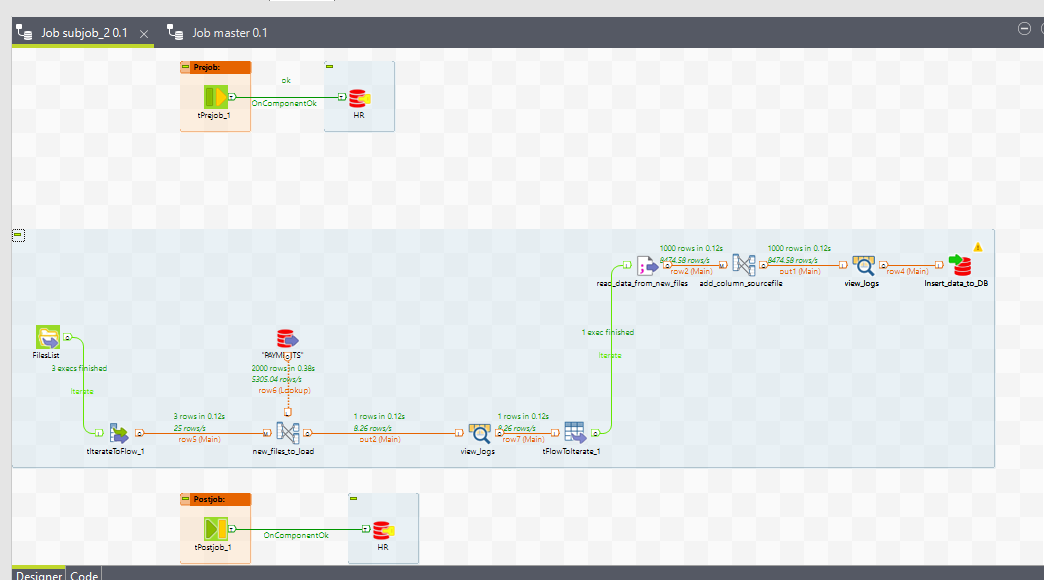
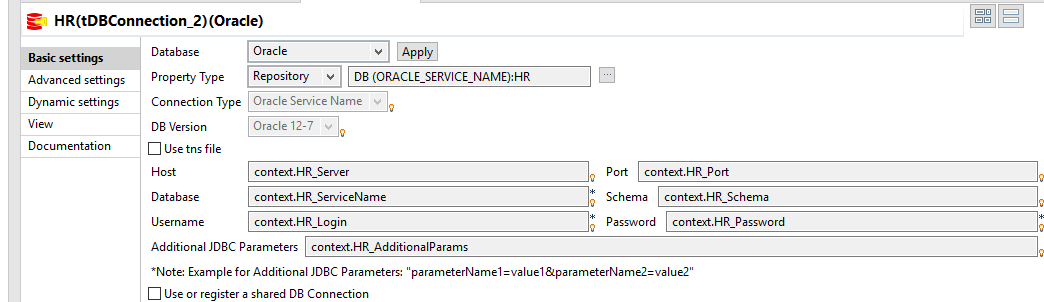
**Subjob components architecture:**



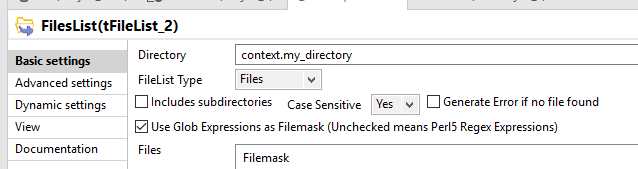
Connection to DB is provided by context group:



1. FilesList:

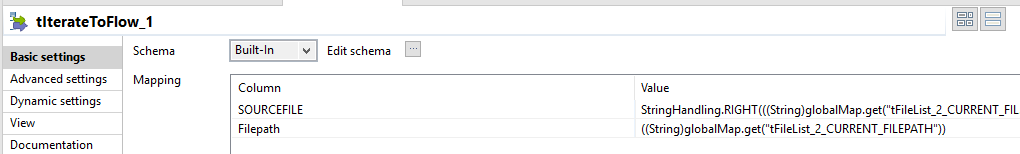
This component tFileList reads all files in specified directory.

Connection to directory is provided by context:



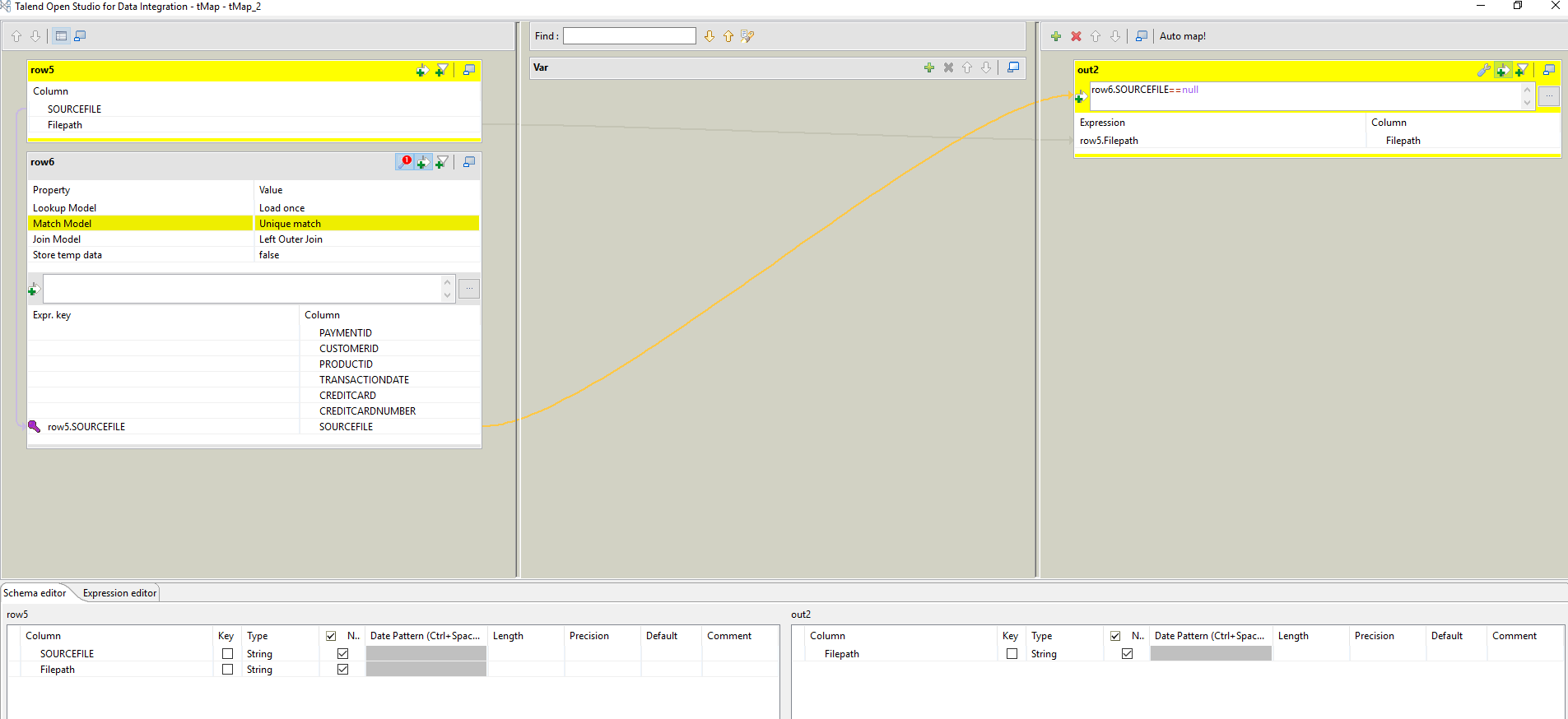
1. tIterateToFlow\_1 component transforms this list of files into dataflow.

Here I set 2 columns – SOURCEFILE (with filename and extension – this info will be upload to oracle table) and filepath (whole path to file):



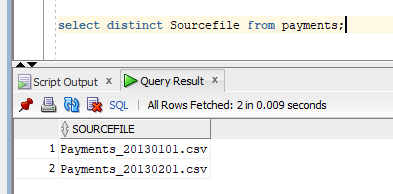
1. new\_files\_to\_load (tMap) and Payments (tOracleInput).

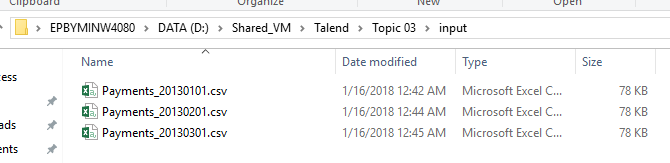
Here I make left join to the previous component data with filtering to get only null values from join:



**Test.**

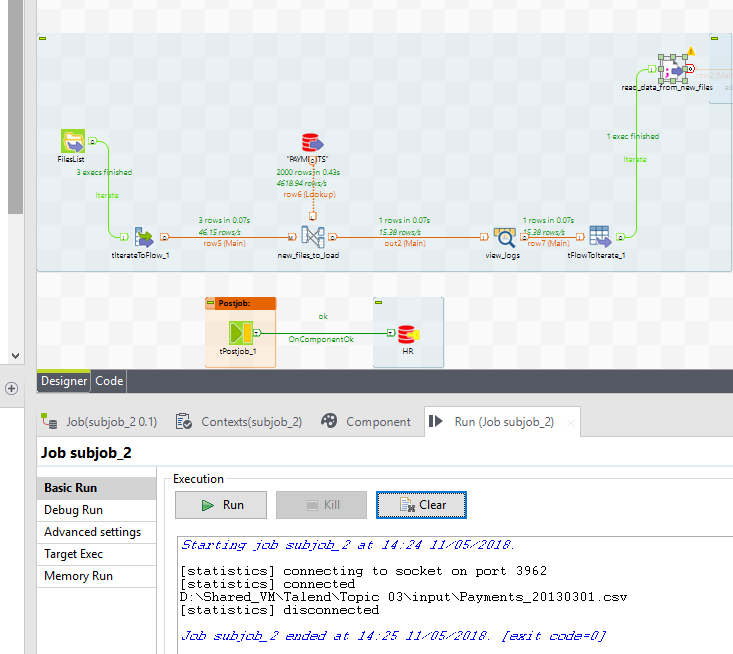
Now in Oracle table I’ve got data from 2 files uploaded:

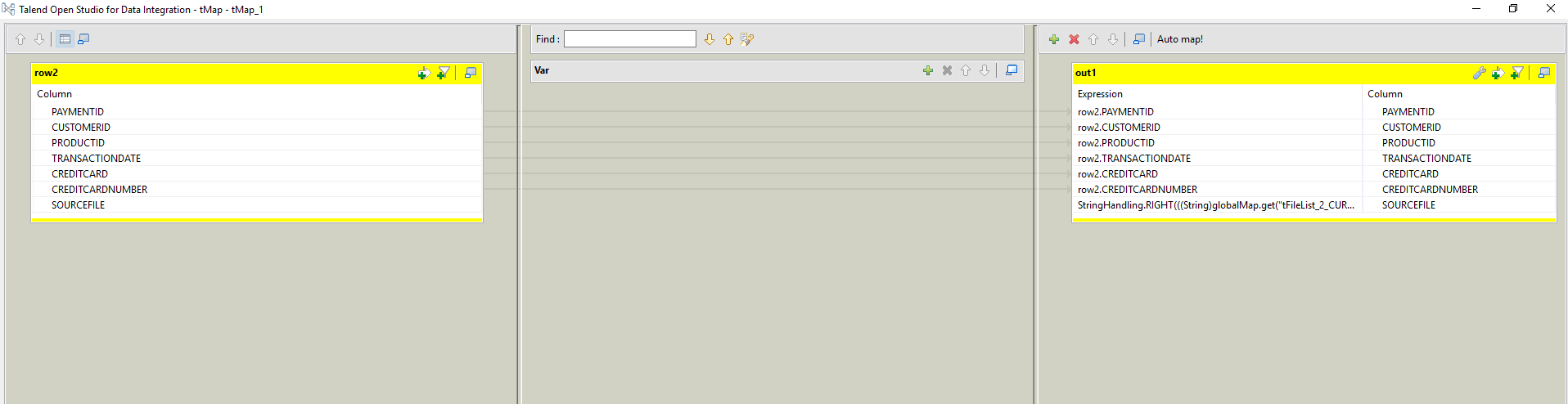


I added 3 file in the directory:  


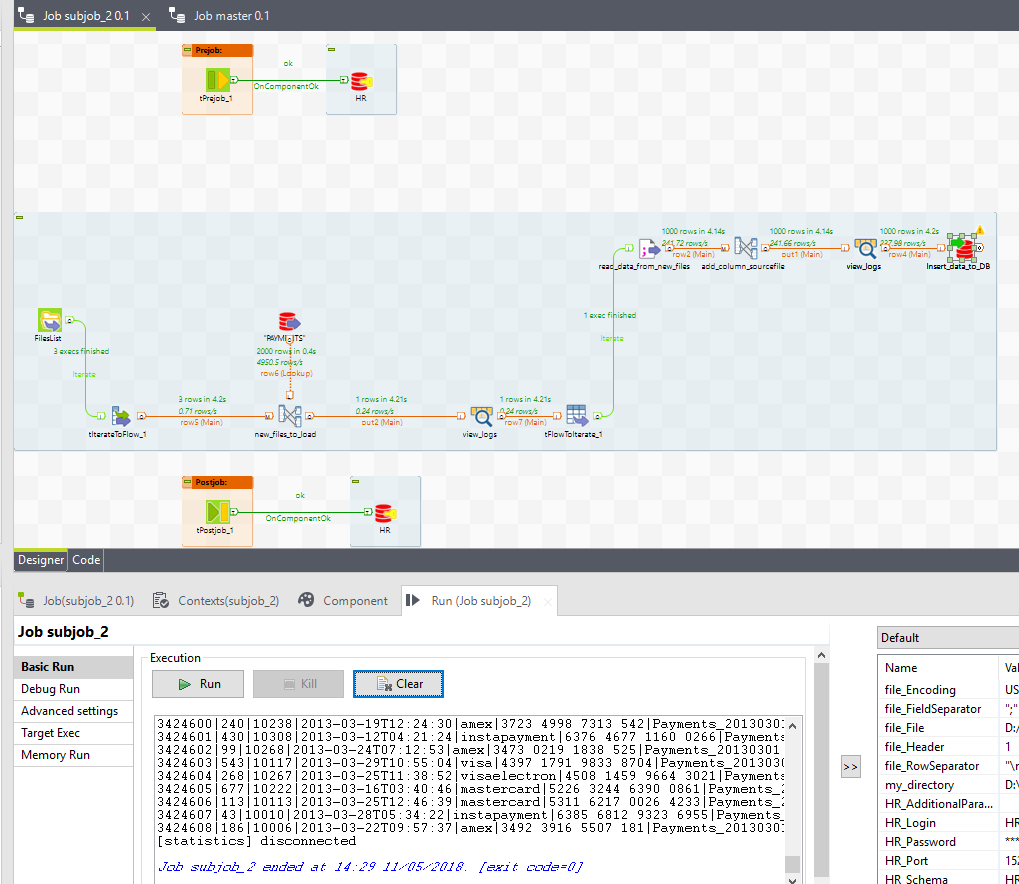
Here I run Talend job and see the expected result that 1 file is new:



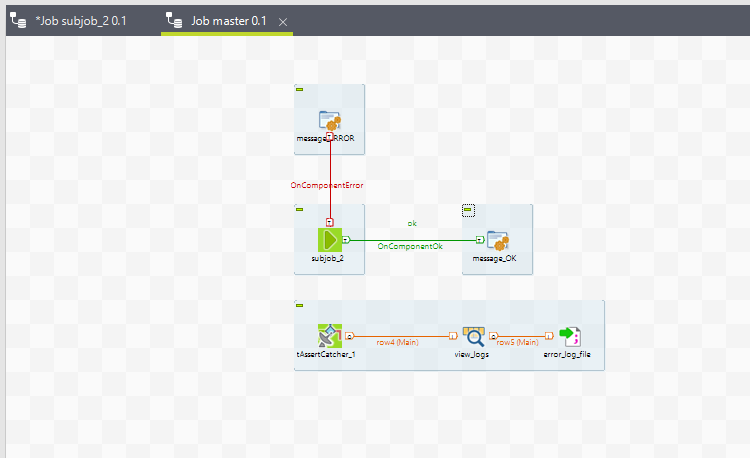
1. tFlowToIterate\_1 is used to transform filepath so that read\_data\_from\_new\_files (tFileInputDelimite) accept it as a directory and read data in file:  
    
2. add\_column\_sourcefile (tMap). Here I added new column with sourcefile name which I compare with files list in directory to upload new files only:



1. Insert\_data\_to\_DB (tOracle\_output). Here I load data from new files to oracle table:



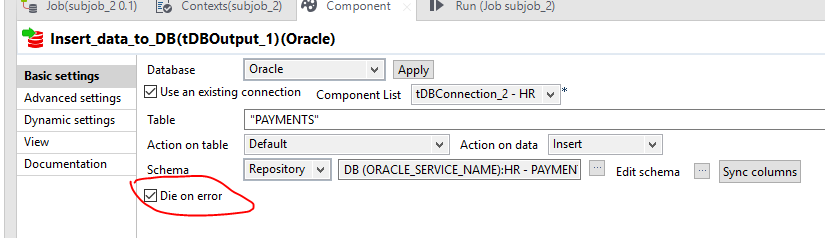
**Masterjob components architecture:**



The components of master job are pretty simple:

Here I’ve got subjob to run,

If subjob fails then I will see in logs Fail! and exceptions will be recorded to csv file using tAssertCatcher. Also I set in subjob tOracleoutput to die on error:



Here I deleted data from oracle table with sourcefile 20130301 and run masterjob to test its’ work and

subjob was launched, data was uploaded:

