

УДК 004.65
ББК 32.972.134
074

074 Осипов Д. Л.

Технологии проектирования баз данных. – М.: ДМК Пресс, 2019. – 498 с.: ил.

ISBN 978-5-97060-737-4

Книга основана на материалах лекций и практических занятий, подготовленных автором и объединяет теоретические основы и практический аспект разработки современных баз данных (БД).

Основная задача издания — предоставить читателю профессиональную методику проектирования БД. Страницы книги проведут читателя по всем этапам жизненного цикла проекта баз данных от момента возникновения идеи разработки программного обеспечения до этапа ввода готового продукта в эксплуатацию, подробно объясняя каждый шаг.

Издание отличается глубиной и ясностью изложения материала, поэтому издание окажется полезным как для студентов и преподавателей ИТ-специальностей, так и для разработчиков баз данных и программистов, стремящихся самостоятельно освоить технологические приемы проектирования современных БД.

УДК 004.65
ББК 32.972.134

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.

Материал, изложенный в данной книге, многократно проверен. Но, поскольку вероятность технических ошибок все равно существует, издательство не может гарантировать абсолютную точность и правильность приводимых сведений. В связи с этим издательство не несет ответственности за возможные ошибки, связанные с использованием книги.

ISBN 978-5-97060-737-4

© Осипов Д. Л., 2019
© Оформление, издание ДМК Пресс, 2019

Оглавление

Введение	12
Соглашения.....	13
Глава 1. Эволюция баз данных	14
Электронные картотеки.....	16
Принцип построения систем файлов.....	17
Недостатки систем файлов	18
Пути устранения недостатков систем файлов.....	22
Что такое база данных?.....	23
Эволюция моделей БД.....	24
Необходимость моделирования	26
Иерархическая модель.....	27
Сетевая модель.....	30
Попытки разработки стандарта БД.....	32
Реляционная модель.....	34
Объектно-ориентированная модель	36
Слабоструктурированные данные	38
Документ-ориентированная модель	38
Резюме.....	39
Вопросы для самопроверки	39
Глава 2. Система управления базами данных.....	41
Функционал СУБД.....	42
Компоненты СУБД.....	45
Системный каталог	48
Архитектурные решения доступа к БД	48
Файл-сервер.....	49
Клиент-сервер	50
Распределенная система.....	54
Резюме.....	55
Вопросы для самопроверки	56
Глава 3. Персонал и пользователи БД.....	57
Администратор данных	59
Администратор базы данных.....	60
Разработчики баз данных.....	61
Прикладные программисты.....	61

4 ❖ Оглавление

Конечные пользователи	62
Резюме.....	63
Вопросы для самопроверки	63
Глава 4. Реляционная модель	65
Сущность и атрибуты.....	66
Тип данных и домен.....	69
Связь.....	71
Отношение.....	73
Ключи.....	76
Целостность данных.....	77
Целостность доменов	78
Целостность сущностей.....	79
Ссылочная целостность.....	80
Корпоративная целостность	80
Реляционная алгебра	81
Резюме.....	87
Вопросы для самопроверки	88
Глава 5. Технология разработки БД	89
Роль БД на предприятии	90
Жизненный цикл базы данных	94
Этап планирования разработки БД.....	96
Этап определения и анализа требований к системе	96
Этап проектирования БД	100
Этап выбора СУБД	104
Этап создания клиентского программного обеспечения.....	105
Этап тестирования и отладки.....	107
Этап реализации.....	109
Этап эксплуатации и сопровождения.....	110
Резюме.....	111
Вопросы для самопроверки	111
Глава 6. Концептуальное проектирование и ER-модель.....	112
Концептуальная модель БД	113
ER-модель.....	113
Типы сущностей и атрибуты.....	114
Связи в ER-модели.....	119
Вариации ER-моделей.....	127
Резюме.....	129
Вопросы для самопроверки	130

Глава 7. Логическое проектирование и нормализация	131
Первая нормальная форма	134
Функциональная зависимость атрибутов	137
Порядок определения первичного ключа	139
Вторая нормальная форма	141
Третья нормальная форма	142
Нормальная форма Бойса-Кодда	145
Четвертая нормальная форма	145
Пятая нормальная форма	147
Финал «гонки» нормальных форм	149
Резюме	149
Вопросы для самопроверки	150
Глава 8. Физическое представление данных	151
Двухуровневая модель хранения данных	151
Представление реляционных данных	153
Поля	154
Записи	156
Блоки	157
Файл	157
Модификация записей	158
Особенности представления объектов	158
Журнальная информация	159
Резюме	160
Вопросы для самопроверки	160
Глава 9. Индексирование	162
Индексы на основе хеширования	164
Индексы на основе В-деревьев	168
Битовые индексы	171
Правила назначения пользовательских индексов	172
Избирательность индекса	173
Резюме	175
Вопросы для самопроверки	175
Глава 10. Безопасность данных	176
Откуда исходят угрозы?	178
Политика безопасности	181
Правила защиты БД	182

Идентификация, аутентификация и авторизация.....	184
Криптографическая защита.....	185
Резервное копирование и восстановление.....	188
Аудит событий безопасности.....	189
Модернизация программного обеспечения.....	191
Безопасный доступ к данным.....	191
Экономическая оправданность.....	192
Резюме.....	192
Вопросы для самопроверки.....	193
Глава 11. Знакомимся с SQL.....	194
Возможности SQL.....	197
Типы данных SQL.....	198
Предопределенные типы.....	199
Непредопределенные типы.....	206
Константы.....	209
Преобразование данных.....	210
Операторы.....	212
Операция присваивания.....	213
Арифметические операторы.....	213
Логические операторы.....	214
Операторы сравнения.....	215
Проверка на неопределенность NULL.....	215
Конкатенация строк.....	216
Встроенные функции.....	216
Резюме.....	217
Вопросы для самопроверки.....	217
Глава 12. Манипулирование данными SQL.....	219
Запрос, инструкция SELECT.....	219
Псевдонимы имен столбцов и таблиц.....	222
Порядок сортировки, ORDER BY.....	223
Условие отбора данных, предложение WHERE.....	224
Агрегирующие функции.....	232
Группировка данных GROUP BY.....	233
Соединение таблиц в запросе SELECT.....	234
Вставка, инструкция INSERT.....	244
Модификация, инструкция UPDATE.....	246
Удаление, инструкция DELETE.....	248
Слияние данных, инструкция MERGE.....	249
Резюме.....	250
Вопросы для самопроверки.....	251

Глава 13. Определение данных средствами SQL	252
Базы данных (схемы)	252
Домены	255
Таблицы	256
Внешние ключи и связи между таблицами	258
Ограничения на значения столбцов	260
Столбец-перечисление	261
Столбец-множество	262
Временные таблицы	263
Модификация таблицы	264
Клонирование и копирование таблиц	265
Индексы	266
Изменение индекса	267
Удаление индекса	268
Представления	269
Изменение представления	272
Удаление представления	273
Модифицируемые представления	273
Резюме	274
Вопросы для самопроверки	275
Глава 14. Процедурный SQL	276
Элементы процедурного SQL	278
Переменные	278
Составной оператор BEGIN..END	281
Условные операторы	281
Циклы	284
Хранимые процедуры и функции	288
Вызов хранимой процедуры	291
Особенности работы с функциями	292
Изменение процедур и функций	294
Удаление процедур и функций	294
Триггеры	294
Контекстные переменные	297
Примеры триггеров	297
Курсоры	302
Примеры курсоров	305
Резюме	309
Вопросы для самопроверки	310
Глава 15. Регулярные выражения в запросах	311
Операторы для регулярных выражений	311

Основы синтаксиса.....	312
Регулярные выражения в запросах.....	319
Резюме.....	320
Вопросы для самопроверки	320
Глава 16. Управление транзакциями	322
Требования к транзакции	323
Состояние транзакции.....	324
Проблемы совместного доступа к данным	325
Управление параллельными транзакциями.....	326
Пессимистический подход	327
Оптимистический подход.....	330
Детализация уровня блокировок	333
Требования стандарта SQL.....	334
Явное управление транзакцией.....	335
Точки сохранения.....	338
Резюме.....	338
Вопросы для самопроверки	339
Глава 17. Определение прав пользователей	340
Идентификатор авторизации.....	341
Объекты защиты	342
Управление наборами привилегий	343
Предоставление привилегий	344
Лишение привилегий	347
Резюме.....	349
Вопросы для самопроверки	349
Глава 18. Интерактивная аналитическая обработка OLAP	350
Требования к OLAP-инструментам	351
Хранилище данных.....	353
OLAP-куб.....	355
Язык многомерных выражений MDX.....	356
Резюме.....	358
Вопросы для самопроверки	358
Глава 19. Расширяемый язык разметки XML	360
Корректность документа XML.....	361
Построение документа XML	361
Элементы документа	362
Атрибуты	363

Пространство имен.....	364
Определение типа документа DTD	367
XML Schemas.....	373
Элементы схемы	376
Атрибуты схемы	385
Подключение XML-схемы к документу.....	386
Поддержка XML в СУБД	386
Резюме.....	387
Вопросы для самопроверки	388
Глава 20. Клиент-серверные БД.....	389
Модель взаимодействия открытых систем.....	389
Клиент-серверные СУБД	393
Модели распределения функций	393
Резюме.....	396
Вопросы для самопроверки	397
Глава 21. Особенности разработки клиента БД.....	398
Выбор языка программирования.....	398
Технология доступа к данным ODBC.....	399
Технология доступа к данным ADO .NET.....	400
Технология доступа к данным FireDAC.....	402
Технология JDBC	404
Интерфейс клиента.....	405
Сколько людей, столько и мнений	406
Пользовательские критерии качества интерфейса	407
Рекомендации по проектированию	408
Резюме.....	410
Вопросы для самопроверки	411
Глава 22. Распределенные БД	412
Предпосылки децентрализации	412
Система управления распределенной базой данных.....	414
Правила распределенных БД от Криса Дейта	415
Аспекты проектирования распределенных БД	416
Фрагментация.....	417
Распределение	419
Репликация	419
Особенности управления системным каталогом.....	420
Распределенные транзакции	420
Преимущества распределенных БД.....	421

Недостатки распределенных БД.....	422
Резюме.....	423
Вопросы для самопроверки	424
Глава 23. Объектно-ориентированная модель данных.....	425
Предпосылки появления модели	425
Преимущества ООБД.....	427
Объектно-ориентированная терминология	428
Абстрагирование.....	430
Инкапсуляция.....	430
Модульность.....	431
Наследование.....	432
Идентификатор объекта	432
Манифест объектно-ориентированных СУБД.....	433
Стандарт ODMG.....	437
Что было сделано на практике?.....	437
Postgres	437
UniSQL.....	438
Cache	439
Versant Object Database.....	439
ObjectStore.....	440
Что пошло не так?.....	440
Недостатки ООБД.....	441
Объектно-реляционные СУБД.....	443
Резюме.....	444
Вопросы для самопроверки	445
Глава 24. Документ-ориентированные БД.....	446
Чем плоха нормализация?.....	446
БД ключ-значение	447
Документ-ориентированные БД.....	448
NoSQL.....	449
Распределенная обработка MapReduce	452
Сегментирование	453
Репликация	455
Когда следует использовать документ-ориентированную модель?	456
Резюме.....	456
Вопросы для самопроверки	457
Глава 25. Большие данные	458
Что такое «большие данные»?.....	459

Принципы работы с большими данными	461
Лямбда-архитектура	461
Apache Hadoop	463
Apache Storm	465
Apache Impala	466
Apache Kafka	466
NewSQL	467
Добыча данных	468
Резюме	469
Вопросы для самопроверки	470
Глава 26. Составление программной документации	471
Виды программных документов	472
Техническое задание	473
Пояснительная записка	475
Эксплуатационные документы	476
Руководство системного программиста	477
Руководство оператора	478
Документация в тексте программы	479
Резюме	480
Вопросы для самопроверки	481
Приложение 1. Модель БД «Склад»	482
Приложение 2. Пример XML-схемы	483
Приложение 3. Стандарты по единой системе программной документации	487
Список литературы	489
Предметный указатель	493