

УДК 004.6:004.42 MongoDB

ББК 32.973.26-018.2

Б71

Бэнкер, Кайл.

Б71 MongoDB в действии / К. Бэнкер ; пер. с англ. А. А. Слинкина. — 2-е изд., эл. — 1 файл pdf : 396 с. — Москва : ДМК Пресс, 2023. — Систем. требования: Adobe Reader XI либо Adobe Digital Editions 4.5 ; экран 10". — Текст : электронный.

ISBN 978-5-89818-344-8

MongoDB — это документо-ориентированная база данных, предназначенная для гибкой, масштабируемой и очень быстрой работы даже при больших объемах данных. При ее проектировании изначально закладывалась высокая доступность, поддержка сложных динамических схем и простое распределение данных по нескольким серверам.

Эта книга представляет собой введение в MongoDB и документо-ориентированную модель данных. Она дает не только общую картину, необходимую разработчику, но и достаточно деталей, чтобы удовлетворить системного инженера. Многочисленные примеры помогут обрести уверенность в области моделирования данных — вопросе, который необычайно важен для разработки ПО. Вам понравится углубленное изложение различных функциональных возможностей, в том числе репликации, автосегментирования и разветвления.

УДК 004.6:004.42 MongoDB

ББК 32.973.26-018.2

Электронное издание на основе печатного издания: MongoDB в действии / К. Бэнкер ; пер. с англ. А. А. Слинкина — Москва : ДМК Пресс, 2014. — 394 с. — ISBN 978-5-97060-057-3. — Текст : непосредственный.

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.

В соответствии со ст. 1299 и 1301 ГК РФ при устранении ограничений, установленных техническими средствами защиты авторских прав, правообладатель вправе требовать от нарушителя возмещения убытков или выплаты компенсации.

ISBN 978-5-89818-344-8

© 2014 by Manning Publications Co.

© Оформление, перевод на русский язык
ДМК Пресс, 2014



ОГЛАВЛЕНИЕ

Предисловие	13
Благодарности.....	15
Об этой книге.....	16
Об иллюстрации на обложке.....	20

ЧАСТЬ 1.

Приступая к работе	21
---------------------------------	-----------

Глава 1. База данных для современной веб..... 22

1.1. Рожденная в облаке	24
1.2. Основные особенности MongoDB	25
1.2.1. Документная модель данных.....	25
1.2.2. Произвольные запросы	28
1.2.3. Вторичные индексы	31
1.2.4. Репликация.....	32
1.2.5. Быстродействие и долговечность	32
1.2.6. Масштабирование	34
1.3. Сервер и инструментальные средства MongoDB	36
1.3.1. Сервер.....	37
1.3.2. JavaScript-оболочка	38
1.3.3. Языковые драйверы.....	39
1.3.4. Командные утилиты	40
1.4. Почему именно MongoDB?	41
1.4.1. Сравнение MongoDB с другими СУБД.....	41
1.4.2. Сценарии использования и примеры реального развертывания	46
1.5. Советы и ограничения.....	48
1.6. Резюме	49

Глава 2. MongoDB сквозь призму

JavaScript-оболочки.....	51
---------------------------------	-----------

2.1. Первое знакомство с оболочкой MongoDB.....	52
---	----

2.1.1. Запуск оболочки	52
2.1.2. Вставка и выборка	52
2.1.3. Обновление документов	55
2.1.4. Удаление данных	57
2.2. Создание индексов и применение их в запросах	58
2.2.1. Создание большой коллекции	58
2.2.2. Индексирование и команда explain()	60
2.3. Основы администрирования	63
2.3.1. Получение информации о базе данных	63
2.3.2. Как работают команды	65
2.4. Получение справки	66
2.5. Резюме	67

Глава 3. Разработка программ для MongoDB... 69

3.1. MongoDB сквозь призму Ruby	70
3.1.1. Установка и подключение к базе	70
3.1.2. Вставка документов на Ruby	71
3.1.3. Запросы и курсоры	73
3.1.4. Обновление и удаление	74
3.1.5. Команды базы данных	75
3.2. Как работают драйверы	76
3.2.1. Генерация идентификатора объекта	76
3.2.2. Формат BSON	78
3.2.3. Передача по сети	80
3.3. Разработка простого приложения	81
3.3.1. Подготовка	82
3.3.2. Сбор данных	83
3.3.3. Визуализация архива	85
3.4. Резюме	88

ЧАСТЬ 2.

Разработка приложения для MongoDB 89

Глава 4. Документо-ориентированные данные 91

4.1. Принципы проектирования схемы	92
4.2. Проектирование модели данных для интернет-магазина ..	94
4.2.1. Товары и категории	95
4.2.2. Пользователи и заказы	100
4.2.3. Отзывы	102
4.3. Технические детали: о базах данных, коллекциях и документах	104
4.3.1. Базы данных	104

4.3.2. Коллекции	107
4.3.3. Документы и вставка.....	111
4.4. Резюме	117

Глава 5. Запросы и агрегирование 118

5.1. Запросы в приложении для интернет-магазина	119
5.1.1. Товары, категории и обзоры	119
5.1.2. Пользователи и заказы	122
5.2. Язык запросов MongoDB.....	124
5.2.1. Селекторы запроса.....	124
5.2.2. Дополнительные средства.....	136
5.3. Агрегирование заказов	139
5.3.1. Группировка отзывов по пользователям	139
5.3.2. Распределение-редукция для заказов по регионам	141
5.4. Агрегирование в деталях.....	143
5.4.1. Максимум и минимум	143
5.4.2. Команда distinct	144
5.4.3. Команда group	144
5.4.4. Map-reduce.....	146
5.5. Резюме	149

Глава 6. Обновление, атомарные операции и удаление 150

6.1. Краткий экскурс в обновление документов	151
6.2. Обновление данных интернет-магазина	154
6.2.1. Товары и категории	154
6.2.2. Отзывы	159
6.2.3. Заказы	160
6.3. Атомарная обработка документа	164
6.3.1. Переходы состояний заказа.....	164
6.3.2. Управление запасами	166
6.4. Технические детали: обновление и удаление в MongoDB.....	171
6.4.1. Типы и параметры операций обновления.....	171
6.4.2. Операторы обновления	173
6.4.3. Команда findAndModify	178
6.4.4. Операции удаления	179
6.4.5. Параллелизм, атомарность и изолированность	179
6.4.6. Замечания о производительности обновления	180
6.5. Резюме	182

ЧАСТЬ 3.

MongoDB – постижение мастерства..... 183

Глава 7. Индексирование и оптимизация запросов	185
7.1. Теория индексирования	186
7.1.1. Мысленный эксперимент	186
7.1.2. Основные понятия индексирования	190
7.1.3. В-деревья	194
7.2. Индексирование на практике	196
7.2.1. Типы индексов	196
7.2.2. Администрирование индексов	199
7.3. Оптимизация запросов	204
7.3.1. Выявление медленных запросов	205
7.3.2. Исследование медленных запросов	208
7.3.3. Образцы запросов	216
7.4. Резюме	218
Глава 8. Репликация	219
8.1. Обзор репликации	219
8.1.1. Почему так важна репликация	220
8.1.2. Сценарии репликации	221
8.2. Наборы реплик	223
8.2.1. Настройка	223
8.2.2. Как работает репликация	229
8.2.3. Администрирование	236
8.3. Репликация типа главный-подчиненный	247
8.4. Драйверы и репликация	247
8.4.1. Подключение и отработка отказов	248
8.4.2. Гарантии записи	250
8.4.3. Масштабирование чтения	252
8.4.4. Тегирование	254
8.5. Резюме	256
Глава 9. Сегментирование	257
9.1. Обзор сегментирования	258
9.1.1. Что такое сегментирование	258
9.1.2. Как работает сегментирование	260
9.2. Тестовый сегментированный кластер	266
9.2.1. Настройка	266
9.2.2. Запись в сегментированный кластер	272
9.3. Индексирование сегментированного кластера и запросы к нему	278
9.3.1. Типы сегментированных запросов	278
9.3.2. Индексирование	283
9.4. Выбор сегментного ключа	284

9.4.1. Неэффективные сегментные ключи	284
9.4.2. Идеальные сегментные ключи	287
9.5. Сегментирование в производственных системах	288
9.5.1. Развертывание и конфигурирование	288
9.5.2. Администрирование	294
9.6. Резюме	300

Глава 10. Развертывание

и администрирование 301

10.1. Развертывание.....	302
10.1.1. Среда развертывания	302
10.1.2. Конфигурирование сервера.....	308
10.1.3. Импорт и экспорт данных.....	310
10.1.4. Безопасность.....	312
10.2. Мониторинг и диагностика	315
10.2.1. Протоколирование.....	315
10.2.2. Средства мониторинга	316
10.2.3. Внешние приложения для мониторинга	320
10.2.4. Диагностические средства (mongosniff, bsondump) ..	321
10.3. Обслуживание.....	322
10.3.1. Резервное копирование и восстановление	322
10.3.2. Сжатие и ремонт	324
10.3.3. Модернизация	326
10.4. Разрешение проблем, связанных с производительностью.....	326
10.4.1. Проверка эффективности индексов и запросов	327
10.4.2. Добавление памяти	328
10.4.3. Повышение производительности дисков	329
10.4.4. Горизонтальное масштабирование.....	330
10.4.5. Обращение к профессионалам	330
10.5. Резюме	330

ПРИЛОЖЕНИЕ А.

Установка..... 332

A.1. Установка	332
A.1.1. MongoDB в Linux	332
A.1.2. MongoDB в Mac OS X	334
A.1.3. MongoDB в Windows.....	336
A.1.4. Компилирование MongoDB из исходного кода.....	337
A.1.5. Поиск и устранение неполадок	337
A.2. Основные конфигурационные параметры.....	339
A.3. Установка Ruby	340

A.3.1. Linux и Mac OS X	340
A.3.2. Windows.....	341

ПРИЛОЖЕНИЕ В.

Паттерны проектирования 342

B.1. Паттерны	342
B.1.1. Вложение или ссылка	342
B.1.2. Связь один-ко-многим	343
B.1.3. Связь многие-ко-многим.....	344
B.1.4. Деревья.....	345
B.1.5. Очереди	348
B.1.6. Динамические атрибуты.....	349
B.1.7. Транзакции.....	351
B.1.8. Локальность и предвычисления.....	352
B.2. Антипаттерны	353
B.2.1. Непродуманное индексирование	353
B.2.2. Смешанные типы.....	354
B.2.3. Коллекции-свалки	354
B.2.4. Большие документы с глубокой вложенностью.....	354
B.2.5. Одна коллекция на каждого пользователя	355
B.2.6. Несегментируемые коллекции	355

ПРИЛОЖЕНИЕ С.

Двоичные данные и GridFS 356

C.1. Хранение простых двоичных объектов	357
C.1.1. Хранение миниатюр	357
C.1.2. Хранение MD5-свертки	358
C.2. GridFS.....	359
C.2.1. GridFS в Ruby.....	359
C.2.2. Доступ к GridFS с помощью mongofiles	362

ПРИЛОЖЕНИЕ D.

MongoDB на PHP, Java и C++ 364

D.1. PHP	365
D.1.1. Документы	365
D.1.2. Подключение.....	365
D.1.3. Пример программы	366
D.2. Java	367
D.2.1. Документы	367
D.2.2. Подключение.....	368
D.2.3. Пример программы.....	368
D.3. C++.....	370

D.3.1. Документы	370
D.3.2. Подключение.....	371
D.3.3. Пример программы	372

ПРИЛОЖЕНИЕ Е.

Пространственные индексы 374

E.1. Основы пространственного индексирования	375
E.2. Более сложные запросы	377
E.3. Составные пространственные индексы.....	378
E.4. Сферическая геометрия	379

Предметный указатель 381