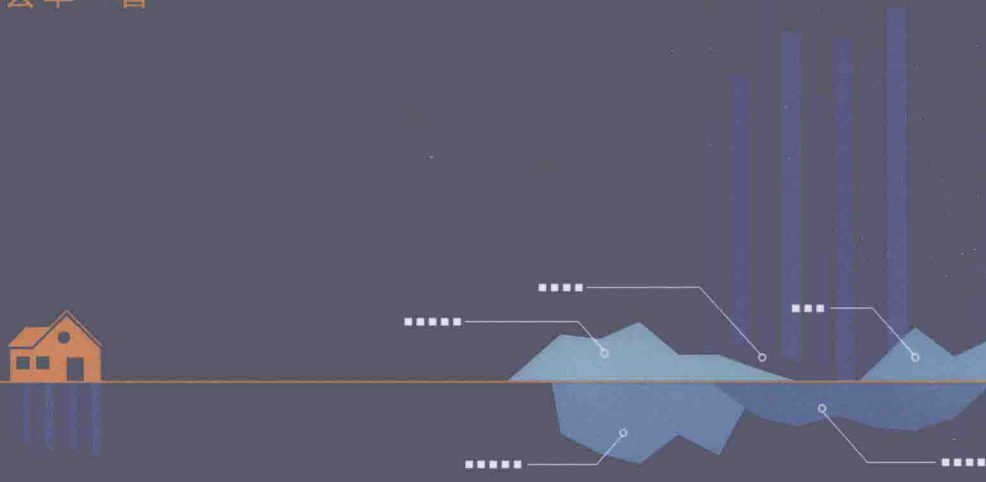
大数据时代
数据仓库技术研究

Techniques Research for Data Warehouse in Big Data

王会举 著

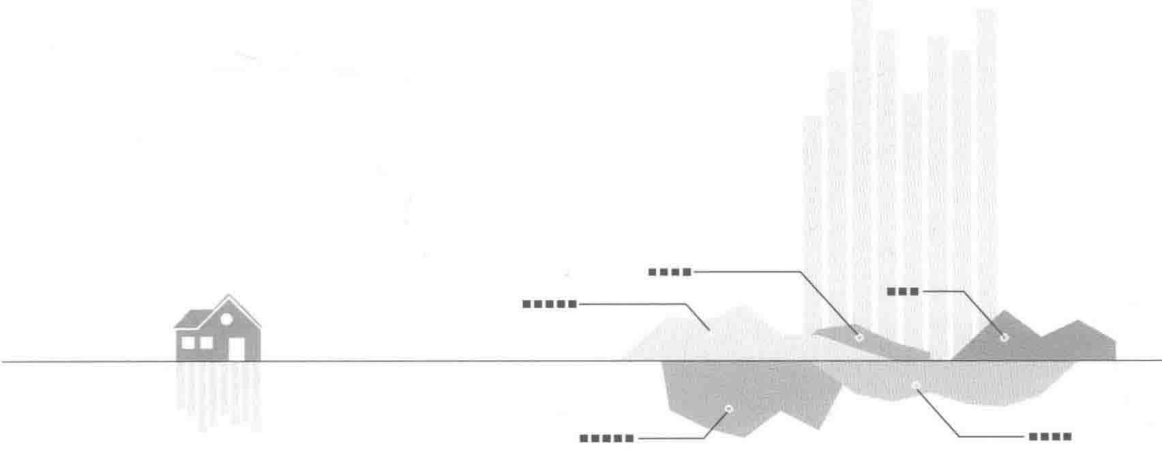


WUHAN UNIVERSITY PRESS
武汉大学出版社

大数据时代
数据仓库技术研究

Techniques Research for Data Warehouse in Big Data Era

王会举　著



WUHAN UNIVERSITY PRESS
武汉大学出版社

图书在版编目(CIP)数据

大数据时代数据仓库技术研究/王会举著. —武汉：武汉大学出版社，2016.12

ISBN 978-7-307-18873-0

Ⅰ.大… Ⅱ.王 … Ⅲ.数据库系统—研究 Ⅳ.TP311.13

中国版本图书馆 CIP 数据核字(2016)第 288743 号

责任编辑：王金龙　　责任校对：李孟潇　　整体设计：韩闻锦

出版发行：**武汉大学出版社**　(430072　武昌　珞珈山)

(电子邮件：cbs22@whu.edu.cn　网址：www.wdp.com.cn)

印刷：虎彩印艺股份有限公司

开本：720×1000　1/16　印张：10.75　字数：156 千字　插页：1

版次：2016 年 12 月第 1 版　2016 年 12 月第 1 次印刷

ISBN 978-7-307-18873-0　定价：29.00 元

前　言

2005 年，笔者受原单位委托，要基于某省第一次全国经济普查数据，用 Cognos 和 SQL Server DTS 为该省某局开发一个数据分析系统，自此便和数据仓库结下了不解之缘。掐指算来，笔者已在这个领域摸爬滚打 12 年有余：4 年多的数据仓库项目开发实施经历，外加 8 年多国内外大数据和数据仓库研究经历。现在呈现在您面前的，便是笔者这 10 多年来对数据仓库的一些思考和总结。

本书以大数据为时代背景，系统分析了传统数据仓库技术当前存在的问题及面临的挑战，并全面深入对比了当前主流的面向大数据的数据仓库解决方案；在此基础上，笔者系统阐述了自己的一系列研究工作，包括两个原型系统 LinearDB 和 Pagrol 的核心技术——面向大数据的数据存储、可扩展且高效的查询处理模型、面向属性图的多维模型构建以及并行属性图多维立方体计算算法设计实现等，内容涵盖了 ROLAP 和 MOLAP 两种主要的 OLAP 实现方式。

本书具有取材新颖、系统性强、内容丰富、手段方法先进等特点，反映了当前大规模数据仓库研究的最新动态和成果，具备一定的学术价值和实用价值。本书可供计算机、信息管理与信息系统、数据仓库/商务智能、大数据分析等相关专业的科研、教学及管理人员参考，也可作为大数据处理相关工程技术人员的参考用书。

本书的撰写离不开以往各位合作者的辛苦付出，尤其是中国人民大学王珊教授、杜小勇教授、周煊副教授、张延松副教授、覃雄派博士、李芙蓉同学、覃左言同学，新加坡国立大学 Tan Kian-Lee 教授、Wang Zhengkui 博士、Fan Qi 博士及美国加州大学圣芭芭拉分校 Divyakant Agrawal 教授、Amr El Abbadi 教授等人的辛勤劳动，

在此表示深切的感谢。在本书撰写过程中，得到了中南财经政法大学信息与安全工程学院老师们的热情鼓励和帮助，作者向他们以及一切支持本项研究工作的人们致以真诚的谢意。

由于大数据处理技术发展迅速，而本书的篇幅有限，因此在取材和论述方面必然有不全面之处，敬请广大读者指正。

本书为中南财经政法大学信息与安全工程学院学术专著基金和中南财经政法大学人才引进科研启动金资助。

王令举

2016年10月

摘　要

大数据时代，数据仓库系统中的管理对象已发生质的变化。一方面，管理对象的规模在持续爆炸式增长着，PB 级的数据仓库已是寻常规模。数据规模的变化，对数据仓库查询处理模式的影响是根本性的。以简单的扫描操作为例，1PB 数据在 50MB/s 的 I/O 速度①下，仅执行一次扫描操作就需要 230 天②。数据“量”的变化，急需技术“质”的更新。另一方面，管理对象呈现出多样化的特点。随着传感器、智能设备以及社交协作技术的飞速发展，现实世界中的数据也变得更加复杂，因为它不仅包含传统的关系型数据，还包含来自网页、社交网络、互联网日志文件等结构化、半结构化和非结构化数据。考虑到关系模型和属性图的紧密关联性③，本书以关系型数据和属性图数据(如 Facebook、万维网等)为例进行研究。

大规模并行处理是大数据处理的有效途径。出于代价的考虑，由中低端硬件构成的大规模机群环境成为大数据分析的主流计算平台。如果利用 6000 块磁盘仍以 50MB/s 的 I/O 速度并行读取 1PB 的数据，整个读取仅需要 1 个小时。运行于大规模机群上的 MapReduce 平台是目前大数据处理的基础平台。

---

① 当前服务器的 I/O 速度一般为 50MB/s 及以下，高端的在 100MB/s 左右。鉴于大数据分析的主流平台一般基于中低端硬件搭建，我们选择一般的 I/O 速度进行计算，此 I/O 速度的选择并不会导致结论性的变化。

② 此处假设有 1 磁盘可以装下 1PB 的数据，磁盘读取时，没有借助预取等高级 I/O 特性。

③ 属性图可看做是关系模型的泛化，它不仅包含节点/边实体属性信息和不同实体型间实体联系信息(如学生与课程间的选课关系)，也包含同一实体型内实体联系信息(如学生间关系)。

新的计算环境和复杂的管理对象挑战着传统数据仓库系统。(1)传统数据仓库扩展能力面临巨大挑战。并行数据库是传统数据数据仓库系统处理海量数据的主流平台，而并行数据库扩展性有限，至多可扩至百级节点规模，导致传统数据仓库系统难以实现大规模可扩展能力。此外，传统数据仓库的可靠性依赖于高端硬件来保证。在中低端硬件构成的计算环境下，原本可靠的硬件变得不再可靠，导致原有的基于高端硬件平台设计的并行 OLAP 查询算法不能适应这种由不可靠计算单元组成的大规模并行计算环境。计算的容错能力也严重限制了传统数据仓库系统的扩展性。(2) 传统数据仓库技术难以应对新的数据类型。以属性图为例，属性图中既包含节点属性信息，也包含节点间边的联系信息和属性信息。而传统的数据仓库技术主要处理节点/边属性信息和不同实体型间实体联系信息，无法处理同一实体型内不同实体间的联系信息即无法处理图结构信息，导致其对属性图难以提供充分有效的 OLAP 分析功能。如何设计新的面向属性图的 OLAP 模型及实现算法，充分挖掘节点/边的属性信息和节点间的联系信息，开发有效的属性图分析功能，是一项亟待解决的工作。

本书主要关注如何基于 MapReduce 平台高效地处理巨量数据上的 OLAP 查询。OLAP 分为 ROLAP(Relational OLAP)和 MOLAP(Multidimensional OLAP)，出于内容的完整性，我们分别基于关系数据和属性图对其进行研究：

1. 面向关系数据的 ROLAP 研究

众所周知，MapReduce 的性能远低于并行数据库，尤其是连接操作，归根结底源于其执行方式：MapReduce 最初是面向单数据集上的扫描操作而设计的，而数据仓库查询往往涉及多个数据集间的连接操作。因而在基于 MapReduce 实现数据仓库查询时，往往需要启动多个 MapReduce 作业，并借助于物化的中间数据将这些作业连接起来，从而导致较高的 I/O 代价和网络传输代价。为解决此问题，我们设计了同时具备 MapReduce 的扩展性和关系数据库的性能新型数据仓库框架及执行引擎。具体研究思路为：基于 Ma-

pReduce 平台，利用关系数据库技术，设计一个同时具备关系数据库性能和 MapReduce 扩展性的新型数据仓库系统。其研究内容包括：

(1)提出“关系化”MapReduce 的思想，即在不改变 MapReduce 扩展性和容错性的前提下，利用关系数据库技术，根据 MapReduce 的执行特点，对其进行优化，以使其接近甚至达到关系数据库的性能。我们基于 MapReduce 平台提出了大规模可扩展的高效数据仓库架构，并从查询执行和数据存储两个关键点进行了深入研究。

(2)设计了面向 MapReduce 平台的高效的新型数据仓库查询执行框架。为了使 OLAP 查询的处理能够适应 MapReduce 框架的“扫描——聚集”处理模型，本书对传统的星形模型(雪花模型)的存储方式及星形(雪花)查询处理模式进行改造，提出了全新的无连接存储模型和 TAMP 执行模型。无连接存储模型基于层次编码技术，将维表层次等关键信息压缩进事实表，使得事实表可以独立地以扫描的方式对数据进行处理，从数据模型层保证了数据计算的独立性；TAMP 执行模型将 OLAP 查询的处理抽象为 Transform、Aggregation、Merge、Postprocess 四个操作，使得 OLAP 查询可被划分为众多可并行执行的独立子任务，从执行层保证了系统的高度可扩展特性。在性能优化方面，本书提出了 scan-index 扫描和跳跃式扫描算法，以尽可能地减少 I/O 访问操作；设计了并行谓词判断、批量谓词判断等优化算法，以加速本地计算速度。同时为了应对因维表更新所导致的层次编码变更问题，提出了多版本共存的数据更新协议。实验表明，原型系统 LinearDB 可以获得较好的扩展性和容错性，其性能比原有的 Hadoop 高出一个数量级。

(3)提出了针对 MapReduce 存储系统的智能型存储模型。MapReduce 依赖于数据文件块的冗余机制来获得较好的容错性，简单起见，每一个冗余块都采用相同的存储模型。这种方式忽略了如下事实：MapReduce 作为大数据处理分析的重要平台，其上运行的任务是多种多样的；不同任务的数据访问模式是不同的，单一存储模型无法适应所有任务。为了能让 MapReduce 同时从多种存储模型中获益，我们提出了智能型数据存储的思想：为同一数据块的不同

备份设计不同的存储方式，如第一个备份存为列存，第二个备份存为 PAX 存储等。对于每一个 MapReduce 任务，智能型存储模型对每一种存储格式的访问代价进行估计，并针对查询特点和当前负载状况，选择访问代价最低的冗余块(存储模型)进行数据的访问。基于该思想，我们以两种列存储模型——纯列存和 PAX 存储(新的存储模型称为 HC 存储模型(Hybrid column-store))为例，进行了实验研究。实验结果表明，HC 存储可以超越单一的 PAX 存储或列式存储，尤其是面对多样的查询负载时。

2. 面向属性图的 MOLAP 研究

属性图上的 MOLAP 立方体计算在大数据时代面临巨大挑战，原因如下：

(1)计算量较大。计算量大不仅仅因为数据规模大，还在于立方体中包含的单位立方体的个数较多。对于节点和边分别有 n 个维度和 m 个维度的属性图来说，在不考虑维度层次的情况下，需要计算 2^{n+m} 个单位立方体。

(2)单位立方体计算代价高。每一个单位立方体都是基于原始数据的一次聚集查询，每一个聚集查询往往涉及事实表和维表、节点信息和边信息的巨量数据连接，因此立方体的计算代价相当于执行 2^{n+m} 次基于巨量数据上的连接查询。

(3)属性图上的多维计算更加复杂。基于属性图的多维数据立方体，连接操作不仅存在于事实表和维表之间，也存在于节点和边之间；属性的聚集操作不仅要作用于属性维度上，也要作用于结构信息上。

为了应对该挑战，针对属性图，我们提出了基于 MapReduce 的 Pagrol 并行属性图 OLAP 系统，以有效且高效地对大型属性图提供有力决策支持。Pagrol 提出了一种新的面向属性图的立方体模型，即超图立方体(Hyper Graph Cube)，这种模型能够按照不同的粒度和层级对属性图进行聚合，以对不同查询提供支持；设计了新型的 OLAP 上卷/下钻操作，支持点/边上维度层次的灵活上卷/下钻操作。基于该模型，Pagrol 实现了基于 MapReduce 的并行图立方

体计算算法——MRGraph-Cubing，并使用了多种有效的优化技术：自成一体的连接策略、立方体分批方法、基于代价的批次分包方法（每个包包含多个批次）等。基于 Facebook 真实数据与人工数据进行的大量实验表明，Pagrol 可行、高效且具有高度可扩展性。

目　录

第1章 绪 论

1.1 研究背景

1.1.1 大数据时代

近年来，“大数据”已广为人知。大数据的热门主要源于两点共识。首先，在过去的20年间，数据产生速度越来越快。据国际数据公司IDC报道，人类产生的数据量正在呈指数级增长，大约每2年翻一番，2020年全球数据量将达到40ZB[1]。其次，大数据中隐藏着巨大的机会和价值，将给许多领域带来变革性的发展。著名管理咨询公司麦肯锡称：“数据已经渗透到当今每一个行业和业务职能领域，成为重要的生产因素。人们对于大数据的挖掘和运用，预示着新一波生产力增长和消费盈余浪潮的到来。”[2]美国政府把大数据称做“未来的新石油”，并制订了《联邦大数据研究与发展战略计划》[3]。我国国务院也发布了《促进大数据发展行动纲要》，将数据放在了战略性资源的高度[4]。各种期刊杂志（如《Nature》和《Science》）、公共媒体（如经济学人、美国时代周刊，美国国家公共广播电台等）充斥了大数据的相关信息。大数据研究已经吸引了产业界、政府和学术界的广泛关注。“大数据时代”已然来临。

大数据泛指大规模、复杂的数据集，因可从中挖掘出有价值的信息而备受关注，但传统方法无法对其进行有效处理和分析。其主要特征可总结为3V，即体量大（volume）、速度快（velocity）、数据类型多样（variety）。体量大指的是数据规模越来越大；速度快指数据产生的速度越来越快，相应地系统处理速度也要求越来越高；数据类型多样指

数据不仅是结构化数据，也有半结构化和非结构化数据。

1.1.2　数据管理技术发展历程

纵观历史，数据管理技术经历了人工管理、文件系统管理、数据库管理系统三个阶段。通过对其发展过程的分析，可以看到有四股力量推动着数据管理技术的发展：硬件技术、软件技术、管理对象和应用需求。由人工管理阶段进入到文件系统阶段，是因为在硬件上出现了随机访问设备，软件上研发出了操作系统和文件系统；由文件系统阶段步入数据库管理系统阶段，是因为管理对象的规模发生了较大变化，数据管理的应用领域越来越广，数据共享的需求越来越强烈。

数据管理技术在这三个阶段间的过渡均用了 10 年左右的时间。数据库管理系统自 20 世纪 60 年代末诞生以来，至今已有 40 年发展历史。在过去的 40 年里，数据管理系统的管理对象、软/硬件平台和用户需求经历了由量的积累到质的转变的过程，导致关系数据库已难以应对。为解决大数据分析的问题，各种 NoSQL 数据库产生并发展起来[5]。当前，数据管理技术已经进入了大数据管理阶段(见图 1-1，计算环境即为软/硬件平台)。

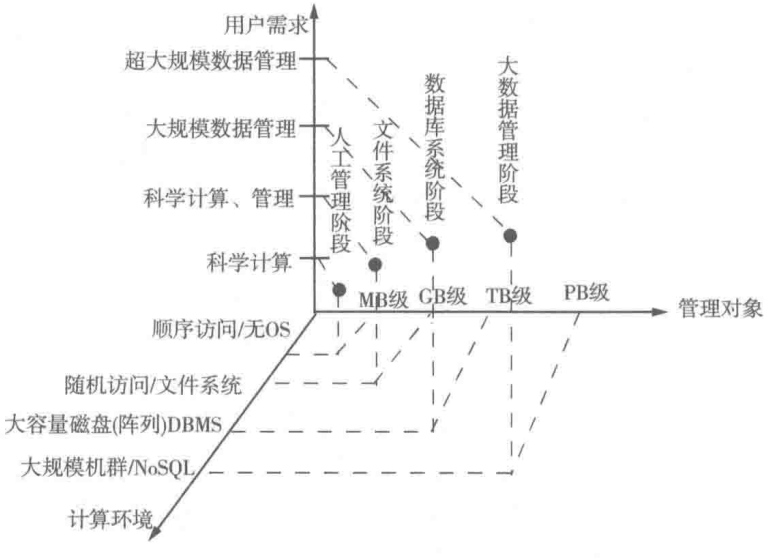


图 1-1　数据管理技术发展过程分析