

Spis treści

Przedmowa	15
Część I. Wprowadzenie do MongoDB	19
1. Wprowadzenie	21
Łatwość użycia	21
Skalowanie	21
Bogactwo funkcji	23
...bez ograniczania prędkości	23
Filozofia	24
2. Zaczynamy	25
Dokumenty	25
Kolekcje	26
Dynamiczne schematy	26
Nazewnictwo	27
Bazy danych	28
Pobieranie i uruchamianie MongoDB	29
Wprowadzenie do powłoki MongoDB	30
Uruchamianie powłoki	31
Klient MongoDB	31
Podstawowe operacje w powłoce	32
Typy danych	34
Proste typy danych	34
Daty	36
Tablice	37
Zagnieżdżone dokumenty	37
Klucz _id i typ ObjectId	38
Używanie powłoki MongoDB	40
Porady przy używaniu powłoki	40
Uruchamianie skryptów w powłoce	41
Tworzenie pliku .mongorc.js	43
Dostosowywanie znaku zachęty	44
Edytowanie złożonych zmiennych	45
Niewygodne nazwy kolekcji	45
3. Tworzenie, aktualizowanie i usuwanie dokumentów	47
Wstawianie dokumentów	47
Metoda insertMany	47

Sprawdzanie poprawności wstawiania	50
Metoda insert	50
Usuwanie dokumentów	51
Metoda drop	52
Aktualizowanie dokumentów	52
Zastępowanie dokumentów	53
Używanie operatorów aktualizacji	54
Operacje typu upsert	64
Aktualizowanie wielu dokumentów	66
Zwracanie zaktualizowanych dokumentów	67

4. Zapytania **71**

Metoda find	71
Wybieranie zwracanych kluczy	72
Ograniczenia	72
Kryteria zapytania	73
Zapytania warunkowe	73
Zapytania LUB	74
Operator \$not	75
Zapytania związane z różnymi typami	75
Typ null	75
Wyrażenia regularne	76
Zapytania dotyczące tablic	77
Zapytania do dokumentów zagnieżdżonych	81
Zapytania \$where	83
Kursory	84
Limity, pominięcia i sortowanie	85
Unikanie pomijania dużej liczby dokumentów	87
Nieśmiertelne kursory	89

Część II. Projektowanie aplikacji **91**

5. Indeksy **93**

Wprowadzenie do indeksów	93
Tworzenie indeksu	95
Wprowadzenie do indeksów złożonych	98
W jaki sposób MongoDB wybiera indeks?	102
Używanie indeksów złożonych	103
W jaki sposób operatory \$ korzystają z indeksów?	120
Indeksowanie obiektów i tablic	130
Liczność indeksu	132
Dane z polecenia explain	132
Kiedy nie używać indeksu?	140
Rodzaje indeksów	141
Indeksy unikalne	141
Indeksy częściowe	143

Administrowanie indeksem	144
Identyfikowanie indeksów	145
Zmienianie indeksów	146
6. Specjalne typy indeksów i kolekcji	147
Indeksy geoprzestrzenne	147
Rodzaje zapytań geoprzestrzennych	148
Używanie indeksów geoprzestrzennych	149
Złożone indeksy geoprzestrzenne	157
Indeksy 2d	157
Indeksy wyszukiwania pełnotekstowego	159
Tworzenie indeksu tekstowego	160
Wyszukiwanie tekstowe	161
Optymalizowanie wyszukiwania pełnotekstowego	163
Wyszukiwanie w innych językach	164
Kolekcje ograniczone	164
Tworzenie kolekcji ograniczonych	166
Wieczne kursory	167
Indeksy o ograniczonym czasie życia	167
Przechowywanie plików za pomocą GridFS	168
Zaczynamy pracę z GridFS — mongofiles	168
Praca z GridFS w sterownikach MongoDB	169
Pod maską	170
7. Wprowadzenie do frameworka agregacji	173
Potoki, etapy i regulatory	173
Praca z etapami — typowe operacje	175
Wyrażenia	179
Etap typu \$project	180
Etap typu \$unwind	185
Wyrażenia tablicowe	192
Akumulatory	196
Używanie akumulatorów w etapie projekcji	197
Wprowadzenie do grupowania	198
Pole _id w etapie grupowania	202
Etap grupowania i etap projekcji	205
Zapisywanie wyników potoku agregacji do kolekcji	208
8. Transakcje	209
Wprowadzenie do transakcji	209
Definicja standardu ACID	209
Jak używać transakcji?	210
Dostosowywanie limitów transakcji w swojej aplikacji	214
Ograniczenia czasowe i limity wielkości protokołu	215
9. Projektowanie aplikacji	217

Projektowanie schematu danych	217
Wzorce projektowe schematów	218
Normalizacja i denormalizacja	221
Przykłady reprezentacji danych	222
Kardynalność	226
Znajomi, obserwujący i inne nieprzyjemności	226
Optymalizowanie manipulacji na danych	229
Usuwanie starych danych	229
Planowanie baz danych i kolekcji	230
Spójność danych	231
Migrowanie schematów	232
Zarządzanie schematami	233
Kiedy nie używać MongoDB?	233
Część III. Replikacja	235
10. Konfigurowanie zbioru replik	237
Wprowadzenie do replikacji	237
Konfigurowanie zbioru replik, część 1.	238
Przemyślenia na temat sieci	239
Przemyślenia na temat bezpieczeństwa	239
Konfigurowanie zbioru replik, część 2.	240
Kontrolowanie replikacji	243
Modyfikowanie konfiguracji zbioru replik	248
Jak zaprojektować zbiór?	250
Jak działa wybieranie serwera podstawowego?	252
Opcje konfiguracji elementów zbioru replik	253
Priorytet	254
Ukrywanie serwerów	254
Arbiter wyborów	255
Tworzenie indeksów	257
11. Komponenty zbioru replik	259
Synchronizacja	259
Synchronizacja początkowa	261
Replikacja	263
Jak radzić sobie z przestarzałymi danymi?	263
Żądania heartbeat	263
Stany elementów zbioru	264
Wybory	265
Cofanie zmian	266
Gdy cofanie operacji się nie powiedzie	269
12. Łączenie aplikacji ze zbiorem replik	271
Zachowania związane z łączeniem klienta ze zbiorem replik	271
Oczekiwanie na replikację operacji zapisu	273

Pozostałe wartości klucza "w"	275
Definiowanie gwarancji dla replikacji	275
Gwarantowany jeden serwer na centrum danych	275
Gwarancja większości nieukrytych serwerów	277
Tworzenie innych gwarancji	277
Wysyłanie żądań odczytu do serwerów wtórnych	278
Problemy ze spójnością danych	278
Problemy z obciążeniem serwerów	279
Kiedy warto korzystać z serwerów wtórnych do odczytywania danych	280
13. Administracja	283
Uruchamianie serwerów w trybie samodzielnym	283
Konfiguracja zbioru replik	284
Tworzenie zbioru replik	284
Zmienianie serwerów w zbiorze	285
Tworzenie większych zbiorów	285
Wymuszanie rekonfiguracji	286
Manipulowanie stanem serwera	286
Przekształcanie serwera podstawowego we wtórny	287
Zapobieganie wyborom	287
Monitorowanie replikacji	287
Odczytywanie statusu	288
Wizualizacja grafu replikacji	291
Pętle replikacji	292
Wyłączanie chainingu	292
Wyliczanie opóźnień	293
Zmiana wielkości protokołu operacji	294
Tworzenie indeksów	295
Replikacja budżetowa	296
Część IV. Sharding	299
14. Wprowadzenie do shardingu	301
Czym jest sharding?	301
Komponenty klastra	302
Sharding w klastrze jednoserwerowym	303
15. Konfigurowanie shardingu	313
Kiedy stosować sharding?	313
Uruchamianie serwerów	314
Serwery konfiguracji	314
Procesy mongos	315
Tworzenie sharda ze zbioru replik	316
Zwiększanie pojemności	319
Dzielenie danych	320
W jaki sposób MongoDB kontroluje dane klastra?	320

Zakresy kawałków	321
Dzielenie kawałków	323
Równoważenie obciążeń	325
Zestawienia	326
Strumienie zmian	326
16. Wybieranie klucza shardingu	329
Mierzenie sposobu używania kolekcji	329
Rozrysowywanie rozdziału danych	330
Rosnące klucze shardingu	330
Klucze shardingu o losowym rozkładzie	333
Klucze shardingu zależne od lokalizacji	335
Strategie kluczy shardingu	336
Haszowany klucz shardingu	336
Haszowane klucze shardingu dla systemu GridFS	338
Strategia węża strażackiego	338
Hotspoty	339
Reguły i wskazówki dotyczące kluczy shardingu	341
Ograniczenia kluczy shardingu	341
Kardynalność kluczy shardingu	343
Kontrolowanie rozdziału danych	343
Używanie klastra z wieloma bazami danych i kolekcjami	343
Sharding manualny	345
17. Administrowanie shardingiem	347
Sprawdzanie aktualnego stanu	347
Przeglądanie podsumowania za pomocą metody sh.status()	347
Przeglądanie informacji o konfiguracji	349
Kontrolowanie połączeń sieciowych	355
Pobieranie statystyk dotyczących połączeń	355
Ograniczanie liczby połączeń	362
Administrowanie serwerem	363
Dodawanie serwerów	363
Zmienianie serwerów w shardzie	363
Usuwanie sharda	364
Równoważenie danych	367
Proces równoważący	367
Zmiana wielkości kawałka	368
Przenoszenie kawałków	369
Wielkie kawałki	371
Odświeżanie konfiguracji	374
Część V. Administrowanie aplikacją	375
18. Kontrolowanie działania aplikacji	377
Przeglądanie aktualnych operacji	377

Wyszukiwanie problematycznych operacji	380
Zatrzymywanie operacji	380
Fałszywe alarmy	381
Zapobieganie powstawaniu fantomowych operacji	381
Używanie profilera systemowego	382
Wyliczanie wielkości	385
Dokumenty	385
Kolekcje	386
Bazy danych	390
Używanie narzędzi mongotop i mongostat	391

19. Wprowadzenie do bezpieczeństwa MongoDB **395**

Uwierzytelnianie i autoryzacja	395
Mechanizmy uwierzytelniania	395
Autoryzacja	396
Stosowanie certyfikatów X.509 do uwierzytelniania serwerów i klientów	398
Samouczek uwierzytelniania i szyfrowania komunikacji w MongoDB	400
Tworzenie centrum certyfikacji	400
Generowanie i podpisywanie certyfikatów serwerów	405
Generowanie i podpisywanie certyfikatów klientów	406
Uruchamianie zbioru replik bez włączonego uwierzytelniania i autoryzacji	406
Tworzenie użytkownika administracyjnego	407
Ponowne uruchomienie zbioru replik z włączonym uwierzytelnianiem i autoryzacją	408

20. Trwałość danych **411**

Trwałość danych na poziomie serwera dzięki mechanizmowi księgowania	411
Trwałość danych na poziomie klastra dzięki opcjom „write concern”	413
Opcje "w" i "wtimeout"	413
Używanie opcji "j" w dokumencie writeConcern	414
Trwałość danych na poziomie klastra dzięki opcjom „read concern”	415
Trwałość danych w transakcjach dzięki opcjom „write concern”	415
Czego MongoDB nie gwarantuje?	416
Poszukiwanie uszkodzeń danych	417

Część VI. Administrowanie serwerem **419**

21. Konfigurowanie MongoDB w środowisku produkcyjnym **421**

Uruchamianie z wiersza poleceń	421
Konfiguracja zapisana w pliku	425
Zatrzymywanie serwera MongoDB	426
Bezpieczeństwo	426
Szyfrowanie danych	427
Połączenia SSL	428
Protokołowanie	428

22. Monitorowanie MongoDB	431
Monitorowanie wykorzystania pamięci	431
Wprowadzenie do pamięci komputerów	431
Kontrolowanie wykorzystania pamięci	432
Kontrolowanie liczby błędów stron	433
Parametr I/O wait	435
Wyliczanie wielkości zbioru roboczego	435
Kilka przykładów zbiorów roboczych	437
Kontrolowanie wydajności	437
Kontrolowanie wolnej przestrzeni na dysku	439
Monitorowanie procesu replikacji	439
 23. Tworzenie kopii zapasowych	 443
Metody tworzenia kopii zapasowych	443
Tworzenie kopii zapasowej serwera	444
Migawka systemu plików	444
Kopiowanie plików z danymi	447
Używanie narzędzia mongodump	449
Szczególny przypadek kopii zapasowej zbioru replik	451
Szczególny przypadek kopii zapasowej klastra shardów	452
Tworzenie kopii zapasowej całego klastra i jej odtwarzanie	453
Tworzenie kopii zapasowej pojedynczego sharda i jej odtwarzanie	453
 24. Wdrożenia MongoDB	 455
Projektowanie systemu	455
Wybieranie nośników danych	455
Zalecane konfiguracje macierzy RAID	456
Procesor	457
System operacyjny	457
Przestrzeń wymiany	458
System plików	458
Wirtualizacja	459
Nadmierne używanie pamięci	459
Tajemnicza pamięć	459
Problemy z dyskami sieciowymi	459
Używanie dysków niesieciowych	461
Konfigurowanie ustawień systemowych	461
Wyłączanie opcji NUMA	461
Opcja readahead	463
Wyłączanie opcji THP	464
Wybieranie algorytmu planisty dysku	465
Wyłączanie śledzenia czasu dostępu	465
Modyfikowanie limitów	466
Konfigurowanie sieci	467
Porządkowanie systemu	469
Synchronizowanie zegarów	469

Proces OOM killer	469
Wyłączanie zadań okresowych	470
A. Instalowanie MongoDB	471
Wybieranie wersji	471
Instalowanie w systemie Windows	472
Instalowanie jako usługa	473
Instalowanie w systemach POSIX (Linux i Mac OS X)	473
Instalowanie za pomocą menedżera pakietów	474
B. Wewnętrzne elementy MongoDB	475
BSON	475
Protokół komunikacji	476
Pliki danych	476
Przestrzenie nazw	478
Mechanizm zapisywania danych WiredTiger	478

oprac. BPK