

Spis treści

Przedmowa	13
Podziękowania	15
O książce	17
CZĘŚĆ I ROZPOCZĘCIE PRACY	21
Rozdział 1. Baza danych dla nowoczesnej sieci WWW	23
1.1. Serwer opracowany na potrzeby internetu	26
1.2. Kluczowe funkcje MongoDB	27
1.2.1. Model danych oparty na dokumencie	27
1.2.2. Zapytania ad hoc	31
1.2.3. Indeksy	31
1.2.4. Replikacja	32
1.2.5. Szybkość działania i niezawodność	33
1.2.6. Skalowanie	35
1.3. Podstawowy serwer MongoDB i jego narzędzia	36
1.3.1. Podstawowy serwer	37
1.3.2. Konsola JavaScript	38
1.3.3. Sterowniki bazy danych	39
1.3.4. Narzędzia powłoki	39
1.4. Dlaczego MongoDB?	40
1.4.1. MongoDB kontra inne bazy danych	41
1.4.2. Przykłady użycia i wdrożeń w środowiskach produkcyjnych	45
1.5. Wskazówki i ograniczenia	47
1.6. Historia MongoDB	49
1.7. Zasoby dodatkowe	52
1.8. Podsumowanie	53
Rozdział 2. MongoDB i konsola JavaScript	55
2.1. Zagłębiamy się w powłokę MongoDB	56
2.1.1. Uruchomienie powłoki	56
2.1.2. Bazy danych, kolekcje i dokumenty	57
2.1.3. Wstawianie i pobieranie danych	58
2.1.4. Uaktualnienie dokumentu	60
2.1.5. Usunięcie danych	64
2.1.6. Inne funkcje powłoki	65

2.2. Tworzenie indeksów i użycie ich w zapytaniach	66
2.2.1. Utworzenie ogromnej kolekcji	66
2.2.2. Indeksowanie i metoda explain()	68
2.3. Podstawowa administracja serwerem bazy danych	73
2.3.1. Pobieranie informacji o bazie danych	73
2.3.2. Jak działają polecenia?	75
2.4. Uzyskiwanie pomocy	76
2.5. Podsumowanie	77

Rozdział 3. Tworzenie programów używających MongoDB **79**

3.1. MongoDB przez pryzmat języka Ruby	80
3.1.1. Instalacja sterownika i nawiązanie połączenia z MongoDB	80
3.1.2. Wstawianie dokumentów za pomocą języka Ruby	82
3.1.3. Zapytania i kursory	83
3.1.4. Operacje uaktualnienia i usunięcia	84
3.1.5. Polecenia bazy danych	85
3.2. Jak działają sterowniki?	86
3.2.1. Generowanie identyfikatora obiektu	87
3.3. Utworzenie prostej aplikacji	89
3.3.1. Konfiguracja	90
3.3.2. Zbieranie danych	91
3.3.3. Wyświetlenie zawartości archiwum	94
3.4. Podsumowanie	96

CZĘŚĆ II TWORZENIE APLIKACJI w MONGODB **99**

Rozdział 4. Dane oparte na dokumentach **101**

4.1. Reguły projektowe schematu	102
4.2. Opracowanie modelu danych dla aplikacji typu e-commerce	104
4.2.1. Podstawy schematu	105
4.2.2. Użytkownicy i zamówienia	109
4.2.3. Opinie	112
4.3. Bazy danych, kolekcje i dokumenty w szczegółach	113
4.3.1. Bazy danych	113
4.3.2. Kolekcje	117
4.3.3. Dokumenty i operacje wstawiania	121
4.4. Podsumowanie	127

Rozdział 5. Tworzenie zapytań **129**

5.1. Zapytania w aplikacji typu e-commerce	130
5.1.1. Produkty, kategorie i opinie o produktach	130
5.1.2. Użytkownicy i zamówienia	133
5.2. Język zapytań w MongoDB	134
5.2.1. Selektory i kryteria zapytania	135
5.2.2. Opcje zapytania	149

5.3. Podsumowanie	152
Rozdział 6. Agregacja	153
6.1. Ogólne omówienie frameworka agregacji	154
6.2. Przykład agregacji w aplikacji typu e-commerce	156
6.2.1. Produkty, kategorie i opinie	157
6.2.2. Użytkownik i zamówienie	164
6.3. Operatory potoku agregacji	168
6.3.1. Operator \$project	168
6.3.2. Operator \$group	169
6.3.3. \$match, \$sort, \$skip i \$limit	171
6.3.4. Operator \$unwind	171
6.3.5. Operator \$out	172
6.4. Modyfikacja dokumentów	172
6.4.1. Funkcje ciągu tekstowego	173
6.4.2. Funkcje arytmetyczne	174
6.4.3. Funkcje daty i godziny	175
6.4.4. Funkcje logiczne	175
6.4.5. Operatory zbioru	176
6.4.6. Pozostałe funkcje	177
6.5. Wydajność działania potoku agregacji	178
6.5.1. Opcje potoku agregacji	179
6.5.2. Funkcja explain() frameworka agregacji	179
6.5.3. Opcja allowDiskUse	183
6.5.4. Opcja cursor w agregacji	184
6.6. Inne możliwości agregacji	185
6.6.1. Funkcje .count() i .distinct()	185
6.6.2. Funkcja modelu MapReduce	185
6.7. Podsumowanie	188
Rozdział 7. Uaktualnienia, operacje niepodzielne i usunięcia	191
7.1. Krótkie omówienie procesu uaktualnienia dokumentu	192
7.1.1. Modyfikacja przez zastąpienie	193
7.1.2. Modyfikacja za pomocą operatora	193
7.1.3. Porównanie obu metod	194
7.1.4. Podjęcie decyzji — zastąpienie kontra operatory	194
7.2. Uaktualnienia w modelu typu e-commerce	196
7.2.1. Produkty i kategorie	196
7.2.2. Opinie o produkcie	201
7.2.3. Zamówienia	203
7.3. Niepodzielne przetwarzanie dokumentu	206
7.3.1. Zmiana stanu zamówienia	207
7.3.2. Zarządzanie produktami	209
7.4. Usunięcia i uaktualnienia w MongoDB w szczegółach	215
7.4.1. Opcje i typy uaktualnień	215

7.4.2. Operatory uaktualnienia	216
7.4.3. Polecenie findAndModify()	225
7.4.4. Usunięcie dokumentu	225
7.4.5. Współbieżność, niepodzielność i izolacja	226
7.4.6. Uwagi dotyczące wydajności uaktualnienia	227
7.5. Przegląd operatorów uaktualnienia	229
7.6. Podsumowanie	230

CZĘŚĆ III ZAAWANSOWANE MONGODB 231

Rozdział 8. Indeksowanie i optymalizacja zapytania 233

8.1. Teoria indeksowania	234
8.1.1. Prosty eksperyment	234
8.1.2. Podstawowe koncepcje indeksowania	238
8.1.3. Struktura B-tree	242
8.2. Indeksowanie w praktyce	244
8.2.1. Typy indeksów	244
8.2.2. Administracja indeksem	249
8.3. Optymalizacja zapytania	255
8.3.1. Identyfikacja wolno wykonywanych zapytań	255
8.3.2. Analiza wolno wykonywanych zapytań	260
8.3.3. Wzorce zapytania	280
8.4. Podsumowanie	282

Rozdział 9. Wyszukiwanie tekstowe 285

9.1. Wyszukiwanie tekstowe — nie tylko dopasowanie wzorca	286
9.1.1. Wyszukiwanie tekstowe kontra dopasowanie wzorca	288
9.1.2. Wyszukiwanie tekstowe kontra wyszukiwanie stron internetowych	288
9.1.3. Wyszukiwanie tekstowe iv MongoDB kontra dedykowane silniki wyszukiwania tekstowego	291
9.2. Pobranie danych katalogu książek Manning	294
9.3. Zdefiniowanie indeksów wyszukiwania tekstowego	296
9.3.1. Wielkość indeksu wyszukiwania tekstowego	297
9.3.2. Przypisanie indeksowi własnej nazwy oraz zindeksowanie wszystkich pól tekstowych kolekcji	298
9.4. Proste wyszukiwanie tekstowe	299
9.4.1. Bardziej zaawansowane operacje wyszukiwania	300
9.4.2. Ocena wyszukiwania tekstowego	302
9.4.3. Sortowanie wyników na podstawie oceny wyszukiwania tekstowego	304
9.5. Wyszukiwanie tekstowe we frameworku agregacji	304
9.5.1. Gdzie jest MongoDB in Action, Second Edition?	306
9.6. Wyszukiwanie tekstowe w innych językach	308
9.6.1. Wskazanie języka w indeksie	309
9.6.2. Określenie języka iv dokumencie	310

9.6.3. Podanie języka w operacji wyszukiwania	311
9.6.4. Dostępne języki	313
9.7. Podsumowanie	314
Rozdział 10. WiredTiger i dołączany silnik magazynu danych	315
10.1. API Pluggable Storage Engine	315
10.1.1. Dlaczego warto używać różnych silników magazynów danych	316
10.2. Silnik WiredTiger	318
10.2.1. Przejście do silnika WiredTiger	318
10.2.2. Migracja bazy danych do WiredTiger	320
10.3. Porównanie z MMAPv1	321
10.3.1. Pliki konfiguracyjne	322
10.3.2. Skrypty wstawiania danych i przeprowadzania testu wydajności	323
10.3.3. Wyniki testów wydajności wstawiania danych	326
10.3.4. Skrypty sprawdzające wydajność operacji odczytu danych	327
10.3.5. Wyniki testów wydajności odczytu danych	329
10.3.6. Podsumowanie testów wydajności	330
10.4. Inne przykłady dołączanych silników magazynów danych	332
10.5. Tematy zaawansowane	333
10.5.1. Jak działa dołączany silnik magazynu danych?	333
10.5.2. Struktura danych	334
10.5.3. Nakładanie blokad	338
10.6. Podsumowanie	338
Rozdział 11. Replikacja	341
11.1. Ogólne omówienie replikacji	342
11.1.1. Dlaczego replikacja ma znaczenie?	342
11.1.2. Przykłady użycia replikacji i jej ograniczenia	344
11.2. Zbiory replik	345
11.2.1. Konfiguracja	346
11.2.2. Jak działa replikacja?	353
11.2.3. Administracja	362
11.3. Sterowniki i replikacja	372
11.3.1. Połączenia i reakcja na wystąpienie awarii	372
11.3.2. Pewność udanego zapisu	375
11.3.3. Skalowanie operacji odczytu	376
11.3.4. Tagi	379
11.4. Podsumowanie	381
Rozdział 12. Skalowanie systemu za pomocą shardingu	383
12.1. Ogólne omówienie shardingu	384
12.1.1. Co to jest sharding?	384
12.1.2. Kiedy należy stosować sharding?	385
12.2. Poznajemy komponenty klastra shardingu	387
12.2.1. Shardy — pamięć masowa dla danych aplikacji	388

12.2.2. Router mongos — przekierowywanie operacji	388
12.2.3. Serwery konfiguracji — przechowywanie metadanych	389
12.3. Rozproszenie danych w klastrze shardingu	389
12.3.1. Sposoby rozpraszania danych w klastrze shardingu	391
12.3.2. Rozproszenie baz danych między shardami	392
12.3.3. Sharding na podstawie kolekcji	392
12.4. Budowa przykładowego klastra shardingu	394
12.4.1. Uruchomienie serwerów mongod i mongos	394
12.4.2. Konfiguracja klastra	397
12.4.3. Kolekcje shardingu	398
12.4.4. Zapis danych w klastrze shardingu	400
12.5. Wykonywanie zapytań i indeksowanie klastra shardingu	406
12.5.1. Routing zapytania	406
12.5.2. Indeksowanie w klastrze shardingu	408
12.5.3. Narzędzie explain() w klastrze shardingu	408
12.5.4. Agregacja w klastrze shardingu	410
12.6. Wybór klucza shardu	411
12.6.1. Brak równowagi podczas wykonywania operacji zapisu (hotspot)	412
12.6.2. Fragmenty niemożliwe do podziału	413
12.6.3. Kiepskie adresowanie (klucz shardu nie znajduje się w zapytaniach)	414
12.6.4. Idealny klucz shardu	415
12.6.5. Nieodłączne kompromisy podczas projektowania (aplikacja klienta poczty)	415
12.7. Sharding w produkcji	418
12.7.1. Provisioning	418
12.7.2. Wdrożenie	421
12.7.3. Obsługa i konserwacja	423
12.8. Podsumowanie	428
Rozdział 13. Wdrożenie i administracja	429
13.1. Sprzęt i provisioning	430
13.1.1. Topologia klastra	430
13.1.2. Środowisko wdrożenia	432
13.1.3. Provisioning	440
13.2. Monitorowanie i diagnostyka	442
13.2.1. Rejestracja danych	442
13.2.2. Polecenia diagnostyczne MongoDB	443
13.2.3. Narzędzia diagnostyczne MongoDB	443
13.2.4. Usługa monitorowania MongoDB	446
13.2.5. Zewnętrzne aplikacje monitorowania	446
13.3. Kopia zapasowa	447
13.3.1. Narzędzia mongodump i mongorestore	447
13.3.2. Kopia zapasowa na podstawie plików danych	448
13.3.3. Kopia zapasowa tworzona za pomocą monitorowania MMS	450

13.4. Zapewnienie bezpieczeństwa	450
13.4.1. Bezpieczne środowisko	450
13.4.2. Szyfrowanie komunikacji sieciowej	451
13.4.3. Uwierzytelnianie	454
13.4.4. Uwierzytelnienie zbioru replik	457
13.4.5. Uwierzytelnianie klastra shardingu	459
13.4.6. Funkcje zabezpieczeń w korporacyjnej wersji MongoDB	459
13.5. Zadania administracyjne	459
13.5.1. Import i eksport danych	460
13.5.2. Naprawa i zmniejszenie ilości miejsca zajmowanego przez pliki danych	461
13.5.3. Uaktualnienie	462
13.6. Rozwiązywanie problemów związanych z wydajnością	463
13.6.1. Zbiór roboczy	463
13.6.2. Nagły spadek wydajności	464
13.6.3. Interakcje zapytań	465
13.6.4. Szukanie profesjonalnej pomocy	466
13.7. Lista rzeczy do sprawdzenia podczas wdrożenia	466
13.8. Podsumowanie	468

DODATKI **469**

Dodatek A. Instalacja **471**

A.1. Instalacja	471
A.1.1. Wdrożenie w środowisku produkcyjnym	471
A.1.2. Architektura 32-bitowa kontra 64-bitowa	472
A.2. MongoDB w systemie Linux	472
A.2.1. Instalacja z użyciem prekompilowanych plików binarnych	472
A.2.2. Użycie menedżera pakietów	473
A.3. MongoDB w systemie OS X	474
A.3.1. Instalacja z użyciem prekompilowanych plików binarnych	474
A.3.2. Użycie menedżera pakietów	475
A.4. MongoDB w Windows	475
A.4.1. Instalacja z użyciem prekompilowanych plików binarnych	476
A.5. Kompilacja MongoDB z kodu źródłowego	477
A.6. Rozwiązywanie problemów	477
A.6.1. Nieprawidłowa architektura	477
A.6.2. Brak katalogu danych	477
A.6.3. Brak uprawnień	478
A.6.4. Brak możliwości dołączenia do portu	478
A.7. Podstawowe opcje konfiguracyjne	478
A.8. Instalacja języka Ruby	480
A.8.1. Systemy Linux i OS X	480
A.8.2. Windows	481

Dodatek B. Wzorce projektowe	483
B.1. Osadzenie kontra odwołanie	483
B.2. Związek typu „jeden do wielu”	483
B.3. Związek typu „wiele do wielu”	485
B.4. Drzewo	486
B.5. Kolejki procesów roboczych	489
B.6. Atrybuty dynamiczne	490
B.7. Transakcje	491
B.8. Lokalizacja i obliczenia wstępne	492
B.9. Antywzorce	493
B.9.1. Niepoprawne indeksowanie	493
B.9.2. Bałagan w typach	494
B.9.3. Kolekcje kubeków	494
B.9.4. Ogromne, głęboko zagnieżdżone dokumenty	494
B.9.5. Jedna kolekcja dla użytkownika	494
B.9.6. Kolekcje niemożliwe do shardingu	495
 Dodatek C. Dane binarne i GridFS	 497
C.1. Przechowywanie prostych obiektów binarnych	497
C.1.1. Przechowywanie miniatury	498
C.1.2. Przechowywanie wartości MD5	498
C.2. GridFS	499
C.2.1. GridFS w języku Ruby	450
C.2.2. GridFS i mongofiles	503
 Skorowidz	 505

oprac. BPK