

# Record Tracking Satellites

Book Sections Reference: 5.3.5 & 12.1.8

Record Tracking Satellites provide the ability to track when business-keys were last seen provided by the source. This is required especially in the context when some system limitation is causing business keys to sometimes disappear, reappear. Record tracking satellites can be used to also highlight the combination of business keys, i.e. they can exist on Link tables as well. However, for Link tables we would propose you consider a Effectivity satellites, as it is probably more suited.

## WhereScape 3D

### Dependent objects

### Sample Repo

- Look at the examples created in 3D repo **wsDVSamples.repo**

### Model Conversion Rules

- 1 additional rule for the generation of Data Vault (**ws3d\_grv - Create Record tracking satellites**)
- 1 additional rule for generation of Load and Stage (**ws3d\_rvls - Record tracking satellite - staging**)

Copy from wsDVSamples.repo or import **RecordTracking\_MCR.xml**

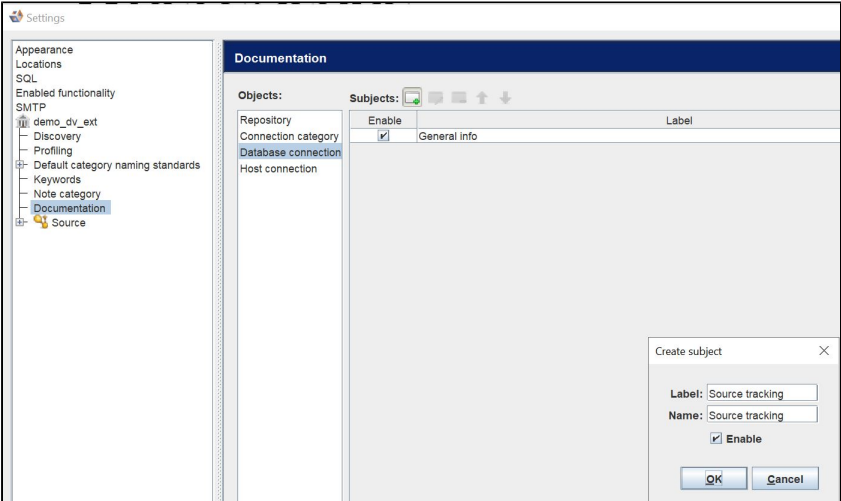
### Template

- 1 additional Query generation template used by the Load and Stage (**ws3d\_dv\_record\_tracking**)

Copy from wsDVSamples.repo or import **RecordTracking\_TEMPLATES.xml**

### Steps to implement

If a particular stage is not mentioned, then you can assume it would follow the standard process of generating a Data Vault.

#	Category	Description
1	Global settings	<p>Record-tracking is enabled on connection level</p> <p>You have to add a documentation subject on "Database connection" - level named "Source tracking"</p> 

2 Connection level

Then set insert into the desired connection an "Y" for enabled - anything else for disabled

The screenshot shows a tree view on the left with 'Connection' expanded to 'Southwind'. A 'Document Southwind' dialog box is open, showing the 'Source tracking' tab. Under 'Source tracking:', the value 'Y' is entered.

3 Data Vault

Add Attribute-Type "dss\_tracking\_flag" to Data Vault

The screenshot shows the 'Attribute types' section in the Data Vault design tool. A table lists various attribute types and their descriptions. The attribute 'dss\_tracking\_flag' is highlighted in yellow.

Attribute Type	Description
Change hash key	A change hash column (low)
Multi-active satellite sequence attribute	The sequence column for M...
Hub surrogate key	A Hub surrogate key column
dss_load_date	A dss column for load date
dss_start_date	A dss column for start date
dss_end_date	A dss column for end date
dss_record_source	A dss column for record sour
dss_update_time	A dss column for last update
dss_create_time	A dss column for create time
dss_version	A dss column for version num
dss_current_flag	A dss column for indicating c
To be deleted	A column that will be remove
Satellite low volatility	To create a Satellite with low
Satellite medium volatility	To create a Satellite with me
Satellite high volatility	To create a Satellite with hig
Satellite transaction	To create a Satellite with tra
Link business key	Foreign key attributes that m
dss_tracking_flag	dss_tracking_flag

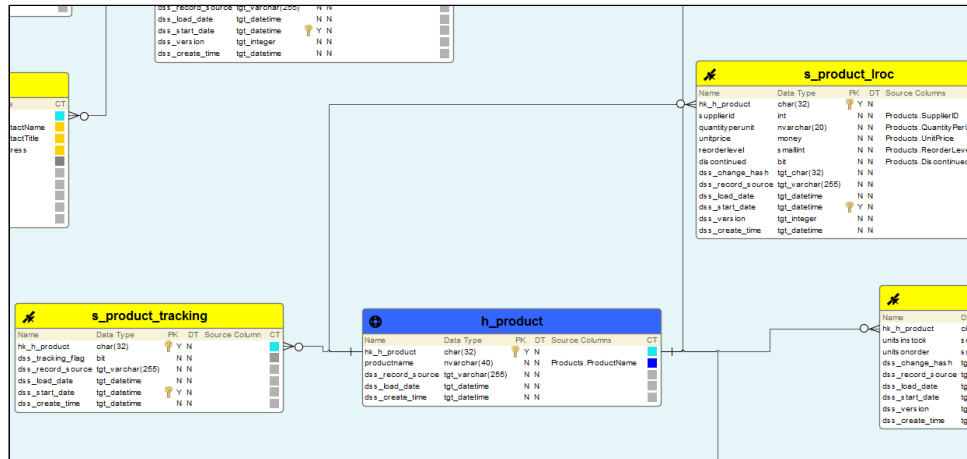
Generate from the Data Vault Design Model with the standard-process enhanced with the following Rules

The screenshot shows the 'Generate data vault' dialog box. Under 'Apply model conversion', four rules are selected in a list:

- ws3d\_grv - Logical to DV Spine (Template)
- ws3d\_grv - Create satellites
- ws3d\_grv - Create Record tracking satellites (Template)
- ws3d\_grv - Hash Key generation (Template)

The 'Create snapshots' checkbox is unchecked. The dialog has '< Back', 'Next >', and 'Cancel' buttons at the bottom.

In the Data Vault Model you should notice that for every hub and link that receives data from the rts-active connection a satellite "\_tracking" has been created:



Things to note:

- The dss\_tracking\_flg indicates with 0 (not seen) or 1 (seen) that with the last loading this key has been seen or not.
- There shouldn't be any other columns (except for the standard-dates, recordsource and the key itself)

4

Load And Stage

Add Attribute-Type "dss\_tracking\_flag" to Load and Stage

Documentation	Change mashi key	A change mashi column (p...
Data vault	Multi-active satellite sequence attribute	The sequence column for M...
General	Hub surrogate key	A Hub surrogate key column
Version	dss_load_date	A dss column for load date
Connection	dss_start_date	A dss column for start date
Discovery	dss_end_date	A dss column for end date
Profiling	dss_record_source	A dss column for record sour
Entity	dss_update_time	A dss column for last update
Attribute	dss_create_time	A dss column for create time
General	dss_version	A dss column for version num
Attribute types	dss_current_flag	A dss column for indicating c
Attribute ratings	To be deleted	A column that will be remov
Data characteristics	Satellite low volatility	To create a Satellite with low
Relationship	Satellite medium volatility	To create a Satellite with med
Constraint	Satellite high volatility	To create a Satellite with high
Index	Satellite transaction	To create a Satellite with tra
Import	Link business key	Foreign key attributes that m
Storage	dss_tracking_flag	dss_tracking_flag
Source Mapping		
Creation		

5 Generate from the Data Vault Design Model with the standard-process enhanced with the following Rules:

Generate load and stage for data vault

Apply model conversion

Select a model conversion

Model conversion:

- ws3d\_rvls - Create initial stages (Template) x
- ws3d\_rvls - Define change hashes on satellite stages (Template) x
- ws3d\_rvls - Define extended properties for hub and link hash key (Template) x
- ws3d\_rvls - Merge and clean up stages (Template) x
- ws3d\_rvls - Create loads (Template) x
- ws3d\_rvls - Record tracking satellite - staging (Template) x
- ws3d\_rvls - Housekeeping (Template) x

Things to note:

- There will be no additional stage for the tracking-satellites as they use the staging from the hub or link.
- The DW-Query has already been created and will be exported to RED - that is why there should be 3D-templates for different target-DWH.
- The DW-Query and the column-transformation will determine the query that will be used in the final script.

# WhereScape RED

## Dependent objects

## Templates

- dv\_perm templates need additional logic to remove the Current Satellite version join in the case of a tracking satellite

## Steps to implement

Look at the sample template **cust\_sqlserver\_proc\_dv\_perm.peb** for an example of what changes are required.

#	Description of change	Snippet Code
1	The dv_perm template for your target platform needs to be adjusted in order to use record tracking satellites	<pre>{%- set isTrackingSatellite = false -%} {%- from table.columns as col where col.name == "dss_tracking_flag" -%} {%} {%- set isTrackingSatellite = true -%} {%- endfrom -%}</pre>
2	This variable should then be used to NOT <b>addSatCurrentVersion</b>	<pre>{%- elseif table.objectType == Types.ObjectType.Satellite and not (isTrackingSatellite or isEffectivitySatellite) %} --&gt;addSatCurrentVersion&lt;-- {{addSatCurrentVersion()}}</pre>
3	This variable should then be used to NOT <b>addSatWhereNotExists</b>	<pre>{%- elseif table.objectType == Types.ObjectType.Satellite and not (isTrackingSatellite or isEffectivitySatellite) %} --&gt;addSatWhereNotExists&lt;-- {{addSatWhereNotExists()}}</pre>

# Effectivity Satellites

Book Sections Reference: 5.3.4: "Its purpose is to track when the **link** is active according to the business and provides begin and end dates for this purposes"

See also: "The Data Vault Guru: A pragmatic guide on building a data vault" Chapter 6.3 Effectivity Satellite (Page 301)

As part of the enablement pack WhereScape ships the functionality to use the "Applied Date" to determine the effectivity of the link. You can choose whether you want a default system generated date to populate "Applied Date" or whether you want to choose no more than 1 date from the source.

In cases where multiple dates or fields from the source describes the temporality of the data we provide an option for a different approach, and not to resolve it during load of Raw Vault, but rather further downstream. See details under our Multi-Temporal sources enablement pack.

## WhereScape 3D

### Dependent objects

### Sample Repo

- Look at the examples created in 3D repo **wsDVSamples**

### Model Conversion Rules

- 1 of 3 generation of Data Vault rule (**ws3d\_grv - Create effectivity satellites**)
- 1 of 3 generation of Data Vault rule (**ws3d\_grv - Prepare effectivity satellites**)
- 1 of 3 generation of Data Vault rule (**ws3d\_grv - Housekeeping effectivity satellites**)
- 1 additional generation of load and stage (**ws3d\_rvls - Create effectivity satellites**)

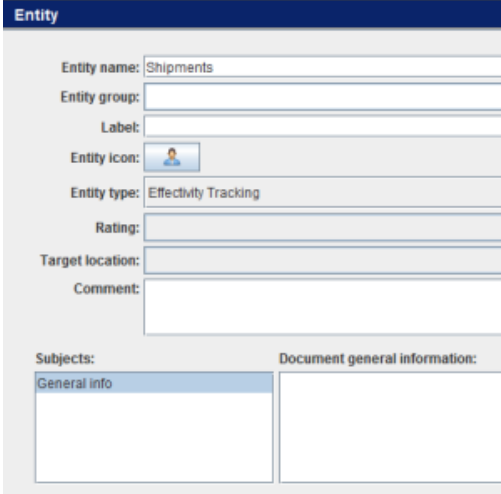
Copy from wsDVSamples or import **EffectivityTracking\_MCR.xml**

### Template

No additional templates are required

### Steps to implement

If a particular stage is not mentioned, then you can assume it would follow the standard process of generating a Data Vault.

#	Category	Description
1	Data Vault Design	<p>For the purpose of this document, the Shipments-table has been chosen to create an Effectivity Satellite upon. In this scenario the OrderID shall be the DrivingKey (dss_driving_key) and the ShippedDate has been declared as Applied Date (dss_applied_date) - which is purely for the example purposes.</p> <p>The entity containing the link business keys has been set to the entity-type "Effectivity Tracking"</p> 

Attributes

Attribute name:

Column name: ShippedDate

Created by: User defined

Label:

Data type: datetime

Size:

Scale:

Attribute type: Applied Date

Rating:

Nullable:

Unique:

Primary key:

Auto increment:

Default value:

Comment: The date that the order was shipped from our factory

Hidden:

Attributes

Attribute name:

Column name: OrderID

Created by: Import from external format

Label:

Data type: int

Size:

Scale:

Attribute type: Link business key; Driving key

Rating:

Nullable:

Unique:

Primary key:

Auto increment:

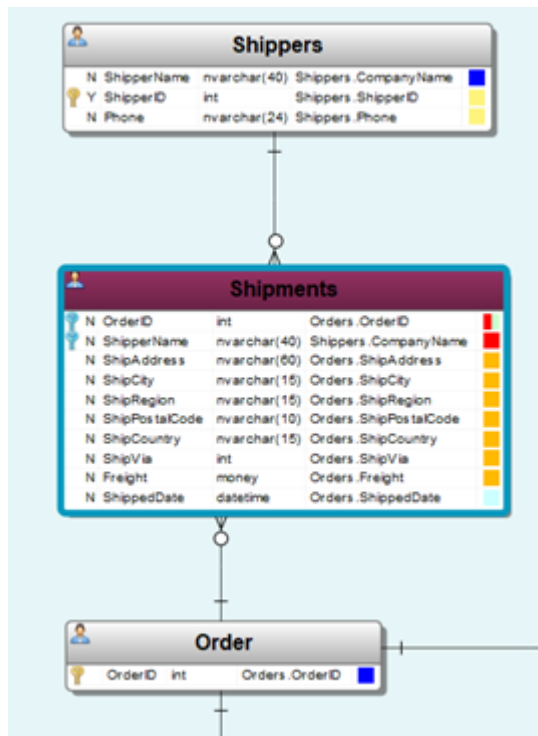
Default value:

Comment: This is the internal order id

Hidden:

Subjects:

Document general information:

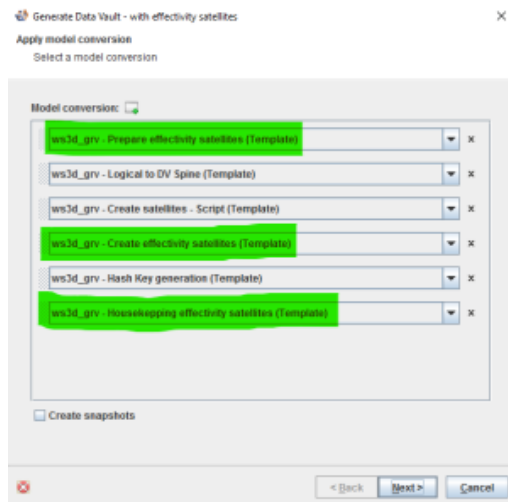


Things to note:

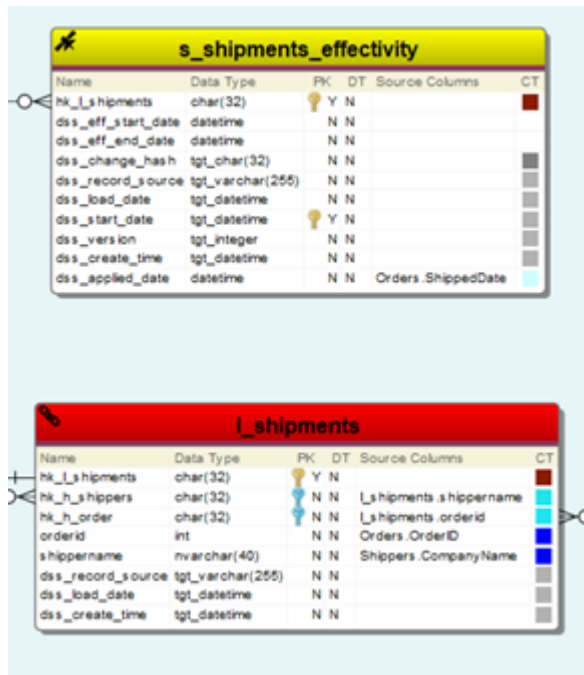
- Attribute types that are being used are:
  - Driving Key (light green)
  - Applied Date (light blue)
  - Entity-Type "Effectivity Tracking" (burgundy)
- The single source of each attribute is shown in the diagram
- Using the dss\_applied\_date is optional - if not used, the dss\_create\_time will be used as source

2 **Data Vault Model**

Generate from the Data Vault Design Model with the standard-process enhanced with the following Rules:




The Data Vault Model should look like this



Things to note:

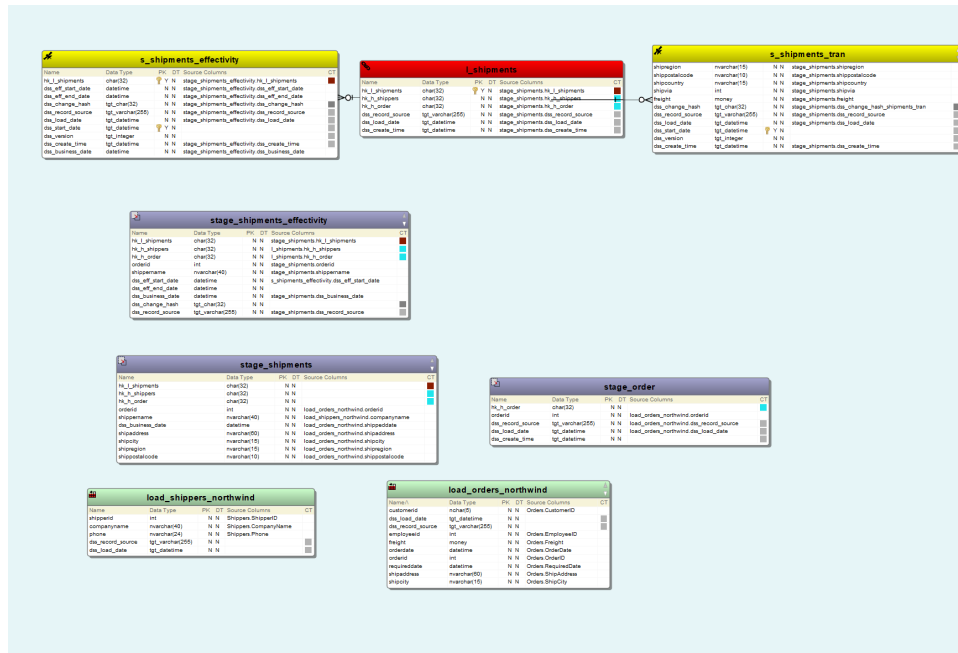
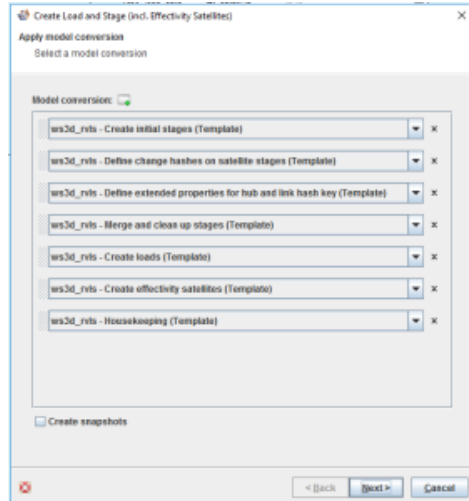
- All of the extra DV control columns have been added to the entities as per DV 2.0 rules
- Hub and Change hash keys have been added where needed.
- The setting of the attribute "Driving Key" has been copied into the comment of the newly created satellite.

Table	
Table name:	s_shipments_effectivity
Schema:	
Database:	
Label:	shipments
Table icon:	
Table type:	Satellite, Effectivity Tracking
Rating:	
Target location:	
Comment:	DrivingKey=>["orderid"]<=>DrivingKey
Subjects:	Document purpose:
<ul style="list-style-type: none"> <li>Purpose</li> <li>Grain</li> <li>Examples</li> <li>Usage</li> <li>Observations</li> </ul>	



3 **Load and Staging Model**

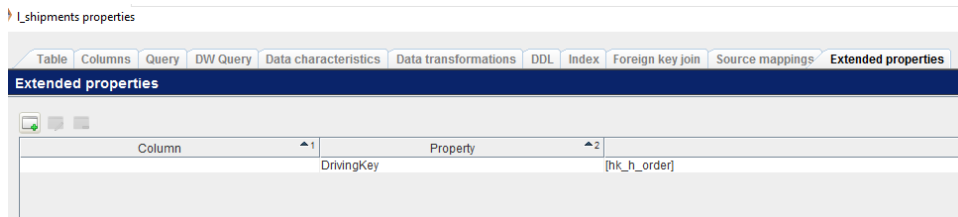
Generate from the Data Vault Design Model with the standard-process enhanced with the following Rules:



Things to note:

- There should be an additional stage for every effective satellite which is sourced mainly from the original stage for the link - but also from the link and the effectivity-satellite itself

The setting "Driving Key" from the comment has now moved to the extended Properties:



4 **Export to RED**

Export the RED Export model version to RED using the Basic Process described previously. The model should deploy to RED without any issues and all Load, Stage and Data Vault objects can be loaded successfully.

# WhereScape RED

\*Currently only target SQL Server is supported

## Dependent objects

### Templates

3 new templates need to be used or the logic incorporated into existing templates

- cust\_sqlserver\_proc\_dv\_stage (Procedure template for Data Vault Stage)
- cust\_sqlserver\_proc\_dv\_stage\_effectivity\_utility
- cust\_sqlserver\_utility\_dv

## Steps to implement

### Via application installation

#	Description of change
1	Unzip file app_tem_SQLDV_EFSAT_202210040854
2	Install application DV_SQL_EFSAT from wsl files in the zipped folder

### Via copying the content from template files

The 3 templates have also been attached separately to look at.

# Multi-Temporal Sources

Source systems often do not store only a single version of the data. There are multiple fields that describe the temporality of the source data. This could be technical dates, business effective dates and even CDC dates (or a combination of all). This leaves a question of how to adequately load the data into the Raw Vault. An option is simply to add all of the fields describing the temporality as additional metadata on the satellites. Then the correct effectivity can be determined downstream as this would consist of some business logic.

## WhereScape 3D

### Dependent objects

#### Sample Repo

- Look at the examples created in 3D repo **wsDVSamples**

#### Model Conversion Rules

- Group Temporal Metadata from Source
  - ws3d\_grv - Add Temporal Metadata to all Satellites
  - ws3d\_rvls - Remove Custom Attribute types

Copy from wsDVSamples or import **Temporal\_Source\_additional\_fields.xml**

#### Template

No additional templates required

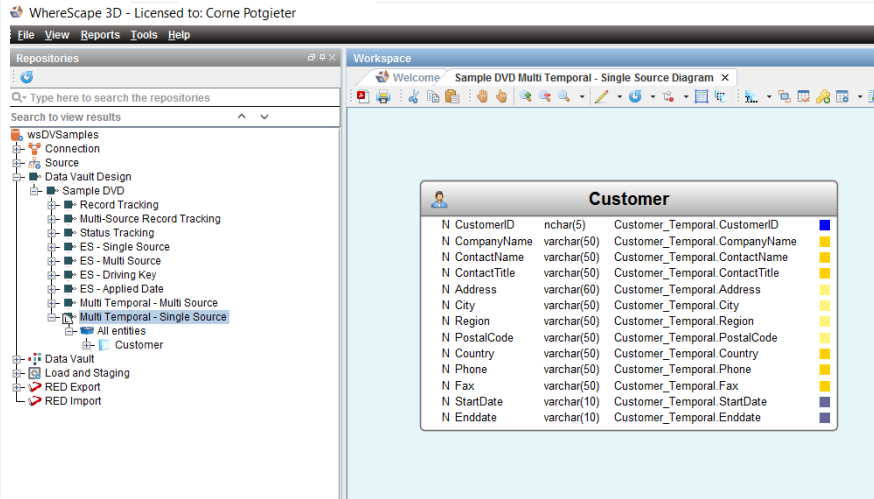
### Steps to implement

If a particular stage is not mentioned, then you can assume it would follow the standard process of generating a Data Vault.

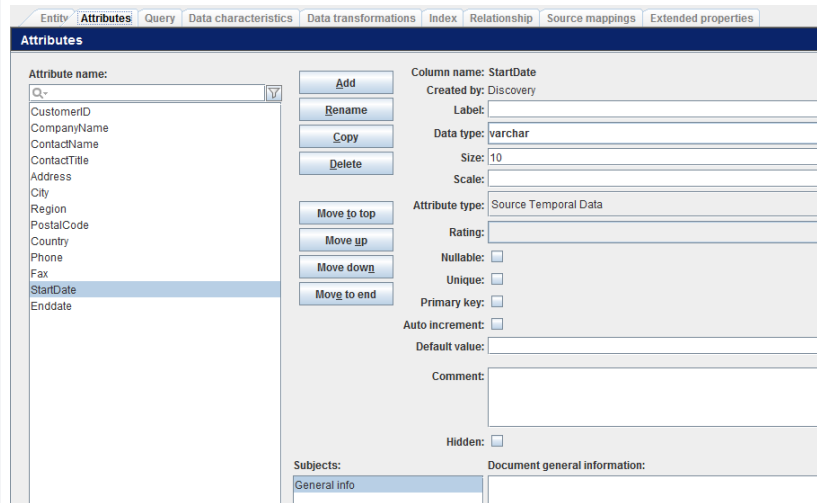
#	Category	Description
---	----------	-------------

1 Data Vault Design

See these multi temporal examples in the Samples repo

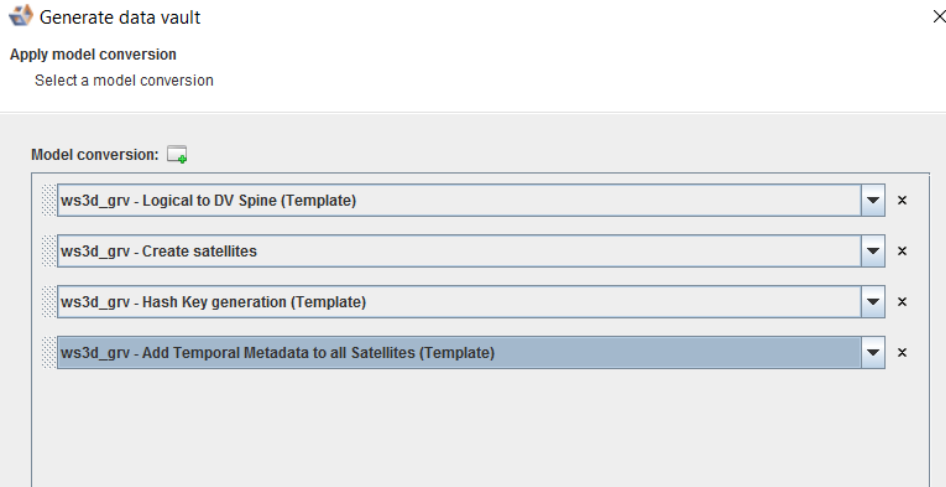


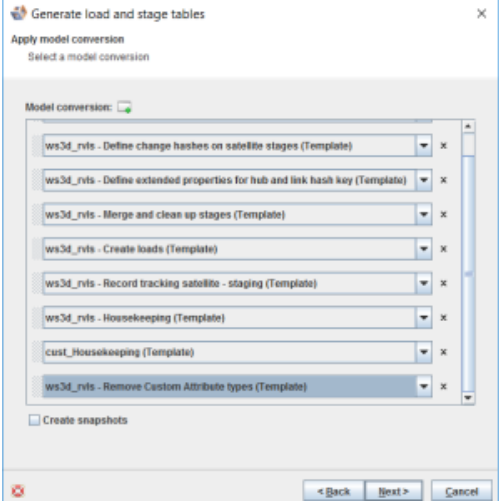
Tag the source fields that describe the temporality or effectivity as "Source Temporal Data"



2 Generating Data Vault

When generating the Data Vault, this additional MCR (ws3d\_grv - Add Temporal Metadata to all Satellites) should be used to copy the fields to all relevant satellites



3	Generating Load and Stage	<p>During the generation of Load and Stage the model conversion rule "ws3d_rvls - Remove Custom Attribute types" should be applied at the end</p>  <p>Prerequisite: Multi-Active Satellite Natural Key Attribute type needs to be an attribute type for Load and Stage. We are reusing this field for deploying to RED to allow for multiple versions of the data to be active based on the source temporality</p>
---	---------------------------	---

## WhereScape RED

No prerequisites are required for RED. The templates should handle the inserts into the Satellites if there is a Multi-active Natural key present. The loads will be aware of the additional keys and will allow multiple versions of the same business key.