HZ BOOKS
华章科技

数据库
技术丛书

Redis 设计与实现

黄健宏 著

The Design and Implementation of Redis

- 由资深 Redis 技术专家撰写，深入了解 Redis 技术内幕的必读之作。
- 从源码角度解析 Redis 的架构设计、实现原理和工作机制，为高效使用 Redis 提供原理性指导。

机械工业出版社
China Machine Press

数据库
技术丛书

Redis
设计与实现

The Design and Implementation of Redis

黄健宏 著

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

Redis设计与实现/黄健宏著. —北京：机械工业出版社，2014.4
（数据库技术丛书）

ISBN 978-7-111-46474-7

I. R… II. 黄… III. 数据库—基本知识 IV. TP311.13

中国版本图书馆CIP数据核字（2014）第079820号

本书全面而完整地讲解了 Redis 的内部机制与实现方式，对 Redis 的大多数单机功能以及所有多机功能的实现原理进行了介绍，展示了这些功能的核心数据结构以及关键的算法思想，图示丰富，描述清晰，并给出大量参考信息。通过阅读本书，读者可以快速、有效地了解 Redis 的内部构造以及运作机制，更好、更高效地使用 Redis。

本书主要分为四大部分。第一部分“数据结构与对象”介绍了 Redis 中的各种对象及其数据结构，并说明这些数据结构如何影响对象的功能和性能。第二部分“单机数据库的实现”对 Redis 实现单机数据库的方法进行了介绍，包括数据库、RDB 持久化、AOF 持久化、事件等。第三部分“多机数据库的实现”对 Redis 的 Sentinel、复制、集群三个多机功能进行了介绍。第四部分“独立功能的实现”对 Redis 中各个相对独立的功能模块进行了介绍，涉及发布与订阅、事务、Lua 脚本、排序、二进制位数组、慢查询日志、监视器等。本书作者专门维护了 www.redisbook.com 网站，提供带有详细注释的 Redis 源代码，以及本书相关的更新内容。

Redis设计与实现

黄健宏 著

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：吴 怡 责任校对：殷 虹
印 刷：藁城市京瑞印刷有限公司 版 次：2014年6月第1版第1次印刷
开 本：186mm×240mm 1/16 印 张：25.25
书 号：ISBN 978-7-111-46474-7 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）88378991 88361066 投稿热线：（010）88379604
购书热线：（010）68326294 88379649 68995259 读者信箱：hzjsj@hzbook.com

前　言

时间回到2011年4月，当时我正在编写一个用户关系模块，这个模块需要实现一个“共同关注”功能，用于计算出两个用户关注了哪些相同的用户。

举个例子，假设huangz关注了peter、tom、jack三个用户，而john关注了peter、tom、bob、david四个用户，那么当huangz访问john的页面时，共同关注功能就会计算并打印出类似“你跟john都关注了peter和tom”这样的信息。

从集合计算的角度来看，共同关注功能本质上就是计算两个用户关注集合的交集，因为交集这个概念是如此的常见，所以我很自然地认为共同关注这个功能可以很容易地实现，但现实却给了我当头一棒：我所使用的关系数据库并不直接支持交集计算操作，要计算两个集合的交集，除了需要对两个数据表执行合并（join）操作之外，还需要对合并的结果执行去重复（distinct）操作，最终导致交集操作的实现变得异常复杂。

是否存在直接支持集合操作的数据库呢？带着这个疑问，我在搜索引擎上面进行查找，并最终发现了Redis。在我看来，Redis正是我想要找的那种数据库——它内置了集合数据类型，并支持对集合执行交集、并集、差集等集合计算操作，其中的交集计算操作可以直接用于实现我想要的共同关注功能。

得益于Redis本身的简单性，以及Redis手册的详尽和完善，我很快学会了怎样使用Redis的集合数据类型，并用它重新实现了整个用户关系模块：重写之后的关系模块不仅代码量更少，速度更快，更重要的是，之前需要使用一段甚至一大段SQL查询才能实现的功能，现在只需要调用一两个Redis命令就能够实现了，整个模块的可读性得到了极大的提高。

自此之后，我开始在越来越多的项目里面使用Redis，与此同时，我对Redis的内部实现也越来越感兴趣，一些问题开始频繁地出现在我的脑海中，比如：

- Redis的五种数据类型分别是由什么数据结构实现的？
- Redis的字符串数据类型既可以存储字符串（比如"hello world"），又可以存储整数和浮点数（比如10086和3.14），甚至是二进制位（使用*SETBIT*等命令），Redis在内部是怎样存储这些值的？
- Redis的一部分命令只能对特定数据类型执行（比如*APPEND*只能对字符串执行，*HSET*

只能对哈希表执行），而另一部分命令却可以对所有数据类型执行（比如*DEL*、*TYPE*和*EXPIRE*），不同的命令在执行时是如何进行类型检查的？Redis在内部是否实现了一个类型系统？

- Redis的数据库是怎样存储各种不同数据类型的键值对的？数据库里面的过期键又是怎样实现自动删除的？
- 除了数据库之外，Redis还拥有发布与订阅、脚本、事务等特性，这些特性又是如何实现的？
- Redis使用什么模型或者模式来处理客户端的命令请求？一条命令请求从发送到返回需要经过什么步骤？

为了找到这些问题的答案，我再次在搜索引擎上面进行查找，可惜的是这次搜索并没有多少收获：Redis还是一个非常年轻的软件，对它的最好介绍就是官方网站上面的文档，但是这些文档主要关注的是怎样使用Redis，而不是介绍Redis的内部实现。另外，网上虽然有一些博客文章对Redis的内部实现进行了介绍，但这些文章要么不齐全（只介绍了Redis中的少数几个特性），要么就写得过于简单（只是一些概述性的文章），要么关注的就是旧版本（比如2.0、2.2或者2.4，而当时的最新版已经是2.6了）。

综合来看，详细而且完整地介绍Redis内部实现的资料，无论是外文还是中文都不存在。意识到这一点之后，我决定自己动手注释Redis的源代码，从中寻找问题的答案，并通过写博客的方式与其他Redis用户分享我的发现。在积累了七八篇Redis源代码注释文章之后，我想如果能将这些博文汇集成书的话，那一定会非常有趣，并且我自己也会从中学到很多知识。于是我在2012年年末开始创作《Redis设计与实现》，并最终于2013年3月8日在互联网发布了本书的第一版。

尽管《Redis设计与实现》第一版顺利发布了，但在我的心目中，这个第一版还是有很多不完善的地方：

- 比如说，因为第一版是我边注释Redis源代码边写的，如果有足够时间让我先完整地注释一遍Redis的源代码，然后再进行写作的话，那么书本在内容方面应该会更为全面。
- 又比如说，第一版只介绍了Redis的内部机制和单机特性，但并没有介绍Redis多机特性，而我认为只有将关于多机特性的介绍也包含进来，这本《Redis设计与实现》才算是真正的完成了。

就在我考虑应该何时编写新版来修复这些缺陷的时候，机械工业出版社的吴怡编辑来信询问我是否有兴趣正式地出版《Redis设计与实现》，能够正式地出版自己写的书一直是我梦寐以求的事情，我找不到任何拒绝这一邀请的理由，就这样，在《Redis设计与实现》第一版发布几天之后，新版《Redis设计与实现》的写作也马不停蹄地开始了。

从2013年3月到2014年1月这11个月间，我重新注释了Redis在unstable分支的源代码（也即是现在的Redis 3.0源代码），重写了《Redis设计与实现》第一版已有的所有章节，并向书中添加了关于二进制位操作（bitop）、排序、复制、Sentinel和集群等主题的新章节，最终完成了这本新版的《Redis 设计与实现》。本书不仅介绍了Redis的内部机制（比如数据库实现、类型系统、事件模型），而且还介绍了大部分Redis单机特性（比如事务、持久化、Lua脚本、排序、二进制位操

作），以及所有Redis多机特性（如复制、Sentinel和集群）。

虽然作者创作本书的初衷只是为了满足自己的好奇心，但了解Redis内部实现的好处并不仅仅在于满足好奇心：通过了解Redis的内部实现，理解每一个特性和命令背后的运作机制，可以帮助我们更高效地使用Redis，避开那些可能会引起性能问题的陷阱。我衷心希望这本新版《Redis设计与实现》能够帮助读者更好地了解Redis，并成为更优秀的Redis使用者。

本书的第一版获得了很多热心读者的反馈，这本新版的很多改进也来源于读者们的意见和建议，因此我将继续在www.RedisBook.com设置disqus论坛（可以不注册直接发贴），欢迎读者随时就这本新版《Redis设计与实现》发表提问、意见、建议、批评、勘误，等等，我会努力地采纳大家的意见，争取在将来写出更好的《Redis设计与实现》，以此来回报大家对本书的支持。

黄健宏（huangz）

2014年3月于清远

致　　谢

我要感谢hoterran 和iammutex 这两位良师益友，他们对我的帮助和支持贯穿整本书从概念萌芽到正式出版的整个阶段，也感谢他们抽出宝贵的时间为本书审稿。

我要感谢吴怡编辑鼓励我创作并出版这本新版《Redis 设计与实现》，以及她在写作过程中对我的悉心指导。

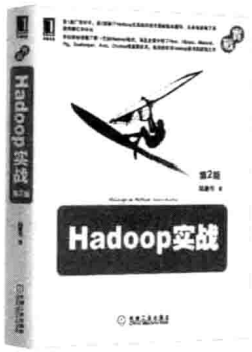
我要感谢TimYang 在百忙之中抽空为本书审稿，并耐心地给出了详细的意见。

我要感谢Redis 之父Salvatore Sanfilippo ，如果不是他创造了Redis 的话，这本书也不会出现了。

我要感谢所有阅读了《Redis 设计与实现》第一版的读者，他们的意见和建议帮助我更好地完成这本新版《Redis 设计与实现》。

最后，我要感谢我的家人和朋友，他们的关怀和鼓励使得本书得以顺利完成。

推荐阅读


Hadoop实战
第2版
Hadoop实战

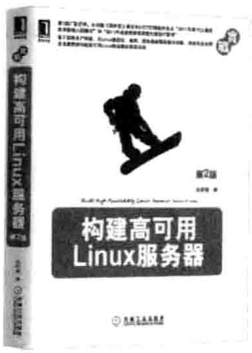

Hadoop技术内幕
深入解析Hadoop Common和
HDFS架构设计与实现原理

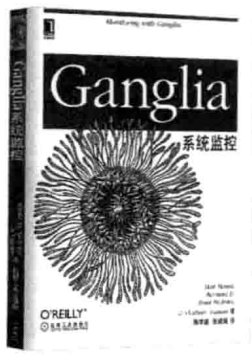

Hadoop技术内幕
深入解析MapReduce架构设计与实现原理


Hadoop技术内幕
深入解析YARN架构设计与实现原理

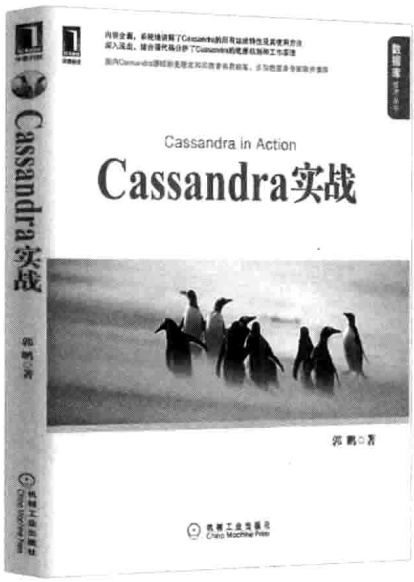

Linux软件管理平台
设计与实现

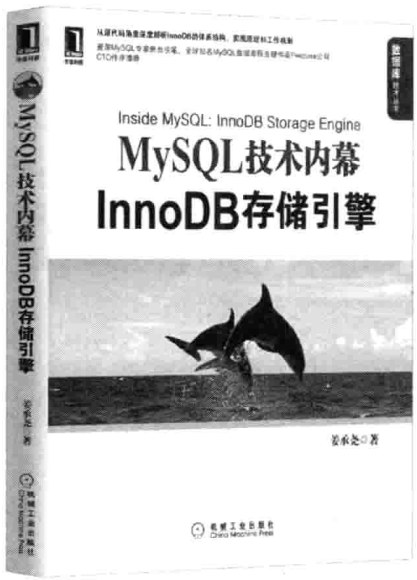

高性能Linux服务器构建实战
运维监控、性能调优与集群应用


构建高可用
Linux服务器


Ganglia
系统监控
O'REILLY

推荐阅读





Cassandra实战

作者：郭鹏 ISBN：978-7-111-34164-2 定价：59.00元

内容全面，系统地讲解了Cassandra的所有功能特性及其使用方法
深入浅出，结合源代码分析了Cassandra的底层机制和工作原理

MySQL技术内幕：InnoDB存储引擎

作者：姜承尧 ISBN：978-7-111-32188-0 定价：69.00元

从源代码角度深度解析InnoDB的体系结构、
实现原理和工作机制资深MySQL专家亲自执笔，
全球知名MySQL数据库服务提供商Percona公司CTO作序推荐

目　录

第二部分 单机数据库的实现

第 1 章
引　言

本书对 Redis 的大多数单机功能以及所有多机功能的实现原理进行了介绍，力图展示这些功能的核心数据结构以及关键的算法思想。

通过阅读本书，读者可以快速、有效地了解 Redis 的内部构造以及运作机制，这些知识可以帮助读者更好地、也更高效地使用 Redis。

为了让本书的内容保持简单并且容易读懂，本书会尽量以高层次的角度来对 Redis 的实现原理进行描述，如果读者只是对 Redis 的实现原理感兴趣，但并不想研究 Redis 的源代码，那么阅读本书就足够了。

另一方面，如果读者打算深入了解 Redis 实现原理的底层细节，本书在 RedisBook.com 提供了一份带有详细注释的 Redis 源代码，读者可以先阅读本书对某一功能的介绍，然后再阅读该功能对应的实现代码，这有助于读者更快地读懂实现代码，也有助于读者更深入地了解该功能的实现原理。

1.1　Redis 版本说明

本书是基于 Redis 2.9——也即是 Redis 3.0 的开发版来编写的，因为 Redis 3.0 的更新主要与 Redis 的多机功能有关，而 Redis 3.0 的单机功能则与 Redis 2.6、Redis 2.8 的单机功能基本相同，所以本书的内容对于使用 Redis 2.6 至 Redis 3.0 的读者来说应该都是有用的。

另外，因为 Redis 通常都是渐进地增加新功能，并且很少会大幅地修改已有的功能，所以本书的大部分内容对于 Redis 3.0 之后的几个版本来说，应该也是有用的。

1.2　章节编排

本书由“数据结构与对象”、“单机数据库的实现”、“多机数据库的实现”、“独立功能的实现”四个部分组成。