



数据存储 技术与实践

查伟 著

DATA STORAGE
THEORY AND PRACTICE



清华大学出版社



数据存储 技术与实践

DATA STORAGE THEORY AND PRACTICE

查伟 著

清华大学出版社
北京

内 容 简 介

本书试图以时间轴的方式描述自 20 世纪 50 年代 IBM 的第一个硬盘出现以来, 整个存储工业及技术的演进。本书内容聚焦于存储的技术演进, 除基本技术的描述外, 重点放在数据中心在云计算及大数据环境下存储新技术出现的背景、新技术架构、原理及实现, 同时本书也把一部分内容放在具体的存储解决方案及实际运维管理中。

作为存储知识技术类图书, 本书可以用于高校信息管理专业的教材或参考书, 也非常适合信息管理专业技术人员、IT 经理人等专业人士阅读参考, 为技术专家、高校学生以及更多愿意了解和投入到存储浪潮中的人们搭建一座桥梁。

本书封面贴有清华大学出版社防伪标签, 无标签者不得销售。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

数据存储技术与实践 / 查伟 著. —北京: 清华大学出版社, 2016
ISBN 978-7-302-44752-8

I. ①数… II. ①查… III. ①数据存储 IV. ①TP333

中国版本图书馆 CIP 数据核字(2016)第 182603 号

责任编辑: 陈 莉 高 岫

封面设计: 周晓亮

版式设计: 方加青

责任校对: 牛艳敏

责任印制: 王静怡

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社 总 机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 三河市中晟雅豪印务有限公司

经 销: 全国新华书店

开 本: 175mm×226mm 印 张: 22.5 字 数: 476 千字

版 次: 2016 年 9 月第 1 版 印 次: 2016 年 9 月第 1 次印刷

印 数: 1~3500

定 价: 58.00 元

产品编号: 063819-01

编委会

(排名不分先后)

唐 鲲 刘嘉程 钟 颢 邓晓阳 姚建业 欧阳戟 张 巍
赖俊臣 雷小松 张程伟 罗庆超 方 新 王道辉 张 秦
赵 波 明 亮 张 雷 董 浩 王克俊 李夫路 梁 胜
段翰聪 楼新远 陈 浩 李 亿 彭 勃 洪 慧 胡善勇

今天，当我们身处存储变革的风暴眼中。俗话说，温故而知新，回顾过去多年来存储技术的演变与发展，梳理存储技术的创新之道，就可以更好地理解和把握当前企业用户的需求，清楚用什么样的方法和工具能够更好地帮助企业从容应对云计算、大数据、移动互联、社交网络带来的新挑战。本书从传统企业存储概念和技术讲起，继而又延伸到云计算、大数据，以及企业用户最关注的数据库存储领域，让读者可以对存储技术的前世今生有一个全面掌握，也有利于指导企业用户当前及未来的存储应用实践。

存储的价值不仅在于“存”这一基础功能。在云计算、大数据已经逐渐普及的今天，不但有软件定义网络，还有数据定义的软件。我们之所以还激情满满地谈论存储，其实关注点并不在存储的性能、功能等指标上，而在于数据的处理、数据挖掘、数据的价值乃至记忆等。

数据是存储的核心。只有看清了这一点，我们才能更加游刃有余地使用分布式存储、软件定义存储、融合与超融合、全闪存等创新的技术和架构，改造和优化存储架构，并使之成为数据价值挖掘的有利工具和平台。本书在详细阐述存储基本概念的基础上，着重介绍云计算、大数据等新应用的需求，存储架构的演进思路和实现方式，这对于推动存储云化，以及大数据存储的广泛应用具有非常积极的意义。

互联网和移动互联网作为信息技术发展的新常态，企业的数字化转型成了当务之急。在这种背景下，企业更应该重新审视自身业务发展与IT之间的关系。以前，IT被当成企业的成本中心，或者后端支撑平台。现在，IT是企业业务发展的直接驱动因素，其本身就是企业业务的重要组成部分。

当前，企业正处于从办公自动化、流程自动化向大数据平台过渡的重要转折时期，传统的企业数据库如何云化才能保证效率的最大化？传统的SAN存储与云存储之间是相互替代还是融合？如何进一步优化大数据存储平台，使之更好地配合数据挖掘和智能分析？这一切都需要新的答案。本书试图探索并给出在云计算、大数据等需求日益迫切的情况下，存储架构

的创新、实施和优化的解决办法。

推动存储变革的因素有很多：从业务需求的角度分析，云计算、大数据的出现，将颠覆传统的技术架构和商业模式，IT以服务的模式交付，对存储的性能、可用性、可扩展性以及经济性等都提出了更高的要求；从技术演进的角度分析，硬件将深度解耦，闪存将取代机械式硬盘，并逐渐成为主流，在这种情况下，融合与超融合解决方案将加快落地，而针对闪存应用的持续优化将成为存储厂商的一项核心工作；从存储市场格局的演变分析，在软件定义、分布式架构、开源技术等因素的共同作用下，存储公司的创新脚步将加快，生态圈经济的影响日益加深，新的存储市场格局正在酝酿之中，而存储的主旋律是开放、开源。

本书将技术理念与应用实践相结合，揭示了存储技术变革的脉络，对于想了解最新存储技术发展，以及如何将创新的存储技术应用于企业实践的人来说大为有益。

中国科学院院士 欧亚科学院院士
中国人工智能学会理事长 中国云计算专家委员会主席
李德毅
2016年5月

在大数据时代，越来越多的人和企业意识到数据作为一种资产的重要性。对数据进行挖掘，以便从中发现更多价值正在成为热门话题；国家“十三五”规划纲要里也提出要把大数据作为基础性战略资源，加快推动数据资源共享和开发应用，助力产业转型升级和社会治理创新。可以说，如何更好地存储、管理和使用数据，正在受到更多人的关注。

与此同时，我们也注意到存储产业正在发生着一场变革。

首先，新的存储产品形态和技术架构开始普及。一方面闪存和固态硬盘(SSD)价格持续降低，成为更多人的选择；另一方面，存储技术也在快速发展，新的接口(如PCIe等)、协议(如NVMe等)、架构(如3D XPoint等)开始出现，使存储设备的带宽、性能和存储容量得到极大提升，为更高性能的服务器设计和更大规模的数据中心应用做好了准备。

其次，新的存储系统以及建立在其上的各种存储架构理念开始出现。除传统的存储阵列之外，全闪存阵列、混合存储或者Server SAN等存储系统开始涌现，从不同的角度满足大规模存储和上层应用对于数据的快速读取需要；融合存储、超融合存储、软件定义存储(SDS)等各种理念，各自都能看到许多厂商的背后支持。

此外，在软件技术和硬件发展的推动下，云计算技术和产业日趋成熟，给包括存储在内的传统硬件产业带来巨大的冲击，在这种背景下，在设计信息系统架构时，尤其是保存企业宝贵的数据资产的存储系统时，如何选择服务提供商，选择何种架构，就成了企业迫切关注的问题。

这个变革期对于中国来说无疑是良好的机遇。一方面，对于新技术的理解和应用，可以给中国企业的信息化建设带来有效帮助；另一方面，在国外传统存储厂商垄断市场多年之后，新技术的出现有利于新的市场竞争格局的塑造，在同一起跑线下，中国存储企业有能力，也应该有信心树立新的标准，推出代表性的产品，参与到国际市场的竞争中去。

存储技术是企业发展的核心，也是在整个国家信息化发展的这个节点上，需要有一本书来帮助读者详细梳理整个存储行业的演进脉络，理清各种概念，使读者能明白在具体的场景

下，有着什么样的选择，以及应该做出何种选择。

所以本书从四个方面对存储技术和实践进行了解读，包括“企业存储概述”、“云存储”、“数据库”和“大数据”。从整体上看，其是一个从基础设施到软件系统，再到上层应用的完整体系。本书不仅勾画了长幅的存储产品乃至技术架构的发展路径，能让读者清晰地了解到这些技术为什么如此发展，以及发展演化之后能应对当前的哪些问题挑战；而且能帮助读者把这些概念具体对应到自己的实际业务应用中去。

本书能以翔实的架构，深入浅出地完成对数据存储技术和产业的梳理工作，背后是编写组雄厚的技术背景和实践经验。编写组的成员，不仅在技术上一直紧跟时代发展的潮流，而且在具体实践过程中也通过产品的设计和推广，推动着中国乃至世界存储产业的发展。可以说，无论是理论分析还是应用实践，本书的编写组都有足够的专业性和权威性。通过本书，我们也能大致研读到作为国内研发投入最大、研发队伍最大的高技术企业，华为在数据存储方面有着怎样的技术理念和业务观点。

值得注意的是，编写组在不同部分都会以一些典型的应用场景对相关内容进行收尾总结，具有较高的参考价值，相信能帮助读者开拓思路。

综上所述，本书适合所有想要了解数据存储技术，或者希望利用数据存储技术发掘更多价值的读者，即便是存储业内的专业人士，也能从本书的阅读中有所收获。期待有更多的人加入到存储行业中来，为中国成为存储行业的引领者而努力。

中国工程院院士
倪光南
2016年5月

物质、能源和信息被视为人类社会生存和发展的三大战略资源。以数字电子计算机发明为标志，信息科技蓬勃发展了70年，而又以20世纪90年代中期互联网商用为标志，近20年的发展尤其迅猛。信息科技及其应用(信息化)广泛并深刻地影响和改变了人类社会，而且这种作用正在加强，无处不在的信息技术的深度应用甚至将重构人类社会！

回顾信息技术发展及信息化进程，过去的信息化有两次大的浪潮，大致以20世纪90年代中期为分界，此前是以单机应用为主要特征的数字化阶段(信息化1.0)，此后是以联网应用为主要特征的网络化阶段(信息化2.0)。随着信息技术的不断低成本化和互联网及其延伸所带来的无处不在的信息技术应用，第三波浪潮信息化浪潮正扑面而来，我们正在开启以数据的深度挖掘及融合应用为主要特征的智慧化阶段(信息化3.0)。

在信息化3.0时代，作为我们获取信息、汲取知识的素材和源泉，数据正成为与我们的工作和生活密不可分的(如阳光、空气和水般)、既具体又重要的“元素”，我们每一个人既是数据的创造者，也是数据的消费者，毫无例外，无时无刻。近年来，随着社会文化日益繁荣、全球经济持续发展以及信息科技不断获得突破，数据的产生、处理、分析和消费高度活跃，围绕数据诞生的新概念和新术语也层出不穷，如数据交易、数字经济、数据变革、大数据等。其中，最热、最火的概念当属大数据，自其出现以来，已产生了巨大的全球效应，大数据蕴藏的巨大潜力被广泛认知，成为全球关注的热点和各国政府的战略选择，正引发新一轮的信息化建设热潮。工业和信息化部苗圩部长指出：“(大数据)将创造下一代互联网生态、下一代创新体系、下一代职业形态以及下一代社会治理结构。”

如今，没有“数据”参与的社会或经济活动，已是不可想象之事。数据源于人类认识自然、改造自然，并推动社会经济发展的各类活动。信息技术推进数据的规范化和格式化，使数据不断升华为信息和知识，最终成为全人类的“数据宝藏”，又被重新投入到新一轮的各类社会经济活动中，创造出更大的价值。

要让数据全面、充分地发挥出其潜在价值，一个看似简单却必要的前提是：数据需要被

妥善保管在可靠、可信、可管理的平台中，数据的完整性和可信度需要得到确保，数据应该能够被随时随地且方便地存取。本书正是针对这些需求和问题而撰写的。该书内容丰富、层次清晰，既回顾了数据存储行业的历史及发展，也探讨了未来的发展趋势和方向。全书以生动的细节、翔实的素材，给读者展示了数据存储行业的前世今生，值得一读。

该书有两个值得称赞的特点：一是将聚焦点放在发现、分析存储行业所存在的问题上，并通过给出解决方案和过程，反映出存储行业近年来所发生的巨大变化；二是围绕当前存储技术的热点话题，在侧重阐述数据存储架构方法的同时，注重结合存储技术实践，使其成为一本可用于企业实际业务的“工具书”和“参考书”。

当今，随着材料技术和信息技术的快速发展，存储技术不断推陈出新，存储行业持续向前发展，希望本书编写组能对存储技术的国内外发展不断总结归纳，将知识和经验源源不断地奉献给存储行业，并惠及专业读者。更希望我国科研机构和企业的专业人士锐意进取，能在世界舞台上推进、引领存储技术的创新变革。

愿本书的出版，不是一个句号，而是一个推进我国乃至世界大数据存储技术和系统创新发展的过程的新起点。是为序。

发展中国家科学院院士
中国科学院院士 梅宏
2016年5月

站在2016年年初的时间节点，我们明显感觉到了存储行业正在发生着一场深刻而又影响广泛的革命。这场革命的起因来自多个方面，其中企业IT部门越来越多地开始计划甚至已经部署的云计算以及随着成本的快速下降而快速普及的Flash存储是推动存储变革的两个最重要推手。存储行业方方面面的重大事件及各大咨询机构发布的预测信息也向这个行业的业内人士及用户展示了这场革命已经到来，我们有必要为这场革命做出充分而又周全的准备及部署：①从巨头们的公司动态看，2015年10月DELL与EMC 670亿美元的收购事件无疑是存储行业历史上最大的事件之一，行业巨头已经明显感觉到了云计算的巨大冲击而采取了抱团取暖、以云供应商及服务商的姿态面对市场与客户；②从业内权威咨询机构的预测看，Wikibon、Gartner、IDC都在其后五年预测中定义了诸如Converge Storage、Hyper Converge Storage类的新形态，并预测这类新形态的增长将远高于传统的外置存储产品，这是前几年在技术上热炒的SDS概念在企业IT的正式落地的预兆；③从近两年各存储巨头推出的产品形态看，全Flash存储、超融合存储、SDS，乃至直接面向第三平台环境下的存储是主要特征，这类存储形态会更强化向上通过软件定义提升在云计算环境下的灵活性、弹性与迅速部署、迅速上线，以及向下构筑更为高效的Flash介质友好与优化的存储层。

因此，我们一直有个想法，通过一本书的形式，向我们的读者较为详细地解析整个存储行业的演进脉络。站在今天的节点上看，它是从哪里来的，原来是为解决什么问题而生的，现在又发生了哪些变化，存储系统需要做出什么样的变化来应对这些变化，在侧重技术的同时，理清这些技术、产品要解决的存储问题，理清技术随着业务诉求变化而变化的脉络。

本书共分为四篇，具体如下所示。

第一篇 企业存储概述。在本篇，通过回顾存储工业60年的演进历史，尤其是自20世纪90年代起开始的近代外置存储阵列的历史，我们试图站在今天的历史节点上给读者呈现一幅较生动的存储工业的演进画面，让读者较为清晰地理清存储演进的原始诉求，通过怎样的技术架构及产品解决这些诉求，进而推动工业的发展；通过历史的分析，为读者理解及思考后

面在新介质及云环境下存储工业的进一步研究打下基础。

第二篇 云存储。随着IT技术的发展，特别是以Amazon、Microsoft、Google为代表的互联网公有云应用，让用户可以像使用水、电、气等基础设施那样的方式使用IT技术，极大地简化客户安装、部署、运维等工作，让客户的IT应用可以快速、敏捷地提供服务，本篇较为系统地介绍了云计算环境的典型分布式存储系统要解决的问题以及其技术架构。

第三篇 数据库。本章以企业到云化演进为出发点，探索数据库从纵向扩展架构向分布式扩展架构的需求变化以及其后对存储产生的影响。

第四篇 大数据。应用集中上云，带来数据进一步集中，这为产品产生智能的数据以及个人决策打下了基础，互联网的大数据分析就此产生；大数据的核心是数据的高效、批量处理，这需要高效的存储系统的支撑，而传统存储基本是为结构化数据而生的，即使是像NAS这样面向非结构化存储，也是为人与人之间的交换和共享而生的，本章将主要介绍大数据存储的挑战及应对技术、产品。

感谢成书过程中各主要成员全力以赴，感谢华为存储与我们共事多年的兄弟姐妹，我们患难与共，一起实现共同的梦想。感谢清华大学出版社在本书撰写过程中提出的宝贵意见。

由于部分技术涉及商业机密，加上笔者水平有限、时间较紧，所以书中难免存在谬误，很多技术点涉及的细节描述得还不够详尽，恳请读者批评指正，可将任何意见和建议发送到邮箱oceanstor@huawei.com。

《数据存储技术与实践》编写组
2016年3月于深圳

第一篇 企业存储概述**第 1 章 存储工业演进历史 3**

- 1.1 高端存储系统的演进 4
- 1.2 中端存储兴起 6
- 1.3 SAN外置存储阵列 7
- 1.4 NAS外置存储阵列 9
- 1.5 块、文件统一存储系统 10
- 1.6 全Flash外置存储阵列 10
- 1.7 Server SAN融合存储 11

第 2 章 存储产品架构演进趋势 13

- 2.1 高端存储架构演进 14
- 2.2 中端存储架构演进 15

第 3 章 企业存储产品架构及关键技术要素简介 17

- 3.1 存储软件技术的发展与演进 18
- 3.2 高端存储产品架构及关键技术要素简介 18
- 3.3 中端存储产品硬件及软件架构分析 23
- 3.4 全Flash架构分析 28
- 3.5 企业存储关键技术及特性 33
- 3.6 案例：OceanStor V3融合存储 39

第二篇 云存储

第4章 云平台及OpenStack开源架构提供的存储服务纵览 47

- 4.1 云平台的存储介绍 48
- 4.2 云平台的存储分类 48
- 4.3 云平台的存储标准 49
- 4.4 OpenStack开源领域的云平台全景 51
- 4.5 本章小结 57

第5章 分布式块存储服务 59

- 5.1 分布式块存储介绍 60
- 5.2 分布式块存储功能和特点 62
- 5.3 业界分布式块存储技术 66
- 5.4 本章小结 78

第6章 分布式对象存储服务 79

- 6.1 分布式对象存储介绍 80
- 6.2 分布式对象存储功能和特点 80
- 6.3 业界分布式对象存储技术 81
- 6.4 本章小结 90

第7章 分布式队列服务 91

- 7.1 分布式队列介绍 92
- 7.2 分布式队列功能和特点 92
- 7.3 业界分布式队列功能和特点 93
- 7.4 本章小结 96

第8章 分布式存储系统的可靠性 97

- 8.1 分布式存储系统可靠性介绍 98
- 8.2 分布式存储系统可靠性关键技术 103
- 8.3 开源实现分析 109

8.4 本章小结 111

第 9 章 分布式存储系统的管理及运维 113

9.1 分布式存储系统管理运维介绍 114

9.2 分布式存储对管理的挑战 114

9.3 分布式存储的管理 116

9.4 分布式存储的维护 118

9.5 本章小结 122

第 10 章 面向云平台的存储安全架构体系 123

10.1 面向云平台的存储安全风险分析 124

10.2 存储系统安全解决方案设计原则 126

10.3 存储系统安全解决方案架构 127

10.4 存储安全解决方案描述 128

10.5 本章小结 131

第 11 章 云平台架构下的企业存储实践 133

11.1 华为看 SDS 134

11.2 SDS 与华为 OceanStor DJ 134

11.3 OceanStor DJ 135

11.4 本章小结 139

第 12 章 案例 141

12.1 银行数据中心：支撑业务快速上线 142

12.2 公安厅警务云：提供自动化存储服务 143

12.3 智慧城市时空云平台：提供便捷数据服务 143

第三篇 数据库

第 13 章 数据库存储技术的演进 149

13.1 数据库系统概述 150

13.2	数据库存储技术的发展	154
13.3	数据库存储的主要形态	160
13.4	数据库存储的关键技术	161
13.5	数据库存储的备份容灾	162
13.6	存储介质对未来数据库存储的影响	164
13.7	本章小结	165
第 14 章 传统数据库系统架构 167		
14.1	传统数据库技术和架构的发展演进	168
14.2	OLTP数据库技术及架构	188
14.3	OLAP数据库技术及架构	193
14.4	MPP DB数据库技术及架构	198
14.5	传统数据库业务发展趋势	203
14.6	本章小结	205
第 15 章 数据库和存储的融合技术与实践 207		
15.1	数据库和存储融合技术概述	208
15.2	数据库和存储融合技术的发展	209
15.3	数据库和存储融合的主要途径	216
15.4	数据库和存储融合技术的实践	217
15.5	数据库和存储融合技术的发展趋势	218
15.6	本章小结	219
第 16 章 新兴数据库技术与实践 221		
16.1	新兴数据库技术和架构发展演进	222
16.2	NoSQL数据库技术及架构	222
16.3	NewSQL数据库技术及架构	226
16.4	内存数据库技术及架构	237
16.5	分布式数据库技术发展趋势	250
16.6	本章小结	251

第 17 章 案例 253

- 17.1 案例一：双活场景下与数据库的案例 254
- 17.2 案例二：中型规模OLAP类型Oracle数据库案例 262

第四篇 大数据

第 18 章 大数据的前世今生 271

- 18.1 大数据概念的由来 272
- 18.2 什么是大数据 276
- 18.3 我们能用大数据来做什么 278
- 18.4 大数据生态体系及其发展趋势 279
- 18.5 本章小结 282

第 19 章 大数据对存储系统的挑战 283

- 19.1 大容量(Volume)的挑战 284
- 19.2 数据多样化(Variety)的挑战 294
- 19.3 性能(Velocity)问题的挑战 295
- 19.4 本章小结 299

第 20 章 大数据存储技术实践 301

- 20.1 数据库领域 302
- 20.2 Apache Hadoop的存储组件 305
- 20.3 基于内存的大数据组件 315
- 20.4 商用Hadoop版本的独特贡献 320
- 20.5 数据湖泊 327
- 20.6 本章小结 335

第 21 章 最佳应用实践案例 337

- 21.1 招商银行利用大数据进行业务创新 338
- 21.2 交通管理领域大数据研判系统 339
- 21.3 电信运营商运用大数据优化业务 341