

完全手册丛书

存储网络
权威性参考手册

Storage Networks: The Complete Reference

存储网络 完全手册

[美] Robert Spalding 著 郭迅华 等译

快速掌握存储网络的
内涵和精髓

展示了各种新兴的技术
标准，如 iSCSI 和
InfiniBand，介绍了存
储网络的未来发展方向

深入理解 SAN 和
NAS 中的品牌兼容
性、存储管理以及基
本 I/O 负载规划问题



电子工业出版社
Publishing House of Electronics Industry
<http://www.phei.com.cn>

完全手册丛书

存储网络完全手册

Storage Networks: The Complete Reference

[美] Robert Spalding 著

郭迅华 等译

电子工业出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

本书是关于存储网络的一本全面指南,全书共分为七个部分,25章。涵盖了包括SAN、NAS、光纤通道、文件系统、虚拟化、数据库存储等各种存储网络技术。通过阅读本书,读者会深入理解存储网络在数据中心中的应用现状,并进而一览无余地领略当今的存储网络革命;本书也将帮助读者认识存储网络中的种种核心概念,包括体系结构、设备、连通性选择和数据组织手段等,并借助这些思想来有效地设计和实施各种复杂的存储网络;还将指导读者熟练地运用两种主要的存储网络模式,即NAS(Network Attached Storage,网络连接存储)和SAN(Storage Area Network,存储区域网络),并使用IP、RAID存储体系以及虚拟化要素对之进行集成。

本书适合于对存储网络感兴趣的专业人士。对于网络工程师、应用系统设计师和IT运营管理人员而言,本书是一本极具价值的参考书,有助于他们更好地理解、规划和维护存储网络。

Robert Spalding: Storage Networks: The Complete Reference. ISBN: 0-07-222476-2

Copyright © 2003 by The McGraw-Hill Companies, Inc.

Original language published by The McGraw-Hill Companies, Inc. All rights reserved. No part of this publication may be reproduced or distributed in any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Simplified chinese translation edition jointly published by McGraw-Hill Education (Asia) Co. and Publishing House of Electronics Industry. Copyright © 2004.

本书中文简体字翻译版由电子工业出版社和美国麦格劳-希尔教育出版(亚洲)公司合作出版。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有 McGraw-Hill 公司激光防伪标签,无标签者不得销售。

版权贸易合同登记号 图字: 01-2002-5890

图书在版编目(CIP)数据

存储网络完全手册/(美)斯波尔丁(Spalding, R.)著;郭迅华等译.-北京:电子工业出版社,2004.5
(完全手册丛书)

书名原文: Storage Networks: The Complete Reference

ISBN 7-5053-9891-1

I. 存... II. ①斯... ②郭... III. 信息存储-计算机网络-技术手册 IV. TP393.0-62

中国版本图书馆CIP数据核字(2004)第038999号

责任编辑:赵红燕 特约编辑:赵宏英

印刷:北京兴华印刷厂

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编:100036

经销:各地新华书店

开本:787×1092 1/16 印张:24 字数:614千字

印次:2004年5月第1次印刷

定价:38.00元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换;若书店售缺,请与本社发行部联系。联系电话:(010)68279077。质量投诉请发邮件至 zltz@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

译者序

自20世纪后期以来,信息技术以一种前所未有的力量冲击着当代商务世界和企业运作模式。尽管人们解读信息技术的角度和方法各不相同,但信息技术作为大势所趋的时代潮流已经得到了普遍接受。21世纪生产力发展水平的重要标志是信息技术的广泛应用,数字化已成为时代的标志。

随着IT应用在范围上的不断拓展和深度上的持续渗透,信息能力的运用以及信息资源的管理已经成为企业至关重要的核心任务之一。信息技术不但是企业赖以生存、发展并建立竞争优势的一种有力工具,同时也是推动企业自身变革和战略提升的一种不可替代的驱动力量。在企业中的三种主要运动形式(物流、活动流、信息流)中,信息流日益显现出了主导地位,因为数字化的企业经营模式使得物流和活动流最终也以信息流的形式为管理者所掌控。从这种意义上来说,信息不但是企业的资产,而且是最具根本意义的战略性资产之一。

数据是信息的基本存在形式,在现代信息系统中居于核心地位,因而,信息资源管理的基础是数据的管理。在数据的采集、存储、加工和运用等各个环节中,存储是一个支撑全局的基础性环节。对于数据的可用性、可靠性、完整性和安全性而言,强大的存储能力是必不可少的。随着应用的深入,企业对数据存储技术和数据管理方法的要求不断提升,这种需求推动着IT行业加快存储技术的推陈出新,并不断更新存储管理的理念。一方面,应用的积累使得企业中的数据量呈现出爆炸性的增长,数据的形式变得日益多样化,图像、音频、视频等多媒体数据进一步加剧了存储量的膨胀。同时,跨时空的高度可用性要求存储体系中具有良好的稳定性以及巨大的吞吐能力。另一方面,分布式信息系统应用的迅速发展进一步增加了存储体系的复杂性,促使IT人员在设计存储架构时必须对多方面的因素加以综合性的考察,并且在存储管理工作中也面临着更多的不确定性和风险。

在这样的形势下,传统的存储技术以及客户机/服务器存储模式已经不堪重负,一场新的存储革命呼之欲出。此时应运而生的存储网络初步显现了强大的威力,得到了企业和厂商的广泛关注。存储网络打破了传统模式中存储设备高度依赖于计算机的限制,以高性能网络为桥梁,将存储能力与计算能力分离开来。这种分离首先解开了束缚存储容量增长的缰绳,使得存储容量不再受到计算机内部体系结构以及服务器处理能力的制约,实现真正意义上的“海量”存储。其次,这种分离使得存储体系成为一种独立的基础设施平台,以整体形式为企业的信息系统应用提供全面的支撑,从而为信息的集成共享奠定了底层基础。第三,这种分离也使得企业的信息系统技术架构得到了进一步的层次化,有力地提升了灵活性和可扩展性,同时也有助于建立纵深的信息安全体系。此外,技术架构的变革也推动着IT管理体系的调整,促使企业内部IT管理方式向着更高效有序的分工协作方向发展。

存储网络的最重要的两种形式是网络连接存储(Network Attached Storage, NAS)和存储区域网络(Storage Area Network, SAN)。NAS模式建立在现有TCP/IP网络的基础上,实施过程较为简便,可以较为迅速地增加存储容量并提升数据共享能力,为初步改造企业存储模式提

供了一条较为便捷的途径。SAN则独立于传统网络之外,以高效率的光纤通道技术为存储建立专门的网络。以SAN为代表的存储网络具有现代数据存储所需要的高容量、高速度、高可用性、高可扩展性、集成性、远程虚拟存储等特性,并通过两个网络的分离充分保证了应用系统的效率。其未来发展的目标是可将系统监控、资源管理、系统配置、安全策略、高可靠性、容量计划及冗余管理等众多功能集于一身的集成式数据分发与检索架构解决方案。存储网络被视为第三次IT革命浪潮,在西方国家方兴未艾。近几年来,数据存储和数据管理问题也日益地为中国企业所重视,存储网络在中国也赢得了广泛的关注。

本书是关于存储网络方面的一本全面的指南。资深技术咨询顾问Robert Spalding先生凭借其对存储领域的深刻认识和丰富的实践经验,全面介绍了存储网络的概念、原理、技术、方案以及管理中的各种问题,其中所讨论的内容对于存储网络的规划、设计、实施和运行等各个环节都具有重要的指导意义。作为一名咨询顾问,作者在本书的写作中充分体现了面向实践的特点,既没有拘泥于技术细节,也没有一味地推崇新技术,而是致力于以一种全面、客观的视角对存储网络的技术特性及其应用和管理加以深入的剖析。其中,有关数据组织方法、工作负荷估算、商务持续性规划、容量规划、安全性管理等方面的探讨都是颇为独到的。这种面向实践、注重管理的写作风格使得本书不但可以用做存储和网络技术人员从事存储网络建设工作时的实用教材和参考资料,对于企业IT决策人员和数据管理人员,同样也是了解存储网络技术、掌握存储管理要点,从而建立高效数据管理策略的有力工具。

2003年,我与肖勇波、王征宇、文进等人共同翻译了John Vacca先生的《存储区域网络精髓》一书。该书出版之后,我陆续收到了一些读者的来信,对相关问题进行了探讨。这些读者中既有从事信息系统建设的IT实践者,也有从事信息技术研究的在校研究生。与这些读者的讨论使我进一步感受到存储网络已经产生的广泛影响以及难以估量的广阔应用前景。因而,我希望借助本书的翻译出版,为这场存储革命在中国的兴起和发展,再增添一份微薄的推动力量。同时我也期待着更多读者的来信交流,欢迎大家通过电子邮件(guoxh@em.tsinghua.edu.cn)与我联系。

本书由郭迅华主译,第七部分(附录)由肖勇波协助翻译。翻译的过程中我们与文进、王征宇进行了频繁的讨论并获得了大量极具价值的参考意见和建议,此外还得到了李飞、李志军、张林鹏、曹庆、任雪茵、刘晓玲、梁冰、冯锦锋、史锐、肖宽等人的大力协助,在此一并表示感谢。

由于译者水平有限,不足之处在所难免,望各位读者批评指正。

序 言

Sean Derrington

META Group 公司

基础设施策略部高级项目总监

越来越多的 IT 组织（IT Organization, ITO）将存储看做是为其现有及待建的应用系统赢得差别化竞争优势的战略性要素。许多 IT 组织都已经认识到了信息在企业中所发挥的关键性作用，并进而清楚地意识到，企业需要有牢不可破的存储基础设施来支持信息管理。我相信任何一个有远见的 IT 组织都不会再将存储产品的选择与服务器的采购混为一谈。这种观念上的分离有助于确存储解决方案不再被单纯地视为新建应用系统的附属品。对于多数组织而言，这将是一种最佳实践的开端，并且随着存储市场上技术、产品的日渐成熟，越来越多的 IT 专业人员深入认识存储的价值并掌握相关的技能，这一潮流也将继续向前推进。尽管目前大多数企业对 IT 方面的投资表现得极为谨慎，但是也有一些领先一步的 IT 组织正着手为存储解决方案构建技术上的基础，同时也在规划和厂商选择方面进行着战略性的准备。

总体说来，存储正在成为与 IP 网络和未来的应用基础设施相类似的一种服务模式。人们将从功能、价格以及复杂性等方面来对存储服务进行分级，并根据需要将之提供给应用系统或是业务流程使用。要想让 IT 组织以这种方式提供存储服务，存储（包括其各个方面，如磁盘、磁带、灾难恢复、容量规划）就必须被作为一种基础设施组件来看待和管理。基础设施组件应该能够在尽可能多的用户（例如应用系统和服务器）之中共用和重用，存储也不例外。这就要求存储在企业的战略规划中成为一个高优先级的考虑因素（而不是事后的补充）。更进一步而言，基础设施的设计和规划团队应该在企业的发展计划中占有一席之地，才能将技术转化为企业的价值。

一般说来，IT 体系结构的设计者所遵循的体系规范和相应的方法论面向的是三到五年（随着厂商和产品的不同而不同）的规划周期，也就是说，其主要考虑的是 36 个月的需求。这就给 IT 组织带来了一个挑战，即如何将存储基础设施的规划和设计方案转化为存储平台的实际运作。对于 IT 组织而言，这是一个关乎全局的关键成功因素，对于存储管理而言更是如此，因为存储产品的容量会持续不断地迅速发展，存储厂商也会持续不断地进行变革调整。可以预见，在组织结构上，到 2004 年，企业中的存储基础设施建设和存储运作的工作仍然会由同一个小组负责。然而更长远一点地看，到 2005 或 2006 年，这两项职责将会分离，形成两个明确独立的小组，就好像目前企业中系统、网络、应用基础设施及运作分别由不同的团队负责一样。基础设施规划的目标是：“确定在短期、中期及长期条件下为企业提供其所要求的应用服务等级时所需基础设施的范围、规模以及设计方案。”进一步而言，“信息系统的首要设计目标，是必须能够支持企业流程的快速变革。同时，用以支撑这种变革的系统和技术基础设施本身，也应该具有根据企业的要求调整、变化的能力。”

自动化存储网络包括存储区域网络（SAN）或是网络连接存储（NAS），其目的是要实现存储应用管理的合理化（减少其中不一致之处），同时实现这些功能的优化集成。在其实施过程中需要对各个层面上的存储软件、硬件以及专业服务进行调整，才能成功完成所需的存储基础设施建设。自动化存储网络将给企业带来敏捷性和生产效率的快速提升，同时使得IT组织能够以较低的成本实现对TB和PB级数据的高效管理。本书将详细介绍存储体系设计中的各种考虑因素、用以满足各种存储体系要求的组件，以及网络存储环境下的各种连接手段。

存储和应用的服务等级协议（SLA）

应用系统和服务器的配置取决于企业的需求，因而是千差万别的。因此，无论是在大型企业还是小型内部工作团队的存储问题中，存储服务质量的评估都是一个需要考虑的问题。到2004或2005年，数据中心中60%的服务器都将与网络存储相连接。与此同时，根据预测，70%以上的存储容量将实现网络化。虽然网络存储建设成本将有适度的增长，但其在管理效率、系统利用率、敏捷性、基础设施和人力资源的节约以及可扩展性等方面所带来的效益将是极为可观的。

共享外部存储资源的服务器也将有助于对存储容量进行网络化的规划，并在总体上提高存储容量的利用率（例如GB/TB级存储、光纤通道光纤网，以及对数据中心的物理空间利用具有重要意义的外部子系统等）。对于一些增长速度较快的应用，可以在达到容量临界值（60%~80%的使用率）时将更多的存储容量指派给相应的服务器。相对于一些内部的存储而言，在这一领域中引入企业级存储并进行容量规划，则具有更为重要的意义。一般而言，建议企业级存储以及其他具有竞争力的存储环境应该保持一个6~9个月的采购周期。总体上说，IT组织在制订计划时，可以假设硬件价格每年下降大约35%（即每季度8%~10%）。同时，根据目前经济全球化的现状来看，应当谨慎地安排长期（12个月以上）采购计划。

存储管理

人力资源一直是数据中心运营成本中最昂贵的一部分，因此，企业存储管理将很有可能成为IT组织降低运作成本的重要途径。随着网络存储环境（容纳磁盘和磁带的SAN）中服务器数量的增长，存储管理成本控制的重要性将会与日俱增。IT组织应当改变过去那种将操作系统和存储管理混为一谈的思维方式。企业级的存储管理将在规模和灵活性方面带来显著的经济效益（例如，提高每个管理员所能够管理的存储容量）。当然，存储管理应该与其他一些运作活动（例如系统管理、网络管理、数据库管理等）保持协调。不过，若想实现卓越的存储管理，其核心问题之一是要有良好的设计目标。

存储管理的提升方向是实现日常性工作的自动化，提高其效率，从而降低其对人力资源的需求。例如，在过去的24个月中，光纤通道光纤网的主要发展方向是用统一的核心功能对各个企业存储厂商的产品进行管理。在多厂商存储管理的问题中，一些重要的功能将会首先得到解决，如存储资源的储备及分配（减少此工作所需的人员数量，并将所耗费的时间从数小时或数天降低到数分钟）、资源管理（各用户或应用系统所使用的存储容量信息）以及拓扑映射（物理或逻辑资源的占用信息）等。

要达到这些目标,关键是要弄清楚哪些运作任务可以实现自动化并且在各种不同的服务器操作系统以及来自不同厂商的存储硬件之间协同调控,从而在存储管理人员数量和存储基础设施规模保持不变(甚至随着时间的推移而有所缩减)的条件下实现存储容量的增长。

为了保证自动化的效率,应该选择那些能够无缝集成的解决方案来支持存储运作。所谓无缝集成,就是要将各种不同的功能(可能来自于不同厂商的产品)纳入到一个应用系统之中,而不是简单地将各个相互独立的自动化程序的图标集中到一个启动面板上。各解决方案都应当能够支持一个统一的中央储存库(永久性数据库)。该储存库实际上是一个“存储管理数据仓库”,它使得IT组织能够充分地利用那些由各个“管理组件”功能所带来的效率提升。进而,倘若能够对这个中央储存库进行更充分的利用,将其中的信息提供(类似于实时数据仓库当中的反馈循环)给其他管理元件或是ISV(独立软件开发商)的应用系统,则其价值将会更为可观。若是不能实现这一层次的集成/整合,那么即使用户付出双倍的努力,也无法实现存储环境的扩展能力(性能、容量等),并且也无法有效地控制存储管理的员工成本。

存储资源的储备和分配将是首先实现自动化的领域之一。不过这需要一定程度上的组织变革(或者至少是功能上的变革),因为就目前而言,这方面的工作通常不是由个人来完成的,而是由特定的部门来负责(例如服务器管理部门、数据库管理部门或是存储管理部门)。存储分配的工作往往需要同时涉及多项由不同人员负责的活动,因而要对其进行简化并不是一件容易的事。它通常会涉及以下几方面的事务:

- 在物理存储子系统中指定或配置正确的LUN。
- 对阵列上所连接的主机数量及位置进行识别和分配。
- 为特定的服务器设置LUN掩码。
- 对光纤通道交换式光纤网进行分区。
- 制定和调整所需的光纤通道安全策略(若可用)。
- 对卷管理程序或文件系统,以及主机上的应用程序或数据库进行及时更新,执行所需的变更操作。
- 确保在存储策略变更时,其他的存储管理应用系统(如备份和恢复、远程复制服务等)能够与之保持一致。

“卓越数据中心”(Center-of-excellence, COE)中的最佳实践向我们展示了这种集成及整合工作所带来的效益,其中最为显著的是花费在运作监管方面的成本得到了大幅度的降低。对于企业而言,这种收益还将在此后逐年显现出来,而不只是一次性的成本节余。这种逐年收益不仅体现在资金上,同时还体现为企业敏捷性的提高和基础设施灵活性的增强。

应用系统可恢复性

应用程序及数据库管理系统(DBMS)的可恢复性是促使人们实施网络存储(主要是基于SAN的存储方式)的最重要因素之一。基于在实践中对完成恢复操作所需的平均时间(这是指从发现故障开始到应用系统重新开始响应事务请求为止所需的时间,而不仅仅是将数据从磁带上回存到服务器所需的时间)的分析,可以利用若干种不同的技术来满足这些要求。快照或是

基于存储卷的复制方式（可以是本地的，也可以是远程的）能够显著地降低恢复时间，将之减少到数分钟，并且无须使用磁带回存（但可能会用到磁带的备份与恢复）。

数据的回存是一项重要且必需的功能，但是信息或数据的“状态”往往会比信息或数据本身更为重要。若是不能将数据恢复到具有一致性的状态，或是在恢复之后，各个相关的数据库中存放的却是不同时间上的数据，那么这种恢复操作就是徒劳的。在复制和恢复的过程中保持对各组件（可能是来自不同厂商的产品）进行一致性的管理，将为存储管理运作带来重要的益处，并且能够在所需的各种不同平台（通常大型机环境除外）之间实现软硬件的合理化备份与恢复。对各存储厂商的快照或卷赋值功能保持一致性的管理，将有可能实现基于策略的自动化，从而在运作上带来更多的成本节约。进一步而言，所有的应用系统并不都是用一个模子印出来的，同样，存储复制软件也是千差万别的。因此，即便是复制软件在过去的36个月中已经取得了长足的进步，各厂商的产品在与操作系统、集群软件、应用系统和数据库的集成方式上，以及所需的静止时间（获取一份应用系统的一致性视图所需的时间）方面都还存在着一定程度的差异。

卓越的数据和媒体中心

总体看来，卓越的数据和媒体中心（data and media center of excellence）存在着三个层面：概念层（策略）、物理层（基础设施）和运作层（运作管理）。每个层面都是由若干功能或子功能元件组成的“超集”。下面将对各个层面中的每个功能以及相关的职位（供参考）进行详细的介绍，以便为读者提供一份有关功能和组织结构的指导性框架。

另外还有一个重要的问题需要注意。尽管我们将要谈到的许多功能类别在大型主机环境当中同样存在，但是直接让精通大型主机的人员来负责管理所有的系统却不见得是一种恰当的做法。若想建立一个高效的COE，就需要在人员、流程以及技术之间取得均衡，而大型主机与非大型主机环境的技术之间往往存在着重大差异。无论如何，我们都必须以一种整体的观念去看待COE，审慎考察以下各方面的问题。

策略

策略是存储管理中最高层次的视角，它对所需的基础设施起着决定性的作用。企业中应当设置一名存储策略总监（storage policy director, SPD）来负责这一工作。策略之间是相互关联的，应该作为一个整体统筹考虑。下面对一些重要的策略进行了分析。

- **服务组合及服务等级协议** SPD应明确说明他们将向各业务单位提供怎样的服务包，并根据业务单位的实际需要，在成本和服务之间做出取舍。
- **数据保护策略** 不同的数据有着不同的存储要求，数据保护机制应根据服务组合的要求进行配置设计。
- **厂商选择策略** IT组织需要在多厂商策略（这种策略可能较为理性，但不具有单一厂商的“套餐”式便利）和单一厂商策略之间做出选择。这也是一个重要的策略问题。在这方面，不同类型的方案之间可能有所差别，但企业必须从战略目标的视角来看待这一问题。从这一角度来看，这两种策略并不总是互斥的。不论采用的是何种策略，兼容性都应该作为一个前提性要求得到严格的保障，以便在最大程度上实现方案的灵活性。

基础设施

在确定了整体存储策略之后,还需要有一套存储基础设施体系架构来支持各种组件的设计和实施,以实现既定的服务目标。基础设施架构中的主要问题包括设备和管理(在功能层次上),以及以下一些重要工作:

- **物理拓扑设计** 在许多组织中,都需要同时使用到 SAN、NAS 以及直接连接存储形式,并对之进行集成,以实现优化的性能或是满足安全性和数据保护的要求。在这类高度复杂的异质环境之中(多协议、多厂商、多功能、跨组织、跨地区),不但需要考察各厂商产品在这种环境中共存并正常运转的能力,还需要考虑其对建设这种环境的支持能力(借助于用户和技术专家的共同努力),从而选择最为合适的平台(包括硬件和软件)以支撑一个具有足够灵活性的基础设施体系,并随着时间的推延、技术的发展进行相应的调整。
- **性能分析建模及监测** 在存储设施体系结构的设计中不但要考虑容纳足够的资源,还需要保证能够实现所需的服务等级。这既涉及到存储组件(如磁盘阵列),也涉及到存储网络(如 FC 光纤网)。在使用 NAS 的情况下,还需要与 IP 网络的管理人员协同工作。
- **安全性** 在存储管理工作中,安全性并没有得到足够的重视。这主要是由于目前容量问题还是人们关心的重点。但安全性也应该被视为存储基础设施体系结构设计的一个重要组成部分,并且与整个企业的安全架构相结合。

与基础设施相关的其他工作还包括物理子系统和磁带库的设计,容错和性能管理,虚拟化, SAN 管理以及企业级的规划、设计和实施。

运作

数据和媒体 COE 的运作与现有实践相比不会有太大的变化,但需要由专职的存储管理人员来负责。从职能上看,系统管理团队仍然应该负责传统的系统管理工作,关键的区别在于存储运作应当开始从系统、数据库及应用程序的管理工作中分离出来。采取何种程度的分离取决于组织的实际情况,大体上应考虑以下三个主要职能的划分:

- **性能管理** 包括性能分析模型的建立和管理(包括基于 FC 的网络)、资产管理、根据各业务单元对系统的使用情况进行计费 and 收费。
- **资源管理** 指资源的分配和供应,即根据体系或应用程序的要求指派存储资源(如 LUN、数据分区、卷等),此外还包括资源配额管理、资源使用管理(例如是按照应用还是按照实际的用户身份进行资源的分配)、资源测量以及容量规划。
- **可用性管理** 处理备份和恢复等操作。备份和恢复应该是存储管理而非系统管理的一部分,其所涉及的工作包括备份和恢复任务的调度、介质管理以及故障排除。高可用性的服务器集群管理(与系统管理相配合)、层级存储管理以及复制等工作也应该在可用性管理的考虑范围之内。

企业的架构、基础设施、运作方式以及业务类型都必须根据实际情况的需要进行适当的协调,数据和媒体 COE 也是如此。在实际工作中可能会发现有些职能或任务并不属于策略、基础

设施和运作的类别范围之内。若想成功地建设一个能够适应变化、可测量的存储环境，团队之间的沟通是一个关键性的要素。

执行要点

IT 组织应当完成以下一些重要工作：

- **对存储硬件和软件进行合理化配置** 存储的各个方面都应该进行合理化配置，从服务器到应用组合都是如此。当然，存储的生命周期也是一个需要考虑的问题，但一套合理化战略（在厂商和配置环境等方面减少差异性）将为企业带来显著的战略价值。
- **建设存储基础设施** IT 组织应当着手对存储资源进行联网（使用 SAN、NAS 等技术并设计备份与恢复的体系架构），对各层次的存储设施（例如内部存储、中等范围存储和企业级存储）和功能软件（如复制软件、服务器集群软件以及备份/恢复软件）进行协调，并且尽可能地实现组件（如光纤通道交换机、主机总线适配器等）的通用化。IT 组织应当致力于寻找那些符合严格的品牌兼容性标准的新型组件，以便在最大程度上实现存储环境的灵活性。
- **优化存储运作** 这包括管理工具以及人员分工方面的合理化和集成。此外，多厂商环境下的统一存储管理正在成为一个广受关注的问题，具有重要的战略导向意义。
- **建立卓越的媒体和数据中心** IT 组织应当按照上一节中所提出的指导思想建立一个卓越的数据和媒体中心。

外包

从 2003 年到 2006 年，随着 IT 组织开始在企业中引入与传统外包业务相类似的存储服务提供模式，对存储服务内容和质量的准确衡量就成为 IT 组织的一项关键成功因素。对于那些既想获取信任，又想证明采取某种存储方式及相关的存储管理分包方式正确性（从而做出明智决策）的 IT 组织而言，这一能力是至关重要的。在考虑衡量手段时，有三类方法可供选择。第一类方法中包括传统的 ROI 和总拥有成本等指标，这些指标常用于高级管理分析，主要是通过对当前及未来的收益和成本进行内部比较，以评判各种策略所能够带来的价值。第二类和第三类方法包括一些技术上和运作上的考察因素，对于某些存储厂商而言，这些衡量尺度并不一定适用，因而这些方法更多的是提供战略上的参考。IT 组织应当尽可能保持衡量尺度的简洁性，将注意力集中在与企业经营密切相关的重点问题上。通常而言，在起步阶段可以使用网络服务外包（如 Web 站点托管）的 SLA 作为参考框架来设计存储服务的外包方案。衡量存储服务的三类方法具体说明如下：

- **内部比较** 主要指标包括每个管理员所能够管理的存储范围、存储利用率、存储周转时间、数据生命周期乘数、数据可用性以及平均恢复时间。
- **技术指标** 主要指标包括存储可用性、存储响应时间、数据迁移效率、故障诊断效率、容量利用率、性能、资源利用率、故障恢复和排错的平均时间及最长时间。
- **运作指标** 主要指标包括最长通报时间、移动/添加/变更的效率、针对项目的 SLA 以及厂商反应能力。

小结

IT组织正在逐渐意识到存储的重要性,特别是对于那些以往不具有企业级存储容量需求的中等规模企业而言更是如此。到2005年或2006年,大多数组织都会兴建存储基础设施并建立存储运作团队,从而形成卓越的媒体和数据中心。因此,即使是困难的经济环境导致许多企业不由自主地采取一些不恰当的单纯战术决策方法,IT组织也应当正确认识存储所具有的战略重要性。IT组织应当执行一项战略评估,寻找那些能够利用多厂商存储能力支持日常业务功能(并实现其自动化)的方法,以及减少那些耗费大量资源的任务(如存储恢复、存储能力的提供和获取)的途径。

进一步而言,随着存储技术的日渐成熟,越来越多的服务器使用了网络化的存储(SAN和NAS),存储已经成为一种根据企业的商务需求(如性能、可用性、变更的灵活性等)而设计的服务,企业若想得到更高规格的服务,就需要付出更高的成本,在这种形势下,IT组织应当能够对存储能力进行准确的衡量。

在过去的二十年中,Robert Spalding一直是存储技术和存储管理领域一位深受尊敬的权威专家。我曾与Robert共事,他第一个阐述了具有灵活适应能力的存储基础设施的重要性,并探讨了建设这种基础设施所需的各种准备工作。基于这些研究,Robert于1995年指出,企业中应当设置首席存储官(Chief Storage Officer)。

这些开创性的思想大多都已渗透在Robert的著作中,尤其是本书中。作为一名关注各种存储技术和存储管理市场的行业分析者,我从个人的经验中认识到,简明的技术性参考手册是极具实用价值的。存储专业人员若是能将此书用做案头的常备资料,定能受益匪浅。我愿与各位共享这一著作。

Sean Derrington是存储技术和存储管理基础设施领域的权威专家,其研究专长包括存储子系统、光纤通道、存储区域网络、网络连接存储、备份/恢复以及灾难恢复架构等。他同时也关注电子商务以及服务于事务处理和分析处理的存储基础设施。Sean在卡内基·梅隆大学获得材料科学与工程学士学位和工程与公共政策学士学位,并于1995年7月加入META Group。读者可通过电子邮件地址 sean.derrington@metagroup.com 与之联系。

致谢

写作一本技术性著作就像盖一座房子,需要许多人的鼓励、支持、指点,以及才华的交汇和勤奋的工作。各个部分大量的工作中只要有任何一项任务未能按照预定的计划完成,本书就不可能顺利地出版。我要向所有在本书的写作和出版工作中各显其能的人们致以深深的谢意。

如果我能够将本书献给某个人的话,那么这个人一定就是鼓励我创作这本书的 Donn Thornton。Donn 以其言传身教让我学会了如何跳出旧有观念的束缚,从客观的角度去理解、聆听,进而去思考事物。本书的进展过程中一直得到了 Kevin Ansell 的帮助,他的经验和见解帮助我度过了许多难关,我也应对他的幽默感和鼓励表示感谢。

McGraw-Hill/Osborne 的选题编辑 Franny Kelly 为本书的出版发挥了重要作用。我要感谢 Franny 对我的信任,这种信任让我得以完成这一艰巨的工作。将众多的材料组织成一本完整的著作是一项繁重的任务,对此我向本书的责任编辑 Jody Mckenzie 致以真诚的谢意。Jody 与那些出色的文字编辑们一同帮助我将许多分散的素材整合成图文并茂的书稿,他们的工作对于本书的完成具有不可替代的作用。

还有许多人给予了我技术上的帮助。造诣非凡的存储咨询顾问 Patty Then 审阅并修正了书中的许多重要技术概念和规范。Patty 在系统、存储基础设施以及数据中心方面的深厚功力为本书提供了 SAN 技术方面的关键性资源。她是一个难得的人才,对于从大型机到开放式系统环境中性能、容量规划和存储运作问题均有丰富的经验。我还要对 Himanshu Dwivedi 表示感谢,他对本书第 25 章中关于存储基础设施安全性问题的讨论做出了重要的贡献。

我还要对富有创造力的技术作家 Ian Spalding 表示感谢,他帮助我将笔记整理成一些关键性的章节,并最终修改定稿。此外,在本书出版之前, Ian 还参与了最后的审校工作。在此我要预祝她正在创作的第一部小说能够取得成功。

最后我要说,如果没有我的妻子、同时也是我的工作伙伴 Artie 的支持、帮助和鼓励,这本书就不可能完成。她在本书的写作过程中给予了我重大的帮助,其贡献是无法用语言来表达的。谢谢!

Robert Spalding

2003 年 2 月

前言

存储网络代表了计算机系统中数据存储和访问方式的一种基本典范。与其他重大技术变革一样,存储网络带来了新的挑战,要求我们必须理解新的存储架构模式,并进而将各种新型解决方案整合成现实的数据中心。本书的研究、设计和写作工作围绕着一个目标,就是要让这本书成为那些已经开始接触(或者打算涉足)这一快速发展并且常常显得艰深难懂的存储工作的IT业界人士手边的通用、实用的参考资料。

在我看来,自从20世纪80年代后期客户机/服务器体系所引发的变革以来,存储网络是计算机体系中最为重要的发展。从那时开始,数据存储需求量的增长就从来没有停止过。客户机/服务器体系的成功应用将服务器组件技术的处理能力推到了新的高度,超过了存储设备及相关的连接组件所能够支持的水平。这种形势促使系统和存储厂商致力于开发更高容量的磁盘和磁带组等产品,而这些产品的生命周期变得越来越短,这清楚地反映出最终用户对在线数据存储容量的需求呈现着指数级的增长。在探索具有快速数据访问能力的高容量设备的过程中,厂商和用户团体的共同努力促成了两种具有重要影响力的标准的出现,即外围组件接口(Peripheral Component Interface, PCI)和小型计算机系统互联(Small Computer Systems Interconnect, SCSI)。在这些标准的基础上,人们力图克服客户机/服务器体系架构中的I/O瓶颈,使得外围组件能够用比以往更为强大的方式与服务器相连接,并允许其他设备充分利用这些连接。

然而,即便是这些革新,也无法有效地应对20世纪90年代以来潮水一般增长的存储需求。那时,整个存储行业以及各个数据中心都在寻求新的途径以实现能够快速访问的大容量存储。这些研究工作的重点是如何将存储连接的手段从直接连接模式转变为网络模式。当存储网络的概念作为一种可行的存储解决方案出现之后,这一领域中便产生了两个重要的变化。首先,一种存储网络模式克服了纯粹UNIX导向的限制,形成了网络连接存储(Network Attached Storage, NAS)模式。其次,对现有技术的综合改进产生了专门用于存储的网络,并进而发展出存储区域网络(Storage Area Network, SAN)模式。不论是SAN还是NAS,存储网络产品在20世纪90年代后期得到了迅猛的发展,对该领域的技术投资也在此期间达到了顶峰。技术的发展和模式的变革总是会在行业中带来新的气象,并为行业中原有的企业和创业者带来新的机遇。存储网络已经突破了产品研发过程中最重大的障碍,并将持续快速地发展,不过今后的研究重点将会是数据中心的价值和集成。其他的一些研究工作也还将继续推动存储技术的进步,这些工作包括iSCSI和连接技术方面的最新进展,如InfiniBand和RapidIO。

不幸的是,存储网络为IT行业所带来的价值尚未充分地转化为对企业的现实贡献,实施和管理上的复杂性使得许多人对之望而却步。有关存储网络应用的问题中最为重要的一个层面是如何综合利用各种不同解决方案的价值。在当前以及可以预见的将来,这一点将是存储网络发展的关键性驱动因素。本书的主要设计目标,就是要让读者理解存储网络,了解存储网络解决方案中的各种组件及其在数据中心中的应用。本书力图成为IT人员的通用参考资料,着重探讨的是从IT的视角学习、理解并应用一种新技术时所遇到的各种实务性问题。像其他的参考书一

样,本书可以有多种使用方式。首先,它可以作为理解存储网络中各种概念的基本参考书。其次,它可以用于对技术应用方式进行设计和定位的工具。最后,本书也可以用于帮助读者完成安装实施工作并过渡到维护和管理阶段时的补充参考。另一方面,在各种正规和非正规的教学活动中,本书或是书中的部分章节也可作为课堂学习的教材。如何使用本书取决于读者现有的计算机知识水平,特别是有关存储系统方面的知识水平。

本书是为那些需要从数据中心的视角理解存储网络及其相关问题的IT人士准备的。书中的例子都是从实践中提炼而成的,探讨这些实例的目的是为了更好地揭示数据中心存储网络的规划、实施和管理过程中的重点问题。

可以相信,企业数据存储将会作为一种重要的基础设施而继续发展,而存储网络将成为这一发展变革的基石,并带给我们面貌一新的未来数据中心。

目 录

第一部分 存储网络的概念

第 1 章	数据存储和数据访问	2
1.1	应用系统设计的挑战	2
1.1.1	应用系统中的非线性性能问题	3
第 2 章	为存储规模和访问能力而战	12
2.1	存储网络的意义	14
2.1.1	问题所在：存储规模	14
2.1.2	问题所在：访问能力	15
2.2	从存储的角度定义商务应用系统	15
2.2.1	维护和支持应用系统	16
2.3	建设商务应用系统	17
2.4	存储网络对商务应用系统的益处	20
2.4.1	商务应用系统数据访问的可扩展性	20
2.4.2	整合商务应用系统数据	20
2.4.3	通过更高的带宽和集中配置提高数据访问速度	21
2.4.4	通过提高存储管理的效率增强商务应用系统的可用性	21
2.5	存储网络对 IT 组织的影响	22
第 3 章	分离存储组件：将存储放到网络上	24
3.1	NAS 思想	24
3.2	NAS 的演化和发展	26
3.3	NAS 运作概述	29
3.3.1	运用 NAS 应对数据访问	30
3.3.2	运用 NAS 应对数据规模	31
3.4	NAS 的不利因素	32
第 4 章	分离存储组件：为存储建立网络	33
4.1	以数据为中心的世界	33
4.2	分布式计算机处理	34
4.2.1	对称多处理系统	34
4.2.2	大规模并行处理系统	35
4.3	分布式 I/O 处理	36

4.3.1	增强的 I/O 协议——光纤通道	36
4.3.2	高速互联——交换式光纤网络	36
4.3.3	“无共享”体系结构	37
4.4	SAN 思想	38
4.5	SAN 的演化和发展	38
4.6	SAN 运作概述	40
4.6.1	针对数据访问的 SAN	40
4.6.2	针对数据规模的 SAN	40
4.7	SAN 的不利因素	41

第二部分 存储基本原理

第 5 章	存储体系架构	44
5.1	存储元件	44
5.1.1	数据存储层级	46
5.2	存储系统	47
5.2.1	典型的存储系统配置	48
第 6 章	设备概览	52
6.1	外围连接组件及相关概念	52
6.2	主机适配器	54
6.3	控制器	55
6.4	磁盘存储	57
6.5	磁盘系统	58
6.5.1	磁盘阵列	58
6.5.2	RAID 磁盘阵列	59
6.6	磁带存储	61
6.6.1	磁带驱动器	62
6.6.2	格式	63
6.6.3	读写机制	63
6.6.4	磁带系统	63
6.6.5	光存储	64
第 7 章	连接方式	66
7.1	连接：搭上总线的“公共汽车”	66
7.1.1	带宽	67
7.1.2	寻址	67
7.1.3	中断功能	68
7.1.4	电子和机械特性	68