HZ BOOKS
华章IT

Data Management in the Cloud
Challenges and Opportunities

大 数 据 管 理 丛 书

云数据管理
挑战与机遇

[美] 迪卫艾肯特·阿格拉沃尔（Divyakant Agrawal）
苏迪皮托·达斯（Sudipto Das）
阿姆鲁·埃尔·阿巴迪（Amr El Abbadi）
著

马友忠 孟小峰 译

机械工业出版社
China Machine Press

大/数/据/管/理/丛/书

Data Management in the Cloud

Challenges and Opportunities

云数据管理

挑战与机遇

迪卫艾肯特·阿格拉沃尔（Divyakant Agrawal）

[美] 苏迪皮托·达斯（Sudipto Das） 著

阿姆鲁·埃尔·阿巴迪（Amr El Abbadi）

马友忠 孟小峰 译

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

云数据管理：挑战与机遇 /（美）迪卫艾肯特·阿格拉沃尔（Divyakant Agrawal）等著；马友忠，孟小峰译 . —北京：机械工业出版社，2017.3
（大数据管理丛书）
书名原文：Data Management in the Cloud: Challenges and Opportunities

ISBN 978-7-111-56327-3

Ⅰ. 云… Ⅱ. ①迪… ②马… ③孟… Ⅲ. 数据管理 Ⅳ. TP274

中国版本图书馆 CIP 数据核字（2017）第 050722 号

本书版权登记号：图字：01-2016-5927

本书共分 7 章。第 1 章介绍了云计算、云数据管理的基本概念，并描述了本书的组织结构；第 2 章主要介绍了分布式数据管理的相关知识，包括分布式系统、P2P 系统、并发控制和分布式数据恢复等；第 3 章对云数据管理的早期研究工作进行了描述，包括不同的键 - 值存储系统在数据模型、数据分布和容错等方面的区别，以及 Bigtable、PNUTS 和 Dynamo 这三个有代表性的键 - 值存储系统的特点；第 4 章介绍了托管数据的事务问题，包括数据托管模式、托管数据的事务执行、数据存储和复制等内容；第 5 章主要介绍了分布式数据事务相关技术；第 6 章讨论了云数据管理中的多租户技术，包括多租户模型、云中的数据库弹性以及云中数据库负载的自动控制；第 7 章对相关经验教训进行了总结，并指出了未来的主要研究方向。

本书适合计算机及相关专业学生学习数据管理和分析使用，也适合对数据管理和分析感兴趣的其他开发人员阅读。

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：李秋荣
印　　刷：北京文昌阁彩色印刷有限责任公司　　版　　次：2017 年 5 月第 1 版第 1 次印刷
开　　本：170mm×242mm　1/16　　印　　张：9.75
书　　号：ISBN 978-7-111-56327-3　　定　　价：69.00 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）88378991　88361066　　投稿热线：（010）88379604
购书热线：（010）68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com

‖ 丛书前言

当下大数据技术发展变化日新月异，大数据应用已经遍及工业和社会生活的方方面面，原有的数据管理理论体系与大数据产业应用之间的差距日益加大，而工业界对于大数据人才的需求却急剧增加。大数据专业人才的培养是新一轮科技较量的基础，高等院校承担着大数据人才培养的重任。因此大数据相关课程将逐渐成为国内高校计算机相关专业的重要课程。但纵观大数据人才培养课程体系尚不尽如人意，多是已有课程的“冷拼盘”，顶多是加点“调料”，原材料没有新鲜感。现阶段无论多么新多么好的人才培养计划，都只能在20世纪六七十年代编写的计算机知识体系上施教，无法把当下大数据带给我们的新思维、新知识传导给学生。

为此我们意识到，缺少基础性工作和原始积累，就难以培养符合工业界需要的大数据复合型和交叉型人才。因此急需在思维和理念方面进行转变，为现有的课程和知识体系按大数据应用需求进行延展和补充，加入新的可以因材施教的知识模块。我们肩负着大数据时代知识更新的使命，每一位学者都有责任和义务去为此“增砖添瓦”。

在此背景下，我们策划和组织了这套大数据管理丛书，希望能够培养数据思维的理念，对原有数据管理知识体系进行完善和补充，面向新的技术热点，

提出新的知识体系 / 知识点，拉近教材体系与大数据应用的距离，为受教者应对现代技术带来的大数据领域的新问题和挑战，扫除障碍。我们相信，假以时日，这些著作汇溪成河，必将对未来大数据人才培养起到“基石”的作用。

丛书定位：面向新形势下的大数据技术发展对人才培养提出的挑战，旨在为学术研究和人才培养提供可供参考的“基石”。虽然是一些不起眼的“砖头瓦块”，但可以为大数据人才培养积累可用的新模块（新素材），弥补原有知识体系与应用问题之前的鸿沟，力图为现有的数据管理知识查漏补缺，聚少成多，最终形成适应大数据技术发展和人才培养的知识体系和教材基础。

丛书特点：丛书借鉴 Morgan & Claypool Publishers 出版的 Synthesis Lectures on Data Management，特色在于选题新颖，短小精湛。选题新颖即面向技术热点，弥补现有知识体系的漏洞和不足（或延伸或补充），内容涵盖大数据管理的理论、方法、技术等诸多方面。短小精湛则不求系统性和完备性，但每本书要自成知识体系，重在阐述基本问题和方法，并辅以例题说明，便于施教。

丛书组织：丛书采用国际学术出版通行的主编负责制，为此特邀中国人民大学孟小峰教授（email：xfmeng@ruc.edu.cn）担任丛书主编，负责丛书的整体规划和选题。责任编辑为机械工业出版社华章分社姚蕾编辑（email：yaolei@hzbook.com）。

当今数据洪流席卷全球，而中国正在努力从数据大国走向数据强国，大数据时代的知识更新和人才培养刻不容缓，虽然我们的力量有限，但聚少成多，积小致巨。因此，我们在设计本套丛书封面的时候，特意选择了清代苏州籍宫廷画家徐扬描绘苏州风物的巨幅长卷画作《姑苏繁华图》（原名《盛世滋生图》）作为底图以表达我们的美好愿景，每本书选取这幅巨卷的一部分，一步步见证和记录数据管理领域的学者在学术研究和工程应用中的探索和实践，最终形成适应大数据技术发展和人才培养的知识图谱，共同谱写出我们这个大数据时代的盛世华章。

在此期望有志于大数据人才培养并具有丰富理论和实践经验的学者和专业人员能够加入到这套书的编写工作中来，共同为中国大数据研究和人才培养贡献自己的智慧和力量，共筑属于我们自己的“时代记忆”。欢迎读者对我们的出版工作提出宝贵意见和建议。

大数据管理丛书

主编：孟小峰

大数据管理概论

孟小峰　编著

2017 年 5 月

异构信息网络挖掘：原理和方法

[美] 孙艺洲（Yizhou Sun） 韩家炜（Jiawei Han） 著

段磊　朱敏　唐常杰　译

2017 年 5 月

大规模元搜索引擎技术

[美] 孟卫一（Weiyi Meng） 於德（Clement T. Yu） 著

朱亮　译

2017 年 5 月

大数据集成

[美] 董欣（Xin Luna Dong） 戴夫士·斯里瓦斯塔瓦（Divesh Sriva-tava） 著

王秋月　杜治娟　王硕　译

2017 年 5 月

短文本数据理解

王仲远　编著

2017 年 5 月

个人数据管理

李玉坤　孟小峰　编著

2017 年 5 月

位置大数据隐私管理

潘晓　霍峥　孟小峰　编著

2017 年 5 月

移动数据挖掘

连德富　张富峥　王英子　袁晶　谢幸　编著

2017 年 5 月

云数据管理：挑战与机遇

[美] 迪卫艾肯特·阿格拉沃尔（Divyakant Agrawal） 苏迪皮托·达斯（Sudipto Das） 阿姆鲁·埃尔·阿巴迪（Amr El Abbadi） 著

马友忠　孟小峰　译

2017 年 5 月

译者序

随着物联网、社交网络、移动互联网等新兴技术和服务的快速普及与应用，数据以前所未有的速度不断增长，人类进入了大数据时代。数据规模的海量性、数据种类的多样性以及数据产生速度的快速性等特点给数据管理带来了巨大挑战。为实现对大规模数据的有效管理，云数据管理技术应运而生。

云数据管理虽然已有十余年的发展历程，但仍存在诸多挑战和发展机遇。本书以面向数据存储和服务于互联网应用的云数据管理系统为主要对象，描述了其中存在的若干关键性挑战。本书共 7 章，第 1 章介绍了云计算、云数据管理的基本概念，对其中面临的关键挑战进行了概述，并描述了本书的组织结构；第 2 章主要介绍了分布式数据管理的相关知识，包括分布式系统、P2P 系统、并发控制和分布式数据恢复等；第 3 章对云数据管理的早期研究工作进行了描述，包括不同的键 – 值存储系统在数据模型、数据分布和容错等方面的区别，以及 Bigtable、PNUTS 和 Dynamo 这三个有代表性的键 – 值存储系统的特点；第 4 章介绍了托管数据的事务问题，包括数据托管模式、托管数据的事务执行、数据存储和复制等内容；第 5 章主要介绍了分布式数据事务相关技术；第 6 章讨论了云数据管理中的多租户技术，包括多租户模型、云中的数据库弹性以及云中数据库负载的自动控制；第 7 章对相关经验教训进行了总结，并指出了未来的主要研究方向。

本书主要由马友忠负责翻译，孟小峰负责统稿和审校。本书于2016年9月译出初稿，责任编辑关敏对初稿进行了认真审核，张瑞玲、刘栋、贾世杰、张永新等也认真阅读初稿，给出了许多宝贵的修改意见。之后由孟小峰、马友忠根据责任编辑和同事提出的意见，逐章进行修改和完善。最后于2017年1月完成定稿。

本书译词主要遵从教科书及相关学术著作、科研论文中的习惯用法，并参考《计算机科学技术名词》等典籍。由于译者能力有限，译文中难免有不当之处，恳请读者批评指正并不吝赐教。如有任何建议或意见，敬请发邮件至ma_youzhong@163.com。

马友忠

2017年1月于洛阳

前 言

大数据和云计算是研究文献和主流媒体中大量使用的两个术语。当我们走进云计算和数据洪流的时代，经常被问到的一个问题是：云数据管理中的新挑战是什么？本书就是由我们寻求回答这个问题发展而来，并使我们自己对这一问题有了更为深入的理解。本书首先介绍了一些初步的综述性论文，这些综述论文总结了适合键－值存储系统的主要设计原则，这些系统如谷歌的 Bigtable、亚马逊的 Dynamo 和雅虎的 PNUTS，通过在一个数据中心或者有可能在世界不同地方的多个数据中心中部署成千上万台服务器来达到前所未有的规模。由于这一领域引起了学术界和工业界越来越多的研究人员的关注，该领域从键－值存储进一步发展到支持更丰富功能的可扩展数据存储，如事务或除简单键－值模型之外的模式。因此，我们将 3 个系统的简单综述在新加坡举办的 VLDB 2010 会议和在瑞典乌普萨拉举办的 EDBT 2011 会议扩展成一个 3 小时长的教程。后来又有很多相关资料的介绍，因为这些教程以及我们对该问题的理解也随时间的推移发生了改变。其间也提出了更多的系统。本书对我们这些年课程的学习以及来自于我们讲座的很多有趣的讨论进行了总结。

与传统数据管理时代事务处理与数据分析系统之间的划分一样，云数据管理也有一个类似的划分。一种是面向数据存储和服务于互联网应用的系统。这些系统与经典的事务处理系统类似，尽管有很多不同之处。另一种是数据

分析系统，类似于数据仓库，通过分析大量数据来从中获得知识和智能。随着企业不断地搜集用户数据，并对来自于多种数据源的数据进行合并，基于MapReduce的系统，如Hadoop及其生态系统，使得数据分析和数据仓库更加大众化。云数据分析方面有几十个开源产品和数百篇相关领域的研究论文，已经成为一个热门的研究领域。因为企业试图从它们的数据库中获得新的见解，从而取得竞争优势，该领域会得到进一步扩展。

我们的研究、分析和调查主要关注于第一类系统，即数据管理和存储系统。因此，本书也主要关注这些系统。本书将深入探讨在设计这些更新密集型系统中存在的挑战，这些更新密集型系统必须对访问数据库小部分数据的查询和更新提供快速响应。在该类中，我们进一步将研究划分成两类系统。在第一类中，挑战在于对系统进行扩展，从而服务于拥有几千个并发请求和数百GB到数百TB频繁访问数据的大型应用。第二类包括这样一种情况，云服务提供商必须有效地服务于数十万个应用程序，每个应用程序的查询负载和资源需求都比较少。

致谢

本书源自于几年前我们试图更好地理解云数据管理设计领域的愿望。结果就有了我们对该设计领域的不断深入的理解。这得益于我们周围有很多人提供了帮助，人数太多，以至于这里无法一一列出。但是，我们想借此机会感谢那些在本书中发挥了重要作用的人。

首先，我们想感谢编辑M. Tamer Özsu，他给了我们写这本书的机会，并在整个过程中为我们提供了持续的支持和反馈。他认真阅读了大量的早期草稿，并给出了很多意见和修正，大大完善了本书。Diane Cerra作为我们的出版商Morgan & Claypool的执行编辑，为我们提供了必要的行政支持。没有来自Tamer和Diane的帮助与支持，本书将无法出版。

本书中的大部分材料都以不同的形式在世界各地的不同地点呈现过。在这

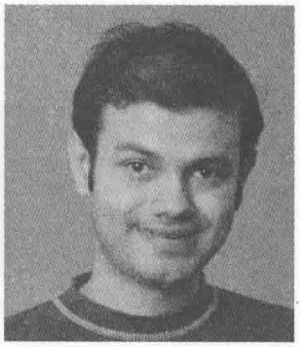
些演示过程中，我们收到了许多与会者的反馈，这些反馈直接或间接地改善了我们的演示，并经常会给我们提供不同的角度。我们非常感谢所有提供这些慷慨反馈的人。我们也从与 Shyam Anthony、Philip Bernstein、Selcuk Candan、Aaron Elmore、Wen-syan Li、Klaus Schauser 和 Junichi Tatemura 的大量讨论中获益匪浅，在此对他们表示感谢。我们还要感谢 2008 ~ 2012 年间学习研究生课程（CMPSC 271 和 CMPSC 274）的所有研究生的贡献。

最后，我们要感谢我们各自的家庭，他们容忍我们为准备本书和相关资料而花费了无数个小时。没有他们的一贯支持和理解，本书也不会有面世的一天。

Divyakant Agrawal、Sudipto Das 和 Amr El Abbadi

作者简介 ‖

迪卫艾肯特·阿格拉沃尔（Divyakant Agrawal）加州大学圣塔芭芭拉分校计算机科学系教授。主要研究方向包括数据库系统、分布式计算、数据仓库和大规模信息系统。他是 ACM 和 IEEE Fellow，在数据库系统、分布式系统、多维索引、数据仓库和云数据管理等领域发表论文 300 余篇。曾任多个国际会议、论坛的程序委员会委员，1993 至 2008 年，任《分布式和并行数据库期刊》（Journal of Distributed and Parallel Databases）编辑，2003 至 2008 年，任《 VLDB Journal 》编辑。他是 ACM SIGMOD 2010 程序委员会主席，多次担任 ACM SIGSPATIAL 会议的大会主席。目前担任《 Journal of Distributed and Parallel Databases 》的主编，ACM TODS 和 IEEE TKDE 编委，VLDB 基金会的受托人。在加州大学圣塔芭芭拉分校工作超过 25 年，培养了 30 多位博士研究生。荣获加州大学圣塔芭芭拉分校杰出指导导师奖。



苏迪皮托·达斯（Sudipto Das） 微软研究院极限计算组（eXtreme Computing Group）研究员。于加州大学圣塔芭芭拉分校获得计算机科学博士。研究兴趣广泛，主要包括可扩展数据管理系统和分布式系统。其研究跨多个领域，如云计算平台的可扩展事务处理系统、针对大数据的高级数据分析系统和多租户数据库系统。在众多著名的数据库相关期刊、会议（如 SIGMOD、VLDB、ICDE、CIDR、MDM 和 SoCC）上发表过著作。在云计算和大数据领域做过多次培训。曾荣获加州大学圣塔芭芭拉分校 2012 年 Lancaster 论文奖、CIDR 2011 最佳论文奖、MDM 2011 最佳论文奖第二名、2012 杰出论文奖，2011 加州大学圣塔芭芭拉分校优秀学生奖和 2006 年 TCS-JU 最佳学生奖。

阿姆鲁·埃尔·阿巴迪（Amr El Abbadi） 加州大学圣塔芭芭拉分校计算机科学系教授。埃及亚历山大大学计算机科学学士，康奈尔大学计算机科学硕士、博士。2007 至 2011 年，担任加州大学圣塔芭芭拉分校计算机科学系主任。他是 ACM 和 AAAS Fellow。曾任多个数据库期刊（包括 VLDB Journal）编辑，多个数据库和分布式系统会议（包括 VLDB 2010、SIGSPATIAL GIS 2010 和 SoCC 2011）的程序委员会主席。2002 至 2008 年，任 VLDB 基金会委员。2007 年，荣获 UCSB Senate 杰出导师奖。在数据库和分布式系统领域发表超过 275 篇论文。

目 录