

Information Storage and Management

Second Edition

**Storing, Managing, and Protecting Digital Information
in Classic, Virtualized, and Cloud Environments**



Information Storage and Management



Information Storage and Management

**Storing, Managing, and Protecting
Digital Information in Classic,
Virtualized, and Cloud Environments**

2nd Edition

Edited by
Somasundaram Gnanasundaram
Alok Shrivastava



John Wiley & Sons, Inc.

Information Storage and Management: Storing, Managing, and Protecting Digital Information in Classic, Virtualized, and Cloud Environments 2nd Edition

Published by
John Wiley & Sons, Inc.
10475 Crosspoint Boulevard
Indianapolis, IN 46256
www.wiley.com

Copyright © 2012 by EMC Corporation
Published by John Wiley & Sons, Inc., Indianapolis, Indiana

Published simultaneously in Canada

ISBN: 978-1-118-09483-9
ISBN: 978-1-118-22347-5 (ebk)
ISBN: 978-1-118-23696-3 (ebk)
ISBN: 978-1-118-26187-3 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Limit of Liability/Disclaimer of Warranty: The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or Web site is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or website may provide or recommendations it may make. Further, readers should be aware that Internet websites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services please contact our Customer Care Department within the United States at (877) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at <http://booksupport.wiley.com>. For more information about Wiley products, visit www.wiley.com.

Library of Congress Control Number: 2012936405

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. Copyright © 1996, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012 EMC Corporation. All Rights Reserved. EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

EMC2, EMC, Data Domain, RSA, EMC Centera, EMC ControlCenter, EMC LifeLine, EMC OnCourse, EMC Proven, EMC Snap, EMC SourceOne, EMC Storage Administrator, Acartus, Access Logix, AdvantEdge, AlphaStor, ApplicationXtender, ArchiveXtender, Atmos, Authentica, Authentic Problems, Automated Resource Manager, AutoStart, AutoSwap, AVALONidm, Avamar, Captiva, Catalog Solution, C-Clip, Celerra, Celerra Replicator, Centera, CenterStage, CentraStar, ClaimPack, ClaimsEditor, CLARiiON, ClientPak, Codebook Correlation Technology, Common Information Model, Configuration Intelligence, Configuresoft, Connectrix, CopyCross, CopyPoint, Dantz, DatabaseXtender, Direct Matrix Architecture, DiskXtender, DiskXtender 2000, Document Sciences, Documentum, elnput, E-Lab, EmailXaminer, EmailXtender, Enginuity, eRoom, Event Explorer, FarPoint, FirstPass, FLARE, FormWare, Geosynchrony, Global File Virtualization, Graphic Visualization, Greenplum, HighRoad, HomeBase, InfoMover, Infoscapes, Infra, InputAccel, InputAccel Express, Invista, Ionix, ISIS, Max Retriever, MediaStor, MirrorView, Navisphere, NetWorker, nLayers, OnAlert, OpenScale, PixTools, Powerlink, PowerPath, PowerSnap, QuickScan, Rainfinity, RepliCare, RepliStor, ResourcePak, Retrospect, RSA, the RSA logo, SafeLine, SAN Advisor, SAN Copy, SAN Manager, Smarts, SnapImage, SnapSure, SnapView, SRDF, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix, Symmetrix DMX, Symmetrix VMAX, TimeFinder, UltraFlex, UltraPoint, UltraScale, Unisphere, VMAX, Vblock, Viewlets, Virtual Matrix, Virtual Matrix Architecture, Virtual Provisioning, VisualSAN, VisualSRM, Voyence, VPLEX, VSAM-Assist, WebXtender, xPression, xPresso, YottaYotta, the EMC logo, and where information lives, are registered trademarks or trademarks of EMC Corporation in the United States and other countries.

All other trademarks used herein are the property of their respective owners.

© Copyright 2012 EMC Corporation. All rights reserved. Published in the USA.

John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.



About the Editors

Somasundaram Gnanasundaram (Somu) is the director at EMC Education Services, leading worldwide industry readiness initiatives. Somu is the architect of EMC's open curriculum, aimed at addressing the knowledge gap that exists in the IT industry in the area of information storage and emerging technologies such as cloud computing. Under his leadership and direction, industry readiness initiatives such as the EMC Academic Alliance program continue to experience significant growth, educating thousands of students worldwide on information storage and management technologies. Key areas of Somu's responsibility include guiding a global team of professionals, identifying and partnering with global IT education providers, and setting the overall direction for EMC's industry readiness initiatives. Prior to his current role, Somu held various managerial and leadership roles within EMC as well as with other leading IT service providers. He holds an undergraduate technology degree from Anna University Chennai, and a Master of Technology degree from the Indian Institute of Technology, Mumbai, India. Somu has been in the IT industry for more than 25 years.

Alok Shrivastava is the senior director at EMC Education Services. Alok is the architect of several of EMC's successful education initiatives, including the industry leading EMC Proven Professional program, industry readiness programs such as EMC's Academic Alliance, and this unique and valuable book on information storage technology. Alok provides vision and leadership to a team of highly talented experts, practitioners, and professionals that develops world-class technical education for EMC's employees, partners, customers, students, and other industry professionals covering technologies such as storage, virtualization, cloud, and big data. Prior to his success in education, Alok built

and led a highly successful team of EMC presales engineers in Asia-Pacific and Japan. Earlier in his career, Alok was a systems manager, storage manager, and backup/restore/disaster recovery consultant working with some of the world's largest data centers and IT installations. He holds dual Master's degrees from the Indian Institute of Technology in Mumbai, India, and the University of Sagar in India. Alok has worked in information storage technology and has held a unique passion for this field for most of his 30-year career in IT.



Credits

Executive Editor

Carol Long

Project Editor

Tom Dinse

Senior Production Editor

Debra Banninger

Copy Editor

San Dee Phillips

Editorial Manager

Mary Beth Wakefield

Freelancer Editorial Manager

Rosemarie Graham

Associate Director of Marketing

David Mayhew

Marketing Manager

Ashley Zurcher

Business Manager

Amy Knies

Production Manager

Tim Tate

**Vice President and Executive
Group Publisher**

Richard Swadley

**Vice President and Executive
Publisher**

Neil Edde

Associate Publisher

Jim Minatel

Project Coordinator, Cover

Katie Crocker

Proofreader

Nancy Carrasco

Indexer

Robert Swanson

Cover Designer

Mallesha Gurram, EMC



Acknowledgments

When we embarked upon the project to develop this book in 2008, the first challenge was to identify a team of subject matter experts covering the vast range of technologies that form the modern information storage infrastructure.

A key factor that continues to work in our favor is that at EMC we have the technologies, the know-how, and many of the best talents in the industry. When we reached out to individual experts, they were as excited as we were about the prospect of publishing a comprehensive book on information storage technology. This was an opportunity to share their expertise with professionals and students worldwide.

This book is the result of efforts and contributions from a number of key EMC organizations led by EMC Education Services and supported by the office of CTO, Global Marketing, and EMC Engineering.

The first edition of the book was published in 2009, and the effort was led by **Ganesh Rajaratnam** of EMC Education Services and **Dr. David Black** of the EMC CTO office. The book continues to be the most popular storage technology book around the world among professionals and students. In addition to its English and e-book editions, it is available in Mandarin, Portuguese, and Russian.

With the emergence of cloud computing and the broad adoption of virtualization technologies by the organizations, we felt it is time to update the content to include information storage in those emerging technologies and also the new developments in the field of information storage. **Ashish Garg** of Education Services led the effort to update content for the second edition of this book. In addition to reviewing the content, **Joe Milardo** and **Nancy Gessler** led the effort of content review with their team of subject matter experts.

We are grateful to the following experts from EMC for their support in developing and reviewing the content for various chapters of this book:

Content contributors:

Rodrigo Alves
Charlie Brooks
Debasish Chakrabarty
Diana Davis
Amit Deshmukh
Michael Dulavitz
Dr. Vanchi Gurumoorthy
Simon Hawkshaw
Anbuselvi Jeyakumar

Sagar Kotekar Patil
Andre Rossouw
Tony Santamaria
Saravanaraj Sridharan
Ganesh Sundaresan
Jim Tracy
Anand Varkar
Dr. Viswanth VS

Content reviewers:

Ronen Artzi
Eric Baize
Greg Baltazar
Edward Bell
Ed Belliveau
Paul Brant
Juergen Busch
Christopher Chaulk
Brian Collins
Juan Cubillos
John Dowd
Roger Dupuis
Deborah Filer
Bala Ganeshan
Jason Gervickas
Jody Goncalves
Jack Harwood

Manoj Kumar
Arthur Johnson
Michelle Lavoie
Tom McGowan
Jeffery Moore
Toby Morral
Wayne Pauley
Peter Popieniuck
Ira Schild
Shashikanth, Punuru
Murugeson Purushothaman
Shekhar Sengupta
Kevin Sheridan
Ed VanSickle
Mike Warner
Ronnie Zubi
Evan Burleigh

We also thank **Mallik Motilal** of EMC for his support in creating all illustrations; **Mallesh Gurram** of EMC for the cover design; and the publisher, **John Wiley & Sons**, for its timely support in bringing this book to the industry.

— **Somasundaram Gnanasundaram**

Director, Education Services, EMC Corporation

— **Alok Shrivastava**

Senior Director, Education Services, EMC Corporation

March 2012



Contents

Foreword	xxvii
Introduction	xxix
Section I Storage System	1
Chapter 1 Introduction to Information Storage	3
1.1 Information Storage	4
1.1.1 Data	4
1.1.2 Types of Data	6
1.1.3 Big Data	7
1.1.4 Information	9
1.1.5 Storage	9
1.2 Evolution of Storage Architecture	9
1.3 Data Center Infrastructure	11
1.3.1 Core Elements of a Data Center	11
1.3.2 Key Characteristics of a Data Center	12
1.3.3 Managing a Data Center	13
1.4 Virtualization and Cloud Computing	14
Summary	15
Chapter 2 Data Center Environment	17
2.1 Application	18
2.2 Database Management System (DBMS)	18
2.3 Host (Compute)	19
2.3.1 Operating System	19
Memory Virtualization	20
2.3.2 Device Driver	20
2.3.3 Volume Manager	20
2.3.4 File System	22
2.3.5 Compute Virtualization	25

2.4 Connectivity	27
2.4.1 Physical Components of Connectivity	27
2.4.2 Interface Protocols	28
IDE/ATA and Serial ATA	28
SCSI and Serial SCSI	29
Fibre Channel	29
Internet Protocol (IP)	29
2.5 Storage	29
2.6 Disk Drive Components	31
2.6.1 Platter	32
2.6.2 Spindle	32
2.6.3 Read/Write Head	32
2.6.4 Actuator Arm Assembly	33
2.6.5 Drive Controller Board	33
2.6.6 Physical Disk Structure	34
2.6.7 Zoned Bit Recording	35
2.6.8 Logical Block Addressing	36
2.7 Disk Drive Performance	36
2.7.1 Disk Service Time	37
Seek Time	37
Rotational Latency	37
Data Transfer Rate	38
2.7.2 Disk I/O Controller Utilization	39
2.8 Host Access to Data	40
2.9 Direct-Attached Storage	41
2.9.1 DAS Benefits and Limitations	42
2.10 Storage Design Based on Application Requirements and Disk Performance	43
2.11 Disk Native Command Queuing	45
2.12 Introduction to Flash Drives	46
2.12.1 Components and Architecture of Flash Drives	47
2.12.2 Features of Enterprise Flash Drives	48
2.13 Concept in Practice: VMware ESXi	48
Summary	49
Chapter 3 Data Protection: RAID	51
3.1 RAID Implementation Methods	52
3.1.1 Software RAID	52
3.1.2 Hardware RAID	52
3.2 RAID Array Components	53
3.3 RAID Techniques	53
3.3.1 Striping	53
3.3.2 Mirroring	55
3.3.3 Parity	55
3.4 RAID Levels	57
3.4.1 RAID 0	57
3.4.2 RAID 1	58

3.4.3 Nested RAID	59
3.4.4 RAID 3	62
3.4.5 RAID 4	63
3.4.6 RAID 5	63
3.4.7 RAID 6	64
3.5 RAID Impact on Disk Performance	64
3.5.1 Application IOPS and RAID Configurations	66
3.6 RAID Comparison	66
3.7 Hot Spares	68
Summary	68
Chapter 4 Intelligent Storage Systems	71
4.1 Components of an Intelligent Storage System	72
4.1.1 Front End	72
4.1.2 Cache	72
Structure of Cache	73
Read Operation with Cache	73
Write Operation with Cache	75
Cache Implementation	75
Cache Management	76
Cache Data Protection	77
4.1.3 Back End	78
4.1.4 Physical Disk	78
4.2 Storage Provisioning	79
4.2.1 Traditional Storage Provisioning	79
LUN Expansion: MetaLUN	80
4.2.2 Virtual Storage Provisioning	82
Comparison between Virtual and Traditional Storage Provisioning	82
Use Cases for Thin and Traditional LUNs	84
4.2.3 LUN Masking	84
4.3 Types of Intelligent Storage Systems	85
4.3.1 High-End Storage Systems	85
4.3.2 Midrange Storage Systems	86
4.4 Concepts in Practice: EMC Symmetrix and VNX	87
4.4.1 EMC Symmetrix Storage Array	87
4.4.2 EMC Symmetrix VMAX Component	88
4.4.3 Symmetrix VMAX Architecture	89
Summary	91
Section II Storage Networking Technologies	93
Chapter 5 Fibre Channel Storage Area Networks	95
5.1 Fibre Channel: Overview	96
5.2 The SAN and Its Evolution	97
5.3 Components of FC SAN	98
5.3.1 Node Ports	99
5.3.2 Cables and Connectors	99

5.3.3 Interconnect Devices	100
5.3.4 SAN Management Software	101
5.4 FC Connectivity	102
5.4.1 Point-to-Point	102
5.4.2 Fibre Channel Arbitrated Loop	102
5.4.3 Fibre Channel Switched Fabric	103
FC-SW Transmission	105
5.5 Switched Fabric Ports	106
5.6 Fibre Channel Architecture	106
5.6.1 Fibre Channel Protocol Stack	107
FC-4 Layer	108
FC-2 Layer	108
FC-1 Layer	108
FC-0 Layer	108
5.6.2 Fibre Channel Addressing	109
5.6.3 World Wide Names	109
5.6.4 FC Frame	110
5.6.5. Structure and Organization of FC Data	112
5.6.6 Flow Control	112
BB_Credit	112
EE_Credit	112
5.6.7 Classes of Service	113
5.7 Fabric Services	113
5.8 Switched Fabric Login Types	114
5.9 Zoning	115
5.9.1 Types of Zoning	116
5.10 FC SAN Topologies	118
5.10.1 Mesh Topology	118
5.10.2 Core-Edge Fabric	119
Benefits and Limitations of Core-Edge Fabric	119
5.11 Virtualization in SAN	122
5.11.1 Block-level Storage Virtualization	122
5.11.2 Virtual SAN (VSAN)	124
5.12 Concepts in Practice: EMC Connectrix and EMC VPLEX	125
5.12.1 EMC Connectrix	125
Connectrix Switches	126
Connectrix Directors	126
Connectrix Multi-purpose Switches	126
Connectrix Management Tools	127
5.12.2 EMC VPLEX	127
VPLEX Family of Products	128
Summary	128
Chapter 6 IP SAN and FCoE	131
6.1 iSCSI	132
6.1.1 Components of iSCSI	132
6.1.2 iSCSI Host Connectivity	133

6.1.3 iSCSI Topologies	133
Native iSCSI Connectivity	133
Bridged iSCSI Connectivity	135
Combining FC and Native iSCSI Connectivity	135
6.1.4 iSCSI Protocol Stack	135
6.1.5 iSCSI PDU	136
6.1.6 iSCSI Discovery	138
6.1.7 iSCSI Names	138
6.1.8 iSCSI Session	140
6.1.9 iSCSI Command Sequencing	141
6.2 FCIP	142
6.2.1 FCIP Protocol Stack	142
6.2.2 FCIP Topology	144
6.2.3 FCIP Performance and Security	144
6.3 FCoE	145
6.3.1 I/O Consolidation Using FCoE	145
6.3.2 Components of an FCoE Network	147
Converged Network Adapter	148
Cables	148
FCoE Switches	149
6.3.3 FCoE Frame Structure	150
FCoE Frame Mapping	151
6.3.4 FCoE Enabling Technologies	152
Priority-Based Flow Control (PFC)	153
Enhanced Transmission Selection (ETS)	154
Congestion Notification (CN)	154
Data Center Bridging Exchange Protocol (DCBX)	154
Summary	155
Chapter 7 Network-Attached Storage	157
7.1 General-Purpose Servers versus NAS Devices	158
7.2 Benefits of NAS	159
7.3 File Systems and Network File Sharing	160
7.3.1 Accessing a File System	160
7.3.2 Network File Sharing	160
7.4 Components of NAS	162
7.5 NAS I/O Operation	163
7.6 NAS Implementations	163
7.6.1 Unified NAS	164
7.6.2 Unified NAS Connectivity	164
7.6.3 Gateway NAS	164
7.6.4 Gateway NAS Connectivity	165
7.6.5 Scale-Out NAS	166
7.6.6 Scale-Out NAS Connectivity	167
7.7 NAS File-Sharing Protocols	168
7.7.1 NFS	169
7.7.2 CIFS	170

7.8 Factors Affecting NAS Performance	171
7.9 File-Level Virtualization	174
7.10 Concepts in Practice: EMC Isilon and EMC VNX Gateway	175
7.10.1 EMC Isilon	175
7.10.2 EMC VNX Gateway	176
Summary	177
Chapter 8 Object-Based and Unified Storage	179
8.1 Object-Based Storage Devices	180
8.1.1 Object-Based Storage Architecture	181
8.1.2 Components of OSD	182
8.1.3 Object Storage and Retrieval in OSD	183
8.1.4 Benefits of Object-Based Storage	184
8.1.5 Common Use Cases for Object-Based Storage	185
8.2 Content-Addressed Storage	187
8.3 CAS Use Cases	188
8.3.1 Healthcare Solution: Storing Patient Studies	188
8.3.2 Finance Solution: Storing Financial Records	189
8.4 Unified Storage	190
8.4.1 Components of Unified Storage	190
Data Access from Unified Storage	192
8.5 Concepts in Practice: EMC Atmos, EMC VNX, and EMC Centera	192
8.5.1 EMC Atmos	193
8.5.2 EMC VNX	194
8.5.3 EMC Centera	195
EMC Centera Architecture	196
Summary	197
Section III Backup, Archive, and Replication	199
Chapter 9 Introduction to Business Continuity	201
9.1 Information Availability	202
9.1.1 Causes of Information Unavailability	202
9.1.2 Consequences of Downtime	203
9.1.3 Measuring Information Availability	204
9.2 BC Terminology	205
9.3 BC Planning Life Cycle	207
9.4 Failure Analysis	210
9.4.1 Single Point of Failure	210
9.4.2 Resolving Single Points of Failure	211
9.4.3 Multipathing Software	212
9.5 Business Impact Analysis	213
9.6 BC Technology Solutions	213
9.7 Concept in Practice: EMC PowerPath	214
9.7.1 PowerPath Features	214
9.7.2 Dynamic Load Balancing	215

I/O Operation without PowerPath	215
I/O Operation with PowerPath	216
9.7.3 Automatic Path Failover	217
Path Failure without PowerPath	218
Path Failover with PowerPath: Active-Active Array	218
Path Failover with PowerPath: Active-Passive Array	219
Summary	221
 Chapter 10 Backup and Archive	 225
10.1 Backup Purpose	226
10.1.1 Disaster Recovery	226
10.1.2 Operational Recovery	226
10.1.3 Archival	226
10.2 Backup Considerations	227
10.3 Backup Granularity	228
10.4 Recovery Considerations	231
10.5 Backup Methods	231
10.6 Backup Architecture	233
10.7 Backup and Restore Operations	234
10.8 Backup Topologies	236
10.9 Backup in NAS Environments	239
10.9.1 Server-Based and Serverless Backup	239
10.9.2 NDMP-Based Backup	240
10.10 Backup Targets	242
10.10.1 Backup to Tape	243
Physical Tape Library	243
Limitations of Tape	245
10.10.2 Backup to Disk	245
10.10.3 Backup to Virtual Tape	246
Virtual Tape Library	246
10.11 Data Deduplication for Backup	249
10.11.1 Data Deduplication Methods	249
10.11.2 Data Deduplication Implementation	250
Source-Based Data Deduplication	250
Target-Based Data Deduplication	250
10.12 Backup in Virtualized Environments	252
10.13 Data Archive	254
10.14 Archiving Solution Architecture	255
10.14.1 Use Case: E-mail Archiving	256
10.14.2 Use Case: File Archiving	257
10.15 Concepts in Practice: EMC NetWorker,	
EMC Avamar, and EMC Data Domain	257
10.15.1 EMC NetWorker	258
10.15.2 EMC Avamar	258
10.15.3 EMC Data Domain	259
Summary	260

Chapter 11	Local Replication	263
11.1	Replication Terminology	264
11.2	Uses of Local Replicas	264
11.3	Replica Consistency	265
11.3.1	Consistency of a Replicated File System	265
11.3.2	Consistency of a Replicated Database	266
11.4	Local Replication Technologies	269
11.4.1	Host-Based Local Replication	269
	LVM-Based Replication	269
	Advantages of LVM-Based Replication	269
	Limitations of LVM-Based Replication	270
	File System Snapshot	271
11.4.2	Storage Array-Based Local Replication	272
	Full-Volume Mirroring	273
	Pointer-Based, Full-Volume Replication	274
	Pointer-Based Virtual Replication	277
11.4.3	Network-Based Local Replication	278
	Continuous Data Protection	279
	CDP Local Replication Operation	280
11.5	Tracking Changes to Source and Replica	281
11.6	Restore and Restart Considerations	282
11.7	Creating Multiple Replicas	283
11.8	Local Replication in a Virtualized Environment	284
11.9	Concepts in Practice: EMC TimeFinder, EMC SnapView, and EMC RecoverPoint	285
11.9.1	EMC TimeFinder	285
	TimeFinder/Clone	286
	TimeFinder/Snap	286
11.9.2	EMC SnapView	286
	SnapView Snapshot	286
	SnapView Clone	287
11.9.3	EMC RecoverPoint	287
	Summary	287
Chapter 12	Remote Replication	289
12.1	Modes of Remote Replication	289
12.2	Remote Replication Technologies	292
12.2.1	Host-Based Remote Replication	292
	LVM-Based Remote Replication	293
	Host-Based Log Shipping	294
12.2.2	Storage Array-Based Remote Replication	295
	Synchronous Replication Mode	295
	Asynchronous Replication Mode	296
	Disk-Buffered Replication Mode	297
12.2.3	Network-Based Remote Replication	298
	CDP Remote Replication	298
12.3	Three-Site Replication	300

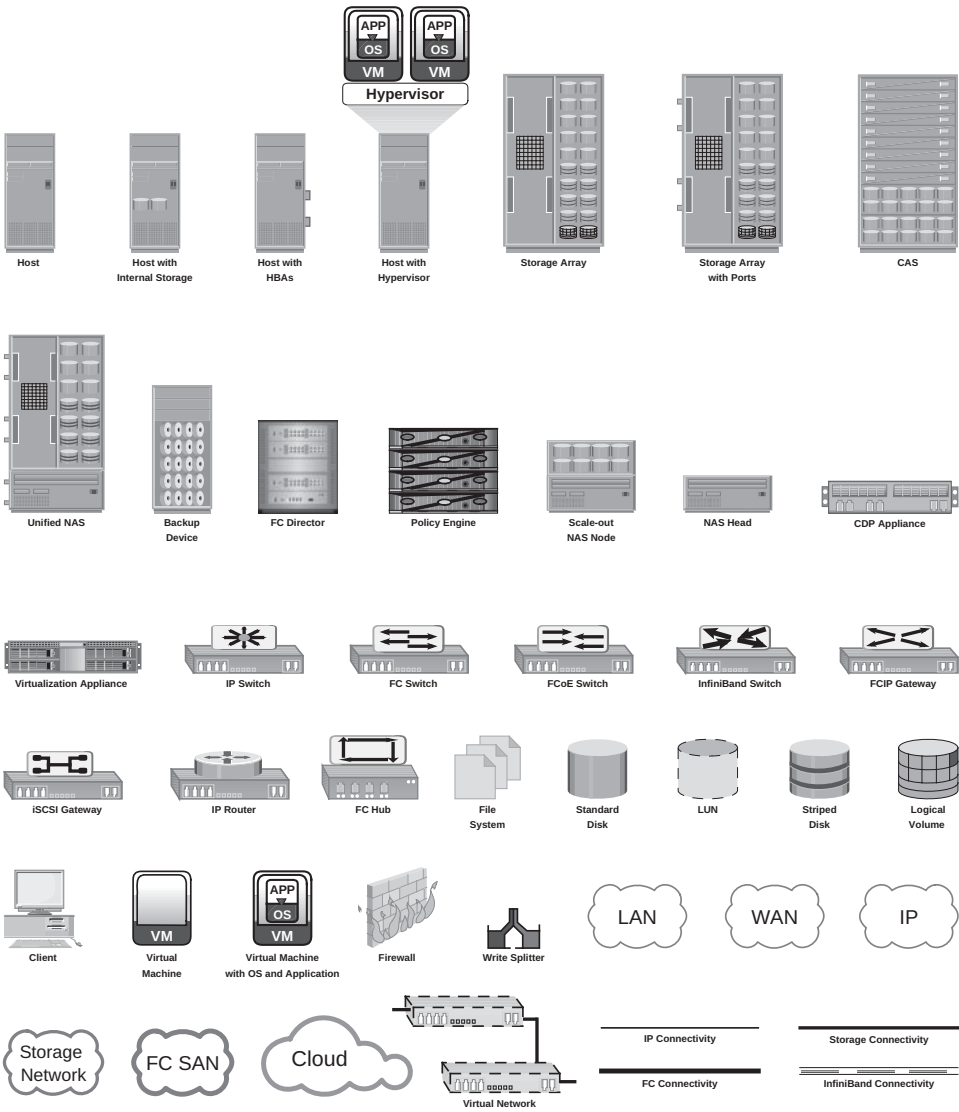
12.3.1 Three-Site Replication — Cascade/Multihop	300
Synchronous + Asynchronous	300
Synchronous + Disk Buffered	302
12.3.2 Three-Site Replication — Triangle/Multitarget	302
12.4 Data Migration Solutions	304
12.5 Remote Replication and Migration in a Virtualized Environment	306
12.6 Concepts in Practice: EMC SRDF, EMC MirrorView, and EMC RecoverPoint	307
12.6.1 EMC SRDF	308
12.6.2 EMC MirrorView	308
12.6.3 EMC RecoverPoint	308
Summary	309
Section IV Cloud Computing	311
Chapter 13 Cloud Computing	313
13.1 Cloud Enabling Technologies	314
13.2 Characteristics of Cloud Computing	314
13.3 Benefits of Cloud Computing	316
13.4 Cloud Service Models	316
13.4.1 Infrastructure-as-a-Service	316
13.4.2 Platform-as-a-Service	317
13.4.3 Software-as-a-Service	318
13.5 Cloud Deployment Models	318
13.5.1 Public Cloud	318
13.5.2 Private Cloud	319
13.5.3 Community Cloud	320
13.5.4 Hybrid Cloud	321
13.6 Cloud Computing Infrastructure	322
13.6.1 Physical Infrastructure	322
13.6.2 Virtual Infrastructure	323
13.6.3 Applications and Platform Software	324
13.6.4 Cloud Management and Service Creation Tools	324
13.7 Cloud Challenges	326
13.7.1 Challenges for Consumers	326
13.7.2 Challenges for Providers	327
13.8 Cloud Adoption Considerations	327
13.9 Concepts in Practice: Vblock	329
Summary	330
Section V Securing and Managing Storage Infrastructure	331
Chapter 14 Securing the Storage Infrastructure	333
14.1 Information Security Framework	334
14.2 Risk Triad	334
14.2.1 Assets	335
14.2.2 Threats	336

14.2.3 Vulnerability	337
14.3 Storage Security Domains	338
14.3.1 Securing the Application Access Domain	339
Controlling User Access to Data	340
Protecting the Storage Infrastructure	341
Data Encryption	342
14.3.2 Securing the Management Access Domain	342
Controlling Administrative Access	344
Protecting the Management Infrastructure	344
14.3.3 Securing Backup, Replication, and Archive	345
14.4 Security Implementations in Storage Networking	346
14.4.1 FC SAN	346
FC SAN Security Architecture	347
Basic SAN Security Mechanisms	347
LUN Masking and Zoning	349
Securing Switch Ports	349
Switch-Wide and Fabric-Wide Access Control	350
Logical Partitioning of a Fabric: Virtual SAN	350
14.4.2 NAS	350
NAS File Sharing: Windows ACLs	351
NAS File Sharing: UNIX Permissions	352
NAS File Sharing: Authentication and Authorization	353
Kerberos	354
Network-Layer Firewalls	355
14.4.3 IP SAN	357
14.5 Securing Storage Infrastructure in Virtualized and Cloud Environments	358
14.5.1 Security Concerns	359
14.5.2 Security Measures	359
Security at the Compute Level	359
Security at the Network Level	360
Security at the Storage Level	361
14.6 Concepts in Practice: RSA and VMware Security Products	361
14.6.1 RSA SecureID	362
14.6.2 RSA Identity and Access Management	362
14.6.3 RSA Data Protection Manager	362
14.6.4 VMware vShield	363
Summary	363
Chapter 15 Managing the Storage Infrastructure	365
15.1 Monitoring the Storage Infrastructure	366
15.1.1 Monitoring Parameters	366
15.1.2 Components Monitored	367
Hosts	367
Storage Network	368
Storage	369

15.1.3 Monitoring Examples	369
Accessibility Monitoring	369
Capacity Monitoring	370
Performance Monitoring	372
Security Monitoring	374
15.1.4 Alerts	375
15.2 Storage Infrastructure Management Activities	376
15.2.1 Availability Management	376
15.2.2 Capacity Management	376
15.2.3 Performance Management	377
15.2.4 Security Management	377
15.2.5 Reporting	378
15.2.6 Storage Infrastructure Management in a Virtualized Environment	378
15.2.7 Storage Management Examples	380
Example 1: Storage Allocation to a New Server/Host	380
Example 2: File System Space Management	381
Example 3: Chargeback Report	382
15.3 Storage Infrastructure Management Challenges	384
15.4 Developing an Ideal Solution	384
15.4.1 Storage Management Initiative	385
15.4.2 Enterprise Management Platform	386
15.5 Information Lifecycle Management	386
15.6 Storage Tiering	388
15.6.1 Intra-Array Storage Tiering	388
15.6.2 Inter-Array Storage Tiering	390
15.7 Concepts in Practice: EMC Infrastructure Management Tools	391
15.7.1 EMC ControlCenter and Prosphere	391
15.7.2 EMC Unisphere	392
15.7.3 EMC Unified Infrastructure Manager (UIM)	393
Summary	393
Appendix A Application I/O Characteristics	395
Random and Sequential	395
Reads and Writes	395
I/O Request Size	396
Appendix B Parallel SCSI	399
SCSI Standards Family	400
SCSI Client-Server Model	401
Parallel SCSI Addressing	402
Appendix C SAN Design Exercises	405
Exercise 1	405
Solution	405
Exercise 2	406
Solution	406

Appendix D Information Availability Exercises	409
Exercise 1	409
Solution	409
Exercise 2	410
Solution	410
Appendix E Network Technologies for Remote Replication	411
DWDM	411
CWDM	412
SONET	412
Appendix F Acronyms and Abbreviations	413
Glossary	427
Index	465

Icons Used In This Book





Foreword

In the two short years since we originally published this book, the world as we've known it has undergone a change of unprecedented magnitude. We are now in a digital era in which the world's information is more than doubling every two years, and in the next decade alone IT departments globally will need to manage 50 times more information while the number of IT professionals available will grow only 1.5 times (IDC Digital Universe Study, sponsored by EMC, June 2011). Virtualization and cloud computing are no longer an option for enterprises but an imperative for survival. And Big Data is creating significant new opportunity for organizations to analyze, act on, and drive new value from their most valuable asset — information — and create competitive advantage.

The Information Technology industry is undergoing a tremendous transformation as a result. The Cloud has introduced radically new technologies, computing models, and disciplines, dramatically changing the way IT is built, run, governed, and consumed. It has created new roles such as cloud technologists and cloud architects to lead this transformation. And it is transforming the IT organization from a back office overseer of infrastructure — with the task of keeping it running — into a key strategic contributor of the business with a focus on delivering IT as a service.

All of these changes demand new core competencies within the IT organization and a new way of thinking about technology in the context of business requirements and strategic objectives — even a new organizational structure within the data center. Information storage and management professionals must build on their existing knowledge and develop additional skills in the technologies most critical to successfully undertaking the complex, multi-year journey to the cloud: virtualization, converged networking, information security, data protection, and data warehousing and analytics, to name a few.

We have revised *Information Storage and Management* to give you an updated perspective and behind-the-scenes view of the new technologies and skills required today to design, implement, manage, optimize, and leverage virtualized infrastructures to achieve the business benefits of the cloud. You will learn from EMC subject matter experts with the most advanced training, certification, and practical experience in the industry.

If you are a storage and information management professional in the midst of virtualizing your data center, or building a robust cloud infrastructure — or if you are simply interested in learning the concept and principles of these new paradigms — transforming your IT skills has never been more critical. And by accelerating your transformation with this book and by taking advantage of the new training and certifications now available to you, you can help close a critical skills gap in the industry, advance your career, and become a valued contributor to your company's growth, sustainability, and profitability.

The challenges in this industry are many — but so are the rewards. Nelson Mandela said, "Education is the most powerful weapon which you can use to change the world." I hope you will make this book a key part of your IT education and professional development — regardless of your current role — and that you will seize this opportunity to help transform yourself and change the world.

Thomas P. Clancy

Vice President, Education Services, EMC Corporation

May 2012