



Отчуждаемые таблицы в PgPro EE

Фролков И.А.

postgrespro.ru



Данных много

- Раньше было много данных
- Сейчас очень много
- Будет еще больше
- Что это за данные?
 - Журналы операций — финансовые, клики, показы, истории болезни и т. д. и т.п.
 - Показания датчиков

Данных много

- Раньше было много данных
- Сейчас очень много
- Будет еще больше
- Что это за данные?
 - Журналы операций — финансовые, клики, показы, истории болезни и т. д. и т.п.
 - Показания датчиков

Это неизменяемые данные

Проблема

- Плохо ли хранить большие объемы неизменяемых данных в БД?
- И да, и нет:
 - **Плохо**, потому что БД — это не только язык запросов, но и мощные средства для обеспечения параллельной работы
 - Раздутый бекап (инкрементальный может помочь)
 - Сложности с добавлением, если данных действительно много
 - PREPARE/EXECUTE, COPY
 - Не забываем про индексы, вакуум и статистику
 - **Хорошо**, потому что все-таки удобный язык запросов и методы доступа

Проблема - и строка, и страница жестко привязаны к создавшей их БД.

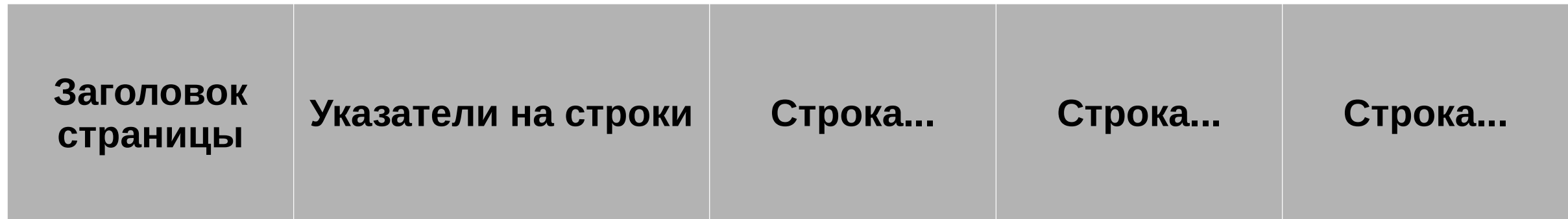
Проблема

- Почему? Определение видимости
 - xmin/xmax — транзакции, которые создали и удалили строку
 - Стало быть, нужен clog — журнал собственно транзакций
- Вывод — нельзя вот так просто взять и перенести таблицу. И индексы. И статистику.
- Не забываем про toast!

Правила определения видимости строки

- xmin/xmax
 - xmin — транзакция, которая создала строку
 - xmax — транзакция, которая строку удалила
- Список активных транзакций
- clog — история транзакций
 - Последовательность бит, указывающих, успешно ли завершилась транзакция.
- И что же, для каждой строки надо посмотреть, активна ли транзакция, создавшая ее и успешно ли завершилась транзакция, строку удалившая?
 - Нет!
 - **VACUUM!!!**

Как выглядит страница



PD_ALL_VISIBLE — все строки на странице видны всем транзакциям
 /include/storage/bufpage.h

Как выглядит строка

xmin	xmax	Прочее	флаги
199233	0	...	XMIN_COMMITTED, HEAP_XMIN_FROZEN, HEAP_XMAX_COMMITTED...

`/include/access/htup_details.h`

Как выглядят таблицы в ФС

- Таблица — это просто группа файлов с определенным именем (pg_class.relfilenode)
- Индекс — это тоже группа файлов с определенным именем
- VM/FSM — и это файлы с определенным именем.

Что надо для переносимой таблицы?

- Создать таблицу
- Вставить данные
- VACUUM FREEZE
- По-хорошему это все — осталось удалить ее из каталога в одной БД и добавить в каталог другой.
- Ну разумеется, есть детали — например, таблица должна быть доступна только на чтение.
- И не забываем о toast!

Выход есть в PgPro EE

- pg_transfer
 - Стандартное расширение
 - Относительно простая процедура:
 - Подготовить таблицу для переноса
 - Создать в целевой БД пустую таблицу с аналогичной структурой — две команды
 - Перенести данные
 - Все
- <https://postgrespro.ru/docs/postgresproee/9.6/pgtransfer.html>

Что это дает?

- Данные могут быть подготовлены на других серверах.
 - Очевидно, серверов может быть много.
- Время собственно добавления данных не зависит от их объема
- С репликацией, конечно, ситуация хуже — надо писать на второй сервер
- Возможно обеспечить хранение архива не в виде дампа целой БД

**Фактически неограниченная скорость
добавления**

Ограничения

- Требуется бинарная совместимость между серверами
- Относительно сложный процесс переноса

Вопросы?

Postgres Professional

<http://postgrespro.ru/>

+7 495 150 06 91

info@postgrespro.ru

The background is a collage of hexagonal tiles in various shades of blue and orange. Some tiles contain abstract patterns like splatters, wavy lines, or dotted grids. A white wavy line graphic is positioned at the bottom center of the page.

postgrespro.ru