

《MongoDB权威指南（第2版）》

图书基本信息

书名：《MongoDB权威指南（第2版）》

13位ISBN编号：9787115341082

10位ISBN编号：7115341087

出版时间：2014-1-1

出版社：人民邮电出版社

作者：霍多罗夫 (Kristina Chodorow)

页数：390

译者：邓强,王明辉

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu000.com

内容概要

MongoDB如何帮你管理通过Web应用收集的海量数据呢？通过这本经典著作全新升级版的权威解读，你会了解面向文档数据库的诸多优点，会发现MongoDB如此稳定、性能优越甚至能够无限水平扩展背后的原因。

本书是一本广受好评的MongoDB权威著作。新版本对上一版进行了全面扩充，可作为数据库开发人员的工作指南，也可作为系统管理人员的进阶指导，还可供项目中其他成员了解MongoDB使用。书中介绍了面向文档的存储方式及利用MongoDB的无模式数据模型处理文档、集合和多个数据库，讲述了如何执行基本的写操作以及各种复杂的条件查询，还介绍了索引、聚合工具以及其他高级查询技术，另外对监控、安全性和身份验证、备份和修复、水平扩展MongoDB数据库等内容也有所涉及。

本书主要内容：

MongoDB核心概念和术语

在不同的安全等级和速度下执行基本的写操作

利用限制、跳过等选项执行复杂的查询，并对查询结果进行排序

基于MongoDB设计应用程序

聚合数据，包括统计不同值的数量，找出不同值，给文档分组，以及使用MapReduce

收集并解释有关集合和数据库的统计信息

在MongoDB中设置副本集和自动故障转移

使用分片横向扩展数据库，并学习这样做对应用的影响

深入介绍监控、安全和身份验证、备份和恢复以及其他管理任务

《MongoDB权威指南（第2版）》

作者简介

作者简介：

Kristina Chodorow

谷歌软件工程师，曾有5年是MongoDB项目的核心成员。她领导了MongoDB的副本集开发，并编写了PHP和Perl驱动程序。

译者简介：

邓强

软件工程师，常年关注互联网行业，对互联网产品和技术兴趣浓厚，Apple粉、Google粉、Amazon粉。职业生涯的前几年一直从事金融行业软件系统开发，后来不顾一切投身互联网。目前在一个新成立的互联网创业小团队任职。Email：dengqiang@outlook.com。

王明辉

同济大学计算机系学生，经常使用MongoDB。知乎网常用id为纳米黑客。

书籍目录

序	XV
前言	XVII
第一部分 MongoDB 介绍	
第1章 MongoDB 简介	3
1.1 易于使用	3
1.2 易于扩展	4
1.3 丰富的功能	4
1.4 卓越的性能	5
1.5 小结	5
第2章 MongoDB 基础知识	7
2.1 文档	7
2.2 集合	8
2.2.1 动态模式	8
2.2.2 命名	9
2.3 数据库	10
2.4 启动MongoDB	11
2.5 MongoDB shell 简介	12
2.5.1 运行shell	12
2.5.2 MongoDB 客户端	13
2.5.3 shell 中的基本操作	14
2.6 数据类型	16
2.6.1 基本数据类型	16
2.6.2 日期	18
2.6.3 数组	18
2.6.4 内嵌文档	19
2.6.5 _id 和ObjectId	20
2.7 使用MongoDB shell	21
2.7.1 shell 小贴士	22
2.7.2 使用shell 执行脚本	23
2.7.3 创建.mongorc.js 文件	25
2.7.4 定制shell 提示	26
2.7.5 编辑复合变量	26
2.7.6 集合命名注意事项	27
第3章 创建、更新和删除文档	29
3.1 插入并保存文档	29
3.1.1 批量插入	29
3.1.2 插入校验	30
3.2 删除文档	31
3.3 更新文档	32
3.3.1 文档替换	32
3.3.2 使用修改器	34
3.3.3 upsert	45
3.3.4 更新多个文档	47
3.3.5 返回被更新的文档	48
3.4 写入安全机制	50
第4章 查询	53
4.1 find 简介	53

4.1.1	指定需要返回的键	54
4.1.2	限制	55
4.2	查询条件	55
4.2.1	查询条件	55
4.2.2	OR 查询	56
4.2.3	\$not	57
4.2.4	条件语义	57
4.3	特定类型的查询	58
4.3.1	null	58
4.3.2	正则表达式	59
4.3.3	查询数组	59
4.3.4	查询内嵌文档	64
4.4	\$where 查询	65
4.5	游标	67
4.5.1	limit、skip 和sort	69
4.5.2	避免使用skip 略过大量结果	70
4.5.3	高级查询选项	72
4.5.4	获取一致结果	73
4.5.5	游标生命周期	75
4.6	数据库命令	75
第二部分 设计应用		
第5章 索引 81		
5.1	索引简介	81
5.1.1	复合索引简介	84
5.1.2	使用复合索引	90
5.1.3	\$ 操作符如何使用索引	91
5.1.4	索引对象和数组	96
5.1.5	索引基数	98
5.2	使用explain() 和hint()	98
5.3	何时不应该使用索引	103
5.4	索引类型	104
5.4.1	唯一索引	104
5.4.2	稀疏索引	106
5.5	索引管理	107
5.5.1	标识索引	108
5.5.2	修改索引	108
第6章 特殊的索引和集合 111		
6.1	固定集合	111
6.1.1	创建固定集合	113
6.1.2	自然排序	113
6.1.3	循环游标	115
6.1.4	没有_id 索引的集合	115
6.2	TTL 索引	116
6.3	全文本索引	116
6.3.1	搜索语法	119
6.3.2	优化全文本搜索	120
6.3.3	在其他语言中搜索	121
6.4	地理空间索引	121
6.4.1	地理空间查询的类型	122

6.4.2	复合地理空间索引	123
6.4.3	2d 索引	123
6.5	使用GridFS 存储文件	125
6.5.1	GridFS 入门	126
6.5.2	在MongoDB 驱动程序中使用GridFS	126
6.5.3	揭开GridFS 的面纱	127
第7章	聚合	129
7.1	聚合框架	129
7.2	管道操作符	131
7.2.1	\$match	132
7.2.2	\$project	132
7.2.3	\$group	137
7.2.4	\$unwind	140
7.2.5	\$sort	141
7.2.6	\$limit	142
7.2.7	\$skip	142
7.2.8	使用管道	142
7.3	MapReduce	143
7.3.1	示例1：找出集合中的所有键	143
7.3.2	示例2：网页分类	145
7.3.3	MongoDB 和MapReduce	146
7.4	聚合命令	148
7.4.1	count	149
7.4.2	distinct	149
7.4.3	group	150
第8章	应用程序设计	155
8.1	范式化与反范式化	155
8.1.1	数据表示的例子	156
8.1.2	基数	159
8.1.3	好友、粉丝，以及其他的麻烦事项	160
8.2	优化数据操作	162
8.2.1	优化文档增长	162
8.2.2	删除旧数据	164
8.3	数据库和集合的设计	164
8.4	一致性管理	165
8.5	模式迁移	166
8.6	不适合使用MongoDB 的场景	167
第三部分	复制	
第9章	创建副本集	171
9.1	复制简介	171
9.2	建立副本集	172
9.3	配置副本集	176
9.3.1	rs 辅助函数	178
9.3.2	网络注意事项	178
9.4	修改副本集配置	178
9.5	设计副本集	180
9.6	成员配置选项	184
9.6.1	选举仲裁者	184
9.6.2	优先级	185

9.6.3	隐藏成员	186
9.6.4	延迟备份节点	187
9.6.5	创建索引	187
第10章	副本集的组成	189
10.1	同步	189
10.1.1	初始化同步	190
10.1.2	处理陈旧数据	193
10.2	心跳	193
10.3	选举	195
10.4	回滚	195
第11章	从应用程序连接副本集	201
11.1	客户端到副本集的连接	201
11.2	等待写入复制	202
11.2.1	可能导致错误的原因	203
11.2.2	"w"的其他值	204
11.3	自定义复制保证规则	204
11.3.1	保证复制到每个数据中心的一台服务器上	204
11.3.2	保证写操作被复制到可见节点中的“大多数”	206
11.3.3	创建其他规则	206
11.4	将读请求发送到备份节点	207
11.4.1	出于一致性考虑	207
11.4.2	出于负载的考虑	208
11.4.3	何时可以从备份节点读取数据	208
第12章	管理	211
12.1	以单机模式启动成员	211
12.2	副本集配置	212
12.2.1	创建副本集	212
12.2.2	修改副本集成员	213
12.2.3	创建比较大的副本集	213
12.2.4	强制重新配置	214
12.3	修改成员状态	215
12.3.1	把主节点变为备份节点	215
12.3.2	阻止选举	215
12.3.3	使用维护模式	215
12.4	监控复制	216
12.4.1	获取状态	216
12.4.2	复制图谱	218
12.4.3	复制循环	220
12.4.4	禁用复制链	220
12.4.5	计算延迟	221
12.4.6	调整oplog大小	222
12.4.7	从延迟备份节点中恢复	223
12.4.8	创建索引	224
12.4.9	在预算有限的情况下进行复制	225
12.4.10	主节点如何跟踪延迟	226
12.5	主从模式	227
12.5.1	从主从模式切换到副本集模式	228
12.5.2	让副本集模仿主从模式的行为	228
第四部分	分片	

第13章	分片	233	
13.1	分片简介	233	
13.2	理解集群的组件	234	
13.3	快速建立一个简单的集群	235	
第14章	配置分片	243	
14.1	何时分片	243	
14.2	启动服务器	244	
14.2.1	配置服务器	244	
14.2.2	mongos 进程	245	
14.2.3	将副本集转换为分片	245	
14.2.4	增加集群容量	247	
14.2.5	数据分片	247	
14.3	MongoDB 如何追踪集群数据	248	
14.3.1	块范围	249	
14.3.2	拆分块	250	
14.4	均衡器	254	
第15章	选择片键	257	
15.1	检查使用情况	257	
15.2	数据分发	258	
15.2.1	升序片键	258	
15.2.2	随机分发的片键	261	
15.2.3	基于位置的片键	262	
15.3	片键策略	263	
15.3.1	散列片键	264	
15.3.2	GridFS 的散列片键	265	
15.3.3	流水策略	266	
15.3.4	多热点	267	
15.4	片键规则和指导方针	270	
15.4.1	片键限制	270	
15.4.2	片键的势	270	
15.5	控制数据分发	270	
15.5.1	对多个数据库和集合使用一个集群	270	
15.5.2	手动分片	272	
第16章	分片管理	275	
16.1	检查集群状态	275	
16.1.1	使用sh.status 查看集群摘要信息	275	
16.1.2	检查配置信息	277	
16.2	查看网络连接	282	
16.2.1	查看连接统计	283	
16.2.2	限制连接数量	283	
16.3	服务器管理	285	
16.3.1	添加服务器	285	
16.3.2	修改分片的服务器	285	
16.3.3	删除分片	286	
16.3.4	修改配置服务器	288	
16.4	数据均衡	289	
16.4.1	均衡器	289	
16.4.2	修改块大小	290	
16.4.3	移动块	290	

16.4.4	特大块	292
16.4.5	刷新配置	295
第五部分 应用管理		
第17章 了解应用的动态 299		
17.1	了解正在进行的操作	299
17.1.1	寻找有问题的操作	301
17.1.2	终止操作的执行	301
17.1.3	假象	302
17.1.4	避免幽灵操作	302
17.2	使用系统分析器	303
17.3	计算空间消耗	305
17.3.1	文档	305
17.3.2	集合	305
17.3.3	数据库	306
17.4	使用mongotop和monogostat	307
第18章 数据管理 311		
18.1	配置身份验证	311
18.1.1	身份验证基本原理	312
18.1.2	配置身份验证	313
18.1.3	身份验证的工作原理	314
18.2	建立和删除索引	315
18.2.1	在独立的服务器上建立索引	315
18.2.2	在副本集上建立索引	315
18.2.3	在分片集群上建立索引	316
18.2.4	删除索引	316
18.2.5	注意内存溢出杀手	316
18.3	预热数据	317
18.3.1	将数据库移至内存	317
18.3.2	将集合移至内存	318
18.3.3	自定义预热	318
18.4	压缩数据	320
18.5	移动集合	321
18.6	预分配数据文件	322
第19章 持久性 323		
19.1	日记系统的用途	323
19.1.1	批量提交写入操作	324
19.1.2	设定提交时间间隔	325
19.2	关闭日记系统	325
19.2.1	替换数据文件	325
19.2.2	修复数据文件	326
19.2.3	关于mongod.lock文件	326
19.2.4	隐蔽的异常退出	327
19.3	MongoDB无法保证的事项	327
19.4	检验数据损坏	327
19.5	副本集中的持久性	329
第六部分 服务器管理		
第20章 启动和停止MongoDB 333		
20.1	从命令行启动	333
20.2	停止MongoDB	336

20.3	安全性	337
20.3.1	数据加密	338
20.3.2	SSL 安全连接	338
20.4	日志	338
第21章	监控MongoDB	341
21.1	监控内存使用状况	341
21.1.1	有关电脑内存的介绍	341
21.1.2	跟踪监测内存使用状况	342
21.1.3	跟踪监测缺页中断	343
21.1.4	减少索引树的脱靶次数	345
21.1.5	IO 延迟	345
21.1.6	跟踪监测后台刷新平均时间	346
21.2	计算工作集的大小	347
21.3	跟踪监测性能状况	349
21.4	监控副本集	352
第22章	备份	355
22.1	对服务器进行备份	355
22.1.1	文件系统快照	355
22.1.2	复制数据文件	356
22.1.3	使用mongodump	357
22.2	对副本集进行备份	359
22.3	对分片集群进行备份	360
22.3.1	备份和恢复整个集群	360
22.3.2	备份和恢复单独的分片	360
22.4	使用mongooplog 进行增量备份	361
第23章	部署MongoDB	363
23.1	设计系统结构	363
23.1.1	选择存储介质	363
23.1.2	推荐的RAID 配置	367
23.1.3	CPU	368
23.1.4	选择操作系统	368
23.1.5	交换空间	369
23.1.6	文件系统	369
23.2	虚拟化	370
23.2.1	禁止内存过度分配	370
23.2.2	神秘的内存	370
23.2.3	处理网络磁盘的IO 问题	371
23.2.4	使用非网络磁盘	372
23.3	系统配置	372
23.3.1	禁用NUMA	372
23.3.2	更智能地预读取数据	375
23.3.3	禁用大内存页面	376
23.3.4	选择一种磁盘调度算法	377
23.3.5	不要记录访问时间	377
23.3.6	修改限制	378
23.4	网络配置	379
23.5	系统管理	381
23.5.1	时钟同步	381
23.5.2	OOM Killer	381

23.5.3	关闭定期任务	382
附录A	安装MongoDB	383
附录B	深入MongoDB	387

精彩短评

- 1、很赞的一本书，配得上权威指南。可初学，可深入，可参考。
- 2、非常细，从入门介绍，使用，设计复制，分片服务器管理与部署监控都在使用层面介绍得非常完整，原理有但略少
- 3、似懂非懂，云里雾里，还需继续钻研。
- 4、新版变厚了一倍啦！（一一）你们快再去读一遍.....
- 5、MongoDB，这本书最好
- 6、太喜欢这本书了，第二版比第一版厚实了很多，作者作为官方驱动开发者，对mongodb了解得非常深入，讲解细致生动，而且此书翻译的非常好，文笔流畅，清晰易懂。适合入门者阅读，也适合高级读者，真是本好书。
- 7、看过第一版的PDF，特地去购买了一本，结果第二版比第一版页码增加了差不多一倍。。对非数据库管理人员来说内容有点太多了，可以当工具书
- 8、MongoDB的优缺点全都讲到了，还是非常不错的。NoSql的时代已经来到，不过对于很多人的转型非常困难。我用了很多年的关系数据库，很多思想很难转变。这才是转型的最大难度。
- 9、翻译很差
- 10、很全，技术不难，很容易读完的一本书。不过，在mapreduce方面的介绍太少太少太少，不知道作者是咋想的。
- 11、不能算完整看过。。但至少看过的地方还挺清楚的。
- 12、比第一版厚实了不少，3.0加入Wired Tiger存储引擎后变化还是蛮大的
- 13、其实只读完了前10章，复制和分片不能实际操作，感觉有点难以读下去。
- 14、很不错的书，从入门到深入，清晰有条理
- 15、图书馆偶然翻到，赶紧学起来
涉及到的是数据库概念那块，和服务器实战联系不大
- 16、参考性的书籍，覆盖面还算广，作为mongodb入门挺适合的
- 17、mongodb 第2版，读完1~9章，剩下的用到的时候再读。
- 18、Mongo简直是黑科技，只看了半本，另外半本应该暂时用不到，等用到时候再看，至于四星原因是--目测它14年的最新版不知道是在干什么，差了Mongo两个版本
- 19、没读完。
适合我这种从零开始学习数据库的人。
- 20、文档操作和索引讲解的很详细，备份还原还有很多不理解的，真正运用还需要在实际中去发现问题，看了最新的官方文档，稳定版已经增加到2.6，权限配置和本书讲解的不太一样了。
- 21、比1厚了不少

《MongoDB权威指南（第2版）》

精彩书评

1、新版变厚了一倍啦！（一 一）你们快再去读一遍.....为什么说我的评论太短了呢.....豆娘你在想啥.....而且是2014年出版的.....hmmm不想凑字了.....如果还不行我就要复制粘贴凑字数了（一 一）

章节试读

1、《MongoDB权威指南（第2版）》的笔记-第70页

用skip略过少量的文档还是不错的。但是要是数量非常多的话，skip就会变得很慢，因为要先找到需要被略过的数据，然后再抛弃这些数据。大多数数据库都会在索引中保存更多的元数据，用于处理skip，但是MongoDB目前还不支持，所以要尽量避免略过太多的数据。通常可以利用上次的结果来计算下一次查询条件项目当中模拟插入了120W条数据，在同一个文档当中单纯查询数据的速度还不错，主要是对查询的文档字段添加了索引，但是对查询结果的前台分页确有问题。具体来说是不设置任何查询条件的时候，会查询出来将近120W条满足条件的结果，使用mongodb的limit()和skip()来取出来第一页前20条数据，这样在后台的java程序当中只是这20条数据占用内存。

代码具体形式类似于用mongodb客户端执行

```
db.feedbackInfo.find(criteria).skip(0).limit(20) 获得第一页0-20条数据
```

```
db.feedbackInfo.find(criteria).skip(20).limit(20) 获得第二页20-40条数据
```

.....

```
db.feedbackInfo.find(criteria).skip(N).limit(20) 获得第二页N-N+20条数据
```

但问题在于随着不断翻页，skip的值N会越来越大，前台的反应越来越慢。很直接的一个表现就是在前台从第一页直接跳转进入最后一页根本反应不过来。

对于这个实际问题，原因就是本书这里所言的skip略过大量结果会带来性能问题，再根源地说是mongodb还不够完善，索引本身还比较简单。

具体的这个分页效率的问题，有两种思路：

第一，等mongodb升级，优化这个skip的执行效率。

第二，不用skip()而实现分页效果。这个思路的基础就是mongodb本身对于where查询和limit()的效率还比较不错，也就是本来分页的那个查询用where和limit速度还可以的前提（一般就是需要建立必要的索引）。假如这个前提不成立，那没法讨论。

本书接下来具体讨论了不使用skip对结果分页的实现例子，这个本质是对信息系统增加一个查询中间量——上次查询的业务数值，在逻辑上承担起跟skip相对等价的功能。

比如说是第一页查询是按照一个日期date值查询，第一次用

```
db.foo.find().sort({"date",-1}).limit(20)
```

而点击下一页的时候，事先将上次查询的date的边界值给传递过去，第二页查询的时候就使用新的find条件查询

```
db.foo.find({"date":{"$gt":latest.date}}); 而后再对查询结果排序即可
```

这种绕过skip的方式评价：

第一，很难比较方便地解决所有的分页问题，简单来说对于使用正则表达式的查询，根本无法通过记录边界条件来实现。

第二，不得不多传递上次查询的那个边界条件，增加了工作量，不够优雅。

第三，只能够解决一页一页往下翻页的问题，如果我要从第1页直接跳到100页，就束手无策

《MongoDB权威指南（第2版）》

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:www.tushu000.com