

# Contents

<b>Contents</b>	<b>vii</b>
<b>Preface</b>	<b>xix</b>
For the Second Edition	xix
From Evan Marcus	xix
From Hal Stern	xxii
Preface from the First Edition	xxiv
From Evan Marcus	xxv
From Hal Stern	xxviii
<b>About the Authors</b>	<b>xxxi</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
Why an Availability Book?	2
Our Approach to the Problem	3
What's Not Here	4
Our Mission	4
The Availability Index	5
Summary	6
Organization of the Book	6
Key Points	8
<b>Chapter 2 What to Measure</b>	<b>9</b>
Measuring Availability	10
The Myth of the Nines	11
Defining Downtime	14
Causes of Downtime	15
What Is Availability?	17
M Is for Mean	18
What's Acceptable?	19

<b>Chapter 5</b>	<b>20 Key High Availability Design Principles</b>	<b>75</b>
	#20: Don't Be Cheap	76
	#19: Assume Nothing	77
	#18: Remove Single Points of Failure (SPOFs)	78
	#17: Enforce Security	79
	#16: Consolidate Your Servers	81
	#15: Watch Your Speed	82
	#14: Enforce Change Control	83
	#13: Document Everything	84
	#12: Employ Service Level Agreements	87
	#11: Plan Ahead	88
	#10: Test Everything	89
	#9: Separate Your Environments	90
	#8: Learn from History	92
	#7: Design for Growth	93
	#6: Choose Mature Software	94
	#5: Choose Mature, Reliable Hardware	95
	#4: Reuse Configurations	97
	#3: Exploit External Resources	98
	#2: One Problem, One Solution	99
	#1: K.I.S.S. (Keep It Simple . . .)	101
	Key Points	104
<b>Chapter 6</b>	<b>Backups and Restores</b>	<b>105</b>
	The Basic Rules for Backups	106
	Do Backups Really Offer High Availability?	108
	What Should Get Backed Up?	109
	Back Up the Backups	110
	Getting Backups Off-Site	110
	Backup Software	111
	Commercial or Homegrown?	111
	Examples of Commercial Backup Software	113
	Commercial Backup Software Features	113
	Backup Performance	115
	Improving Backup Performance:	
	Find the Bottleneck	118
	Solving for Performance	122
	Backup Styles	125
	Incremental Backups	126
	Incremental Backups of Databases	130
	Shrinking Backup Windows	130
	Hot Backups	131
	Have Less Data, Save More Time (and Space)	132
	Hierarchical Storage Management	132
	Archives	134
	Synthetic Fulls	134

Use More Hardware	135
Host-Free Backups	135
Third-Mirror Breakoff	136
Sophisticated Software Features	138
Copy-on-Write Snapshots	138
Multiplexed Backups	140
Fast and Flash Backup	141
Handling Backup Tapes and Data	141
General Backup Security	144
Restores	145
Disk Space Requirements for Restores	146
Summary	147
Key Points	148
<b>Chapter 7 Highly Available Data Management</b>	<b>149</b>
Four Fundamental Truths	150
Likelihood of Failure of Disks	150
Data on Disks	151
Protecting Data	151
Ensuring Data Accessibility	151
Six Independent Layers of Data Storage and Management	152
Disk Hardware and Connectivity Terminology	153
SCSI	153
Fibre Channel	156
Multipathing	157
Multihosting	157
Disk Array	157
Hot Swapping	158
Logical Units (LUNs) and Volumes	158
JBOD (Just a Bunch of Disks)	158
Hot Spares	158
Write Cache	159
Storage Area Network (SAN)	159
RAID Technology	161
RAID Levels	161
RAID-0: Striping	161
RAID-1: Mirroring	162
Combining RAID-0 and RAID-1	163
RAID-2: Hamming Encoding	167
RAID-3, -4, and -5: Parity RAID	167
Other RAID Variants	169
Hardware RAID	170
Disk Arrays	173
Software RAID	175
Logical Volume Management	176
Disk Space and Filesystems	176
Large Disks or Small Disks?	178
What Happens When a LUN Fills Up?	179



	Managing Disk and Volume Availability	180
	Filesystem Recovery	181
	Key Points	182
<b>Chapter 8</b>	<b>SAN, NAS, and Virtualization</b>	<b>183</b>
	Storage Area Networks (SANs)	184
	Why SANs?	186
	Storage Centralization and Consolidation	186
	Sharing Data	187
	Reduced Network Loads	188
	More Efficient Backups	188
	A Brief SAN Hardware Primer	189
	Network-Attached Storage (NAS)	190
	SAN or NAS: Which Is Better?	191
	Storage Virtualization	196
	Why Use Virtual Storage?	197
	Types of Storage Virtualization	198
	Filesystem Virtualization	198
	Block Virtualization	198
	Virtualization and Quality of Service	200
	Key Points	202
<b>Chapter 9</b>	<b>Networking</b>	<b>203</b>
	Network Failure Taxonomy	204
	Network Reliability Challenges	205
	Network Failure Modes	207
	Physical Device Failures	208
	IP Level Failures	209
	IP Address Configuration	209
	Routing Information	210
	Congestion-Induced Failures	211
	Network Traffic Congestion	211
	Design and Operations Guidelines	213
	Building Redundant Networks	214
	Virtual IP Addresses	215
	Redundant Network Connections	216
	Redundant Network Attach	217
	Multiple Network Attach	217
	Interface Trunking	219
	Configuring Multiple Networks	220
	IP Routing Redundancy	223
	Dynamic Route Recovery	224
	Static Route Recovery with VRRP	225
	Routing Recovery Guidelines	226
	Choosing Your Network Recovery Model	227
	Load Balancing and Network Redirection	228
	Round-Robin DNS	228
	Network Redirection	229
	Dynamic IP Addresses	232

Network Service Reliability	232
Network Service Dependencies	233
Hardening Core Services	236
Denial-of-Service Attacks	237
Key Points	240
<b>Chapter 10 Data Centers and the Local Environment</b>	<b>241</b>
Data Centers	242
Data Center Racks	244
Advantages and Disadvantages to Data Center Racks	244
The China Syndrome Test	247
Balancing Security and Access	247
Data Center Tours	248
Off-Site Hosting Facilities	250
Electricity	252
UPS	253
Backup Generators	254
Cabling	255
Cooling and Environmental Issues	257
System Naming Conventions	259
Key Points	261
<b>Chapter 11 People and Processes</b>	<b>263</b>
System Management and Modifications	264
Maintenance Plans and Processes	265
System Modifications	266
Things to Aim For	266
Software Patches	268
Spare Parts Policies	269
Preventative Maintenance	270
Vendor Management	271
Choosing Key Vendors	271
Working with Your Vendors	274
The Vendor's Role in System Recovery	275
Service and Support	275
Escalation	276
Vendor Integration	276
Vendor Consulting Services	277
Security	277
Data Center Security	279
Viruses and Worms	280
Documentation	280
The Audience for Documentation	282
Documentation and Security	283
Reviewing Documentation	284
System Administrators	284
Internal Escalation	287
Trouble Ticketing	289
Key Points	290



<b>Chapter 12</b>	<b>Clients and Consumers</b>	<b>291</b>
	Hardening Enterprise Clients	292
	Client Backup	292
	Client Provisioning	294
	Thin Clients	296
	Tolerating Data Service Failures	296
	Fileserver Client Recovery	297
	NFS Soft Mounts	297
	Automounter Tricks	298
	Database Application Recovery	299
	Web Client Recovery	301
	Key Points	302
<b>Chapter 13</b>	<b>Application Design</b>	<b>303</b>
	Application Recovery Overview	304
	Application Failure Modes	305
	Application Recovery Techniques	306
	Kinder, Gentler Failures	308
	Application Recovery from System Failures	309
	Virtual Memory Exhaustion	309
	I/O Errors	310
	Database Application Reconnection	311
	Network Connectivity	312
	Restarting Network Services	313
	Network Congestion, Retransmission, and Timeouts	314
	Internal Application Failures	316
	Memory Access Faults	317
	Memory Corruption and Recovery	318
	Hanging Processes	319
	Developer Hygiene	319
	Return Value Checks	320
	Boundary Condition Checks	322
	Value-Based Security	323
	Logging Support	324
	Process Replication	326
	Redundant Service Processes	326
	Process State Multicast	327
	Checkpointing	329
	Assume Nothing, Manage Everything	330
	Key Points	331
<b>Chapter 14</b>	<b>Data and Web Services</b>	<b>333</b>
	Network File System Services	334
	Detecting RPC Failures	334
	NFS Server Constraints	336
	Inside an NFS Failover	337
	Optimizing NFS Recovery	337
	File Locking	339
	Stale File Handles	341

Database Servers	342
Managing Recovery Time	343
Database Probes	343
Database Restarts	344
Surviving Corruption	346
Unsafe at Any (High) Speed	347
Transaction Size and Checkpointing	347
Parallel Databases	348
Redundancy and Availability	349
Multiple Instances versus Bigger Instances	350
Web-Based Services Reliability	351
Web Server Farms	352
Application Servers	353
Directory Servers	356
Web Services Standards	357
Key Points	359
<b>Chapter 15 Local Clustering and Failover</b>	<b>361</b>
A Brief and Incomplete History of Clustering	362
Server Failures and Failover	365
Logical, Application-centric Thinking	367
Failover Requirements	369
Servers	372
Differences among Servers	372
Failing Over between Incompatible Servers	373
Networks	374
Heartbeat Networks	374
Public Networks	377
Administrative Networks	381
Disks	381
Private Disks	381
Shared Disks	382
Placing Critical Applications on Disks	384
Applications	385
Larger Clusters	385
Key Points	386
<b>Chapter 16 Failover Management and Issues</b>	<b>387</b>
Failover Management Software (FMS)	388
Component Monitoring	389
Who Performs a Test, and Other Component Monitoring Issues	391
When Component Tests Fail	392
Time to Manual Failover	393
Homemade Failover Software or Commercial Software?	395
Commercial Failover Management Software	397
When Good Failovers Go Bad	398
Split-Brain Syndrome	398
Causes and Remedies of Split-Brain Syndrome	400
Undesirable Failovers	404



Verification and Testing	404
State Transition Diagrams	405
Testing the Works	407
Managing Failovers	408
System Monitoring	408
Consoles	409
Utilities	410
Time Matters	410
Other Clustering Topics	411
Replicated Data Clusters	411
Distance between Clusters	413
Load-Balancing Clusters and Failover	413
Key Points	414
<b>Chapter 17 Failover Configurations</b>	<b>415</b>
Two-Node Failover Configurations	416
Active-Passive Failover	416
Active-Passive Issues and Considerations	417
How Can I Use the Standby Server?	418
Active-Active Failover	421
Active-Active or Active-Passive?	424
Service Group Failover	425
Larger Cluster Configurations	426
N-to-1 Clusters	426
N-Plus-1 Clusters	428
How Large Should Clusters Be?	430
Key Points	431
<b>Chapter 18 Data Replication</b>	<b>433</b>
What Is Replication?	434
Why Replicate?	435
Two Categories of Replication Types	435
Four Latency-Based Types of Replication	435
Latency-Based Type 1: Synchronous Replication	436
Latency-Based Type 2: Asynchronous Replication	438
Latency-Based Type 3: Semi-Synchronous Replication	439
Latency-Based Type 4: Periodic, or Batch-Style, Replication	439
Five Initiator-Based Types of Replication	441
Initiator-Based Type 1: Hardware-Based Replication	441
Initiator-Based Type 2: Software-Based Replication	443
Initiator-Based Type 3: Filesystem-Based Replication	444
Initiator-Based Type 4: Application-Based Replication	450
Initiator-Based Type 5: Transaction Processing Monitors	454
Other Thoughts on Replication	458
SANs: Another Way to Replicate	458
More than One Destination	459
Remote Application Failover	462
Key Points	463



<b>Chapter 19</b>	<b>Virtual Machines and Resource Management</b>	<b>465</b>
	Partitions and Domains: System-Level VMs	466
	Containers and Jails: OS Level VMs	468
	Resource Management	469
	Key Points	471
<b>Chapter 20</b>	<b>The Disaster Recovery Plan</b>	<b>473</b>
	Should You Worry about DR?	474
	Three Primary Goals of a DR Plan	475
	Health and Protection of the Employees	475
	The Survival of the Enterprise	476
	The Continuity of the Enterprise	476
	What Goes into a Good DR Plan	476
	Preparing to Build the DR Plan	477
	Choosing a DR Site	484
	Physical Location	484
	Considerations in Selecting DR Sites	485
	Other Options	486
	DR Site Security	487
	How Long Will You Be There?	488
	Distributing the DR Plan	488
	What Goes into a DR Plan	488
	So What Should You Do?	490
	The Plan's Audience	490
	Timelines	492
	Team Assignments	493
	Assigning People	493
	Management's Role	494
	How Many Different Plans?	495
	Shared DR Sites	496
	Equipping the DR Site	498
	Is Your Plan Any Good?	500
	Qualities of a Good Exercise	500
	Planning for an Exercise	501
	Possible Exercise Limitations	503
	Make It More Realistic	503
	Ideas for an Exercise Scenario	504
	After the Exercise	507
	Three Types of Exercises	507
	Complete Drill	507
	Tabletop Drill	508
	Phone Chain Drill	508
	The Effects of a Disaster on People	509
	Typical Responses to Disasters	509
	What Can the Enterprise Do to Help?	510
	Key Points	512

<b>Chapter 21    A Resilient Enterprise*</b>	<b>513</b>
The New York Board of Trade	514
The First Time	516
No Way for a Major Exchange to Operate	517
Y2K Preparation	520
September 11, 2001	523
Getting Back to Work	525
Chaotic Trading Environment	528
Improvements to the DR Site	531
New Data Center	532
The New Trading Facility	533
Future Disaster Recovery Plans	534
The Technology	535
The Outcry for Open Outcry	535
Modernizing the Open Outcry Process	536
The Effects on the People	538
Summary	539
<b>Chapter 22    A Brief Look Ahead</b>	<b>541</b>
iSCSI	541
InfiniBand	542
Global Filesystem Undo	543
Grid Computing	545
Blade Computing	547
Global Storage Repository	548
Autonomic and Policy-Based Computing	549
Intermediation	551
Software Quality and Byzantine Reliability	552
Business Continuity	553
Key Points	554
<b>Chapter 23    Parting Shots</b>	<b>555</b>
How We Got Here	555
<b>Index</b>	<b>559</b>