

EMC²
where information lives

信息存储与管理

数字信息的存储、管理和保护

G.Somasundaram Alok Shrivastava 著

罗英伟 汪晓林 尹冬生 黄剑 刘凯 路新喜 冉崇海 王永康 译

EMC Education Services

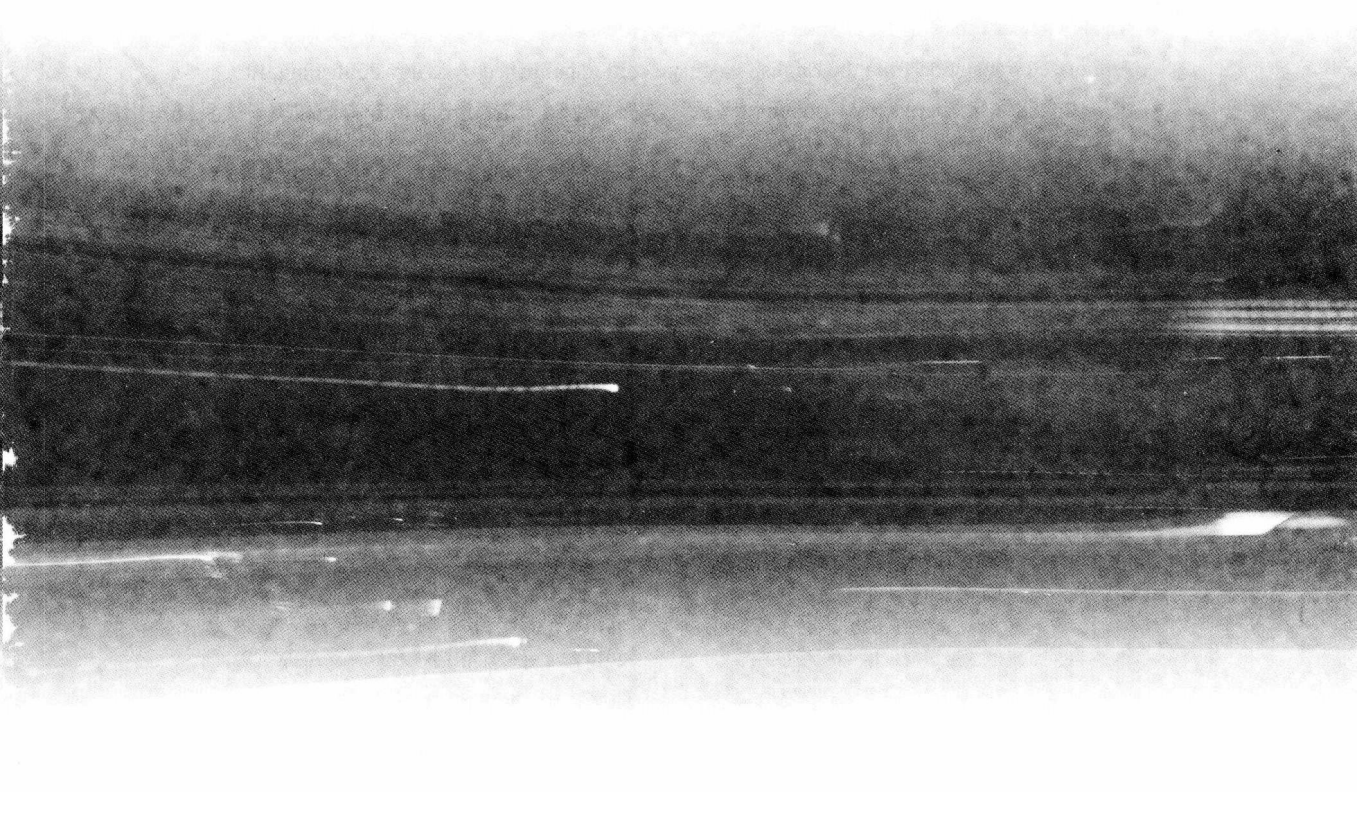
 人民邮电出版社
POSTS & TELECOM PRESS

信息存储与管理

数字信息的存储、管理和保护

G.Somasundaram Alok Shrivastava 著

罗英伟 汪晓林 尹冬生 黄剑 刘凯 路新喜 冉崇海 王永康 译



图书在版编目 (C I P) 数据

信息存储与管理：数字信息的存储、管理和保护 /
(美) 萨曼达 (Somasundaram, G.), (美) 希瓦史塔瓦
(Shrivastava, A.) 著；罗英伟等译. — 北京：人民邮
电出版社, 2010. 1
ISBN 978-7-115-21748-6

I. ①信… II. ①萨… ②希… ③罗… III. ①信息管
理—研究 IV. ①G203

中国版本图书馆CIP数据核字(2009)第216679号

版权声明

Information Storage and Management

Copyright © 2009 by EMC Corporation

All rights reserved. This translation published under license.

Authorized translation from the English language edition published by Wiley Publishing, Inc.

本书中文简体字版由 Wiley Publishing 公司授权人民邮电出版社出版，专有版权属于人民邮电出版社。

信息存储与管理：数字信息的存储、管理和保护

◆ 著 G.Somasundaram Alok Shrivastava
译 罗英伟 汪晓林 尹冬生 黄剑 刘凯
路新喜 冉崇海 王永康

责任编辑 俞彬

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京画中画印刷有限公司印刷

◆ 开本：800×1000 1/16

印张：30.25

字数：638千字

印数：1—4000册

2010年1月第1版

2010年1月北京第1次印刷


著作权合同登记号 图字：01-2009-5777号

ISBN 978-7-115-21748-6

定价：88.00元

读者服务热线：(010)67132705 印装质量热线：(010)67129223

反盗版热线：(010)67171154



内容提要

信息的管理和安全对于企业的成功至关重要，它已经发展成为一门高度成熟和复杂的 IT 支柱产业和学科，本书是一本系统全面的存储技术书籍。

本书所涵盖的概念、原理和方案部署理念贯穿了整个信息存储和管理技术的范畴，本书包含 4 个部分，主要阐述以下方面的内容：数据存储和数据管理面临的挑战，智能化存储系统，存储网络（EC-SAN、IP-SAN、NAS），备份、恢复和归档（包括 CAS），业务连续性和冗灾，存储安全和虚拟化，存储基础架构的管理和监控等。

本书适合作为高校信息管理专业的教材或参考书，同时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。

作者简介

G.Somasundaram (Somu)毕业于孟买的印度理工学院。他具有 22 年的 IT 从业经验，其中后 10 年一直效力于 EMC 公司。他目前是 EMC 全球服务部的主管，引领着全球业界的培训创新。Somu 是 EMC 存储开放课程的架构师，该课程旨在填补 IT 业界关于存储知识缺失的鸿沟。在他的率领和指导下，EMC 学习伙伴（EMC Learning Partner）、学院联盟（EMC Academic Alliance）等业界培训创新项目如火如荼地持续发展壮大，为全球信息存储和管理技术领域培养了成千上万的优秀学生。Somu 所负责的关键领域包括指导全球专家小组、发掘全球 IT 教育提供商并建立合作以及为 EMC 的业界培训创新制定总路线。在此之前，Somu 曾担任 EMC 以及其他 IT 领军厂商的多个管理及领导职位。

Alok Shrivastava 是 EMC 全球服务部门的资深主管，他从 2003 年开始致力于教育领域。Alok 是 EMC 许多成功教育创新计划的架构师，如业界领先的 EMC 专家认证项目、EMC 学院联盟培训项目以及这本关于信息存储技术的独特而宝贵的书籍。Alok 以其卓识远见领导着一个由高水平专家组成的团队，共同为 EMC 员工、合作伙伴、客户和其他业界专业人士提供世界一流的技术教育。在此之前，Alok 曾在亚太地区和日本组建并率领了一个富有成就的 EMC 售前工程师团队。Alok 在早期还曾作为系统管理员、存储管理员以及备份和灾难恢复顾问，效力于全球许多大型数据中心。他获得了印度理工学院（Indian Institute of Technology）和印度 Sagar 大学（University of Sagar）的双硕士学位。在超过 25 年的 IT 从业生涯中，Alok 一直对信息存储技术领域怀着情有独钟的热情。

中文版序

两年前，我在考虑 EMC 中国五年发展规划时曾经提出，在人才培养方面，希望到 2011 年，参与 EMC 学院联盟计划、开设《信息基础架构技术》课程的高校达到 50 所。现在看来，这个计划太保守了。我们只用了一年多时间，在 2008 上半年就已经达到了这一目标。

为什么会有这么快的发展速度呢？除了我们的团队有较强的执行力外，更为重要的原因是，中国对于信息管理人才有巨大的需求，推动各高等院校迅速行动，开设《信息基础架构技术》课程。

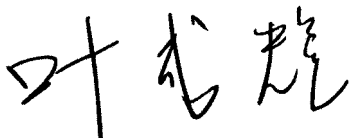
根据国际数据公司（IDC）的研究，全球数字信息量每年以 60% 的速度增长，全球数字信息有 70% 是由个人创造的，其中的 85% 需要由机构来管理。中国的增长比全球平均速度还要快。中国是全球人口最多的国家，互联网用户数、手机用户数都名列全球第一，PC 市场名列全球第二，MP3/MP4、DVD、数码相机、数码照相机、数字电视等信息终端设备广泛被采用，这些原因都造成了数据信息量的加速膨胀。这就意味着，我们需要大量的信息管理专业技术人才。而之前的课程体系在这方面几乎空白，其结果是，人才缺口严重。

几年前，EMC 就认识到，信息管理专业人才匮乏将是一个全球性、长期性的行业问题。作为全球信息基础架构技术和解决方案的领导者，EMC 感到责无旁贷。为此，EMC 与全球顶尖的教育机构合作开发了《信息基础架构技术》课件。作为一门开放式的课程，它着眼于行业基础理论和基础知识，并以最新产品为实例，理论与实际相结合，让学生一出校门就能投入实际工作。这一课程一经推出，就受到全球

众多合作高校师生的欢迎。

在之后的教育实践中，我们深刻感受到，只有课件、没有教材的状况，不利于学生学习和掌握相关知识。于是，EMC 组织专家抓紧时间编写，2009 年 3 月在美国出版了英文版教材 *Information Storage and Management*。在英文版出版的同时，我们也在第一时间将它翻译成中文，并得到出版社的大力支持与合作，让《信息存储与管理：数字信息的存储、管理和保护》成为该教材在全球的第一个非英语版本。

我们相信，《信息存储与管理：数字信息的存储、管理和保护》这本书对于开设“信息基础架构技术”课程的高校师生，对于各位有志于从事信息管理工作的朋友，以及已经在从事信息管理工作、希望系统地学习相关知识的朋友，都非常有用。我们希望它能帮助你的职业生涯快速进步，帮助你所在的机构提高信息管理效率和效益。

Handwritten signature of Ye Weigang in black ink, consisting of three characters: 叶, 威, 刚.

EMC 大中华区总裁

原 书 序

美国著名的思想家、文学家和诗人 **Ralph Waldo Emerson** 曾经说过：智慧永恒不变的特点就是在平凡中发现奇迹。而今，无数平凡的奇迹正充盈在我们身边，事实上我们根本无法忽视它们的存在。我们每个人都拥有许多电子“小玩具”，如数码相机、数字摄像机、移动电话、能访问成千上万网站的高速计算机，以及即时聊天工具、社交网站、搜索引擎、音乐下载等，不胜枚举。所有这些例子都有一个共同点：它们都会产生大量的数据。我们所生活的时代不仅仅是一个信息时代，而且是一个信息爆炸的时代。在这个时代里，爆炸的信息形成了一个数字的宇宙，它呼唤着技术的革新，渴求着能够对信息存储与管理的解决方案实行高效管理、支撑和优化的新一代专家。

我们今天所面临的挑战究竟有多大呢？让我们对它有一个直观的认识吧：我们在一年内所建立、采集、复制的数字信息，就相当于古往今来所有书籍所容纳信息的数百万倍！信息是商业活动中最重要的资本。为了挖掘信息的巨大潜能，我们必须将它们有效地、聪明地存储、保护和管理起来，使之能够被访问、搜索、共享，并最终发挥效用。

我们正处在一场大风暴中，一切都在急速地增长：信息、成本以及能够对这些信息实行有效存储和管理的熟练专家——然而他们的数量还远未能满足激增的需求。IT 管理员首当其冲的顾虑就是如何应对存储的增长，因为企业并不能简单地依靠购买更大更好的“箱子”来存储它们的数据。IT 管理员不仅仅要在存储技术的预算上精打细算，还要考虑到能耗、安装、管理及维护成本等诸多因素。此外，IT 管

理员们还需要聘用更多员工，因为他们非常缺少有丰富经验的、受过存储技术系统教育的专家们来帮助他们管理和优化数据。

我还未曾在市场上见到一本全面而深入介绍存储和管理数据的各种技术的书籍。作为业界的领航者，我们具备的专业背景及实践经验将有助于填补这个空白。这本书能够让读者对信息存储与管理技术的台前幕后有一个深入的认识。读者将从中学到数据存储在哪、它们是如何被管理的以及如何能帮助企业提高盈利额。

如果你的职业正是存储及信息基础设施管理，那么你将注定是一个日新月异领域的先锋。在这个领域里，全新的挑战总是与极大的回报并存。

无论你当前在 IT 领域扮演什么角色，本书都将在你的 IT 藏书和职业发展中发挥着不可或缺的作用。

Thomas P. Clancy

EMC 公司教育部门副总裁

2009 年 3 月



前 言

信息的存储是信息技术的中流砥柱。无论是个人用户还是企业 IT 用户，他们时时刻刻都在产生着海量的数字信息。我们需要对这些信息进行存储、保护、优化和管理。

曾几何时，人们对信息存储的认识还仅仅是连接到电脑主机背后的那一摞摞用于保存数据的磁盘或磁带。即使在今天，也只有那些存储业内人士才深知信息存储技术对于 IT 基础设施的可用性、高性能、完整性和信息优化所发挥的至关重要的作用。近 20 年来，信息存储技术得到了长足的发展，各种新兴的尖端技术为数字信息的存储、管理、互联、保护、安全、共享和优化提供了丰富多彩的解决方案。

伴随着信息的爆炸式增长以及日益复杂化的产品与解决方案的涌现，人们对信息存储专家的需求也在不断扩大。在如何雇佣和培训信息存储领域的专业人才方面，IT 管理员们正面临着严峻的挑战。

尽管许多一流大学已经开始在其计算机或信息技术专业的教学计划中加入了专门讲授存储技术的课程，但遗憾的是，当今的许多 IT 专家们，甚至包括那些有着多年经验的专家，都未曾获益于这种正规教育。因此，包括应用、系统、数据库及网络管理等多个领域在内的资深专家们，对于存储技术对各自领域所产生的影响并不具备一个统一的认知基础。

我们编写本书的目的在于：使学生们和专业人士们对存储技术的各个环节形成一个完整全面的认识与理解。尽管本书的实例部分采用的是 EMC 公司的产品，但是，读者通过本书形成的对于技术概念及其基本原理的理解，必将有助其轻松地掌握其他公司的相关产品。

本书分为 4 个部分，共 16 章。其中，高级专题是基于其前面章节所掌握的内容深化而来的。

第 1 部分，存储系统：这 4 章讲述了信息增长及其带来的挑战，给出了一个存储系统及其运行环境，回顾了存储技术的发展历程，并介绍了智能存储系统。

第 2 部分，存储网络技术和虚拟化：这 6 章涵盖了 SCSI 和光纤通道架构、直连存储 (Direct-attached Storage, DAS)、存储区域网 (Storage Area Networks, SAN)、网络互联存储 (Network-attached Storage, NAS)、IP 存储区域网 (Internet Protocol SAN, IP-SAN)、内容寻址存储 (Content-addressed Storage, CAS) 以及存储虚拟化。

第 3 部分，业务连续性：这 4 章介绍了业务连续性、备份及恢复、本地数据副本以及远程数据副本。

第 4 部分，存储安全与管理：这 2 章涵盖了存储安全、存储基础设施的监控和管理。

此外，本书的网站还提供了一些最新的教辅资源和阅读材料，有兴趣的读者可以访问“<http://education.EMC.com/ismbook>”，以获取更详尽的信息。

EMC 学院联盟

我们真诚地邀请有意教授《信息存储与管理：数字信息的存储、管理和保护》课程的大专院校参与 EMC 学院联盟计划。本计划为教学机构提供了广泛而全面的支持，包括教学辅助、教员培训、学生实践项目等。具体内容请访问“<http://education.EMC.com/academicalliance>”。

EMC 认证专家证书



本书是 EMC 信息存储及管理专家认证考试 (E20-001) 的辅导教材。EMC 专家认证是验证读者所学知识并提升其在信息技术产业信誉度的金牌认证项目。有关此认证的更多信息以及如何获取考试练习，请访问“<http://education.EMC.com>”。

致 谢

当我们开始本书的编撰时，首当其冲的挑战就是如何寻找一个对构成现代信息存储基础设施的广泛技术有着全面了解的专家队伍。

激发并支持我们完成本书的一个关键因素就在于：在 EMC，我们有着丰富的技术和经验，以及许许多多业界最优秀的人才。和这些专家交流时，他们都和我们一样对出版这本全面涵盖信息存储技术的书籍的前景倍感兴奋，因为这是使他们能够与全球的专家和学生们分享他们专业知识的一个难得的契机。

本书是在 EMC 教育服务部的指导下，在 CTO 办公室、全球市场部、EMC 工程部的支持下，由 EMC 许多关键部门的努力和奉献所凝聚的结晶。

本书的初稿是由来自 EMC 教育的服务部的 Ganesh Rajaratnam 根据自己的调研和专业知识，并率领其他相关专家共同努力而完成的。来自 EMC CTO 办公室的 David Black 博士，利用其宝贵的时间审阅了全书的内容，并就本书所囊括的许多关键问题提出了中肯的建议。

下列来自 EMC 教育服务部门的专家们参与编撰了本书的某些章节，我们在此向他们表示衷心的感谢：

Rodrigo Alves

Anbuselvi Jeyakumar

Charlie Brooks

Sagar Kotekar Patil

Debasish Chakrabarty

Andre Rossouw

Diana Davis

Tony Santamaria

Amit Deshmukh	Saravanaraj Sridharan
Michael Dulavitz	Ganesh Sundaresan
Ashish Garg	Anand Varkar
Dr. Vanchi Gurumoorthy	Dr. Viswanth VS
Simon Hawkshaw	Jim Tracy

以下专家在编撰的各阶段全面地审阅了本书，并提出了许多宝贵的反馈和指导：

Ronen Artzi	Jack Harwood
Eric Baize	Arthur Johnson
Greg Baltazar	Michelle Lavoie
Edward Bell	Tom McGowan
Christopher Chaulk	Jeffery Moore
Roger Dupuis	Toby Morral
Deborah Filer	Peter Popieniuck
Bala Ganeshan	Kevin Sheridan
Jason Gervickas	Ed VanSickle
Nancy Gessler	Jody Goncalves

我们还要感谢 NIIT 有限公司为本书初稿提供的帮助，感谢 EMC 的 **Muthaiah Thiagarajan** 和 DreaMarT Interactive Pvt.有限公司为本书制作了所有的插图，感谢本书的出版商 John Wiley & Sons 为将本书出版付印所做的不懈努力。

—— **G. Somasundaram**, EMC 公司教育服务部主管

—— **Alok Shrivastava**, EMC 公司教育服务部资深主管

2009 年 3 月

本书中文版编校人员

翻译人员名单

罗英伟 北京大学信息科学技术学院 教授
汪晓林 北京大学信息科学技术学院 副教授
尹冬生 中山大学信息科学与技术学院 讲师
黄 剑 中山大学信息科学与技术学院 副教授
刘 凯 西安电子科技大学计算机学院 副教授
路新喜 北京航空航天大学软件学院 信息化办公室主任
冉崇海 EMC 企业系统工程部 企业系统工程师
王永康 EMC 大中华技术解决方案中心 教育推广经理

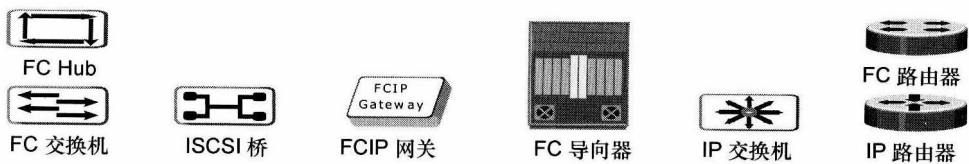
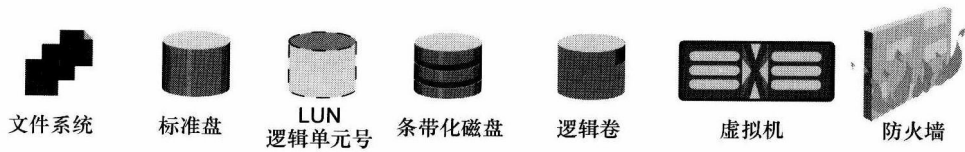
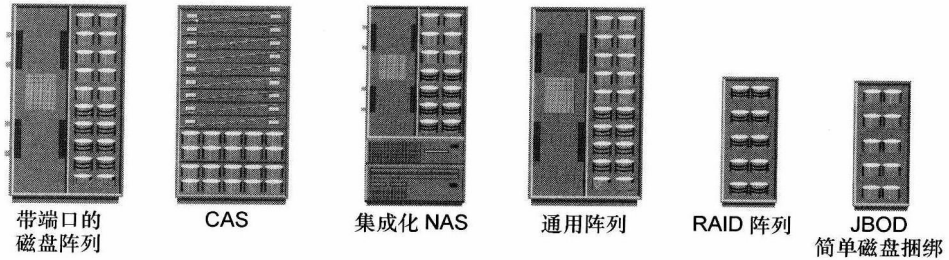
易安信电脑系统（中国）有限公司 易安信中国卓越研发集团

审校

审校人员名单（排名按翻译及审校顺序）

周恣旻 EMC 卓越研发中心软件工程师
周海忠 EMC 卓越研发中心高级软件工程师
潘家铭 EMC 卓越研发中心高级软件工程师
周应超 EMC 卓越研发中心高级软件工程师
汤海鹰 EMC 卓越研发中心软件工程师
张 渊 EMC 卓越研发中心高级软件工程师
杨舜华 EMC 解决方案验证中心高级顾问
张文天 EMC 解决方案验证中心助理顾问
时 祎 EMC 全球解决方案中心高级解决方案工程师
柴可夫 EMC 卓越研发中心软件工程师
赵军平 EMC 卓越研发中心高级软件工程师
鲁凤成 EMC 卓越研发中心高级软件工程师
张 鹏 EMC 卓越研发中心高级软件工程师
葛 琳 EMC 大中华技术解决方案中心高级技术顾问
吕国强 EMC 大中华技术解决方案中心技术顾问
王永康 EMC 大中华技术解决方案中心教育推广经理

本书图标 说明



目 录

第 1 部分 存储系统

第 1 章 信息存储与管理的介绍	3
1.1 信息存储	5
1.1.1 数据	5
1.1.2 数据类型	7
1.1.3 信息	7
1.1.4 存储	8
1.2 存储技术和架构的发展	9
1.3 数据中心基础设施	10
1.3.1 核心部件	10
1.3.2 数据中心部件的关键需求	11
1.3.3 管理存储基础设施	13
1.4 信息管理中的关键挑战	14
1.5 信息生命周期	14
1.5.1 信息生命周期管理	15
1.5.2 ILM 实现	16
1.5.3 ILM 的优点	17
小结	18
第 2 章 存储系统环境	21
2.1 存储系统环境的组成	21
2.1.1 主机	22
2.1.2 连接	24
2.1.3 存储设备	26
2.2 磁盘驱动部件	27

2.2.1	盘片	28
2.2.2	主轴	28
2.2.3	读写头	28
2.2.4	驱动臂装置	29
2.2.5	控制器	29
2.2.6	物理磁盘的结构	30
2.2.7	分区位记录	31
2.2.8	逻辑块寻址	32
2.3	磁盘驱动器的性能	33
2.4	影响磁盘性能的基本准则	35
2.5	主机的逻辑部件	38
2.5.1	操作系统	39
2.5.2	设备驱动	39
2.5.3	卷管理器	39
2.5.4	文件系统	41
2.5.5	应用	44
2.6	应用程序的需求和磁盘的性能	45
	小结	48
第 3 章	数据保护: RAID	51
3.1	RAID 的实现	52
3.1.1	软件 RAID	52
3.1.2	硬件 RAID	52
3.2	RAID 阵列的组成	53
3.3	RAID 级别	54
3.3.1	分条 (Striping)	54
3.3.2	数据镜像	55
3.3.3	奇偶校验	56
3.3.4	RAID 0	57
3.3.5	RAID 1	57
3.3.6	嵌套 RAID	59
3.3.7	RAID 3	59
3.3.8	RAID 4	61
3.3.9	RAID 5	62
3.3.10	RAID 6	62