Зачем еще 64-битные значения?

Teodor Sigaev

www.postgrespro.ru

Postgres PROFESSIONAL

О чем это я?

- История — откуда есть всё пошло
- Что случилось сейчас?
- Как это исправить?

Postgres PROFESSIONAL

XID & MVCC

- Счетчик транзакций. Монотонный.
- Record/Tuple/Row — xmin & xmax
- Commit/Rollback не трогает Record/Tuple/Row
  - Старые версии в таблице. Vacuum и всё такое.

Postgres PROFESSIONAL

Видима ли запись?

```
Bool is_visible(tuple) {
  if (tuple.xmin is committed) {
    if (tuple.xmax is not committed) {
      return true;
    }
  }
  return false;
}
```

XID & MVCC

- Дорого узнавать статус транзакций — сделаем hint bit
- Кто проставляет hint bit?
  - Vacuum
  - Select?! (ой, а у меня read-only запрос насилует диск)

История, эпоха 1

- Счетчик транзакций — 32 бита. 2^32 — это сколько?
- Это много… Очень много…
- 8 байт в заголовке tuple — xmin & xmax

История, эпоха 1

- А что если 2^32 это не очень много?
- Вам не повезло
- Ответ сообщества!

История, эпоха 2

- Wraparound!
- Не сброс счетчика!
- Кольцо в алгебре помните?

История, эпоха 2

- -2^31 от текущего — прошлое
- 2^31 от текущего — будущее

История, эпоха 2

- Э, а как узнать, есть ли очень старые транзакции?
- Vacuum freeze

История, эпоха 2

- Мы живем в этой эпохе
- В ней все хорошо
  - Есть автовакуум
  - Есть freeze map
  - Счастье?

Postgres PROFESSIONAL

Проблемы эпохи 2

- Вакуум не такая уж и дешевая операция
  - Чем реже гоняем, тем она тяжелее. Баланс найти нелегко
  - Автовакуум надо улучшать — может долго не приходить
- А что с темпом?
  - Мощности выросли, диски быстрее — wraparound раз в неделю уже не фантастика (сравните с эпохой 1!)
  - DBA как искусник
- Кто-нибудь в проде видит 32-х битные системы?

64 бита хватит всем!

- XID — 64 bits
- 2^64 = 18446744073709551616 ≈ 2e19
- 2^64 / 2^30 ≈ 50 млн лет

64 бита хватит всем!

- Ура?!
- А как же заголовок тупла? Теперь 16 байт. А еще cmin/cmax…
- Фокус на следующем слайде

64 бита хватит всем!

- Храним только младшие 32 бита
- Вернулись, с чего начали?

64 бита хватит всем!

- Старшие биты — общие на страницу
- Ограничение — 4 млрд разных транзакций на странице (hint bits)
- Да, вакуум. Но он всё равно нужен.
- Есть еще загвоздки в индексах

Будущее, эпоха 3

- 64 бита хватит всем
- Ушел wraparound
- 32-х битные системы — проблема
  - Нет атомарных чтений. Решаемо, но есть плата
- Оперирование бОльшими данными — не проблема
- Подрезание clog

Наступление эпохи 3

- Помогите!
  - Тестами
  - Самим тестированием
  - И не только функциональным, производительность тоже.
  - И не только x86*
  - Ревью!
- Еще б CSN… Но об этом — в следующий раз

Postgres PROFESSIONAL

Спасибо за внимание!

teodor@postgrespro.ru