BEYOND REDUNDANCY

How Geographic Redundancy Can Improve Service Availability and Reliability of Computer-based Systems

ERIC BAUER · RANDEE ADAMS · DANIEL EUSTACE

WILEY

IEEE
IEEE PRESS

BEYOND REDUNDANCY

IEEE Press
445 Hoes Lane
Piscataway, NJ 08854

IEEE Press Editorial Board
Lajos Hanzo, *Editor in Chief*

R. Abhari
J. Anderson
G. W. Arnold
F. Canavero
M. El-Hawary
B-M. Haemmerli
M. Lanzerotti
D. Jacobson
O. P. Malik
S. Nahavandi
T. Samad
G. Zobrist

Kenneth Moore, *Director of IEEE Book and Information Services (BIS)*

Technical Reviewers
Xuemei Zhang
Network Design and Performance Analysis Division, AT&T Labs

Kim W. Tracy
Northeastern Illinois University

BEYOND REDUNDANCY

How Geographic Redundancy Can Improve Service Availability and Reliability of Computer-Based Systems

Eric Bauer
Randee Adams
Daniel Eustace

IEEE PRESS

WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

Copyright © 2012 by Institute of Electrical and Electronics Engineers. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

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 Section 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, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. 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: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data
Bauer, Eric.
Beyond redundancy : how geographic redundancy can improve service availability and reliability of computer-based systems / Eric Bauer, Randee Adams, Daniel Eustace.
p. cm.
ISBN 978-1-118-03829-1 (hardback)
1. Computer input-output equipment–Reliability. 2. Computer networks–Reliability.
3. Redundancy (Engineering) I. Adams, Randee. II. Eustace, Daniel. III. Title.
TK7887.5.B395 2011
004.6–dc22
2011008324

oBook ISBN: 978-1-118-10491-0
ePDF ISBN: 978-1-118-10492-7
ePub ISBN: 978-1-118-10493-4

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To our families for their encouragement and support:
Eric's wife Sandy and children Lauren and Mark
Randee's husband Scott and son Ryan
Dan's wife Helen and daughters Christie and Chelsea

CONTENTS

Figures **xv**

Tables **xix**

Equations **xxi**

Preface and Acknowledgments **xxiii**

Audience xxiv

Organization xxiv

Acknowledgments xxvi

PART 1 BASICS **1**

1 SERVICE, RISK, AND BUSINESS CONTINUITY **3**

1.1 Service Criticality and Availability Expectations 3

1.2 The Eight-Ingredient Model 4

1.3 Catastrophic Failures and Geographic Redundancy 7

1.4 Geographically Separated Recovery Site 11

1.5 Managing Risk 12

1.5.1 Risk Identification 13

1.5.2 Risk Treatments 13

1.6 Business Continuity Planning 14

1.7 Disaster Recovery Planning 15

1.8 Human Factors 17

1.9 Recovery Objectives 17

1.10 Disaster Recovery Strategies 18

2 SERVICE AVAILABILITY AND SERVICE RELIABILITY **20**

2.1 Availability and Reliability 20

2.1.1 Service Availability 20

2.1.2 Service Reliability 21

2.1.3 Reliability, Availability, and Failures 22

2.2 Measuring Service Availability 25

2.2.1 Total and Partial Outages 26

2.2.2 Minimum Chargeable Disruption Duration 27

2.2.3 Outage Attributability 28

2.2.4 Systems and Network Elements 30

2.2.5 Service Impact and Element Impact Outages 30

2.2.6 Treatment of Planned Events 32

2.3 Measuring Service Reliability 33

PART 2 MODELING AND ANALYSIS OF REDUNDANCY 35

3 UNDERSTANDING REDUNDANCY 37

3.1 Types of Redundancy 37

3.1.1 Simplex Configuration 39

3.1.2 Redundancy 41

3.1.3 Single Point of Failure 43

3.2 Modeling Availability of Internal Redundancy 44

3.2.1 Modeling Active–Active Redundancy 45

3.2.2 Modeling Active Standby Redundancy 49

3.2.3 Service Availability Comparison 51

3.3 Evaluating High-Availability Mechanisms 52

3.3.1 Recovery Time Objective (or Nominal Outage Duration) 54

3.3.2 Recovery Point Objective 54

3.3.3 Nominal Success Probability 55

3.3.4 Capital Expense 55

3.3.5 Operating Expense 56

3.3.6 Discussion 56

4 OVERVIEW OF EXTERNAL REDUNDANCY 59

4.1 Generic External Redundancy Model 59

4.1.1 Failure Detection 64

4.1.2 Triggering Recovery Action 66

4.1.3 Traffic Redirection 67

4.1.4 Service Context Preservation 71

4.1.5 Graceful Service Migration 74

4.2 Technical Distinctions between Georedundancy and Co-Located Redundancy 74

4.3 Manual Graceful Switchover and Switchback 75

5 EXTERNAL REDUNDANCY STRATEGY OPTIONS 77

5.1 Redundancy Strategies 77
5.2 Data Recovery Strategies 79
5.3 External Recovery Strategies 80
5.4 Manually Controlled Recovery 81
5.4.1 Manually Controlled Example: Provisioning System for a Database 83
5.4.2 Manually Controlled Example: Performance Management Systems 83
5.5 System-Driven Recovery 83
5.5.1 System-Driven Recovery Examples 85
5.6 Client-Initiated Recovery 85
5.6.1 Client-Initiated Recovery Overview 86
5.6.2 Failure Detection by Client 88
5.6.3 Client-Initiated Recovery Example: Automatic Teller Machine (ATM) 95
5.6.4 Client-Initiated Recovery Example: A Web Browser Querying a Web Server 96
5.6.5 Client-Initiated Recovery Example: A Pool of DNS Servers 97

6 MODELING SERVICE AVAILABILITY WITH EXTERNAL SYSTEM REDUNDANCY 98

6.1 The Simplistic Answer 98
6.2 Framing Service Availability of Standalone Systems 99
6.3 Generic Markov Availability Model of Georedundant Recovery 103
6.3.1 Simplifying Assumptions 103
6.3.2 Standalone High-Availability Model 104
6.3.3 Manually Controlled Georedundant Recovery 107
6.3.4 System-Driven Georedundant Recovery 110
6.3.5 Client-Initiated Georedundancy Recovery 111
6.3.6 Complex Georedundant Recovery 113
6.3.7 Comparing the Generic Georedundancy Model to the Simplistic Model 114
6.4 Solving the Generic Georedundancy Model 115
6.4.1 Manually Controlled Georedundant Recovery Model 118
6.4.2 System-Driven Georedundant Recovery Model 120
6.4.3 Client-Initiated Georedundant Recovery Model 120
6.4.4 Conclusion 121

6.5 Practical Modeling of Georedundancy 121
6.5.1 Practical Modeling of Manually Controlled External System Recovery 122
6.5.2 Practical Modeling of System-Driven Georedundant Recovery 124
6.5.3 Practical Modeling of Client-Initiated Recovery 125
6.6 Estimating Availability Benefit for Planned Activities 130
6.7 Estimating Availability Benefit for Disasters 131

7 UNDERSTANDING RECOVERY TIMING PARAMETERS 133

7.1 Detecting Implicit Failures 134
7.1.1 Understanding and Optimizing T_{TIMEOUT} 134
7.1.2 Understanding and Optimizing $T_{\mathrm{KEEPALIVE}}$ 137
7.1.3 Understanding and Optimizing T_{CLIENT} 139
7.1.4 Timer Impact on Service Reliability 140
7.2 Understanding and Optimizing RTO 141
7.2.1 RTO for Manually Controlled Recovery 141
7.2.2 RTO for System-Driven Recovery 143
7.2.3 RTO for Client-Initiated Recovery 145
7.2.4 Comparing External Redundancy Strategies 146

8 CASE STUDY OF CLIENT-INITIATED RECOVERY 147

8.1 Overview of DNS 147
8.2 Mapping DNS onto Practical Client-Initiated Recovery Model 148
8.2.1 Modeling Normal Operation 150
8.2.2 Modeling Server Failure 151
8.2.3 Modeling Timeout Failure 151
8.2.4 Modeling Abnormal Server Failure 153
8.2.5 Modeling Multiple Server Failure 154
8.3 Estimating Input Parameters 154
8.3.1 Server Failure Rate 155
8.3.2 F_{EXPLICIT} Parameter 158
8.3.3 C_{CLIENT} Parameter 158
8.3.4 μ_{TIMEOUT} Parameter 160
8.3.5 $\mu_{\mathrm{CLIENTSFD}}$ Parameter 160
8.3.6 μ_{CLIENT} Parameter 161
8.3.7 $A_{\mathrm{CLUSTER-1}}$ Parameter 162
8.3.8 F_{CLIENT} Parameter 162
8.3.9 μ_{GRECOVER} and $\mu_{\mathrm{MIGRATION}}$ Parameters 164

8.3.10 μ_{DUPLEX} Parameter 165
8.3.11 Parameter Summary 165
8.4 Predicted Results 165
8.4.1 Sensitivity Analysis 168
8.5 Discussion of Predicted Results 172

9 SOLUTION AND CLUSTER RECOVERY 174
9.1 Understanding Solutions 174
9.1.1 Solution Users 175
9.1.2 Solution Architecture 176
9.2 Estimating Solution Availability 177
9.3 Cluster versus Element Recovery 179
9.4 Element Failure and Cluster Recovery Case Study 182
9.5 Comparing Element and Cluster Recovery 186
9.5.1 Failure Detection 186
9.5.2 Triggering Recovery Action 186
9.5.3 Traffic Redirection 186
9.5.4 Service Context Preservation 187
9.5.5 Graceful Migration 187
9.6 Modeling Cluster Recovery 187
9.6.1 Cluster Recovery Modeling Parameters 190
9.6.2 Estimating $\lambda_{\text{SUPERELEMENT}}$ 193
9.6.3 Example of Super Element Recovery Modeling 196

PART 3 RECOMMENDATIONS 201

10 GEOREDUNDANCY STRATEGY 203
10.1 Why Support Multiple Sites? 203
10.2 Recovery Realms 204
10.2.1 Choosing Site Locations 205
10.3 Recovery Strategies 206
10.4 Limp-Along Architectures 207
10.5 Site Redundancy Options 208
10.5.1 Standby Sites 208
10.5.2 $N + K$ Load Sharing 211
10.5.3 Discussion 215
10.6 Virtualization, Cloud Computing, and Standby Sites 216
10.7 Recommended Design Methodology 217

11 MAXIMIZING SERVICE AVAILABILITY VIA GEOREDUNDANCY 219
11.1 Theoretically Optimal External Redundancy 219
11.2 Practically Optimal Recovery Strategies 220
11.2.1 Internal versus External Redundancy 220
11.2.2 Client-Initiated Recovery as Optimal External Recovery Strategy 222
11.2.3 Multi-Site Strategy 223
11.2.4 Active–Active Server Operation 223
11.2.5 Optimizing Timeout and Retry Parameters 224
11.2.6 Rapid Relogon 225
11.2.7 Rapid Context Restoration 225
11.2.8 Automatic Switchback 226
11.2.9 Overload Control 226
11.2.10 Network Element versus Cluster-Level Recovery 226
11.3 Other Considerations 228
11.3.1 Architecting to Facilitate Planned Maintenance Activities 228
11.3.2 Procedural Considerations 229

12 GEOREDUNDANCY REQUIREMENTS 230
12.1 Internal Redundancy Requirements 230
12.1.1 Standalone Network Element Redundancy Requirements 231
12.1.2 Basic Solution Redundancy Requirements 232
12.2 External Redundancy Requirements 233
12.3 Manually Controlled Redundancy Requirements 235
12.3.1 Manual Failover Requirements 235
12.3.2 Graceful Switchover Requirements 236
12.3.3 Switchback Requirements 236
12.4 Automatic External Recovery Requirements 237
12.4.1 System-Driven Recovery 237
12.4.2 Client-Initiated Recovery 239
12.5 Operational Requirements 242

13 GEOREDUNDANCY TESTING 243
13.1 Georedundancy Testing Strategy 243
13.1.1 Network Element Level Testing 244
13.1.2 End-to-End Testing 245
13.1.3 Deployment Testing 245
13.1.4 Operational Testing 246

13.2 Test Cases for External Redundancy 246
13.3 Verifying Georedundancy Requirements 247
13.3.1 Test Cases for Standalone Elements 248
13.3.2 Test Cases for Manually Controlled Recovery 248
13.3.3 Test Cases for System-Driven Recovery 249
13.3.4 Test Cases for Client-Initiated Recovery 250
13.3.5 Test Cases at the Solution Level 252
13.3.6 Test cases for Operational Testing 253
13.4 Summary 254

14 SOLUTION GEOREDUNDANCY CASE STUDY 256

14.1 The Hypothetical Solution 256
14.1.1 Key Quality Indicators 258
14.2 Standalone Solution Analysis 259
14.2.1 Construct Reliability Block Diagrams 260
14.2.2 Network Element Configuration in Standalone Solution 260
14.2.3 Service Availability Offered by Standalone Solution 261
14.2.4 Discussion of Standalone Solution Analysis 261
14.3 Georedundant Solution Analysis 263
14.3.1 Identify Factors Constraining Recovery Realm Design 263
14.3.2 Define Recovery Realms 264
14.3.3 Define Recovery Strategies 264
14.3.4 Set Recovery Objectives 265
14.3.5 Architecting Site Redundancy 268
14.4 Availability of the Georedundant Solution 269
14.5 Requirements of Hypothetical Solution 269
14.5.1 Internal Redundancy Requirements 269
14.5.2 External Redundancy Requirements 272
14.5.3 Manual Failover Requirements 273
14.5.4 Automatic External Recovery Requirements 274
14.5.5 Operational Requirements 277
14.6 Testing of Hypothetical Solution 277
14.6.1 Testing Strategy 277
14.6.2 Standalone Network Element Testing 278
14.6.3 Automatic External Recovery Requirements Testing 279
14.6.4 End-to-End Testing 282
14.6.5 Deployment Testing 283
14.6.6 Operational Testing 284

Summary 285
Appendix: Markov Modeling of Service Availability 292
Acronyms 296
References 298
About the Authors 300
Index 302

FIGURES

Figure 1.1. The Eight-Ingredient Model 5
Figure 1.2. Canonical Disaster Recovery Scenario 16
Figure 1.3. Recovery Time Objective and Recovery Point Objective 18
Figure 2.1. Canonical Service Impact Timeline 23
Figure 2.2. Reliability, Availability, and Quality of Service 24
Figure 2.3. Outage Downtime in Standalone NE Deployment 31
Figure 2.4. Outage Downtime with Redundant NE Deployment 32
Figure 3.1. Types of System Redundancy 39
Figure 3.2. State Transition Diagram of Simplex System 40
Figure 3.3. Availability Improvement Strategies for a Simplex System 41
Figure 3.4. Reliability Block Diagram of Redundant Pair 42
Figure 3.5. Service Availability of Active–Standby Redundant Pair 43
Figure 3.6. Sample Reliability Block Diagram 44
Figure 3.7. Sample Standalone Redundant System 45
Figure 3.8. Active–Active Markov Availability Model 45
Figure 3.9. Active–Active Markov Availability Model with Formulas 47
Figure 3.10. Active–Standby Markov Availability Model 50
Figure 3.11. Simplex Model with Mathematical Formulas 51
Figure 3.12. Outage Duration 54
Figure 3.13. Outage Duration for Disasters 55
Figure 4.1. Generic High Availability Model 60
Figure 4.2. Stable Service Delivery Path Across Generic Model 61
Figure 4.3. Degenerate Generic Model Without Element "C" 61
Figure 4.4. Failure Scenario in Generic Model 62
Figure 4.5. Recovery Scenario in Generic Model 62
Figure 4.6. Generic High Availability Model with Load Sharing 63
Figure 4.7. Georedundancy Using DNS SRV Records 70
Figure 5.1. Client-Initiated Recovery Scenario 87
Figure 5.2. Typical Client Processing Logic for Standalone Server 91
Figure 5.3. Generic Client-Initiated Recovery Logic with Redundant Servers 92
Figure 5.4. Session States Seen by a SIP Client "A" in Client-Initiated Recovery 93

Figure 5.5. Session States Seen by “A” in Client-Initiated Recovery Without Registration 95
Figure 5.6. Browser Query to Web Server 96
Figure 6.1. Generic High Availability Model 100
Figure 6.2. Sample Unavailability Contribution for Active–Standby Redundancy 100
Figure 6.3. Simplified Standalone High-Availability Downtime Model 104
Figure 6.4. General Georedundant Manual Recovery Markov Transition Diagram 107
Figure 6.5. System-Driven Georedundant Recovery Markov Transition Diagram 110
Figure 6.6. Client-Initiated Georedundant Recovery Markov Transition Diagram 111
Figure 6.7. Client-Initiated and System-Driven Georedundant Recovery 113
Figure 6.8. Overlaying Generic Georedundancy Model Onto Simplistic Model 115
Figure 6.9. Outage Durations for Sample System with internal Redundancy 123
Figure 6.10. Simplified Client-Initiated Recovery Markov Model 126
Figure 6.11. Modified Client-Initiated Recovery Model 130
Figure 6.12. Estimating Service Unavailability During Manual Disaster Recovery 132
Figure 7.1. Standard Protocol Timeout 135
Figure 7.2. Adaptive Protocol Timeout 136
Figure 7.3. Timeout with Multiple Parallel Requests 136
Figure 7.4. Client/Server Keepalive 138
Figure 7.5. Manually Controlled Recovery Timeline 142
Figure 7.6. System-Driven Recovery Timeline 143
Figure 7.7. Client-Initiated Recovery Timeline 145
Figure 8.1. Generic Model of DNS 149
Figure 8.2. Practical Client-Initiated Model for DNS 150
Figure 8.3. Modeling Normal DNS Operation 151
Figure 8.4. Modeling Server Failure 152
Figure 8.5. Modeling Timeout Failure 152
Figure 8.6. Modeling Abnormal Server Failure 154
Figure 8.7. Modeling Multiple Server Failures 155
Figure 8.8. Simplified Client-Initiated Recovery Model with Formulas 155
Figure 8.9. Critical Failure Rate Parameter 156
Figure 8.10. Failure Rate as a Function of MTTR 158
Figure 8.11. F_{EXPLICIT} Parameter 159
Figure 8.12. C_{CLIENT} Parameter 159
Figure 8.13. μ_{TIMEOUT} Parameter 160
Figure 8.14. $\mu_{\text{CLIENTSFD}}$ Parameter 161
Figure 8.15. μ_{CLIENT} Parameter 162
Figure 8.16. $A_{\text{CLUSTER}-1}$ Parameter 163
Figure 8.17. F_{CLIENT} Parameter 163

Figure 8.18. μ_{GRECOVER} and $\mu_{\text{MIGRATION}}$ Parameters 164
Figure 8.19. μ_{DUPLEX} Parameter 165
Figure 8.20. Sensitivity of Critical Failures per Year 169
Figure 8.21. Sensitivity of C_{CLIENT} 170
Figure 8.22. Sensitivity of $\mu_{\text{GRECOVERY}}$ 170
Figure 8.23. Sensitivity of μ_{DUPLEX} 171
Figure 8.24. Sensitivity of μ_{CLIENT} 172
Figure 9.1. External interfaces to Hypothetical Solution 175
Figure 9.2. Network Diagram of Hypothetical Solution 177
Figure 9.3. External interfaces to Hypothetical Solution 177
Figure 9.4. Reliability Block Diagram for End User Service of Hypothetical Solution 178
Figure 9.5. Reliability Block Diagram for User Provisioning Service of Hypothetical Solution 178
Figure 9.6. Downtime Expectation for End User Service of Sample Solution 178
Figure 9.7. Generic Redundant Cluster Model 179
Figure 9.8. Element Recovery upon Element Failure 180
Figure 9.9. Cluster Recovery upon Element Failure 180
Figure 9.10. Reliability Block Diagram for Georedundant Redundant Hypothetical Solution 182
Figure 9.11. Client Recovery for a Frontend Server Failure 183
Figure 9.12. Element Recovery of Security Server in Sample Solution 183
Figure 9.13. Cluster Recovery of Database Server in Sample Solution 184
Figure 9.14. Modeling Super Element Recovery 188
Figure 9.15. Client-Initiated Super Element Cluster Recovery Model 188
Figure 9.16. Active–Active States That Can Be Mitigated by Client-Initiated Recovery 194
Figure 9.17. Active–Active State Transitions That Can Activate Client-Initiated Recovery 195
Figure 10.1. Georedundancy with a Standby Site 209
Figure 10.2. Spare Site Georedundancy After a Failure 210
Figure 10.3. Georedundancy with $N + K$ Load Sharing 212
Figure 10.4. $N + K$ Load Sharing Georedundancy After a Failure 213
Figure 10.5. $N + K$ Load Sharing with 1 + 1 Redundant Elements 214
Figure 10.6. $N + K$ Load Sharing with 1 + 1 Redundancy After a Failure 215
Figure 14.1. Hypothetical Solution Architecture 257
Figure 14.2. Reliability Block Diagram for End User Service of Hypothetical Solution 260
Figure 14.3. Reliability Block Diagram for User Provisioning Service of Sample Solution 260
Figure 14.4. Estimating Service Availability Experienced by End Users 263
Figure S1. Generic High Availability Model 286
Figure S2. Visualization of Sample Generic Modeling Results 288
Figure A1. Andrey Markov (1856–1922). 293
Figure A2. Simple Availability Transition Diagram 293

TABLES

Table 1.1. FAA's Availability Expectations by Service Thread Criticality 5
Table 2.1. Service Availability and Downtime Ratings 21
Table 3.1. Sample Input Parameters for Active–Active Model 48
Table 3.2. Probability of Time Spent in Each Active–Active State 48
Table 3.3. Sample Additional Input Parameters for Active–Standby Model 50
Table 3.4. Probability of Time Spent in Each Active–Standby State 51
Table 3.5. Probability of Time Spent in Each Simplex State 52
Table 3.6. Nominal Downtime Predictions for Different Redundancy Schemes 53
Table 6.1. Mitigation of Service Downtime by Georedundancy Recovery Category 102
Table 6.2. Nominal Modeling Parameters and Values 117
Table 6.3. Nominal Modeling Results Summary 119
Table 7.1. Nominal RTO Times for Each External Redundancy Strategy 146
Table 8.1. DNS RCODE Return Code Values 148
Table 8.2. Deducing Server Status from DNS RCODE 149
Table 8.3. Modeling Input Parameters 156
Table 8.4. DNS RCODEs by C_{CLIENT} 160
Table 8.5. DNS Client-Initiated Recovery Modeling Parameters 166
Table 8.6. Predicted Service Availability for DNS 167
Table 9.1. Comparison of Client-Initiated Recovery Parameter Values 191
Table 9.2. Input Parameters for Standalone Active–Active Model 197
Table 9.3. Time Spent in Each Standalone Active–Active State 197
Table 9.4. Sample Cluster Client-Initiated Recovery Modeling Parameters 198
Table 9.5. Client-Initiated Recovery Prediction 199
Table 9.6. Solution Downtime Prediction 200
Table 14.1. Service Criticality by External Solution Interface 259
Table 14.2. Availability Expectations for Hypothetical Solution Elements 262
Table 14.3. Redundancy Schemes for the Solution Elements 270

EQUATIONS

Equation 2.1. Service Availability Equation 20
Equation 2.2. Defects per Million Equation 22
Equation 2.3. Converting Service Reliability to DPM 22
Equation 3.1. Availability of Simplex System Formula 40
Equation 3.2. Availability of Active–Standby System Formula 50
Equation 6.1. Simplistic Redundancy Availability Formula 99
Equation 6.2. Maximum Feasible Availability Benefit of System-Driven Recovery 125
Equation 7.1. General Equation for RTO 141
Equation 8.1. Availability as a Function of MTBF and MTTR 157
Equation 8.2. MTBF as a Function of Availability and MTTR 157
Equation 8.3. Failure Rate as a Function of Availability and MTTR 157
Equation 9.1. Estimating $\lambda_{\text{SUPERELEMENT}}$ 195

PREFACE AND ACKNOWLEDGMENTS

The best practice for mitigating the risk of site destruction, denial, or unavailability causing disastrous loss of critical services is to deploy redundant systems in a geographically separated location; this practice is called geographic redundancy or georedundancy. Enterprises deploying a geographically redundant system may spend significantly more than when deploying a standalone configuration up front, and will have higher ongoing operating expenses to maintain the geographically separated redundant recovery site and system. While the business continuity benefits of georedundancy are easy to understand, the feasible and likely service availability benefits of georedundancy are not generally well understood. This book considers the high-level question of what service availability improvement is feasible and likely with georedundancy. The emphasis is on system availability of IP-based applications. WAN availability is briefly mentioned where applicable, but is not factored into any of the modeling. The service availability benefit is characterized both for product attributable failures, as well as for nonproduct attributable failures, such as site disasters. Human factors are also taken into consideration as they relate to procedural downtime. Furthermore, this book considers architectural and operational topics, such as: whether it is better to only do a georedundancy failover for a failed element or for the entire cluster of elements that contains the failed element; whether georedundancy can/should be used to reduce planned downtime for activities such as hardware growth and software upgrade; what availability-related georedundancy requirements should apply to each network element and to clusters of elements; and what network element- and cluster-level testing is appropriate to assure expected service availability benefits of georedundancy.

This book considers the range of IP-based information and communication technology (ICT) systems that are typically deployed in enterprise data centers and telecom central offices. The term “enterprise” is used to refer to the service provider or enterprise operating the system, “supplier” is used to refer to the organization that develops and tests the system, and “user” is used for the human or system that uses the system. In some cases, “enterprise,” “supplier,” and “user” may all be part of the same larger organization (e.g., system that is developed, tested and operated by the IT department of a larger, and used by employees of the organization), but often two or all three of these parties are in different organizations.

The term network element refers to a system device, entity, or node including all relevant hardware and/or software components deployed at one location providing a particular primary function; an instance of a domain name system (DNS) server is a network element. A system is "*a collection of components organized to accomplish a specific function or set of functions*" (*IEEE Standard Glossary*, 1991); a pool of DNS servers is an example of system. A solution is an integrated suite of network elements that can provide multiple primary functions; a customer care center that may include functionality, such as call handling facilities, web servers, and billing servers, is an example of a solution. With improvements in technology and hardware capacity, the distinction between these terms often blurs, since a single server could perform all of the functionality required of the solution and might be considered a network element. The more general term "external redundancy" is used to encompass both traditional geographic redundancy in which redundant system instances are physically separated to minimize the risk of a single catastrophic event impacting both instances, as well as the situation in which redundant system instances are physically co-located. While physically co-located systems do not mitigate the risk of catastrophic site failure, they can mitigate the risk of system failures. External redundancy is contrasted with internal redundancy in which the redundancy is confined to a single element instance. For example, a RAID array is a common example of internal redundancy because the software running on the element or the RAID hardware assures that disk failures are detected and mitigated without disrupting user service. If each element requires a dedicated RAID array and an enterprise chooses to deploy a pair of elements for redundancy, then those elements could either be co-located in a single facility or installed in separate, presumably geographically distant, facilities. Both co-located and geographically separated configurations are considered "externally redundant," as the redundancy encompasses multiple element instances. Elements can be deployed with no redundancy, internal redundancy, external redundancy, or hybrid arrangements. This book discusses internal redundancy but focuses on external redundancy arrangements.

AUDIENCE

This book is written for network architects and designers, maintenance and operations engineers, and decision makers in IT organizations at enterprises who are considering or have deployed georedundant systems. This book is also written for system architects, system engineers, developers, testers, and others (including technical sales and support staff) involved in the development of systems supporting external redundancy and solutions considering system redundancy. This book is also written for reliability engineers and others who model service availability of systems that include external redundancy, including georedundancy.

ORGANIZATION

The book is organized to enable different audiences to easily access the information they are most interested in. Part 1, "Basics," gives background on georedundancy and service availability, and is suitable for all readers. Part 2, "Modeling and Analysis of Redundancy," gives technical and mathematical details of service availability modeling of georedundant configurations, and thus is most suitable for reliability engineers and others with deeper mathematical interest in the topic. Part 3 'Recommendations' offers specific recommendations on architecture, design, specification, testing, and analysis of georedundant configurations. The recommendations section ends with Chapter 15 which offers a summary of the material. Most readers will focus on Parts 1 and 3; reliability engineers will focus on Parts 2 and 3; and readers looking for a high-level summary can focus on Chapter 15, "Summary."

Part 1—Basics, contains the following chapters:

- *"Service, Risk, and Business Continuity"* reviews risk management, business continuity and disaster recovery in the context of service availability of critical systems.
- *"Service Availability and Service Reliability"* reviews the concepts of service availability and service reliability, including how these key metrics are measured in the field.

Part 2—Modeling and Analysis of Redundancy contains the following chapters:

- *"Understanding Redundancy"* factors redundancy into three broad categories: simplex (no redundancy), internal system redundancy, and external system redundancy (including co-located and geographically separated configurations). The fundamentals of high-availability mechanisms and modeling of availability improvements from internal redundancy are covered. Criteria for evaluating high-availability mechanisms are also given.
- *"Overview of External Redundancy"* reviews the key techniques and mechanisms that support failure detection and recovery that enable internal and external redundancy. This chapter also reviews the technical differences between local (co-located) and geographically separated redundancy.
- *"External Redundancy Strategy Options"* reviews the three fundamental system-level external redundancy strategies that are used today: manually controlled, system-driven, and client-initiated recovery. Case studies are given to illustrate how these techniques can be integrated to achieve highly available and reliable systems.
- *"Modeling Service Availability with External System Redundancy"* presents mathematical modeling of the service availability benefit of the three external redundancy strategies. First, a generic model that roughly covers all external redundancy strategies is presented to highlight the differences between the recovery strategies; then more practical strategy specific models are presented and analyzed.
- *"Understanding Recovery Timing Parameters"* details how key recovery-related timing parameters used in the mathematical modeling of the previous chapter should be set to optimize the recovery time for the various external redundancy strategies.
- *"Case Study of Client-Initiated Recovery"* uses a domain name system (DNS) cluster as an example of client-initiated recovery to illustrate the concepts and models discussed earlier in this section.

- *"Solution and Cluster Recovery"* considers how clusters of network elements organized into solutions delivering sophisticated services to enterprises and their customers can be recovered together, and discusses the potential benefits of cluster recovery compared to recovery of individual elements.

Part 3—Recommendations contains the following chapters

- *"Georedundancy Strategy"* reviews considerations when engineering the number of sites to deploy a solution across to assure acceptable quality service is highly available to users.
- *"Maximizing Service Availability via Georedundancy"* reviews the architectural characteristics that can maximize the service availability benefit of external system redundancy.
- *"Georedundancy Requirements"* lists sample redundancy requirements for enterprise IT organizations to consider when specifying critical services.
- *"Georedundancy Testing"* discusses how the verifiable requirements of the "Georedundancy Requirements" chapter should be tested across the integration, system validation, deployment/installation, and operational lifecycle phases.
- *"Solution Georedundancy Case Study"* discusses analysis, architecture, design, specification, and testing of a hypothetical solution.
- *"Summary"* reviews the feasible improvements in service availability that can be practically achieved by properly configuring solutions and redundant systems.

Since many readers will not be familiar with the principles of Markov modeling of service availability used in this book, a basic overview of Markov modeling of service availability is included as an appendix.

ACKNOWLEDGMENTS

The authors acknowledge Chuck Salisbury for his diligent early work to understand the service availability benefits of georedundancy. The authors are also grateful for Ron Santos' expert input on DNS. Bill Baker provided extensive comments and shared his valuable insights on this subject. Doug Kimber provided detail and thoughtful review, and the technical reviewers provided excellent feedback that led us to improve the content and flow of the book. Anil Macwan provided guidance on procedural reliability considerations. Ted Lach and Chun Chan provided expert input on several subtle reliability items. Michael Liem provided valuable feedback.

Eric Bauer
Randee Adams
Daniel Eustace

PART 1

BASICS