



Russ White  
Jeff Tantsura

# NAVIGATING NETWORK COMPLEXITY

Next-generation Routing with SDN,  
Service Virtualization, and Service Chaining

FREE SAMPLE CHAPTER

SHARE WITH OTHERS



# Navigating Network Complexity

*This page intentionally left blank*

---

---

# Navigating Network Complexity

Next-generation Routing with SDN,  
Service Virtualization, and  
Service Chaining

Russ White  
Jeff Tantsura

◆ Addison-Wesley  
800 East 96th Street  
Indianapolis, Indiana 46240 USA

## **Navigating Network Complexity**

### **Next-generation Routing with SDN, Service Virtualization, and Service Chaining**

Copyright © 2016 Pearson Education, Inc.

Published by: Addison Wesley

All rights reserved. Printed in the United States of America. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, 200 Old Tappan Road, Old Tappan, New Jersey 07675, or you may fax your request to (201) 236-3290.

Text printed in the United States on recycled paper at RR Donnelley, Crawfordsville, IN

First Printing November 2015

Library of Congress Cataloging-in-Publication Number: 2015950654

ISBN-13: 978-0-13-398935-9

ISBN-10: 0-13-398935-6

#### **Trademarks**

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed with initial capital letters or in all capitals.

#### **Warning and Disclaimer**

The author and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

#### **Special Sales**

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at [corpsales@pearsoned.com](mailto:corpsales@pearsoned.com) or (800) 382-3419.

For government sales inquiries, please contact [governmentsales@pearsoned.com](mailto:governmentsales@pearsoned.com).

For questions about sales outside the U.S., please contact [international@pearsoned.com](mailto:international@pearsoned.com).

Visit us on the Web: [informit.com/aw](http://informit.com/aw)

#### **Publisher**

Paul Boger

#### **Associate Publisher**

David Dusthimer

#### **Executive Editor**

Brett Bartow

#### **Senior Development Editor**

Christopher Cleveland

#### **Managing Editor**

Sandra Schroeder

#### **Project Editor**

Mandie Frank

#### **Copy Editor**

Cenveo® Publisher Services

#### **Technical Editors**

Ignas Bagdonas

Jon Mitchell

#### **Editorial Assistant**

Vanessa Evans

#### **Cover Designer**

Alan Clements

#### **Composition**

Cenveo® Publisher Services

#### **Indexer**

Cenveo® Publisher Services

# About the Authors



**Russ White** began his network engineering career installing terminal emulation cards and inverse multiplexers in the United States Air Force. In 1996, he moved to Raleigh, N.C., to join Cisco Systems in the Technical Assistance Center (TAC) routing protocols team. From TAC, Russ moved to the global escalation team, and then into engineering, and finally into sales as a Distinguished Architect. He is currently a network architect working in the area of network complexity and large scale design, a member of the IETF Routing Area Directorate, an active speaker and writer, and active in the Internet Society. He holds CCIE #2637, CCDE 2007:001, the CCAr, a Masters in Information Technology from Capella University, and a Masters in Christian Ministry from Shepherds Theological Seminary. He lives in Oak Island, N.C., with his wife and two children, and is currently a P.h.D student at Southeastern Baptist Theological Seminary.



**Jeff Tantsura** started his network engineering career in early 1990s at a small ISP as system/network administrator, later working for bigger ISPs where he was responsible for network design and architecture, vendor selection.

Currently Jeff is heading Technology Strategy Routing at Ericsson as well as chairing IETF Routing Working Group.

Jeff holds MSc in Computer Science and Systems Analysis from University of Georgia and Executive Certificate of Business Excellence from Haas School of Business, Berkeley.

He also holds CCIE R&S #11416 and Ericsson Certified Expert IP Networking #8

Jeff lives in Palo Alto, CA, with his wife and youngest child.

## **About the Technical Reviewers**

### **Ignas Bagdonas**

Ignas Bagdonas has been involved in the network engineering field for last two decades, covering operations, deployment, design, architecture, development, and standardization aspects. He has worked on multiple large SP and enterprise networks worldwide, participated in many of the world's first technology deployments, and has been involved with building community awareness via conferences, seminars, and workshops. His current focus covers end-to-end network architecture evolution and new emerging technologies. Ignas holds Cisco CCDE and CCIE certifications.

### **Jon Mitchell**

Jon Mitchell, CCIE No. 15953, is a network engineer in Google's Technical Infrastructure organization where he works on their global backbone. Prior to Google, Jon has worked in roles of network architecture at Microsoft, systems engineering at Cisco Systems, network architecture and engineering at AOL, and network engineering at Loudcloud for the last 15 years that he has been in the networking industry. He is also an active IETF participant in various routing area working groups. Through all of these roles, Jon has always had a passion for working on large-scale problems and solving them through simplification and automation. When Jon is not thinking about networking, he enjoys many other passions such as hiking, running, supporting clean water, microbreweries, and spending time with his wife and four children.

# Dedications

*Russ White: This book is dedicated to Bekah and Hannah. Thank you for sticking with your grumpy old dad through thick and thin.*

*Jeff Tantsura: This book is dedicated to my family: Marina, Ilia, Miriam, and Davy—thank you for your support!*



# Acknowledgments

## **Russ White:**

I would like to thank the many people who have taught me networking through the years, including Denise Fishburne, Don Slice, Alvaro Retana, Robert Raszuk, and a host of others—too many to name. I would also like to thank Dr. Doug Bookman for driving me to be a better thinker, Dr. Will Coberly for driving me to be a better writer, and Dr. Larry Pettegrew for driving me just to be a better person. Finally, I'd like to thank Greg Ferro and Ethan Banks for inspiring me to start writing again.

## **Jeff Tantsura:**

I would like to thank the people who taught and helped me through the years to understand networking better: Acee Lindem, Tony Przygienda, Tony Li, Jakob Heitz and many others, too many to help.

Special thanks to my co-author Russ for inspiring me!

# Contents at a Glance

Introduction.....	xvii
Chapter 1: Defining Complexity .....	1
Chapter 2: Components of Complexity.....	25
Chapter 3: Measuring Network Complexity .....	47
Chapter 4: Operational Complexity.....	61
Chapter 5: Design Complexity.....	79
Chapter 6: Managing Design Complexity .....	113
Chapter 7: Protocol Complexity .....	137
Chapter 8: How Complex Systems Fail .....	163
Chapter 9: Programmable Networks .....	185
Chapter 10: Programmable Network Complexity .....	215
Chapter 11: Service Virtualization and Service Chaining .....	233
Chapter 12: Virtualization and Complexity .....	249
Chapter 13: Complexity and the Cloud.....	267
Chapter 14: A Simple Ending.....	279
Index .....	289

# Contents

<b>Introduction</b> .....	<b>xvii</b>
How This Book Is Organized .....	xvii
<b>Chapter 1: Defining Complexity</b> .....	<b>1</b>
What Is Complexity? .....	3
Anything I Don't Understand .....	3
Anything with a Lot of Parts .....	5
Anything for Which There Is More State Than Required to Achieve a Goal .....	9
Unintended Consequences .....	11
Why So Much Complexity? .....	13
Future Extensions versus New Protocols .....	16
Unexpected Errors .....	17
Why Not Build Infinitely Complex Systems? .....	18
Quick, Cheap, and High Quality: Choose Two .....	20
Consistency, Availability, and Partition Tolerance: Choose Two .....	21
Journey into the Center of Complexity .....	22
<b>Chapter 2: Components of Complexity</b> .....	<b>25</b>
Network Convergence .....	26
Path Vector: A BGP Example .....	26
Distance Vector: An EIGRP Example .....	28
Link State: OSPF and IS-IS Convergence .....	30
State .....	31
Amount of Information .....	32
An Example of State Failure in the Real World .....	33
Final Thoughts on State .....	34
Speed .....	35
The Network That Never Converges .....	35

The Flapping Link .....	37
Final Thoughts on Speed .....	38
Surface .....	38
The Hourglass Model .....	41
Optimization .....	44
A Final Thought .....	46
<b>Chapter 3: Measuring Network Complexity .....</b>	<b>47</b>
Some Measures of Network Complexity .....	49
Network Complexity Index .....	49
Modeling Design Complexity .....	51
NetComplex .....	53
Organized Complexity .....	55
Is This a Waste of Time? .....	58
A Final Thought .....	58
<b>Chapter 4: Operational Complexity .....</b>	<b>61</b>
Exploring the Problem Space .....	61
The Cost of Human Interaction with the System .....	61
Policy Dispersion versus Optimal Traffic Handling .....	66
Solving the Management Complexity Problem .....	69
Automation as a Solution to Management Complexity ....	69
Modularity as a Solution to Management Complexity ....	72
Protocol Complexity versus Management Complexity ....	74
A Final Thought .....	77
<b>Chapter 5: Design Complexity .....</b>	<b>79</b>
Control Plane State versus Stretch .....	81
Aggregation versus Stretch .....	83
Traffic Engineering .....	85
State versus Stretch: Some Final Thoughts .....	87
Topology versus Speed of Convergence .....	88
Ring Topology Convergence .....	88
Redundancy versus Resilience .....	90
Topology versus Speed of Convergence:	
Some Final Thoughts .....	93
Fast Convergence versus Complexity .....	94

Improving Convergence with Intelligent Timers:	
Talk Faster .....	96
Removing Timers from Convergence: Precompute .....	99
Working around Topology: Tunneling to the Loop-Free	
Alternate .....	101
Some Final Thoughts on Fast Convergence .....	104
Virtualization versus Design Complexity .....	106
Functional Separation .....	108
Forwarding Plane Complexity .....	109
Control Plane Complexity .....	109
Shared Fate Risk Groups .....	111
A Final Thought .....	111
<b>Chapter 6: Managing Design Complexity .....</b>	<b>113</b>
Modularity .....	113
Uniformity .....	114
Interchangeable Modules .....	120
How Modularity Attacks the Complexity Problem .....	121
Information Hiding .....	122
Aggregation .....	122
Failure Domains and Information Hiding .....	126
Final Thoughts on Information Hiding .....	128
Models .....	129
Waterfall .....	129
Places in the Network .....	131
Hierarchical .....	132
UML .....	134
A Final Thought .....	136
<b>Chapter 7: Protocol Complexity .....</b>	<b>137</b>
Flexibility versus Complexity: OSPF versus IS-IS .....	138
Layering versus Protocol Complexity .....	141
The Seven-Layer Model .....	143
The Four-Layer Model .....	146
The Iterative Model .....	147
Protocol Stacks and Design .....	148
Protocol Complexity versus Design Complexity .....	149
Microloops and Fast Reroute .....	149

EIGRP and the Design Conundrum .....	158
Final Thoughts on Protocol Complexity .....	162
<b>Chapter 8: How Complex Systems Fail .....</b>	<b>163</b>
Feedback Loops .....	164
Positive Feedback Loops in Network Engineering .....	169
Speed, State, and Surface: Stability in the Network Control Plane .....	174
Shared Fate .....	177
Virtual Circuits .....	177
TCP Synchronization as a Shared Fate Problem .....	179
A Final Thought .....	181
Thoughts on Root Cause Analysis .....	181
Engineering Skills and Failure Management .....	182
<b>Chapter 9: Programmable Networks .....</b>	<b>185</b>
Drivers and Definition .....	186
Business Drivers .....	186
The Ebb and Flow of Centralization .....	188
Defining Network Programmability .....	191
Use Cases for Programmable Networks .....	193
Bandwidth Calendaring .....	193
Software-Defined Perimeter .....	196
Programmable Network Interfaces .....	200
The Programmable Network Landscape .....	201
OpenFlow .....	202
YANG .....	204
Path Computation Element Protocol .....	207
Interface to the Routing System .....	210
A Final Thought .....	212
<b>Chapter 10: Programmable Network Complexity .....</b>	<b>215</b>
The Subsidiarity Principle .....	216
Policy Management .....	217
Policy Dispersion .....	220
Policy Consistency .....	222
Policy Complexity .....	223
Surface and the Programmable Network .....	224

Impact on Failure Domains .....	226
Wide Area Failure Domains .....	227
Data Center Failure Domains .....	228
Application to Control Plane Failure Domain .....	229
Controller to Controller Failure Domain .....	229
Final Thoughts on Failure Domains .....	229
A Final Thought .....	230
<b>Chapter 11: Service Virtualization and Service Chaining .....</b>	<b>233</b>
Network Function Virtualization .....	234
NFV: A Use Case .....	236
Service Chaining .....	242
Service Function Chaining .....	243
Segment Routing .....	245
A Final Thought .....	248
<b>Chapter 12: Virtualization and Complexity .....</b>	<b>249</b>
Policy Dispersion and Network Virtualization .....	250
State and Service Chaining .....	253
State and Optimization .....	254
Surface and Policy Interaction .....	255
Surface and Policy Proxies .....	255
Other Design Considerations .....	256
Coupling and Failure Domains .....	257
Troubleshooting .....	260
The Orchestration Effect .....	262
Managing Complexity .....	264
A Final Thought .....	266
<b>Chapter 13: Complexity and the Cloud .....</b>	<b>267</b>
Where Does the Complexity Live? .....	268
Cloud Centric .....	269
Vendor Centric .....	270
Network Centric .....	271
Is There a “Right Way?” .....	271
Centralize What? .....	272

Cloudy Complications .....	273
Security .....	273
Data Portability .....	276
A Final Thought .....	277
<b>Chapter 14: A Simple Ending .....</b>	<b>279</b>
Defining Complexity .....	279
Difficult to Understand .....	280
Unintended Consequences .....	280
Large Numbers of Interacting Parts .....	280
What Makes Something “Too Complex”? .....	281
Complexity Is a Tradeoff .....	282
Modeling Complexity .....	284
Managing Complexity in the Real World .....	286
Don’t Ignore Complexity .....	286
Find a Model to Contain the Complexity .....	287
A Final Thought .....	288
<b>Index .....</b>	<b>289</b>



*This page intentionally left blank*

# Introduction

Every engineer, no matter what type of engineering they do, face complexity almost constantly. The faster, cheaper, higher quality triad are the constant companion of everything the engineer does. Sometimes, though, the complexity isn't so obvious. For instance, how long will the software project take? Two hours, two weeks, or too long is the common reply—another encounter with complexity.

While research into complexity theory has proceeded apace in the scientific and mathematical worlds, the application of the theories and ideas to more practical engineering problems, particularly in a way that the “average engineer” can read and understand, simply hasn't kept pace. This book aims to fill that gap for network engineering.

In a move that will probably be disappointing to readers with a math degree, this book is devoid of the elegant mathematical models complexity theorists are working with. Rather, the focus is on the practical application of the theoretical constructs of complexity. Instead of being focused on the “pure theory” and math, this book is focused on the practical application of the ideas being investigated by complexity theorists.

This book, then, is targeted at the “average” network engineer who wants to gain an understanding of why particular common constructs work the way they do, such as hierarchical design, aggregation, and protocol layering. By getting behind these widely accepted ways of designing a network, exposing the reasons for the tradeoffs designed into the system in each case, the authors hope the reader learns to take lessons learned in one area and apply them to other areas. After reading this book, network engineers should begin to understand why hierarchical design and layering, for instance, work, and how to see and change the tradeoffs in more specific situations.

## How This Book Is Organized

This book begins, in the first chapter, with a high level view of complexity theory. This is not deep theory, but it is rather designed to provide a hands-on view

of complexity without diving into heavy math (or any math at all). The second chapter provides an overview of various attempts at measuring complexity in a network, including some of the problems plaguing each attempt. The third chapter proposes a model of complexity that will be used throughout the rest of the book.

Chapter 4 through 7 consider examples of complexity in network and protocol design. The point these chapters attempt to drive home is how the models and concepts in the first three chapters. Chapter 8 is something of an important detour into the world of complex system failure in network engineering terms. The rest of the book is dedicated to first providing an overview of three new technologies network engineers face at the time of writing, and then an analysis of each one in terms of complexity tradeoffs.

Note, Documents referenced with the form draft-xxx are IETF works in progress, and therefore do not have stable document names or uniform resource locators (URLs). Because of these limitations, explicit references have not been given, but rather just the title of the document and the document name. To find these documents , please perform a document search at the IETF website ([ietf.org](http://ietf.org)).

## Chapter 3

---

---

# Measuring Network Complexity

Given these four fundamental aspects of complexity—state, speed, surface, and optimization—it only makes sense to measure these four points and generate a single number describing the overall complexity of a given design and deployment structure. It would be nice if there were some way to examine a proposed network design, or a proposed change to a network design, and be able to assign actual numbers to the complexity of each component so the complexity can be compared to any potential gain in performance, or the loss of complexity in one area can be compared to the gain in complexity in another. If it were only that simple.

As it turns out, the effort to measure complexity is, itself, quite complex.

Two problems rise to the surface when examining the problem of measuring and quantifying a network toward gaining an understanding of the overall system complexity. *First*, there is the sheer amount of information available. Given the current push toward big data analytics, and the ability to measure thousands to millions of interactions and data mining to discover important trends and artifacts, shouldn't something the size of an average network be an easy problem? Consider some of the various points of measurement just in trying to understand the interaction between the data flowing through each point in the network and the queuing mechanisms used to handle that traffic. This might include things such as:

- The amount of data flowing through each point in the network, including the input and output queue of each forwarding device.
- The depth and state of each queue of each forwarding device in the network.

- The source, destination, and other header information of each packet forwarded through the network.
- The number of packets dropped by each forwarding device, including the reason why they were dropped (tail drop, packet error, filtered, filtering rule, etc.).

Considering that the measurements themselves must pass through the network—and the measurements can easily contain more traffic than the measured traffic—the problems with measuring everything should quickly become apparent. How can you separate the measurement from the measured if the measurement is being carried on the same channel as what you are measuring? Added to this challenge are the states of each individual control plane system, and the components of those systems—things like the memory and processor utilization of each forwarding device, the state of each adjacency between each pair of devices participating in the control plane, and the flow of each reachability advertisement within the control plane. To make measuring the system complexity even more complex, the interactions between the systems must also somehow be taken into account—things like the impact of reachability information on the distribution and application of policy, any interdependencies between parallel control planes in terms of reachability information and system resources, and interactions between overlay and underlay control planes. Measuring not only the systems but also the interactions between the systems quickly becomes an intractable problem.

When measuring a system to understand its complexity level, some sort of sampling must take place. Sampling necessarily means that some information must be left out—which, in turn, means that any measurement of complexity along these lines is necessarily an abstract representation of the complexity, rather than a measure of the complexity itself.

To top all of this complexity off, there is very little agreement on the set of things to measure to create even an accurate abstract representation of the complexity of a network.

There is a second problem looming on the horizon just past this first one—a problem that's not so obvious, and actually makes the problem of measuring network complexity intractable. Network design represents ordered (or intentional or organized—these three terms are often used interchangeably) complexity, rather than unordered complexity. While data analytics deals with unordered data well enough, ordered complexity is an entirely different problem set.

Let's begin by examining some methods proposed to measure network complexity, and then consider ordered versus unordered complexity. Finally, several realms of complexity will be examined that will lead to practical applications.

---

## Some Measures of Network Complexity

The difficulty of the task hasn't stopped researchers from attempting to measure network complexity. Quite the opposite—there are a number of methods that have been tried over the years. Each of these methods has contributed useful thinking to the problem space, and can actually be used to provide some insight into what network complexity looks like. Overall, though, none of these measurements will truly provide a complete view of the complexity of a network.

Let's look at three examples of network complexity measurements to get a feel for the space.

### Network Complexity Index

The Network Complexity Index is described in “A Network Complexity Index for Networks of Networks”<sup>1</sup> by Bailey and Grossman (commonly called the NCI). The general idea is to tackle describing network complexity in two steps:

- Break the network down into subnetworks. As described in the original paper:

*Given a network  $N$ , we first divide the network into smaller sub-networks  $C[1], \dots, C[j], \dots, C[p]$  with the property that two nodes selected at random from the sub-network  $C[i]$  are more likely to be connected to each other than two nodes selected at random from outside the sub-network ( $N \setminus C$ ).*

- Compute the complexity based on the size and number of the subnetworks. Again, as described in the original paper:

*Given the sub-communities of the network  $N$ , let  $X[j]$  denote the size of the  $j$  largest sub-community, so that the sequence  $X[1], \dots, X[p]$  is in decreasing order. In general, different communities may have the same size. We define the network complexity index  $B(N)$  of the network  $N$  as the solution of the following equation:  $B(N) = \text{Max } j, X[j]^j$*

The equation given is a standard statistic used in evaluating the importance of scientific research known as the H-index. The H-index determines the impact of

---

1. Stewart Bailey and Robert L. Grossman, “A Network Complexity Index for Networks of Networks” (Infoblox, 2013), n.p., <https://web.archive.org/web/20131001093751/http://flowforwarding.org/docs/Bailey%20-%20Grossman%20article%20on%20network%20complexity.pdf>.

a particular piece of research by evaluating the number of citations of the work in a way that is similar to a web page search index using the number of links to a page to determine the importance or relevance of that page.

Seen this way, the NCI attempts to combine the connectivity within a network with the number of nodes within a network:

- The more the subcommunities, the more connection points there must be between these subcommunities, and hence the more complex the connection graph must be.
- The larger the subcommunities, the more nodes there are within the network; this again impacts the implied connectivity graph of the network

The size and scope of the connectivity graph, in turn, impacts the way information flows within the network, which also relates to the complexity of the network.

### *What the NCI Does Well*

The NCI does a good job of producing a single number that describes the size and shape of a network in terms of nodes and communities, in turn implying the scope and complexity of the network interconnections. This single number, computed over time, can help network managers and designers understand the growth of a network in terms other than sheer size.

### *What the NCI Doesn't Do*

From a network engineer's perspective, there are several practical problems with using the NCI as a single measure of network complexity. *First*, this isn't something you're going to compute on a napkin while you're eating dinner, or do rough calculations in your head around.<sup>2</sup> This is a math heavy computation that requires automated tools to compute. *Second*, other than measuring the growth and interconnectedness of a topology, it's hard to see how and where the NCI is useful in the real world. There's no obvious way to reduce network complexity as measured by the NCI other than reducing the number and size of the subcommunities in the network.

This second objection, however, leads to another shortcoming of the NCI: it doesn't really measure the complexity network operators interact with. It's quite

---

2. In fact, an entire project called Tapestry was built around measuring the NCI by gathering configurations automatically and running them through a processor. The project can be found on GitHub at <https://github.com/FlowForwarding/tapestry>.

common, in the real world, to find very large networks supporting only a few workloads that have been heavily optimized for that workload, and hence are not very complex from an engineer's point of view. It's also quite common, in the real world, to find small networks with a very diverse workload, and hence cannot be optimized for a single workload. These networks are more complex than their size indicates—the NCI would likely underestimate the complexity of these networks.

So what does the NCI miss? Just those pieces of network architecture that designers deal with most of the time, such as:

- Policy, expressed through configuration, metrics, protocols, and other methods
- Resilience, expressed through the amount of redundancy, fast convergence mechanisms, and other highly complex design components

So while the NCI is useful, it doesn't capture all the complexity of a single network in a way that can be usefully applied to real-world networks.

## Modeling Design Complexity

In a set of slides presented to the Internet Research Task Force's Network Complexity Research Group, a group of researchers described a model for measuring and describing the complexity of enterprise routing design.<sup>3</sup> The process of measurement is as follows:

1. Decompose the network design into individual pieces, implemented as individual configuration components (across all devices in the network).
2. Build a network of connections between these individual components.
3. Measure this network to determine the complexity of the configuration.

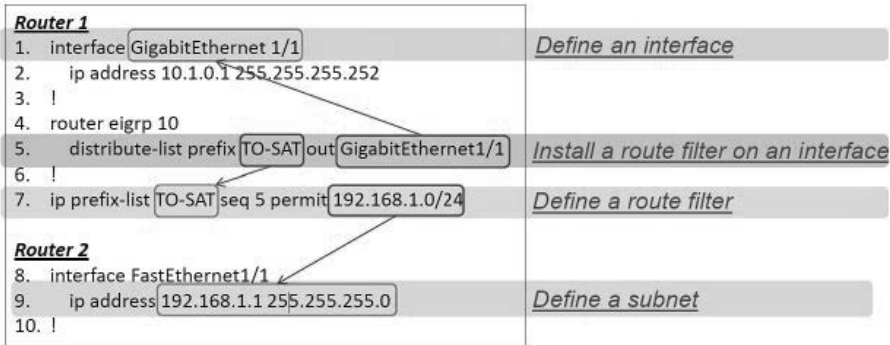
Figure 3.1 is taken from these slides,<sup>4</sup> illustrating the linkages between the various configuration components.

---

3. Xin Sun, Sanjay G. Rao, and G.Xie Geoffrey, "Modeling Complexity of Enterprise Routing Design" (IRTF NCRG, November 5, 2012), n.p., accessed October 5, 2014, <http://www.ietf.org/proceedings/85/slides/slides-85-ncrg-0.pdf>.

4. Ibid.





**Figure 3.1** Evaluating Linkages in a Network Configuration

The more items configured, and the more dense the interconnection between the configurations (particularly between boxes), the more complex the design is determined to be. By examining these factors, the process yields a single number describing the complexity of the design. The presentation and the papers written by the same authors extend this concept to determining if the implemented design matches the design intent, given the design intent is stated in a way amenable to the process.

### *What Modeling Design Complexity Does Well*

The concept of measuring not just the lines of configuration, but the interconnections between the lines of configuration, is extremely powerful. There is little doubt that much of complexity comes from the interrelated configurations spread throughout a network required to implement a single policy or to make a single protocol or process work. Looking at the interactions, or network of configurations, also takes the number of lines of configuration somewhat out of the picture as a measure of complexity. Because some devices can express a policy in a few simple lines, while others require a lot of configuration to express the same policy, this is a useful outcome.

### *What Modeling Design Complexity Doesn't Do*

At the same time, however, the interconnections between lines of configuration can fall prey to the same problems just counting the number of lines of configuration can fall prey to—an entire policy might be represented by a single line of configuration on one device, while requiring a number of lines of policy on another. For instance, on Cisco IOS Software, the command *remove-private-as* is used to remove any private AS numbers in a BGP route advertisement. This

single command essentially replaces a set of configuration commands that would necessarily be interconnected, such as a filter and an application of that filter to a particular set of BGP peers. Both configurations are valid and perform the same set of actions, but they would appear to have completely different complexity levels according to the measure as it's described. Further complicating the situation, different BGP implementations might use different sets of command to perform the same action, making one configuration appear more complex than another, although they're both implementing the same policy.

Another failing of the measure described above is that it's not always obvious what pieces fit together to make a policy. For instance, a configuration removing private AS numbers on every eBGP speaker in an AS might not appear to be related within the measurement; there is no specific point at which these multiple configurations overlap or interact in a way that's obvious, *unless you know the intent of the configuration*. Thus some policies might easily be missed as they consist of configurations with no obvious point at which they tie together.

Finally, it's difficult to assess how a single configuration used to implement multiple policies would be managed in this measure of network complexity—and yet, this is one of the thorniest problems to manage from a complexity standpoint, as this is precisely one of those difficult to manage interaction surfaces between otherwise unrelated policy implementations. How do the various policies measured interact? On this point, modeling design complexity is silent.

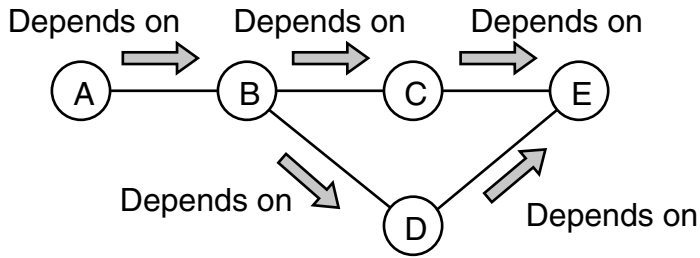
## NetComplex

As previously discussed, the NCI measures complexity based on scale and perceived subcomponents; modeling design rates complexity on the network of interconnected lines of configuration. What about measuring complexity based on the amount of work needed to keep the distributed database that represents the network topology synchronized across all the devices participating in the control plane? This is precisely what NetComplex does. As Chun, Ratnasamy, and Kohler state:

*We conjecture that the complexity particular to networked systems arises from the need to ensure state is kept in sync with its distributed dependencies. The metric we develop in this paper reflects this viewpoint and we illustrate several systems for which this dependency centric approach appears to appropriately reflect system complexity.<sup>5</sup>*

---

5. Byung-Gon Chun, Sylvia Ratnasamy, and Eddie Kohler, "NetComplex: A Complexity Metric for Networked System Designs" (5th Usenix Symposium on Networked Systems Design and Implementation NSDI 2008, April 2008), n.p., accessed October 5, 2014, <http://berkeley.intel-research.net/sylvia/netcomp.pdf>.



**Figure 3.2** *Dependency and Complexity in NetComplex*

NetComplex evaluates the chain of dependent states in the network, assigns a metric to each dependency, and then calculates a single complexity measure based on these assigned metrics. Figure 3.2 illustrates dependency and complexity in NetComplex.

In this figure:

- Routers C and D depend on E to obtain a correct view of the network beyond Router E.
- Router B depends on Routers C, D, and E to obtain a correct view of the network beyond Router E.
- Router A depends on Routers B, C, D, and E to obtain a correct view of the network beyond Router E.

Hence, Router A “accumulates” the complexity of synchronization of information originating beyond E through the entire network. Through these dependencies, Router A is said to be linked to the remaining routers in the network. By examining these links, and combining them with the local state, the complexity of keeping the entire control plane synchronized can be given a single metric.

### *What NetComplex Does Well*

By focusing on the amount of state and the way state is carried through the network, NetComplex does a good job of describing the complexity of a control plane. Based on this, NetComplex is useful for determining the additional complexity required to carry source routing information through the network, and forward based on this source routing information. Another place where NetComplex would be useful is in putting a metric on the additional state information required to forward traffic on a per flow, rather than per destination/virtual topology basis.

### *What NetComplex Doesn't Do*

NetComplex, however, is focused on the control plane within a single administrative or failure domain. There is no way, for instance, to account for the information hidden through route aggregation, nor to differentiate between topology information and reachability (such as what happens at a link state flooding domain boundary). NetComplex doesn't work with policies, nor policy implementation; nor does it deal with traffic flows, subnetwork scale, or network density.

---

## Organized Complexity

Three different measures of network complexity have been examined at this point: NCI, modeling design complexity, and NetComplex. Each of these attempts to measure, in some way, at least some component of the four realms of network complexity—state, speed, surface, and optimization. None of them, however, measure everything in any one of these three domains, and none of them even come close to measuring overall network complexity. Why? The problem isn't just the ability to measure and process all the information needed to produce a single complexity number, it's embedded in the problem of network complexity itself.

Imagine, for a moment, a pool table with a set of balls on it. These specific balls are (at least nearly) perfect in their resilience, so they lose only infinitely small amounts of energy when they strike another object, and the bumpers on the sides of the table are designed in much the same way. There are no pockets in this table, either, so there is no place for the balls to leave the table. Now, place the balls on the table in some random distribution, and then strike one so it starts a chain reaction. The result will be a statistically random set of movements, each ball moving about the table, striking another ball or a bumper, retaining most of its energy, and then moving in a straight line in some other direction.

This particular problem is ripe for statistical regression analysis, or any other form of analysis data science can provide. The data scientist can tell you, based on a set of derived formulas, how often one ball will strike another, how long the system will take to run out of energy, what patterns will form in the randomly moving balls at what time—and many other things. Data science excels at finding patterns in seemingly random bits of data. In fact, it is often found that the data set must be larger to make an accurate prediction; the larger the data set, the more accurate characterization of it can be made, and the more accurate the predictions about the state of that data at some specific point in the future will be.

But let's change the situation somewhat. Let's take the same pool table, the same balls—all the same physical conditions. Only this time, someone has pre-planned the position and movement of every ball such that no two balls strike one another, even though they are all in motion. In fact, the movement of every ball is identical throughout the entire time the balls are in motion.

What can data science tell us about this particular situation? *Nothing*.

Simple observation can tell us which ball will be where at any point in time. Simple observation might even be able to provide a formula telling us where there will be clumps of balls on the table, or near misses. But statistical analysis cannot go much beyond a few simple facts here. What's more interesting is that statistical analysis cannot tell us what the point is in having these balls arranged just this way.

This is the problem of *organized complexity*.

As Warren Weaver noted in 1948:

*This new method of dealing with disorganized complexity, so powerful an advance over the earlier two-variable methods, leaves a great field untouched. One is tempted to oversimplify, and say that scientific methodology went from one extreme to the other—from two variables to an astronomical number—and left untouched a great middle region. The importance of this middle region, moreover, does not depend primarily on the fact that the number of variables involved is moderate—large compared to two, but small compared to the number of atoms in a pinch of salt. The problems in this middle region, in fact, will often involve a considerable number of variables. The really important characteristic of the problems of this middle region, which science has as yet little explored or conquered, lies in the fact that these problems, as contrasted with the disorganized situations with which statistics can cope, show the essential feature of organization. In fact, one can refer to this group of problems as those of organized complexity.<sup>6</sup>*

This field of organized complexity exactly describes the situation engineers face in looking at computer networks. No matter what angle a computer network is approached from, the problem is both complex and organized.

- Protocols are designed with a specific set of goals in mind, a specific mind-set about how the problems approached should be solved, and a set of tradeoffs between current optimal use, future flexibility, supportability, and ease of implementation.
- Applications that run on top of a network are designed with a specific set of goals in mind.

---

6. Warren Weaver, "Science and Complexity," *American Scientist* 36 (1948): 539.

- Control planes that provide the metadata that make a computer network work are designed with a specific set of goals in mind.
- Protocols that carry information through the network, at every level, are designed with a specific set of goals in mind.

No matter which system within computer network is considered—from protocols to design to applications to metadata—each one was designed with a specific set of goals, a specific mindset about how to solve the problems at hand, and a specific set of tradeoffs. Some of these might be implicit, rather than explicit, but they are, nonetheless, intentional goals or targets.

A network is not just a single system that exhibits organized complexity, but a lot of different interlocking systems, each of which exhibits organized complexity, and all of which combined exhibit a set of goals as well (perhaps a more ephemeral set of goals, such as “making the business grow,” but a set of goals nonetheless).

### **A Philosophical Aside**

Within the realm of philosophy, there are those who believe that there is no such thing as organized complexity. Instead, what appears to be organized complexity is simply the result of emergence within any physical system once it becomes complex enough—that organization is somehow “built in” to the natural order, or into the way matter itself is formed and interacts. This school of thought believes that any and all actions can be traced back to some physical cause (for instance, that humans do not actually make decisions as much as decisions happen to humans). Whatever the reader’s stand on this topic (and it is outside the scope of this book to argue the philosophical questions here), the practical result, in terms of network architecture, is: it doesn’t matter. Networks are designed by people to solve a particular set of problems; no matter what is “behind” these designs, we must, to understand computer networks and their designs, get to the “why.” Why did someone design this in this particular way? Why did someone make that particular tradeoff?

Network complexity, then, cannot simply be measured, computed, and “solved,” in the traditional sense. Even everything could be measured in a single network, and even if all the information gathered through such measurement could be processed in a way that made some sense, it would still not be possible to fully express the complexity of a computer network in all its myriad parts—in essence because there is no way to measure or express intent.

---

## Is This a Waste of Time?

None of this means it is wasting time to attempt to measure network complexity. What it does mean, however, is that the problem must be approached with a large dose of humility. Engineers need to be very careful about understanding the tradeoffs being made in every part of the design, and very intentional in remembering that there are limits to accurately predicting the outcome of any particular design decision.

Instead of “giving up,” engineers should do everything possible to understand the complexity, to contain it, to minimize it, and to make intelligent tradeoffs—but there isn’t, and won’t ever be a silver bullet for complexity. As explained in Chapter 1, “Defining Complexity,” there are sets of three out of which only two can be chosen, and there are curves where increasing complexity in one axis to solve a particular problem actually causes problems in another axis.

Measuring and managing complexity is not wasting time unless you believe you can actually *solve the problem*—because the “problem” cannot be “solved.”

---

## A Final Thought

This investigation of complexity has so far concluded:

- Complexity is necessary to solve difficult problems, particularly in the area of robust design.
- Complexity beyond a certain level actually causes brittleness—robust yet fragile.
- Complexity is difficult (or perhaps impossible) to measure in any meaningful way at the systemic level.
- There are a number of classes of problems where it is impossible to resolve for more than two of three goals (such as fast, cheap, high quality).

Given this set of points, it might seem like this is the end of the road. Network engineers are reliant on something that cannot be effectively measured—and measurement is always the first step in controlling and managing a problem set. There will never, in the end, be a single number or formula that can describe network complexity. Should we simply put on our pirate hats and proclaim, “Abandon hope all ye who enter here”? Or is there some way out of this corner?

There is, in fact, a reasonable way to approach complexity in the real world. Rather than trying to find an absolute “measure of complexity,” or find some algorithm that will “solve” complexity, it’s possible to construct a heuristic, or a method of looking at the problem set that will enable a path to a solution. The heuristic, in this case, is a two-part process.

First, expose the complexity tradeoffs inherent in network design. Exposing these tradeoffs will help engineers make intelligent choices about what is being gained, and what is being lost, when choosing any particular solution to a particular problem set. Every problem cannot be solved equally well; any solution applied at one point will increase complexity somewhere else.

To put it in other terms, network *engineers need to learn to be intentional about complexity.*

The next chapter will begin looking at three specific realms of complexity—operational, design, and protocol. In each of these cases, several places where designers must make tradeoffs to illustrate the process of bringing complexity out into the open will be considered. The closer engineers get to making intentional decisions about complexity when designing and managing networks, the more likely meeting the real-world demands placed on networks will be possible.



*This page intentionally left blank*

# Index

## A

abstraction  
  leaky. *See* leaky abstractions  
  orchestration system, 263

aggregation  
  concept, 83  
  information hiding, 122–126  
  as leaky abstraction, 125, 126  
  *vs.* stretch, 83–85, 83

Alderson, David L., 14, 41

API. *See* Application Programming Interface (API)

apocryphal story of the engineer and the hammer, 65–66

application layer  
  four-layer model, 147  
  seven-layer model, 145

Application Programming Interface (API), 9, 143

Application Specific Integrated Circuit (ASIC), 235

application support, policy and, 218

*The Art of Network Architecture* (Cisco Press), 44

ASIC. *See* Application Specific Integrated Circuit (ASIC)

Asynchronous Transfer Mode (ATM), 106

ATM. *See* Asynchronous Transfer Mode (ATM)

auditing, cloud provider, 274

automation  
  abstractions of network state, 71  
  brittleness, 70–71  
  as solution to management complexity, 69–72

availability  
  calculating, 90, 91  
  *vs.* redundancy, 90

## B

bandwidth calendaring, 188, 193–196

bandwidth utilization, 193

BGP. *See* Border Gateway Protocol (BGP)

Border Gateway Protocol (BGP), 9

amount of information carried in, 32

convergence, 26–28  
  formula for, 35  
  propagating a new destination, 35–36

layering, on OSPF, 6

OSFP and, coupling, 257–258

virtualization, 110–111

Brewer, Eric, 21

business continuity, policy and, 218

business drivers, programmable networks, 186–188

## C

capacity *vs.* growth over time, 239, 240

Capital Expenditures (CAPEX), 188

CAP theorem, 21–22, 283

centralization  
  decentralization *vs.*, 189–190  
  ebb and flow of, 188–191

cheap solution. *See* quick, cheap, and high quality conundrum

choke points, 134

Chun, Byung-Gon, 53

Cisco Technical Assistance Center, 235

Clarke, Arthur C., 3

clean solution, 3

CLNP. *See* Connectionless Networking Protocol (CLNP)

Clos, Charles, 114

Clos fabric, 114–115

cloud centric model, 269–270

cloud computing, 267

cloud services  
  centralizing, 272–273  
  cloud centric deployment, 269–270  
  complications, 273–276  
  model, 268–272  
  network centric model, 271  
  overview, 267  
  vendor centric model, 270–271

cohesion principle, 263

command-line interfaces (CLIs), 192–193

- complexity
    - components, 25–46
    - defining, 3–5, 279–282
    - measuring, 47–59
    - negative responses to, 2
    - numbers of moving parts, 5–9, 280
    - overview, 1–3
    - perceptions, 3–11
    - positive responses to, 2
    - reactions to, 2
    - robustness and, 281
    - routing protocols interaction, 6
      - as tradeoff, 20, 282–283
      - understanding, 3–5, 280
      - unintended consequences, 11–13, 280
  - complex solution, 3
  - computing, 272
  - Connectionless Networking Protocol (CLNP), 143
  - continue in segment routing, 247
  - controller to controller failure domain, 229
  - control plane
    - failure domain, 229
    - microloops, 172–174
    - packet loops in, 170–174
    - policy, 219
  - control plane complexity, virtualization, 109
  - control plane state *vs.* stretch, 81–87. *See also* stretch
    - aggregation, 83–85
    - traffic engineering, 85–87
  - convergence, 26–31
    - distance vector, 28–29
    - link state, 30–31
    - path vector, 26–28
    - ring topologies, 88–90
  - coupling and failure domains, 257–260
- D**
- database
    - accessible, 21
    - consistent, 21
    - tolerating partitions, 22
  - Data Center (DC) edge switch, 243
  - data center failure domains, 228–229
  - data link layer, 144
  - data models, 204
  - data portability, 276
  - data science, 55
  - data scientist, 55
  - DC edge router, 244–245
  - decentralization *vs.* centralization, 189–190
  - default free zone, 37
  - Denial of Service (DoS) attacks, 198
  - depth of interaction surfaces, 39–40
  - Designated Router (DR), 140
  - Desktop Virtualization (DV), 273
  - Diffusing Update Algorithm (DUAL), 160, 161
  - Dijkstra, Edsger, 260, 261
  - direct to server attacks, 198
  - distance vector protocols. *See also* Enhanced Interior Gateway Routing Protocol (EIGRP)
    - amount of information carried in, 32
    - convergence, 28–29
      - ring topologies, 88
  - distributed systems, 21
  - Doyle, John C., 14, 41
  - DR. *See* Designated Router (DR)
  - DUAL. *See* Diffusing Update Algorithm (DUAL)
  - Dyche, Jill, 187
- E**
- ebb and flow of centralization, 188–191
  - EIGRP. *See* Enhanced Interior Gateway Routing Protocol (EIGRP)
  - End Systems (ES), 138
  - end-to-end principle, 216. *See also* subsidiarity principle
  - Enhanced Interior Gateway Routing Protocol (EIGRP)
    - concept, 95
    - convergence failure, 160–161
    - convergence process, 28–29
    - design conundrum, 158–162
    - fast convergence, 95–96
    - large-scale environments, 161–162
    - large-scale manual deployments, 63–64
    - stuck in active process, 162
    - stuck in active timer, 161
  - error detection, parity bit and, 17–18
  - event-driven detection, 45–46
  - eXtensible Markup Language (XML). *See* XML
- F**
- failure domains
    - controller to controller, 229
    - control plane, 229
    - coupling and, 257–260
    - data center, 228–229

- and information hiding, 126–128
- programmable networks, 226–230
- wide area, 227–228
- failure theory, 163
- fast convergence, 283
  - intelligent timers, 96–99
  - removing timers from, 99–101
- feedback loops, 164
  - negative feedback loop, 164–166
  - positive feedback loop, 166–174
- flexible design, 217–218
- FIB. *See* Forwarding Information Base (FIB)
- fixed length encoding, OSPF, 139
- fixed packets, OSPF, 139
- flapping links, 37–38
- focused layered functionality, hierarchical models, 133
- focused policy points, hierarchical models, 134
- Forwarding Information Base (FIB), 200. *See also* interfaces, programmable networks
- forwarding plane complexity, 109
- four-layer model, 146–147
  - application layer, 147
  - Internet layer, 147
  - link layer, 146
  - transport layer, 147
- functional separation, virtualization, 108–109

## G

- Garcia, J. J., 152
- Goldberg, Rube. *See* Rube Goldberg Machine
- goto* statements, 260, 261
- Graphical User Interfaces, 192
- growth *vs.* capacity over time, 239, 240

## H

- hack, 4
- hardware, uniformity, 114–116
- hardware abstraction layer, 116
- hiding information. *See* information hiding
- hierarchical models, 132–134
  - basic design, 133
  - focused layered functionality, 133
  - focused policy points, 134
  - modularity and, 121
- hourglass model, 41–44, 42
  - speed, 43
  - state, 43
  - surfaces, 43
- HTML, 135, 205
- human interactions, 61–63

- HyperText Markup Language (HTML).  
*See* HTML

## I

- I2RS. *See* Interface to the Routing System (I2RS)
- IaaS. *See* infrastructure as a service (IaaS)
- IGRP. *See* Interior Gateway Routing Protocol (IGRP)
- information economy, 186–187
- information hiding, 122–129
  - aggregation, 122–126, 123
  - failure domains, 126–128
- information models, 204
- infrastructure as a service (IaaS), 272
- insider attack, 197
- intelligence/analytics, 272
- intelligent timers, convergence with, 96–99
- interaction surfaces, 38–41
  - breadth of, 39–40
  - depth of, 39–40
  - modularization and, 122
  - overlapping interactions, 40–41
- interfaces, programmable networks, 200–201
  - FIB, 200
  - information contained by, 200
  - northbound, 200
  - RIB, 200, 201
  - switching path, 201
- Interface to the Routing System (I2RS), 210–212
  - architecture, 211–212
  - asynchronous, filtered events, 212
  - distributed routing protocols, 212
  - ephemeral state, 212
  - multiple simultaneous asynchronous operations, 211–212
- Interior Gateway Routing Protocol (IGRP), 63–64
- Intermediate Systems (IS), 138
- Intermediate System-to-Intermediate System (IS-IS), 9
  - convergence, 30–31
  - information encoding, 140
  - OSPF *vs.*, 138–141
  - TLV, 139, 140–141
- Internet layer, 147
- Internet Research Task Force, 51
- inventory, interfaces and, 200
- iterative model, 147–148

## J

- Joel on Software*, 117

**K**

Kohler, Eddie, 53

**L**

Label Distribution Protocol (LDP), 158

Label Switched Path (LSP), 75

LAG. *See* Link Aggregation Groups (LAG)

law of leaky abstractions, 117. *See also* leaky abstractions

layering

four-layer model, 146–147

iterative model, 147–148

protocol complexity *vs.*, 141–149

seven-layer model, 43–44, 143–146

LDP. *See* Label Distribution Protocol (LDP)

leaky abstractions, 76, 117–118

aggregation *as*, 125, 126

LFA. *See* Loop Free Alternates (LFA)

Li, Tony, 185

Link Aggregation Groups (LAG), 93

link layer, four-layer model, 146

Link State Advertisement (LSA), 95, 140

Link State Packet (LSP), 139

link state protocols. *See also* Intermediate

System-to-Intermediate System (IS-IS);

Open Shortest Path First (OSPF)

amount of information carried in, 33

control plane traffic flooding, 92–93

convergence, 30–31

ring topologies, 89

Linux kernel code, 261

local information, local control, 216

Loop Free Alternates (LFA), 152–154

remote, 157–158

tunneling to, 101–104

LSA. *See* Link State Advertisement (LSA)

LSA generation timer, 96–98

**M**

management complexity, 69–76

automation *as* solution to, 69–72

modularity *as* solution to, 72–74

protocol complexity *vs.*, 74–76

Mandelbulb, 5–6

Maximum Transmission Unit (MTU), 139

Mean Time Between Mistakes (MTBM), 63, 187

measuring complexity, 47–58

modeling design complexity, 51–53

NetComplex, 53–55

Network Complexity Index, 49–51

organized complexity, 55–57

overview, 47–48

microloops

control plane, 172–174

and fast reroute, 152–158

LFA. *See* Loop Free Alternates (LFA)

NotVia, 155–157

Remote LFA, 157–158

minimum route advertisement interval. *See* MRAI (minimum route advertisement interval)

modeling design complexity, 51–53. *See also* measuring complexity

models, 129–136

hierarchical, 132–134

PIN, 131–132

UML, 134–136

waterfall, 129–130

modularity, 113–122

attacking problems, 121–122

interchangeable modules, 120–121

*as* solution to management complexity, 72–74

taken to extremes, 73–74

uniformity. *See* uniformity

monoculture failures, 114

MPLS, 106, 118, 119–120, 157

MRAI (minimum route advertisement interval), 35, 36

MTBM. *See* Mean Time Between Mistakes (MTBM)

Multiprotocol Label Switching (MPLS). *See* MPLS

mutual redistribution, 171–172

**N**

negative feedback loop, 164–166

NetComplex, 53–55. *See also* measuring complexity

NETCONF, 206, 207

Network Address Translation (NAT), 234, 247

network centric model, 271

Network Complexity Index, 49–51. *See also* measuring complexity

Network Complexity Research Group, 51

network convergence. *See* convergence

Network Function Virtualization (NFV), 234–241

use case, 236–241

network layer, 144–145

Network Time Protocol (NTP), 8

Network Translation, 234

- NFV. *See* Network Function Virtualization (NFV)
- northbound interfaces, 200
- NorVia, 155–157
- NTP. *See* Network Time Protocol (NTP)
- NVGRE, 244
- O**
- obscure code, 3
- Occam's Razor, 10
- OpenFlow, 202–203
- Open Shortest Path First (OSPF), 9, 95
  - aggregated reachability, 139
  - BGP and, coupling, 257–258
  - BGP layering on, 6
  - convergence, 30–31
  - event-driven detection and, 45–46
  - fixed length encoding, 139
  - fixed packets, 139
  - goals, 139
  - information encoding, 140
  - IS-IS *vs.*, 138–141
- Open Systems Interconnect (OSI) protocol
  - stack, 138
- operational complexity, 61–77
- Operational Expenditures (OPEX), 62
- optimal traffic handling *vs.* policy dispersion, 66–69
- optimization, 44–46
  - concept, 44–45
  - event-driven detection, 45–46
  - state and, 254
  - timer-driven detection, 45
- orchestration effect, 262–263. *See also* virtualization
- organized complexity, 55–57. *See also* measuring complexity
- OSI-based networks, 138–139
- overlapping interactions, 40–41. *See also* interaction surfaces
- P**
- PaaS. *See* platform as a service (PaaS)
- packet classification, 256
- packet inspection, 255, 256
- packet loops, 168–170. *See also* positive feedback loop
  - in control plane, 170–174
- parity bit
  - and error detection, 17–18
  - packet processing, 18
- Path Computation Client (PCC), 207
- Path Computation Element Protocol (PCEP), 207–210
  - MPLS labels, 209–210
  - OpenFlow *vs.*, 209
- path vector convergence, 26–28. *See also* Border Gateway Protocol (BGP)
- PCC. *See* Path Computation Client (PCC)
- PCEP. *See* Path Computation Element Protocol (PCEP)
- physical layer, 144
- PIN. *See* Places in the Network (PIN)
- PIX firewall, 234–235
- PIX-PL, 235
- Places in the Network (PIN), 131–132
- platform as a service (PaaS), 273
- policy, 82
  - business requirements and, 217–218
  - control plane, 219
  - described, 219–220
  - management, 217–220
- Policy Based Routing, 82
- policy complexity, programmable network, 223–224
- policy consistency, programmable network, 222–223
- policy dispersion
  - vs.* optimal traffic handling, 66–69
  - programmable network complexity, 220–222
  - virtualization, 250–256
- policy proxies, 255–256
- positive feedback loop, 166–168
  - microloops, 172–174
  - mutual redistribution, 171
  - packet loops, 168–170
- predictability, 262
- presentation layer, 145
- problem space, 61–69
  - human interaction with system, 61–63
  - policy dispersion *vs.* optimal traffic handling, 66–69
  - troubleshooting network failure, 65–66
- programmable networks
  - business drivers, 186–188
  - defining, 191–193
  - drivers for, 186–191
  - ebb and flow of centralization, 188–191
  - failure domain, 226–230
  - interfaces, 200–201
  - landscape, 201–212

programmable networks (*continued*)  
 overview, 185–186  
 policy complexity, 223–224  
 policy consistency, 222–223  
 policy dispersion, 220–222  
 policy management, 217–220  
 software-defined perimeter, 196–199  
 surface, 224–226  
 use cases for, 193–199

protocol complexity  
 design complexity *vs.*, 149–162  
 fast reroute, 152–158  
 flexibility *vs.*, 138–141  
 layering *vs.*, 141–149. *See also* layering  
 management complexity *vs.*, 74–76

protocol stacks, 141–142

**Q**  
 quality solution. *See* quick, cheap, and high quality conundrum  
 quality conundrum  
 quick, cheap, and high quality conundrum,  
 20–21, 283

**R**  
 Ratnasamy, Sylvia, 53  
 redistribution between two routing  
 protocols, 171–172  
 redundancy *vs.* resilience, 90–93  
 Rekhter, Yakov, 185  
 Remote LFA, 157–158  
 REpresentational State Transfer (REST)  
 interface, 207  
 resilience, redundancy *vs.*, 90–93  
 Resource Reservation Protocol (RSVP), 17  
 RESTCONF, 207  
 return on investment, 218  
 RIB. *See* Routing Information Base (RIB)  
 ring topologies, 88–90  
 distance vector protocol, 88  
 fast reroute, 155  
 link state protocol, 89  
 using, reasons for, 89  
 RIP. *See* Routing Information Protocol (RIP)  
 robustness and complexity, 281  
 robust yet fragile, 19  
 root cause analysis, troubleshooting and, 65  
 Routing Information Base (RIB), 200, 201  
 Routing Information Protocol (RIP), 95  
 RSVP. *See* Resource Reservation Protocol (RSVP)  
 Rube Goldberg Machine, 9–11

**S**  
 SaaS. *See* software as a service (SaaS)  
 Saltzer, J. H., 216  
 scale out *vs.* scale up, 238–239  
 virtualization, 263–264  
 screen scraping, 192–193  
 secure information, policy and, 218  
 security  
 as castle, 196  
 centralized data, 274–275  
 cloud-based services, 273–276  
 cross contamination, 275  
 encryption, 276  
 physical media management, 275–276  
 virtual topologies, 107

segment routing, 245–248  
 concept, 245–246  
 continue in, 247  
 example, 246  
 MPLS label, 246, 247  
 MPLS LSPs, 248  
 packet, 247  
 traffic, 247

self-documenting code, 4  
 service chaining, 242–248  
 initial service, 242, 243  
 segment routing, 245–248  
 SFC, 243–245  
 state and, 253

Service Function Chaining (SFC), 243–245  
 architecture of, 244  
 description, 243  
 example, 244  
 specifications, 244  
 tunneling protocols, 244–245

session layer, 145  
 seven-layer model, 43–44, 143–146  
 application layer, 145  
 data link layer, 144  
 Ethernet, 145  
 IP, 145  
 network layer, 144–145  
 physical layer, 144  
 presentation layer, 145  
 session layer, 145  
 TCP, 146  
 transport layer, 145

SFC. *See* Service Function Chaining (SFC)  
 shared fate  
 TCP synchronization, 179–180  
 virtual circuits, 177–178

- Shared Risk Link Groups (SRLG), 111, 177, 178
  - Shortest Path Tree (SPT), 95
  - single responsibility principle, 263
  - software as a service (SaaS), 273
  - software-defined networks (SDN), 185
  - software-defined perimeter, 196–199
  - software environment, 272
  - Source Packet Routing in the Network (SPRING), 245, 246
  - southbound interface, 200–201
  - spaghetti code, 261
  - spaghetti transport system, 118–120
  - Spanning Tree failure, 174–175
  - speed, 35–38
    - flapping links, 37–38
    - network never converging, 35–37
  - Spolsky, Joel, 117
  - SPT. *See* Shortest Path Tree (SPT)
  - state, 31–34
    - amount of information, 32–33
    - and optimization, 254
    - real world failures, 33–34
    - and service chaining, 253
    - vs.* stretch, 87
    - stretch *vs.* *See* stretch
  - static state, 35
  - storage, 272
  - storage as a service, 272
  - stretch
    - aggregation *vs.*, 83–85
    - concept, 81
    - control plane state *vs.*, 81–87
    - increasing, 82
    - measuring, 82
    - state *vs.*, 87
    - traffic engineering, 85–87
  - stuck in active process, of EIGRP, 162
  - stuck in active timer, EIGRP, 161
  - subsidiarity principle, 216–217
  - Super Bowl (1996), 11
  - surface, 38–41. *See also* interaction surfaces
    - policy interaction, 255
    - policy proxies, 255–256
    - programmable network, 224–226
- T**
- TCP synchronization, shared fate problem as, 179–180
  - technical debt, 13
  - telemetry, 200
  - temporary fix, 65
  - temporary workaround, 4
  - three-dimensional Mandelbulb, 5–6
  - thunk, 116
  - timer-driven detection, 45
  - timers, convergence with
    - intelligent, 96–99
    - removing, 99–101
  - topology
    - aggregation of information, 124–125
    - interfaces and, 200
    - ring topologies
    - vs.* speed of convergence, 88–94
    - triangles, 89–90
  - ToR switch, 245
  - “Towards Robust Distributed Systems” (Brewer), 21
  - traceroute, 74–76
  - traffic engineering, 85–87
  - Transmission Control Protocol (TCP), 7, 8
    - internal state, 8
    - leaky abstractions, 117–118
    - NTP and, 8
  - transport layer
    - four-layer model, 147
    - seven-layer model, 145
  - transport system, 118–120
  - triangles, 89–90
  - troubleshooting, virtualization, 260–262
  - troubleshooting network failure, 65–66
    - problem identification, 65
    - problem remediation, 65
    - root cause analysis, 65
  - TTL, 76
  - tunneled fast reroute mechanisms, 103–104
  - tunneling protocols, SFC, 244–245
  - Turing curve, 22, 23
  - Type-Length-Values (TLV), 4, 139, 140–141
    - encodings, 14
    - header, 15
    - on-the-wire bandwidth, 17
    - vs.* optimally structured packet formats, 15–16
- U**
- UDP. *See* User Datagram Protocol (UDP)
  - Unified Modeling Language (UML), 134–136
  - uniformity, 114–120
    - control and management, 116
    - hardware, 114–116
    - transport system, 118–120
    - vendor, 114



unintended consequences, 11–13, 280  
User Datagram Protocol (UDP), 135

## V

vendor, uniformity, 114  
vendor centric model, 270–271  
vendor lock-in, 114  
virtual circuits, 177–178  
Virtual Extensible Local Area Network (VXLAN). *See* VXLAN  
virtualization  
  BGP, 110–111  
  control plane complexity, 109  
  design complexity *vs.*, 106–111  
  design considerations, 256–262  
  forwarding plane complexity, 109  
  functional separation, 108–109  
  managing complexity, 264–265  
  NFV, 234–241  
  orchestration effect, 262–263  
  overview, 249–250  
  policy dispersion, 250–256  
  segment routing, 245–248  
  service chaining. *See* service chaining  
  SFC, 243–245  
  SRLG, 111  
  state and optimization, 254

  state and service chaining, 253  
  surface and policy interaction, 255  
  surface and policy proxies, 255–256  
virtual switch (VSwitch), 245  
VXLAN, 118, 119, 244

## W

waterfall model, 129–130  
Weaver, Warren, 56  
web application, UML for, 134–136  
Weighted Random Early Detection (WRED), 180  
wide area failure domains, 227–228  
WRED. *See* Weighted Random Early Detection (WRED)

## X

XML, 205  
XYZ protocol, 141–142  
XYZ protocol stack, 141–142

## Y

YANG, 204–207  
  description, 204, 205  
  specification, 204

*This page intentionally left blank*



## REGISTER YOUR PRODUCT at [informit.com/register](http://informit.com/register) Access Additional Benefits and SAVE 35% on Your Next Purchase

- Download available product updates.
- Access bonus material when applicable.
- Receive exclusive offers on new editions and related products.  
(Just check the box to hear from us when setting up your account.)
- Get a coupon for 35% for your next purchase, valid for 30 days. Your code will be available in your InformIT cart. (You will also find it in the Manage Codes section of your account page.)

Registration benefits vary by product. Benefits will be listed on your account page under Registered Products.

---

### InformIT.com—The Trusted Technology Learning Source

InformIT is the online home of information technology brands at Pearson, the world's foremost education company. At InformIT.com you can

- Shop our books, eBooks, software, and video training.
- Take advantage of our special offers and promotions ([informit.com/promotions](http://informit.com/promotions)).
- Sign up for special offers and content newsletters ([informit.com/newsletters](http://informit.com/newsletters)).
- Read free articles and blogs by information technology experts.
- Access thousands of free chapters and video lessons.

Connect with InformIT—Visit [informit.com/community](http://informit.com/community)

Learn about InformIT community events and programs.



**informIT.com**

the trusted technology learning source

Addison-Wesley • Cisco Press • IBM Press • Microsoft Press • Pearson IT Certification • Prentice Hall • Que • Sams • VMware Press



# ciscopress.com: Your Cisco Certification and Networking Learning Resource

Subscribe to the monthly Cisco Press newsletter to be the first to learn about new releases and special promotions.

Visit [ciscopress.com/newsletters](http://ciscopress.com/newsletters).

While you are visiting, check out the offerings available at your finger tips.

–Free Podcasts from experts:

- OnNetworking
- OnCertification
- OnSecurity



View them at [ciscopress.com/podcasts](http://ciscopress.com/podcasts).

–Read the latest author articles and sample chapters at [ciscopress.com/articles](http://ciscopress.com/articles).

–Bookmark the Certification Reference Guide available through our partner site at [informit.com/certguide](http://informit.com/certguide).

Connect with Cisco Press authors and editors via Facebook and Twitter, visit [informit.com/socialconnect](http://informit.com/socialconnect).



Connect, Engage, Collaborate

## The Award Winning Cisco Support Community

Attend and Participate in Events

Ask the Experts

Live Webcasts

Knowledge Sharing

Documents

Blogs

Videos

Top Contributor Programs

Cisco Designated VIP

Hall of Fame

Spotlight Awards

Multi-Language Support



<https://supportforums.cisco.com>