

- 硅谷顶尖IT公司的经验结晶
- 华人科技精英的原创奉献
- 传递知识的快车道
- 跨越鸿沟的桥梁

硅 谷 精 英 科 技 丛 书

网络存储技术

赵文辉 徐俊 周加林 李晨 著



清华大学出版社

硅谷顶尖IT公司的经验结晶

华人科技精英的原创奉献

传递知识的快车道

跨越鸿沟的桥梁

硅 谷 精 英 科 技 丛 书

网络存储技术

赵文辉
徐俊
周加林
李晨
著

清华大学出版社
北京

内 容 简 介

存储技术是最近几年 IT 行业最热门的技术之一。本书全面深入地介绍了存储技术的基本原理、体系结构和设计存储系统的方法及其实例分析。内容包括:存储模型的发展历史和分类;存储技术中的基础技术——独立冗余磁盘阵列和 SCSI 技术;文件系统的基本知识;联网存储系统的原理;光纤通道协议;存储区域网络的概念、特点、体系结构及主要产品;光纤通道交换机的工作原理和组织结构;存储管理;存储虚拟化;基于 IP 通信的存储技术中的两种基础技术——iSCSI 和 FCIP;存储产品的概况;最后给出了几个存储系统的设计实例。

本书既可以供计算机、通信、电子、信息、自动化等相关专业的科研人员、研究生和大学高年级学生作为教学参考书,也可以供计算机信息系统开发及应用的工程技术人员参考。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

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。

本书防伪标签采用清华大学核研院专有核径迹膜防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

网络存储技术/赵文辉等著. —北京:清华大学出版社,2005.3

(硅谷精英科技丛书)

ISBN 7-302-10222-8

I. 网… II. 赵… III. 计算机网络—信息存储 IV. TP393.0

中国版本图书馆 CIP 数据核字(2004)第 140841

出 版 者:清华大学出版社

<http://www.tup.com.cn>

社 总 机:010-62770175

地 址:北京清华大学学研大厦

邮 编:100084

客 户 服 务:010-62776969

责任编辑:薛 慧

封面设计:常雪影

印 装 者:清华大学印刷厂

发 行 者:新华书店总店北京发行所

开 本:185×260 印张:16.75 字数:382千字

版 次:2005年3月第1版 2005年3月第1次印刷

书 号:ISBN 7-302-10222-8/TP·1106

印 数:1~5000

定 价:29.00元

Silicon Valley Elite Science =====
===== *and Technology Series*

硅谷精英科技丛书编委会

荣誉顾问 Stephen Lu ([美]南加州大学, 大尉派克主席教授)
王其文 (北京大学, 光华管理学院副院长, 教授)

主 编 蔡 剑 (北京大学)

副主编 赵文辉 (Cisco) 刘 为 (Oracle)

编 委 (按姓氏笔画为序)

王 炜 (Oracle)	刘 为 (Oracle)
李 晨 (Cisco)	陈清华 (Cisco)
林成江 (BEA)	周伟华 (Oracle)
周加林 (Cisco)	赵 刚 (Microsoft)
赵文辉 (Cisco)	赵武钢 (Oracle)
徐 俊 (Cisco)	康 蔚 (Ultratech)
景 楠 (USC)	蔡 剑 (北京大学)

Silicon Valley Elite Science =====
===== *and Technology Series*

Preface —●

The market economy over the past twenty years has set free the potentials of Chinese society and unleashed the talents of Chinese people. Modern Chinese companies are moving up the technology ladder at an unprecedented pace among all nations. The rapid developments of computer industries in China, for example, will put the country on a fast track to become one of the world's largest IT producers and consumers. Last year, Chinese universities granted near five million science and engineering college degrees, a number that is approaching the total for the United States. The Chinese government has been sponsoring several large software industry parks, which will soon become a major contributor in China's new economy. With all these amazing developments continue, the global industrial landscape and leadership will for sure be permanently recast.

There are two critical requirements that can help to further expedite the development process of IT industries in China. First, we should have the right vision to target at those promising, high value added technologies that can result in big market successes to Chinese enterprises. Second, we must learn directly from the global elite sources, which can provide not only the state-of-the-art knowledge but also the hands-on experiences of the leading technologies. While facing with amazing progresses, we must admit that there are still knowledge gaps when Chinese IT professionals try to grasp and understand those rapidly advancing technologies. With traditional approaches of studying and training that have to go through translations and editing, the fast-evolving knowledge often has aged when they reach the Chinese professionals. Therefore, it is valuable to explore a "fast lane" for knowledge transfer and build a "bridge over the gaps". The "*Silicon Valley*

Elite Science and Technology Series” is indeed such a fast lane and bridge to promptly bring advanced knowledge and industrial experiences from the global sources to the Chinese communities.

As the center of the world’s high-tech developments, Silicon Valley in the United States has been very successful in generating break-through technologies, which become the foundation for the leading IT industries. Many overseas Chinese scientists and engineers have played very critical roles in both R&D and entrepreneurship to build successful high-tech companies at Silicon Valley. While enjoying their career successes, they also wish to assume active roles in the economic and technology developments in their homeland. When Dr. Jian Cai, a former student of mine, and his colleagues from the CO-YES (Chinese Overseas Young Elite Society) group invited me to serve as an honorary adviser of the series, I was very moved by their mission and devotion. I am glad to support them with my 20 years of experience as a high-tech entrepreneur and academician. This book series is the first comprehensive collection that is written in Chinese by Chinese engineers and researchers from the leading Silicon Valley IT companies (e. g. Sun, Oracle, Cisco). It covers many important subjects, such as web application design, database systems, storage network, etc. The authors’ years of R&D experiences in the western world, their combined Chinese and American advanced education, and their strong devotion to their homeland will certainly bring unique value to the Chinese IT professionals. I am sure that many Chinese technology elites will enjoy, and benefit from, this book series as China continues to move to become the “Silicon Valley of the World”.

Stephen C-Y. Lu, Ph.D.

David Packard Chair Professor
University of Southern California
Los Angeles, California

April 2003

硅谷精英科技丛书

序

市场经济在过去 20 年解放了中国社会的发展潜力，也使中国人的天分得到了充分发挥。中国的现代企业正在以其他国家望尘莫及的速度攀登科技阶梯。例如，计算机行业在中国的迅速发展将使这个国家快速成为世界上最大的 IT 生产商和消费者之一。去年，中国大学培养了约 500 万个科学和工程专业的大学生，这个数字已经接近美国。中国政府也正在大力发展几个大软件产业园以提供新的经济增长点。如果所有这些惊人的高速发展持续下去，全球工业格局和领导潮流会产生本质的巨变。

要进一步加速中国 IT 产业的发展，有两个要素需要被重视。首先，我们应该着眼那些新兴的、能对中国企业提供高增值并实现大规模市场成功的技术。其次，我们必须直接地向全球的顶尖资源学习，这些资源不但能够提供世界先进水平的知识而且能够介绍有关主导技术的实践经验。当科技发展一日千里时，我们必须承认在中国 IT 人才试图迅速地了解和掌握先进技术的过程中存在着“知识鸿沟”。传统的学习和培训方法依赖翻译和编辑的手段，使用这些传统方法，当快速不断更新的知识流传到中国时往往有些过时。因此，建立知识传递的“快车道”和“跨越鸿沟的桥梁”是很有价值的。“硅谷精英科技丛书”正是这样一条快车道和一座桥梁，它及时地将全球的先进知识和工业界经验带给中国。

作为世界高科技发展的中心，美国硅谷非常成功地研发出很多突破性技术，这些技术成为领先的 IT 产业的基础。许多海外的中国科学家和工程师在硅谷高科技公司的研发和管理中扮演了重要的角色。在事业上取得成功的同时，他们也希望积极参与祖国的经济和技术发展。当我先前的学生蔡剑博士和他来自 CO-YES 学会的同事们邀请了我担当这套丛书的名誉顾问时，我为他们的使命感和奉献精神而感动。我很高兴能以我作为一名高科技企业家和学者的 20 年经验帮助他们。这套丛书是首部由来自硅谷顶尖 IT 公司（如

Silicon Valley

Sun, Oracle, Cisco) 的中国工程师和研究员用中文编写的系列书籍。它包括了许多重要题目, 譬如网络应用设计、数据库系统、存储网络, 等等。作者们在西方多年的研发经验, 他们所受的中美高等教育, 以及他们对祖国的强烈的奉献精神将为中国的 IT 专业人才带来独到的价值。我肯定, 许多中国的技术精英将会喜欢这套丛书并从中受益, 因为中国正在不断努力成为“世界的硅谷”。

Stephen C-Y. Lu (卢志扬), 博士
大尉派克主席教授
美国南加州大学
洛杉矶, 加州

2003 年 4 月

前言



网络存储技术是最近几年高科技行业最热门的技术之一。随着计算机技术和网络技术的发展,越来越多的信息被数据化。这些数据化的信息需要能长时间保存,并且能快速方便地检索。电子商务、电子政务等信息化技术的推广对数据的存储容量、速度以及安全提出了更高的要求。存储技术也从本地存储发展到网络存储。网络存储技术正处在高速发展的阶段。本书深入介绍了存储技术的基本原理、体系结构和设计存储系统的方法以及实例分析。

本书适合于用到或将要用到存储技术的信息系统开发或应用的工程技术人员,希望深入了解存储技术的技术人员和大专院校学生。随着社会的信息化,越来越多的数据需要存储。对于大型信息系统,网络存储将是必然的选择方案。比较大的企业信息系统或电子商务系统、政府的电子政务系统等都需要存储系统来存储大量的数据。对于从事或希望从事电子商务、电子政务等信息系统开发、系统集成以及维护的人员,了解网络存储技术能帮助他们设计扩展性好、更适合未来发展的系统。

第1章介绍了存储模型的发展历史和分类。大致介绍了内嵌式存储系统、直接存储系统、联网存储系统和存储区域网络这四类存储模型的特点。第2章简单介绍数据存储系统中的三种物理设备:存储设备,存储设备接口技术和存储互联设备。

第3、4章介绍了数据存储技术中的基础技术:独立冗余磁盘阵列和SCSI技术。独立冗余磁盘阵列是存储系统中用于存储数据的最常见设备。SCSI技术是数据存储设备与数据操作者之间最常用的通信协议,是其他数据存储协议的基础。

第5章是关于文件系统的基本知识。文件系统以文件的形式隐藏了存储设备的物理细节。它很多是应用程序存取存储设备中数据的接口。本章介绍了常见的Windows和Unix文件系统,特别是网络文件系统。而网络文件系统是联网存储系统的基石。

第6章详细介绍了联网存储系统的原理、系统结构和设计原则,并给出了联网存储系统的一些设计范例。最后还介绍了联网存储系统的一些最新发展方向。

第7章详细介绍了存储技术中的另一常见基础通信协议：光纤通道协议。光纤通道协议是存储区域网系统中的基本通信协议。它为存储设备和需要存取数据的服务器提供了一个高速数据交换通道。本章介绍了光纤通道协议中 FC-0, FC-1, FC-2, FC-3 和 FC-4 层的详细协议内容,同时也讲述了光纤通道协议中的流控制、不同的服务类型和光纤通道登录等概念。

第8章首先介绍了存储区域网系统的概念和它的特点。然后详细介绍了存储区域网的体系结构,分别介绍了其中的服务器连接器件、存储网络连接器件和存储软件。这一章中还提到存储区域网设计中常用的几种典型拓扑结构和它的一些应用场合。最后还介绍了存储区域网的主要产品,以及选用这些产品时的一些考虑因素。存储区域网中的关键设备是光纤通道交换机。

第9章介绍了光纤通道交换机的工作原理和组织结构,同时也介绍了存储区域网系统中的光纤通道路由选择算法。在第9章的最后介绍了市场上的几种主要光纤通道交换机产品。

存储系统的管理和虚拟化是存储技术的热点。第10章介绍了存储管理的概念、技术基础和体系结构。同时也讲述了存储管理的三种管理功能:存储网络管理、存储设备管理和数据管理。最后还介绍了存储管理的一些应用。第11章介绍了存储虚拟化的概念和特点。讲解了存储虚拟化的体系结构和实现方式,同时也提到了存储虚拟化的实际应用。

第12、13章分别介绍了基于IP通信的存储技术中的两种基础技术:iSCSI和FCIP。这两章详细介绍了iSCSI和FCIP协议的内容及其特点。

第14章介绍了存储产品的概况,包括存储协议和标准的发展规律趋势、存储交换机、磁盘阵列等产品的市场概况。

第15章介绍了存储系统的一些设计要点,给出了几个实际的数据存储系统的设计和实现案例。

希望读者在读完本书后,对存储技术的原理以及应用能有更深入的了解。希望本书能对读者设计和实现存储系统有一定的帮助作用。

作者

登录“酷业网”(www.coolyes.com.cn)学习本书课程

- 作者讲述
- 多媒体课件
- 在线讨论
- 相关习题
- 教师 ppt
- 演示程序

目录

第 1 章 存储模型的发展	1
1.1 内嵌式存储系统	1
1.2 直接存储系统	2
1.3 联网存储系统	3
1.4 存储区域网络	4
1.5 全球存储产品市场概貌	5
第 2 章 存储技术原理简介	8
2.1 存储设备	8
2.1.1 磁带库	8
2.1.2 光盘存储设备	9
2.1.3 磁盘存储设备	9
2.2 存储设备接口技术	10
2.2.1 SCSI	10
2.2.2 iSCSI	11
2.2.3 光纤通道	12
2.3 SAN 互联设备	13
2.3.1 SAN 交换机	13
2.3.2 SAN 路由器	13
第 3 章 独立冗余磁盘阵列(RAID)	14
3.1 RAID 的内部 I/O 结构	15
3.2 RAID 的基本特性	16
3.2.1 RAID 容量的扩展性	16
3.2.2 RAID 的可管理性	17
3.2.3 RAID 的高性能	17
3.2.4 RAID 的可靠性	18
3.3 RAID 分区及数据组织	19

3.3.1	分区和阵列	19
3.3.2	分块和分条	20
3.4	RAID 级别	20
3.4.1	RAID 0	21
3.4.2	RAID 1	21
3.4.3	RAID 2	21
3.4.4	RAID 3	22
3.4.5	RAID 4	22
3.4.6	RAID 5	22
3.4.7	RAID 6	23
3.4.8	RAID 组合级别	23
第 4 章	SCSI	25
4.1	SCSI 简介	26
4.1.1	SCSI 标准与标准机构	26
4.1.2	SCSI 协议框架	27
4.1.3	SCSI 发展现状	28
4.2	SCSI 物理层协议标准	30
4.2.1	SCSI 的电信号方式	30
4.3	SCSI 系统与设备	30
4.3.1	SCSI 域	30
4.3.2	SCSI 设备控制器	31
4.4	SCSI 分层通信协议结构	32
4.5	SCSI 基础——信号与地址	33
4.5.1	SCSI 总线信号	33
4.5.2	SCSI 的编址和 LUN	34
4.6	SCSI 总线的各个阶段	35
4.6.1	总线空置阶段	35
4.6.2	总线仲裁与总线仲裁阶段	36
4.6.3	SCSI 的选择阶段	38
4.6.4	总线重选阶段	39
4.6.5	命令、数据、状态和通信阶段	39
4.7	同步和异步信息传输	40
4.7.1	异步传输	40
4.7.2	同步传输	41
4.7.3	时钟方式	42
4.8	其他技术细节	43
4.8.1	SCSI 连接头	43

4.8.2	总线终止技术	43
4.9	SCSI 传输层协议	44
4.9.1	命令描述块(CDB)	44
4.9.2	逻辑单元	46
4.10	总体描述	47
4.10.1	数据读	47
4.10.2	数据和文件的写操作	48
4.10.3	读写的系统整体描述	49
4.11	SCSI 应用	49
4.11.1	OS 内核	49
4.11.2	数据库引擎和存储管理系统	50
4.12	SCSI 的持续发展	50
4.12.1	为什么要串行 SCSI	51
4.12.2	串行 SCSI 的概念	51
第 5 章	文件系统	52
5.1	文件系统介绍	52
5.1.1	文件命名规则	54
5.1.2	文件系统和相关的系统	54
5.2	操作系统的文件系统	55
5.2.1	FAT	55
5.2.2	NTFS	58
5.2.3	其他文件系统	61
5.2.4	各个文件系统的文件共享——虚拟文件系统	62
5.3	网络文件系统	62
5.3.1	NFS 文件系统	63
5.3.2	CIFS 文件系统	70
第 6 章	NAS	77
6.1	NAS 介绍	77
6.1.1	NAS 与 DAS 的比较	78
6.1.2	NAS 设备	80
6.1.3	文件器	81
6.1.4	NAS 与网络	82
6.1.5	NAS 运行	82
6.1.6	NAS 优点总结	83
6.2	NAS 设备设计	85
6.2.1	操作系统	86

网络 存储 技术	6.2.2 应用软件和协议	86
	6.2.3 网络协议和网络结构	87
	6.3 网络设计与存储设计	87
	6.3.1 局域网和广域网	88
	6.3.2 文件的集中化过程	88
	6.3.3 网络存储的可靠性	89
	6.4 NAS 管理	90
	6.5 NAS 典型产品介绍	92
	6.6 NAS 范例——网页	95
	6.6.1 介绍	95
	6.6.2 基本操作	95
	6.6.3 文件复制	96
	6.6.4 测试	97
	6.6.5 CGI 脚本	97
	6.7 网络存储的新发展	98
	第 7 章 光纤通道	100
	7.1 光纤通道概述	100
	7.2 光纤通道结构	102
	7.3 光纤通道端口类型	103
	7.4 光纤通道的拓扑结构	103
	7.4.1 点到点直接连接	103
	7.4.2 环型结构	104
	7.4.3 交换结构	108
	7.5 光纤通道的协议堆栈	109
	7.5.1 FC-0 层	111
	7.5.2 FC-1 层	112
	7.5.3 FC-2 层	114
	7.5.4 FC-3 层	116
	7.5.5 FC-4 映射层	117
	7.6 流控制	118
	7.7 服务类型	119
	7.8 光纤通道地址	122
	7.9 光纤通道登录	123
	7.9.1 光纤通道通用服务	124
	7.9.2 FCIP	124
	第 8 章 存储区域网	126
	8.1 什么是存储区域网(SAN)	126

8.2	为什么需要存储区域网	127
8.3	SAN 的体系结构	129
8.3.1	服务器连接器件	129
8.3.2	存储网络连接器件	131
8.3.3	存储软件	132
8.3.4	SAN 拓扑结构	134
8.3.5	存储区域网的应用	137
8.4	存储区域网的主要产品介绍	142
8.4.1	存储设备	143
8.4.2	存储路由器	144
8.4.3	主机总线适配卡	145
8.4.4	光纤通道交换机	147
第 9 章	光纤通道交换机	148
9.1	光纤通道交换机基本工作原理	149
9.2	光纤通道交换机结构	150
9.2.1	端口类型	151
9.2.2	光纤通道交换机地址	153
9.2.3	域名服务器	154
9.2.4	主交换机	155
9.3	光纤通道路由选择	155
9.3.1	HELLO 协议	156
9.3.2	拓扑结构数据库	156
9.3.3	最短路径优先算法	158
9.3.4	自治区域	158
9.3.5	存储分区	159
9.4	存储交换机产品介绍	160
第 10 章	存储管理	166
10.1	存储管理的发展历史	167
10.2	存储管理技术基础	168
10.3	存储管理功能	169
10.3.1	存储网络管理	169
10.3.2	存储设备管理	171
10.3.3	数据管理	172
10.4	存储管理体系结构	172
10.5	存储管理应用	178

网络 存储 技术	第 11 章 存储虚拟化	180
	11.1 存储虚拟化的优点	181
	11.2 存储虚拟化的体系结构	182
	11.3 存储虚拟化的实现	183
	11.4 存储虚拟化应用	185
	第 12 章 iSCSI 协议	191
	12.1 什么是 iSCSI	191
	12.2 iSCSI 的优点	192
	12.3 iSCSI 概述及术语简介	195
	12.3.1 iSCSI 概论	195
12.3.2 协议层(layer)和进程(session)	196	
12.3.3 序号和 iSCSI 标号	196	
12.3.4 iSCSI 的登录	197	
12.3.5 iSCSI 的全功能阶段	197	
12.3.6 数据的传输	197	
12.3.7 iSCSI 连接的中断	197	
12.4 iSCSI 组成整体结构	198	
12.4.1 SCSI 到 iSCSI 对应的模型	199	
12.4.2 iSCSI 命名和地址方法	200	
12.4.3 iSCSI 的协议数据单元(PDU)	203	
12.4.4 iSCSI 的错误处理及错误恢复	208	
12.4.5 iSCSI 错误	211	
12.5 iSCSI 安全	212	
12.5.1 输入端的认证技术	213	
12.5.2 CHAP 认证	213	
12.5.3 IPSec 简介	214	
第 13 章 iFCP 协议	217	
13.1 什么是 iFCP	217	
13.2 iFCP 出现的背景	218	
13.3 iFCP 和 FCIP 的异同	219	
13.4 iFCP 的网络模型	220	
13.5 iFCP 的通信模型	221	
13.6 iFCP 的协议整体结构	222	
13.6.1 iFCP 的数据包结构	222	
13.6.2 iSNS	222	
13.6.3 网关区域的角色	224	

13.7 iFCP 网关的应用例子	226
第 14 章 存储产品概况	229
14.1 协议及标准化方面的趋势	230
14.2 交换机市场	230
14.3 磁盘阵列市场	232
14.4 下一步：全方位的智能存储	232
第 15 章 存储区域网设计和案例	235
15.1 SAN 设计要点	235
15.2 案例 1：数据备份/恢复	236
15.3 案例 2：IP 存储区域网	239
15.4 案例 3：存储虚拟化	242
15.5 SAN 设计总结	246
参考文献	248