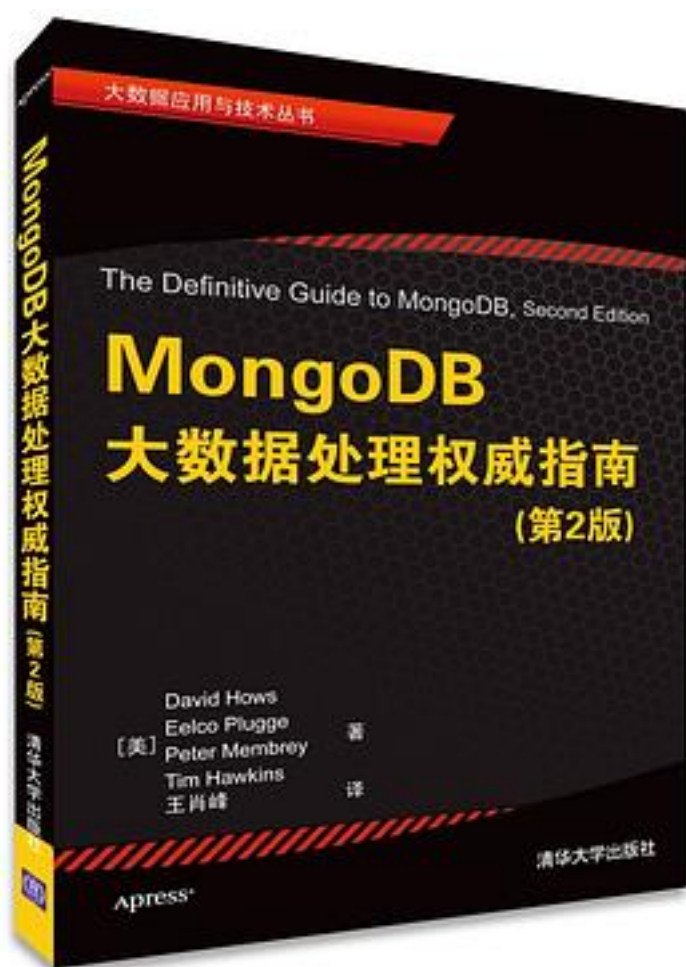


MongoDB大数据处理权威指南（第2版）



[MongoDB大数据处理权威指南（第2版）_下载链接1](#)

著者:(美) 豪斯(Hows, D.)

出版者:清华大学出版社

出版时间:2015-1

装帧:平装

isbn:9787302383666

本书根据MongoDB的最新版本做了相应更新，其中包含MongoDB n最新的所有特性，包括版本2.2中引入的聚集框架和版本2.4中引入的哈希索引。Mong

oDB是最流行的“大数据” NoSQL数据库技术，并且正处于发展中。来自10gen的David Hows以及具有丰富MongoDB开发经验的Peter Membrey和Eelco Plugge，组成了本书的专家团队，他们在本书中分享了自己的专业知识和经验，帮助你了解成为MongoDB专家所需的所有知识。

作者介绍:

David

Howes以优异的成绩毕业于澳大利亚新南威尔士州的卧龙岗大学。他第一次接触计算机，是在尝试不花钱的情况下改进家庭PC的性能。这促使他加入IT行业，David曾经担任过系统管理员、性能工程师、软件开发者、解决方案架构师和数据库工程师等职务。David也曾徒劳地尝试过多年足球，并且他的咖啡杯上写着“Grumble Bum”。

Peter Membrey是一位特许IT研究员，他拥有15年使用Linux和开源解决方案解决现实问题的经验。从17岁起他就是一位红帽认证工程师，并且他也有幸在Red Hat工作过，并编写了几本开源解决方案相关的书籍。他拥有利物浦大学的信息安全硕士学位，目前是香港理工大学的博士生，他的研究方向包括云计算、大数据、科学和安全。他与自己出色的妻子Sara和儿子Kaydyn一起居住在香港。尽管他的世界语一直不错，但广东话却在不断退步。

Eelco Plugge是一位年轻的IET/BSC专家，他对IT安全领域非常感兴趣。在21岁时他就成为一位加密专家，目前从事移动设备管理行业，偶尔也会编写一些书籍。Eelco最近拿到了利物浦大学的计算机安全硕士学位，并持有多项专业证书。他对Linux、网络安全和加密技术都抱有极大的热情。Eelco与他的年轻家庭一起居住在荷兰。他喜欢寿司，讨厌过于复杂的事情。Tim

Hawkins曾在1993年创建了世界上第一个在线分类广告门户网站loot.com，之后负责管理雅虎欧盟的许多非媒体属性产品，例如搜索、本地搜索、邮件、消息和社交网络。他目前正在管理美国主要电子零售商的一个大型离岸团队，负责开发和部署下一代电子商务应用。他喜欢帽子，讨厌复杂性。

目录: 第 I 部分 MongoDB基础

第1章 MongoDB简介 3

1.1 了解MongoDB哲学 3

1.1.1 使用正确的工具处理正确的工作 3

1.1.2 天生缺少对事务的支持 5

1.1.3 JSON和MongoDB 5

1.1.4 采用非关系的方式 7

1.1.5 选择性能还是特性 8

1.1.6 在任何地方均可运行数据库 8

1.2 将所有组合在一起 9

1.2.1 生成或创建键 9

1.2.2 使用键和值 9

1.2.3 实现集合 10

1.2.4 了解数据库 10

1.3 了解特性列表 10

1.3.1 使用面向文档存储(BSON) 11

1.3.2 支持动态查询 11

1.3.3 为文档创建索引 12

1.3.4 使用地理空间索引 13

1.3.5 分析查询 13

1.3.6 就地更新信息 13

1.3.7 存储二进制数据 14

- 1.3.8 复制数据 14
- 1.3.9 实施分片 14
- 1.3.10 使用映射和归约函数 15
- 1.3.11 全新的聚合框架 15
- 1.4 获取帮助 15
 - 1.4.1 访问网站 15
 - 1.4.2 与MongoDB开发者沟通 16
 - 1.4.3 剪切和粘贴MongoDB代码 16
 - 1.4.4 在Google小组中寻找解决方案 16
 - 1.4.5 利用JIRA跟踪系统 16
- 1.5 小结 16
- 第2章 安装MongoDB 17
 - 2.1 选择版本 17
 - 2.2 在系统中安装MongoDB 18
 - 2.2.1 在Linux中安装MongoDB 18
 - 2.2.2 在Windows中安装MongoDB 19
 - 2.3 运行MongoDB 20
 - 2.3.1 先决条件 20
 - 2.3.2 研究安装目录布局 20
 - 2.3.3 使用MongoDB shell 21
 - 2.4 添加额外的驱动 22
 - 2.4.1 安装PHP驱动 22
 - 2.4.2 确认PHP安装正确 25
 - 2.4.3 安装Python驱动 26
 - 2.4.4 确认PyMongo安装正确 28
 - 2.5 小结 29
- 第3章 数据模型 31
 - 3.1 设计数据库 31
 - 3.1.1 集合的更多细节 32
 - 3.1.2 使用文档 33
 - 3.1.3 在文档中内嵌或引用信息 34
 - 3.1.4 创建_id字段 35
 - 3.2 构建索引 36
 - 3.3 使用地理空间索引 37
 - 3.4 在真实世界中使用MongoDB 42
 - 3.5 小结 42
- 第4章 使用数据 43
 - 4.1 浏览数据库 43
 - 4.2 在集合中插入数据 44
 - 4.3 查询数据 45
 - 4.3.1 使用点号 47
 - 4.3.2 使用函数sort、limit和skip 47
 - 4.3.3 使用固定集合、自然顺序和\$natural 48
 - 4.3.4 获取单个文档 50
 - 4.3.5 使用聚集命令 50
 - 4.3.6 使用条件操作符 52
 - 4.3.7 使用正则表达式 59
 - 4.4 更新数据 59
 - 4.4.1 使用update()更新 60
 - 4.4.2 使用save()命令实现upsert 60
 - 4.4.3 自动更新信息 60
 - 4.4.4 指定匹配数组的位置 64
 - 4.4.5 原子操作 65

- 4.4.6 以原子的方式修改和返回文档 66
- 4.5 重命名集合 67
- 4.6 删除数据 67
- 4.7 引用数据库 68
 - 4.7.1 手动引用数据 68
 - 4.7.2 使用DBRef引用数据 70
- 4.8 使用索引相关的函数 71
 - 4.8.1 学习索引相关的命令 72
 - 4.8.2 强制使用某个索引查询数据 72
 - 4.8.3 限制查询匹配 73
- 4.9 小结 74
- 第5章 GridFS 75
 - 5.1 背景 75
 - 5.2 使用GridFS 76
 - 5.3 开始使用命令行工具 76
 - 5.3.1 使用_id键 77
 - 5.3.2 使用文件名 77
 - 5.3.3 决定文件长度 78
 - 5.3.4 使用块大小 78
 - 5.3.5 处理上传日期 79
 - 5.3.6 生成文件的哈希值 79
 - 5.4 查看MongoDB中的数据 79
 - 5.4.1 使用搜索命令 81
 - 5.4.2 删除 81
 - 5.4.3 从MongoDB中获取文件 82
 - 5.4.4 mongofiles命令小结 82
 - 5.5 使用Python 82
 - 5.5.1 连接数据库 83
 - 5.5.2 访问单词 83
 - 5.6 在MongoDB中添加文件 84
 - 5.7 从GridFS中读取文件 84
 - 5.8 删除文件 85
 - 5.9 小结 85
- 第II部分 在开发中使用MongoDB
- 第6章 PHP和MongoDB 89
 - 6.1 比较MongoDB和PHP中的文档 89
 - 6.2 MongoDB类 90
 - 6.3 连接和断开连接 91
 - 6.4 插入数据 92
 - 6.5 查询数据 94
 - 6.5.1 返回单个文档 94
 - 6.5.2 列出所有文档 95
 - 6.5.3 使用查询操作符 96
 - 6.5.4 查询特定信息 96
 - 6.5.5 排序、限制和忽略数据项 97
 - 6.5.6 统计匹配结果的数目 98
 - 6.5.7 使用聚集框架对数组分组 98
 - 6.5.8 使用hint()函数指定索引 99
 - 6.5.9 使用条件操作符重新定义查询 100
 - 6.5.10 正则表达式 106
 - 6.6 使用PHP修改数据 106
 - 6.6.1 使用update()函数更新数据 106
 - 6.6.2 节省更新操作的时间 108

- 6.6.3 使用\$since增加特定键的值 108
- 6.6.4 使用\$set修改键值 109
- 6.6.5 使用\$unset删除字段 109
- 6.6.6 使用\$rename重命名字段 110
- 6.6.7 在更新/插入期间使用\$setOnInsert修改键值 110
- 6.6.8 使用\$push向指定字段中添加值 111
- 6.6.9 使用\$push和\$each向某个键中添加多个值 111
- 6.6.10 使用\$addToSet将数据添加到数组中 112
- 6.6.11 使用\$pop从数组中删除元素 113
- 6.6.12 使用\$pull删除所有指定值 113
- 6.6.13 同时删除多个元素 113
- 6.6.14 使用save()函数更新/插入数据 114
- 6.6.15 以原子的方式修改文档 115
- 6.7 删除数据 116
- 6.8 DBRef 118
- 6.9 GridFS和PHP驱动 120
- 6.9.1 存储文件 120
- 6.9.2 在已存储的文件中添加元数据 121
- 6.9.3 获取文件 121
- 6.9.4 删除数据 122
- 6.10 小结 122
- 第7章 Python和MongoDB 123
- 7.1 在Python中使用文档 123
- 7.2 使用PyMongo模块 124
- 7.3 连接和断开 124
- 7.4 插入数据 125
- 7.5 搜索数据 126
- 7.5.1 搜索单个文档 126
- 7.5.2 搜索多个文档 127
- 7.5.3 使用点操作符 128
- 7.5.4 返回字段 128
- 7.5.5 使用sort()、limit()和skip()简化查询 129
- 7.5.6 聚集查询 130
- 7.5.7 使用hint()指定索引 132
- 7.5.8 使用条件操作符重定义查询 133
- 7.5.9 使用正则表达式执行搜索 137
- 7.6 修改数据 138
- 7.6.1 更新数据 139
- 7.6.2 修改操作符 140
- 7.6.3 使用save()函数快速保存文档 144
- 7.6.4 以原子的方式修改文档 145
- 7.6.5 使用参数 145
- 7.7 删除数据 146
- 7.8 在两个文档之间创建链接 146
- 7.9 小结 149
- 第8章 高级查询 151
- 8.1 文本搜索 151
- 8.1.1 文本搜索的代价和限制 152
- 8.1.2 启用文本索引 152
- 8.1.3 使用文本索引 153

- 8.1.4 其他语言中的文本索引 158
- 8.1.5 文本索引的复合索引 159
- 8.2 聚集框架 161
 - 8.2.1 \$group 162
 - 8.2.2 \$sum 162
 - 8.2.3 \$limit 163
 - 8.2.4 \$match 164
 - 8.2.5 \$sort 165
 - 8.2.6 \$unwind 166
 - 8.2.7 \$project 167
 - 8.2.8 \$skip 168
- 8.3 MapReduce 170
 - 8.3.1 MapReduce的工作方式 170
 - 8.3.2 创建测试文档 170
 - 8.3.3 使用map函数 170
 - 8.3.4 高级MapReduce 173
 - 8.3.5 调试MapReduce 174
- 8.4 小结 175
- 第III部分 使用MongoDB处理大数据
- 第9章 数据库管理 179
 - 9.1 使用管理工具 179
 - 9.1.1 mongo——MongoDB控制台 180
 - 9.1.2 使用第三方管理工具 180
 - 9.2 备份MongoDB服务器 180
 - 9.2.1 创建第一个备份 180
 - 9.2.2 备份单个数据库 182
 - 9.2.3 备份单个集合 182
 - 9.3 深入学习备份 183
 - 9.4 恢复单个数据库或集合 184
 - 9.4.1 恢复单个数据库 185
 - 9.4.2 恢复单个集合 185
 - 9.5 自动化备份 185
 - 9.5.1 使用本地数据存储 185
 - 9.5.2 使用远端数据存储(基于云) 187
 - 9.6 备份大数据库 188
 - 9.6.1 使用隐藏的辅助服务器备份数据 189
 - 9.6.2 使用日志文件系统创建快照 189
 - 9.6.3 使用卷管理器时的磁盘布局 191
 - 9.7 将数据导入MongoDB 191
 - 9.8 从MongoDB导出数据 192
 - 9.9 通过限制对MongoDB服务器的访问保护数据安全 194
 - 9.10 使用认证保护服务器 194
 - 9.10.1 添加admin用户 195
 - 9.10.2 启用认证 195
 - 9.10.3 在mongo控制台中执行认证 195
 - 9.10.4 MongoDB用户角色 196
 - 9.10.5 代理凭证 197
 - 9.10.6 修改用户凭证 197
 - 9.10.7 添加只读用户 198
 - 9.10.8 删除用户 198
 - 9.10.9 在PHP应用中认证连接 198
 - 9.11 管理服务器 199

- 9.11.1 启动服务器 199
- 9.11.2 重新配置服务器 200
- 9.11.3 获得服务器版本 201
- 9.11.4 获得服务器状态 201
- 9.11.5 关闭服务器 203
- 9.12 使用MongoDB日志文件 203
- 9.13 验证和修复数据 204
 - 9.13.1 修复服务器 204
 - 9.13.2 验证单个集合 206
 - 9.13.3 修复集合验证错误 207
- 9.14 升级MongoDB 209
- 9.15 MongoDB的滚动升级 209
- 9.16 监控MongoDB 209
- 9.17 使用MongoDB管理服务 211
- 9.18 小结 211
- 第10章 优化 213
 - 10.1 优化服务器硬件 213
 - 10.1.1 了解MongoDB使用内存的方式 213
 - 10.1.2 了解工作集大小 214
 - 10.1.3 选择正确的数据库服务器硬件 214
 - 10.2 评估查询性能 214
 - 10.2.1 MongoDB分析器 215
 - 10.2.2 使用explain()分析特定的查询 217
 - 10.2.3 使用分析器和explain()优化查询 218
 - 10.3 管理索引 222
 - 10.3.1 显示索引 222
 - 10.3.2 创建简单的索引 223
 - 10.3.3 创建复合索引 223
 - 10.4 指定索引选项 224
 - 10.4.1 使用{background: true}在后台创建索引 224
 - 10.4.2 使用{unique: true}创建唯一键索引 225
 - 10.4.3 使用{dropdups: true}自动去重 225
 - 10.4.4 使用{sparse: true}创建稀疏索引 226
 - 10.4.5 TTL索引 226
 - 10.4.6 文本索引 227
 - 10.4.7 删除索引 227
 - 10.4.8 重建集合索引 227
 - 10.5 MongoDB选择索引的方式 228
 - 10.6 使用hint()强制使用特定的索引 229
 - 10.7 优化小对象的存储 229
 - 10.8 小结 230
- 第11章 复制 231
 - 11.1 MongoDB复制的目标 231
 - 11.1.1 改善可扩展性 231
 - 11.1.2 改善持久性/可靠性 232
 - 11.1.3 提供隔离性 232

- 11.2 复制基础 232
 - 11.2.1 主服务器的定义 233
 - 11.2.2 辅助服务器的定义 233
 - 11.2.3 仲裁服务器的定义 234
- 11.3 深入学习oplog 234
- 11.4 实现复制集 235
 - 11.4.1 创建复制集 236
 - 11.4.2 启动服务器成员 237
 - 11.4.3 向复制集中添加服务器 238
 - 11.4.4 添加仲裁服务器 241
 - 11.4.5 复制集链 242
 - 11.4.6 管理复制集 242
 - 11.4.7 为复制集成员设置选项 246
 - 11.4.8 从应用连接到复制集 248
- 11.5 小结 251
- 第12章 分片 253
 - 12.1 了解分片的需求 253
 - 12.2 对数据进行水平和垂直分区 254
 - 12.2.1 对数据进行垂直分区 254
 - 12.2.2 对数据进行水平分区 254
 - 12.3 分析一个简单的分片场景 255
 - 12.4 使用MongoDB实现分片 256
 - 12.4.1 创建分片设置 258
 - 12.4.2 在集群中添加新的分片 260
 - 12.4.3 从集群中移除分片服务器 262
 - 12.4.4 确定连接的方式 263
 - 12.4.5 列出分片服务器的状态 263
 - 12.4.6 使用复制集实现分片 264
 - 12.5 均衡器 264
 - 12.6 哈希片键 266
 - 12.7 标签分片 267
 - 12.8 小结 269
- • • • • (收起)

[MongoDB大数据处理权威指南（第2版）](#) [下载链接1](#)

标签

mongodb

编程

大数据

计算机

python

正在阅读中，现在给我的感觉是内容感不够，翻译效果略待提高。

我的书架

评论

从教你安装就能看出来非常一般，入门读物水平，翻译也不怎么好，很多话不通顺

书很不错，翻译也很到位。Mongodb涉及到的也都讲了。如果想要学习Mongodb，推荐。

偏重实践而缺少原理，用来做大数据收集处理等方面能做个查询手册。

翻完了 两天 入门书籍 感觉很乱…翻译也不好 错别字很多 想看的可以看英文版 已经出第三版了 这个版本还是2.4测试… 哎 还要看一遍第三版… 有人吐槽书名翻译英文就有大数据处理这几个字的 apress好像都是入门书 估计卖得好

当工具书，不够深入

不适合入门

[MongoDB大数据处理权威指南（第2版）_下载链接1](#)

书评

随手翻一下，到237页，看到原文中的一句命令如下：\$ mongod --dbpath /db/active1/data --port 27021 --replSettestset 看到问题在哪里了没？
仔细看看！新手是不可能知道的。我照着命令打了，提示说未定义的参数 --replSettestset！！！ 然后我在命令中间加了个空格，这个...

相对RDBMS，MongoDB本身不是太复杂的数据库，这本书可以很轻松让你了解关于MongoDB全面的知识，从GridFS到replica sets,从Php、Python如何连接MongoDB到Shards；这是第二版,覆盖比较新的2.4版本.

[MongoDB大数据处理权威指南（第2版）_下载链接1](#)