

DATA SOLUTIONS DE SPIRLET

EDW DV Generator (2.4) for

SAP Sybase PowerDesigner

A DATA SOLUTIONS DE SPIRLET White Paper

Revised October 2015

All information also available on <http://www.data-solutions-de-spirlet.com>



DATA SOLUTIONS DE SPIRLET

Executive Summary

In September 2013, we release our first version of “EDW DV Generator for SAP Sybase PowerDesigner ”.

This white paper goes a bit deeper and describes the new capabilities of the latest edition of “EDW DV Generator Extension for SAP Sybase PowerDesigner ”:

- **(New 2.3)** : Data movement models gives you lineage to the field level between all physical models
- More comfort in parameters settings
- Better interface in PowerDesigner
- Editions/modes available :
 - Generator of physical data models only
 - Generator of physical data models and data movement models
 - **New (2.3)** Generator of physical data models and data movement models for SAP Data Services (as documented and supported by SAP)
 - Orchestration (synchronized) model
 - Offline service: you submit your model and we make the job for you
 - **New (2.4)** : customize generated data types for generated columns

Why such solution

Initial idea of this product was suggested by several usual questions when building data warehouse:

- How to elaborate data vault data model from classical model using automation
- How to generate all steps (extraction, staging and final data vault model) considering classical technical aspects without drowning end user
- How to generate full traceability (mapping) between all models to keep full impact and lineage powerdesigner's capabilities
- How to pre-optimize data model considering target physical database type
- How to generate complete synchronized flows between source models, extraction models, staging model and final data vault model
- How to generate fast prototyping or final models, including SQL generation for **pre-optimized** physical data models **in few minutes** while manual work takes weeks or months
- How to generate fast coding while manual work takes weeks or months



DATA SOLUTIONS DE SPIRLET

Our solution

« EDW DV Generator » is our solution permitting generating a Data Vault data model from original physical data model:

- Several PowerDesigner extension models to support analysis and automation
- Automatic model generation
 - Pre-optimized physical data models for extraction, staging and final model including mappings which permit full traceability between models and SQL generation
 - Data movement models showing all components and how to sequence data load (edition with DMM generation)
 - A final Data Movement Model integrating all steps
 - **NEW (2.2) for DMM generation edition:** support data movement model at component level permitting complete development documentation in few seconds
 - **NEW (2.3) for DMM generation edition:** support data movement model at field level permitting complete development code (with current limitations from SAP PowerDesigner itself) and generation of Data Services XML file in few seconds
 - **NEW (2.4) :** now permit customization of data types used for generated columns.

If you have a doubt on our solution capabilities, contact us at sales@dstds-it.net and we will provide you complete example.

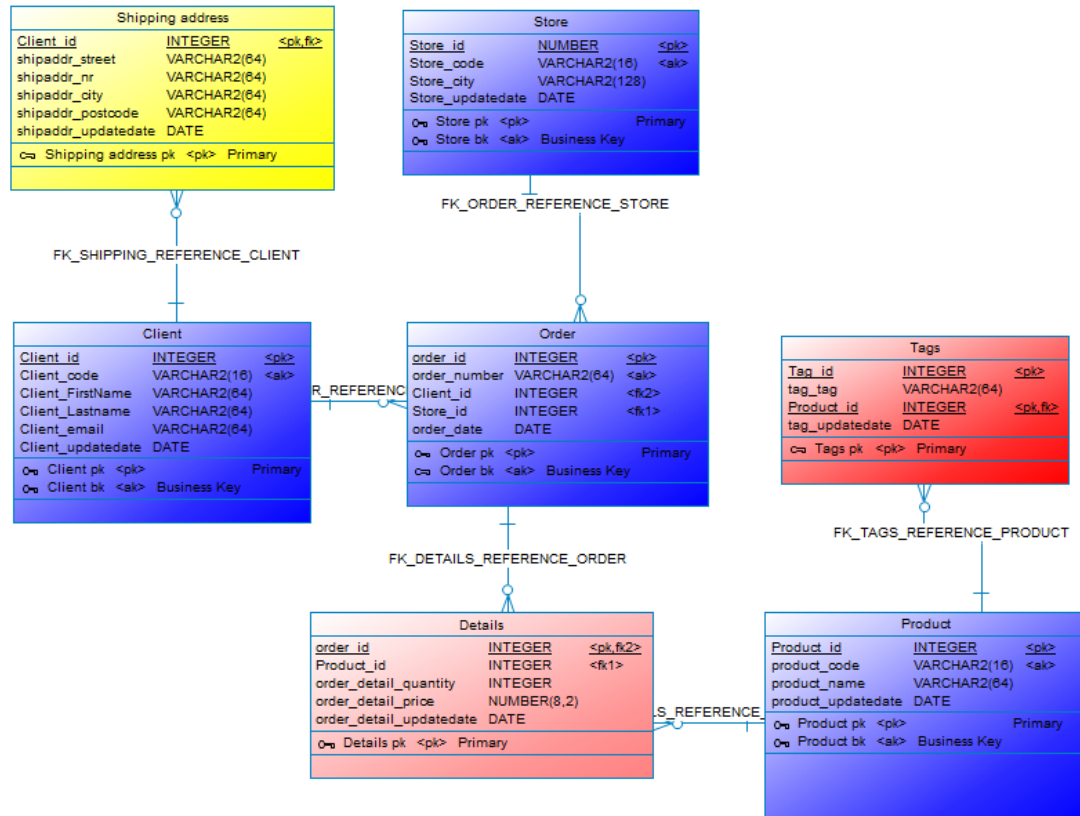


DATA SOLUTIONS DE SPIRLET

Quick description of our approach

1. Source analysis:

As usual in Business Intelligence area, all sources must be well understood from Business and Technical point of view. First choose a source model which is usually representative of an operational model (for example classical ERP/CRM or extractor model):



We provide extended definition to permit identification of Business keys which will be used during automation:

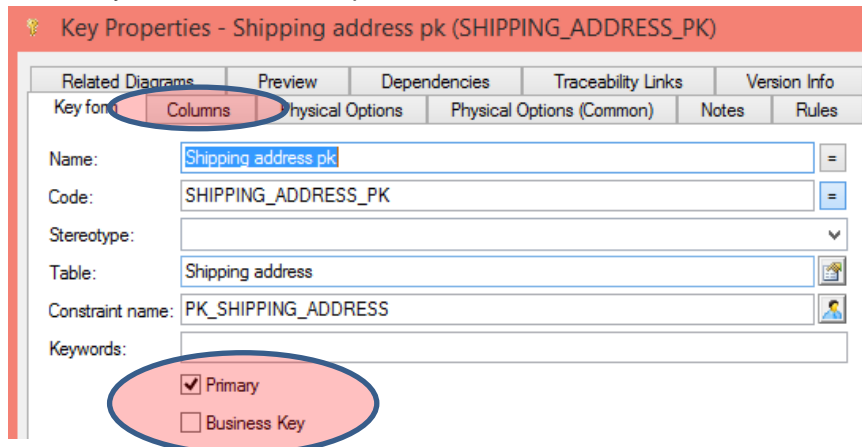
Column Properties dialog box showing the following options:

- Business Key
- LastUpdateDate
- to be deleted post generation
- To be extracted
- to be promoted as PK

DATA SOLUTIONS DE SPIRLET

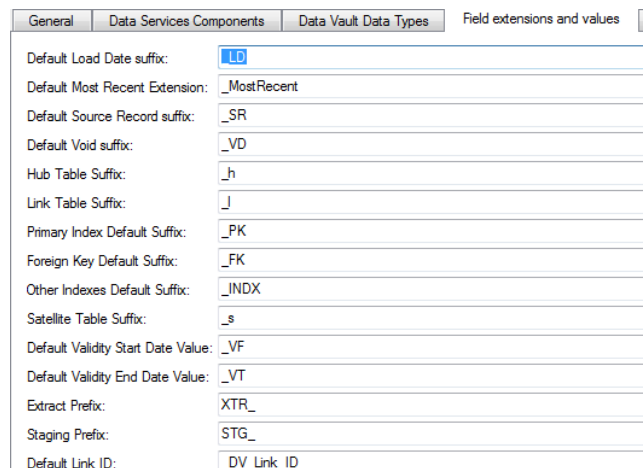
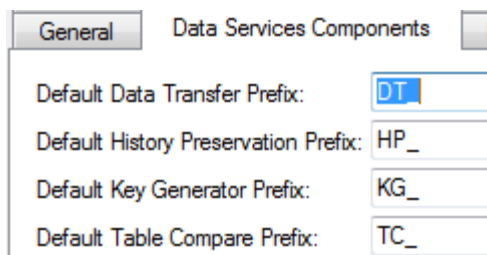
2. Complex (multi-fields) primary and business keys are identified:

In the previous editions, Business keys were single-column objects; with this new edition, business keys can be more complex even if it is not often the case:



3. Many options available at model level, in multiple forms, to customize model generation:

- Naming prefix convention for
 - extraction and staging areas
 - fields used for validity period
 - primary keys
 - hub, satellite and link tables
 - Default identifiers
 - Most recent field



DATA SOLUTIONS DE SPIRLET

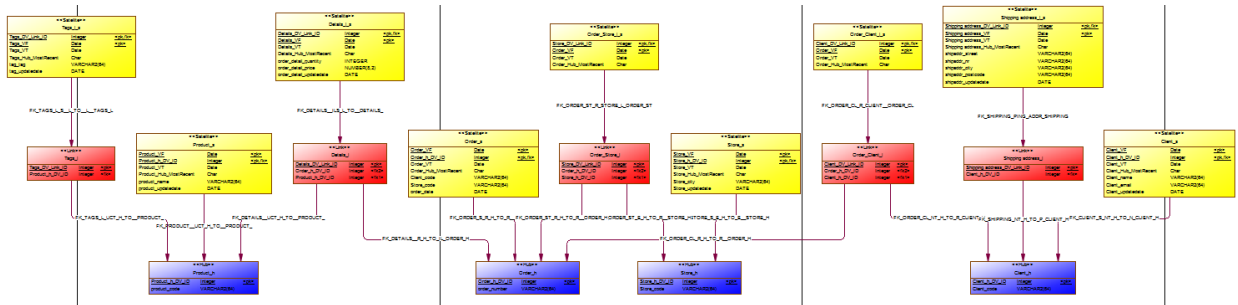
- Customization of data types used for generated fields:

General	Data Services Components	Data Vault Data Types
Most Recent Data Type:		<input type="text" value="VARCHAR2(1)"/>
Identifier Data Type:		<input type="text" value="VARCHAR2(128)"/>
Validity End Date Data Type:		<input type="text" value="DATE"/>
Validity Start Date Data Type:		<input type="text" value="DATE"/>
Void Data Type:		<input type="text" value="VARCHAR2(1)"/>
Default Source Record suffix Data Type:		<input type="text" value="VARCHAR2(255)"/>

Please note that incorrect data types may hang the generator.

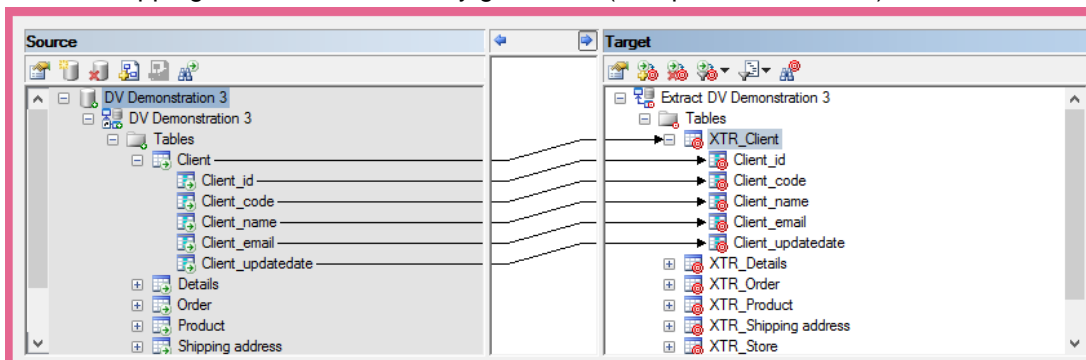
DATA SOLUTIONS DE SPIRLET

- 4. Target model:
 - a. Model Generation:

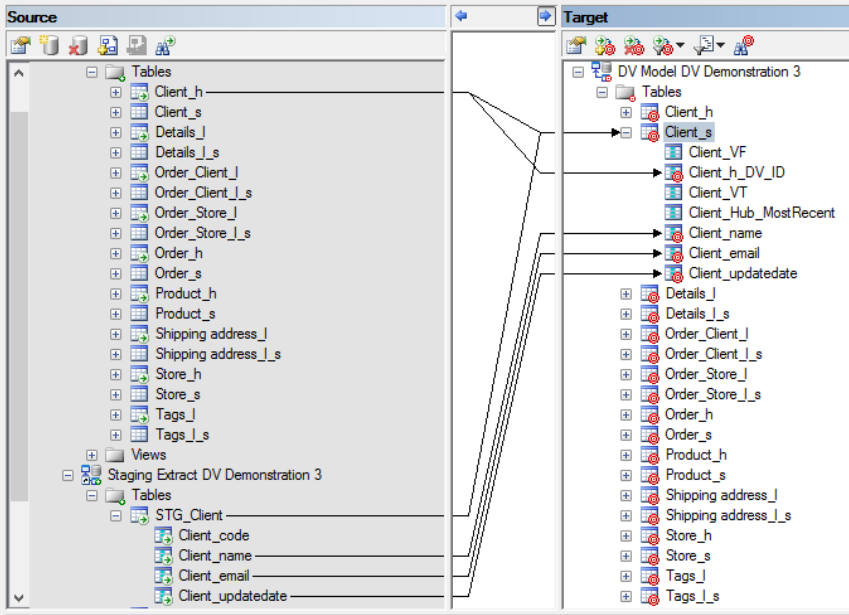


Target model will be automatically draw in a well understandable layout with color support to quickly visualize final objects:

- b. Each model is pre-optimized using specific extended model (one extended model is available for each supported database type)
- c. Mapping between tables is fully generated (except technical fields)

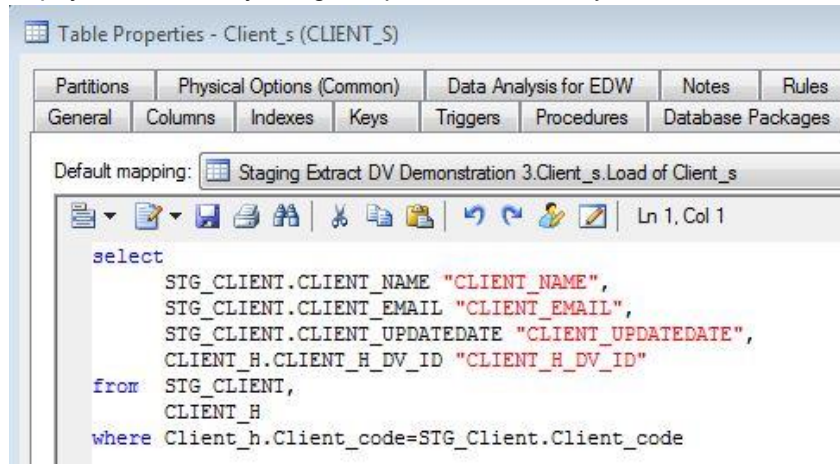


DATA SOLUTIONS DE SPIRLET

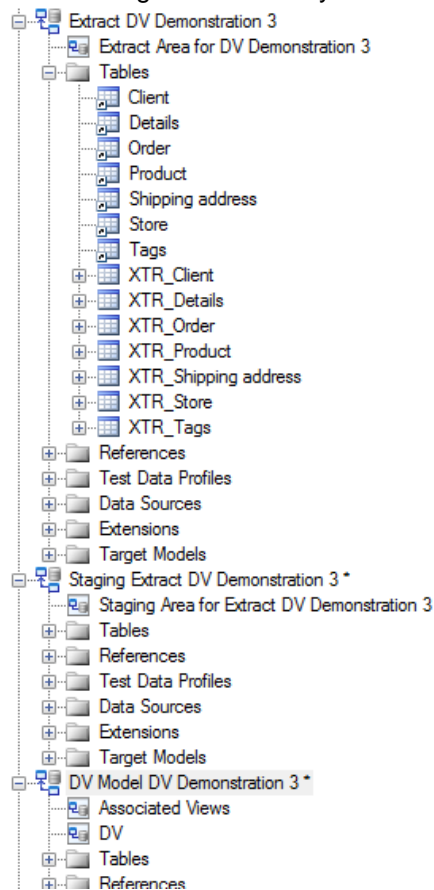


DATA SOLUTIONS DE SPIRLET

In physical models, you'll get implicit SQL to load your model:



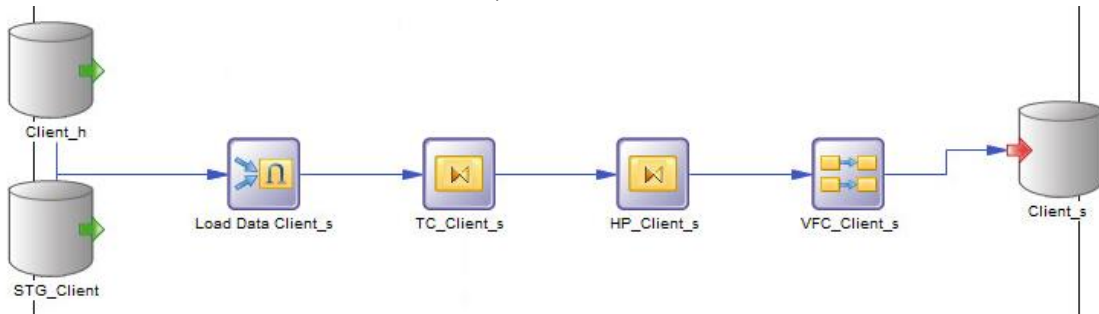
- d. Each intermediate model (extract and staging) can be directly generated from PowerDesigner without any modification



DATA SOLUTIONS DE SPIRLET

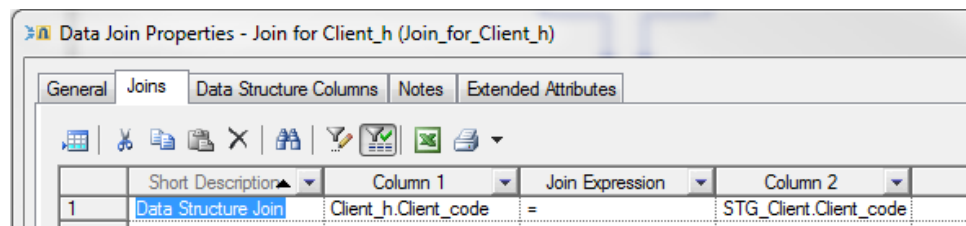
- e. For data movement models, you'll get all necessary information to describe flows to be build and detailed till database objects components (tables and views) and fields (see next):

For each target component, you'll get complete flow description (now corresponding to Data Services 4.2 extended model)

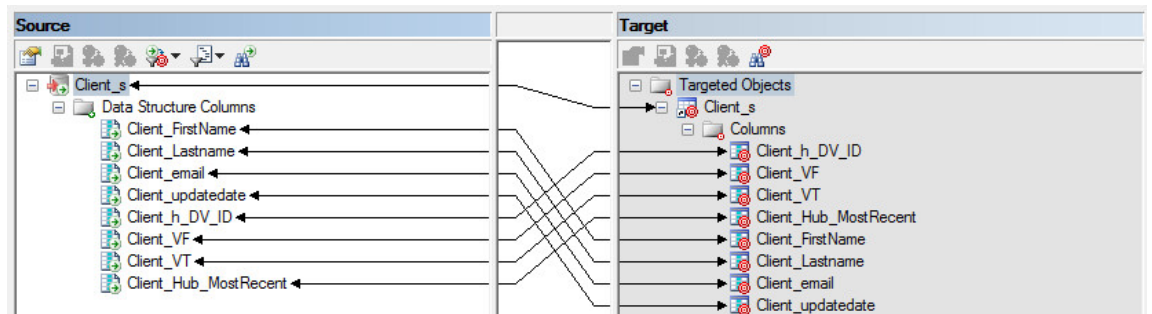


- f. **New (2.2): include complete field details**

- Join elements

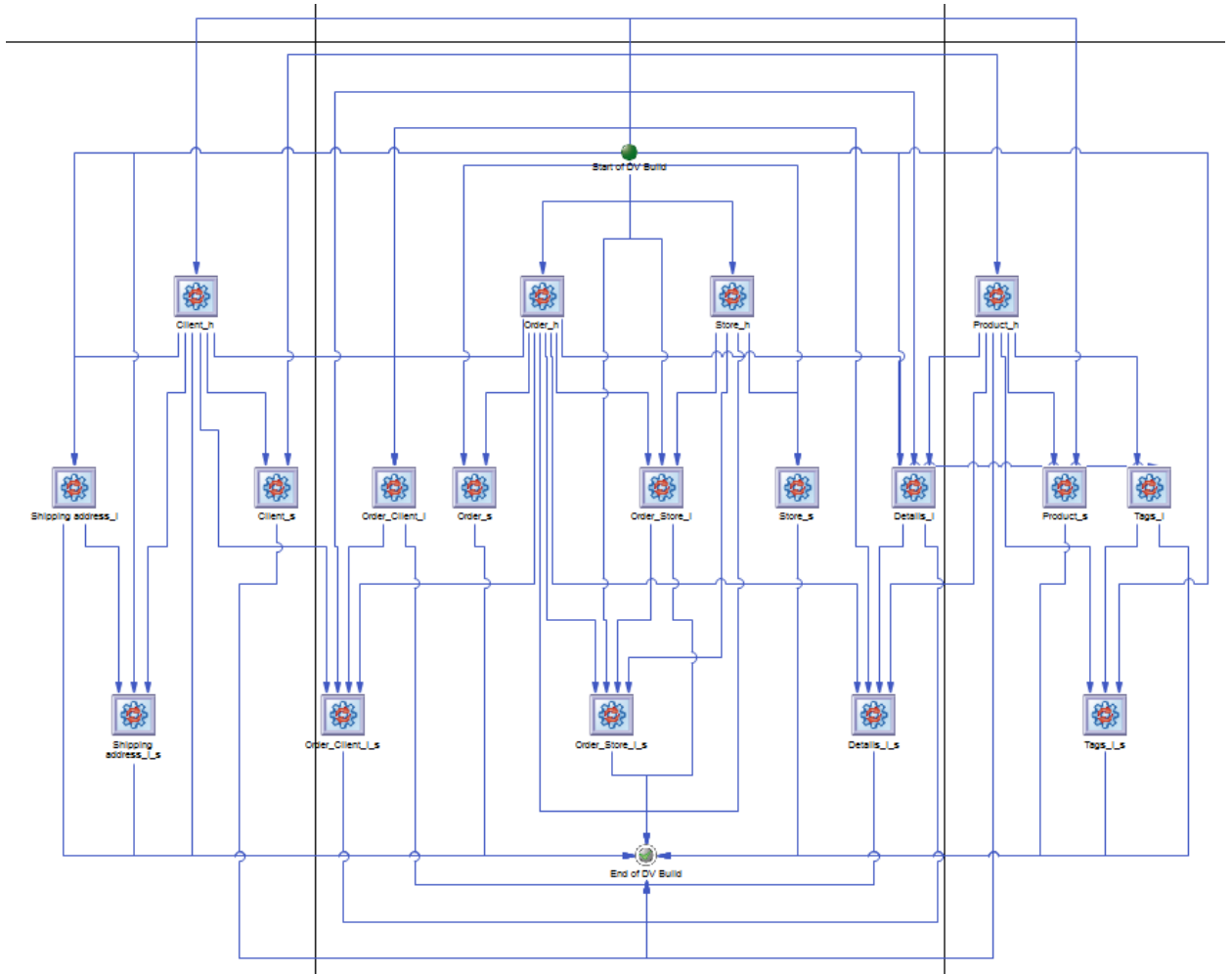


- All mappings including technical fields (new 2.3)



DATA SOLUTIONS DE SPIRLET

- g. In addition you'll get synchronization model to load a particular physical model (not for SAP Data Services 4.2)



DATA SOLUTIONS DE SPIRLET

In summary

	<i>Classical working</i>	<i>Our Solution</i>
<i>Time to design</i>	Classically weeks or months	Immediate
<i>Production of models</i>	Manual	Fully automated (after Business Model)
<i>Time to produce physical models</i>	Classically weeks or months	Models are generated in <u>minutes</u>
<i>Linage between physical models</i>	Often inexistent or manually defined	Fully automated
<i>Production of data movement models</i>	Often absent or made in other tools (visio, powerpoint) ...)	Fully integrated and provide information for external scheduler
<i>Time to produce data movement model</i>	Days or weeks	Produced in minutes
<i>Linage between physical models</i>	Often inexistent or manually defined	Fully automated
<i>Data Service Code</i>	Depends on the model size but can be several man months	Fully automated (*)

(*) we provide a "basic" generation but we can also provide on-site services in order to customize our product to your needs.