

Performer Migration - Technical Documentation

Description

This data pipeline consists of two main Talend jobs that orchestrate the SAP data ingestion and the transformation of business dimensions and facts. It is part of a broader data ecosystem built on GCP BigQuery, following a layered architecture (STG > ODS > DM), and includes SCD2 dimensional modeling, procedural calculations, and KPI-specific file generation.

Tools

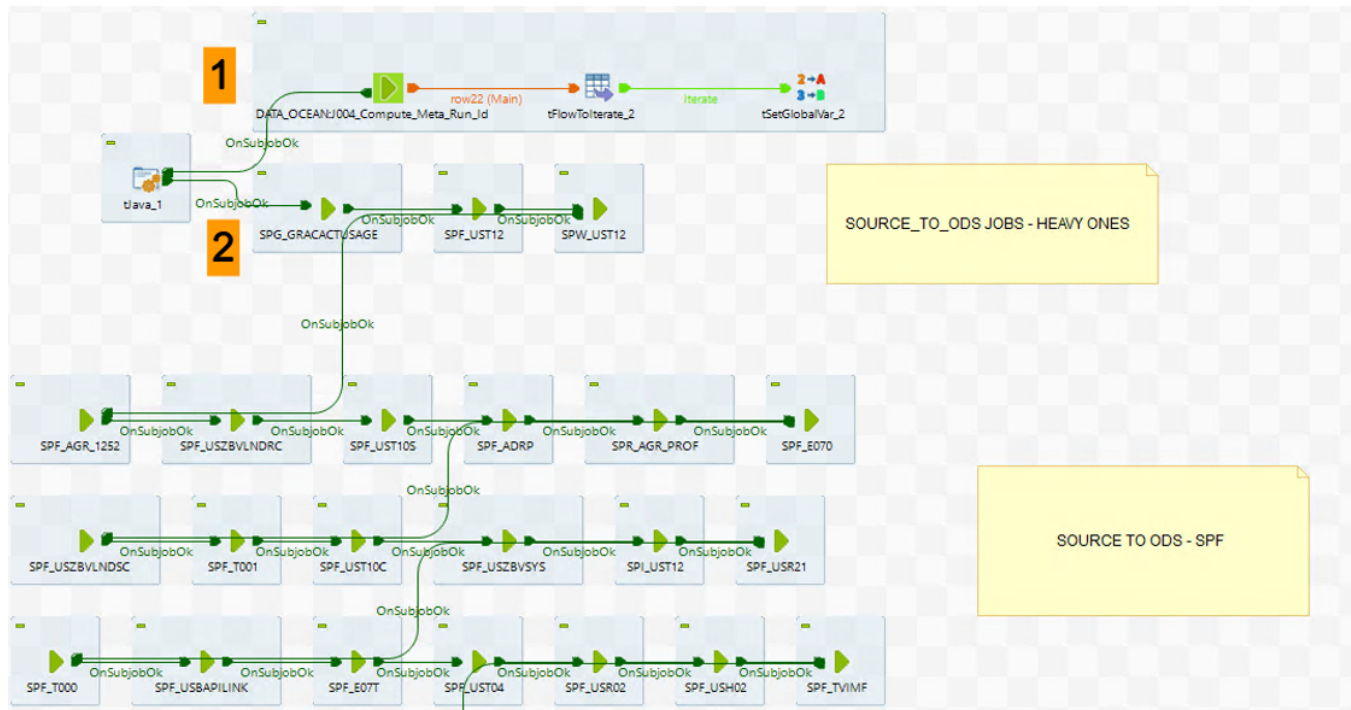
- **ETL Tool:** Talend
- **Environment:** GCP (BigQuery)

Jobs

1. 0010_PERFORMER_ORCHESTRATION_SAP_INGESTION — SAP Data Extraction

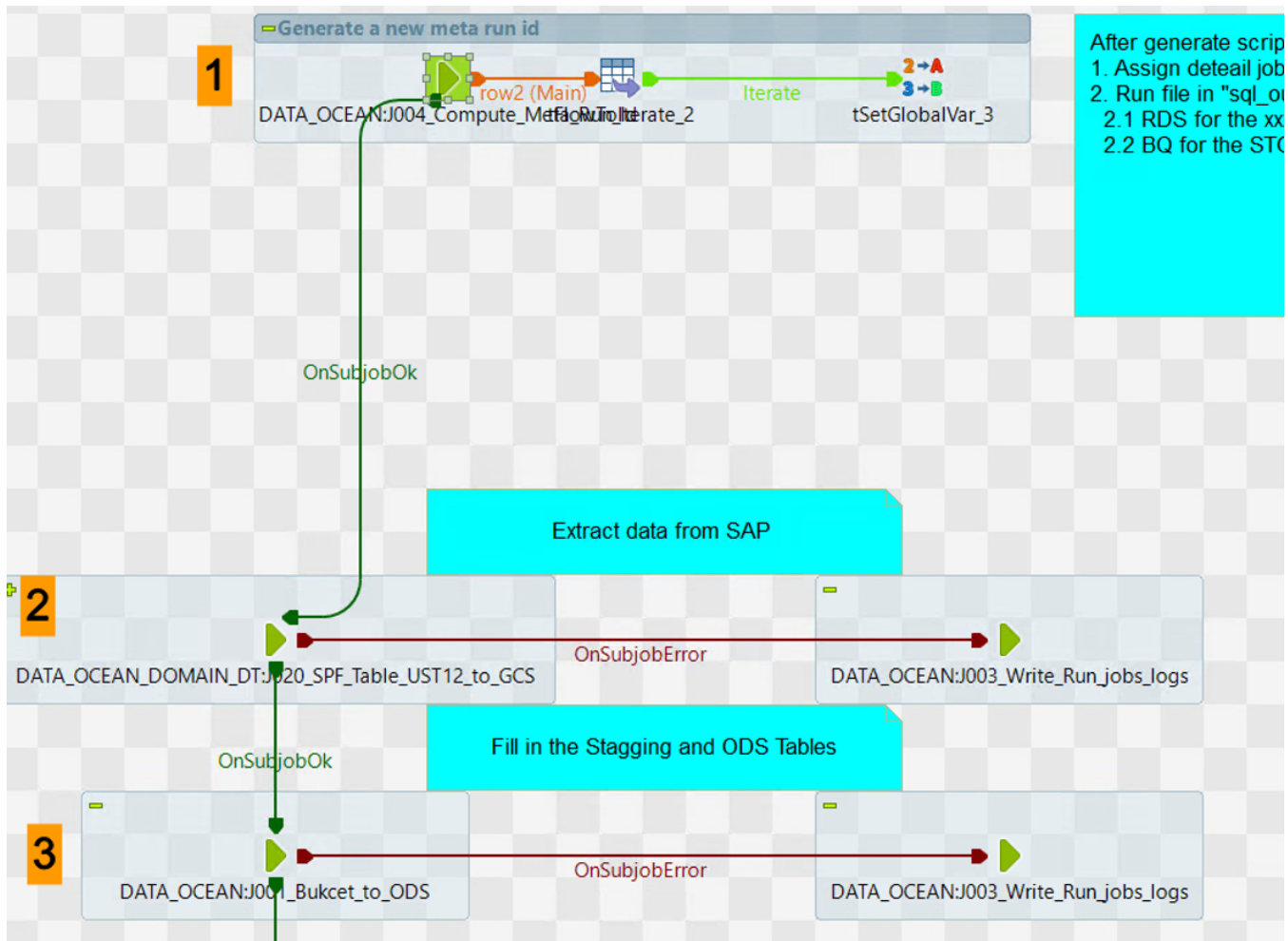
- Extracts raw data from SAP ECC tables for multiple systems (e.g., PF1_20, PI1_20, WP1_400).
- Each table in each system is managed by an individual Talend job ("1 job per table/system").
- The extracted data is ingested into the following layers:
 - STG: Staging
 - ODS: Operational Data Store
 - DM: Data Mart

Job Design:



1. Define variable and Generate meta_run_id
2. Call reference job to extract from SAP to GCP.

Job Design of Full Load table tRunJob:



1. Define variable and Generate meta_run_id
2. Call reference job to extract from SAP and move into Staging Bucket GCP **cs-ew1-prj-data-dm-dt-{env}-staging**
3. Move data from Staging bucket to STG dataset and then ODS.

Job Design of Incremental Load table tRunJob:

Execution Steps

1. Load variables from parameters database
2. Set statistics and execution metadata (meta_run_id, date for filenames)
3. Compare current DIM_KPI data with source data in CFG_KPI
4. Apply SCD Type 2 logic to detect and process changes
5. Catch and handle exceptions gracefully
6. Delete obsolete rows before performing upsert
7. Insert updated records into DIM_KPI
8. Call logging jobs to register execution
9. Send failure notifications by email (custom handling for line breaks and special characters)

Special Notes

- Step 9 diverges from the default job template due to issues handling line breaks and special characters — a customized component is used instead.
- Template Reference: [Solvay SCD2 Template](#)

Job Description: J012_BKP_FCT_ITGC_DETL_DM_to_GCS

This job implements a business process to export data from the fact table DM.FCT_ITGC_DETL to a Google Cloud Storage (GCS) folder. The export occurs only on the first day of each month, and it filters and extracts only data from the last day of the previous month.

Technology

- ETL Tool: Talend
- Platform: GCP (Google Cloud Platform)
- Project: prj-data-performer-dash-dev
- Source Table: DM.FCT_ITGC_DETL
- Destination: GCS Bucket: cs-ew1-prj-data-performer-dash-dev-staging/ITGC
- Format: Columnar (CSV file output)

Data Flow

DM.FCT_ITGC_DETL Local CSV File GCS (ITGC folder)

Execution Steps

1. Set variables
2. Check if the current date is the 1st day of the month
 - **IF yes:** Proceed with data extraction and export
 - **IF no:** Exit gracefully with a warning log
1. Load data from BigQuery to local file
2. Upload local file to GCS bucket (/ITGC)
3. Delete local file after upload
4. Log execution metadata and status

Special Notes

- This job uses the following template:
[Solvay Template – J012_MSSQL_TO_GCS](#)
- The file saved in GCS is structured under:
[gs://cs-ew1-prj-data-performer-dash-dev-staging/ITGC/<file_name>.csv](#)

Access Rights

- **GCP Project:** prj-data-performer-dash-dev
- **GCS Bucket:** cs-ew1-prj-data-performer-dash-dev-staging

Format

- **Output Format:** CSV
- **Storage:** Columnar format in BigQuery
- **Approximate File Size:** 1.5 MB

Assessment

- Data exported must match exactly the contents of the DM.FCT_ITGC_DETL table (rows and values).

Load Strategy

Task	Description
1.1 Full Load	Exports last day of previous month
1.3 Reloading	Not applicable
1.4 Schedule Plan	Daily , but logic restricts run to 1st day of the month only
1.5 Execution Time	~1 minute
Criticality	Low (to be confirmed)

Data Flow

SAP Systems/GSheets Talend ETL (Job per Table/System) STG (prj-data-dm-dt-{env}) ODS (prj-data-dm-dt-{env})

DM (prj-data-performer-dash-{env}) Procedures Facts KPI Output Files

Procedures Flows TO BE MADE

Layers & Tables

STG Tables

Project: prj-data-dm-dt-{env}

STG_SPF_0000_0000_F001_F_D_ADRP
STG_SPF_0000_0000_F001_F_D_AGR_1252
STG_SPF_0000_0000_F001_F_D_AGR_AGRS
STG_SPF_0000_0000_F001_F_D_AGR_DEFINE
STG_SPF_0000_0000_F001_F_D_AGR_PROF
STG_SPF_0000_0000_F001_F_D_AGR_TCODES
STG_SPF_0000_0000_F001_F_D_AGR_USERS
STG_SPF_0000_0000_F001_F_D_E070
STG_SPF_0000_0000_F001_F_D_E07T
STG_SPF_0000_0000_F001_F_D_T000
STG_SPF_0000_0000_F001_F_D_T001
STG_SPF_0000_0000_F001_F_D_TVIMF
STG_SPF_0000_0000_F001_F_D_USBAPILINK
STG_SPF_0000_0000_F001_F_D_USH02
STG_SPF_0000_0000_F001_F_D_USR02
STG_SPF_0000_0000_F001_F_D_USR21
STG_SPF_0000_0000_F001_F_D_UST04
STG_SPF_0000_0000_F001_F_D_UST10C
STG_SPF_0000_0000_F001_F_D_UST10S
STG_SPF_0000_0000_F001_F_D_UST12
STG_SPF_0000_0000_F001_F_D_USZBVLNDR

STG_SPF_0000_0000_F001_F_D_USZBVLNDSC
STG_SPF_0000_0000_F001_F_D_USZBVSYS
STG_SPF_0000_0000_F003_I_D_usr02_pf1
STG_SPG_0000_0000_F001_F_D_GRACACCESSRULE
STG_SPG_0000_0000_F001_F_D_GRACACTRULE
STG_SPG_0000_0000_F001_F_M_GRACROLEPRMVL
STG_SPG_0000_0000_F001_I_M_GRACACTUSAGE
STG_SPI_0000_0000_F001_F_D_ADRP
STG_SPI_0000_0000_F001_F_D_AGR_1252
STG_SPI_0000_0000_F001_F_D_AGR_AGRS
STG_SPI_0000_0000_F001_F_D_AGR_DEFINE
STG_SPI_0000_0000_F001_F_D_AGR_PROF
STG_SPI_0000_0000_F001_F_D_AGR_TCODES
STG_SPI_0000_0000_F001_F_D_AGR_USERS
STG_SPI_0000_0000_F001_F_D_E070
STG_SPI_0000_0000_F001_F_D_E07T
STG_SPI_0000_0000_F001_F_D_T000
STG_SPI_0000_0000_F001_F_D_T001
STG_SPI_0000_0000_F001_F_D_TVIMF
STG_SPI_0000_0000_F001_F_D_USBAPILINK
STG_SPI_0000_0000_F001_F_D_USH02
STG_SPI_0000_0000_F001_F_D_USR02
STG_SPI_0000_0000_F001_F_D_USR21
STG_SPI_0000_0000_F001_F_D_UST04
STG_SPI_0000_0000_F001_F_D_UST10C
STG_SPI_0000_0000_F001_F_D_UST10S
STG_SPI_0000_0000_F001_F_D_UST12
STG_SPI_0000_0000_F001_F_D_USZBVLNDRC
STG_SPI_0000_0000_F001_F_D_USZBVLNDSC
STG_SPI_0000_0000_F001_F_D_USZBVSYS
STG_SPI_0000_0000_F001_I_D_usr02_pi1
STG_SPW_0000_0000_F001_F_D_ADRP
STG_SPW_0000_0000_F001_F_D_AGR_1252
STG_SPW_0000_0000_F001_F_D_AGR_AGRS
STG_SPW_0000_0000_F001_F_D_AGR_DEFINE
STG_SPW_0000_0000_F001_F_D_AGR_PROF
STG_SPW_0000_0000_F001_F_D_AGR_TCODES
STG_SPW_0000_0000_F001_F_D_AGR_USERS
STG_SPW_0000_0000_F001_F_D_E070
STG_SPW_0000_0000_F001_F_D_E07T
STG_SPW_0000_0000_F001_F_D_T000

STG_SPW_0000_0000_F001_F_D_T001
STG_SPW_0000_0000_F001_F_D_TVIMF
STG_SPW_0000_0000_F001_F_D_USBAPILINK
STG_SPW_0000_0000_F001_F_D_USH02
STG_SPW_0000_0000_F001_F_D_USR02
STG_SPW_0000_0000_F001_F_D_USR21
STG_SPW_0000_0000_F001_F_D_UST04
STG_SPW_0000_0000_F001_F_D_UST10C
STG_SPW_0000_0000_F001_F_D_UST10S
STG_SPW_0000_0000_F001_F_D_UST12
STG_SPW_0000_0000_F001_F_D_USZBVLNDRC
STG_SPW_0000_0000_F001_F_D_USZBVLNDSC
STG_SPW_0000_0000_F001_F_D_USZBVSYS
STG_SPW_0000_0000_F004_I_D_usr02_wp1

ODS Tables

Project: prj-data-dm-dt-{env}

ODS_SPF_0000_F001_F_D_ADRP
ODS_SPF_0000_F001_F_D_AGR_1252
ODS_SPF_0000_F001_F_D_AGR_AGRS
ODS_SPF_0000_F001_F_D_AGR_DEFINE
ODS_SPF_0000_F001_F_D_AGR_PROF
ODS_SPF_0000_F001_F_D_AGR_TCODES
ODS_SPF_0000_F001_F_D_AGR_USERS
ODS_SPF_0000_F001_F_D_E070
ODS_SPF_0000_F001_F_D_E07T
ODS_SPF_0000_F001_F_D_T000
ODS_SPF_0000_F001_F_D_T001
ODS_SPF_0000_F001_F_D_TVIMF
ODS_SPF_0000_F001_F_D_USBAPILINK
ODS_SPF_0000_F001_F_D_USH02
ODS_SPF_0000_F001_F_D_USR02
ODS_SPF_0000_F001_F_D_USR21
ODS_SPF_0000_F001_F_D_UST04
ODS_SPF_0000_F001_F_D_UST10C
ODS_SPF_0000_F001_F_D_UST10S
ODS_SPF_0000_F001_F_D_UST12
ODS_SPF_0000_F001_F_D_USZBVLNDRC
ODS_SPF_0000_F001_F_D_USZBVLNDSC
ODS_SPF_0000_F001_F_D_USZBVSYS

ODS_SPF_0000_F003_I_D_usr02_pf1
ODS_SPG_0000_F001_F_D_GRACACCESSRULE
ODS_SPG_0000_F001_F_D_GRACACTRULE
ODS_SPG_0000_F001_F_M_GRACROLEPRMVL
ODS_SPG_0000_F001_I_M_GRACACTUSAGE
ODS_SPI_0000_F001_F_D_ADRP
ODS_SPI_0000_F001_F_D_AGR_1252
ODS_SPI_0000_F001_F_D_AGR_AGRS
ODS_SPI_0000_F001_F_D_AGR_DEFINE
ODS_SPI_0000_F001_F_D_AGR_PROF
ODS_SPI_0000_F001_F_D_AGR_TCODES
ODS_SPI_0000_F001_F_D_AGR_USERS
ODS_SPI_0000_F001_F_D_E070
ODS_SPI_0000_F001_F_D_E07T
ODS_SPI_0000_F001_F_D_T000
ODS_SPI_0000_F001_F_D_T001
ODS_SPI_0000_F001_F_D_TVIMF
ODS_SPI_0000_F001_F_D_USBAPILINK
ODS_SPI_0000_F001_F_D_USH02
ODS_SPI_0000_F001_F_D_USR02
ODS_SPI_0000_F001_F_D_USR21
ODS_SPI_0000_F001_F_D_UST04
ODS_SPI_0000_F001_F_D_UST10C
ODS_SPI_0000_F001_F_D_UST10S
ODS_SPI_0000_F001_F_D_UST12
ODS_SPI_0000_F001_F_D_USZBVLNDRC
ODS_SPI_0000_F001_F_D_USZBVLNDSC
ODS_SPI_0000_F001_F_D_USZBVSYS
ODS_SPI_0000_F001_I_D_usr02_pi1
ODS_SPW_0000_F001_F_D_ADRP
ODS_SPW_0000_F001_F_D_AGR_1252
ODS_SPW_0000_F001_F_D_AGR_AGRS
ODS_SPW_0000_F001_F_D_AGR_DEFINE
ODS_SPW_0000_F001_F_D_AGR_PROF
ODS_SPW_0000_F001_F_D_AGR_TCODES
ODS_SPW_0000_F001_F_D_AGR_USERS
ODS_SPW_0000_F001_F_D_E070
ODS_SPW_0000_F001_F_D_E07T
ODS_SPW_0000_F001_F_D_T000
ODS_SPW_0000_F001_F_D_T001
ODS_SPW_0000_F001_F_D_TVIMF

ODS_SPW_0000_F001_F_D_USBAPILINK
ODS_SPW_0000_F001_F_D_USH02
ODS_SPW_0000_F001_F_D_USR02
ODS_SPW_0000_F001_F_D_USR21
ODS_SPW_0000_F001_F_D_UST04
ODS_SPW_0000_F001_F_D_UST10C
ODS_SPW_0000_F001_F_D_UST10S
ODS_SPW_0000_F001_F_D_UST12
ODS_SPW_0000_F001_F_D_USZBVLNDRC
ODS_SPW_0000_F001_F_D_USZBVLNDSC
ODS_SPW_0000_F001_F_D_USZBVSYS
ODS_SPW_0000_F004_I_D_usr02_wp1

DM Tables & Procedures

Project: prj-data-performer-dash-{env}

- **Fact Tables:**
 - a. FCT_ITGC_AGG
 - b. FCT_ITGC_DETL
 - c. FCT_ITGC_MNTH
- **Procedures:**
 - 1. RT_DM_EXECUTE_FCT_ITCG_AGG
 - 2. RT_DM_EXECUTE_FCT_ITCG_DETL
 - 3. RT_DM_EXECUTE_FCT_ITCG_MNTH
- **Dimension Table:** DIM_KPI

Views

ODS Views (project: prj-data-performer-dash-{env}) **Dataset:** ODS_DataOcean

- V_ODS_ADRP
- V_ODS_ADRP
- V_ODS_AGR_1252
- V_ODS_AGR_AGRS
- V_ODS_AGR_DEFINE
- V_ODS_AGR_PROF
- V_ODS_AGR_TCODES
- V_ODS_AGR_USERS
- V_ODS_E070
- V_ODS_E07T
- V_ODS_GRACACCESSRULE
- V_ODS_GRACACTRULE
- V_ODS_GRACACTUSAGE
- V_ODS_GRACMGMTACTUSAGE
- V_ODS_GRACROLEPRMVL
- V_ODS_GRACSODRISK
- V_ODS_GRACSODRISKFUNC
- V_ODS_GRACUSERPRMVL
- V_ODS_GRACXTOBJ
- V_ODS_T000
- V_ODS_T001
- V_ODS_TVIMF
- V_ODS_USBAPILINK
- V_ODS_USH02
- V_ODS_USR02
- V_ODS_USR21
- V_ODS_UST04
- V_ODS_UST10C
- V_ODS_UST10S

- V_ODS_UST12
- V_ODS_USZBVLNDRRC
- V_ODS_USZBVLNDSC
- V_ODS_USZBVSYS

DM Views

- V_DATE
- V_SOURCE_SYSTEM

WDL Temporary Tables

- temp_FinalResult
- temp_finalResult_agg

Schedule Plan

- **Name:** PL_PER_ORCHESTRATION
- **Frequency:** Daily at 02:00 AM UTC
- **Estimated Execution Time:** 4 to 5 hours
- **Criticality:** High

Sizing (Sample Record Counts)

Table Name	Record Count
STG_SPF_0000_0000_F001_F_D_ADRP	601571
STG_SPF_0000_0000_F001_F_D_AGR_1252	2432293
STG_SPF_0000_0000_F001_F_D_AGR_AGRS	8445
STG_SPF_0000_0000_F001_F_D_AGR_DEFINE	31397
STG_SPF_0000_0000_F001_F_D_AGR_PROF	55295
STG_SPF_0000_0000_F001_F_D_AGR_TCODES	312178
STG_SPF_0000_0000_F001_F_D_AGR_USERS	735459
STG_SPF_0000_0000_F001_F_D_E070	228368
STG_SPF_0000_0000_F001_F_D_E07T	228225
STG_SPF_0000_0000_F001_F_D_T000	8
STG_SPF_0000_0000_F001_F_D_T001	1625
STG_SPF_0000_0000_F001_F_D_TVIMF	16823
STG_SPF_0000_0000_F001_F_D_USBAPILINK	0
STG_SPF_0000_0000_F001_F_D_USH02	981553
STG_SPF_0000_0000_F001_F_D_USR02	17638
STG_SPF_0000_0000_F001_F_D_USR21	17639
STG_SPF_0000_0000_F001_F_D_UST04	749400
STG_SPF_0000_0000_F001_F_D_UST10C	1475
STG_SPF_0000_0000_F001_F_D_UST10S	1067999
STG_SPF_0000_0000_F001_F_D_UST12	16521722
STG_SPF_0000_0000_F001_F_D_USZBVLNDRRC	10
STG_SPF_0000_0000_F001_F_D_USZBVLNDSC	1

STG_SPF_0000_0000_F001_F_D_USZBVSYS	19557
STG_SPF_0000_0000_F003_I_D_usr02_pf1	4656
STG_SPG_0000_0000_F001_F_D_GRACACCESSRULE	0
STG_SPG_0000_0000_F001_F_D_GRACACTRULE	152835
STG_SPG_0000_0000_F001_F_M_GRACROLEPRMVL	43993886
STG_SPG_0000_0000_F001_I_M_GRACACTUSAGE	212014
STG_SPI_0000_0000_F001_F_D_ADRP	235621
STG_SPI_0000_0000_F001_F_D_AGR_1252	9223
STG_SPI_0000_0000_F001_F_D_AGR_AGRS	1261
STG_SPI_0000_0000_F001_F_D_AGR_DEFINE	4482
STG_SPI_0000_0000_F001_F_D_AGR_PROF	1760
STG_SPI_0000_0000_F001_F_D_AGR_TCODES	93198
STG_SPI_0000_0000_F001_F_D_AGR_USERS	24051
STG_SPI_0000_0000_F001_F_D_E070	69826
STG_SPI_0000_0000_F001_F_D_E07T	69793
STG_SPI_0000_0000_F001_F_D_T000	4
STG_SPI_0000_0000_F001_F_D_T001	1678
STG_SPI_0000_0000_F001_F_D_TVIMF	13064
STG_SPI_0000_0000_F001_F_D_USBAPILINK	0
STG_SPI_0000_0000_F001_F_D_USH02	139320
STG_SPI_0000_0000_F001_F_D_USR02	2937
STG_SPI_0000_0000_F001_F_D_USR21	2953
STG_SPI_0000_0000_F001_F_D_UST04	27142
STG_SPI_0000_0000_F001_F_D_UST10C	1471
STG_SPI_0000_0000_F001_F_D_UST10S	44641
STG_SPI_0000_0000_F001_F_D_UST12	248202
STG_SPI_0000_0000_F001_F_D_USZBVLNDRC	0
STG_SPI_0000_0000_F001_F_D_USZBVLNDSC	0
STG_SPI_0000_0000_F001_F_D_USZBVSYS	0
STG_SPI_0000_0000_F001_I_D_usr02_pi1	469
STG_SPW_0000_0000_F001_F_D_ADRP	553560
STG_SPW_0000_0000_F001_F_D_AGR_1252	7214947
STG_SPW_0000_0000_F001_F_D_AGR_AGRS	178122
STG_SPW_0000_0000_F001_F_D_AGR_DEFINE	0
STG_SPW_0000_0000_F001_F_D_AGR_PROF	68235
STG_SPW_0000_0000_F001_F_D_AGR_TCODES	2727611
STG_SPW_0000_0000_F001_F_D_AGR_USERS	1694985
STG_SPW_0000_0000_F001_F_D_E070	168658
STG_SPW_0000_0000_F001_F_D_E07T	168647
STG_SPW_0000_0000_F001_F_D_T000	4
STG_SPW_0000_0000_F001_F_D_T001	234

STG_SPW_0000_0000_F001_F_D_TVIMF	17243
STG_SPW_0000_0000_F001_F_D_USBAPILINK	0
STG_SPW_0000_0000_F001_F_D_USH02	1938976
STG_SPW_0000_0000_F001_F_D_USR02	18326
STG_SPW_0000_0000_F001_F_D_USR21	18326
STG_SPW_0000_0000_F001_F_D_UST04	1550910
STG_SPW_0000_0000_F001_F_D_UST10C	1364
STG_SPW_0000_0000_F001_F_D_UST10S	1052568
STG_SPW_0000_0000_F001_F_D_UST12	17548032
STG_SPW_0000_0000_F001_F_D_USZBVLNDRC	0
STG_SPW_0000_0000_F001_F_D_USZBVLNDSC	0
STG_SPW_0000_0000_F001_F_D_USZBVSYS	0
STG_SPW_0000_0000_F004_I_D_usr02_wp1	5738

Access Rights

- prj-data-dm-dt-{env}
- prj-data-performer-dash-{env}
- prj-data-dm-common-{env}

Reload Instructions

- To reload data, simply re-run the corresponding job (O010_ORCHESTRATION_SAP_INGESTION or O010_ORCHESTRATION_DIM_PROC).

Logging Query

```
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 ('O010_PERFORMER_ORCHESTRATION_SAP_INGESTION', 'O010_PERFORMER_ORCHESTRATION_DIM_PROC')
order by job.meta_start_date desc;
```