

Задание 5. Шардирование

Реализован функционал

- Отправка сообщения пользователю (метод `/dialog/{user_id}/send` из [спецификации](#))
- Получение диалога между двумя пользователями (метод `/dialog/{user_id}/list` из [спецификации](#))

Шардирование

Обеспечить горизонтальное масштабирование хранилищ на запись с помощью шардинга.

В качестве БД используется Citus.

Докер файл для поднятия кластера доступен в github <https://github.com/eaahome/HL-arch/blob/master/HighLoadDialogs/src/main/docker/docker-compose.yml>

```
CREATE TABLE IF NOT EXISTS public."dialog" (  
    id character varying NOT NULL default gen_random_uuid(),  
    "text" character varying NOT NULL,  
    src_user_id character varying NOT NULL,  
    dst_user_id character varying NOT NULL,  
    created_at TIMESTAMPTZ DEFAULT Now(),  
    dist_key character varying default gen_random_uuid()  
);  
  
SELECT create_distributed_table('dialog', 'dist_key');
```

Предусмотреть возможность решардинга

```
alter system set wal_level = logical;  
SELECT run_command_on_workers('alter system set wal_level = logical');
```

После применения настроек рестартим кластер.

После добавления-удаления воркеров и применения настроек выше запускаем решардинг:

```
SELECT citus_rebalance_start();
```

Решение вопроса "тяжелых" пользователей

Предусмотреть:

- “Эффект Леди Гаги” (один пользователь пишет сильно больше среднего)

Если пользователь пишет больше среднего, т.е. пишет сразу во многих диалогах (с многими пользователями), то нужен композитный ключ, который разделит переписку на конкретные диалоги, например

```
dist_key = 'src_user_id' + 'dst_user_id'
```

Это позволит один диалог держать на одном шарде, что ускорит выборку одного диалога целиком, и при этом распределит разные диалоги на разные шарды.

См. метод [calculateDistributionKey](#).

Если же кроме описанной активности добавляется еще и большая активность внутри конкретного диалога, то есть смысл добавить к ключу еще дату.

```
dist_key = 'src_user_id' + 'dst_user_id' + date()
```

Это позволит распределить еще и сам диалог по разным шардам, при этом дневной диалог все еще будет на одном шарде, что позволит дневную выборку сделать быстро, но при этом один и тот же диалог за разные дни уже будет распределен по разным шардам, что позволит распределить нагрузку более равномерно.

Создание композитного ключа предлагается делать на стороне приложения, чтобы обойти ограничения, которые накладываются на ключ шардирования (например: "Distribution column must not use GENERATED ALWAYS AS (...) STORED", "cannot use column reference in DEFAULT expression").