



NoSQL

[美] Joe Celko 著
王春生 范东来 译

Joe Celko's Complete Guide to NoSQL
What Every SQL Professional Needs to Know about Nonrelational Databases

权威指南

 中国工信出版集团

 人民邮电出版社
POSTS & TELECOM PRESS

NoSQL

Joe Celko's Complete Guide to NoSQL
What Every SQL Professional Needs to Know about Nonrelational Databases

权威指南

[美] Joe Celko 著
王春生 范东来 译

人民邮电出版社
北 京

图书在版编目(CIP)数据

NoSQL权威指南 / (美) 乔·塞科 (Joe Celko) 著 ;
王春生, 范东来译. — 北京 : 人民邮电出版社, 2016. 7
书名原文: Joe Celko's Complete Guide to NoSQL:
What Every SQL Professional Needs to Know about
Non-Relational Databases
ISBN 978-7-115-42787-8

I. ①N… II. ①乔… ②王… ③范… III. ①数据库
系统—指南 IV. ①TP311.138-62

中国版本图书馆CIP数据核字(2016)第134766号

内 容 提 要

本书是根据作者进行培训和开发的经验编写的 NoSQL 权威指南, 是了解在什么场景、什么时候、为什么 NoSQL 的好处超过 SQL 的理想书籍。通过本书, 读者可以对 SQL 的缺点多于好处的场景有一个完整的理解, 而后更好地确定何时使用 NoSQL 技术可以获得最大的好处, 对列式数据库、流式数据库和图数据库的利弊有更深入的理解, 并在备受欢迎的 SQL 专家 Joe Celko 的指导下顺利过渡到 NoSQL。

本书会详细介绍 NoSQL 常见的数据库的历史、技术原理及其优缺点, 这些数据库包括列式数据库、图数据库、流式数据库、键值数据库、文本数据库、地理信息数据库、指纹数据库、分析型数据库、多值数据库以及层次数据库等。

本书适合所有对 NoSQL 数据库感兴趣的技术人员阅读。

-
- ◆ 著 [美] Joe Celko
 - 译 王春生 范东来
 - 责任编辑 杨海玲
 - 责任印制 焦志炜
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
 - 邮编 100164 电子邮件 315@ptpress.com.cn
 - 网址 <http://www.ptpress.com.cn>
 - 北京天宇星印刷厂印刷
 - ◆ 开本: 800×1000 1/16
 - 印张: 12.5 2016 年 7 月第 1 版
 - 字数: 242 千字 2016 年 7 月北京第 1 次印刷
 - 著作权合同登记号 图字: 01-2014-8591 号
-

定价: 45.00 元

读者服务热线: (010)81055410 印装质量热线: (010)81055316
反盗版热线: (010)81055315

版 权 声 明



Joe Celko's *Complete Guide to NoSQL: What Every SQL Professional Needs to know about Non-Relational Databases*, First Edition

Joe Celko

ISBN: 9780124071926

Copyright 2014 by Elsevier Inc.. All rights reserved.

Authorized Simplified Chinese translation edition published by Elsevier (Singapore) Pte Ltd. and Posts and Telecom Press

Copyright 2016 by Elsevier (Singapore) Pte Ltd.

All rights reserved.

Published in China by Posts and Telecom Press under special arrangement with Elsevier (Singapore) Pte Ltd.. This edition is authorized for sale in China only, excluding Hong Kong, Macau and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书简体中文版由 Elsevier (Singapore) Pte Ltd. 授予人民邮电出版社在中国大陆地区（不包括香港、澳门以及台湾地区）出版与发行。未经许可之出口，视为违反著作权法，将受法律之制裁。

本书封底贴有 Elsevier 防伪标签，无标签者不得销售。

对本书的赞誉



“如果你遇到了通过 SQL 模型无法解决的问题，或者你想学习数据管理方面的知识，尤其是 NoSQL 方面，那么 Joe Celko 的书会是你理想的选择。”

——Jeff Garbus，飞鹰咨询合伙人

译者序



如果说某种技术是以“No”开头的，那么这种技术想必是会很有趣的。NoSQL 正好属于此类。当一个个有趣的名字（Cassandra、Dynamo、MongoDB 等）从我们眼前掠过时，我们才发现，“No”并不是想要否定什么，而是给了我们更多的选择。

1970 年，Codd 博士提出了关系模型，并给出了“Codd 十二定律”，自此关系数据库横空出世，逐渐占领了主流市场。得益于成熟的 SQL 标准，直到今日关系数据库仍然长盛不衰。SQL 也成为数据库中的必修课。关系模型也逐渐超越了层次模型和网络模型。

从 2003 年开始，Google 公司的 3 篇论文极大地推动了云计算、大数据技术的发展，而这些技术的发展也反过来刺激了大数据应用的落地。人们逐渐发现关系数据库在某些场景下越来越力不从心，高可用的需求越来越强烈，有些时候性能比一致性更加重要，不是每个查询都需要做连接……就在这时 NoSQL 如一股春风吹进数据库世界，但其实我们回头看看，Berkeley DB 早在 1991 年就已经问世，并且已经内置到了 BSD 里。如今，NoSQL 产品层出不穷，架构、实现原理都差异较大，并且其应对的业务场景也是大相径庭的。这给初学者带来了一些困难。那么，学习 NoSQL 的第一步就是了解它们。

如果你是 NoSQL 系统的支持者，那么你会很自然地想要去了解它，如果你是 NoSQL 系统的反对者，那么你就必须了解它。谁能帮你完成这个任务呢？Joe Celko 无疑是很好的入选，作为数据库专家，他的《SQL 权威指南》在国内已经多次再版，是很多人的数据库启蒙教材，另外他还参与了 SQL-89 和 SQL-92 标准的制定，另著有一系列数据处理的数据库相关的书籍，是一位高产的作者。

本书中，除了常规的键值数据库，对于地理信息数据库、流式数据库、指纹数据库甚至分析型数据库都有涉猎。本书并不关注 NoSQL 数据库产品本身，但是它能帮你理清这些产品之间的关系与特性，使你在根据业务场景构建自己的架构时更加得心应手。

NoSQL 从诞生之日起其初衷就不是取代关系数据库，而是作为关系数据库的一种有益的补充，不要因为它改变了现有秩序而对它敬而远之。可以预见的是，NoSQL 系



统和关系数据库将共存下去。读完这本书，我们会发现，No 有时并不是 “no”，还可以是 “not only”。

范东来

2016 年 5 月于成都

译者简介



范东来 北京航空航天大学硕士，成都数联铭品（BBD）科技有限公司大数据技术部负责人，大数据平台架构师，著有《Hadoop 海量数据处理》，译有《解读 NoSQL》，极客学院布道师，研究方向为并行图挖掘、模式分类。

王春生 网名“平凡的香草”，典型的“完美主义+强迫症+现实主义”综合体，追求完美并苛刻。新浪第一个大规模数据库平台的设计者和开发者，先后担任过 DBA、架构师、技术总监等职位，对数据库、网络、系统等环节的应用和开发均有丰富经验，曾参与翻译《Puppet 实战手册》。微博@平凡的香草。

引言



“没有什么会比引入新秩序更难，因为创新者必须要面对那些在旧环境中已经做得很好的对手，以及那些在新环境中做得很好的冷漠者。”

——*Niccolo Machiavelli*^①

在过去的几十年，我已经通过 Elsevier/Morgan Kaufmann 出版社出版了一系列的书，这些书几乎全部是关于 SQL 和 RDBMS 的。而这本书对行业媒体中所谓的大数据、新 SQL 或 NoSQL（我们这些极客非常喜欢流行语）做了一些概述。第一个创造或挖掘了新名词的专栏作家或博主很可能会在维基百科上有对其（新名词）进行的阐述，甚至可能会写成一本书来讨论。

由于 SQL 模型是事实上的主导数据库模型，因此任何其他模型都会被认为是挑战者。但是，我们应该用什么流行词来表示这种新的模型呢？我们已经在 SQL 中存放数 PB 的数据很多年了，因此，“大数据”这个词看起来并不合适。而且，SQL 每五年会引入一个新的 ANSI/ISO 标准并发布，并不是一夜之间就从“旧 SQL”变为“新 SQL”的。“新 SQL”这个名词让我想起了新可口可乐（New Coke^②），新可口可乐并没有激发用户的信心，也没有获得成功。

在当下的所有流行语中，我觉得“NoSQL”是最好的，我把它读作“N. O. SQL”（就是“not only SQL”的简写），而不是其常被读作的“no SQL”。“no SQL”意味着在过去 40 多年中，数据库技术做得非常不好，但事实并非如此！很多使用 SQL 的人，尤其是我，在处理 IT 工作时，越来越坚信俗语中说的“手里有锤子的人会把眼中的一切都认为是钉子^③”，但做 IT 工作并不是用锤子盖房子。

我们使用的一些数据库工具已经存在了几十年，甚至比 RDBMS 还要早。当然，由于技术发展使然，产生了很多新的工具。打开工具箱时，要考虑工具的所有可用选项，以及如何用它们完成任务。

① 尼科洛·马基雅弗利，意大利新兴资产阶级思想政治家，历史学家（http://en.wikipedia.org/wiki/Niccol%C3%B2_Machiavelli）。——译者注

② “kid with a hammer who thinks every problem is a nail”，参见 https://en.wikipedia.org/wiki/New_Coke。——译者注

③ 参见 http://en.wiktionary.org/wiki/if_all_you_have_is_a_hammer_everything_looks_like_a_nail。——译者注



这本定位为概述的书可以带大家快速浏览一下你可能从没听说过或是已经忘记的老技术。然后，我们再研究“新玩意”，了解它为什么会存在。

我对硬件或深入研究某些特定软件不是那么感兴趣。这本书没有足够的篇幅讲这些，你可以针对自己感兴趣的或自己项目中某个聚焦的细节，阅读一些对应的书籍。请把这本书看成是百货公司的商品目录，通过它，你可以获得一些灵感，或是学到一些新的流行语。

请发送更正和意见至 jcelko212@earthlink.net，并在本书的配套网站 (<http://elsevierdirect.com/v2/companion.jsp?ISBN=9780124071926>) 中寻找反馈。

下面是你可能期待在这本书中能找到的内容的快速导览。

第 1 章：NoSQL 和事务处理。将作业队列读入大型计算机仍然是大量商业数据处理采用的方式。事务处理模型通过使用新的 ETL 工具来加载数据库，完成批处理作业。我们需要了解批处理和事务处理这两种模型，以及它们在新技术中如何使用。

第 2 章：列式数据库。列式数据库使用传统的结构化数据，并经常运行某些版本的 SQL，不同之处在于它们存储数据的方式。传统的面向行的方式被将数据以列的方式存放替代，并将数据重新组装成人们在 RDBMS 模型中熟悉的行。由于这些列，每个列都被定义为一种且仅一种数据类型或域空间，因此，它们可以被压缩，并且可以分布于多个存储系统中，如 RAID。

第 3 章：图数据库。图数据库是基于图论的，图论是离散数学的一个分支。图数据库基于实体之间的关系建立模型，而不是根据这些属性的值做计算和聚合，以及根据这些值进行检索。

第 4 章：MapReduce 模型。在 IT 行业媒体上一般称为 NoSQL 或大数据的领域中，MapReduce 模型是最流行的。它用于在大型文件系统中并行地快速检索大量数据。这些系统在按照这样的速度和数量进行处理时，较少考虑数据完整性。基本操作是简单的，很少优化。但是，很多应用都愿意为此做出权衡。

第 5 章：流式数据库和复杂事件。关系模型和之前的传统数据库系统假定在查询时表是静态的，并且其结果也是一个静态表。但流式数据库是建立在不断流动的数据的模型之上的，可以认为是数据实时流动的河流或管道。流式数据库中最著名的例子是要在一秒内完成交易的股票软件和商品交易软件。这种系统必须根据这一流中的事件采取行动。

第 6 章：键值存储。从本质上来说，键值存储是键值对 (<键>, <值>) 的集合，可以理解为一类简单的数组。在每个集合内部，键都是唯一的。键可以是任意可以检验相等性的数据类型。这是 MapReduce 系列的一种形式，但其性能取决于键的设计方法。因此，散列成为很重要的技术。

模式与无模式。SQL 和所有以前的数据库都使用“模式”(schema)来定义它们的结构、

约束、默认值等。但是，使用和维护模式是有开销的。“无模式”的方式将所有的数据完整性（如果有的话！）放在应用程序中进行处理。但是，表现层却没有办法知道它们将被返回什么样的数据。这些系统是为检索优化的，SQL 系统的安全性和查询能力被更好的可扩展性和检索性能替代。

第 7 章：文本数据库。最重要的业务数据并不在数据库或文件中，而是在文本中。这些数据可能是合同/契约、担保条款、信件、手册和参考资料。文本天然就是模糊并且笨重的，传统的数据需要被精确和紧凑地编码。起初，文本数据库只能找到文件，但是随着算法的改进，我们已经能够理解阅读和文本的要点。

第 8 章：地图数据。地理信息系统（Geographic information system, GIS）是一个存储地理、地理空间或者时空数据的数据库。它不仅是一种地图制图。我们不仅是试图在地图上定位某样东西，还试图找出某个地区内这些东西的数量、密度和内容，在一段时间中的改变，等等。

第 9 章：大数据和云计算。大数据这个术语是由 Forrester Research 公司在白皮书中随着 4V 的流行语提出的，大数据的 4V 指的是：规模（volume）、速度（velocity）、多样性（variety）和可变性（variability）。大数据已经应用在目前我们已经讨论过并且试图协调混用多种数据库模型的场景下。

第 10 章：生物特征、指纹和专业数据库。生物特征还没有开始商用。它们用一个生物实体而非商业实体来识别一个人。我们处在一个医学和法制的世界。到最后，当安全性成为一个议题时，生物特征才可能会转到商用领域，并且我们希望安全地交易隐私信息。

第 11 章：分析型数据库。传统 SQL 数据库被用于联机事务处理。它的目标是为日常业务应用提供支持。联机分析处理数据库是构建在 OLTP 数据之上的，但是这种模型的目标是运行处理数据聚合的查询而非独立的事务。它是分析型而非事务型的。

第 12 章：多值数据库或 NFNF 数据库。RDBMS 是基于第一范式的，它假设数据以标量值的形式保存在行中的列里，这些记录有着相同的结构。多值模型允许用于表内嵌到列里。它们有一个不被 SQL 程序员所熟知的小众市场。这种数据模型有一种代数理论，听起来就像关系型模型。

第 13 章：层次数据库系统和网络数据库系统。IMS 和 IDMS 是最重要的前关系型技术，直到今天还在广泛使用。事实上，有一个好机会，IMS 仍然比 SQL 数据库保存有更多的商业数据。这些产品仍然在银行业、保险业和大型机上的大型商业应用里“繁忙工作”，它们使用 COBOL。它们对那种变化不大而又需要移动大量数据的场景非常适合。因为如此多的数据仍然存在于 IMS 和 IDBS 中，所以你必须至少了解分层数据库系统和网络数据库系统的基础知识，以便从中得到数据再用 NoSQL 工具保存。

作者简介



Joe Celko 为 ANSI/ISO SQL 标准委员会工作了 10 年，并且参与编写了 SQL-89 和 SQL-92 标准。

Celko 先生是 Elsevier/Morgan Kaufmann 出版的一系列关于 SQL 和 RDBMS 图书的作者。他是得克萨斯奥斯汀的一位独立咨询师。他已经为计算机行业和学术出版写过 1200 多个专栏，其中大多数都是关于处理数据和数据库方面的。

目 录



第 1 章 NoSQL 和事务处理	1
简介	1
1.1 批处理中的数据库事务处理	1
1.2 磁盘处理中的事务处理	2
1.3 ACID	2
1.4 悲观并发详解	4
1.4.1 隔离级别	5
1.4.2 私有的隔离级别	7
1.5 CAP 定理	8
1.6 BASE	9
1.7 服务器端一致性	10
1.8 错误处理	11
1.9 为什么在这些场景下 SQL 不能发挥作用	11
总结思考	12
参考文献	12
第 2 章 列式数据库	13
简介	13
2.1 列式数据库的历史	14
2.2 技术原理	19
2.3 查询优化	19
2.4 多用户和硬件	19
2.5 执行一个 ALTER 语句	20
2.6 数据仓库和列式数据库	21
总结思考	21
参考文献	22



第3章 图数据库	23
简介	23
3.1 图论基础	24
3.1.1 节点	24
3.1.2 边	25
3.1.3 图的结构	25
3.2 RDBMS 与图数据库	26
3.3 凯文·贝肯问题的六度	26
3.3.1 通用图的邻接表模型	27
3.3.2 通用图的覆盖路径模型	30
3.3.3 真实数据的复杂关系	32
3.4 顶点覆盖	34
3.5 图编程工具	35
3.5.1 图数据库	36
3.5.2 图数据库语言	36
总结思考	39
参考文献	39
第4章 MapReduce 模型	41
简介	41
4.1 Hadoop 分布式文件系统	43
4.2 查询语言	43
4.2.1 Pig Latin	43
4.2.2 Hive 和其他工具	52
总结思考	54
参考文献	54
第5章 流式数据库和复杂事件	55
简介	55
5.1 代并发模型	56
5.1.1 乐观并发	56
5.1.2 乐观并发下的隔离级别	56
5.2 复杂事件处理	58

5.2.1 与事件处理相关的术语	59
5.2.2 事件处理与状态更改约束	61
5.2.3 事件处理与 Petri 网	62
5.3 商业产品	63
5.3.1 StreamBase	63
5.3.2 Kx	66
总结思考	68
参考文献	68

第6章 键值存储 69

简介	69
6.1 模式与无模式	69
6.2 查询与检索	70
6.3 “键”的处理	70
6.3.1 Berkeley DB	71
6.3.2 通过树索引或散列访问	71
6.4 “值”的处理	71
6.4.1 任意字节数组	72
6.4.2 已知结构的小文件	72
6.5 产品	73
总结思考	75

第7章 文本数据库 77

简介	77
7.1 经典文档管理系统	77
7.1.1 文件索引和存储	78
7.1.2 关键字和题内关键字	78
7.1.3 行业标准	79
7.2 文本挖掘与理解	83
7.2.1 语义与语法	83
7.2.2 语义网	84
7.3 语言问题	85
7.3.1 Unicode 和 ISO 标准	86
7.3.2 机器翻译	86



总结思考	87
参考文献	88

第8章 地图数据 89

简介	89
8.1 GIS 查询	90
8.1.1 简单位置	90
8.1.2 简单距离	91
8.1.3 在一个区域中查找数量、密度和内容	91
8.1.4 邻近关系	91
8.1.5 时间关系	91
8.2 定位	92
8.2.1 经度和纬度	92
8.2.2 层次三角网格	93
8.2.3 街道地址	95
8.2.4 邮政编号	96
8.2.5 ZIP 编码	96
8.2.6 加拿大邮政编码	96
8.2.7 英国邮政编码	97
8.3 GIS 的 SQL 扩展	99
总结思考	99
参考文献	100

第9章 大数据和云计算 101

简介	101
9.1 对大数据和云计算的疑问	102
9.1.1 云计算仅是一种时尚	103
9.1.2 云计算没有内部数据服务器那么安全	103
9.1.3 云计算代价高昂	103
9.1.4 云计算太复杂	103
9.1.5 云计算对大公司才有意义	104
9.1.6 只是技术上的改变	104
9.1.7 如果网络中断, 云计算将毫无用处	105
9.2 大数据和数据挖掘	105

9.2.1 用于非传统分析的大数据	106
9.2.2 系统合并的大数据	107
总结思考	107
参考文献	108

第10章 生物特征、指纹和专业数据库 109

简介	109
10.1 原生生物特征	110
10.2 指纹	111
10.2.1 分类	112
10.2.2 匹配	113
10.2.3 NIST 标准	113
10.3 DNA 识别	115
基本原则和技术	116
10.4 面部数据库	117
10.4.1 历史	118
10.4.2 谁在使用面部数据库	119
10.4.3 它有多好	120
总结思考	121
参考文献	121

第11章 分析型数据库 123

简介	123
11.1 数据立方体	123
11.2 Codd 博士的 OLAP 规则	124
11.2.1 Codd 博士理论的基础特性	124
11.2.2 独有特性	126
11.2.3 报表特性	126
11.2.4 维度控制	127
11.3 MOLAP	127
11.4 ROLAP	128
11.5 HOLAP	128
11.6 OLAP 查询语言	128
11.7 SQL 中的聚合操作符	129