HZ BOOKS
华章IT

PEARSON

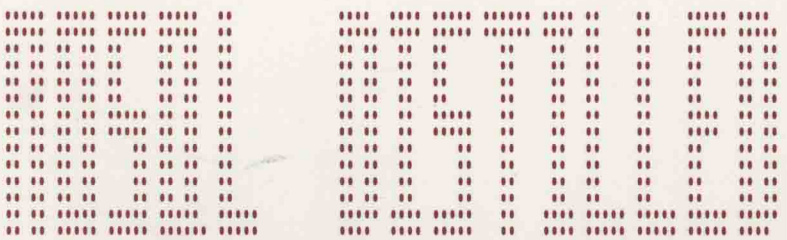
世界级软件开发大师和软件开发“教父”Martin Fowler与Jolt生产效率大奖图书作者
Pramod J. Sadalage最新力作，权威性毋庸置疑

全方位比较关系型数据库与NoSQL数据库的异同，详细讲解4大主流NoSQL数据库的优劣势、用法和适用场合，深入探讨实现NoSQL数据库系统的各种细节

经典原版书库

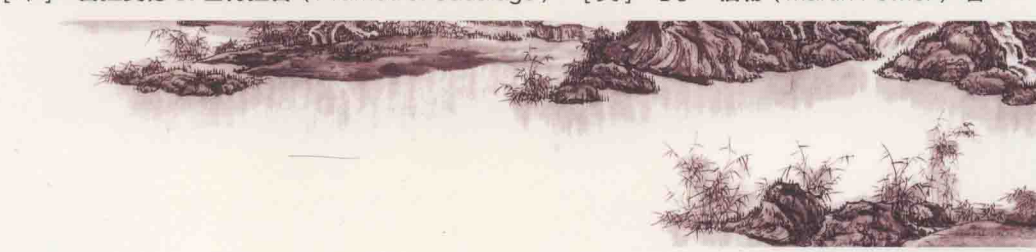
NoSQL精粹

（英文版）



A BRIEF GUIDE TO THE EMERGING WORLD OF POLYGLOT PERSISTENCE

[印] 普拉莫德 J. 塞得拉吉（Pramod J. Sadalage） [美] 马丁·福勒（Martin Fowler）著



机械工业出版社
China Machine Press

经典原版书库

NoSQL精粹

（英文版）

[印] 普拉莫德 J. 塞得拉吉（Pramod J. Sadalage）
[美] 马丁·福勒（Martin Fowler） 著

A BRIEF GUIDE TO THE EMERGING WORLD OF
POLYGLOT PERSISTENCE

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

NoSQL 精粹（英文版）/（印）塞得拉吉（Sadalage, P. J.），（美）福勒（Fowler, M.）著．—北京：机械工业出版社，2015.10
（经典原版书库）
书名原文：NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence

ISBN 978-7-111-51800-6

I. N… II. ① 塞… ② 福… III. 数据库系统－英文 IV. TP311.138

中国版本图书馆 CIP 数据核字（2015）第 243901 号

本书版权登记号：图字：01-2015-5189

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：藁城市京瑞印刷有限公司	版 次：2016 年 1 月第 1 版第 1 次印刷
开 本：170mm × 242mm 1/16	印 张：11
书 号：ISBN 978-7-111-51800-6	定 价：49.00 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

华章IT
HZBOOKS | Information Technology

前　言

我们已经在企业级计算领域研究了20余年，编程语言、架构、平台、软件开发流程等技术都在改变，然而这期间有一件事却一直没变，那就是：大家依然使用关系型数据库来存储数据。虽说也出现了一些挑战关系型数据库的产品，而且有的还在某些领域成功了，但是总体来说，留给架构师的数据存储问题仍然是选择使用哪款关系型数据库的问题。

稳定性在此领域颇受重视。企业的数据比程序存储的时间要长很多（至少大家都是这么说的。当然啦，我们也见过许多非常老的程序）。拥有一个既稳定，又容易理解，而且还能让许多应用程序编程平台访问的数据库，是非常有价值的。

不过，关系型数据库现在碰上新对手了，它的名字叫NoSQL。由于我们需要处理的数据量越来越大，必须以商用服务器集群来构建大型硬件平台，因此NoSQL就应运而生了。这也使大家要再次考虑那个存在已久的难题，即代码如何才能同关系型数据库良好地结合起来。

“NoSQL”这个词的定义是非常不明确的。它泛指那些最近诞生的非关系型数据库，诸如Cassandra、MongoDB、Neo4J和Riak等。它们主张使用无模式（schemaless）的数据，可以运行在集群环境中，并且能够牺牲传统数据库所具备的一致性，以换取另外一些有用的特性。NoSQL的倡导者声称，使用它们可以构建出性能更高、扩展度更好且更易编程的系统。

这会不会敲响了关系型数据库即将灭亡的第一声警钟呢？还是说NoSQL要抢走数据库领域的头把交椅？我们的回答是：“这两种情况都不会出现。”关系型数据库是一个非常强大的工具，我们希望能长时间使用下去；然而大家也要看到一场深远的变革，那就是：关系型数据库不再是唯一的选择了。我们认为，数据库领域正进入混合持久化（Polyglot Persistence）时代，由企业乃至个人研发的应用程序，可以使用多种技术来管理数据。因此架构师需要熟悉这些技术，并且能根据不同的需求做出适当的选择。若非如此，笔者怎会花那么多时间和精力来写这

本书呢？

本书给诸位读者提供足够多的信息，协助大家在以后的研发过程中思考：项目是否真的值得使用 NoSQL 数据库。每个项目都是不同的，我们不可能写出一个简单的决策树，用它来选出合适的数据存储方式。与之相反，本书力求讲解大量的背景知识，以便大家了解 NoSQL 的工作原理，这样的话，你不用在互联网上四处寻找，就能够做出适合自己项目的决定了。笔者刻意将本书写得很短，以便读者能够快速阅览它。虽说本书不会回答各种具体问题，但是，它可以帮你缩小考虑的范围，让你明白自己当前应该提出哪些问题。

NoSQL 数据库为何引人关注

我们来看一下大家选用 NoSQL 数据库的两个主要原因。

- 应用程序的开发效率。在很多应用程序的开发过程中，大量精力和时间都放在了内存（in-memory）数据结构和关系型数据库之间的映射上面。NoSQL 数据库可以提供一种更加符合应用程序需求的数据模型，从而简化了数据交互，减少了所需编写、调试并修改的代码量。
- 大规模的数据。企业所重视的是，数据库要能够快速获取并处理数据。他们发现，即便关系型数据库能达成这一目标，其成本也很高。主要原因在于，关系型数据库是为独立运行的计算机而设计的，但是现在大家通常使用由更小、更廉价的计算机所组成的集群来计算数据，这样更实惠些。许多 NoSQL 数据库正是为集群环境而设计，因此它们更适合大数据量的应用场景。

本书内容

本书分为两个部分。第一部分主要讲述核心概念，让读者能够判断出 NoSQL 数据库是否适合自己，并且了解各种 NoSQL 数据库之间的差别。第二部分更加专注于实现 NoSQL 数据库系统。

第 1 章解释了 NoSQL 发展如此迅速的原因：由于需要处理的数据量越来越多，所以大型系统的扩展方式，由原来在单一计算机上的纵向扩展，转变为在计算机集群上的横向扩展。这也印证了许多 NoSQL 数据库的数据模型所具备的一

个重要特性，那就是：可以把内容密切相关的数据组织成一种丰富的结构，并将其显式存储起来，以便作为一个单元（unit）来访问。本书中，我们将这种类型的结构称为聚合（aggregate）。

第 2 章描述了在 NoSQL 领域的三种主要数据模型中，如何体现“聚合”这一概念。这三种数据库模型是：“键值模型”（key-value，参见 2.2 节），“文档模型”（document，参见 2.2 节）和“列族模型”（column family，参见 2.3 节）。聚合为许多种应用提供了一个自然的交互单元，既改善了集群的运行状况，又使编写程序来访问数据库变得更为容易。第 3 章转到聚合的缺点上面：难以处理位于不同聚合的实体之间的关系（参见 3.1 节）。这自然就引出了图数据库（参见 3.2 节），它是一个不属于面向聚合（aggregated-oriented）阵营的 NoSQL 数据模型。我们也会讲到 NoSQL 数据库的共同特性：它们都是以“无模式”的形式来操作的（参见 3.3 节）。模式的这种特性确实提供了更大的灵活性，但是它并不像大家想象的那么万能。

在讲完 NoSQL 数据模型方面的内容之后，我们接下来要讲分布模型。第 4 章描述了数据库如何在集群中分布数据。这个问题又细分为“分片”（sharding，参见 4.2 节）和“复制”（replication），复制方式可以是“主从复制”（master-slave replication，参见 4.3 节）或者“对等复制”（peer-to-peer replication，参见 4.4 节）。了解完分布模型的概念后，接下来要讲“一致性”（consistency）问题。与关系型数据库相比，NoSQL 数据库在一致性方面提供了更多选择，这么做是因为 NoSQL 要更好地支持集群。于是，第 5 章谈到了更新与读取操作对一致性的影响（分别参见 5.1 节和 5.2 节），如何在一致性与持久性之间进行仲裁（参见 5.5 节），以及如何放宽对持久性的约束以提升其他特性（参见 5.4 节）。如果之前听过 NoSQL，那么就应该听过“CAP 定理”（The CAP Theorem）。5.3 节中介绍了 CAP 定理的相关知识，告诉大家如何根据该理论来权衡一致性与其他特性。

前面这些章节主要侧重于如何分布数据并保持其一致性，接下来的两章讨论了要完成这项工作所需的一些重要工具。第 6 章讲述了版本戳（version stamp），它用来记录数据库的内容变更，并且可以检测数据是否一致。第 7 章概述了“映射 - 化简”（Map-Reduce）操作，这种计算方式很适合在集群中组织并行计算，因而也适用于 NoSQL 系统。

讲完这些概念后，我们针对以下 4 种数据库各举一些例子，来演示如何实现上述概念。第 8 章使用 Riak 来演示“键值数据库”，第 9 章使用 MongoDB 作为“文

档数据库”的示例，第 10 章选用 Cassandra 来探讨“列族数据库”，第 11 章选择了 Neo4J 作为“图数据库”的示例。此处必须强调：要想全面学习数据库，只依靠这些章节是不够的。因为除此之外还有很多内容，没办法写在这本书中，而且还有更多东西必须尝试之后才能学会。本书选择这些示例，并不是建议大家在工作中使用它们，其目的是让读者知道数据的各种存储方式，明白不同的数据库技术如何使用前面提到的概念。读者会看到这些数据库系统都需要何种程序代码，并且简单了解使用它们时所应遵循的开发思路。

有些人经常会觉得：因为 NoSQL 数据库没有模式，所以在应用程序的生命期中，可以毫无困难地改变其数据结构。本书不同意此观点，因为无模式的数据库其实隐含了一种模式，在实现数据结构变更时，也必须修改其规则。所以，第 12 章解释了数据如何在强模式与无模式系统之间迁移。

所有这一切都清楚地表明：NoSQL 不是独立存在的，也不会取代关系型数据库。第 13 章着眼于混合持久化领域的发展趋势：多种数据存储方式将共存，有时甚至会存在于同一个应用中。第 14 章将大家的视野扩展至本书之外，在混合持久化领域中，考虑一些前面没有涉及的技术。

掌握了前面所讲的全部内容之后，读者就应该明白如何选择合适的数据存储技术了。所以最后一章（第 15 章）提供了一些选择数据库时可以参考的建议。笔者认为，有两个关键因素：找到一种高效的编程模型，其数据存储模型要非常符合待开发的应用程序，并且确保其获取数据的效率与弹性均符合开发者的需求。从 NoSQL 诞生之初，我们就担心没有一套定义明确的流程可以遵循，现在，你仍然需要结合自己的需求，来验证自己所选择的数据库技术是否合适。

本书只是个简要的概述，所以笔者一直在尽力压缩篇幅。我们精选了自己认为最重要的信息，这部分内容读者就不必再去找了。如果打算认真研究这些技术，那就需要进一步研读本书以外的知识了，不过，我们还是希望本书能为你的探索之路开个好头。

还需要强调的是：计算机领域中的这些技术是日新月异的，存储技术的某些重要方面在不断变化，每年都会出现新的特性与新的数据库。笔者投入了巨大的精力来专门讲述概念，因为就算底层技术变了，对这些概念的理解也依然有价值。我们非常确信，本书所讲的大部分概念都会历久绵长，但绝不能保证所有概念都会如此。

谁应该阅读本书

如果正在考虑选用某种形式的NoSQL数据库，那就应该阅读本书。选用NoSQL的原因可能是你打算做一个新的项目，也可能是既有项目遭遇瓶颈，所以要将其数据库迁移到NoSQL数据库上。

本书致力于给读者提供足够的信息，以判断自己所选的NoSQL技术是否符合需求，如果符合的话，应该深入研究哪些工具。我们设想本书的主要读者是架构师或技术主管，然而那些想大概了解这门新技术的软件管理人员也可以阅读本书。此外，对于想大概了解这项技术的开发人员来说，这也是本很好的入门读物。

本书不讲编程细节，也不去部署某个特定的数据库，那些内容留待更为专业的教材来写吧。我们还严格限制了本书的篇幅。笔者认为，这种书应该在坐飞机的时候读：它不会回答你提出的所有问题，但却会激发你提出一堆好问题来。

若是之前已经深入研究了NoSQL领域，那么本书可能不会增加你的知识储备。不过，它仍然有助于你将之前学到的东西解释给别人听。把围绕着NoSQL的争论理解清楚是很重要的，尤其当你要劝说别人在项目中也采用NoSQL技术时更是如此。

本书要讲的数据库类型

本书遵循常见的分类方式，也就是按照数据模型来划分各种NoSQL数据库。下表列出了4种数据模型，以及归属于每种数据模型的数据库。这份列表并不完整，其中只列出了较为常见的数据库。撰写本书时，在http://nosql-database.org与http://nosql.mypopescu.com/kb/nosql都可查阅到更为完整的列表。每个分类中，以斜体标出的数据库，都会在相关章节中作为范例来讲解。

数据模型	范例数据库	数据模型	范例数据库
键值（参见第8章）	BerkeleyDB LevelDB Memcached Project Voldemort Redis *Riak*	文档（参见第9章）	CouchDB *MongoDB* OrientDB RavenDB Terrastore
列族（参见第10章）	Amazon SimpleDB *Cassandra* HBase Hypertable	图（参见第11章）	FlockDB HyperGraphDB Infinite Graph *Neo4J* OrientDB

这样划分的目的是从每一类数据库中，选出一个最有代表性的工具来讲。尽管每个分类下列出的那些数据库各不相同，不可像这样一概而论，但是，书中提到的那些具体示例，其实大多数情况下也适用于此分类中的其他数据库。我们会从“键值数据库”“文档数据库”“列族数据库”和“图数据库”这4类中各选一个作为范例，此外，在必要时，还会提到可以满足某个特定功能的其他产品。

按数据模型来分类是可行的，但却失之武断。不同数据模型之间的界限往往是模糊的，比如键值和文档数据库（参见2.2节）之间的区别就不是很明显。许多数据库并不能明确地归入某一类。例如，OrientDB称自己既是文档数据库又是图数据库。

致谢

首先感谢ThoughtWorks的诸位同仁，在过去的几年中，很多同事在交付的项目中应用了NoSQL。笔者写作本书的动机主要来源于他们的经验，而这些经验亦是能印证NoSQL技术价值的实用信息。目前为止通过使用NoSQL数据存储积累了一些有益的经验，基于这些经验，我们认为：NoSQL是一项重要的数据存储技术，它正引发该领域内的一场重大变革。

我们也要感谢举办公开讲座、发表文章和博客来分享NoSQL使用心得的各种社群。若是大家都不愿意与同行分享研究成果的话，那么许多软件开发领域的发展就不为人知了。特别感谢谷歌及亚马逊的BigTable和Dynamo技术规范论文，它们对NoSQL的发展影响深远。也要感谢为开源NoSQL数据库的开发提供赞助及技术贡献的公司。这一次发生在数据存储领域的变革，与以往相比有一个较为有趣的差别：NoSQL的发展深度植根于开源工作。

特别感谢ThoughtWorks公司给予笔者时间来写作本书。我们两个大约同一时间加入ThoughtWorks，并且在这里工作了10余年。ThoughtWorks对我们来说一直是个非常友好的大家庭，同时也是知识和实践的来源。在这个良好的环境中，大家可以公开分享各自所学的知识，这与传统的系统交付公司（System Delivery Organization）非常不同。

Bethany Anders-Beck、Ilias Bartolini、Tim Berglund、Duncan Craig、Paul Duvall、Oren Eini、Perryn Fowler、Michael Hunger、Eric Kascic、Joshua Kerievsky、Anand Krishnaswamy、Bobby Norton、Ade Oshineye、Thiyagu

Palanisamy、Prasanna Pendse、Dan Pritchett、David Rice、Mike Roberts、Marko Rodriquez、Andrew Slocum、Toby Tripp、Steve Vinoski、Dean Wampler、Jim Webber 和 Wee Witthawaskul 审阅了本书初稿，并提出了改进建议。

此外，Pramod 要感谢绍姆堡图书馆（Schaumburg Library）提供的一流服务和安静的写作空间；感谢爱女 Arhana 和 Arula，你们知道爸爸到图书馆是为了写书，而没有带你们同去；感谢爱妻 Rupali，你给了我巨大的支持和帮助，让我能够集中精力完成本书。

Contents