

contents

preface xvii
acknowledgments xix
about this book xxi
about the cover illustration xxiv

PART 1 GETTING STARTED.....1

1 *A database for the modern web* 3

- 1.1 Built for the internet 5
- 1.2 MongoDB's key features 6
 - Document data model* 6 ▪ *Ad hoc queries* 10
 - Indexes* 10 ▪ *Replication* 11 ▪ *Speed and durability* 12
 - Scaling* 14
- 1.3 MongoDB's core server and tools 15
 - Core server* 16 ▪ *JavaScript shell* 16 ▪ *Database drivers* 17
 - Command-line tools* 18
- 1.4 Why MongoDB? 18
 - MongoDB versus other databases* 19 ▪ *Use cases and production deployments* 22
- 1.5 Tips and limitations 24
- 1.6 History of MongoDB 25

1.7 Additional resources 27

1.8 Summary 28

2 *MongoDB through the JavaScript shell* 29

2.1 Diving into the MongoDB shell 30

Starting the shell 30 ▪ *Databases, collections, and documents* 31

Inserts and queries 32 ▪ *Updating documents* 34

Deleting data 38 ▪ *Other shell features* 38

2.2 Creating and querying with indexes 39

Creating a large collection 39 ▪ *Indexing and explain()* 41

2.3 Basic administration 46

Getting database information 46 ▪ *How commands work* 48

2.4 Getting help 49

2.5 Summary 51

3 *Writing programs using MongoDB* 52

3.1 MongoDB through the Ruby lens 53

Installing and connecting 53 ▪ *Inserting documents in Ruby* 55

Queries and cursors 56 ▪ *Updates and deletes* 57

Database commands 58

3.2 How the drivers work 59

Object ID generation 59

3.3 Building a simple application 61

Setting up 61 ▪ *Gathering data* 62 ▪ *Viewing the archive* 65

3.4 Summary 69

PART 2 APPLICATION DEVELOPMENT IN MONGODB71

4 *Document-oriented data* 73

4.1 Principles of schema design 74

4.2 Designing an e-commerce data model 75

Schema basics 76 ▪ *Users and orders* 80 ▪ *Reviews* 83

4.3 Nuts and bolts: On databases, collections,
and documents 84

Databases 84 ▪ *Collections* 87 ▪ *Documents and insertion* 92

4.4 Summary 96

- ## 5 *Constructing queries* 98
- 5.1 E-commerce queries 99
 - Products, categories, and reviews* 99 ▪ *Users and orders* 101
 - 5.2 MongoDB's query language 103
 - Query criteria and selectors* 103 ▪ *Query options* 117
 - 5.3 Summary 119
- ## 6 *Aggregation* 120
- 6.1 Aggregation framework overview 121
 - 6.2 E-commerce aggregation example 123
 - Products, categories, and reviews* 125
 - User and order* 132
 - 6.3 Aggregation pipeline operators 135
 - \$project* 136 ▪ *\$group* 136 ▪ *\$match*, *\$sort*, *\$skip*, *\$limit* 138 ▪ *\$unwind* 139 ▪ *\$out* 139
 - 6.4 Reshaping documents 140
 - String functions* 141 ▪ *Arithmetic functions* 142
 - Date functions* 142 ▪ *Logical functions* 143
 - Set Operators* 144 ▪ *Miscellaneous functions* 145
 - 6.5 Understanding aggregation pipeline performance 146
 - Aggregation pipeline options* 147 ▪ *The aggregation framework's explain() function* 147 ▪ *allowDiskUse option* 151
 - Aggregation cursor option* 151
 - 6.6 Other aggregation capabilities 152
 - .count() and .distinct()* 153 ▪ *map-reduce* 153
 - 6.7 Summary 156
- ## 7 *Updates, atomic operations, and deletes* 157
- 7.1 A brief tour of document updates 158
 - Modify by replacement* 159 ▪ *Modify by operator* 159
 - Both methods compared* 160 ▪ *Deciding: replacement vs. operators* 160
 - 7.2 E-commerce updates 162
 - Products and categories* 162 ▪ *Reviews* 167 ▪ *Orders* 168
 - 7.3 Atomic document processing 171
 - Order state transitions* 172 ▪ *Inventory management* 174

- 7.4 Nuts and bolts: MongoDB updates and deletes 179
 - Update types and options* 179 ▪ *Update operators* 181
 - The findAndModify command* 188 ▪ *Deletes* 189
 - Concurrency, atomicity, and isolation* 190
 - Update performance notes* 191
- 7.5 Reviewing update operators 192
- 7.6 Summary 193

PART 3 MONGODB MASTERY 195

8 *Indexing and query optimization* 197

- 8.1 Indexing theory 198
 - A thought experiment* 198 ▪ *Core indexing concepts* 201
 - B-trees* 205
- 8.2 Indexing in practice 207
 - Index types* 207 ▪ *Index administration* 211
- 8.3 Query optimization 216
 - Identifying slow queries* 217 ▪ *Examining slow queries* 221
 - Query patterns* 241
- 8.4 Summary 243

9 *Text search* 244

- 9.1 Text searches—not just pattern matching 245
 - Text searches vs. pattern matching* 246 ▪ *Text searches vs. web page searches* 247 ▪ *MongoDB text search vs. dedicated text search engines* 250
- 9.2 Manning book catalog data download 253
- 9.3 Defining text search indexes 255
 - Text index size* 255 ▪ *Assigning an index name and indexing all text fields in a collection* 256
- 9.4 Basic text search 257
 - More complex searches* 259 ▪ *Text search scores* 261
 - Sorting results by text search score* 262
- 9.5 Aggregation framework text search 263
 - Where's MongoDB in Action, Second Edition?* 265

- 9.6 Text search languages 267
 - Specifying language in the index* 267
 - *Specifying the language in the document* 269
 - *Specifying the language in a search* 269
 - Available languages* 271
- 9.7 Summary 272

10 **WiredTiger and pluggable storage** 273

- 10.1 Pluggable Storage Engine API 273
 - Why use different storages engines?* 274
- 10.2 WiredTiger 275
 - Switching to WiredTiger* 276
 - *Migrating your database to WiredTiger* 277
- 10.3 Comparison with MMAPv1 278
 - Configuration files* 279
 - *Insertion script and benchmark script* 281
 - *Insertion benchmark results* 283
 - Read performance scripts* 285
 - *Read performance results* 286
 - Benchmark conclusion* 288
- 10.4 Other examples of pluggable storage engines 289
- 10.5 Advanced topics 290
 - How does a pluggable storage engine work?* 290
 - Data structure* 292
 - *Locking* 294
- 10.6 Summary 295

11 **Replication** 296

- 11.1 Replication overview 297
 - Why replication matters* 297
 - *Replication use cases and limitations* 298
- 11.2 Replica sets 300
 - Setup* 300
 - *How replication works* 307
 - Administration* 314
- 11.3 Drivers and replication 324
 - Connections and failover* 324
 - *Write concern* 327
 - Read scaling* 328
 - *Tagging* 330
- 11.4 Summary 332

- ## 12 *Scaling your system with sharding* 333
- 12.1 Sharding overview 334
 - What is sharding?* 334
 - When should you shard?* 335
 - 12.2 Understanding components of a sharded cluster 336
 - Shards: storage of application data* 337
 - Mongos router: router of operations* 338
 - Config servers: storage of metadata* 338
 - 12.3 Distributing data in a sharded cluster 339
 - Ways data can be distributed in a sharded cluster* 340
 - Distributing databases to shards* 341
 - Sharding within collections* 341
 - 12.4 Building a sample shard cluster 343
 - Starting the mongod and mongos servers* 343
 - Configuring the cluster* 346
 - Sharding collections* 347
 - Writing to a sharded cluster* 349
 - 12.5 Querying and indexing a shard cluster 355
 - Query routing* 355
 - Indexing in a sharded cluster* 356
 - The explain() tool in a sharded cluster* 357
 - Aggregation in a sharded cluster* 359
 - 12.6 Choosing a shard key 359
 - Imbalanced writes (hotspots)* 360
 - Unsplittable chunks (coarse granularity)* 362
 - Poor targeting (shard key not present in queries)* 362
 - Ideal shard keys* 363
 - Inherent design trade-offs (email application)* 364
 - 12.7 Sharding in production 365
 - Provisioning* 366
 - Deployment* 369
 - Maintenance* 370
 - 12.8 Summary 375
- ## 13 *Deployment and administration* 376
- 13.1 Hardware and provisioning 377
 - Cluster topology* 377
 - Deployment environment* 378
 - Provisioning* 385
 - 13.2 Monitoring and diagnostics 386
 - Logging* 387
 - MongoDB diagnostic commands* 387
 - MongoDB diagnostic tools* 388
 - MongoDB Monitoring Service* 390
 - External monitoring applications* 390
 - 13.3 Backups 391
 - mongodump and mongorestore* 391
 - Data file-based backups* 392
 - MMS backups* 393

13.4	Security	394
	<i>Secure environments</i>	394
	▪ <i>Network encryption</i>	395
	<i>Authentication</i>	397
	▪ <i>Replica set authentication</i>	401
	<i>Sharding authentication</i>	402
	▪ <i>Enterprise security features</i>	402
13.5	Administrative tasks	402
	<i>Data imports and exports</i>	402
	▪ <i>Compaction and repair</i>	403
	<i>Upgrading</i>	405
13.6	Performance troubleshooting	405
	<i>Working set</i>	406
	▪ <i>Performance cliff</i>	407
	<i>Query interactions</i>	407
	▪ <i>Seek professional assistance</i>	408
13.7	Deployment checklist	408
13.8	Summary	410
<i>appendix A</i>	<i>Installation</i>	<i>411</i>
<i>appendix B</i>	<i>Design patterns</i>	<i>421</i>
<i>appendix C</i>	<i>Binary data and GridFS</i>	<i>433</i>
	<i>index</i>	<i>441</i>