IRM Talend Jobs

Below are the list of Talend jobs which were developed as part of IRM .

Talend Project : DATA_OCEAN_DOMAIN_PROCUREMENT

GCP Project : prj-data-dm-procurement-prod

Schedule : All this jobs are scheduled to run weekly (yet the moment its configured to run on Monday's , but can be changed based on the need)

Job Explanation: Available at the bottom of the page.

Source SAP System	Source Table	Load Type	Plan name	Task Name	STG Table	ODS Table
PF1	EINA	Full	PL_SPF_F001_F_W_EINA_TO_BQ	F001_SPF_F001_F_W_EINA_TO_BQ	STG_SPF_0000_0000_F001_F_W_EINA	ODS_SPF_0000_F001_F_W_EINA
	EINE	Full	PL_SPF_F001_F_W_EINE_TO_BQ	F001_SPF_F001_F_W_EINE_TO_BQ	STG_SPF_0000_0000_F001_F_W_EINE	ODS_SPF_0000_F001_F_W_EINE
	EKKO	Incremental	PL_SPF_F001_I_W_EKKO_TO_BQ	F001_SPF_F001_I_W_EKKO_TO_BQ	STG_SPF_0000_0000_F001_I_W_EKKO	ODS_SPF_0000_F001_I_W_EKKO
	ISEG	Incremental	PL_SPF_F001_I_W_ISEG_TO_BQ	F001_SPF_F001_I_W_ISEG_TO_BQ	STG_SPF_0000_0000_F001_I_W_ISEG	ODS_SPF_0000_F001_I_W_ISEG
	LFB1	Full	PL_SPF_F001_F_W_LFB1_TO_BQ	F001_SPF_F001_F_W_LFB1_TO_BQ	STG_SPF_0000_0000_F001_F_W_LFB1	ODS_SPF_0000_F001_F_W_LFB1
	LFM1	Full	PL_SPF_F001_F_W_LFM1_TO_BQ	F001_SPF_F001_F_W_LFM1_TO_BQ	STG_SPF_0000_0000_F001_F_W_LFM1	ODS_SPF_0000_F001_F_W_LFM1
	LFM2	Full	PL_SPF_F001_F_W_LFM2_TO_BQ	F001_SPF_F001_F_W_LFM2_TO_BQ	STG_SPF_0000_0000_F001_F_W_LFM2	ODS_SPF_0000_F001_F_W_LFM2
	MARA	Full	PL_SPF_F001_F_W_MARA_TO_BQ	F001_SPF_F001_F_W_MARA_TO_BQ	STG_SPF_0000_0000_F001_F_W_MARA	ODS_SPF_0000_F001_F_W_MARA
	MARC	Full	PL_SPF_F001_F_W_MARC_TO_BQ	F001_SPF_F001_F_W_MARC_TO_BQ	STG_SPF_0000_0000_F001_F_W_MARC	ODS_SPF_0000_F001_F_W_MARC
	MARD	Full	PL_SPF_F001_F_W_MARD_TO_BQ	F001_SPF_F001_F_W_MARD_TO_BQ	STG_SPF_0000_0000_F001_F_W_MARD	ODS_SPF_0000_F001_F_W_MARD
	MAST	Full	PL_SPF_F001_F_W_MAST_TO_BQ	F001_SPF_F001_F_W_MAST_TO_BQ	STG_SPF_0000_0000_F001_F_W_MAST	ODS_SPF_0000_F001_F_W_MAST
	MCHB	Incremental	PL_SPF_F001_I_W_MCHB_TO_BQ	F001_SPF_F001_I_W_MCHB_TO_BQ	STG_SPF_0000_0000_F001_I_W_MCHB	ODS_SPF_0000_F001_I_W_MCHB
	MKAL	Full	PL_SPF_F001_F_W_MKAL_TO_BQ	F001_SPF_F001_F_W_MKAL_TO_BQ	STG_SPF_0000_0000_F001_F_W_MKAL	ODS_SPF_0000_F001_F_W_MKAL
	MKPF	Incremental	PL_SPF_F001_I_W_MKPF_TO_BQ	F001_SPF_F001_I_W_MKPF_TO_BQ	STG_SPF_0000_0000_F001_I_W_MKPF	ODS_SPF_0000_F001_I_W_MKPF
	MSEG	Incremental	PL_SPF_F001_I_W_MSEG_TO_BQ	F001_SPF_F001_I_W_MSEG_TO_BQ	STG_SPF_0000_0000_F001_I_W_MSEG	ODS_SPF_0000_F001_I_W_MSEG
	STPO	Full	PL_SPF_F001_F_W_STPO_TO_BQ	F001_SPF_F001_F_W_STPO_TO_BQ	STG_SPF_0000_0000_F001_F_W_STPO	ODS_SPF_0000_F001_F_W_STPO
	T024	Full	PL_SPF_F001_F_W_T024_TO_BQ	F001_SPF_F001_F_W_T024_TO_BQ	STG_SPF_0000_0000_F001_F_W_T024	ODS_SPF_0000_F001_F_W_T024
	ASMD	Full	PL_SPF_F001_F_W_ASMD_TO_BQ	F001_SPF_F001_F_W_ASMD_TO_BQ	STG_SPF_0000_0000_F001_F_W_ASMD	ODS_SPF_0000_F001_F_W_ASMD
	EBAN	Incremental	PL_SPF_F001_I_W_EBAN_TO_BQ	F001_SPF_F001_I_W_EBAN_TO_BQ	STG_SPF_0000_0000_F001_I_W_EBAN	ODS_SPF_0000_F001_I_W_EBAN
	EKPO	Incremental	PL_SPF_F001_I_W_EKPO_TO_BQ	F001_SPF_F001_I_W_EKPO_TO_BQ	STG_SPF_0000_0000_F001_I_W_EKPO	ODS_SPF_0000_F001_I_W_EKPO
	ESLH	Full	PL_SPF_F001_F_W_ESLH_TO_BQ	F001_SPF_F001_F_W_ESLH_TO_BQ	STG_SPF_0000_0000_F001_F_W_ESLH	ODS_SPF_0000_F001_F_W_ESLH
	ESLL *	Full	PL_SPF_F001_F_W_ESLL_TO_BQ	F001_SPF_F001_F_W_ESLL_TO_BQ	STG_SPF_0000_0000_F001_F_W_ESLL	ODS_SPF_0000_F001_F_W_ESLL
	ESSR	Incremental	PL_SPF_F001_I_W_ESSR_TO_BQ	F001_SPF_F001_I_W_ESSR_TO_BQ	STG_SPF_0000_0000_F001_I_W_ESSR	ODS_SPF_0000_F001_I_W_ESSR
	MBEW	Full	PL_SPF_F001_F_W_MBEW_TO_BQ	F001_SPF_F001_F_W_MBEW_TO_BQ	STG_SPF_0000_0000_F001_F_W_MBEW	ODS_SPF_0000_F001_F_W_MBEW

	T024W	Full	PL_SPF_F001_F_W_T024W_TO_BQ	F001_SPF_F001_F_W_T024W_TO_BQ	STG_SPF_0000_0000_F001_F_W_T024W	ODS_SPF_0000_F001_F_W_T024W
	MAKT	Full	PL_SPF_F001_F_W_MAKT_TO_BQ	F001_SPF_F001_F_W_MAKT_TO_BQ	STG_SPF_0000_0000_F001_F_W_MAKT	ODS_SPF_0000_F001_F_W_MAKT
	T156T	Full	PL_SPF_F001_F_W_T156T_TO_BQ	F001_SPF_F001_F_W_T156T_TO_BQ	STG_SPF_0000_0000_F001_F_W_T156T	ODS_SPF_0000_F001_F_W_T156T
	T001W	Full	PL_SPF_F001_F_W_T001W_TO_BQ	F001_SPF_F001_F_W_T001W_TO_BQ	STG_SPF_0000_0000_F001_F_W_T001W	ODS_SPF_0000_F001_F_W_T001W
	T001L	Full	PL_SPF_F001_F_W_T001L_TO_BQ	F001_SPF_F001_F_W_T001L_TO_BQ	STG_SPF_0000_0000_F001_F_W_T001L	ODS_SPF_0000_F001_F_W_T001L
	T001	Full	PL_SPF_F001_F_W_T001_TO_BQ	F001_SPF_F001_F_W_T001_TO_BQ	STG_SPF_0000_0000_F001_F_W_T001	ODS_SPF_0000_F001_F_W_T001
	T064T	Full	PL_SPF_F001_F_W_T064T_TO_BQ	F001_SPF_F001_F_W_T064T_TO_BQ	STG_SPF_0000_0000_F001_F_W_T064T	ODS_SPF_0000_F001_F_W_T064T

Source SAP System	Source Table	Load Type	Plan name	Task Name	STG Table	ODS Table
WP1	EINA	Full	PL_SPW_F001_F_W_EINA_TO_BQ	F001_SPW_F001_F_W_EINA_TO_BQ	STG_SPW_0000_0000_F001_F_W_EINA	ODS_SPW_0000_F001_F_W_EINA
	EINE	Full	PL_SPW_F001_F_W_EINE_TO_BQ	F001_SPW_F001_F_W_EINE_TO_BQ	STG_SPW_0000_0000_F001_F_W_EINE	ODS_SPW_0000_F001_F_W_EINE
	EKKO	Incremental	PL_SPW_F001_I_W_EKKO_TO_BQ	F001_SPW_F001_I_W_EKKO_TO_BQ	STG_SPW_0000_0000_F001_I_W_EKKO	ODS_SPW_0000_F001_I_W_EKKO
	ISEG	Incremental	PL_SPW_F001_I_W_ISEG_TO_BQ	F001_SPW_F001_I_W_ISEG_TO_BQ	STG_SPW_0000_0000_F001_I_W_ISEG	ODS_SPW_0000_F001_I_W_ISEG
	LFB1	Full	PL_SPW_F001_F_W_LFB1_TO_BQ	F001_SPW_F001_F_W_LFB1_TO_BQ	STG_SPW_0000_0000_F001_F_W_LFB1	ODS_SPW_0000_F001_F_W_LFB1
	LFM1	Full	PL_SPW_F001_F_W_LFM1_TO_BQ	F001_SPW_F001_F_W_LFM1_TO_BQ	STG_SPW_0000_0000_F001_F_W_LFM1	ODS_SPW_0000_F001_F_W_LFM1
	LFM2	Full	PL_SPW_F001_F_W_LFM2_TO_BQ	F001_SPW_F001_F_W_LFM2_TO_BQ	STG_SPW_0000_0000_F001_F_W_LFM2	ODS_SPW_0000_F001_F_W_LFM2
	MARA	Full	PL_SPW_F001_F_W_MARA_TO_BQ	F001_SPW_F001_F_W_MARA_TO_BQ	STG_SPW_0000_0000_F001_F_W_MARA	ODS_SPW_0000_F001_F_W_MARA
	MARC	Full	PL_SPW_F001_F_W_MARC_TO_BQ	F001_SPW_F001_F_W_MARC_TO_BQ	STG_SPW_0000_0000_F001_F_W_MARC	ODS_SPW_0000_F001_F_W_MARC
	MARD	Full	PL_SPW_F001_F_W_MARD_TO_BQ	F001_SPW_F001_F_W_MARD_TO_BQ	STG_SPW_0000_0000_F001_F_W_MARD	ODS_SPW_0000_F001_F_W_MARD
	MAST	Full	PL_SPW_F001_F_W_MAST_TO_BQ	F001_SPW_F001_F_W_MAST_TO_BQ	STG_SPW_0000_0000_F001_F_W_MAST	ODS_SPW_0000_F001_F_W_MAST
	MCHB	Incremental	PL_SPW_F001_I_W_MCHB_TO_BQ	F001_SPW_F001_I_W_MCHB_TO_BQ	STG_SPW_0000_0000_F001_I_W_MCHB	ODS_SPW_0000_F001_I_W_MCHB
	MKAL	Full	PL_SPW_F001_F_W_MKAL_TO_BQ	F001_SPW_F001_F_W_MKAL_TO_BQ	STG_SPW_0000_0000_F001_F_W_MKAL	ODS_SPW_0000_F001_F_W_MKAL
	MKPF	Incremental	PL_SPW_F001_I_W_MKPF_TO_BQ	F001_SPW_F001_I_W_MKPF_TO_BQ	STG_SPW_0000_0000_F001_I_W_MKPF	ODS_SPW_0000_F001_I_W_MKPF
	MSEG	Incremental	PL_SPW_F001_I_W_MSEG_TO_BQ	F001_SPW_F001_I_W_MSEG_TO_BQ	STG_SPW_0000_0000_F001_I_W_MSEG	ODS_SPW_0000_F001_I_W_MSEG
	STPO	Full	PL_SPW_F001_F_W_STPO_TO_BQ	F001_SPW_F001_F_W_STPO_TO_BQ	STG_SPW_0000_0000_F001_F_W_STPO	ODS_SPW_0000_F001_F_W_STPO
	T024	Full	PL_SPW_F001_F_W_T024_TO_BQ	F001_SPW_F001_F_W_T024_TO_BQ	STG_SPW_0000_0000_F001_F_W_T024	ODS_SPW_0000_F001_F_W_T024
	ASMD	Full	PL_SPW_F001_F_W_ASMD_TO_BQ	F001_SPW_F001_F_W_ASMD_TO_BQ	STG_SPW_0000_0000_F001_F_W_ASMD	ODS_SPW_0000_F001_F_W_ASMD
	EBAN	Incremental	PL_SPW_F001_I_W_EBAN_TO_BQ	F001_SPW_F001_I_W_EBAN_TO_BQ	STG_SPW_0000_0000_F001_I_W_EBAN	ODS_SPW_0000_F001_I_W_EBAN
	EKPO	Incremental	PL_SPW_F001_I_W_EKPO_TO_BQ	F001_SPW_F001_I_W_EKPO_TO_BQ	STG_SPW_0000_0000_F001_I_W_EKPO	ODS_SPW_0000_F001_I_W_EKPO
	ESLH	Full	PL_SPW_F001_F_W_ESLH_TO_BQ	F001_SPW_F001_F_W_ESLH_TO_BQ	STG_SPW_0000_0000_F001_F_W_ESLH	ODS_SPW_0000_F001_F_W_ESLH

ESLL	Full	PL_SPW_F001_F_W_ESLL_TO_BQ	F001_SPW_F001_F_W_ESLL_TO_BQ	STG_SPW_0000_0000_F001_F_W_ESLL	ODS_SPW_0000_F001_F_W_ESLL
ESSR	Incremental	PL_SPW_F001_I_W_ESSR_TO_BQ	F001_SPW_F001_I_W_ESSR_TO_BQ	STG_SPW_0000_0000_F001_I_W_ESSR	ODS_SPW_0000_F001_I_W_ESSR
MBEW	Full	PL_SPW_F001_F_W_MBEW_TO_BQ	F001_SPW_F001_F_W_MBEW_TO_BQ	STG_SPW_0000_0000_F001_F_W_MBEW	ODS_SPW_0000_F001_F_W_MBEW
T024W	Full	PL_SPW_F001_F_W_T024W_TO_BQ	F001_SPW_F001_F_W_T024W_TO_BQ	STG_SPW_0000_0000_F001_F_W_T024W	ODS_SPW_0000_F001_F_W_T024W
MAKT	Full	PL_SPW_F001_F_W_MAKT_TO_BQ	F001_SPW_F001_F_W_MAKT_TO_BQ	STG_SPW_0000_0000_F001_F_W_MAKT	ODS_SPW_0000_F001_F_W_MAKT
T156T	Full	PL_SPW_F001_F_W_T156T_TO_BQ	F001_SPW_F001_F_W_T156T_TO_BQ	STG_SPW_0000_0000_F001_F_W_T156T	ODS_SPW_0000_F001_F_W_T156T
T001W	Full	PL_SPW_F001_F_W_T001W_TO_BQ	F001_SPW_F001_F_W_T001W_TO_BQ	STG_SPW_0000_0000_F001_F_W_T001W	ODS_SPW_0000_F001_F_W_T001W
T001L	Full	PL_SPW_F001_F_W_T001L_TO_BQ	F001_SPW_F001_F_W_T001L_TO_BQ	STG_SPW_0000_0000_F001_F_W_T001L	ODS_SPW_0000_F001_F_W_T001L
T001	Full	PL_SPW_F001_F_W_T001_TO_BQ	F001_SPW_F001_F_W_T001_TO_BQ	STG_SPW_0000_0000_F001_F_W_T001	ODS_SPW_0000_F001_F_W_T001
T064T	Full	PL_SPW_F001_F_W_T064T_TO_BQ	F001_SPW_F001_F_W_T064T_TO_BQ	STG_SPW_0000_0000_F001_F_W_T064T	ODS_SPW_0000_F001_F_W_T064T

We are using the Data Ocean standard jobs to extract and load the data into the STG/ODS tables. As we are dealing with Full and incremental type of loading so there is a little variance in the jobs, which is explained below:

Full Load

For the explanation, I am using F001_SPF_F001_F_W_ESLL_TO_BQ which is a full load job. Job parameters like STG table, ODS table , Bucket details , filters etc are maintained in the RDS parameters and its not hardcoded in the job, if required it can be changed any time in the RDS itself.

Parameters Snapshot

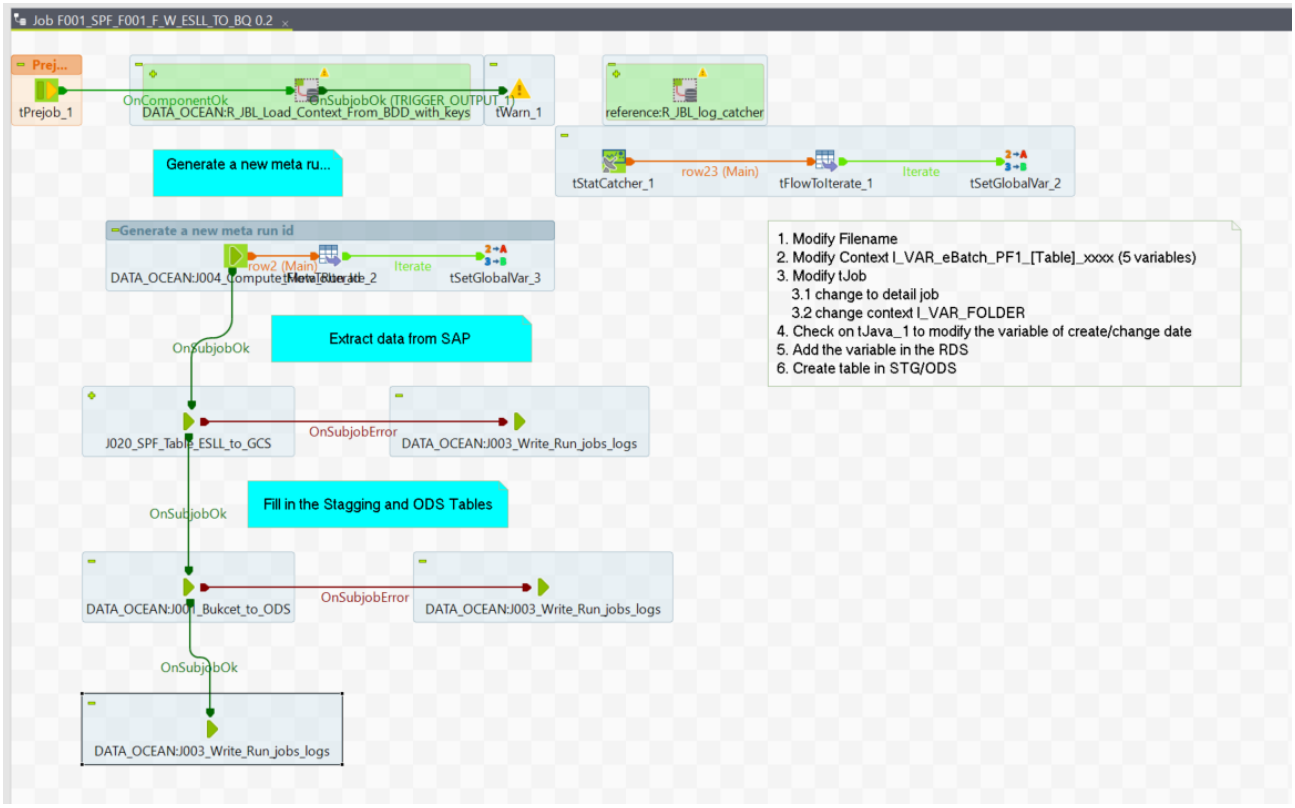
	env	key	value	project	LastModifiedOn	CreatedOn	Comment	key_type
1	DEV	I_VAR_PF1_ESLL_additional_filter		DATA_OCEAN_DOMAIN_PROCUREMENT	2024-04-26 10:26:22.317	2024-04-26 10:26:22.317		[NULL]
2	DEV	I_VAR_PF1_ESLL_BACKET	SPF_PCR_0000_0000_F001_F_W_ESLL	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-04-26 10:26:22.317	2024-04-26 10:26:22.317		[NULL]
3	DEV	I_VAR_PF1_ESLL_BQ_Table_ODS	ODS_SPF_0000_F001_F_W_ESLL	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-04-26 10:26:22.317	2024-04-26 10:26:22.317		[NULL]
4	DEV	I_VAR_PF1_ESLL_BQ_Table_STG	STG_SPF_0000_0000_F001_F_W_ESLL	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-04-26 10:26:22.317	2024-04-26 10:26:22.317		[NULL]
5	DEV	I_VAR_PF1_ESLL_INC_LOAD	PF1_ESLL	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-04-26 10:26:22.317	2024-04-26 10:26:22.317		[NULL]

- All parameters are self explanatory except the additional_filter which is used to define the type of the load i.e. blank means its a full refresh and "incremental" mean its a incremental load type. See the below example for PF1_MSEG which is a incremental load type, so we used "incremental" in the additional_filter parameter.

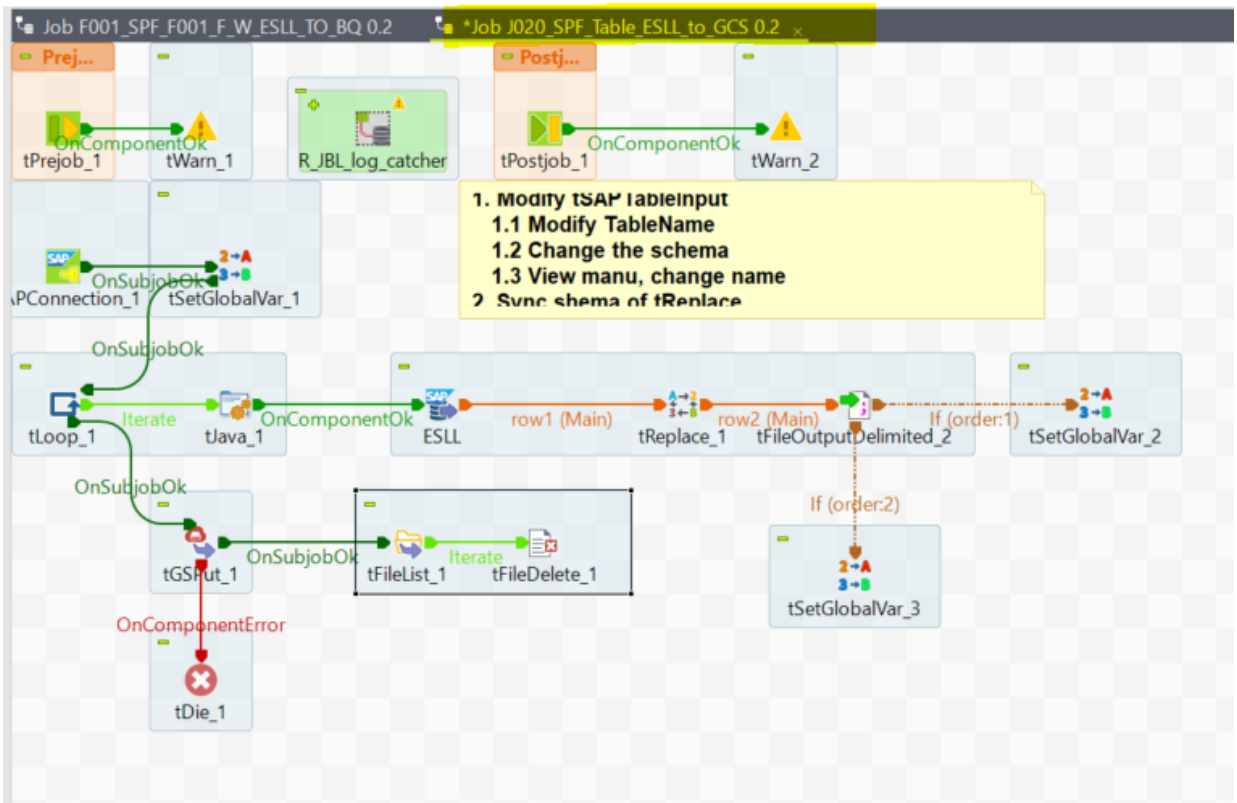
env	key	value	project	LastModifiedOn	CreatedOn	Comment	key_type
DEV	I_VAR_PF1_MSEG_additional_filter	incremental	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.703	2024-08-22 02:15:22.703	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_BACKET	SPF_PCR_0000_0000_F001_I_W_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_BQ_Table_ODS	ODS_SPF_0000_F001_I_W_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_BQ_Table_STG	STG_SPF_0000_0000_F001_I_W_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_INC_LOAD	PF1_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]

Job Details

Below is the snapshot of the main Flow job, where we are using all the Data Ocean standard jobs , except "SAP to GCS" extraction based on each source.



J020_SPF_Table_ESLL_to_GCS Job which is used to extract the data from the SAP table and load into the Google Cloud Storage, as I said all the parameters are configured in the RDS and no hardcoding of any values in the Talend job.



****
****

Incremental Load

For the explanation, I am using F001_SPF_F001_I_W_MSEG_TO_BQ which is a incremental load job. Job parameters like STG table, ODS table , Bucket details , filters etc are maintained in the RDS parameters and its not hardcoded in the job, if required it can be changed any time in the RDS itself.

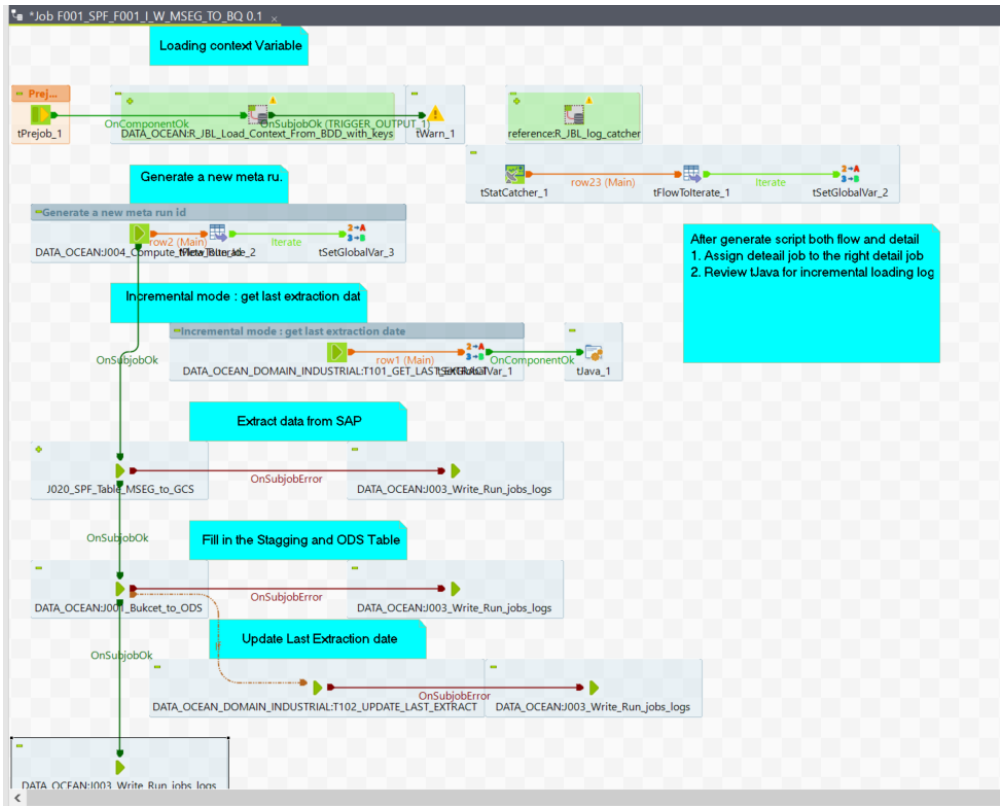
Parameters Snapshot

- All parameters are self explanatory except the additional_filter which is used to define the type of the load i.e. blank means its a full refresh and "incremental" mean its a incremental load type. See the below example for PF1_MSEG which is a incremental load type, so we used "incremental" in the additional_filter parameter.

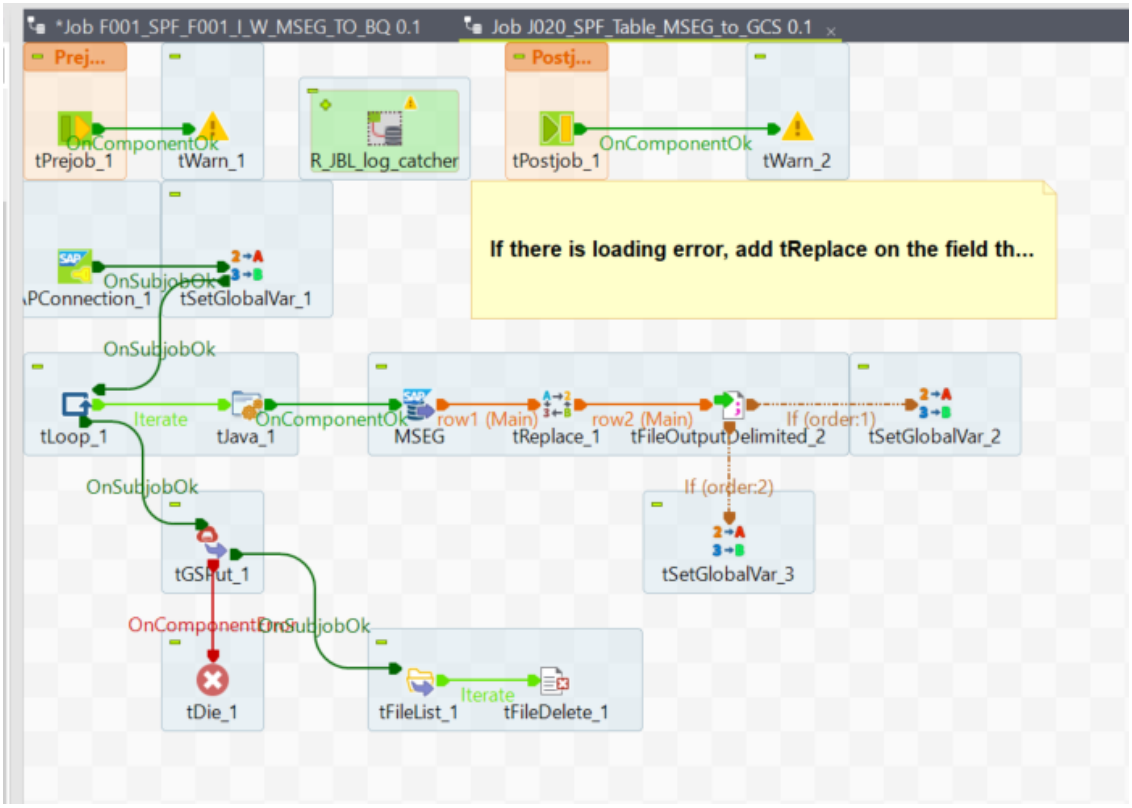
env	key	value	project	LastModifiedOn	CreatedOn	Comment	key_type
DEV	I_VAR_PF1_MSEG_additional_filter	incremental	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.703	2024-08-22 02:15:22.703	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_BACKET	SPF_PCR_0000_0000_F001_I_W_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_BQ_Table_ODS	ODS_SPF_0000_F001_I_W_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_BQ_Table_STG	STG_SPF_0000_0000_F001_I_W_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]
DEV	I_VAR_PF1_MSEG_INC_LOAD	PF1_MSEG	DATA_OCEAN_DOMAIN_PROCUREMENT	2024-08-22 02:15:22.707	2024-08-22 02:15:22.707	[NULL]	[NULL]

Job Details

Below is the snapshot of the main Flow job, where we are using all the Data Ocean standard jobs , except "SAP to GCS" extraction based on each source,

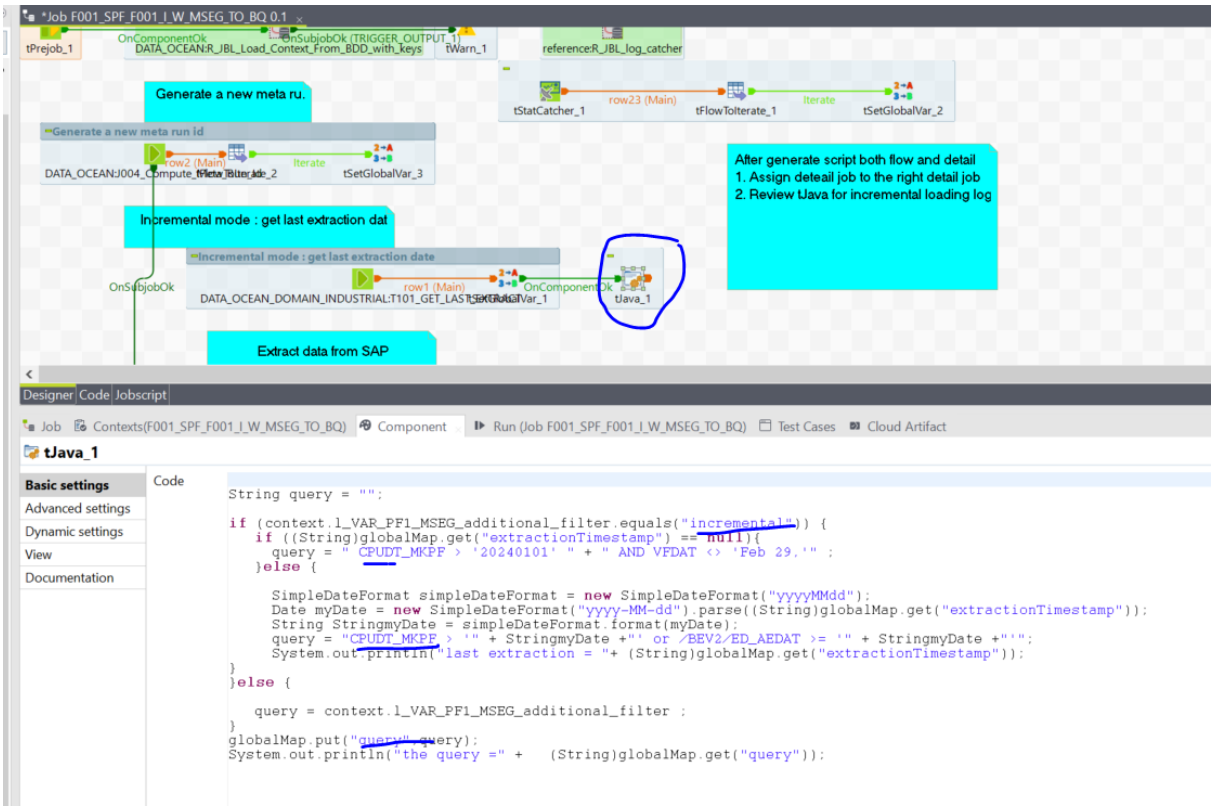


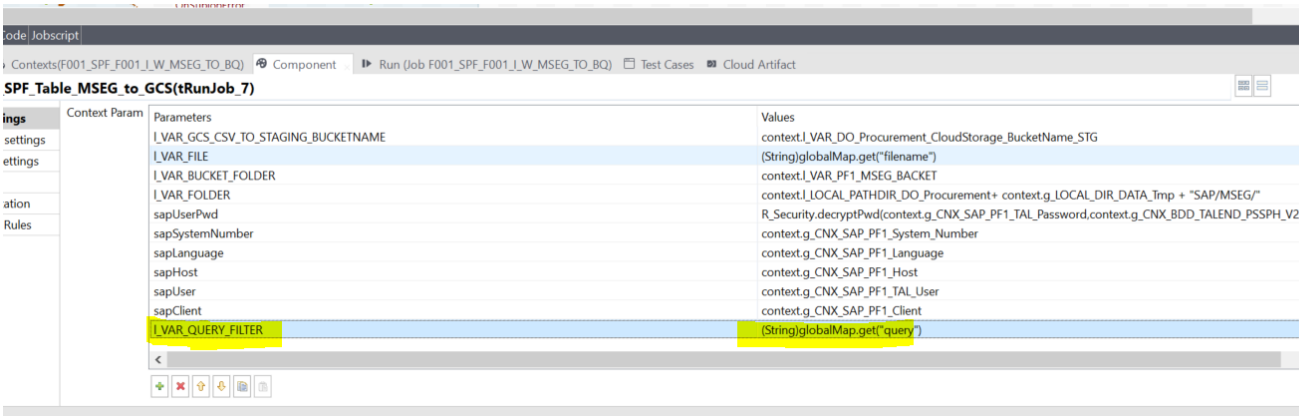
J020_SPF_Table_MSEG_to_GCS Job which is used to extract the data from the SAP table and load into the Google Cloud Storage, as I said all the parameters are configured in the RDS and no hardcoding of any values in the Talend job.



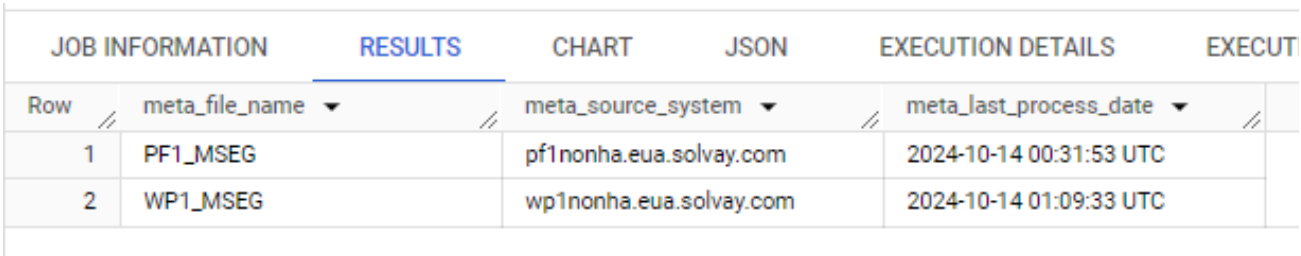
The only difference between the Full load and incremental load is incremental load will have additional sub jobs to deal with last execution datatime, this can be used to fetch the delta records. These jobs are standard jobs and the parameters values are sent from the main flow job

- T101_GET_LAST_EXTRACT : This job takes the I_VAR_PF1_MSEG_INC_LOAD parameter value and query to in the prj-data-dm-procurement-prod.STG.incremental_loading table to fetch the value in the meta_last_process_date column, which stores the last executed datetime from the delta column. And this value is being used in the tJava_1 where we construct the query which is passed to the extraction job in the below parameters





- T102_UPDATE_LAST_EXTRACT : This job is used to get the max("meta_stg_insert_date") from the STG table and update the prj-data-dm-procurement-prod.STG.incremental_loading table for that SOURCE_TABLE name. If you look at the below snapshot, where the value in meta_last_process_date tells the date & time till when we have extracted the records from the source table. In the next run, the above job will use this value and construct the filter condition so the job will extract the records which are inserted after this datetime.



Row	meta_file_name	meta_source_system	meta_last_process_date
1	PF1_MSEG	pf1nonha.eua.solvay.com	2024-10-14 00:31:53 UTC
2	WP1_MSEG	wp1nonha.eua.solvay.com	2024-10-14 01:09:33 UTC