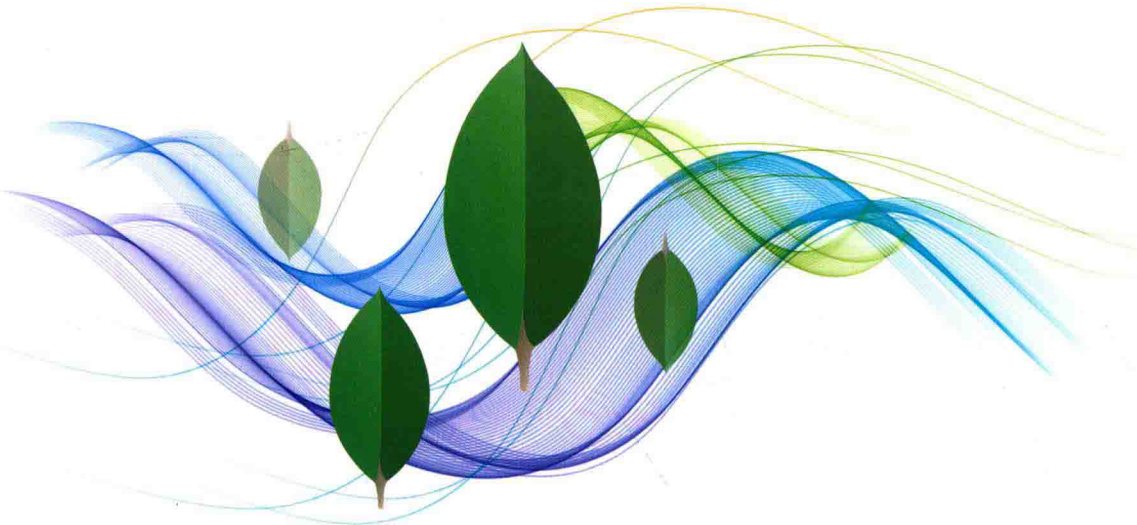
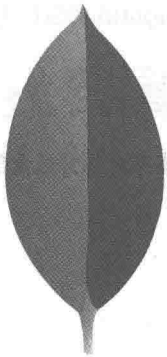
Broadview®
www.broadview.com.cn



MongoDB 运维实战

张甦 贺磊 / 著

中国工信出版集团
電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn



MongoDB 运维实战

张甦 贺磊 / 著

電子工業出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

MongoDB 自 2009 年推出以来，历经了近十年的发展，在这十年中，数据库领域可谓经历了翻天覆地的变化。

本书深入剖析 MongoDB 新旧版本的特性，结合生产案例详细讲解 MongoDB 的常见故障；引领学习 MongoDB 索引，以便更好地掌握 MongoDB 性能调优技巧；描述备份恢复的重要性，让读者掌握 MongoDB 备份恢复技巧；充分利用 MongoDB 的水平扩展能力，详解 MongoDB 复制集、分片架构环境；最后讲解如何使用 PMM 性能监控平台，做好线上 MongoDB 的监控工作。

图书在版编目（CIP）数据

MongoDB 运维实战 / 张甦，贺磊著. —北京：电子工业出版社，2018.9
ISBN 978-7-121-34989-8

Ⅰ. ①M… Ⅱ. ①张… ②贺… Ⅲ. ①关系数据库系统 Ⅳ. ①TP311.138

中国版本图书馆 CIP 数据核字（2018）第 207727 号

责任编辑：陈晓猛
印　　刷：三河市君旺印务有限公司
装　　订：三河市君旺印务有限公司
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　　邮编：100036
开　　本：787×980　1/16　印张：14.25　字数：273.6 千字
版　　次：2018 年 9 月第 1 版
印　　次：2018 年 9 月第 1 次印刷
定　　价：69.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

推荐序 1

找到属于你自己的那束光

张甦老师的新书出版，邀我来写几句话，下面谈谈我对数据库领域变革的一点观察和理解。

MongoDB 自 2009 年推出以来，转眼已经近十年，这十年间，正是数据库领域风起云涌的 十年。在同样的时间进程中，阿里巴巴在 2008 年提出去 IOE 理念，推动了中国由互联网至传统行业的数据应用的深刻变革。

我曾经在 2015 年的 DTCC 数据库大会上提出，我们今天已经进入了“后 IOE 时代”，这个时代的典型特征就是“百花齐放”，数据库新产品的不断涌现，为我们带来了新的可能，不同场景可以有多种不同的产品和解决方案，用户因此获得了“自由”。

而 MongoDB 的出现，因其面向文档，具有 Schema Free 等灵活优势，让用户在管理文档、日志，以及基于社交、物流等场景有了一个更好的选择，于是其市场经历了快速的增长，并扛起了 NoSQL 的大旗，也因此在 DB-Engines 的数据库流行度排行榜中，MongoDB 荣膺 2013 和 2014 的年度数据库。

2017 年 10 月，MongoDB 在纳斯达克上市，成为今天市值 30 亿美元的数据库公司，这不得不说是近代数据库历史上的一个巨大成功。而相应地，另外一个更受欢迎的开源数据库 MySQL，几经辗转成为 Oracle 的囊中之物，原因何在？

最近在做一个 MongDB 的迁移，将数据库的存储引擎从 MMAPv1 更换为 WiredTiger。同时回顾了一下历史， 2014 年 MongoDB 收购了 WiredTiger 公司，WiredTiger 为其开发了一个专用版本的存储引擎，今天成为 MongoD 的默认存储引擎，我们不得不钦佩 MongoDB 的英明之处。对比一下 MySQL 的发展历程，当 MySQL 的最佳存储引擎 InnoDB 被 Oracle 釜底抽薪收购（2006 年）之后，MySQL 最后被 SUN 收购（2008 年），辗转落入 Oracle 之手（2009 年），而

自 2009 年 MySQL 5.5 开始， InnoDB 就成为 MySQL 默认存储引擎。

决定一个产品成败的是技术，而决定一家公司成败的，往往是视野。

张甦老师的学习和成长，兼具技术和视野，他不断学习研究和砥砺，使自己获得了深厚的技术体验，而选择 MySQL 和 MongoDB 入行，更可见他对于开源的信心。他此前出版的《MySQL 王者晋级之路》，深受读者欢迎，而 MongoDB 更是指引他向前的“一束光”。现在这束光放射开来，希望能够让更多读者见证 MongoDB 的光彩和未来。

我也祝福走在技术道路上的每一位朋友，能早日找到指引自己的那束光！

盖国强，云和恩墨创始人，Oracle ACE 总监

推荐序 2

近两年，随着互联网的迅猛发展，Oracle 之外的各种开源数据库迅速崛起，数据库领域呈现出了百花齐放、百家争鸣的局面。MongoDB 就是其中最为夺目的一朵开源之花。在 DB-Engines 公布的 2018 年 2 月的数据库排名中，MongoDB 成为榜单中涨幅最大的一个，上涨了 5.47 个百分点，位列榜单第 5 名。

MongoDB 是可以应用于各种规模的企业、各个行业及各类应用程序的开源数据库。MongoDB 能够使企业的数据库更具敏捷性和可扩展性，各种规模的企业都可以通过使用 MongoDB 来创建新的应用。该公司是 NoSQL 数据库技术领域的知名公司。MongoDB 采用分布式基础架构，并且深受移动应用和 Web 应用开发者的欢迎。此外，MongoDB 还是一个基于文件的数据库。在 MongoDB 中，数据被编码成能够兼容多种不同数据格式的文件。MongoDB 的流行程度是显而易见的，目前其应用在全球范围内的下载次数已经突破了 1000 万次。简单来说，使用 MongoDB 能够提高与客户之间的工作效率，缩短产品上线时间，以及降低企业成本。

MongoDB 是一个基于分布式文件存储的数据库，旨在为企业提供可扩展的高性能数据存储解决方案。其数据库特点是高性能、易部署、易使用、储存方便。首先，它支持的数据结构非常松散，是与 JSON 相类似的 BSON 文档格式，因此可以存储比较复杂的数据类型，且该格式文档比较易读、高效。其次，MongoDB 支持的查询语言非常强大，其语法有点类似于面向对象的查询语言，几乎可以实现类似关系数据库单表查询的绝大部分功能，包括传统数据库的功能如二级索引、完整的查询系统等，而且还支持对数据建立索引。此外，MongoDB 的数据可实现复制和故障恢复，还具有在云端的伸缩性，支持水平的数据库集群延伸。

本书作者张甦先生在开源数据库领域深耕细作多年，技术功底扎实，实战经验丰富，对 MongoDB、MySQL 等数据库都有深入的研究，在各大开源数据库技术论坛、网站上颇有名气。

更为难得的是，张甦先生拥有丰富的数据库培训经验，擅长深入浅出地讲解技术问题，能够用简单明了的语言阐述复杂疑难的技术故障。

在数据库领域从业多年，我见过很多技术牛人，但是很多人仅限于自己牛却不擅长将技术分享给更多的人，他们的价值总是有限的，而且是局限于自身的，只有像张甦先生这样能做技术、能讲技术，还能写技术的人才能将技术的价值无限放大，让更多的人受益。

本书是两位作者多年工作和研究经验的总结，语言简单明了，实战案例丰富。本书对工作中常见的各种故障给出了作者的分析和解决方案，具有非常高的实用价值，无论是刚开始学习MongoDB 的小白，还是有一定工作经验的老 DBA，都可以从中受益。

张甦先生是一位专业的 DBA，也是一位勤奋的写作者，他的上一本书《MySQL 王者晋级之路》才出版不久，我又收到了《MongoDB 运维实战》这本书的手稿，在繁忙的工作之余，还有这样的勤奋和毅力，真是令人钦佩！希望张甦先生能够写出更多更优秀的技术书籍，帮助更多即将走上 DBA 之路的伙伴。

侯圣文，Oracle ACE 总监、教育专家，恩墨学院院长

自序

张甦自序

我出生在北京，小时候的梦想是成为一名足球运动员，可以驰骋于赛场，为自己喜欢的北京国安队效力。从来没有想过会从事数据库相关的工作，更不会想到今后自己会写书。继《MySQL 王者晋级之路》之后，《MongoDB 运维实战》是我写的第二本书。好朋友跟我开玩笑地说，这可能就是一个最美丽的错误。今后可能会出版更多与技术相关的书，就让这个美丽的错误一直延续下去！

在自己近十年的技术生涯中，需要感谢的人真的太多了。并不是说因为现在做出点小成绩，或是因为出书了，就要开始写感谢的话了。在我状态处于最低谷、最迷茫的阶段，是我的那些贵人、前辈和挚友把一直在黑暗中行走的我拉了出来，他们的帮助就好比是一束光，指引着我看清未来的方向。在这条道路上所经历的各种辛酸，只有你们最能理解我！

能写完这本《MongoDB 运维实战》，首先要感谢我的好兄弟贺磊，我们一拍即合。经历了无数个日日夜夜的编写、修改、再编写，目的就是把多年累积下来的工作经验梳理成完整的知识体系分享给大家，让大家在工作中少走弯路，快速达到自己的预期目标。

我们可能永远成为不了那些大腕、明星，我们有的只是一颗本本分分做人、踏踏实实做事、研究技术的心。最后说一句：雷霆雨露，俱是天恩，感谢老天赐予我们的各种艰辛和磨难。正因为经历了这些，才让我们更加珍惜现在来之不易的幸福，使我们更好地去努力奋斗，争取让自己的家庭更好，让我们的父母和孩子为我们所付出的努力而感到骄傲和自豪（送给所有努力工作的兄弟姐妹们）。

望广大读者多提宝贵意见。更希望大家可以花些时间，认真品味 MongoDB 给我们带来的简单快乐。

贺磊自序

在北京这个大城市，时光飞逝，有 4 个人对我有至关重要的影响，我想在这里对他们表示感谢。

首先感谢我的好友张甦先生，是他邀请我一起撰写这本《MongoDB 运维实战》，如果没有张甦先生，可能这些内容会一直躺在我的个人笔记里，或者零零散散地写在我的博客上，并没有成体系。

我一直以来都崇尚分享，我相信分享能够提升自己，因此一有时间，我就毫无保留地在博客上撰写一些实战案例。但零散的博客和写书完全不同，写书不是博客随笔，要有严谨的思路和错误校验，如果没有张甦先生，我可能没有机会将这些知识以书面的形式分享给大家，在他的帮助下，才得此机遇，圆了出书的梦。

其次要感谢我的爱人李爱璇女士，无论在生活中，还是工作中，她对我都是无条件地支持，让我以饱满的精神状态面对工作。多少个下班后的夜晚，是她默默在背后支持我写书，让我非常感动，真心感谢她。

还要感谢卓汝林先生和潘友飞先生，在我入职小米以来，是他们带着我学习、工作，让我的技能水平有了巨大的进步。

MongoDB 的最新版本已经到了 4.0 版本，而目前市面上的书大多数还停留在 2.6 版本和 3.0 版本。MongoDB 在每个新版本里加入了诸多的新特性，本书结合笔者职业生涯中的众多案例，为您一一列出前因后果和处理办法，也希望对得起“实战”二字。在实战中分析、在案例中学习是本书的要义所在，也希望读者看过此书后能有所收获，这是对笔者最大的慰藉。

前言

随着大数据时代的到来及技术的不断发展，以及互联网 Web 2.0 的兴起，传统的关系型数据库在应付超大规模和高并发的 SNS 类型的 Web 2.0 纯动态网站时已经显得力不从心了，暴露了很多难以解决的问题，而非关系型数据库则由于其本身的特点得到了非常迅速的发展。NoSQL 领域首屈一指的就是“芒果”数据库，即大名鼎鼎的 MongoDB。我是从第一个 GA 版本开始接触 MongoDB 的，它在我最孤独、最寂寞的时候，陪伴着我一路成长。倘若我一直在黑暗中行走，那么 MongoDB 就是那一束光，指引着我未来前进的方向。

写此书的目的

我把 MySQL 和 MongoDB 当作自己的两个“孩子”一样看待，一直想把多年运维数据库的经验分享出来，前不久已经出版了一本 MySQL 的著作《MySQL 王者晋级之路》，收到的读者反馈都不错，读者说书很实用，看完之后收获很大。其实写书的真正目的就是为了让大家可以系统地进行学习，少走一些作者在工作中走过的弯路。我的一些学生和朋友经常和我抱怨：“我们公司有一个项目准备用 MongoDB，但还得从头学习，网上针对 MongoDB 的资料也不多，关键还不知道从何学起，MongoDB 实战的书籍也偏少”。这类问题不止一个人和我说过。正因为有了这样的需求，我们才有了想投入全部精力、认真地去写一本有关 MongoDB 实战方面图书的冲动。希望《MongoDB 运维实战》能够真正地帮助大家解决在学习 MongoDB 数据库过程中的诸多疑惑，敲开大数据运维的门。

如何阅读本书

第 1、2 章主要介绍 MongoDB 3.4 和 MongoDB 3.6 这两个版本的新特性，以复制集架构和分片架构作为整体切入点。MongoDB 版本更新到 3.X 之后发生了巨变，进入了一个新的时代。在引入 wiredtiger 存储引擎之后，实现了文档级别的锁，提高了并发性。该引擎支持压缩，节约了存储成本，具有更简单高效的高可用架构，维护起来更加轻松。这让我们对 MongoDB 4.X 时代更加期待。

第 3 章是本书中一道亮丽的风景线，是 MongoDB 的实战案例分析部分。详细介绍 oplog 大小引发的从库宕机、副本集延迟突然增大到上万秒、最大连接数限制等问题的处理过程和思路。

第 4 章是 MongoDB 的性能调优部分。从索引角度出发，通过各类索引的使用，包括配合执行计划的查看来梳理性能问题。

第 5 章介绍 MongoDB 备份与恢复，主要以逻辑备份和物理备份两种方式来进行讲解，其中会演练 oplog replay 的过程。

第 6 章是高可用架构集群管理。核心的两个部分就是复制集和分片架构，包括副本集、分片架构原理、成员类型、实战安装部署过程，以及多种实现方式和管理维护架构中遇到的诸多问题，最后还会介绍升级 MongoDB 架构版本的注意事项。

第 7 章介绍 MongoDB 的监控。主要介绍 PMM，它是一款能够监控 MySQL、MongoDB 性能的开源平台。本章会讲解 server 组件和 client 组件及其安装过程。

第 8 章是 MongoDB 的常用命令。本章列举一些我们在 MongoDB 的运维过程中常用的命令，帮助刚接触 NoSQL 领域的读者快速上手生产环境中 MongoDB 的运维工作，为自己的公司和老板排忧解难。

致谢

在《MySQL 王者晋级之路》中已经说了太多感谢的话。这次先要感谢我的好兄弟贺磊，我们一拍即合，经历了多个日夜，坚持了一年的时间完成了这本 MongoDB 的图书。其次，还是要感谢电子工业出版社编辑陈晓猛先生的耐心指导与对我的支持。

这是我出版的第二本数据库相关的书籍了。第一本是关系型数据库方向的图书，本书是 NoSQL 方向的，未来计划出版一本人工智能方向的图书。道行尚浅，还需努力，希望广大读者多提宝贵意见。更希望大家可以花些时间，认真品味 MongoDB 给运维带来的简单快乐。

张甦

目录

第 1 章 MongoDB 3.4 新特性

我们在学习掌握一项技术的时候，总会在某个时间段遇到学习瓶颈，遇到瓶颈之后，会有一种失去激情不想学习的感觉，研究技术就会存在这样的问题。一成不变的技术知识，时间长了，我们都会觉得枯燥无味、索然无趣。这时就需要新鲜的血液进入我们的神经中枢，激发我们的求知欲。所以第 1、2 章就是给大家“打鸡血”，详解 MongoDB 3.4 和 MongoDB 3.6 的新特性，让大家感受前所未有的新鲜感，开始我们 MongoDB 的学习之旅。

1.1 复制集（Replica Set）

Default Journaling Behavior of majority Write Concern

配置复制集时，增加 writeConcernMajorityJournalDefault 选项，默认为 true，即当指定 WriteConcern 为 majority 时，数据写到大多数节点并且 journal 成功刷盘后，才向客户端确认成功；如果为 false，则数据写到大多数节点的内存时就向客户端确认。

writeConcernMajorityJournalDefault 参数的具体的默认值与 protocolVersion 相关，当 protocolVersion =1 时，writeConcernMajorityJournalDefault 默认为 true；当 protocolVersion =0 时，writeConcernMajorityJournalDefault 默认为 false。

protocolVersion 从 3.2 版本起默认为 1，老版本的 MongoDB protocolVersion 默认为 0。protocolVersion 为 0 的复制集成员不能加入 protocolVersion=1 的副本集。

MongoDB 3.6 版本已经弃用了 protocolVersion=0。

已弃用的 protocolVersion =0 在低于 3.6 版本的 MongoDB 中都可以配置 protocolVersion=0。