

УДК 004.655
ББК 32.973.26-018.2
Д40

Джуба С., Волков А.
Д40 Изучаем PostgreSQL 10 / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2019. – 400 с.: ил.

ISBN 978-5-97060-643-8

Прочитав книгу, вы будете хорошо понимать основы PostgreSQL 10 и обладать навыками, необходимыми для разработки эффективных решений с применением базы данных. Это хорошее пособие для близкого знакомства с PostgreSQL. С той или иной степенью полноты оно охватывает практически все вопросы, с которыми встречается разработчик и администратор, начинающий профессионально работать с этой СУБД.

Издание рекомендовано ведущими разработчиками PostgreSQL в России, оно будет полезно как начинающим разработчикам, так и действующим администраторам этой СУБД.

УДК 004.655
ББК 32.973.26-018.2

Authorized Russian translation of the English edition of Learning PostgreSQL 10, 2nd ed. ISBN 9781788392013 © 2017 Packt Publishing.

This translation is published and sold by permission of Packt Publishing Ltd., which owns or controls all rights to publish and sell the same.

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

ISBN 978-1-78839-201-3 (анг.)
ISBN 978-5-97060-643-8 (рус.)

Copyright © 2017 Packt Publishing
© Оформление, издание, перевод, ДМК Пресс, 2019

Содержание

Благодарность от редакции	5
Об авторах	13
О рецензентах	15
Предисловие	16
Глава 1. Реляционные базы данных.....	22
Системы управления базами данных	22
Историческая справка	22
Категории баз данных	23
Базы данных NoSQL.....	23
Реляционные и объектно-реляционные базы данных	25
Свойства ACID.....	26
Язык SQL	26
Понятия реляционной модели	27
Реляционная алгебра.....	33
Операции выборки и проекции	34
Операция переименования	36
Теоретико-множественные операции	36
Операция декартова произведения	36
Моделирование данных	38
Виды моделей данных.....	38
Модель сущность-связь.....	39
UML-диаграммы классов.....	44
Резюме.....	44
Глава 2. PostgreSQL в действии	46
Обзор PostgreSQL.....	46
История PostgreSQL.....	46
Преимущества PostgreSQL	47
Применения PostgreSQL.....	48
Истории успеха	49
Ответвления.....	49
Архитектура PostgreSQL.....	50
Сообщество PostgreSQL.....	52
Возможности PostgreSQL	52

Репликация	52
Безопасность.....	53
Расширения.....	54
Возможности NoSQL.....	55
Адаптеры внешних данных	56
Производительность	57
Установка PostgreSQL	58
Установка PostgreSQL с помощью менеджера пакетов APT	59
Установка PostgreSQL в Windows.....	63
Клиенты PostgreSQL	64
Резюме	69

Глава 3. Основные строительные блоки PostgreSQL..... 71

Кодирование базы данных.....	71
Соглашение об именовании объектов базы данных.....	71
Идентификаторы в PostgreSQL.....	72
Документация.....	73
Система управления версиями.....	73
Средство миграции базы данных	74
Иерархия объектов в PostgreSQL	74
Шаблонные базы данных	74
Пользовательские базы данных	75
Роли	76
Табличное пространство.....	77
Шаблонные процедурные языки.....	78
Параметры	78
Взаимодействия с объектами PostgreSQL верхнего уровня	80
Компоненты базы данных PostgreSQL	81
Схема	81
Применение схем	82
Таблица.....	83
Встроенные типы данных	84
База данных сайта торговли автомобилями	91
Резюме	94

Глава 4. Дополнительные строительные блоки PostgreSQL 96

Представления	96
Синтаксис определения представления	98
Категории представлений.....	99
Материализованные представления.....	99
Обновляемые представления	100
Индексы.....	102
Синтаксис создания индекса	103

8 ❖ Содержание

Избирательность индекса	103
Типы индексов	106
Категории индексов	106
Рекомендации по работе с индексами	108
Функции	109
Встроенные языки программирования PostgreSQL	110
Создание функции на языке C	110
Применение функций	112
Зависимости между функциями	112
Категории функций в PostgreSQL	113
Анонимные функции в PostgreSQL	114
Пользовательские типы данных	114
Триггеры и правила	118
Правила	118
Триггеры	120
Резюме	127
Глава 5. Язык SQL	129
Основы SQL	129
Лексическая структура SQL	131
Запрос данных командой SELECT	134
Структура запроса SELECT	134
Список выборки	136
Фраза FROM	142
Фраза WHERE	148
Группировка и агрегирование	152
Упорядочение и ограничение количества результатов	155
Подзапросы	156
Теоретико-множественные операции – UNION, EXCEPT, INTERSECT	158
Значения NULL	159
Изменение данных в базе	162
Команда INSERT	162
Команда UPDATE	164
Команда DELETE	166
Команда TRUNCATE	167
Резюме	167
Глава 6. Дополнительные сведения о написании запросов	168
Общие табличные выражения	168
СТЕ как средство повторного использования SQL-кода	170
Рекурсивные и иерархические запросы	172
Изменение данных сразу в нескольких таблицах	176
Оконные функции	178

Определение окна	179
Фраза WINDOW	180
Использование оконных функций	181
Оконные функции с группировкой и агрегированием	183
Продвинутые методы работы с SQL	184
Выборка первых записей	184
Извлечение выборочных данных	185
Функции, возвращающие множества	186
Латеральные подзапросы	189
Дополнительные средства группировки	191
Дополнительные виды агрегирования	193
Резюме	195

Глава 7. Серверное программирование на PL/pgSQL

Сравнение языков SQL и PL/pgSQL	196
Параметры функций в PostgreSQL	197
Параметры функций, относящиеся к авторизации	197
Параметры функции, относящиеся к планировщику	199
Параметры функции, относящиеся к конфигурации	202
Команды управления в PL/pgSQL	203
Объявления	203
Присваивание	205
Условные команды	207
Итерирование	209
Возврат из функции	212
Предопределенные переменные в функциях	215
Обработка исключений	216
Динамический SQL	218
Динамическое выполнение команд DDL	218
Динамическое выполнение команд DML	219
Динамический SQL и кеширование	220
Рекомендации по использованию динамического SQL	220
Резюме	222

Глава 8. OLAP и хранилища данных

Оперативная аналитическая обработка	224
Извлечение, преобразование и загрузка	225
Моделирование данных для OLAP	228
Агрегирование	230
Секционирование	231
Параллельные запросы	235
Просмотр только индексов	236
Резюме	238

Глава 9. За пределами традиционных типов данных	239
Массивы	240
Функции и операторы массивов	243
Доступ к элементам массива и их модификация	244
Индексирование массивов	245
Хранилище ключей и значений	246
Индексирование hstore	248
Структура данных JSON	249
JSON и XML	249
Типы данных JSON в PostgreSQL	250
Доступ к объектам типа JSON и их модификация	250
Индексирование JSON-документов	252
Реализация REST-совместимого интерфейса к PostgreSQL	253
Полнотекстовый поиск в PostgreSQL	257
Типы данных tsquery и tsvector	257
Сопоставление с образцом	258
Полнотекстовые индексы	260
Резюме	261
Глава 10. Транзакции и управление параллельным доступом	262
Транзакции	262
Транзакции и свойства ACID	263
Транзакции и конкурентность	264
Уровни изоляции транзакций	267
Явная блокировка	272
Блокировка на уровне таблиц	273
Блокировка на уровне строк	276
Взаимоблокировки	277
Рекомендательные блокировки	278
Резюме	279
Глава 11. Безопасность в PostgreSQL	281
Аутентификация в PostgreSQL	281
Файл pg_hba.conf	283
Прослушиваемые адреса	284
Рекомендации по аутентификации	284
Привилегии доступа по умолчанию	285
Система ролей и прокси-аутентификация	286
Уровни безопасности в PostgreSQL	288
Безопасность на уровне базы данных	288
Безопасность на уровне схемы	289

Безопасность на уровне таблицы	289
Безопасность на уровне столбца	290
Безопасность на уровне строк	290
Шифрование данных	293
Шифрование паролей полей в PostgreSQL	293
Расширение pgcrypto	293
Резюме	297
Глава 12. Каталог PostgreSQL	298
Системный каталог	298
Системный каталог для администраторов	301
Получение версии кластера баз данных и клиентских программ	301
Завершение и отмена пользовательского сеанса	301
Задание и получение параметров кластера баз данных	302
Получение размера базы данных и объекта базы данных	304
Очистка базы данных	305
Очистка данных в базе	308
Оптимизация производительности	310
Избирательная выгрузка	311
Резюме	314
Глава 13. Оптимизация производительности базы данных	315
Настройка конфигурационных параметров PostgreSQL	316
Максимальное количество подключений	316
Параметры памяти	316
Параметры жесткого диска	317
Параметры планировщика	317
Эталонное тестирование вам в помощь	318
Оптимизация производительности записи	318
Оптимизация производительности чтения	321
План выполнения и команда EXPLAIN	322
Обнаружение проблем в планах выполнения запросов	326
Типичные ошибки при написании запросов	329
Избыточные операции	329
Индексы отсутствуют или построены не так	329
Использование CTE без необходимости	333
Использование процедурного языка PL/pgSQL	333
Межстолбцовая корреляция	334
Секционирование таблиц	336
Недостатки механизма исключения в силу ограничений	336
Переписывание запросов	337
Резюме	338

Глава 14. Тестирование	339
Автономное тестирование	339
Специфика автономного тестирования в базе данных	340
Фреймворки юнит-тестирования	343
Различие схем	345
Интерфейсы абстрагирования базы данных	346
Отличия в данных	347
Тестирование производительности	350
Резюме	352
Глава 15. PostgreSQL в приложениях на Python	353
Python DB API 2.0	354
Низкоуровневый доступ к базе данных с помощью psycopg2	355
Соединение с базой данных	357
Пул соединений	358
Выполнение SQL-команд	359
Чтение данных из базы	361
Команда COPY	361
Асинхронный доступ	362
Альтернативные драйверы для PostgreSQL	363
pg8000	363
asyncpg	364
SQLAlchemy – библиотека объектно-реляционного отображения	366
Основные компоненты SQLAlchemy	367
Подключение к базе и выборка данных с помощью языка SQL Expression	367
ORM	369
Резюме	373
Глава 16. Масштабируемость	374
Проблема масштабируемости и теорема CAP	375
Репликация данных в PostgreSQL	377
Журнал транзакций	377
Физическая репликация	378
Логическая репликация	384
Применение репликации для масштабирования PostgreSQL	387
Масштабирование на большое количество запросов	388
Разделение данных	389
Масштабирование с ростом числа подключений	391
Резюме	392
Предметный указатель	394