

Sharding

1. Реализация

Для практики решила не использовать готовое решение, а реализовать шардирование на стороне приложения. Подняла два новых инстанса postgres. Реализовала простое шардирование по ключу.

На основном хосте базы (нешардированная) есть таблица dialog, в которой лежат dialog_id и айди пользователей участвующих в диалоге. dialog_id - ключ шардирования. Чтобы получить номер шарда в приложении беру остаток от деления на кол-во шардов.

Запись dialog/{user_id}/send

Беру шард под номером dialog_id%shardCount и пишу туда

Чтение dialog/list

Получаю все диалоги (dialog_id) из основной нешардированной базы. Делаю параллельные запросы отдельно на каждый шард и склеиваю результаты, сортирую по диалогам и времени отправления сообщения.

2. Решардинг

- В сервисе создаем новый маппинг номера шарда с его соединением
- Запись:
 - Используем новый маппинг шардов
- Чтение:
 - Читаем из нового и старого маппинга шардов. Для текущей задачи с сообщениями достаточно склеить все полученные данные. Если происходит перемещение данных из старого шарда на новый, то при наличии одинаковой записи в обоих шардах, в приоритете будет запись из нового шарда.
- Перенос данных
 - Далее вручную будем собирать данные, которые лежат не на своем шарде после решардинга. Переливаем их на нужный шард. Удаляем со старого шарда.