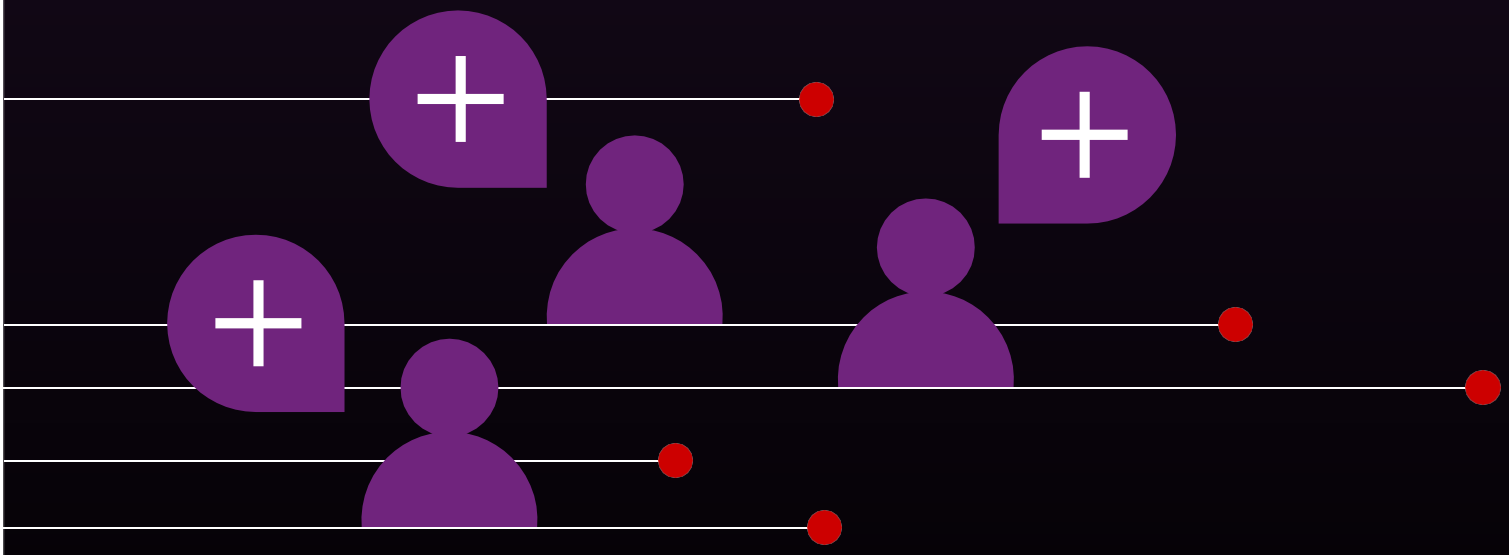
VAULTSPEED

VaultSpeed Studio

SCD2 dimension on a HUB



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 on a single HUB based on the transaction timestamp.

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.

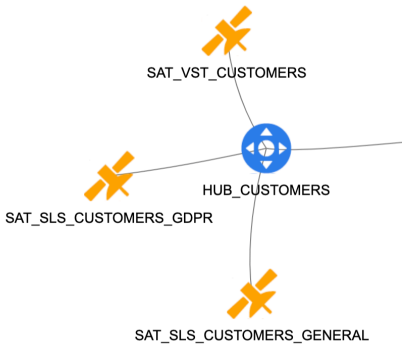
This template is designed for the following characteristics in the use-case:

- DIMENSION creation based on a HUB
- Slowly Changing Type 2 Dimension
- Versions in the dimension will be compressed
- PIT table must exist on the HUB
- 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.

In the Raw Data Vault model this is a HUB_CUSTOMER.



Components of the implementation

- PIT table on the HUB of choice
- 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

PIT table is created on the HUB

The Dimension is a SCD2 type dimension, for this we need to have a detailed PIT implementation on the HUB based on the TRANS_TIMESTAMP

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

Object Name (Linked) ↑
HUB_CUSTOMERS

Create Signature objects

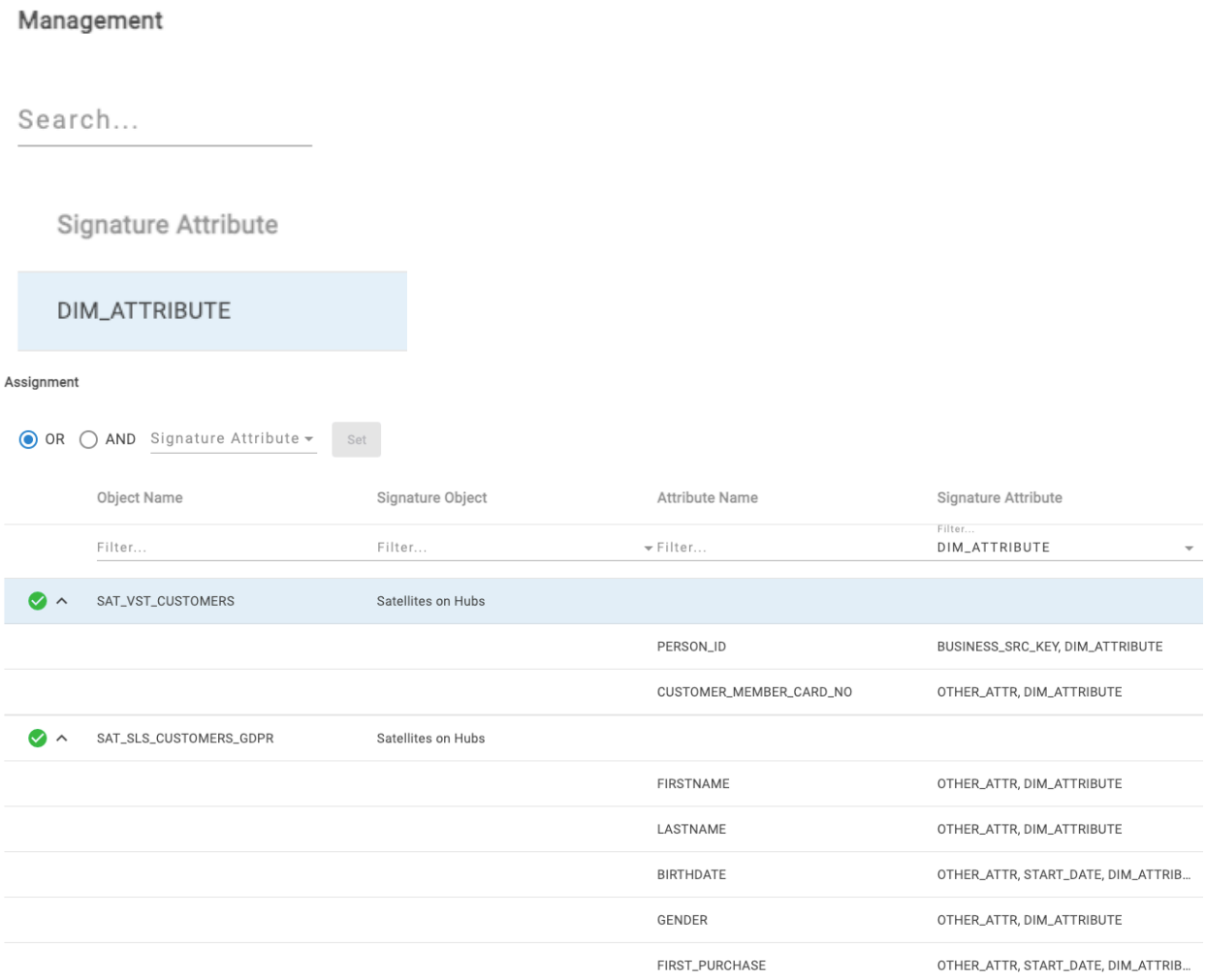
The PIT that is used in the implementation must receive the Signature object PIT_DETAIL_TRANS.

Create the Signature object PIT_DETAIL_TRANS

Signature objects are assigned

Assign the signature object PIT_DETAIL_TRANS to each PIT that has been created with the intention of having the SCD2 time setup implemented.

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 <signature object name> references with your chosen name.

In the example we use SCD2_DIM_HUB, because this is a SCD Type 2 dimension on a HUB template

SCD2_TRANS_DIM_ON_HUB	SCD Type 2 Dimension base...	SCD2_DIM_HUB	DIM	VIEW	ALL	DATA_WAREHOUSE	Snowflake	Hubs

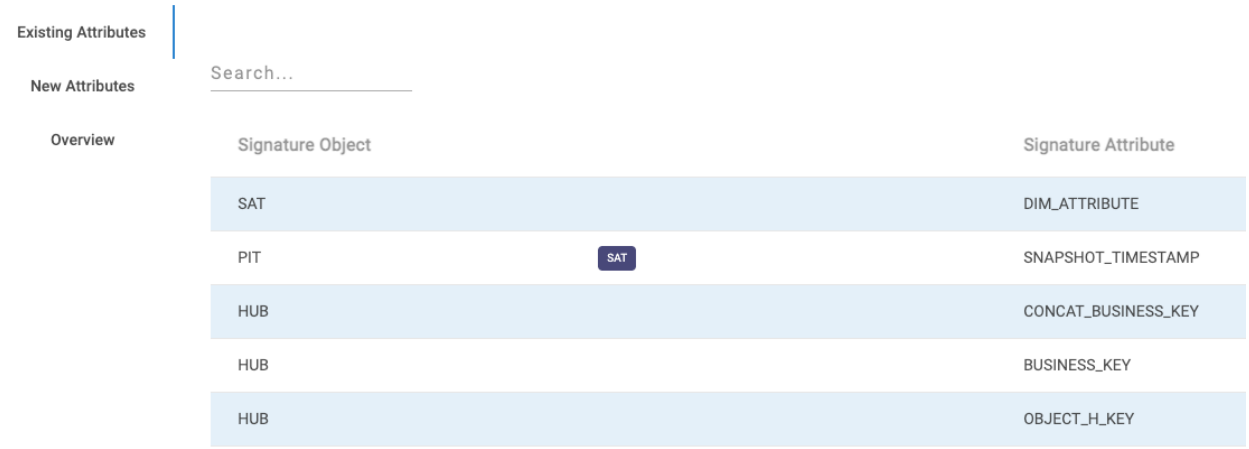
Ensure you have this use case's latest version from the following location: https://github.com/Vaultspeed/studio-templates .

Each of our templates has its folder where the code resides.

Take the dim_etl_template.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 attributes for the:

- calculated snapshot
- calculated end snapshot
- dim_<entity>_hkey

Be careful; there are 2 _ between snapshot and timestamp.

Unique	Name	Prefix	Suffix	Signature Object	Signature Attribute	Data Type
☑	snapshot__timestamp					TIMESTAMP_TZ
☑	end_snapshot__timestamp					TIMESTAMP_TZ
☐		DIM		HUB	OBJECT_H_KEY	VARCHAR

Fill in the Dependency

Define on which HUB table this template must be implemented.

Example: