



Cisco Intelligent WAN (IWAN)

Brad Edgeworth, CCIE No. 31574

Jean Marc Barozet

David Prall, CCIE No. 6508

Anthony Lockhart

Nir Ben-Dvora

Cisco Intelligent WAN (IWAN)

Brad Edgeworth, CCIE No. 31574

Jean-Marc Barozet

David Prall, CCIE No. 6508

Anthony Lockhart

Nir Ben-Dvora

Cisco Press

800 East 96th Street

Indianapolis, Indiana 46240 USA

Cisco Intelligent WAN (IWAN)

Brad Edgeworth, Jean-Marc Barozet, David Prall, Anthony Lockhart, Nir Ben-Dvora

Copyright © 2017 Cisco Systems, Inc.

Published by:

Cisco Press

800 East 96th Street

Indianapolis, IN 46240 USA

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the publisher, except for the inclusion of brief quotations in a review.

Printed in the United States of America

First Printing November 2016

Library of Congress Control Number: 2016954607

ISBN-13: 978-1-58714-463-9

ISBN-10: 1-58714-463-8

Warning and Disclaimer

This book is designed to provide information about the Cisco Intelligent WAN (IWAN) and Software Defined Wide Area Networking (SD-WAN). Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an “as is” basis. The authors, Cisco Press, and Cisco Systems, Inc. shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc., cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

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 or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact intlcs@pearson.com.

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers' feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through email at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Editor-in-Chief: Mark Taub

Production Line Manager: Brett Bartow

Business Operation Manager, Cisco Press:
Ronald Fligge

Acquisitions Editor: Michelle Newcomb

Managing Editor: Sandra Schroeder

Development Editor: Ellie Bru

Senior Project Editor: Tracey Croom

Copy Editor: Barbara Wood

Technical Editor(s): Denise Fishburne,
Tom Kunath

Editorial Assistant: Vanessa Evans

Cover Designer: Chuti Prasertsith

Composition: codeMantra

Indexer: Lisa Stumpf

Proofreader: H.S. Rupa



Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.



CCDE, CCENT, Cisco Eos, Cisco HealthPresence, the Cisco logo, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, Cisco TelePresence, Cisco WebEx, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play and Learn and Cisco Store are service marks; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCOIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, Gigamon, HomeLink, Internet Quotient, IOS, iPhone, iQuick Study, IronPort, the IronPort logo, LightStream, Linksys, MediaTone, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0812R)

About the Authors

Brad Edgeworth, CCIE No. 31574 (R&S & SP), is a Systems Engineer at Cisco Systems. Brad is a distinguished speaker at Cisco Live where he has presented on a variety of topics. Before joining Cisco, Brad worked as a network architect and consultant for various Fortune 500 companies. Brad's expertise is based on enterprise and service provider environments with an emphasis on architectural and operational simplicity. Brad holds a Bachelor of Arts degree in computer systems management from St. Edward's University in Austin, TX. Brad can be found on Twitter @BradEdgeworth.

Jean-Marc Barozet is a Principal Engineer with the Intelligent WAN (IWAN) Product Management team, helping to architect and lead the Cisco SD-WAN solution. Jean-Marc has more than 25 years of experience in the networking industry and has been with Cisco for more than 19 years. He previously was Consulting Systems Engineer in the Cisco sales organization in France. Before joining Cisco, Jean-Marc worked as a consulting engineer for Alcatel Business Systems. He works closely with the largest Cisco customers to design complex large-scale networks in both the enterprise and service provider verticals. He holds an engineering degree from the National Institute of Applied Sciences of Lyon (INSA). Jean-Marc can be found on Twitter @jbarozet.

David Prall, CCIE No. 6508 (R/S, SP, & Security), is a Communications Architect on the Enterprise Networking Technical Strategy team. David has been with Cisco more than 16 years. He previously held Consulting Systems Engineer and Systems Engineer positions within the US federal area supporting civilian agencies. He is currently focused on the Cisco Intelligent WAN (IWAN), primarily complex routing and switching designs to include design, deployment, and troubleshooting of large networks using IOS, IOS XE, and IOS XR. Areas of expertise include IPv6, multicast, MPLS, fast convergence, and quality of service. He holds a Bachelor of Science degree in computer science from George Washington University. David works out of the Herndon, VA, office.

Anthony Lockhart is a Technical Marketing Engineer for the Enterprise Networking Group at Cisco Systems focusing on WAN optimization. Before joining Cisco, Anthony worked as a network engineer and consultant for various companies over his 19-year IT career. He lives in Kentucky with his wife, Colleen, and they enjoy traveling and sightseeing. Anthony received a bachelor's degree in history and sociology from Houston Baptist University in Houston, TX.

Nir Ben-Dvora is a Technical Leader, Engineering for the Core Software Group at Cisco Systems. Nir has more than 25 years of experience in software development, management, and architecture. Nir has been with Cisco for more than 17 years working as a software architect for various network services products in the areas of security, media, application recognition, visibility, and control. In recent years, Nir has been working as a Software Architect for the Application Visibility and Control (AVC) solution and as part of the Intelligent WAN (IWAN) architecture team. Nir owns multiple patents on network services and is a coauthor of IETF RFC 6759. Nir holds a Master of Science degree in electrical engineering and business management from Tel Aviv University and always likes to learn about and experiment with new technologies. Nir lives in Herzliya, Israel, with his wife, Naama, and their three boys.

About the Technical Reviewers

Tom Kunath, CCIE No. 1679 (R&S), is a Solutions Architect in Cisco Advanced Services, working with Cisco customers to plan, design, and implement large-scale, complex networks. Tom has more than 20 years of experience in the networking industry and in recent years has been exclusively focused on SD-WAN and the Cisco IWAN solution, helping to shape the architecture and lead successful deployments on several early adopter networks. Before joining Cisco, Tom worked at Juniper Networks' Professional Services Group as a Resident Engineer supporting several service provider IP and MPLS backbones, and prior to that as a Principal Consultant at International Network Services (INS). Tom is a frequent speaker on the Cisco Live circuit and is the coauthor of the 2011 Cisco Press Publication *Enterprise Network Testing*.

Denise "Fish" Fishburne, CCDE No. 20090014, CCIE No. 2639 (R&S, SNA), is an Engineer and Team Lead with the Customer Proof of Concept Lab (CPOC) in North Carolina. Fish is a geek who absolutely adores learning and passing it on. She works on many technologies in the CPOC, but her primary technical strength is troubleshooting. Fish has been with Cisco since 1996 and CPOC since 2001 and has been a regular speaker at Networkers/Cisco Live since 2006. Cisco Live is a huge passion for Fish! As such, in 2009, she got even more deeply involved with it by becoming a Cisco Live Session Group Manager. Look for Fish swimming in the bits and bytes all around you, or just go to www.NetworkingWithFish.com.

Dedications

Brad Edgeworth:

This book is dedicated to my family. To my wife, Tanya: Thank you for your support and patience. To my daughter, Teagan: Go after your dreams; anything is possible. And to my dog, Louie: I hope you are chasing tennis balls in heaven. I miss you.

Jean-Marc Barozet:

This book is dedicated to my supportive wife, Laurence—thank you for always being there for me—and to my beloved children, Amélie, Marie, Julien, and Pierre.

David Prall:

I would like to thank my wife, Crystal, and six children, Robbie, Cullen, Abby, Braden, Mackenzie, and Hudson, who allowed me to sit at their sporting and scouting events while working on my laptop.

Anthony Lockhart:

This book is dedicated to my family. To my wife, Colleen: Thank you for always being there for me and putting up with my insane work hours. To my lovely children, Naaman, Taylor, Keegan, and Collin: I love you! And a special shout-out to my stepdaughter, Sienna: What's up?

Nir Ben-Dvora:

Dedicated with love to my wife, Naama, and our three boys, Assaf, Aylon, and Elad, who inspire me in whatever I do.

Acknowledgments

Thank you to our technical editors, Denise and Tom, for making this a better product by pointing out our mistakes and showing us a different perspective for presenting various concepts.

Special thanks go to the Cisco Press team for their assistance and insight throughout this project.

Many people within Cisco have provided feedback, suggestions, and support to make this a great book. Thanks to all who have helped in the process, especially Steven Allspach, Sarel Altshuler, Shashikant Bhadoria, David Block, Bryan Byrne, Patrick Charretour, Gary Davis, Murali Erraguntala, Kelly Fleshner, Mani Ganesan, Jason Gooley, Amos Graitzer, Lior Katzri, Guy Keinan, Arshad Khan, Mike Korenbaum, Manish Kumar, Steve Moore, Noah Ofek, Craig Smith, Sharon Sturdy, Mike Sullenberger, Scott Van de Houten, David Yaron, Tina Yu, and the authors' management teams.

We are extremely grateful to be around all of these people who want to see this knowledge shared with others.

Contents at a Glance

Foreword xxix

Introduction xxxi

Part I Introduction to IWAN

Chapter 1 Evolution of the WAN 1

Part II Transport Independent Design

Chapter 2 Transport Independence 15

Chapter 3 Dynamic Multipoint VPN 35

Chapter 4 Intelligent WAN (IWAN) Routing 109

Chapter 5 Securing DMVPN Tunnels and Routers 219

Part III Intelligent Path Control

Chapter 6 Application Recognition 287

Chapter 7 Introduction to Performance Routing (PfR) 327

Chapter 8 PfR Provisioning 359

Chapter 9 PfR Monitoring 411

Chapter 10 Application Visibility 459

Part IV Application Optimization

Chapter 11 Introduction to Application Optimization 509

Chapter 12 Cisco Wide Area Application Services (WAAS) 537

Chapter 13 Deploying Application Optimizations 581

Part V QoS

Chapter 14 Intelligent WAN Quality of Service (QoS) 623

Part VI Direct Internet Access

Chapter 15 Direct Internet Access (DIA) 671

Part VII Migration

Chapter 16 Deploying Cisco Intelligent WAN 723

Part VIII Conclusion

Chapter 17 Conclusion and Looking Forward 755

Appendix A Dynamic Multipoint VPN Redundancy Models 759

Appendix B IPv6 Dynamic Multipoint VPN 763

Index 779

Contents

Foreword	xxix
Introduction	xxxix

Part I Introduction to IWAN

Chapter 1 Evolution of the WAN 1

WAN Connectivity	1
Leased Circuits	1
Internet	2
Multiprotocol Label Switching VPNs (MPLS VPNs)	3
Increasing Demands on Enterprise WANs	3
Server Virtualization and Consolidation	4
Cloud-Based Services	4
Collaboration Services	4
Bring Your Own Device (BYOD)	5
Guest Internet Access	5
Quality of Service for the WAN	6
Branch Internet Connectivity and Security	6
Centralized Internet Access	7
Distributed Internet Access	8
Cisco Intelligent WAN	8
Transport Independence	8
Intelligent Path Control	9
Application Optimization	10
Secure Connectivity	11
<i>Zone-Based Firewall</i>	11
<i>Cloud Web Security</i>	12
Software-Defined Networking (SDN) and Software-Defined WAN (SD-WAN)	12
Summary	13

Part II Transport Independent Design

Chapter 2 Transport Independence 15

WAN Transport Technologies	15
Dial-Up	15
Leased Circuits	16

Virtual Circuits	16
Peer-to-Peer Networks	17
Broadband Networks	18
Cellular Wireless Networks	19
Virtual Private Networks (VPNs)	20
<i>Remote Access VPN</i>	20
<i>Site-to-Site VPN Tunnels</i>	21
<i>Hub-and-Spoke Topology</i>	21
<i>Full-Mesh Topology</i>	22
Multiprotocol Label Switching (MPLS) VPNs	23
<i>Layer 2 VPN (L2VPN)</i>	23
<i>Layer 3 VPN (L3VPN)</i>	24
<i>MPLS VPNs and Encryption</i>	25
Link Oversubscription on Multipoint Topologies	25
Dynamic Multipoint VPN (DMVPN)	26
Benefits of Transport Independence	28
Managing Bandwidth Cost	30
Leveraging the Internet	31
Intelligent WAN Transport Models	32
Summary	33
Chapter 3 Dynamic Multipoint VPN	35
Generic Routing Encapsulation (GRE) Tunnels	36
GRE Tunnel Configuration	37
GRE Example Configuration	40
Next Hop Resolution Protocol (NHRP)	42
Dynamic Multipoint VPN (DMVPN)	44
Phase 1: Spoke-to-Hub	45
Phase 2: Spoke-to-Spoke	45
Phase 3: Hierarchical Tree Spoke-to-Spoke	45
DMVPN Configuration	48
DMVPN Hub Configuration	48
DMVPN Spoke Configuration for DMVPN Phase 1 (Point-to-Point)	50
Viewing DMVPN Tunnel Status	54
Viewing the NHRP Cache	56
DMVPN Configuration for Phase 3 DMVPN (Multipoint)	61

Spoke-to-Spoke Communication	64
Forming Spoke-to-Spoke Tunnels	64
NHRP Route Table Manipulation	70
NHRP Route Table Manipulation with Summarization	72
Problems with Overlay Networks	76
Recursive Routing Problems	76
Outbound Interface Selection	77
Front-Door Virtual Route Forwarding (FVRF)	78
<i>Configuring Front-Door VRF (FVRF)</i>	79
<i>FVRF Static Routes</i>	80
<i>Verifying Connectivity on an FVRF Interface</i>	80
<i>Viewing the VRF Routing Table</i>	81
IP NHRP Authentication	82
Unique IP NHRP Registration	82
DMVPN Failure Detection and High Availability	84
NHRP Redundancy	85
NHRP Traffic Statistics	88
DMVPN Tunnel Health Monitoring	89
DMVPN Dual-Hub and Dual-Cloud Designs	89
IWAN DMVPN Sample Configurations	92
Sample IWAN DMVPN Transport Models	100
Backup Connectivity via Cellular Modem	103
Enhanced Object Tracking (EOT)	103
Embedded Event Manager	104
IWAN DMVPN Guidelines	105
Troubleshooting Tips	106
Summary	107
Further Reading	108

Chapter 4 Intelligent WAN (IWAN) Routing 109

Routing Protocol Overview	109
Topology	112
WAN Routing Principles	114
Multihomed Branch Routing	114
Route Summarization	117
Traffic Engineering for DMVPN and PfR	120

EIGRP for IWAN	122
Base Configuration	123
Verification of EIGRP Neighbor Adjacencies	128
EIGRP Stub Sites on Spokes	129
EIGRP Summarization	133
EIGRP Traffic Steering	137
Complete EIGRP Configuration	140
Advanced EIGRP Site Selection	147
Border Gateway Protocol (BGP)	151
BGP Routing Logic	151
Base Configuration	153
BGP Neighbor Sessions	153
Default Route Advertisement into BGP	159
Routes Learned via DMVPN Tunnel Are Always Preferred	161
Branch Router Configuration	163
<i>Single-Router Branch Sites</i>	163
<i>Multiple-Router Branch Sites</i>	164
Changing BGP Administrative Distance	168
Route Advertisement on DMVPN Hub Routers	169
<i>DMVPN Hub LAN Connectivity Health Check</i>	170
<i>BGP Route Advertisement on Hub Routers</i>	173
<i>BGP Route Filtering</i>	175
<i>Redistribution of BGP into OSPF</i>	178
Traffic Steering	180
Complete BGP Configuration	183
Advanced BGP Site Selection	195
FVRF Transport Routing	199
Multicast Routing	200
Multicast Distribution Trees	200
<i>Source Trees</i>	200
<i>Shared Trees</i>	201
Rendezvous Points	201
Protocol Independent Multicast (PIM)	201
Source Specific Multicast (SSM)	201
Multicast Routing Table	202
IWAN Multicast Configuration	202

Hub-to-Spoke Multicast Stream	205
Spoke-to-Spoke Multicast Traffic	209
<i>Modify the SPT Threshold</i>	212
<i>Modify the Multicast Routing Table</i>	214
Summary	217
Further Reading	217
Chapter 5 Securing DMVPN Tunnels and Routers	219
Elements of Secure Transport	220
IPsec Fundamentals	222
Security Protocols	223
<i>Authentication Header</i>	223
<i>Encapsulating Security Payload (ESP)</i>	223
Key Management	223
Security Associations	224
ESP Modes	224
<i>DMVPN without IPsec</i>	225
<i>DMVPN with IPsec in Transport Mode</i>	225
<i>DMVPN with IPsec in Tunnel Mode</i>	226
IPsec Tunnel Protection	226
Pre-shared Key Authentication	226
<i>IKEv2 Keyring</i>	227
<i>IKEv2 Profile</i>	228
<i>IPsec Transform Set</i>	230
<i>IPsec Profile</i>	232
<i>Encrypting the Tunnel Interface</i>	233
<i>IPsec Packet Replay Protection</i>	234
<i>Dead Peer Detection</i>	234
<i>NAT Keepalives</i>	235
<i>Complete Configuration</i>	235
Verification of Encryption on IPsec Tunnels	236
Private Key Infrastructure (PKI)	239
<i>IOS Certificate Authority (CA) Server</i>	241
<i>DMVPN Hub PKI Trustpoints</i>	246
<i>DMVPN Branch PKI Trustpoints</i>	252
<i>PKI IPsec Protection Configurations</i>	256
<i>Certificate Registration with Out-of-Band Management Tunnel</i>	258

IKEv2 Protection	262
Basic IOS CA Management	263
Securing Routers That Connect to the Internet	264
Access Control Lists (ACLs)	264
Zone-Based Firewalls (ZBFWs)	266
<i>Self</i>	267
<i>Default</i>	267
<i>ZBFW Configuration</i>	268
Control Plane Policing (CoPP)	275
IOS Embedded Packet Capture (EPC)	275
IOS XE Embedded Packet Capture	277
Analyzing and Creating the CoPP Policy	278
Device Hardening	284
Summary	286
Further Reading	286

Part III Intelligent Path Control

Chapter 6 Application Recognition 287

What Is Application Recognition?	287
What Are the Benefits of Application Recognition?	288
NBAR2 Application Recognition	288
NBAR2 Application ID, Attributes, and Extracted Fields	289
NBAR2 Application ID	289
NBAR2 Application Attributes	290
NBAR2 Layer 7 Extracted Fields	293
NBAR2 Operation and Functions	293
Phases of Application Recognition	295
<i>First Packet Classification</i>	295
<i>Multistage Classification</i>	295
<i>Final Classification</i>	296
<i>Further Tracking</i>	296
NBAR2 Engine and Best-Practice Configuration	296
<i>Multipacket Engine</i>	297
<i>DNS Engine</i>	297
<i>DNS Authoritative Source (DNS-AS) Engine</i>	297
<i>DNS Classification by Domain</i>	300

<i>Control and Data Bundling Engine</i>	301
<i>Behavioral and Statistical Engine</i>	301
<i>Layer 3, Layer 4, and Sockets Engine</i>	301
<i>Transport Hierarchy</i>	301
<i>Subclassification</i>	302
Custom Applications and Attributes	303
Auto-learn Traffic Analysis Engine	303
Traffic Auto-customization	305
Manual Application Customization	305
<i>HTTP Customization</i>	306
<i>SSL Customization</i>	306
<i>DNS Customization</i>	307
<i>Composite Customization</i>	307
<i>Layer 3/Layer 4 Customization</i>	308
<i>Byte Offset Customization</i>	308
Manual Application Attributes Customization	308
NBAR2 State with Regard to Device High Availability	310
Encrypted Traffic	310
NBAR2 Interoperability with Other Services	310
NBAR2 Protocol Discovery	311
Enabling NBAR2 Protocol Discovery	311
Displaying NBAR2 Protocol Discovery Statistics	311
Clearing NBAR2 Protocol Discovery Statistics	312
NBAR2 Visibility Dashboard	313
NBAR2 Protocol Packs	314
Release and Download of NBAR2 Protocol Packs	314
NBAR2 Protocol Pack License	315
Application Customization	315
NBAR2 Protocol Pack Types	315
NBAR2 Protocol Pack States	315
Identifying the NBAR2 Software Version	315
Verifying the Active NBAR2 Protocol Pack	316
Loading an NBAR2 Protocol Pack	316
NBAR2 Taxonomy File	318
Protocol Pack Auto Update	318
<i>Protocol Pack Configuration Server</i>	318

	<i>Protocol Pack Source Server</i>	318
	Validation and Troubleshooting	322
	Verify the Software Version	322
	Check the Device License	322
	Verifying That NBAR2 Is Enabled	322
	Verifying the Active NBAR2 Protocol Pack	323
	Checking That Policies Are Applied Correctly	323
	Reading Protocol Discovery Statistics	324
	Granular Traffic Statistics	324
	Discovering Generic and Unknown Traffic	324
	Verifying the Number of Flows	325
	Summary	325
	Further Reading	325
Chapter 7	Introduction to Performance Routing (PfR)	327
	Performance Routing (PfR)	328
	Simplified Routing over a Transport-Independent Design	328
	“Classic” Path Control Used in Routing Protocols	329
	Path Control with Policy-Based Routing	330
	Intelligent Path Control—Performance Routing	332
	Introduction to PfRv3	334
	Introduction to the IWAN Domain	335
	IWAN Sites	337
	Device Components and Roles	339
	IWAN Peering	340
	Parent Route Lookups	342
	Intelligent Path Control Principles	343
	PfR Policies	343
	Site Discovery	343
	Site Prefix Database	345
	PfR Enterprise Prefixes	346
	WAN Interface Discovery	346
	<i>Hub and Transit Sites</i>	347
	<i>Branch Sites</i>	347
	Channel	348
	Smart Probes	350
	Traffic Class	350

- Path Selection 351
 - Direction from Central Sites (Hub and Transit) to Spokes* 351
 - Direction from Spoke to Central Sites (Hub and Transit)* 351
- Performance Monitoring 353
- Threshold Crossing Alert (TCA) 355
- Path Enforcement 356

- Summary 356
- Further Reading 357

Chapter 8 PfR Provisioning 359

- IWAN Domain 360
- Topology 360
 - Overlay Routing 363
 - Advertising Site Local Subnets* 363
 - Advertising the Same Subnets* 364
 - Traffic Engineering for PfR 366
 - PfR Components 367
- PfR Configuration 369
 - Master Controller Configuration 369
 - Hub Site MC Configuration* 369
 - Transit Site MC Configuration* 371
 - Branch Site MC Configuration* 372
 - MC Status Verification* 374
 - BR Configuration 377
 - Transit BR Configuration* 377
 - Branch Site BR Configuration* 381
 - BR Status Verification* 382
 - NetFlow Exports 384
 - Domain Policies 386
 - Performance Policies* 386
 - Load-Balancing Policy* 391
 - Path Preference Policies* 392
 - Quick Monitor* 394
 - Hub Site Master Controller Settings* 395
 - Hub, Transit, or Branch Site Specific MC Settings* 395
- Complete Configuration 396

Advanced Parameters	399
Unreachable Timer	399
Smart Probes Ports	400
Transit Site Affinity	400
Path Selection	401
Routing—Candidate Next Hops	401
Routing—No Transit Site Preference	401
Routing—Site Preference	403
PfR Path Preference	406
PfR Transit Site Preference	407
Using Transit Site Preference and Path Preference	408
Summary	409
Further Reading	410
Chapter 9 PfR Monitoring	411
Topology	412
Checking the Hub Site	413
Check the Routing Table	413
Checking the Hub MC	415
Checking the Hub BRs	417
Verification of Remote MC SAF Peering with the Hub MC	418
Checking the Transit Site	422
Check the Branch Site	423
Check the Routing Table	423
Check Branch MC Status	424
Check the Branch BR	429
Monitoring Operations	435
Routing Table	435
Monitor the Site Prefix	436
Monitor Traffic Classes	438
Monitor Channels	444
Transit Site Preference	450
<i>With Transit Site Affinity Enabled (by Default)</i>	454
<i>With Transit Site Affinity Disabled (Configured)</i>	455
Summary	456
Further Reading	457

Chapter 10 Application Visibility 459

Application Visibility Fundamentals	459
Overview	460
Components	460
Flows	462
<i>Observation Point</i>	464
<i>Flow Direction</i>	464
<i>Source/Destination IP Versus Connection</i>	464
Performance Metrics	465
Application Response Time Metrics	466
Media Metrics	467
Web Statistics	468
<i>HTTP Host</i>	469
<i>URI Statistics</i>	469
Flexible NetFlow	470
Flexible NetFlow Overview	470
Configuration Principles	470
<i>Create a Flexible NetFlow Flow Record</i>	471
<i>Create a Flow Exporter</i>	472
<i>Create a Flow Monitor</i>	474
<i>Apply a Flow Monitor to the WAN</i>	475
Flexible NetFlow for Application Visibility	478
<i>Use Case 1: Flow Statistics</i>	478
<i>Use Case 2: Application Client/Server Statistics</i>	478
<i>Use Case 3: Application Usage</i>	479
Monitoring NetFlow Data	479
<i>View Raw Data Directly on the Router</i>	479
<i>View Reports on NetFlow Collectors</i>	484
Flexible NetFlow Summary	484
Evolution to Performance Monitor	485
Principles	485
Performance Monitor Configuration Principles	487
Easy Performance Monitor (ezPM)	492
<i>Application Statistics Profile</i>	493
<i>Application Performance Profile</i>	493
<i>Application Experience Profile</i>	494

ezPM Configuration Steps	494
Monitoring Performance Monitor	499
Metrics Export	499
Flow Record, NetFlow v9, and IPFIX	499
Terminology	500
NetFlow Version 9 Packet Header Format (RFC 3954)	502
IPFIX Packet Header Format (RFC 7011)	502
Monitoring Exports	502
Monitoring Performance Collection on Network Management Systems	504
Deployment Considerations	505
Performance Routing	505
Interoperability with WAAS	505
Summary	507
Further Reading	507

Part IV Application Optimization

Chapter 11 Introduction to Application Optimization 509

Application Behavior	510
Bandwidth	512
Latency	514
<i>Application Latency</i>	514
<i>Network Latency</i>	515
Cisco Wide Area Application Services (WAAS)	516
Cisco WAAS Architecture	517
<i>Application Optimizers</i>	518
<i>Configuration Management System</i>	519
<i>Data Redundancy Elimination (DRE) with Scheduler</i>	519
<i>Storage</i>	519
<i>Network I/O</i>	519
<i>Interception and Flow Management</i>	519
TCP Optimization	520
<i>TCP Windows Scaling</i>	521
<i>TCP Initial Window Size Maximization</i>	521
<i>Increased Buffering</i>	521
<i>Selective Acknowledgment (SACK)</i>	522
<i>Binary Increase Congestion (BIC) TCP</i>	522

Caching and Compression	522
Compression	523
<i>Data Redundancy Elimination (DRE)</i>	523
<i>Unified Data Store</i>	526
<i>Lempel-Ziv (LZ) Compression</i>	527
Object Caching	528
Application-Specific Acceleration	528
Microsoft Exchange Application Optimization	529
HTTP Application Optimization	530
SharePoint Application Optimization	530
SSL Application Optimization	530
Citrix Application Optimization	531
CIFS Application Optimization	532
SMB Application Optimization	533
NFS Acceleration	534
Akamai Connect	534
<i>Transparent Cache</i>	535
<i>Akamai Connected Cache</i>	535
<i>Dynamic URL HTTP Cache (Over-the-Top Cache)</i>	535
<i>Content Prepositioning for Enhanced End-User Experience</i>	535
Summary	536
Further Reading	536

Chapter 12 Cisco Wide Area Application Services (WAAS) 537

Cisco WAAS Architecture	537
Central Management Subsystem	539
Interface Manager	539
Monitoring Facilities and Alarms	539
Network Interception and Bypass Manager	540
Application Traffic Policy Engine	540
Disk Encryption	542
Cisco WAAS Platforms	542
Router-Integrated Network Modules	543
Appliances	543
<i>WAVE Model 294</i>	543
<i>WAVE Model 594</i>	543
<i>WAVE Model 694</i>	546
<i>WAVE Model 7541</i>	546

<i>WAVE Model 7571</i>	546
<i>WAVE Model 8541</i>	546
<i>Interception Modules</i>	547
<i>Virtual WAAS</i>	547
ISR-WAAS	549
<i>Architecture</i>	549
<i>Sizing</i>	550
WAAS Performance and Scalability Metrics	553
WAAS Design and Performance Metrics	553
Device Memory	553
Disk Capacity	554
Number of Optimized TCP Connections	555
WAN Bandwidth and LAN Throughput	556
Number of Peers and Fan-out Each	558
Central Manager Sizing	559
Licensing	560
Cisco WAAS Operational Modes	560
Transparent Mode	561
Directed Mode	561
Interception Techniques and Protocols	561
Web Cache Communication Protocol	562
<i>WCCP Service Groups</i>	562
<i>Forwarding and Return Methods</i>	563
<i>Load Distribution</i>	564
<i>Failure Detection</i>	565
<i>Flow Protection</i>	565
<i>Scalability</i>	565
<i>Redirect Lists</i>	566
<i>Service Group Placement</i>	566
<i>Egress Methods</i>	567
Policy-Based Routing (PBR)	567
Inline Interception	569
AppNav Overview	570
<i>AppNav Cluster Components</i>	572
<i>Class Maps</i>	572
<i>AppNav Policies</i>	573
<i>AppNav Site Versus Application Affinity</i>	573

AppNav IOM	573
<i>AppNav Controller Deployment Models</i>	573
<i>AppNav Controller Interface Modules</i>	574
<i>AppNav IOM Interfaces</i>	575
<i>Guidelines and Limitations</i>	575
AppNav-XE	576
Advantages of Using the AppNav-XE Component	576
Guidelines and Limitations	577
WAAS Interception Network Integration Best Practices	578
Summary	578
Further Reading	579
Chapter 13 Deploying Application Optimizations	581
GBI: Saving WAN Bandwidth and Replicating Data	582
WAN Optimization Solution	583
Deploying Cisco WAAS	584
WAAS Data Center Deployment	584
<i>GBI Data Centers</i>	584
<i>Data Center Device Selection and Placement</i>	585
Primary Central Manager	587
<i>Initial Primary Central Manager Configuration</i>	587
<i>Configuring the Primary Central Manager's NTP Settings</i>	590
<i>Configuring the Primary Central Manager's DNS Settings</i>	590
<i>Configuring WAAS Group Settings</i>	591
<i>Device Group Basic Settings</i>	592
Standby Central Manager	592
<i>Standby Central Manager's Configuration</i>	593
AppNav-XE	595
Initial GBI AppNav-XE Deployment	595
Deploying a Data Center Cluster	600
Deploying a Separate Node Group and Policy for Replication	605
Deploying a New Policy for Data Center Replication	610
GBI Branch Deployment	615
Branch 1 Sizing	615
Branch 1 Deployment	615
Branch 12 Sizing	618
Branch 12 WAAS Deployment	618
Summary	621

Part V QoS

Chapter 14 Intelligent WAN Quality of Service (QoS) 623

- QoS Overview 624
- Ingress QoS NBAR-Based Classification 626
- Ingress LAN Policy Maps 629
- Egress QoS DSCP-Based Classification 630
- Egress QoS Policy Map 631
- Hierarchical QoS 633
- DMVPN Per-Tunnel QoS 640
 - Per-Tunnel QoS Tunnel Markings 641
 - Bandwidth-Based QoS Policies 643
 - Bandwidth Remaining QoS Policies 644
 - Subrate Physical Interface QoS Policies 648
 - Association of Per-Tunnel QoS Policies 649
 - Per-Tunnel QoS Verification 650
 - Per-Tunnel QoS Caveats 658
- QoS and IPSec Packet Replay Protection 660
- Complete QoS Configuration 661
- Summary 669
- Further Reading 669

Part VI Direct Internet Access

Chapter 15 Direct Internet Access (DIA) 671

- Guest Internet Access 673
 - Dynamic Host Configuration Protocol (DHCP) 676
 - Network Address Translation (NAT) 678
 - Verification of NAT 680
 - Zone-Based Firewall (ZBFW) Guest Access 680
 - Verification of ZBFW for Guest Access 684
- Guest Access Quality of Service (QoS) 685
- Guest Access Web-Based Acceptable Use Policy 688
 - Guest Network Consent 688
 - Guest Authentication 692
- Internal User Access 697
- Fully Specified Static Default Route 698
- Verification of Internet Connectivity 699

Network Address Translation (NAT)	704
Policy-Based Routing (PBR)	706
Internal Access Zone-Based Firewall (ZBFW)	708
Cloud Web Security (CWS)	711
Baseline Configuration	712
Outbound Proxy	717
WAAS and WCCP Redirect	720
Prevention of Internal Traffic Leakage to the Internet	720
Summary	721
References in this Chapter	722

Part VII Migration

Chapter 16 Deploying Cisco Intelligent WAN 723

Pre-Migration Tasks	723
Document the Existing WAN	724
Network Traffic Analysis	724
Proof of Concept	724
Finalize the Design	725
Migration Overview	725
IWAN Routing Design Review	726
EIGRP for the IWAN and the LAN	726
BGP for the IWAN and an IGP (OSPF) for the LAN	727
Routing Design During Migration	727
Deploying DMVPN Hub Routers	728
Migrating the Branch Routers	734
Migrating a Single-Router Site with One Transport	735
Migrating a Single-Router Site with Multiple Transports	737
Migrating a Dual-Router Site with Multiple Transports	739
Post-Migration Tasks	740
Migrating from a Dual MPLS to a Hybrid IWAN Model	742
Migrating IPsec Tunnels	744
PfR Deployment	746
Testing the Migration Plan	752
Summary	752
Further Reading	753

Part VIII Conclusion**Chapter 17 Conclusion and Looking Forward 755**

Intelligent WAN Today 755

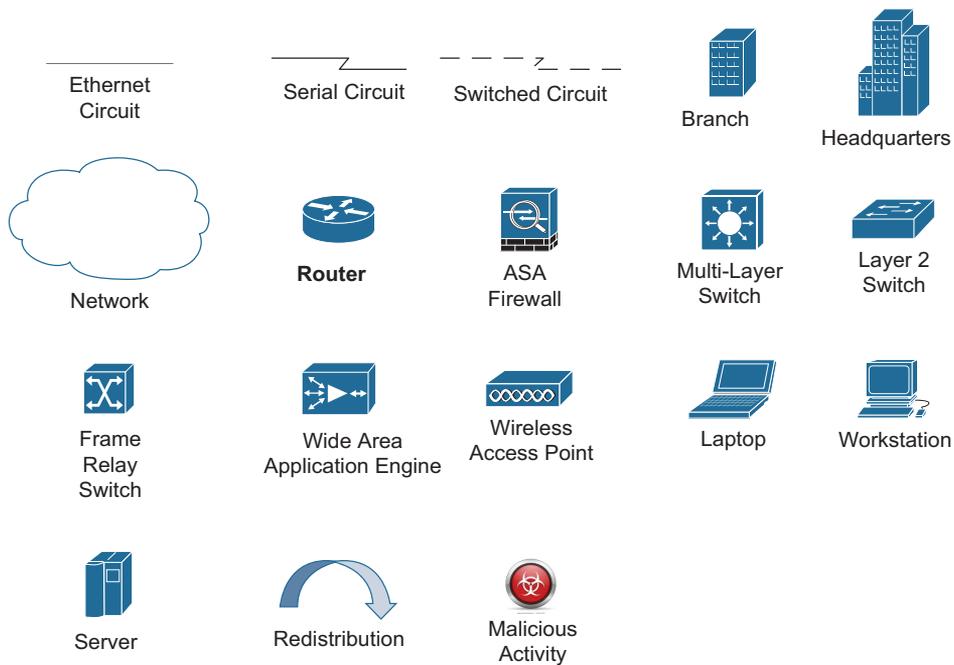
Intelligent WAN Architecture 756

Intelligent WAN Tomorrow 756

Appendix A Dynamic Multipoint VPN Redundancy Models 759**Appendix B IPv6 Dynamic Multipoint VPN 763**

Index 779

Icons Used in This Book



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the IOS Command Reference. The Command Reference describes these conventions as follows:

- **Boldface** indicates commands and keywords that are entered literally as shown. In actual configuration examples and output (not general command syntax), boldface indicates commands that are manually input by the user (such as a **show** command).
- *Italic* indicates arguments for which you supply actual values.
- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets ([]) indicate an optional element.
- Braces ({ }) indicate a required choice.
- Braces within brackets ([{ }]) indicate a required choice within an optional element.

Foreword

The world is changing fast. And demands on the network are growing exponentially. More than ever before, businesses need technology to provide speed, flexibility, and information in a cost-effective manner across their systems and processes. The Cisco Intelligent WAN (IWAN) helps companies in any market segment connect the lifeblood of their organization—their branch locations—to business value located anywhere on the network. Whether your branch is a retail store, a healthcare clinic, or a remote office, branches are a critical component of business. These are the places where organizations interface with customers and other citizens, where most business intelligence is acquired, and where the bulk of employees work. It's crucial that the branch play a large role in any organization's plans for digitization and value.

As the leader of the Cisco Systems Engineering team, I have the privilege of working with the best networking professionals in the industry. Working on the front lines of the customer relationship, our SE teams are uniquely positioned to provide feedback from our vast customer base back to the Cisco innovation engine, our development team. Cisco has thousands of systems engineers globally working with our customers every day, and they gain great insights into the top issues facing IT and our customers' businesses in general. The feedback collected from our customers and Systems Engineering team led to the development of IWAN.

In many traditional WAN implementations, customers, vendors, suppliers, and employees who are located in the branch often cannot receive optimal service, and their capabilities are limited. The Cisco IWAN allows IT to remove those limitations by enabling intelligence on the WAN. With IWAN's ability to simplify VPNs and allow more control, applications such as guest Internet traffic, public cloud services, and partner cloud applications can be offloaded immediately with the appropriate quality of service levels. And with visibility to the application level, applications that are dependent upon data center connectivity can perform better. Last, given the need for all these use cases to be secure, you will see the value of IWAN in providing secure connectivity for your applications while providing better service and improved performance.

This book was written by an all-star team, including Brad Edgeworth, one of the key leaders in our Systems Engineering organization. Holding multiple CCIE certifications, this team of contributing authors present at both internal and external events, which means they can explain the technology and how it helps businesses. Their depth of experience and knowledge is demonstrated in this book as they address IWAN, its features, benefits, and implementation, and provide readers insight into the top issues facing IT: security, flexibility, application visibility, and ease of use. These are the most important issues facing the WAN and IT in general.

The Cisco IWAN solution helps businesses achieve their goals, and this book will help IT departments get the most out of these solutions. The book describes IWAN and its implementation in an easy-to-understand format that will allow network professionals to take full advantage of this solution in their environments. In doing so, it will allow those IT professionals to deliver tremendous business value to their organizations. At Cisco, we believe that technology can truly help businesses define their strategy and value in the market. And we believe that IT can help deliver that value through speed, agility, and responsiveness to their customers and their businesses.

Michael Koons

VP Systems Engineering and Technology,

Cisco Systems

Introduction

The Cisco Intelligent WAN (IWAN) enables organization to deliver an uncompromised experience over any WAN transport. With the Cisco IWAN architecture, organizations can provide more bandwidth to their branch office connections using cost-effective WAN transports without affecting performance, security, or reliability.

The authors' goal was to provide a multifunction self-study book that explains the technologies used in the IWAN architecture that would allow the reader to successfully deploy the technology. Concepts are explained in a modular structure so that the reader can learn the logic and configuration associated with a specific feature. The authors provide real-world use cases that will influence the design of your IWAN network.

Knowledge learned from this book can be used for deploying IWAN via CLI or other Cisco management tools such as Cisco Prime Infrastructure or Application Policy Infrastructure Controller Enterprise Module (APIC-EM).

Who Should Read This Book?

This book is for network engineers, architects, and consultants who want to learn more about WAN networks and the Cisco IWAN architecture and the technical components that increase the effectiveness of the WAN. Readers should have a fundamental understanding of IP routing.

How This Book Is Organized

Although this book can be read cover to cover, it is designed to be flexible and allow you to easily move between chapters and sections of chapters so that you can focus on just the material that you need.

Part I of the book provides an overview of the evolution of the WAN.

- **Chapter 1, “Evolution of the WAN”:** This chapter explains the reasons for increased demand on the WAN and why the WAN has become more critical to businesses in any market vertical. The chapter provides an introduction to Cisco Intelligent WAN (IWAN) and how it enhances user experiences while lowering operational costs.

Part II of the book explains transport independence through the deployment of Dynamic Multipoint VPN (DMVPN).

- **Chapter 2, “Transport Independence”:** This chapter explains the history of WAN technologies and the current technologies available to network architects. Dynamic Multipoint VPN (DMVPN) is explained along with the benefits that it provides over other VPN technologies.
- **Chapter 3, “Dynamic Multipoint VPN”:** This chapter explains the basic concepts of DMVPN and walks the user from a simple topology to a dual-hub, dual-cloud topology. The chapter explains the interaction that NHRP has with DMVPN because that is a vital component of the routing architecture.

- **Chapter 4, “Intelligent WAN (IWAN) Routing”:** This chapter explains why EIGRP and BGP are selected for the IWAN routing protocols and how to configure them. In addition to explaining the logic for the routing protocol configuration, multicast routing is explained.
- **Chapter 5, “Securing DMVPN Tunnels and Routers”:** This chapter examines the vulnerabilities of a network and the steps that can be taken to secure the WAN. It explains IPsec DMVPN tunnel protection using pre-shared keys and PKI infrastructure. In addition, the hardening of the router is performed through the deployment of Zone-Based Firewall (ZBFW) and Control Plane Policing (CoPP).

Part III of the book explains how to deploy intelligent routing in the WAN.

- **Chapter 6, “Application Recognition”:** This chapter examines how an application can be identified through the use of traditional ports and through deep packet inspection. Application classification is essential for proper QoS policies and intelligent routing policies.
- **Chapter 7, “Introduction to Performance Routing (PfR)”:** This chapter discusses the need for intelligent routing and a brief evolution of Cisco Performance Routing (PfR). The chapter also explains vital concepts involving master controllers (MCs) and border routers (BRs) and how they operate in PfR version 3.
- **Chapter 8, “PfR Provisioning”:** This chapter explains how PfRv3 can be configured and deployed in a topology.
- **Chapter 9, “PfR Monitoring”:** This chapter explains how PfR can be examined to verify that it is operating optimally.
- **Chapter 10, “Application Visibility”:** This chapter discusses how PfR can view and collect application performance on the WAN.

Part IV of the book discusses and explains how application optimization integrates into the IWAN architecture.

- **Chapter 11, “Introduction to Application Optimization”:** This chapter covers the fundamentals of application optimization and how it can accelerate application responsiveness while reducing demand on the current WAN.
- **Chapter 12, “Cisco Wide Area Application Services (WAAS)”:** This chapter explains the Cisco WAAS architecture and methods that it can be inserted into a network. In addition, it explains how the environment can be sized appropriately for current and future capacity.
- **Chapter 13, “Deploying Application Optimizations”:** This chapter explains how the various components of WAAS can be configured for the IWAN architecture.

Part V of the book explains the specific aspects of QoS for the WAN.

- **Chapter 14, “Intelligent WAN Quality of Service (QoS)”:** This chapter explains NBAR-based QoS policies, Per-Tunnel QoS policy, and other changes that should be made to accommodate the IWAN architecture.

Part VI of the book discusses direct Internet access and how it can reduce operational costs while maintaining a consistent security policy.

- **Chapter 15, “Direct Internet Access (DIA)”**: This chapter explains how direct Internet access can save operational costs while providing additional services at branch sites. The chapter explains how ZBFW or Cisco Cloud Web Security can be deployed to provide a consistent security policy to branch network users.

Part VII of the book explains how IWAN can be deployed.

- **Chapter 16, “Deploying Cisco Intelligent WAN”**: This chapter provides an overview of the steps needed to successfully migrate an existing WAN to Cisco Intelligent WAN.

The book ends with a closing perspective on the future of the Cisco software-defined WAN (SD-WAN) and the management tools that are being released by Cisco.

Learning in a Lab Environment

This book contains new features and concepts that should be tested in a lab environment first. Cisco VIRL (Virtual Internet Routing Lab) provides a scalable, extensible network design and simulation environment that includes several Cisco Network Operating System virtual machines (IOSv, IOS-XRv, CSR 1000V, NX-OSv, IOSvL2, and ASA v) and has the ability to integrate with third-party vendor virtual machines or external network devices.

The authors will be releasing a VIRL topology file so that readers can learn the technologies as they are explained in the book. More information about VIRL can be found at <http://virl.cisco.com>.

Additional Reading

The authors tried to keep the size of the book manageable while providing only necessary information about the topics involved. Readers who require additional reference material may find the following books to be a great supplementary resource for the topics in this book:

- Bollapragada, Vijay, Mohamed Khalid, and Scott Wainner. *IPSec VPN Design*. Indianapolis: Cisco Press, 2005. Print.
- Edgeworth, Brad, Aaron Foss, and Ramiro Garza Rios. *IP Routing on Cisco IOS, IOS XE, and IOS XR*. Indianapolis: Cisco Press, 2014. Print.
- Karamanian, Andre, Srinivas Tenneti, and Francois Dessart. *PKI Uncovered: Certificate-Based Security Solutions for Next-Generation Networks*. Indianapolis: Cisco Press, 2011. Print.
- Seils, Zach, Joel Christner, and Nancy Jin. *Deploying Cisco Wide Area Application Services*. Indianapolis: Cisco Press, 2008. Print.
- Szigeti, Tim, Robert Barton, Christina Hattingh, and Kenneth Briley Jr. *End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks, Second Edition*. Indianapolis: Cisco Press, 2013. Print.

Introduction to Performance Routing (PfR)

This chapter covers the following topics:

- Performance Routing (PfR)
- Introduction to the IWAN domain
- Intelligent path control principles

Bandwidth cost, WAN latency, and lack of bandwidth availability all contribute to the complexities of running an efficient and cost-effective network that meets the unique, application-heavy workloads of today's enterprise organizations. But as the volume of content and applications traveling across the network grows exponentially, organizations must optimize their WAN investments.

Cisco *Performance Routing (PfR)* is the IWAN intelligent path control component that can help administrators to accomplish the following:

- Augment the WAN with additional bandwidth to including lower-cost connectivity options such as the Internet
- Realize the cost benefits of provider flexibility and the ability to choose different transport technologies (such as MPLS L3VPN, VPLS, or the Internet)
- Offload the corporate WAN with highly secure direct Internet access
- Improve application performance and availability based upon an application's performance requirements
- Protect critical applications from fluctuating WAN performance

Performance Routing (PfR)

Cisco Performance Routing (PfR) improves application delivery and WAN efficiency. PfR dynamically controls data packet forwarding decisions by looking at application type, performance, policies, and path status. PfR protects business applications from fluctuating WAN performance while intelligently load-balancing traffic over the best-performing path based on the application policy.

Simplified Routing over a Transport-Independent Design

One of the critical IWAN components and also a key design decision was to architect the next-generation WAN around a *transport-independent design (TID)*. The choice of DMVPN was extensively explained in Chapter 2, “Transport Independence.” This overlay approach allows the use of a single routing protocol over the WAN and greatly simplifies the routing decision process and Performance Routing in multiple ways, two of the main ones being

- Simplified reachability information
- Single routing domain

The first benefit of this overlay approach is simplified reachability information.

The traditional routing protocols were designed to solve the endpoint reachability problem in a hop-by-hop destination-only forwarding environment of unknown topology. The routing protocols choose only the best path based on statically assigned cost. There are a few exceptions where the network path used can be somewhat engineered. Some routing protocols can select a path that is not the shortest one (*BGP, MPLS traffic engineering [TE]*).

Designing deterministic routing behavior is difficult with multiple transport providers but is much simpler thanks to DMVPN. The DMVPN network topology is flat, and it is consistent because it is an overlay network that masks the network complexity underneath. This approach simplifies the logical view of the network and minimizes fundamental topology changes. Logically, only reachability to the next hop across the WAN can change.

An overlay network’s routing information is very simple: a set of destination prefixes, and a set of potential transport next hops for each destination. As a result, PfR just needs a *mapping service* that stores and serves all resolved forwarding states for connectivity per overlay network. Each forwarding state contains destination prefix, next hop (overlay IP address), and corresponding transport address.

The second benefit of using overlay networks is the single routing domain design. In traditional hybrid designs, it is common to have two (or more) routing domains:

- One routing domain for the primary path over MPLS—EBGP, static, or default routes
- One routing domain on the secondary path over the Internet—EIGRP, IBGP, or floating static routes

The complexity increases when routes are exchanged between the multiple routing domains, which can lead to suboptimal routing or routing loops. Using DMVPN for all WAN transports allows the use of a single routing protocol for all paths regardless of the transport choice. Whether the topology is dual hybrid (MPLS plus Internet) or dual Internet (two Internet paths), the routing configuration remains exactly the same, meaning that if there is a change in how your provider chooses to deliver connectivity, or you wish to add or change a provider underneath the DMVPN, the investment in your WAN routing architecture is secure.

EIGRP and IBGP are the best routing protocol options today with DMVPN.

After routing connectivity is established, PfR enters the picture and provides the advanced path control in IWAN. PfR is not a replacement for the routing protocol and never will be. As an adjunct, PfR uses the next-hop information from the routing protocol and overrides it based on real-time performance and link utilization ratio. This next-hop information per destination prefix is critical for PfR to work correctly and is a critical element in the routing design. Having a single routing domain and a very basic mapping service requirement has greatly simplified PfR interaction with the routing protocol.

“Classic” Path Control Used in Routing Protocols

Path control, commonly referred to as “traffic engineering,” is the process of choosing the network path on which traffic is sent. The simplest form is trivial: send all traffic down the primary path unless the path goes down; in that case, send everything through the backup path.

Figure 7-1 illustrates the concept where R31 (branch) sends traffic to R11 (headquarters). When R31’s link to the MPLS provider fails, traffic is sent through the Internet.

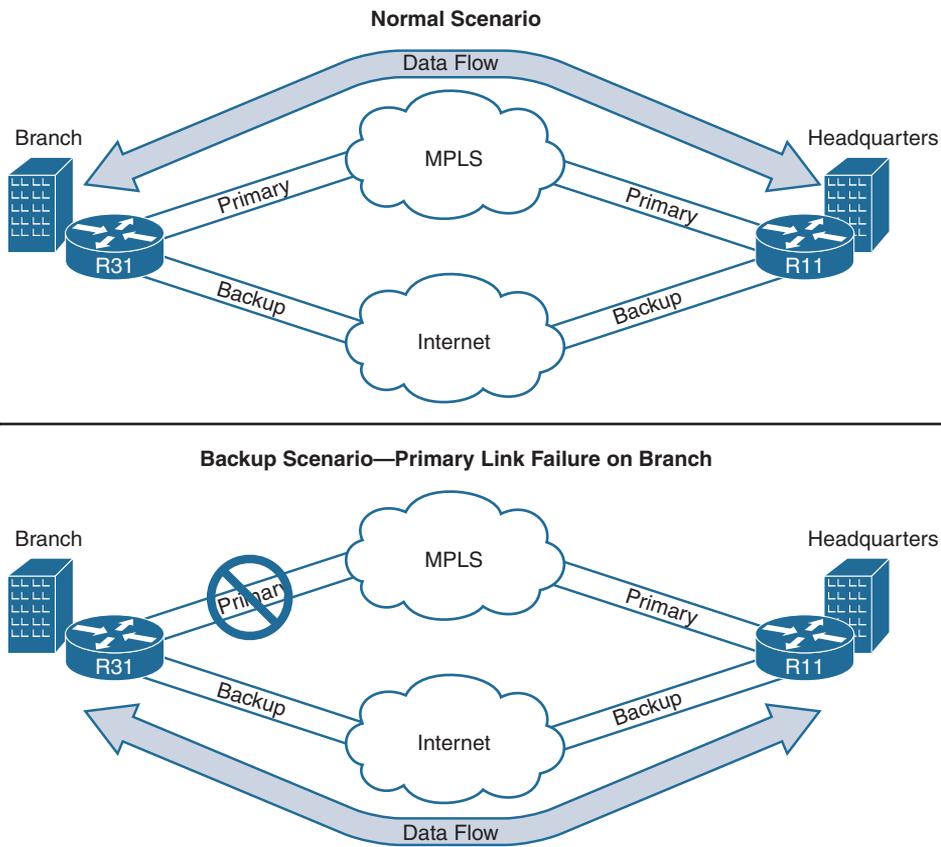


Figure 7-1 Traffic Flow over Primary and Backup Links

This approach has two main drawbacks:

- Traffic is forwarded over a single path regardless of the application type, performance, or bandwidth issues.
- The backup path is used only when the primary link goes down and not when there is performance degradation or brownouts over the primary path because the routing protocol peers are usually still up and running and do not detect such performance issues.

Path Control with Policy-Based Routing

The next level of path control lets the administrator specify categories of traffic to send on a specific path as long as that path remains up. One of the most common options is the use of *policy-based routing (PBR)*, routing based on DSCP values:

- DSCP values that are mapped to critical business applications and voice/video types of applications are assigned a next hop that is over the preferred path.
- DSCP values that are mapped to best-effort applications or applications that do not suffer from performance degradation are assigned a next hop over the secondary path.

However, this approach is not intelligent and does not take into account the dynamic behavior of the network. Routing protocols have keepalive timers that can determine if the next hop is available, but they cannot determine when the path selected suffers from degraded performance, and the system cannot compensate.

Figure 7-2 illustrates the situation where R31 (branch) sends traffic to R11 (headquarters). When R31's path across the MPLS provider experiences performance issues, traffic continues to be sent through the MPLS backbone. PBR alone is unaware of any performance problems. An additional mechanism is needed to detect events like these, such as the use of IP SLA probes.

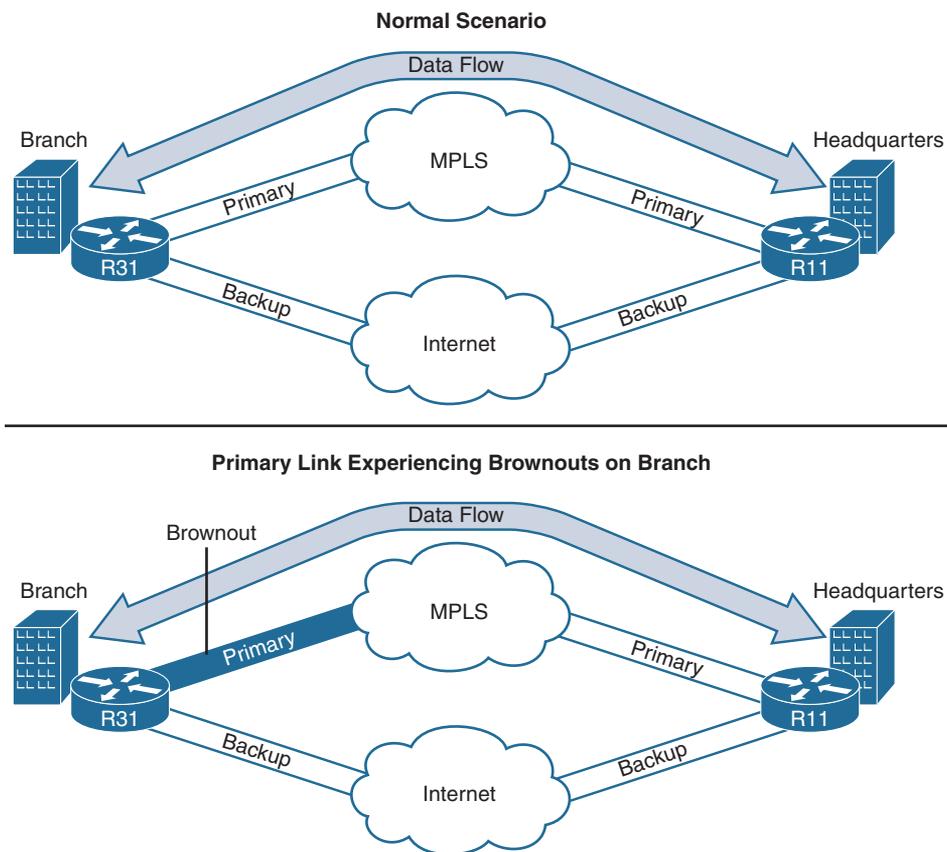


Figure 7-2 PBR's Inability to Detect Problematic Links

Intelligent Path Control—Performance Routing

Classic routing protocols or path control with PBR cannot detect performance issues and fall back affected traffic to an alternative path. Intelligent path control solves this problem by monitoring actual application performance on the path that the applications are traversing, and by directing traffic to the appropriate path based on these real-time performance measurements.

When the current path experiences performance degradation, Cisco intelligent path control moves the affected flows according to user-defined policies.

Figure 7-3 illustrates the situation where R31 sends traffic to R11. When R31's path across the MPLS provider experiences performance issues, only affected traffic is sent to the Internet path. The choice of traffic to fall back is based on defined policies. For example, voice or business application flows are forwarded over the secondary path, whereas best-effort traffic remains on the MPLS path.

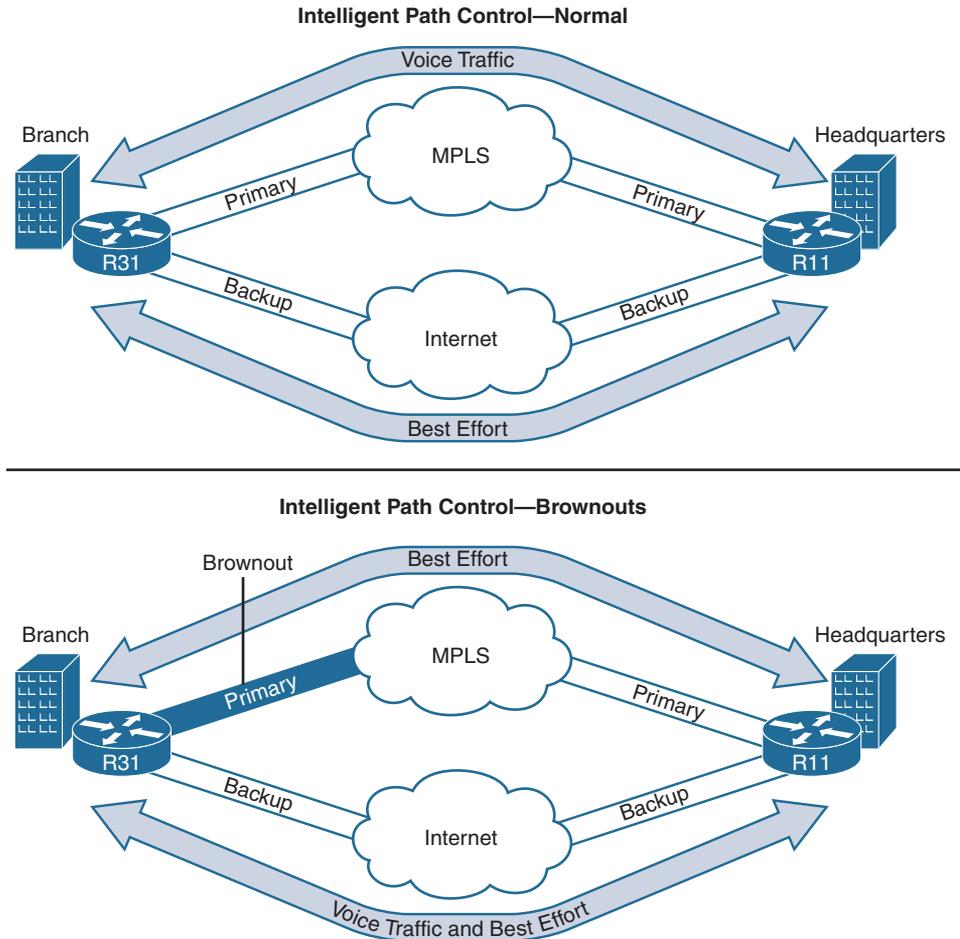


Figure 7-3 Traffic Flow over Multiple Links with Cisco Intelligent Path Control

Advanced path control should include the following:

- Detection of issues such as delay, loss, jitter, and defined path preference before the associated application is impacted.
- Passive performance measurement based on real user traffic when available and passively monitored on existing WAN edge routers. This helps support SLAs to protect critical traffic.
- Efficient load distribution across the WAN links for medium-priority and best-effort traffic.
- Effective reaction to any network outages before they can affect users or other aspects of the network. These include *blackouts* that cause a complete loss of connectivity as well as *brownouts* that are network slowdowns caused by path degradation along the route to the destination. Although blackouts can be detected easily, brownouts are much more challenging to track and are usually responsible for bad user experience.
- Application-based policies that are designed to support the specific performance needs of applications (for example, point of sale, enterprise resource planning [ERP], and so on).
- Low WAN overhead to ensure that control traffic is not contributing to overall traffic issues.
- Easy management options, including a single point of administration and the ability to scale without a stacked deployment.

Cisco Performance Routing (PfR), part of Cisco IOS software, provides intelligent path control in IWAN and complements traditional routing technologies by using the intelligence of a Cisco IOS infrastructure to improve application performance and availability.

As explained before, PfR is not a replacement for the routing protocols but instead runs alongside of them to glean the next hop per destination prefix. PfR has APIs with NHRP, BGP, EIGRP, and the routing table to request information. It can monitor and then modify the path selected for each application based on advanced criteria, such as reachability, delay, loss, and jitter. PfR intelligently load-balances the remainder of the traffic among available paths based on the tunnel bandwidth utilization ratio.

Note The routing table, known as the *routing information base (RIB)*, is built from dynamic routing protocols and static and directly connected routes. The routing table is referred to as the RIB throughout the rest of this chapter.

Cisco PfR has evolved and improved over several releases with a focus on simplicity, ease of deployment, and scalability. Table 7-1 provides a list of features that have evolved with each version of PfR.

Table 7-1 *Evolution of PfR Versions and Features*

Version	Features
PfR/Optimized Edge Routing (OER)	Internet edge Basic WAN Provisioning per site per policy Thousands of lines of configuration
PfRv2	Policy simplification App path selection Scale 500 sites Tens of lines of configuration
PfRv3	Centralized provisioning Application Visibility Control (AVC) infrastructure VRF awareness Scale 2000 sites Hub configuration only Multiple data centers Multiple next hops per DMVPN network

Introduction to PfRv3

Performance Routing Version 3 (PfRv3) is the latest generation of the original PfR created more than ten years ago. PfRv3 focuses on ease of use and scalability to make it easy to transition to an intelligent network with PfR. It uses one-touch provisioning with multisite coordination to simplify its configuration and deployment from previous versions of PfR. PfRv3 is a DSCP- and application-based policy-driven framework that provides multisite path control optimization and is bandwidth aware for WAN- and cloud-based applications. PfRv3 is tightly integrated with existing AVC components such as Performance Monitor, QoS, and NBAR2.

PfR is composed of devices performing several roles, which are *master controller (MC)* and *border router (BR)*. The MC serves as the control plane of PfR, and the BR is the forwarding plane which selects the path based on MC decisions.

Note The MC and BR are components of the IOS software features on WAN routers.

Figure 7-4 illustrates the mechanics of PfRv3. Traffic policies are defined based on DSCP values or application names. Policies can state requirements and preferences for applications and path selection. A sample policy can state that voice traffic uses preferred path MPLS unless delay is above 200 ms. PfR learns the traffic, then starts measuring the bandwidth and performance characteristics. Then the MC makes a decision by comparing the real-time metrics with the policies and instructs the BRs to use the appropriate path.

Note The BRs automatically build a tunnel (known as an *auto-tunnel*) between other BRs at a site. If the MC instructs a BR to redirect traffic to a different BR, traffic is forwarded across the auto-tunnel to reach the other BR.

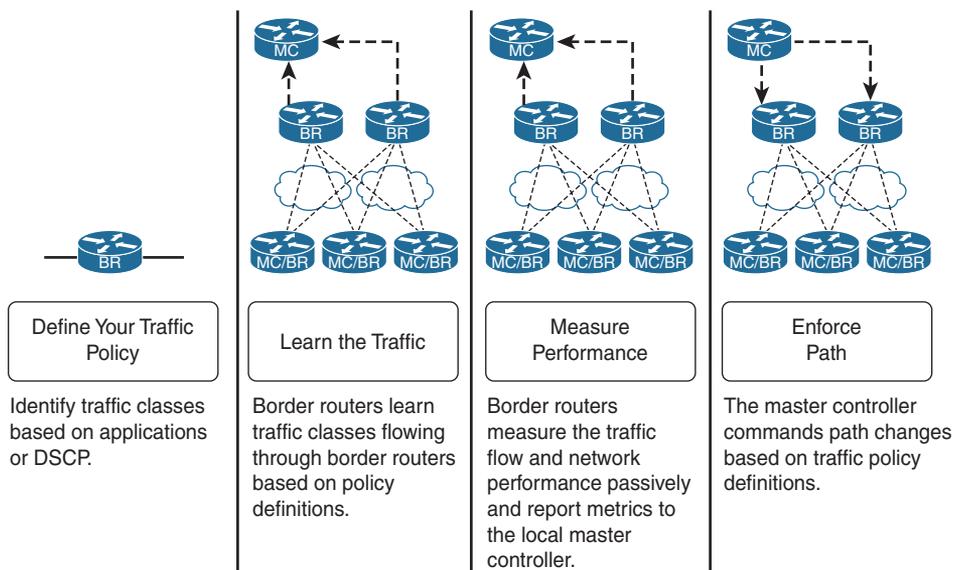


Figure 7-4 *Mechanics of PfRv3*

Note The first iteration of PfRv3 was introduced in summer 2014 with IOS 15.4(3)M and IOS XE 3.13.

Introduction to the IWAN Domain

An IWAN domain is a collection of sites that share the same set of policies and are managed by the same logical PfR *domain controller*. Each site runs PfR and gets its path control configuration and policies from the logical IWAN domain controller through

the IWAN peering service. At each site, an MC is the local decision maker and controls the BRs responsible for performance measurement and path enforcement. The IWAN domain can be an entire enterprise WAN, a particular region, and so forth.

The key point for PfRv3 is that provisioning is fully centralized at a logical domain controller, whereas path control decisions and enforcement are fully distributed within the sites that are in the domain, making the solution very scalable.

Figure 7-5 shows a typical IWAN domain with central and branch sites. R10, R20, R31, R41, and R51 are all MCs for their corresponding sites and retrieve their configuration from the logical domain controller. R11, R12, R21, R22, R31, R41, R51, and R52 are all BRs that report back to their local MC. Notice that R31, R41, and R51 operate as both the MC and the BR for their sites.

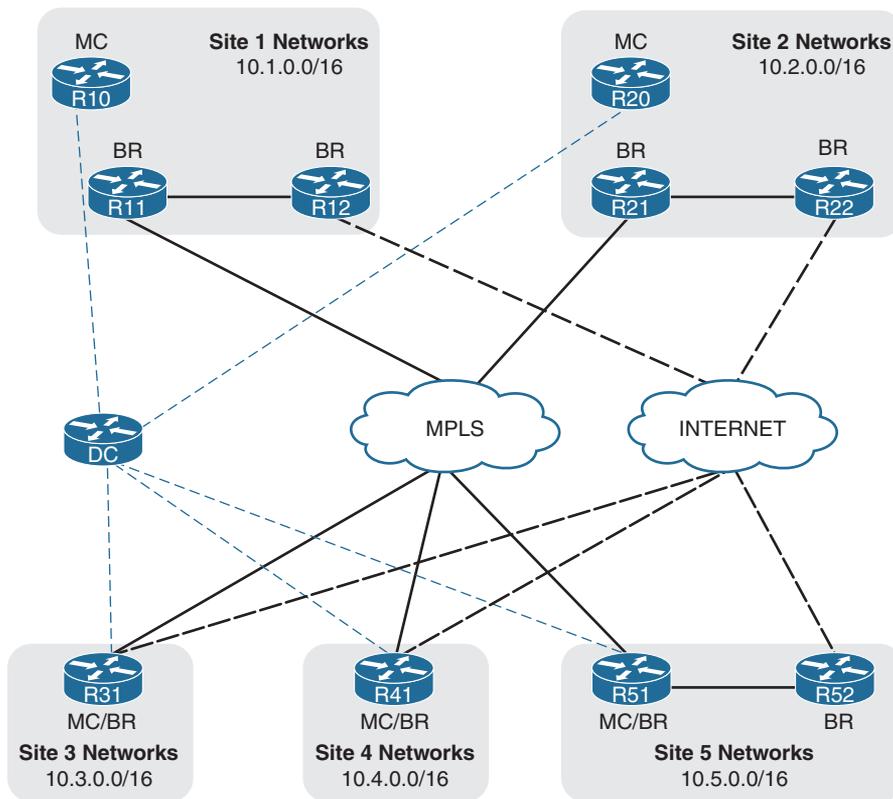


Figure 7-5 *IWAN Domain Concepts*

Note In the remainder of this book, all references to PfR mean PfRv3.

IWAN Sites

An IWAN domain includes a mandatory hub site, optional transit sites, as well as branch sites. Each site has a unique identifier called a *site ID* that is derived from the loopback address of the local MC.

Central and headquarters sites play a significant role in PfR and are called IWAN *Points of Presence (POPs)*. Each site has a unique identifier called a POP ID. These sites house DMVPN hub routers and therefore provide the following traffic flows (streams):

- Traditional DMVPN spoke-to-hub connectivity.
- Spoke-to-hub-to-spoke connectivity until DMVPN spoke-to-spoke tunnels establish.
- Connectivity through NHS chaining until DMVPN spoke-to-spoke tunnels establish.
- Transit connectivity to another site via a data center interconnect (DCI) or shared data center network segment. In essence, these sites act as *transit sites* for the traffic crossing them. Imagine in Figure 7-5 that R31 goes through R21 to reach a network that resides in Site 1. R21 does not terminate the traffic at the local site; it provides transit connectivity to Site 1 via the DCI.
- Data centers may or may not be colocated with the hub site. To elaborate further, some hub sites contain data centers whereas other hub sites do not contain data centers (such as outsourced colocation cages).

Hub site

- The logical domain controller functions reside on this site's MC.
- Only one hub site exists per IWAN domain because of the uniqueness of the logical domain controller's presence. The MC for this site is known as the Hub MC, thereby making this site the hub site.
- MCs from all other sites (transit or branch) connect to the Hub MC for PfR configuration and policies.
- A POP ID of 0 is automatically assigned to a hub site.
- A hub site may contain all other properties of a transit site as defined below.

Transit sites

- Transit sites are located in an enterprise central site, headquarters, or carrier-neutral facilities.
- They provide transit connectivity to access servers in the data centers or for spoke-to-spoke traffic.
- A data center may or may not be colocated with the transit site. A data center can be reached via a transit site.

- A POP ID is configured for each transit site. This POP ID has to be unique in the domain.
- The local MC (known as a *Transit MC*) peers with the Hub MC (domain controller) to get its policies and to monitor configuration and timers.

Branch sites

- These are always DMVPN spokes and are stub sites where traffic transit is not allowed.
- The local Branch MC peers with the logical domain controller (Hub MC) to get its policies and monitoring guidelines.

Figure 7-6 shows the IWAN sites in a domain with two central sites (one is defined as the hub site and the other as a transit site). R10, R11, and R12 belong to the hub site, and R20, R21, and R22 belong to a transit site. R31, R41, R51, and R52 belong to a branch site. The dotted lines represent the site's local MC peering with the Hub MC.

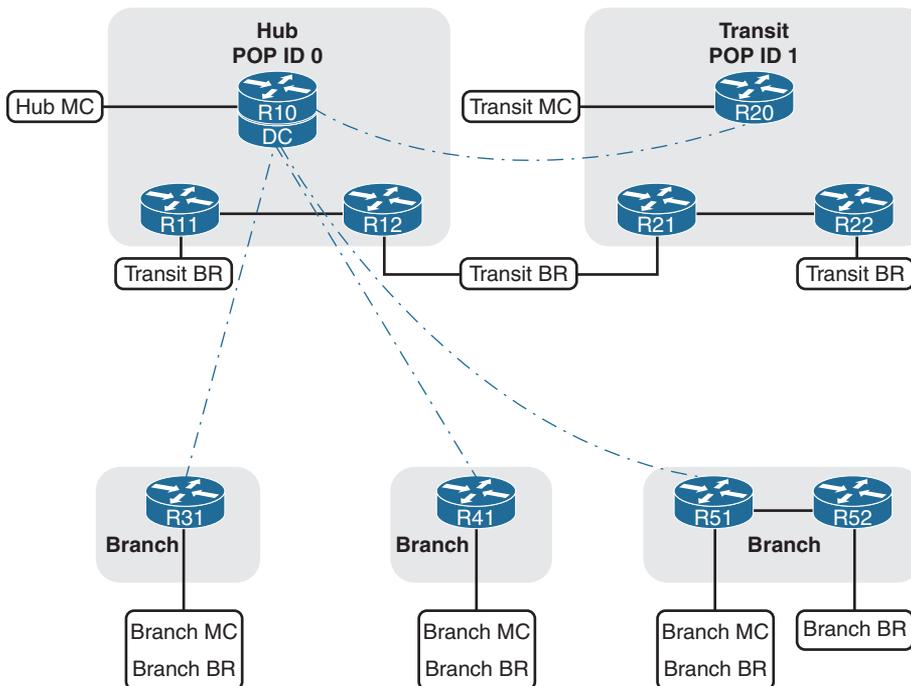


Figure 7-6 IWAN Domain Hub and Transit Sites

Device Components and Roles

The PfR architecture consists of two major Cisco IOS components, a master controller (MC) and a border router (BR). The MC is a policy decision point where policies are defined and applied to various traffic classes (TCs) that traverse the BR systems. The MC can be configured to learn and control TCs on the network:

- **Border routers (BRs)** are in the data-forwarding path. BRs collect data from their Performance Monitor cache and smart probe results, provide a degree of aggregation of this information, and influence the packet forwarding path as directed by the site local MC to manage router traffic.
- **The master controller (MC)** is the policy decision maker. At a large site, such as a data center or campus, the MC is a dedicated (physical or logical) router. For smaller branch locations, the MC is typically colocated (configured) on the same platform as the BR. As a general rule, large locations manage more network prefixes and applications than a branch deployment, thus consuming more CPU and memory resources for the MC function. Therefore, it is a good design practice to dedicate a chassis for the MC at large sites.

Each site in the PfR domain must include a local MC and at least one BR.

The branch typically manages fewer active network prefixes and applications. Because of the costs associated with dedicating a chassis at each branch, the network manager can colocate the local MC and BR on the same router platform. CPU and memory utilization should be monitored on platforms operating as MCs, and if utilization is high, the network manager should consider an MC platform with a higher-capacity CPU and memory. The local MC communicates with BRs and the Hub MC over an authenticated TCP socket but has no requirement for populating its own IP routing table with anything more than a route to reach the Hub MC and local BRs.

PfR is an intelligent path selection technology and requires

- At least two external interfaces under the control of PfR
- At least one internal interface under the control of PfR
- At least one configured BR
 - If only one BR is configured, both external interfaces are attached to the single BR.
 - If more than one BR is configured, two or more external interfaces are configured across these BRs.

The BR, therefore, owns external links, or exit points; they may be logical (tunnel interfaces) or physical links (serial, Ethernet, and so on). With the IWAN prescriptive design, external interfaces are always logical DMVPN tunnels.

A device can fill five different roles in an IWAN domain:

- **Hub MC:** This is the MC at the hub site. It acts as MC for the site, makes optimization decisions for that site, and provides the path control policies for all the other MCs. The Hub MC contains the logical PfR domain controller role.
- **Transit MC:** This is the MC at a transit site that makes optimization decision for those sites. There is no policy configuration on Transit MCs because they receive their policies from the Hub MC.
- **Branch MC:** The Branch MC is the MC for branch sites that makes optimization decisions for those sites. There is no policy configuration on Branch MCs because they receive their policies from the Hub MC.
- **Transit BR:** The Transit BR is the BR at a hub or transit site. The WAN interface terminates in the BRs. PfR is enabled on these interfaces. At the time of this writing, only one WAN interface is supported on a Transit BR. This limitation is overcome by using multiple BR devices.

Note Some Cisco documentation may refer to a Transit BR as a Hub BR, but the two function identically because transit site capabilities were included in a later release of PfR.

- **Branch BR:** The Branch BR resides at the branch site and forwards traffic based on the decisions of the Branch MC. The only PfR configuration is the identification of the Branch MC and setting its role as a BR. The WAN interface that terminates on the device is detected automatically.

The PfR Hub MC is currently supported only on the IOS and IOS XE operating systems.

IWAN Peering

PfR uses an IWAN peering service between the MCs and BRs which is based on a publish/subscribe architecture. The current IWAN peering service uses Cisco SAF to distribute information between sites, including but not limited to

- Learned site prefix
- PfR policies
- Performance Monitor information

The IWAN peering service provides an environment for service advertisement and discovery in a network. It is made up of two primary elements: client and forwarder.

- An IWAN peering service client is a producer (advertises to the network), a consumer of services (requests a service from the network), or both.
- An IWAN peering service SAF forwarder receives services advertised by clients, distributes the services reliably through the network, and makes services available to clients.

- An IWAN peering service client needs to send a register message to a forwarder before it is able to advertise (publish) or request (subscribe to) services.

The IWAN peering service also adopts a logical unicast topology to implement the peering system. Each instance that joins the IWAN peering service serves as both a client and a forwarder:

- The Hub MC listens for unicast packets for advertisements or publications from Transit MCs, Branch MCs, and local BRs.
- The Transit MC peers with the Hub MC and listens to its local BRs.
- The Branch MC peers with the Hub MC and listens to its local BRs.
- BRs always peer with their local MC.

Figure 7-7 illustrates the IWAN peering service with the policies advertised from the Hub MC, the advertisement of monitors, and the exchange of site prefixes.

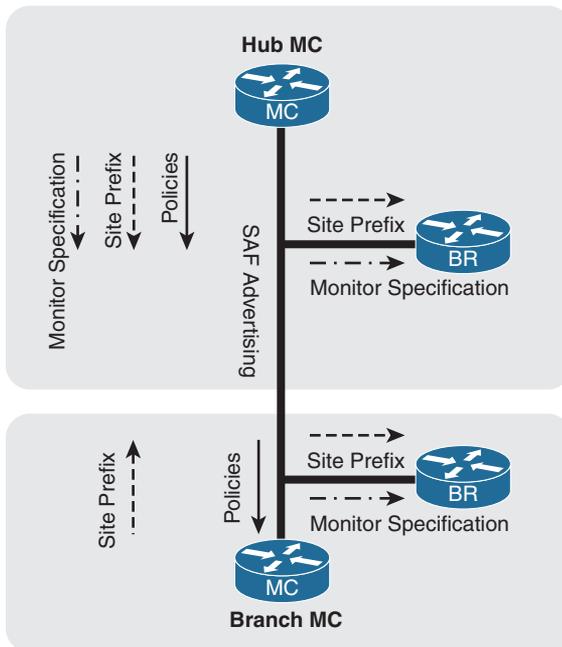


Figure 7-7 IWAN SAF Peering Service

SAF is automatically configured when PfR is enabled on a site. SAF dynamically discovers and establishes a peering as defined previously. The Hub MC advertises all policies and monitoring configuration to all the sites. Every site is responsible for advertising its own site prefix information to other sites in the domain.

Each instance must use an interface with an IP address that is reachable (routed) through the network to join in the IWAN peering system. PfRv3 requires that this address be a loopback address. It is critical that all these loopback addresses be reachable across the IWAN domain.

Parent Route Lookups

PfR uses the concept of a *parent route lookup* which refers to locating all the paths that a packet can take to a specific network destination regardless of the best-path calculation. The parent route lookup is performed so that PfR can monitor all paths and thereby prevent network traffic from being blackholed because the BRs have only summary routes in their routing table. PfR has direct API accessibility into EIGRP and BGP and can identify all the paths available for a prefix regardless of whether alternative paths were installed into the RIB.

PfR requires a parent route for every WAN path (primary, secondary, and so on) for PfR to work effectively. PfR searches the following locations in the order listed to locate all the paths for a destination:

1. NHRP cache (when spoke-to-spoke direct tunnels are established)
2. BGP table (where applicable)
3. EIGRP topology table (where applicable)
4. Static routes (where applicable)
5. RIB. Only one path is selected by default. In order for multiple paths to be selected, the same routing protocol must find both paths to be equal. This is known as equal-cost multipathing (ECMP).

Note If a protocol other than EIGRP or BGP is used, all the paths have to be ECMP in the RIB. Without ECMP in the RIB, PfR cannot identify alternative paths, and that hinders PfR's effectiveness.

The following logic is used for parent route lookups:

- The parent route lookup is done during channel creation (see the following section, "Intelligent Path Control Principles," for more information).
- For PfR Internet-bound traffic, the parent route lookup is done every time traffic is controlled.

In a typical IWAN design, BGP or EIGRP is configured to make sure MPLS is the preferred path and the Internet the backup path. Therefore, for any destination prefix, MPLS is the only available path in the RIB. But PfR looks into the BGP or EIGRP table

and knows if the Internet is also a possible path and can use it for traffic forwarding in a loop-free manner.

Intelligent Path Control Principles

PfR is able to provide intelligent path control and visibility into applications by integrating with the Cisco Performance Monitoring Agent available on the WAN edge (BR) routers. Performance metrics are passively collected based on user traffic and include bandwidth, one-way delay, jitter, and loss.

PfR Policies

PfR policies are global to the IWAN domain and are configured on the Hub MC, then distributed to all MCs via the IWAN peering system. Policies can be defined per DSCP or per application name.

Branch and Transit MCs also receive the Cisco Performance Monitor instance definition, and they can instruct the local BRs to configure Performance Monitors over the WAN interfaces with the appropriate thresholds.

PfR policies are divided into three main groups:

- **Administrative policies:** These policies define path preference definition, path of last resort, and zero SLA used to minimize control traffic on a metered interface.
- **Performance policies:** These policies define thresholds for delay, loss, and jitter for user-defined DSCP values or application names.
- **Load-balancing policy:** Load balancing can be enabled or disabled globally, or it can be enabled for specific network tunnels. In addition, load balancing can provide specific path preference (for example, the primary path can be INET01 and INET02 with a fallback of MPLS01 and MPLS02).

Site Discovery

PfRv3 was designed to simplify the configuration and deployment of branch sites. The configuration is kept to a minimum and includes the IP address of the Hub MC. All MCs connect to the Hub MC in a hub-and-spoke topology.

When a Branch or Transit MC starts:

- It uses the loopback address of the local MC as its site ID.
- It registers with the Hub MC, providing its site ID, then starts building the IWAN peering with the Hub MC to get all information needed to perform path control. That includes policies and Performance Monitor definitions.
- The Hub MC advertises the site ID information for all sites to all its Branch or Transit MC clients.

At the end of this phase, all MCs have a site prefix database that contains the site ID for every site in the IWAN domain.

Note The site ID is based on the local MC loopback address and is a critical piece of PfR. Routing for MC addresses must be carefully designed to ensure that this address is correctly advertised across all available paths.

Figure 7-8 shows the IWAN peering between all MCs and the Hub MC. R10 is the Hub MC for this topology. R20, R31, R41, and R51 peer with R10. This is the initial phase for site discovery.

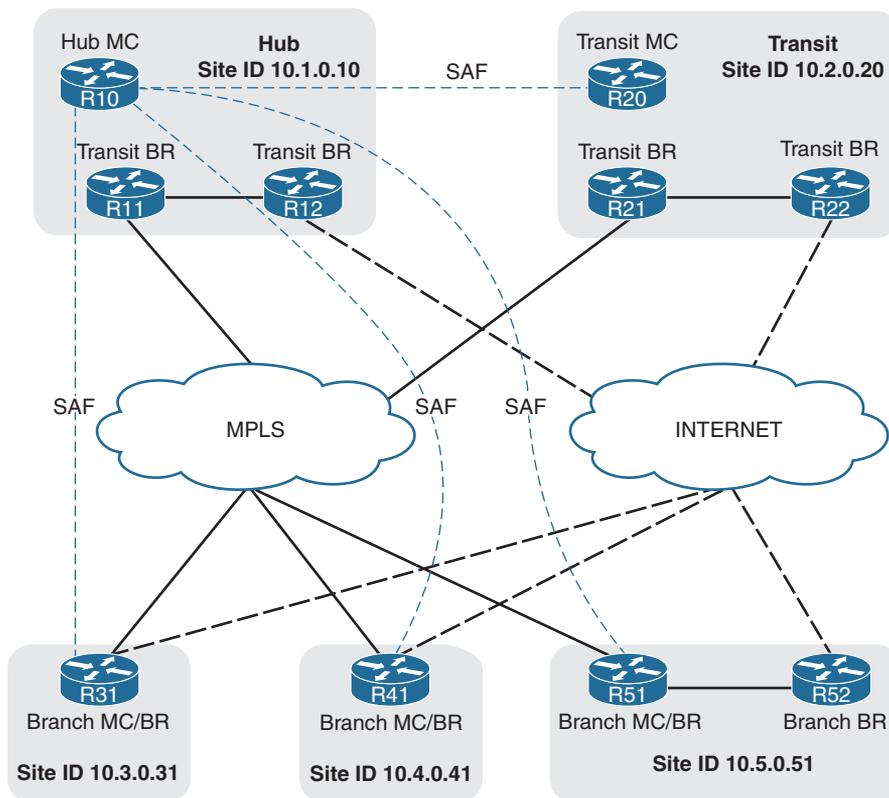


Figure 7-8 Demonstration of IWAN Peering to the Domain Controller

Site Prefix Database

PfR maintains a topology that contains all the network prefixes and their associated site IDs. A site prefix is the combination of a network and the site ID for the network prefix attached to that router. The PfR topology table is known as the *site prefix database* and is a vital component of PfR. The site prefix database resides on the MCs and BRs. The site prefix database located on the MC learns and manages the site prefixes and their origins from both local egress flow and advertisements from remote MC peers. The site prefix database located at a BR learns/manages the site prefixes and their origins only from the advertisements from remote peers. The site prefix database is organized as a longest prefix matching *tree* for efficient search.

Table 7-2 provides the site prefix database on all MCs and BRs for the IWAN domain shown in Figure 7-8. It provides a mapping between a destination prefix and a destination site.

Table 7-2 *Site Prefix Database for an IWAN Domain*

Site Name	Site Identifier	Site Prefix
Site 1	10.1.0.10	10.1.0.0/16
Site 1	10.1.0.10	172.16.1.0/24
Site 2	10.2.0.20	10.2.0.0/16
Site 3	10.2.0.31	10.3.3.0/24
Site 4	10.4.0.41	10.4.4.0/24
Site 5	10.5.0.51	10.5.5.0/24

Note The site prefix database can contain multiple network prefixes per site and is not limited to just one. A second entry was added to the table for Site 1 to display the concept.

In order to learn from advertisements via the peering infrastructure from remote peers, every MC and BR subscribes to the site prefix subservice of the PfR peering service. MCs publish and receive site prefixes. BRs only receive site prefixes. An MC publishes the list of site prefixes learned from local egress flows by encoding the site prefixes and their origins into a message. This message can be received by all the other MCs and BRs that subscribe to the peering service. The message is then decoded and added to the site prefix databases at those MCs and BRs. Site prefixes will be explained in more detail in Chapter 8, “PfR Provisioning.”

Note Site prefixes are dynamically learned at branch sites but must be statically defined at hub and transit sites. The branch site prefixes can be statically defined too.

PfR Enterprise Prefixes

The enterprise-prefix prefix list defines the boundary for all the internal enterprise prefixes. A prefix that is not from the enterprise-prefix prefix list is considered a PfR Internet prefix. PfR does not monitor performance (delay, jitter, byte loss, or packet loss) for network traffic.

In Figure 7-9, all the network prefixes for remote sites (Sites 3, 4, and 5) have been dynamically learned. The central sites (Site 1 and Site 2) have been statically configured. The enterprise-prefix prefix list has been configured to include all the network prefixes in each of the sites so that PfR can monitor performance.

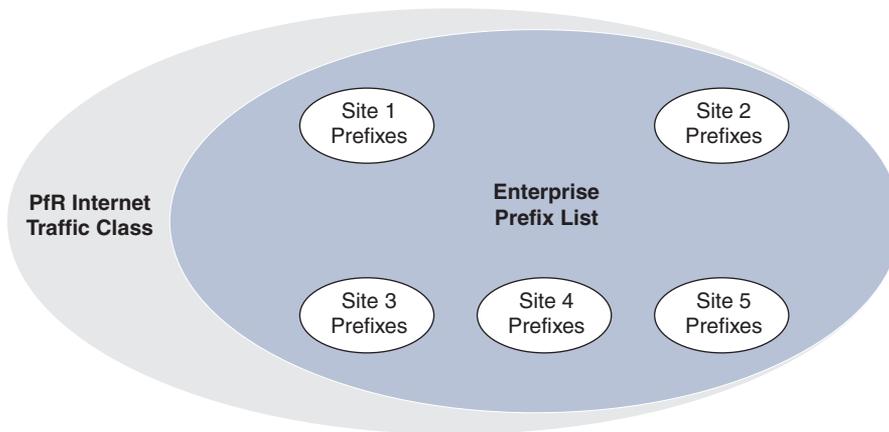


Figure 7-9 *PfR Site and Enterprise Prefixes*

Note In centralized Internet access models, in order for PfR to monitor performance to Internet-based services (email hosting and so forth), the hosting network prefix must be assigned to the enterprise-prefix prefix list. In addition, the hosting network is added to all the site prefix lists for sites that provide Internet connectivity.

WAN Interface Discovery

Border router WAN interfaces are connected to different SPs and have to be defined or discovered by PfR. This definition creates the relationship between the SPs and the administrative policies based on the path name in PfR. A typical example is to define an MPLS-VPN path as the preferred one for all business applications and the Internet-based path as a fallback path when there is a performance issue on the primary.

Hub and Transit Sites

In a PfR domain, a *path name* and a *path identifier* need to be configured for every WAN interface (DMVPN tunnel) on the hub site and all transit sites:

- The *path name* uniquely identifies a transport network. For example, this book uses a primary transport network called MPLS for the MPLS-based transport and a secondary transport network called INET for the Internet-based transport.
- The *path identifier* uniquely identifies a path on a site. This book uses path-id 1 for DMVPN tunnel 100 connected to MPLS and path-id 2 for tunnel 200 connected to INET.

IWAN supports multiple BRs for the same DMVPN network on the hub and transit sites only. The path identifier has been introduced in PfR to be able to track every BR individually.

Every BR on a hub or transit site periodically sends a *discovery packet* with path information to every discovered site. The discovery packets are created with the following default parameters:

- **Source IP address:** Local MC IP address
- **Destination IP address:** Remote site ID (remote MC IP address)
- **Source port:** 18000
- **Destination port:** 19000

Branch Sites

WAN interfaces are automatically discovered on Branch BRs. There is no need to configure the transport names over the WAN interfaces.

When a BR on a branch site receives a *discovery probe* from a central site (hub or transit site):

- It extracts the path name and path identifier information from the probe payload.
- It stores the mapping between the WAN interface and the path name.
- It sends the interface name, path name, and path identifier information to the local MC.
- The local MC knows that a new WAN interface is available and also knows that a BR is available on that path with the path identifier.

The BR associates the tunnel with the correct path information, enables the Performance Monitors, collects performance metrics, collects site prefix information, and identifies traffic that can be controlled.

This discovery process simplifies the deployment of PfR.

Channel

Channels are logical entities used to measure path performance per DSCP between two sites. A channel is created based on real traffic observed on BRs and is based upon a unique combination of factors such as interface, site, next hop, and path. Channels are based on real user traffic or synthetic traffic generated by the BRs called smart probes. A channel is added every time a new DSCP, interface, or site is added to the prefix database or when a new smart probe is received. A channel is a logical construct in PfR and is used to keep track of next-hop reachability and collect the performance metrics per DSCP.

Note In the IWAN 2.1 architecture, multiple next-hop capability was added so that PfR could monitor a path taken through a transit site. A channel is actually created per next hop. In topologies that include a transit site, a channel is created for every next hop to the destination prefix to monitor performance.

Figure 7-10 illustrates the channel creation over the MPLS path for DSCP EF. Every channel is used to track the next-hop availability and collect the performance metrics for the associated DSCP and destination site.

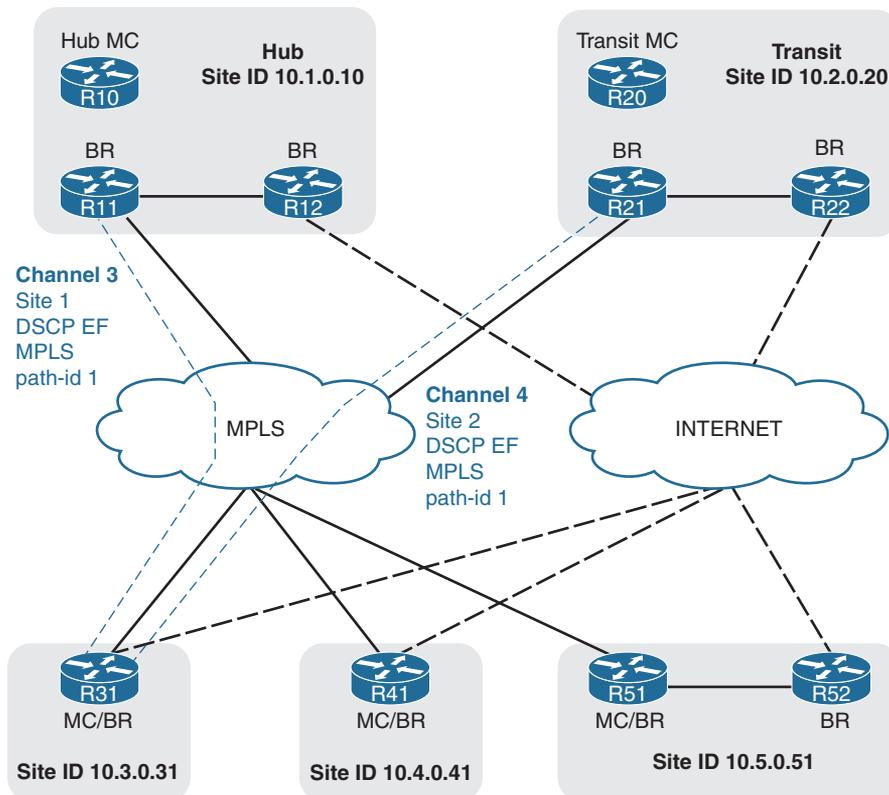


Figure 7-10 Channel Creation for Monitoring Performance Metrics

When a channel needs to be created on a path, PfR creates corresponding channels for any alternative paths to the same destination. This allows PfR to keep track of the performance for the destination prefix and DSCP for every DMVPN network. Channels are deemed active or standby based on the routing decisions and PfR policies.

Multiple BRs can sit on a hub or transit site connected to the same DMVPN network. DMVPN hub routers function as NHRP NHSs for DMVPN and are the BRs for PfR. PfR supports multiple next-hop addresses for hub and transit sites only but limits each of the BRs to hosting only one DMVPN tunnel. This limitation is overcome by placing multiple BRs into a hub or transit site.

The combination of multiple next hops and transit sites creates a high level of availability. A destination prefix can be available across multiple central sites and multiple BRs. For example, if a next hop connected on the preferred path DMVPN tunnel 100 (MPLS) experiences delays, PfR is able to fail over to the other next hop available for DMVPN tunnel 100 that is connected to a different router. This avoids failing over to a less preferred path using DMVPN tunnel 200, which uses the Internet as a transport.

Figure 7-11 illustrates a branch with DSCP EF packets flowing to a hub or transit site that has two BRs connected to the MPLS DMVPN tunnel. Each path has the same path name (MPLS) and a unique path identifier (path-id 1 and path-id 2). If BR1 experiences performance issues, PfR fails over the affected traffic to BR2 over the same preferred path MPLS.

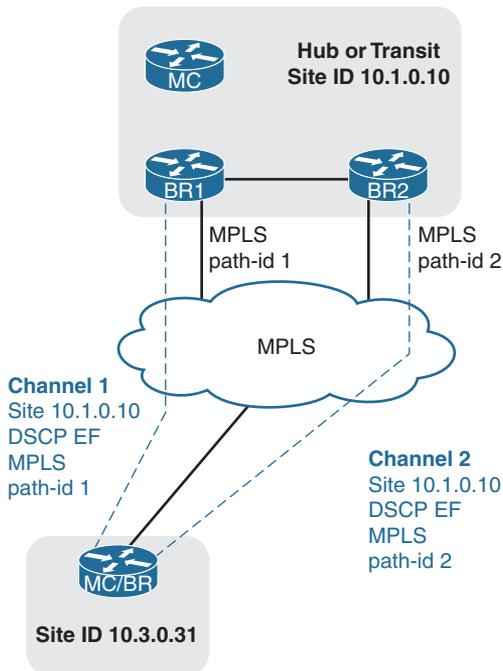


Figure 7-11 Channels per Next Hop

A parent route lookup is done during channel creation. PfR first checks to see if there is an NHRP shortcut route available; if not, it then checks for parent routes in the order of BGP, EIGRP, static, and RIB. If at any point an NHRP shortcut route appears, PfR selects that and relinquishes using the parent route from one of the routing protocols. This behavior allows PfR to dynamically measure and utilize DMVPN shortcut paths to protect site-to-site traffic according to the defined policies as well.

A channel is deemed *reachable* if the following happens:

- Traffic is received from the remote site.
- An unreachable event is not received for two monitor intervals.

A channel is declared *unreachable* in both directions in the following circumstances:

- No packets are received since the last unreachable time from the peer, as detected by the BR. This unreachable timer is defined as one second by default and can be tuned if needed.
- The MC receives an unreachable event from a remote BR. The MC notifies the local BR to make the channel unreachable.

When a channel becomes unreachable, it is processed through the threshold crossing alert (TCA) messages, which will be described later in the chapter.

Smart Probes

Smart probes are synthetic packets that are generated from a BR and are primarily used for WAN interface discovery, delay calculation, and performance metric collection for standby channels. This synthetic traffic is generated only when real traffic is not present, except for periodic packets for one-way-delay measurement. The probes (RTP packets) are sent over the channels to the sites that have been discovered.

Controlled traffic is sent at periodic intervals:

- **Periodic probes:** Periodic packets are sent to compute one-way delay. These probes are sent at regular intervals whether actual traffic is present or not. By default this is one-third of the monitoring interval (the default is 30 seconds), so by default periodic probes are sent every 10 seconds.
- **On-demand probes:** These packets are sent only when there is no traffic on a channel. Twenty packets per second are generated per channel. As soon as user traffic is detected on a channel, the BR stops sending on-demand probes.

Traffic Class

PfR manages aggregations of flows called *traffic classes (TCs)*. A traffic class is an aggregation of flows going to the same destination prefix, with the same DSCP or application name (if application-based policies are used).

Traffic classes are learned on the BR by monitoring a WAN interface's egress traffic. This is based on a Performance Monitor instance applied on the external interface.

Traffic classes are divided into two groups:

- **Performance TCs:** These are any TCs with performance metrics defined (delay, loss, jitter).
- **Non-performance TCs:** The default group, these are the TCs that do not have any defined performance metrics (delay, loss, jitter), that is, TCs that do have any match statements in the policy definition on the Hub MC.

For every TC, the PfR route control maintains a list of active channels (current exits) and standby channels.

Note Real-time load balancing affects only non-performance TCs. PfR moves default TCs between paths to keep bandwidth utilization within the boundaries of a predefined ratio. For performance TCs, new TCs use the least loaded path. After a traffic class is established, it stays on the path defined, unless that path becomes out of policy.

Path Selection

Path and next-hop selection in PfR depends on the routing design in conjunction with the PfR policies. From a central site (hub and transit) to a branch site, there is only one possible next hop per path. From a branch site to a central site, multiple next hops can be available and may span multiple sites. PfR has to make a choice among all next hops available to reach the destination prefix of the traffic to control.

Direction from Central Sites (Hub and Transit) to Spokes

Each central site is a distinct site by itself and controls only traffic toward the spoke on the WAN paths to that site. PfR does not redirect traffic between central sites across the DCI or WAN core to reach a remote site. If the WAN design requires that all the links be considered from POP to spoke, use a single MC to control all BRs from both central sites.

Direction from Spoke to Central Sites (Hub and Transit)

The path selection from BR to a central site router can vary based on the overall network design. The following sections provide more information on PfR's path selection process.

Active/Standby Next Hop

The spoke considers all the paths (multiple next hops) toward the central sites and maintains a list of active/standby candidate next hops per prefix and interface. The concept of *active* and *standby* next hops is based on the routing best metric to gather information about the preferred POP for a given prefix. If the best metric for a given prefix is on a specific central site, all the next hops on that site for all the paths are

tagged as *active* (only for that prefix). A next hop in a given list is considered to have a best metric based on the following metrics/criteria:

- Advertised mask length
- BGP weight and local preference
- EIGRP feasible distance (FD) and successor FD

Transit Site Affinity

Transit Site Affinity (also called POP Preference) is used in the context of a multiple-transit-site deployment with the same set of prefixes advertised from multiple central sites. Some branches prefer a specific transit site over the other sites. The affinity of a branch to a transit site is configured by altering the routing metrics for prefix advertisements to the branch from the transit site. If one of the central sites advertising a specific prefix has the best next hop, the entire site is preferred over the other sites for all TCs to this destination prefix. Transit site preference is a higher-priority filter and takes precedence over path preference. The Transit Site Affinity feature was introduced in Cisco IWAN 2.1.

Path Preference

During Policy Decision Point (PDP), the exits are first sorted on the available bandwidth, Transit Site Affinity, and then a third sort algorithm that places all primary path preferences in the front of the list followed by fallback preferences. A common deployment use case is to define a primary path (MPLS) and a fallback path (INET). During PDP, MPLS is selected as the primary channel, and if INET is within policy it is selected as the fallback.

- With path preference configured, PfR first considers all the links belonging to the preferred path (that is, it includes the active and the standby links belonging to the preferred path) and then uses the fallback provider links.
- Without path preference configured, PfR gives preference to the active channels and then the standby channels (active/standby is per prefix) with respect to the performance and policy decisions.

Note Active/standby tagging happens whether Transit Site Affinity is enabled or disabled. The active and standby channels (per prefix) may span central sites if they advertise the same prefix. Spoke routers use a hash to choose the active channel.

Transit Site Affinity and Path Preference Usage

Transit Site Affinity and path preference are used in combination to influence the next-hop selection per TC. For example, this book uses a topology with two central sites (Site 1 and Site 2) and two paths (MPLS and INET). Both central sites advertise the same prefix (10.10.0.0/16 as an example), and Site 1 has the best next hop for that prefix

(R11 advertises 10.10.0.0/16 with the highest BGP local preference). Enabling Transit Site Affinity and defining a path preference with MPLS as the primary and INET as the fallback path, the BR identifies the following routers (in order) for the next hop:

1. R11 is the primary next hop for TCs with 10.10.0.0/16 as the destination prefix
2. Then R12 (same site, because of Transit Site Affinity)
3. Then R21 (Site 2, because of path preference)
4. Then R22

Performance Monitoring

The PfR monitoring system interacts with the IOS component called *Performance Monitor* to achieve the following tasks:

- Learning site prefixes and applications
- Collecting and analyzing performance metrics per DSCP
- Generating threshold crossing alerts
- Generating out-of-policy report

Performance Monitor is a common infrastructure within Cisco IOS that passively collects performance metrics, number of packets, number of bytes, statistics, and more within the router. In addition, Performance Monitor organizes the metrics, formats them, and makes the information accessible and presentable based upon user needs. Performance Monitor provides a central repository for other components to access these metrics.

PfR is a client of Performance Monitor, and through the performance monitoring metrics, PfR builds a database from that information and uses it to make an appropriate path decision. When a BR component is enabled on a device, PfR configures and activates three *Performance Monitor instances (PMIs)* over all discovered WAN interfaces of branch sites, or over all configured WAN interfaces of hub or transit sites. Enablement of PMI on these interfaces is dynamic and completely automated by PfR. This configuration does not appear in the startup or running configuration file.

The PMIs are

- **Monitor 1:** Site prefix learning (egress direction)
- **Monitor 2:** Egress aggregate bandwidth per traffic class
- **Monitor 3:** Performance measurements (ingress direction)

Monitor 3 contains two monitors: one dedicated to the business and media applications where failover time is critical (called *quick monitor*), and one allocated to the default traffic.

PfR policies are applied either to an application definition or to DSCP. Performance is measured only per DSCP because SPs can differentiate traffic only based on DSCP and not based on application.

Performance is measured between two sites where there is user traffic. This could be between hub and a spoke, or between two spokes; the mechanism remains the same.

- The egress aggregate monitor instance captures the number of bytes and packets per TC on egress on the source site. This provides the bandwidth utilization per TC.
- The ingress per DSCP monitor instance collects the performance metrics per DSCP (channel) on ingress on the destination site. Policies are applied to either application or DSCP. However, performance is measured per DSCP because SPs differentiate traffic only based on DSCP and not based on discovered application definitions. All TCs that have the same DSCP value get the same QoS treatment from the provider, and therefore there is no real need to collect performance metrics per application-based TC.

PfR passively collects metrics based on real user traffic and collects metrics on alternative paths too. The source MC then instructs the BR connected to the secondary paths to generate smart probes to the destination site. The PMI on the remote site collects statistics in the same way it would for actual user traffic. Thus, the health of a secondary path is known prior to being used for application traffic, and PfR can choose the best of the available paths on an application or DSCP basis.

Figure 7-12 illustrates PfR performance measurement with network traffic flowing from left to right. On the ingress BRs (BRs on the right), PfR monitors performance per channel. On the egress BRs (BRs on the left), PfR collects the bandwidth per TC. Metrics are collected from the user traffic on the active path and based on smart probes on the standby paths.

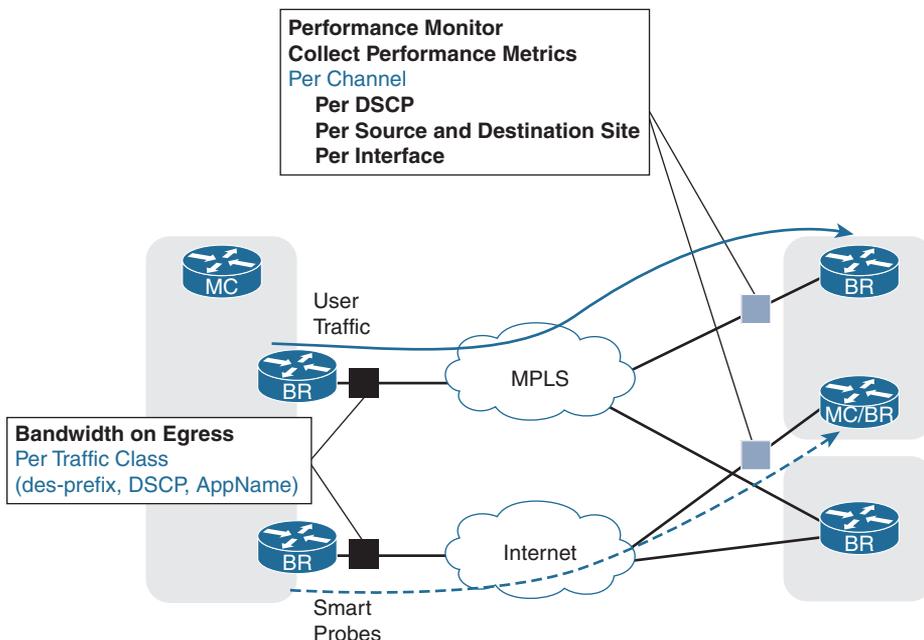


Figure 7-12 PfR Performance Measurement via Performance Monitor

Note Smart probes are not IP SLA probes. Smart probes are directly forged in the data plane from the source BR and discarded on the destination BR after performance metrics are collected.

Threshold Crossing Alert (TCA)

Threshold crossing alert (TCA) notifications are alerts for when network traffic exceeds a set threshold for a specific PfR policy. TCAs are generated from the PMI attached to the BR's ingress WAN interfaces and smart probes. Figure 7-13 displays a TCA being raised on the destination BR.

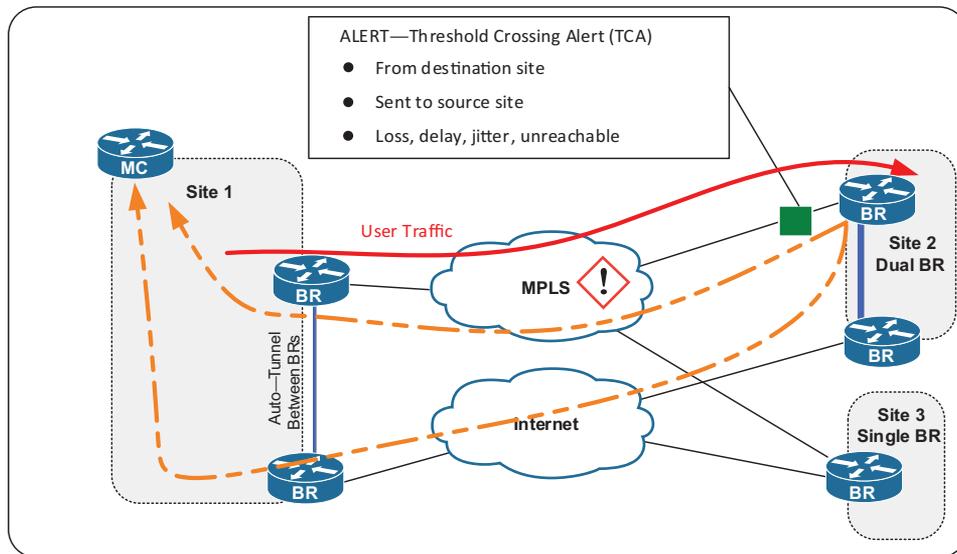


Figure 7-13 *Threshold Crossing Alert (TCA)*

Threshold crossing alerts are managed on both the destination BR and source MC for the following scenarios:

- The destination BR receives performance TCA notifications from the PMI, which monitors the ingress traffic statistics and reports TCA alerts when threshold crossing events occur.
- The BR forwards the performance TCA notifications to the MC on the source site that actually generates the traffic. This source MC is selected from the site prefix database based on the source prefix of the traffic. TCA notifications are transmitted via multiple paths for reliable delivery.

- The source MC receives the TCA notifications from the destination BR and translates the TCA notifications (that contain performance statistics) to an out-of-policy (OOP) event for the corresponding channel.
- The source MC waits for the TCA processing delay time for all the notifications to arrive, then starts processing the TCA. The processing involves selecting TCs that are affected by the TCA and moving them to an alternative path.

Path Enforcement

PfR uses the Route Control Enforcement module for optimal traffic redirection and path enforcement. This module performs lookups and reroutes traffic similarly to policy-based routing but without using an ACL. The MC makes path decisions for every unique TC. The MC picks the next hop for a TC's path and instructs the local BR how to forward packets within that TC.

Because of how path enforcement is implemented, the next hop has to be directly connected to each BR. When there are multiple BRs on a site, PfR sets up an mGRE tunnel between all of them to accommodate path enforcement. Every time a WAN exit point is discovered or an *up/down* interface notification is sent to the MC, the MC sends this notification to all other BRs in the site. An endpoint is added to the mGRE tunnel pointing toward this BR as a result.

When packets are received on the LAN side of a BR, the route control functionality determines if it must exit via a local WAN interface or via another BR. If the next hop is via another BR, the packet is sent out on the tunnel toward that BR. Thus the packet arrives at the destination BR within the same site. Route control gets the packet, looks at the channel identifier, and selects the outgoing interface. The packet is then sent out of this interface across the WAN.

Summary

This chapter provided a thorough overview of Cisco intelligent path control, which is a core pillar of the Cisco IWAN architecture and is based upon Performance Routing (PfR). The following chapters will expand upon these theories while explaining the configuration of PfR.

PfR provides the following benefits for a WAN architecture:

- Maximizes WAN bandwidth utilization
- Protects applications from performance degradation
- Uses passive monitoring to track application performance across the WAN
- Enables the Internet as a viable WAN transport
- Provides multisite coordination to simplify network-wide provisioning

- Provides an application-based policy-driven framework that is tightly integrated with existing Performance Monitor components
- Provides a smart and scalable multisite solution to enforce application SLAs while optimizing network resource utilization

PfRv3 is the third-generation multisite-aware bandwidth and path control/optimization solution for WAN- and cloud-based applications and is available now on Cisco *Integrated Services Router (ISR)* Generation 2 series, ISR-4000 Series, and CSR 1000V and ASR 1000 Series routers.

Further Reading

Cisco. “Performance Routing Version 3.” www.cisco.com.

Cisco. “PfRv3 Transit Site Support.” www.cisco.com.