

<<七周七数据库>>

图书基本信息

书名：<<七周七数据库>>

13位ISBN编号：9787115312242

10位ISBN编号：7115312249

出版时间：2013-5-30

出版时间：人民邮电出版社

作者：[美]Eric Redmond

译者：王海鹏,田思源,王晨

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## <<七周七数据库>>

### 内容概要

如今，我们要面对和使用的数据正在变得越来越庞大和复杂。

如果说数据是新的石油。

那么数据库就是油田、炼油厂、钻井和油泵。

作为一名现代的软件开发者，我们需要了解数据管理的新领域，既包括rdbms，也包括nosql。

《七周七数据库》遵循《七周七语言》的写作风格和体例，带领你学习和了解当令最热门的开源数据库。

在简单的介绍之后，本书分章介绍了7种数据库。

这些数据库分别属于5种不同的数据库风格，但每种数据库都有自己保存数据和看待世界的方式。

它们依次是postgresql、riak、apache hbase、mongodb、apache couchdb、neo4j和redis。

本书将深入每一种数据库，介绍它们的优势和不足，以及如何选取一种最符合你的应用需求的数据库。

《七周七数据库》适合数据库架构师、数据库管理员，以及想要了解和学习各种nosql数据库技术的程序员阅读。

本书将帮助读者了解、选择和应用这些数据库，从而更好地发挥日益增长的大数据的能力。

《七周七数据库》适用于数据库架构师、想要学习nosql数据库技术的程序员、数据库管理员。

## 作者简介

作者：（美国）Eric Redmond（美国）Jim R.Wilson 译者：王海鹏 田思源 王晨Eric Redmond，在软件行业有超过15年的经验，曾在财富500强公司、政府和许多创业公司工作过。

他是一个程序员、国际演讲者，也是几个技术团体的积极组织者。

Jim R.Wilson，从13岁就开始编程。

他曾在医疗、搜索和市场营销行业的公司中担任过工程师和Web专家。

他从2007年开始思考NoSQL数据库，并为大规模的开源项目（如MediaWiki和HBase）贡献了代码。

他和妻子Ruth及两个小孩住在马萨诸塞州的利特尔顿，经常在本地的JavaScript和NoSQL会议上演讲。

## 书籍目录

第1章 概述 1 1.1 从一个问题开始 1 1.2 风格 2 1.2.1 关系数据库 3 1.2.2 键—值数据库 3 1.2.3 列型数据库 4 1.2.4 文档型数据库 5 1.2.5 图数据库 5 1.2.6 混合使用多种数据库 6 1.3 前进和提升 6 第2章 PostgreSQL 7 2.1 这就是Post—greS—Q—L 7 2.2 第1天：关系、CRUD和联接 8 2.2.1 从SQL开始 10 2.2.2 使用表 10 2.2.3 使用联接的查询 13 2.2.4 外联接 15 2.2.5 使用索引快速查找 16 2.2.6 第1天总结 18 2.2.7 第1天作业 18 2.3 第2天：高级查询、代码和规则 19 2.3.1 聚合函数 19 2.3.2 分组 20 2.3.3 窗口函数 21 2.3.4 事务 22 2.3.5 存储过程 24 2.3.6 触发器 26 2.3.7 视图 27 2.3.8 规则是什么 28 2.3.9 联表分析 30 2.3.10 第2天总结 32 2.4 第3天：全文检索和多维查询 32 2.4.1 模糊搜索 34 2.4.2 SQL标准的字符串匹配 34 2.4.3 字符串相似比较算法 levenshtein 35 2.4.4 三连词 36 2.4.5 全文检索 37 2.4.6 组合使用字符串匹配方法 42 2.4.7 把电影风格表示成多维超立方体 42 2.4.8 第3天总结 45 2.5 总结 46 2.5.1 PostgreSQL的优点 46 2.5.2 PostgreSQL的缺点 47 2.5.3 结束语 47 第3章 Riak 48 3.1 Riak喜欢Web 48 3.2 第1天：CRUD、链接和MIME 49 3.2.1 REST是最棒的（或用cURL时） 51 3.2.2 将值放于桶中 52 3.2.3 链接 54 3.2.4 Riak的MIME类型 58 3.2.5 第1天总结 58 3.3 第2天：Mapreduce和服务器集群 59 3.3.1 填充脚本 59 3.3.2 mapreduce介绍 60 3.3.3 Riak中的mapreduce 63 3.3.4 关于一致性和持久性 69 3.3.5 第2天总结 75 3.4 第3天：解决冲突和扩展Riak 76 3.4.1 以向量时钟解决冲突 76 3.4.2 扩展Riak 83 3.4.3 第3天总结 87 3.5 总结 87 3.5.1 Riak的优点 88 3.5.2 Riak的缺点 88 3.5.3 Riak之于CAP 88 3.5.4 结束语 89 第4章 HBase 90 4.1 介绍HBase 91 4.2 第1天：CRUD和表管理 91 4.2.1 配置HBase 92 4.2.2 HBase的shell 93 4.2.3 创建表 93 4.2.4 插入、更新和读取数据 95 4.2.5 修改表 96 4.2.6 通过编程方式添加数据 99 4.2.7 第1天总结 100 4.3 第2天：处理大数据 101 4.3.1 导入数据，调用脚本 101 4.3.2 流式XML 102 4.3.3 流式维基百科 103 4.3.4 压缩和Bloom过滤器 105 4.3.5 开始 106 4.3.6 区域和监控磁盘使用简介 107 4.3.7 区域的问讯 108 4.3.8 扫描一个表来建立另一个表 111 4.3.9 构建扫描程序 112 4.3.10 运行脚本 114 4.3.11 检查输出 114 4.3.12 第2天总结 116 4.4 第3天：放入云端 117 4.4.1 开发Thrift协议的HBase应用 117 4.4.2 Whirr简介 121 4.4.3 设置EC2 121 4.4.4 准备Whirr 122 4.4.5 配置集群 122 4.4.6 启动集群 123 4.4.7 连接集群 124 4.4.8 销毁集群 125 4.4.9 第3天总结 125 4.5 总结 126 4.5.1 HBase的优点 126 4.5.2 HBase的缺点 127 4.5.3 HBase on CAP 127 4.5.4 结束语 128 第5章 MongoDB 129 5.1 其大无比 129 5.2 第1天：CRUD和嵌套 130 5.2.1 命令行的乐趣 131 5.2.2 Mongo的更多有趣内容 134 5.2.3 深入挖掘 136 5.2.4 更新 140 5.2.5 引用 142 5.2.6 删除 143 5.2.7 用代码来读取 144 5.2.8 第1天总结 145 5.3 第2天：索引、分组和mapreduce 146 5.3.1 索引：如果还不够快 146 5.3.2 聚合查询 150 5.3.3 服务器端命令 152 5.3.4 mapreduce（以及Finalize） 155 5.3.5 第2天总结 159 5.4 第3天：副本集、分片、地理空间和GridFS 159 5.4.1 副本集 159 5.4.2 偶数节点的问题 162 5.4.3 分片 163 5.4.4 地理空间查询 165 5.4.5 GridFS 167 5.4.6 第3天总结 168 5.5 总结 168 5.5.1 Mongo的优点 168 5.5.2 Mongo的缺点 169 5.5.3 结束语 169 第6章 CouchDB 170 6.1 在沙发上放松 170 6.2 第1天：CRUD、Futon与cURL Redux 171 6.2.1 享受Futon 171 6.2.2 用cURL执行基于REST的CRUD操作 175 6.2.3 用GET读取文档 175 6.2.4 用POST新建文档 176 6.2.5 用PUT更新文档 177 6.2.6 用DELETE移除文档 178 6.2.7 第1天总结 178 6.3 第2天：创建 / 查询视图 179 6.3.1 通过视图访问文档 179 6.3.2 编写你的第一个视图 181 6.3.3 将视图另存为“设计文档” 183 6.3.4 由Name查找Artists 184 6.3.5 由name查找albums 184 6.3.6 查询自定义的Artist与Album视图 185 6.3.7 使用Ruby将数据导入CouchDB 187 6.3.8 第2天总结 192 6.4 第3天：进阶视图、ChangesAPI以及复制数据 192 6.4.1 用规约器创建进阶视图 193 6.4.2 规约器调用详解 195 6.4.3 监控CouchDB的变化 196 6.4.4 连续监控变化 201 6.4.5 在CouchDB中复制数据 204 6.4.6 第3天总结 208 6.5 总结 209 6.5.1 CouchDB的优点 209 6.5.2 CouchDB的缺点 209 6.5.3 结束语 209 第7章 Neo4j 211 7.1 Neo4j，白板友好的数据库 211 7.2 第1天：图、Groovy和CRUD 212 7.2.1 Neo4j之Web接口 213 7.2.2 通过Gremlin操作Neo4j 215 7.2.3 pipe的威力 218 7.2.4 Pipeline与顶点 221 7.2.5 无模式的社会性数据 221 7.2.6 垫脚石 223 7.2.7 引入Groovy 225 7.2.8 特定领域的步骤 227 7.2.9 更新、删除与完成 229 7.2.10 第1天总结 230 7.3 第2天：REST、索引与算法 230 7.3.1 引入REST 230 7.3.2 用REST新建节点与关系 231 7.3.3 查找路径 232 7.3.4 索引 233 7.3.5 REST与Gremlin 235 7.3.6 大数据 235 7.3.7 功能全面的算法 237 7.3.8 第2天总结 242 7.4 第3天：分布式高可用性 243 7.4.1 事务 243 7.4.2 高可用性 244 7.4.3 HA集群 245 7.4.4 备份 250 7.4.5 第3天总结 251 7.5 总结 251 7.5.1 Neo4j的优点 251 7.5.2 Neo4j的缺点 252 7.5.3 Neo4j之于CAP 252 7.5.4 结束语 252 第8章 Redis 254 8.1 数据结构服务器存储库 254 8.2 第1天：CRUD与数据类型 255

<<七周七数据库>>

8.2.1 入门指南 255 8.2.2 事务 257 8.2.3 复杂数据类型 258 8.2.4 到期 265 8.2.5 数据库命名空间 266 8.2.6 更多命令 267 8.3 第2天：高级用法、分布 268 8.3.1 一个简单的接口 268 8.3.2 服务器信息 271 8.3.3 Redis配置 272 8.3.4 AOF ( appendonlyfile ) 274 8.3.5 主从复制 276 8.3.6 数据转储 277 8.3.7 Redis集群 279 8.3.8 Bloom过滤器 280 8.3.9 SETBIT和GETBIT 282 8.3.10 第2天总结 283 8.4 第3天：与其他数据库合作 284 8.4.1 多持久并存服务 284 8.4.2 数据填充 285 8.4.3 关系存储 291 8.4.4 服务 293 8.4.5 第3天总结 296 8.5 总结 297 8.5.1 Redis的优点 297 8.5.2 Redis的缺点 297 8.5.3 结束语 298 第9章 结束语 299 9.1 类型终极版 299 9.1.1 关系型 299 9.1.2 键—值存储库 300 9.1.3 列型 301 9.1.4 文档型 301 9.1.5 图 302 9.2 选择 303 9.3 我们将走向哪里 303 附录A 数据库概述表 305 附录B CAP定理 307

## 章节摘录

版权页：插图：Redis就像润滑油。

润滑油通常用于润滑系统的各个运转部件，通过减少摩擦，保持它们运行顺畅，并加快其整体功能。不论哪种系统构造，加点润滑油很可能有改善。

有时候，只要明智地使用Redis，就能搞定你的问题。

Redis (REmote Dictionary Service，远程字典服务)最早发布于2009年，是一个简单易用的键值对存储库，带有一套成熟的命令。

若论速度，很多数据库难出其右。

它读取速度快，写入速度更快，根据某些基准测试，每秒可处理高达10万次SET操作。

Redis的创始人是萨尔瓦托·勒圣菲利波 (Salvatore Sanfilippo)，他把该项目称为“数据结构服务器”，以反映其对复杂数据类型及其他功能的细致处理。

这不只是一个超快的键值对存储系统，学习它，将使我们对现代数据库的了解更加完整。

8.1数据结构服务器存储库 将Redis精确归类可能有点困难。

当然，从基本层面上说，它是一个键值对存储库。

但这种简单的说法并不全面。

虽然Redis没有达到文档型数据库的程度，但它支持高级的数据结构。

它支持基于集合的查询操作，但不支持关系数据库中同样的粒度或类型。

当然，它很快，为了速度而在持久性方面作出了让步。

Redis是高级数据结构服务器，此外，它也是阻塞队列（或栈）和发布-订阅系统。

它支持可配置的到期策略、持久性级别，以及复制选项。

所有这些使得Redis不仅是某类数据库中的一员，更是有用的数据结构算法和程序的工具包。

Redis有丰富的客户端库，在许多编程语言中可以很方便地使用它。

它不仅易用，还是一种乐趣。

如果说API是程序员的用户体验，那么在现代艺术博物馆中Redis应该和Mac·Cube放在一起。

在第1天和第2天，我们将探讨Redis的功能、约定和配置。

与往常一样，从简单的CRUD操作开始，我们会很快转向更高级的操作，涉及更强大的数据结构：列表、哈希表、集合和有序集合。

我们将创建事务，并操作数据有效期的特征。

我们会用Redis创建一个简单的消息队列，并探讨其发布-订阅功能。

然后，我们将深入探讨Redis的配置和复制选项，学习如何在数据持久性和速度之间，取得适合应用程序的平衡。

数据库常常彼此配合使用，这种趋势在增加。

把Redis安排在本书的最后介绍，这样我们就可以用这种方式使用它。

在第3天，我们将构建我们的终极系统，一个功能丰富的多数据库的音乐解决方案，包括Redis、CouchDB、Neo4j和Postgres，用Node.js将它们结合在一起。

## <<七周七数据库>>

### 媒体关注与评论

本书针对现代数据库给出了很好的结构性概述，而且它并没有过于深入。

它也不应该过于深入，因为它给出了你根据自己需求选择数据库所需的所有知识。

—— Help Net Security如果你需要使用或考虑使用任何一种与传统关系型数据库有所不同的数据库技术，而且你不知道应该如何从为数众多的选项中做出选择，那么，本书将帮助你了解这一领域，并且帮助你根据当前的需求更好地做出选择。

郑重推荐这本书！

—— Matthew Helmke.net阅读这本书，就像去迪斯尼乐园的荒野地带。

有转弯，有曲折，你永远不知道下一个拐角是什么，但是非常有趣。

—— Jeffrey Newman如下的故事是完美的。

周五，你会安装并运行一种新的数据库。

周六，你会看到在日常使用中它是什么样子。

到了周日，你已经学到了一些技巧，甚至可能让专家感到惊讶！

而下一周，你会跳到另一个数据库，并且又是另一番乐趣。

—— Ian Dees 《Using JRuby》作者这本书针对几种关键数据库给出了很好的概览。

它将大幅度地提升你对数据建模的选择能力和实践技能。

如果你希望别人对你在数据库方面的能力抱以七倍的羡慕，那么，请阅读这本书。

—— Sean Copenhaver, backgroundchecks.com 首席程序员本书是迄今为止对现代数据库最好的实质性概述。

它不同于我读过的大量教程、博客文章和文档，它教给我：为什么我要使用每种类型的数据库，以及应该以何种方式使用它们，从而更容易理解和保留信息。

阅读这本书是一种乐趣。

—— Loren Sands-Ramshaw, 美国国防部软件工程师这是我所见到的介绍CouchDB最好的书籍之一。

—— Jan Lehnardt Apache CouchDB的开发者和作者如果更理想的话，在这本书的封面上可以印上：“包含你想知道的关于数据库的一切知识，而且你离不开这些知识”。

说实话，这本书会卖的很好。

—— Nic Williams博士, Engine Yard的技术副总裁本书很好地介绍了现代数据库设计和实施的各个方面。

无论你是新手还是专家，即便是只花一天的时间阅读一章，你也会拓展技能，受益匪浅。

—— Jerry Sievert, Daily Insight Group工程总监

## <<七周七数据库>>

### 编辑推荐

《七周七数据库》编辑推荐：试想，周五你安装并运行了一个新的数据库。

周六，你会看到在日常使用中它是什么样子。

到了周日，你已经学到了一些技巧，甚至可能让专家们感到惊讶！

而下一周，你又会跳到另一个数据库。

Ian Dees告诉我们这样的效率是有可能的《七周七数据库》言简意赅、直入主题，深入每一种数据库，介绍它们的优势和不足，以及如何选取一种最符合你的应用需求的数据库，探究当前流行和具有代表性的七种数据库技术。

你准备好了吗？



## <<七周七数据库>>

### 名人推荐

《七周七数据库》针对现代数据库给出了很好的结构性概述，而且它并没有过于深入。它也不应该过于深入，因为它给出了你根据自己需求选择数据库所需的所有知识。

——Help Net Security如果你需要使用或考虑使用任何一种与传统关系型数据库有所不同的数据库技术，而且你不知道应该如何从为数众多的选项中做出选择，那么，《七周七数据库》将帮助你了解这一领域，并且帮助你根据当前的需求更好地做出选择。

郑重推荐这本书！

——Matthew Helmke.net阅读《七周七数据库》，就像去迪斯尼乐园的荒野地带。有转弯，有曲折，你永远不知道下一个拐角是什么，但是非常有趣。

——Jeffrey Newman如下的故事是完美的。

周五，你会安装并运行一种新的数据库。

周六，你会看到在日常使用中它是什么样子。

到了周日，你已经学到了一些技巧，甚至可能让专家感到惊讶！

而下一周，你会跳到另一个数据库，并且又是另一番乐趣。

——Ian Dees 《Using JRuby》作者《七周七数据库》针对几种关键数据库给出了很好的概览。它将大幅度地提升你对数据建模的选择能力和实践技能。

如果你希望别人对你在数据库方面的能力抱以七倍的羡慕，那么，请阅读这本书。

——Sean Copenhaver, backgroundchecks网 首席程序员《七周七数据库》是迄今为止对现代数据库最好的实质性概述。

它不同于我读过的大量教程、博客文章和文档，它教给我：为什么我要使用每种类型的数据库，以及应该以何种方式使用它们，从而更容易理解和保留信息。

阅读这本书是一种乐趣。

——Loren Sands-Ramshaw, 美国国防部软件工程师《七周七数据库》是我所见到的介绍CouchDB最好的书籍之一。

——Jan Lehnardt Apache CouchDB的开发者和作者如果更理想的话，在《七周七数据库》的封面上可以印上：“包含你想知道的关于数据库的一切知识，而且你离不开这些知识”。

说实话，这本书会卖的很好。

——Nic Williams博士, Engine Yard的技术副总裁《七周七数据库》很好地介绍了现代数据库设计和实施的各个方面。

无论你是新手还是专家，即便是只花一天的时间阅读一章，你也会拓展技能，受益匪浅。

——Jerry Sievert, Daily Insight Group工程总监

#### 版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>