

深入云计算

MongoDB

管理与开发实战详解

© 邹贵金 著

书中源代码下载地址:

<http://www.tdpress.com/zyzx/tsscflwj>

精准的内容梯度安排

遵循读者学习习惯和 MongoDB 数据库技术应用实践, 合理安排图书内容; 精挑细选经典实例, 嵌入完善的代码注释。

丰富的实用经验阐述

将多年开发经验融入其中, 读者在全面掌握 MongoDB 应用和开发技术的同时更能快速分析和解决实际问题的能力。

Getting you the Best Books!

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

深入浅出

MongoDB

管理与开发实战详解

◎ 邹贵金 著

Getting you the Best Books!

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

作为基于分布式文件存储的数据库，在目前的云计算实践中，MongoDB 炙手可热。本书系统全面的介绍了 MongoDB 开发、管理、维护和性能优化等方方面面。详细而深入，对 MongoDB 的开发和管理方法进行了详细的讲解，也对 MongoDB 的工作机制进行了深入的探讨。注重实战，通过实际中的案例为读者讲解使用 MongoDB 时遇到的各种问题，并给出了解决方案。

本书旨在帮助云计算初学者迅速掌握 MongoDB 数据库，提升读者在云计算实践中的应用和开发能力。同时本书极强的系统性和大量翔实的案例对于有一定基础的中高级用户有非常好的参考价值。

图书在版编目（CIP）数据

深入云计算：MongoDB 管理与开发实战详解 / 邹贵金著. — 北京：中国铁道出版社，2013.6
ISBN 978-7-113-16228-3

I. ①深… II. ①邹… III. ①关系数据库系统 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2013)第 055415 号

书 名：深入云计算：MongoDB 管理与开发实战详解
作 者：邹贵金 著

责任编辑：荆 波
封面设计：付 巍
责任印制：赵星辰

读者服务热线：010-63560056
特邀编辑：赵树刚

出版发行：中国铁道出版社（北京市西城区右安门西街 8 号 邮政编码：100054）
印 刷：北京铭成印刷有限公司
版 次：2013 年 6 月第 1 版 2013 年 6 月第 1 次印刷
开 本：787mm×1 092mm 1/16 印张：24.75 字数：595 千
书 号：ISBN 978-7-113-16228-3
定 价：59.80 元

版权所有 侵权必究

凡购买铁道版图书，如有印制质量问题，请与本社发行部联系调换。

NoSQL 数据库与传统的关系型数据库相比,具有操作简单、完全免费、源码公开、随时下载等特点,并可以用于各种商业目的。这使 NoSQL 产品广泛应用于各种大型门户网站和专业网站,大大降低了运营成本。

2010 年,随着互联网 Web 2.0 网站的兴起, NoSQL 在国内掀起一阵热潮,其中风头最劲的莫过于 MongoDB 了。越来越多的业界公司已经将 MongoDB 投入实际的生产环境,很多创业团队也将 MongoDB 作为自己的首选数据库,创造出非常多的移动互联网应用。

MongoDB 的文档模型自由灵活,可以让用户在开发过程中畅顺无比。对于大数据量、高并发、弱事务的互联网应用, MongoDB 可以应对自如。MongoDB 内置的水平扩展机制提供了从百万到十亿级别的数据量处理能力,完全可以满足 Web 2.0 和移动互联网的数据存储需求,其开箱即用的特性也大大降低了中小型网站的运维成本。

本书由浅入深,全面、系统地介绍了 MongoDB 基础、应用、管理、性能优化、数据库的架构、环境搭建实例、编程实例等内容。书中的每一章都提供了大量的实例代码,以方便读者进行练习和学习。每个例程都经过精挑细选,具有很强的针对性,适合各个阶段的读者学习。本书既注重基础知识,又非常注重实践,读者可以快速上手并迅速提高。通过学习本书的内容,读者不仅可以全面掌握 MongoDB 的应用,还可以获得快速分析和解决实际问题的能力,从而能够在最短的时间内,以最好的效果来解决问题,提升工作效率。

本书的特点

1. 每章都提供了大量的实例代码

为了便于读者练习和学习,每章都提供了大量的实例代码。

2. 结构合理,内容全面、系统

本书详细介绍了 MongoDB 基础、应用、管理、性能优化、数据库的架构、环境搭建实例、编程实例等内容,将实际项目开发经验贯穿于全书,思想和内容都非常丰富。根据读者的学习习惯和内容的梯度合理安排,更加适合读者学习。

3. 叙述翔实,例程丰富

本书有详细的例程,每个例子都经过精挑细选,有很强的针对性。书中的程序都有完整的代码,而且代码非常简洁和高效,便于读者学习和调试。读者也可以直接重用这些代码来解决自己的问题。

4. 实际项目相结合

本书提供有项目环境搭建实例、Java 编程实例、C#编程实例,供读者练习和学习。

本书内容体系

本书共 16 章，分为 6 篇，各篇对应的章节和具体内容介绍如下：

第 1 篇包括第 1~3 章，主要介绍 MongoDB 的基础知识，对 NoSQL 的介绍，讲解 MongoDB 的文档模型，特点，文档增加、修改和删除等基本操作。学习本章之后，读者可以对 MongoDB 进行简单的操作。

第 2 篇包括第 4~7 章，主要介绍 MongoDB 的高级技术，例如高级查询、高级更新、高级功能的应用，游标，存储过程，Capped 集合，GridFS 存储文件，MapReduce 统计。学习本章之后，读者可以熟练的使用 MongoDB 来完成日常的业务需求。

第 3 篇包括第 8 章，主要介绍 MongoDB 常用的运维管理工具、各种命令管理以及访问控制等方面的技术，看完第三篇内容就能成为 MongoDB DBA，因为这一篇包含了 MongoDB DBA 所需的全部技能。

第 4 篇包括第 9 章和第 10 章，详细的介绍 MongoDB 的索引、优化，以及性能监控等方面的只是，本篇内容将使读者具备对 MongoDB 进行调优的能力。

第 5 篇包括第 11~13 章，重点的讲解了从主从复制、复制集、sharding 分片、复制集+分片等高可用架构的细节，掌握了第五篇内容就可以独立设计出 MongoDB 应用系统的架构。

第 6 篇包括第 14~16 章，分别讲解了如何使用 Java 和 C#来操作 MongoDB 数据库，掌握了第 6 篇内容后就可以将 Java 和 C#开发技术与 MongoDB 数据库结合起来，从而完成特定应用系统的开发。

本书读者对象

- MongoDB 初学者；
- 准备扩展 MongoDB 知识的开发人员；
- 有兴趣了解 MongoDB 的 DBA 或者数据库应用程序编程人员；
- MongoDB 数据库爱好者；
- 正在学习数据库课程的计算机相关专业的学生；
- 利用 MongoDB 数据库开发应用的技术人员；
- 大中专院校的学生和老师；
- 相关培训学校的学员。

本书作者

本书作者邹贵金具备多年信息系统项目开发经历，曾参与人力资源管理系统、深圳市招调工系统、电信 C 网项目、广东电信项目、辽宁移动项目、湖南联通网项目、辽宁移动域名改造和风格改造项目、辽宁移动二次确认项目、电信集团 UDB 统一认证项目、移动 OMP 项目（PortalONE）、专利检索项目、信托非资本市场投资管理系统、意健险承保系统等大型项目的开发和设计，积累了丰富的项目和管理经验。精通 MongoDB、Oracle、SQL Server、MySQL、DB2 等大型数据库的运行维护和管理。

作为国内 NoSQL 领域的早期践行者之一，对 NoSQL 数据库中的 MongoDB 有着丰富的实战经验。云时代的来临，数据库至关重要，NoSQL 最终将脱颖而出，到底 NoSQL 有哪些优势！本书以当今 NoSQL 数据库中最热门的 MongoDB 为例给大家讲解。

特别感谢

在本书编写过程中，得到了家人、同事和朋友的支持和帮助。在此表示感谢！

在此感谢我的妻子（刘学昕）、父母、家人一直以来对我的支持，还要感谢郭现杰一直以来对我的指导和帮助，感谢所有帮助过我的人。由于时间仓促，笔者水平有限，书中难免存在遗漏和不足之处，恳请广大读者提出宝贵意见。

QQ: 422245310

微博: <http://weibo.com/zouguijin>

作 者

2013 年 2 月

第 1 篇 基础篇

第 1 章 MongoDB 简介

1.1	关系型数据库简介	2
1.2	关系型数据库面临的问题	3
1.3	NoSQL 的崛起	5
1.4	MongoDB 是如何解决这些问题的	6
1.5	初识 MongoDB	8
1.5.1	MongoDB 的特点	8
1.5.2	数据模型	9
1.5.3	扩展性	11
1.5.4	功能性	11
1.5.5	速度快	12
1.5.6	简便的管理	13
1.5.7	使用场合	13
1.6	本章小结	14

第 2 章 MongoDB 基本原理与安装

2.1	数据库结构	15
2.2	文档	17
2.3	集合	18
2.3.1	集合的无模式	19
2.3.2	集合的命名	19
2.4	MongoDB 数据类型	20
2.4.1	基本数据类型	20
2.4.2	数字类型	21

2.4.3	日期类型	21
2.4.4	数组类型	22
2.4.5	内嵌文档类型	22
2.4.6	_id 键和 ObjectId 对象	22
2.5	MongoDB 的下载和安装	24
2.5.1	MongoDB 的下载	24
2.5.2	在 Windows 平台下的下载与安装	25
2.5.3	在 Linux 平台下的下载与安装	28
2.6	MongoDB shell 的使用	31
2.6.1	启动 Shell	31
2.6.2	使用 Shell 对 MongoDB 的基本操作	32
2.6.3	使用 Shell 的诀窍	33
2.6.4	特殊的集合名	34
2.7	本章小结	35

第 3 章 文档的增加、修改及删除

3.1	插入并保存文档	36
3.1.1	插入的原理和作用	37
3.1.2	批量插入	37
3.2	删除文档	38
3.3	修改文档	39
3.3.1	整个文档的替换	39
3.3.2	使用修改器	42
3.3.3	upsert 和 save 更新	56
3.3.4	修改多个文档	60
3.3.5	修改文档并返回修改后的文档	63
3.4	数据库响应	67
3.4.1	安全操作	67
3.4.2	捕获异常	67
3.5	客户端请求和 MongoDB 数据库连接	68
3.6	本章小结	69

第2篇 应用篇

第4章 查询

4.1 find 简介	72
4.1.1 返回指定的键	73
4.1.2 find 查询限制	74
4.2 条件操作符	75
4.2.1 \$all 匹配所有	78
4.2.2 \$exists 判断字段是否存在	79
4.2.3 null 值处理	80
4.2.4 \$mod 取模运算	81
4.2.5 \$ne 不等于	82
4.2.6 \$in 包含	83
4.2.7 \$nin 不包含	83
4.2.8 \$size 数组元素个数	84
4.2.9 正则表达式匹配	84
4.2.10 Javascript 查询和\$where 查询	86
4.2.11 count 查询记录条数	88
4.2.12 limit 限制返回记录数	89
4.2.13 skip 限制返回记录的起点	89
4.2.14 sort 排序	90
4.2.15 分页查询	92
4.2.16 随机抽取文档	96
4.3 distinct 找出给定键所有不同的值	97
4.4 group 分组	98
4.4.1 使用完成器	105
4.4.2 将函数作为键使用	107
4.5 游标	110
4.6 存储过程	112
4.7 本章小结	113

第 5 章 Capped 集合

5.1 特性	114
5.2 使用和约束	114
5.3 应用	114
5.3.1 创建 capped collection	115
5.3.2 限制 capped collection 中对象个数	115
5.4 注意事项	117
5.5 本章小结	117

第 6 章 GridFS 存储文件

6.1 为什么要用 GridFS	118
6.2 如何实现海量存储	118
6.3 语言支持	118
6.4 简单介绍	119
6.5 命令行工具	119
6.6 内部原理	122
6.7 本章小结	124

第 7 章 MapReduce 统计

7.1 Map 函数	126
7.2 Reduce 函数	127
7.3 结果存储	127
7.4 对 Reduce 函数结果进一步处理	129
7.5 其他控制细节	130
7.6 本章小结	131

第 3 篇 管理篇

第 8 章 管理

8.1 启动和停止 MongoDB	134
8.1.1 使用命令行启动	134

8.1.2	配置文件	136
8.1.3	Daemon 方式启动	137
8.1.4	mongod 参数说明	138
8.1.5	停止数据库	139
8.2	访问控制	142
8.2.1	绑定 IP 内网地址访问 MongoDB 服务	142
8.2.2	设置监听端口	142
8.2.3	使用用户名和口令登录	143
8.3	命令行操作	147
8.3.1	通过 eval 参数执行指定语句	147
8.3.2	执行指定文件中的内容	147
8.4	进程控制	149
8.4.1	查看活动进程	149
8.4.2	结束进程	151
8.5	监控	152
8.5.1	使用管理接口	152
8.5.2	serverStatus	154
8.5.3	mongostat	156
8.5.4	第三方插件	157
8.6	数据导出 mongoexport	157
8.6.1	常用导出方法	158
8.6.2	导出 CSV 格式的文件	159
8.7	数据导入 mongoimport	161
8.7.1	导入 JSON 数据	161
8.7.2	导入 CSV 数据	162
8.8	数据备份和修复	163
8.8.1	数据文件备份	163
8.8.2	数据备份 mongodump	163
8.8.3	数据恢复 mongorestore	165
8.8.4	fsync 和锁	167
8.8.5	从属备份	168
8.8.6	修复	168
8.9	本章小结	169

第4篇 性能篇

第9章 索引

9.1 索引简介	171
9.1.1 基础索引	171
9.1.2 文档索引	173
9.1.3 组合索引	175
9.1.4 唯一索引	177
9.1.5 强制使用索引	178
9.1.6 扩展索引	180
9.1.7 索引内嵌文档中的键	180
9.1.8 为排序创建索引	181
9.1.9 索引名称	181
9.1.10 删除索引	182
9.2 explain 执行计划	185
9.3 优化器 profiler	191
9.3.1 开启 profiler 功能	191
9.3.2 查询 profiler 日志	192
9.4 索引管理	192
9.5 本章小结	193

第10章 性能优化

10.1 优化方案	195
10.1.1 优化方案 1: 创建索引	195
10.1.2 优化方案 2: 限定返回结果条数	195
10.1.3 优化方案 3: 只查询使用到的字段，而不查询所有字段	196
10.1.4 优化方案 4: 采用 capped 集合	196
10.1.5 优化方案 5: 采用 Server Side Code Execution	196
10.1.6 优化方案 6: hint	196
10.1.7 优化方案 7: 采用 profiling	197
10.2 性能监控	197

10.2.1	mongosniff 底层监控	197
10.2.2	mongostat 查看运行中的实例统计信息	199
10.2.3	db.serverStatus 查看实例运行状态	200
10.2.4	db.stats 查看数据库状态	203
10.2.5	第三方工具	205
10.3	本章小结	205

第 5 篇 架构篇

第 11 章 复制集

11.1	主从复制	208
11.1.1	选项	210
11.1.2	添加及删除源	211
11.2	复制集	212
11.2.1	部署复制集	213
11.2.2	初始化复制集	214
11.2.3	复制集中的节点	215
11.2.4	故障切换和活跃节点选举	216
11.3	主从配置信息	217
11.4	管理维护复制集	219
11.4.1	读写分离	219
11.4.2	故障转移	220
11.4.3	增减节点	222
11.5	工作原理	228
11.5.1	主从操作日志 oplog	229
11.5.2	同步	230
11.5.3	复制状态和本地数据库	230
11.5.4	阻塞复制	231
11.6	管理	232
11.6.1	诊断	232
11.6.2	变更 oplog 的大小	232
11.6.3	复制的认证问题	233

11.7 本章小结	233
-----------------	-----

第 12 章 sharding 分片

12.1 分片简介	234
12.2 MongoDB 中的自动分片	234
12.3 片键	235
12.3.1 将已有的集合分片	236
12.3.2 递增片键还是随机片键	236
12.3.3 片键对操作的影响	236
12.4 建立分片	237
12.4.1 启动 Config Server 配置服务器	238
12.4.2 启动 mongos 路由	238
12.4.3 启动 Shard Server 服务器	239
12.4.4 配置 Sharding	239
12.4.5 验证 Sharding 正常工作	240
12.5 管理维护 Sharding	242
12.5.1 列出所有的 Shard Server	242
12.5.2 查看 Sharding 信息	242
12.5.3 判断是否是 Sharding	243
12.5.4 对现有的集合进行 Sharding	243
12.5.5 新增 Shard Server	245
12.5.6 移除 Shard Server	247
12.6 本章小结	249

第 13 章 复制集+shardin 分片体

13.1 创建数据目录	251
13.2 配置复制集	251
13.2.1 配置 shard1 所用到的复制集	251
13.2.2 配置 shard2 所用到的复制集	252
13.3 配置多台 Config Server	253
13.4 配置多台 mongos	253

13.5 配置 Shard Cluster	254
13.6 验证 Sharding 正常工作	254
13.7 管理分片	255
13.7.1 配置集合	255
13.7.2 分片命令	257
13.8 本章小结	258

第 6 篇 实例篇

第 14 章 实际项目搭建 MongoDB 环境

14.1 环境搭建实例	260
14.1.1 服务器信息	260
14.1.2 分片结果表	260
14.1.3 复制集+sharding 分片架构图	261
14.1.4 MongoDB 环境搭建步骤	261
14.2 遇到的问题	281
14.3 本章小结	282

第 15 章 Java 对 MongoDB 的基本操作实例

15.1 Java 快速入门	283
15.1.1 安装 Java 驱动程序和开发环境	283
15.1.2 访问控制	285
15.1.3 Java 对 MongoDB 数据库的基本操作	286
15.1.4 Java 驱动的一致性	292
15.2 Java 常用操作	293
15.2.1 Java 对 MongoDB 的操作实例	293
15.2.2 对用户的操作	297
15.2.3 对集合的操作	300
15.2.4 对索引的操作	310
15.3 Java 高级查询	315
15.3.1 通过游标获取所有的文档	315
15.3.2 比较运算符	316
15.3.3 逻辑运算符	327

15.3.4	正则表达式查询	328
15.3.5	skip 跳过查询	330
15.4	Java 操作 GridFS	331
15.4.1	上传文件	331
15.4.2	查询数据	333
15.4.3	下载文件	334
15.4.4	删除数据	336
15.5	本章小结	337

第 16 章 C#对 MongoDB 的基本操作实例

16.1	C#快速入门	338
16.1.1	下载驱动和配置开发环境	338
16.1.2	访问控制	340
16.1.3	C#对数据库的基础操作	341
16.2	C#驱动一致性	345
16.3	C#常用操作	346
16.3.1	对 MongoDB 的操作实例	346
16.3.2	对用户的操作	349
16.3.3	对集合的操作	352
16.3.4	对索引的操作	359
16.4	C#高级查询	362
16.4.1	判断列名是否存在	362
16.4.2	比较运算符	364
16.4.3	逻辑运算符	370
16.4.4	正则表达式查询	372
16.4.5	跳过查询	373
16.5	C#操作 GridFS	374
16.5.1	上传文件	374
16.5.2	查询数据	375
16.5.3	下载文件	376
16.5.4	删除数据	377
16.6	本章小结	378

第1篇 基础篇



首先简单介绍了关系型数据库,从关系型数据库所遇到的问题讲到 NoSQL 的崛起,然后讲述了 NoSQL 是如何解决这些问题的。MongoDB 是 NoSQL 中最热门的一种数据库,本篇主要讲解了 MongoDB 的基本操作和技巧,包括 MongoDB 的特点、数据模型、扩展性、功能性、管理、使用场合等等。然后更深入的讲解了 MongoDB 的数据结构、文档、集合、MongoDB shell 的使用。还有文档的增加、修改及删除。