

Оглавление

О книге	17
Глава 1. Введение	23
Часть I. Изоляция и многоверсионность	45
Глава 2. Изоляция	47
Глава 3. Страницы и версии строк	75
Глава 4. Снимки данных	97
Глава 5. Внутространичная очистка и hot-обновления	111
Глава 6. Очистка и автоочистка	125
Глава 7. Заморозка	150
Глава 8. Перестроение таблиц и индексов	165
Часть II. Буферный кеш и журнал	177
Глава 9. Буферный кеш	179
Глава 10. Журнал предзаписи	201
Глава 11. Режимы журнала	224
Часть III. Блокировки	243
Глава 12. Блокировки отношений	245
Глава 13. Блокировки строк	258
Глава 14. Блокировки разных объектов	283
Глава 15. Блокировки в памяти	295
Часть IV. Выполнение запросов	305
Глава 16. Этапы выполнения запросов	307
Глава 17. Статистика	333
Глава 18. Табличные методы доступа	358
Глава 19. Индексные методы доступа	381
Глава 20. Индексное сканирование	401
Глава 21. Вложенный цикл	426

Оглавление

Глава 22. Хеширование	446
Глава 23. Сортировка и слияние	472
Часть V. Типы индексов	499
Глава 24. Хеш-индекс	501
Глава 25. В-дерево	513
Глава 26. Индекс GiST	540
Глава 27. Индекс SP-GiST	574
Глава 28. Индекс GIN	598
Глава 29. Индекс BRIN	628
 Заключение	 656
Предметный указатель	657

Содержание

О книге	17
Глава 1. Введение	23
1.1. Организация данных	23
Базы данных	23
Системный каталог	24
Схемы	25
Табличные пространства	26
Отношения	27
Слои и файлы	28
Страницы	33
TOAST	33
1.2. Процессы и память	39
1.3. Клиенты и клиент-серверный протокол	41
 Часть I. Изоляция и многоверсионность	 45
Глава 2. Изоляция	47
2.1. Согласованность	47
2.2. Уровни изоляции и аномалии в стандарте SQL	49
Потерянное обновление	50
Грязное чтение и Read Uncommitted	50
Неповторяющееся чтение и Read Committed	51
Фантомное чтение и Repeatable Read	51
Отсутствие аномалий и Serializable	52
Почему именно эти аномалии?	52
2.3. Уровни изоляции в PostgreSQL	54
Read Committed	55
Repeatable Read	63
Serializable	70
2.4. Какой уровень изоляции использовать?	73

Глава 3. Страницы и версии строк	75
3.1. Структура страниц	75
Заголовок страницы	75
Специальная область	76
Версии строк	76
Указатели на версии строк	77
Свободное место	78
3.2. Структура версий строк	78
3.3. Выполнение операций над версиями строк	80
Вставка	81
Фиксация	85
Удаление	87
Отмена	88
Обновление	88
3.4. Индексы	89
3.5. TOAST	90
3.6. Виртуальные транзакции	91
3.7. Вложенные транзакции	92
Точки сохранения	92
Ошибки и атомарность операций	94
Глава 4. Снимки данных	97
4.1. Что такое снимок данных	97
4.2. Видимость версий строк в снимке	98
4.3. Из чего состоит снимок	99
4.4. Видимость собственных изменений	104
4.5. Горизонт транзакции	105
4.6. Снимок данных для системного каталога	108
4.7. Экспорт снимка данных	109
Глава 5. Внутривстраничная очистка и hot-обновления	111
5.1. Внутривстраничная очистка	111
5.2. Hot-обновления	115
5.3. Внутривстраничная очистка при hot-обновлениях	119
5.4. Разрыв hot-цепочки	121
5.5. Внутривстраничная очистка индексов	123

Глава 6. Очистка и автоочистка	125
6.1. Очистка вручную	125
6.2. Еще раз о горизонте базы данных	128
6.3. Этапы выполнения очистки	131
Сканирование таблицы	131
Очистка индексов	132
Очистка таблицы	133
Усечение таблицы	133
6.4. Анализ	134
6.5. Автоматическая очистка и анализ	134
Устройство автоочистки	135
Какие таблицы требуют очистки	136
Какие таблицы требуют анализа	138
Автоочистка в действии	139
6.6. Регулирование нагрузки	143
Управление интенсивностью обычной очистки	144
Управление интенсивностью автоочистки	144
6.7. Мониторинг очистки	145
Отслеживание выполнения ручной очистки	145
Отслеживание выполнения автоочистки	148
Глава 7. Заморозка	150
7.1. Переполнение счетчика транзакций	150
7.2. Заморозка версий и правила видимости	151
7.3. Управление заморозкой	154
Минимальный возраст для заморозки	155
Возраст для агрессивной заморозки	157
Возраст для аварийного срабатывания автоочистки	159
Возраст для приоритетного режима заморозки	161
7.4. Заморозка вручную	162
Очистка с заморозкой	162
Заморозка при загрузке	162
Глава 8. Перестроение таблиц и индексов	165
8.1. Полная очистка	165
Необходимость	165
Оценка плотности информации	166

Заморозка	170
8.2. Другие способы перестроения	171
Аналоги полной очистки	171
Перестроение без долгих блокировок	172
8.3. Профилактика	173
Читающие запросы	173
Обновление данных	173
Часть II. Буферный кеш и журнал	177
Глава 9. Буферный кеш	179
9.1. Кеширование	179
9.2. Устройство буферного кеша	180
9.3. Попадание в кеш	182
9.4. Промах кеша	187
Поиск буфера и вытеснение	188
9.5. Массовое вытеснение	190
9.6. Настройка размера	194
9.7. Прогрев кеша	197
9.8. Локальный кеш	199
Глава 10. Журнал предзаписи	201
10.1. Журналирование	201
10.2. Устройство журнала	203
Логическая структура	203
Физическая структура	206
10.3. Контрольная точка	208
10.4. Восстановление	213
10.5. Фоновая запись	216
10.6. Настройка	217
Настройка контрольной точки	217
Настройка фоновой записи	220
Мониторинг	221
Глава 11. Режимы журнала	224
11.1. Производительность	224

11.2. Надежность	228
Кеширование	229
Повреждение данных	230
Неатомарность записи	232
11.3. Уровни журнала	236
Minimal	237
Replica	239
Logical	241
 Часть III. Блокировки	 243
 Глава 12. Блокировки отношений	 245
12.1. Общие сведения о блокировках	245
12.2. Тяжелые блокировки	248
12.3. Блокировки номеров транзакций	249
12.4. Блокировки отношений	251
12.5. Очередь ожидания	254
 Глава 13. Блокировки строк	 258
13.1. Устройство	258
13.2. Режимы блокировки строки	259
Исключительные режимы	259
Разделяемые режимы	261
13.3. Мультитранзакции	262
13.4. Очередь ожидания	264
Исключительные режимы	264
Разделяемые режимы	271
13.5. Блокировка без ожидания	274
13.6. Взаимоблокировки	276
Взаимоблокировка при обновлении строк	278
Взаимоблокировка двух команд UPDATE	279
 Глава 14. Блокировки разных объектов	 283
14.1. Блокировки неотношений	283
14.2. Блокировки расширения отношения	285
14.3. Блокировки страниц	286

14.4. Рекомендательные блокировки	286
14.5. Предикатные блокировки	288
Глава 15. Блокировки в памяти	295
15.1. Спин-блокировки	295
15.2. Легкие блокировки	296
15.3. Примеры	296
Буферный кеш	296
Буферы журнала предзаписи	298
Кеш SLRU	299
15.4. Мониторинг ожиданий	300
15.5. Семплирование	302
Часть IV. Выполнение запросов	305
Глава 16. Этапы выполнения запросов	307
16.1. Демонстрационная база данных	307
16.2. Простой протокол запросов	310
Разбор	310
Трансформация	312
Планирование	314
Исполнение	325
16.3. Расширенный протокол запросов	327
Подготовка	327
Привязка параметров	328
Планирование и исполнение	329
Получение результатов	332
Глава 17. Статистика	333
17.1. Базовая статистика	333
17.2. Неопределенные значения	337
17.3. Уникальные значения	338
17.4. Наиболее частые значения	340
17.5. Гистограмма	343
17.6. Статистика для нескалярных типов данных	347
17.7. Средний размер поля	348

17.8. Корреляция	348
17.9. Статистика по выражению	349
Расширенная статистика по выражению	350
Статистика для индекса по выражению	351
17.10. Многовариантная статистика	352
Функциональные зависимости между столбцами	352
Многовариантное число различных значений	354
Многовариантные списки частых значений	356
Глава 18. Табличные методы доступа	358
18.1. Подключаемые движки хранения	358
18.2. Последовательное сканирование	361
Оценка стоимости	362
18.3. Параллельные планы выполнения	366
18.4. Параллельное последовательное сканирование	367
Оценка стоимости	367
18.5. Ограничения параллельного выполнения	371
Количество рабочих процессов	371
Нераспараллеливаемые запросы	375
Ограниченно распараллеливаемые запросы	376
Глава 19. Индексные методы доступа	381
19.1. Индексы и расширяемость	381
19.2. Классы и семейства операторов	384
Класс операторов	384
Семейство операторов	389
19.3. Интерфейс механизма индексирования	391
Свойства метода доступа	392
Свойства индекса	396
Свойства столбцов	397
Глава 20. Индексное сканирование	401
20.1. Простое индексное сканирование	401
Оценка стоимости	402
Хороший случай: высокая корреляция	403
Плохой случай: низкая корреляция	406

20.2. Сканирование только индекса	409
Include-индексы	412
20.3. Сканирование по битовой карте	414
Точность карты	415
Действия с битовыми картами	417
Оценка стоимости	418
20.4. Параллельные версии индексного сканирования	422
20.5. Сравнение методов доступа	424
Глава 21. Вложенный цикл	426
21.1. Виды и способы соединений	426
21.2. Соединение вложенным циклом	428
Декартово произведение	428
Параметризованное соединение	432
Кеширование (мемоизация) строк	436
Внешние соединения	440
Анти- и полусоединения	441
Неэквисоединения	443
Параллельный режим	444
Глава 22. Хеширование	446
22.1. Соединение хешированием	446
Однопроходное соединение хешированием	446
Двухпроходное соединение хешированием	452
Динамические корректировки плана	455
Соединение хешированием в параллельных планах	459
Параллельное однопроходное хеш-соединение	460
Параллельное двухпроходное хеш-соединение	462
Модификации	465
22.2. Группировка и уникальные значения	467
Глава 23. Сортировка и слияние	472
23.1. Соединение слиянием	472
Слияние отсортированных наборов	472
Параллельный режим	476
Модификации	477

23.2. Сортировка	478
Быстрая сортировка	480
Частичная пирамидальная сортировка	481
Внешняя сортировка	483
Инкрементальная сортировка	487
Параллельный режим	489
23.3. Группировка и уникальные значения	491
23.4. Сравнение способов соединения	494
Часть V. Типы индексов	499
Глава 24. Хеш-индекс	501
24.1. Общий принцип	501
24.2. Страничная организация	502
24.3. Класс операторов	509
24.4. Свойства	510
Свойства метода доступа	510
Свойства индекса	511
Свойства столбцов	512
Глава 25. В-дерево	513
25.1. Общий принцип	513
25.2. Поиск и вставка	514
Поиск по равенству	514
Поиск по неравенству	516
Поиск по диапазону	517
Вставка	517
25.3. Страничная организация	519
Компактное хранение дубликатов	523
Компактное хранение внутренних индексных записей	525
25.4. Класс операторов	526
Семантика сравнения	526
Сортировка и составные индексы	532
25.5. Свойства	537
Свойства метода доступа	537
Свойства индекса	538
Свойства столбцов	539

Глава 26. Индекс GiST	540
26.1. Общий принцип	540
26.2. R-дерево для точек	542
Страничная организация	545
Класс операторов	546
Поиск вхождения в область	548
Поиск ближайших соседей	550
Вставка	555
Ограничение исключения	556
Свойства	559
26.3. RD-дерево для полнотекстового поиска	562
Про полнотекстовый поиск	562
Индексация tsvector	563
Свойства	571
26.4. Другие типы данных	571
Глава 27. Индекс SP-GiST	574
27.1. Общий принцип	574
27.2. Дерево квадрантов для точек	576
Класс операторов	577
Страничная организация	581
Поиск	582
Вставка	583
Свойства	586
27.3. К-мерные деревья для точек	588
27.4. Префиксное дерево для строк	590
Класс операторов	591
Поиск	592
Вставка	593
Свойства	595
27.5. Другие типы данных	596
Глава 28. Индекс GIN	598
28.1. Общий принцип	598
28.2. Индекс для полнотекстового поиска	599
Страничная организация	601
Класс операторов	603

Поиск	605
Частые и редкие лексемы	606
Вставка	610
Ограничение выборки	612
Свойства	613
Ограничения GIN и RUM-индекс	615
28.3. Индекс для триграмм	616
28.4. Индекс для массивов	618
28.5. Индекс для JSON	622
Класс операторов jsonb_ops	622
Класс операторов jsonb_path_ops	625
28.6. Другие типы данных	627
Глава 29. Индекс BRIN	628
29.1. Общий принцип	628
29.2. Пример	629
29.3. Страничная организация	631
29.4. Поиск	633
29.5. Обновление сводной информации	634
Вставка значений	634
Обобщение зоны	635
29.6. Диапазоны значений (minmax)	636
Выбор столбцов для индексирования	637
Размер зоны и эффективность поиска	638
Свойства	642
29.7. Мультидиапазоны значений (minmax-multi)	645
29.8. Охватывающие значения (inclusion)	648
29.9. Фильтры Блума (bloom)	651
Заключение	656
Предметный указатель	657

О книге

— До чего же это все просто! — воскликнул Шпунтик. — А я где-то читал, что писателю нужен какой-то вымысел, замысел...

— Э, замысел! — нетерпеливо перебил его Смекайло. — Это только в книгах так пишется, что нужен замысел, а попробуй задумай что-нибудь, когда все уже и без тебя задумано! Что ни возьми — все уже было.

Николай Носов, *Приключения Незнайки и его друзей*

Для кого эта книга

Эта книга для тех, кого не устраивает работа с базой данных как с черным ящиком. Если вы любознательны, не довольствуетесь авторитетными советами и хотите во всем разобраться сами — нам по пути.

Я ориентируюсь на читателей, имеющих определенный опыт использования PostgreSQL и хотя бы в общих чертах представляющих себе, что к чему. Для совсем новичков текст будет тяжеловат. Например, я ни слова не скажу о том, как устанавливать сервер, вводить команды в `psql` или изменять конфигурационные параметры.

Надеюсь, что книга будет полезной и тем, кто хорошо знаком с устройством другой СУБД, но переходит на PostgreSQL и хочет разобраться в отличиях. Несколько лет назад такая книга сэкономила бы мне много времени. Именно поэтому я ее в конце концов и написал.

Чего нет в книге

Эта книга — не сборник рецептов. На все случаи жизни готовых решений не напасешься, а понимание внутренней механики сложной системы дает

возможность критически переосмысливать чужой опыт и делать свои собственные выводы. Поэтому я и объясняю такие подробности устройства, знание которых на первый взгляд не имеет практического смысла.

Но эта книга и не учебник. Она углубляется в одни области (более интересные мне самому) и обходит стороной другие. Если вы изучаете SQL, обратите внимание на учебник Евгения Моргунова *PostgreSQL. Основы языка SQL*¹, а необходимый теоретический фундамент даст книга Бориса Новикова *Основы технологий баз данных*².

Называться справочником эта книга тоже не претендует. Я старался быть точным, но у меня не было цели заменить книгой документацию, поэтому я легко опускал непринципиальные, на мой взгляд, подробности. В любой непонятной ситуации читайте документацию.

Еще эта книга не учит разрабатывать ядро PostgreSQL. Я не предполагаю у читателя знания языка C и ориентируюсь на администраторов и прикладных разработчиков. Хотя и ссылаюсь постоянно на исходный код, из которого можно узнать столько подробностей, сколько душе угодно, и даже больше.

Что в книге есть

Во вводной главе без особых деталей я даю основные понятия, на которые опирается все дальнейшее повествование. Я предполагаю, что вы не почерпнете из этой главы практически ничего нового, но все-таки включаю ее для полноты картины. К тому же она может пригодиться тем, кто переходит с других СУБД.

Первая часть книги посвящена вопросам согласованности и изоляции, которые я сперва рассматриваю с позиции пользователя (какие уровни изоляции существуют и чем это грозит), а затем с точки зрения внутреннего устройства. Для этого мне приходится погрузиться в детали реализации многоверсионности и изоляции на основе снимков данных. Особенно много внимания требует процедура очистки неактуальных версий строк.

¹ postgrespro.ru/education/books/sqlprimer.

² postgrespro.ru/education/books/dbtech.

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

e-Univers.ru