, High, Major, Medium, Anomaly, Critical and Standard which get from variable I_VAR_IND_DASH_customer_complaint_severity_mapping
 - iv. site_cal to have all plant and gbu from V_sfc_case sfc cross join with severity_mapping and calendar to have template table that have all period, gbu, plant and serivity
- b. Create an aggregate internal table of customer complaint (cc_aggr) from V_sfc_case sfc
- c. Create an aggregate internal table of delivery (del_aggr) from V_BWH_no_of_delivery_orders
- d. Select the final data by calculate rolling 12 months
3. Delete all data in DM.FACT_customer_complaints
4. Load data from point 2 to FACT_customer_complaints
5. In case of error, email inform DataOps (I_VAR_IND_DASH_email)
6. Update log

Access rights

It is required to access BW query via Xtract.

and access to Salesforce

Source

BW

BW query = QVBW_QRY_MVSDSO57_0002

Xtract job = TALEND_[ENV]_QVBW_QRY_MVSDSO57_0002

SFC (<https://login.salesforce.com/services/Soap/u/52.0>)

Condition that select from SFC

```
FROM Case
WHERE CreatedDate >= 2020-01-01T00:00:00Z
and PO2_CASE_Organization__c='ECO'
```

Control variable to connect

- I_LOCAL_Domain_Marketing_SF_USER
- I_LOCAL_Domain_Marketing_SF_PASSWORD
- I_LOCAL_Domain_Marketing_SF_TOKEN
- I_LOCAL_Domain_Marketing_SF_endpoint

Format

columnar format

Destination

DataOcean

- Bucket = cs-ew1-prj-data-dm-marketing-dev-staging/
 - TALEND_DEV_QVBW_QRY_MVSDSO57_0002
 - Case
- DataOcean GCP = prj-data-dm-marketing-[env]
- STG Table names =
 - STG_BWH_0000_0000_F001_F_M_qvbw_qry_mvdsso57_0002
 - STG_SFC_0000_0000_F001_F_M_case
- ODS Table names =
 - ODS_BWH_0000_F001_F_M_qvbw_qry_mvdsso57_0002 V_ODS_BWH_0000_F001_F_M_qvbw_qry_mvdsso57_0002 VB_BWH_no_of_delivery_orders
 - ODS_SFC_0000_F001_F_M_case V_ODS_SFC_0000_F001_F_M_case VB_sfc_case
- DS_prj_industrial_dash
 - V_BWH_no_of_delivery_orders
 - V_sfc_case

Product

- GCP = prj-data-industrial-dash-[env]
- DataOcean
 - V_BWH_no_of_delivery_orders
 - DataOcean.V_sfc_case
- DM
 - FACT_customer_complaints
- DPL
 - V_FACT_customer_complaints
 - V_BWH_no_of_delivery_orders
 - V_sfc_case

Format

columnar format

Sizing

- STG_BWH_0000_0000_F001_F_M_qvbw_qry_mvdsso57_0002
maximum 6 months. around 52,200 records / month, Xtract job spend around 2 min / month
- STG_SFC_0000_0000_F001_F_M_case around 18,576 records

Assessment

Data same as BW query and SFC

Loading

1.1 Incremental Load

N/A

1.2 Full load

Delivery Source to ODS : F001_QVBW_QRY_MVSDSO57_0002_to_ODS (WS_DATA_OCEAN_DOMAIN_MARKETING)

It is kind of mix between full and incremental since we will load last 3 months on daily and full 2 years load for the full load on weekend, which control by

I_VAR_XTRACT_PARA_TALEND_QVBW_QRY_MVSDSO57_0002 = last3month, it will load last 3 months including current month on 0CALMONTH

I_VAR_FULL_LOAD_WEEKDAY_QVBW_QRY_MVSDSO57_0002 = the day of the week (1 = Monday , 7 = Sunday) to do the full load. If we need to do the full load on Sat, this parameter should be 6.

Case (number of complaint) Source to ODS : F020_SFC_case_customer_complaint_to_ODS(WS_DATA_OCEAN_DOMAIN_MARKETING)

1.3. Reloading data

Delivery

- I_VAR_XTRACT_PARA_TALEND_QVBW_QRY_MVSDSO57_0002 = &YYYYMM_Start=202301&YYYYMM_End=202305. Maximum should not over 6 months otherwise Xtract will be out of memory and run F001_QVBW_QRY_MVSDSO57_0002_to_ODS (WS_DATA_OCEAN_DOMAIN_MARKETING)

Note: maximum reload should not over than 7 months, it can cause Xtract out of memory

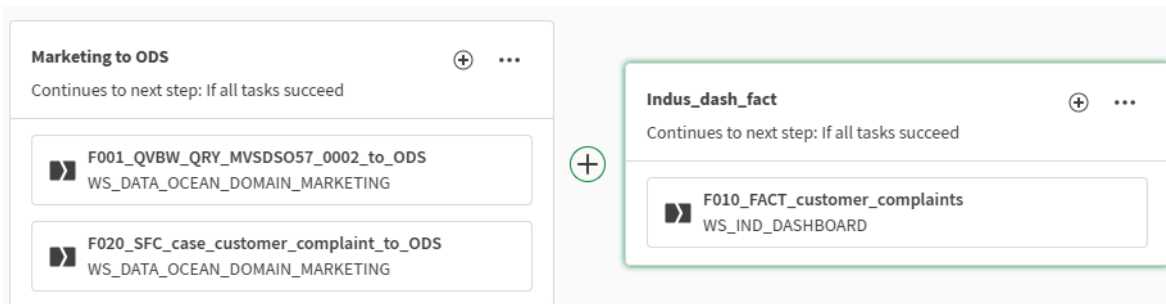
Case (number of complaint)

- Just run the F020_case_customer_complaint_to_ODS (WS_DATA_OCEAN_DOMAIN_MARKETING) again.

FACT_customer_complaints

- Just run the F010_FACT_customer_complaints (WS_IND_DASHBOARD) again.

1.4 Plan to schedule



It is scheduled by plans below on WS_IND_DASH

- PL_IND_DASH_customer_complaint Monthly on 1st at 05:00 CET

This plan will load on both Marketing and Indus dashboard for customer complaints

1.5 Timing

Delivery 10 minutes from source to ODS (current month)

Case 3- 10 minutes from source to ODS (full)

Criticality

Low?

Logging

Check the loading records on prj-data-dm-marketing-[dev]

```
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_QVBW_QRY_MVSDSO57_0002_to_ODS_MAIN', 'F001_QVBW_QRY_MVSDSO57_0002_to_ODS', 'F020_SFC_case_customer_complaint_to_ODS')
and meta_start_date > DATE_SUB ( CURRENT_TIMESTAMP () , INTERVAL 1 DAY )
order by job.meta_start_date desc
```