
Contents

1	Introduction	1
1.1	Audience	2
1.2	Roadmap of This Book	4
1.3	Real-World Examples	8
2	Elementary Concepts	13
2.1	Business Issues	14
2.1.1	Business Continuity as the Overall Goal	16
2.1.2	Regulatory Compliance and Risk Management	16
2.2	System and Outage Categorization	17
2.3	High Availability – Handling Minor Outages	22
2.3.1	Availability	24
2.3.2	Reliability	25
2.3.3	Serviceability	25
2.4	Disaster Recovery – Handling Major Outages	26
2.5	Quantifying Availability: 99.9...% and Reality	29
2.6	Service Level Agreements	31
2.7	Basic Approach: Robustness and Redundancy	34
2.8	Layered Solution with Multiple Precautions	38
2.9	Summary	39
3	Architecture	41
3.1	Objectives	45
3.2	Conceptual Model	48
3.3	System Model	51
4	System Design	55
4.1	Base Concepts	55
4.1.1	System Stack	56
4.1.2	Redundancy and Replication	61
4.1.3	Robustness and Simplicity	74

4.1.4	Virtualization	77
4.2	Solution Roadmap	78
4.2.1	List Failure Scenarios	79
4.2.2	Evaluate Failure Scenarios	82
4.2.3	Map Scenarios to Requirements	82
4.2.4	Design Solution	85
4.2.5	Review Selected Solution Against Scenarios	86
4.3	System Solution Patterns	86
4.3.1	System Implementation Process	87
4.3.2	Systems for All Process Steps	87
4.3.3	Use Case: SAP Server	89
5	Hardware	99
5.1	Components and Computer Systems	104
5.2	Disk Storage	108
5.2.1	Raid – Redundant Array of Independent Disks	109
5.2.2	Storage Systems	119
5.2.3	SAN vs. NAS	124
5.2.4	Journaling Is Essential for High Availability	125
5.3	Virtualization of Resources	126
5.4	Vendor Selection and Purchasing Decisions	128
5.5	System Installation	132
5.6	System Maintenance and Operations	139
5.7	Making Our Own Statistics	142
6	Operating Systems	149
6.1	Failover Clusters	151
6.1.1	How Does It Work?	157
6.1.2	Failover Cluster Implementation Experiences	166
6.2	Load-Balancing Clusters	176
6.2.1	Load-Balancing Approaches	178
6.2.2	Target Selection for Load Balancing	181
6.3	Cluster and Server Consolidation	183
6.3.1	Virtualization and Moore’s Law	183
6.3.2	Host Virtualization	184
7	Databases and Middleware	189
7.1	Middleware Categories	191
7.2	Database Servers	193
7.2.1	High-Availability Options for Database Servers	199
7.2.2	Disaster Recovery for Databases	204
7.3	Web Servers	205
7.4	Application Servers	208
7.5	Messaging Servers	213

8	Applications	215
8.1	Integration in a Cluster on the Operating System Level	217
8.2	High Availability Through Middleware	223
8.3	High Availability From Scratch	225
8.4	Code Quality Is Important	227
8.5	Testing for High Availability	229
9	Infrastructure	233
9.1	Network	234
9.1.1	Network Devices	238
9.1.2	LAN Segments	240
9.1.3	Default Gateway	248
9.1.4	Routing in LANs and WANs	252
9.1.5	Firewalls and Network Address Translation	258
9.1.6	Network Design for Disaster Recovery	264
9.2	Infrastructure Services	267
9.2.1	Dynamic Host Configuration Protocol (DHCP)	267
9.2.2	Domain Name Service (DNS)	271
9.2.3	Directory Server	276
9.3	Backup and Restore	283
9.4	Monitoring	284
10	Disaster Recovery	287
10.1	Concepts	289
10.2	Approach	291
10.3	Conceptual Design	292
10.3.1	Scenarios for Major Outages	293
10.3.2	Disaster-Recovery Scope	295
10.3.3	Primary and Disaster-Recovery Sites	297
10.3.4	State Synchronization	298
10.3.5	Shared System, Hot or Cold Standby	300
10.3.6	Time to Recovery – Failback to the Primary Site	303
10.4	Solutions	305
10.4.1	Metro Cluster	306
10.4.2	Fast Restore	309
10.4.3	Application-Level or Middleware-Level Clustering	309
10.4.4	Application Data Mirroring	310
10.4.5	Disk Mirroring	317
10.4.6	Matching Configuration Changes	317
10.5	Disaster-Recovery Tests	318
10.5.1	Test Goals and Categories	319
10.5.2	Organizational Test Context	321
10.5.3	Quality Characteristics	322
10.6	Holistic View – What Is Needed Besides Technology?	322
10.6.1	Command Center and War Room	323

10.6.2	Disaster-Recovery Emergency Pack	323
10.7	A Prototypical Disaster-Recovery Project	324
10.7.1	System Identification – the Primary Site	326
10.7.2	Business Requirements and Project Goals	331
10.7.3	Business View	333
10.7.4	System Design	336
10.7.5	Implementation	345
10.8	Failover to Disaster-Recovery Site or Disaster-Recovery Systems	351
10.8.1	General Approach	351
10.8.2	Example Checklist for a Database Disaster-Recovery Server	355
10.8.3	Failback to the Primary System	357
A	Reliability Calculations and Statistics	359
A.1	Mathematical Basics	360
A.2	Mean Time Between Failures and Annual Failure Rate ...	362
A.3	Redundancy and Probability of Failures	363
A.4	Raid Configurations	365
A.5	Example Calculations	372
A.6	Reliability over Time – the Bathtub Curve	374
B	Data Centers	377
B.1	Room Installation	378
B.2	Heat and Fire Control	381
B.3	Power Control	384
B.4	Computer Setup	386
C	Service Support Processes	387
C.1	Incident Management	388
C.2	Problem Management	389
C.3	Configuration Management	391
C.4	Change Management	394
C.5	Release Management	395
C.6	Information Gathering and Reporting.....	397
	References	399
	Index	401