

Add12 Split file



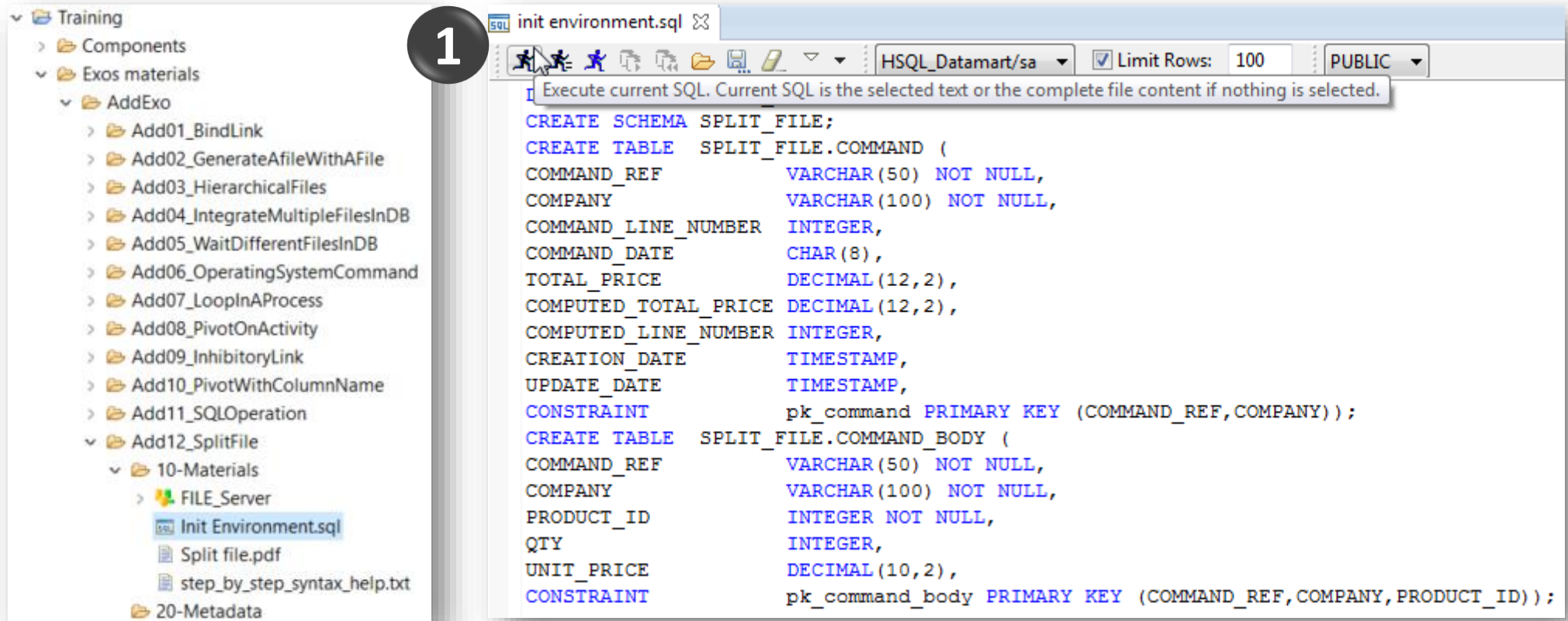
Related articles in Stambia.org

- <http://www.stambia.org/doc/200-technology-articles/file/transformation-scripting/420-transformation-scripting>
- <http://www.stambia.org/doc/130-technology-articles/file/multirecord-files/432-reading-only-a-specified-range-of-records>

- ❖ The Objective is to be able to load in database a file that has header data in the first record, body data (n rows) and footer data on the last record

command_ref	line_number	company	product_id	product_qty	product_unit_price	command_date	total_price
1	1245XYZ125	4	Stambia				
2	1245	10	15.45				
3	4567	1	157.28				
4	7895	5	15.00				
5	8452	16	0.45				
6	20170614	393.98					

- ❖ 1 - Execute the SQL file (Init Environment.sql) on DATAMART connection to create the required schema and tables



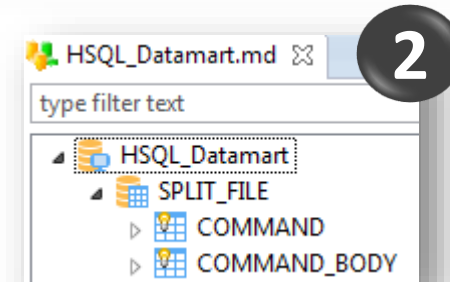
The screenshot shows a SQL IDE interface. On the left, a tree view under 'Training' shows a folder 'Add12_SplitFile' containing '10-Materials', which in turn contains 'FILE_Server' and 'Init Environment.sql'. The 'Init Environment.sql' file is selected. On the right, the SQL editor shows the following code:

```

CREATE SCHEMA SPLIT_FILE;
CREATE TABLE SPLIT_FILE.COMMAND (
  COMMAND_REF          VARCHAR(50) NOT NULL,
  COMPANY              VARCHAR(100) NOT NULL,
  COMMAND_LINE_NUMBER  INTEGER,
  COMMAND_DATE         CHAR(8),
  TOTAL_PRICE          DECIMAL(12,2),
  COMPUTED_TOTAL_PRICE DECIMAL(12,2),
  COMPUTED_LINE_NUMBER INTEGER,
  CREATION_DATE        TIMESTAMP,
  UPDATE_DATE          TIMESTAMP,
  CONSTRAINT           pk_command PRIMARY KEY (COMMAND_REF,COMPANY));
CREATE TABLE SPLIT_FILE.COMMAND_BODY (
  COMMAND_REF          VARCHAR(50) NOT NULL,
  COMPANY              VARCHAR(100) NOT NULL,
  PRODUCT_ID           INTEGER NOT NULL,
  QTY                  INTEGER,
  UNIT_PRICE           DECIMAL(10,2),
  CONSTRAINT           pk_command_body PRIMARY KEY (COMMAND_REF,COMPANY,PRODUCT_ID));
  
```

A toolbar at the top of the editor includes icons for running SQL, saving, and other functions. A tooltip is visible over the run icon, stating: "Execute current SQL. Current SQL is the selected text or the complete file content if nothing is selected." The toolbar also shows the connection 'HSQL_Datamart/sa', a 'Limit Rows' dropdown set to 100, and a 'PUBLIC' dropdown.

- ❖ 2 - Reverse the schema and the tables



The screenshot shows a database explorer window titled 'HSQL_Datamart.md'. It displays a tree view of the database structure. Under the 'HSQL_Datamart' node, there is a folder 'SPLIT_FILE'. Inside 'SPLIT_FILE', there are two tables: 'COMMAND' and 'COMMAND_BODY'. A large red number '2' is overlaid on the right side of the image.

❖ 3 - Reverse a first time the file (FileToSplit.txt)

- Define the 3 columns

File Wizard

Step 1/2
Set file properties

Properties

Name: FileToSplit_header

Physical Name: FileToSplit.txt Browse

Type: DELIMITED

Charset Name:

Line separator: \n \n 0D0A

Field separator: ; 3B

String delimiter:

Decimal separator: . 2E

Heading Lines To Skip: 0

Ending Lines To Skip: 0

Line Size:

Header Line Position: 0

Preview

Row limit: 100 Refresh

? < Back Next > Finish Cancel

3

File Wizard

Step 2/2
Reverse file fields

Reverse

Helper

Row Limit (analysing): 1000 Reverse Helper

Reverse options

☒ Name ☒ Physical Name ☒ Size ☒ Type ☒ Link position/size

Size Management

String min size: 50 safety margin: 50 ☒ determine datatype

Numeric min size: 12 safety margin: 0

Name	Position	Size	Physical Size	Type
COMMAND_REF	1	50		String
LINE_NUMBER	2	5		Numeric
COMPANY	3	100		String

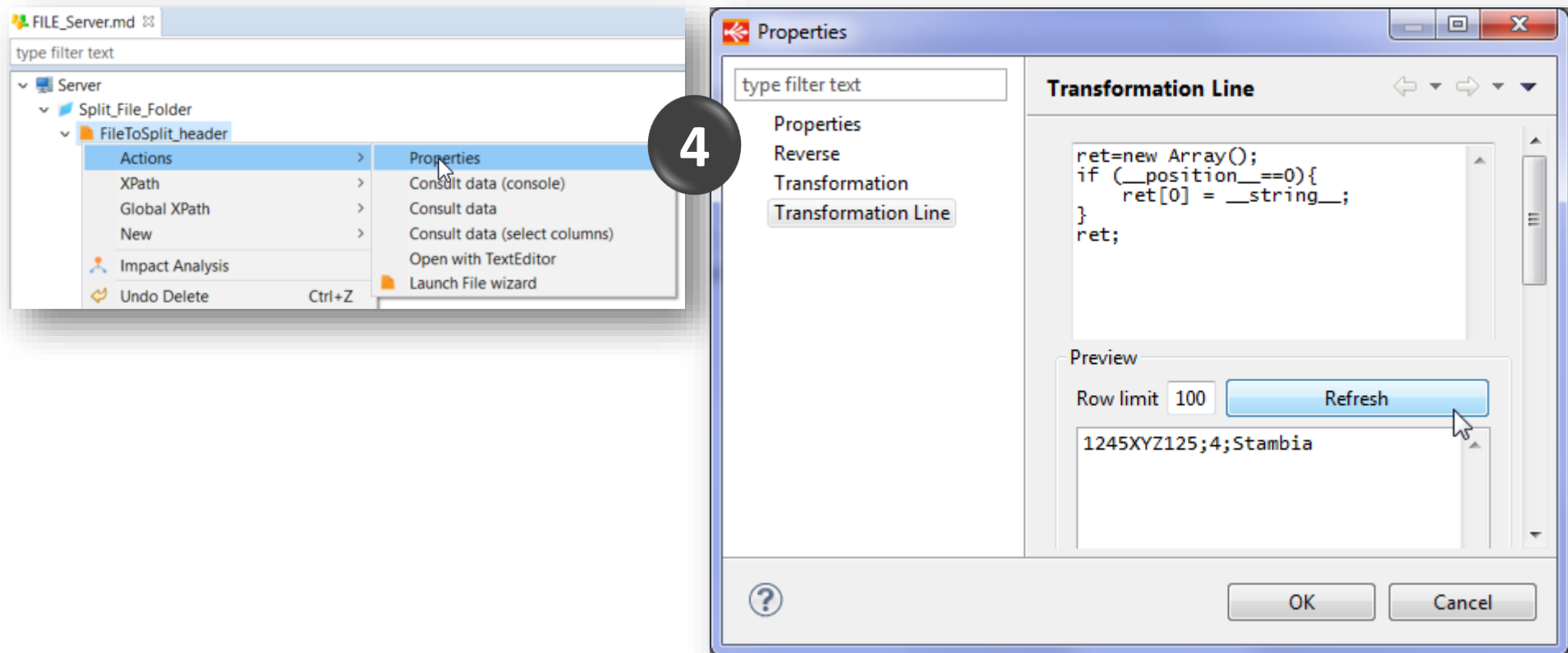
Refresh ☐ Auto-refresh Preview

? < Back Next > Finish Cancel

FILE_Server

- Server
 - Split_File_Folder
 - FileToSplit_header
 - COMMAND_REF
 - LINE_NUMBER
 - COMPANY

- ❖ 4 - Right click on the file, action > Properties
 - Add the following JavaScript code in Transformation Line to select only the first line
 - Push on Refresh button to check the result



❖ 5 - Reverse a second time the file (a new file hit) and define the 3 columns

File Wizard
Step 1/2
Set file properties

Name: FileToSplit_lines

Physical Name: FileToSplit.txt Browse

Type: DELIMITED

Charset Name:

Line separator: \r\n 0D0A

Field separator: ; 3B

String delimiter:

Decimal separator: . 2E

Heading Lines To Skip: 1

Ending Lines To Skip: 1

Line Size:

Header Line Position: 0

Preview

Row limit: 100 Refresh

```
1245;10;15.45
4567;1;157.28
7895;5;15.00
8452;16;0.45
```

? < Back Next > Finish Cancel

5

FILE_Server.md

type filter text

- Server
 - Split_File_Folder
 - FileToSplit_header
 - FileToSplit_lines
 - PRODUCT_ID
 - QTY
 - UNIT_PRICE

File Wizard
Step 2/2
Reverse file fields

Helper

Row Limit (analysing) 1000 Reverse Helper

Reverse options

☒ Name ☒ Physical Name ☒ Size ☒ Type ☒ Link position/size

Size Management

String min size 50 safety margin 50 ☒ determine datatype

Numeric min size 12 safety margin 0

Name	Position	Size	Physi...	Type	Decimal
PRODUCT_ID	1	10		Numeric	
QTY	2	10		Numeric	
UNIT_PRICE	3	10		Numeric	2

Preview

Row Limit 100 Refresh ☐ Auto-refresh Preview

```
1245 10 15.45
4567 1 157.28
7895 5 15.00
8452 16 0.45
```

? < Back Next > Finish Cancel

- ❖ 6 - Reverse a 3rd time the file (a new file hit) and define the 2 columns
 - Add a record named “LastRecord”
 - Add a filter under the record and set First Line property to “-1” value on Line Range
 - Means that the file will begin at the last line

The screenshot shows the Stambia interface with a tree view on the left and a configuration panel on the right. A large number '6' is overlaid on the tree view, indicating the step number.

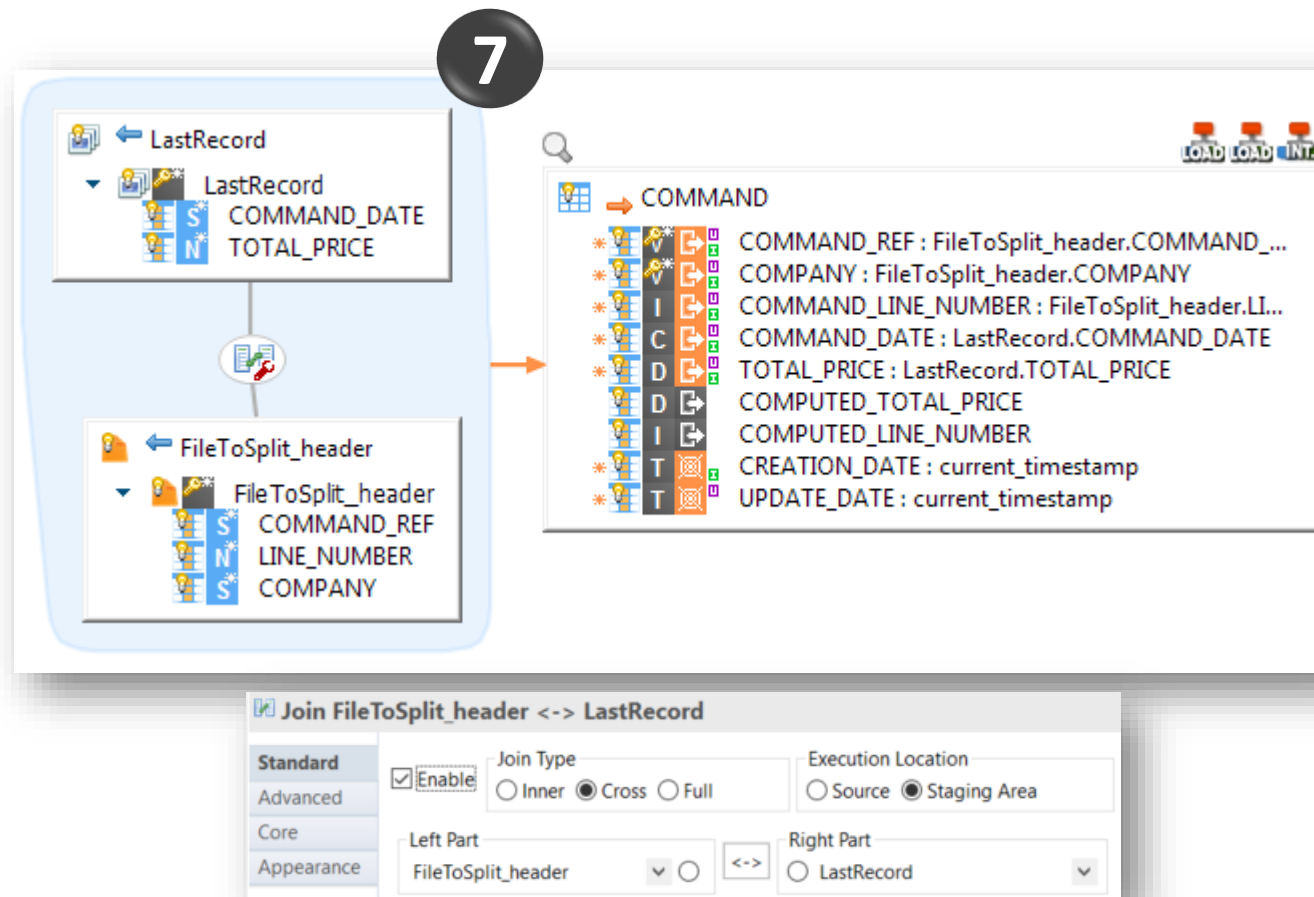
Tree View:

- Server
 - Split_File_Folder
 - FileToSplit_footer
 - LastRecord
 - OnlyLastLine** (selected)
 - COMMAND_DATE
 - TOTAL_PRICE
 - FileToSplit_header
 - FileToSplit_lines

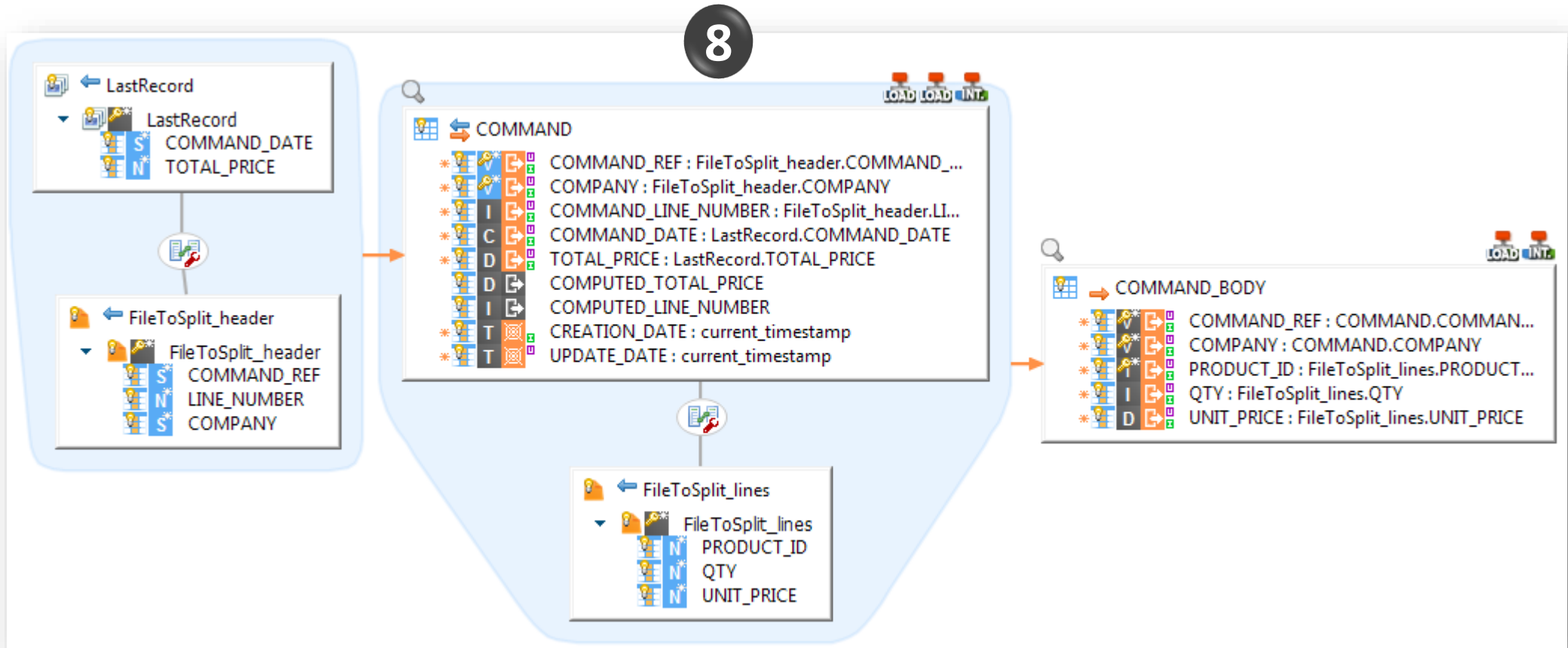
Configuration Panel (OnlyLastLine):

- Name:** OnlyLastLine
- Data Criteria:**
 - Test value: []
 - Starts At: []
 - Length: []
 - ☐ Not
 - Operator: Equals
 - ☐ Ignore Case
- Line Range:**
 - First Line: -1
 - Last Line: []
- ☐ Exclude
- Ancestor Level Context: []
- Description: []

- ❖ 7 - Create a mapping to load data in database
 - Do a cross join between header file and last record to load COMMAND table



- ❖ 8 - Update the mapping to load COMMAND_BODY table
 - Do a cross join between lines file and COMMAND table

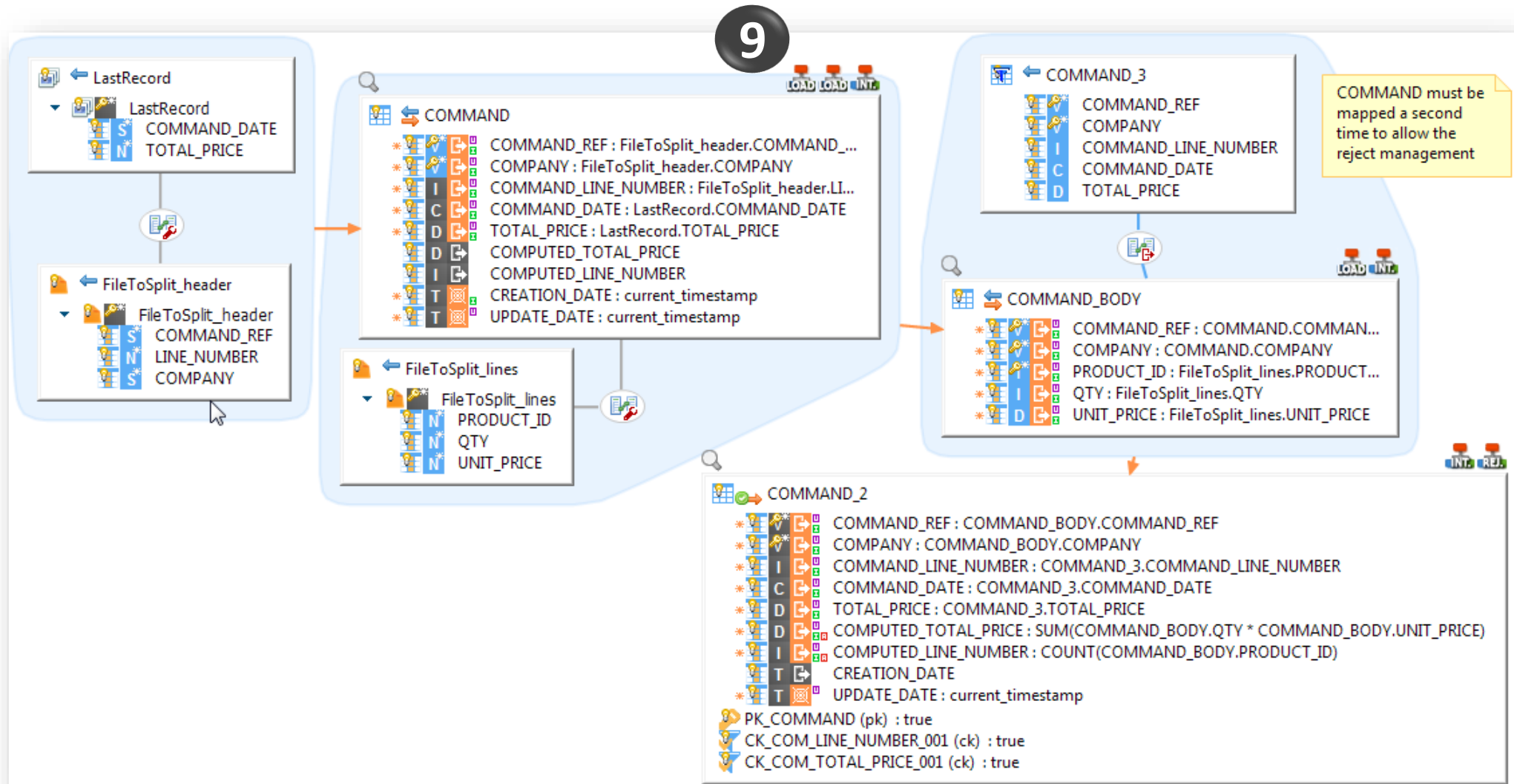


Join FileToSplit_lines <-> COMMAND

Standard	<input checked="" type="checkbox"/> Enable	Join Type <input type="radio"/> Inner <input checked="" type="radio"/> Cross <input type="radio"/> Full	Execution Location <input type="radio"/> Source <input checked="" type="radio"/> Staging Area
Advanced			
Core			
Appearance	Left Part FileToSplit_lines	Right Part COMMAND	

❖ 9 - Update the mapping to check the COMMAND table (rejects)

- Create two conditions
- Compute TOTAL_PRICE and LINE_NUMBER



❖ The result for the two tables:

```
select * from SPLIT_FILE.COMMAND
```

[select * from SPLIT_FL...] X Messages

COMMAND_REF	COMPANY	COMMAND_LINE_NUMBER	COMMAND_DATE	TOTAL_PRICE	COMPUTED_TOTAL_PRICE	COMPUTED_LINE_NUMBER	CREATION_DATE	UPDATE_DATE
1245XYZ125	Stambia	4	20170614	393,98	393,98	4	2017-06-15 10:14:33:662	2017-06-15 10:14:33:794

```
select * from SPLIT_FILE.COMMAND_BODY
```

[select * from SPLIT_FL...] X Messages

COMMAND_REF	COMPANY	PRODUCT_ID	QTY	UNIT_PRICE
1245XYZ125	Stambia	1245	10	15,45
1245XYZ125	Stambia	4567	1	157,28
1245XYZ125	Stambia	7895	5	15,00
1245XYZ125	Stambia	8452	16	0,45

❖ And with the error file (FileToSplit_err.txt) :

```
select * from SPLIT_FILE.R_COMMAND
```

[select * from SPLIT_FL...] X Messages

SESSION_ID	REJECT_MODE	REJECT_MESSAGE	REJECT_CODE	REJECTED_BY	RULE_NAME	RULE_TYPE	RULE_SEVERITY	REJECT_DATE
ac140a02015cabb18e0fcc0106ae42c	F	Price issue	COM_001	Split files and load data in DB (_BFc5cFGdEee7Oa2LiZLdYA)	CK_COM_TOTAL_PRICE_001	CK	REJECT	2017-06-15 14:29:45:565
ac140a02015cabb18e0fcc0106ae42c	F	Line number issue	COM_002	Split files and load data in DB (_BFc5cFGdEee7Oa2LiZLdYA)	CK_COM_LINE_NUMBER_001	CK	REJECT	2017-06-15 14:29:45:571

COMMAND_REF	COMPANY	COMMAND_LINE_NUMBER	COMMAND_DATE	TOTAL_PRICE	COMPUTED_TOTAL_PRICE	COMPUTED_LINE_NUMBER
1245XYZ125	Stambia	5	20170614	400,00	393,98	4
1245XYZ125	Stambia	5	20170614	400,00	393,98	4

- ❖ You can also check another metadata already defined :

```
AnotherFileToSplit.txt
-- I have two lines
-- not used
1245XYZ125;5;Stambia
1234ABC578;6;Stambia
1;12
2;42
3;45
4;87
5;98
6;784
7;956
8;24
9;465
10;456
11;965
```

```
Training
├── Components
├── Exos materials
├── Files_In
│   ├── Company countries
│   ├── Countries
│   ├── Excel_Files
│   ├── fileControl
│   ├── Heterogeneous files to load in DB
│   ├── HierarchicalFiles
│   ├── Json
│   ├── MultiGenFiles
│   ├── Multiple_Files
│   ├── Reference_Files
│   └── Split_Files
│       ├── AnotherFileToSplit.txt
│       ├── FileToSplit_err.txt
│       └── FileToSplit.txt
```

```
Training
├── Components
├── Exos materials
│   └── AddExo
│       ├── Add01_BindLink
│       ├── Add02_GenerateAfileWithAFile
│       ├── Add03_HierarchicalFiles
│       ├── Add04_IntegrateMultipleFilesInDB
│       ├── Add05_WaitDifferentFilesInDB
│       ├── Add06_OperatingSystemCommand
│       ├── Add07_LoopInAProcess
│       ├── Add08_PivotOnActivity
│       ├── Add09_InhibitoryLink
│       ├── Add10_PivotWithColumnName
│       ├── Add11_SQLOperation
│       └── Add12_SplitFile
│           ├── 10-Materials
│           │   ├── Another FILE_Server
│           │   ├── FILE_Server
│           │   ├── Init Environment.sql
│           │   └── step_by_step_syntax_help.txt
│           └── 20-Metadata
```

- ❖ And consult the transformation expression to retrieve only the line 3 and 4 with 2 lines to skip :

Properties

type filter text

Properties

Reverse

Transformation

Transformation Line

Transformation

```

i=0;
nblines=0;
do{
    /* read the next character */
    i=__in__.read();

    /* if a character is found */
    if (i>-1){
        ch=String.fromCharCode(i);

        /* if this is a carriage return, increment the counter */
        if (ch == '\n'){
            nblines++;
        }

        i=ch.charCodeAt();
        __out__.write(i);
    }
}

```

Preview

Row limit 100 Refresh

1245XYZ125;5;Stambia
1234ABC578;6;Stambia