Сущность: товарный учёт продуктового магазина

Статусная модель:

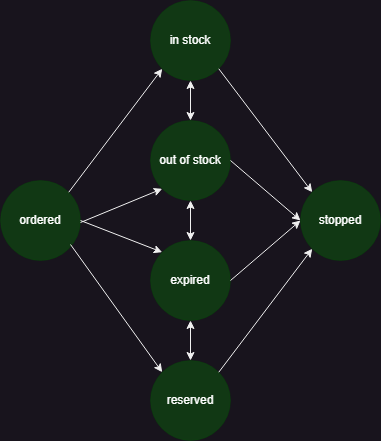
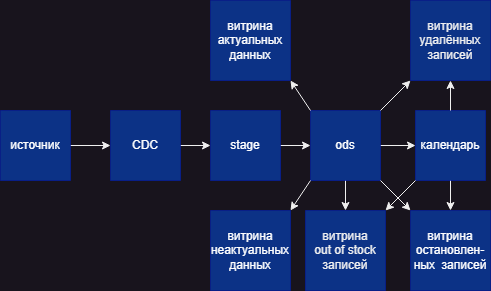


Схема решения:



Сгенерированные данные поступают в источник; далее с помощью триггера заполняется слой CDC данными из источника. В слое stage происходит загрузка данных из CDC, которые отсутствуют в ods (слой очищается на каждой итерации). В слое ods выполняется загрузка данных из stage. Заполнение календаря основано на данных из ods. Дополнительные витрины используют данные из ods и календаря.

**Схема oltp\_src\_system**

oltp\_src\_system.tracking\_products – таблица-источник, которая содержит синтетические данные.

oltp\_src\_system.create\_tracking() – функция, генерирующая новые записи.

oltp\_src\_system.update\_existed\_tracking() – функция, обновляющая случайные существующие записи.

oltp\_src\_system.delete\_existed\_tracking() – функция, удаляющая случайные существующие записи.

tracking\_products\_changes – триггер, реагирующий на delete, update и insert в таблице oltp\_src\_system.tracking\_products.

**Схема oltp\_cdc\_src\_system**

oltp\_cdc\_src\_system.cdc\_tracking\_products\_information – таблица, которая содержит лог изменений CDC. Строится на триггерах таблицы источника oltp\_src\_system.tracking\_products.

oltp\_cdc\_src\_system.cdc\_tracking\_products\_information() – функция, заполняющая слой CDC данными в зависимости от триггера.

**Схема dwh\_stage**

dwh\_stage.tracking\_products\_information\_src – таблица, которая содержит инкремент данных (разница между данными в CDC и ODS).

dwh\_stage.load\_from\_cdc\_tracking\_products\_information() – функция, заполняющая слой stage данными из CDC, которых нет в ODS. Очищает таблицу перед каждой загрузкой.

**Схема dwh\_ods**

dwh\_ods.tracking\_products\_information\_hist – таблица, которая содержит все версии всех строк таблицы источника.

dwh\_ods.dim\_date – аналитический справочник календарь.

dwh\_ods.load\_from\_tracking\_products\_information\_src() – функция, заполняющая слой ODS данными из stage.

dwh\_ods.load\_dim\_date() – функция, заполняющая календарь по записям из ODS. Очищает таблицу перед каждой загрузкой.

**Схема report**

report.tracking\_products\_information\_actual – таблица, которая содержит крайнее актуальное состояние данных в системе источнике.

report.tracking\_products\_information\_outdated – таблица, которая содержит неактуальное состояние данных в системе источнике.

report.tracking\_products\_information\_out\_of\_stock – таблица, которая содержит записи, имеющие статус out of stock. Витрина использует календарь и предназначена для анализа продуктов, которых нет в наличии (к какому идентификатору календаря относится запись; продукты закончились в выходные или в будние дни).

report.tracking\_products\_information\_stopped – таблица, которая содержит записи, имеющие конечный статус stopped. Витрина использует календарь и предназначена для анализа продуктов, поставка которых остановлена.

report.tracking\_products\_information\_deleted – таблица, которая содержит удалённые записи. Витрина использует календарь и предназначена для анализа продуктов, данные о которых были удалены.

report.load\_tracking\_products\_information\_actual() – функция, заполняющая таблицу актуальных записей данными из ODS.

report.load\_tracking\_products\_information\_outdated() – функция, заполняющая таблицу неактуальных записей данными из ODS.

report.load\_tracking\_products\_information\_out\_of\_stock() – функция, заполняющая таблицу записей, статус которых out of stock, данными из ODS.

report.load\_tracking\_products\_information\_stopped() – функция, заполняющая таблицу остановленных записей данными из ODS.

report.load\_tracking\_products\_information\_deleted() – функция, заполняющая таблицу удалённых записей данными из ODS.

Алгоритм добавления дополнительных таблиц-источников в слое oltp\_src\_system:

1. Настроить функции генератора синтетических данных на новый источник;
2. Добавить триггер на операции с данными в новом источнике для заполнения слоя CDC;

Дальнейшие функции не потребуют изменений, так как данные из нескольких источников поступают в один слой CDC.

Алгоритм действителен для данных, имеющих один тип и набор параметров в разных источниках.

Для добавления данных, имеющих иную структуру, необходимо реализовать новые функции на основе существующих.