# Кейс 3.

# Разработка прототипа хранилища данных

## Оглавление

2
2
2
3
3
4
4
4
4
4
5
5
6
6
6
6
6
7
8
9
9
15

# Описание и назначение полей таблиц базы данных

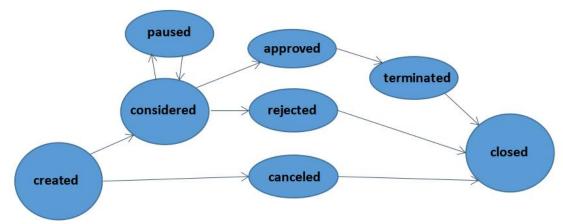
## OLTP Src System - Источник

Таблица со статусами заявок на получение премиальной карты обслуживания Cxema «oltp src system».

Таблица «card application data»

Наименование поля	Назначение
Id	Бизнес ключ (глобальный номер заявки)
ID_client	ID клиента (FK)
Create_dttm	Время создания задачи
Update_dttm	Время обновления задачи
Status_nm	Наименование статуса

- 1. Клиент может подать заявку на выпуск премиальной карты, одобрение поступает при соблюдении ряда условий.
- 2. До момента, пока заявка не принята в обработку, клиент может отменить запрос.
- 3. При рассмотрении заявки могут быть состояния технической паузы, после которого обязательно следует период рассмотрения.
- 4. Итогом рассмотрения заяви может быть подтверждение либо отклонение.
- 5. Если заявку на выдачу подтвердили, то следующий шаг выполнение заявки.
- 6. Из состояний отмена, отклонено и выполнено статус заявки переходит в закрыта.



Статусная модель

# OLTP CDC - логирование изменений источника

Таблица логирования изменений таблицы источника «card\_application\_data». Заполняется по триггеру на Insert, Update, Delete.

Схема «oltp\_cdc\_src\_system».

Таблица «cdc card application data changes»

Наименование поля	Назначение
Id	Бизнес ключ (глобальный номер заявки)
ID_client	ID клиента (FK) - в данном кейсе связь пустая (второй таблицы не создано)
Create_dttm	Время создания задачи
Update_dttm	Время обновления задачи
Status_nm	Наименование статуса
operation	Код операции (I, U, D). Код операции заносится в соответствии с выполненным действием
updated_dttm	Время срабатывания триггера и занесения данных (лог) в СОС

# DWH STAGE - первый слой прототипа хранилища данных

Таблица, получающая данные из CDC (логирование изменений таблицы источника), сформированные до момента запуска алгоритма записи в хранилище. Схема **«dwh stage»**.

Таблица «card application data dwh src»

Наименование поля	Назначение
Id	Бизнес ключ (глобальный номер заявки)
ID_client	ID клиента (FK) - в данном кейсе связь пустая (второй таблицы не создано)
Create_dttm	Время создания задачи
Status_nm	Наименование статуса
operation	Код операции (I, U, D). Код операции заносится в соответствии с выполненным действием
updated_dttm	Время срабатывания триггера и занесения данных (лог) в ССС
hash	Хеш-значение сгенерированное по строке таблицы

# DWH ODS - детальный слой хранилища

Таблица с историей изменения и состояния всех данных источника, валидные на момент проведения последней загрузки.

Схема «dwh ods».

Таблица «card application data hist»

Наименование поля	Назначение
Id	Бизнес ключ (глобальный номер заявки)
ID_client	ID клиента (FK) - в данном кейсе связь пустая (второй таблицы не создано)
Create_dttm	Время создания задачи
Status_nm	Наименование статуса
operation	Код операции (I, U, D). Код операции заносится в соответствии с выполненным действием
updated_dttm	Время срабатывания триггера и занесения данных (лог) в ССС
hash	Хеш-значение сгенерированное по строке таблицы
valid_from_dttm	Время старта актуальности записи. Соответствует времени занесения лога в CDC
valid_to_dttm	Время конца актуальности данных. Устанавливается либо в техническую бесконечность, либо для задач 'U' и 'D' на секунду меньше, чем дата загрузки (now() - interval '1 second')
deleted_flg	Флаг, что данная запись была удалена в источнике. "Y" / "N"
deleted_dttm	Время , когда было проведено удаление в источнике. Соответствует времени занесения лога в CDC

# Описание алгоритмов формирования данных в источнике и захвата данных в хранилище

#### Функция генерации появления данных в источнике

Название функции - oltp\_src\_system.create\_tasks().

Назначение функции - генерация случайной записи для внесения в таблицу источник. Для генерации номера заявки для формирования аргумента старта лля **generate\_series** сначала оценивается максимальное значение поля id и к этому значению прибавляется единица. Может появиться несколько заявок (рандом до 5). Шаг инкремента - 1.

**id** - Номер заявки - генерация последовательной серии значений, исходя из имеющегося наибольшего значения ID.

**id\_client** -рандомное число до 10000, поскольку округление к наименьшему целому, то +1 для избежания ситуации 0 - floor(random()\*(10000)+1).

status\_nm - 'created' - Установка статуса "Создано".

create dttm - Дата создания заявки - now().

**update dttm** -Дата обновления - в данном случае равна дате создания now().

## Функция обновления данных в источнике

Название функции - oltp src system.update existed task().

Для реализации алгоритма случайного выбора действий объявлена переменная status var := floor(random()\*(15) + 1).

В зависимости от полученного значения попадаем в блок обработки (попробовала несколько разграничить вероятность попадания в разные блоки), где генерируется выборочная метка на изменение:

- status\_var in (1, 8, 15) Выборочные заявки со статусом 'created' могут перейти в статус 'considered' (рассматривается) или 'paused' (на паузе например, нужны уточнения).
- status\_var in (2) Заявка со статусом 'considered' ("на рассмотрении") может перейти в состояние 'paused'.
- status\_var in (3, 13) Заявка 'paused' перед переходом в другое состояние должна снова быть 'considered'.
- status\_var in (6, 10, 12) Рассматриваемые заявки 'considered' могут быть либо 'отклонены' 'rejected' либо 'подтверждены' 'approved'.
- status\_var = 5 Созданная, но еще не обработанная заявка ('created'), может быть 'отменена' 'canceled'.
- status\_var in (4, 9, 14) Заявки со статусом 'подтверждено'- 'approved' должны обработаться 'terminated'.
- status\_var in (7, 11, 16) Заявки со статусом 'отказано', 'обработано' или 'отменено' (*'rejected'*, *'terminated'*, *'canceled'*) должны стать 'закрыты' *'closed'*.

#### Функция удаления данных из источника

Название функции - oltp\_src\_system.delete\_existed\_task(). Генерируется random-метка на удаление.

# Отслеживание изменений - формирование лог изменений СДС

Функция oltp\_src\_system.load\_to\_cdc\_from\_card\_application\_data\_changes()

Алгоритм строится на обработке срабатывания триггера таблицы источника. После выполнения целевого действия (добавление, изменение, удаление данных в источнике) по триггеру вызывается функция

oltp\_src\_system.load\_to\_cdc\_from\_card\_application\_data\_changes(). В таблицу, накапливающую логи oltp\_cdc\_src\_system.cdc\_card\_application\_data\_changes добавляется запись с дополнительными полями - тип проведенной операции Operation\_type (Insert - значение 'I'/ Update - значение 'U'/ Delete значение - 'D'). И время внесения данной записи в таблицу, которое соответствует настоящему времени updated\_dttm = now(). Данное поле в дальнейшем рассматривается как время актуального изменения состояния записей таблицы источника.

# Функция переноса данных из слоя oltp\_cdc\_src\_system в первый слой dwh stage

Функция dwh stage.load from cdc src system card application data to stage()

Сначала производится очистка значений таблицы dwh\_stage.card\_application\_data\_dwh\_src с помощью команды truncate table. Далее данные из таблицы логов oltp\_cdc\_src\_system.cdc\_card\_application\_data\_changes вставляются в таблицу dwh\_stage.card\_application\_data\_dwh\_src. При этом в таблицеприемнике для каждой строки записи генерируется hash-значение, при подсчете которого учитываются все поля, кроме типа операции (operation\_type). Поле hash используется в дальнейшем при идентификации задачи. Данное исключение сделано для того, чтобы при удалении записи не получить новый хеш.

Необходимо удостовериться, что подключено расширение pgcrypto;

## Функция записи данных из слоя dwh stage в детальный слой dwh ods

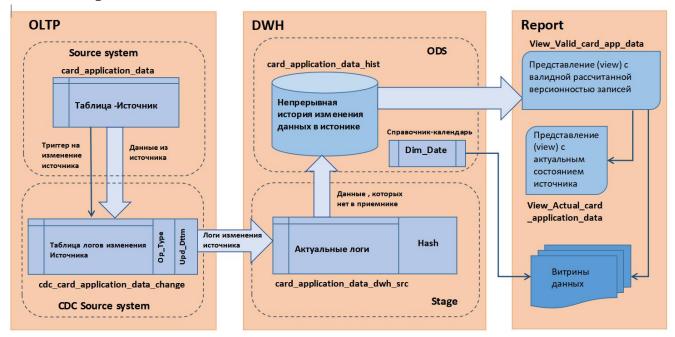
Данная функция **dwh\_ods.load\_card\_application\_data\_hist()** выполняет заливку данных, которых еще нет в хранилище, либо вносит сведения об изменении / удалении (из источника) данных о заявке, которые уже внесенны на хранение.

При записи данных происходит расчет и заполнение полей valid\_from\_dttm, valid\_to\_dttm, deleted\_flg, deleted\_dttm. Для задач, которые уже есть в ODS (условие dst.id = src.id) и сведения о их изменении есть в STAGE (произошедшие изменения в источнике залогированы системой CDC к данному моменту) (условие src.operation\_type in ('U', 'D')) изменяется значение в поле valid to dttm = now() - interval '1 second.

Далее в таблицу ODS добавляются записи, которых еще нет (в таблицу ODS отстувует записи с таким хеш-значанием). В качестве даты начала устанавливается залогированное значение появления записи (valid\_from\_dttm устанавливается равным updated dttm, которое получено при логировании CDC).

В качестве даты валидности окончания версии valid\_to\_dttm устанавливается техническая бесконечность (2999/12/31). По условию типа операции для добавляемых строк (является ли operation\_type = "D") устанавливается флаг удаленной записи deleted\_flg ("Y" / "N"). При этом для задач, которые были удалены будет дополнительно зафиксировано время удаления deleted\_dttm - равно времени обновления записи updated dttm.

# Схема решения



#### Маппинги S2T

Маппинги приложены в файле st71 case 3 mapping S2T.xlsx

# Слой Report

# Представление с валидной историей изменений источника

Название - report.View\_Valid\_card\_app\_data.

Формируется представление как копия данных их детального слоя хранилища на данный момент (на основании таблицы dwh\_ods.card\_application\_data\_hist), но производится корректировка даты окончания валидности истории записей в источнике. Для этого используется оконная функция, устанавливающая время valid\_to\_dttm либо на 1 секунду меньше, чем старт следующей записи с тем же id, но большей датой обновления, либо устанавливается техническая бесконечность для крайней записи.

Для просмотра необходимо запустить команду: select \* from report. View\_Valid\_card\_app\_data; Поскольку в DAG данное решение по просмотру не реализовано.

# Представление, содержащее информацию о крайнем актуальном состоянии источника

Название - report. View Actual card application data.

При формировании представления из родительского представления report.View\_Valid\_card\_app\_data с использованием оконной функции выбираются записи у которых самая старшая дата начала валидности valid\_from\_dttm и при этом operation\_type != 'D' - чтобы исключить удаленные записи, если они были самыми последними для задачи.

Для просмотра необходимо запустить команду: select \* from report. View\_Actual\_card\_application\_data. Поскольку в DAG данное решение по просмотру не реализовано.

# Задание на формирование отчетов и описание таблиц

**Уточнение**: В разработанном решении каждая задача может в один день перейти во множество статусов. Поэтому при расчете витрин такие задачи учитываются в один день в разных состояния.

1. Подготовить отчет о количестве задач, получивших каждый из возможных статусов в каждый день, заполнения и корректировки источника.

```
Название функции генерации отчета - report.tasks_start_in_state_count(); Название таблицы для сохранения отчета - tasks_start_in_state Поля таблицы: date_day - дата, status_nm - статус задач, tasks count - количество задач.
```

2. Подготовить отчет о количестве задач, <u>находящихся</u> в каждом из возможных статусов каждый день. В функции создания отчета предусмотреть связь с аналитическим справочником-календарем, для возможности добавления выбора данных по условиям. Так же предусмотреть возможность формирования выборки за определенный период.

```
Название функции генерации отчета - report.tasks_in_state_count() Название таблицы для отчета - tasks in state
```

```
Поля:
```

```
date_key - ключ из аналитического справочника-календаря, date_actual - дата, status_nm - статус задач, tasks count - количество задач.
```

3. Подготовить отчет о количестве задач, <u>находящихся</u> в каждом из возможных статусов которые были <u>удалены</u>. В функции создания отчета предусмотреть связь с аналитическим справочником-календарем, для возможности добавления выбора данных по условиям. Так же предусмотреть возможность формирования выборки за определенный период.

```
Название функции генерации отчета - report.deleted_tasks_count()
Название таблицы для отчета - deleted_tasks_count
Поля:
date_key - ключ из аналитического справочника-календаря,
date_actual - дата,
status_nm - статус задач,
tasks_count - количество задач.
```

4. Подготовить отчет, конкретизирующий ID задач, их статус и дату удаления из источника. В функции создания отчета предусмотреть связь с аналитическим справочником-календарем, для возможности добавления выбора данных по условиям. Так же предусмотреть возможность формирования выборки за определенный период.

```
Название функции генерации отчета - report.deleted_tasks_in_state()
Название таблицы для отчета - deleted_tasks_in_state
Поля:
date_key - ключ из аналитического справочника-календаря,
date_actual - дата,
tasks_id - количество задач,
status nm - статус задач.
```

# Алгоритмы формирования витрин

1. Отчет о количестве задач, получивших каждый из возможных статусов в каждый день, заполнения и корректировки источника.

Название функции генерации отчета - report.tasks start in state count();

Название таблицы для сохранения отчета - tasks\_start\_in\_state

#### Переменные функции:

 $var\_period\_from$  /\* Начало периода для построения витрины ОТ \*/  $var\_period\_to$  /\* Конец периода для построения витрины ДО \*/

#### Алгоритм:

Шаг 1. Очистка таблицы - tasks start in state.

Шаг 2. Для сгруппированных по дате valid\_from\_dttm и статусу задачи status\_nm данных, из представления View\_Valid\_card\_app\_data, содержащего валидную историю изменения данных источника, подсчитать количество ID задач для каждого статуса каждого дня.

**Шаг 3.** Накладывается ограничение по периоду выборки с помощью переменных **var\_period\_to**, а так же исключаются из подсчета строки с установленным флагом **deleted flg.** 

2. Отчет о количестве задач, находящихся в каждом из возможных статусов каждый день. В функции создания отчета предусмотреть связь с аналитическим справочником-календарем, для возможности добавления выбора данных по условиям. Так же предусмотреть возможность формирования выборки за определенный период.

Название функции генерации отчета - report.tasks\_in\_state\_count()

Название таблицы для отчета - tasks in state

#### Переменные функции:

var\_period\_from /\* Начало периода для построения витрины ОТ \*/
var\_period\_to /\* Конец периода для построения витрины ДО \*/

#### Алгоритм:

Шаг 1. Очистка таблицы - tasks in state count.

Шаг 2. Создается СТЕ tabl.

- Создается связь «для каждого дня все возможные статусы задачи».
- Подсчитывается количество задач, не находящихся в состоянии удалено ( deleted\_flg = "N") с актуальной датой валидности на день подсчета.

**Шаг 3.** Накладывается ограничение переменными функции **var\_period\_from** и **var\_period\_to** периода. А так же исключается выборка в итоговую таблицу null значений task count.

3. Отчет о количестве задач, находящихся в каждом из возможных статусов которые были удалены.

Название функции генерации отчета - report.deleted tasks count()

Название таблицы для отчета - deleted tasks count

#### Переменные функции:

var\_period\_from /\* Начало периода для построения витрины ОТ \*/ var period to /\* Конец периода для построения витрины ДО \*/

#### Алгоритм:

Шаг 1. Очистка таблицы - deleted tasks count.

**Шаг 2.** Из соединения по дате (inner join) представления **View\_Valid\_card\_app\_data** со справочником-календарем **DimDate**, отбираются поля Ключ и Дата (из календаря), статус (status\_nm) из представления и рассчитывается количество задач (count(ID)), имеющих флаг удаления задачи **deleted flg = "Y"**.

**Шаг 3.** Накладывается ограничение на вывод дней, когда удаленных задач не было, а также ограничивается период ( var period from и var period to)

4. Отчет, конкретизирующий ID задач, их статус и дату удаления из источника с наложением на календарь.

Название функции генерации отчета - report.deleted tasks in state()

Название таблицы для отчета - deleted tasks in state

## Переменные функции:

var\_period\_from /\* Начало периода для построения витрины ОТ \*/
var\_period\_to /\* Конец периода для построения витрины ДО \*/

#### Алгоритм:

Шаг 1. Очистка таблицы - deleted tasks in state.

**Шаг 2.** Из соединения по дате (inner join) представления **View\_Valid\_card\_app\_data** со справочником-календарем **DimDate**, отбираются поля Ключ и Дата (из календаря) и Номер задачи (ID) и статус (status\_nm) из представления для периода, ограниченного переменными функции **var\_period\_from** и **var\_period\_to** и при этом флаг удаления задачи **deleted flg** установлен в "Y".

# Тест-кейс для прототипа

Уточнение: Все команды по запуску функций и выборке данных подготовлены и размещены в конце файла со скриптом sql - st71\_case\_3\_control\_script.sql, приложенным к отчету.

## Загрузка и работа на тестовых данных

Для тестирования подготовлены тестовые данные и функции их внесения на уровень OLTP.

#### Описание тестовых данных:

Подготовлена история изменения для 10 задач.

#### Учитываются следующие случаи:

- ✓ Задача создана и отменена на следующий день (задачи 1, 5). Окончательное состояние задач Canseled.
- ✓ Задача создается и переходит в следующий статус в один и тот же день (задача 2). Окончательное состояние задачи Closed.
- ✓ Задача переходит по статусам в разные дни и удаляется ( задача 3, дата удаления 7.05.2023). Окончательное состояние задачи Closed. DELETE.
- ✓ Задача создана, но долго не переходит в иные состояния (задача 4). Окончательное состояние задачи Created.
- ✓ Задача Отклонена и закрыта ( задача 6). Окончательное состояние задачи Closed.
- ✓ Задача Одобрена, выполнена и закрыта (задача 7). Окончательное состояние задачи Closed.
- ✓ Задача создана и удалена в статусе created (задачи 8 и 9, даты удаления 16. 05.2023). Окончательное состояние задачи Created. DELETE.
- ✓ Задача создана, принята в рассмотрение и удалена (задача 10, дата удаления 16.05.2023). Окончательное состояние задачи Considered. DELETE.

#### Информация по удаленным задачам:

id 💌	id_clien 🔻	status_nm	create_dttm	update_dttm ▼	oteration_type
3	30	closed	03.05.2023	07.05.2023	D
8	80	created	15.05.2023	16.05.2023	D
9	90	created	15.05.2023	16.05.2023	D
10	170	considered	15.05.2023	16.05.2023	D

#### Шаги для проведения теста:

1. Подготовка пустого хранилища

```
truncate table dwh_stage.card_application_data_dwh_src; /* Очистка Stage*/truncate table dwh_ods.card_application_data_hist; /* Очистка ODS*/
```

- 2. Заполнение таблицы источника данными запуск функции загрузки данных в источник. select oltp\_src\_system.load\_src\_for\_test();
- Просмотреть состояние источника:
   select \* from oltp\_src\_system.card\_application\_data cad;

Результат загрузки данных в источник:

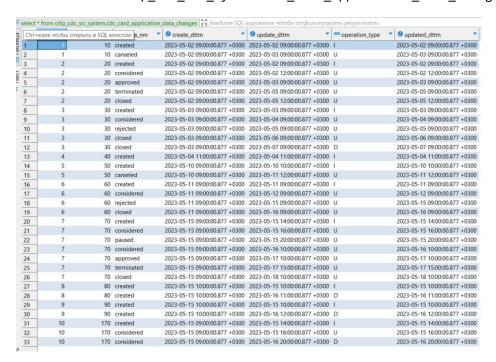
	123 id	*	123 id_client	noc status_nm	② create_dttm ▼	<pre>update_dttm</pre>
1		1	10	canseled	2023-05-02 09:00:00.877 +0300	2023-05-03 09:00:00.877 +0300
2		2	20	closed	2023-05-02 09:00:00.877 +0300	2023-05-05 12:00:00.877 +0300
3		4	40	created	2023-05-04 11:00:00.877 +0300	2023-05-04 11:00:00.877 +0300
4		5	50	canseled	2023-05-10 09:00:00.877 +0300	2023-05-11 12:00:00.877 +0300
5		6	60	closed	2023-05-11 09:00:00.877 +0300	2023-05-16 09:00:00.877 +0300
6		7	70	closed	2023-05-15 09:00:00.877 +0300	2023-05-18 10:00:00.877 +0300

4. При заполнении источника сработает триггер на Insert о вставке записей в источник и произойдет запись логов в CDC. Но Исторических изменений этих данных в CDC нет. Поэтому необходимо запустить функцию загрузки CDC подготовленными историческими логами. Функция load\_cdc\_for\_test() проведет обнуление информации в CDC и запись подготовленной истории.

```
select oltp_cdc_src_system.load_cdc_for_test();
```

5. Просмотр данных в СОС

select \* from oltp\_cdc\_src\_system.cdc\_card\_application\_data\_changes;

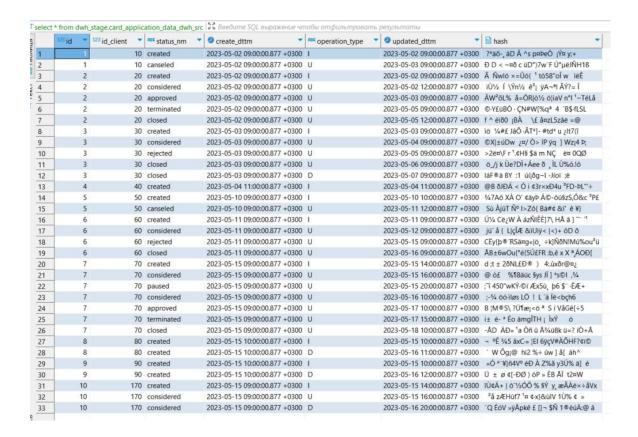


6. Для не автоматизированного запуска наполнения хранилища выполняем команды:

```
select dwh_stage.load_from_cdc_src_system_card_application_data_to_stage();
```

7. Просмотр записей в Stage слое хранилища. Сформирован хеш для каждой записи.

```
select * from dwh stage.card application data dwh src;
```



8. Просмотреть, что в детальный слой данные занесены. Для удаленных записей выставлены флаги. И заполнены даты удаления.

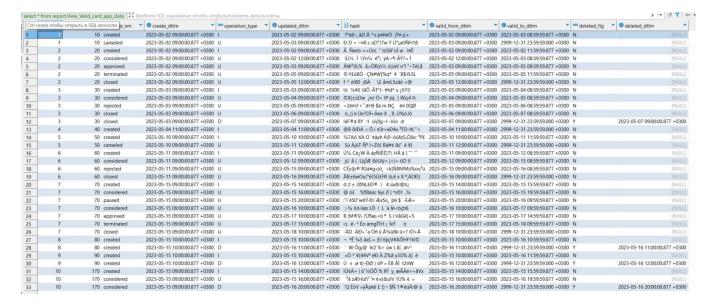
select \* from dwh ods.card application data hist order by create dttm;



Даты окончания валидности записей ограничиваются либо технической бесконечностью, либо временем загрузки в ODS хранилища. Восстановление историчности изменения будет проводится на следующем шаге.

9. Загрузить для просмотра представление на основе данных из ODS. Обратить внимание, что в представлении скорректированы даты версионности для каждой задачи. Восстановлена валидность записей.

select \* from report.View\_Valid\_card\_app\_data;

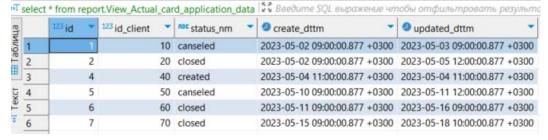


Задание 1. Получить крайнее актуальное состояние данных в системе источнике

Выполнить команду:

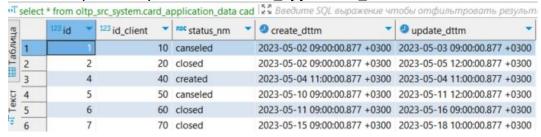
select \* from report.View\_Actual\_card\_application\_data;

Содержимое представления View Actual card application data:



Сравнить с содержимым источника

Содержимое представления card application data:



РЕЗУЛЬТАТ: Содержимое совпадает. Сбор данных работает корректно.

#### Задание 2. Подготовить несколько витрин данных

**Уточнение**: В разработанном решении каждая задача может в один день перейти во множество статусов. Поэтому при расчете витрин такие задачи учитываются в один день в разных состояния.

1. Запустить функцию подготовки отчета о количестве задач, получивших каждый из возможных статусов по дням.

select report.tasks\_start\_in\_state\_count();

Просмотр отчета **select** \* **from** report.tasks\_start\_in\_state\_count;

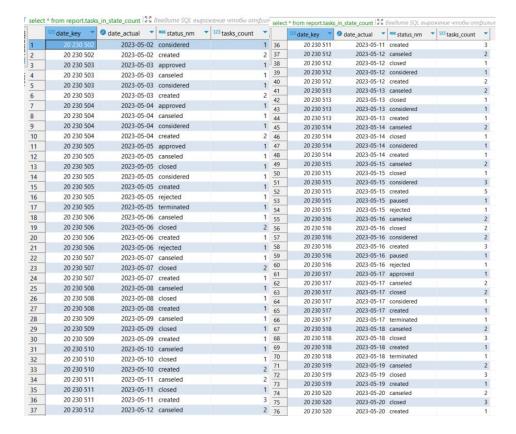


*Результат* - данные корректны, поскольку некоторые задачи переходят в один и тот же день в несколько статусов.

2. Запустить функцию подготовки отчета о количестве задач, находящихся в каждом из возможных статусов по дням.

```
select report.tasks_in_state_count();
```

Просмотр отчета **select** \* **from** report.tasks\_in\_state\_count;

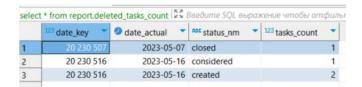


Результат - данные корректны.

3. Запустить функцию подготовки отчета о количестве удаленных задач, для каждого из возможных статусов по дням.

select report.deleted\_tasks\_count();

Просмотр отчета **select** \* **from** report.deleted\_tasks\_count;

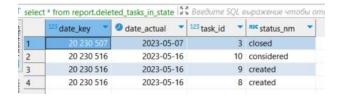


Результат - данные корректны.

4. Запустить функцию подготовки отчета об идентификаторах и статусе удаленных задач по дням. (мы знаем из предидущего отчета, что удалено 4 задачи, посмотрим их идентификаторы и статусы).

select report.deleted\_tasks\_in\_state();

Просмотр отчета **select** \* **from** report.deleted\_tasks\_in\_state;



Результат - данные корректны.

#### Запуск генерации данных

last\_run\_flg\_dummy

а) ручной запуск функций генерации новых данных

```
/*Функции РАНДОМ-генерации данных для тестирования*/
select oltp_src_system.create_tasks(); /* генерация новых задач*/
select oltp_src_system.update_existed_task(); /* изменение случайных существующих*/
select oltp_src_system.delete_existed_task(); /* удаление случайных задач*/
```

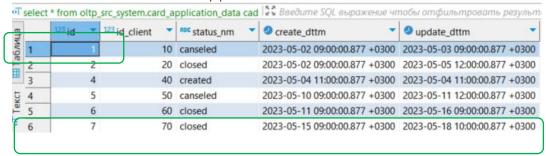


При окончании работы DAG/либо ручном запуске некоторых функций генерации данных выполнить команды для просмотра отчетов.

```
select * from report.tasks_start_in_state_count; /* ВИТРИНА -Количество задач, получивших данный статус в данный день */
select * from report.tasks_in_state_count; /* ВИТРИНА - Количество задач, находящихся в данном статус в данный день*/
select * from report.deleted_tasks_count; /* ВИТРИНА - Количество удаленных задач, находящихся в данном статусе в данный день*/
select * from report.deleted_tasks_in_state; /* ВИТРИНА - Задачи, удаленные в данном статусе в данный день*/
```

В зависимости от того, как отработает рандомайзер генерации и обновления данных результаты будут отличаться. Могут быть случаи, что данные не обновляются или не удаляются несколько циклов запуска. Далее показано как проверить данные.

Состояние источника ДО теста:



Просмотрим состояние источника и представления , построенного по данным из хранилища, после теста

select \* from report.View\_Actual\_card\_application\_data;

Представление с актуальными данными источника

_	127 id 127 id_c	lient 🐣	noc status_nm *	<pre>create_dttm</pre>	updated_dttm
1	2	20	closed	2023-05-02 09:00:00.877 +0300	2023-05-05 12:00:00.877 +0300
_	4	40	created	2023-05-04 11:00:00.877 +0300	2023-05-04 11:00:00.877 +0300
3	5	50	canseled	2023-05-10 09:00:00.877 +0300	2023-05-11 12:00:00.877 +0300
4	6	60	closed	2023-05-11 09:00:00.877 +0300	2023-05-16 09:00:00.877 +0300
5	7	70	closed	2023-05-15 09:00:00.877 +0300	2023-05-18 10:00:00.877 +0300
5	8	3 130	created	2023-05-20 22:08:56.616 +0300	2023-05-20 22:08:56.616 +0300
7	9	3 474	created	2023-05-20 22:08:56.616 +0300	2023-05-20 22:08:56.616 +0300

select \* from oltp\_src\_system.card\_application\_data cad;

Содержимое источника ПОСЛЕ теста

	123 id	123 id_client	noc status_nm	create_dttm	update_dttm **
1	2	20	closed	2023-05-02 09:00:00.877 +0300	2023-05-05 12:00:00.877 +0300
2	4	40	created	2023-05-04 11:00:00.877 +0300	2023-05-04 11:00:00.877 +0300
3	5	50	canseled	2023-05-10 09:00:00.877 +0300	2023-05-11 12:00:00.877 +0300
4	6	60	closed	2023-05-11 09:00:00.877 +0300	2023-05-16 09:00:00.877 +0300
5	7	70	closed	2023-05-15 09:00:00.877 +0300	2023-05-18 10:00:00.877 +0300
6	8	3 130	created	2023-05-20 22:08:56.616 +0300	2023-05-20 22:08:56.616 +0300
7	9	3 474	created	2023-05-20 22:08:56.616 +0300	2023-05-20 22:08:56.616 +0300

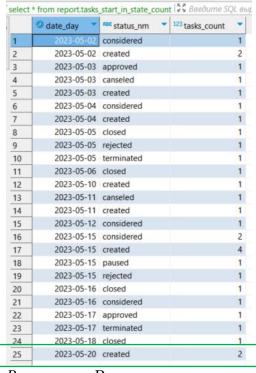
Итого, после выполнения функций обновления данных в источнике находится **три** задачи в статусе **closed**, **одна** задача в статусе **canseled** и **три** задачи в статусе **created**. Было **добавлено две** задачи. Была **удалена одна** задача в статусе **canseled** (**ID=1**)

Необходимо проверить, правильно ли рассчитаны витрины статусов и удаленных задач.

Витрина «Количество задач, получивших статус по дням».

```
select * from report.tasks_start_in_state_count;
```

20/05/2023 года при подготовке описания процесса тестирования было (рандомно) создано и размещено в источник 2 задачи.

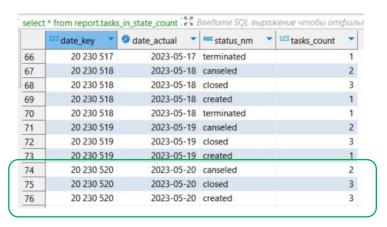


Результат - Витрина рассчитана верно.

Витрина «Количество задач, имеющих определенный статус по дням».

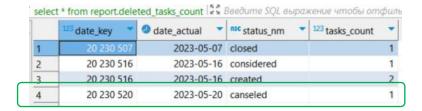
```
select * from report.tasks_in_state_count;
```

*Результат* - по состоянию на 20/05/2023 в статусе **canseled** в источнике находится **2** задачи, в статусе **created** - **3** задачи и в статусе **Closed** - **3** задачи. *Витрина рассчитана верно*.



Витрина «Количество задач, которые были удалены в каждом статусе по дням».

select \* from report.deleted\_tasks\_count;



Результат - при подготовке теста 20/05/2023 была удалена одна задача в состоянии canseled.

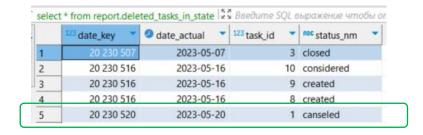
Витрина рассчитана верно.

Теперь можем конкретизировать какие именно задачи удалены.

Витрина «Идентификаторы и статусы удаленных задач по дням».

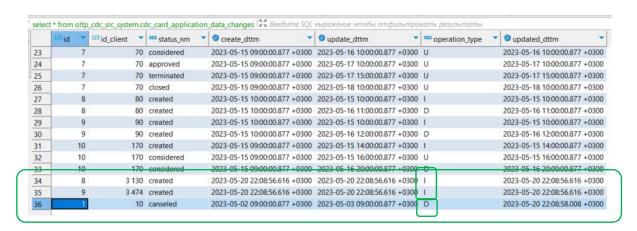
select \* from report.deleted\_tasks\_in\_state;

Pезультат - При подготовке теста 20/05/2023 была удалена одна задача (ID = 1) в состоянии canseled. Витрина рассчитана верно.



Так же можно проверить слои хранилища, и как данные об изменениях помещались в слои.

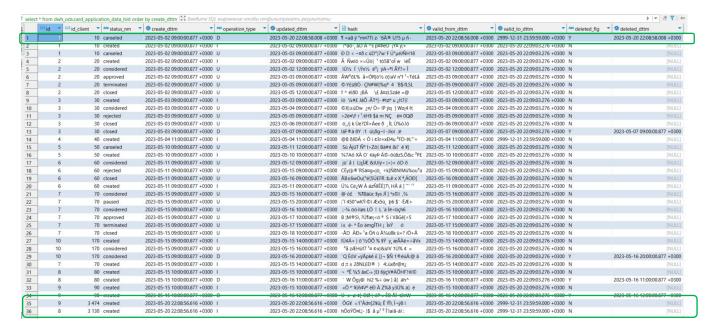
Таблица логов изменения источника CDC select \* from oltp\_cdc\_src\_system.cdc\_card\_application\_data\_changes;



Результат: изменения внесены корректно.

Просмотр содержимого приемника слоя ODS

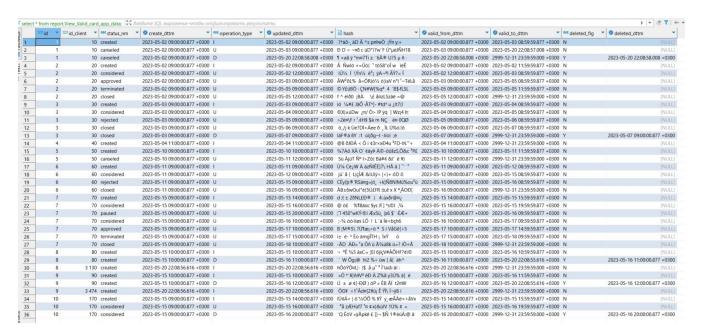
select \* from dwh ods.card application data hist order by create dttm;



*Результат*: Данные внесены. Техничесое поле **valid\_from\_dttm** заполнено в соответствии с временем добаления логов в CDC. При первом внесении записи техническое поле **valid\_to\_dttm** устанавливается в техническую бесконечность. При последующих изменениях записей об обновлении в техническое поле **valid\_to\_dttm** будет вноситься время обновления приемника.

Проверка пересчета валидности дат в слое Report

select \* from report.View\_Actual\_card\_application\_data;



*Результат*: Данные отражаются корректно. Версионность историчности данных для поля **valid\_to\_dttm** восстановлена.