

Designing Storage Area Networks

A Practical Reference for

Implementing Fibre Channel and IP SANs

Second Edition

存储区域网络设计

—— 实现光纤通道和IP SAN的实用指南

(第二版)

[美] Tom Clark 著

邓劲生 李宝峰 李 蕾 等译



存储技术丛书

存储区域网络设计

——实现光纤通道和 IP SAN 的实用指南

(第二版)

Designing Storage Area Networks
A Practical Reference for Implementing Fibre Channel and IP SANs
Second Edition

[美] Tom Clark 著

邓劲生 李宝峰 李蕾 等译

电子工业出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书全面系统地介绍了基于光纤通道和IP技术的存储区域网络的设计方法。书中首先对存储和网络概念做了一些基础性的介绍,然后对光纤通道结构内部的工作模式进行讲解,并介绍了光纤通道SAN拓扑结构和一些产品。随后对IP SAN技术进行分析,并介绍了一些IP SAN产品。接着本书讨论了各种SAN的硬件和软件、SAN的故障隔离、SAN的管理、存储虚拟化等内容,并给出多个SAN的应用实例以便进行分析,最后简单讨论了SAN技术未来的发展趋势。附录包含SAN资源、厂商、SNIA 共享存储模型、SNIA 存储网络术语辞典等内容。

本书内容简洁明快,实用性强,能够使读者在了解光纤通道和IP SAN基本技术的基础上,较快地掌握SAN的设计方法。本书适合网络设计工程师、存储系统设计师、网络管理员等技术人员,以及对SAN这一领域感兴趣的读者阅读。

Simplified Chinese edition Copyright © 2005 by PEARSON EDUCATION ASIA LIMITED and Publishing House of Electronics Industry.

Designing Storage Area Networks: A Practical Reference for Implementing Fibre Channel and IP SANs, 2nd Edition, ISBN: 0321136500 by Tom Clark. Copyright © 2003.

All Rights Reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Addison-Wesley.

This edition is authorized for sale only in the People's Republic of China (excluding the Special Administrative Region of Hong Kong and Macau).

本书中文简体字翻译版由电子工业出版社和Pearson Education培生教育出版亚洲有限公司合作出版。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Pearson Education培生教育出版集团激光防伪标签,无标签者不得销售。

版权贸易合同登记号 图字:01-2003-4994

图书在版编目(CIP)数据

存储区域网络设计——实现光纤通道和IP SAN的实用指南(第二版)/(美)克拉克(Clark, T.)著;
邓劲生等译. -北京:电子工业出版社, 2005.1

(存储技术丛书)

书名原文: Designing Storage Area Networks: A Practical Reference for Implementing Fibre Channel and IP SANs, 2nd Edition

ISBN 7-121-00563-8

I. 存… II. ①克… ②邓… III. 信息存储-计算机网络-设计 IV. TP393.0

中国版本图书馆CIP数据核字(2004)第118715号

责任编辑:杜闽燕

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:24.5 字数:627千字

印 次:2005年1月第1次印刷

定 价:45.00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 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

译者序

今天,随着信息技术的突飞猛进,商业运转和学术研究中产生了越来越多的海量数据,从而给IT界带来了前所未有的存储“危机”。人们开始使用大容量的磁盘,希望能够缓解一部分负荷,然而使用传统的技术很难将足够数量的磁盘连接到文件和程序服务器上。于是人们又寄希望于分布式文件系统等传统网络存储技术,但是要想使其带宽跟上存储数据的要求又成了一个难题。这些问题使得基本存储结构的性能与用户需求之间的矛盾日显突出。如今必须超越传统的服务器/存储器互连的技术局限,寻找新的存储解决之道。

存储区域网络(SAN)应运而生,成为人们解决带宽、容量和管理问题的首选方案。SAN是一种基于光纤通道或IP标准的结构,它提供了吉比特级别的速度、长距离的传输,并且支持网络扩展到数百万台设备。尽管SAN这个古老的名词在较早的电子计算机和微型计算机领域就已经存在,但存储网络的概念却是一直到光纤通道产品面世才真正进入到主流应用中来的。随着技术的发展,基于千兆位以太网的SAN更加引人注目,得到了更为广泛的网络厂商的支持。相比目前的传统存储模式,SAN代表着一种更先进的技术,因为它在服务器和存储器之间引入了新的网络化的产品和概念。

本书的第一版于1999年冬出版,是国外关于SAN技术的一本经典著述,并多次重印。经过三年多的发展,无论是基于光纤通道还是基于IP的SAN,都有了长足进步。一些概念和提法已经过时,一些新的模型被提了出来,一些新的标准得以通过并且工业化,市场也正掀起风起云涌的波涛。Tom Clark先生在2002年推出“IP SANs: A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks”一书之后,继而对本书于2003年进行了全面重写,以反映最新的技术趋势和市场现状。

与第一版相比,第二版在提纲上保留了原有的标题框架,但内容却是彻头彻尾地全面改写,引入了SNIA共享存储模型并贯彻始终,以全新的视角和独立于厂商的角度,看待近几年技术和市场的变革,以其二十多年的IT经历笑看风云变幻。本书新增的附录占据了非常大的篇幅,几乎占到本书厚度的一半,更为读者充分理解正文内容提供了相当翔实的文献和资料。无论是对于初次涉及SAN领域的读者,还是那些使用SAN技术解决实际问题的人,这都是一本很好的参考书。

目前,随着对网络计算和存储环境需求的不断增强,译者所在的“并行与分布处理”国家重点实验室下属课题组目前正在进行“光互连高速存储器”关键技术的研究与实现。继翻译并出版本书第一版和“IP SANs: A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks”之后,我们再次有幸翻译本书第二版。在这三本书的前后翻译过程中,我们也随着作者的思路,清晰地看到整个领域的发展过程,感到受益匪浅。对于与作者多次E-mail来往,并承蒙馈赠亲笔签名样书,我们深感荣幸。

本书翻译工作由姜南组织，初稿由邓劲生、李宝峰、李蕾、余磊、姜南、方兴等同志完成，罗金平、陈曙晖、汪东、徐峰、胡勇、郑艳波、周建飞等人也做了大量的工作，最后由邓劲生统稿。在翻译过程中，张银福研究员以其极大热情审阅了部分关键性章节，并提出了许多指导性意见。SNIA 中国的于震先生也一直关注我们的工作，并给予了很大的支持。

尽管我们为此付出了许多努力，但由于存储网络是一个正在兴起的全新领域，译者涉足该领域的时间也不长，加上翻译时间仓促，笔头疏忽和理解错误在所难免，欢迎大家批评指正，以求共同进步。

前 言

本书是对前一版本内容的扩展和更新,该书仅用二百多页第一次为存储区域网络(SAN)技术提供了简要概览,对于厂商和客户来说都适合作为介绍和培训用的教材。在过去这三年里,存储行业中发生了迅猛的市场增长和技术发展,新的功能和技术要求已经改变了SAN这一领域。除了光纤通道性能和交换技术得以更进一步地发展之外,还引入了基于TCP/IP和千兆位以太网的SAN,并且正在浮现的存储虚拟化也使客户在处理其存储需求时有更多的选择。虽然这些技术动态的部分内容已经在作者的第二本书“IP SANs: A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks”中有所覆盖,本书却是第一次针对实现人员在使用当前技术来设计一个SAN解决方案时所需要考虑的实际问题而展开阐述的。

存储区域网络(SAN)现在已经被认为是满足公共机构和企业关键数据存储需求的优先之选。SAN在实现用户价值方面的成功,体现在所有的主流解决方案提供商都采用这项技术作为其主推服务器和存储方案。比如IBM、HP、Sun、Dell和其他解决方案提供商,现在都提供认证的SAN配置以优化存储数据的性能、可用性和备份。存储厂商,如EMC、Ditachi Data Systems、XIOtech等,都在其主流产品中提供了SAN接口。虽然在当前基于SAN的解决方案所占有的市场还只有整个存储市场的20%,SAN却有望在几年内成为市场的主流。

正在发展的存储网络代表了两种不同技术的融合,两者都有其各自的重点、词汇和背景。在过去的二十年里,存储的发展比较缓慢,尽管得到了体积更小容量更大的磁盘和更快的通道性能,但是进展仍然局限于主机与存储设备之间固定的主从关系。与之相反的是,网络技术的迅猛发展却早已颠覆了先前的体系结构和关系,正如Internet所代表的,产生了创造力无限和丰富多彩的世界。将相对保守的存储与更为活跃的网络结合,所产生的背景冲突却更有利于双方发展。数据存储从数据通信的局外人,转变成了投资者、厂商和技术专家所关心的热点。网络获得了无穷的高容量数据支持,而数据得以更大的扩展并调整了在企业中的重要性。数据还获得了新的机遇,可以创建存储专用的接口和功能。

将存储和网络集成到一项新的技术中去,为客户和厂商都提出了挑战和机遇。存储资源的管理员和经理们现在都必须接受网络的概念和术语,如编址、路由和网络收敛;而网络架构师和管理员也必须学习以前所不熟悉的LUN、JBOD和RAID级别。因此本书的读者面较为广泛,IT经理、管理员、顾问人员和技术员工都可从中学习存储和网络的实现。另外,存储网络还产生了一些新的概念,也出现了一些以前在存储或网络中都未曾有过的问题。正文中还讨论了在一个开放式网络环境中存储资源的放置所产生的不同效果,这些内容在应用研究中也有所涉及。

虽然存储网络对于处理存储数据的急速增长来说是一项很好的技术,但是现有的棘手的互操作性和管理问题,也使得它备受指责。SAN相关的某些问题,只是新的技术在试图开拓新的应用背景时所产生的副产品;然而遗憾的是,另外的某些问题,则是厂商在试图获取市场份额时产生的。克服互操作性和管理复杂性,是SAN技术得以更广泛应用的需求;现在已有多个行

业机构正在试图处理这些障碍。虽然这些努力可能需要好几年才能有所成效，本书也进行了讨论。

在编写本书第一版的时候，SAN的底层构造还只有光纤通道。作为第一个成功的吉比特串行传输技术，光纤通道所开创的信号传输和数据编码机制后来被千兆位以太网所采用。如果千兆位以太网首先出现的话，恐怕它早已超越光纤通道成为存储网络传输的主流了。虽然以太网在数据通信市场上占据了不可撼动的统治地位，但是它对于存储网络来说还只是个迟到者，与风头正盛的光纤通道的冲突才刚刚开始。虽然这场竞争使得厂商产生了激烈的争论，但毕竟光纤通道和以太网都只是简单的底层结构。本书对于光纤通道和基于以太网的SAN都给予技术细节的讨论，希望客户能够自行做出决定看看到底哪种传输方式最适合自己。销售SAN解决方案的主流厂商，都计划提供光纤通道和基于IP的两种产品，以满足不同客户的需要。

直接负责设计和实现SAN的人们，迫切希望关于产品容量和互操作性更详细的信息，比如哪个版本的特定备份应用程序能够适用于哪种数据库或哪种操作系统。仅仅给出本书写作的这么点时间，是不可能面面俱到地覆盖这些有用信息的。即使本书中牵涉到这些内容，也会马上过时并且毫无价值。SNIA（Storage Networking Industry Association，存储网络行业协会）正在建立最终用户组织，以维护客户问题和解决方案的在线数据库。作者建议大家搜索更多详细的SAN解决方案信息，加入到这个SNIA组织的活动中，以构建将使更多人受益的实践资源。

设计一个存储网络，首先就要回答一个基本问题：它的应用是什么？客户毕竟不会去花费数百万美元在存储阵列、磁带子系统、交换机、服务器和缆线上，却仅仅是为了技术本身。应用是SAN发展的源动力，并且应用随着机构和企业的需求不同而完全不同。虽然厂商可以准备好全套的解决方案来满足通常的需求（比如存储合并或磁带备份），但是SAN架构师应该注意不要被厂商所左右，而是去驱动厂商来解决自己的特定应用需求。本书试图讲述SAN技术高效解决的最为常见的应用，但这些都只是一个开始，存储网络将会在更为广阔的现实世界应用中施展拳脚。

本书的组织结构

本书将会带领你学习一系列的概念，最终得到理解存储网络基础设施和应用的一些基础知识。本书不需要读者以前学习过SAN的相关知识，同时也试图在基础和高级的内容之间保持平衡。这种混合将为希望了解它的读者提供足够的技术细节，又能为希望更深层次地理解SAN技术的人们提供有针对性的概览。

前面两章为理解共享存储的中心概念提供了一个框架。第1章简述了SNIA共享存储模型，它清晰地描述了存储应用和底层基础设施的基本层次。SNIA共享存储模型对于理解SAN体系结构并定位于相关上层应用需求非常有用。它还解释了一个特定的解决方案如何更好地满足业务需求，从而有利于调整SAN结构以更好地管理。第2章提供了存储和网络这两个概念的概览，并解释二者是如何融合而产生一个新的技术来解决数据存储问题的。

由于光纤通道首先出现在存储网络领域，并且一直占据着很大的市场份额，因此接下来的三章提供了关于光纤通道协议、拓扑和产品的技术讨论。第3章概述了底层物理传输、协议和编址。第4章概述了光纤通道拓扑，重点在于当前大部分SAN解决方案所提供的光纤通道光纤

网交换式服务。第5章描述了光纤通道厂商开发的产品，包括服务器、存储设备、互连以及其他组件。总的来说，它们为SAN架构师在设计高效的解决方案时提供了丰富的工具箱。

从2000年开始，将存储数据在主流的TCP/IP网络上进行传输的企图，产生了IP存储网络体系结构和产品。这些产品进一步增强了SAN设计人员在构造共享存储配置上的能力，可以单独使用IP，或者将IP和光纤通道集成为一个异构的存储网络。第6章概述了IP SAN的协议和问题，第7章描述了在过去两年多时间里市场上所出现的一些IP存储产品。

第8章概述了由SAN基础设施支撑的特定存储应用。服务器集群用以获取高可用的数据访问，磁带备份用以对数据进行存档，以及通过磁盘镜像进行数据复制，尽管这些存储应用原本只为高级客户业务应用程序和数据库服务，但现在却都是SAN解决方案中很常见的一部分。

然而，SAN在带来这一切的同时，势必也带来了某些痛苦。虽然厂商都在试图进一步调试和验证SAN配置，但是在安装或接下来的操作中必然会引发一些问题。第9章解释了一些基础的故障排除技术和工具，用于确定并隔离存储网络问题。

第10章讨论了SAN管理及其特有的挑战。共享存储的管理人员，必须将传输的管理以及数据在磁盘或磁带上进行放置的管理结合起来。建立一个统一的管理是非常困难的，即使软件厂商已经掌握了足够的资源来处理SAN管理。但是，CIM（Common Information Model，通用信息模型）已经进入这个行业中，可能会为存储网络的综合管理提供一个框架。

第11章试图描述尚未定型的一个东西，叫做“存储虚拟化”。存储虚拟化在某种意义上被厂商的市场宣传过度炒作，但是对于简化存储管理并为SAN解决方案打开更为广阔的市场来说，它是一项可行的技术。该章覆盖了存储虚拟化的理论能力，以及在当前市场产品中可用的极为有限的功能。

各种机构和企业，正在为互不相同的应用而实现SAN。其中的一些应用在第12章中进行了讨论，提供了在实施和管理SAN时客户可能会遇到的潜在问题。在实用中，光纤通道和IP存储产品的结合是可行的，尤其是在处理城域或广域SAN配置时。

第13章讨论了SAN技术中迄今仍然没有解决的一些问题，包括互操作性、管理以及使用主流网络的SAN收敛。对于工程师而言，这些问题都只是必须掌握的挑战而已，市场的竞争并不能够阻挡大势所趋。

作为SAN设计的总结，第14章是对存储网络前景的思考。这些前景的承诺，与SAN以前的发展一样地难以捉摸，从新的存储方向（比如虚拟化和基于IP的SAN）的快速发展，到新的主机互连（如InfiniBand）的到来。

本书的附录中包含了大量额外材料，有助于读者获得关于存储网络的参考和背景信息。SNIA也以SNIA共享存储模型和SNIA术语辞典的形式，做了一些非常有价值的工作。作者希望存储网络技术的厂商和客户都采用SNIA的工作，并且如果可能的话，也加入到SNIA活动中。客户与技术人员的目标越是接近，可行的解决方案就会越早递交到客户手中。

致谢

在编写本书的新版本时，我受益于在过去十年里SAN开发过程中成千上万的人们的辛勤劳动。解释一项技术的原理，与那些定义并生成不同解决方案的辛苦相比，要相对容易得多。因此我首先希望感谢技术人员、工程师、架构师、施工人员、系统工程师、技术支持人员、最终

用户还有市场人员，是他们共同构造了存储网络的产品和应用。存储网络是一项用于保证信息访问的技术，这样才能提高我们的通用环境。

我要感谢编辑 Karen Gettman 和 Addison-Wesley 的 Emily Frey，是他们在管理这个项目、认真校对并编辑手稿。没有他们在处理出版问题上始终如一的支持和努力，这项工作就不可能完成。

我还要感谢我的朋友，他们审阅文稿，并对手稿提出了建议和修正。这里感谢 Larry Krantz、Charles Monia、Milan Merhar、Bob Snead 和 Gary Orenstein，他们不辞劳苦地审读，并从技术和语句组织等方面都提出了改进。

我要感谢我的妻子兼第一线编辑 Lou Clark，感谢她在生活和工作中给予我的关怀和支持。

本书包含很多对存储网络行业协会的参考，我是站在它的肩膀上完成这项工作的。SNIA 的发起成立者主要是存储和存储网络厂商，团队主要来自成员公司的志愿者。SNIA 以其雄厚的技术和丰富的经验，致力于确保存储网络成为完善并可信的解决方案。为新的存储技术确定标准需求、遵循的标准、互操作性、管理、培训以及其他工作，需要持久地关注和研究；而负责这些任务的人却很少被 IT 业界所熟知。因此我要感谢所有的志愿者，他们花费了大量的时间和精力在 SNIA 技术工作组、委员会、论坛和技术委员会；我同样要感谢他们所在的公司和研究机构。

下面还要感谢 Network Associates 的 Jim Vale，他最早认识到 IP 网络对于 SAN 的重要性，并将存储的思想带入了主流 IT 领域。

最后，我要感谢那些客户，他们对 SAN 解决方案在实施时的优势和弱点都提出了很有价值的建议。我还要特别感谢 SNIA Customer Executive Council 和 SNIA Customer Advisory Council 的成员，尤其是 Carlson 公司的 Gary Johnson 和 Mark Price，以及微软的 John Alley、Brad Basden（和他的朋友们）和 Mark Licata。

Tom Clark

于西雅图

designing.sans@verizon.net

目 录

第 1 章 绪论	1
1.1 使用 SNIA 共享存储模型	1
1.2 范例: Carlson 公司	6
1.3 本书概览	9
1.4 小结	11
第 2 章 存储和网络概念	12
2.1 服务器前方的网络	12
2.1.1 串行传输	15
2.1.2 访问方法	16
2.1.3 编址	16
2.1.4 数据打包	16
2.1.5 报文路由	17
2.1.6 上层协议支持	17
2.2 SCSI 体系结构	17
2.3 并行 SCSI 总线	20
2.4 网络附接存储	22
2.5 服务器后方的网络	23
2.6 小结	24
第 3 章 光纤通道内部解析	26
3.1 光纤通道层次	26
3.2 1 Gbps 和 2 Gbps 传输	27
3.3 物理层的选择	28
3.4 数据编码	30
3.5 有序集	32
3.6 组帧协议	33
3.7 服务类	34
3.8 流量控制	35
3.9 命名和编址的约定	36
3.10 小结	38
第 4 章 光纤通道 SAN 拓扑结构	40
4.1 点到点	40
4.2 仲裁环	41

4.2.1	环的物理拓扑	41
4.2.2	环的寻址	43
4.2.3	环的初始化	45
4.2.4	端口登录	49
4.2.5	环端口状态机	50
4.2.6	仲裁	50
4.2.7	仲裁环的非广播特性	52
4.2.8	仲裁环设计时需要考虑的问题	53
4.3	光纤网	56
4.3.1	光纤网登录	59
4.3.2	简单名字服务器	60
4.3.3	状态转变通告	60
4.3.4	专用环的支持	61
4.3.5	光纤网分区	62
4.4	建立扩展的 SAN	63
4.4.1	E_Port 标准化	64
4.4.2	主交换机选择	65
4.5	光纤网和环	66
4.6	小结	67
第 5 章	光纤通道产品	69
5.1	收发器	69
5.2	主机总线适配器	71
5.3	光纤通道 RAID	74
5.4	光纤通道 JBOD	76
5.5	仲裁环 Hub	78
5.5.1	星形拓扑结构仲裁环	78
5.5.2	Hub 的体系结构	79
5.5.3	不可管理 Hub	81
5.5.4	可管理 Hub	81
5.6	交换式 Hub	83
5.7	光纤网交换机	84
5.7.1	部门级光纤网交换机	85
5.7.2	光纤通道骨干交换机	86
5.8	Fibre Channel-to-SCSI 桥接器	87
5.9	光纤通道扩展产品	88
5.9.1	使用 DWDM 的光纤通道扩展	88
5.9.2	使用 IP 隧道的光纤通道扩展	89
5.9.3	光纤通道 WAN 桥接	90
5.10	小结	91

第 6 章	IP SAN 技术	94
6.1	以太网和 TCP/IP	94
6.1.1	千兆位以太网传输	94
6.1.2	TCP/IP	99
6.2	本地 IP 存储协议	104
6.2.1	Internet 光纤通道协议	104
6.2.2	Internet SCSI	106
6.3	IP SAN 的发现	109
6.3.1	服务定位协议 (SLP)	109
6.3.2	Internet 存储名字服务器	110
6.4	IP SAN 的服务质量	112
6.5	IP SAN 的安全	113
6.6	广域存储网络	114
6.7	小结	117
第 7 章	IP SAN 产品	120
7.1	千兆位以太网交换机	120
7.2	IP 路由器	122
7.3	iSCSI 适配卡	122
7.4	iSCSI 存储设备	124
7.5	IP 存储网关	125
7.6	iSCSI-to-SCSI 桥接器	127
7.7	iSNS 服务器	128
7.8	小结	129
第 8 章	SAN 软件产品	131
8.1	服务器集群	131
8.2	磁带备份	132
8.3	数据复制	134
8.4	分布式文件系统和文件共享	135
8.5	小结	136
第 9 章	SAN 的故障隔离	138
9.1	简单故障隔离技术	138
9.2	光纤通道分析器	140
9.3	iSCSI 网络分析器	142
9.4	性能工具	143
9.5	小结	143
第 10 章	SAN 的管理	145
10.1	存储网络管理	145

10.1.1	in-band 管理	146
10.1.2	out-of-band 管理	147
10.1.3	SNMP	148
10.1.4	HTTP	149
10.1.5	Telnet	150
10.1.6	存储网络管理中的问题	151
10.2	存储资源管理	151
10.3	存储管理	152
10.4	存储、系统以及企业级管理的集成	152
10.5	通用接口模型	153
10.6	小结	154
第 11 章	存储虚拟化	156
11.1	什么是存储虚拟化	156
11.2	in-band 和 out-of-band 虚拟化	159
11.3	基于主机的存储虚拟化	160
11.4	基于 SAN 互连的存储虚拟化	160
11.5	基于存储的虚拟化	162
11.6	多厂商存储虚拟化	163
11.7	文件系统和 NAS 虚拟化	163
11.8	磁带虚拟化	164
11.9	虚拟化和数据存储公用设施	165
11.10	小结	166
第 12 章	应用研究	168
12.1	视频后期编辑	168
12.2	印前处理	169
12.3	LAN-free 和 Server-free 磁带备份	170
12.4	服务器集群	174
12.5	存储合并	175
12.6	Internet 服务提供商	177
12.7	园区存储网络	179
12.8	远程磁带遥连	180
12.9	灾难恢复	181
12.10	小结	183
第 13 章	SAN 的问题	185
13.1	标准化	185
13.2	互操作性	186
13.3	管理	187

13.4	融合	188
13.5	小结	189
第 14 章	SAN 的未来	190
14.1	将 SAN 集成到主流网络中	190
14.2	共享存储的普遍性	191
14.3	虚拟化	192
14.4	人的因素	192
14.5	有贡献的技术	193
14.6	小结	194
附录 A	SAN 资源	195
附录 B	SAN 和相关厂商	197
附录 C	标准化过程	202
附录 D	存储网络行业协会	204
附录 E	SNIA 共享存储模型	210
附录 F	SNIA 存储网络术语辞典	239
参考文献		376

第1章 绪 论

存储区域网络的设计,需要对依赖存储资源的高层应用以及满足应用需求的存储系统的体系结构都有深入的理解。尽管特定应用的需求会因客户的不同而不同,但是能够帮助存储架构师对存储方案进行取舍的框架结构已经开发出来。SNIA 的 SSM (Shared Storage Model, 共享存储模型) 为理解主机应用、存储网络和存储设备之间的关系提供了指导。

1.1 使用 SNIA 共享存储模型

在多个服务器或工作站之间共享存储资源,需要一个连接目标机器和源机器的对等 (peer-to-peer) 网络。网络的构成及其上传输的存储数据类型因结构而异。一般来说,共享存储体系结构分为 SAN (Storage Area Network, 存储区域网络) 以及 NAS (Network Attached Storage, 网络附接存储)。对于 SAN 来说,网络基础设施可能是光纤通道 (Fibre Channel) 或千兆位以太网 (Gigabit Ethernet), 其上传输的存储数据类型为 SCSI (Small Computer System Interface, 小型计算机系统接口) 块数据。对于 NAS 来说,典型的网络基础设施则是以太网 (快速以太网或千兆位以太网), 其上传输的存储数据类型是基于文件的。在最抽象的层次上, SAN 和 NAS 的共同特性是二者都允许存储资源在多个发起方之间共享,而不论数据是基于块的还是基于文件的。

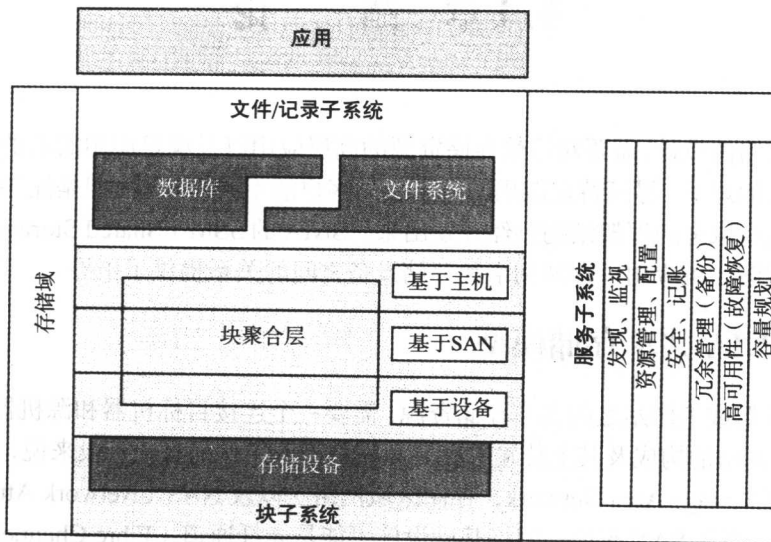
对于工程师和签署存储设备购置协议的经理们来说,理解 DAS (Direct Attached Storage, 直接附接存储)、SAN 和 NAS 方案的不同角色,并将它们应用于统一的 IT 存储战略中,这是一项挑战。SNIA 共享存储模型提供了一个有用的框架,有助于理解高层应用以及支持它们的存储基础设施之间的关系。SNIA 具有将当前存储配置映射成建议方案的能力,从而能够帮助澄清存储架构师试图阐述的问题,同时也为接下来的需求分析和方案实施创建了一个框架。

如图 1.1 所示,共享存储模型为运行在服务器和主机上的用户应用,与底层的存储域 (storage domain) 建立了普遍的联系。应用可能支持诸如处理在线事务、数据库挖掘或 Web 服务器等的用户行为。存储特定的应用,比如管理、备份、集群服务以及磁盘对磁盘的数据复制等,包含在存储域的服务子系统 (service subsystem) 中。这样模型就将处于高层的终端用户或商业应用,与监测、支持底层存储域的次要应用区分开了。

存储域细分为三个范畴:文件/记录子系统、块聚合层以及块子系统。文件/记录子系统 (file/record subsystem) 是高层应用与存储资源之间的界面。数据库应用 (如 SQL Server 及 Oracle) 使用记录格式作为处理单元,而大多数其他应用则希望使用文件作为处理单元。

对于高层应用来说,有用的信息无论是作为记录还是文件出现,两种格式最终都以连续的数据字节存储到磁盘或磁带上,这些连续的数据字节称为数据块 (block)。数据块的大小,以及将记录或文件映射到块上的方式,都可以因系统而异。在任何情况下,存储域都需要一些方法,以便将数据块与恰当的记录或文件描述符关联起来。这种功能由块聚集 (block aggregation) 层进行刻画,它既可以基于主机系统,又可以基于存储网络或存储设备。通过相关的记录或文

件进行识别,数据块从物理存储设备中被读出或写入,如共享存储模型中块子系统 (block subsystem) 所示的那样。



Copyright © 2000, Storage Networking Industry Association (work in progress)

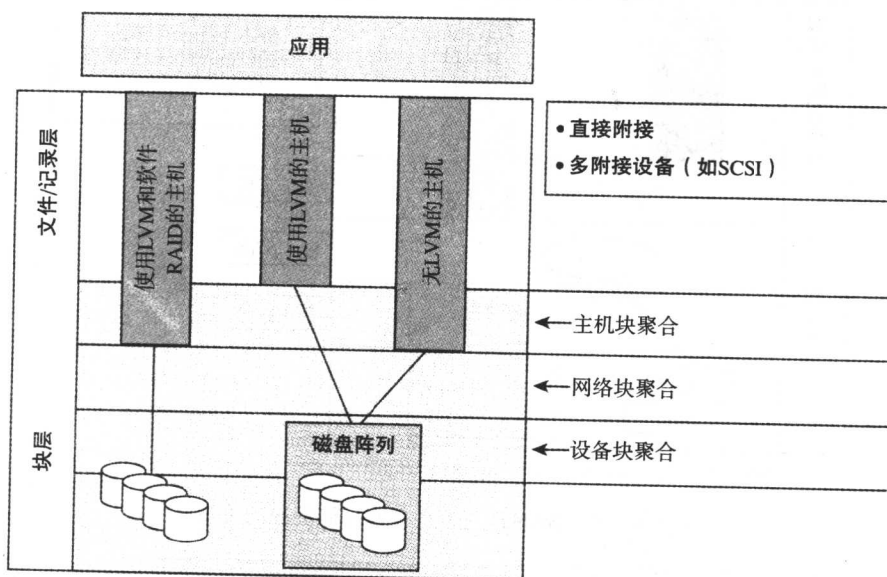
图 1.1 SNIA 共享存储模型概览

同属于存储域,但作为一个辅助子系统分开的,是服务子系统 (service subsystem)。它包含一系列存储相关的功能,包括管理、安全、备份、可用性维护以及容量规划等。这些服务既可以作为存储产品所集成的功能,又可以作为监测和管理存储资源的独立软件。某些特殊的块聚集方式可能需要专门的管理服务。以NAS设备为例,它需要一类独立类型的应用的服务来完成与其兄弟 SAN 不尽相同的备份功能。

在创建能够可靠集成文件/记录、块聚集以及块子系统的可行共享存储方案方面,已经投资进行了大量的工程研究。例如,保证数据在高速串行链路上传输完整性块数据的吉比特传输,就花费了数年来进行标准的开发以及验证测试工作。直到这些基础设施问题得以解决,服务子系统的开发才变得可行。因此,管理、安全以及其他一些辅助服务的建设也滞后于共享存储网络的发展。今天,由于缺乏增强的管理服务功能导致抑制了共享存储在市场上的进一步扩张,从而存储行业的焦点也从基础设施转移到使其易于安装和支持的共享存储方案的辅助服务上来。在使用共享存储模型来配置自有的存储方案时,可以期望服务子系统中出现新的产品来支持管理、安全以及其他一些功能。

以层次化的共享存储模型作为指导,现在可以插入服务器和其他存储组件,来对直接附接、SAN 以及 NAS 配置进行清晰地区分。如图 1.2 所示, DAS 通过并行 SCSI 电缆从服务器扩展到目标存储设备。尽管共享存储预计在最近几年内将逐步取代 DAS,但它仍然是目前最常见的存储配置。在本例中,图 1.2 左边展示了一个主机上具有 LVM (Logical Volume Management, 逻辑卷管理) 以及软件 RAID (Redundant Array of Independent Disks, 冗余磁盘阵列) 的服务器。当服务器从应用程序接受到需要写盘的信息时,软件 RAID 将数据库分块并分散到块层的多个磁盘上。这里主机完成了块聚集功能,而软件 RAID 执行了将数据分块存放于多个磁盘上的机

制。逻辑卷管理器以卷（例如 M: 驱动器）、目录以及子目录的方式，为高层应用提供了一致的数据视图。



Copyright © 2000, Storage Networking Industry Association (work in progress)

图 1.2 在 SNIA 共享存储模型中的直接附接存储

在图 1.2 的右边，图中的服务器有一个通过 SCSI 连接的集成 RAID 控制器的磁盘阵列。在此例中，主机就不用通过软件 RAID 来对数据进行分条（striping）了。磁盘阵列自身完成此项功能。因此，图中的块和块聚合层有所重叠。

SAN 改变了服务器与目标存储器之间的关系，如图 1.3 所示。服务器与存储器通过对等网络而非并行 SCSI 电缆连接起来。与直接附接存储相同，逻辑卷管理、软件 RAID 以及硬件 RAID 依然完成各自的功能，而服务器和存储设备之间的互连则允许任何服务器和存储设备之间进行连接。图中由中间区域的 SCSI 电缆表示出来的服务器，对存储资源独有的所有权不再是必需的。可以根据意愿将存储资源分配给指定的服务器，并可以动态地改变服务器和存储设备之间的关系以适应应用的需求。

将直接附接的连接换成更为灵活的网络互连，产生了新的存储解决方案。例如，可以对存储进行合并，在磁带和磁盘间共享资源，以及将多个服务器进行集群，以获得高可用性。根据所使用的 SAN 拓扑结构，还可以将网络化的存储扩展至更大数量的服务器和存储设备。例如，在一独立的 SAN 中，一个大型磁盘阵列可以支持几十台到数百台服务器，从而将阵列的成本分摊至更多的主机上，并可以高效地进行存储管理。

SAN 底层基础设施可以是光纤通道或千兆位以太网，也可以通过适当的 IP 存储交换机将二者结合使用。为了准确地表达客户实际的配置情况，该模型能够刻画服务多种高层应用的多个 SAN，并提供对普通或特殊存储资源的连接，如图 1.4 所示。在异构环境中，区分哪些 SAN 仅使用光纤通道光纤网（fabric）、哪些 SAN 牵涉到混合的光纤通道以及 IP 组件，是很有用处的。