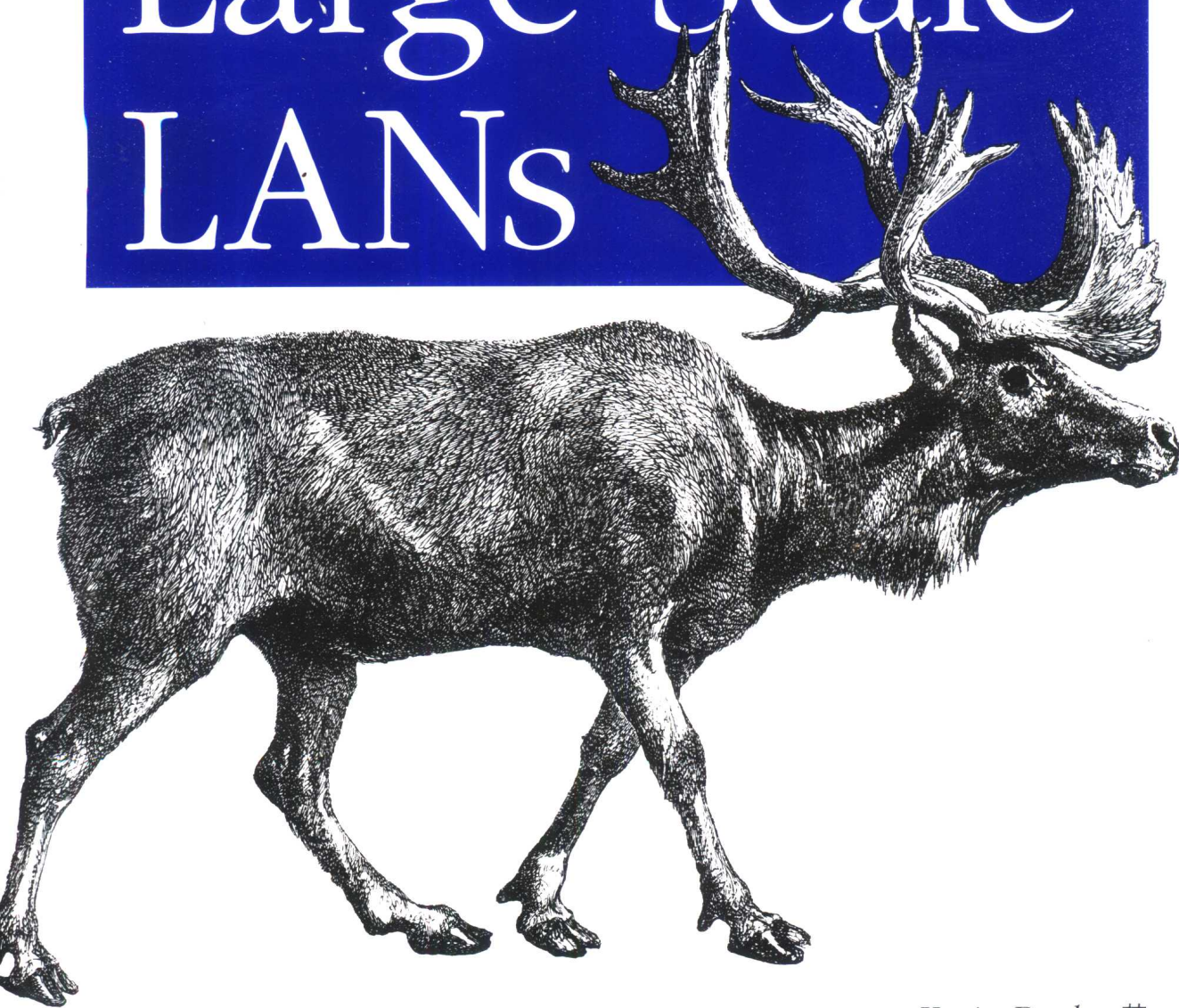


大规模局域网设计(影印版)

Designing

Large-Scale LANs



O'REILLY®



Kevin Dooley 著

清华大学出版社

大规模局域网设计(影印版)

Designing Large-Scale LANs



O'REILLY®

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

O'Reilly & Associates, Inc. 授权清华大学出版社出版

清华大学出版社

图书在版编目 (CIP) 数据

大规模局域网设计: / (美) 杜莱 (Dooley, K.) 著 — 影印版. — 北京: 清华大学出版社, 2002.11

书名原文: Designing Large-Scale LANs

ISBN 7-302-05972-1

I . 大 ... II . 杜 ... III . 局部网络 - 基本知识 - 英文 IV . TP393.1

中国版本图书馆 CIP 数据核字 (2002) 第 078739 号

北京市版权局著作权合同登记

图字: 01-2002-4633 号

©2002 by O'Reilly & Associates, Inc.

Reprint of the English Edition, jointly published by O'Reilly & Associates, Inc. and TsingHua University Press, 2002. Authorized reprint of the original English edition, 2002 O'Reilly & Associates, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly & Associates, Inc. 出版 2002。

英文影印版由清华大学出版社出版 2002。此影印版的出版和销售得到出版权和销售权的所有者 —— O'Reilly & Associates, Inc. 的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式重制。

本书封面贴有清华大学出版社激光防伪标签, 无标签者不得销售。

书 名 / 大规模局域网设计 (影印版)

书 号 / ISBN 7-302-05972-1/TP · 3557

责任编辑 / 冯志强

封面设计 / Ellie Volckhausen, 张健

出版发行 / 清华大学出版社 (www.tup.tsinghua.edu.cn)

地 址 / 北京清华大学学研大厦 (邮政编码 100084)

经 销 / 各地新华书店

印 刷 / 北京艺辉印刷有限公司

开 本 / 787 毫米 × 980 毫米 16 开本 25.25 印张

版 次 / 2002 年 11 月第一版 2002 年 11 月第一次印刷

印 数 / 0001-1500 册

定 价 / 105.00 元 (册)

O'Reilly & Associates 公司介绍

O'Reilly & Associates 公司是世界上在 UNIX、X、Internet 和其他开放系统图书领域具有领导地位的出版公司，同时是联机出版的先锋。

从最畅销的*The Whole Internet User's Guide & Catalog* (被纽约公共图书馆评为20世纪最重要的50本书之一)到GNN(最早的Internet门户和商业网站),再到WebSite (第一个桌面PC的Web服务器软件), O'Reilly & Associates 一直处于Internet发展的最前沿。

许多书店的反馈表明, O'Reilly & Associates 是最稳定的计算机图书出版商 —— 每一本书都一版再版。与大多数计算机图书出版商相比, O'Reilly & Associates 公司具有深厚的计算机专业背景, 这使得 O'Reilly & Associates 形成了一个非常不同于其他出版商的出版方针。O'Reilly & Associates 所有的编辑人员以前都是程序员, 或者是顶尖级的技术专家。O'Reilly & Associates 还有许多固定的作者群体 —— 他们本身是相关领域的技术专家、咨询专家, 而现在编写著作, O'Reilly & Associates 依靠他们及时地推出图书。因为 O'Reilly & Associates 紧密地与计算机业界联系着, 所以 O'Reilly & Associates 知道市场上真正需要什么图书。

出版说明

计算机网络与通信技术的成熟和广泛应用,以及 Internet 与 Web 的迅速发展,为人类的工业生产、商业活动和日常生活都带来了巨大的影响。网络与通信技术在我国的很多领域也已经广泛应用,并且取得了巨大的效益。然而,该领域的技术创新的速度之快也是有目共睹的。为了帮助国内技术人员和网络管理人员在第一时间掌握国外最新的技术,清华大学出版社引进了美国 O'Reilly & Associates 公司的一批在计算机网络理论 and 应用方面代表前沿技术领域的著作,以飨读者。本套丛书采用影印版的形式,力求与国外图书“同步”出版,“原汁原味”地展现给读者各种权威技术理论和技术术语,适合于相关行业的高级技术人员、科研机构研究人员和高校教师阅读。

首批图书包括以下几种:

- 《Web 安全与电子商务》
- 《无线 Java 入门》
- 《Web 缓存》
- 《BEEP 权威指南》
- 《802.11 无线网络权威指南》
- 《大规模局域网设计》
- 《IP 路由》
- 《DNS 与 BIND》

Preface

Computer networks are like any complex engineering project. A small network can be slapped together quite successfully with minimal experience. But a larger network requires careful thought and planning. As with many types of engineering projects, this planning and design phase is best served by an organized and disciplined design philosophy. The trouble with design is that it is difficult to differentiate between personal or near-religious biases and sound ecumenical strategies that can result in better usability, stability, security, and manageability.

Everyone has religious biases when it comes to network design. This is because most networks are so complex that a feeling of black magic falls over anybody trying to understand them. They tend to be too large and too intricate to hold in your mind all at once. So when some particular incantation appears to work miracles, it is adopted as an article of faith. And when a vendor's equipment (or support engineer) saves the day in some important way, it can turn into a blanket belief in that vendor as savior.

So, in the interests of making plain my assumptions and biases, let me explain right from the start that I am a network agnostic. I have used equipment from most of the major vendors, and I believe that every individual piece of gear has its pluses and minuses. I prefer to use the gear that is right for the job, rather than expressing a blind devotion to one or another. So this book is vendor neutral.

I will discuss some proprietary protocols and standards because these are often the best for a particular situation. But in general I will try to lead the reader towards open industry standards: I believe that it is unwise to lock your technology budget to one particular vendor.

In the mainframe-computing era, many firms spent large amounts of money on one company's equipment. Then they found that this required them to continue spending their hardware budget with that company unless they wanted to abandon their initial investment. All incremental upgrades merely reinforced their dependency on

this one vendor. This was fine unless another manufacturer came along with gear that would be better (cheaper, faster, more scalable, etc.) for important business requirements of the company. It is wise to avoid the “fork-lift” upgrade where the entire infrastructure has to be replaced simultaneously to improve performance.

In practice, most LANs are multivendor hybrids. This may be by design or by chance. In many cases a best-of-breed philosophy has been adopted so that a particular type of Ethernet switch is used in the wiring closets, another type at the backbone, with routers from another vendor, while ATM switches and long-haul equipment are provided by still other vendors. In other cases, the multivendor nature of the network is more of an historical accident than intention. And there are also cases where all or nearly all of the network hardware comes from the same manufacturer. If this is the case, then the choice should be made consciously, based on solid technical and business reasons. Having stated my biases here, I leave the reader to make these decisions freely.

Because computer networks are large and complex engineering projects, they should be designed carefully and deliberately. There are many important questions to ask about how a network should function and what purposes it needs to serve. And there are even more questions to ask about how best to meet these objectives. This book will serve as a guide to this process.

Audience

This book is intended for anybody who needs to build or maintain a large-scale network. It is not a theoretical book for classroom use, and it isn’t intended to help programmers with designing applications. Instead it is a hands-on set of rules, guiding principles, and useful tips for people who build networks.

So it should be useful for network-operations people who need to understand the overall logic of their network. It should also be helpful to engineers who need to think about upgrading parts of an existing network in a logical way. And it is particularly relevant for network designers and architects.

In short, this is the book that I always wanted to read when I was starting to work on large networks.

Organization

The book begins in Chapter 1, *Networking Objectives*, by discussing the most important question of all: why build a network in the first place? The answers to this question shape everything that follows. This first chapter also reviews several of the basic networking concepts used throughout the book.

Chapter 2, *Elements of Reliability*, discusses reliability in networks, the factors that lead to a solid network. This includes discussion of how to find the single points of failure in your network. It also includes an important section on how to evaluate the stability of your network and its components statistically.

Chapter 3, *Design Types*, describes many of the most successful design types and their strengths and weaknesses. This chapter should help you to decide on the large-scale shape of the network you build. It includes many ideas for ensuring both reliability and scalability in a large-scale network.

Chapter 4, *Local Area Network Technologies*, delves into the technologies commonly used on LANs. This is intended as a guide to the network designer and implementer, rather than the engineer building the hardware. There are many other books that provide that higher level of detail, which is beyond the scope of this book. The intent here is to provide the information that a network designer needs to make appropriate decisions.

The same is true of Chapter 5, *IP*, which begins a three-chapter discussion of the various Layer 3 and 4 protocols that are commonly used on LANs, focusing in particular on TCP/IP (see Chapter 6, *IP Dynamic Routing*) and IPX (see Chapter 7, *IPX*). At the same time, these chapters look at good ways of implementing networks based on these protocols, as well as appropriate dynamic routing protocols.

Then in Chapter 8, *Elements of Efficiency*, I turn to efficiency. What is meant by efficiency in a network? How is it achieved? In particular, I discuss how to implement Quality of Service (QoS) through a network to ensure that low-priority traffic doesn't interfere with the delivery of important data.

Chapter 9, *Network Management*, is devoted to network management. But, rather than looking at how to manage a network, this section focuses on how a network's design can make it either easier or more difficult to manage. There are several key design decisions that have implications on how the network will later be managed. Since manageability is one of the keys to reliability, it is important to design networks so that they can be managed effectively.

Chapter 10, *Special Topics*, discusses other important considerations that may not be relevant to every network. These include issues of LAN security, designing for multicast capabilities, and inclusion of the new IPv6 protocol.

The Appendix describes in some mathematical detail how to combine statistical probabilities. This is important for estimating failure rates in a complex network. This information is particularly used in Chapter 2 in the section on calculating "Mean time between failures."

The Glossary is a listing of networking terms. Networking has unfortunately become bogged down with highly technical jargon that often makes it difficult to understand things that are actually relatively simple. So this glossary is included to help prevent that from being a problem.

Conventions Used in This Book

I have used the following formatting conventions in this book:

Italic

Used for emphasis and the first use of technical terms, as well as email and URL addresses.

Constant Width

Used for MAC and IP addresses.

Comments and Questions

Please address comments and questions concerning this book to the publisher:

O'Reilly & Associates, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472
(800) 998-9938 (in the United States or Canada)
(707) 829-0515 (international/local)
(707) 829-0104 (fax)

There is a web page for this book, which lists errata, examples, or any additional information. You can access this page at:

<http://www.oreilly.com/catalog/lgscalelans/>

To comment or ask technical questions about this book, send email to:

bookquestions@oreilly.com

For more information about books, conferences, Resource Centers, and the O'Reilly Network, see the O'Reilly web site at:

<http://www.oreilly.com>

Acknowledgments

There are several people who were instrumental in helping me get this book completed. Most important was my editor, Jim Sumser. Jim was extremely encouraging, which was important because of the aggressive schedule that I wanted to follow. He also really helped me organize and present these (often scattered) ideas in a rational way.

I'd like to thank my technical reviewers, who each pointed out more than one serious blunder. Andre Paree-Huff made a number of encouraging comments about the book. Ian J. Brown, CCIE# 3372, showed his vast depth of technical knowledge in his comments, which allowed me to make significant improvements to the quality of

the information presented. And Scott M. Ballew offered many important structural, as well as technical, suggestions for which I am grateful.

In addition, I'd like to thank the following people at O'Reilly: Sheryl Avruch, Mary Brady, Claire Cloutier, Emma Colby, Tatiana Apandi Diaz, Jeffrey Holcomb, Mihaela Maier, Jessamyn Read, Ann Schirmer, Ellie Volckhausen, Melanie Wang, Rachel Wheeler, and Sue Willing.

And, I am particularly grateful to Sherry Biscope for her unconditional support and encouragement throughout this project.

About the Author

Kevin Dooley earned his Ph.D. in physics from the University of Toronto in 1993, having financed much of his schooling as a network and system administrator. He then immediately started a career in networking and has never looked back. He now lives in Toronto, Canada, where he works as a consultant and runs a consulting company called Manageable Networks (<http://www.manageablenetworks.com>). In this capacity he has designed and implemented large-scale Local and Wide Area Networks for some of Canada's largest companies.

In his spare time, Kevin is a jazz pianist and composer. He has a particular interest in chamber music and has served on the board of directors of a chamber music society. One day he plans to try his hand at writing film scores.

Colophon

Our look is the result of reader comments, our own experimentation, and feedback from distribution channels. Distinctive covers complement our distinctive approach to technical topics, breathing personality and life into potentially dry subjects.

The animal on the cover of *Designing Large-Scale LANs* is a reindeer. Reindeer, or caribou, can be found in the arctic tundra, the mountain tundra, and the northern forests of North America, Russia, and Scandinavia. Though the wild caribou and the domesticated reindeer are members of the same species, North America is the only place where a naming distinction is made between the two. Reindeer were first domesticated in Europe and Asia about 5,000 years ago. Most domesticated reindeer come from that same stock, as more recent attempts to tame wild caribou have been unsuccessful. Once domesticated, reindeer were used as food, for their hides, and for transportation (some were saddled and ridden, while others pulled sleds).

Reindeer (or caribou) are the only members of the deer family in which both sexes grow antlers. After mating, adult bulls shed their antlers around November or December, while cows and young often carry their antlers throughout the entire winter. During growth, the antlers have a fuzzy covering, or *velvet*, which contains blood vessels that carry nutrients.

In addition to their antlers, reindeer have lateral hooves that allow their feet to spread on snow or soft ground. The hooves also act as paddles, making them excellent swimmers. Due to breeding, the colors of reindeer vary from white and gray to brown and black. Colors often vary even within the same herd.

Jeffrey Holcomb was the production editor and proofreader for *Designing Large-Scale LANs*. Ann Schirmer was the copyeditor. Sheryl Avruch, Claire Cloutier, Tatiana Apandi Diaz, and Sue Willing provided proofreading assistance. Rachel Wheeler and Mary Brady provided quality control. Phillip Dangler, Darren Kelly, Edie Shapiro, and Sarah Sherman provided production assistance. Nancy Crumpton wrote the index.

Ellie Volckhausen designed the cover of this book, based on a series design by Edie Freedman. The cover image is an illustration from the *Illustrated Natural History: Mammalia*. Emma Colby produced the cover layout with QuarkXPress 4.1 using Adobe's ITC Garamond font.

Melanie Wang designed the interior layout, based on a series design by David Futato. Mihaela Maier converted the files from Microsoft Word to FrameMaker 5.5.6 using tools created by Mike Sierra. The text font is Linotype Birka; the heading font is Adobe Myriad Condensed; and the code font is LucasFont's TheSans Mono Condensed. The illustrations that appear in the book were produced by Robert Romano and Jessamyn Read using Macromedia FreeHand 9 and Adobe Photoshop 6. The tip and warning icons were drawn by Christopher Bing. This colophon was written by Linley Dolby.

Whenever possible, our books use a durable and flexible lay-flat binding.

Table of Contents

Preface	xi
1. Networking Objectives	1
Business Requirements	1
OSI Protocol Stack Model	5
Routing Versus Bridging	10
Top-Down Design Philosophy	12
2. Elements of Reliability	14
Defining Reliability	14
Redundancy	17
Failure Modes	44
3. Design Types	50
Basic Topologies	50
Reliability Mechanisms	61
VLANs	66
Toward Larger Topologies	70
Hierarchical Design	81
Implementing Reliability	113
Large-Scale LAN Topologies	115
4. Local Area Network Technologies	126
Selecting Appropriate LAN Technology	126
Ethernet and Fast Ethernet	130
Token Ring	141
Gigabit and 10 Gigabit Ethernet	146
ATM	149

FDDI	154
Wireless	155
Firewalls and Gateways	160
Structured Cabling	162
5. IP	169
IP-Addressing Basics	170
IP-Address Classes	173
ARP and ICMP	175
Network Address Translation	177
Multiple Subnet Broadcast	179
General IP Design Strategies	182
DNS and DHCP	191
6. IP Dynamic Routing	193
Static Routing	194
Types of Dynamic Routing Protocols	196
RIP	197
IGRP and EIGRP	206
OSPF	213
BGP	228
7. IPX	234
Dynamic Routing	236
General IPX Design Strategies	242
8. Elements of Efficiency	247
Using Equipment Features Effectively	248
Hop Counts	249
MTU Throughout the Network	250
Bottlenecks and Congestion	252
Filtering	253
Quality of Service and Traffic Shaping	254
9. Network Management	273
Network-Management Components	273
Designing a Manageable Network	275
SNMP	280
Management Problems	288

10. Special Topics	306
IP Multicast Networks	306
IPv6	319
Security	333
Appendix: Combining Probabilities	341
Glossary	344
Bibliography	364
Index	369

Networking Objectives

The American architect Louis Henry Sullivan described his design philosophy with the simple statement “form follows function.” By this credo he meant that a structure’s physical layout and design should reflect as precisely as possible how this structure will be used. Every door and window is where it is for a reason.

He was talking about building skyscrapers, but this philosophy is perhaps even more useful for network design. Where building designs often include purely esthetic features to make them more beautiful to look at, every element of a good network design should serve some well-defined purpose. There are no gargoyles or frescos in a well-designed network.

The location and configuration of every piece of equipment and every protocol must be carefully optimized to create a network that fulfills the ultimate purposes for which it was designed. Any sense of esthetics in network design comes from its simplicity and reliability. The network is most beautiful when it is invisible to the end user.

So the task of designing a network begins with a thorough study of the required functions. And the form will follow from these business requirements.

Business Requirements

This is the single most important question to answer when starting a network design: why do you want to build a network? It sounds a little silly, but frequently people seem confused about this point. Often they start building a network for some completely valid and useful reason and then get bogged down in technical details that have little or nothing to do with the real objectives. It is important to always keep these real objectives in mind throughout the process of designing, implementing, and operating a network.

Too often people build networks based on technological, rather than business, considerations. Even if the resulting network fulfills business requirements, it will usually be much more expensive to implement than is necessary.

If you are building a network for somebody else, then they must have some reason why they want this done. Make sure you understand what the real reasons are. Too often user specifications are made in terms of technology. Technology has very little to do with business requirements. They may say that they need a Frame Relay WAN, or that they need switched 100Mbps Ethernet to every desk. You wanted them to tell you why they needed these things. They told you they needed a solution, but they didn't tell you what problem you were solving.

It's true that they may have the best solution, but even that is hard to know without understanding the problem. I will call these underlying reasons for building the network "business requirements." But I want to use a very loose definition for the word "business." There are many reasons for building a network, and only some of them have anything to do with business in the narrow sense of the word. Networks can be built for academic reasons, or research, or for government. There are networks in arts organizations and charities. Some networks have been built to allow a group of friends to play computer games. And there are networks that were built just because the builders wanted to try out some cool new technology, but this can probably be included in the education category.

What's important is that there is always a good reason to justify spending the money. And once the money is spent, it's important to make sure that the result actually satisfies those requirements. Networks cost money to build, and large networks cost large amounts of money.

Money

So the first step in any network design is always to sit down and list the requirements. If one of the requirements is to save money by allowing people to do some task faster and more efficiently, then it is critical to understand how much money is saved.

Money is one of the most important design constraints on any network. Money forms the upper limit to what can be accomplished, balancing against the "as fast as possible" requirement pushing up from below. How much money do they expect the network to save them? How much money do they expect it will make for them? If you spend more money building this network than it's going to save (or make) for the organization, then it has failed to meet this critical business objective. Perhaps neither of these questions is directly relevant. But in that case, somebody is still paying the bill, so how much money are they willing to spend?

Geography

Geography is the second major requirement to understand. Where are the users? Where are the services they want to access? How are the users organized geographically? By geography I mean physical location on whatever scale is relevant. This