

CISCO SYSTEMS



Traffic Engineering with MPLS

Design, configure, and manage MPLS TE
to optimize network performance

ciscopress.com

Eric Osborne, CCIE® #4122
Ajay Simha, CCIE #2970

Traffic Engineering with MPLS

Eric Osborne

Ajay Simha

Copyright © 2003 Cisco Systems, Inc.

Published by:

Cisco Press

201 West 103rd Street

Indianapolis, IN 46290 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 or 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 1 2 3 4 5 6 7 8 9 0

First Printing July 2002

Library of Congress Cataloging-in-Publication Number: 200-1086632

ISBN: 1-58705-031-5

Warning and Disclaimer

This book is designed to provide information about Multiprotocol Label Switching Traffic Engineering (MPLS TE). 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 and 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.

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 of the professional technical community.

Reader 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 e-mail at feedback@ciscopress.com. Please be sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

You can find additional information about this book, any errata, and Appendix B at www.ciscopress.com/1587050315.

Publisher
Editor-In-Chief
Cisco Systems Management

Executive Editor
Production Manager
Development Editor
Project Editor
Copy Editor
Technical Editors

Team Coordinator
Book Designer
Cover Designer
Compositor
Indexer

John Wait
John Kane
Michael Hakkert
Tom Geitner
Brett Bartow
Patrick Kanouse
Christopher Cleveland
Eric T. Schroeder
Gayle Johnson
Jim Guichard
Alexander Marhold
Jean-Philippe Vasseur
Tammi Ross
Gina Rexrode
Louisa Klucznik
Amy Parker
Tim Wright



Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems Europe
11 Rue Camille Desmoulins
92782 Issy-les-Moulineaux
Cedex 9
France
<http://www-europe.cisco.com>
Tel: 33 1 58 04 60 00
Fax: 33 1 58 04 61 00

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems Australia,
Pty., Ltd.
Level 17, 99 Walker Street
North Sydney
NSW 2059 Australia
<http://www.cisco.com>
Tel: +61 2 8448 7100
Fax: +61 2 9957 4350

Cisco Systems has more than 200 offices in the following countries. Addresses, phone numbers, and fax numbers are listed on the Cisco Web site at www.cisco.com/go/offices

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China • Colombia • Costa Rica • Croatia • Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong • Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden • Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

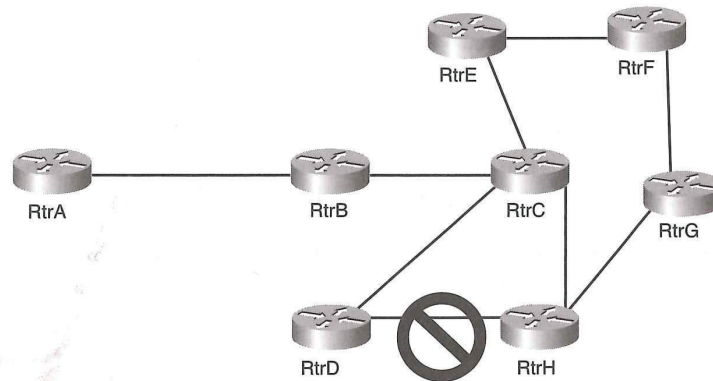
Copyright © 2000, Cisco Systems, Inc. All rights reserved. Access Registrar, AccessPath, Are You Ready, ATM Director, Browse with Me, CCDA, CCDE, CCDP, CCIE, CCNA, CCNP, CCSI, CD-PAC, CiscoLink, the Cisco NetWorks logo, the Cisco Powered Network logo, Cisco Systems Networking Academy, Fast Step, FireRunner, Follow Me Browsing, FormShare, GigaStack, IGX, Intelligence in the Optical Core, Internet Quotient, IP/VC, iQ Breakthrough, iQ Expertise, iQ FastTrack, iQuick Study, iQ Readiness Scorecard, The iQ Logo, Kernel Proxy, MGX, Natural Network Viewer, Network Registrar, the Networkers logo, Packet, PIX, Point and Click Internetworking, Policy Builder, RateMUX, ReyMaster, ReyView, ScriptShare, Secure Script, Shop with Me, SlideCast, SMARTnet, SVX, TrafficDirector, TransPath, VianDirector, Voice LAN, Wavelength Router, Workgroup Director, and Workgroup Stack are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn, Empowering the Internet Generation, are service marks of Cisco Systems, Inc.; and Aironet, ASIST, BPIX, Catalyst, Cisco, the Cisco Certified Internetwork Expert Logo, Cisco IOS, the Cisco IOS logo, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Collision Free, Enterprise/Solver, EtherChannel, EtherSwitch, FastHub, FastLink, FastPAD, IOS, IPTV, IPX, LightStream, LightSwitch, MICA, NetRanger, Post-Routing, Pre-Routing, Registrar, StratView Plus, Stratix, SwitchProbe, TeleRouter, are registered trademarks of Cisco Systems, Inc. or its affiliates in the U.S. and certain other countries.

All other brands, names, or trademarks mentioned in this document or Web site 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. (0010R)

Tunnel Reoptimization

What happens if, while a tunnel is up, a path appears that is a new best path from head to tail? Consider the network shown in Figure 4-7.

Figure 4-7 New Path Is a More Optimal TE Tunnel Candidate



In Figure 4-7:

- All links start with 100-Mbps reservable bandwidth.
- Router A and Router D both want to build 60-Mbps tunnels to Router H.
- The link between Router D and Router H is down.

Assume that the following sequence of events happens:

- 1 Router D builds a tunnel D→C→H.
- 2 Router A builds a tunnel A→B→C→E→F→G→H.
- 3 Router D reduces its D→C→H bandwidth reservation to 30 Mbps, either by configuration or because of something like auto bandwidth adjustment.

At this point, the best possible path Router A could take to get to Router H is A→B→C→H. But Router A already has a tunnel established. Should Router A change its tunnel over to the B→C→H path? What if Router D needs to increase its tunnel bandwidth again? How long should Router A wait before cutting over to the best path? You can ask all kinds of questions along this line.

When a router looks to see if there is a better path for tunnels that are already up, this is known as *reoptimization*. Four things affect reoptimization:

- Periodic reoptimization
- Manual reoptimization
- Event-driven reoptimization
- Lockdown

exclude address; it's
t path that uses both

th submode.

fied index.

or delete.

s.

g path option. You can
:(through **append-after**
index number). **no index**

when selecting a path
is listed; they are tried in
Router A to Router E, as
en path 4 if path 3 is
ally path 1, you could

nel calculation be **dynamic**:

ensures that if it's at all
a dynamic path option isn't
uld be useful if a link were
options wouldn't need to be
path.

It's important to understand that this reoptimization has nothing to do with the behavior when a tunnel goes down. If a tunnel goes down, you don't wait for the reoptimization timer to kick in before trying to find a better path for that tunnel; that calculation is done immediately.

Finally, realize that RSV-TE has what's called *make-before-break*, so that the act of making a new tunnel reservation does not disturb any existing reservations for that tunnel. This is covered in more detail in the "Resource Reservation Protocol (RSVP)" section.

Periodic Reoptimization

Cisco implements a periodic reoptimization timer, which can be configured on a global basis. After a tunnel comes up, an attempt is made to find a better path for it, given the tunnel's configured constraints. By default, this happens once an hour; this timer can be configured through the command **mpls traffic-eng tunnels reoptimize timers frequency 0-604800**. 0-604800 is how often (in seconds) Cisco IOS Software looks for a better path for a tunnel.

Setting this timer to 0 means that tunnels are never reoptimized after they are up; on the other end of the scale, 604,800 is the number of seconds in 168 hours, or exactly one week. The default reoptimization timer is 3600 seconds, or one hour.

It's important to note that the reoptimization timer, although it can be configured only globally, is kept on a per-tunnel basis. What does this mean? Suppose you have 20 different tunnels, T1 through T20, that come up 2 minutes after each other (T1 comes up at 00:00, T2 at 00:02, T3 at 00:04, and so on). T20 comes up at 00:40. Twenty minutes after that, the global reoptimization timer for T1 kicks in and tries to find a better path, but only for T1. T20 is not reoptimized until it has been up for an hour, or when the clock reads 1:40.

Manual Reoptimization

What if you know there's been a change in your network, and you don't want to wait for that tunnel's reoptimization timer to kick in before finding a better path? You can use the enable-level command **mpls traffic-eng reoptimize [tunnel-name]** to force the router to reoptimize a specific tunnel at any time.

Event-Driven Reoptimization

Finally, consider the link between RtrD and RtrH in Figure 4-7. If that link comes up, should RtrD reoptimize its D→H tunnel so that this tunnel flows over the directly connected link? Probably. But there are scenarios in which a link coming up should *not* trigger a reoptimization.

nothing to do with the behavior
t wait for the reoptimization timer
el; that calculation is done

re-break, so that the act of making
ervations for that tunnel. This is
ocol (RSVP)" section.

can be configured on a global
a better path for it, given the
nce an hour; this timer can be
s **reoptimize timers frequency**
Software looks for a better path

nized after they are up; on the
168 hours, or exactly one week.
our.

h it can be configured only
? Suppose you have 20 different
other (T1 comes up at 00:00,
. Twenty minutes after that, the
a better path, but only for T1.
when the clock reads 1:40.

nd you don't want to wait for
better path? You can use the
-name] to force the router to

4-7. If that link comes up,
lows over the directly
link coming up should *not*

One scenario in which an up link should not trigger a reoptimization is if the link is flapping and the IGP timers aren't enough to hide the flapping from the network. Reoptimizing onto a flapping link and then calculating a path around that link when it fails again constantly changes the tunnel's path characteristics. Although RSVP's make-before-break leads to minimal data loss along this path, a constantly changing tunnel path could lead to large delay variation, which isn't good for either data (TCP) or voice.

For this reason, tunnels are not automatically reoptimized when a new link comes up. Tunnels are reoptimized onto a link that has just come up when the tunnel reoptimization timer goes off, but not until then. You can change this behavior with the following global configuration:

```
mpls traffic-eng reoptimize events link-up
```

Lockdown

You might have some tunnels that you don't ever want to reoptimize. You can specify this on a per-tunnel basis using the **lockdown** option in the path option specification:

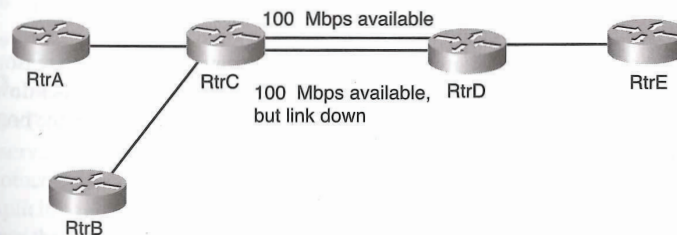
```
tunnel mpls traffic-eng path-option preference {dynamic | explicit name |  
identifier id>} {lockdown}
```

The reason for putting lockdown on a path option as opposed to a tunnel itself is that there might be some paths you never want to leave as long as the links underneath those paths are up, but if the underlying links are down, you want to fall back to another path until the preferred path is up.

To make that last sentence clearer, an example is in order.

In Figure 4-8, each link has 100 Mbps reservable to begin with.

Figure 4-8 Sample Network for Lockdown Example



At some point in time, two tunnels were established; at that time, the lower link between RtrC and RtrD was down. At some point after these tunnels came up, the lower link also came up. A 60-Mbps tunnel from RtrA to RtrE crosses the top C→D link, and a 30-Mbps tunnel from RtrB to RtrE crosses the same link. Figure 4-9 shows this.