

Workshop: Data Integration SCD using Talend

Name: Shreya Pattabhiraman Iyer

NUID: 001300102

1) Talend job for EmployeeDeptChangesv2

Talend Real-time Big Data Platform (8.0.1.20211109_1610) | Local_Project (Connection: Local)

File Edit View Search Window Help

Feature Manager Support

Repository LOCAL: Local_Project

- Job Designs
 - Standard
 - demo
 - demo_20220304
 - EmployeeDeptChangesv2_20220402
 - sf_films_location
 - Contexts
 - Code
 - SQL Templates
 - Metadata
 - Db Connections
 - AdventureWorks2019 0.1
 - AdventureWorksDW2019 0.1
 - Chinook 0.1
 - dmg7370_output 0.1
 - Queries
 - Synonym schemas
 - Table schemas
 - View schemas
 - File delimited
 - EmployeeDeptChangesv2_20220402
 - EmployeeDeptChangesv2_2022042 1.0
 - EmployeeDeptChangesv3_20220402
 - metadata
 - Columns(5)
 - san_francisco_films
 - Street_Tree_List_tsv 0.1
 - File positional
 - File regex
 - File XML
 - File Excel
 - File Idif

Job Job_EmployeeDeptChangesLogRow.v3 0.1

Job Job_EmployeeDeptChangesLogRow 0.1

EmployeeDeptChangev2

311 rows in 0.14s
2205.67 rows/s
deptlogrows (Main)

290 rows in 0.85s
341.58 rows/s
vempllookup (Lookup)

311 rows in 0.39s
807.79 rows/s
putLogrow (Main order2)

EmployeeDeptChangesv2_metadata

tMap_1

vEmployee

EmployeeDeptHistorySCDv2

tLogRow_1

Designer Code

Job Job_EmployeeDeptChangesLogRow

Basic Run

Debug Run

Advanced settings

Target Exec

Memory Run

Execution

Run Kill Clear

Starting job Job_EmployeeDeptChangesLogRow at 17:12 02/04/2022.
[statistics] connecting to socket on port 3922
[statistics] connected
[statistics] disconnected
Job Job_EmployeeDeptChangesLogRow ended at 17:12 02/04/2022. [Exit code

Default

Name	Value
------	-------

sort

Favorites

Recently Used

- tSortRow
- tSplitRow
- tLogRow
- tMap
- tBigQueryInput
- tFileInputDelimited

Big Data

- Google BigQuery
- Google Storage
- Hive

Business Intelligence

Business

Cloud

- Amazon
- Azure
- Box
- Dropbox
- Google Storage
- Google
- Salesforce
- ServiceNow
- tCloudStart
- tCloudStop

Custom Code

Data Quality

Databases

DotNET

ELT

- Combined SQL
- Connections
- Map
- SQLTemplate

2) Talend job for EmployeeDeptChangev3

Talend Real-time Big Data Platform (8.0.1.20211109_1610) | Local_Project (Connection: Local)

File Edit View Search Window Help

Feature Manager Support

Repository LOCAL: Local_Project

- Job Designs
 - Standard
 - demo
 - demo_20220304
 - EmployeeDeptChangesv2_20220402
 - sf_films_location
 - Contexts
 - Code
 - SQL Templates
 - Metadata
 - Db Connections
 - AdventureWorks2019 0.1
 - AdventureWorksDW2019 0.1
 - Chinook 0.1
 - dmg7370_output 0.1
 - Queries
 - Synonym schemas
 - Table schemas
 - View schemas
 - File delimited
 - EmployeeDeptChangesv2_20220402
 - EmployeeDeptChangesv3_2022 1.0
 - EmployeeDeptChangesv3_20220402
 - metadata
 - Columns(5)
 - san_francisco_films
 - Street_Tree_List_tsv 0.1
 - File positional
 - File regex
 - File XML
 - File Excel
 - File Idif

Job Job_EmployeeDeptChangesLogRow.v3 0.1

Job Job_EmployeeDeptChangesLogRow 0.1

Truncate

EmployeeDeptHistorySCDv2

OnSubJobOk

SCD Dept History

EmployeeDeptChangesv2_metadata

tSortRow_1

312 rows in 0.02s
15600 rows/s
getFile (Main)

312 rows in 0.06s
5288.14 rows/s
deptlogrows (Main)

290 rows in 0.27s
1062.27 rows/s
vempllookup (Lookup)

312 rows in 0.14s
2151.72 rows/s
putLogrow (Main order2)

vEmployee

EmployeeDeptHistorySCDv2

tLogRow_1

Designer Code

Job Job_EmployeeDeptChangesLogRow_v3

Basic Run

Debug Run

Advanced settings

Target Exec

Memory Run

Execution

Run Kill Clear

Starting job Job_EmployeeDeptChangesLogRow_v3 at 15:56 02/04/2022.
[statistics] connecting to socket on port 4036
[statistics] connected
[statistics] disconnected
Job Job_EmployeeDeptChangesLogRow_v3 ended at 15:57 02/04/2022. [Exit code

Default

Name	Value
------	-------

sort

Favorites

Recently Used

- tSortRow
- tSplitRow
- tLogRow
- tMap
- tBigQueryInput
- tFileInputDelimited

Big Data

- Google BigQuery
- Google Storage
- Hive

Business Intelligence

Business

Cloud

- Amazon
- Azure
- Box
- Dropbox
- Google Storage
- Google
- Salesforce
- ServiceNow
- tCloudStart
- tCloudStop

Custom Code

Data Quality

Databases

DotNET

ELT

- Combined SQL
- Connections
- Map
- SQLTemplate