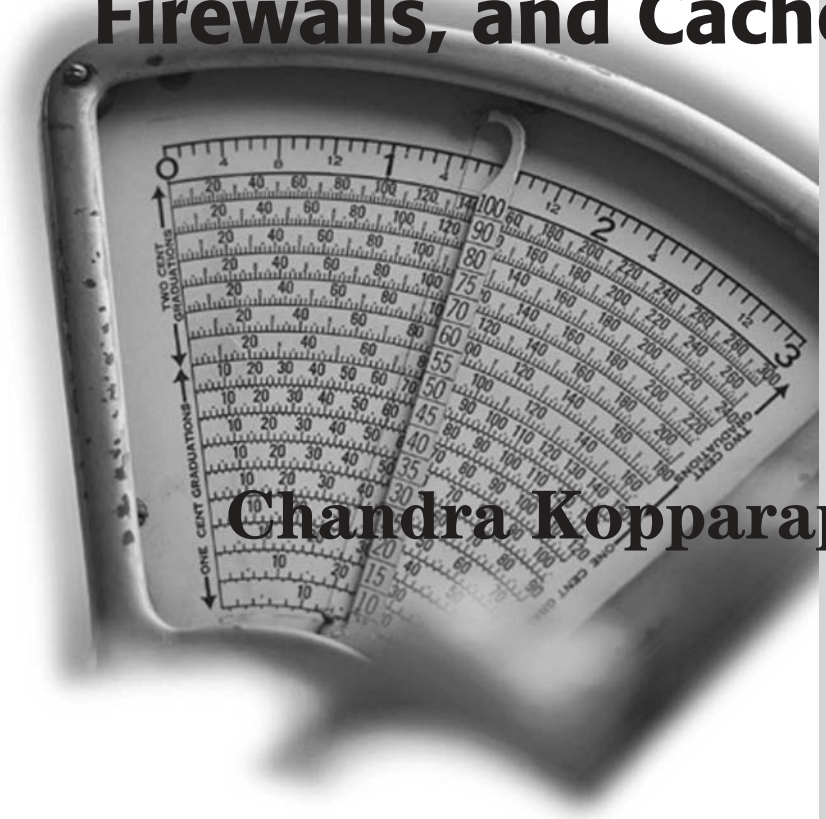
Load Balancing Servers, Firewalls, and Caches



Chandra Kopparapu

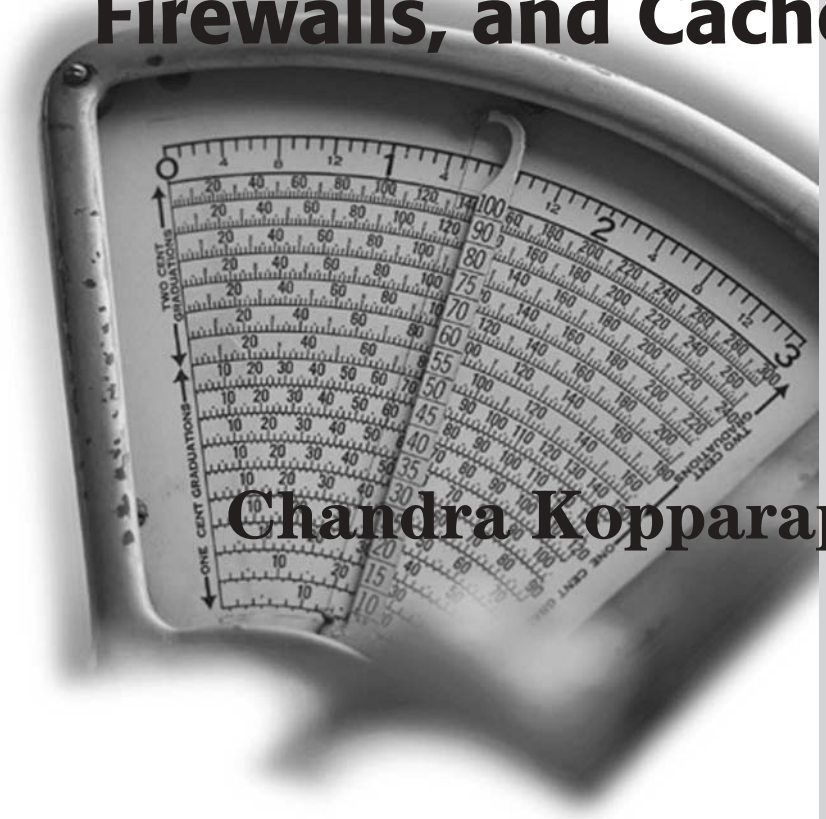
Wiley Computer Publishing

John Wiley & Sons, Inc.

NEW YORK • CHICHESTER • WEINHEIM • BRISBANE • SINGAPORE • TORONTO

Load Balancing Servers, Firewalls, and Caches

Load Balancing Servers, Firewalls, and Caches



Chandra Kopparapu

Wiley Computer Publishing

John Wiley & Sons, Inc.

NEW YORK • CHICHESTER • WEINHEIM • BRISBANE • SINGAPORE • TORONTO

Publisher: Robert Ipsen

Editor: Carol A. Long

Developmental Editor: Adaobi Obi

Managing Editor: Micheline Frederick

Text Design & Composition: Interactive Composition Corporation

This book is printed on acid-free paper. ♾

Published by John Wiley & Sons, Inc.

Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Kopparapu, Chandra.
Load balancing servers, firewalls, and caches / Chandra Kopparapu.
p. cm
Includes bibliographical references and index.
ISBN 0-471-41550-2 (cloth : alk. paper)
1. Client/server computing. 2. Firewalls (Computer security) I. Title.

QA76.9.C55 K67 2001
004.6–dc21

2001046757

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To my beloved daughters,
Divya and Nitya,
who bring so much joy to my life.

ACKNOWLEDGMENTS

First and foremost, my gratitude goes to my family. Without the support and understanding of my wife and encouragement from my parents, this book would not have been completed.

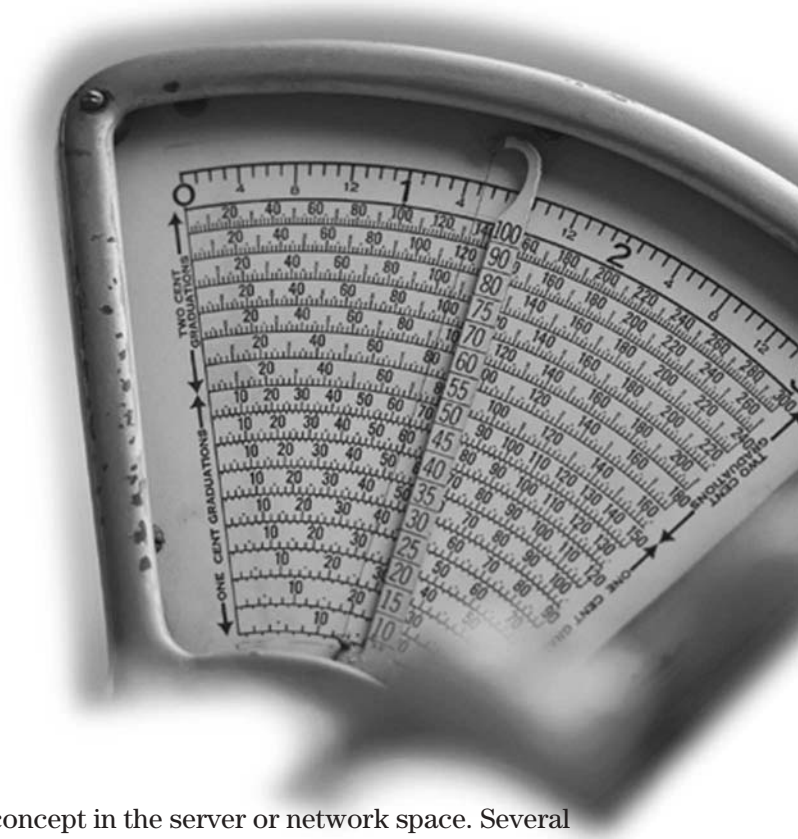
Rajkumar Jalan, principal architect for load balancers at Foundry Networks, was of invaluable help to me in understanding many load-balancing concepts when I was new to this technology. Many thanks go to Matthew Naugle, systems engineer at Foundry Networks, for encouraging me to write this book, giving me valuable feedback, and reviewing some of the chapters. Matt patiently spent countless hours with me, discussing several high-availability designs, and contributed valuable insight based on several customers he worked with. Terry Rolon, who used to work as a systems engineer at Foundry Networks, was also particularly helpful to me in coming up to speed on load-balancing products and network designs.

I would like to thank Mark Hoover of Acuitive Consulting for his thorough review and valuable analysis on Chapters 1, 2, 3, and 9. Mark has been very closely involved with the evolution of load-balancing products as an industry consultant and guided some load-balancing vendors in their early days. Many thanks to Brian Jacoby from America Online, who reviewed many of the chapters in this book from a customer perspective and provided valuable feedback.

Countless thanks to my colleagues at Foundry Networks, who worked with me over the last few years in advancing load-balancing product functionality and designing customer networks. I worked with many developers, systems engineers, customers, and technical support engineers to gain valuable insight into how load balancers are deployed and used by customers. Special thanks to Srini Ramadurai, David Cheung, Joe Tomasello, Ivy Hsu, Ron Szeto, and Ritesh Rekhi for helping me understand various aspects of load balancing functionality. I would also like to thank Ken Cheng, VP of Marketing at Foundry, for being supportive of this effort, and Bobby Johnson, Foundry's CEO, for giving me the opportunity to work with Foundry's load-balancing product line.

CHAPTER 1

Introduction



Load balancing is not a new concept in the server or network space. Several products perform different types of load balancing. For example, routers can distribute traffic across multiple paths to the same destination, balancing the load across different network resources. A server load balancer, on the other hand, distributes traffic among server resources rather than network resources. While load balancers started with simple load balancing, they soon evolved to perform a variety of functions: load balancing, traffic engineering, and intelligent traffic switching. Load balancers can perform sophisticated health checks on servers, applications, and content to improve availability and manageability. Because load balancers are deployed as the front end of a server farm, they also protect the servers from malicious users, and enhance security. Based on information in the IP packets or content in application requests, load balancers make intelligent decisions to direct the traffic appropriately—to the right data center, server, firewall, cache, or application.