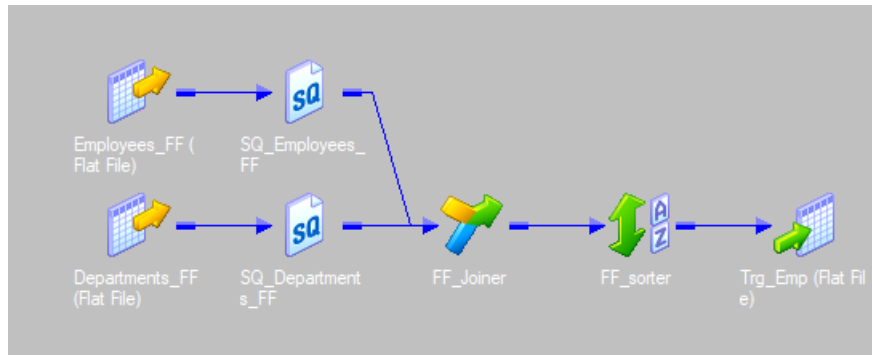


## Designer Window

### Design Mapping



### Create Transformation

#### sorter

Edit Transformations

Transformation | Ports | Properties | Metadata Extensions

Select transformation: **FF\_sorter**

Transformation type: **Sorter**

Transformation Attribute	Value
Sorter Cache Size	Auto
Case Sensitive	<input checked="" type="checkbox"/>
Work Directory	\$PMTempDir
Distinct	<input checked="" type="checkbox"/>
Tracing Level	Normal
Null Treated Low	<input type="checkbox"/>
Transformation Scope	All Input

**Distinct**  
Indicates if distinct records are outputted. Once distinct is selected, all columns will be selected.

OK Cancel Apply

#### Joiner

Edit Transformations

Transformation | Ports | Properties | Condition | Metadata Extensions

Select transformation: **FF\_Joiner**

Transformation type: **Joiner**

Transformation Attribute	Value
Cache Directory	\$PMTCacheDir
Join Condition	DEPARTMENT_ID1 = DEPARTMENT_ID
Join Type	Master Outer Join
Null ordering in master	Null is Highest Value
Null ordering in detail	Null is Highest Value
Tracing Level	Normal
Joiner Data Cache Size	Auto
Joiner Index Cache Size	Auto
Sorter Input #	1

**Join Type**  
Type of Join

OK Cancel Apply Help

## Workflow Manager

### Design Workflow



### Configure connection

Edit Tasks

General | Properties | Config Object | Mapping | Components | Metadata Extensions

Select task: **s\_F2F**

Task type: **Session**

- Start Page
- Connections
- Memory Properties
- Files, Directories and Com
- Sources
  - SQ SQ\_Employees\_FF
  - SQ SQ\_Departments\_FF
- Targets
  - Trg\_Emp
- Transformations
  - FF\_Joiner
  - FF\_sorter

**Targets**

Instance	Writers
Trg_Emp	File Writer

**Connections**

Type	Value	Connections
Trg_Emp - Connection	None	Connection

**Properties** Set File Show Session Level Properties

Attribute	Value
Trg_Emp - File Writer	
Merge Type	No Merge
Merge File Directory	\$PMTTargetFileDir\
Merge File Name	trg_emp1.out
Append if Exists	<input type="checkbox"/>

**Trg\_Emp - File Writer**

OK Cancel Apply Help

Workflow Monitor

Run Properties

Connec	ETL QA	Workflow Run	Start Time	Completion Time	Status
wf_F2F	wf_F2F	wf_F2F	21-02-2024 18:58:05	In progress	Running
wf_F2F	wf_F2F	s_F2F	21-02-2024 18:58:05	21-02-2024 18:58:06	Succeeded
wf_F2F	wf_F2F	wf_F2F	21-02-2024 18:32:11	21-02-2024 18:32:15	Succeeded
wf_F2F	wf_F2F	wf_F2F	21-02-2024 18:32:11	21-02-2024 18:32:13	Succeeded

Gantt Chart Task View

s\_F2F [21-02-2024 18:58:05]

Task Details

Source/Target Statistics

Transformation Name	Node	Applied Rows	Affected Rows	Rejected Rows	Throughput (Rows/Sec)	Throughput (Rows)
sq_SQ_Departments_FF	node01	27	27	0	27	783
sq_SQ_Employees_FF	node01	109	109	0	109	3597
Trg_Emp	node01	107	107	0	107	7276

Session Logs

SESSION LOAD SUMMARY						
INFO	21-02-2024 18:58:06	node01	DIRECTOR	TM_6252	Source Load Summary.	
INFO	21-02-2024 18:58:06	node01	DIRECTOR	CMN_1740	Table: [SQ_Departments_FF] (Instance Name: [SQ_Departments_FF])	
					Output Rows [27], Affected Rows [27], Applied Rows [27], Rejected Rows [0]	
INFO	21-02-2024 18:58:06	node01	DIRECTOR	CMN_1740	Table: [SQ_Employees_FF] (Instance Name: [SQ_Employees_FF])	
					Output Rows [109], Affected Rows [109], Applied Rows [109], Rejected Rows [0]	
INFO	21-02-2024 18:58:06	node01	DIRECTOR	TM_6253	Target Load Summary.	
INFO	21-02-2024 18:58:06	node01	DIRECTOR	CMN_1740	Table: [Trg_Emp] (Instance Name: [Trg_Emp])	
					Output Rows [107], Affected Rows [107], Applied Rows [107], Rejected Rows [0]	
INFO	21-02-2024 18:58:06	node01	DIRECTOR	TM_6023		
INFO	21-02-2024 18:58:06	node01	DIRECTOR	TM_6020	Session [s_F2F] completed at [Wed Feb 21 18:58:06 2024].	

Source Vs Target

Source  
Employees Flat File

```
F:\VARMA\IT_Transition\ETL\F2F>bash
varma@DESKTOP-0E3HS0A:/mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Employees.csv
EMPLOYEE_ID,FULL_NAME,DEPARTMENT_ID
198,DonaldOConnell,50
199,DouglasGrant,50
200,JenniferWhalen,10
201,MichaelHartstein,20
202,PatFay,20
203,SusanMavris,40
204,HermannBaer,70
```

Departments Flat File

```
varma@DESKTOP-0E3HS0A:/mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Departments.csv
"DEPARTMENT_ID","DEPARTMENT_NAME"
10,"Administration"
20,"Marketing"
30,"Purchasing"
40,"Human Resources"
50,"Shipping"
60,"IT"
70,"Public Relations"
80,"Sales"
90,"Executive"
100,"Finance"
```

Target

Trg\_Emp Flat File

```
varma@DESKTOP-0E3HS0A:/mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Trg_Emp.csv
#EMPLOYEE_ID,FULL_NAME,DEPARTMENT_ID,DEPARTMENT_NAME
100,"StevenKing",90,"Executive"
101,"NeenaKochhar",90,"Executive"
102,"LexDe Haan",90,"Executive"
103,"AlexanderHunold",60,"IT"
104,"BruceErnst",60,"IT"
105,"DavidAustin",60,"IT"
106,"ValliPataballa",60,"IT"
107,"DianaLorentz",60,"IT"
108,"NancyGreenberg",100,"Finance"
109,"DanielFaviet",100,"Finance"
110,"JohnChen",100,"Finance"
```

## Positive Scenarios

### Count of Records in Source vs Target

```
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Employees.csv | wc -l
110
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Departments.csv | wc -l
28
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Trg_Emp.csv | wc -l
108
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$
```

### Duplicates In Source vs Target

```
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Employees.csv | sort | uniq -d
151,DavidBernstein,80
152,PeterHall,80
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Trg_Emp.csv | sort | uniq -d
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$
```

### Sorter & Joiner Transformation in Target Files

Employee Details in Ascending Order

Department Name in Target (Joined from Departments File)

```
varma@DESKTOP-0E3HS0A: /mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Trg_Emp.csv | head
#EMPLOYEE_ID,FULL_NAME,DEPARTMENT_ID,DEPARTMENT_NAME
100,"StevenKing",90,"Executive"
101,"NeenaKochhar",90,"Executive"
102,"LexDe Haan",90,"Executive"
103,"AlexanderHunold",60,"IT"
104,"BruceErnst",60,"IT"
105,"DavidAustin",60,"IT"
106,"ValliPataballa",60,"IT"
107,"DianaLorentz",60,"IT"
108,"NancyGreenberg",100,"Finance"
```

## Negative Scenarios

### Invalid Datatype Delimiter Missing

```
EMPLOYEE_ID,FULL_NAME,DEPARTMENT_ID
198 DonaldOConnell,50
199a,DouglasGrant,50
VAR,JenniferWhalen,10
```

```
F:\VARMA\IT_Transition\ETL\F2F>bash
varma@DESKTOP-0E3HS0A:/mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Employees.csv |wc -l
110
varma@DESKTOP-0E3HS0A:/mnt/f/VARMA/IT_Transition/ETL/F2F$ cat Trg_Emp.csv | wc -l
105
varma@DESKTOP-0E3HS0A:/mnt/f/VARMA/IT_Transition/ETL/F2F$
```

Lesser Rows in Target in comparison to source

1. Delimiter Not specified in a row
2. Alpha numeric entry for Employee ID, which has to be numeric
3. String instead of Numeric for Employee ID

### Rejected Rows

Transformation Name	Node	Applied Rows	Affected Rows	Rejected Rows	Throughput (Rows/Sec)
SQ_Departments_FF	node01	27	27	0	27
SQ_Employees_FF	node01	106	106	3	106
Trg_Emp	node01	104	104	0	104

Session ID	Time	Node	Director	TM	Summary
INFO	21-02-2024 22:59:21	node01	DIRECTOR	TM_6252	Source Load Summary.
INFO	21-02-2024 22:59:21	node01	DIRECTOR	CMN_1740	Table: [SQ_Employees_FF] (Instance Name: [SQ_Employees_FF]) Output Rows [106], Affected Rows [106], Applied Rows [106], Rejected Rows [3]
INFO	21-02-2024 22:59:21	node01	DIRECTOR	CMN_1740	Table: [SQ_Departments_FF] (Instance Name: [SQ_Departments_FF]) Output Rows [27], Affected Rows [27], Applied Rows [27], Rejected Rows [0]
INFO	21-02-2024 22:59:21	node01	DIRECTOR	TM_6253	Target Load Summary.
INFO	21-02-2024 22:59:21	node01	DIRECTOR	CMN_1740	Table: [Trg_Emp] (Instance Name: [Trg_Emp]) Output Rows [104], Affected Rows [104], Applied Rows [104], Rejected Rows [0]
INFO	21-02-2024 22:59:21	node01	DIRECTOR	TM_6023	

Flat file : Simple data sources, where data is stored in text format.

Types of Flat files:

1. Delimited flat File: Data is separated with delimiter (ex: CSV)
2. Fixed width Flat File: Where fields are defined on fixed width

## Design & Development

Import source and target details to the Designer (Both are Flat files)

Design Mapping from source to target with necessary transformation

Joiner Transformation used to get data from Multiple sources

Sorter transformation is used to sort the data and distinct properties defined to get unique records in the target.

## Data Validation

Check Run properties and session logs for the detailed report

Validate source and target with Unix Commands

Positive Scenarios:

1. Check Record counts in Source and Target
2. Check for Duplicates in Target
3. Check Transformation logic applied in Target
4. Use Diff Command to check for truncation of Data

## Negative Scenarios

1. Check for Different Delimiter than specified in SQ

2. Change Datatype in the column entries
3. Error Handling Mechanism (Ex: if the file is absent)
4. Process with zero byte files (File having No data)