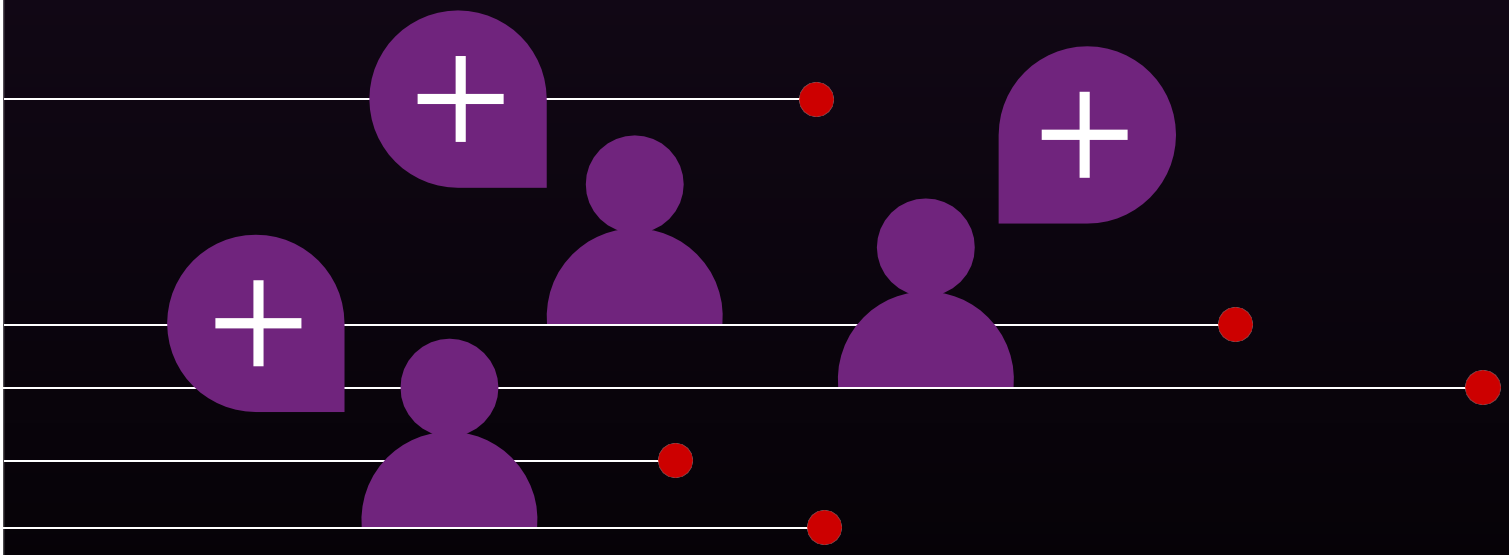
VAULTSPEED

VaultSpeed Studio

SCD2 dimension on a BRIDGE



USE CASE FOR THIS TEMPLATE

This document describes how a user can configure a template in the VaultSpeed Studio to generate an SCD Type 2 dimension based on an insert-only RDV by using a bridge table with multiple HUB's. In the setup there is the possibility to limit the fields which should show up in the dimension. The generated query will also remove unnecessary records (no changes in the combination of selected fields).

Before you explore and use this example template, ensure that you understand the example SQL attached to understand what the template does and that it covers your needs.

This template is designed for:

- DIMENSION creation based on a BRIDGE
- The BRIDGE represents 1 hierarchy in a dimension with max 5 levels (eg product → product type → product segment). With the outer ends of the bridge being the lowest and highest level of that hierarchy
- Slowly Changing Type 2 Dimension
- Versions in the dimension will be compressed
- PIT tables must exist on EACH entry table on the BRIDGE
- MAIN_HUB must be defined
- Only the change date of the indicated LNK tables with the SO DIM_ON_BRIDGE_LNK_TIME_SELECTION will be used in defining the SCD2 time slices
- Dimension_hkey will be calculated based on the BK's of the MAIN_HUB in combination with the date

Example

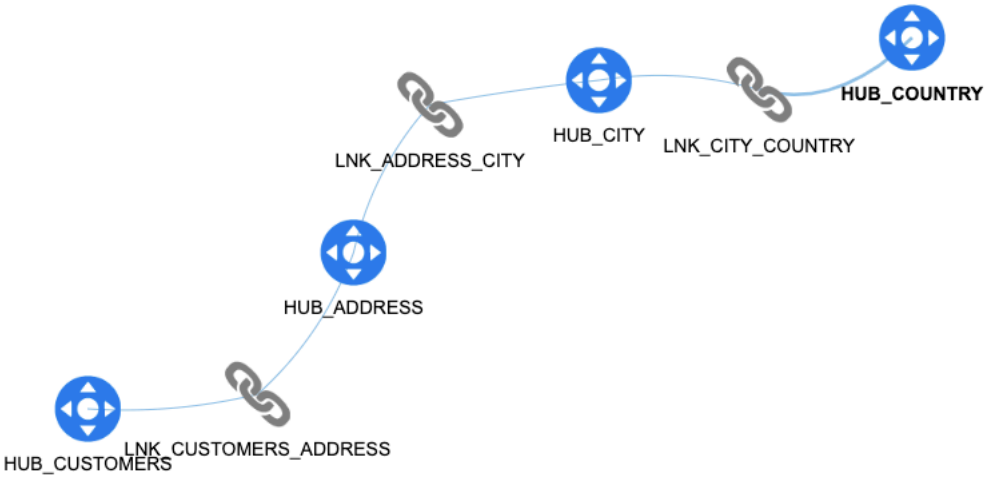
The example that is used in this document is based on a customer and its address information. In the Raw Data Vault model this is a HUB_CUSTOMER with a Link to the HUB_ADDRESSES, country and city.

Components of the implementation

- Standard business vault BRIDGE table
- PIT table on all HUB and LNK objects that exist in the BRIDGE
- Signature objects
- Assign Signature objects to the correct tables
- Create Signature attribute type and assigned in the requested fields
- Create the Template
- Create Target definition
- Fill in the Dependency

Implement a Standard business vault BRIDGE table

In the open release of the Business Vault, create the bridge table. For use within the template, this bridge doesn't need to have the hash key created or the Business keys of the HUBs added.



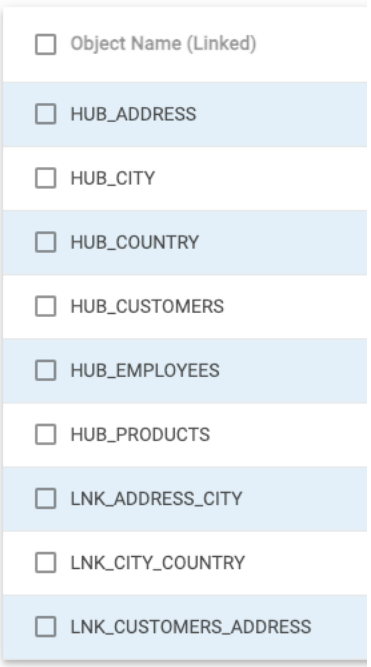
PIT table is created on all HUB and LNK objects that exist in the BRIDGE.

The Dimension is a SCD2 type dimension, for this we need to have a detailed PIT implementation on all HUBs and LNKs that exist in the BRIDGE that has been created.

The PIT setup is like the following:

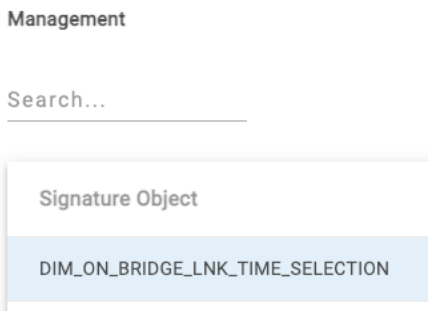
PIT Name	PIT Type ↑	Snapshot Interval	Interval Unit	Timestamp Signature Attribute
DETAIL_TRANS	detail			TRANS_TIMESTAMP

After that, the PIT is applied on the HUBs and LNKs of the example:


Object Name (Linked)
HUB_ADDRESS
HUB_CITY
HUB_COUNTRY
HUB_CUSTOMERS
HUB_EMPLOYEES
HUB_PRODUCTS
LNK_ADDRESS_CITY
LNK_CITY_COUNTRY
LNK_CUSTOMERS_ADDRESS

Create Signature objects

In order for the template to know which tables to use for which purpose we need to create some specific Signature Objects.



Signature objects are assigned

MAIN_HUB Indicate the driving table of the dimension. Meaning the lowest level of your multi-level dimension setup

Object Name	Object Type	Signature Object
Filter...	Filter...	Filter... MAIN_HUB
HUB_CUSTOMERS	Hubs	HUB, MAIN_HUB

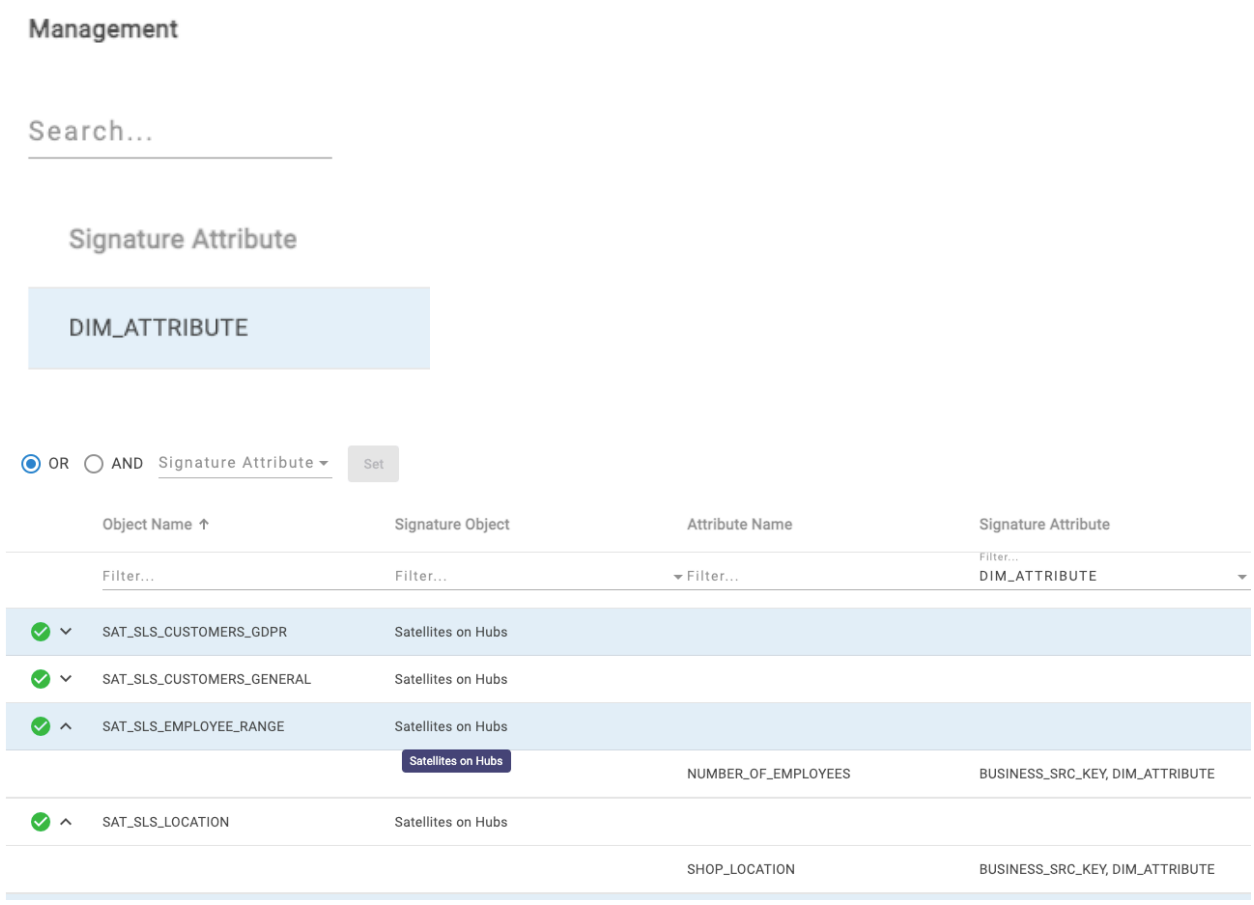
PIT_DETAIL_TRANS Indicating which PIT tables must be used in the setup of the date ranges

PIT_DETAIL_TRANS_CUSTOMERS	Point in Time table	PIT, PIT_DETAIL_TRANS
PIT_DETAIL_TRANS_ADDRESS	Point in Time table	PIT, PIT_DETAIL_TRANS
PIT_DETAIL_TRANS_CITY	Point in Time table	PIT, PIT_DETAIL_TRANS
PIT_DETAIL_TRANS_ADDRESS_CITY	Point in Time table	PIT, PIT_DETAIL_TRANS
PIT_DETAIL_TRANS_COUNTRY	Point in Time table	PIT, PIT_DETAIL_TRANS
PIT_DETAIL_TRANS_CUSTOMERS_ADDRESS	Point in Time table	PIT, PIT_DETAIL_TRANS
PIT_DETAIL_TRANS_CITY_COUNTRY	Point in Time table	PIT, PIT_DETAIL_TRANS

DIM_ON_BRIDGE_LNK_TIME_SELECTION Indicating from which LNK table that is directly connected with the main_hub we want to take the time slices for the SCD2

Object Name ↓	Object Type	Signature Object
Filter... shops	Filter... Relationship Based Links ×	Filter...
LNK_SHOPS_LOCATION	Relationship Based Links	LNK, DIM_ON_BRIDGE_LNK_TIME_SELECTION
LNK_SHOPS_EMPLOYEERANGE	Relationship Based Links	LNK, DIM_ON_BRIDGE_LNK_TIME_SELECTION

Create Signature attribute type and flag the usage in the requested fields



Create the template

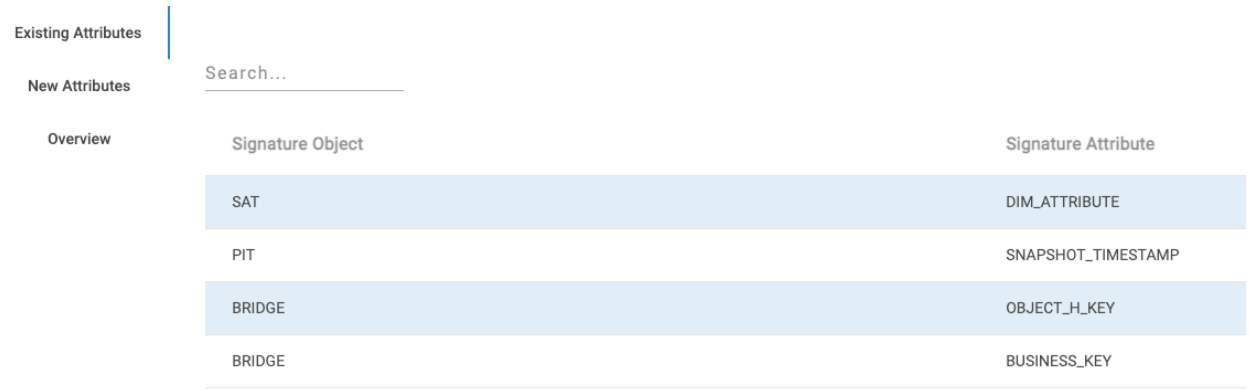
Template definition, according to the standard that you want to use. Important element here is the Signature Object naming. That name comes back in the .dvt file that contains the definition of the dimension template. If you have multiple implementations, you might want to go for a different name. But then you will need to adapt the template file accordingly and replace all the DIMB references with your chosen name.

Name	Description	Signature Object	Prefix	Suffix ↑	Object Type	Load Type	Signature Schema	Database Type	Folder Name	Base Type
DIM_ON_BRIDGE	Dimension on Bridge table	DIMB	DIM		VIEW	ALL	INFORMATION_MARTS	Snowflake		Bridge table

Take the dimension_view_on_bridge_multi_etl.dvt file and upload it for the ETL template of this view.

Fill in the Target Definition

The target definition is very specific to the template:



Add new attribute for the calculated snapshot and end snapshot. Attention, to avoid overlap with the existing snapshot_timestamp there is a double _!

Unique	Name	Prefix	Suffix	Signature Object	Signature Attribute	Data Type
	snapshot__timestamp					TIMESTAMP_TZ
	end_snapshot__timestamp					TIMESTAMP_TZ

Fill in the Dependency

Define on which BRIDGE table this template must be implemented.