

MySQL DBA 修炼之道

陈晓勇 著

数据库技术专家撰写，多年数据库领域的经验结晶。实战性强，从架构、调优、运维、开发、测试等多个角度对MySQL管理和维护进行了全方位的归纳和总结，包含大量来自实际生产环境的经典案例，并深入阐述MySQL DBA进阶实战的技巧和方法！



数据库
技术丛书

MySQL DBA 修炼之道

陈晓勇 著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

MySQL DBA 修炼之道 / 陈晓勇著. —北京: 机械工业出版社, 2017.1
(数据库技术丛书)

ISBN 978-7-111-55841-5

I.M… II. 陈… III. 关系数据库系统 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2017) 第 009058 号

MySQL DBA 修炼之道

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 张梦玲

责任校对: 董纪丽

印刷: 北京市荣盛彩色印刷有限公司

版次: 2017 年 2 月第 1 版第 1 次印刷

开本: 186mm × 240mm 1/16

印张: 31.75

书号: ISBN 978-7-111-55841-5

定价: 79.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88379426 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzit@hzbook.com

版权所有 · 侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

华章IT
HZBOOKS | Information Technology



将本书献给三岁的女儿陈观之。

Foreword 推荐序

我之前也看过很多数据库相关的图书，但是没有一本能像这本书一样，让我读起来感觉那么轻松愉快，读完后觉得必须要收藏一本作为案头必备。

本书的作者是互联网一线的数据库开发、运维专家，书中的内容是其对 10 多年工作中所遇问题的思考和总结，围绕着 MySQL 徐徐展开，犹如庖丁解牛，对 MySQL 的核心逻辑解释得相当清晰和透彻。本书以一个数据库专家的视角，解析其观察到的方方面面，内容涉及“业务系统设计”“测试体系”“运维管理”等。本书的很多内容已经不仅仅是从一个 DBA 的角度出发，更多的是从一个系统架构师和运维管理者的角度来思考问题。读完全书，你将会对整个研发、运维体系的相关领域都有一个概要的认识。这种提纲挈领的架构，对于某个知识领域的学习是非常有价值的。

之所以说本书读起来令人轻松愉悦，是因为书中提及的很多问题都是我所关心的，而笔者均以很简练的语言给予了回答和梳理，让人理解起来非常清晰、不费劲。我边看边忍不住想，这风格分明就是 UC 内部的培训资料嘛，有很强烈的亲切感。

本书对实战中的很多问题，都给出了详细的解题思路，方案成熟、观点中肯，体现了对技术应有的严谨和敬畏，我相信对从事 DBA 工作的很多技术人员来说，本书具有非常重要的参考价值。

毫不夸张地说，MySQL 开源项目推动了整个互联网产品的发展。我们从中获益不少，同时也深刻体会到自由分享精神对社会进步的贡献。从晓勇写的这本书中，我也能感受到这一分享理念。我非常赞赏这种分享精神，也希望更多的技术人员都能有此回报社会的情怀。

20 年前，互联网刚刚起步，工程师是靠掌握一批指令和娴熟的操作来执行运维工作的。现如今，开发和运维体系已经渐趋成熟，不少企业更是将基础运维工作交给云服务厂商，研

发和运维人员得以从烦琐的技术细节中解放出来，从而更专注于业务分析和产品设计，这个进步是巨大的。

往后看，我们正从 IT 时代过渡到 DT 时代。在 DT 企业中，工程师使用贝叶斯变换和机器学习来操作数据，就好比当初使用“if()... else()”来编写程序一样，巨大的技术变革正在来临。

在这风起云涌之际，技术让我们再次感受到年轻和无知，希望我们能从 MySQL 出发，保持旺盛的好奇心和探索精神，迈向下一个崭新的时代。

梁捷 (Jack)

UC 联合创始人，神马搜索总裁

2016 年 12 月

为什么要写本书

本书主要讲述 MySQL DBA 的必备技能，包括 MySQL 的安装部署、开发、测试、监控和运维，此外，读者还可从中学习到系统架构的一些知识。

我从业 10 多年，先是在传统行业做开发工程师、系统管理员、Oracle DBA，2008 年因为机缘巧合投身互联网，开始从事 MySQL 运维工作。相对于成熟的商业数据库，MySQL 缺乏高质量的技术文档和图书，我在接触 MySQL 的过程中，也感觉市面上的相关图书还存在一些不足，难以系统化地学习 MySQL。

从一名 Oracle DBA 转型为一名 MySQL DBA，从传统领域转投到互联网公司，即便我之前有丰富的经验，在学习 MySQL 的过程中也仍然走了一些弯路。成为一名 MySQL DBA 并不难，但成为一名高水平的 MySQL DBA 则需要时间、知识、技能、经验和意识的积累。

我在学习 MySQL 的过程中，有时会去看技术论坛，或者通过 MSN 群等聊天工具咨询他人一些问题，也得到过一些朋友的帮助。国内存在一批高素质的 MySQL DBA，但由于各种现实因素，有心写一本关于 MySQL DBA 实战的书的人很少，所以市面上缺乏高质量的相关图书不足为奇。2013 年年初，华章公司的策划编辑杨绣国找到我，说希望我能写一本关于 MySQL 的书，我当时很犹豫，虽然我有时会在网上回答一些问题，也定期撰写个人博客，但是，写一本书，对于我来说，是一个艰巨的任务。经过一些交流，我慢慢明确了自己的想法，其实我一直是想写一本书的，既然我对市面上的相关图书不太满意，那么就自己写一本吧，当时我唯一欠缺的是写作经验以及时间。

我写这本书的目的是想做一个尝试——引领感兴趣的读者进入 MySQL 数据库运维领域。

国内互联网行业正在高速发展，迫切需要大量的 MySQL 人才，希望这本书可以帮助一些读者顺利进入数据库领域。而且，我也想将自己的一些心得分享给读者，希望热爱数据库技术的同行们在工作中少走弯路。

在技术领域工作多年后，文字写作对于我来说其实已经很陌生了，弗朗西斯·培根说过，“阅读使人充实，谈论使人机敏，写作使人精确”。在本书的写作的过程中，其实我自己也获得了许多，不仅学到了更多的知识，对于自己的精神也是一种洗礼。写作真的是一种积极而富有价值的创作，我们只有正确地掌握所讲述的内容，才能为言行思想带来正能量。

希望在这个世界上，有越来越多的人愿意分享，且能享受分享的乐趣。

读者对象

本书的主要读者是 MySQL DBA，在现实中，许多公司并没有配备专职的数据库维护人员，数据库的维护工作往往也是由开发工程师和系统管理员负责的，因此这本书也适用于他们。

这是一本偏向实战的技术书籍，不会过多地涉及技术的细节和原理，我会尽量直接地给出解决方案；本书除了讲 MySQL 技能，还花了大量篇幅讲述架构；本书不仅讲述技术，也讲述技术之外的一些运维管理规则。对数据库的使用、维护和管理感兴趣的运维工程师、架构师、运维经理、开发工程师、测试工程师都可以将本书作为参考图书，而了解其他领域会有助于你的职业发展。

本书也适合希望转行到数据库运维领域的人士。许多人想从事 IT 工作，但当下时间宝贵，要想进入一个行业或改变职业方向，往往会花费巨大的时间成本，所以这本书将尽量做到简单、易懂，以节省大家的学习成本。

如何阅读本书

本书将分为 5 个部分，分别从入门、开发、测试、运维、性能与架构这几个方面来介绍 MySQL 的使用。对于初次接触 MySQL 的读者，建议按照章节顺序逐步学习。对于已经有一定经验的读者，则可以选择自己感兴趣的篇章，跳过自己已经熟悉的内容。

第一部分讲述了 MySQL 的基础架构、权限机制、常用的存储引擎、复制架构、安装及常用命令等知识。如果读者是初次接触 MySQL，那么可能还需要在这一部分上花一些时间。

在掌握 Linux 和 MySQL 的基本使用方法之后，就可以开始第二部分的学习了。

第二部分将介绍 MySQL 数据库开发相关的基础知识和技巧。基础知识包括关系数据模型、字符集、常用的 SQL 语法、范式、索引和事务等。由于开发的领域很广，所以本部分仅仅选取了一些常用的技巧分享给大家。最后会结合实际生产，提供一份开发规范供大家参考。

第三部分介绍了数据库基准测试所需要的理论知识和常用的测试工具。本部分将介绍一个 MySQL 的基准测试模型。

第四部分介绍了 MySQL 运维工作的各项职责：监控、复制、迁移、升级、备份和恢复，然后通过一些案例向读者传授一些维护技巧及处理问题的方法。读者还将学习到规模化运维 MySQL 的一些知识和规则。

第五部分介绍了性能调优的一些理论知识，以及从应用程序到数据库，再到存储等各个环节的优化。由于架构和性能优化密切相关，本部分也介绍了一些 MySQL DBA 需要熟悉的架构优化知识。初次接触 MySQL 的读者对于架构优化的内容可能会感到难以理解，但随着经验的增长，再理解这些内容将不会再有问题。

本书假设读者已经对软硬件有了一定的认识，掌握了一门脚本语言，并且对 Unix 或 Linux 有一定的使用经验，对于数据库有了基本的认识。阅读本书时，读者不需要预先准备好上述的所有知识，但需要有意识地在阅读本书之外不断地补充自己的基础知识。我会对以上内容做深入的讲解，但如果读者有基础会更好，好的基础有利于快速吸收知识和深入思考问题。如果读者还不会使用 Linux 和编写 Shell 脚本，那么，建议尽快搭建一个学习环境。

由于 DBA 需要和研发、测试、产品、运营、监控等团队进行合作，所以对于相关领域所涉及的数据库知识，本书也会做一些介绍。但是，由于经验侧重的关系，本书将主要从 DBA 的角度来讲述这些知识和技能。

本书主要基于 MySQL 官方 5.1 版本写作，这也是目前最流行的版本，我会补充 MySQL 最新版本的少许内容，但跟踪 MySQL 新版本更合适的策略是关注官方发布的新特性说明、新版本的文档手册，跟踪业内专家的技术博客和社交媒体等。

通过阅读本书，读者可以学到 MySQL 的许多知识，但是仅通过阅读是难以获得技能和经验的。读者需要有一个适合自己的 MySQL 测试环境，并能够不断地思考和实践自己的想法，这样才能够掌握技能，并得到属于自己的经验。

勘误和支持

由于作者的水平有限，写作时间也很仓促，书中难免存在一些错误或不准确的地方，如有不妥之处，恳请读者批评指正。为此，我特意创建了在线支持页面 <http://www.db110.com/>。你可以将书中的错误发布在勘误表页面，若遇到任何问题，也可以访问 Q&A 页面，我将尽量在线上为你提供最满意的解答。书中的全部源文件都将发布在这个网站上。如果你有更多的宝贵意见，也欢迎你发送邮件至我的邮箱 ucgary@gmail.com，很期待听到你们的真挚反馈。

致谢

感谢机械工业出版社华章公司的策划编辑杨绣国的努力工作，没有她的投入和耐心，就不可能有本书的面世。本书写作的时间较长，我有时会充满愧疚，是杨绣国编辑的包容和鼓励，最终引导我顺利完成全部书稿。

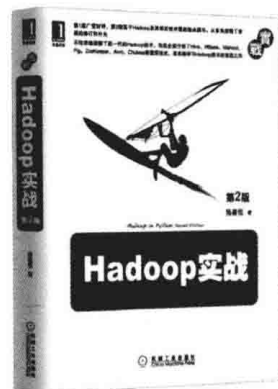
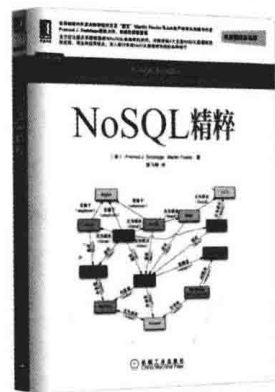
感谢 UC 的旧同事，和你们的共事，是我职业生涯最宝贵的财富，我将一直铭记在心。

最后，我要感谢我的家人和朋友，是你们的支持，让我能够坚持下来。

陈晓勇 (Gary Chen)

中国，长沙，2016 年 12 月

推荐阅读



推荐序
前 言

第一部分 入门篇

第 1 章 理解 MySQL 2

1.1 MySQL 介绍 2

1.1.1 应用领域和适用场景 2

1.1.2 为什么那么多公司和 机构选择使用 MySQL 3

1.1.3 MySQL 的优势是什么, 它解决了什么问题 3

1.2 MySQL 的基础架构和版本 4

1.2.1 软件架构中数据库的定位 4

1.2.2 MySQL 的基础架构 5

1.2.3 MySQL 的版本及特性 6

1.2.4 MySQL 的开发周期 8

1.3 查询执行过程概述 9

1.4 MySQL 权限 10

1.4.1 MySQL 权限机制 10

1.4.2 赋予权限和回收权限 10

1.5 长连接、短连接、连接池 10

1.5.1 短连接 10

1.5.2 长连接 11

1.5.3 连接池 12

1.5.4 持久连接和连接池的区别 12

1.6 存储引擎简介 12

1.6.1 InnoDB 引擎 12

1.6.2 MyISAM 引擎 13

1.6.3 MEMORY 存储引擎 13

1.6.4 ARCHIVE 存储引擎 14

1.6.5 选择合适的引擎 14

1.6.6 选择何种平台 14

1.7 MySQL 复制架构 15

1.8 一些基础概念 15

第 2 章 MySQL 安装部署和入门 17

2.1 如何选择 MySQL 版本 17

2.2 官方版本的安装 18

2.2.1 二进制包的安装 18

2.2.2	源码编译安装	20
2.3	其他 MySQL 分支的安装	21
2.4	安装 InnoDB Plugin	21
2.5	常用命令	22
2.5.1	使用 mysql 命令	22
2.5.2	使用 mysqladmin 命令	26
2.5.3	使用 mysqldump 命令	26
2.6	MySQL 的主要参数设置	27

第二部分 开发篇

第 3 章 开发基础

3.1	相关基础概念	30
3.2	数据模型	33
3.2.1	关系数据模型介绍	33
3.2.2	实体-关系建模	34
3.2.3	其他数据模型	35
3.3	SQL 基础	38
3.3.1	变量	39
3.3.2	保留字	40
3.3.3	MySQL 注释	42
3.3.4	数据类型	43
3.3.5	函数	51
3.3.6	操作符及优先级	53
3.3.7	MySQL 示例	
	employees 数据库	54
3.3.8	SQL 语法	55
3.4	PHP 开发	69
3.4.1	概述	69
3.4.2	客户端访问过程	69
3.4.3	开发工具	72

3.4.4	操作数据	73
3.4.5	PHP 数据库开发建议	78
3.5	索引	78
3.5.1	索引介绍	78
3.5.2	使用索引的场景 及注意事项	79
3.5.3	索引的错误用法	82
3.5.4	如何使用 EXPLAIN 工具	83
3.5.5	优化索引的方法学	90
3.6	ID 主键	91
3.7	字符集和国际化支持	93
3.7.1	什么是字符集	93
3.7.2	国际化支持	93
3.7.3	字符集设置	94

第 4 章 开发进阶

4.1	范式和反范式	98
4.1.1	范式	98
4.1.2	反范式	102
4.2	权限机制和安全	104
4.2.1	MySQL 访问权限系统	104
4.2.2	强化安全	106
4.2.3	SQL 注入	109
4.3	慢查询日志	112
4.3.1	查看慢查询日志	112
4.3.2	使用工具分析慢查询日志	114
4.4	应用程序性能管理	118
4.4.1	为什么需要性能管理	118
4.4.2	应用性能管理概述	119
4.4.3	应用性能管理的关注点	120
4.4.4	具体应用	120

4.5	数据库设计	122	5.3	处理重复值	165
4.5.1	逻辑设计	122	5.4	分页算法	167
4.5.2	物理设计	123	5.5	处理 NULL 值	167
4.6	导入导出数据	127	5.6	存储 URL 地址	169
4.6.1	规则简介	127	5.7	归档历史数据	169
4.6.2	使用 mysqldump 导出, 使用 mysql 导入	128	5.8	使用数据库存储图片	170
4.6.3	使用 SELECT INTO OUTFILE 命令导出数据	130	5.9	多表 UPDATE	170
4.6.4	使用 LOAD DATA 导入数据	130	5.10	生成全局唯一 ID	171
4.6.5	用 mysqlimport 工具导入	132	5.11	使用 SQL 生成升级 SQL	171
4.6.6	用 mysql 程序的 批处理模式导出	132	第 6 章 查询优化		172
4.6.7	用 split 切割文件, 加速导入数据	133	6.1	基础知识	172
4.7	事务和锁	133	6.1.1	查询优化的常用策略	172
4.7.1	概述	133	6.1.2	优化器介绍	173
4.7.2	MyISAM 的表锁	134	6.1.3	MySQL 的连接机制	175
4.7.3	事务定义和隔离级别	135	6.2	各种语句优化	176
4.7.4	InnoDB 的行锁	137	6.2.1	连接的优化	176
4.8	死锁	141	6.2.2	GROUP BY、DISTINCT、 ORDER BY 语句优化	177
4.9	其他特性	142	6.2.3	优化子查询	178
4.9.1	临时表	142	6.2.4	优化 limit 子句	179
4.9.2	分区表	142	6.2.5	优化 IN 列表	180
4.9.3	存储过程、触发器、外键	147	6.2.6	优化 UNION	181
4.9.4	视图	159	6.2.7	优化带有 BLOB、 TEXT 类型字段的查询	181
第 5 章 开发技巧		161	6.2.8	filesort 的优化	181
5.1	存储树形数据	161	6.2.9	优化 SQL_CALC_ FOUND_ROWS	183
5.2	转换字符集	163	6.2.10	优化临时表	183
			6.3	OLAP 业务优化	184

第7章 研发规范	187
7.1 命名约定	187
7.2 索引	188
7.3 表设计	189
7.4 SQL 语句	190
7.5 SQL 脚本	191
7.6 数据架构的建议	192
7.7 开发环境、测试环境的 配置参数建议	193
7.8 数据规划表	193
7.9 其他规范	194

第三部分 测试篇

第8章 测试基础	196
8.1 基础概念	196
8.2 性能测试的目的	197
8.3 基准测试	197
8.4 性能 / 基准测试的步骤	199
8.5 测试的注意事项	199
第9章 测试实践	201
9.1 硬件测试	201
9.1.1 概述	201
9.1.2 CPU 测试	202
9.1.3 内存测试	203
9.1.4 I/O 测试	203
9.1.5 网络测试	207
9.2 MySQL 测试	207
9.2.1 概述	207

9.2.2 常用测试工具的 介绍和使用	208
9.2.3 MySQL 基准测试模型	211
9.3 应用数据库性能测试	221

第四部分 运维篇

第10章 基础知识	224
10.1 文件和 I/O 管理	224
10.1.1 MySQL 日志文件	224
10.1.2 InnoDB 数据文件和 日志文件	228
10.1.3 临时文件	232
10.1.4 MySQL 套接字文件	232
10.2 MySQL 如何进行灾难恢复	233
10.3 变量设置、配置文件和 主要参数	234
10.3.1 概述	234
10.3.2 如何设置参数、变量	235
10.3.3 配置文件的读取顺序	235
10.3.4 环境变量、配置文件、 命令行选项的优先级	237
10.3.5 配置文件详述	237
10.3.6 配置文件示例	238
10.4 MySQL Query Cache 和 优化器	239
10.5 SHOW INNODB STATUS 解析	241
第11章 MySQL 的监控	250
11.1 非数据库的监控	250

11.1.1	开源监控工具 / 平台	250	12.5	跨 IDC 复制	297
11.1.2	编写程序来收集信息	251	12.6	多主复制	298
11.2	数据库的监控	251	12.7	延时复制	298
11.2.1	数据库服务的 基本监控方式	251	12.8	半同步复制	299
11.2.2	应该收集的信息和 收集方法	252	12.9	在线搭建从库	299
11.2.3	MySQL 需要关注的 参数及状态变量	260	12.9.1	操作系统下对 打包文件配置主从	299
11.3	数据库监控的实现	263	12.9.2	利用 mysqldump 制作从库	301
11.3.1	Nagios	263	12.10	配置日志服务器	303
11.3.2	swatch	263	12.11	常见的复制问题及处理方法	305
11.3.3	Cacti	267	12.11.1	跳过复制错误	305
11.3.4	如何打造一个强大的 监控系统	274	12.11.2	临时表和复制	305
11.4	数据库监控的可视化	275	12.11.3	内存表和复制	306
11.4.1	折线图	276	12.11.4	主库宕机重新启动成功, 但复制关系中断	306
11.4.2	散点图	276	12.11.5	主库宕机重启不成功	307
11.4.3	热图	277	12.11.6	多个从库的 server-id 相同	307
第 12 章	MySQL 复制	279	12.11.7	锁定导致的复制延时	307
12.1	基础知识	279	12.11.8	对 MyISAM 引擎的表恢复数据	307
12.1.1	原理及注意事项	279	12.11.9	如何彻底清除 Slave 设置	308
12.1.2	常用命令	281	12.11.10	网络异常导致的 复制延时	308
12.1.3	参数设置	287	第 13 章	迁移、升级、备份、 恢复数据库	309
12.1.4	配置文件	290	13.1	升级	309
12.1.5	复制模式	292	13.1.1	升级表结构或变更数据	309
12.1.6	复制兼容性	294			
12.2	配置主从复制	294			
12.3	配置主主复制	296			
12.4	配置级联复制、环形复制	297			