

华章程序员书库

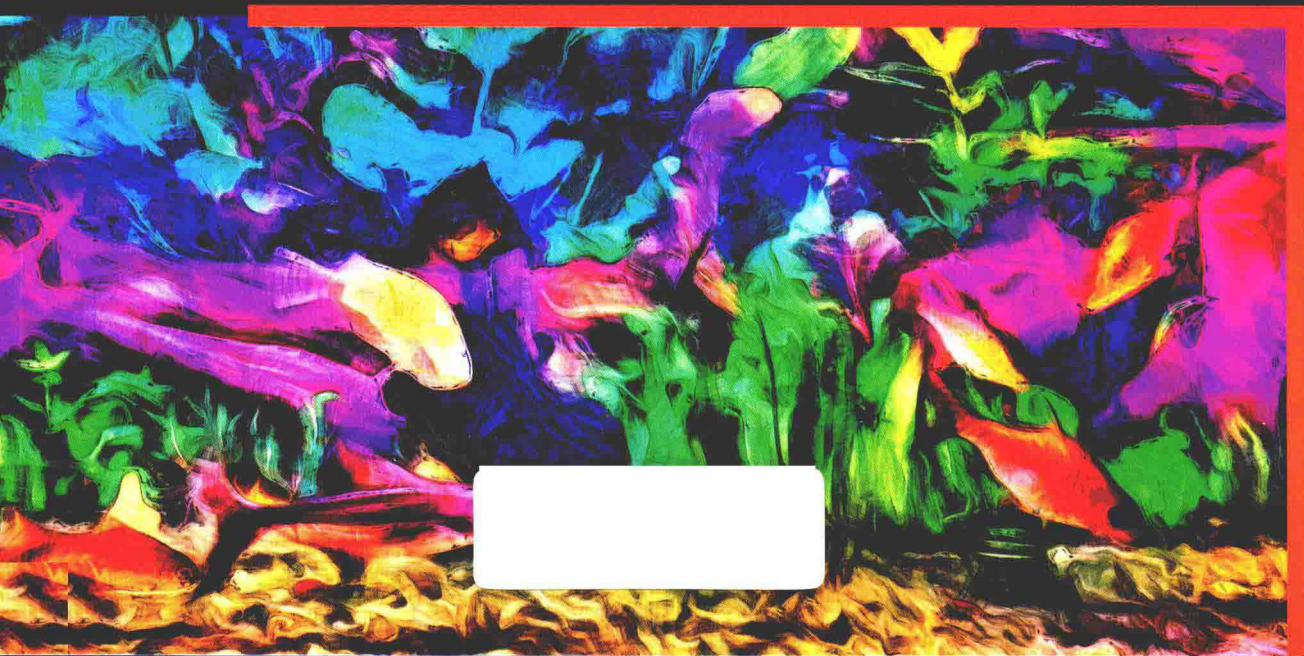
PEARSON

HZ BOOKS
华章IT

NoSQL实践指南

基本原则、设计准则及实用技巧

NoSQL for Mere Mortals



[美] 丹·苏利文 (Dan Sullivan) 著

爱飞翔 译



机械工业出版社
China Machine Press

华章程序员书库



NoSQL for Mere Mortals

NoSQL实践指南

基本原则、设计准则及实用技巧

[美] 丹·苏利文 (Dan Sullivan) 著

爱飞翔 译



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

NoSQL 实践指南：基本原则、设计准则及实用技巧 / (美) 苏利文 (Sullivan, D.) 著；爱飞翔译. —北京：机械工业出版社，2016.3

(华章程序员书库)

书名原文：NoSQL for Mere Mortals

ISBN 978-7-111-53117-3

I.N… II. ①苏… ②爱… III. 数据库系统—指南 IV. TP311.138-62

中国版本图书馆 CIP 数据核字 (2016) 第 040639 号

本书版权登记号：图字：01-2015-7580

Authorized translation from the English language edition, entitled NoSQL for Mere Mortals, 9780134023212 by Dan Sullivan, published by Pearson Education, Inc., Copyright © 2015.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

Chinese simplified language edition published by Pearson Education Asia Ltd., and China Machine Press Copyright © 2016.

本书中文简体字版由 Pearson Education (培生教育出版集团) 授权机械工业出版社在中华人民共和国境内 (不包括中国台湾地区和香港、澳门特别行政区) 独家出版发行。未经出版者书面许可，不得以任何方式抄袭、复制或节录本书中的任何部分。

本书封底贴有 Pearson Education (培生教育出版集团) 激光防伪标签，无标签者不得销售。

NoSQL 实践指南

基本原则、设计准则及实用技巧

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：关 敏

责任校对：董纪丽

印 刷：北京市荣盛彩色印刷有限公司

版 次：2016 年 3 月第 1 版第 1 次印刷

开 本：186mm × 240mm 1/16

印 张：20.75

书 号：ISBN 978-7-111-53117-3

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

近些年来，随着 Web 应用程序、云计算以及大数据的兴起，NoSQL 已经逐渐取得了与关系型数据库相对等的地位。

和计算机领域中的其他技术一样，数据库技术的发展在某些方面也是由需求带动的。当前各种应用程序所使用的数据，以及技术人员和商务人员对这些数据所进行的分析，都比关系型数据库占主导地位的那个时代更加丰富。NoSQL 技术之所以会引发关注，部分原因就在于它有时能够更好地应对这些灵活多变的数据。

在灵活应对数据这个大前提之下，不同的程序又有不同的侧重点。有些程序适合使用轻量级的数据库，有些程序强调对数据的搜索功能，有些程序需要处理海量的信息，还有一些程序则关注数据之间的联系。于是，NoSQL 数据库也形成了四大门类。尽管它们之间的界限并不十分严格，但还是有必要了解这四种产品各自的优点和缺点，以便选出最适合当前任务的数据库。怎样根据程序需求和产品特性选出合适的数据库，正是本书所要讲解的核心问题。

此外，作者也反复强调，NoSQL 一词之中的 No 并不是对关系型数据库的简单否定，而是对它的一种传承和创新。关系型数据库的某些特征、查询语法以及设计思路，同样可以移用到 NoSQL 数据库之中。作者在全书的多个章节里面，都对这两大体系进行了对比，这些内容或许能给大家理解 NoSQL 的用法带来一些启发。

讲解数据库技术的《For Mere Mortals》书系一贯以通俗易懂的风格而著称，本书也是如此，凡是对 NoSQL 感兴趣的人都不妨一读。本书前 2 章直接指出了数据库与需求之间的关系，并概括了 NoSQL 与传统数据库的异同，以及 4 种主要的 NoSQL 产品所具备的特征。其后 12 章分为 4 组，每组内的 3 章分别讲解某种 NoSQL 数据库的基础知识、专业术语以及设计技巧。讲解设计技巧时，作者以货运公司所面对的数据管理需求为例，分 4 个案例来演示 4 种 NoSQL 数据库的用法，这 4 种用法合起来就是最后一章所描绘的那幅 NoSQL

数据库全景图。了解 NoSQL 技术的全景之后，读者可以根据本书所提供的参考资料，继续研读其他专著。

因为这是一本通俗读物，所以除了一些难于翻译的术语之外，一般术语都尽量译为中文。同一个术语有 2、3 种译法的，译文会采用较易理解的一种，并在括号内给出其他称呼，译法过于纷繁的请参见书后的术语表。

翻译本书的过程中，我得到了机械工业出版社华章公司诸位编辑和工作人员的帮助，在此深表谢意。

由于译者水平有限，错误与疏漏之处，请大家发邮件至 eastarstormlee@gmail.com，或访问 github.com/jeffreybaoshenlee/zh-translation-errata-nosql/issues 留言，给我以批评和指教。该网页还有中英文词汇对照表以供参考。

爱飞翔

Preface 前言

“人生的进步，凭的是勇气，而不是适应能力。”

——亨利·米勒 (Henry Miller)

在工作中我们难免会谈及数据。每个人都关注个人数据的隐私权，每家公司都竭力防止网络入侵者盗取数据，而政府部门和商业机构则都对数据非常感兴趣。为了描述数据，IT 分析师想出了很多令人眼花缭乱的专业名词，如大数据 (Big Data)、流数据 (streaming data)、高速数据 (high-velocity data) 和非结构化数据 (unstructured data) 等。另外，描述数据存储方式的术语也层出不穷，如数据库 (database)、数据存储 (data store)、数据仓库 (data warehouse)、数据湖泊 (data lake) 等。有人甚至还造出了数据沼泽 (data swamp) 这个词。

当其他人都在热烈讨论数据的时候，有一些人需要去收集数据、处理数据、分析数据并管理数据。本书就是为这部分人而写的。

NoSQL 数据库是为了应对需求而诞生的。原来的那些数据管理工具在过去几十年里一直都能非常好地应对相关的需求，但现在，它们却无法适应新兴的网络应用程序了。目前的数据管理工具所要应对的最大挑战，并不是令成千上万的商务专家能够使用公司内部数据库，而是要使 Google、Amazon、Facebook 及 Yahoo! 这样的公司能够应对百万级别的网络用户。

从理论上来说，关系型数据模型还是非常坚实的，过去它能够满足我们对数据的需求，而现在，我们则需要对其做出一些改进。网络信息采集 (Web crawling) 和在线购物管理 (online shopping cart management) 等专门的网络程序，推动了业界对原有数据模型的改进，同时也催生了一些非关系型数据库，如键值数据库、文档数据库、列族数据库及图数据库等。原有的关系型数据库仍然有其作用，它们并不会为 NoSQL 数据库所取代。这些

NoSQL 数据库只是提供了一些额外的选项，以满足用户对性能和功能的不同需求。

本书是 NoSQL 数据库的入门指南，它旨在使读者了解何时应该使用 NoSQL 数据库，何时不应使用它们。而在本书中，笔者最想告诉大家的则是怎样根据自己的数据管理需求来明智地使用 NoSQL 数据库。

与本书有关的幻灯片、章节习题及教学指导手册，可以从 Pearson 的 Instructor Resource Center (IRC) 获取，网址是 pearsonhighered.com。

Acknowledgements 致 谢

虽然封面上的作者名字只有一个，但本书却是多人协作的成果。感谢 Joan Murray 编辑构思了这本书的出版计划，并邀请我来执笔。有很多令人尊敬的作者及出版业专家都曾参与 For Mere Mortals 书系的写作及发行工作，能与他们同列，笔者深感荣幸。

Tonya Simpson 耐心而专业的工作使得本书由粗略的初稿变为精美的成品。感谢 Pearson 公司的 Sondra Scott、Cindy Teeters 及 Mark Renfrow 在成书过程中对笔者提供的帮助。感谢本书文字编辑 Karen Annett 所付出的诸多努力。

感谢技术编辑 Theodor Richardson，他甚为周到地处理了书中的细节问题。

在写书过程中，家人始终给我以坚定的支持。

岳父 Bill Aiken 是我的头号粉丝，他总是给我源源不断的鼓励。

我的孩子 Nicole、Charles、Kevin 以及他们的伴侣 Katie 和 Sara，也给了我很多支持，感谢你们。

特别要感谢两个儿子 Nicholas 和 James。Nicholas 阅读了每一章，并且像阅读教科书那样完成了各章的习题。他指出了书中的不足，并提出了改进建议，使我能够更好地把相关知识点解释清楚。James 是一位专业的技术写作者，他帮我撰写了与图数据库有关的章节，并且毫不犹豫地自己的日程安排中拿出一些时间帮助我完成本书，使相关章节的质量得以改善。

我在过去 30 年中所取得的专业成就和个人成就，都得益于妻子 Katherine 始终不渝的支持和关爱，此书也不例外。有些事情，表面上是我独自完成的，但实际上却是两人共同的杰作，这一点外人很难知晓，甚至可能根本就不会意识到。这本书可以算是我们人生旅途中的又一件作品吧。

Dan Sullivan

俄勒冈州波特兰市

2015 年

引言 *Introduction*

“每当我觉得自己刚刚领会目前这种生活方式时，人生却又有了新的变化。”

——Hugh Prather

数据库的发展过程和电视一样。曾经有一段时间，我们只能在有限的几个频道之间切换，而那些频道所放映的节目彼此都没有太大区别。但现在已经发生了变化。现在的电视不再像从前那样，只能在几个台之间观看大同小异的节目。与之类似，现在的数据管理系统也不再只是关系型数据库的代名词。

PostgreSQL、MySQL、Oracle、Microsoft SQL Server 和 IBM DB2 等名字，IT 人士都应该非常熟悉，甚至连数据管理领域之外的一些从业者也了解这些产品。在过去几十年中，数据管理人员一直使用关系型数据库来完成工作。它们既能应对包裹追踪及账户余额管理等商业需求，又能满足真菌研究及人类疾病研究等科研需求。关系型数据库可以按照某种逻辑来划分数据，并使得用户能够轻松地获取这些数据。这种数据库的一项重要特性就是，无论数据库内部正在发生多少变化，它都能给各位用户提供协调一致的数据视图。

数据库领域内的很多人都感觉自己似乎已经理解了数据库的用法，可在这个时候，情况却发生了变化，实际上，是因特网发生了变化。因特网的前身是美国军方赞助的 ARPANET 网络，它起初是个学术合作平台，后来逐渐变为商用平台和个人应用平台，网络中的数据量和数据种类也持续增大。除了要管理账户余额，我们还要用计算机来阅读最近的新闻、帮助自己完成家庭作业，并汇总与新电影有关的评论等。很多人不仅要通过计算机与家人取得联系、与同事相互协作，而且还要通过它来接受职业教育并谋求职业发展。

既然与数据管理有关的需求已经发生了巨大的变化，那自然就会诞生许多全新的数据管理方式。最新一代的数据管理工具统称为 NoSQL 数据库。NoSQL 这个名字并没有从正

面描述这些数据管理系统究竟是什么，而是从反面描述了这些系统不是什么。之所以采用这种描述形式，是为了强调它与关系型数据库之间的对比，关系型数据库曾经占据着统治地位，那些数据库使用一种名叫 SQL 的语言。

NoSQL 数据库可以分为四大类：键值数据库、文档数据库、列族数据库和图数据库。（以搜索为导向的数据管理系统，如 Solr 及 Elasticsearch 等，有时也算作 NoSQL 数据库中的一员，但它们不在本书的讨论范围之内。）

键值数据库采用一种简单的模型来存放数据，用户可以根据标识符来查找某条数据，这个标识符叫做键（key），待查询的数据叫做值（value）。20 世纪 90 年代中期发布的 BerkleyDB 是一款早期的键值数据库，对于那种不宜使用关系型数据库的应用程序来说，可以使用这种数据库。

文档数据库对键值数据库的概念做了扩充，它把键值组划分到一种名为文档（document）的逻辑结构之中。文档数据库是性能较高且较为灵活的数据管理系统，越来越多的人用它来完成各类数据管理任务。

列族数据库从表面上看与关系型数据库相似。首次实现列族数据库这一概念的产品是 Google 公司的 BigTable，该产品的名称既暗示了它与关系型数据库之间的联系，又点出了列族数据库的核心数据结构，那就是 table（表）。列族数据库适用于某些数据量较大且要求较为严苛的数据密集型应用程序。

图数据库最适合对网状数据进行建模，也就是对那种相互之间有联系的数据进行建模。计算机之间的通信以及用户之间的交互，都可以采用图数据库来建模。

在数据库系统的研究与开发领域，可能会呈现出两类数据库并行的局面。一方面，我们已经有了较为成熟且适用范围较广的关系型数据库，它能够很好地处理很多数据管理问题，但另一方面，我们还有 NoSQL 数据库可供选择，其中包括早已开始研发的键值数据库，以及最近才设计出来的文档数据库、列族数据库及图数据库等产品。

这种局面会令我们在做技术决策的时候感到更加困难。本书就是想要降低大家在选择数据库时所面临的困难。读过本书之后，你应该对各种 NoSQL 数据库都有相当的了解，并且明白它们分别适用于哪些场合。

大家要知道，NoSQL 数据库变化得非常迅速。阅读本书的时候，你经常使用的 NoSQL 数据库可能已经添加了一些书中没有提到的功能。尤其要注意的是，NoSQL 数据库对事务处理的支持会越来越丰富。不同的数据管理系统会用不同的方式来处理事务，而这正是能够区分各种数据管理系统的一项重要特征。（你可能还不太熟悉事务处理这一概念，没关系，继续读下去，你就会明白它的意思。）

本书面对的读者

凡是对 NoSQL 数据库的使用方式感兴趣的人都适合阅读本书。数据库开发新手、有经验的关系型数据建模者以及 NoSQL 开发老手都可以在书中发现有用的内容。

数据库开发新手可以通过本书开头几章学到数据管理系统的基本原则和设计准则，而且还能了解数据管理的历史。为什么要谈这些历史呢？因为我们都知道，历史总是在不停地重演。

整本书都会在 NoSQL 数据库与关系型数据库之间进行对比。如果你在关系型数据库的设计方面有较多的经验，那么可以通过这些对比来迅速地掌握 NoSQL 数据库的特性，并了解这些特性所具备的价值。

对于已经学会某些 NoSQL 数据库用法的读者来说，本书可以帮你了解其他类型的 NoSQL 数据库。键值数据库和文档数据库运用得非常广泛，但如果你还没有遇到过列族数据库或图数据，那么本书可以帮你熟悉它们。

如果你已经能够顺畅地使用各种 NoSQL 数据库，但是还想了解这些分布式系统的内部机理，那么可以把本书当成研究的起点。通过本书，大家会熟悉 quorum、Bloom 过滤器以及 anti-entropy（反熵）等实现层面的特性。书中谈到这些问题时，会给出一些参考资料以供大家深入研究。

在 NoSQL 数据库的开发文档中能够找到的内容，本书是不会重复的。如果想知道如何向数据库中插入一条数据，那么最好的办法就是直接查阅文档。而如果想获得较为深入的解说，想知道某个数据库的优点和缺点，想了解使用该数据的最佳方式，那么就需要阅读像本书这样的教程了。本书是对数据库文档的一种补充，而不是要完全取代那些文档。

本书的目标

本书是写给对数据感兴趣的读者看的，笔者想要告诉大家，应该如何使用 NoSQL 数据库来解决问题。本书并不要求读者是经验比较丰富的数据库专家，只要你能够熟练地运用 Excel 就可以开始学习书中的内容了。

通过本书，不仅能够了解 NoSQL 数据库，而且还能知道如何运用各种设计原则及实践技巧来应对数据管理方面的需求。本书将会展示 NoSQL 数据库管理系统的内部细节，以便解释分布式数据库系统的工作原理，并告诉大家如何构建易于扩展且较为可靠的应用程序，同时还会指出构建此类程序时所应避免的一些做法。

本书的特色就在于实用。书中的全部内容都是为了告诉大家如何用 NoSQL 数据库来解

决实际的问题。尽管有些地方会提到一些计算机科学理论，但那完全是为了更加详细地解释某些重要的话题。若是对相关的理论比较熟悉，则可以跳过那些文字。

如何阅读本书

对于刚接触数据库系统的读者来说，可以先阅读第 1 章和第 2 章。有了这些准备知识，就可以更好地阅读后续章节了。

如果对关系型数据库及其前身比较熟悉，那么可以跳过第 1 章。如果已经用过 NoSQL 数据库，那么可以跳过第 2 章。不过，由于第 2 章里面完整地讲述了 4 种主要的 NoSQL 数据库类型，所以至少应该略读一下该章中的某些内容，以便了解自己不太熟悉的那几种类型。

每位读者都应该阅读本书的第二部分，因为书中的其他各部分都会引用这一部分的内容。第三、四、五部分可以按照任意顺序来阅读，不过后面的部分会引用早前章节的某些内容。为了更好地理解每一种 NoSQL 数据库，大家应该把第二、三、四、五部分中的每一章都读一遍，这 4 个部分各含有 3 章内容。

第 15 章应该是在熟悉了本书其他章节的内容之后再阅读的。不过，如果你已经对那些内容有了充分的了解，那么可以跳过本章之前与 NoSQL 数据库有关的部分。对于想知道如何在各种 NoSQL 数据库之间做出选择的读者来说，一定要阅读第 15 章。

本书的结构

下面概述每一部分及每一章的内容。

第一部分

NoSQL 数据库并不是凭空产生的。这一部分介绍了与关系型数据库及早前的数据管理系统有关的一些背景知识。

第 1 章 介绍了关系型数据库及其早前的数据管理系统，同时还讨论了当前的数据管理系统所要面对的各项需求。这些需求可以通过 NoSQL 数据库所提供的各种方案来解决。

第 2 章 讲解了数据库的重要功能以及实现分布式数据库时所面临的困难，同时还讨论了不同类型的数据库在相关因素上所做的权衡。该章还简单地介绍了后续章节中的一些研究案例，这些案例描述了各种 NoSQL 数据库的实际应用方式。

第二部分

这一部分讲述了键值数据库的用法以及使用时所应避免的问题。

第 3 章 概述了键值数据库，它是一种最为简单的 NoSQL 数据库。

第 4 章 列出了理解键值数据库的结构和功能所需的专业词汇。

第 5 章 讲解了键值数据库的设计原则、局限性以及这种数据库所用的设计模式。该章还包含研究案例，以描述键值数据库的实际用法。

第三部分

这一部分讲解了使用范围较为广泛的文档数据库，并给出了一些建议，告诉大家如何利用文档数据库来有效地实现相关的应用程序。

第 6 章 描述了文档数据库的基本特征，并介绍了无模式数据库这一概念，同时还讲解了文档数据库的一些基本操作。

第 7 章 使读者能够熟悉文档数据库所用到的一些专业词汇。

第 8 章 深入讨论了规范化 (normalization) 和去规范化 (denormalization) 的优势、针对变化的文档所做的规划、编制索引的技巧，以及常见的设计模式。该章最后有个研究案例，描述了怎样用文档数据库来实现一款商务程序。

第四部分

这一部分介绍了大数据应用程序，并讲述了列族数据库所能应对的需求。

第 9 章 描述了 Google 的 BigTable 数据库所用的设计方式，对比了键值数据库、文档数据库以及列族数据库之间的差别，讲解了列族数据库的架构。

第 10 章 介绍了列族数据库的专用词汇。如果你一直不太明白 anti-entropy (反熵) 的意思，那就应该看看这一章。

第 11 章 介绍了如何设计表格、如何编制索引，以及如何为数据分区，同时还讲解了怎样应对与大数据有关的需求。

第五部分

这一部分讲解了图数据库，并给出了一些特别适合用图数据库来解决的用例。

第 12 章 介绍了图与网络的建模，并讲解了图数据库的优势。

第 13 章 介绍了图论 (graph theory) 中的一些专用词汇。图论是数学的一个分支，它为图数据库提供了理论支持。

第 14 章 讲解了图数据库的设计技巧、设计时所应避免的陷阱，以及图数据库的查询方式。该章最后给出了一个案例，告诉大家如何用图数据库来解决商务问题。

第六部分

这一部分告诉大家如何运用前面所学的内容。

第 15 章 根据前面各章所讲的内容提出了一些建议，告诉大家在为应用程序选择数据库时，需要考虑哪些因素。

第七部分

附录 A 列出了每章末尾的复习题，以及这些习题的答案。

附录 B 列出了许多（但不是全部的）NoSQL 数据库，其中很多数据库都是开源或免费使用的。

术语表中列出了本书所提到的 NoSQL 术语，以及这些术语的定义。

目 录 *Contents*

译者序	
前 言	
致 谢	
引 言	

第一部分 入 门

第1章 用不同的数据库应对不同的需求	2
1.1 关系型数据库的设计	3
1.2 早期的数据管理系统	4
1.2.1 平面文件数据管理系统	4
1.2.2 分层数据模型系统	7
1.2.3 网络数据管理系统	9
1.2.4 早期数据管理系统小结	11
1.3 引发变革的关系型数据库	11
1.4 催生 NoSQL 数据库（非单纯的 SQL 数据库 / 非 SQL 数据库）的动因	18
1.4.1 可伸缩性	18
1.4.2 成本开销	19
1.4.3 灵活性	19
1.4.4 可用性	20
1.5 小结	20

1.6	案例研究	21
1.7	复习题	21
1.8	引用资料	22
1.9	参考书目	22
第2章	NoSQL数据库的种类	23
2.1	分布式数据库的数据管理	24
2.1.1	持久地存储数据	25
2.1.2	维护数据的一致性	25
2.1.3	确保数据的可用性	26
2.1.4	在响应时间、一致性与持久性之间寻求平衡	29
2.1.5	CAP 定理：一致性、可用性及分区保护性	31
2.2	ACID 与 BASE	32
2.2.1	ACID：原子性、一致性、隔离性、持久性	33
2.2.2	BASE：基本可用、软状态、最终一致	33
2.2.3	最终一致性的类型	34
2.3	四种 NoSQL 数据库类型	35
2.3.1	键值对数据库	36
2.3.2	文档数据库	40
2.3.3	列族数据库	42
2.3.4	图数据库	43
2.4	小结	45
2.5	复习题	46
2.6	引用资料	47
2.7	参考书目	47

第二部分 键值数据库

第3章	键值数据库简介	50
3.1	从数组到键值数据库的演变	51

3.1.1	数组 (带着辅助轮的键值数据存储机制)·····	51
3.1.2	关联数组 (拿掉辅助轮的键值数据存储机制)·····	52
3.1.3	缓存 (加装了齿轮的自行车)·····	53
3.1.4	内存和磁盘中的键值数据库 (从自行车到机动车)·····	55
3.2	键值数据库的重要特性·····	56
3.2.1	简洁: 开发者不需要设计复杂的数据模型·····	56
3.2.2	速度: 越快越好·····	57
3.2.3	易于缩放: 随时应对访问量的变化·····	59
3.3	键: 不只是无意义的标识符·····	63
3.3.1	如何构造键名·····	63
3.3.2	通过键来定位相关的值·····	65
3.4	值: 存放任意数据·····	68
3.4.1	值不一定要有明确的类型·····	68
3.4.2	对值进行搜索时的一些限制·····	69
3.5	小结·····	70
3.6	复习题·····	71
3.7	引用资料·····	71
3.8	参考书目·····	72
第4章	键值数据库的术语 ·····	73
4.1	与键值数据库的数据建模有关的术语·····	74
4.1.1	键·····	75
4.1.2	值·····	77
4.1.3	命名空间·····	77
4.1.4	分区·····	78
4.1.5	分区键·····	80
4.1.6	无纲要的模型·····	80
4.2	与键值数据库的架构有关的术语·····	81
4.2.1	集群·····	81
4.2.2	环·····	82
4.2.3	复制·····	84