

# Cisco Reader Comment Card

## General Information

- 1 Years of networking experience \_\_\_\_\_ Years of experience with Cisco products \_\_\_\_\_
- 2 I have these network types:  LAN  Backbone  WAN  
 Other: \_\_\_\_\_
- 3 I have these Cisco products:  Switches  Routers  
 Other: Specify model(s) \_\_\_\_\_
- 4 I perform these types of tasks:  H/W Install and/or Maintenance  S/W Config  
 Network Management  Other: \_\_\_\_\_
- 5 I use these types of documentation:  H/W Install  H/W Config  S/W Config  
 Command Reference  Quick Reference  Release Notes  Online Help  
 Other: \_\_\_\_\_
- 6 I access this information through: \_\_\_\_\_% Cisco Connection Online (CCO) \_\_\_\_\_% CD-ROM  
\_\_\_\_\_% Printed docs \_\_\_\_\_% Other: \_\_\_\_\_
- 7 Which method do you prefer? \_\_\_\_\_
- 8 I use the following three product features the most:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

## Document Information

Document Title: Cisco MPLS Controller Software Configuration Guide

Part Number: 78-10672-01

S/W Release (if applicable): 9.3.10

On a scale of 1–5 (5 being the best) please let us know how we rate in the following areas:

\_\_\_\_\_ The document was written at my  
technical level of understanding.

\_\_\_\_\_ The information was accurate.

\_\_\_\_\_ The document was complete.

\_\_\_\_\_ The information I wanted was easy to find.

\_\_\_\_\_ The information was well organized.

\_\_\_\_\_ The information I found was useful to my job.

Please comment on our lowest score(s):  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

## Mailing Information

Company Name

Date

Contact Name

Job Title

Mailing Address  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

City

State/Province

ZIP/Postal Code

Country

Phone ( )

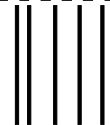
Extension

Fax ( )

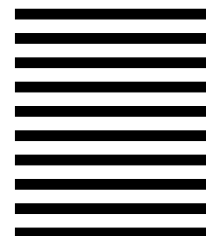
E-mail

Can we contact you further concerning our documentation?  Yes  No

You can also send us your comments by e-mail to [bug-doc@cisco.com](mailto:bug-doc@cisco.com), or fax your comments to us at (408) 527-8089.



NO POSTAGE  
NECESSARY  
IF MAILED  
IN THE  
UNITED STATES



**BUSINESS REPLY MAIL**  
FIRST-CLASS MAIL PERMIT NO. 4631 SAN JOSE CA

POSTAGE WILL BE PAID BY ADDRESSEE

ATTN DOCUMENT RESOURCE CONNECTION  
**CISCO SYSTEMS INC**  
170 WEST TASMAN DRIVE  
SAN JOSE CA 95134-9883





## Cisco MPLS Controller Software Configuration Guide

Release 9.3.10  
May, 2001

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

Customer Order Number: DOC-7811658  
Text Part Number: 78-11658-01, Rev. B0

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Access Registrar, AccessPath, Any to Any, Are You Ready, AtmDirector, Browse with Me, CCDA, CCDE, CCDP, CCIE, CCNA, CCNP, CCSI, CD-PAC, the Cisco logo, Cisco Certified Internetwork Expert logo, *CiscoLink*, the Cisco Management Connection logo, the Cisco NetWorks logo, the Cisco Powered Network logo, Cisco Systems Capital, the Cisco Systems Capital logo, Cisco Systems Networking Academy, the Cisco Systems Networking Academy logo, the Cisco Technologies logo, Fast Step, FireRunner, Follow Me Browsing, FormShare, GigaStack, IGX, Intelligence in the Optical Core, Internet Quotient, IP/VC, IQ Breakthrough, IQ Expertise, IQ FastTrack, IQ Readiness Scorecard, The IQ Logo, Kernel Proxy, MGX, Natural Network Viewer, NetSonar, Network Registrar, the Networkers logo, *Packet*, PIX, Point and Click Internetworking, Policy Builder, Precept, RateMux, ReyMaster, ReyView, ScriptShare, Secure Script, Shop with Me, SlideCast, SMARTnet, SVX, *The Cell*, TrafficDirector, TransPath, VlanDirector, Voice LAN, Wavelength Router, Workgroup Director, and Workgroup Stack are trademarks; Changing the Way We Work, Live, Play, and Learn, Empowering the Internet Generation, The Internet Economy, and The New Internet Economy are service marks; and Aironet, ASIST, BPX, Catalyst, Cisco, Cisco IOS, the Cisco IOS logo, Cisco Systems, the Cisco Systems logo, the Cisco Systems Cisco Press logo, CollisionFree, Enterprise/Solver, EtherChannel, EtherSwitch, FastHub, FastLink, FastPAD, FastSwitch, GeoTel, IOS, IP/TV, IPX, LightStream, LightSwitch, MICA, NetRanger, Post-Routing, Pre-Routing, Registrar, StrataView Plus, Stratm, TeleRouter, and VCO are registered trademarks of Cisco Systems, Inc. or its affiliates in the U.S. and certain other countries. All other trademarks mentioned in this document 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. (0005R)

*Cisco MPLS Controller Software Configuration Guide*

Copyright © 2000, Cisco Systems, Inc.

All rights reserved. Printed in USA.



## **Preface xvii**

Documentation CD-ROM	xvii
Related Documentation	xviii
Previous Cisco WAN Switch Product Names	xix
MPLS and Tag Terminology	xix
Terms Specific to MPLS	xx
Conventions	xxi
Cisco Connection Online	xxii
Documentation CD-ROM	xxii

---

## CHAPTER 1

### **Introduction to MPLS 1-1**

What is MPLS?	1-1
Label Switching Features	1-2
Label Switching Benefits	1-3
MPLS Compared to Other IP-over-ATM Schemes	1-4
Problems of Running IP Routing over An ATM Network without MPLS	1-5
MPLS Network Structure	1-6
MPLS Applications	1-7
MPLS Virtual Private Network	1-7
Intranet and Extranet VPNs	1-8
MPLS VPN Features	1-8
MPLS VPN Benefits	1-10
References	1-11

---

## CHAPTER 2

### **Integrating MPLS with IP and ATM 2-1**

Why Integrate IP with ATM?	2-1
Structure of An IP+ATM Switch	2-3
Use of IP+ATM	2-5
Routing on ATM Switches	2-6
Building Internets on ATM	2-6
Label Switching Operation at Layer 3	2-7
Forwarding Component	2-8

- Control Component 2-9
- MPLS Elements in An ATM WAN 2-10
  - Forwarding Via ATM Switches 2-11
  - Control Via ATM Switches 2-12
  - Cell Interleave Problem 2-13
  - Virtual Circuit Merge-Capable Switches 2-14
  - Label VC Connections and Cross-Connects 2-15
- Label Switch Controllers 2-16
  - BPX 8650 Label Switch Router: Controlling a BPX 8600 with An LSC 2-17
  - IP+ATM Capability 2-18
- An ATM MPLS Point of Presence 2-20
  - Using an LSC as An Edge LSR 2-21
  - Using An Access Switch in An ATM MPLS PoP 2-22
  - A Fully Integrated PoP 2-23
- Dual Backbones: Traditional ATM and ATM MPS or Packet-Over-SONET 2-23
- Virtual Private Networks 2-25
  - Route Distinguisher 2-26
  - Forwarding in a Cisco Virtual Private Network Service 2-27
  - Control in a Cisco MPLS+BGP Virtual Private Network Service 2-28
  - Attributes of Cisco MPLS+BGP Virtual Private Networks 2-30
    - Privacy and Security 2-30
    - Customer Independence 2-31
    - Scalability and Stability 2-31
    - Management 2-32
- Migrating MPLS into a Traditional ATM Network 2-33

CHAPTER 3

**Designing MPLS for ATM 3-1**

- Structures for MPLS Networks 3-1
  - Simple Packet-based MPLS 3-2
  - ATM MPLS with Router-based Edge LSRs 3-2
  - Mixed ATM and Packet-based MPLS 3-3
  - ATM MPLS with Separate Access Devices 3-3
  - ATM MPLS with Integrated IP+ATM Access Devices 3-3
  - ATM MPLS Using Traditional ATM Switches 3-4
  - Dual Backbones 3-5
- Choosing Cisco MPLS Equipment for ATM 3-6

Choosing ATM MPLS Edge Equipment	3-6
Choosing ATM Label Switch Routers	3-9
Label Switch Routers Not Based on ATM Switches	3-11
Designing MPLS Networks	3-11
Points of Presence Structures	3-12
Single ATM Edge LSR	3-12
Multiple Edge LSRs and An ATM LSR	3-12
Edge LSR PoP with BPX 8650 and MGX 8220 Access Concentrators	3-14
Cisco 6400 and MGX 8850 Edge LSRs	3-14
Stand-Alone ATM LSRs	3-15
Dimensioning An MPLS Network's Links	3-15
Redundant Pairs of ATM Links	3-23
IP Routing in An MPLS Network	3-24
MPLS-Specific IP Routing Issues	3-27
Dimensioning MPLS Label VC Space	3-29
Destinations	3-29
LVCs Used Per Link and VC Merge	3-30
Design Calculations: Edge LSRs	3-31
Edge LSR Examples	3-33
Design Calculations: ATM LSRs with VC Merge	3-35
ATM LSRs with VC Merge: Example 1	3-36
ATM LSRs with VC Merge: Example 2	3-36
ATM LSRs with VC Merge: Example 3	3-37
Design Calculations: ATM LSRs without VC Merge	3-37
ATM LSRs without VC Merge, with One CoS: Example 1	3-38
ATM LSRs without VC Merge, with Two CoS: Example 2	3-39
Additional Example Considerations	3-39
Internet Routing Tables	3-39
Traffic Engineering	3-40
VP Tunnels	3-40
Alternative Calculations	3-40
Ongoing Network Design	3-41

**Quality of Service in MPLS Networks** 4-1

MPLS QoS with IP+ATM Overview	4-1
Best Effort Traffic and IP QoS Requirements	4-3

- Effects of Connectionless Traffic 4-3
- Specifying QoS for Connectionless Service 4-5
- The Differential Services Approach to Quality of Service 4-6
  - Contracts for Access Bandwidths 4-6
  - Using Best-Effort Traffic to Help Guarantee Bandwidths 4-8
- Modeling Network Traffic Flows to Meet Service Level Agreements 4-10
- A Recommended Process for Estimating and Modeling Traffic 4-12
  - Engineering DiffServ Per-Hop Behaviors 4-13
  - DiffServ Classes and Cisco IP+ATM Switches 4-15
  - Service-Level Agreements Using DiffServ 4-17
  - Sample Service Level Agreement Using the Two-Class Model 4-18
  - Sample Service Level Agreement with Provision for Real-Time Traffic 4-20
  - Adding a New Site 4-21
  - What If There Isn't Much Best-Effort Traffic in My Network? 4-21
    - Standardization 4-22
  - The Differential Services Approach to Quality of Service: Summary 4-22
- MPLS Traffic Engineering 4-23
- More Stringent Quality of Service in IP+ATM Networks 4-25
- Quality of Service for MPLS VPNs 4-26
- Discard Policies 4-28
- Delay Limits 4-33
- Alternative Service Types 4-33

CHAPTER 5

- Configuring MPLS with the BPX Switch and the 6400/7200/7500 Routers 5-1**
  - Introduction 5-2
  - Equipment and Software Requirements 5-2
  - Configuration Preview 5-3
  - Initial Setup of MPLS Switching 5-6
  - Configuration for BPX Switch Portions of the BPX 8650 ATM-LSRs 5-7
    - Command Syntax Summary for BPX Portion of MPLS Configuration 5-7
    - Configuration for BPX 1 Portion of ATM-LSR-1 5-8
    - Configuration for BPX 2 Portion of ATM-LSR-2 5-10
  - Configuration for LSC 1 and LSC 2 Portions of the BPX 8650 5-12
    - Configuration for LSC1 Portion of ATM-LSR-1 5-12
    - Configuration for LSC2 Portion of ATM-LSR-2 5-14
  - Configuration for Edge Label Switch Routers, LSR-A and LSR-B 5-15



Configuration of Cisco 7500 as An Edge Router, Edge LSR-A	5-16
Configuration of Cisco 7500 as An Edge Router, Edge LSR-C	5-16
MPLS Configures LVCs According to the Routing Protocol	5-17
Testing the MPLS Network Configuration	5-18
Useful LSC Commands	5-18
Checking the BPX Extended ATM Interfaces	5-18
Basic Router Configuration	5-23
Accessing the Router Command-Line Interface	5-23
Booting the Router for the First Time	5-23
Configuring the Router for the First Time	5-23
Using the System Configuration Dialog	5-24
Configuring Port Adapter Interfaces	5-27
Preparing to Configure Port Adapter Interfaces	5-27
Identifying Chassis Slot, Port Adapter Slot, and Interface Port Numbers	5-27
Configuring ATM Interfaces	5-28
Other Router Interfaces	5-29
Checking the Configuration	5-29
Using Show Commands to Verify the New Interface Status	5-29
Using Show Commands to Display Interface Information	5-29
Cisco Show Interfaces Command	5-30
Using the ping Command	5-31
Using Configuration Mode	5-32
Cisco IOS Software Basics	5-33
Cisco IOS Modes of Operation	5-33
Getting Context-Sensitive Help	5-35
Saving Configuration Changes	5-35

## CHAPTER 6

<b>MPLS CoS with BPX 8650</b>	<b>6-1</b>
MPLS CoS Overview	6-1
Related Documents	6-2
Prerequisites	6-2
MPLS CoS in An IP+ATM Network	6-3
ATM CoS Service Templates and Qbins on the BPX 8650	6-5
Initial Setup of LVCs	6-6
Service Template Qbins	6-6
MPLS CoS over IP+ATM Operation	6-8

- Configuration Example **6-9**
  - BPX Configurations **6-10**
    - BPX1 **6-10**
    - BPX2 **6-10**
  - LSC Configurations **6-11**
    - LSC1 **6-11**
    - LSC2 **6-11**
  - Edge LSR Configurations **6-12**
    - LSR1 **6-12**
    - LSR2 **6-13**
    - BPX1/BPX1 **6-13**
    - LSC1 and LSC2 **6-13**
    - LSR1 and LSR2 **6-14**

CHAPTER 7

**MPLS VPNS with BPX 8650 7-1**

- Introduction: MPLS-Enabled VPNS **7-1**
  - MPLS Labeling Criteria **7-3**
  - Quality of Service **7-3**
  - Security **7-4**
  - Manageability **7-5**
  - Scalability **7-5**
- MPLS VPNS over IP+ATM Backbones **7-5**
  - Built-In VPN Visibility **7-6**
  - BGP Protocol **7-7**
  - Virtual Routing/Forwarding **7-10**
  - VPN Route-Target Communities **7-10**
  - IBGP Distribution of VPN Routing Information **7-10**
  - Label Forwarding **7-11**
- Configuration Example **7-11**
  - Configuring the BPX 8650 ATM LSR **7-12**
  - Configuring VRFs **7-12**
  - Configuring BGP's **7-13**
  - Configuring Import and Export Routes **7-13**
  - Verifying VPN Operation **7-14**
  - Configuration Example **7-15**
- Command List **7-16**

---

**CHAPTER 8**

<b>MPLS Redundancy for IP+ATM Networks</b>	<b>8-1</b>
What Is LSC Redundancy	8-1
Benefits of LSC Redundancy	8-2
LSC Redundancy Allows Different Software Versions	8-2
LSC Redundancy Does Not Use Shared States or Databases	8-3
LSC Redundancy Lets You Use Different Hardware	8-3
LSC Redundancy Provides An Easy Migration from Stand-alone LSCs to Redundant LSCs	8-3
LSC Redundancy Allows Configuration Changes in a Live Network	8-3
LSC Redundancy Provides Fast Reroute in IP+ATM Networks	8-3
LSC Redundancy Architecture	8-4
Operational Modes	8-5
LSC Hot Redundancy	8-5
How the LSC, ATM Switch, and VSI Work Together	8-7
Implementing LSC Redundancy	8-8
Partitioning the Resources of the ATM Switch	8-8
Implementing the Parallel VSI Model	8-9
Adding Interface Redundancy	8-9
Implementing Hot LSC Redundancy	8-10
Sample LSC Redundancy Configuration	8-11
Connections to BPX1	8-12
Connections to BPX2	8-12
BPX1 Resource Parameter Settings	8-12
LER1 Configuration File	8-16
LSC1 Configuration File	8-17
LSC2 Configuration File	8-18

---

**GLOSSARY**

---

**INDEX**





<i>Figure 1-1</i>	Typical MPLS Network Structure	<b>1-6</b>
<i>Figure 2-1</i>	IP over ATM	<b>2-2</b>
<i>Figure 2-2</i>	Structural Elements of IP+ATM Switches	<b>2-4</b>
<i>Figure 2-3</i>	An IP+ATM Multiservice Network	<b>2-7</b>
<i>Figure 2-4</i>	Label Forwarding Information Base in An IP Packet Environment	<b>2-8</b>
<i>Figure 2-5</i>	Downstream Label Allocation	<b>2-9</b>
<i>Figure 2-6</i>	MPLS Elements in An ATM Network	<b>2-10</b>
<i>Figure 2-7</i>	Label Forwarding Information Base in An ATM Environment	<b>2-11</b>
<i>Figure 2-8</i>	Downstream On-Demand Label Allocation, Ordered Mode	<b>2-13</b>
<i>Figure 2-9</i>	Problem of Cell Interleave	<b>2-14</b>
<i>Figure 2-10</i>	VC Merge	<b>2-15</b>
<i>Figure 2-11</i>	Interconnecting ATM Label Switch Routers	<b>2-16</b>
<i>Figure 2-12</i>	Label Switch Controller Locations	<b>2-17</b>
<i>Figure 2-13</i>	Connecting a BPX 8650 and Label Switch Controller	<b>2-18</b>
<i>Figure 2-14</i>	Comparing MPLS, PNNI, and IP+ATM Switches	<b>2-19</b>
<i>Figure 2-15</i>	Comparing MPLS, PNNI, and IP+ATM Networks	<b>2-20</b>
<i>Figure 2-16</i>	An ATM MPLS Point of Presence (PoP)	<b>2-21</b>
<i>Figure 2-17</i>	An ATM MPLS PoP with Combined LSC and Edge Device	<b>2-22</b>
<i>Figure 2-18</i>	Using an Access Switch or Concentrator in An ATM MPLS PoP	<b>2-22</b>
<i>Figure 2-19</i>	MGX 8800 as An Integrated ATM MPLS PoP	<b>2-23</b>
<i>Figure 2-20</i>	Supporting IP+ATM Services Using Dual Backbones	<b>2-24</b>
<i>Figure 2-21</i>	Evolution of ATM MPLS Networks to Dual Backbones	<b>2-25</b>
<i>Figure 2-22</i>	Many Virtual Private Networks Provided by One Network	<b>2-26</b>
<i>Figure 2-23</i>	Providing Virtual Private Network Services Using An MPLS Network	<b>2-27</b>
<i>Figure 2-24</i>	Forwarding Packets in a Cisco MPLS Virtual Private Network Service	<b>2-28</b>
<i>Figure 2-25</i>	Control Functions in a Cisco MPLS Virtual Private Network Service	<b>2-29</b>
<i>Figure 2-26</i>	Management Operations: Adding a Site to a VPN	<b>2-32</b>
<i>Figure 2-27</i>	Migrating MPLS over a Traditional ATM Cloud	<b>2-34</b>
<i>Figure 3-1</i>	Typical MPLS Network Structure	<b>3-2</b>
<i>Figure 3-2</i>	Devices in MPLS Networks, Part One	<b>3-4</b>
<i>Figure 3-3</i>	Devices in MPLS Networks, Part Two	<b>3-5</b>

<i>Figure 3-4</i>	Point of Presence Structures for ATM MPLS Networks	<b>3-13</b>
<i>Figure 3-5</i>	Sample Network in Australia: PoP and Total Access Topologies	<b>3-17</b>
<i>Figure 3-6</i>	Sample Network in Australia	<b>3-20</b>
<i>Figure 3-7</i>	Network Design Example: Calculating Link Bandwidths	<b>3-22</b>
<i>Figure 3-8</i>	Viewpoints of An ATM MPLS Network	<b>3-24</b>
<i>Figure 3-9</i>	Routing Viewpoints in An ATM MPLS Network	<b>3-26</b>
<i>Figure 3-10</i>	Multiple Routing Areas and Summarization in An ATM MPLS Network	<b>3-28</b>
<i>Figure 3-11</i>	Label VC Requirements	<b>3-29</b>
<i>Figure 3-12</i>	Destination-Prefixes in An MPLS Network (or Any Other IP Network)	<b>3-30</b>
<i>Figure 3-13</i>	LVCs to Each Destination	<b>3-31</b>
<i>Figure 4-1</i>	How Connectionless Traffic Drives Meshing	<b>4-4</b>
<i>Figure 4-2</i>	Specifying Bandwidths for An IP Service	<b>4-5</b>
<i>Figure 4-3</i>	Cisco Committed Access Rate Policers	<b>4-7</b>
<i>Figure 4-4</i>	Using CAR on Customer Premises	<b>4-8</b>
<i>Figure 4-5</i>	Ensuring Access to Bandwidth Using Differentiated Services	<b>4-9</b>
<i>Figure 4-6</i>	Refining Estimates of Network Loads	<b>4-11</b>
<i>Figure 4-7</i>	Estimating Network Loads Per-Hop Behavior	<b>4-14</b>
<i>Figure 4-8</i>	Per-VC Service and Class of Service in ATM Switches	<b>4-15</b>
<i>Figure 4-9</i>	Per-VC Service with VC Merge	<b>4-16</b>
<i>Figure 4-10</i>	Committed Delivery in An IP Network	<b>4-18</b>
<i>Figure 4-11</i>	Reoptimization of Traffic Using MPLS Traffic Engineering	<b>4-24</b>
<i>Figure 4-12</i>	Reserved Point-to-Point Bandwidths in MPLS Networks	<b>4-26</b>
<i>Figure 4-13</i>	Quality of Service in Virtual Private Networks	<b>4-27</b>
<i>Figure 4-14</i>	Providing Bandwidth to Specific Users and Applications in Virtual Private Networks	<b>4-28</b>
<i>Figure 4-15</i>	Discard Policies	<b>4-30</b>
<i>Figure 4-16</i>	Example of Combining Weighted Fair Queueing and Differential Discards	<b>4-31</b>
<i>Figure 4-17</i>	Effects of Combining Weighted Fair Queueing and Differential Discards	<b>4-32</b>
<i>Figure 5-1</i>	High-Level View of Configuration of An MPLS Network	<b>5-4</b>
<i>Figure 5-2</i>	Label Swapping Detail	<b>5-5</b>
<i>Figure 5-3</i>	Simplified Example of Configuring An MPLS Network	<b>5-6</b>
<i>Figure 5-4</i>	Example of LVCs in An MPLS Switched Network	<b>5-17</b>
<i>Figure 6-1</i>	Multiple LVCs for IP QoS Services	<b>6-3</b>
<i>Figure 6-2</i>	Example of Multiple LVCs CoS with BPX 8650s	<b>6-5</b>
<i>Figure 6-3</i>	Service Template and Associated Qbin Selection	<b>6-7</b>
<i>Figure 6-4</i>	MPLS CoS over IP+ ATM with BPX 8650 LSRs	<b>6-8</b>

<i>Figure 6-5</i>	Configuration Example for MPLS CoS with BPX 8650 LSRs	<b>6-10</b>
<i>Figure 7-1</i>	VPN Network	<b>7-2</b>
<i>Figure 7-2</i>	Benefits of MPLS Labels	<b>7-3</b>
<i>Figure 7-3</i>	MPLS VPNs in Cisco IP+ATM Network	<b>7-6</b>
<i>Figure 7-4</i>	VPN-IP Address Format	<b>7-7</b>
<i>Figure 7-5</i>	VPN with Service Provider Backbone	<b>7-8</b>
<i>Figure 7-6</i>	Using MPLS to Build VPNs	<b>7-9</b>
<i>Figure 8-1</i>	LSC Redundancy with Physically Separate Trunks	<b>8-6</b>
<i>Figure 8-2</i>	LSC Redundancy with Shared Trunks	<b>8-7</b>
<i>Figure 8-3</i>	XtagATM Interfaces	<b>8-9</b>
<i>Figure 8-4</i>	Interface Redundancy	<b>8-10</b>
<i>Figure 8-5</i>	Topology for Sample Hot Redundancy Configuration	<b>8-11</b>







<i>Table 3-1</i>	Choosing MPLS Edge Equipment for ATM MPLS Networks	<b>3-7</b>
<i>Table 3-2</i>	Choosing ATM LSRs	<b>3-10</b>
<i>Table 3-3</i>	Network Example: Unidirectional Traffic Matrix	<b>3-18</b>
<i>Table 3-4</i>	Network Example: Approximate Bidirectional Traffic Flows	<b>3-18</b>
<i>Table 3-5</i>	Checking the LVC Limits of Edge LSR	<b>3-32</b>
<i>Table 3-6</i>	Cisco ATM Edge LSRs and LVC Capacity	<b>3-33</b>
<i>Table 3-7</i>	Checking the LVC Limits of ATM LSRs with VC Merge	<b>3-35</b>
<i>Table 3-8</i>	Cisco ATM LSRs and LVC Capacity, If VC Merge Is Used	<b>3-35</b>
<i>Table 3-9</i>	Checking the LVC Limits of ATM LSRs without VC Merge	<b>3-38</b>
<i>Table 3-10</i>	Cisco ATM LSRs and LVC Capacity, If VC Merge Is Not Used	<b>3-38</b>
<i>Table 5-1</i>	Cisco IOS Operating Modes	<b>5-34</b>
<i>Table 6-1</i>	CoS Services and Features	<b>6-2</b>
<i>Table 6-2</i>	Type of Service and Related CoS	<b>6-4</b>
<i>Table 6-3</i>	Class of Service and Relative Bandwidth Weighting	<b>6-9</b>
<i>Table 6-4</i>	Class of Service and Relative Bandwidth Weighting Setup	<b>6-9</b>





## Preface

---

Multiprotocol Label Switching (MPLS) is an improved method for forwarding packets through a network.

This guide:

- introduces the MPLS technology
- explains its benefits
- presents the foundations of MPLS network design
- provides specific MPLS configuration instructions for Cisco BPX 8600 series switches
- explains how to configure redundant MPLS switch controllers for added robustness

The intended audience is network administrators and technicians interested in a thorough introduction to label switching and Cisco's MPLS implementation. It is also intended for those performing initial BPX configuration for MPLS. Both the installers and the network administrator should be familiar with BPX network operation and modern WAN concepts.

## Documentation CD-ROM

Cisco documentation and additional literature are available in the CD-ROM package that ships with your product. Because the Documentation CD-ROM is updated monthly, it might be more current than printed documentation.

To order additional copies of the Documentation CD-ROM, contact your local sales representative or call Cisco Customer Service. The CD-ROM package is available as a single package or as an annual subscription.

You can also access Cisco documentation on the World Wide Web at:

<http://www.cisco.com>

<http://www-china.cisco.com>

<http://www-europe.cisco.com>.

If you are reading Cisco product documentation on the World Wide Web, you can submit comments electronically. Click **Feedback** in the toolbar and select **Documentation**. After you complete the form, click **Submit** to send it to Cisco. We appreciate your comments.

## Related Documentation

The following Cisco publications contain additional information related to the operation of the BPX switch and associated equipment in a Cisco WAN switching network:

<i>Cisco BPX 8600 Series Installation and Configuration</i> DOC-7810674=	Provides a general description and technical details of the BPX broadband switch.
<i>Cisco IGX 8400 Series Reference</i> DOC-7810706=	Provides a general description and technical details of the IGX multiband switch.
<i>Update to the Cisco IGX 8400 Series Reference Guide</i> DOC-78-11029=	Provides update information about new features in the 9.3.10 Switch Software release that apply to the IGX 8400 switch. Use this update document in conjunction with the Cisco IGX 8400 Series Reference, 9.3.05 Switch Software release documentation on the IGX 8400 switch.
<i>Cisco IGX 8400 Installation and Configuration</i> DOC-7810722=	Provides installation instructions for the IGX multiband switch.
<i>Update to the Cisco WAN Switching Command Reference Guide</i> DOC-7810703=	Provides update information about new features contained in the 9.3.10 Switch Software release that apply to both BPX and IGX switches documented in the WAN Switching Command Reference. Use this update document in conjunction with <i>Cisco WAN Switching Command Reference, Release 9.3.05</i> .
<i>Cisco WAN Switching Command Reference</i> DOC-7810703=	Provides detailed information on the general command line interface commands.
<i>Cisco WAN Switching SuperUser Command Reference</i> DOC-7810702=	Provides detailed information on the command line interface commands requiring SuperUser access authorization.
<i>Cisco MPLS Installation and Configuration</i> DOC-7810672=	Provides information on a method for forwarding packets through a network.
<i>WAN CiscoView for the IGX 8400 Switches</i> DOC-7810669=	Provides instructions for using WAN CiscoView for the IGX 8400.
<i>WAN CiscoView for the BPX 8600 Switches</i> DOC-7810670=	Provides instructions for using WAN CiscoView for the BPX 8600.
<i>Cisco WAN Manager Installation Guide for Solaris, Release 10</i> DOC-7810308=	Provides procedures for installing Release 10 of the Cisco WAN Manager (CWM) network management system on Solaris systems.
<i>Cisco WAN Manager User's Guide, Release 10</i> DOC-7810658=	Provides procedures for using Release 10 of the Cisco WAN Manager (CWM) network management system.

<i>Cisco WAN Manager SNMP Proxy Agent Guide</i> DOC-7810786=	Provides information about the Cisco WAN Manager Simple Network Management Protocol (SNMP) Service Agent components and capabilities.
<i>Cisco WAN Manager Database Interface Guide</i> DOC-7810785=	Provides the information to gain direct access to the Cisco WAN Manager Informix OnLine database that is used to store information about the elements within your network.

## Previous Cisco WAN Switch Product Names

The Cisco WAN Switching products were once known by older names

Old Name	New Name
Any switch in the BPX switch family (Cisco BPX <sup>®</sup> 8620 broadband switch and Cisco BPX <sup>®</sup> 8650 broadband switch)	A Cisco BPX <sup>®</sup> 8600 series broadband switch
The BPX Service Node switch	The Cisco BPX <sup>®</sup> 8620 broadband switch
The BPX switch as a tag-switched controller	The Cisco BPX <sup>®</sup> 8650 broadband switch
The AXIS shelf	The Cisco MGX <sup>™</sup> 8220 edge concentrator
Any switch in the IGX switch family (IGX 8, IGX 16, and IGX 32 wide-area switches)	The Cisco IGX <sup>™</sup> 8400 series multiband switch
The IGX 8 switch	The Cisco IGX <sup>™</sup> 8410 multiband switch
The IGX 16 switch	The Cisco IGX <sup>™</sup> 8430 multiband switch.
Cisco StrataView Plus <sup>®</sup>	Cisco WAN Manager <sup>®</sup> (CWM)

## MPLS and Tag Terminology

Multiprotocol Label Switching (MPLS) is a standardized version of Cisco's original Tag Switching proposal. MPLS and Tag Switching are identical in principle, and nearly identical in operation.


In this document, the term "label switching" and "MPLS" are used interchangeably.

This document uses Label Switching terminology rather than the, now obsolete, Tag Switching terminology. The following table shows the new and old terms.

An exception is the term "Tag Distribution Protocol." (TDP). TDP and the MPLS Label Distribution Protocol (LDP) are nearly identical in general function, but use different message formats and some different procedures.

The following table documents the change from tag switching terms to MPLS terms.

Old Designation	New Designation
Tag Switching	MPLS, Multiprotocol Label Switching
Tag (short for Tag Switching)	MPLS
Tag (item or packet)	Label

Old Designation	New Designation
TDP (Tag Distribution Protocol)	LDP (Label Distribution Protocol)
	 <p><b>Note</b> Cisco TDP and LDP (MPLS Label Distribution Protocol) are nearly identical in function, but use incompatible message formats and some different procedures. Cisco will be changing from TDP to a fully compliant LDP.</p>
Tag Switched	Label Switched
TFIB (Tag Forwarding Information Base)	LFIB (Label Forwarding Information Base)
TSR (Tag Switch Router)	LSR (Label Switch Router)
TSC (Tag Switch Controller)	LSC (Label Switch Controller)
ATM-TSR	ATM-LSR (ATM Label Switch Router, such as, BPX 8650)
TVC (Tag VC, Tag Virtual Circuit)	LVC (Label VC, Label Virtual Circuit)
TSP (Tag-Switched Path)	LSP (Label-Switched Path)
TCR (Tag Core Router)	LSR (Label Switching Router)
XTag ATM (extended Tag ATM port)	XmplsATM (extended MPLS ATM port)

## Terms Specific to MPLS

These terms are unique to discussions of MPLS technology:

Term	Definition
Edge Label Switch Router (LSR)	The term “Label Edge Router” is not used. The equivalent term “Edge LSR” is technically more correct.
ATM MPLS	“ATM MPLS” is the form of MPLS that runs in networks with ATM switches that do MPLS switching. More specifically, it is the form of MPLS where each different label on a link is represented by a different VC.
Packet-based MPLS	Packet-based MPLS means the form of MPLS that runs in networks that do not use ATM MPLS. More specifically, it is the form of MPLS where labels are carried as an extra header attached to each packet. Packet-based MPLS is also known as non-ATM MPLS, frame-based MPLS, and router-based MPLS. The term “Frame-based MPLS” is not used in this document, as it seems to imply Frame Relay, but packet-based MPLS does not necessarily have anything to do with Frame Relay.

Term	Definition
Packet-based LSR	A Packet-based LSR is a device that manipulates whole packets rather than cells. A router running packet-based MPLS is a packet-based LSR. An ATM Edge LSR is also a type of packet-based LSR.
Traditional ATM	Traditional ATM switches and networks do not use ATM MPLS. Traditional ATM networks may support packet-based MPLS traffic within Permanent Virtual Circuits (PVCs). A traditional ATM switch can support ATM MPLS within a Permanent Virtual Path (PVP), which acts a virtual trunk. In any case, the traditional ATM switches do not actually perform Multiprotocol Label Switching—though they might be used to support tunnels through which MPLS packets are carried.

## Conventions

Command descriptions use these conventions:

- Commands and keywords are in **boldface**.
- Arguments for which you supply values are in *italics*.
- Required command arguments are inside angle brackets (< >).
- Optional command arguments are in square brackets ([ ]).
- Alternative keywords are separated by vertical bars ( | ).

Examples use these conventions:

- Terminal sessions and information the system displays are in `screen font`.
- Information you enter is in **boldface screen font**.
- Nonprinting characters, such as passwords, are in angle brackets (< >).
- Default responses to system prompts are in square brackets ([ ]).



### Note

Means you should *take note*. Notes contain important suggestions or references to materials not contained in the current body of text.



### Caution

Means *reader be careful*. In this situation, you might do something that could result in equipment damage or loss of data.

# Cisco Connection Online

Cisco Connection Online (CCO) is Cisco Systems' primary, real-time support channel. Maintenance customers and partners can self-register on CCO to obtain additional information and services.

Available 24 hours a day, 7 days a week, CCO provides a wealth of standard and value-added services to Cisco's customers and business partners. CCO services include product information, product documentation, software updates, release notes, technical tips, the Bug Navigator, configuration notes, brochures, descriptions of service offerings, and download access to public and authorized files.

CCO serves a wide variety of users through two interfaces that are updated and enhanced simultaneously: a character-based version and a multimedia version that resides on the World Wide Web (WWW). The character-based CCO supports Zmodem, Kermit, Xmodem, FTP, and Internet e-mail, and it is excellent for quick access to information over lower bandwidths. The WWW version of CCO provides richly formatted documents with photographs, figures, graphics, and video, as well as hyperlinks to related information.

You can access CCO in the following ways:

- WWW: <http://www.cisco.com>
- WWW: <http://www-europe.cisco.com>
- WWW: <http://www-china.cisco.com>
- Telnet: [cco.cisco.com](http://cco.cisco.com)
- Modem: From North America, 408 526-8070; from Europe, 33 1 64 46 40 82. Use the following terminal settings: VT100 emulation; databits: 8; parity: none; stop bits: 1; and connection rates up to 28.8 kbps.

For a copy of CCO's Frequently Asked Questions (FAQ), contact:  
[cco-help@cisco.com](mailto:cco-help@cisco.com).

For additional information, contact:  
[cco-team@cisco.com](mailto:cco-team@cisco.com).

If you are a network administrator and need personal technical assistance with a Cisco product that is under warranty or covered by a maintenance contract, contact Cisco's Technical Assistance Center (TAC) at:  
800 553-2447  
408 526-7209, or  
[tac@cisco.com](mailto:tac@cisco.com).

To obtain general information about Cisco Systems, Cisco products, or upgrades, contact:  
800 553-6387  
408 526-7208, or  
[cs-rep@cisco.com](mailto:cs-rep@cisco.com).

## Documentation CD-ROM

Cisco documentation and additional literature are available in a CD-ROM package that ships with your product. The Documentation CD-ROM, a member of the Cisco Connection Family, is updated monthly and might be more current than printed documentation.

To order additional copies of the Documentation CD-ROM, contact your local sales representative or call customer service. The CD-ROM package is available as a single package or as an annual subscription. You can also access Cisco documentation on the World Wide Web at these sites:



- <http://www.cisco.com>
- <http://www-china.cisco.com>
- <http://www-europe.cisco.com>.

If you are reading Cisco product documentation on the World Wide Web, you can submit comments electronically. Click **Feedback** in the toolbar and select **Documentation**. After you complete the form, click **Submit** to send it to Cisco. We appreciate your comments.





# Introduction to MPLS

---

This chapter is an overview of Multiprotocol Label Switching (MPLS), highlighting MPLS in ATM networks and packet-based networks. It concentrates on the fundamentals of MPLS network design that apply to all ATM MPLS networks, including those supporting VPNs and traffic engineering.

- What is MPLS?
- Label Switching Features
- Label Switching Benefits
- MPLS Compared to Other IP-over-ATM Schemes
- MPLS Network Structure
- MPLS Applications
- MPLS Virtual Private Network
- References

## What is MPLS?

Multiprotocol Label Switching (MPLS) is a high-performance method for forwarding packets (frames) through a network. It enables routers at the edge of a network to apply simple labels to packets (frames). ATM switches or existing routers in the network core can switch packets according to the labels with minimal lookup overhead.

The BPX® 8650 is an IP+ATM switch that provides ATM-based broadband services and integrates Cisco IOS® software via Cisco 7200 series routers to deliver Multiprotocol Label Switching (MPLS) services.

MPLS integrates the performance and traffic management capabilities of Data Link Layer 2 with the scalability and flexibility of Network Layer 3 routing. It is applicable to networks using any Layer 2 switching, but has particular advantages when applied to ATM networks. It integrates IP routing with ATM switching to offer scalable IP-over-ATM networks.

In contrast to label switching, conventional Layer 3 IP routing is based on the exchange of network reachability information. As a packet traverses the network, each router extracts all the information relevant to forwarding from the Layer 3 header. This information is then used as an index for a routing table lookup to determine the packet's next hop. This is repeated at each router across a network. At each hop in the network, the optimal forwarding of a packet must be again determined.

Conventional IP packet forwarding has several limitations. It has limited capability to deal with addressing information beyond just the destination IP address carried on the packet. Because all traffic to the same IP destination-prefix is usually treated similarly, various difficulties arise. For example, it becomes difficult to perform traffic engineering on IP networks. Also, IP packet forwarding does not easily take into account extra addressing-related information such as Virtual Private Network membership.

The main concept of MPLS is to include a *label* on each packet.

Packets or cells are assigned short, fixed-length labels. Switching entities perform table lookups based on these simple labels to determine where data should be forwarded.

The label summarizes essential information about routing the packet:

- Destination
- Precedence
- Virtual Private Network membership
- Quality of Service (QoS) information from RSVP
- The route for the packet, as chosen by traffic engineering (TE)

With Label Switching the complete analysis of the Layer 3 header is performed only once: at the edge label switch router (LSR), which is located at each edge of the network. At this location, the Layer 3 header is mapped into a fixed-length label, called a label.

At each router across the network, only the label need be examined in the incoming cell or packet in order to send the cell or packet on its way across the network. At the other end of the network, an Edge LSR swaps the label out for the appropriate header data linked to that label.

A key result of this arrangement is that forwarding decisions based on some or all of these different sources of information can be achieved by means of a single table lookup from a fixed-length label. For this reason, label switching makes it feasible for routers and switches to make forwarding decisions based upon multiple destination addresses.

Label switching integrates switching and routing functions, combining the reachability information provided by the router function, plus the traffic engineering benefits achieved by the optimizing capabilities of switches. These benefits are described in more detail in the next section.

## Label Switching Features

MPLS, in conjunction with other standard technologies, offers many features critical for service providers:

- MPLS, in combination with the standard IP routing protocols OSPF or IS-IS, provides full, highly scalable support of IP routing within an ATM infrastructure.
- MPLS, in combination with the Border Gateway Protocol (BGP), provides support for highly scalable IP Virtual Private Network (VPN) services. IP VPN services are an invaluable development in provider networks, giving enterprise customers a service that meets their needs for private, connectionless delivery of IP services.
- Service-Level Agreements may be provided in a form suitable for connectionless traffic. Cisco networks assist the process of providing Service-Level Agreements by supporting MPLS in combination with forthcoming DiffServ standard. Along with supporting Virtual Private Networks, the ability to offer Service-Level Agreements suitable for IP traffic is a critical requirement to meet new demand for IP services.

Cisco IP+ATM networks fully support all relevant IP routing protocols and MPLS, while fully supporting traditional ATM services. MPLS and IP routing can readily be introduced to traditional ATM networks by using PVP or PVC tunnels, as MPLS-capable switches are continuously introduced.

Cisco IP+ATM switches allow carriers to continue to meet their existing demand for virtual circuit services while adding optimized support for critically important new services: IP and IP Virtual Private Networks. Furthermore, Cisco supports all of the standards relevant to carrier-class IP services: MPLS, the Multiprotocol Border Gateway Protocol, other standard routing protocols, and MPLS Traffic Engineering.

## Label Switching Benefits

MPLS offers many advantages over traditional IP-over-ATM.

When integrated with ATM switches, label switching takes advantage of switch hardware optimized to take advantage of the fixed length of ATM cells and to switch the cells at high speeds. For multiservice networks, label switching enables the BPX switch to provide ATM, Frame Relay, IP Internet service, and IP Virtual Private Network service all on a single platform in a highly scalable way. Support of all these services on a common platform provides operational cost savings and simplifies provisioning for multiservice providers.

For Internet service providers (ISPs) using ATM switches at the core of their networks, label switching enables the Cisco BPX 8600 series, the 8540 Multiservice Switch Router, and other Cisco ATM switches to provide a more scalable and manageable networking solution than overlaying IP over an ATM network. Label switching avoids the scalability problem of too many router peers and provides support for a hierarchical structure within an ISPs network.

These MPLS benefits are analyzed in greater detail:

- **Integration**

When applied to ATM, MPLS integrates IP and ATM functionality rather than overlaying IP on ATM. This makes the ATM infrastructure visible to IP routing and removes the need for approximate mappings between IP and ATM features. MPLS does not need ATM addressing and routing techniques such as PNNI, although these can be used in parallel if required.

- **Higher Reliability**

In Wide Area Networks (WANs) with ATM infrastructures, MPLS is an easy solution for integrating routed protocols with ATM. Traditional IP over ATM involves setting up a mesh of Permanent Virtual Circuits (PVCs) between routers around an ATM cloud, and the Next Hop Resolution Protocol (NHRP) achieves a similar result with switched virtual circuits (SVCs). But there are a number of problems with this approach, all arising from the method that the PVC links between routers are overlaid on the ATM network. This makes the ATM network structure invisible to the routers. A single ATM link failure could make several router-to-router links fail, creating problems with large amounts of routing update traffic and subsequent processing. (See Problems of Running IP Routing over An ATM Network without MPLS, page 1-5)

- **Better Efficiency**

Without extensive tuning of routing weights, all PVCs are seen by IP routing as single-hop paths with the same cost. This might lead to inefficient routing in the ATM network.

- **Direct Classes of Service Implementation**

When used with ATM hardware, MPLS makes use of the ATM queueing and buffering capabilities to provide different Classes of Service (CoS). This allows direct support of IP Precedence and CoS on ATM switches without complex translations to the ATM Forum Service Classes.

- VPN Scalability and Manageability**  
 MPLS can make IP Virtual Private Network services highly scalable and very easy to manage. Virtual Private Network services are an important service for providing enterprises with private IP networks within their infrastructures. When an ISP offers a VPN service, the carrier supports many individual VPNs on a single infrastructure. With an MPLS backbone, VPN information can be processed only at the ingress and exit points, with MPLS labels carrying packets across a shared backbone to their correct exit point. In addition to MPLS, the Multiprotocol Border Gateway Protocol (BGP) is used to deal with information about the VPNs. The combination of MPLS and Multiprotocol BGP makes MPLS-based VPN services easier to manage, with straightforward operations to manage VPN sites and VPN membership. It also makes MPLS-based VPN services extremely scalable, with one network able to support hundreds of thousands of VPNs.
- Reduces Control Load on Network Cores; More Robust**  
 VPN services demonstrate how MPLS supports a hierarchy of routing knowledge. Additionally, you can isolate Internet routing tables from service provider network cores. Like VPN data, MPLS allows access to the Internet routing table only at the ingress and exit points of a service provider network. With MPLS, transit traffic entering at the edge of the provider's autonomous system can be given labels that are associated with specific exit points. As a result, internal transit routers and switches need only process the connectivity with the provider's edge routers, shielding the core devices from the overwhelming route signaling volume exchanged in the Internet. This separation of interior routes from full Internet routes also provides better fault isolation and improved stability.
- Traffic Engineering Capabilities**  
 Other benefits of MPLS include traffic engineering (TE) capabilities needed for the efficient use of network resources. Traffic engineering enables you to shift the traffic load from overutilized portions to underutilized portions of the network, according to traffic destination, traffic type, traffic load, time of day, and so on.

## MPLS Compared to Other IP-over-ATM Schemes

In ATM networks, MPLS allows ATM switches to directly support IP services, giving maximum efficiency compared to other approaches. Traditional IP-over-ATM connects routers over Permanent Virtual Circuits (PVC).

Cisco also supports an alternative IP-over-ATM scheme called Multiprotocol over ATM (MPOA), which uses the Next Hop Resolution Protocol (NHRP). Unlike MPLS, MPOA overlays IP-over-ATM rather than fully integrating them. Although they do not share many of the advantages of MPLS in the WAN, MPOA and NHRP are cost-effective technologies for interconnecting nearby emulated LANs (ELANs) at high speeds. MPOA and similar proprietary approaches carry IP traffic over Switched Virtual Circuits (SVC). Traditional IP over ATM, MPOA, and proprietary approaches all have similar disadvantages:

- It is difficult to offer some types of IP services on the networks. For example, IP Class of Service cannot be offered natively by traditional ATM switches, and must be offered by translation to quite different ATM Forum Quality of Service concepts.
- Where IP services are offered, they are difficult to administer. Two levels of routing must be administered: IP routing (via OSPF or EIGRP or similar) and PNNI or similar routing for ATM. MPOA requires additional administration. Service translations, for example IP Class of Service to ATM Quality of Service, also require administration.
- IP services can be quite inefficient over ATM networks. For example, IP Multicast over ATM networks is difficult to achieve on a large scale due to the interaction of multicast routing, multicast group membership processing and ATM VC maintenance.

- There can be scaling limitations and/or dangerous interactions between IP routing (OSPF, and so on) and the ATM network, leading to unstable networks. Traditional IP over ATM can lead to storms of IP routing updates and subsequent network meltdown, if more than 30 OSPF routers are connected in a full mesh over PVCs. MPOA is unsafe when connecting routers to each other, and is intended only to connect hosts to routers or hosts to hosts. (See below.)
- IP services require a substantial implementation and management effort. For example, an MPOA implementation requires PNNI, SVC signaling, ATM ARP, an ATM ARP server, NHRP, and a NHRP server, in addition to AAL5, IP routing (OSPF, and so on) and an IPv4 stack.

MPLS in ATM networks avoid all of these disadvantages.

## Problems of Running IP Routing over An ATM Network without MPLS

If  $N$  number of routers are running OSPF and are connected in a full mesh over ATM PVCs, a single physical ATM link failure may result in ATM-layer rerouting of a large number of PVCs. If this takes too long, or if the ATM network cannot reroute PVCs at all, a large number of PVCs effectively fails.

The number of PVCs involved may be of the same order magnitude as  $N$ , and even  $N^2$  in some cases. In any case, it is likely to be seen by  $O(N)$  routers, where “ $O(N)$ ” means “a number proportional to  $N$ ”. So, a single ATM link failure will cause each of  $O(N)$  routers to send a link state advertisement (LSA) of size (at least)  $O(N)$  to  $(N-1)$  neighbors. Thus a single event in the ATM network results in  $O(N^3)$  to  $O(N^4)$  traffic.

When a router receives an LSA, it must immediately recalculate its routing table because it must not forward packets based on old routing information. The processor load caused by a storm of routing updates might cause the routers to drop or not send keep-alive packets, which appears to the neighboring routers as further link failures. These lead to further LSAs being sent, which perpetuates the problem.

The net result is that a full mesh network can go persistently unstable after a single network event.

This critical failure occurs because the routers do not see the state of the ATM links and switches directly. IS-IS has somewhat better performance than OSPF in full mesh conditions because IS-IS has more sophisticated flooding capabilities (these capabilities, specifically the ability to pace flooding and block flooding on some interfaces, are also becoming available on OSPF). However this does not address the underlying problem.

The solution is to enable IP routing to directly see the state of ATM links, which is what is done by ATM MPLS.

MPLS also addresses a different problem that arises when the ATM network runs PNNI routing: the basic conflict between routing protocols. PNNI routing at the ATM layer can make decisions that conflict with OSPF or similar routing at the IP layer. These conflicting decisions can lead to persistent loops. (See the NHRP Protocol Applicability Statement, RFC2333, for more on this. Further investigation on router-to-router NHRP at the IETF revealed that router-to-router NHRP was not practical.)

The only reliable solution to this problem is to use the same routing protocol at the IP layer and ATM layer. This is exactly what MPLS does in ATM networks.

# MPLS Network Structure

A typical structure for Multiprotocol Label Switching networks used by providers (carriers or ISPs) is shown in Figure 1-1.

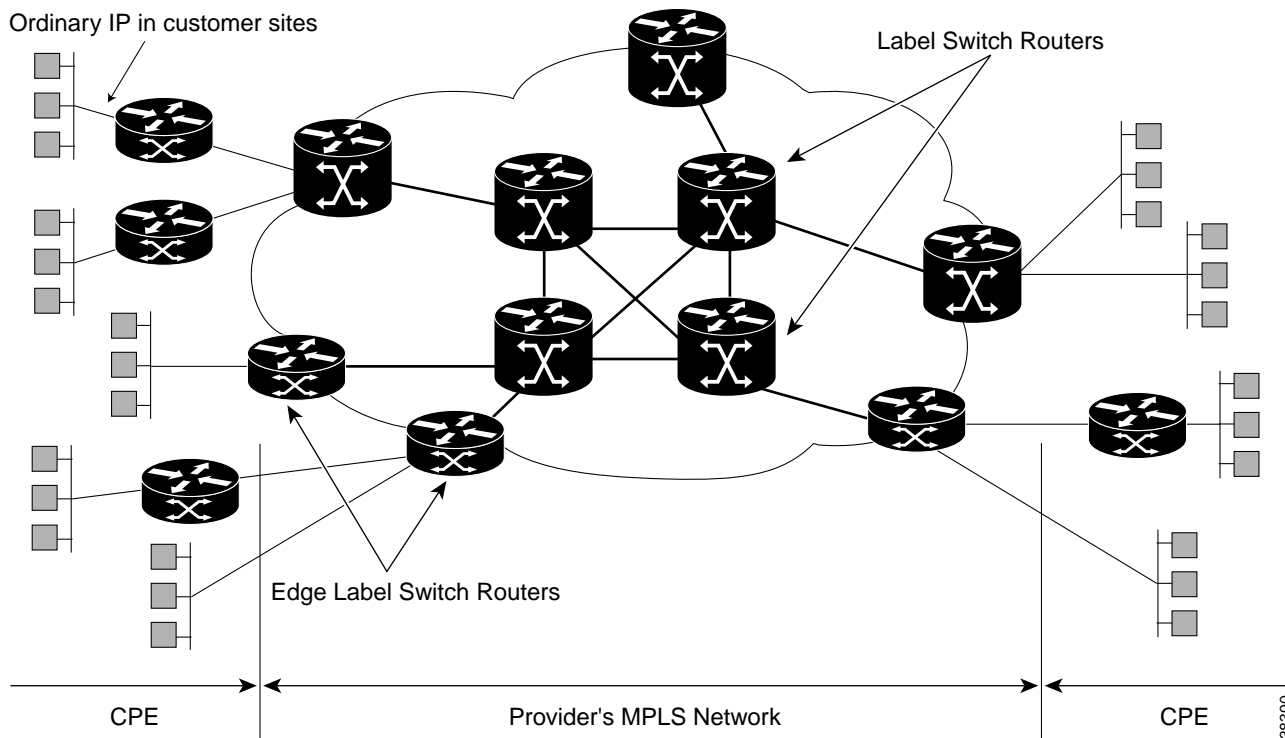
The basic elements in a label switching network are:

- Edge Label Switch Routers**  
 Edge Label Switch Routers are located at the boundaries of a network, performing value-added network layer services and applying labels to packets. These devices can be either routers, such as the Cisco 7500, or multilayer LAN switches, such as the Cisco Catalyst 5000.
- ATM Label Switch Routers**  
 These devices switch labeled packets or cells based on the labels. ATM Label Switch Routers may also support full Layer 3 routing or Layer 2 switching in addition to label switching. Examples of ATM LSRs include the Cisco 6400, the Cisco 8540 Multiservice Switch Router, Cisco BPX 8650, and Cisco 7500.
- Label Distribution Protocol**  
 The Label Distribution Protocol (LDP) is used in conjunction with standard network layer routing protocols to distribute label information between devices in a label switched network.

An MPLS network consists of Edge Label Switch Routers (Edge LSRs) around a core of Label Switch Routers (LSRs). Customer sites are connected to the provider MPLS network.

Typically there are several hundred customer sites per Edge LSR. The Customer Premises Equipment (CPE) runs ordinary IP forwarding but usually does not run MPLS. If the CPE does run MPLS, it uses it independently of the provider.

Figure 1-1 Typical MPLS Network Structure





It is important to note that the Edge LSRs are part of the provider network and are controlled by the provider. The Edge LSRs are critical to network operation and are not intended to be CPE under any circumstances. The provider may locate and manage routers at customer sites, but these are running ordinary IP and are outside the MPLS network.

## MPLS Applications

MPLS networks as shown in Figure 1-1 have three main applications. Typically, two or all three of these capabilities are used simultaneously:

- **IP+ATM Integration**

MPLS fully integrates IP services directly on ATM switches. The IP routing and LDP software resides directly on ATM switches. Thus MPLS allows ATM switches to optimally support IP multicast, IP Class of Service, RSVP, and Virtual Private Networks (see below). Optimal integration of IP+ATM means that MPLS is far more scalable and far less complex than overlay schemes like MPOA, CSI, and IP Navigator.

- **IP Virtual Private Network (VPN) Services**

A VPN service is the infrastructure of a managed intranet or extranet service offered by a provider to many corporate customers. These are often massive IP networks. MPLS, in combination with the Border Gateway Protocol (BGP), allows one provider network to support thousands of customer's VPNs. In this way, MPLS with BGP offers a very flexible, scalable, and manageable way of providing VPN services on both ATM and packet-based equipment. Even on small provider's networks, the flexibility and manageability of MPLS+BGP VPN services are a major benefit.

- **IP Explicit Routing and Traffic Engineering (TE)**

An important problem in current IP networks is the lack of ability to finely adjust IP traffic flows to make best use of available network bandwidth. Also absent are related capabilities to send selected flows down selected paths, for example, to select protected trunks for particular classes of traffic. MPLS uses Label Switched Paths (LSPs), a type of lightweight VC. These can be set up on both ATM and packet-based equipment. The IP Traffic Engineering capability of MPLS uses special LSPs to finely adjust IP traffic flows.

The next section summarizes label switching operations in various network services.

## MPLS Virtual Private Network

MPLS Virtual Private Networks (VPN) deliver enterprise-scale connectivity deployed on a shared infrastructure with the same policies enjoyed in a private network. A VPN can be built on the Internet or on a service provider's IP, Frame Relay, or ATM infrastructure. Businesses that run their intranets over a VPN service enjoy the same security, prioritization, reliability, and manageability as they do in their own private networks.

VPNs based on IP can extend intranets over wide-area links to remote offices, mobile users, and telecommuters. They can support extranets linking business partners, customers, and suppliers to provide better customer satisfaction and reduced manufacturing costs. VPNs can also connect communities of interest, providing a secure forum for common topics of discussion.

New IP-based services such as videoconferencing, packet telephony, distance learning, and information-rich applications offer businesses the promise of improved productivity at reduced costs. As these networked applications become more prevalent, businesses increasingly look to their service providers for intelligent services based on a rich set of controls that go beyond transport to optimize the

delivery of applications end to end. Today organizations want their applications to traverse a network in a secure, prioritized environment, and they want the opportunity to reduce costs, improve connectivity, and gain access to networking expertise.

## Intranet and Extranet VPNs

Intranet VPN services link employees, telecommuters, mobile workers, remote offices, and so on, to each other with the same privacy as a private network.

Extranet VPN services link suppliers, partners, customers, or communities of interest over a shared infrastructure with the same policies as a private network.

Cisco provides a range of ATM- and IP-based choices for deploying large-scale intranet and extranet VPN services, including Multiprotocol Label Switching (MPLS)-based services, which provide secure, business-quality VPN solutions that scale to support tens of thousands of VPN customers over IP or IP+ATM networks.

A VPN built with MPLS affords broad scalability and flexibility across any IP, IP+ATM, or multivendor backbone. MPLS forwards packets using labels. The VPN identifier in the label isolates traffic to a specific VPN. In contrast with IP tunnel and virtual-circuit architectures, MPLS-based VPNs enable connectionless routing within each VPN community. Service providers can easily scale their services to support tens of thousands of VPNs on the same infrastructure, with full QoS benefits across IP and ATM environments.

Cisco MPLS-based VPN solutions are supported on its IP+ATM WAN switch platforms including the BPX 8650 and MGX families, and on its high-end router platforms such as the Cisco 12000 series GSR.

## MPLS VPN Features

The VPN feature for MPLS Switching allows a Cisco IOS network to deploy scalable IPv4 Layer 3 VPN backbone services. MPLS Switching VPNs provide essential characteristics and features that service providers require to deploy scalable VPNs and build the foundation to deliver these value-added services:

### Performance

When MPLS VPNs are set up using ATM LSRs such as the BPX 8650, the capabilities of scalable connectionless service of IP are combined with the performance and traffic management capabilities of ATM.

### Connectionless Service

A significant technical advantage of MPLS VPNs is connectionless service. The Internet owes its success to its basic technology, TCP/IP, built on the packet-based, connectionless network paradigm. This means that no prior action is necessary to establish communication between hosts, making it easy for two parties to communicate.

To establish privacy in a connectionless IP environment, current VPN solutions impose a connection-oriented, point-to-point overlay on the network. Even if it runs over a connectionless network, today's VPN cannot take advantage of the ease of connectivity and multiple services available in connectionless networks.

By creating a connectionless MPLS VPN, tunnels and encryption are not required for network privacy, thus eliminating significant complexity.

## Centralized Service

Building VPNs in Layer 3 has the additional advantage of allowing delivery of targeted services to a group of users represented by a VPN.

A VPN must give service providers more than a mechanism for privately connecting users to intranet services. It must also provide a way to flexibly deliver value-added services to targeted customers. Scalability is critical, because customers want to use services privately in their intranets and extranets.

Because MPLS Switching VPNs are seen as private intranets, it's easy to leverage new IP services:

- multicast
- Quality of Service
- telephony support within a VPN
- centralized services such as content and Web hosting to a VPN

Now myriad combinations of specialized services can be customized for individual customers, for example, a service that combines IP multicast with a low-latency service class to enable videoconferencing within an intranet.

## Scalability

Scalability is the major deficiency of VPNs created using connection-oriented, point-to-point overlays, Frame Relay, or ATM VCs. Specifically, connection-oriented VPNs require a full  $N^2$  mesh of connections between customer sites to support any-to-any communication.

MPLS-based VPNs instead use the peer model and Layer 3 connectionless architecture to leverage a highly scalable VPN solution. The peer model requires a customer site to make peer connection with only one provider edge (PE) router as opposed to all other CPE or customer edge (CE) routers that are members of the VPN. The connectionless architecture allows the creation of VPNs in Layer 3, eliminating the need for tunnels or VCs.

Other scalability capabilities of MPLS Switching VPNs are due to the partitioning of VPN routes between PE routers and the further partitioning of VPN and IGP routes between PE routers and provider (P) routers in a core network. PE routers must maintain VPN routes for those VPNs who are members. P routers do not maintain any VPN routes. This increases the scalability of the providers core and insures that no one device is a scalability bottleneck.

## Security

MPLS Switching VPNs offer the same level of security as connection-oriented VPNs. Packets from one VPN will not inadvertently go to another VPN. Security is provided at the edge and core of a provider network:

- at the edge, security ensures that packets received from a customer are placed on the correct VPN
- at the backbone, VPN traffic is kept separate

Malicious spoofing of a provider edge (PE) router is nearly impossible because the packets received from customers are IP packets. These IP packets must be received on a particular interface or subinterface to be uniquely identified with a VPN label.

## Easy to Create

To take full advantage of VPNs, it must be easy to create new VPNs and user communities. Because MPLS VPNs are connectionless, no specific point-to-point connection maps or topologies are required.

Now it is easy to add sites to intranets and extranets and to easily form closed user groups. Managing VPNs in this manner enables membership of any given site in multiple VPNs, maximizing flexibility in building intranets and extranets.

## Flexible Addressing

To make a VPN service more accessible, users should be able to design their own addressing plan, independent of addressing plans for other VPN customers supported by a common service provider.

Many organizations use private address spaces, as defined in RFC 1918 today, and do not want to undertake the time and expense of implementing registered IP addresses to enable intranet connectivity. MPLS VPNs allow customers to continue to use their present address spaces without network address translation (NAT) by providing a public and private view of the address.

If two VPNs want to communicate and both have overlapping addresses, that communication requires NAT at one endpoint. This enables customers to use their own unregistered private addresses and communicate freely across a public IP network.

## Integrated Class of Service (CoS) Support

CoS is an essential ingredient of an IP VPN because it provides the ability to address two fundamental VPN requirements:

- predictable performance and policy implementation
- support for multiple Classes of Service in an MPLS Switching VPN

Network traffic is classified and labeled at the edge of the network before traffic is aggregated according to policies defined by subscribers and implemented by the provider and transported across the provider core. Traffic at the edge and core of the network can then be differentiated into different classes by drop probability or delay.

## Straightforward Migration

For service providers to quickly deploy these VPN services, a straightforward migration path is required. MPLS VPNs are unique because they can be built over multiple network architectures, including IP, ATM, Frame Relay, and hybrid networks.

Migration for the end customer is also simplified because there is no requirement to support MPLS on the customer edge (CE) router and no modifications are required to a customer's intranet.

## MPLS VPN Benefits

- A platform for rapid deployment of additional value-added IP services, including intranets, extranets, voice, multimedia, and network commerce
- Privacy and security equal to Layer 2 VPNs by limiting the distribution of a VPN's routes to only those routers that are members of the VPN
- Seamless integration with customer intranets
- Increased scalability over current VPN implementations, with thousands of sites per VPN and hundreds of thousands of VPNs per service provider
- IP Class of Service (CoS), with support for multiple Classes of Service and priorities within a VPN, as well as between VPNs
- Easy management of VPN membership and easy provisioning of new VPNs for rapid deployment

- Scalable any-to-any connectivity for extended intranets and extranets that encompass multiple businesses
- MPLS enables business IP services
  - VPNs with strong SLAs for QoS
  - privacy and QoS of ATM without tunneling or encryption
  - enabled by Cisco's unique combination of MPLS and open standards routing
- Lower operating costs
  - enables low-cost managed services to increase SP market share
  - increases profits though lower marginal cost for new services
  - network establishes VPN connectivity; no provisioning
  - build once/sell many; single routing image for all VPNs
- The first transport-independent VPN
  - universal VPN: one VPN, any access/transport: dial, xDSL, ATM, and so on
  - service delivery independent of transport/access technology
- Simpler to use
  - VPN managed by the service provider
  - transparent support for private IP addresses
  - multiple QoS service classes to implement business net policy
- Revenue and growth
  - revenue from today's transport services, growth from IP
- Business IP services enabled by MPLS/IOS
  - MPLS brings IOS to service provider ATM networks
  - MPLS is the new industry standard for bringing IP and ATM together
- Seamless service delivery
  - wide breadth of services; circuit emulation to IP VPNs
  - single pipe; multiple services (any service, any port)
- lower cost of operation and competitive advantages
  - ROI, TTM, economies of a multiservice network

## References

- The Cisco “IP+ATM Solutions” page at <http://www.cisco.com/go/ipatm> has links to press releases, brochures, white papers and other information. Use the links on the left-hand side of the page.
- The OSPF version 2 specification is <http://www.ietf.org/rfc/rfc2328.txt>
- The “IS-IS for Routing in TCP/IP and Dual Environments” specification is <http://www.ietf.org/rfc/rfc1195.txt>

- IETF documents on MPLS are at <http://www.ietf.org/html.charters/mpls-charter.html>. The most important documents are:
  - “MPLS Architecture” draft-ietf-mpls-arch-05.txt
  - “MPLS Label Stack Encodings” draft-ietf-mpls-label-encaps-04.txt
  - “MPLS using LDP and ATM VC Switching” draft-ietf-mpls-atm-02.txt
  - “LDP Specification” draft-ietf-mpls-ldp-05.txt
  - “MPLS Support of Differentiated Services by ATM LSRs and Frame Relay LSRs” draft-ietf-mpls-diff-ext-01.txt
- Other IETF documents on Differentiated Services are at <http://www.ietf.org/html.charters/diffserv-charter.html>
- The most important IETF documents on the Border Gateway Protocol are:
  - “A Border Gateway Protocol 4 (BGP-4)” <http://www.ietf.org/rfc/rfc1771.txt>
  - “Multiprotocol Extensions for BGP-4” <http://www.ietf.org/rfc/rfc2283.txt>
  - A further informational document shows how BGP can be used to support VPNs: “BGP/MPLS VPNs,” RFC 2457, <http://www.ietf.org/rfc/rfc2547.txt>
- The following books on routing, MPLS and related topics are very useful:
  - Halabi, B., *Internet Routing Architectures*, Cisco Press, 1997.
  - Metz, C., *IP Switching Protocols and Architectures*, McGraw-Hill, 1999
  - Rekhter, et al., *Switching in IP Networks*, Morgan Kaufmann, 1998
- Useful magazine articles are:
  - Feldman, et al., “Evolution of Multiprotocol Label Switching,” *IEEE Communications Magazine*, Vol. 36, No. 5, May 1998
  - Metz, C., “Ingredients for Better Routing: Read the Label,” *IEEE Internet Computing*, Sept/Oct. 1998
- Archives on MPLS and related technologies:
  - <http://infonet.aist-nara.ac.jp/member/nori-d/mlr/>
  - <http://dcn.soongsil.ac.kr/~jinsuh/home-mpls.html>



## Integrating MPLS with IP and ATM

---

One of the most important applications of MPLS is in IP+ATM networks. “IP+ATM” is Cisco’s trade name for equipment that simultaneously supports traditional ATM services (PVCs, SVCs, SPVCs, PVPs, and so on) and optimized IP transport using MPLS.

These networks offer traditional ATM and Frame Relay services while providing optimized IP support using ATM MPLS. MPLS also brings important new services, such as IP Virtual Private Networks, to both IP+ATM networks and router networks.

This chapter explains how MPLS integrates IP into ATM networks:

- Why Integrate IP with ATM?
- Structure of An IP+ATM Switch
- Routing on ATM Switches
- Building Internets on ATM
- MPLS Elements in An ATM WAN
- Label Switch Controllers
- An ATM MPLS Point of Presence
- Dual Backbones: Traditional ATM and ATM MPS or Packet-Over-SONET
- Virtual Private Networks
- Migrating MPLS into a Traditional ATM Network

### Why Integrate IP with ATM?

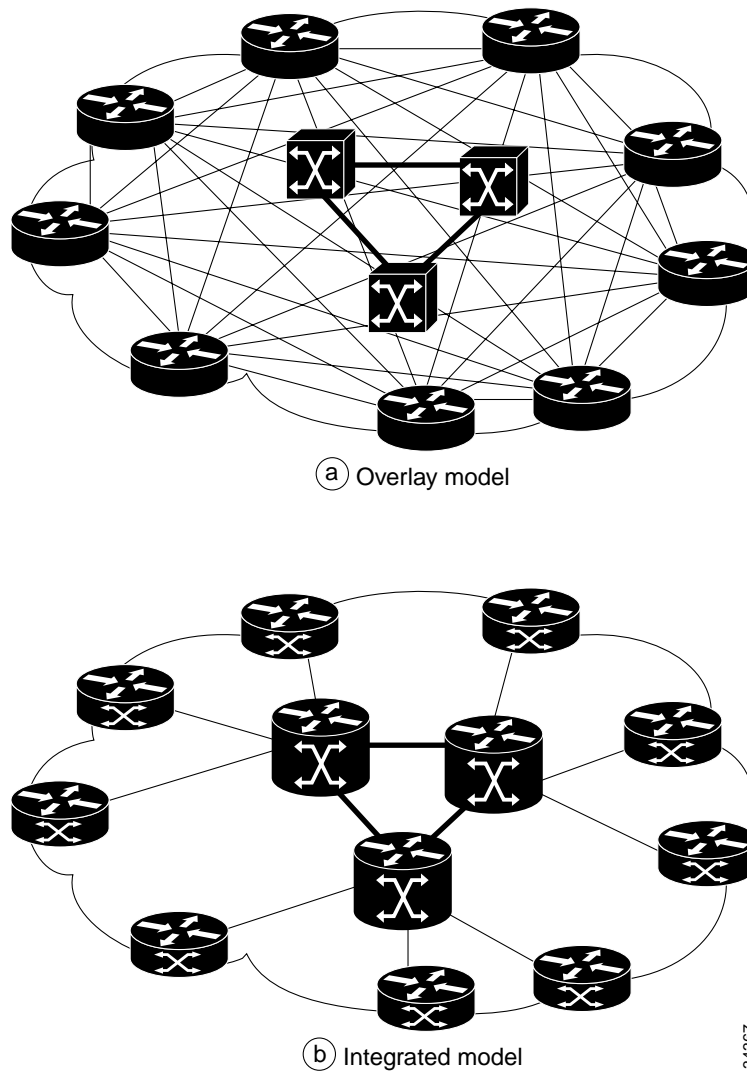
Today IP routing protocols typically run on top of ATM or Frame Relay with little integration. ISPs, for example, build ATM or Frame Relay cores inside their routed networks; these cores are used to build pipes between the routed edges.

In other words, two IP-routed networks are connected together using Permanent Virtual Circuits (PVCs) across an ATM or Frame Relay cloud. This creates an overlay model that is neither scalable nor manageable (Figure 2-1, Topology a), primarily because all routers on the cloud become IP neighbors.

This method also uses network resources inefficiently because the ATM links are invisible to IP routing. This means, for example, that a PVC using many hops will be used by IP routing just as readily as a single-hop PVC, because both PVCs are each a single IP hop.

Another problem with traditional networks results from routing protocols, such as OSPF, that do not perform well on large, fully meshed clouds due to the link state update duplication and the large number of neighbor state machines that have to be maintained. The route oscillation caused by circuit failures can exceed router CPU use and cause an indeterministic route convergence behavior. Experience has shown that this becomes a problem with a full mesh larger than 20 routers.

Figure 2-1 IP over ATM



MPLS solves the meshing problem by eliminating the notion of an ATM cloud. With MPLS, the ATM links are treated as IP links and each ATM switch can become an IP routing peer as in the integrate model, Figure 2-1, Topology (b).

By implementing IP intelligence into the ATM switches, designers eliminate the overlay of IP links on ATM and make a one-to-one mapping between them. This resolves most IP scalability problems.



In addition, this integration of the layers results in a distributed routing/switching model that takes advantage of the wealth of capabilities offered in each layer:

- The router part is needed to make use of the routing algorithms such as OSPF and BGP4 for exchanging reachability information and calculating paths.
- The MPLS part is needed to translate that reachability information into elements that can be understood by the switches.

The switching part uses advanced hardware capabilities to switch data at wire speed.

## Structure of An IP+ATM Switch

The concept of an IP+ATM switch is shown in Figure 2-2. A single switch contains two logically separate switches:

- An MPLS ATM label switch router (LSR) optimized for IP transport
- A traditional ATM PVC/SVC switch

Each trunk can support both PVCs (or SVCs, and so on), and MPLS Label VCs (LVCs).

Although an IP+ATM switch contains logically separate switches, it physically contains only one switch. However it contains two (or more) separate sets of control software. One set of software controls PVCs, SVCs, and so on, and the other control software controls MPLS. These controllers act independently, allowing the single physical switch to act as two (or more) virtual switches.

In switches such as the BPX 8650 and MGX 8850, this independent control is implemented by using a the Virtual Switch Interface (VSI). The VSI allows two or more separate controllers to independently control a single switch, as shown in Figure 2-2 Topology (b).

The MPLS control software is physically located in a label switch controller (LSC).

In the BPX 8650, the LSC is a device separate from the main switch shelf.

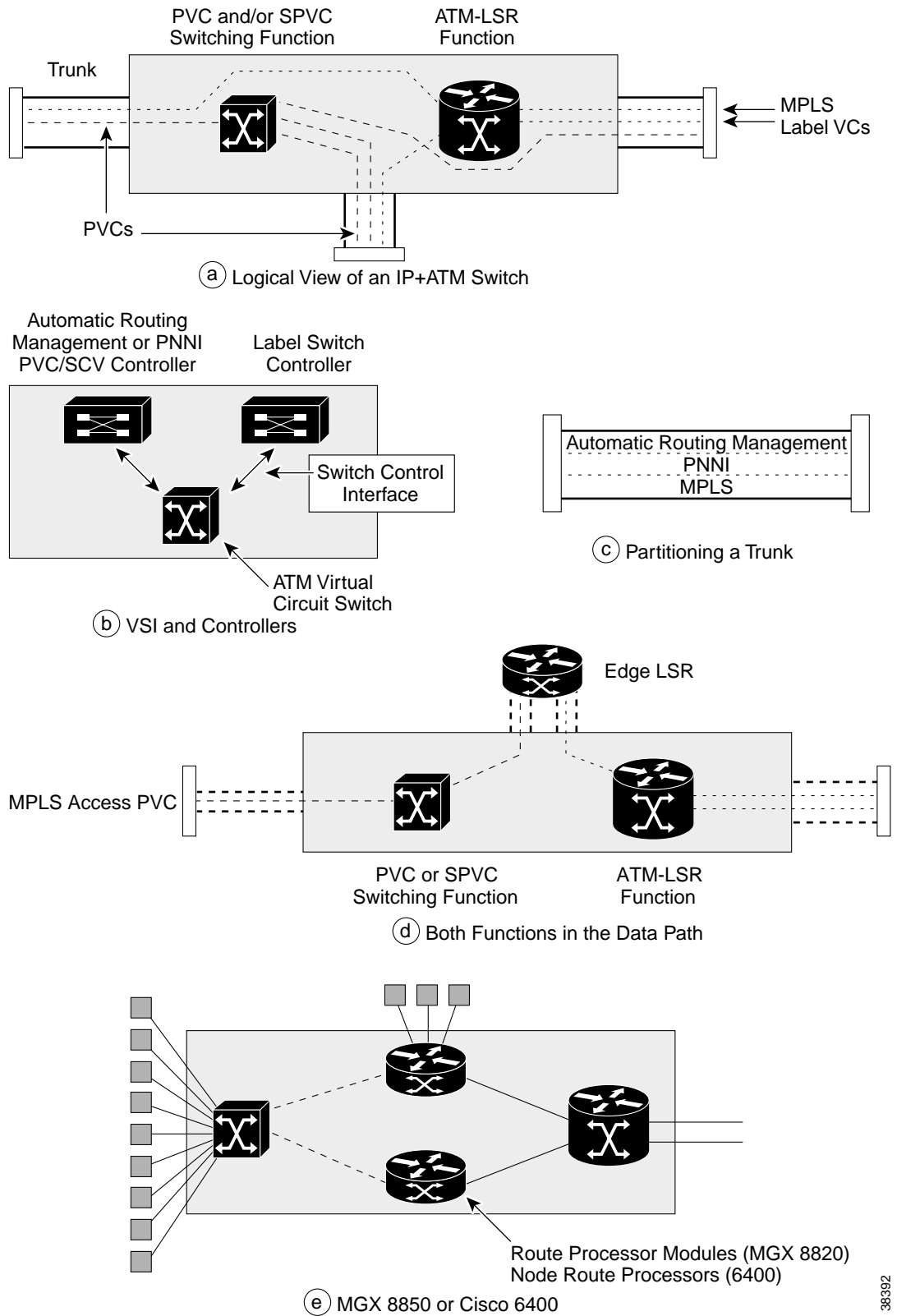
In the MGX 8850, the LSC is based on a Route Processor Module (RPM) in the switch shelf itself.

Other VSI controllers may be software running on the switch control card. In the case of the BPX 8650 and MGX 8850, AutoRoute software, which controls PVCs, runs on the switch control card.

PNNI control may be added to the BPX 8650 as a separate controller on the Service Expansion Shelf (SES).

The LS1010 and 8540 MSR implement similar functionality to the VSI using internal software interfaces.

Figure 2-2 Structural Elements of IP+ATM Switches



38392

To ensure that the control planes can act independently, the VSI slave processes in the switch must allocate resources to the different control planes (MPLS or PNNI). In the BPX 8650, resources for AutoRoute PVCs are reserved in a similar way.

The resources partitioned in the different control planes are:

- **VPI/VCI space on trunks**  
Each control plane gets a range of VPIs to use.
- **Bandwidth**  
Each control plane is guaranteed a certain bandwidth for Connection Admission Control (CAC) purposes. With soft partitioning, there can be a pool of bandwidth shared between control planes for CAC purposes. Even with hard partitioning, spare bandwidth unused by a control plane is available on a cell-by-cell basis to other control planes.
- **Traffic queues**  
One of the keys to IP+ATM is that MPLS traffic gets different traffic queues on the switch than the PVC and SVC traffic. This means that MPLS traffic can be handled by queues that directly support the MPLS “Class of Service” concept. The alternative is manually configured translations to ATM forum service types. The need for these translations is one of the main disadvantages of IP-over-ATM schemes apart from MPLS, and IP+ATM avoids this disadvantage.

Part of the configuration process for IP+ATM switches is the assignment of these resources to the different control planes, which involves creating different “partitions” of link resources for the different control planes, as shown in Figure 2-2 Topology (c).

## Use of IP+ATM

IP+ATM can be used to offer MPLS services, along with PVC and SVC services, all on the same network. This means that all (or many) switches in the network act as both ATM LSRs and traditional ATM switches, as in Figure 2-2 Topology (a).

The traditional ATM services can also be used in conjunction with an MPLS service. Figure 2-2 Topology (d) shows the use of a PVC to connect ordinary IP traffic from customer site through to an ATM Edge LSR. A PVC used in this fashion is called an “MPLS Access PVC.”

Other PVCs are “traditional PVCs” as part of a traditional end-to-end PVC service. The traffic from the Edge LSR can then be fed back through the ATM LSR function in the same switch that supports the MPLS Access PVC, or alternatively through a different switch. In any case, the end-to-end data path for IP traffic can include both MPLS Access PVCs and MPLS Label VCs.

An integrated IP+ATM edge switch, such as the MGX 8850 or Cisco 6400, contains ATM LSR function, as well as traditional access switch and PVC switching function. In addition, the Edge LSR function is also integrated into the device. This shown in Topology (e).

In the MGX 8850, routing function is supported by Route Processor Modules (RPMs). Node Route Processor (NRP) modules are used in the Cisco 6400. Each RPM or NRP acts as an Edge LSR.

In the MGX 8850, one of the RPMs will simultaneously act as an LSC and an Edge LSR.

## Routing on ATM Switches

You can implement routing on ATM switches by either integrating the routing engine inside the switch or by using separate routing controllers (a router).

The integrated solution runs routing and MPLS software on the switch control processor. This is done on the Cisco LS1010, Catalyst 5500, and 8540 MSR ATM switches. The controller model makes use of a separate router that controls the switch hardware. This separate router is called a label switch controller (LSC). The LSC can be either a routing card in the switch shelf or an external router. The LSC will handle all the IP functionality and would interact with the switch via either the backplane (for a router card) or an external control interface. The first label switch controller offered by Cisco is an external controller for the BPX 8650 platform. The MGX 8800 will use an LSC running on the Route Processor Module (RPM) in the switch shelf.

From an outcome point of view, both the integrated and controller methods will deliver the same result—which is providing an intelligent way of integrating routing and ATM.

The differences lie in the new operational model of using separate controllers rather than integrated solutions. The controller model has the advantage of separating the services into separate logical entities, each having a roadmap that does not interfere with the other.

In other words, if an external router controls a BPX 8650 switch running PNNI and SVC services, you can perform an IP MPLS upgrade without disturbing the operation of the PNNI and SVC services. In the WAN space, this is an attractive functionality.

One other benefit is that the controller itself can be fully utilized for other purposes. In the case of the LSC + BPX 8650 model, the router controlling the BPX 8650 can be utilized as a regular packet-based router offering IP and other Layer 3 services and still controlling the ATM switch hardware.

On the other hand, the cost of an external controller, both in hardware and use of a port on the switch, might not be justified in all situations. Because of this, the integrated approach on the LS1010 may be preferable.

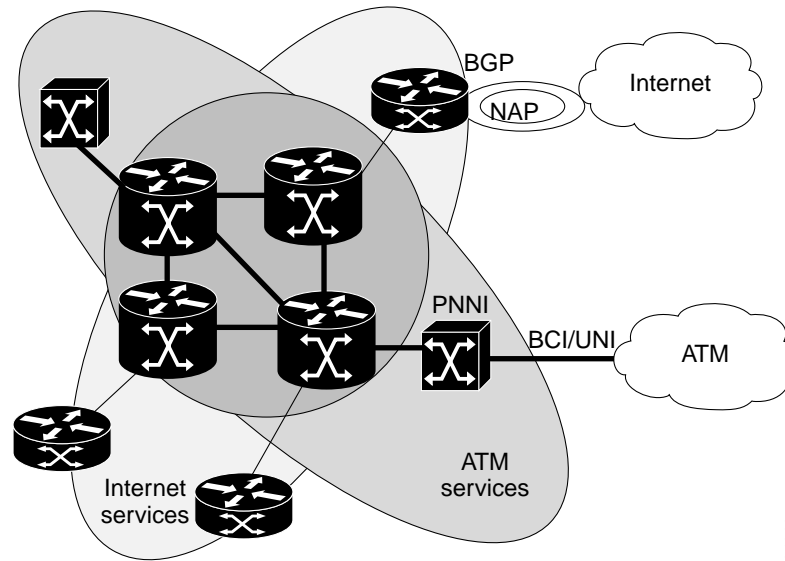
## Building Internets on ATM

The Internet is a collection of service providers offering IP services to their customers, all interconnected either directly or via high-speed network access points (NAPs). The NAPs are usually managed by a dedicated provider acting as a point of contact for coordination and connectivity purposes.

Each ISP maintains multiple Points of Presence (PoPs) that serve as concentration points for customer connectivity in multiple regions. PoPs can be interconnected via an ATM infrastructure or via direct high-speed, leased-line connections. Currently, ISPs use the Border Gateway Protocol version 4 (BGP4) for the purposes of interdomain connectivity. BGP4 offers a wide range of capabilities in segmenting providers' networks and offering routing policies that define the providers' administrative and political boundaries.

Integrating ATM infrastructures into the Internet model is as simple as providing an IP continuity between the ATM network and the rest of the IP world. IP is integrated over ATM inside the AS using MPLS, and the AS is connected to the rest of the Internet via BGP, as illustrated in Figure 2-3.

Figure 2-3 An IP+ATM Multiservice Network



This is done by using IP+ATM, where the ATM switches can continue to operate according to the ATM Forum and ITU-T standards while running MPLS in parallel. This means that other network applications such as PNNI, SVC, and AutoRoute can still operate independently of the MPLS application offering routed services.

Figure 2-3 shows an IP+ATM network offering both Internet and ATM services, illustrated by the two shaded areas. The VPI/VCI space on the multiservice switches (where the two shaded areas overlap) is divided between MPLS and the ATM Forum services. The MPLS network is connected to other ISPs via the Network Access Points (NAPs). The ATM network is running PNNI internally and is connected to other service providers.

Note that there is no interaction whatsoever between PNNI and MPLS except that they both share the ATM switch resources and link bandwidths. The MPLS implementations on the LS1010, 8540 MSR, BPX 8650 and MGX 8850 switches allow control of the allocation of resources between MPLS and PNNI.

The subsequent sections explain how MPLS operates on whole packets as opposed to ATM cells and then continue to define the signaling required to achieve MPLS over ATM

## Label Switching Operation at Layer 3

Label switching relies on two major components:

- **Forwarding**  
Forwarding uses the label information to perform packet forwarding.
- **Control**  
The control component maintains the correct label-forwarding information along with a group of interconnected label switches.

This section covers these elements as they are used in ordinary MPLS packet forwarding. Advanced services, including Traffic Engineering, Virtual Private Networks, and Class of Service, are considered later.

## Forwarding Component

The forwarding component is based on label swapping.

When a label switch router based on router hardware (such as a 7500 or 12000-series router) receives a packet with a label, the label is used as an index in a Label Forwarding Information Base (LFIB). Each entry in the LFIB consists of an incoming label and one or more subentries of the form:

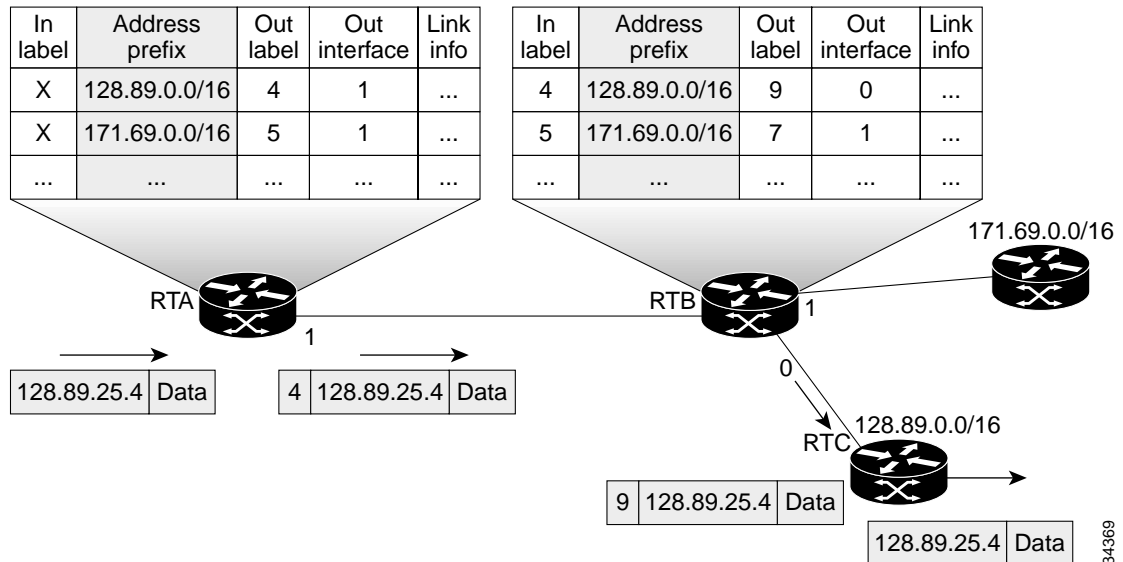
```
<outgoing label, outgoing interface, outgoing link level information>
```

For each sub-entry, the label switch replaces the incoming label with the outgoing label and sends the packet on its way over the outgoing interface with the corresponding link-level information.

Figure 2-4 shows an example of label switching:

1. An unlabeled IP packet with destination 128.89.25.4 arrives at Router A (RTA).
2. RTA checks its LFIB and matches the destination with prefix 128.89.0.0/16. (The /16 denotes 16 network masking bits per the Classless Interdomain Routing (CIDR) standard.)
3. The packet is labeled with an outgoing label of 4 and sent toward its next hop RTB.
4. RTB receives the packet with an incoming label of 4 that it uses as an index to the LFIB.
5. The incoming label of 4 is swapped with outgoing label 9.
6. The packet is sent out over interface 0 with the appropriate Layer 2 information (such as MAC address) according to the LFIB. (Note that RTB did not have to do any prefix IP lookup based on the destination as was done by RTA. Instead, RTB used the label information to do the label forwarding.)
7. When the packet arrives at RTC, it removes the label from the packet and forwards it as an unlabeled IP packet.

**Figure 2-4 Label Forwarding Information Base in An IP Packet Environment**



34369

## Control Component

The control component of MPLS consists of IP routing protocols (typically OSPF or IS-IS) running in conjunction with MPLS label allocation and maintenance procedures. The control component is responsible for setting up label forwarding paths along IP routes and then distributing these label bindings to the label switches. The control component will also maintain accuracy for the paths given topology changes that might occur.

The label distribution protocol (LDP) is a major part of the control component. LDP establishes peer sessions between label switches and exchanges the labels needed by the forwarding function.

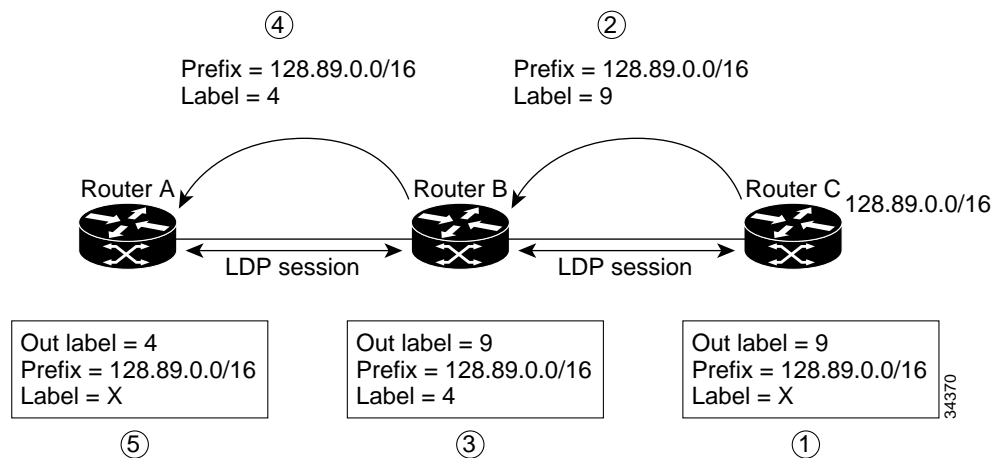
The OSPF or IS-IS routing protocol runs in the normal way, automatically creating forwarding tables in each MPLS label switch router. The MPLS Label Distribution Protocol (LDP) is linked to the routing protocols and works in parallel with them. Based on the routing information provided by OSPF or IS-IS, LDP exchanges the labels needed by the forwarding function.

In a packet environment, LDP is used in a downstream label allocation scheme (see Figure 2-5) which works like this:

1. For each route in its routing table, the label switch router allocates a label and creates an entry in its Label Forwarding Information Base (LFIB) with the incoming label set to the allocated label.
2. The label switch router then advertises the binding between the label (incoming) it created and the route to other adjacent label switch routers.
3. When a label switch router receives label binding information for a route and that information was originated by the next hop for that route, the switch places the label into the outgoing label of the LFIB entry associated with the route.

This creates the binding between the outgoing label and the route.

**Figure 2-5 Downstream Label Allocation**

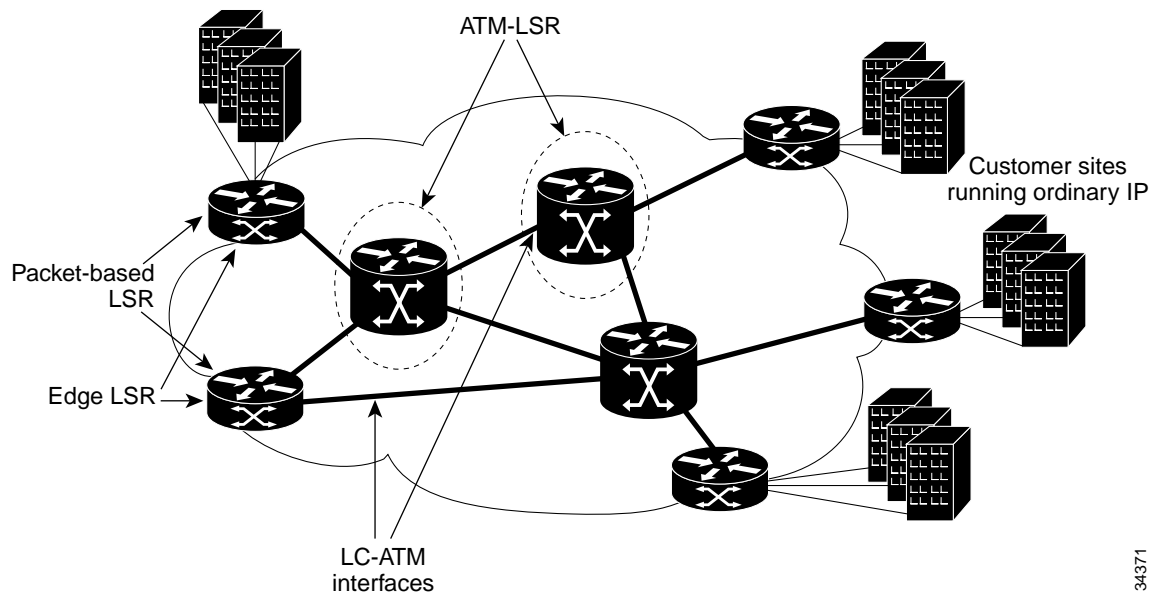


## MPLS Elements in An ATM WAN

This section defines multiple MPLS elements. Figure 2-6 illustrates these elements in a network environment:

- **Label Switch Router (LSR)**  
A device that implements the MPLS control and forwarding components.
- **Label-Controlled ATM (LC-ATM) interface**  
An ATM interface controlled by the MPLS control component. Cells traversing such an interface carry labels in the VCI field of a user-selected range of VPIs. The control component could be integrated in the switch or on an outside controller.
- **ATM-LSR**  
A LSR based on an ATM switch. It has LC-ATM interfaces.
- **Packet-based LSR**  
A LSR that forwards complete packets, between its interfaces. A packet-based LSR can have zero or more LC-ATM interfaces. Packet-based LSRs typically consist of MPLS software running on ordinary router platforms, such as the Cisco 3600, 4700, 6400, 7200, or 7500 series. Sometimes there are some hardware features specifically for MPLS, as on the Cisco 12000 series.
- **ATM Edge LSR**  
A packet-based LSR connected to the ATM-LSR cloud via LC-ATM interfaces. The ATM Edge LSR has the function of adding labels to unlabeled packets and stripping labels from labelled packets. Note that Edge LSRs are part of the same service provider network as the ATM-LSRs. Edge LSRs are not intended to be customer premises equipment or customer located equipment.

Figure 2-6 MPLS Elements in An ATM Network



34371



Similar to MPLS in a packet environment, MPLS in an ATM environment consists of a forwarding component and a control component:

- **Forwarding Component**

In an ATM environment, the label switching forwarding function is carried out identically to traditional switching. The label information needed for label switching can be carried in the VCI field within one or a small number of VPs. The labels act as the VCIs.

- **Control Component**

For the control component over ATM networks, a label distribution protocol is used to bind VCIs to IP routes. The switch also has to participate in IP routing protocols such as OSPF, BGP, and RSVP.

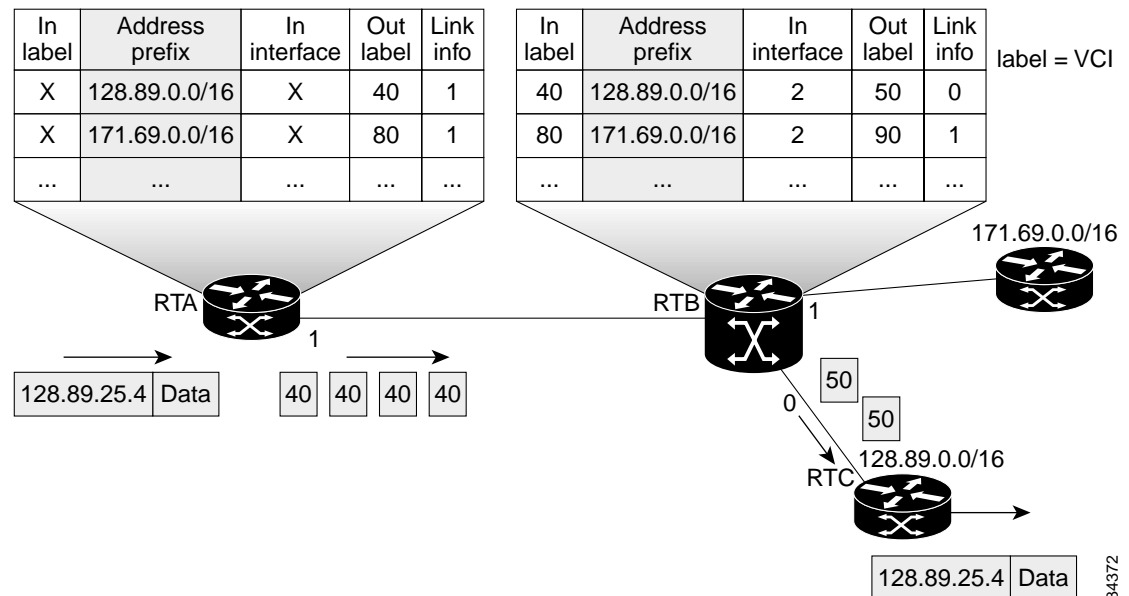
## Forwarding Via ATM Switches

Whereas the ATM forwarding operation is based upon switching cells by swapping VCIs and VPIs, in an ATM environment the MPLS forwarding function is done by the normal switch operation. The label information needed for MPLS can be carried in the VCI field within one or a small number of VPs.

Figure 2-7 shows the forwarding operation of an ATM switch in which the labels are designated VCIs:

1. An unlabeled IP packet with destination 128.89.25.4 arrives at RTA (edge label switch router A).
2. RTA checks its LFIB and matches the destination with prefix 128.89.0.0/16.
3. RTA converts the AAL5 frame to cells and sends the frame out as a sequence of cells on VCI 40.
4. RTB (an ATM LSR controlled by a label switch controller) performs a normal switching operation by switching incoming cells on interface 2/VCI 40 to interface 0/VCI 50.

Figure 2-7 Label Forwarding Information Base in An ATM Environment



## Control Via ATM Switches

The control component for MPLS over ATM is similar to router-based MPLS. A standard IP routing protocol such as OSPF or IS-IS runs in the network, alongside the Label Distribution Protocol (LDP). LDP is needed to bind VCIs to IP routes.

ATM-LSRs use the downstream-on-demand allocating mechanism. Each ATM-LSR maintains a forwarding information base (FIB) that contains a list of all IP routes that the ATM-LSR uses. This control function is handled by the routing engine function which is either embedded in the switch or runs on an outside controller.

For each route in its forwarding information base, the edge ATM LSR identifies the next hop for a route. It then issues a request via LDP to the next hop for a label binding for that route.

When the next hop ATM-LSR receives the route, it allocates a label, creates an entry in its LFIB with the incoming label changed to the allocated outgoing label.

The next action depends on whether the label allocation is in an *independent mode* or a *ordered mode*:

- **Independent Mode**

The ATM-LSR will immediately return the binding between the incoming label and the route to the LSR that sent the request. However, this could mean that it is not immediately able to forward labeled packets as they arrive, because the ATM-LSR might not yet have an outgoing label/VCI for the route.

The LSR that initiated the request receives the binding information, it creates an entry in its LFIB, and sets the outgoing label in the entry to the value received from the next hop. The next hop ATM LSR then repeats the process, sending a binding request to its next hop, and the process continues until all label bindings along the path are allocated.

- **Ordered Mode**

The ATM-LSR does not immediately return the binding, but waits until it has an outgoing label. The next hop LSR sends a new binding request to its next hop, and the process repeats until the destination ATM Edge LSR is reached. It then returns a label binding to the previous ATM-LSR, causing it to return a label binding, and so on until all the label bindings along the path are established.

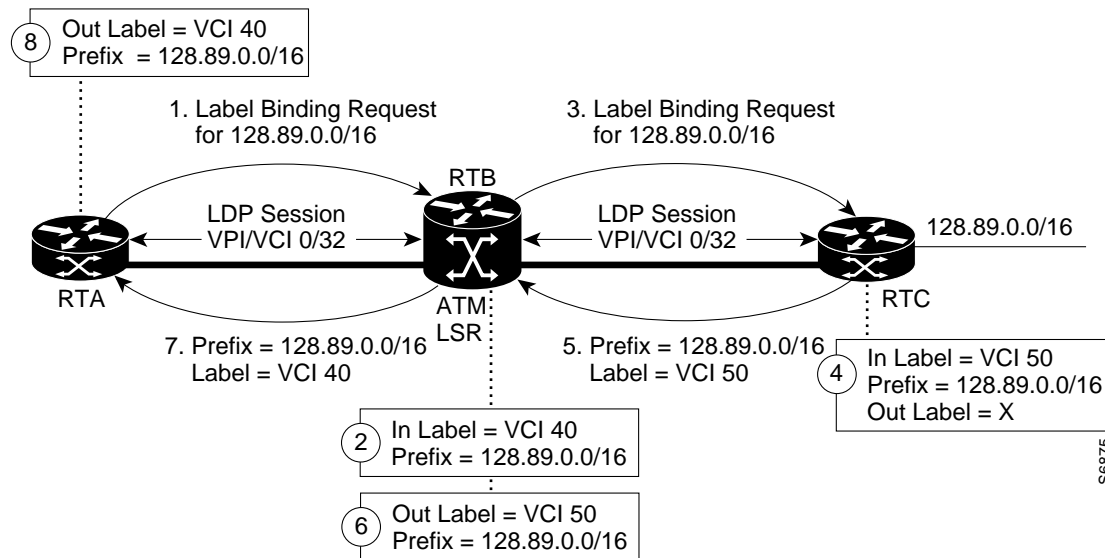
Figure 2-8 shows an ordered allocation. ATM Edge LSR RTA is an IP routing peer to ATM-LSR RTB. In turn, ATM-LSR RTB is an IP routing peer to ATM-LSR-RTC. IP routing updates are exchanged over VPI/VCI 0/32 between RTA-RTB and RTB-RTC.

For example:

1. RTA sends a label binding request toward RTB to bind prefix 128.89.0.0/16 to a specific VCI.
2. RTB allocates VCI 40 and creates an entry in its LFIB with VCI 40 as the incoming label.
3. RTB then sends a bind request toward RTC.
4. RTC issues VCI 50 as a label.
5. RTC sends a reply to RTB with the binding between prefix 128.89.0.0/16 and the VSI 50 label.
6. RTB sets the outgoing label to VCI 50.
7. RTB sends a reply to RTA with the binding between prefix 128.89.0.0/16 and the VCI 40 label.
8. RTA then creates an entry in its LFIB and sets the outgoing label to VCI 40.

Independent mode operation is similar to that shown in Figure 2-8, except that the events labeled 7 and 8 in the figure may occur concurrently with event 3.

Figure 2-8 Downstream On-Demand Label Allocation, Ordered Mode



MPLS networks can use traditional ATM equipment. They are usually built as a step toward introducing MPLS to an existing ATM network.

Traditional ATM switches can be used in three ways.

- Backhauling, when the access device is remote from the Edge LSR. The access device is connected to the Edge LSR by PVCs switched through an ATM network.
- Tunneling through ATM switches between an Edge LSR and an ATM LSR. The Edge LSR does not need to be adjacent to an ATM-LSR, but can be connected through an ATM network.
- Tunneling through ATM switches between ATM LSRs. The core network uses traditional ATM switches as well as ATM LSRs.

These uses of traditional ATM equipment have disadvantages and must be used with care.

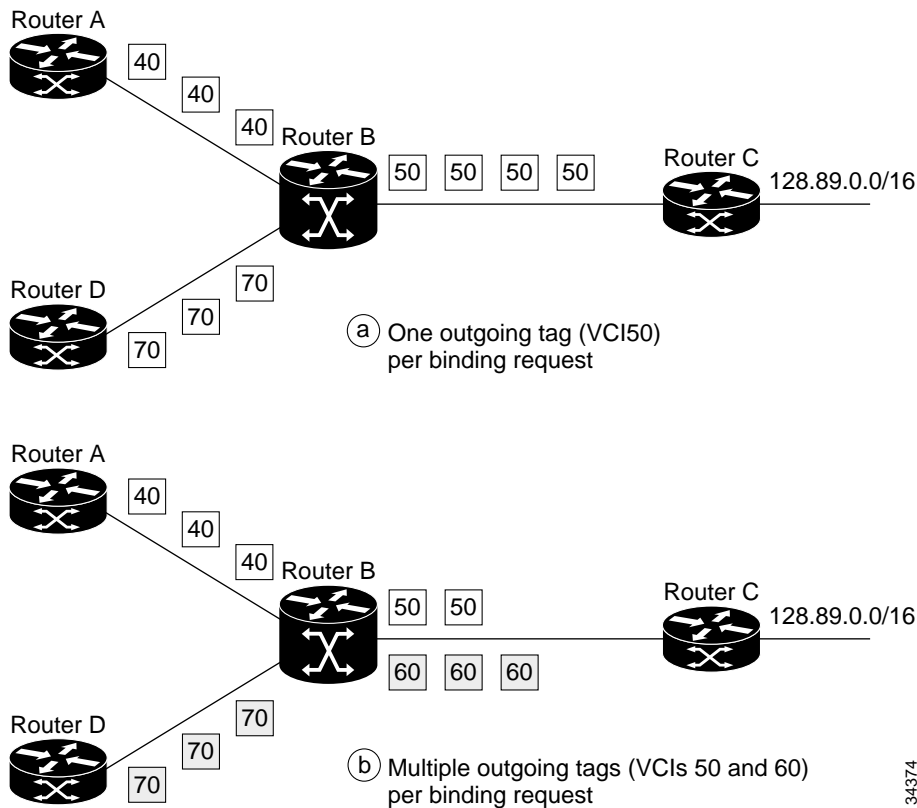
## Cell Interleave Problem

The problem of having multiple sources transmitting data to the same destination creates some challenges with label VC allocation over ATM.

An ATM-LSR receiving binding requests from different upstream neighbors toward the same prefix will have to request multiple outbound labels from its downstream neighbor. If the ATM-LSR were to allocate only one outgoing VCI, then cells from different AAL5 frames would potentially be interleaved and dropped at the receiving end.

Allocating different outbound VCIs for the same destination will ensure that cells will be received in order. This is illustrated in Figure 2-9.

Figure 2-9 Problem of Cell Interleave



34374

Figure 2-9 Topology (a) shows a hypothetical situation. RTB has received two different binding requests for prefix 128.89.0.0/16 from RTA and RTD. Hence, RTB will create two entries in its LFIB and assign incoming labels for each request. In this example, RTB has assigned VCI 40 for RTA and VCI 70 for RTD. In case RTB doesn't already have an outbound label for the prefix, RTB will send a binding request toward RTC and will get assigned VCI 50 as an outbound label. As a result, cells arriving from RTA and RTD on VCIs 40 and 70 would be sent over VCI 50 and would potentially get interleaved causing AAL5 frames to be discarded.

Figure 2-9 Topology (b) shows the same scenario with the difference that RTB has now requested two outgoing labels for prefix 128.89.0.0/16. RTB will get assigned two VCIs 50 and 60. Cells from RTA will be switched using cross-connect (40, 50) and cells from RTD will be switched using cross-connect (70, 60). As such, complete noninterleaved AAL5 frames will be received at the destination. This is how MPLS supports switches that do not have "VC Merge" capability.

## Virtual Circuit Merge-Capable Switches

To better utilize the VC space, VC merge may be implemented. VC merge allows the switch to transmit cells coming from different VCIs over the same outgoing VCI toward the same destination. In other words, it allows multipoint-to-point connections.

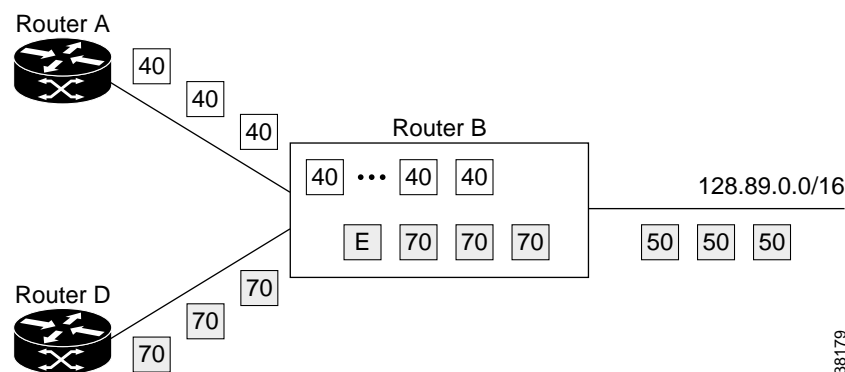
VC merge is accomplished by queuing complete AAL5 frames in input buffers until the end of frame has been received. The cells from the same AAL5 frame are all transmitted before sending cells from any other frame. This requires sufficient buffering capabilities inside the switch, but no more buffering than is required in IP networks.

The small additional delay caused by VC merge is of little concern, because VC merge is designed for IP traffic and need not be used for delay-sensitive traffic. IP traffic has good delay tolerance compared to some other traffic that might be carried on an ATM network.

VC merge is illustrated in Figure 2-10. Here, RTA and RTD are sending traffic toward prefix 128.89.0.0/16. RTB has a single outbound VCI 50 bound to that prefix. Cells coming over VCIs 40 and 70 are buffered in separate queues of RTB until complete AAL5 frames have been formed.

In this example, an end of frame has been detected over VCI 70 and the complete frame has been transmitted over VCI 50. An end of frame has not been detected for cells coming over VCI 40, and these cells are held back in the input buffer. This solves the cell interleave problem and minimizes VC usage.

Figure 2-10 VC Merge



## Label VC Connections and Cross-Connects

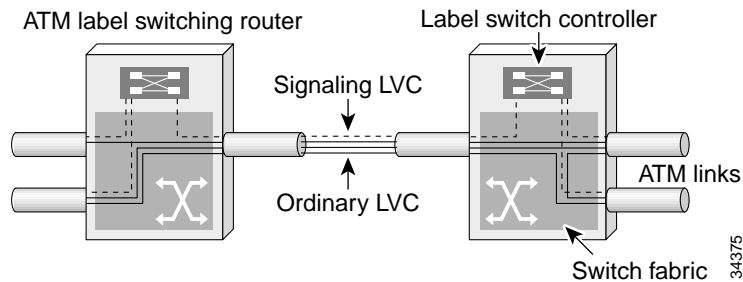
A link between two ATM LSRs, or between an ATM Edge LSR and an ATM LSR, is an ordinary ATM link. Because ATM MPLS uses the VCI fields of a few separate VPIs to carry a label, each label on a link corresponds to a different virtual circuit (VC).

These VCs are called label virtual circuits (LVCs). LVCs are neither switched virtual circuits (SVCs) nor permanent virtual circuits (PVCs), and are set up using LDP instead of ATM Forum signaling protocols. LVCs, PVCs, and SVCs may all be used on the same link, though they use different parts of the VPI/VCI space.

As illustrated in Figure 2-11, at least two distinct types of LVCs are used on each link:

- Signaling PVC**  
 This VC is used to carry IP packets that are reassembled and examined at each ATM LSR. It is used to carry routing information (BGP, OSPF, IS-IS, and so on) and LDP. It might also be used to carry management traffic, such as Simple Network Management Protocol (SNMP) traffic or Internet Control Message Protocol (ICMP) traffic. By default, this VC has VPI and VCI (0, 32).
- Ordinary LVCs**  
 These carry label-switched data. Packets on ordinary LVCs are cross-connected by ATM LSRs without being reassembled. On each link, all ordinary LVCs are within the same virtual path (VP) or small set of VPs.

Figure 2-11 Interconnecting ATM Label Switch Routers



An ATM LSR differs from an ordinary ATM switch in the way connections are set up. Normally an ATM connection is set up by control software running a connection routing protocol such as PNNI or AutoRoute. In ATM MPLS, a piece of software called a label switch controller is used.

## Label Switch Controllers

A label switch controller (LSC) is part of an ATM LSR. The LSC runs an IP routing protocol such as OSPF or IS-IS, in addition to MPLS software. The IP routing software maintains knowledge of the layout of the network. Using this information, LDP establishes labels (such as VCs) on links connected to the ATM LSR.

When the LSC has established incoming and outgoing labels for the same route in its LFIB, it then instructs the switch fabric to set up a connection with the parameters (incoming interface, incoming label VCI, outgoing interface, outgoing label VCI).

Figure 2-12 shows three possible locations for the LSC:

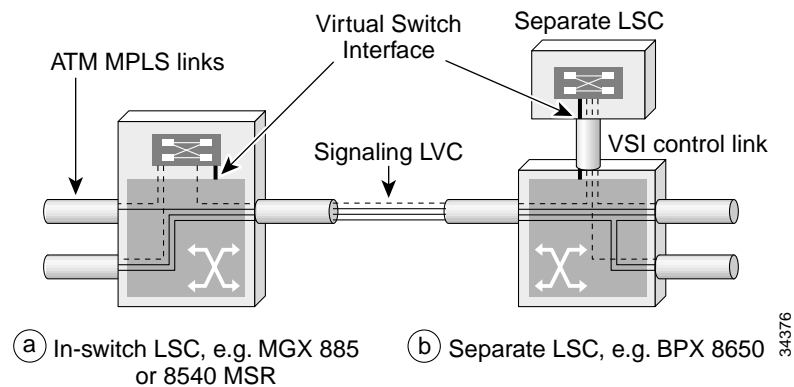
- Switch control card**  
 LSC software might run on board the ATM switch, on the main control card. In an LS1010, Catalyst 5500 or 8500 ATM-LSR, the LSC software runs on the main control card, the ATM Switch Processor.
- Another card in the switch shelf**  
 The software might run on board the ATM switch on a card separate from the main switch control card. In the MGX 8800, a Route Processor Module (RPM) card in the switch is the LSC.
- Separate hardware**  
 The LSC may also be a separate piece of hardware. A Cisco BPX 8650 IP+ATM switch consists of a BPX 8600 ATM switch shelf and an LSC based on a Cisco 7200 series router. The LSC and switch are interconnected by a switch control link. For the BPX 8650, the switch control link is an ATM link. This link is used in a different way to the other ATM interfaces on the LSR—it will be used to connect the signaling PVCs from all other interfaces on the switch to the LSC, but it will often not carry any data.

A LSC sets up connections in the switch fabric by way of a switch control interface. In the case of the LS1010, Catalyst 5500 or 8500, this is an internal interface within Cisco IOS®. In the case of the BPX 8650 and MGX 8850, an external switch control interface is used.

## BPX 8650 Label Switch Router: Controlling a BPX 8600 with An LSC

Figure 2-12 Topology (b) shows how a LSC is connected in a BPX 8650 switch. The physical connection between the LSC and the BPX-series ATM switch shelf is the VSI control link, which is an ATM link. The VSI control link could be an STM-1 link, connected to one port of a four-port or eight-port broadband switching module (BXM) STM-1 card.

**Figure 2-12 Label Switch Controller Locations**



Enabling LSC control of a switch requires:

- Declaring that an ATM interface on the LSC is a Label Switch Control interface
- Enabling a port on the switch as a control interface

The data connections between the LSC and the switch shelf consist of two sets of VCs:

- **Signaling PVCs**

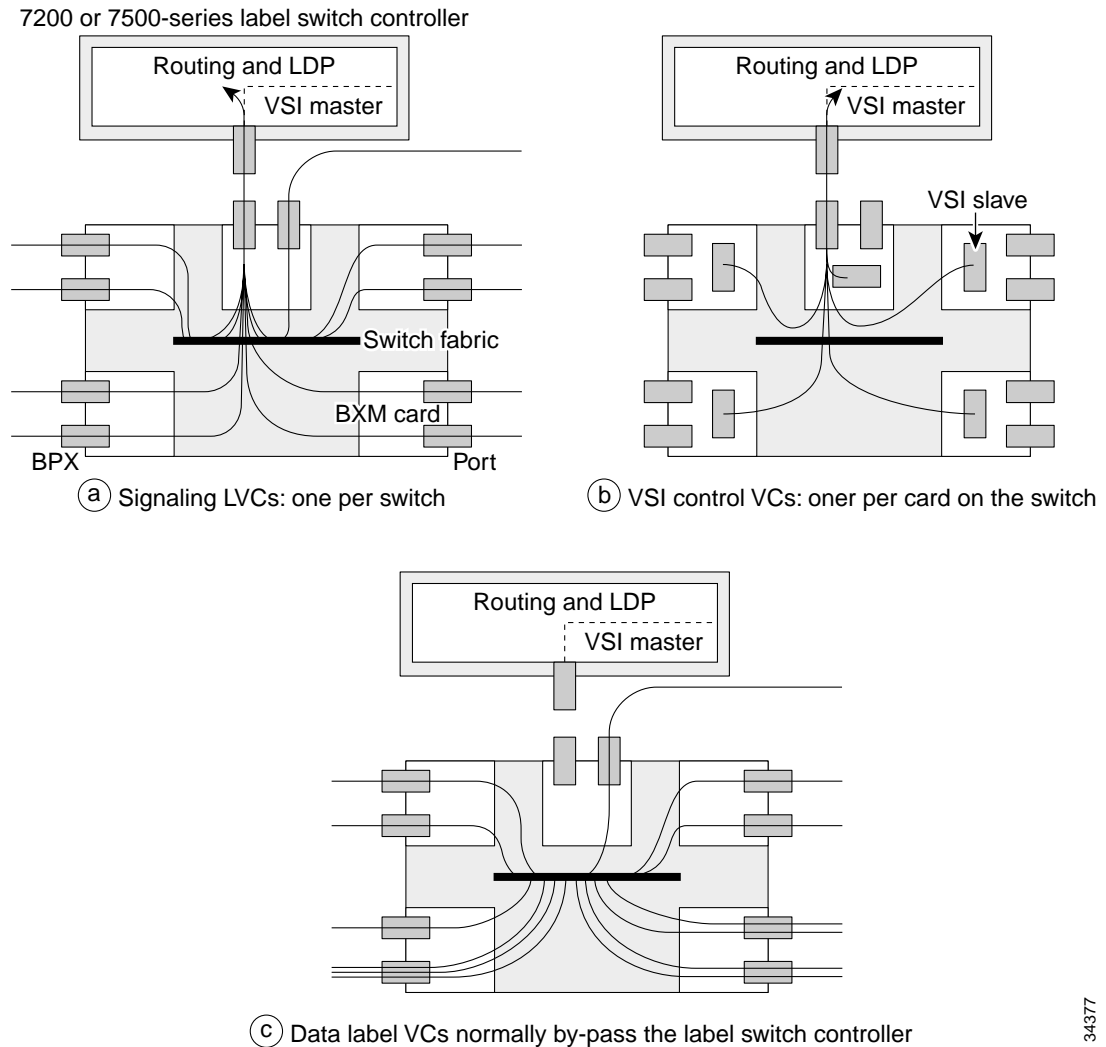
The signaling label VCs from every interface of the ATM switch must be connected through to the LSC shown in Figure 2-12 Topology (a). The signaling PVC on each interface is on VPI and VCI (0, 32) by default, but will generally be cross-connected to a different VCI on the switch control link. This VCI is chosen by the LSC software, which requests the setup of the cross-connects as part of its initialization.

- **Switch Control VCs**

The LSC uses an interface control protocol to discover the port configuration of the switch and make switch connections. This protocol operates using VCs connected to each port card, called the external control VCs shown in Figure 2-12 Topology (b). There may be up to 12 of these, one for each BXM port card in the BPX 8650. These are set up automatically if external control is enabled.

Using the infrastructure of signaling PVCs and external control VCs, the LSC can establish label bindings with the neighboring ATM Edge LSRs, and consequently request the set up of LVC cross-connects in the switch. Most data LVCs by-pass the LSC, as shown in Figure 2-13 Topology (c).

Figure 2-13 Connecting a BPX 8650 and Label Switch Controller



34377

An LSC can simultaneously act as an Edge LSR (see Using an LSC as An Edge LSR, page 2-21). Because it can act as an Edge LSR, an LSC will always have a set of LVCs terminating on it. However most LVCs in a typical ATM-LSR configuration will by-pass the LSC. In addition, the LVCs terminating on the LSC will not be used much, unless the Edge LSR function of the LSC is used.

## IP+ATM Capability

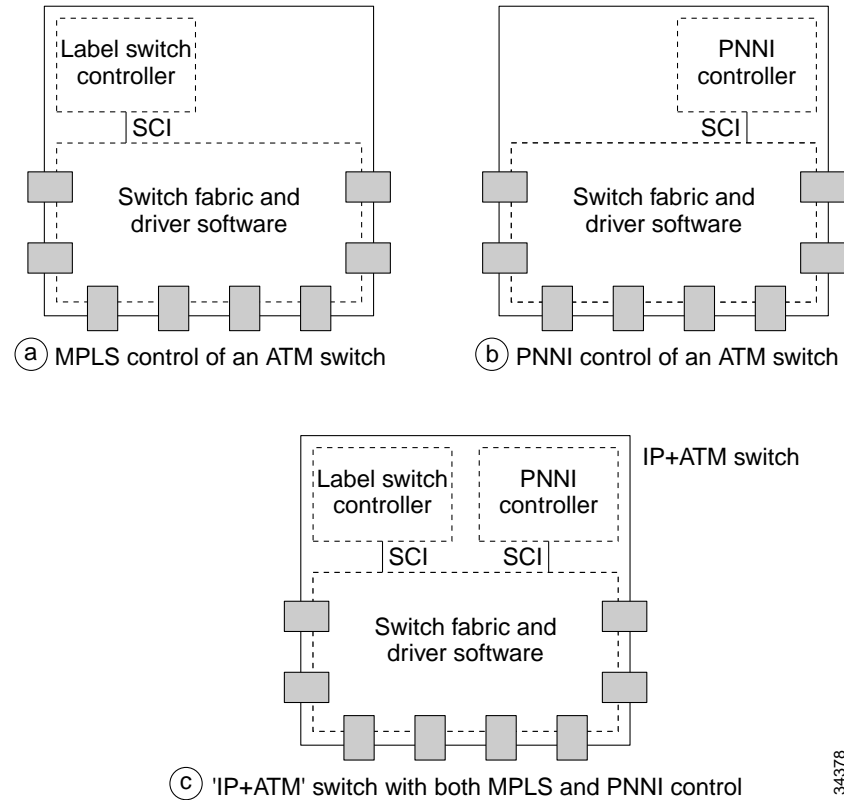
The preceding sections show a label switch controller can be added to an ATM switch to give it MPLS capability and how an ATM network can be used to simultaneously provide MPLS service and traditional ATM switching.

The key to doing this is the IP+ATM capability of Cisco ATM switches. Figure 2-14 Part (a) shows an ATM switch with an MPLS label switch controller.



PNNI control software can be connected to an ATM switch in the same way as shown in Figure 2-14 Part (b). Cisco IP+ATM switches allow an LSC and a PNNI controller to be simultaneously connected to the same switch (Figure 2-14 Part (c)). This means that the same switch can support both optimized IP services using MPLS, as well as traditional ATM services using PNNI.

**Figure 2-14 Comparing MPLS, PNNI, and IP+ATM Switches**

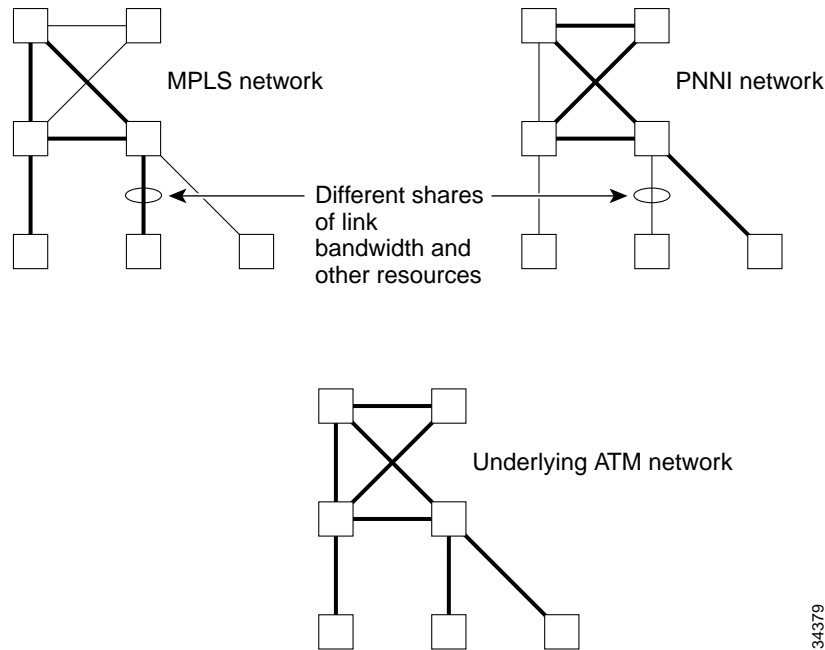


A Cisco IP+ATM network physically consists of ordinary ATM switches and links. As part of the initial configuration of the network, the operator assigns resources of the ATM network to PNNI and MPLS:

- Bandwidth on links
- VPI/VCI space on links
- VC connection table spaces
- Traffic management

As illustrated in Figure 2-15, this partitioning of resources is quite flexible, with arbitrary divisions of resources between the different control planes. You may define fixed allocations of resources to the control planes or pool link bandwidth or connection table spaces shared between the control planes.

Figure 2-15 Comparing MPLS, PNNI, and IP+ATM Networks



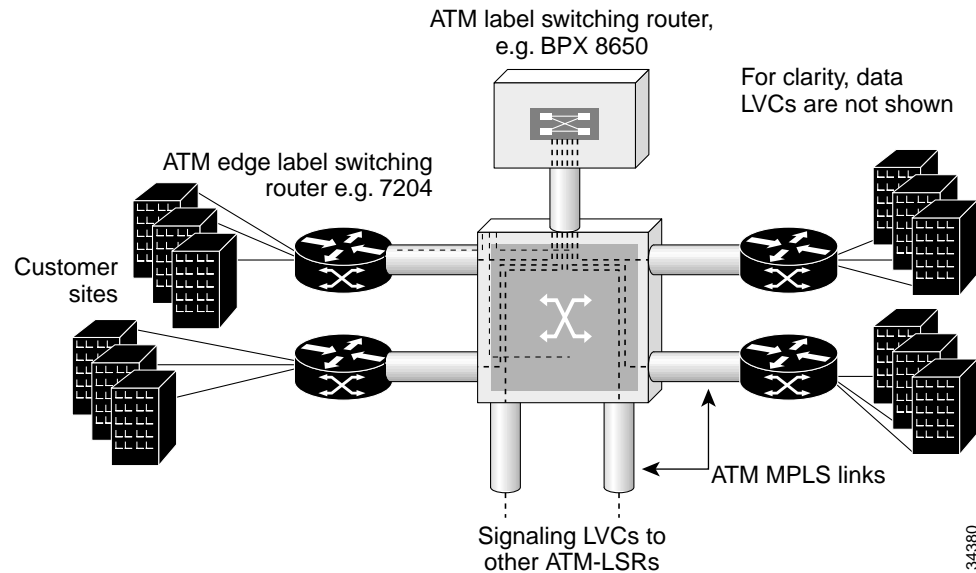
The concept of controllers independently controlling a switch extends to more than two controllers. Cisco IP+ATM switches have the ability to support four or more control planes.

## An ATM MPLS Point of Presence

An example ATM MPLS Point of Presence (PoP) is shown in Figure 2-16. This consists of an ATM-LSR and several edge routers connected to customer sites. Signaling PVCs connect every Edge LSR in the PoP to the LSC in the ATM-LSR. In this example, each of the edge routers is connected to only one ATM-LSR, although this is not necessary.

34379

Figure 2-16 An ATM MPLS Point of Presence (PoP)



## Using an LSC as An Edge LSR

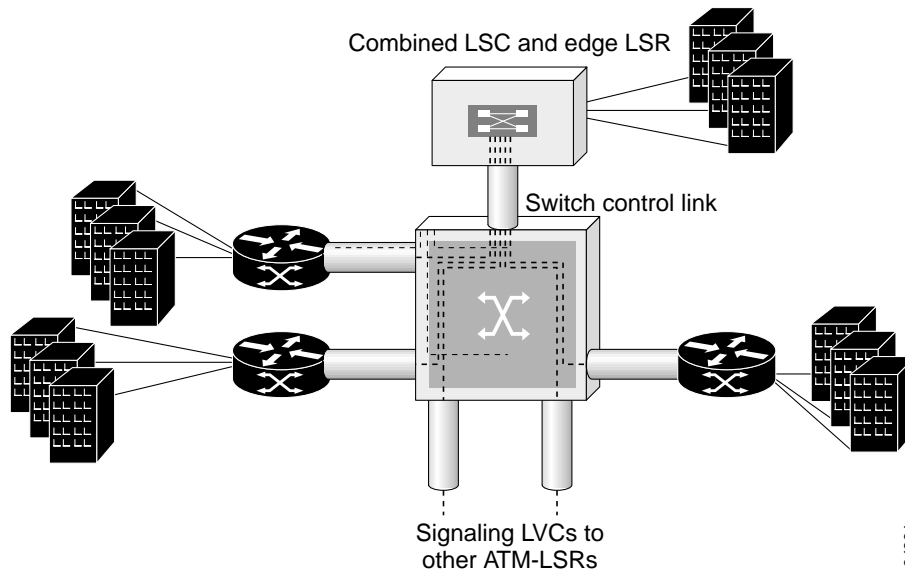


### Note

If the ATM switch used is a BPX 8650, then the PoP shown in Figure 2-16 requires an LSC in addition to the Edge LSRs. However, an extra device is not always required.

In LSRs such as the BPX 8650 Cisco LSR, which uses a separate LSC, the LSC can also act as an ATM Edge LSR. This means that if a BPX 8600 ATM-LSR is added to an existing PoP, an extra router is not necessarily required, because one of the existing routers can provide LSC functionality. A PoP using an edge router as LSC is shown in Figure 2-17.

Figure 2-17 An ATM MPLS PoP with Combined LSC and Edge Device



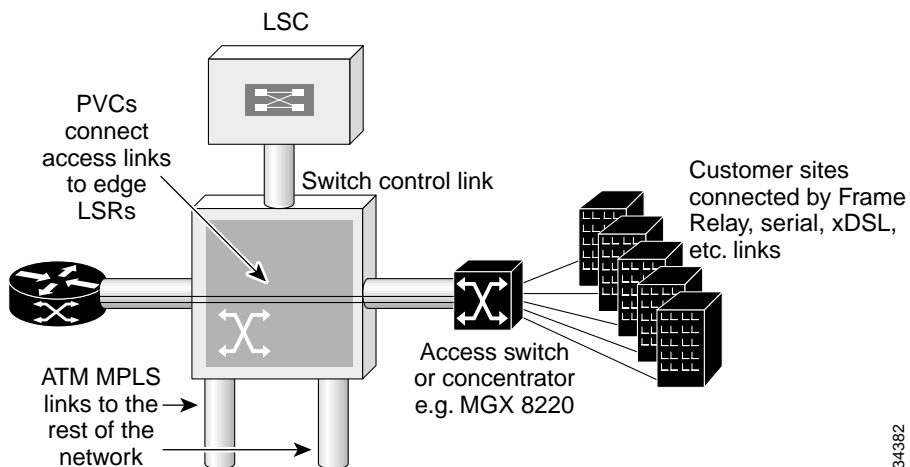
34381

## Using An Access Switch in An ATM MPLS PoP

In the previous two sample PoPs, the customer access lines have been connected directly to the Edge LSRs.

Figure 2-18 shows another alternative: to bring the customer lines into the PoP by way of an access switch or concentrator such as the MGX 8220. Customer lines are connected to the access switch, and then “backhauled” to the Edge LSRs by way of PVCs.

Figure 2-18 Using an Access Switch or Concentrator in An ATM MPLS PoP



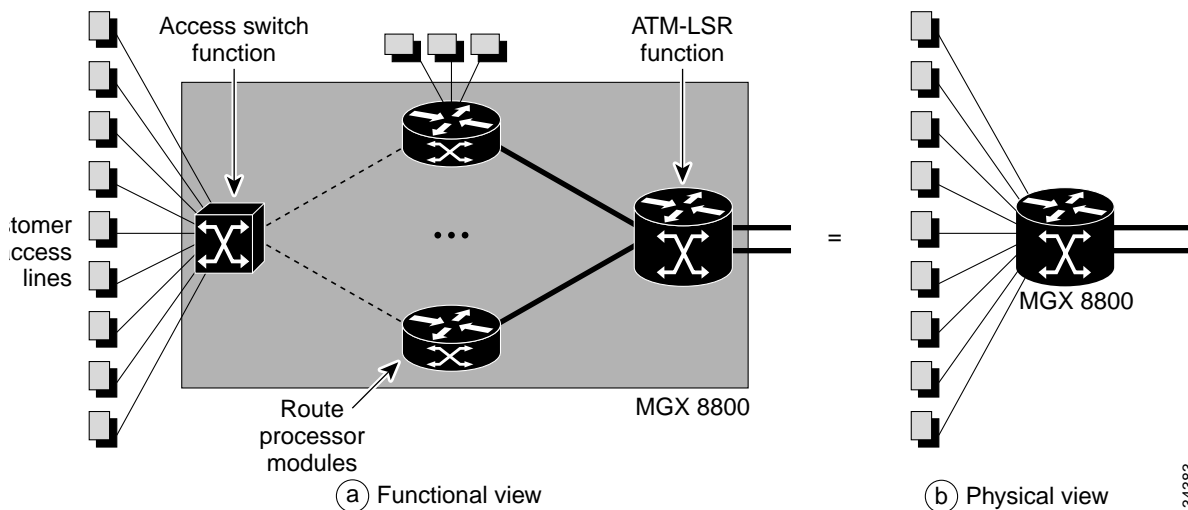
34382

## A Fully Integrated PoP

Yet another alternative is to use a single device that integrates an ATM-LSR, Edge LSRs, and an access switch, as shown in Figure 2-19 (a). The MGX 8800 IP+ATM switch does this.

In the MGX 8850, router cards called Route Processor Modules (RPMs) act as Edge LSRs. One of the RPMs will also act as a label switch controller, giving ATM-LSR function to the MGX 8800. All these functions are combined in a single chassis, shown in Figure 2-19 (b). The 6400 access switch has similar capabilities.

Figure 2-19 MGX 8800 as An Integrated ATM MPLS PoP



## Dual Backbones: Traditional ATM and ATM MPS or Packet-Over-SONET

The networks shown so far have combined MPLS and traditional ATM services on the same switches. Providers may wish to build a new MPLS infrastructure (either ATM MPLS or packet-based MPLS) alongside an existing ATM infrastructure, using two separate backbones in the networks: one for traditional ATM services and one for IP services.

Cisco IP+ATM edge devices support dual backbones in a very flexible way, allowing access to both the MPLS network and services, and the old ATM network, even from a single access link. A single IP+ATM edge switch will support full Edge LSR and LSC functions, as well as traditional ATM SPVC and SVC services.

From a functional perspective, ATM MPLS services and traditional ATM services are then carried separately. However with a Cisco IP+ATM backbone, the separate MPLS and Traditional ATM switching functions are carried on the same switch backbone. This is illustrated in Figure 2-20 (b).

Another alternative, shown in Figure 2-20 (c), is to use physically separate backbones for MPLS services and traditional ATM services. The MPLS backbone can use ATM-LSRs, or alternatively router-based LSRs such as the Cisco 7500 series, or Cisco 12000 series Gigabit Switch Routers. Nearly any link types can be used in the MPLS backbone, including packet-over-SONET/SDH, or packet-over-WDM, if desired.

Cisco IP+ATM edge switches connect to dual backbones readily. Even from a single link to an end-customer site, some DLCIs or VCIs can connect into the MPLS-based IP services, while others connect into traditional ATM or Frame Relay SPVC services.

Figure 2-20 Supporting IP+ATM Services Using Dual Backbones

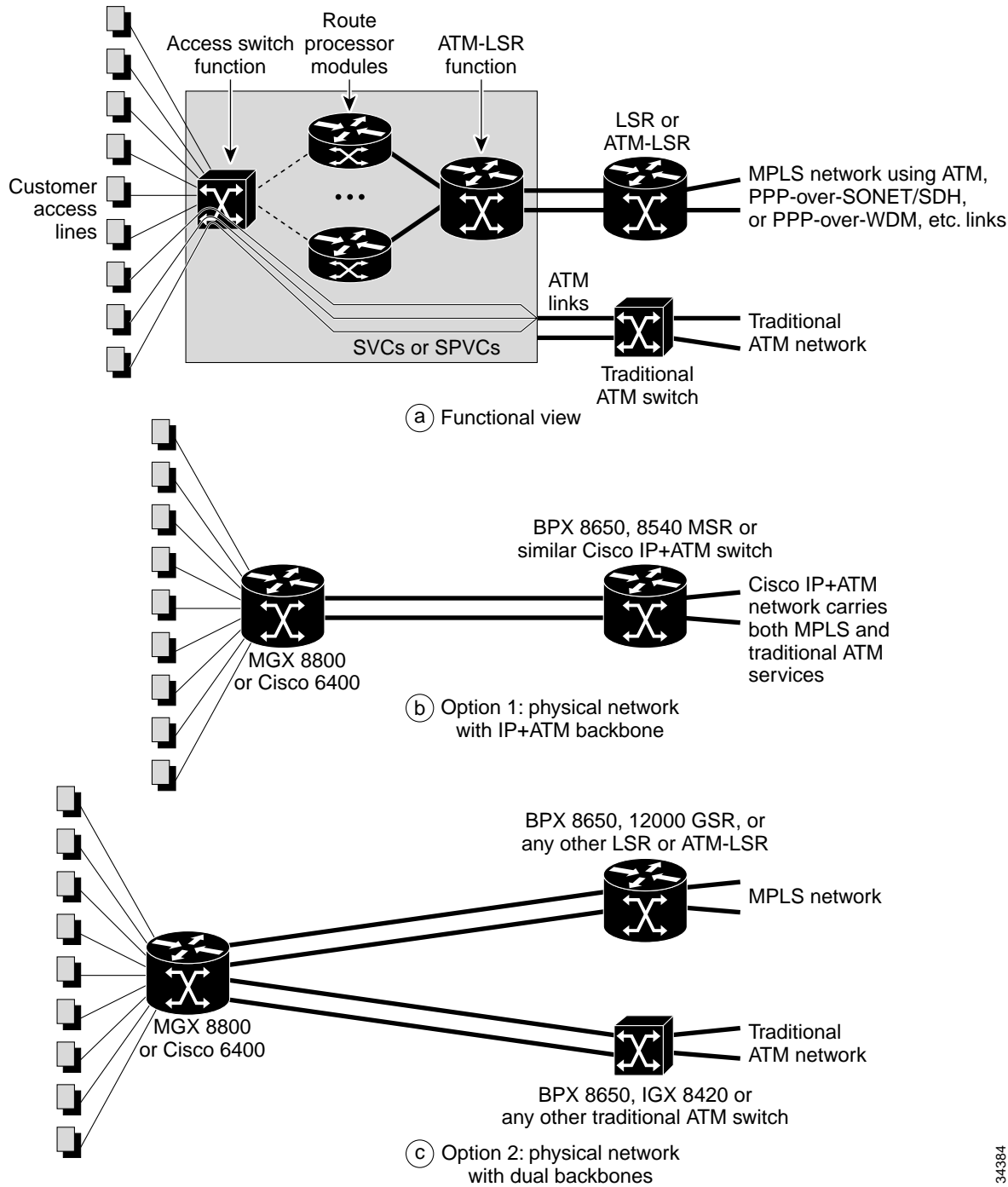
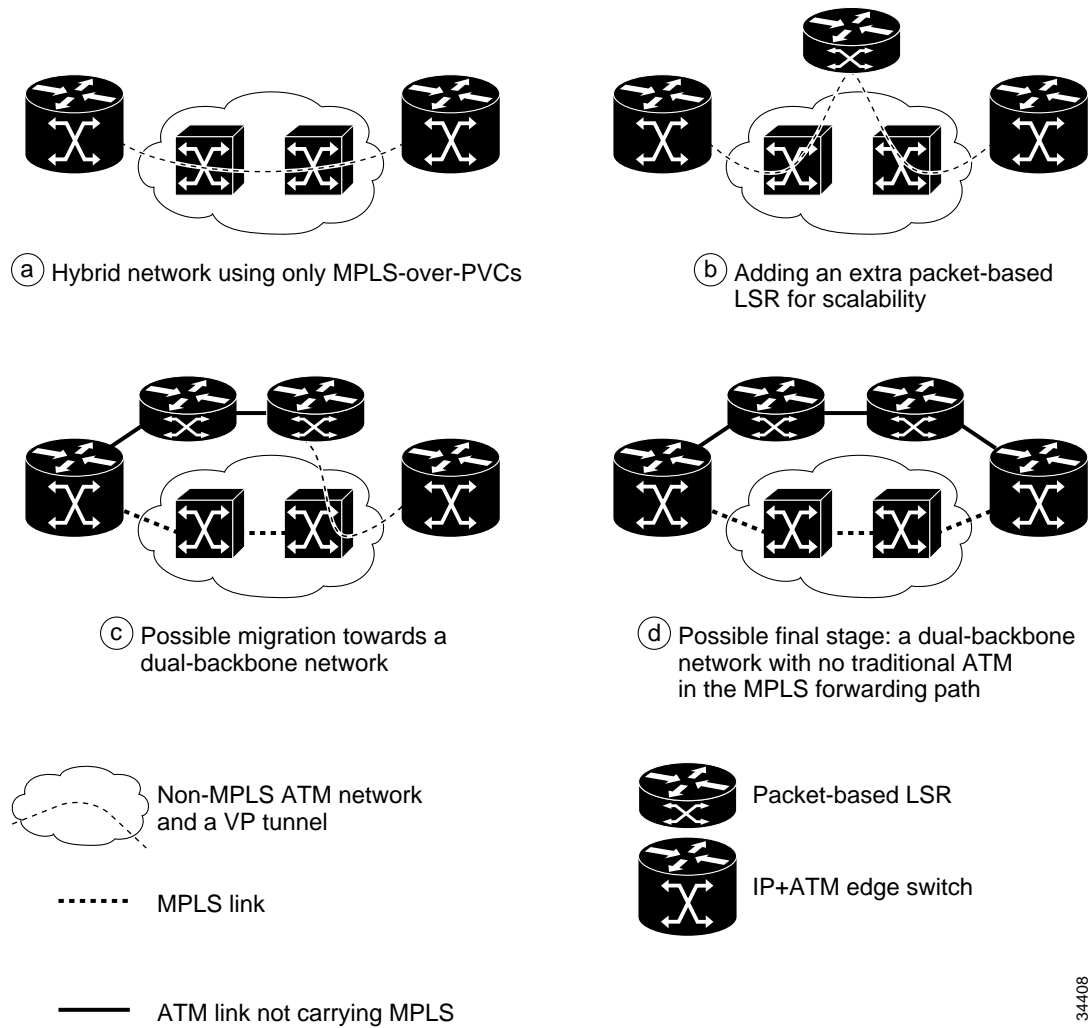


Figure 2-21 shows one strategy for evolving hybrid networks toward a dual backbone solution.

34384

Figure 2-21 Evolution of ATM MPLS Networks to Dual Backbones



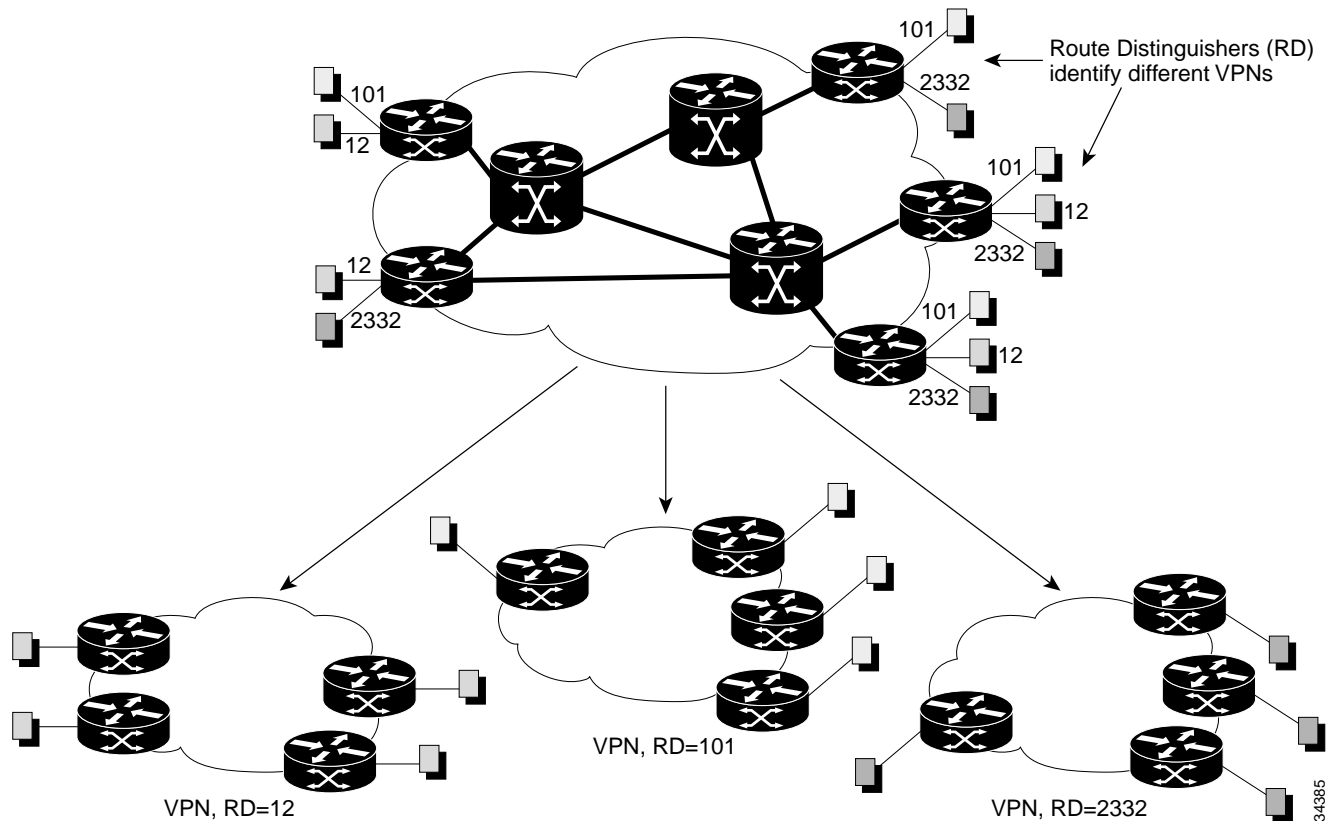
34408

## Virtual Private Networks

Figure 2-22 shows the concept of an IP Virtual Private Network (VPN) service. One service provider network supports many different IP Virtual Private Networks. Each VPN appears to its users as a private network, separate from all other networks.

Within each VPN, there is any-to-any connectivity: each site can send IP packets directly to any other site in the VPN, without having to go through a central site. A simple use of a VPN is an intranet, forming the wide-area IP network of a corporation. Another use is an extranet linking several corporations.

Figure 2-22 Many Virtual Private Networks Provided by One Network



In a VPN service based on Cisco MPLS, each individual VPN is identified by a Route Distinguisher (RD). In Figure 2-22, three different VPNs are shown with Route Distinguishers 12, 101, and 2332. Many more than three can be supported by a single network. Cisco MPLS VPNs networks initially support thousands of VPNs, and later this is extendable to hundreds of thousands, and even millions if required.

## Route Distinguisher

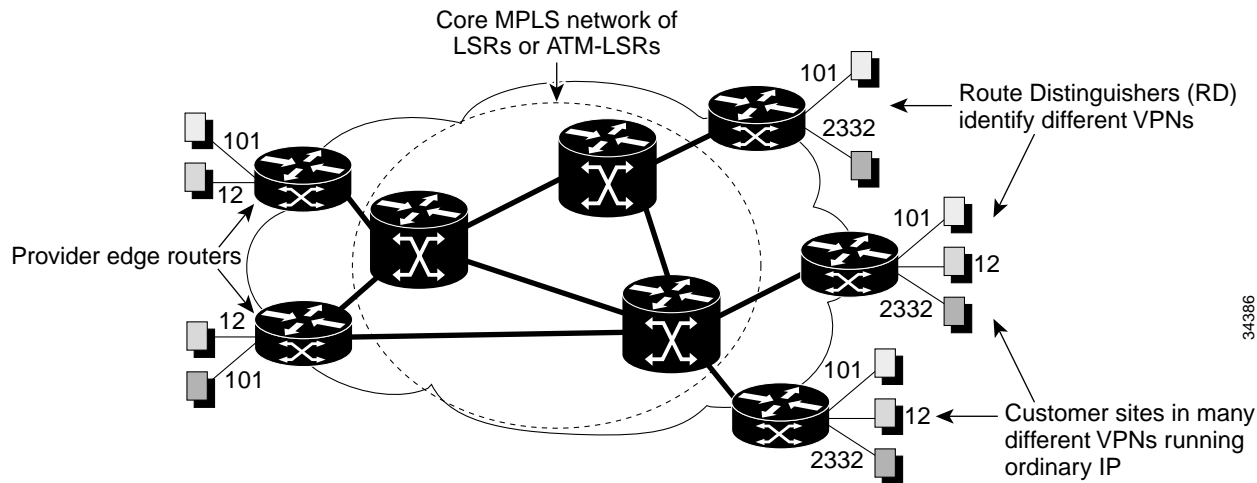
Figure 2-23 shows more detail on how a Cisco MPLS network supports VPNs. An ordinary MPLS network is surrounded by Provider Edge (PE) routers, which are Edge LSRs that support VPN functions. The MPLS network in combination with the Provider Edge routers forms a network supporting VPN services. Customer sites are connected to the Provider Edge routers in any of the ways supported by any MPLS network: Frame Relay, ATM, xDSL, PPP, and so on. Any Cisco device that can act as an Edge LSR can act as a PE router, including the MGX 8800.

More sophisticated cases are possible by means of a Route Distinguisher. A Route Distinguisher (RD) identifies a separate IP address family. It is possible for a single RD to be shared by several intranets and one or more extranets. A separate attribute called the Target VPN controls which addresses are reachable from which VPNs.

However, in a simple intranet service, each VPN does have a unique RD, and the Target VPN can equal the RD. The Route Distinguisher and Target VPN attribute together perform roughly the function of the VPN Identifier. They also extend the VPN Identifier concept from intranets to extremely flexible groupings of sites in extranets.



Figure 2-23 Providing Virtual Private Network Services Using An MPLS Network



In a Cisco MPLS VPN service, the customer sites run ordinary IP. They do not need to run MPLS, or IPsec or other special VPN functions.

At the PE router, a route distinguisher is associated with each link to a customer site. The links may be physical links, PPP links, individual Frame Relay or ATM virtual circuits, xDSL links, or other options. The Route Distinguisher is configured at the PE router as part of the set-up of a VPN site. It is not configured on the customer equipment and is not visible to the customer.

As with other MPLS networks, functions in MPLS VPN equipment can be divided into a forwarding component and a control component:

	Core Network Functions	Additional Provider Edge Router Functions
Forwarding Component	Label switching	Label switching with VPN support
Control Component	Routing: OSPF or IS-IS Signalling: LDP	Routing: Multiprotocol BGP 4 Signalling: Multiprotocol BGP 4

The core network uses ordinary MPLS forwarding and control. The Provider Edge routers have some additional functions to support VPN services.

## Forwarding in a Cisco Virtual Private Network Service

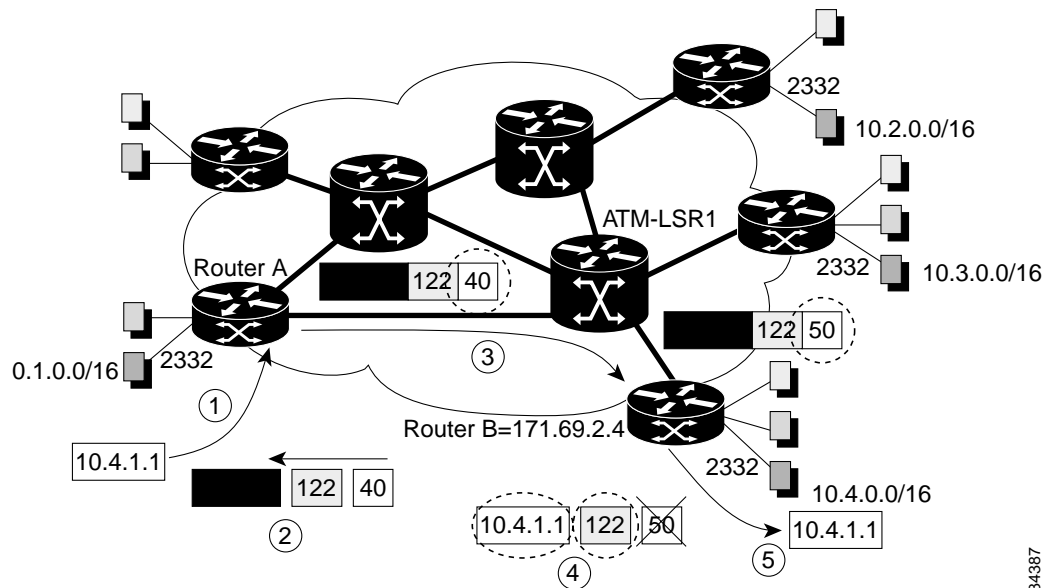
Figure 2-24 shows how packets are forwarded in a Cisco MPLS+BGP VPN service.

1. Packets arrive from a particular customer site. In this example, the site is in the VPN with RD=2332, and is attached to PE router RTA. The packet that arrives from the customer site is an ordinary IP packet, with destination IP address 10.4.1.1.
2. RTA looks up its VPN forwarding table, shown in Figure 2-25. RTA gets two different labels to put on the packet. The inner label, which has value 122 in this example, is carried in a header encapsulated along with the rest of the IP packet. The inner label carries information specific to the

virtual private network with RD=2332. The outer label, value 40, is an ordinary MPLS label that tells the rest of the network that the packet is to be delivered to RTB, IP address 171.69.2.4. The label value 40 was bound to 171.69.2.4/32 using ordinary LDP procedures described earlier.

- The packet is sent on to the core of the network, which performs ordinary label switching. In this example, ATM-LSR1 swaps label 40 with label 50 while forwarding the packet on towards RTB. Labels 40 and 50 are the labels for 171.69.2.4/32 on different links.
- When RTB receives the packet, it ignores the outer label, because it corresponds to RTB's own IP address (171.69.2.4). It then looks up the inner label, 122, in a table (see Figure 2-25). In this case, the inner label corresponds to RD 2332. It then looks at the IP address on the packet, and finds that the packet is destined to RD=2332, IP address=10.4.1.1.
- RTB finds that RD=2332, IP address=10.4.1.1 is on a site directly connected to RTB by an ordinary IP link. It discards the labels and forwards the ordinary IP packet on to that site.

Figure 2-24 Forwarding Packets in a Cisco MPLS Virtual Private Network Service



Note that the IP addresses 10.4.1.1 and 171.69.2.4 have different scope. The IP address 10.4.1.1 is part of the VPN with route distinguisher 2332. Only packets inside that VPN will be able to reach this address. There may be other IP addresses 10.4.1.1 in other VPNs, but they are distinguished by having different RDs.

Address 171.69.2.4 is part of the service provider's backbone and is not part of any VPN. It is probably a Registered IP address from a range issued by the Internet Assigned Numbers Authority.

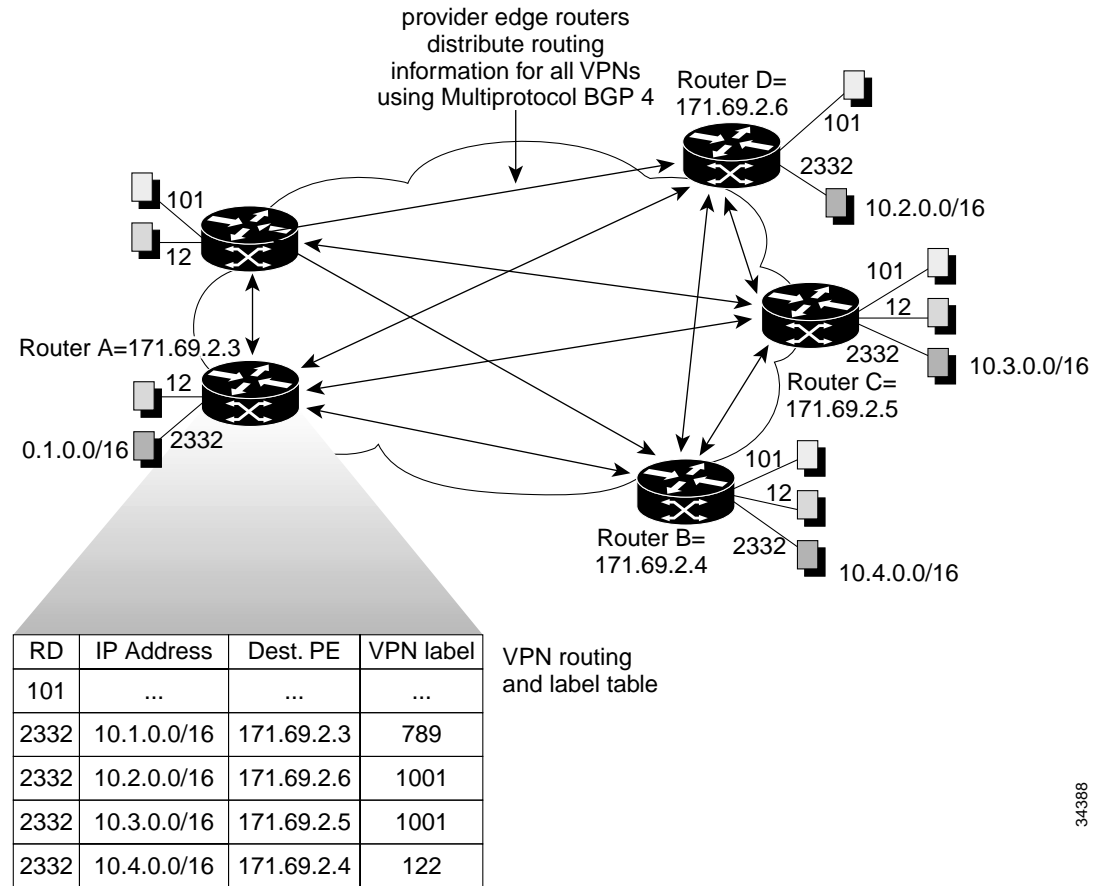
## Control in a Cisco MPLS+BGP Virtual Private Network Service

Critical to the function and scalability of Virtual Private Network services is the use of the Border Gateway Protocol (BGP) version 4.

Figure 2-25 shows the extra control functions in a Cisco MPLS+BGP Virtual Private Network service. The PE routers communicate with each other by using the Border Gateway Protocol (BGP) version 4, with Multiprotocol Extensions.

BGP v4 is the standard routing protocol for communicating Internet routing tables. Multiprotocol BGP v4 is a standard extension to BGP v4 (RFC 2283). The PE routers use Multiprotocol BGP v4 to send VPN routing information to each other. This allows each PE router to build a complete VPN routing table. Part of a VPN routing table is shown in Figure 2-25.

Figure 2-25 Control Functions in a Cisco MPLS Virtual Private Network Service



The VPN works like this:

1. PE router RTA knows that it has a site with RD=2332, because it has been configured with this information.
2. PE router RTA learns that an IP address range (RD=2332, Destination-Prefix=10.1.0.0/16) is reachable there. It gets this information either by running an IP routing protocol (such as RIP or OSPF) out to the customer site, or by static configuration. RTA records the address range (RD=2332, Destination-Prefix=10.1.0.0/16), and also picks a label value, say 789, to correspond to this route.
3. RTA tells all the other PEs about this route and label using multiprotocol BGP.
4. RTA records that its own IP address, 171.69.2.3, is the destination PE router for this address.  
This example is simplified; it ignores the Target VPN attributes. This attribute is sent with each route along with the Route Distinguisher and is also recorded in the label tables.
5. Similarly, RTB learns that the route (RD=2332, Destination-Prefix=10.4.0.0/16) is on a link directly connected to RTB.

34388

6. RTB picks a label, such as 122 to correspond to this route.
7. RTB then advertises this route to all other PE routers, which record the route and label in their tables, along with RTB's IP address (171.69.2.4).

This is how, in Figure 2-25, RTA knew to put label 122 on a packet that matched the route (RD=2332, Destination-Prefix=10.4.0.0/16). RTC and RTD will also send routes for RD=2332.

In this way, each PE router gets access to complete routing information for all the RDs it supports.

Note that the assignment of VPN labels does not need to be coordinated across different PE routers. For example, in Figure 2-25, label 1001 from RTC (171.69.2.5) is different from label 1001 from RTD (171.69.2.6), because the labels are applied to packets sent to different PE routers. The outer label shown in Figure 2-25 ensure that the each packet reaches the correct PE router.

## Attributes of Cisco MPLS+BGP Virtual Private Networks

### Privacy and Security

- Cisco MPLS+BGP virtual private networks have the same level of privacy as Frame Relay networks. In a Frame Relay network, each packet carries a label, called a DLCI, which ensures correct delivery of the packet, provided the network is configured correctly. Because MPLS+BGP VPNs label each packet with destination information, they achieve the same level of privacy as Frame Relay networks.
- In Cisco MPLS+BGP virtual private network services, routing information is kept private. Route distribution is controlled by the Route Distinguisher. (The Target VPN attribute can also be involved.) This keeps all customers' routing information separate, and ensure that customers can reach only legal addresses for their VPN. If routing protocols (RIP, OSPF, and so on) run out to customers' sites, then each customer gets access only to routes with the correct Route Distinguisher.
- Spoofing is impossible. Because MPLS labels are applied by the Provider's edge router, a customer cannot pretend to be a member of another VPN by applying different labels. This would cause the PE routers simply to discard the packets. Only ordinary IP packets are accepted from customer sites, and the network will forward them as normal for the customer's VPN. (A later extension may allow MPLS packets to be accepted from customer sites. When this is supported, any label on a packet received from a customer link must correspond to an IP address within the customer VPN. A packet with an illegal label will be discarded.)
- Denial-of-service attacks may be prevented by using mechanisms described in "Migrating MPLS into a Traditional ATM Network" section on page 2-33. Briefly, any packets sent in excess of the customer's traffic contract can be marked and carried at lower precedence than in-contract traffic. If this is done, then excess out-of-contract traffic from any source cannot adversely affect service for any traffic compliant with its traffic contract.

In addition, authentication and encryption can be used to give even greater security.



#### Note

Most Frame Relay customers are satisfied with Frame Relay privacy and do not encrypt their traffic. The same can be expected of MPLS+BGP virtual private network customers.

## Customer Independence

- Customers can use any IP addresses. Because the provider network keeps customers' IP address ranges separate with Route Distinguishers, each customer may use any IP addresses in their network. Many customers may use identical IP addresses. Customers can use registered addresses, Private addresses (such as IP network 10.x), or even completely illegal unregistered addresses.
- Customer equipment does not run MPLS or any special features. Any IP-capable equipment can be used at customer sites. The customer sites do not need to run "VPN Routers" with IPsec or any other special equipment. However Cisco will allow IPsec to be used to extend a VPN to a remote customer site attached to another provider's network. In this case, an IPsec tunnel will run between the customer site and the PE router.

There might be advantages to running Cisco routers at customer sites. Optional Frame Relay congestion-prevention features are available if the CPE is Cisco routers. Cisco CPE routers have industry-leading capability for setting and marking the desired Class of Service for IP packets.

- Customers automatically get to send any-to-any traffic within their virtual private network. There is no need to backhaul traffic to a central site in the VPN.

## Scalability and Stability

- Routing distribution is highly scalable. In Cisco MPLS+BGP virtual private network services, a single set of BGP peerings between PEs is used, irrespective of how many different VPNs are supported. This compares to alternative, unscalable techniques where a different routing protocol runs in the provider backbone for each VPN. In addition, existing BGP techniques such as Route Reflectors can be used to help VPN route distribution scale even further. Initially, Cisco MPLS+BGP virtual private network services support tens of thousands of VPNs and over 100,000 routes. This will be extended to hundreds of thousands of VPNs and millions of routes, and there is no reason why it cannot be extended even further. In addition, BGP has techniques such as "route flap damping" which prevents badly behaving customer sites from affecting BGP stability. Note also that BGP has been the only routing protocol with the scalability and stability to support networks the size of the Internet. The current Internet could not exist without BGP.
- Packet forwarding is highly scalable and stable. Cisco MPLS+BGP virtual private network services use two-level labelling. Only the top-level label is seen by the core MPLS network. This means that the number of VCs used in the network depends only on the number of PE routers used.
  - MPLS LVCs are not created for individual customer flows, which makes packet forwarding highly scalable.
  - The core label switch routers have no knowledge of VPNs and do not respond to any changes in VPNs, which makes packet forwarding highly stable.

## Management

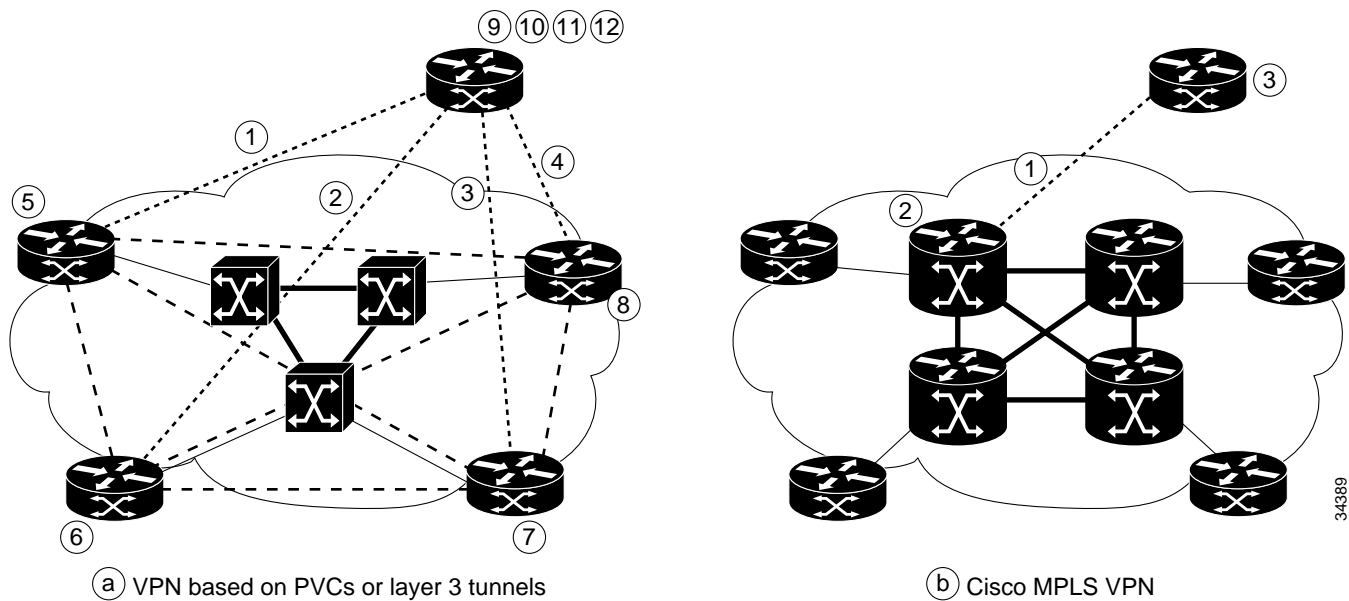
Cisco MPLS+BGP virtual private networks have significant management advantages. These are important even for smaller carriers and service providers, to whom the scalability advantages might not be important:

- The provider network uses only one set of BGP peerings, no matter how many VPNs it supports. This results in simple administration of routing in the network.
- Adding new VPNs and sites is very straightforward.

For example, Figure 2-26 shows the steps necessary to add an existing site to a small VPN with full connectivity. With VPNs based on Frame Relay or ATM PVCs, or IPsec or similar Layer 3 tunnels, the number of steps is dependent on the number of sites already in the VPN.

In the example in Figure 2-26 Topology (a), there are four existing sites. Consequently, configuration steps 1-4 are the set-up of PVCs or Layer 3 tunnels. Steps 5-8 involve configuring the endpoints of these links at the existing customer sites. Steps 9-12 are the configuration at the tunnel endpoints at the new site. There may be more configuration steps at all routers for configuring IP routing to the new site. Obviously, far more configuration steps are required if there are tens or more sites in a VPN.

Figure 2-26 Management Operations: Adding a Site to a VPN



- With Cisco MPLS+BGP virtual private network services, configuration is very much simpler, irrespective of the number of sites in a VPN. There are three steps:
  - a. establish the link to the new site
  - b. configure the link at the PE router (with R, and enabling a routing protocol or static address)
  - c. configure the link at the customer site (enabling a routing protocol or default route).

BGP will then automatically distribute reachability information and labels for the new site and full communication to the new VPN site will be established. This distribution requires no manual configuration and will typically take a minute or less.

In this way, Cisco MPLS+BGP virtual private network services have extreme simplicity of management built into the technology that runs the service itself.

## Migrating MPLS into a Traditional ATM Network

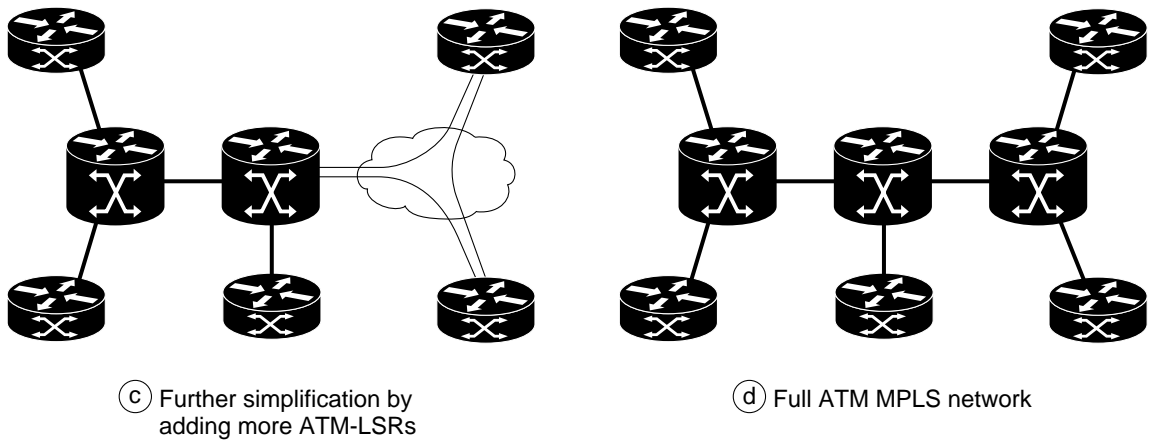
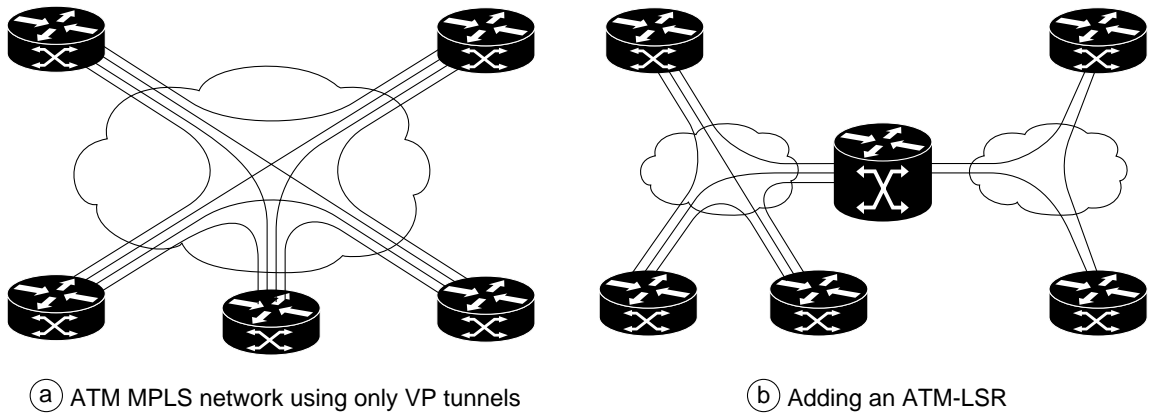
MPLS can be deployed into a traditional ATM network gradually, starting with just a single pair of ATM-LSRs in an otherwise purely ATM network. MPLS can be deployed across non-MPLS-capable switches using VP connections through the traditional ATM switches.

These VP connections are called VP tunnels because they allow MPLS to “tunnel” through traditional ATM switches. VP tunnels provide for easy migration to a full MPLS integration, although they do have several disadvantages. A possible migration strategy for introducing MPLS in an existing ATM network is shown in Figure 2-27.

- Figure 2-27 Topology (a) shows a starting position with routers connected by PVPs through an ATM cloud. This has most of the disadvantages of traditional IP-over-ATM networks, notably poor router peering scaling and poor bandwidth efficiency. However, it can support MPLS VPN services.
- Deploying some ATM-LSRs in the network as shown in Figure 2-27 Topology (b) and Topology (c) improves scalability: the number of PVPs to each Edge LSR may be reduced to one (two if dual-homed), and in some cases to zero, if an Edge LSR is adjacent to an ATM-LSR. ATM-LSRs can be interconnected with ordinary ATM switches in a variety of ways. Careful deployment of ATM-LSRs and PVPs can be used to make the PVP mesh more closely match the ATM link topology and hence improve the bandwidth efficiency.
- Because ATM-LSRs can also support traditional ATM services, ordinary ATM switches may be phased out, leading to full deployment of ATM MPLS, as shown in Figure 2-27 Topology (d). The use of PVPs is no longer required.

An alternative to using PVPs is to use PVCs. These work only between with Edge LSRs, and not with ATM-LSRs. In other words, the network of Figure 2-27 Topology (a) could use PVCs instead of PVPs, but none of the other networks in Figure 2-27 could use PVCs.

Figure 2-27 Migrating MPLS over a Traditional ATM Cloud



— ATM MPLS link

☁ Non-MPLS ATM network and a VP tunnel

34390





## Designing MPLS for ATM

---

This chapter details these steps in the process of designing an MPLS network:

- Structures for MPLS Networks
- Choosing Cisco MPLS Equipment for ATM
- Designing MPLS Networks
- Dimensioning An MPLS Network's Links
- IP Routing in An MPLS Network
- Dimensioning MPLS Label VC Space
- Ongoing Network Design

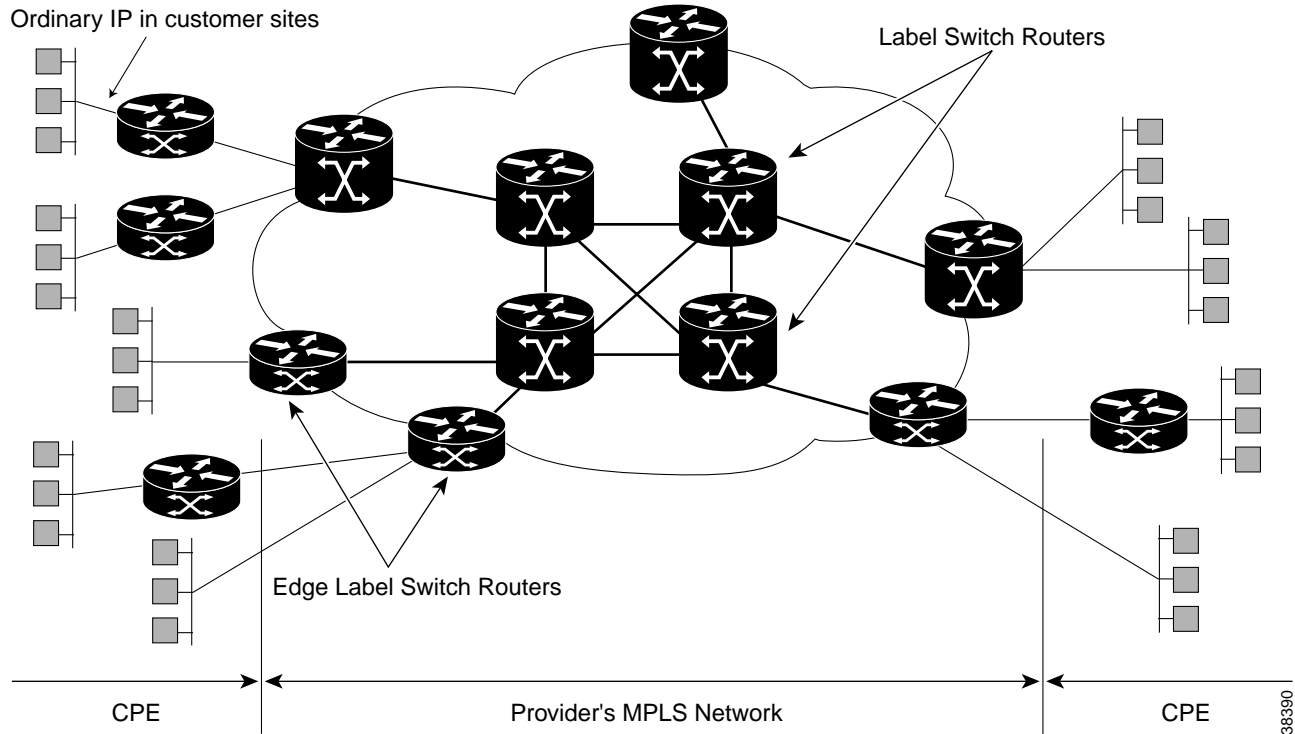
Additional design steps are required for Class-of-Service, MPLS VPNs, traffic engineering, and other IP services.

### Structures for MPLS Networks

A typical structure for MPLS provider networks (carriers or ISPs) is shown in Figure 3-1. An MPLS network consists of edge label switch routers (Edge LSRs) around a core of label switch routers (LSRs). Customer sites are connected to the provider MPLS network. Figure 3-1 shows nine customer sites and six Edge LSRs, but typically there are several hundred customer sites per Edge LSR.

The customer premises equipment (CPE) runs ordinary IP forwarding and normally does not run MPLS. If the CPE does run MPLS, it uses it independently from the provider. It is important to note that the Edge LSRs are part of the provider network and are controlled by the provider. The Edge LSRs are critical to network operation and are not intended to be CPE under any circumstances. The provider might locate and manage routers at the customer sites but these should be running ordinary IP outside the MPLS network itself.

Figure 3-1 Typical MPLS Network Structure



For details on mixed MPLS and IP-over-ATM networks, see Chapter 2, “Integrating MPLS with IP and ATM.”

## Simple Packet-based MPLS

The simplest MPLS network structure is shown in Figure 3-2 Topology (a). This structure applies to router-only networks that might use MPLS for supporting VPN services or IP Traffic Engineering. In this structure, customer sites are connected directly to router-based Edge LSRs. The Edge LSRs are connected to other LSRs that are also based on router platforms.

The routers are interconnected by virtually any sort of link: serial, Ethernet, packet-over-SONET, and so on, and packets with MPLS headers. The routers involved could be 6400, 7200, 7500, or 12000 series Gigabit Switch Routers. Midrange routers (3600 and 4700 series) could be used in lower-bandwidth applications.

In a variant of this structure, the point-to-point links between the routers are actually ATM PVCs. These may be used during migration to ATM MPLS.

## ATM MPLS with Router-based Edge LSRs

The simplest ATM MPLS network structure is shown in Figure 3-2 Topology (b). As with the previous case, customer sites are connected directly to router-based Edge LSRs, typically 6400, 7200, or 7500 series routers. The Edge LSRs are connected by ATM links to the core ATM LSRs. The ATM LSRs may be BPX 8650 IP+ATM Switches, LS1010, 8500 MSR, and later other ATM switches such as the MGX 8800 with PXM-45.

The ATM switches carry packets with ATM MPLS labels; this means that on each ATM link there is a different MPLS Label VC (LVC) for each label.

## Mixed ATM and Packet-based MPLS

It is possible to mix ATM MPLS and packet-based MPLS on one network. A simple example of this is shown in Figure 3-2, Topology (c). In a network such as this, some links run packet-based MPLS, and some links run ATM MPLS. The devices that interface between packet-based MPLS and ATM MPLS are the same routers that act as ATM Edge LSRs: anything from a 3600 up to a 12000.

## ATM MPLS with Separate Access Devices

ATM MPLS networks with router-based Edge LSRs may also use separate access devices, as shown in Figure 3-2 Topology (d). This will occur when access is required through a device that does not support MPLS services. There are three common situations where this is required:

- Access is required to both IP services and ATM PVC services through an access device that does not support MPLS. The most common example of this is the MGX 8220.
- The access device does not yet have software that supports MPLS.
- By means of a separate access device, you can support higher densities of low-bandwidth access lines than would be possible by simply using an Edge LSR.

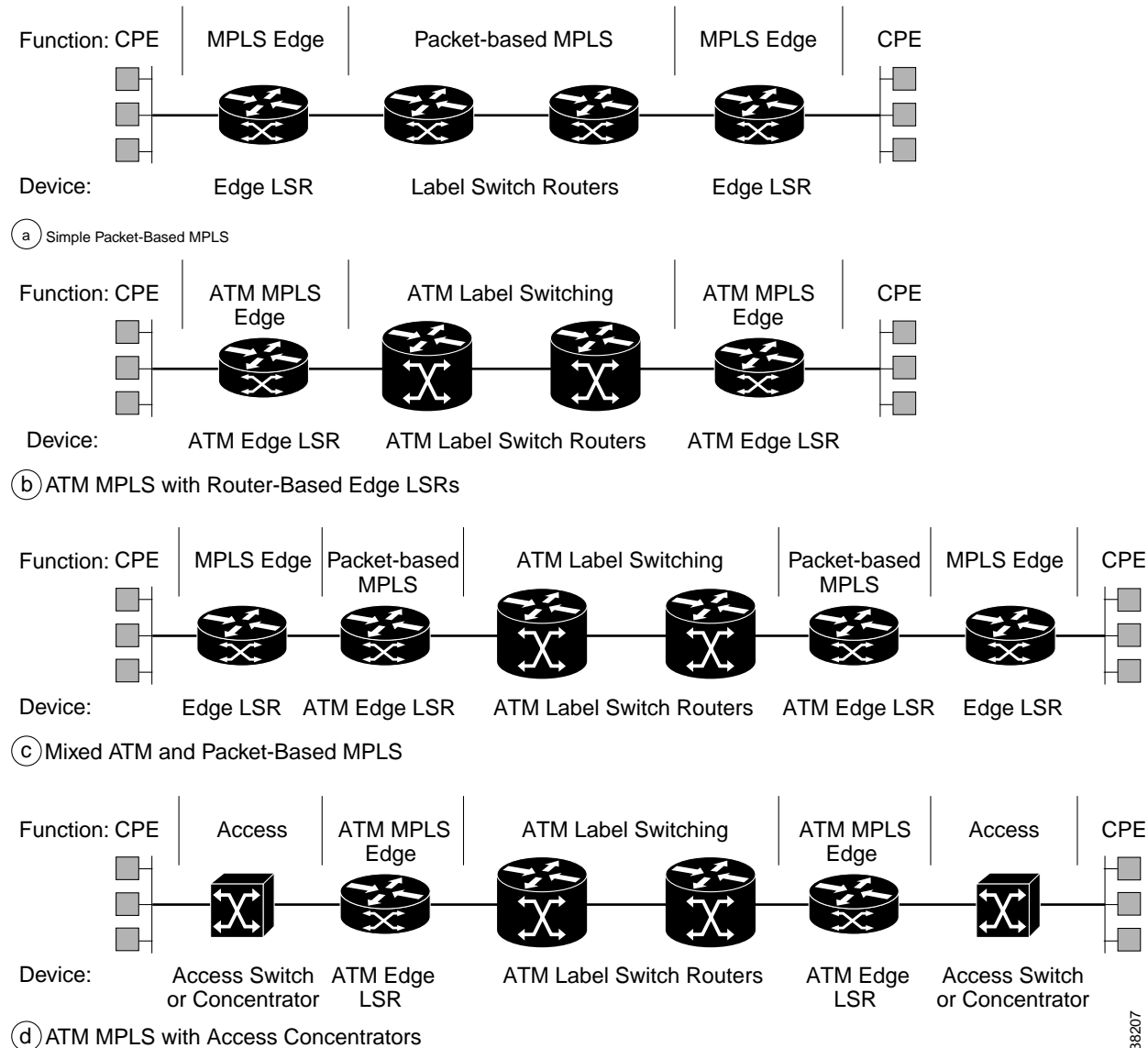
Customer traffic is carried through the access device to the Edge LSR. Between the access device and the Edge LSR, there is a different logical link for each customer. This may be a Frame Relay or ATM PVC, or a PPP link.

## ATM MPLS with Integrated IP+ATM Access Devices

The previous type of network can be simplified if the access device supports Edge LSR function as well as Frame Relay, ATM, or other access services. This is shown in Figure 3-2 Topology (b).

In the case of IP+ATM edge switches, a single device gives access to both MPLS services and PVC or SVC services. IP+ATM edge switches include the BPX 8680, MGX 8850, and 6400 universal access concentrator.

Figure 3-2 Devices in MPLS Networks, Part One



## ATM MPLS Using Traditional ATM Switches

MPLS networks can use traditional ATM equipment as a migration step in introducing MPLS to an existing ATM network. Traditional ATM switches can be used in three ways, as shown in Figure 3-3 Topology (f).

- Backhauling, when the access device is remote from the Edge LSR. The access device is connected to the Edge LSR by PVCs switched through an ATM network.
- Tunnelling through ATM switches between an Edge LSR and an ATM LSR. In the case, the Edge LSR does not need to be adjacent to an ATM LSR, but can be connected through an ATM network.

- Tunelling through ATM switches between ATM LSRs. In this case, the core network uses traditional ATM switches as well as ATM switches.

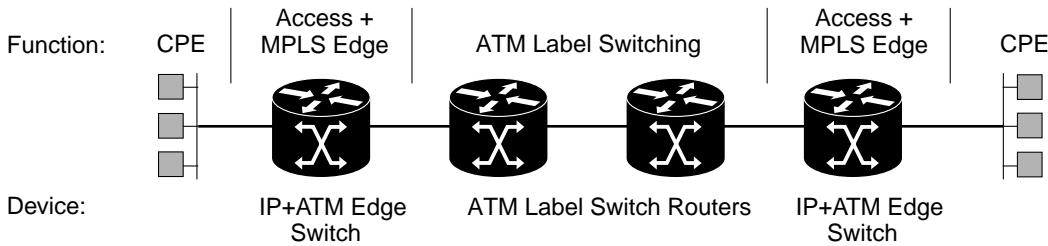
These uses of traditional ATM equipment have disadvantages and must be used with care.

## Dual Backbones

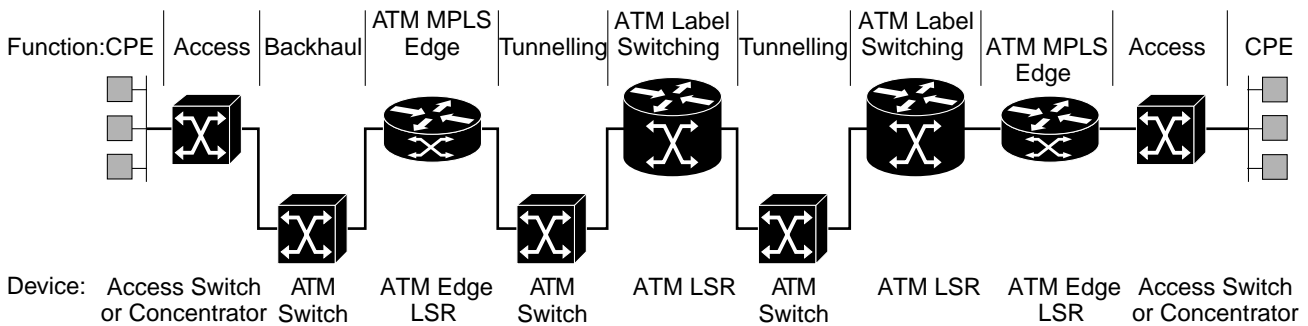
Providers may want to keep an existing ATM infrastructure while building a new MPLS infrastructure (either ATM MPLS or packet-based MPLS) alongside the old infrastructure. Cisco IP+ATM edge devices support this well, allowing customers to access both the MPLS network and services, along with the old ATM network, even from a single access link. This is shown in Figure 3-3 Topology (g).

The IP+ATM access devices can be any of those that can be used in Figure 3-3 Topology (e). The network in Figure 3-3 Topology (g) supports the same functions and services as Figure 3-3 Topology (e), but the Topology (g) network requires more equipment.

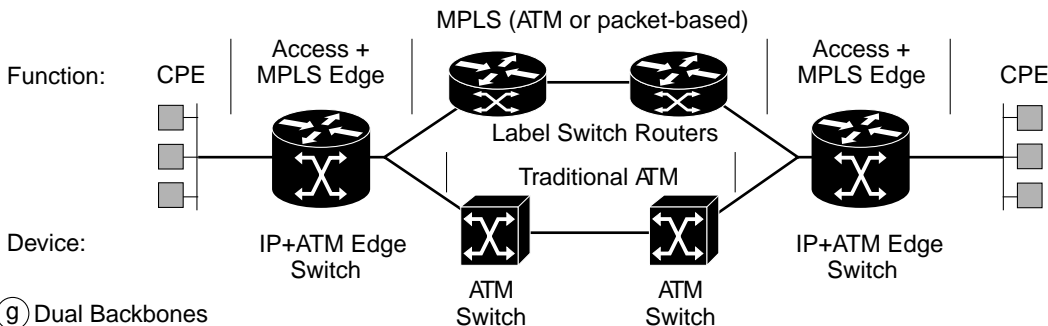
Figure 3-3 Devices in MPLS Networks, Part Two



(e) ATM MPLS with Integrated ATM Edge Device



(f) ATM MPLS with Backhaul and Tunnelling



(g) Dual Backbones

38208

Also see: Dual Backbones: Traditional ATM and ATM MPS or Packet-over-SONET, page 2-23.

# Choosing Cisco MPLS Equipment for ATM

## Choosing ATM MPLS Edge Equipment

There are four main considerations when choosing ATM MPLS edge equipment:

1. Type of services to be offered: IP+ATM, that is, end-to-end PVC and SVC services as well as IP services, or just IP
2. Type of access lines
3. Number of access lines
4. Requirements for redundancy and reliability. Key issues are:
  - Whether the equipment can minimize (warm standby) or completely prevent (hot standby) disruption to data flow in the case of software or hardware failure. Hot standby means zero or almost zero (under 1 second) interruption to end-to-end data flows in the case of equipment failure, with no rerouting beyond the failed equipment.
  - Whether individual components, such as port cards can be hot-swapped.

Redundancy levels for MPLS edge devices can be classified as:

- **None**  
The edge device has no redundancy features. The network must rely on rerouting for reliability. Customer sites must be dual-homed to two different access devices to ensure reliable service.
- **Processor redundancy**  
The edge device has a redundant pair of processors and backplanes with warm or hot standby. Because port cards are not in redundant pairs, the network must still rely on rerouting for reliability. Customer sites must still be dual-homed to ensure reliable service in case of port card failure, but they may be dual-homed to two different cards on the same access device.
- **Full redundancy**  
The edge device has redundant processors and port cards. Customer sites will receive reliable service from the edge equipment even if they are single-homed.

Equipment recommendations based on these requirements are shown in Table 3-1. Cisco will support MPLS on most or all mid-range and high-end routers, most ATM switches, and most access products with routing capability.

Table 3-1 Choosing MPLS Edge Equipment for ATM MPLS Networks


Equipment	Type of Services	Access Lines	Redundancy Support	Comments
3600 router	IP only	Relatively small numbers of async, modem, serial/Frame Relay, 10 Mbps Ethernet, ISDN BRI & PRI, HSSI, E1/T1 serial, Fast Ethernet, OC-3/STM-1 ATM, voice interfaces, and others.	None	Small number of LVCs supported on ATM cards will lead to limitations on MPLS network size. Not recommended for provider ATM MPLS networks.
4700 router	IP only	Relatively small numbers of serial/Frame Relay, 10 Mbps Ethernet, ISDN BRI, E1/T1 serial, Fast Ethernet, E3, T3 or OC-3/STM-1 ATM, and others.	None	Small number of LVCs supported on ATM cards will lead to limitations on MPLS network size. Not recommended for provider ATM MPLS networks.
7200 router	IP only	Serial/Frame Relay up to E1/T1, 10 Mbps Ethernet and Fast Ethernet, ISDN BRI, HSSI, high-speed serial, E3, T3 or OC-3/STM-1 ATM, packet-over-SONET/SDH and others.	None	Minimum recommended for provider networks. PA-A2 CES-ATM port adaptors do not currently support MPLS.
7505, 7507, 7513, or 7576 router	IP only	Serial/Frame Relay or ISDN up to E1/T1, 10 Mbps Ethernet, Fast Ethernet, and Gigabit Ethernet, HSSI, high-speed serial, ATM, packet-over-SONET/SDH and others.	Warm-Standby Processor Redundancy with dual RSPs.	
12008/12012	IP only	<p>POSIP and ATM at OC-3 to OC-48 rates, and Gigabit Ethernet.</p> <p> <b>Note</b> The highest ATM bandwidth density supported by the 12000 series port cards is 1 x OC-12 per slot. Since all traffic in an ATM Edge LSR must go through an ATM interface into the ATM MPLS network, this relatively low ATM bandwidth density of the 12000 limits its capacity as an ATM Edge LSR.</p>	Warm-Standby Processor Redundancy	Suitable for high-speed peerings between providers.
Catalyst 5500 with Route Switch Modules	IP+ATM	10 Mbps and Fast Ethernet, E3, T3, OC-3/STM-1 and OC-12/STM-4 ATM, and others.	None	The Cat 5500 is primarily a LAN switch, but also has limited Edge LSR capability. The Cat 5500 may be connected only to an ATM MPLS network by tunnelling.

Table 3-1 Choosing MPLS Edge Equipment for ATM MPLS Networks (continued)

Equipment	Type of Services	Access Lines	Redundancy Support	Comments
6400	IP + ATM	ATM at E3/T3 to STM-4 rates, also Ethernet and Fast Ethernet.	Warm-Standby Processor Redundancy	MPLS support is not yet shipping.
MGX 8850	IP + ATM	High numbers of 56K/64K Frame Relay, T1/E1 Frame Relay, channelized, and ATM, and higher-speed Frame Relay, serial and channelized T3.	Full Warm-to-Hot Standby.	At FCS, the MGX 8800 will have hot-standby 1:N redundancy capability for customer access lines, hot standby control for PVCs and hot-standby trunks. 1:N warm standby redundancy for RPMs is scheduled for release in CY2000,
BPX 8650	IP + ATM	High numbers of 56K/64K Frame Relay, T1/E1 Frame Relay, channelized, and ATM, ATM at E3/T3 to STM-4 rates, and others.	Excellent redundancy in general, but there is a single point of failure for Edge LSR function.	(See BPX 8680.)
BPX 8680	IP + ATM	High numbers of 56K/64K Frame Relay, T1/E1 ATM, Frame Relay and channelized. Also ATM at E3/T3 to STM-4 rates, and others. Extra 6400, 7200, or 7500 routers (or "label switch controller" packages) may be required to act as Edge LSRs. E3/T3 or faster ATM access lines are used. If IP service is to be supported for large numbers of ATM links at T3/E3 rates and above, it is more cost-effective to use separate stand-alone routers.	Full Warm-to-Hot Standby (but with FCS limitations).	<p>BXM trunk cards must be used. BXM cards are required. MPLS is not supported on BNI cards, except if the BNI cards are used as Feeder Trunks.</p> <p>BCC cards must be BCC3-64 or later. BCC4 cards are strongly recommended.</p> <p>The BPX 8680 can include up to 16 MGX 8850 shelves, with redundancy features as described above. Full redundancy for the combined device relies on redundancy for the label switch controller for the BPX 8600 shelf. A redundant configuration using two simultaneously active controllers is supported in BPX software release 9.3.</p>

There are some practical issues to bear in mind when examining the product specifications for products to be used as ATM Edge LSRs:

- When considering the number of access lines supported, take into account the card slots used by the ATM MPLS interfaces.



For example, a Cisco 7206 router has 6 card slots, and can nominally support 48 Ethernet ports (or 8 per slot). However, when using the Cisco 7206 router as an ATM Edge LSR, at least one slot must be used for an ATM interface. So, the actual Ethernet port capacity of a Cisco 7206 ATM Edge LSR is 40 Ethernet ports.

(See [http://www.cisco.com/warp/public/cc/cisco/mkt/core/7200/prodlit/c7200\\_ds.htm](http://www.cisco.com/warp/public/cc/cisco/mkt/core/7200/prodlit/c7200_ds.htm))

- Some Edge LSRs use card slots for cards performing the Edge LSR function. These must be considered.

For example, the Cisco 6400 Universal Access Concentrator has 8 card slots, but when it acts as an Edge LSR, at least one of these slots must be used for a Node Route Processor card, and not a line card. (See [http://www.cisco.com/warp/public/cc/cisco/mkt/access/dslaggr/prodlit/6400\\_ds.htm](http://www.cisco.com/warp/public/cc/cisco/mkt/access/dslaggr/prodlit/6400_ds.htm))

- Some Edge LSRs can deal with a throughput of MPLS traffic that varies with the number of processing cards.

For example, the Node Route Processor (NRP) card in a Cisco 6400 Universal Access Concentrator can handle roughly 150 Mbps full-duplex of MPLS edge traffic. Assuming a typical activity fraction of 25 percent, a Cisco 6400 with one NRP can handle four OC-3/STM-1 access lines; whereas a Cisco 6400 with two NRPs can handle eight OC-3/STM-1 lines.

The number of slots taken by processor cards also must be considered when calculating the number of access lines supported for any particular configuration.

## Choosing ATM Label Switch Routers

There are five main considerations when choosing ATM LSRs:

1. Type of trunks
2. Number of trunks
3. Number of connections supported
4. Whether VC Merge is required
5. Requirements for redundancy and reliability

Equipment recommendations based on these requirements are shown in Table 3-2. MPLS will be supported on most ATM switches. Consequently, more ATM LSRs will become available in the future. In addition, any traditional ATM switch can be used in a Cisco MPLS network if tunnelling is used, subject to significant limitations.



### Note

---

If redundant pairs of trunks are used, the number of trunks supported will be half the numbers shown.

---

Table 3-2 Choosing ATM LSRs

Equipment	Type and Number of ATM Trunks	Number of Connections Supported	VC Merge Supported	Redundancy Support	Comments
MGX 8850 with PXM1	4xOC-3/STM1	8K. 16K full-duplex connection legs are supported on the PXM card of the MGX 8850. If both legs of all connections are on the PXM card, then 8K connections are supported.	No	Full Warm-to-Hot Standby	The MGX 8850 is intended primarily to be an Edge LSR, but will also has limited ATM LSR capability. This capability will not be available until late in CY 2000.
LS1010	32 x T1/E1 with Inverse Multiplexing over ATM (IMA), 32 x T3/E3, 32 x OC-3/STM-1, 8 x OC-12/STM-4	64K	Yes	None	
6400	16 x T3/E3, 16 x OC-3/ATM-1, 8x OC-12/STM-4	64K	Yes	Warm-Standby Processor Redundancy	ATM LSR support on this switch is not yet shipping.
BPX 8650	144 x T3/E3, 96 x OC-3/STM-1, 24 x OC-12/STM-4	192K	VC Merge is supported in BPX Software Release 9.3, with Enhanced BXM cards.	Some redundancy features at FCS, full redundancy is possible later.	All MPLS interfaces must be on BXM cards. BCC cards must be BCC3-64 or later. BCC4 cards are strongly recommended. The BPX 8650 has supports hot standby trunks and switching fabrics. Full redundancy relies on redundancy for the label switch controller for the BPX 8600 shelf. A redundant configuration using two simultaneously active controllers is supported in BPX Software Release 9.3.

Table 3-2 Choosing ATM LSRs (continued)

Equipment	Type and Number of ATM Trunks	Number of Connections Supported	VC Merge Supported	Redundancy Support	Comments
8540 MSR	64 x T1/E1 with IMA, 64 x T3/E3, 64 x OC-3/STM-1, 32 x OC-12/STM-4, 8 x OC-48/STM-16	256K	Yes	Warm-Standby Processor Redundancy	Some trunk cards are not yet shipping.
MGX 8800 with PXM-45 card(s)	192 x T3/E3, 144 x OC-3/STM-1, 48 x OC-12/STM-4, 12 x OC-48/STM-16	384K	Yes	Full Warm-to-Hot Standby	PXM-45 cards are not yet shipping

## Label Switch Routers Not Based on ATM Switches

LSRs other than ATM switches may be also be used. In particular, the following routers may be used as LSRs:

- 3600 and 4700 series routers (low-bandwidth applications only)
- 6400, 7200, and 7500 series routers
- 12000 series Gigabit Switch Routers

Using these routers, MPLS may be supported over virtually any link type: ATM, packet-over-SONET, Ethernet, and so on. Router-based LSRs do not support native ATM virtual circuit connections, and all except the 12000 series have relatively low capacity compared to Cisco ATM LSRs.

## Designing MPLS Networks

The goal of designing an MPLS network prior to installation is to produce a network that will operate reliably and optimally. Because of the inherently connectionless nature of IP traffic, customers will not be able to tell a carrier exactly which traffic they want to send where. Because of this, it is not possible to perfectly design a network ahead of time.

This section considers these initial design steps result in a working network:

1. Design Points of Presence.
2. Dimension backbone links in the network.
3. Design IP routing.
4. Dimension MPLS Label VC space.
5. Refine the design once the network is operational. This final design step is an ongoing process of optimizing the network design.

## Points of Presence Structures

The design of Points of Presence (PoP) for an ATM MPLS network is constrained by:

- The choice of access line types and equipment for a network.
- Location of PoPs, which is largely determined by where the cities are.
- The population of user sites surrounding each location.

Some typical PoP designs are shown in Figure 3-4.

## Single ATM Edge LSR

Where a single Edge LSR device is sufficient for supporting the number and types of access lines in a Point of Presence (PoP) location, then the simple structure shown in Figure 3-4 Topology (a) is sufficient. Numerous access lines (typically tens or hundreds) are brought into a single Edge LSR, which is connected to the rest of the ATM MPLS network. Numbers and types of access lines supported by single Edge LSRs are described in Table 3-1.

## Multiple Edge LSRs and An ATM LSR

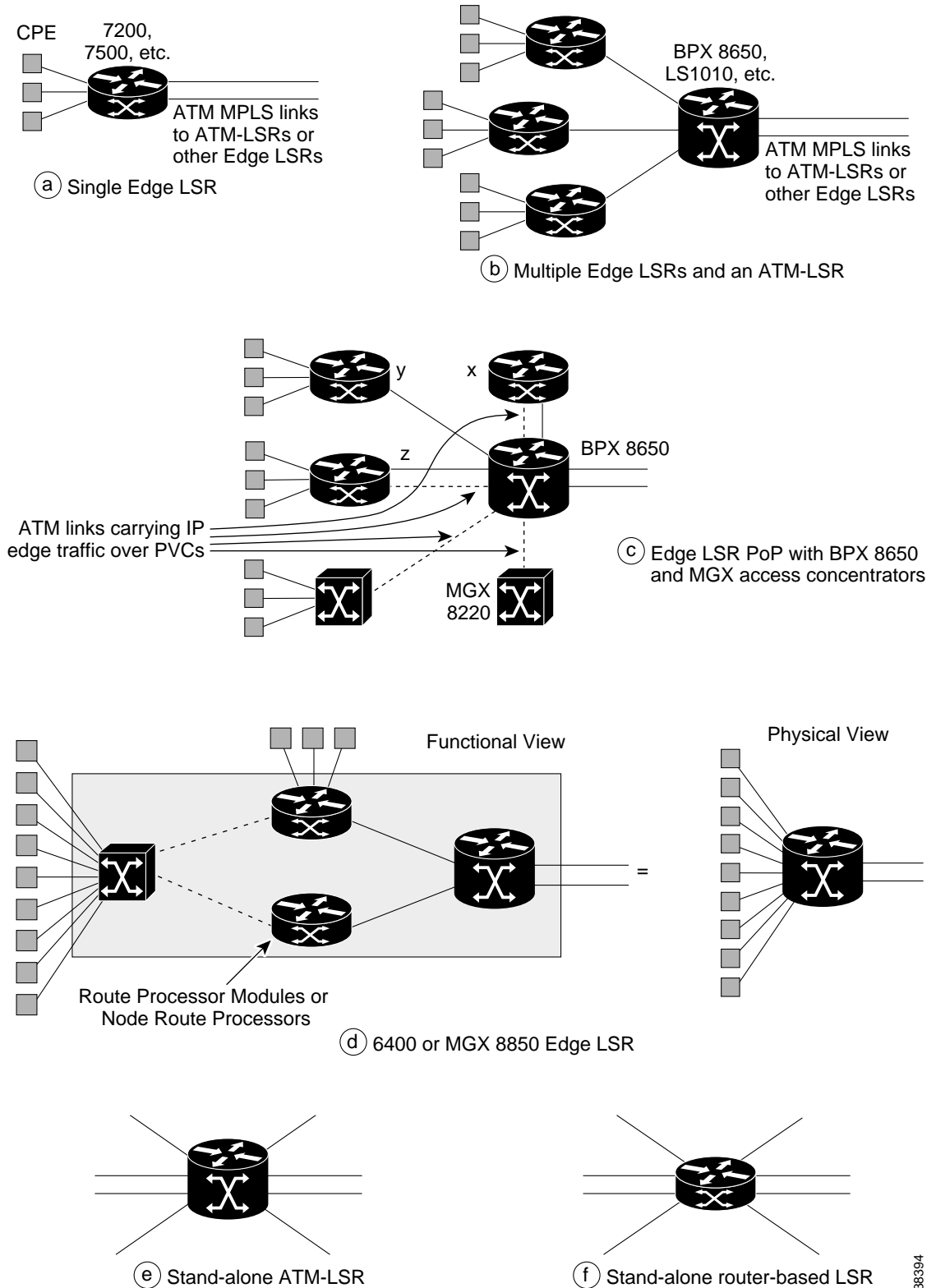
A PoP may require more than one Edge LSR because of a large number of access lines to be supported at that location. Alternatively, different types of Edge LSR might be required because of different types of access lines to be supported. Where there are several Edge LSRs in a PoP, it makes sense to also include an ATM LSR. This is shown in Figure 3-4 Topology (b).

The ATM LSR:

- Locally switches traffic going between different Edge LSRs in the PoP.
- Concentrates traffic going from the PoP onto a single set of ATM MPLS links. The alternative would be either separate sets of links to all Edge LSRs, or using one Edge LSR to carry traffic to the others.
- Improves scalability of routing. Only one set of IP routing protocol (such as OSPF) peerings are required from the ATM LSR to other points in the MPLS network. Without the ATM LSR, separate peerings would be required from all Edge LSRs.

Depending on reliability requirements, redundant pairs of links would be used between the Edge LSRs and the ATM LSR.

Figure 3-4 Point of Presence Structures for ATM MPLS Networks



38394

## Edge LSR PoP with BPX 8650 and MGX 8220 Access Concentrators

An extension to the previous model is to use traditional access concentrators in addition to Edge LSRs and an ATM LSRs. Circumstances where this is appropriate are discussed in “Structures for MPLS Networks” on page 1.

One example of this type of PoP uses MGX 8220 access shelves, 6400, 7200, or 7500 Edge LSRs, and BPX 8650. This is shown in Figure 3-4 Topology (c). IP traffic from access concentrators is carried in ATM PVCs to a Edge LSRs. These may be carried through the same BPX 8650 that acts as an ATM LSR because this is an IP+ATM switch.

Edge LSRs in such a PoP may have three different types of configuration:

- Typically, one or more Edge LSRs will be dedicated to dealing with IP traffic on the edge PVCs. Router X in Figure 3-4 Topology (c) has this configuration. It must have at least two ATM interfaces: one for access PVCs (from the access concentrator) and one for ATM MPLS traffic.
- An Edge LSR dealing with access PVCs may also have customer access lines directly terminating on it. Router Z in Figure 3-4 Topology (c) does this.
- There may also be Edge LSRs in the PoP that do not handle access PVCs at all, and have only directly-connected access lines. Router Y has this configuration.

Note that the label switch controller (LSC) in the BPX 8650 can act as an Edge LSR simultaneously while performing its LSC function. However, use of an LSC as an Edge LSR is not recommended for providers who consider the separation of MPLS control functions from data forwarding functions to be important. As an Edge LSR, an LSC can perform any of the three functions discussed above.

The number of Edge LSRs required in the PoP depends on:

- The total number of access lines
- The total bandwidth of the access lines, calculated from the average utilization. For example, if the sum of the access lines bandwidths were 1 Gbps, the utilization might not exceed 500 Mbps.

The capacity of a 6400, 7200, or 7500 router running MPLS edge functions is roughly the same as its ordinary IP capacity using Cisco Enhanced Forwarding (CEF). For example, a 7200 router with an NPE 200 processor can support close to 200 Mbps of MPLS edge traffic at normal IP packet sizes.

If additional edge functions such as CAR and WRED are used, then performance may be affected. In these circumstances, the performance of the routers should be verified for the particular combination of features to be used. Note also that voice-over-IP packets are very short and will have significantly lower throughput than indicated here.

## Cisco 6400 and MGX 8850 Edge LSRs

The MGX 8850 and Cisco 6400 integrate the functions described in the previous example into a single device, illustrated in Figure 3-4 Topology (d). It consists of:

- A multiservice access concentrator with various types of Frame Relay and ATM access lines, as well as circuit emulation lines. Voice access capability and other types of access lines will be added later.
- One or more Edge LSRs. Each Edge LSR is a route processor module (RPM) card in the case of the MGX 8850, or a node route processor (NRP) card in the case of the 6400. The number of RPMs or NRPs required to act as Edge LSRs depends on:
  - The total number of access lines

- The total bandwidth of the access lines, downrated according to the utilization. For example, if the sum of the access lines bandwidth were 1 Gbps, the utilization might not exceed 500 Mbps.
- An RPM with an NPE150 processor that can support MPLS edge function for 700 access lines. It will support close to 150 Mbps of MPLS edge traffic, at normal IP packet sizes. Again, the performance should be verified with the particular combination of edge functions to be used in addition to MPLS edge function.
- The NRP capabilities are similar to those of the RPM.
- An ATM LSR. In the MGX 8850, one of the RPM cards acts as a label switch controller. It may perform both LSC function and Edge LSR function simultaneously, if desired. Use of an RPM for simultaneous LSC and Edge LSR is not recommended for providers who consider the separation of MPLS control functions from data forwarding functions to be important. In the Cisco 6400, the main Node Switch Processor also acts an LSC.

The 6400 and MGX 8850 also have IP+ATM capability.

## Stand-Alone ATM LSR s

Some sites in a network may have a pure switching role. In an ATM MPLS network, these sites will consist of a single ATM LSR, as shown in Figure 3-4 Topology (e), or possibly a redundant pair of ATM LSRs. The ATM LSR would typically be a BPX 8650.

In some networks, it may make sense to use a router-based LSR instead, as shown in Figure 3-4 Topology (f). A 7500 or 12000 series router may be suitable for this application. Core LSRs usually have Edge LSR capability as well. For example, the BPX 8650 has limited Edge LSR capability as part of its label switch controller.

## Dimensioning An MPLS Network's Links

This section is a guideline to dimensioning the links in a small MPLS network, but it is not the only way of doing this. Different providers will have their own unique procedures for designing and running networks, but they all will be similar to this.

It is not the purpose of initial design to produce a perfect network. Several approximations are made in this process leading to a network that works but is not necessarily optimal. The last step in the design process is to optimize the network, discussed in the “Ongoing Network Design” section on page 3-42.

As an illustrative example, see Figure 3-5, which is a large network distributed across Australia.

1. Design edge points of presence and their layout.

The first step in MPLS network design is to choose the size, type, and layout of the points of presence according to the considerations described above.

The edge PoPs shown in Figure 3-5 Topology (a) are chosen based on the estimated customer link demand shown in Figure 3-5 Topology (b). BPX 8600-based Edge LSR PoPs (which consist of several MGX 8800 shelves and a BPX 8650 for aggregation) are used in Sydney and Melbourne, which have the largest link bandwidths and number of links in this example. An MGX 8800 is used in Brisbane. Adelaide and Perth are smaller centers that can be adequately served by router-based PoPs.

2. Estimate traffic from each point of presence.

Based on the total access line bandwidths, an estimate on the total traffic sent from customers into each PoP can be made. A busy-period estimate should be used, such as of the rate during the busiest minute of the day. This is to ensure adequate dimensioning.

A conservative estimate would be the total of the access line bandwidths at the PoP, as in Figure 3-5 Topology (b). However it will often be reasonable to take a somewhat lower estimate, such as 50 percent of the total access bandwidth, as shown in Figure 3-6 Topology (c)

3. Estimate the traffic matrix.

The exact process for this step will vary from network to network. In Australia, for example, the two main business centers are Sydney and Melbourne, with Sydney being slightly larger. In a large MPLS network for interstate business IP traffic, a reasonable first approximation may be that 50 percent of traffic will go to Sydney, 40 percent to Melbourne, 5 percent to Brisbane, and 2.5 percent to Adelaide and Perth respectively.

An existing service provider would probably already have estimates for traffic patterns for their region. Based on the estimated traffic distribution percentages and the total PoP traffic from Step 2, a traffic matrix can be estimated. The traffic matrix for this example is in Table 3-3.

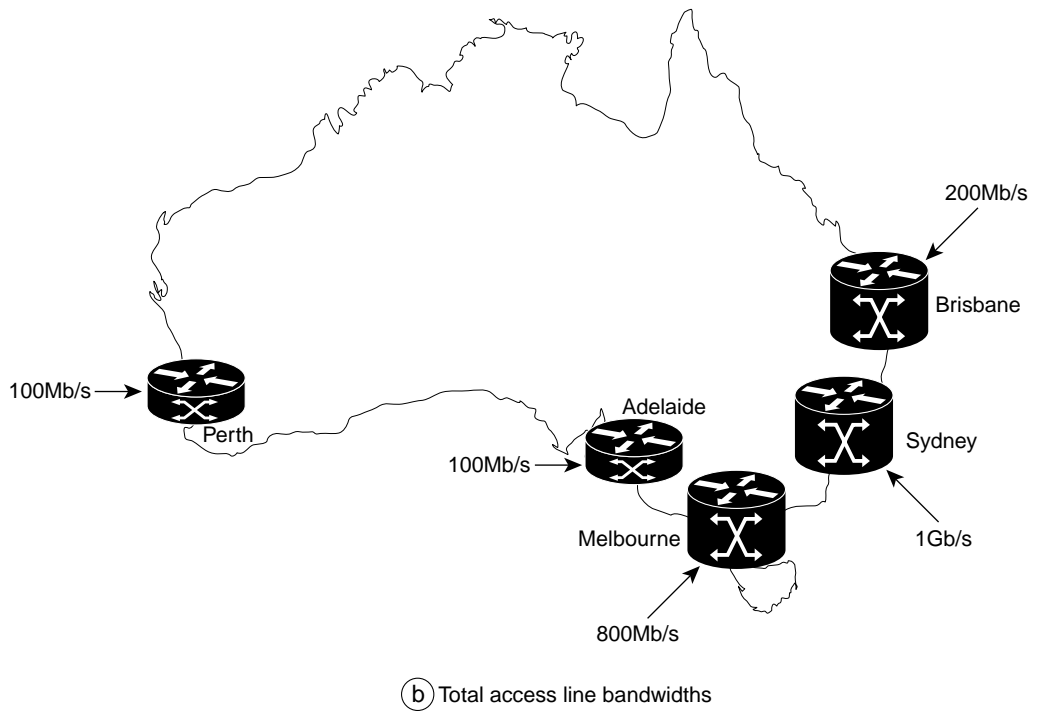
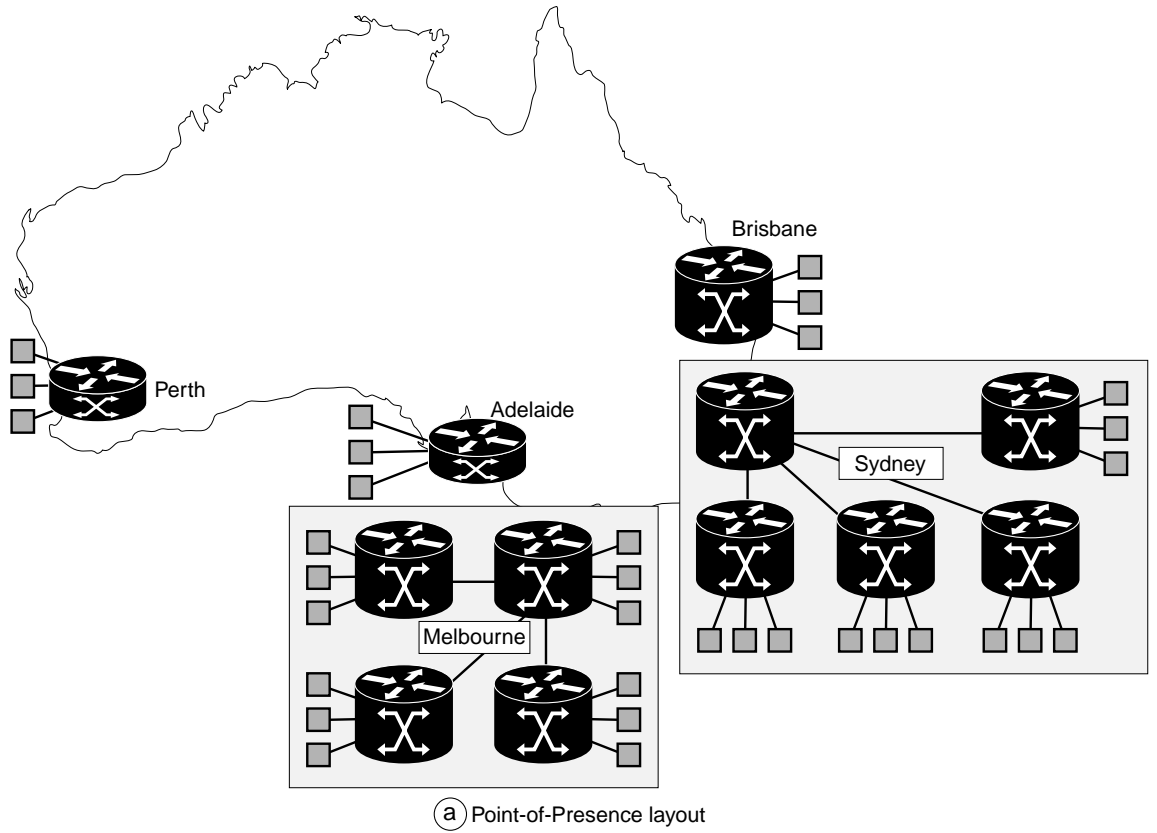
In a typical network, this matrix will be very roughly symmetrical. For example, in Table 3-3, the traffic from Sydney to Adelaide is 12.5 Mbps, but the traffic from Adelaide to Sydney is 25 Mbps. If the traffic were more asymmetrical than about 2:1 or 3:1, then there might be an error in traffic estimates or modeling.

4. Estimate bidirectional traffic flows.

In IP networks, traffic from x to y will often flows along the same path (but in the reverse direction) as traffic from y to x. Although this can be overridden by numerous routing protocol features, it may be useful to assume that this will happen, particularly in small networks. Because of this, it may be easier to use bidirectional traffic flows rather than unidirectional flows in an initial network design.



Figure 3-5 Sample Network in Australia: PoP and Total Access Topologies



38208

**Table 3-3 Network Example: Unidirectional Traffic Matrix**

Traffic Destination	Distribution Percentage	Traffic Source				
		Adelaide	Brisbane	Melbourne	Perth	Sydney
Adelaide	2.5%	1.25	2.5	10	1.25	12.5
Brisbane	5%	2.5	5	20	2.5	25
Melbourne	40%	20	40	160	20	200
Perth	2.5%	1.25	2.5	10	1.25	12.5
Sydney	50%	25	50	200	25	250
Total	100%	50 Mbps	100 Mbps	400 Mbps	50 Mbps	500 Mbps

The estimated bidirectional flows for the sample network are shown in Table 3-4. The bidirectional traffic bandwidth between Adelaide and Sydney, for example, is taken to be 25 Mbps, which is the maximum of the unidirectional bandwidth from Sydney to Adelaide (12.5 Mbps) and the bandwidth from Adelaide to Sydney (25 Mbps). Forming bidirectional flows in this way will tend to slightly overestimate the traffic in the network. This is useful as a conservative first approximation.

**Table 3-4 Network Example: Approximate Bidirectional Traffic Flows**

	Adelaide	Brisbane	Melbourne	Perth	Sydney
Adelaide	1.25				
Brisbane	2.5	5			
Melbourne	20	40	160		
Perth	1.25	2.5	20	1.25	
Sydney	25	50	200	25	250

5. Design the layout of the backbone network.

The layout of the backbone will involve consideration of:

- Geographic layout, that is, good locations for nodes
- Network-level redundancy, that is, having multiple paths to each destination
- Redundancy of trunks

The network layout chosen in this example is shown in Figure 3-6 Topology (d). Many layouts are possible. The one chosen consists of a combination of a partial ring, linking adjacent nodes at the outside of the network, and a star, connecting nodes back to an extra ATM LSR in the core of the network. This provides a good degree of network-level redundancy, with at least two paths between each pair of nodes.

The extra ATM LSR node is placed in Bourke, a town roughly equidistant from four of the five customer PoPs. With a good degree of network-level redundancy, it is not essential to have redundant trunks, because it is possible to reroute MPLS Label VCs.

In traditional connection-oriented networks, rerouting of virtual circuits is a last resort to be used only when all other redundancy mechanisms have failed. This is because it inevitably involves disruption of customer traffic for many seconds or minutes as all circuits are rerouted. In traditional IP networks, rerouting is a much less severe issue, as packet flows can be switched from one link to another almost instantaneously, once the IP routing protocol has converged.

MPLS networks lie between these two extremes. Rerouting in MPLS networks is particularly feasible if VC Merge is used, for two reasons: VC merge reduces the number of VCs which are used in the network and it reduces the scope of changes required in connections when rerouting does occur.

In this example, most trunks are chosen to be non-redundant for economy. A redundant pair of trunks is used for the link which is expected to have the heaviest utilization, namely between Sydney and Melbourne.

6. Estimate link flows based on the estimated traffic.

This involves calculating the paths taken by traffic through the network backbone. With IP routing protocols such as OSPF, this is a straightforward procedure because the IP traffic will follow the minimum-hop path unless administrative costs are used. Where there are two or more minimum-hop paths, the traffic will be approximately balanced across them.

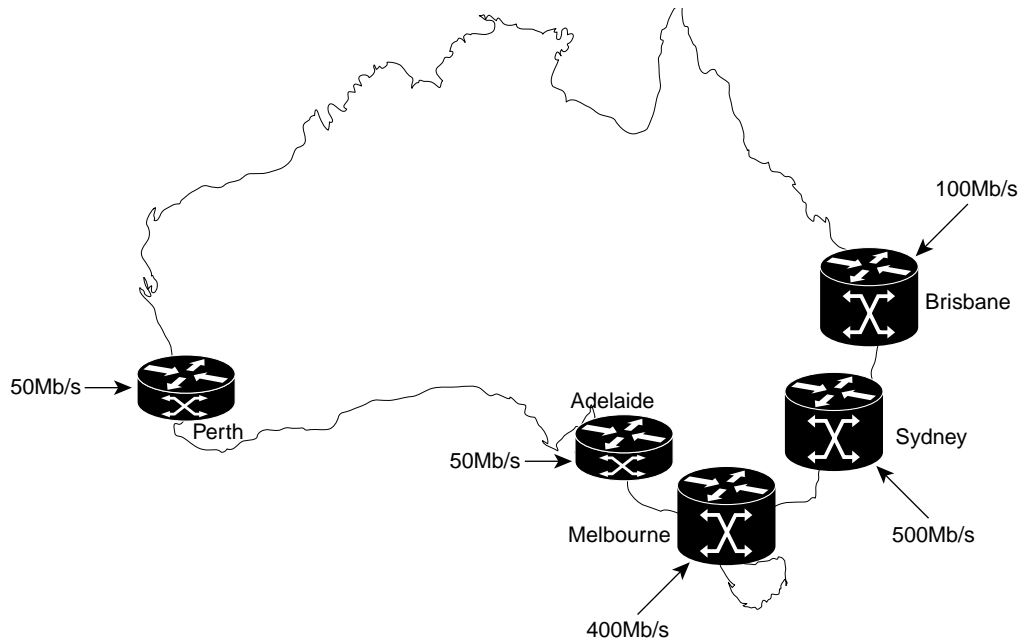
The process of calculation of link flows for the traffic in Table 3-4 is shown in Figure 3-7 (a), and the totals in Figure 3-7 (b)

7. Assign link capacities.

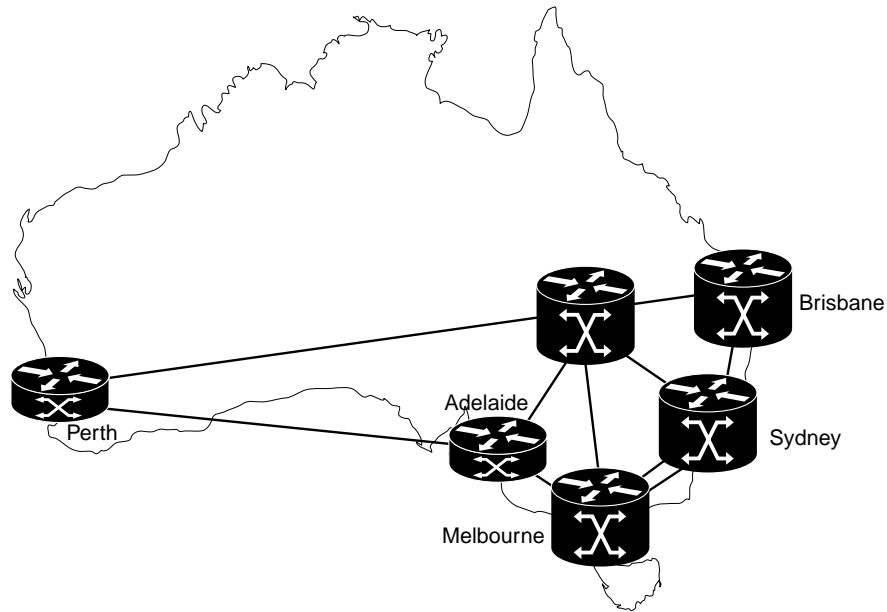
Based on the estimated link flows, link capacities can be assigned to the links in the network. This would generally involve choosing the next standard link size (T3/E3, STM-1, and so on) larger than the link flows just calculated. This is illustrated in Figure 3-6 (c).

8. Adjust for redundancy.

Figure 3-6 Sample Network in Australia



(c) Peak traffic sent from each PoP



(d) Network link design

Where redundant trunks are not used, and the network relies on rerouting for reliability, it may be necessary to adjust link bandwidths to ensure that there is sufficient capacity to deal with link failures.

For example, if the link labelled “r” in Figure 3-7 Topology (c) fails, then links “s” and “t” will need to carry some or all of its traffic. The load on these links would then exceed E3 rates, therefore an STM-1 link (or multiple E3 links) would be required for each of links “s” and “t.”

Similarly, if link “u” failed, then link “v” would require more than E3 capacity. Also, if link “y” failed then the offered load for “w” would exceed E3. So, the final allocation of link bandwidths is as shown in Figure 3-7 Topology (d).

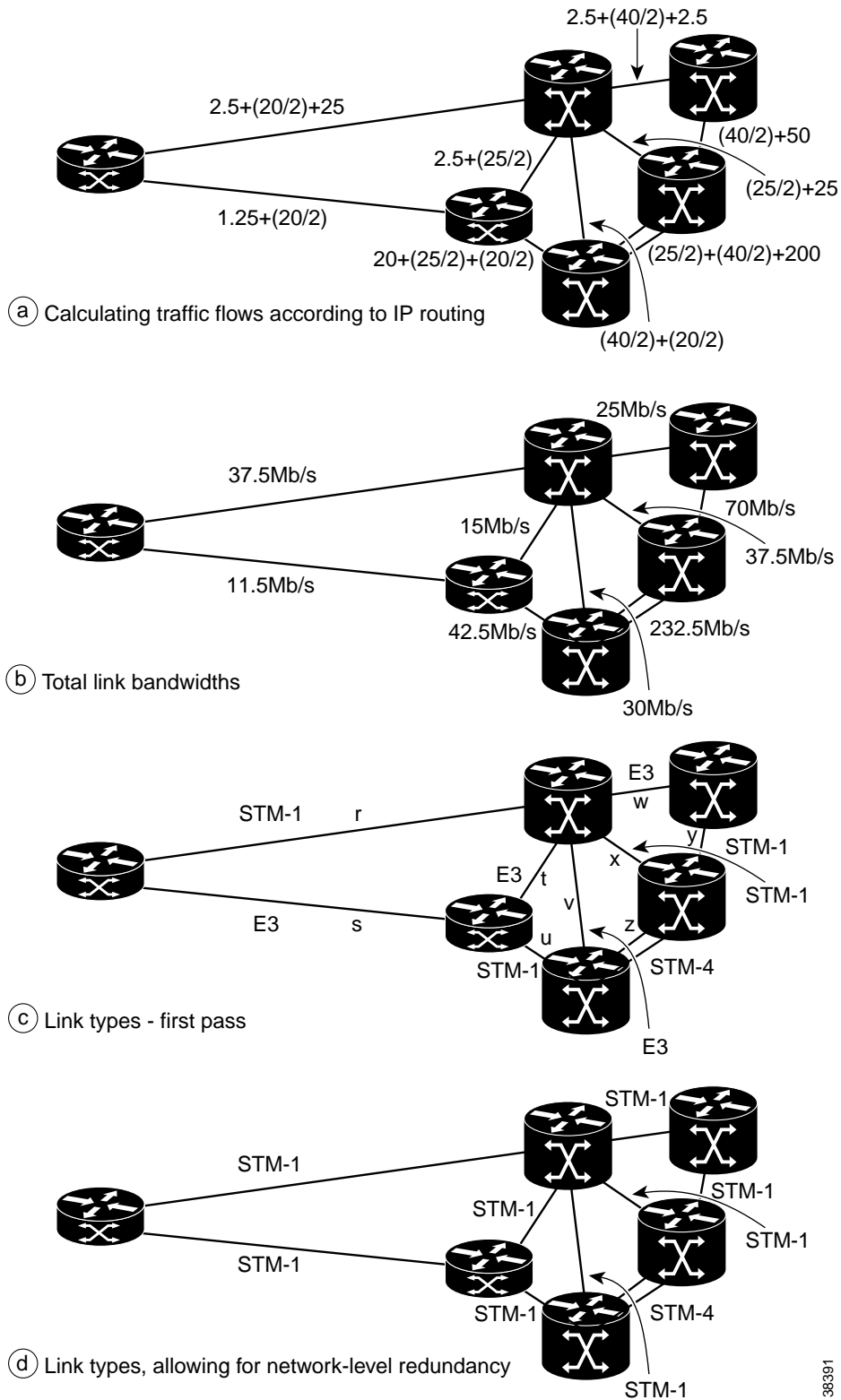
9. Check whether the selected equipment is adequate.

This involves checking in Table 3-1 and Table 3-2 to see whether the selected PoP equipment can support the number and size of links chosen in the network design. The network in this case would pass this check.

Assume the Melbourne PoP had used an MGX 8850 instead of a BPX 8680. This PoP needs two STM-4 links and two STM-1 links, which is not yet supported on an MGX 8850. So, in this case, the PoP would need to be redesigned by using a BPX 8680 instead of an MGX 8850.

Note that any such redesigns, if required, are a relatively minor issue. Many different types of Cisco equipment can be used in ATM MPLS PoPs, and it will usually be found that a PoP can be built to meet the requirements of one location, simply by using combinations of equipment used at other locations. A BPX 8680, for example, combines several MGX 8850 shelves.

Figure 3-7 Network Design Example: Calculating Link Bandwidths



38391

## Redundant Pairs of ATM Links

There are three main ways of achieving change-over for a redundant pair of ATM links:

### 1. Data link-level change-over

This is the normal link redundancy mechanism in traditional ATM networks. Change-over occurs because of physical and Data Link monitoring in the ATM switch or ATM LSR hardware. The switch hardware also typically sets up a copy of all virtual circuit state on the backup link. In a switch with data link-level redundancy, any single link failure will typically result in close to zero data loss on any virtual circuits.

In addition, the network layers and routing (IP or PNNI, and so on) will not be affected by the link failure, or even be aware that it has occurred. However with data link redundancy, the backup link is not available to carry data except in the case of failure of the main link.

Depending on how it is implemented, SONET Automatic Protection Switching (APS) may be a form of data link redundancy. On the MGX 8850 and BPX 8650, SONET APS change-overs result in no change to the interfaces as seen by connection routing and no loss of connection state.

### 2. Inverse multiplexing over ATM (IMA)

IMA carries distributes data over a group of links by distributing cells across the links in round-robin fashion. It offers both data-link level load sharing across links, and redundancy. If one of the links in a group fails, cells are no longer sent on that one link, but the others are still used. IMA is available only for low-speed links—groups of T1 or E1 links.

### 3. Parallel links with network-layer change-over

In this case, a redundant pair of trunks is used, but data-link layer protection is not used at all, and all connection change-over takes place at the network layer. IP or PNNI routing is aware of all link failures and reacts to them.

This is not particularly good for connection-oriented traffic, but works well with IP routing and MPLS. With OSPF equal-cost-multipath or similar, OSPF will choose to balance traffic for every route across both links in a pair of links. This causes a pair of MPLS Label LVCs to be set up for each destination, one per link. If one of the links fails, IP routing will simply divert traffic onto the other, already-established, LVCs. If VC merge is used, this will require no more MPLS signaling, and could be achieved in one second or so.

The advantage of this method is that it allows the bandwidth in the backup trunk to be used, allowing more best-effort traffic to be carried in the network. SONET APS change-overs, as implemented on the non-MSSBU Cisco equipment, are a form of parallel link redundancy, but without the capability of equal-cost multipath to set up back-up links.

Where a redundant link is required, recommendations for use of these modes are:

- **IP+ATM networks**

These should use Inverse Multiplexing for redundancy for low-speed trunks, and otherwise use Data Link redundancy. This avoids costly reroutes of the connection-oriented traffic.

- **Pure ATM MPLS networks**

These should use Inverse Multiplexing for redundancy for low-speed trunks. Otherwise, if the network uses VC merge, parallel links with network-layer change-over should be used, in order to make the full network capacity available for use. Finally, if VC merge is not available, Data Link redundancy should be used.

# IP Routing in An MPLS Network

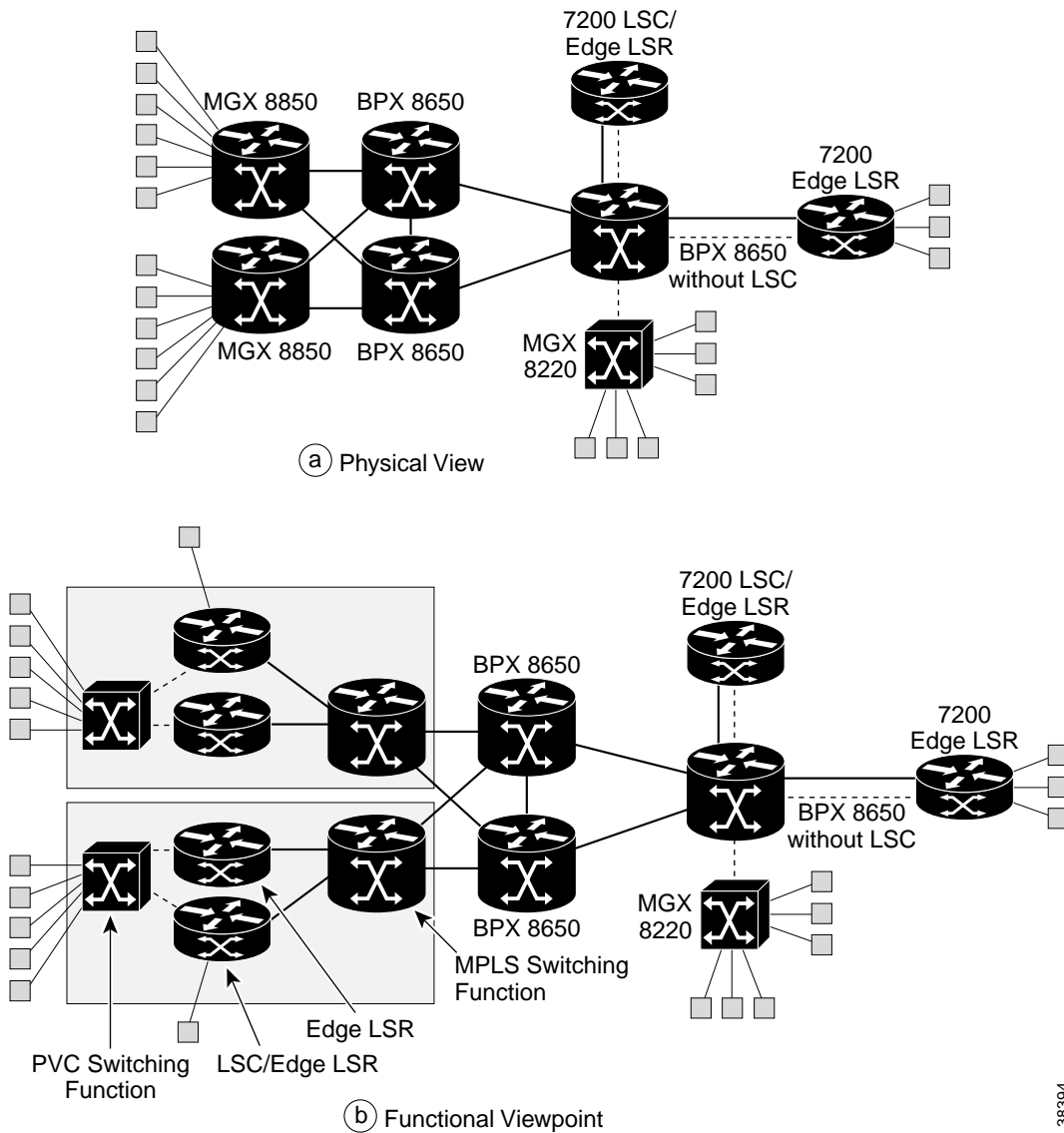
MPLS uses ordinary IP routing protocols—OSPF, IS-IS, and so on—to determine the routes for IP traffic and LVCs. Every LSR runs ordinary IP routing protocols in the same way that ordinary IP routers do.

An important implication of this is that OSPF (or IS-IS, and so on) “sees” an MPLS network as being exactly like an ordinary router network. It is possible to have various viewpoints of an ATM MPLS network:

- **Physical viewpoint**

This viewpoint represents the physical devices and links in a network. An example is shown in Figure 3-8 Topology (a).

Figure 3-8 Viewpoints of An ATM MPLS Network



38394



- **Functional viewpoint**

Where a product has several functions, these can be shown separately. For example, the MGX 8850s in each include two separate Edge LSRs, which are shown separately in Figure 3-8. In addition, it is useful to think of the PVC switching function of an MGX 8850 to be separate from the MPLS switching function. It is sometimes useful to consider the label switch controller (LSC) function in an ATM LSR as being separate from the switching function. This is particularly true if the LSC is also acting as an Edge LSR.

- **Routing viewpoint**

This viewpoint shows the network as it is seen by an IP routing protocol. An example of deriving it is shown in Figure 3-9 Topologies (a) and (b).

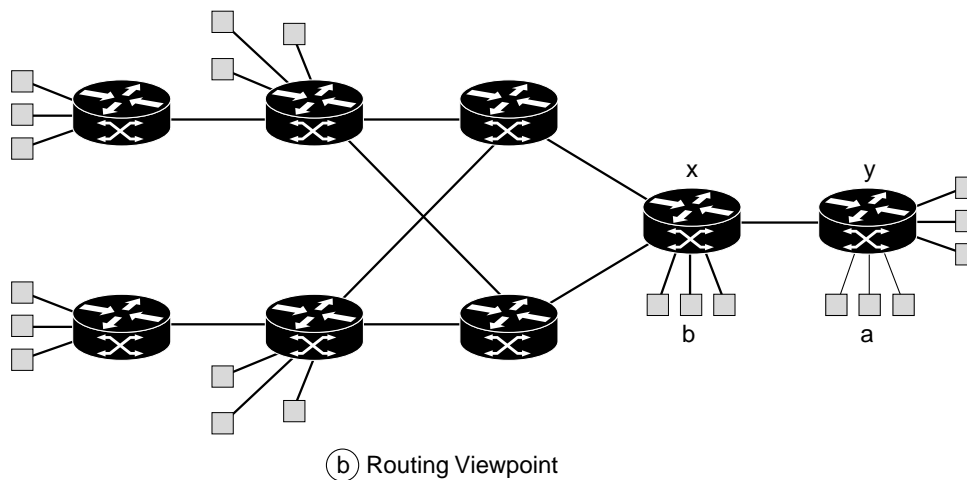
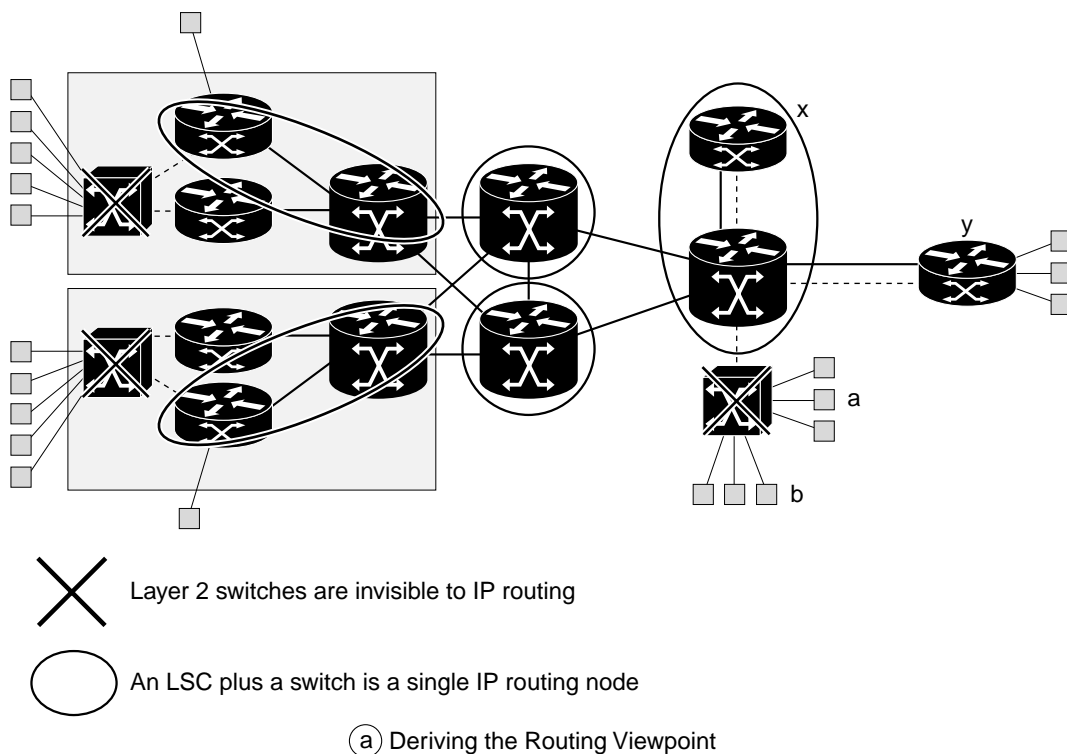
- Layer 2 PVC switches and PVC switching functions are invisible to IP routing. If a customer site is connected to a router by a PVC, then the PVC is a one-hop direct connection from an IP routing perspective. See for example the sites labelled “a” Figure 3-9 (a), and assume that these are connected by PVCs to Edge LSR “y.” Then, in the routing viewpoint, the sites are directly adjacent to router “y.”
- A label switch controller and a switch together form a single routing node.

Using these rules, the routing viewpoint of an MPLS network can be derived. This is shown in Figure 3-9 Topology (b).

Designing IP routing in an MPLS network is almost exactly the same process as designing IP routing for an ordinary IP network. By looking at the routing viewpoint, a network can be divided into areas, route summarization can be designed, and so on.

There are several design guides for IP routing on <http://www.in-cons.cisco.com/~dblack/design-guides/>. Cisco training partners provide a number of courses on IP routing design, and your Cisco account or Jumpstart team will be able to provide other assistance. Also, see the book “Internet Routing Architectures,” mentioned in the Introduction. A few routing issues specific to MPLS networks are considered in the next section.

Figure 3-9 Routing Viewpoints in An ATM MPLS Network



383396

## MPLS-Specific IP Routing Issues

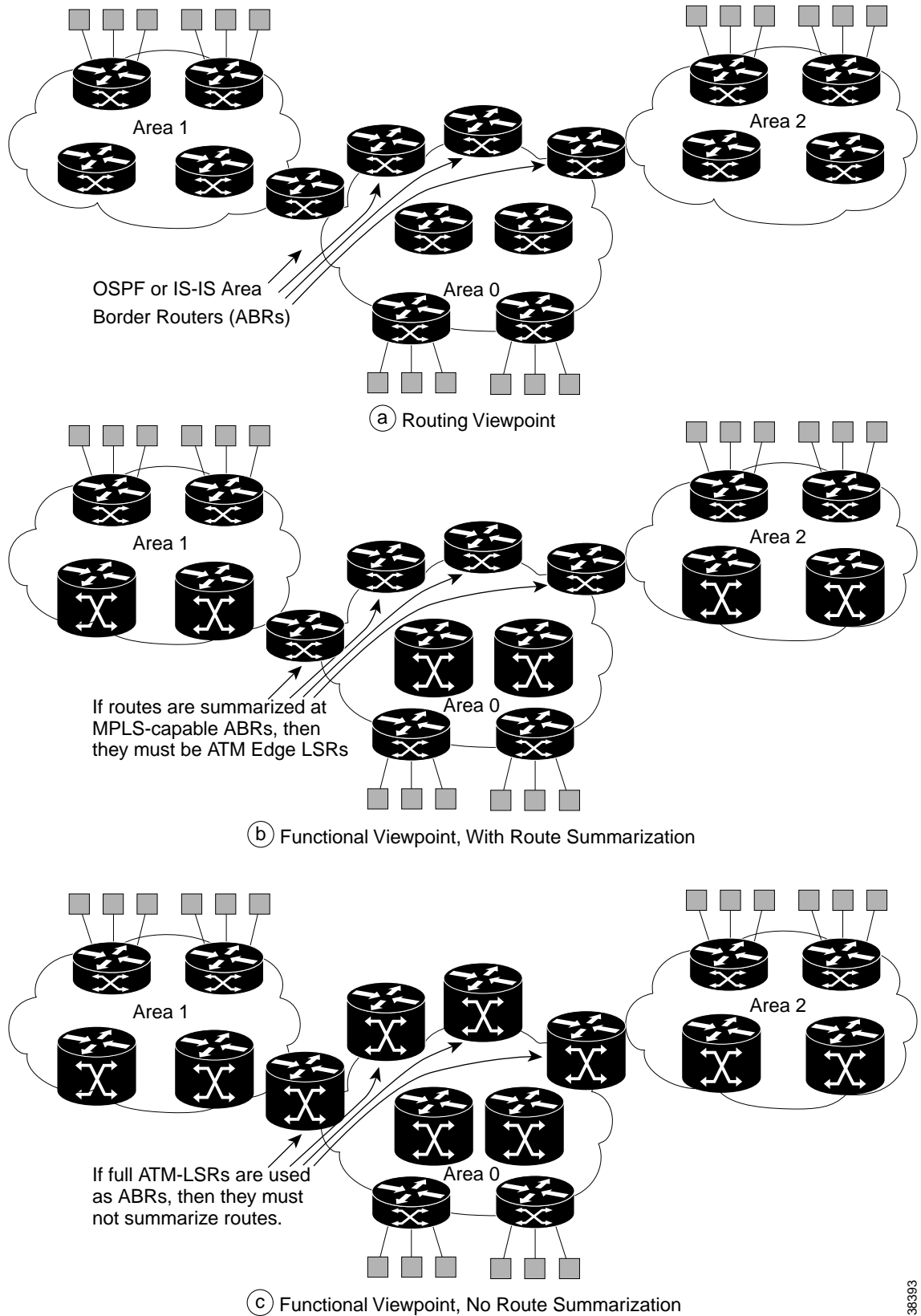
- The interior routing protocol used in MPLS backbones should be either OSPF or IS-IS. IS-IS is currently supported on most Cisco MPLS equipment, but not on the LS1010 and 8540 MSR. EIGRP can also be used, but it will not work with an advanced MPLS-based IP traffic engineering feature called Routing with Resource Reservations (RRR). RRR requires a link-state routing protocol, that is OSPF or IS-IS. (Both the OSPF and IS-IS Working Groups at the IETF are working on extensions to support IP Traffic Engineering, and these capabilities will be used by RRR.) Because EIGRP is a distance-vector routing protocol, it will not work with RRR. IGRP or RIP also will work with MPLS but not RRR, and are not recommended. Note that RRR is sometimes referred to loosely as “MPLS Traffic Engineering,” but is actually a specific type of MPLS traffic engineering.
- Use unnumbered IP links where possible. This reduces the number of IP destinations known to the routers, and hence reduces the number of LVCs used in the network. This is also discussed under “Dimensioning MPLS Label VC Space” section on page 3-29.
- Route summarization must not be done at an ATM LSR. Multiple OSPF or IS-IS areas can be used in an ATM MPLS network, as shown in Figure 3-10. An ATM LSR may be used as an OSPF or IS-IS Area Border Router (ABR), but only if no summarization is done at the Area Border routers. In Figure 3-10 Topology (c), this means that the address prefixes known in all the areas must be the same. An ABR in Figure 3-10 Topology (c) may not, for example, summarize reachability for 1.1.1.0/24, 1.1.2.0/24 and 1.1.3.0/24 with a single route for 1.1.0.0/16. If route summarization is required in an ATM MPLS network, it must be done at an ATM Edge LSR, as shown in Figure 3-10 Topology (b).
- The previous rule also applies to Autonomous Systems and BGP 4. An ATM LSR may not be a BGP Autonomous System Boundary Router, but an ATM Edge LSR may.
- Routing with Resource Reservations (RRR) works best in backbones which contain a single OSPF or IS-IS Area. RRR may not be used in multiple area networks where the Area Border Routers are ATM LSRs. This restriction will be eased in a later version of RRR.
- It is currently impossible to use RRR or TE in a network where the ABRs are ATM LSRs. A future RRR release will allow TE/RRR tunnels to pass through Area Border Routers, enabling RRR to be used in networks where the ABRs are ATM LSRs.
- Route summarization may not be done in the interior of an MPLS VPN network. The interior of a MPLS network supporting VPNs may have multiple OSPF or IS-IS areas, but summarization should not be used.

The restrictions on summarization exist because summarization stops some types of label-switched paths being set up end-to-end.

For example, assume that an ABR summarizes reachability for 1.1.1.0/24, 1.1.2.0/24 and 1.1.3.0/24 with a single route for 1.1.0.0/16. Now assume that a packet with IP address 1.1.1.23 arrives with a label for 1.1.0.0/16. The ABR cannot label-switch the packet. It must look past the label and examine the IP address to find that the packet should go on to 1.1.1.0/24.

Because ATM LSRs cannot examine IP addresses, they may not do IP route summarization. Some ATM LSRs, such as the BPX 8650, have a limited ability to examine IP addresses by sending the packets to the Edge LSR function in their label switch controller. However this can be done only for a small minority of the traffic flowing through the ATM LSR.

Figure 3-10 Multiple Routing Areas and Summarization in An ATM MPLS Network



38393

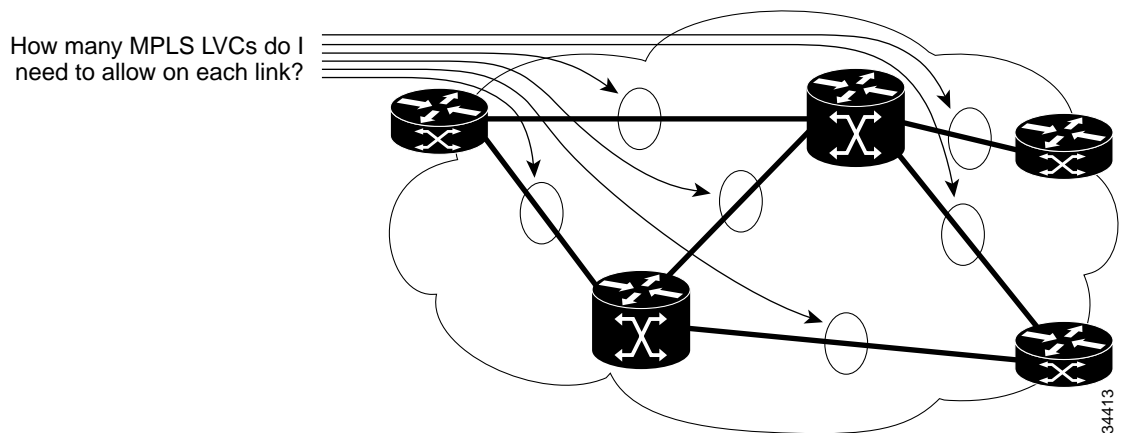
## Dimensioning MPLS Label VC Space

This chapter has shown how many of the issues of designing MPLS networks are similar to those of designing ordinary IP networks. One important exception to this is the dimensioning of MPLS LVC requirements on each link. This design problem is illustrated in Figure 3-11.

In order to complete the design of an ATM MPLS network, a sufficient number of VCs must be reserved for use as LVCs on each link. This can be a problem because any ATM switch can support only a certain number of active VCs. This is particularly important if there are multiple ATM services—MPLS, PNNI, and so on—sharing the resources of the links in an IP+ATM network.

The design problem is to determine the number of LVCs required.

**Figure 3-11 Label VC Requirements**



The required number of LVCs depends on:

- The number of IP destinations in the network
- The relationship between destinations and LVCs
- Whether VC merge is used
- The paths chosen by IP routing

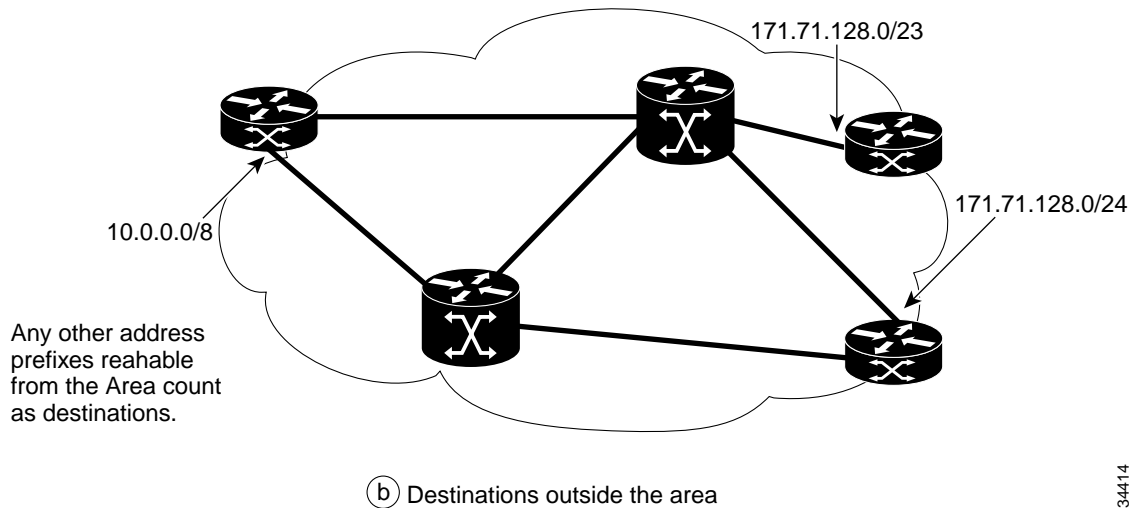
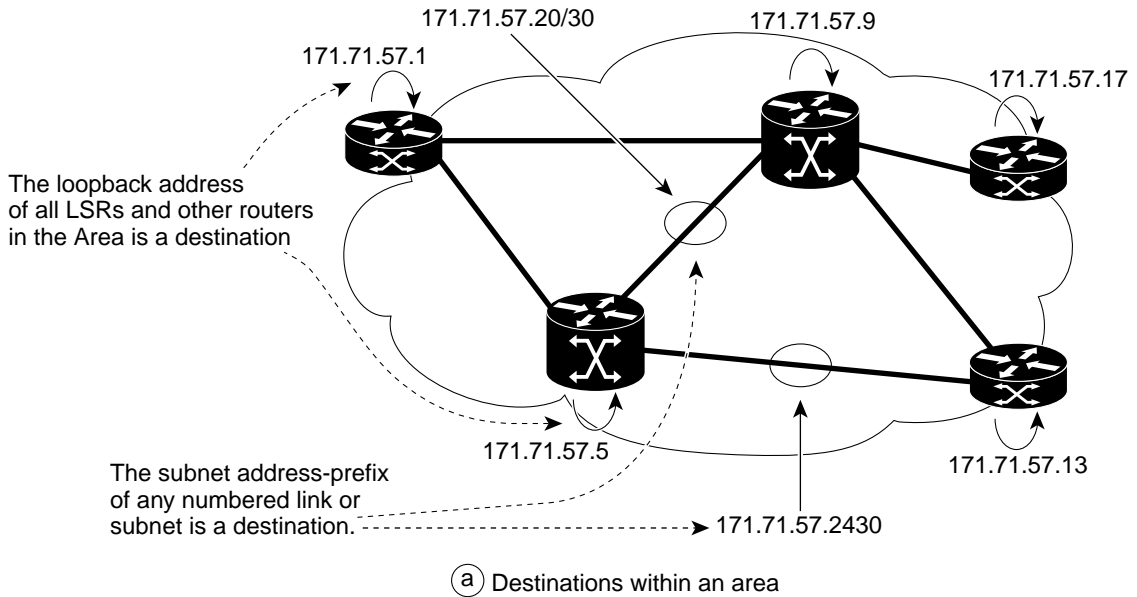
## Destinations

The number of LVCs used in a particular area of a network depends on the number of IP destination-prefixes known in that area. This follows the normal rules for an IP network:

- The loopback address of all LSRs and other routers in the area is a destination prefix.
- The subnet address-prefix of any numbered point-to-point link, or any other subnet, is a destination prefix. Because of this, it is best to use unnumbered links in MPLS networks.
- Any other address prefixes advertised into the area must be counted as well. Note that in MPLS VPN networks, this does not apply to VPN customers' addresses. VPN customers' destination prefixes are not advertised into the core of the network. This is one of the keys to scalability of MPLS VPNs. If many addresses are summarized into a single address at the area border router (or autonomous system border router), then this counts as a single destination prefix.

These destination prefix rules are shown in Figure 3-12.

**Figure 3-12 Destination-Prefixes in An MPLS Network (or Any Other IP Network)**



34414

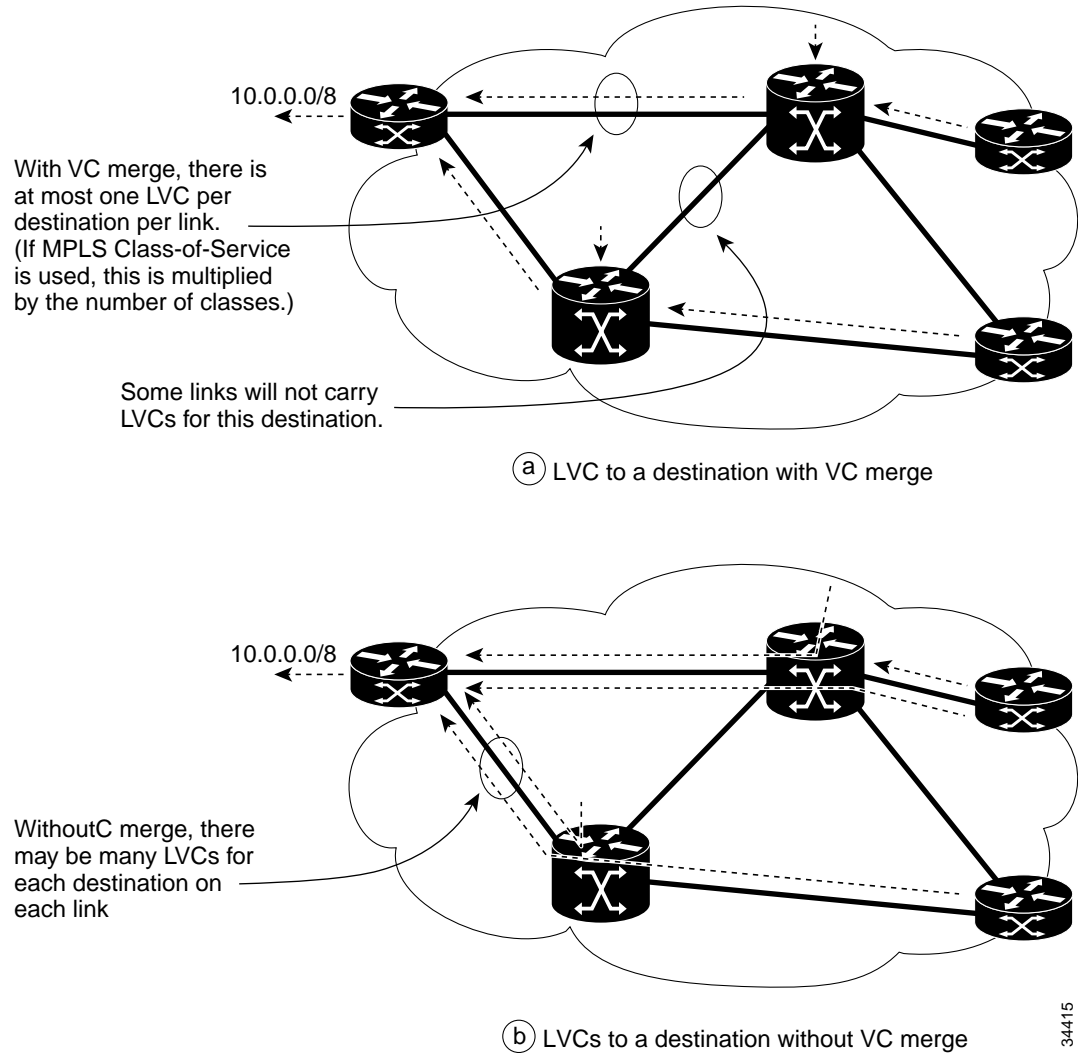
## LVCs Used Per Link and VC Merge

Each ATM Edge LSR and each label switch controller will ask a neighboring MPLS node for LVCs for the destination-prefixes it knows about. If MPLS Class of Service is used, it may ask for up to four LVCs for each destination-prefix.

The requests for LVCs flow through the network according to the paths chosen by IP routing. With VC merge, the LVCs to each destination will be merged at each ATM LSR. This means that on each link, there is at most one LVC per destination in the Area. This is shown in Figure 3-13 Topology (a).

If MPLS Class-of-Service is used, then this is multiplied by the number of classes. If VC merge is not used, there may be many more LVCs.

**Figure 3-13 LVCs to Each Destination**



## Design Calculations: Edge LSRs

For ATM Edge LSRs, the number of LVCs used per link depends on whether VC merge is being used in the network.

### Equation 1

Let  $d$  be the number of destination-prefixes known in an area, and  $c$  be the number of Classes of Service used in the network. If VC merge is used, then the number of LVCs used per link  $l$  satisfies

$$l \leq d$$

**Equation 2**

If VC merge is not being used in the network, there are three dependencies:

- the number of LSCs in the area
- the number of Edge LSRs in the area
- how many destinations are directly reached through the Edge LSR in question

If  $d_e$  is the number of destinations reachable through a particular ATM Edge LSR (this will often equal 1, due to summarization) and the total number of ATM Edge LSRs and LSCs in the area is  $n$ , then the number of LVCs used per link satisfies

$$l \leq (d - d_e) + cnd_e$$

**Equation 3**

A simpler equation applies in the particular case where all these conditions exist:

- VC merge is not used
- There is one destination prefix per Edge LSR or LSC
- All links are unnumbered
- There are no address prefixes from outside the area

These conditions will often apply in the core of MPLS networks supporting VPNs, but not using VC Merge.

The number of LVCs used per link on the ATM Edge LSR in this case is given by:

$$l < 2cn$$

One of the preceding three equations is then used to check whether a sufficient number of LVCs is available on the equipment, as shown in Table 3-5. Table 3-6 shows the LVC capacity of Cisco ATM Edge LSR interfaces.

**Table 3-5 Checking the LVC Limits of Edge LSR**

Device	Situation	Key Parameter	Equation
Edge LSR	The network uses VC merge.	Number of active VCs supported per ATM link.	Equation 1
Edge LSR	The network does not use VC merge; there is one destination-prefix per LSR or Edge LSR, all links are unnumbered, and there are no out-of-area routes.	Number of active VCs supported per ATM link.	Equation 3
Edge LSR	The network does not use VC merge, all other situations apply.	Number of active VCs supported per ATM link.	Equation 2



**Table 3-6 Cisco ATM Edge LSRs and LVC Capacity**

Device	Interface Hardware	Number of Active LVCs supported	Notes
3600	NM-1A ATM Network Modules	1024	
4700	NP-1A ATM Network Processor Module	1023	
7200, 7500	PA-A1 or standard ATM port adaptor	2048	
Catalyst 5500, 7200, 7500	PA-A3 ATM port adaptor.	4096	
6400	Node Route Processor (NRP)	2048	Capacity is reduced by one LVC for each active PVC that terminates on the NRP.
MGX 8850 IP+ATM switch	Route Processor Module (RPM)	4096	Capacity is reduced by one LVC for each active PVC that terminates on the RPM. In addition, the PXM is limited to 16K LVCs. This is unlikely to be a problem unless more than three RPMs are used in an MGX 8850 shelf.
12000 series routers	4xOC-3 ATM Line Card	2047	The 2047 active VCs are shared between all four ports. Network capacity is reduced by one destination prefix for every second and subsequent route chosen for each destination according to equal-cost multipath routing, if the extra routes are on the same card.
12000 series routers	1xOC-12 ATM line card	2047	

## Edge LSR Examples

Consider a network where VC merge is being used and one Class of Service is being used. If the Edge LSRs are all 7200 series routers with PA-A3 port adaptors, then how many IP destination prefixes can safely be supported in the area?

VC merge is being used, so Table 3-5 indicates that Equation 1 should be used. One Class of Service is being used, so  $c = 1$ . Table 3-6 states that the PA-A3 port adaptor supports 4096 LVCs, so  $l = 4096$ . Substituting these into Equation 1 gives

$$4096 \leq d$$

Which is equivalent to:  $d \geq 4096$

This means that 4096 destination-prefixes are guaranteed to be supported within the area, provided that the ATM LSRs do not impose a tighter limit (this discussion considers only the Edge LSRs).

Consider a network where VC merge is not used, and four Classes of Service are being used. This network is the core of an MPLS VPN service, and there is one destination-prefix per LSR or Edge LSR. All links are unnumbered. No out-of-area routes are injected into the interior routing protocol. The Edge LSRs are 7200 and 7500 series routers with a mixture of PA-A1 and PA-A3 ATM port adaptors.

What is the largest number of LSRs that can be used if the network consists of a single area? Assume that the ATM LSRs support a sufficiently large number of LVCs.

According to the conditions given, Table 3-5 indicates that Equation 3 should be used. Four Classes of Service are used, so  $c + 4$ . The Table 3-6 shows ATM interfaces each support 2048 or 4096 LVCs. The interfaces with 2048 LVCs give a tighter restriction, so use  $l = 2048$ . Substituting these into Equation 3 gives

$$2048 < 2(4)n$$

Which is equivalent to:  $n > 256$

This means that a maximum of 256 LSRs (Edge LSRs or ATM LSRs) may be used in the area, provided that the IP routing protocol supports that many routers in an area.

Consider a network where:

- VC merge is not used
- 4 Classes of Service are used
- This network is the core of an MPLS VPN service
- There is one destination prefix per ATM LSR or Edge LSR
- All links are unnumbered
- The network has multiple areas and there are at most 100 ATM LSRs and LSRs in each area
- All the Edge LSRs' ATM port adaptors are PA-A3
- The ATM LSRs support a sufficiently large number of LVCs

How many LSRs can be used in the entire network?

There are multiple areas and so there will be out-of-area routes in each area. Table 3-5 indicates that Equation 2 should be used in this case. Four Classes of Service are used, so  $c = 4$ . Table 3-6 gives  $l = 4096$ . There are at most 100 ATM LSRs and LSRs in each area, so use  $n = 100$ .

Observe that there is one route per Edge LSR. In the worst case, all out-of-area routes are accessed through a single LSR—this concentrates the LVC requirements on the links to that single LSR. In this case,  $d_e = (d - 100)$ . Substituting these into Equation 2 gives

$$4096 \leq 4(100) + 4(100)(d - 100)$$

$$(d - 100) \geq 3696 / 400$$

which is equivalent to:  $d \geq 109$

This means that only 109 LSRs can be used in the network.

By comparison with the previous example, we can see that using multiple areas can have major disadvantages in ATM MPLS networks without VC merge. These examples indicated that ATM LSRs without VC merge typically cannot be used in networks of larger than a few hundred nodes. An alternative that works around these limitations is to use the same switches, but to use MPLS-over-PVCs instead of ATM MPLS.

## Design Calculations: ATM LSRs with VC Merge

With VC merge, the LVCs to each destination will be merged at each ATM LSR. This means that there is at most one LVC per destination on each link, as shown in Figure 3-13 Topology (a). If MPLS Class-of-Service is used, then this is multiplied by the number of classes.

### Equation 4

If  $d$  is the number of destination-prefixes known in an area, and  $c$  is the number of Classes of Service used, then the number of LVCs used per link  $l$  satisfies

$$l < cd$$

### Equation 5

Another important issue in switches that support VC merge is the number of LVCs that must be merged together in the switch,  $m$ . This depends on the number of links into the switch  $k$ . The limit is:

$$m < cd(k - 1)$$

These equations are then used to check whether a sufficient number of LVCs is available on the equipment, as shown in Table 3-7. Both equations must be checked.

**Table 3-7 Checking the LVC Limits of ATM LSRs with VC Merge**

Device	Key Parameters	Check Against
ATM LSR with VC Merge	<ol style="list-style-type: none"> <li>1. Number of active VCs supported per ATM link.</li> <li>2. Number of merging LVCs supported per switch, or per port card, whichever is applicable to the switch architecture.</li> </ol>	Equation 4 Equation 5

In general, a per-switch limit applies to shared-memory switches such as the LS1010 or 8540 MSR. A per-port card limit applies to crossbar switches such as the BPX 8650.

In Table 3-8, the numbers of active LVCs supported are maximums; the actual limits will be dependent on configurations. In a BPX 8650, for example, the actual number of active LVCs supported per link must be down-rated by a minimum of 270 lines per interface if AutoRoute is enabled on that interface. On all switches, the VC space reserved for PVCs, SVCs, and so on, must be subtracted from the available VC space.

**Table 3-8 Cisco ATM LSRs and LVC Capacity, If VC Merge Is Used**

Device	Interface Hardware	Number of Active LVCs Supported	Number of Active Merging LVCs Supported
LS1010	Any ATM port hardware.	4096 per OC-3 port, 16K per OC-12 port, 16K per OC-48 port	64K per switch
6400	Any ATM port hardware.	4096 per OC-3 port, 16K per OC-12 port, 16K per OC-48 port	256K per switch
8540 MSR	Any ATM port adaptors.	4096 per OC-3 port, 16K per OC-12 port, 16K per OC-48 port	256K per switch

**Table 3-8 Cisco ATM LSRs and LVC Capacity, If VC Merge Is Used (continued)**

Device	Interface Hardware	Number of Active LVCs Supported	Number of Active Merging LVCs Supported
BPX 8650 or 8680	BXM-E cards	32K per BXM, shared amongst up to 12 interfaces	32K per BXM, with a maximum of 16K per port on OC-3 BXM cards and 2xOC-12 BXM cards. T3/E3 BXM cards and 1xOC-12 BXM cards have a limit of 32K per port.
MGX 8800 with PXM-45	AXSM cards	128K per AXSM, shared amongst up to 16 interfaces	128K per AXSM

## ATM LSRs with VC Merge: Example 1

A network uses BPX 8650 ATM LSRs with VC merge. Two Classes of Service are used. Each BPX 8650 has 4x1-port OC-12/STM-4 BXM cards, with each port used to link to another ATM LSR or Edge LSR.

What limit do these ATM LSRs put on the number of IP-destination-prefixes that can be supported inside an area?

Table 3-7 shows that both Equation 4 and Equation 5 must be checked. Two Classes of Service are used, so  $c = 2$ . Each switch has four ports, so  $k = 4$ . Looking up the BPX 8650 in Table 3-8 shows that BXM cards support 32K active LVCs. In this case, each BXM card has one port, so each link supports 32K LVCs, or  $l = 32768$ . Table 3-8 shows that 32K LVCs can be merged into a 1-port OC-12 BXM card, so  $m = 32768$ .

Substituting these parameters into Equation 4 gives

$$32768 < 2d$$

which is equivalent to:  $d > 16k$

Substituting the parameters into Equation 5 gives

$$32768 < 2d(4 - 1)$$

which is equivalent to:  $d > 5461$

The limit from Equation 5 is tighter, which means that the limit imposed by the ATM LSRs is 5461 destination-prefixes in the area. (The Edge LSRs might impose a tighter limit.)

## ATM LSRs with VC Merge: Example 2

A network uses 8540 MSR ATM LSRs with VC merge. Four Classes of Service are used. Each 8540 MSR has 8xOC-3/STM-1 with each port used to link to another ATM LSR or Edge LSR.

What limit do these ATM LSRs put on the number of IP-destination-prefixes that can be supported inside an area?

Table 3-7 shows that both Equation 4 and Equation 5 must be checked. Four Classes of Service are used, so  $c = 4$ . Each switch has eight ports, so  $k = 8$ . Looking up the 8540 MSR in Table 3-8 shows that OC-3 port cards support 4096 LVCs, or  $l = 4096$ . Similarly, Table 3-8 shows that the 8540 MSR supports 256K merging VCs, so  $m = 262144$ .

Substituting these parameters into Equation 4 gives

$$4096 < 4d$$

which is equivalent to:  $d > 1024$

Substituting the parameters into Equation 5 gives

$$262144 < 4d(8 - 1)$$

which is equivalent to:  $d > 9362$

The limit from Equation 4 is tighter, which means that the limit imposed by the ATM LSRs is 1024 destination-prefixes in the area. (The Edge LSRs might impose a tighter limit.)

## ATM LSRs with VC Merge: Example 3

A network uses BPX 8650 ATM LSRs with VC merge. Four Classes of Service are used. Each BPX 8650 has eight ports, on 2x4-port OC-3/STM-1 BXM cards, with each port used to link to another ATM LSR or Edge LSR.

What limit do these ATM LSRs put on the number of IP-destination-prefixes that can be supported inside an area?

Table 3-7 shows that both Equation 4 and Equation 5 must be checked. Four Classes of Service are used, so  $c = 4$ . Each switch has eight ports, so  $k = 8$ . Looking up the BPX 8650 in Table 3-8 shows that BXM cards support 32K active LVCs. In this case, each BXM card has four ports, so we can assume that each link supports 32K/4 LVCs, or  $l = 8192$ . Similarly, Table 3-8 shows that 4xOC-3 BXM cards can support 32K merging VCs, with a maximum of 16K per port. The worst case is when all LVCs try to merge into the same port, so use so  $m = 16384$ .

Substituting these parameters into Equation 4 gives

$$8192 < 4d$$

which is equivalent to:  $d > 2048$

Substituting the parameters into Equation 5 gives

$$16384 < 4d(8 - 1)$$

which is equivalent to:  $d > 585$

The limit from Equation 5 is tighter, which means that the limit imposed by the ATM LSRs is 585 destination-prefixes in the area. (The Edge LSRs might impose a tighter limit.)

## Design Calculations: ATM LSRs without VC Merge

### Equation 6

Without VC merge, there may be many VCs per destination on each link, as shown in Figure 3-13 Topology (b). If the total number of ATM Edge LSRs and LSCs in the area is  $n$ , then there may be up to  $c(n-1)$  LVCs per destination on each link.

The number of LVCs used per link will then satisfy

$$l < cd(n - 1)$$

## Equation 7

A tighter limit applies in the particular case where VC merge is not used, and there is one destination prefix per Edge LSR or LSC, and all links are unnumbered, and there are no address prefixes from outside the area. These conditions will often apply in the core of MPLS networks supporting VPNs but not using VC Merge.

The number of LVCs used in this case is given by:

$$1 \leq (n^2 / 2)$$

Either Equation 6 or Equation 7 is then used to check whether a sufficient number of LVCs is available on the equipment, as shown in Table 3-9. Table 3-10 shows the limits of Cisco ATM LSRs without VC merge capability.

**Table 3-9** Checking the LVC Limits of ATM LSRs without VC Merge

Device	Situation	Key Parameter	Check Against
ATM LSR without VC Merge	There is one destination-prefix per LSR or Edge LSR, all links are unnumbered, and there are no out-of-area routes.	Number of active VCs supported per ATM link.	Equation 3
ATM LSR without VC Merge	All other situations.	Number of active VCs supported per ATM link.	Equation 2

**Table 3-10** Cisco ATM LSRs and LVC Capacity, If VC Merge Is Not Used

Device	Interface Hardware	Number of Active LVCs Supported per Link
BPX 8650	Older BXM cards (or pre-9.3.x software)	16K per BXM, shared amongst up to 12 interfaces



**Note**

The numbers of active LVCs per link are maximums; the actual limits will be dependent on configurations. In a BPX 8650, for example, the actual number of active LVCs supported per link must be down-rated by a minimum of 270 lines per interface if AutoRoute is enabled on that interface. On all switches, the VC space reserved for PVCs, SVCs, and so on must be subtracted from the available VC space.

## ATM LSRs without VC Merge, with One CoS: Example 1

A network uses BPX 8650 ATM LSRs without VC merge. One Class of Service is used. Each BPX 8650 has 2x4-port OC-3/STM-1 BXM cards, with each port used to link to another ATM LSR or Edge LSR. There is one destination-prefix per LSR or Edge LSR. All links in the area are unnumbered, and there are no out-of-area routes known.

What limit do these ATM LSRs put on the number of LSRs or Edge LSRs that can be supported inside an area?

Table 3-9 shows that Equation 7 should be checked in this case. One Class of Service is used, so  $c = 1$ . Looking up the BPX 8650 in Table 3-8 shows that BXM cards support 16K active LVCs. In this case, each BXM card has four ports, so each link supports  $16K/4$  LVCs, or  $l = 4096$ .

Substituting these parameters into Equation 7 gives

$$4096 \leq 1(n^2 / 2)$$

$$(n^2 \geq 8192)$$

which is equivalent to:  $n > 90$

This means that the limit imposed by the ATM LSRs is 90 LSRs or Edge LSRs.

These examples indicate that ATM LSRs without VC merge typically cannot be used in networks of larger than a few hundred nodes. An alternative that works around these limitations is to use the same switches, but to use MPLS-over-PVCs instead of ATM MPLS.

## ATM LSRs without VC Merge, with Two CoS: Example 2

A network uses BPX 8650 ATM LSRs without VC merge. Two Classes of Service are used. Each BPX 8650 has 4x1-port OC-12/STM-4 BXM cards, with each port used to link to another ATM LSR or Edge LSR.

What limit do these ATM LSRs put on the number of IP-destination-prefixes that can be supported inside an area?

No information is given on the relationship between devices and routes, so Table 3-9 shows that Equation 6 should be checked. Two Classes of Service are used, so  $c = 2$ . Looking up the BPX 8650 in Table 3-8 shows that BXM cards support 16K active LVCs. In this case, each BXM card has one port, so each link supports 16K LVCs, or  $l = 16384$ .

Substituting these parameters into Equation 6 gives

$$16384 < 2d(n-1)$$

$$d(n-1) > 8192$$

Because  $n$  the number of LSRs in the area has not been given, this question cannot be explicitly answered as stated. However, if it is assumed that  $n = 50$  this would give an indicative value of  $d > 167$ .

In other words, the number of destination prefixes that may be supported depends on the number of LSRs in the area with, for example, a limit of 167 destination prefixes if there are 50 LSRs in the area.

## Additional Example Considerations

The following sections provide further considerations about the preceding examples.

### Internet Routing Tables

The limits on destination prefixes indicated in the previous examples are much smaller than the size of the Internet backbone routing table, which is about 70,000 routes. Despite this, ATM MPLS can still be used in networks with full Internet routing, by use of an MPLS feature known as BGP Next-Hop Labelling.

BGP Next-Hop Labelling allows BGP Autonomous System Boundary Routers (ASBRs) to exchange the full Internet routing table with each other by way of BGP, while readvertising only a limited subset of these addresses (or none at all) into the Interior routing protocol (OSPF or ISIS) Areas through which they are connected. Because only a limited set of destination-prefixes is known on OSPF or IS-IS in the MPLS network, the limits discussed here are sufficient even though they are much smaller than the Internet routing table.

Cisco MPLS Virtual Private Networks (VPNs) extend the BGP Next-Hop Labelling technique to deal with address from many different customers' VPNs.

## Traffic Engineering

The limits shown in Equations 4 to 7 apply when MPLS Traffic Engineering is not used. If Traffic Engineering is used, then one LVC will be used for each Traffic Engineering Tunnel on each link, in addition to the limits shown above.

## VP Tunnels

VP Tunnels involve several logical links terminating on a single physical interface on an LSR or ATM LSR. When VP Tunnels terminate on an interface, the LVCs on all VP Tunnels must be taken into account. For example, if four VP Tunnels terminate on a logical interface that supports 4000 LVCs, then an average of only 1000 LVCs will be available per VP Tunnel.

## Alternative Calculations

The limits shown in Equations 4 to 7 can be quite loose. The number of LVCs actually used on a particular link may be much less than these limits suggest, particularly if VC merge is not being used. However, it is difficult to calculate exactly how many will be required. This depends on the exact shape and state of the network, and the exact paths chosen by IP routing. If this can be analyzed, taking account of such things as failed links and multipath routing, then fewer LVCs could be safely reserved on each link: a very complex process. In any case, the limits shown above will be safe.

A network that has run out of cross-connects is basically not functioning. There is no way to control which LVCs are created and which are not. This may result in permanent black holes or overloaded LSCs. The network should be designed from the start so that it will not run out of connections.

Two different behaviors might occur when a link runs out of LVCs, depending on the type of traffic:

- **Send IP Traffic on Default LVCs**

The first case applies to ordinary IP traffic that would otherwise be carried with a single MPLS label. An Edge LSR deals with LVC set-up failure by sending the ordinary IP traffic on the default (0,32) LVCs that are otherwise used for TDP/LDP signaling. This traffic is then forwarded around the resource-starved links by the LSC processors, and not the ATM switch fabrics.

This will not affect network stability, provided that MPLS Class of Service is used to give precedence to routing and TDP/LDP traffic. However, if large amounts of user traffic are sent on the (0,32) LVCs, then this user traffic will experience poor performance when it exceeds the LSCs' packet-forwarding capacity.

- **Discard Packets on Failed Paths**

The second case applies to advanced MPLS services such as VPNs. In this case, the packets that would otherwise use the failed paths are discarded. For these services, it is impossible to correctly deliver a packet unless it is fully labelled. This discarding of packets to certain destinations results in routing "black holes."

In case of LVC starvation, it is usually impossible to predict which LVCs are created and which are not. This is because LVCs are created roughly in the order in which routing converges. This is, in turn, dependent on random factors such as which links fail, and the exact timing of link failures compared to routing protocol update timers.

In other words, it is very important to design the network and allocate switch resources so that a sufficient number of LVCs are available on each link.



# Ongoing Network Design

Network design is an ongoing process. Once an ATM-MPLS network is deployed, ongoing design activities are required to:

- Verify the assumptions used in the initial design
- Adjust the network as new customers and PoPs are added

The ongoing process will involve the following steps:

1. Measure actual PoP and link traffic, and compare against
  - The predicted traffic
  - The link capacities
2. Based on the comparison between predicted and actual traffic, some modeling assumptions might be changed. For example, the traffic distribution across nodes might be different than that initially predicted. Review the initial design and dimensioning if the modeling assumptions are changed. Lower and higher-bandwidth links might be required.
3. As customers are added, and as traffic increases, review the initial design to:
  - Add new Edge LSRs to PoPs
  - Add new links to the network
  - Adjust routing
  - Check for sufficient LVC allocation on links





## Quality of Service in MPLS Networks

---

This chapter considers the role of the Quality of Service (QoS) architecture in designing MPLS-based IP+ATM networks. A summary example is provided for configuring BPX 8650 ATM LSRs, their associated LSCs (6400, 7200, or 7500 series), and Edge Label Switch Routers:

- MPLS QoS with IP+ATM Overview
- The Differential Services Approach to Quality of Service
- Modeling Network Traffic Flows to Meet Service Level Agreements
- A Recommended Process for Estimating and Modeling Traffic
- MPLS Traffic Engineering
- More Stringent Quality of Service in IP+ATM Networks
- Quality of Service for MPLS VPNs
- Discard Policies
- Delay Limits
- Alternative Service Types

For configuration procedures for BPX 8650 ATM LSRs and their Edge Label Switch Routers, see Chapter 6, “MPLS CoS with BPX 8650.”

For additional information, refer to Cisco 6400, 7200, or 7500 series router and MPLS-related IOS documentation. Refer to 9.3 Release notes for supported features.

### MPLS QoS with IP+ATM Overview

As part of their VPN services, service providers may wish to offer premium services defined by Service Level Agreements (SLAs) to expedite traffic from certain customers or applications. Quality of Service (QoS) in IP networks gives devices the intelligence to preferentially handle traffic as dictated by network policy. QoS mechanisms give network managers the ability to control the mix of bandwidth, delay, jitter, and packet loss in the network.

QoS is not a device feature; it is an end-to-end system architecture. A robust QoS solution includes a variety of technologies that interoperate to deliver scalable, media-independent services throughout the network, with system-wide performance-monitoring capabilities.

The actual deployment of QoS in a network requires a division of labor for greatest efficiency. Because QoS requires intensive processing, the Cisco model distributes QoS duties between edge and core devices that could be multilayer switches or routers. Edge devices do most of the processor-intensive

work, performing application recognition to identify flows and classify packets according to unique customer policies. Edge devices also provide bandwidth management. Core devices expedite forwarding while enforcing QoS levels assigned at the edge.

MPLS-enabled networks make use of Cisco IOS QoS features to build an end-to-end QoS architecture:

- **IP Precedence**  
This feature uses three bits in the IP header to indicate the service class of a packet (up to eight classes). This value is set at the edge and enforced in the core. In IP+ATM networks, different labels are used to indicate precedence levels.
- **Committed Access Rate (CAR)**  
CAR manages bandwidth allocation for certain traffic types. To enforce customer network policies, managers can configure multiple Layer 3 thresholds based on the desired parameters, such as application or protocol. If a flow exceeds a given threshold, managers can provision a variety of responses, from dropping excess packets to sending them at a lower service class.
- **Weighted Random Early Detection (WRED)**  
This feature prevents network congestion by detecting and slowing flows (according to service class) before congestion occurs.
- **Class-Based Weighted Fair Queuing (CBWFQ)**  
This feature provides the ability to reorder packets and control latency at the edge and in the core. By assigning different weights to different service classes, a switch can manage buffering and bandwidth for each service class. Because weights are relative and not absolute, under-utilized resources can be shared between service classes for optimal bandwidth efficiency.

The key to an effective, network-wide IP QoS plan is scalability. Applying QoS on a flow-by-flow basis is not practical because of the huge numbers of IP traffic flows in carrier-sized networks. A scalable way to provide higher levels of service quality with minimal loss in granularity is to implement multiple service classes, or Classes of Service (CoSs).

For example, a service provider network may implement three service classes: a high-priority, low-latency class, a guaranteed-delivery “mission-critical” service, and a low-priority “best-effort” class. Subscribers can use the mix of services that suits their needs. Some subscribers may wish to use a guaranteed-delivery, low-latency service for their videoconferencing applications, and best-effort service for e-mail traffic.

MPLS makes it possible to apply scalable QoS across very large routed networks and Layer 3 IP QoS in ATM networks because providers can designate sets of labels that correspond to service classes.

In routed networks, MPLS-enabled QoS substantially reduces processing throughout the core for optimal performance. In ATM networks, MPLS makes end-to-end Layer 3-type services possible.

Traditional ATM and Frame Relay networks implement CoS with point-to-point virtual circuits, but this is not scalable because of high provisioning and management overhead. Placing traffic into service classes at the edge enables providers to engineer and manage classes throughout the network. If service providers manage networks based on service classes, rather than point-to-point connections, they can substantially reduce the amount of detail they must track and increase efficiency without losing functionality.

Compared to per-circuit management, MPLS-enabled CoS in ATM networks provides virtually all the benefits of point-to-point meshes with far less complexity. Using MPLS to establish IP CoS in ATM networks eliminates per-VC configuration. The entire network is easier to provision and engineer.

For much IP traffic, the requirements for quality of service are fundamentally weaker and more flexible than requirements for traditional virtual-circuit-switched data traffic. In order to be competitive, providers' IP networks must provide Service-Level Agreements (SLAs) of an appropriate form for IP traffic with relatively weak Quality of Service (QoS) requirements at low cost. The networks must also provide stronger QoS for certain traffic.

Cisco MPLS networks have unique flexibility for meeting all requirements for IP QoS:

- They support Classes of Service (CoS). Service providers might offer Service Level Agreements (SLAs) for some Classes of Service, giving agreed quality-of-service levels, averaged over periods of minutes or hours. These can be engineered and provided at low cost.
- Stronger Quality-of-Service (QoS) guarantees may be provided for traffic that is critically important or otherwise intolerant of varying QoS.
- They provide scalable and easy-to-manage ways of providing quality of service to Virtual Private Networks and individual users and sites within VPNs.

## Best Effort Traffic and IP QoS Requirements

A major reason for the success of the Internet is that IP treats connectivity as being the fundamental requirement of a communications network. The Internet has been successful in its the goal of allowing any host to communicate with any other, without setting up virtual circuits, reserving bandwidths, or performing any other actions with high overhead or costs.

In the Internet, considerations such as QoS are not treated as being as important as allowing any-to-any connectivity. Because of this philosophy, IP traffic is, in general, extremely tolerant of varying QoS. The typical World Wide Web user will not care whether a Web page downloads at 100 Kbps or 5 Kbps.

TCP, the transport protocol for over half of IP traffic, automatically adjusts to the available end-to-end bandwidth and loss. UDP has generally been used only by applications that are tolerant of packet loss. Although there are now IP applications that do require stronger QoS, most IP traffic still requires only a very loose guarantee of connectivity, meaning that the available bandwidth between a given pair of IP addresses should be at least a few hundred bits per second, and that the delay be no more than a couple of seconds. This traffic merely requires “best effort” from the provider networks.

These requirements are stated in very loose terms because there is no hard minimum QoS required for TCP/IP traffic. TCP/IP adapts to QoS even worse than the figures suggest but the average user is likely to find the resulting application performance to be frustratingly poor.

Certain IP traffic does require better QoS guarantees, particularly Voice-over-IP and similar real-time applications. Good QoS ensures good application performance. Some users may prefer to migrate in the longer term towards the networks that give, at reasonable cost, the best QoS for “best effort” traffic. Despite this, it is likely that a large proportion of the traffic in any MPLS network will continue to be very tolerant of widely varying QoS in day-to-day operations.

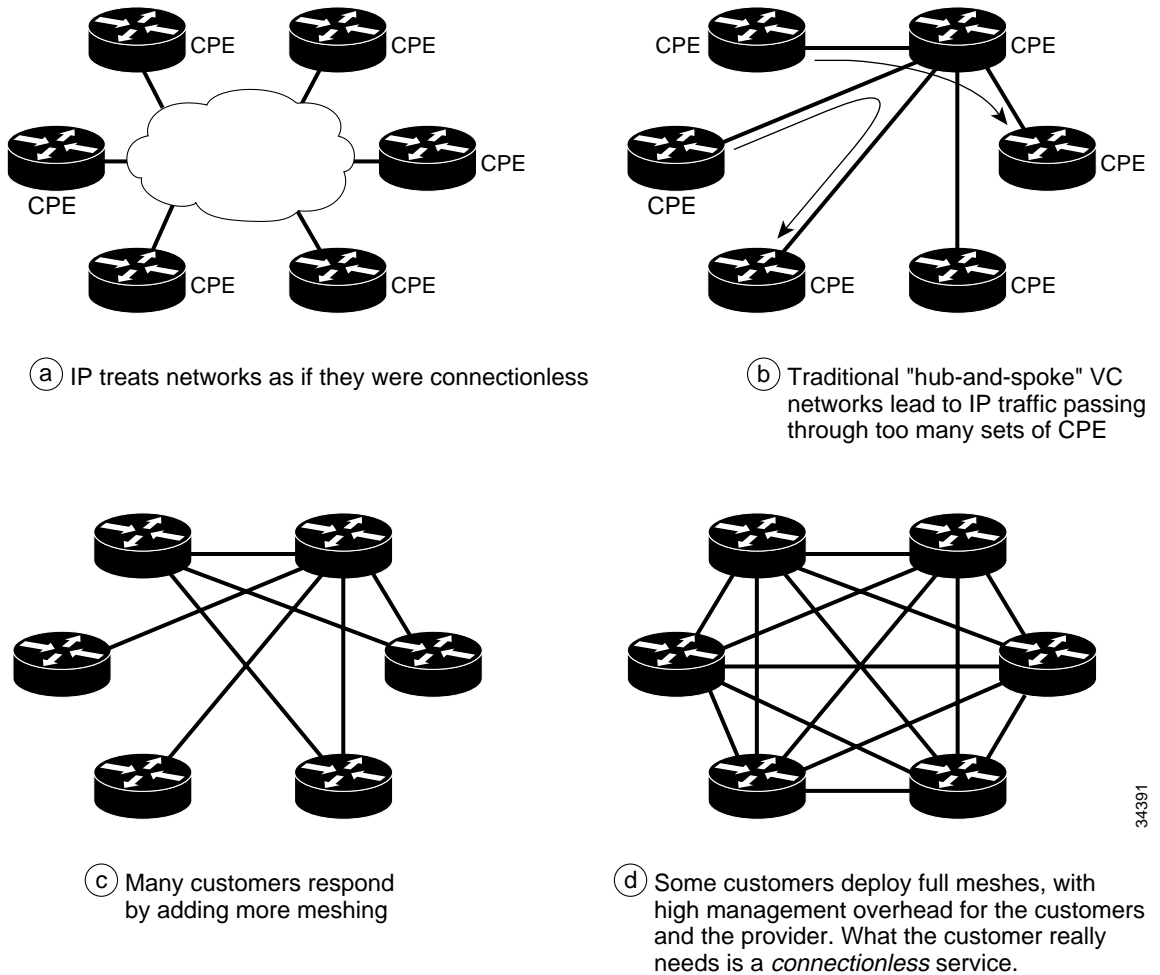
## Effects of Connectionless Traffic

Another important aspect of IP traffic is that it is connectionless. While this is obvious, its effects on traffic patterns is not as obvious.

Because IP is connectionless, IP applications have extreme flexibility in location. Companies are finding that their departments are setting up Web servers and file servers away from their traditional head-office centers. In other words, because IP applications treat networks as being connectionless, the traffic in WAN links in corporate IP networks tends to become more meshed and any-to-any in nature.

This does not fit well with traditional hub-and-spoke virtual circuit connectivity. See Figure 4-1 for a comparison of topologies. Hub-and-spoke architectures lead to traffic passing through intermediate CPE in order to get to its destination. This is undesirable if you must pay for traffic to go across virtual circuits twice. It also wastes bandwidth at the hub sites and means that you must manage routing.

Figure 4-1 How Connectionless Traffic Drives Meshing



34391

In response, many customers are adding more meshing to their virtual circuit networks. The virtual circuit connectivity increasingly reflects the any-to-any connectionless nature of IP traffic.

However, carrying connectionless traffic on increasingly complex meshes of virtual circuits is a quite inadequate solution, for several reasons:

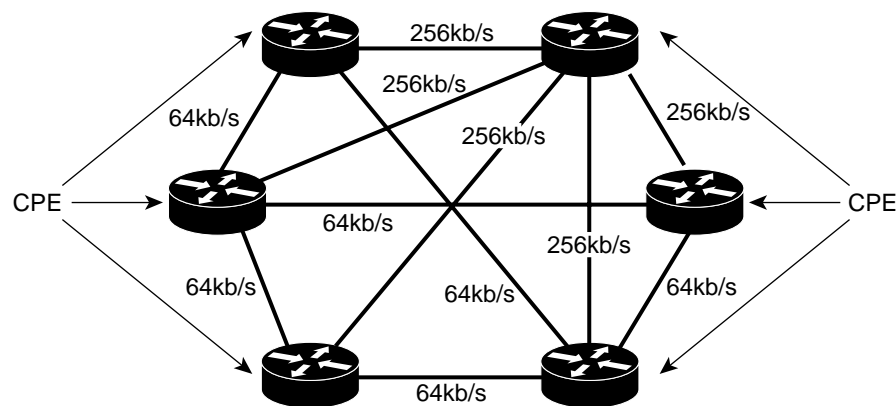
- The complexity of CPE management increases with increasing routing complexity.
- Even if the provider network is able to support all the extra virtual circuits required for meshes, the operational overheads of managing VC meshes become extreme.
- It becomes difficult to determine reasonable QoS parameters for all the PVCs.

The underlying problem is that IP traffic does not naturally fit with a connection-oriented service from a provider. For maximum efficiency and lowest cost, providers must offer a connectionless service to customers. Market research [CIMI Corporation, 1998] suggests that over 50 percent of current demand for IP Virtual Private Network services is unmet. The lack of true connectionless IP services offered by carriers is an important reason why this demand is unmet. Connectionless MPLS Virtual Private Network services simplify management for connectionless IP services.

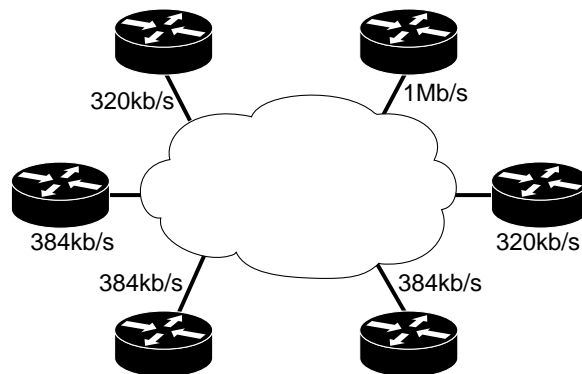
## Specifying QoS for Connectionless Service

In traditional virtual circuit networks such as Frame Relay networks, Committed Information Rates (CIR) must be specified for every link, as shown in Figure 4-2 Topology (a). As networks become more meshed, this becomes more difficult. For example, a full-mesh network of 100 sites would require 9900 separate CIRs to be provisioned. It is obviously unreasonable to dimension any-to-any networks in this way. (However you might want to specify only a few specific origin-destination bandwidths in an otherwise connectionless network.)

Figure 4-2 Specifying Bandwidths for An IP Service



- (a) The traditional Frame Relay model of QoS specification: point-to-point Committed Rates



- (b) In connectionless IP networks, specification of committed access bandwidths is more meaningful

34392

Aside from being connectionless, there are other reasons why providing QoS for IP traffic is fundamentally different from providing QoS in connection-oriented networks. Connection-oriented QoS is based on the premise that most traffic has QoS requirements that must almost always be met in order to provide adequate performance. Most IP applications, on the other hand, are tolerant of widely varying bandwidth; they can tolerate periods of seconds or more of high loss and are usually extremely tolerant of delay and delay variance.

Because of this fundamental difference, traditional connection-oriented QoS tools of connection admission control and per-VC QoS guarantees are an unnecessary overhead for most IP traffic. They are also difficult or impossible to use without a fully specified matrix of traffic requirements; this is an unnatural requirement for an IP service.

Internet services already use a quite different model of QoS specification. As shown in Figure 4-2 Topology (b), Internet users subscribe to a service by specifying access bandwidths for each of their sites. They do not specify a full matrix of bandwidths or any connection-oriented information. This is the natural way of structuring QoS demands for any connectionless IP service because the access bandwidth requirements are easily estimated in proportion to the number of hosts or servers at each site.

Providers who offer IP Service Level Agreements without using traditional connection-oriented QoS methods have an important advantage. They do not have to deal with the equipment costs and operational overheads for connection-oriented QoS, and closely meet requirements for Quality of Service agreements suitable for connectionless networks. Providing connectionless Service Level Agreements is an important requirement for meeting current and future demand for IP services.

## The Differential Services Approach to Quality of Service

### Contracts for Access Bandwidths

Even though a full matrix of point-to-point bandwidths is not normally specified for MPLS networks, usage parameter control or policing can still be applied to constrain use of network resources. This is important because it provides a basis for service providers' traffic planning to meet Service Level Agreements.

In Figure 4-2 Topology (b), a customer contracts for a certain access bandwidth at each site. In a simple case, they could be restricted to this bandwidth by setting the access line's data link. However, more complex access contracts are possible and desirable.

An IP-layer policing function called Committed Access Rate (CAR) is available in Cisco routers. CAR acts independently on each customer access link, or on each virtual circuit on a channelized access link.

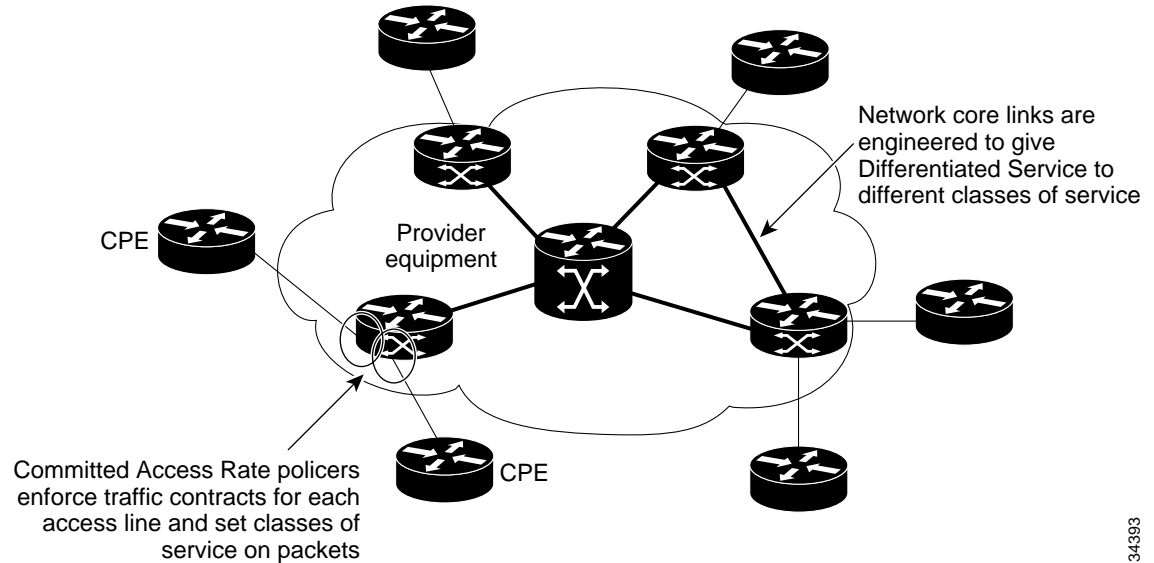
This use of CAR is shown in Figure 4-3. When used on Edge LSRs or other provider access routers, CAR both enforces traffic contracts and marks packets according to the traffic contract. For example, a simple CAR contract may specify that a user site gets:

- 64 Kbps of traffic carried as premium traffic
- The remaining traffic, up to a total site bandwidth of 256 Kbps of traffic, is carried as best effort.

This explanation is simplified for clarity. Typical IP Precedence classes for premium and best-effort classes mentioned here would be 4 and 0, respectively. Typical Differentiated Services (DiffServ) classes for traffic would be AF12 for premium and AF11 for best effort.



Figure 4-3 Cisco Committed Access Rate Policers



34393

Far more sophisticated contracts are possible. Another possible example contract for a site:

- The first 56 Kbps is carried in class Premium 2
- The next 128 Kbps is carried class Premium 1
- The remaining traffic up to a maximum of 1 Mbps is carried as best effort
- Any SMTP and FTP traffic is carried as best effort and not premium

The last point illustrates an important advantage of policing at the IP layer. CAR is able to take into account IP header information. In this example, it is used to specify that certain types of IP traffic that are very tolerant of varying QoS are automatically carried in the best-effort class and not counted against the limits for premium traffic.

CAR enforces the bandwidth contracts by using token bucket policers, which permit burstiness in a short time-scale, while limiting rates in a longer time-frame. Traffic classes are marked on the IP packets admitted into the provider network.

CAR sets IPv4 Precedence bits on packets. (The meaning of these bits will be changed in the forthcoming Differentiated Services standards from the IETF, but DiffServ is backward-compatible with the original IPv4 formats. CAR will be fully compliant with the new meanings of the DiffServ DS bits on IP packets.)



**Note** Two different acronyms are used for Differentiated Services and both are commonly used in other documents. “DiffServ” is used most commonly, and refers to Differentiated Services in general. “DS” is the name given specifically to the bits in the IP headers used by DiffServ.

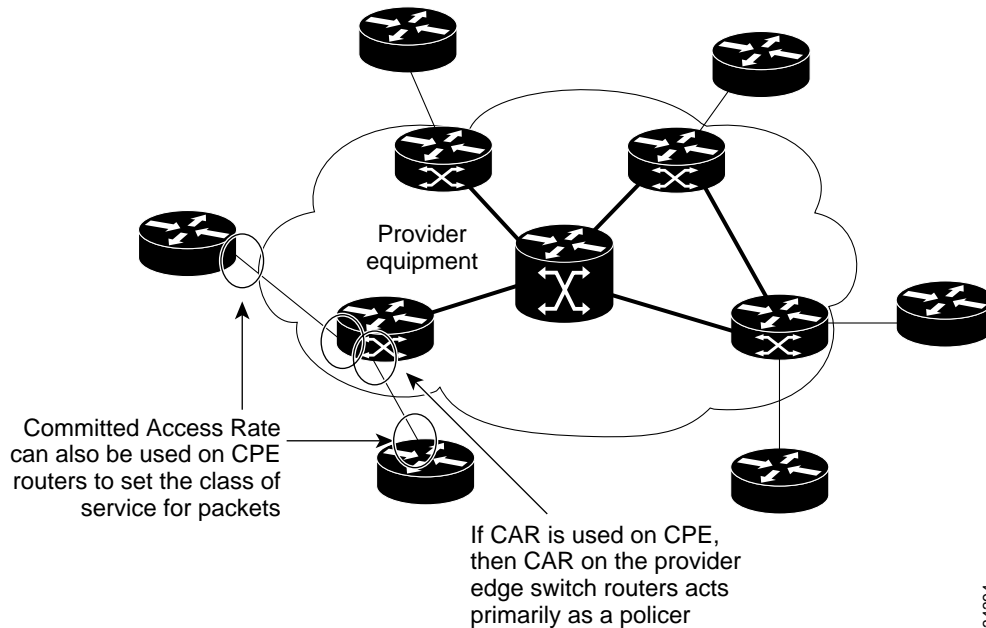
CAR is compliant with the DiffServ architecture, which requires technologies like CAR to mark precedence on IP packets.

The core of the network supports different Differentiated Services classes. In MPLS networks, after CAR has been used to mark Class of Service (CoS) using the Precedence or DS bits on IP packets, the IP packets are sent on different LVCs according to their CoS. There is, in general, a different LVC for each Class of Service.

In other words, CAR sets IP Classes of Service, which are then supported by MPLS.

A service provider can use CAR to both police IP traffic and mark CoS on IP packets using Precedence or DS bits. An alternative is to choose the CoS for packets according to your organization's own policies. CAR can be used on Customer Premised Equipment (CPE) to do this, and will allow customers to coordinate CoS assignments using Directory-Enabled Networking. If CAR is used on CPE to preset CoS, then CAR on the Edge LSRs acts purely as a policer. The use of CAR on both CPE and Edge LSRs is shown in Figure 4-4.

Figure 4-4 Using CAR on Customer Premises



## Using Best-Effort Traffic to Help Guarantee Bandwidths

The presence of much best-effort traffic in a network produces important advantages:

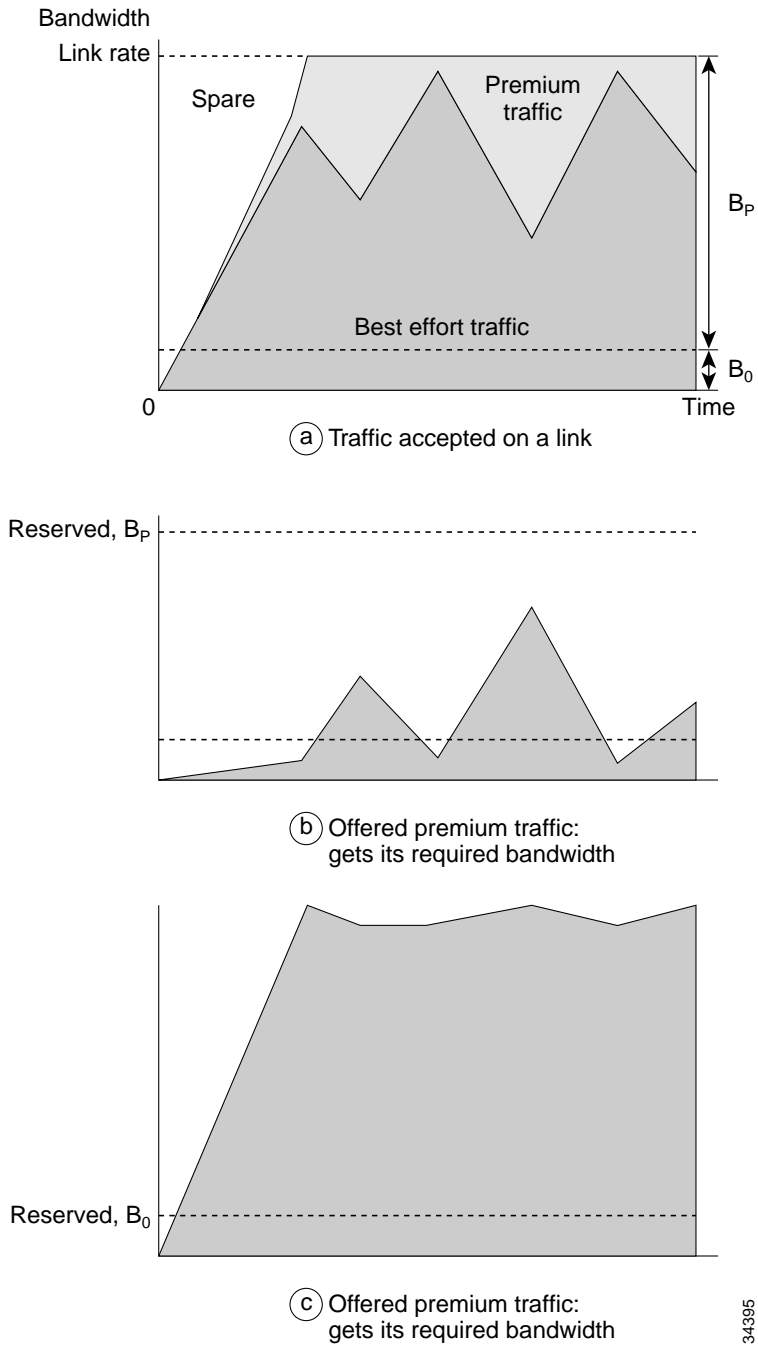
- It is likely that a relatively small amount of traffic in an IP network will require guarantees of QoS and that the rest of the traffic will require best effort.
- If the majority of traffic is best effort, good QoS can be ensured simply by giving precedence to the higher classes of traffic. The traffic requiring good QoS can be given the ability to push best-effort traffic out of the way, in order to provide good QoS for the premium traffic.

This explains why best-effort traffic is an advantage in providing quality of service: it can be used to cushion premium traffic, ensuring that the premium traffic gets premium service. This is the key to the Differentiated Services approach to service quality.

Figure 4-5 shows how Differentiated Services can work to ensure that premium traffic gets good service. The network operator allocates bandwidth to two Classes of Service on a particular link:

- Best-effort traffic is allocated  $B_0$
- Premium traffic is allocated  $B_p$

Figure 4-5 Ensuring Access to Bandwidth Using Differentiated Services



Note that  $B_p$  is greater than the estimated bandwidth of premium traffic. This means that the premium traffic gains access to its required bandwidth even if the actual requirement is much greater than the estimated requirement. The best-effort traffic is guaranteed little—after all, it is best-effort traffic. This means that the best-effort traffic may be denied bandwidth to meet the requirements of premium traffic. This occurs most of the time; the best-effort traffic does not get all the bandwidth it could use. On the other hand, all the premium traffic is carried.

Using this method, Differentiated Services can give excellent quality of service to premium traffic provided that:

- A substantial amount of traffic is best-effort traffic.
- A rough estimate can be made of the premium traffic load on the link, so that a larger value  $B_p$  can be allocated in proportion.  $B_0$  is calculated as being the link rate minus  $B_p$ .

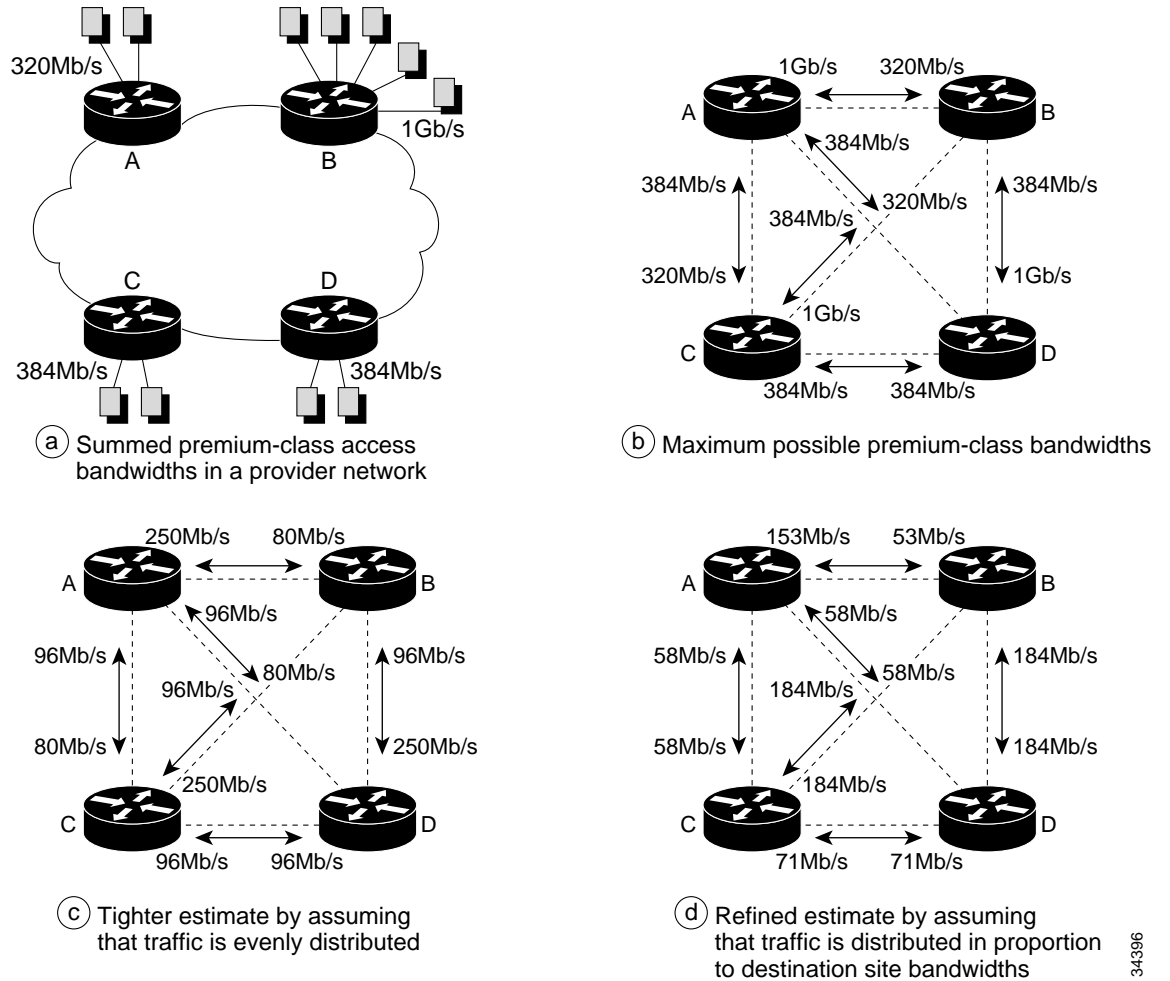
The first condition is likely to hold in future networks because almost all IP traffic today is best-effort and because IP traffic loads are growing much faster than other traffic, such as circuit-switched voice. Unusual cases where this doesn't hold are discussed in "What If There Isn't Much Best-Effort Traffic in My Network?" section on page 4-21. The second condition is quite easy to meet, and this is discussed next.

This example used only one class of premium traffic as well as best-effort traffic. However, multiple classes of premium traffic can be supported in the same way, provided that the two conditions are met.

## Modeling Network Traffic Flows to Meet Service Level Agreements

Now consider the engineering steps required for DiffServ network to provide Service Level Agreements (SLA). We've seen how CAR or similar technologies can be used to enforce access rate contracts for sites. This means that at each Edge LSR, it is easy to calculate the sum of allowed premium-class access bandwidths. An example is shown in Figure 4-6.

Figure 4-6 Refining Estimates of Network Loads



In order to engineer the core network to ensure delivery of premium-class packets, an estimate of the actual distribution of premium traffic is required, even though customers do not specify traffic matrices.

Note that the traffic estimates are required only for aggregate flows between Edge LSRs. For example, consider a provider network with 100 Edge LSRs serving 200,000 customer sites. For this network, traffic estimates are required for the flows between the 100 Edge LSRs, and not for the individual flows between the 200,000 customer sites. Also note that the estimates do not have to be exact.

It is trivial to calculate the maximum possible traffic flows for each origin-destination pair. In Figure 4-6 Topology (a), for example, the sum of the premium-class access bandwidths at Edge LSR A is 320 Mbps. Thus, the maximum possible flow of premium class traffic from A to any other Edge LSR is 320 Mbps, all illustrated in Figure 4-6 (b).

34396

This “maximum possible” is an extreme over-estimate: it is impossible for a given source LSR to send its full rate of premium traffic to each and every other Edge LSR simultaneously, but this is what is assumed by this traffic estimate. Despite this, the “maximum possible” traffic matrix may be useful in some circumstances:

- During trials and early production phases of introducing an MPLS network, it may be appropriate to dimension the network according to the “maximum possible” traffic to reduce the risks associated with more efficient traffic models. The early phases would be used to gather statistics about actual traffic distributions.
- In networks with very small amounts of premium traffic compared to best-effort traffic, the inefficiencies of the “maximum possible” traffic model may be irrelevant, because the best effort traffic will be able to use all the bandwidth over-allocated to premium traffic. In this case, there is no need to dimension the network according to a more realistic traffic model.

Beyond these exceptions, it will usually be better to use a more realistic traffic model. A more realistic traffic model might assume that traffic is distributed evenly among all possible destinations.

In the example in Figure 4-6, there are four Edge LSRs, and the summed premium-class access bandwidth at Edge LSR A is 320 Mbps. In this evenly distributed estimate, (320/4) Mbps is sent from A to all the Edge LSRs. (Edge LSR A can send traffic to itself. This is realistic because it represents corporate VPNs that have more than one VPN site attached to Edge LSR A. In general, there will be thousands of sites in hundreds of VPNs attached to A.)

Similarly, Edge LSR B sends (1000/4) Mbps to every Edge LSR, and so on. This is shown in Figure 4-6 Topology (c).

This method can be refined further by assuming that traffic is sent in proportion to the receiving sites’ access bandwidths. This means that a site with many large access lines is assumed to receive more traffic than a site with fewer, smaller access lines.

In Figure 4-6 (d), the sum of the access bandwidths at nodes A, B, C, and D are in the proportion 320:1000:384:384. Consider the 1000 Mbps of premium-class traffic entering site B. According to these proportions, the 1000 Mbps traffic from node B is divided among nodes A, B, C, and D as 153, 479, 184, and 184 Mbps. The traffic from the other nodes is divided similarly.

Dividing traffic by this method has the property that it leads to balanced estimates of loads, that is, the traffic in one direction between a given pair of nodes is equal to the traffic in the other direction. This “proportional to PoP size” estimated traffic is quite realistic and likely to be useful in many networks.

The setting of link or circuit parameters according to these traffic estimates is covered in following sections.

## A Recommended Process for Estimating and Modeling Traffic

The traffic model used to dimension the network does not have to be exact. Recall from the earlier discussion that premium traffic is first estimated, then actual allocations of bandwidth to premium traffic can be made in proportion to the estimates. But actual allocations should be larger than the estimates to allow a margin of safety.

Here is a recommended process for dealing with traffic estimates and modeling:

- 
- Step 1** In trials and early production phases, use the “maximum possible traffic” model to dimension premium-class bandwidth allocations. This ensures that all premium-class traffic will be delivered under all conditions except equipment or link failures.

- Step 2** During the early production phases, measure the actual origin-destination bandwidths. Cisco Netflow is an excellent tool for collecting these statistics. Alternatively, packet or cell counts can be collected for all the backbone links. Compare this traffic to an estimate derived using the “proportional to PoP bandwidth” estimate. It will usually be found that the “proportional to PoP bandwidth” estimate is a good approximation to the real traffic. In unusual cases, some other model may need to be developed. In any case, a traffic model should be selected for the network.
- Step 3** Change the dimensioning of premium-class bandwidths to agree with the chosen model, but initially over-allocate premium-class bandwidths by 100 percent or more to allow for unexpected variations. Recall that this over-allocation does not usually result in wasted bandwidth because best-effort traffic can typically make use of spare bandwidth.
- Step 4** Continue to collect statistics and compare them to the traffic model, modifying the model as appropriate. As confidence in the traffic model grows, reduce the amount of over-allocation of premium-class bandwidths.

This process is familiar to many network engineers. Aside from the particular traffic models used, it is identical to the process used in the introduction of over-subscribed Frame Relay and ATM networks.

---

## Engineering DiffServ Per-Hop Behaviors

DiffServ networks use queueing technologies such as Weighted Fair Queueing (WFQ) to provide differential service to the different Classes of Service (CoS). Link-by-link engineering of WFQ parameters is the approach suggested by the IETF DiffServ Working Group.

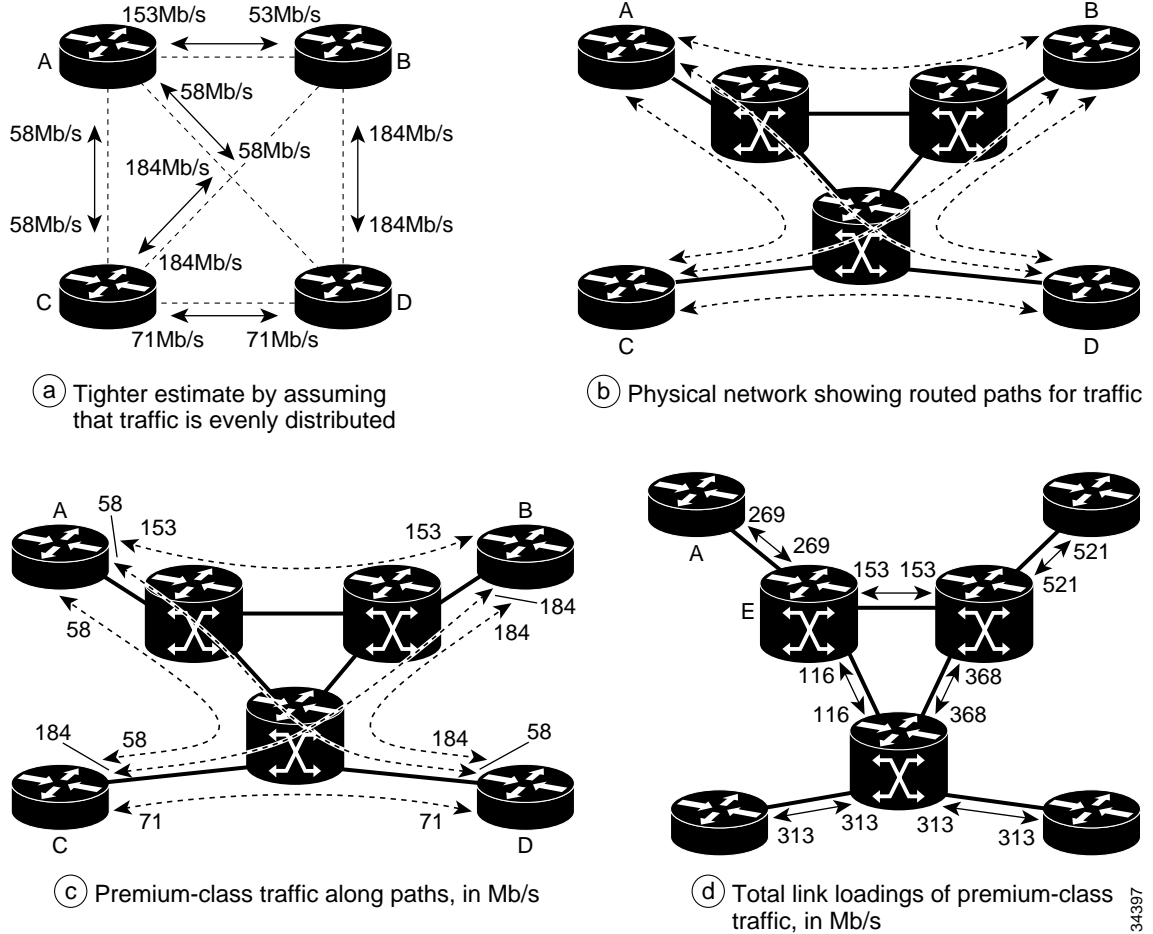
The treatment of a particular CoS on a particular link (or “hop”), using technologies such as Weighted Fair Queueing, is referred to as a per-hop behavior (PHB). Cisco supports engineering of per-hop behaviors on links in both ATM MPLS and packet-based MPLS networks, as well as ordinary IP networks. The principles are the same in all network types, although there are differences in the way CoS information is carried in packets for different networks.

As a prerequisite for setting the parameters for PHBs, the estimated demand for traffic of different PHBs must be derived for each hop from the estimated traffic matrices.

For example, Figure 4-7, topology (a) shows the traffic demand for a network where one class of premium traffic is carried in addition to best-effort traffic. Topology (b) shows the physical network, including the links and core LSRs which will carry the traffic. Topology (b) also shows the routes that will be chosen for IP routing in the network. IP routing protocols such as OSPF and IS-IS normally chose the shortest possible route from a given origin to a given destination. For complex networks, a tool such as Cisco Netsys will be helpful to review and analyze the routes used in a network.

The traffic matrix and routing information together specify the bandwidth used along various paths as shown in Topology (c). From this information, it is straightforward to sum the total bandwidths on each link.

Figure 4-7 Estimating Network Loads Per-Hop Behavior



For example, there are three components of the premium-class traffic flow in the link from A to E: 153 Mbps A->B, 58 Mbps A->C, and 58 Mbps A->D. The sum of these is 269 Mbps, and this is the premium-class bandwidth requirement on the link A->E shown in Topology (d). The other premium-class bandwidths are calculated similarly.

The end result of this calculation are values that should be allocated for premium-class bandwidth on each link. In this example, there has been one class of premium traffic and the premium-class bandwidth reservation on link A->E should be at least 269 Mbps. Referring back to Figure 4-5, this means that the Weighted Fair Queueing bandwidth  $B_p$  assigned on the link A->E should be at least 269 Mbps.

This example uses one class of premium traffic. The same process can be used for multiple classes of premium traffic with the bandwidth requirements for each class calculated as described here.

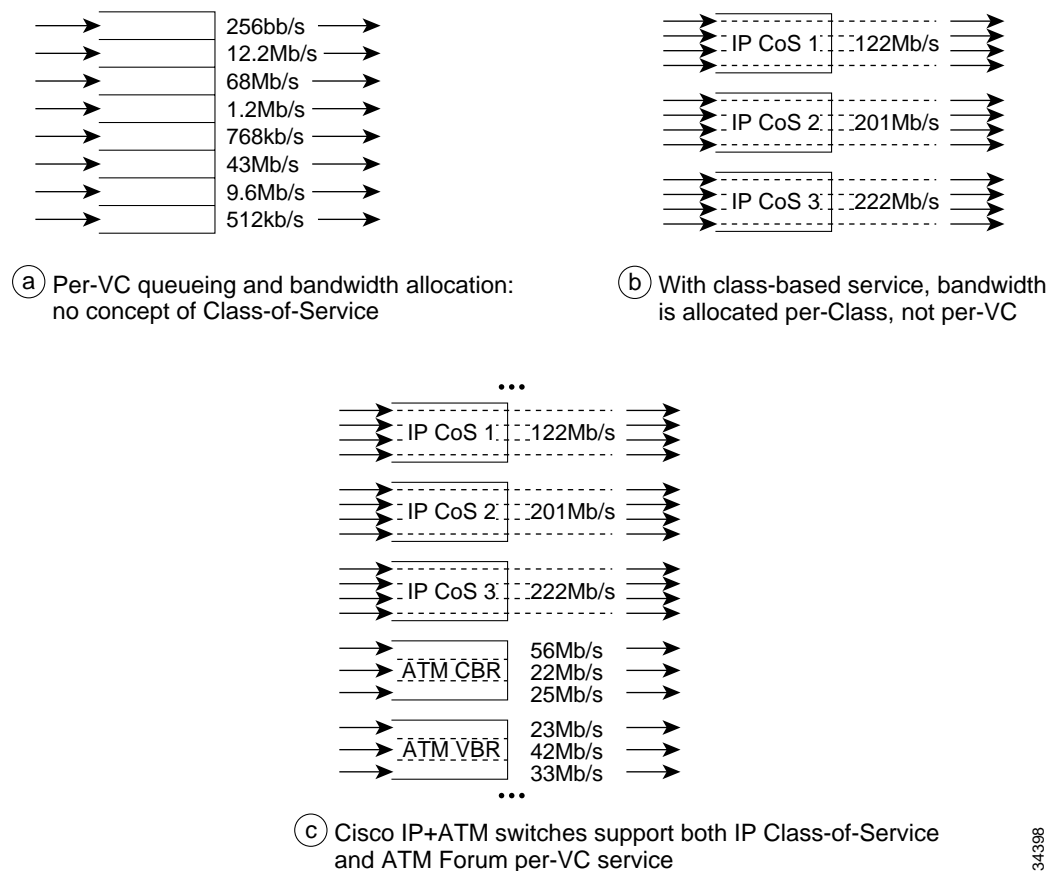


## DiffServ Classes and Cisco IP+ATM Switches

The preceding discussion has shown how engineering of DiffServ networks leads to specifications of required bandwidths for various Classes of Service on various links of the network. This is quite different from traditional per-VC bandwidth management in ATM networks.

The per-VC bandwidth concept used for CBR, VBR, and other ATM Forum traffic management types is illustrated in Figure 4-8 (a). Per-VC bandwidth management requires that bandwidths be specified for every origin-destination flow. If per-VC bandwidth management is used in conjunction with approximate estimates of origin-destination flows, bandwidth will be distributed unfairly. This means that per-VC bandwidth management is less useful in connectionless networks than the DiffServ mechanisms.

**Figure 4-8 Per-VC Service and Class of Service in ATM Switches**



344398

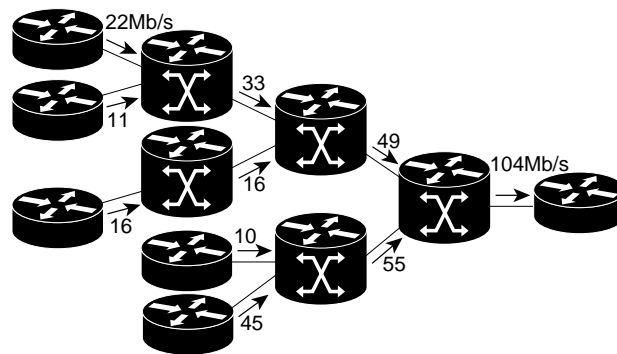
Per-VC bandwidth management is quite different from DiffServ concepts of CoS and Per-Hop Behavior. There is no simple, straightforward way to support DiffServ using ATM Forum traffic management types. Because of this, Cisco does not attempt to use ATM Forum traffic management types to support DiffServ. Instead, class-based queueing is used.

As shown in Figure 4-8 (b), class-based queueing involves a separate queue in the ATM switch for each CoS. Cells from all LVCs of each CoS are queued in a single queue for that CoS. The bandwidth parameters of a CoS on a link are set directly on the CoS queue. The only parameter signalled for each

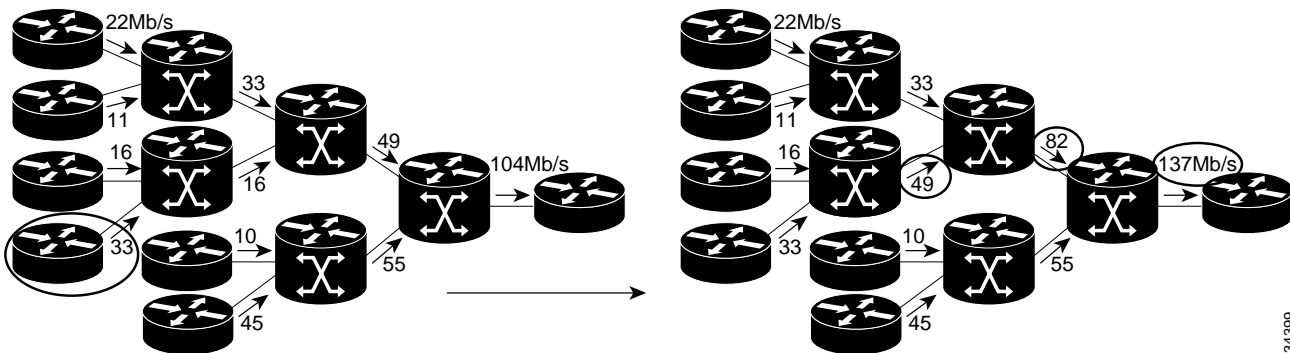
LVC is the Class of Service for the LVC. This means that the ATM MPLS control component is used unchanged, except that multiple LVCs are set up for each destination: one LVC per destination per Class of Service.

Cisco IP+ATM switches support DiffServ for MPLS traffic, alongside ATM Forum Traffic Management types for PVCs and SVCs. Each DiffServ or ATM Forum Traffic Management type gets its own “Class of Service Buffer.” Per-VC queuing can be used in addition to the class of Class of Service buffers and this is done for ATM Forum Traffic Management types. Weighted Fair Queuing is used to assign bandwidths to the IP Class of Service buffers. This means that the IP classes share bandwidth as illustrated in Figure 4-5.

Figure 4-9 Per-VC Service with VC Merge



(a) Hypothetical network using per-VC bandwidth allocation and VC merge



(b) Any change or addition in edge bandwidth leads to ripple of changes through network: per-VC bandwidth allocation negates the signaling scalability advantages of VC merge

Using class-based queuing instead of per-VC queuing for the IP traffic has several advantages:

- The number of parameters programmed into the network is much smaller with class-based queuing: if a network has  $N$  nodes, the number of parameters required is proportional to  $N^2$  with per-VC queuing, but proportional to  $N$  with class-based queuing.

- Class-based queueing is fairer, given approximate information. This is important because engineering of an IP network is based on estimates and models of customer traffic. With class-based queueing, premium-class traffic from any origin to any destination gets preferential access to any premium-class bandwidth left spare from other origin-destination pairs. This is much harder to achieve if bandwidths are assigned to individual origin-destination LVCs.
- Class-based queueing can be used on any link types. Link types include those that do not support virtual circuits: PPP-over-SDH and WDM. Use of class-based queueing helps make a network flexible and open to future changes in technology without major changes in operations, administration, and management. Cisco already makes switch-routers with ATM, PPP-over-SDH, and WDM interfaces.
- Class-based queueing works better with VC merge than per-VC queueing. In fact, per-VC queueing negates the advantages of VC merge in improving signaling scale. Refer to Figure 4-9. If per-LVC queueing were used, each LVC in the tree of LVCs merging to a given destination would need a bandwidth assigned to it according to the sum of bandwidth requirements merging in from other branches. If any addition or change were made to the bandwidths of the merging VCs, this would create a ripple of signaling through the network. This negates one of the important advantages of VC merge, namely that VC merge removes the requirement for end-to-end signaling for most LVCs.
- Even if class-based queueing is used, changes to class-based bandwidths will be required as bandwidth requirements change. However these can be dealt with as a network provisioning issue on a time-frame of at least hours or days. Class-based queueing does not require the real-time QoS signaling overheads of per-VC queueing. Furthermore, the granularity of changes with class-based queueing is per-link; with per-VC queueing, the granularity is per-VC. This is another example of how class-based queueing is more scalable.

For these reasons, Cisco strongly recommends that networks supporting IP services are engineered using class-based queueing.

## Service-Level Agreements Using DiffServ

This section has covered:

- Enforcement of access contracts at the edge of a network using Cisco CAR
- Using the access contracts as a basis for modeling traffic
- Refinement of traffic models based on operation of a network
- Setting of the links' queueing parameters according to the traffic models

These steps can lead to guarantees of packet delivery:

- Stronger guarantees of packet delivery can be given if "maximum possible" traffic models described above are used
- Otherwise, Service-Level Agreements are statistical, and are based on the accuracy of the traffic modeling process

Unless "maximum possible" traffic models are used, the meeting of SLAs is reliant on monitoring network performance over time, reacting to unexpected traffic events, and modifying traffic models.

An important set of statistics to monitor is the amount of premium-class traffic per link. If premium-class traffic is nearing its allocated bandwidth, then there is a danger that SLAs might not be met. However, because this process is reactive, it is important that SLAs are structured to permit the

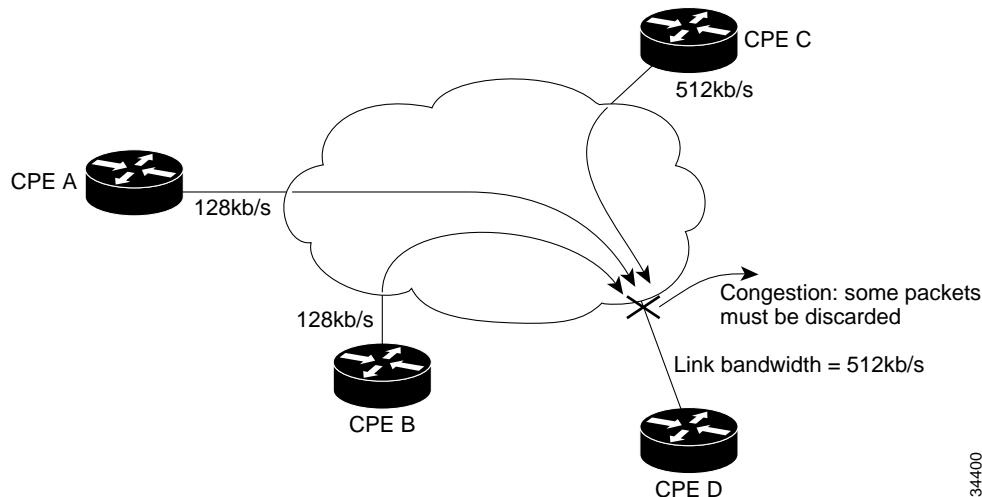
provider to have time to react; in other words they should allow for periods of poor QoS. It is very important to bear in mind that most IP traffic is very tolerant of varying QoS; customers who understand this will understand that a quite weak SLA is satisfactory even for premium traffic.

Another important aspect of SLAs for IP traffic is the nature of commitment. In a point-to-point network, it is quite easy to define a Committed Information Rate between two sites. In an IP network, the nature of commitment is different.

Consider Figure 4-10. Even if sites A, B, and C are transmitting packets within their premium-rate access contracts, not all of their packets will be delivered. They are sending packets to D at total of 768 Kbps, which exceeds the link bandwidth to D: 512 Kbps. The packet loss rate in this example will be roughly 33 percent even though the access contracts have been met.

It is important for IP SLAs to emphasize this, and specify that a packet is not committed for delivery unless it both meets the access contracts, and is sent within the possible receive rate at the destination. This means that if a customer chooses to send 768 Kbps of traffic to a site with a link bandwidth of 512 Kbps, then the resulting loss is the customer's responsibility. (The customer may deal with this by buying more bandwidth on the link to site D.)

**Figure 4-10 Committed Delivery in An IP Network**



A simple way of structuring SLAs for IP traffic is to use two classes of traffic. This results in a traffic contract that is similar to a Frame Relay committed information rate.

## Sample Service Level Agreement Using the Two-Class Model

The examples in this section show a type of Service Level Agreement that a service provider might offer to its customers for IP traffic. Because the meeting of such a Service Level Agreement depends in part on use of appropriate planning and monitoring processes by a service provider, the service levels described are illustrative only and Cisco offers no guarantee that any particular network will be able to meet the service levels described.

1. The first 64 Kbps of traffic sent from a customer site each second is committed. (This definition is slightly simplified for clarity. A real contract might specify that traffic is measured by a token bucket policer, and specify the token bucket parameters.) Any traffic that otherwise satisfies the definition of committed traffic, but is sent so that the sum of the committed bandwidths sent to the

34400

receiver is greater than the receiver's link rate, is counted against the 64 Kbps of committed traffic, but is treated as best-effort traffic for the purpose of clauses 3 to 10. Any packet fully in compliance with this clause is referred to as a committed packet.

2. Excess traffic, up to a total site bandwidth of 256 Kbps of traffic, will be accepted by the network with no guarantee of delivery. This is referred to as "best-effort" traffic or best-effort packets. (The contract could specify that any traffic in excess of 256 Kbps would be discarded, but this limit would typically be automatically enforced using a link rate of 256 Kbps.)
3. Within each month, at least 99 percent of committed packets from this site will be delivered.
4. For each month: during the 1-hour period in which the lowest proportion of committed packets is delivered by the network, at least 90 percent of committed packets from this site will be delivered.
5. 99.9 percent of committed packets delivered will be within 250 milliseconds of being accepted.
6. There is no guarantee that any best-effort traffic will be delivered.
7. 99 percent of best-effort packets delivered will be within 1 second of being accepted.
8. Of all the packets delivered, not more than 0.1 percent will be delivered out of order in which they are received.
9. Of all the packets delivered, not more than 1 in  $10^6$  will have an error introduced by the network.
10. (Further clauses will specify costs, penalties if the clauses above are not met, and so on.)

The first two clauses define committed and best-effort packets, and the allowable rates for both. Clause 1 excludes traffic that cannot possibly be delivered as discussed previously.

Clauses 3 and 4 provide realistic assurance of delivery of packets. Clauses 3 and 4 together allow for relatively poor performance for periods of hours each month, while still assuring adequate delivery performance during bad periods—such a period could indicate some need for improvement in the providers's traffic modeling. Note that a provider must provision a network according to an accurate or suitably over-estimated traffic model in order to have confidence in meeting an SLA such as this. Alternatively, it is safer to provision a network using a "maximum possible traffic" model, as discussed earlier. This would typically be done early in network deployment.

Clauses 5 and 7 specify loose delay bounds, which are straightforward to provide. This will be discussed in more detail in "Delay Limits" section on page 4-33.

Clause 8 is an important provision for IP traffic, namely that packets are delivered in order. TCP, UDP and other transport protocols can deal with IP packet misordering, but large amounts of misordering can lead to poor transport-layer performance. Fortunately, the queueing technologies used on Cisco equipment ensures packet ordering within each Class of Service, except during rerouting.

It is also desirable that premium and best-effort traffic is carried in the same queues, in order to ensure packet ordering across both classes. This is quite possible on Cisco equipment, and preference is given to the premium traffic by way of discard policies. This is discussed in more detail in "Discard Policies" section on page 4-28.

Because rerouting is a rare and short-lived event caused only by link or equipment outages, or manual routing changes, it is easy to give a strong assurance of packet ordering. Modern telecommunications networks, including those using Cisco equipment, rarely introduce bit errors. An error provision such as Clause 9 is easy to meet.

Readers familiar with QoS guarantees and SLAs for ATM and Frame Relay traffic may be surprised by the weakness of the guarantees suggested in the example. The numbers suggested are entirely reasonable for IP traffic, because most IP traffic is extremely tolerant of varying QoS. The SLA is also incomparably better than the SLAs offered on most public IP networks today: most public IP networks offer no SLAs at all, and the Internet has been widely successful despite this. Providers who offer SLAs of the strength suggested above will capture a large untapped market: customers requiring moderate

guarantees for a truly connectionless IP network, but who do not wish to specify all point-to-point bandwidth requirements. Traditional point-to-point QoS guarantees are over-engineered for the needs of most IP traffic.

Some IP traffic requires stronger SLAs. Real-time IP traffic is considered in the next example. Stronger, Frame Relay-quality SLAs may be required for a small minority of IP traffic, and these are considered in “More Stringent Quality of Service in IP+ATM Networks” section on page 4-25.

## Sample Service Level Agreement with Provision for Real-Time Traffic

This section considers the types of Service Level Agreement a service provider might offer to its customers for IP traffic. Because the meeting of such a Service Level Agreement depends in part on use of appropriate planning and monitoring processes by a service provider, the service levels described in this paper are illustrative only and Cisco offers no guarantee that any particular network will be able to meet the service levels described.

1. The CPE may indicate that traffic is real time by setting the IP Precedence field on each packet to a value greater than 0. Any packets with precedence value greater than 0 will be treated as real time, and 64 Kbps of real-time traffic will be accepted. Any real-time traffic in excess of 64 Kbps will be discarded. Furthermore, any traffic that otherwise satisfies this clause, but is sent so that the sum of real-time bandwidths sent to the receiver is greater than the receiver’s link rate, is counted against the 64 Kbps of real-time traffic, but is treated as best-effort traffic for the purpose of clauses 4 to 9.
2. The first 256 Kbps of non-real-time traffic sent from a customer site each second is committed. Any traffic that otherwise satisfies this clause, but is sent so that the sum of the real-time and committed bandwidths sent to the receiver is greater than the receiver’s link rate, is counted against the 256 Kbps of committed traffic, but is treated as best-effort traffic for the purpose of clauses 4 to 9.
3. More traffic up to a total site bandwidth of 1024 Kbps will be accepted by the network with no guarantee of delivery. This is referred to as best-effort traffic.
4. Within each calendar month, at least 99.9 percent of real-time packets from this site will be delivered.
5. For each calendar month: during the 1-hour period, in which the lowest proportion of real-time packets is delivered by the network, at least 99 percent of real-time packets from this site will be delivered.
6. Within each calendar month, at least 99 percent of committed packets from this site will be delivered.
7. For each calendar month: during the 1-hour period in which the lowest proportion of committed packets is delivered by the network, at least 90 percent of committed packets from this site will be delivered.
8. 99.9 percent of real-time packets which are delivered will be delivered within 100 milliseconds of being accepted.
9. (Other clauses as per the previous example.)

If more than two traffic classes are supported, it is necessary to identify the preferred traffic class in some way. Clause 1 defines a means of doing this. Alternatively, the provider equipment could detect the CoS by using the contents of the IP packet headers. Cisco CAR allows for this.

Note that IP Precedence 5 is the default class of Voice Over IP packets sent from Cisco equipment, and that clause 1 supports any other equipment that marks real-time packets with a precedence greater than 0. Clause 1 strictly limits the real-time traffic that a site may send, and this helps in the engineering of the network to support the real-time traffic class.

In order to meet the relatively strong delivery agreements specified in Clauses 4 and 5, it might be necessary to use a “maximum possible traffic” model for the real-time traffic. The remaining clauses are similar to the previous example.

A provider must provision a network according to an accurate or suitably conservative traffic model in order to have confidence in meeting an SLA such as this. Alternatively, it is safer to provision both the real-time and committed traffic using a “maximum possible traffic” model, as discussed in “Modeling Network Traffic Flows to Meet Service Level Agreements” section on page 4-10.

## Adding a New Site

Because IP services are connectionless, normal processes of connection admission control are not useful. The equivalent service admission control is a network management operation.

Consider the process of adding of a new customer site or set of sites. A typical service admission process would follow this procedure:

- 
- Step 1** As an ongoing monitoring operation, maintain a matrix  $A$  of the available bandwidth for each CoS between each pair of Edge LSRs. This should consider engineering rules, for example that 25 percent of the bandwidth for each CoS on each link should be left unallocated as a statistical reserve. Also maintain a matrix  $R$  of bandwidth reserved for services that have been allowed (see Step 3) but not yet activated.
  - Step 2** Form a model of the traffic introduced by the new service. Traffic modeling is discussed in “Modeling Network Traffic Flows to Meet Service Level Agreements” section on page 4-10. This results in a matrix  $N$ .
  - Step 3** Compare the new traffic  $N$  to the available bandwidth ( $A-R$ ). If there is sufficient available bandwidth, that is  $(A-R-N) > 0$ , then allow the new service and increase the reserved bandwidth  $R$  by the new traffic  $N$ . Otherwise, refer the new service request to be dealt with by increasing the provisioning of the network. If increased provisioning isn’t possible, the service request must be rejected.
  - Step 4** After the service is activated, allow some time (weeks or more) for the customer to start using the service. Then decrease the reserved traffic  $R$  by  $N$ , and deal with any further changes in customers’ use of bandwidth by the normal engineering processes.

This process is based on straightforward statistics collection and calculations and could be added to an existing Operations Support System. In a similar manner to connection admission control for point-to-point services, Service Admission Control will ensure that new services will receive their desired quality of service.

---

## What If There Isn’t Much Best-Effort Traffic in My Network?

A prevalence of best-effort traffic assists with the engineering of DiffServ networks by allowing for very good QoS for premium traffic in the presence of approximate traffic estimates and without wasting bandwidth. Best-effort traffic helps cushion premium traffic. If best-effort traffic is not prevalent, it is still possible to use the engineering techniques described previously but there is an increased need for accuracy in traffic models.

These measures are helpful:

- Use initial, maximum-possible traffic models for a longer period while statistics on actual traffic patterns are gathered.
- Use MPLS Traffic Engineering.

## Standardization

CAR and the use of Diffserv discussed in “The Differential Services Approach to Quality of Service” section on page 4-6 is based on the forthcoming MPLS and Differential Services standards from the IETF. Cisco’s implementation of these technologies is either fully compliant with the standards (to the extent they are complete), or is compliant with the older IP Precedence definitions and will be upgraded to comply with DiffSev.

Relevant IETF documents are:

- “A Framework for Differentiated Services” draft-ietf-diffserv-framework-02.txt
- “An Architecture for Differentiated Services,” RFC 2457
- “Assured Forwarding PHB Group” RFC 2587
- “MPLS Support of Differentiated Services by ATM LSRs and Frame Relay LSRs,” draft-ietf-mpls-diff-ext-00.txt
- Request for Comment (RFC) documents are available from <http://www.ietf.org/rfc.html>
- Internet Drafts are available from <http://www.ietf.org/ID.html>

Some of these documents are Working Group Internet Drafts, works in progress at IETF Working Groups. Other Internet Drafts are referred to as Individual Internet Drafts. Individual Internet Drafts have no status in the standardization process, except as proposals from an individual or company. Individual Internet Drafts are easily recognized because “ietf” is not part of their name. Working Group Internet Drafts, on the other hand, are called “draft-ietf-” followed by the name of the Working Group, such as “mpls” or “diffserv,” followed by the title of the draft.

## The Differential Services Approach to Quality of Service: Summary

Good Quality of Service can be provided to connectionless IP traffic, on MPLS networks in particular. The process of doing this involves several steps:

- Enforcement of access contracts at the edge of a network using Cisco CAR
- Using the access contracts as a basis for modeling traffic
- Optional refinement of traffic models based on operation of a network
- Setting of the links’ queuing parameters according to the traffic models
- Offering SLAs of an appropriate form and strength for a connectionless IP service
- Service admission control



# MPLS Traffic Engineering

The forming and measuring of traffic models is an important part in the providing of good Quality of Service for connectionless traffic. Cisco is developing tools to assist this process. The first of these, MPLS Traffic Engineering, is currently available for Cisco router-based MPLS equipment and will be extended to ATM-LSRs. MPLS Traffic Engineering works by automatically measuring the actual traffic loads on the links of a network and then adjusts the routing of traffic to make best use of the available bandwidth.

There are several other uses for MPLS Traffic Engineering. It provides support for a full range of operational requirements in IP networks, all related to the choosing of routes for traffic:

- To fit IP traffic to available link bandwidths to make best use of the links.
- To prepare for operational events like taking a link out of service by shifting traffic off the link ahead of time.
- To deal with unexpected events like a sudden surge of traffic to a particular destination, which would involve spreading the traffic surge around several alternative paths.
- To select specific routes, such as protected SONET links, for certain traffic.

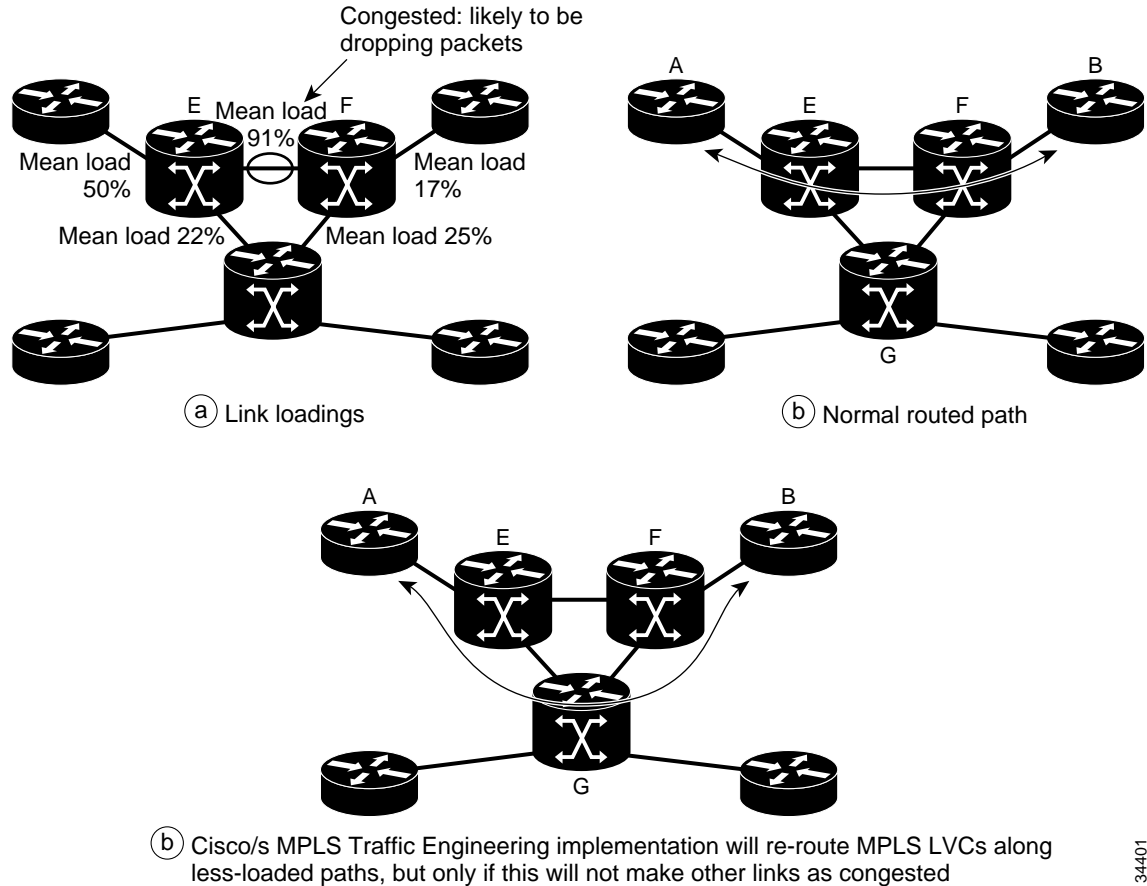
Cisco's implementation of MPLS Traffic Engineering works in this way:

- The network operator identifies candidate origin-destination traffic streams for adjustment by traffic engineering.
- At each Label Switch Router, the time-averaged load on each link is measured by Cisco MPLS Traffic Engineering software.
- The link loading information is distributed using the IP routing protocol used inside the network, specifically OSPF or IS-IS. Either OSPF or IS-IS is required, as these protocols are link-state routing protocols that base their routing decisions on a map of the state of the links in the network. Distance vector routing protocols such as RIP and EIGRP will not work for this particular application.
- Using optimization techniques, Cisco's MPLS Traffic Engineering software attempts to find alternative routes for the candidate streams of traffic so that the percentage loads are minimized. This re-optimization occurs at configurable intervals, and the default is one hour. Instability and oscillations can result if the re-optimization interval is set too small, but the network will not react to unexpected shifts in traffic if it is too great. One hour is a reasonable compromise. This means that traffic is routed so that it sees the lightest possible loads on the links it traverses.
- If a stream of traffic is to be sent on a different route to the normal OSPF or IS-IS route, the traffic engineering software sets up an MPLS Label-Switched Path along the alternate route. This path is known as a traffic engineering (TE) tunnel. Careful checks are made to ensure that the TE tunnel does not lead to loops. Specifically, a check is made to ensure that the exit of the tunnel is closer to the destination than the ingress. This integrity check is repeated periodically and a tunnel is disabled if it fails the check, or fails to pass traffic.

Optimizing traffic routing using MPLS Traffic Engineering is illustrated in Figure 4-11. In Topology (a) the mean load on the link between nodes E and F is 91 percent. At this load there is imminent danger that packet loss will occur, if it is not occurring already. MPLS traffic engineering will find candidate streams to reroute away from that link, if possible.

For example, the LSPs between Edge LSRs A and B may be carrying significant amounts of traffic. MPLS Traffic Engineering attempts to find an alternative route for one or more LSPs so that the load on the (E, F) link is reduced without increasing the loads on other links to a similarly dangerous level. So, depending on the bandwidth on the LSPs, it may be possible to solve the congestion by rerouting traffic away from (E, F) along a different path.

Figure 4-11 Reoptimization of Traffic Using MPLS Traffic Engineering



34401

Cisco's implementation of MPLS Traffic Engineering acts automatically to spread loads around a network's links as evenly as possible. It acts to minimize loads even if links are not currently overloaded. In this way, Cisco MPLS Traffic Engineering actively prevents overloads wherever possible.

MPLS Traffic engineering is based on measuring actual link loads. Existing QoS-aware routing protocols such as the ATM Forum's PNNI are less useful for this application because they are based on signaling of subscribed loads, rather than measurement of actual bandwidth loads. As previously noted, measurement of actual traffic is an important part of automating QoS in IP networks.

PNNI is not a good routing protocol for IP traffic for other reasons. Most notable is that there is no simple way to make PNNI make routing decisions in conjunction with routing information from standard IP routing protocols, namely OSPF and IS-IS. (There was a proposal called "IPNNI" to use PNNI as an IP routing protocol instead of OSPF or IS-IS. This failed because ISPs and carriers' Internet groups had no intention of replacing their existing OSPF or IS-IS infrastructures. Note that OSPF, in particular, is far more widely used than PNNI. IS-IS is also used by about half of the largest ISPs, and by several carriers.)

MPLS traffic engineering based on OSPF or IS-IS overcomes these limitations, and supports measurement-based engineering of connectionless traffic. In addition, it can support PNNI-style point-to-point bandwidth reservations where required, as discussed next.

Note that these issues are relevant only for routing for MPLS LSPs. Cisco IP+ATM switches use a full-featured PNNI implementation for traditional ATM connections: SVCs, SPVCs, and so on. MPLS and traditional ATM connections can be operated on the same links.

The candidate flows for adjustment using MPLS Traffic Engineering must be identified ahead of time. It will normally be found that a relatively small percentages of possible origin-destination traffic streams account for a large proportion of traffic.

For example, a network with 100 Edge LSRs has 9900 possible origin-destination streams, but it would typically be found that 90 percent or more of the traffic in the network might be on a few hundred of these. This means that MPLS Traffic Engineering can successfully optimize the traffic in a network by optimizing routing for a relatively small number of candidate origin-destination streams. The other streams will be left to follow the routes chosen normally by OSPF or IS-IS. Note that this does not imply lower quality of service for the non-candidate streams. These still benefit from the low link loads and even distribution of traffic provided by MPLS Traffic Engineering.

The signaling used in Cisco's implementation of MPLS Traffic Engineering, is compliant with a traffic engineering method approved by the MPLS Working Group. Extensions to carry link loading information in the IS-IS and OSPF routing protocols are works in progress at the IS-IS and OSPF Working Groups at the IETF.

## More Stringent Quality of Service in IP+ATM Networks

Previous discussions have described provisioning of SLAs for connectionless traffic in the absence of subscribed, point-to-point bandwidths. In some cases users will want harder PVC-like Quality of Service guarantees for traffic between certain sites. This may be required for critical business applications, disaster recovery, and so on. Cisco IP+ATM can meet these requirements in two ways.

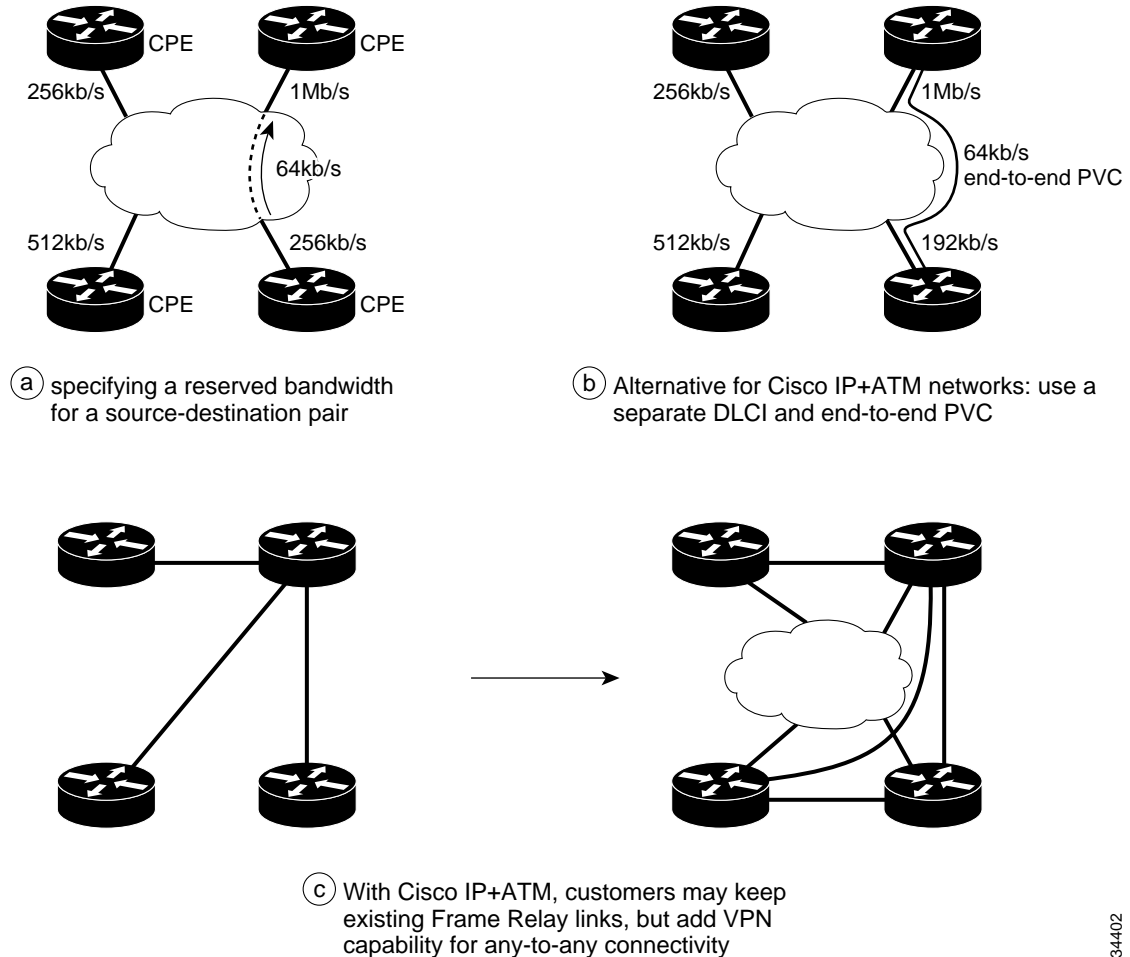
The full MPLS solution is shown in Figure 4-12, Topology (a). In this solution, MPLS Traffic Engineering routes a Label-Switched Path (LSP) with a specific, reserved bandwidth between two Edge LSRs. This LSP is reserved for traffic between two particular customer sites. (LSPs may be aggregated to help scalability—this is discussed in “Quality of Service for MPLS VPNs” below.) Per-VC queueing will be used for this LSP. Because MPLS Traffic Engineering routes this LSP based on a reserved bandwidth, rather than a measured bandwidth, these point-to-point LSPs will give QoS equivalent to switched permanent virtual circuits (SPVCs). This allows IP SLAs to be extended to include traffic reservations between specific sites.

Connectionless IP services require a service admission control (SAC) process rather than a traditional connection admission control (CAC) process. The point-to-point LSP services, on the other hand, use traditional CAC—the network will reject the connection if insufficient bandwidth is available.

There is an alternative to point-to-point MPLS links, which achieves the same results. With Cisco IP+ATM networks, a single customer site with a single access link can access both a traditional end-to-end PVC and a connectionless IP service as shown in Figure 4-12, Topology (b).

The IP service can be used to deliver connectionless SLAs and the PVC will deliver the extreme reliability already present for Frame Relay and ATM services on Cisco IP+ATM switches. The IP+ATM method is likely to be widely used for a long time because customers, such as banks, who already have a Frame Relay or ATM service for existing transaction processing might not want to shift away from Frame Relay or ATM for that traffic. If a customer's established business processes are running on long-used technology, it is quite reasonable for the customer to want to keep that existing infrastructure. Cisco IP+ATM solutions allow existing Frame Relay and ATM connectivity to be retained, while IP services are introduced for any-to-any connectivity for the new IP traffic, as shown in Figure 4-12 Topology (c). The switches in the network carry both traditional Frame Relay and ATM services, as well as IP services.

Figure 4-12 Reserved Point-to-Point Bandwidths in MPLS Networks



34402

## Quality of Service for MPLS VPNs

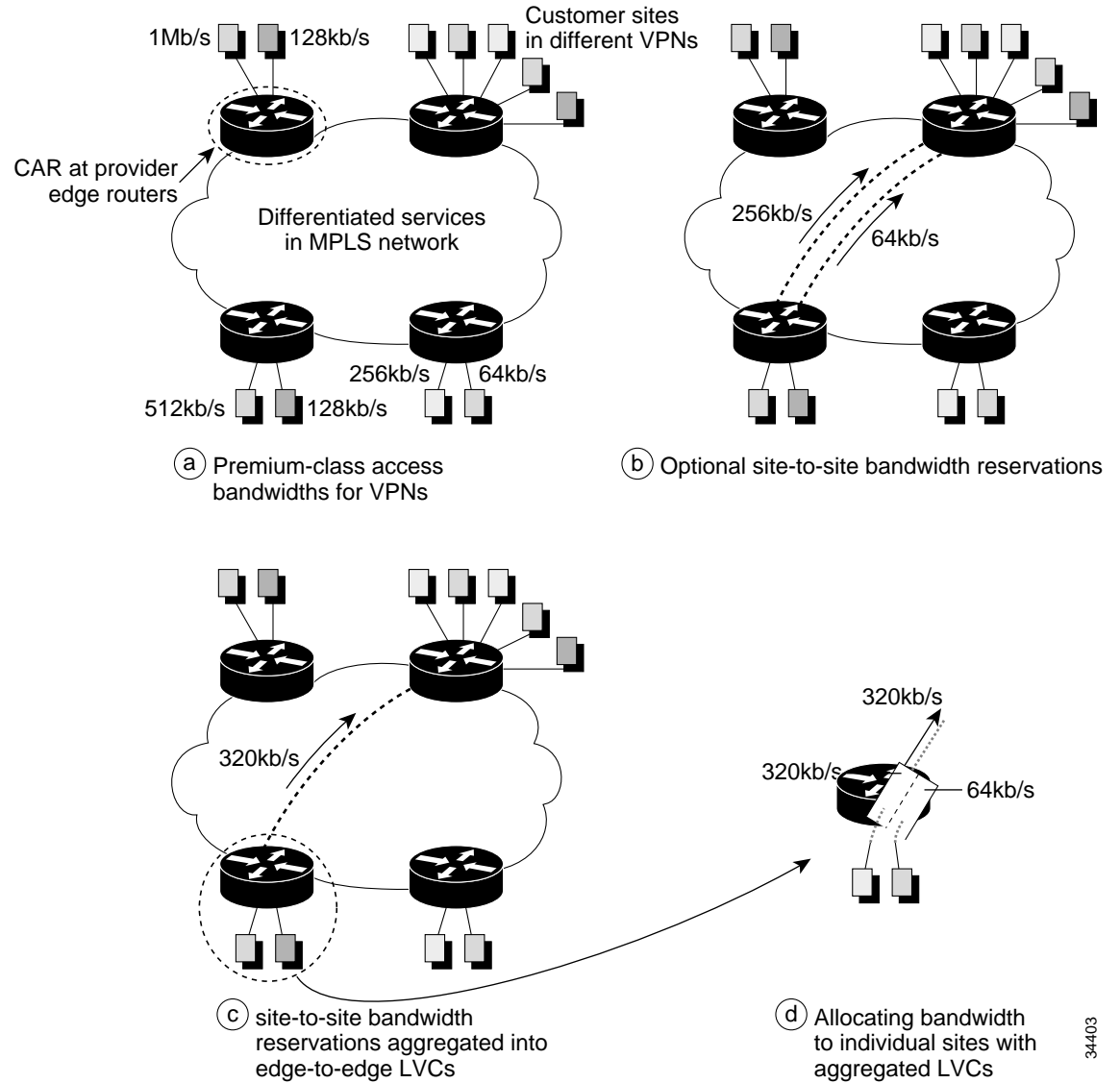
MPLS VPNs have the same QoS options as any other MPLS networks. Sites in VPNs can subscribe to specified rates of specified Classes of Service and the provider can offer connectionless Service Level Agreements for those classes. Cisco committed access rate (CAR) is used at Provider Edge Routers (PER) to enforce traffic contracts. Differentiated Services are used in the network core.

One of the advantages of Cisco MPLS VPNs is that the core LSRs do not have any knowledge or state for individual VPNs. This has advantages for class-based service and fairness. In particular, if premium traffic has precedence over best-effort traffic, then this applies irrespective of which VPNs are the sources of the premium and best-effort traffic.

Site-to-site bandwidth reservations could be used with MPLS VPNs, as shown in Figure 4-13 Topology (b). If there are two separate point-to-point LSPs with the same originating and destination Provider Edge routers, then these may be aggregated into a single LSP as shown in Figure 4-13 Topology (c).

This helps preserve the scalability advantages of MPLS VPNs, specifically the absence of a per-VPN state in the network core. When point-to-point LSPs are aggregated, Weighted Fair Queueing (WFQ) is used to ensure that each site gets its correct share of the aggregated bandwidth, as shown in Topology (d). This means that the QoS achieved with the aggregated LSPs is equivalent to that which is achieved with separate LSPs.

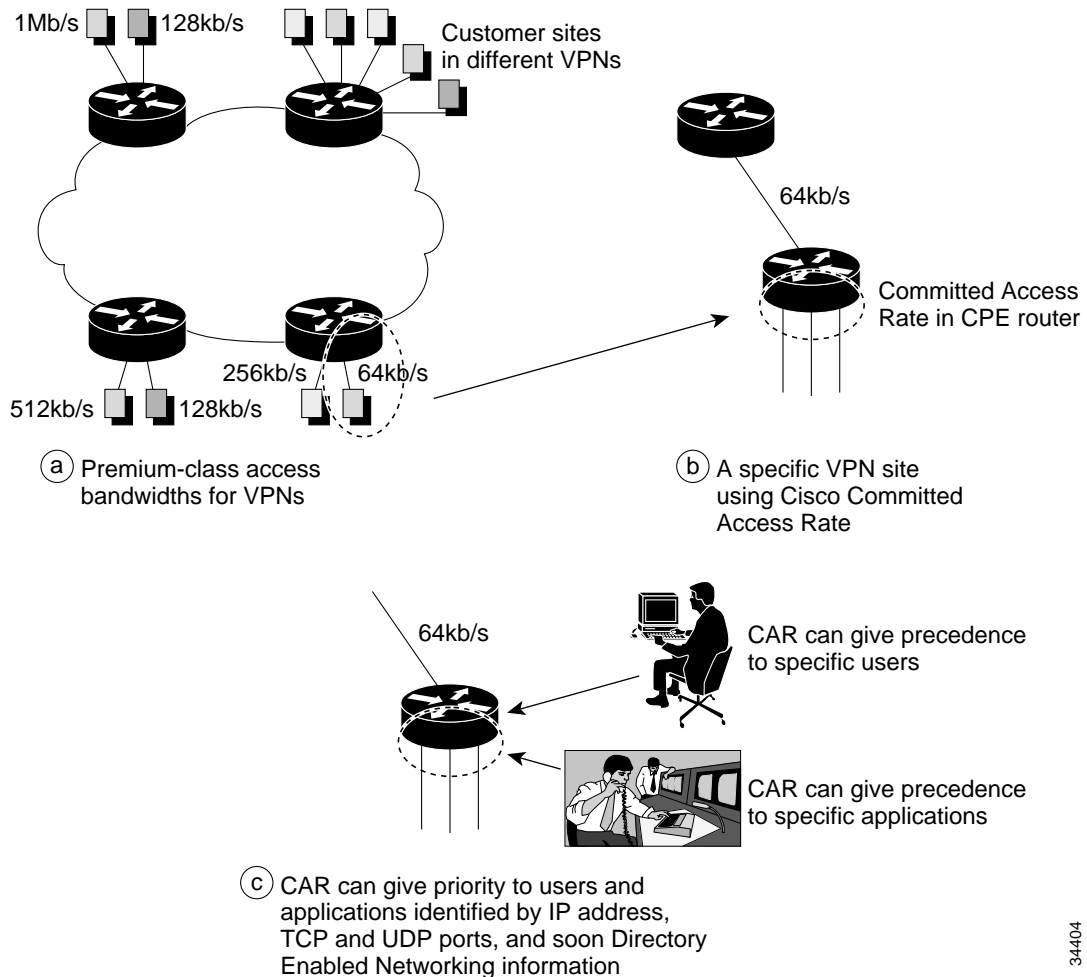
Figure 4-13 Quality of Service in Virtual Private Networks



34403

Sometimes it is desirable to provision bandwidth within VPNs to specific users and applications. This can be achieved by running Cisco CAR on CPE routers. CAR can be used to give precedence or specific bandwidths to specific users or applications, as chosen by the customer. This is illustrated in Figure 4-14.

Figure 4-14 Providing Bandwidth to Specific Users and Applications in Virtual Private Networks



## Discard Policies

This discussion of Classes of Service has described how bandwidth is reserved for different Classes of Service using class-based Weighted Fair Queueing (WFQ). Class-based WFQ is particularly useful for differentiating between Classes of Service, but there are other options. Queues are used at the ingress to each link in a network to ensure efficient use of the link and appropriately differentiation of service.

In a normal queueing system, packets are accepted into a queue up to a discard threshold. Past the discard threshold, 100 percent of arriving packets are discarded. This discard characteristic is shown in Figure 4-15, Topology (a).

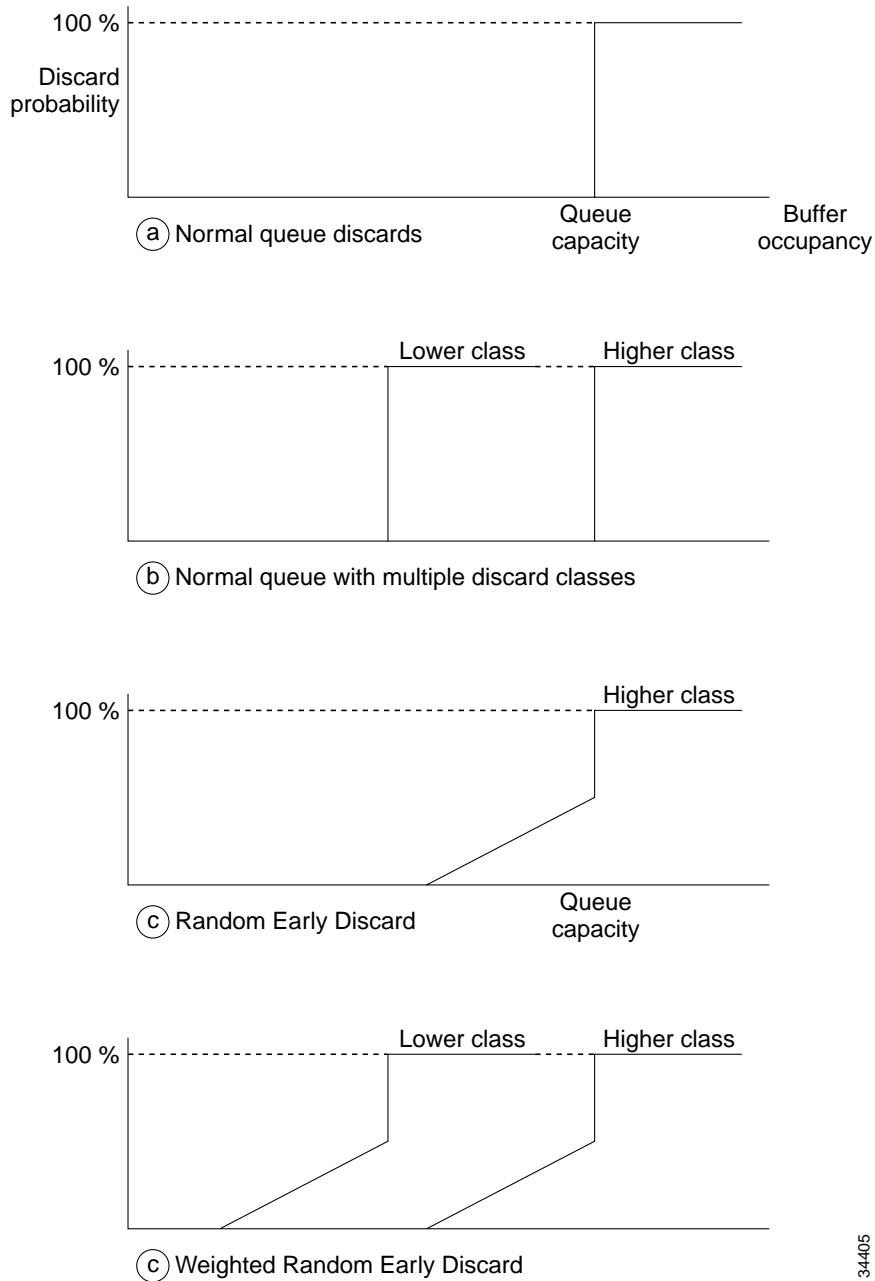
An alternative is to have a single queue with more than one discard class, as shown in Topology (b). This is used in Cisco IP+ATM switches to give Cell Loss Priority (CLP): cells with their CLP bit set are discarded at a lower queue occupancy than cells without their CLP bits set. This gives discard priority to cells without their CLP bits set. CLP-bit setting will enable Cisco Edge LSRs to work in conjunction with CLP on Cisco ATM-LSRs.

An enhancement to hard discard thresholds is to use Random Early Discard (RED). With RED, some packets are randomly discarded below the main discard threshold, shown in Figure 4-15 (c). This has two main advantages:

- Because packets are randomly discarded, there is an easing of any problems with unfair discarding, which can occur when bursty traffic mixes with non-bursty traffic.
- The early discard triggers TCP/IP flow control to reduce flow on TCP/IP streams. This helps prevent congestion and ensures that queues will not actually reach their full discard threshold.

Irrespective of whether WRED is used, there is some suggestion that UDP traffic can get unfairly good QoS compared to TCP traffic because of TCP's behavior of backing off in the presence of packet loss. If this is found to be a problem in practice, it may be solved by transmitting TCP and UDP traffic in different Classes of Service.

Figure 4-15 Discard Policies



As with simple discards, there may be multiple RED characteristics for a single queue to deal with multiple different Classes of Service. This is known as Weighted RED (WRED). Weighted RED is illustrated in Figure 4-15, diagram (d). Cisco routers offer weighted RED, with up to eight discard classes per queue.

Either multiple discard thresholds or WRED can be combined with weighted fair queueing. In the example shown in Figure 4-16:

- Two different queues feed into a single link: one for real-time traffic, and one for normal data traffic.

34405



- Real-time traffic has reserved access to bandwidth  $B_r$  using weighted fair queueing.
- There are two sub-classes of real-time traffic: In-Contract and Out-of-Contract. In-Contract real-time traffic has discard priority over Out-of-Contract real-time traffic. This means:
  - In-Contract real-time traffic has reserved access to bandwidth  $B_r$
  - Out-of-Contract real-time traffic has access to any part of  $B_r$  left unused by the In-Contract traffic. It otherwise has no guarantee of bandwidth whatsoever.
- Normal data traffic has reserved access to bandwidth  $B_c$  using weighted fair queueing.
- There are two sub-classes of normal data traffic: committed and best effort. Committed traffic has discard priority over best-effort traffic. This means:
  - Committed traffic has reserved access to bandwidth  $B_c$
  - Best-effort traffic has reserved access to any part of  $B_c$  left unused by the committed traffic. It otherwise has no guarantee of bandwidth whatsoever.

**Figure 4-16 Example of Combining Weighted Fair Queueing and Differential Discards**

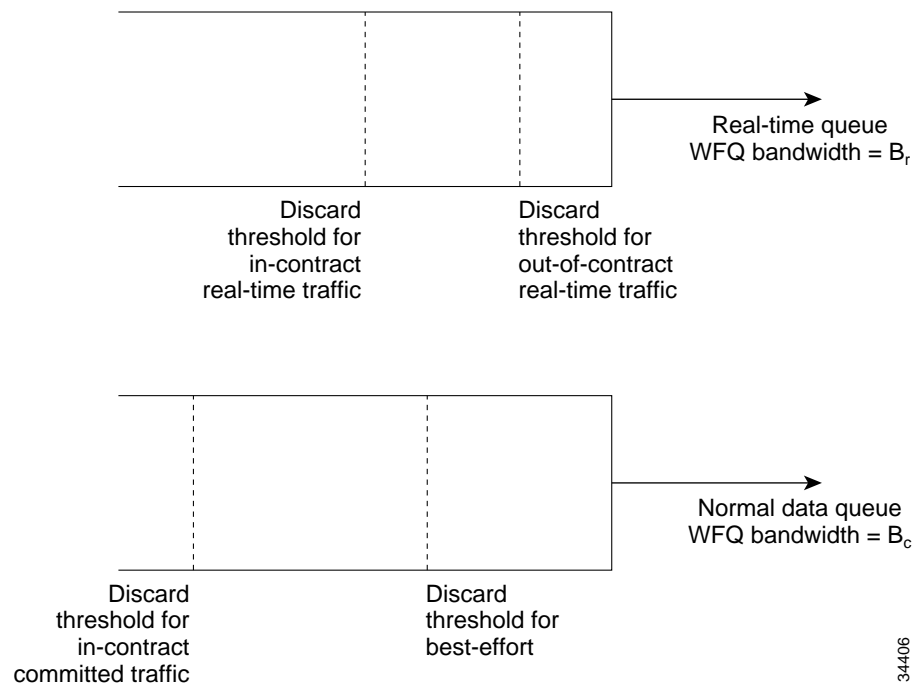
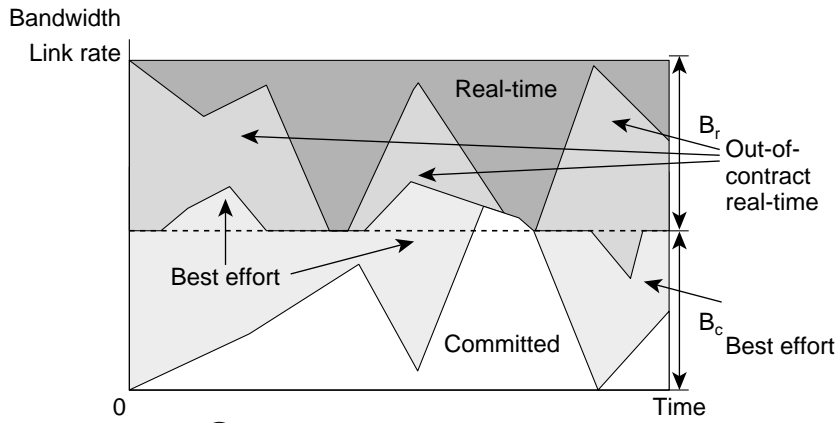
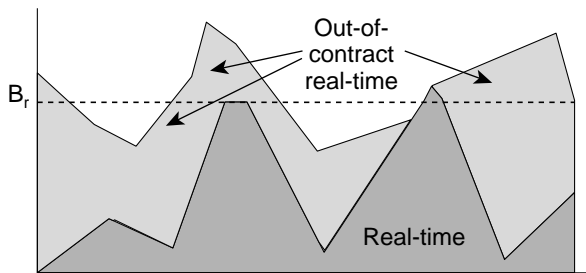


Figure 4-17 shows the effects of offering various mixtures of traffic to the queues shown in Figure 4-16. real-time traffic is has access to  $B_r$  if it needs it, and normal data traffic is has access to  $B_c$ . However, sometimes the out-of-contract real-time traffic or the best-effort traffic gets no bandwidth whatsoever.

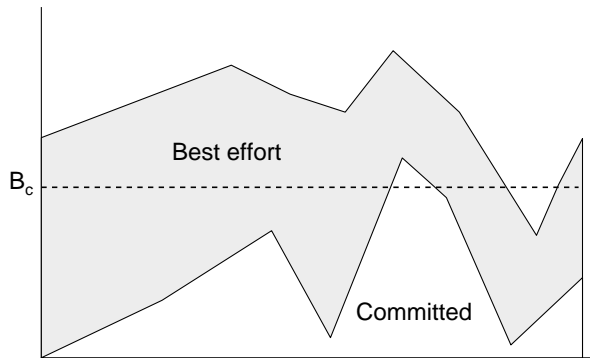
Figure 4-17 Effects of Combining Weighted Fair Queueing and Differential Discards



(a) Traffic accepted on a link



(b) Offered real-time traffic



(c) Offered committed and best-effort traffic

34407

This combination of weighted fair queueing and multiple discard thresholds may be useful for two reasons:

- It may be desirable to give zero bandwidth guarantees to lower classes.
- When two classes are in the same queue, but have different discard thresholds, then packet ordering is guaranteed. If best-effort traffic were in a different queue to committed traffic, then best-effort traffic would often be delivered out of order with respect to committed traffic. Because packet ordering is an important part of IP SLAs, it is important to queue best-effort traffic in the same queue as committed traffic.

## Delay Limits

Delay can be limited by appropriate setting of discard thresholds.

For example, if real-time traffic has a reserved service rate of  $B_r$ , and the discard threshold for real-time traffic is set to  $D_r$ , then the maximum delay at that queue is  $(D_r/B_r)$ .

All discard thresholds in Cisco IP+ATM equipment may be adjusted by the network operator. Real-time IP applications are normally quite tolerant of delay jitter, so it is not clear whether specific engineering for delay jitter is required in MPLS networks. If necessary, low jitter can be ensured by normal traffic engineering methods involving over-allocation of bandwidth to real-time traffic.

A further option in Cisco IP+ATM equipment is to give real-time IP traffic priority access to spare bandwidth, ahead of any other weighted fair queueing classes irrespective of their weight. This is currently available in Cisco hardware and used for ATM Forum CBR and VBR traffic.

In Cisco IP+ATM networks, this over-allocation will typically not result in wasted bandwidth, as other classes have access to bandwidth left unused by real-time traffic.

## Alternative Service Types

In Cisco IP+ATM equipment, QoS is provided for MPLS LVCs with DiffServ Assured Forwarding (AF) classes are supported using class-based weighted fair queueing. There are no per-VC bandwidth allocations and per-VC queueing is not used as it is inconsistent with the DiffServ model.

There is a default configuration. A network operator could override these mappings by using a feature on IP+ATM switches called “Configurable Service Templates.” However, this is not recommended because DiffServ Assured Forwarding assumes the use of class-based service and does not signal any bandwidth parameters. The DiffServ QoS model is quite different from ATM Forum per-VC QoS.

Although the default QoS classes may be overridden with other service classes, including ATM Forum Traffic Management classes, the default mappings have been carefully chosen to be the most appropriate.





## Configuring MPLS with the BPX Switch and the 6400/7200/7500 Routers

---

This chapter provides information for configuring BPX switches and associated label switching controllers along with edge routers for Multiprotocol Label Switching (MPLS) operation.

Procedures are provided for initial configuration of a router and its various interfaces, including ATM and Ethernet interfaces:

- Introduction
- Equipment and Software Requirements
- Configuration Preview
- Initial Setup of MPLS Switching
- Configuration for BPX Switch Portions of the BPX 8650 ATM-LSRs
- Configuration for LSC 1 and LSC 2 Portions of the BPX 8650
- Configuration for Edge Label Switch Routers, LSR-A and LSR-B
- MPLS Configures LVCs According to the Routing Protocol
- Testing the MPLS Network Configuration
- Basic Router Configuration
- Configuring Port Adapter Interfaces
- Checking the Configuration
- Using Configuration Mode
- Cisco IOS Software Basics
- Getting Context-Sensitive Help
- Saving Configuration Changes

For further information regarding the Cisco 6400, 7200, or 7500 series, detailed software configuration information is provided in the Cisco IOS configuration guide and Cisco IOS command reference publications, which are available on the Cisco Documentation CD-ROM.

# Introduction

Configuring the MPLS network consists of setting up ATM router/switches for MPLS. This requires configuring the MPLS controller function on the router entity and the controlled (slave) function on the switch entity of each node.

In the example given here for BPX MPLS nodes (BPX 8650 ATM-LSRs):

- Each MPLS node comprises:
  - a Cisco 6400 or 7200 or 7500 router
  - a BPX switch shelf
- A Cisco 6400 or 7200 or 7500 router provides the controlling function to the BPX switch shelf.

When MPLS is running in the network, the routing protocol (such as OSPF) determines the paths through the MPLS switch network from every Edge Label Switch Router (LSR) to every IP destination. Based on this routing information, MPLS automatically sets up a Label VC (LVC) along each path by using the Label Distribution Protocol (LDP).

Consider packets arriving at the edge of the MPLS network with a particular destination IP address:

1. Labels are applied to these packets at the Edge LSR.
2. The resulting ATM cells are forwarded along the appropriate LVC path through the network using label swapping at each label switch until the far-end Edge LSR is reached.
3. The far-end Edge LSR removes the label, rebuilds the frame, and forwards the IP packet to its LAN destination.

## Equipment and Software Requirements

- BPX 8650
  - BCC-3-64, BCC-4-64, BCC-4-128
- BXM FW
- LSC Router:
  - 7200 Series Router with NPE-150, NPE-200, or 7200VXR processor
  - 7500 Series Router with RSP-2 or RSP-4 processor
  - Cisco 6400
  - 32 MB minimum, 64 MB recommended memory
- IOS:
  - 12.0T(5) or later, IP-only release recommended
- SWSW:
  - 9.2.10 or later

# Configuration Preview

Setting up label switching on a node involves is essentially a three-step process:

1. Configuring BPX switch
  - a. BPX switch (label switch slaves) configuration
  - b. Router (label switch controller) configuration of router extended ATM interfaces on the BPX for tag switching
2. Setting up edge routers (can include setting up policies, and so on)
3. IP routing (typically OSPF or IS-IS) automatically discovers the network topology
4. MPLS automatically sets up LVCs across the network

Figure 5-1 shows a high-level view of an MPLS network. The packets destined for 204.129.33.127 could be real time video, while the packets destined for 204.133.44.129 could be from data files.

Once IP routing and MPLS have been set up on the nodes as shown in Figure 5-1, (ATM-LSR 1 through ATM-LSR 5, Edge LSR\_A, Edge LSR\_B, and Edge LSR\_C), automatic network discovery is thereby enabled. Then MPLS will automatically set up LVCs across the network. At each ATM LSR (label switch), label swapping is used to transport the cells across the previously set up LVC paths.

(“Label swapping” is a name for VCI switching, the underlying capability of an ATM switch.)

At the Edge LSRs, labels are added to incoming IP packets and labels are removed from outgoing packets. Figure 5-1 shows IP packets with host destination 204.129.33.127 transported as labeled ATM cells across LVC 1, and IP packets with host destination 204.133.44.129 transported as labeled ATM cells across LVC 2.

IP addresses shown are for illustrative purposes only and are assumed to be isolated from external networks. Check with your Network Administrator for appropriate IP addresses for your network.

Figure 5-1 High-Level View of Configuration of An MPLS Network

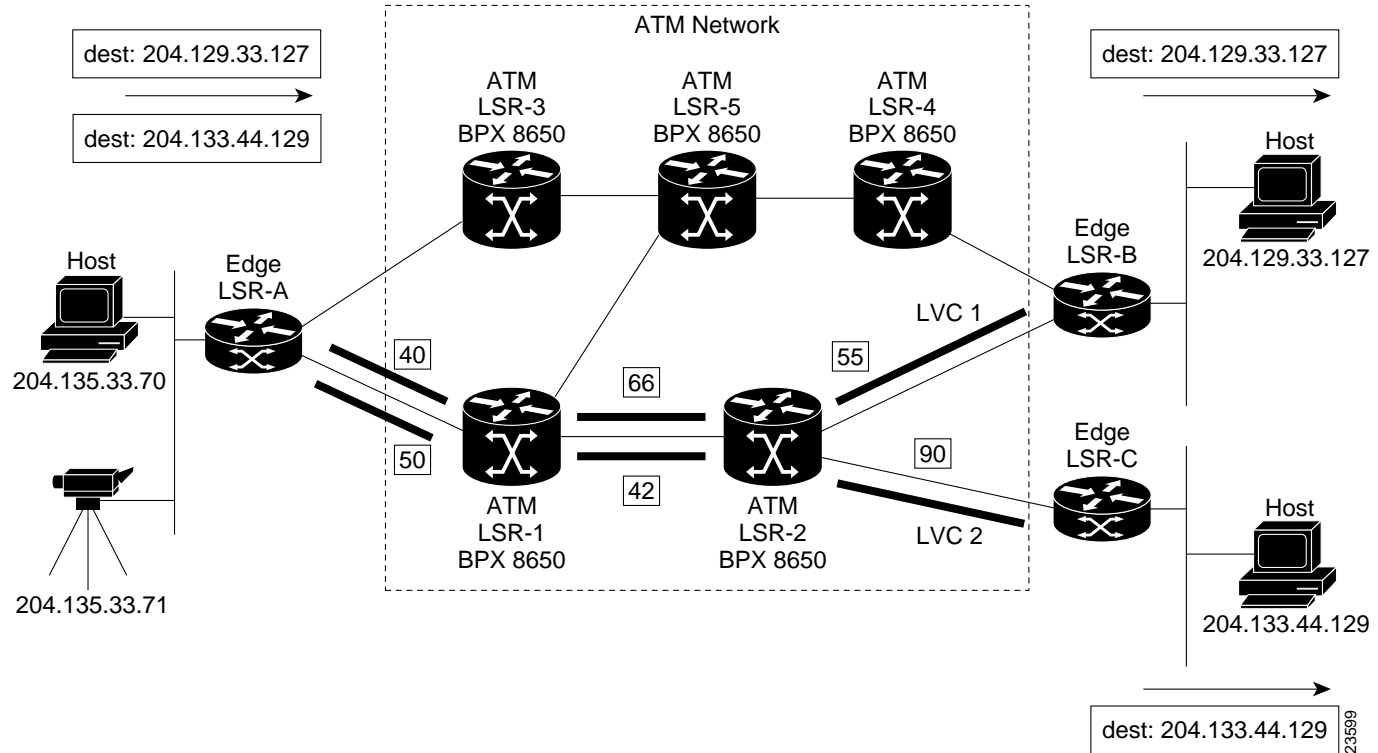
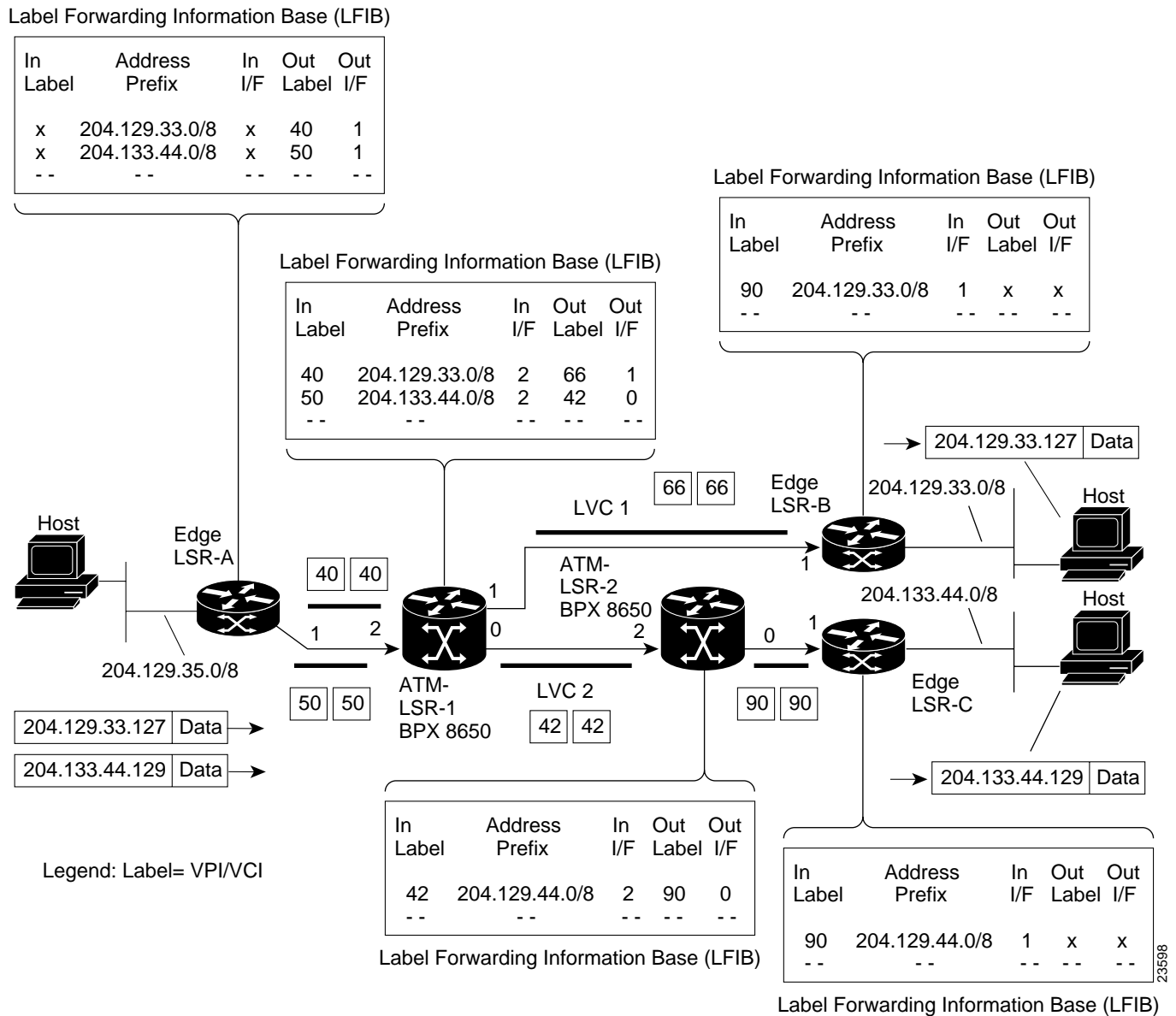


Figure 5-2 is a detailed diagram showing the MPLS label swapping that might take place in the transportation of the IP packets in the form of ATM cells across the network on the LVC1 and LVC2 virtual circuits:

1. An unlabeled IP packet with destination 204.133.44.129 arrives at Edge Label Switching Router (LSR-A).
2. Edge LSR-A checks its label forwarding information base (LFIB) and matches the destination with prefix 204.133.44.0/8.
3. LSR-A converts the AAL5 frame to cells and sends the frame out as a sequence of cells on 1/VCI 50.
4. ATM-LSR-1 (which is a BPX 8650 Label Switch Router) is controlled by a Label Switch Controller (6400, 7200, or 7500 router). The controller has an LFIB that was established by IP routing and MPLS signaling. At the time the LFIB entries were established, the controller used the VC switching information in the LFIB to establish VC connections in the switch. In this case, the incoming cells on interface 2/VCI 50 are switched to outgoing interface 0/VCI 42. The cell-by-cell switching is invisible to the controller, because the traffic is carried only by the switch and does not pass through the controller.
5. Similarly, at ATM-LSR-2, the incoming cells on interface 2/VCI 42 are switched to outgoing interface 0/VCI 90, according to the LFIB.
6. Edge LSR-C receives the incoming cells on incoming interface 1/VCI 90, checks its LFIB, converts the ATM cells back to an AAL5 frame, then to an IP packet, and then sends the outgoing packet onto its LAN destination 204.133.44.129.



Figure 5-2 Label Swapping Detail



23598

## Initial Setup of MPLS Switching

This section provides an example of configuring BPX 8650 MPLS label switches (ATM-LSRs) for MPLS switching of IP packets through an ATM network, along with configuration for 6400/7200/7500 routers for use as Label Edge Switch Routers (Edge LSRs) at the edges of the network.

The example in this section describes the configuration of:

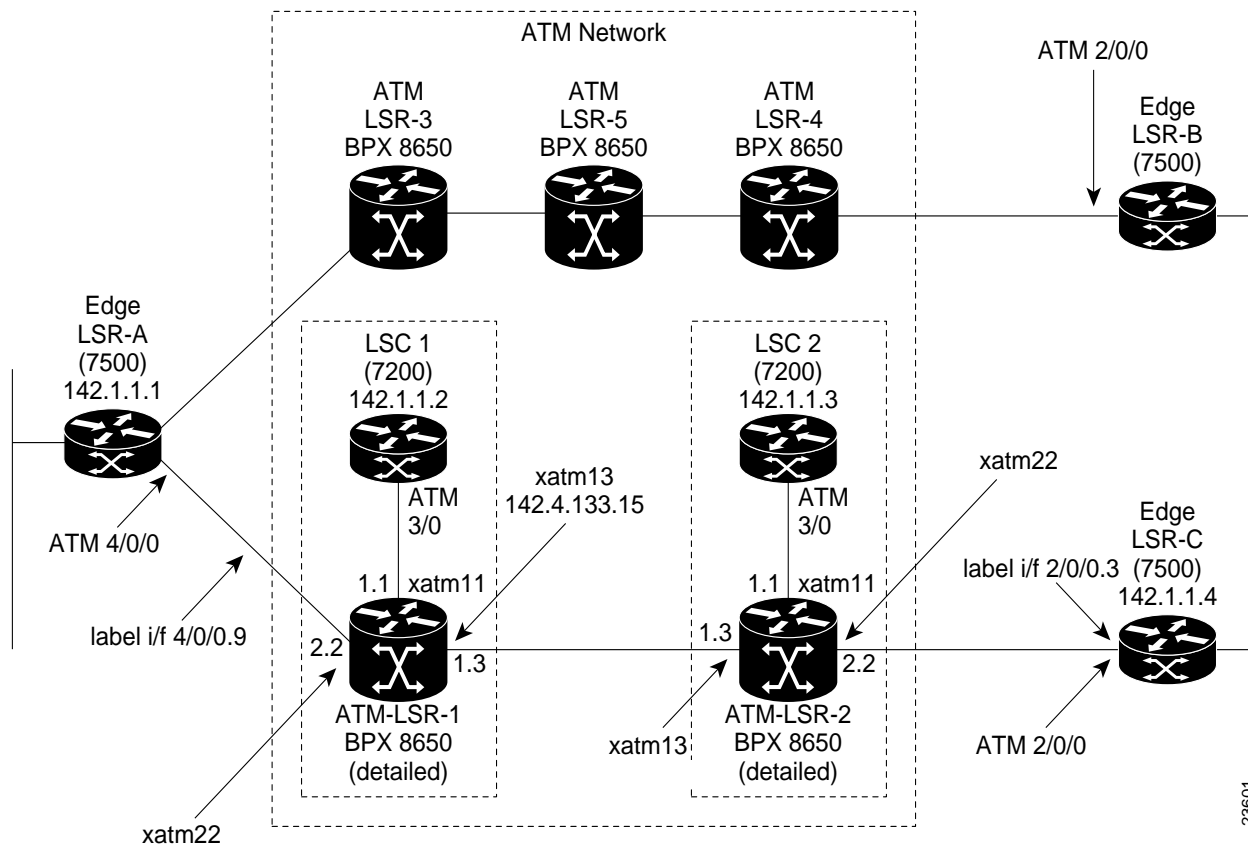
- Edge LSR-A (7500 router)
- Edge LSR-C (7500 router)
- ATM LSR-1 (BPX 8650 switch and controller)
- ATM LSR-2 (BPX 8650 switch and controller) as shown in Figure 5-3

The configuration of ATM LSR-3, ATM LSR-4, and ATM LSR-5 is not detailed, but would be performed in a similar manner to that for ATM LSR-1 and ATM LSR-2. Also, the configuration of Edge LSR-B (7500 router) would be similar to that for Edge LSR-A and LSR-C.

The configuration of a BPX 8650 ATM-LSR consists of two parts:

- Configuring the BPX switch
- Configuring the associated label switch controller (Cisco 6400/7200/7500 routers)

Figure 5-3 Simplified Example of Configuring An MPLS Network



# Configuration for BPX Switch Portions of the BPX 8650 ATM-LSRs

The BPX nodes must be set up and configured in the ATM network, including links to other nodes, and so on. Following this, they may be configured for MPLS operation.

To configure the BPX nodes for operation, you set up a virtual switch interface and associated partition by using the **cnfrsrc** command.

You link the 6400, 7200, or 7500 router to the BPX by using the **addshelf** command to allow the router's label switch controller function to control the MPLS operation of a node.

You may distribute the resources of the partition between the associated ports. Resources include bandwidth, VPI range, and number of logical connection spaces (LCNs). The VPIs are of local significance, so they do not have to be the same for each port in a node, but it is generally convenient from a tracking standpoint to keep them the same for a given BPX node.

In this example, it is assumed that a single external controller per node is supported, so that the partition chosen is always 1.

## Command Syntax Summary for BPX Portion of MPLS Configuration

Syntax for associated commands, **cnfrsrc**, **cnfqbin**, **addshelf** are:

```
cnfrsrc slot.port.{virtual trk} maxpvcLens maxpvcbw [Edit parms ? y/n] partitionID
e/d minvSilcns maxvSilcns vsistartvpi vsiendvpi vsiminbw vsimaxbw
{if you enter "y", to Edit parms?}
```

```
cnfrsrc slot.port.{virtual trk} maxpvcLens maxpvcbw [Edit parms ? y/n]
{accepts defaults if you enter "n" to Edit parms}
```



```
cnfqbin <slot.port> <Qbin_#> <e/d> y/n <Qbin discard_thr> <Low EPD threshold> <CLPhi>
<EFCTI_thr>
{If you enter "n" to not accept template values}
```

```
cnfqbin <slot.port.[virtual trk]> <Qbin_#> <e/d> y/n
{If you enter "y" to accept template values.}
```

```
addshelf <slot.port [virtual trk]> <device-type> <control ID> <control partition ID>
```

## Configuration for BPX 1 Portion of ATM-LSR-1

To configure the BPX 8650 label switch routers, ATM-LSR-1 and ATM-LSR-2:

	Command	Description
Step 1	Check card status:  <b>dspecds</b>	Display status of all cards, BXM cards that you are configuring should be “Standby” or “Active”. If not perform a hard reset, “resetcd 1 h”, resets card 1, for example.
Step 2	Check card connection capabilities:  <b>dspecd 1</b>  Chnls:16320, PG[1] :7048, PG[2] : 7048 PG [1] : 1, 2 PG [2] : 3, 4  <b>dspecd 2</b>  Chnls:16320, PG[1] :7048, PG[2] : 7048 PG [1] : 1, 2 PG [2] : 3, 4	This example shows that ports 1 and 2 together have a total of 7048 connections or “channels” available for use. Ports 1 and 2 form a port group (PG). Similarly, ports 3 and 4 are a port group with a limit of 7048 connections. Unless there is a good reason to do otherwise, it is best to leave many of the LCNs as spares. In this example, we will allocate 1500 LCNs to MPLS on each port using the <b>cnfrsrc</b> command.
Step 3	Enable BXM interfaces:  <b>uptrk 1.1</b>  <b>uptrk 1.3</b>  <b>uptrk 2.2</b>	In this example, trunk 1.1 is the link to the LSC controller, and trunks 1.3 and 2.2 are set up for use by LVCs.   <b>Note</b> A BXM interface is a “trunk” if it connects to another switch or MGX 8220 feeder. The VSI connection to an LSC is also a “trunk.” Other interfaces are ports, typically to service interfaces. Although interfaces 1.3 and 2.2 are configured as trunks, in this example, MPLS would also work if they were configured as ports. However, if links between BPX nodes carry PVC connections as well as MPLS, they must be trunks.   <b>Note</b> The <b>uptrk</b> and related commands are of form <b>uptrk &lt;slot.port. [&lt;virtual trk]&gt;</b> , so if you are configuring a virtual trunk the uptrk command for example, would be of the form, <b>uptrk 1.1.1</b> , <b>uptrk 1.1.2</b> , and so on. Either ports or trunks can be active simultaneously on the same BXM.

Command	Description
<p>Step 4 Configure VSI partitions on the BXM interfaces:</p> <pre><b>cnfrsrc 1.1 256 26000 y 1 e 512 1500 240 255 105000 105000</b></pre> <p>or if entered individually:</p> <pre><b>cnfrsrc 1.1</b> <b>256</b> {PVC LCNs, accept default value} <b>26000</b> <b>y</b> {to edit VSI parameters} <b>1</b> {partition} <b>e</b> {enable partition} <b>512</b> {VSI min LCNs} <b>1500</b> {VSI max LCNs} <b>240</b> {VSI starting VPI} <b>255</b> {VSI ending VPI} <b>105000</b> {VSI min bandwidth} <b>105000</b> {VSI max bandwidth}</pre> <p>Repeat for BXM interfaces 1.3 and 2.2</p> <pre><b>cnfrsrc 1.3 256 26000 y 1 e 512 1500 240 255 105000 105000</b> <b>cnfrsrc 2.2 256 26000 y 1 e 512 1500 240 255 105000 105000</b></pre>	<p>PVC LCNs: [256] default value. Reserves space on this link for 256 AutoRoute PVCs (LCNs = Logical Connection Numbers).</p> <p>Three VSI partitions are supported, numbered 1 through 3. If in doubt, use partition 1 for MPLS.</p> <p>VSI min LCNs: 512 and VSI max LCNs: 1500. Guarantees that MPLS can set up 512 LVCs on this link, but is allowed to use up to 1500, subject to availability of LCNs.</p> <p>VSI starting VPI: 240 and VSI ending VPI: 255. Reserves VPIs in the range of 240-255 for MPLS. Only one VP is really required, but a few more can be reserved for future use. AutoRoute uses a VPI range starting at 0, so MPLS should use higher values. It is best to always avoid using VPIs "0" and "1" for MPLS on the BPX 8650.</p> <p>VPIs are locally significant. In this example, 240 is shown as the starting VPI for each port. A different value could be used for each of the three ports, 1.1, 1.3, and 2.2. But at each end of a trunk, such as, between port 1.3 on ATM LSR-1 and port 1.3 on ATM LSR-2, the same VPI must be assigned.</p> <p>VSI minimum bandwidth: 105000 and VSI maximum 105000. Guarantees that MPLS can use 105000 cps (about 40 Mbps) on this link. More can be allocated if required.</p> <p>MPLS will never use an assigned bandwidth greater than its minimum bandwidth. The maximum bandwidth sets the maximum that can be pre-allocated to connections of the specified partition. The maximum bandwidth is a maximum for connection admission purposes only. It can be exceeded on a cell-by-cell basis if: a) there are bursts of traffic on connections in this partition, and b) bandwidth is left unallocated, or unused or in connections in other partitions.</p> <p>All MPLS connections on a link share the guaranteed minimum bandwidth for MPLS. There is no point in setting the maximum bandwidth greater than the minimum for MPLS. Irrespective of the maximum setting, MPLS connections can burst into spare bandwidth in other partitions, as just noted.</p> <p>PVC maximum bandwidth: 26000. Guarantees that PVCs can always use up to 26000 cells per second (about 10 Mbps) on this link.</p>

	Command	Description
Step 5	<p>Enable MPLS queues on BXM:</p> <p><b>dsqbin 1.1 10</b></p> <p>and verify that it matches the following:</p> <pre>Qbin Database 1.1 on BXM qbin 10 Qbin State: Enable Qbin discard threshold: 65536 EPD threshold: 95% High CLP threshold: 100% EFCI threshold: 40%</pre> <p>If configuration is not correct, enter</p> <p><b>cnfqbin 1.1 10 e n 65536 95 100 40</b></p> <p>Repeat as necessary for BXM interfaces 1.3 and 2.2:</p> <p><b>cnfqbin 1.3 10 e n 65536 95 100 40</b></p> <p><b>cnfqbin 2.2 10 e n 65536 95 100 40</b></p>	MPLS CoS uses Qbins 10-14.
Step 6	<p>Enable the VSI control interface:</p> <p><b>addshelf 1.1 vsi 1 1</b> {link to controller, ID = 1, partition = 1}</p>	<p>The first “1” after “vsi” is the VSI controller ID, which must be set the same on both the BPX 8650 and the LSC. The default controller ID on the LSC is “1”.</p> <p>The second “1” after “vsi” indicates that this is a controller for partition 1.</p>

## Configuration for BPX 2 Portion of ATM-LSR-2

Proceed with configuration as follows:

	Command	Description
Step 1	<p>Check card status:</p> <p><b>dspcds</b></p>	Display status of all cards, BXM cards that you are configuring should be “Standby” or “Active”. If not perform a hard reset, “resetcd 1 h”, resets card 1, for example.
Step 2	<p>Check card connection capabilities:</p> <p><b>dspcd 1</b></p> <pre>Chnls:16320, PG[1] :7048, PG[2] : 7048 PG [1] : 1, 2 PG [2] : 3, 4</pre> <p><b>dspcd 2</b></p> <pre>Chnls:16320, PG[1] :7048, PG[2] : 7048 PG [1] : 1, 2 PG [2] : 3, 4</pre>	This example shows that ports 1 and 2 together have a total of 7048 connections or “channels” available for use. Ports 1 and 2 form a port group (PG). Similarly, ports 3 and 4 are a port group with a limit of 7048 connections. Unless there is a good reason to do otherwise, it is best to leave many of the LCNs as spares. In this example, we will allocate 1500 LCNs to MPLS on each port using the <b>cnfrsrc</b> command.

Command	Description
<p>Step 3 Enable BXM interfaces:</p> <pre> <b>uptrk 1.1</b> <b>uptrk 1.3</b> <b>uptrk 2.2</b> </pre>	<p>In this example, trunk 1.1 is the link to the LSC controller, and trunks 1.3 and 2.2 are being set up as cross-connects for use by LVCs.</p>
<p>Step 4 Configure VSI partitions on the BXM interfaces:</p> <pre> <b>cnfrsrc 1.1 256 26000 y 1 e 512 1500 240 255 105000</b> <b>105000</b> </pre> <p>or if entered individually:</p> <pre> <b>cnfrsrc 1.1</b> <b>256</b> {PVC LCNs, accept default value} <b>26000</b> <b>y</b> {to edit VSI parameters} <b>1</b> {partition} <b>e</b> {enable partition} <b>512</b> {VSI min LCNs} <b>1500</b> {VSI max LCNs} <b>240</b> {VSI starting VPI} <b>255</b> {VSI ending VPI} <b>105000</b> {VSI min bandwidth} <b>105000</b> {VSI max bandwidth} </pre> <p>Repeat for BXM interfaces 1.3 and 2.2</p> <pre> <b>cnfrsrc 1.3 256 26000 y 1 e 512 1500 240 255 105000</b> <b>105000</b> <b>cnfrsrc 2.2 256 26000 y 1 e 512 1500 240 255 105000</b> <b>105000</b> </pre>	

	Command	Description
Step 5	<p>Enable MPLS queues on BXM:</p> <p><b>dspqbin 1.1 10</b></p> <p>and verify that it matches the following:</p> <pre> Qbin Database 1.1 on BXM qbin 10 Qbin State: Enable Qbin discard threshold: 65536 EPD threshold: 95% High CLP threshold: 100% EFCI threshold: 40% </pre> <p>If configuration is not correct, enter</p> <p><b>cnfqbin 1.1 10 e n 65536 95 100 40</b></p> <p>Repeat as necessary for BXM interfaces 1.3 and 2.2:</p> <p><b>cnfqbin 1.3 10 e n 65536 95 100 40</b></p> <p><b>cnfqbin 2.2 10 e n 65536 95 100 40</b></p>	MPLS CoS uses Qbins 10-14.
Step 6	<p>Enable the VSI control interface:</p> <p><b>addshelf 1.1 vsi 1 1</b> {link to controller, ID = 1, partition = 1</p>	<p>The first “1” after “vsi” is the VSI controller ID, which must be set the same on both the BPX 8650 and the LSC. The default controller ID on the LSC is “1”.</p> <p>The second “1” after “vsi” is the partition ID that indicates this is a controller for partition 1.</p>

## Configuration for LSC 1 and LSC 2 Portions of the BPX 8650

Before configuring the routers for the label switch (MPLS) controlling function, it is necessary to perform the initial router configuration.

As part of this configuration, it is necessary to enable and configure the ATM Adapter interface as described in “Configuring ATM Interfaces” section on page 5-28.

Then the extended ATM interface can be set up for Label Switching, and the BPX ports configured by the router as extended ATM ports of the router physical ATM interface according to the following procedures for LSC1 and LSC2.

### Configuration for LSC1 Portion of ATM-LSR-1

	Command	Description
	Preliminary	
Step 1	Router LSC1(config)# ip routing	Enable IP routing protocol.
Step 2	Router LSC1(config)# ip cef switch	Enable Cisco express forwarding protocol.
Step 3	interface Loopback0	Define a loopback address. This is an IP address for the LSC itself, and not for a link on the LSC or BPX.



	Command	Description
Step 4	ip address 142.1.1.2 255.255.255.255	Assigning IP address to Loopback0. It is important that all loopback addresses in an MPLS network are host addresses, that is, with a mask of 255.255.255.255. Using a shorter mask can prevent MPLS-based VPN services from working correctly.
Step 5	Router LSC1(config)# interface ATM3/0	Enable physical interface link to BPX.
Step 6	Router LSC1(config-if)# no ip address	
Step 7	Router LSC1(config-if)# tag-control-protocol vsi [controller ID]	Enable router ATM port ATM3/0 as tag switching controller. Controller ID default is 1, optional values up to 32 for BPX.
	<b>Setting up interslave control link</b>	
Step 8	Router LSC1(config-if)# interface XtagATM13	Interslave link on 1.3 port of BPX (port 3 of BXM in slot 1). This is an extended port of the router ATM3/0 port.
Step 9	Router LSC1(config-if)# extended-port ATM3/0 bpx 1.3	Binding extended port xtagATM13 to BPX slave port 1.3.
Step 10	Router LSC1(config-if)# ip unnumbered Loopback0	Make xtagATM[13/22] an unnumbered IP link, using Loopback0's IP address as a substitute for the link IP address. The interfaces in an ATM MPLS network should usually be unnumbered. This reduces the number of IP destination-prefixes in the routing table, and hence reduces the number of labels and LVCs used in the network.
Step 11	Router LSC1(config-if)# tag-switching ip	Enable MPLS for xtag interface xtagATM13.
	<b>Setting up interslave port</b>	
Step 12	Router LSC1(config-if)# interface XtagATM22	Interslave link on 2.2 port of BPX (port 2 of BXM in slot 2). This is an extended port of the router ATM3/0 port.
Step 13	Router LSC1(config-if)# extended-port ATM3/0 bpx 2.2	Binding extended port xtagATM22 to BPX slave port 2.2
Step 14	Router LSC1(config-if)# ip unnumbered Loopback0	Make xtagATM[13/22] an unnumbered IP link, using Loopback0's IP address as a substitute for the link IP address. The interfaces in an ATM MPLS network should usually be unnumbered. This reduces the number of IP destination-prefixes in the routing table, and hence reduces the number of labels and LVCs used in the network.
Step 15	Router LSC1(config-if)# tag-switching ip	Enable MPLS for xtag interface xtagATM22.
Step 16	Router LSC1 (config-if)# exit	

	Command	Description
	<b>Configuring routing protocol</b>	
Step 17	Router LSC1 (config-if)# Router OSPF 5	Configuring Open Shortest Path First (OSPF) routing protocol or Enhanced Interior Gateway Routing Protocol (EIGRP).
Step 18	Router LSC1 (config-if)# network 142.1.0.0 0.0.255.255 area 0	Setting up OSPF routing and assigning a process ID of 5 which is locally significant. The ID may be chosen from a wide range of available process ID up to approximately 32,000.

## Configuration for LSC2 Portion of ATM-LSR-2

	Command	Description
	<b>Preliminary</b>	
Step 1	Router LSC2(config)# ip routing	Enable IP routing protocol.
Step 2	Router LSC2(config)# ip cef switch	Enable Cisco express forwarding protocol.
Step 3	interface Loopback0	Define a loopback address. This is an IP address for the LSC itself, and not for a link on the LSC or BPX.
Step 4	ip address 142.1.1.3 255.255.255.255	Assigning IP address to Loopback0. It is important that all loopback addresses in an MPLS network are host addresses, that is, with a mask of 255.255.255.255. Using a shorter mask can prevent MPLS-based VPN services from working correctly.
Step 5	Router LSC2(config)# interface ATM3/0	Enable physical interface link to BPX.
Step 6	Router LSC2(config-if)# no ip address	
Step 7	Router LSC2(config-if)# tag-control-protocol vsi [controller ID]	Enable router ATM port ATM3/0 as tag switching controller. Controller ID default is 1, optional values up to 32 for BPX.
	<b>Setting up interslave control link</b>	
Step 8	Router LSC2(config-if)# interface XtagATM13	Interslave link on 1.3 port of BPX (port 3 of BXM in slot 1). This is an extended port of the router ATM3/0 port.
Step 9	Router LSC2(config-if)# extended-port ATM3/0 bpx 1.3	Binding extended port xtagATM13 to BPX slave port 1.3.
Step 10	Router LSC2(config-if)# ip unnumbered Loopback0	Make xtagATM[13/22] an unnumbered IP link, using Loopback0's IP address as a substitute for the link IP address. The interfaces in an ATM MPLS network should usually be unnumbered. This reduces the number of IP destination-prefixes in the routing table, and hence reduces the number of labels and LVCs used in the network.
Step 11	Router LSC2(config-if)# tag-switching ip	Enable MPLS for xtag interface xtagATM1.

	Command	Description
	<b>Setting up interslave port</b>	
Step 12	Router LSC2(config-if)# interface XtagATM22	Interslave link on 2.2 port of BPX (port 2 os BXM in slot 2). This is an extended port of the router ATM3/0 port.
Step 13	Router LSC2(config-if)# extended-port ATM3/0 bpx 2.2	Binding extended port xtagATM22 to bpx slave port 2.
Step 14	Router LSC2(config-if)# ip unnumbered Loopback0	Make xtagATM[13/22] an unnumbered IP link, using Loopback0's IP address as a substitute for the link IP address. The interfaces in an ATM MPLS network should usually be unnumbered. This reduces the number of IP destination-prefixes in the routing table, and hence reduces the number of labels and LVCs used in the network.
Step 15	Router LSC2(config-if)# tag-switching ip	Enable MPLS for xtag interface xtagATM22.
Step 16	Router LSC2 (config-if)# exit	
	<b>Configuring routing protocol</b>	
Step 17	Router LSC2 (config-if)# Router OSPF 5	Configuring Open Shortest Path First (OSPF) routing protocol or Enhanced Interior Gateway Routing Protocol (EIGRP).
Step 18	Router LSC2 (config-router)# network 142.1.1.3 0.0.255.255 area 0	Setting up OSPF routing and assigning a process ID of 5 which is locally significant. You may choose the ID from a wide range of available process IDs up to approximately 32,000.

## Configuration for Edge Label Switch Routers, LSR-A and LSR-B

Before configuring the routers for the MPLS controlling function, it is necessary to perform the initial router configuration.

As part of this configuration, you must enable and configure the ATM Adapter interface as described in “Configuring ATM Interfaces” section on page 5-28.

Then you can set up the extended ATM interface for MPLS, and the BPX ports configured by the router as extended ATM ports of the router physical ATM interface according to the following procedures for LSR-A and LSR-C.

To configure the 7500 routers performing as label edge routers, use the procedures in the following tables.

## Configuration of Cisco 7500 as An Edge Router, Edge LSR-A

	Command	Description
Step 1	Router LSR-A (config)# ip routing	Enable IP routing protocol.
Step 2	Router LSR-A(config)# ip cef distributed switch	Enable label switching for ATM subinterface.
Step 3	Router LSR-A(config)# interface Loopback0	Define a loopback address.
Step 4	Router LSR-A(config)# ip address 142.1.1.1 255.255.255.255	Assigning IP address to Loopback0.
Step 5	Router LSR-A(config)# interface ATM4/0/0	
Step 6	Router LSR-A(config-if)# no ip address	
Step 7	Router LSR-A(config-if)# interface ATM4/0/0.9 label switching	Interface can be basically any number within range limits ATM4/0/0.1, ATM 4/0/0.2, and so on.
Step 8	Router LSR-A(config-if)# ip unnumbered Loopback0	Network 142.1.0.0 0.0.255.255 area 0.
Step 9	Router LSR-A(config-if)# tag-switching ip	
	<b>Configuring routing protocol</b>	
Step 10	Router LSR-A (config-if)# Router OSPF 5	Configuring Open Shortest Path First (OSPF) routing protocol or Enhanced Interior Gateway Routing Protocol (EIGRP).
Step 11	Router LSR-A (config-router)# network 142.1.0.0 0.0.255.255 area 0	Setting up OSPF routing and assigning a process ID of 5 which is locally significant. The ID may be chosen from a wide range of available process IDs up to approximately 32,000.

## Configuration of Cisco 7500 as An Edge Router, Edge LSR-C

	Command	Description
Step 1	Router LSR-C (config)# ip routing	Enable IP routing protocol.
Step 2	Router LSR-C(config)# ip cef distributed switch	Enable label switching for ATM subinterface.
Step 3	Router LSR-C(config)# interface Loopback0	Define a loopback address.
Step 4	Router LSR-C(config)# ip address 142.1.1.4. 255.255.255.255	Assigning IP address to Loopback0.
Step 5	Router LSR-C(config)# interface ATM2/0/0	
Step 6	Router LSR-C(config-if)# no ip address	
Step 7	Router LSR-C(config-if)# interface ATM2/0/0.3 tag-switching	
Step 8	Router LSR-C(config-if)# ip unnumbered Loopback0	Network 142.1.1.4. 0.0.255.255 area 0.
Step 9	Router LSR-C(config-if)# tag-switching ip	



# Testing the MPLS Network Configuration

Preliminary testing of the MPLS network consists of:

- Checking VSI status
- Checking the MPLS interfaces
- Checking the MPLS discovery process

## Useful LSC Commands

The following are some of the useful LSC (also referred to as TSC) commands for monitoring and troubleshooting an MPLS network:

```
show controllers VSI descriptor [descriptor]
```

```
show tag int
```

```
show tag tdp disc
```

For a complete description of these LSC commands refer to the related IOS MPLS documentation:

- Label Switching on Cisco 7000 Family
- Label Switch Controller

## Checking the BPX Extended ATM Interfaces

Use the following procedure as a quick checkout of the tag switching configuration and operation with respect to the BPX switch, for example ATM LSR-1:

- Step 1** Wait a while. Then check whether the controller sees the interfaces correctly; on LSC1, for example, enter the following command:

Command	Description
Router LSC1# show controllers VSI ATM3/0	Shows VSI information for extended ATM interfaces.

The sample output for ATM-LSC-1 (BPX 8650 shelf) is:



**Note** Check the LSC online documentation for the most current information.

```

Phys desc: 1.1
Log intf: 0x00040100 (0.4.1.0)
Interface: slave control port
IF status: n/a                    IFC state: ACTIVE
Min VPI: 0                        Maximum cell rate: 10000
Max VPI: 10                       Available channels: xxx
Min VCI: 0                        Available cell rate (forward): xxxxxxx
Max VCI: 65535                   Available cell rate (backward): xxxxxxx

Phys desc: 1.3
Log intf: 0x00040200 (0.4.2.0)
Interface: ExtTagATM13
IF status: up                    IFC state: ACTIVE
Min VPI: 0                        Maximum cell rate: 10000
Max VPI: 10                       Available channels: xxx
Min VCI: 0                        Available cell rate (forward): xxxxxxx
Max VCI: 65535                   Available cell rate (backward): xxxxxxx

Phys desc: 2.2
Log intf: 0x00040300 (0.4.3.0)
Interface: ExtTagATM22
IF status: up                    IFC state: ACTIVE
Min VPI: 0                        Maximum cell rate: 10000
Max VPI: 10                       Available channels: xxx
Min VCI: 0                        Available cell rate (forward): xxxxxxx
Max VCI: 65535                   Available cell rate (backward): xxxxxxx
-----

```

**Step 2** If there are no interfaces present, first check that card 1 is up, by using this command on the BPX switch:

```
dspcds
```

and, if the card is not up, in this example BXM in slot 1 of the BPX shelf:

```
resetcd 1 h
```

and/or remove the card to get it to reset if necessary.



**Note** This example assumes that the controller is connected to card 1 on the switch. Substitute a different card number, as applicable.

**Step 3** Check the trunk status with the following command:

```
dsptrks
```

The **dsptrks** screen for ATM-LSR-1 should show the 1.1, 1.3, and 2.2 MPLS interfaces, with the “Other End” of 1.1 reading “VSI (VSI)”. Here’s a typical **dsptrks** screen:

```
n4          TN      SuperUser      BPX 15      9.3.10      August 4 2000 16:45 PST

TRK   Type   Current Line Alarm Status      Other End
2.1   OC3     Clear - OK                               j4a/2.1
3.1   E3      Clear - OK                               j6c (AXIS)
5.1   E3      Clear - OK                               j6a/5.2
5.2   E3      Clear - OK                               j3b/3
5.3   E3      Clear - OK                               j5c (IPX/AF)
6.1   T3      Clear - OK                               j4a/4.1
6.2   T3      Clear - OK                               j3b/4
1.1   OC3     Clear - OK                               VSI (VSI)
1.3   OC3     Clear - OK
2.2   OC3     Clear - OK
```

Last Command: dsptrks

Next Command:

#### Step 4 Enter the **dsptime** command.

**dsptime**

The resulting screens should show trunk 1.1 (link to LSC on ATM-LSR-1) as type VSI. Here’s typical **dsptime** screen:

```
n4          TN      SuperUser      BPX 15      9.3.10      August 4 2000 16:46 PST

                                BPX Interface Shelf Information

Trunk   Name      Type      Alarm
3.1     j6c      AXIS      MIN
5.3     j5c      IPX/AF    MIN
1.1     VSI      VSI       OK
```

Last Command: dsptime

Next Command:



**Step 5** Enter the **dsprsrc** command:

```
dsprsrc 1.1 1
```

The resulting screen should show these settings:

```
n4          TN      SuperUser      BPX 15      9.3.10      August 4 2000      16:47 PST
```

```
Port/Trunk : 1.1
```

```
Maximum PVC LCNS:          256      Maximum PVC Bandwidth:105000
```

```
Min Lcn(1) : 0 Min Lcn(2) : 0
```

```
Partition 1
```

```
Partition State :          Enabled
```

```
Minimum VSI LCNS:          512
```

```
Maximum VSI LCNS:          1500
```

```
Start VSI VPI:             240
```

```
End VSI VPI :              255
```

```
Minimum VSI Bandwidth :    26000      Maximum VSI Bandwidth :          105000
```

```
Last Command: dsprsrc 1.1 1
```

```
Next Command:
```

**Step 6** Enter the **dspqbin** command:

```
dspqbin 1.1 10
```

The resulting screen should show these settings:

```
n4          TN      SuperUser      BPX 15      9.3.10      August 4 2000      16:48 PST
```

```
Qbin Database 1.1 on EXM qbin 10
```

```
Qbin State:                Enabled
```

```
Minimum Bandwidth:         0
```

```
Qbin Discard threshold:    65536
```

```
Low CLP threshold:         95%
```

```
High CLP threshold:        100%
```

```
EFCI threshold:            40%
```

```
Last Command: dspqbin 1.1 10
```

```
Next Command:
```

**Step 7** If interfaces 1.3 and 2.2 are present, but not enabled, perform the previous debugging steps for interfaces 1.3 and 2.2 instead of 1.1, except for the **dspsnode** command, which does not show anything useful pertaining to ports 1.3 and 2.2.

- Step 8** Try a ping on the label switch connections. If the ping doesn't work, but all the label switching and routing configuration looks correct, check that the LSC has found the VSI interfaces correctly by entering the following command at the LSC:

Command	Description
Router LSC1# show tag int	Shows the label interfaces.

If the interfaces are not shown, recheck the configuration of port 1.1 on the BPX switch as described in the previous steps.

- Step 9** If the VSI interfaces are shown, but are down, check whether the LSRs connected to the BPX switch show that the lines are up. If not, check such items as cabling and connections.
- Step 10** If the LSCs and BPX switches show the interfaces are up, but the LSC doesn't show this, enter the following command on the LSC:

```
Router LSC1# reload
```

If the “show tag int” command shows that the interfaces are up, but the ping doesn't work, enter the follow command at the LSC:

```
Router LSC1# tag tdp disc
```

The resulting display should show something similar to this:

```
Local TDP Identifier:
 30.30.30.30:0
TDP Discovery Sources:
  Interfaces:
    ExtTagATM1.3:  xmit/recv
    ExtTagATM2.2:  xmit/recv
-----
```

- Step 11** If the interfaces on the display show “xmit” and not “xmit/recv”, then the LSC is sending LDP messages, but not getting responses. Enter this command on the neighboring LSRs.

```
Router LSC1# tag tdp disc
```

If resulting displays also show “xmit” and not “xmit/recv”, then one of two things is likely:

- The LSC is not able to set up VSI connections
  - The LSC is able to set up VSI connections, but cells are not transferred because they cannot get into a queue
- Step 12** Check the VSI configuration on the switch again, for interfaces 1.1, 1.3, and 2.2, paying particular attention to:
- maximum bandwidths are at least a few thousands cells per second
  - Qbins are enabled
  - all Qbin thresholds are non-zero



**Note** VSI partitioning and resources must be set up correctly on the interface connected to the LSC, interface 1.1 in this example, as well as interfaces connected to other label switching devices.

# Basic Router Configuration

This section provides basic configuration information for the Cisco 6400, 7200, or 7500 routers used as the Label Switch Controller for the BPX 8650:

- Accessing the Router Command-Line Interface
- Booting the Router for the First Time
- Configuring the Router for the First Time

## Accessing the Router Command-Line Interface

To configure a router, you must access its command line interface (CLI).

If you will be configuring the router on-site, connect a console terminal (an ASCII terminal or a PC running terminal emulation software) to the console port on the router.

For remote access, connect a modem to the auxiliary port on the router.

## Booting the Router for the First Time

Each time you turn on power to the router, it goes through the following boot sequence:

1. The router goes through power-on self-test diagnostics to verify basic operation of the CPU, memory, and interfaces.
2. The system bootstrap software (boot image) executes and searches for a valid Cisco IOS image. The factory-default setting for the configuration register is 0x2102, which indicates that the router should attempt to load a Cisco IOS image from Flash memory.
3. If after five attempts a valid Cisco IOS image is not found in Flash memory, the Cisco router reverts to boot ROM mode (which is used to install or upgrade a Cisco IOS image).
4. If a valid Cisco IOS image is found, then the Cisco router searches for a valid configuration file.
5. If a valid configuration file is not found in NVRAM, the Cisco router runs the System Configuration Dialog so you can configure it manually. For normal router operation, there must be a valid Cisco IOS image in Flash memory and a configuration file in NVRAM.

The first time you boot the router, you need to configure the router interfaces and then save the configuration to a file in NVRAM. Proceed to the next section, “Configuring the Router for the First Time,” for configuration instructions.

## Configuring the Router for the First Time

You can configure the Cisco router by using one of the procedures described in this section:

- Using the System Configuration Dialog  
Recommended if you are not familiar with Cisco IOS commands.
- Using Configuration Mode  
Recommended if you are familiar with Cisco IOS commands.

- Using Auto Install  
Recommended for automatic installation if another router running Cisco IOS software is installed on the network. This configuration method must be set up by someone with experience using Cisco IOS software.

**Timesaver**


---

Obtain the correct network addresses from your system administrator or consult your network plan to determine correct addresses before you begin to configure the router.

---

Use the procedure that best meets the needs of your network configuration and level of experience using Cisco IOS software.

If you use configuration mode or AutoInstall to configure the router and you would like a quick review of the Cisco IOS software, refer to the section “Cisco IOS Software Basics” later in this chapter. Otherwise, proceed to the next section, “Using the System Configuration Dialog.”

## Using the System Configuration Dialog

If your router does not have a configuration (setup) file and you are not using AutoInstall, the router will automatically start the setup command facility. An interactive dialog called the System Configuration Dialog appears on the console screen. This dialog helps you navigate through the configuration process by prompting you for the configuration information necessary for the Cisco router to operate.

**Note**


---

Many prompts in the System Configuration Dialog include default answers, which are included in square brackets following the question. To accept a default answer, press **Return**; otherwise, enter your response.

---

This section gives a sample configuration using the System Configuration Dialog. When you are configuring your router, respond as appropriate for your network.

At any time during the System Configuration Dialog, you can request help by entering a question mark (?) at a prompt.

Before proceeding with the System Configuration Dialog, obtain from your system administrator:

- the node addresses
- the number of bits in the subnet field (if applicable) of the Ethernet and synchronous serial ports

To configure the router by using the System Configuration Dialog:

- Step 1** Connect a console terminal or modem to the router and power ON the router.
- Step 2** Wait about 30 seconds for messages to be displayed, corresponding to the Cisco IOS release and feature set you selected. The screen displays in this section are for reference only and might not exactly reflect the screen displays on your console:

```
System Bootstrap, Version 11.1(13)CA, EARLY DEPLOYMENT RELEASE SOFTWARE (fc1)
BOOTFLASH: 7200 Software (C7200-BOOT-M), Version 11.1(24)CC, EARLY DEPLOYMENT RELEASE
SOFTWARE (fc1)
```

```
cisco 7206 (NPE200) processor with 122880K/8192K bytes of memory.
R5000 CPU at 200Mhz, Implementation 35, Rev 2.1, 512KB L2 Cache
6 slot midplane, Version 1.3
```

#### Restricted Rights Legend

Use, duplication, or disclosure by the Government is subject to restrictions as set forth in subparagraph (c) of the Commercial Computer Software - Restricted Rights clause at FAR sec. 52.227-19 and subparagraph (c) (1) (ii) of the Rights in Technical Data and Computer Software clause at DFARS sec. 252.227-7013.

Cisco Systems, Inc.  
170 West Tasman Drive  
San Jose, California 95134-1706

```
Cisco Internetwork Operating System Software
IOS (tm) 7200 Software (C7200-P-M), Version 12.0(5.0.2)T2, MAINTENANCE INTERIM SOFTWARE
Copyright (c) 1986-1999 by cisco Systems, Inc.
Compiled Sun 11-Jul-99 08:26 by kpma
Image text-base: 0x60008900, data-base: 0x60D64000
```

```
4 Ethernet/IEEE 802.3 interface(s)
1 FastEthernet/IEEE 802.3 interface(s)
1 ATM network interface(s)
125K bytes of non-volatile configuration memory.
4096K bytes of packet SRAM memory.
```

```
20480K bytes of Flash PCMCIA card at slot 0 (Sector size 128K).
107520K bytes of ATA PCMCIA card at slot 1 (Sector size 512 bytes).
4096K bytes of Flash internal SIMM (Sector size 256K).
Configuration register is 0x102
```

--- System Configuration Dialog ---

At any point you may enter a question mark '?' for help.  
Use ctrl-c to abort configuration dialog at any prompt.  
Default settings are in square brackets '['].  
Would you like to enter the initial configuration dialog? [yes]:

- Step 3** Press **Return** or enter **yes** to begin the configuration process.

**Step 4** When the System Configuration Dialog asks whether you want to view the current interface summary, press **Return** or enter **yes**:

First, would you like to see the current interface summary? **yes**

Any interface listed with OK? value "NO" does not have a valid configuration

Interface	IP-Address	OK?	Method	Status	Protocol
Ethernet0	unassigned	NO	unset	up	down
Serial0	unassigned	NO	unset	down	down
TokenRing0	unassigned	NO	unset	reset	down
ATM 0	unassigned	NO	unset	reset	down

**Step 5** Configure the global parameters. Here's a typical configuration:

Enter host name [7200router]: **aries**

**Step 6** Next, you are prompted to enter an enable secret password. There are two types of privileged-level passwords:

- Enable secret password (a secure, encrypted password).
- Enable password (a less secure, nonencrypted password).

The enable password is used when the enable secret password does not exist. For maximum security, be sure the passwords are different. If you enter the same password for both, the Cisco router will accept your entry, but will display a warning message indicating that you should enter a different password.

**Step 7** Enter an enable secret password:

The enable secret is a one-way cryptographic secret used instead of the enable password when it exists.

Enter enable secret: **orca**

The enable password is used when there is no enable secret and when using older software and some boot images.

**Step 8** Enter the enable and virtual terminal passwords:

Enter enable password: **xxxx**

Enter virtual terminal password: **yyyy**

**Step 9** Press **Return** to accept Simple Network Management Protocol management, or enter **no** to refuse it:

Configure SNMP Network Management? [yes]: **no**

**Step 10** In this example, the Cisco router is configured for AppleTalk, IP, MPLS, and Internetwork Packet Exchange. Configure the appropriate protocols for your router:

```
Configure Vines? [no]:
Configure LAT? [no]:
Configure AppleTalk? [no]:
Multizone networks? [no]: yes
Configure DECnet? [no]:
Configure IP? [yes]:
Configure MPLS? [no]: yes
Configure IGRP routing? [yes]: no
Your IGRP autonomous system number [1]: 15
Configure CLNS? [no]:
Configure bridging? [no]:
Configure IPX? [no]:
Configure XNS? [no]:
Configure Apollo? [no]:
```

**Note**

It is recommended that an MPLS network use either OSPF or IS-IS routing as its routing protocol. EIGRP will also work, but it does not support the useful MPLS feature Traffic Engineering. IGRP and RIP protocols are not recommended.

## Configuring Port Adapter Interfaces

Here is an overview of the procedure:

1. Make port adapter cable connections and complete basic configuration on the router.
2. Configure the applicable port adapter interfaces on the router, Ethernet, Fast Ethernet, ATM, FDDI, and so on.
3. Configure the router for MPLS operation.
4. Add permanent virtual circuits (PVCs) as applicable.

## Preparing to Configure Port Adapter Interfaces

If you want to configure interfaces in a new Cisco 6400, 7200, or 7500 series router, or if you want to change the configuration of an existing interface, be prepared with the information you will need:

- Protocols you plan to route on each new interface.
- Internet protocol (IP) addresses if you plan to configure the interfaces for IP routing.
- The types of interfaces that will be used.

The **configure** command requires privileged-level access to the EXEC command interpreter, which usually requires a password. Contact your system administrator if necessary to obtain EXEC-level access.

## Identifying Chassis Slot, Port Adapter Slot, and Interface Port Numbers

You will need to identify chassis slot, port adapter slot, and interface port numbers on the 6400, 7200, or 7500 Series routers for all port adapter interface types.

Physical port addresses specify the actual physical location of each interface port, regardless of the type.

You can identify port adapter interface ports by physically checking the slot/interface port location on the 7200 or 7500 Series routers, or by using the **show** commands to display information about a specific interface or all interfaces.

## Configuring ATM Interfaces

This section provides the procedure for a basic interface configuration.

Press the **Return** key after each step unless otherwise noted. At any time you can exit the privileged level and return to the user level by entering **disable** at the prompt as follows:

```
Cisco 7200 Router# disable
```

```
Cisco 7200 Router>
```

To perform a basic configuration:

- 
- Step 1** At the privileged-level prompt, enter configuration mode and specify that the console terminal will be the source of the configuration subcommands:

```
Cisco 7200 Router# configure terminal  
Enter configuration commands, one per line. End with CNTL/Z.  
Cisco 7200 Router (config)#
```

- Step 2** At the prompt, enter:

- a. the subcommand **interface** to specify the interface to be configured
- b. then **atm** to specify port adapter type
- c. then *slot/port* (port adapter slot number and interface port number)

This example is the 1/0 interface of the ATM port adapter in a 7200 series router:

```
Cisco 7200 Router (config)# interface switch atm 1/0
```

- Step 3** If IP routing is enabled on the system, you can assign an IP address and subnet mask to the interface with the **ip address** configuration subcommand:

```
Cisco 7200 Router (config-if)# ip address 224.135.128.44 255.255.255.0
```

- Step 4** Add any additional configuration subcommands required to enable routing protocols and set the interface characteristics.

- Step 5** Change the shutdown state to UP and enable the interface:

```
Cisco 7200 Router (config-if)# no shutdown
```

- Step 6** Repeat Step 2 through Step 5 to configure additional interfaces as required.

- Step 7** When you have completed the configuration, press **Ctrl-Z** to exit configuration mode.

- Step 8** Write the new configuration to nonvolatile memory:

```
Cisco Router 7200# copy running-config startup-config  
[OK]  
Cisco Router 7200#
```




---

**Note** If you are going to unattach/reconfigure the ATM interface cable, use the **shutdown** command prior to this action. After re-attaching the ATM interface cable, use the **no shutdown** command to bring the ATM interface into an up state.

---



## Other Router Interfaces

The router has other interfaces for carrying IP traffic. Refer to the Cisco 7200 or 7500 series router documentation, as applicable.

# Checking the Configuration

After configuring the new interface, use:

- the **show** commands to display the status of the new interface or all interfaces
- the **ping** command to check connectivity

## Using Show Commands to Verify the New Interface Status

This procedure uses **show** commands to verify that the new interfaces are configured and operating correctly:

- 
- Step 1** Use the **show version** command to display the system hardware configuration. Ensure that the list includes the new interfaces.
  - Step 2** Display all the current port adapters and their interfaces by using the **show controllers** command. Verify that the new port adapter appears in the correct slot.
  - Step 3** Specify one of the new interfaces by using the **show interfaces port adapter type slot/interface** command.
    - a. Verify that the first line of the display specifies the interface with the correct slot number.
    - b. Verify that the interface and line protocol are in the correct state: up or down.
  - Step 4** Display the protocols configured for the entire system and specific interfaces with the **show protocols** command. If necessary, return to configuration mode to add or remove protocol routing on the system or specific interfaces.
  - Step 5** Display the running configuration file with the **show running-config** command.
  - Step 6** Display the configuration stored in NVRAM using the **show startup-config** command.
  - Step 7** Verify that the configuration is accurate for the system and each interface.

If the interface is down and you configured it as up, or if the displays indicate that the hardware is not functioning properly, ensure that the network interface is properly connected and terminated. If you still have problems bringing the interface up, contact a service representative for assistance.

---

## Using Show Commands to Display Interface Information

To display information about a specific interface, use the **show interfaces** command with the interface type and port address in the format **show interfaces [type slot/port]** for the Cisco router.

## Cisco Show Interfaces Command

Here is an example of how the **show interfaces** [*type slot/port*] command displays status information (including the physical slot and port address) for the interfaces you specify. (Interfaces are administratively shut down until you enable them.)

```
Cisco 7200 Router 3# sh int e 2/0
Ethernet2/0 is administratively down, line protocol is down
  Hardware is AmdP2 Ethernet, address is x.x.x.x (bia 0000.0ca5.2389)
  MTU 1500 bytes, BW 10000 Kbit, DLY 1000 usec, rely 255/255, load 1/255
  Encapsulation ARPA, loopback not set, keepalive set (10 sec)
(display text omitted)
```

When running the **show interfaces** *type slot/port* command, use arguments such as the interface type (Ethernet, and so on), slot, and the port number (slot/port) to display information about a specific Ethernet 10BASE-T interface only.

The **show version** (or **show hardware**) command displays the configuration of the system hardware (the number of each port adapter type installed), the software version, the names and sources of configuration files, and the boot images. Here's an example of the **show version** command:

```
7200 router 1>show version
Cisco Internetwork Operating System Software
IOS (tm) 7200 Software (C7200-P-M), Version 12.0(5.0.2)T2,  MAINTENANCE INTERIM SOFTWARE
Copyright (c) 1986-1999 by cisco Systems, Inc.
Compiled Sun 11-Jul-99 08:26 by kpma
Image text-base: 0x60008900, data-base: 0x60D64000

ROM: System Bootstrap, Version 11.1(13)CA, EARLY DEPLOYMENT RELEASE SOFTWARE (fc1)
BOOTFLASH: 7200 Software (C7200-BOOT-M), Version 11.1(24)CC, EARLY DEPLOYMENT RELEASE
SOFTWARE (fc1)

7200 router 1 uptime is 2 weeks, 2 hours, 38 minutes

System returned to ROM by reload
System image file is "tftp://173.xx.xx.xx/c7200-p-mz.120-5.0.2.T2"

cisco 7206 (NPE200) processor with 122880K/8192K bytes of memory.
R5000 CPU at 200Mhz, Implementation 35, Rev 2.1, 512KB L2 Cache
6 slot midplane, Version 1.3

Last reset from power-on
X.25 software, Version 3.0.0.
4 Ethernet/IEEE 802.3 interface(s)
1 FastEthernet/IEEE 802.3 interface(s)
1 ATM network interface(s)
125K bytes of non-volatile configuration memory.
4096K bytes of packet SRAM memory.

20480K bytes of Flash PCMCIA card at slot 0 (Sector size 128K).
107520K bytes of ATA PCMCIA card at slot 1 (Sector size 512 bytes).
4096K bytes of Flash internal SIMM (Sector size 256K).
Configuration register is 0x102

7200 router 1>
```

To determine which type of port adapter is installed in your system, use the **show diag** command. Specific port adapter information is displayed:

```
7200 router 1>show diag
Slot 0:
Fast-ethernet on C7200 I/O card with MII or RJ45 port adapter, 1 port
Port adapter is analyzed
Port adapter insertion time 2w0d ago
EEPROM contents at hardware discovery:
Hardware revision 1.3          Board revision C0
Serial number 12635836       Part number 73-2956-02
Test history 0x0             RMA number 00-00-00
EEPROM format version 1
EEPROM contents (hex):
  0x20: 01 83 01 03 00 C0 CE BC 49 0B 8C 02 00 00 00 00
  0x30: 60 00 00 00 99 05 10 00 00 FF FF FF FF FF FF FF

Slot 3:
Ethernet port adapter, 4 ports
Port adapter is analyzed
Port adapter insertion time 2w0d ago
EEPROM contents at hardware discovery:
Hardware revision 1.14       Board revision A0
Serial number 12275103      Part number 73-1556-08
Test history 0x0            RMA number 00-00-00
EEPROM format version 1
EEPROM contents (hex):
  0x20: 01 02 01 0E 00 BB 4D 9F 49 06 14 08 00 00 00 00
  0x30: 50 00 00 00 99 03 30 00 FF FF FF FF FF FF FF FF

Slot 6:
ATM WAN DS3 port adapter, 1 port
Port adapter is analyzed
Port adapter insertion time 2w0d ago
EEPROM contents at hardware discovery:
Hardware revision 2.0        Board revision A0
Serial number 14077539      Part number 73-2432-04
Test history 0x0            RMA number 00-00-00
EEPROM format version 1
EEPROM contents (hex):
  0x20: 01 5B 02 00 00 D6 CE 63 49 09 80 04 00 00 00 00
  0x30: 50 00 00 00 99 04 26 00 FF FF FF FF FF FF FF FF

7200 router 1>
```

Proceed to the “Using the ping Command” section to verify that each interface port is functioning properly.

## Using the ping Command

The *packet internet groper* (**ping**) command allows you to verify that an interface port is functioning properly and to check the path between a specific port and connected devices at various locations on the network. After you verify that the system has booted successfully and is operational, you can use **ping** to verify the status of interface ports.

The **ping** command sends an echo request out to a remote device at an IP address that you specify. After sending a series of signals, the command waits a specified time for the remote device to echo the signals. Each returned signal is displayed as an exclamation point (!) on the console terminal; each

signal that is not returned before the specified time-out is displayed as a period (.). A series of exclamation points (!!!!) indicates a good connection; a series of periods (.....) or the messages [timed out] or [failed] indicate that the connection failed.

Here is a successful **ping** command to a remote server with the address 1.1.1.10:

```
Cisco 7200 Router # ping 1.1.1.10 <Return>
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echoes to 1.1.1.10, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/15/64 ms
Cisco 7200 Router #
```

If the connection fails, verify that you have the correct IP address for the server and that the server is active (powered on), then repeat the **ping** command.

## Using Configuration Mode

You can configure the 7200 router manually if you prefer not to use AutoInstall or the prompt-driven System Configuration Dialog.

Refer to the section “Cisco IOS Software Basics” later in this chapter for basic information about Cisco IOS software, getting context-sensitive help, and saving configuration changes.

To configure the Cisco 7200 router manually:

- 
- Step 1** Connect a console terminal.
  - Step 2** Power ON the Cisco 7200 router.
  - Step 3** When you are prompted to enter the initial dialog, enter **no** to go into the normal operating mode of the Cisco 7200 router:

```
Would you like to enter the initial dialog? [yes]: no
```

- Step 4** After a few seconds you will see the user EXEC prompt (Router>). By default, the host name is Router, but the prompt will match the current host name. In the following examples, the host name is **aries**.
- Step 5** Enter the **enable** command to enter enable mode. You can make only configuration changes in enable mode:

```
Router > enable
```

The prompt will change to the privileged EXEC (enable) prompt:

```
7200 Router aries#
```

- Step 6** Enter the **configure terminal** command at the enable prompt to enter configuration mode:

```
Router# config terminal
```

You can now enter any changes you want to the configuration. You will probably want to perform these tasks:

- a. Assign a host name for the Cisco 7200 router by using the **hostname** command.
- b. Enter an enable secret password by using the **enable password** command.
- c. Assign addresses to the interfaces by using the **protocol address** command.
- d. Specify which protocols to support on the interfaces.

Refer to the Cisco IOS configuration guide and command reference publications for more information about the commands you can use to configure the 7200 or 7500 series routers.

**Step 7** When you finish configuring the router, enter the **exit** command until you return to the privileged EXEC prompt (7200 router aries#).

**Step 8** To save the configuration changes to NVRAM, enter the **copy running-config startup-config** command at the privileged EXEC prompt:

```
7200 router aries# copy running-config startup-config
*****
```

The Cisco router is now configured and will boot with the configuration you entered.

---

## Cisco IOS Software Basics

The section provides basic information about the Cisco IOS software:

- Cisco IOS Modes of Operation
- Getting Context-Sensitive Help

## Cisco IOS Modes of Operation

Cisco IOS software provides access to several different command modes. Each command mode provides a different group of related commands.

For security purposes, Cisco IOS software provides two levels of access to commands:

- **user**  
The unprivileged user mode is called user EXEC mode.
- **privileged**  
The privileged mode is called privileged EXEC mode and requires a password.

The commands available in user EXEC mode are a subset of the commands available in privileged EXEC mode.

Table 5-1 describes some of the most commonly used modes, how to enter the modes, and the resulting prompts. The prompt helps you identify which mode you are in and, therefore, which commands are available to you.

Table 5-1 Cisco IOS Operating Modes

Mode of Operation	Usage	How to Enter the Mode	Prompt
User EXEC	User EXEC commands allow you to connect to remote devices, change terminal settings on a temporary basis, perform basic tests, and list system information. The EXEC commands available at the user level are a subset of those available at the privileged level.	Log in.	7200 Router
Privileged EXEC	Privileged EXEC commands set operating parameters. The privileged command set includes those commands contained in user EXEC mode, and also the <b>configure</b> command through which you can access the remaining command modes. Privileged EXEC mode also includes high-level testing commands, such as <b>debug</b> .	From user EXEC mode, enter the <b>enable</b> EXEC command.	7200 Router#
Global configuration	Global configuration commands apply to features that affect the system as a whole.	From global configuration mode, enter the <b>configure</b> privileged EXEC command.	7200 Router#(config)#
Interface configuration	Interface configuration commands modify the operation of an interface such as an ATM, Ethernet, or serial port. Many features are enabled on a per-interface basis. Interface configuration commands always follow an interface global configuration command, which defines the interface type.	From global configuration mode, enter the <b>interface type number</b> command. For example, enter the <b>interface serial 0</b> command to configure the serial 0 interface.	7200 Router#(config-if)#
ROM monitor	ROM monitor commands are used to perform low-level diagnostics. You can also use the ROM monitor commands to recover from a system failure and stop the boot process in a specific operating environment.	From privileged EXEC mode, enter the <b>reload</b> EXEC command. Press Break during the first 60 seconds while the system is booting.	>

Almost every configuration command also has a **no** form. In general, use the **no** form to disable a feature or function. Use the command without the keyword **no** to reenable a disabled feature or to enable a feature that is disabled by default.

For example, IP routing is enabled by default. To disable IP routing, enter the **no ip routing** command and enter **ip routing** to reenable it. The Cisco IOS software command reference publication provides the complete syntax for the configuration commands and describes what the **no** form of a command does.

## Getting Context-Sensitive Help

In any command mode, you can get a list of available commands by entering a question mark (?).

```
7200 Router> ?
```

To obtain a list of commands that begin with a particular character sequence, type in those characters followed immediately by the question mark (?). Do not include a space. This form of help is called word help, because it completes a word for you.

```
7200 Router# co?
configure connect copy
```

To list keywords or arguments, enter a question mark in place of a keyword or argument. Include a space before the question mark. This form of help is called command syntax help, because it reminds you which keywords or arguments are applicable based on the command, keywords, and arguments you have already entered.

```
7200 Router# configure ?
memory      Configure from NV memory
network     Configure from a TFTP network host
terminal    Configure from the terminal
<cr>
```

You can also abbreviate commands and keywords by entering just enough characters to make the command unique from other commands. For example, you can abbreviate the **show** command to **sh**.

## Saving Configuration Changes

Whenever you make changes to the Cisco 7200 router configuration, you must save the changes to memory so they will not be lost if there is a system reload or power outage.

There are two types of configuration files:

- **The running (current operating) configuration**  
The running configuration is stored in RAM.
- **The startup configuration**  
The startup configuration is stored in NVRAM.

To display the current running configuration, enter the **show running-config** command.

To save the current running configuration to the startup configuration file in NVRAM, enter the **copy running-config startup-config** command:

```
7200 Router> enable
7200 Router# copy running-config startup-config
```

To display the startup configuration, enter the **show startup-config** command.

To write the startup configuration to the running configuration, enter the **copy startup-config running-config** command.

```
7200 Router> enable
7200 Router# copy startup-config running-config
```







## MPLS CoS with BPX 8650

---

This chapter describes MPLS Class of Service (CoS) with the use of the BPX 8650 ATM Label Switch Router (ATM LSR). A summary example is provided for configuring BPX 8650 ATM LSRs, their associated LSCs (6400, 7200, or 7500 series), and Edge Label Switch Routers:

- MPLS CoS Overview
- MPLS CoS in An IP+ATM Network
- ATM CoS Service Templates and Qbins on the BPX 8650
- MPLS CoS over IP+ATM Operation
- Configuration Example

For an overview of design issues, see Chapter 3, “Quality of Service in MPLS Networks.”

For additional information, refer to Cisco 6400, 7200, or 7500 series router and MPLS-related IOS documentation. Refer to release notes for supported features.

### MPLS CoS Overview

The MPLS CoS feature enables network administrators to provide differentiated types of service across an MPLS Switching network. Differentiated service satisfies a range of requirements by supplying the particular kind of service specified for each packet by its CoS. Service can be specified in different ways—for example, through use of the IP precedence bit settings in IP packets or in source and destination addresses.

The MPLS CoS feature can be used optionally with MPLS Virtual Private Networks. MPLS CoS can also be used in any MPLS switching network.

In supplying differentiated service, MPLS CoS offers packet classification, congestion avoidance, and congestion management. Table 6-1 lists these functions and the means by which they are delivered.

**Table 6-1 CoS Services and Features**

Service	CoS Function	Description
Packet classification	Committed access rate (CAR). Packets are classified at the edge of the network before labels are assigned.	CAR uses the type of service (TOS) bits in the IP header to classify packets according to input and output transmission rates. CAR is often configured on interfaces at the edge of a network in order to control traffic into or out of the network. You can use CAR classification commands to classify or reclassify a packet.
Congestion avoidance	Weighted random early detection (WRED). Packet classes are differentiated based on drop probability.	WRED monitors network traffic, trying to anticipate and prevent congestion at common network and internetwork bottlenecks. WRED can selectively discard lower priority traffic when an interface begins to get congested. It can also provide differentiated performance characteristics for different Classes of Service.
Congestion management	Weighted fair queueing (WFQ). Packet classes are differentiated based on bandwidth and bounded delay.	WFQ is an automated scheduling system that provides fair bandwidth allocation to all network traffic. WFQ classifies traffic into conversations and uses weights (priorities) to determine how much bandwidth each conversation is allocated, relative to other conversations.

MPLS CoS lets you duplicate Cisco IOS IP CoS (Layer 3) features as closely as possible in MPLS switching devices, including Label Switching Routers (LSRs), Edge LSRs, and ATM label switching routers (ATM LSRs). MPLS CoS functions map nearly one-for-one to IP CoS functions on all interface types.

## Related Documents

For more information on configuration of the CoS functions (CAR, WRED, and WFQ), refer to the *Cisco IOS Class of Service for Tag Switching Feature Guide*, and the *Cisco IOS Quality of Service Solutions Configuration Guide*.

For complete command syntax information for CAR, WRED, and WFQ, refer to the *Cisco IOS Quality of Service Solutions Command Reference*.

For additional information on BPX 8650 CLI commands, refer to the *Cisco WAN Switch Command Reference*.

## Prerequisites

To use the MPLS CoS feature, your network must be running these Cisco IOS features:

- CEF switching in every MPLS-enabled router
- MPLS
- ATM functionality

Also, the BPX 8650 must have:

- the appropriate switch software associated with the Cisco IOS
- the appropriate firmware loaded in the associated BXM cards

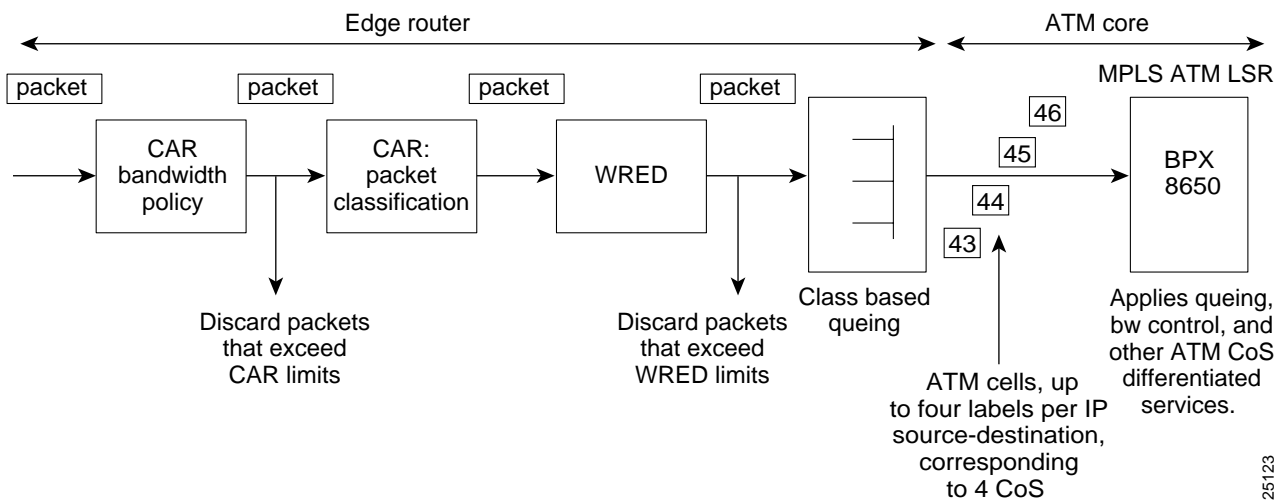
# MPLS CoS in An IP+ATM Network

In IP+ATM networks, MPLS uses predefined sets of labels for each service class, so switches automatically know which traffic requires priority queuing. A different label is used per destination to designate each service class (see Figure 6-1).

There can be up to four labels per IP source-destination. Using these labels, core LSRs implement class-based WFQ to allocate specific amounts of bandwidth and buffer to each service class. Cells are queued by class to implement latency guarantees.

On a Cisco IP+ATM LSR, the weights assigned to each service class are relative, not absolute. The switch can therefore borrow unused bandwidth from one class and allocate it to other classes according to weight. This scenario enables very efficient bandwidth utilization. The class-based WFQ solution ensures that customer traffic is sent whenever unused bandwidth is available, whereas ordinary ATM VCs drop cells in oversubscribed classes even when bandwidth is available.

Figure 6-1 Multiple LVCs for IP QoS Services



Packets have their precedence bits in the Type of Service field of the IP header, set at either the host or an intermediate router, which could be the Edge Label Switch Router (LSR). The precedence bits define a Class of Service (CoS) 0-3, corresponding for to premium, standard, available, or control, for example.

To establish CoS operation when the BPX and the associated LSC router (6400, 7200, or 7500 series) are initially configured, the binding type assigned each LVC interface on the BPX is configured to be multiple LVCs.

Then under the routing protocol (OSPF, for example), four LVCs are set up across the network for each IP source to destination requirement. Depending on the precedence bits set in the packets that are received by the Edge LSR, the packet ATM cells that are sent to the ATM LSR will be one four classes (as determined by the cell label, that is, VPI.VCI). Furthermore, two subclasses are distinguishable within each class by the use of the cell loss priority (CLP) bit in the cells.

Then the ATM LSR performs a MPLS data table look-up and assigns the appropriate template Class of Service template and Qbin characteristics. The default mapping for CoS is listed in Table 6-2.

Two IP Types of Service (ToS) are carried in each MPLS CoS. In the default mapping, both ToS 0 and ToS 4 are carried in the so-called “Available” VC. The two types are distinguished by the CLP bit. In the default mapping, cells of IP packets with IP ToS 0 are carried with CLP=1, whereas cells of IP packets with IP ToS 4 are carried with CLP=0. This means that in times of congestion in the switches, packets with ToS 0 will be discarded, while ToS 4 packets will not be discarded unless the congestion becomes particularly severe.

**Table 6-2** *Type of Service and Related CoS*

Class of Service Mapping	Class of Service	IP ToS
Available (see note)	0	ToS 0/4
Standard	1	ToS 1/5
Premium	2	ToS 2/6
Control	3	ToS 3/7



**Note**

The name "Available" means that, in a typical configuration (but not the default configuration), this Class of Service will have a fairly low guaranteed bandwidth; hence it will mostly use bandwidth that has been left available when traffic of other classes is not currently using its full allocation of bandwidth. "Available" does not mean Available Bit Rate (ABR). ABR is not currently used for MPLS CoS on the BPX 8650.

Figure 6-2 shows an example of IP traffic across an ATM core consisting of BPX 8650 ATM LSRs. The host is sending two types of traffic across the network, interactive video and non-time-critical data. Because multiple LVCs have automatically been generated for all IP source-destination paths, traffic for each source destination is assigned to one of four LVCs, based on the precedence bit setting in the IP packet header.

In this case, the video traffic might be assigned to the premium CoS, and transmitted across the network starting with the cell label “51” out of the Edge LSR-A, and continuing across the network with the cell label “91” applied to the Edge LSR-C. In each BPX 8650 ATM LSR, the cells are processed with the pre-assigned bandwidth, queuing, and other ATM QoS functions suitable to “premium” traffic.

In a similar fashion, low-priority data traffic cells with the same IP source-destination might be assigned label “53 out of Edge LSR-A and arrive at Edge LSR-C with the label “93”, receiving pre-assigned bandwidth, queuing and other ATM QoS functions suitable to “available” traffic.



When ATM cells arrive from the Edge LSR at the BXM port with one of four CoS labels, they receive CoS handling based on that label. A table look-up is performed, and the cells are processed, based on their connection classification. Based on its label, a cell receives the ATM differentiated service associated with its template type, (MPLS1 template) and service type (for example, label cos2 bw), plus associated Qbin characteristics and other associated ATM parameters.

## Initial Setup of LVCs

The service template contains two classes of data:

- **Connection Parameters**  
These parameters are necessary to establish a connection (that is, per LVC) and includes entries such as UPC actions, various bandwidth related items, per LVC thresholds, and so on.
- **CoS Configuration**  
These data items are required to configure the associated Class of Service buffers (Qbins) that provide CoS support.

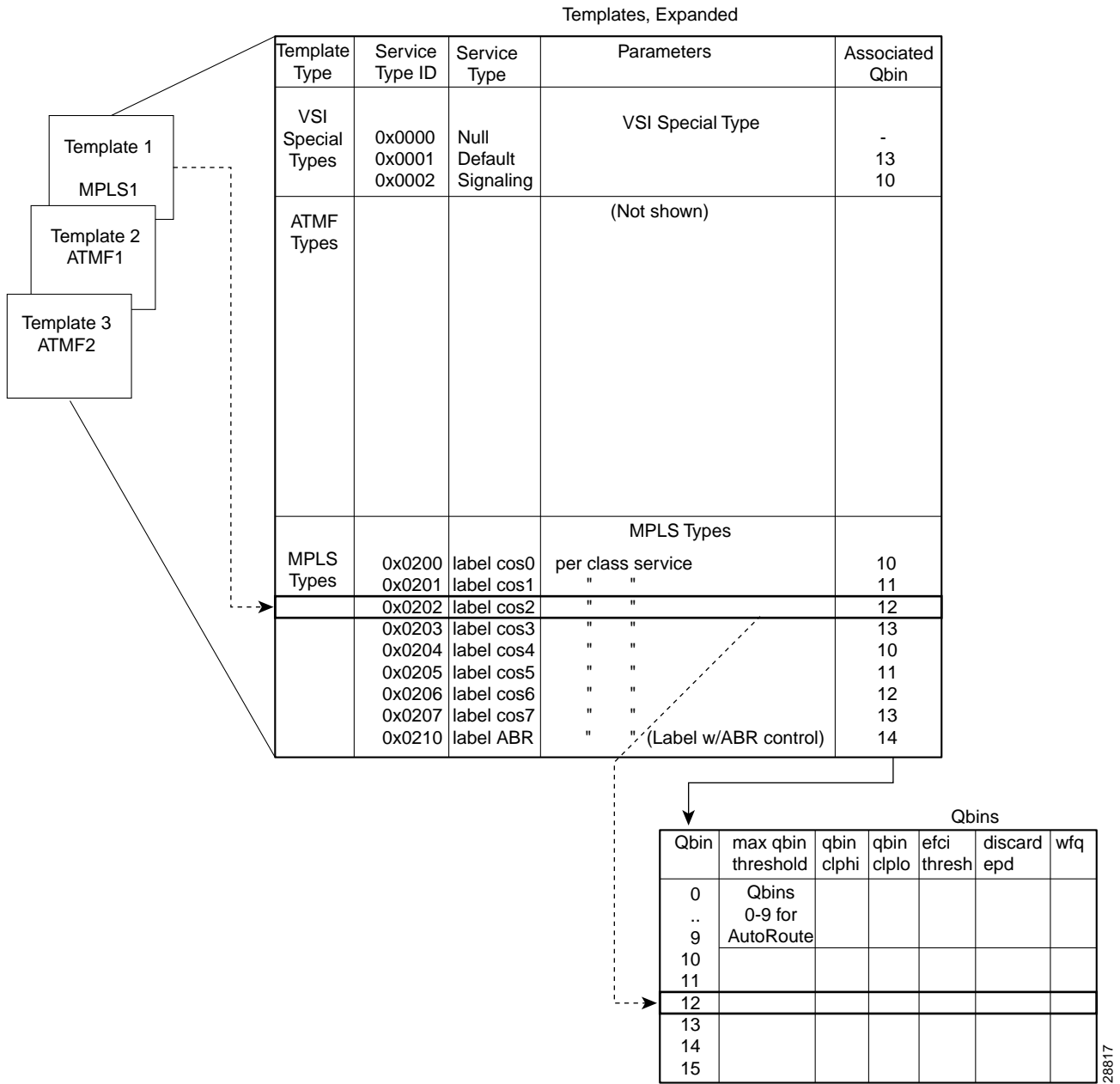
When a connection setup request is received from the VSI master in the Label Switch Controller, the VSI slave (in the BXM, for example) uses the service type identifier to index into a Service Class Template database (Figure 6-3) containing extended parameter settings for connections matching that index. The slave uses these values to complete the connection setup and program the hardware.

## Service Template Qbins

When you use the **upport** or **uptrk** command to activate an interface on the BXM card, the default service template, which is MPLS1, is assigned to the interface (Figure 6-3). Each template table row includes an entry that defines the Qbin to be used for that Class of Service. This mapping defines a relationship between the template and the interface Qbin's configuration.

Qbin templates are used only with Qbins that are available to VSI partitions, namely Qbins 10 through 15. Qbins 10 through 15 are used by the VSI on interfaces configured as trunks or ports. The rest of the Qbins (0-9) are reserved for and configured by AutoRoute.

Figure 6-3 Service Template and Associated Qbin Selection



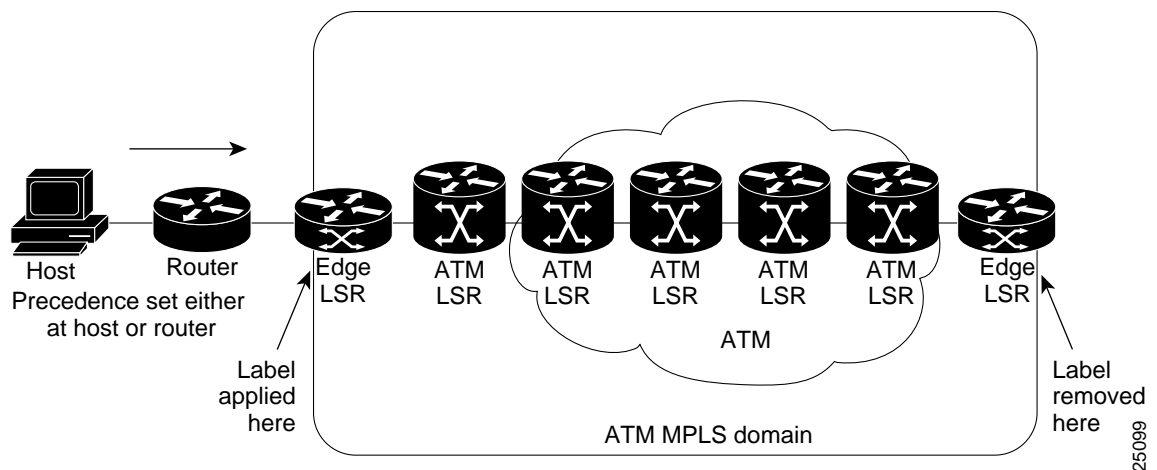
## MPLS CoS over IP+ATM Operation

In a typical operation for MPLS CoS, a packet makes its way from the host on the left side of a network, through the network of conventional routers, label edge routers (LERs), Edge LSRs, and ATM LSRs such as a BPX 8650.

As the packet progresses, basic functions are applied to it, as shown in Figure 6-4:

1. Set the IP Type of Service (ToS) for a packet in the host (or router).
2. In the Edge LSR, label the packet by putting it on a label-VC. There is a choice of up to four label-VCs to each IP destination-prefix, for different CoS. Choose one based on the IP ToS in the packet.
3. Apply ATM CoS bandwidth and queuing to ATM cells based on their Class of Service in the ATM LSR (BPX 8650, for example).
4. At the Edge LSR, receive the packet from the label VC, discard the label information, and forward the IP packet with appropriate ToS towards its destination (Edge LSR).

Figure 6-4 MPLS CoS over IP+ ATM with BPX 8650 LSRs



### Note

In the figure, the functions are shown being performed by separate entities. In general, one or more functions can be performed by the same entity. For example, the setting of precedence and labeling could all be performed in a single label edge router if the host were not participating.

The preceding discussion applies to MPLS networks where the entire network runs ATM MPLS. MPLS CoS also works in networks using a mixture of ATM MPLS and packet-based MPLS. For more information, see the MPLS chapter in the *Cisco IOS Switching Services Configuration Guide*.



## Configuration Example

There are four default policy types for MPLS CoS as shown in Table 6-3 with default relative bandwidth per **xtagatm** interface.

**Table 6-3 Class of Service and Relative Bandwidth Weighting**

Class of Service Mapping	Class of Service	IP ToS	Default Bandwidth Weight
Available	0	ToS 0/4	99
Standard	1	ToS 1/5	0
Premium	2	ToS 2/6	0
Control	3	ToS 3/7	1

The relative bandwidth weights determine the proportion of bandwidth available to MPLS, which is given to each Class of Service on each link. If a CoS does not use the bandwidth given to it, then the bandwidth may be shared among the other CoSs.

The Control CoS is important to guarantee a good quality of service for MPLS control traffic. For this reason, it is desirable to reserve a small amount of bandwidth for the Control CoS as shown in Table 6-4.

**Table 6-4 Class of Service and Relative Bandwidth Weighting Setup**

Class of Service Mapping	Class of Service	IP ToS	Bandwidth Weight
Available	0	ToS 0/4	49
Standard	1	ToS 1/5	50
Premium	2	ToS 2/6	0
Control	3	ToS 3/7	1

To verify an **xtagatm** interface after configuration on the LSC, run this command:

```
show xtagatm cos-bandwidth-allocation xtagatmxx
```

where *xx* is the interface number. The maximum value for CoS bandwidth is 100.

The setup for the configuration example is shown in Figure 6-5.



- zero for premium (2/6)
- zero for control (3/7) Class of Service.

Once xtagatm interface has been defined for each LSC, execute the command:

```
show xtagatm cos-bandwidth-allocation xtagatmxx
```

where xx is interface number. Verify that default relative bandwidth is properly assigned in percentage value. The maximum value for CoS bandwidth is 100.

## LSC Configurations

### LSC1

```
LSC1l-1#config t
LSC1(config)#int atm1/0//LSC1LSC1 control port
LSC1(config-if)#no shut
LSC1(config-if)#tag-control-protocol vsi
LSC1(config-if)#exit

LSC1(config)#int xtagatm12//LSR1 port 1.2
LSC1(config-if)#extended-port atm1/0 bpx 1.2
LSC1(config-if)#tag-switching ip
LSC1(config-if)#tag-switching atm cos available 49
LSC1(config-if)#tag-switching atm cos standard 50
LSC1(config-if)#tag-switching atm cos premium 0
LSC1(config-if)#tag-switching atm cos control 1
LSC1(config-if)#ip unnumbered loopback0
LSC1(config-if)#exit

LSC1(config)#int xtagatm13//LSR1 port 1.3
LSC1(config-if)#extended-port atm1/0 bpx 1.3
LSC1(config-if)#tag-switching ip
LSC1(config-if)#tag-switching atm cos available 49
LSC1(config-if)#tag-switching atm cos standard 50
LSC1(config-if)#tag-switching atm cos premium 0
LSC1(config-if)#tag-switching atm cos control 1
LSC1(config-if)#ip unnumbered loopback0
LSC1(config-if)#exit

LSC1(config)#int loopback0//configure loopback0 interface
LSC1(config-if)#ip address 200.200.200.1 255.255.255.255
LSC1(config-if)#exit

LSC1(config)#ip routing//enable IP routing
LSC1(config)#ip cef//enable Cisco Express Forwarding Protocol
LSC1(config)#router ospf 10
LSC1(config-router)#network 200.200.200.1 0.0.0.0 area 0
LSC1(config-router)#end
```

### LSC2

```
LSC2#config t
LSC2(config)#int atm2/0//LSC2 control port
LSC2(config-if)#no shut
LSC2(config-if)#tag-control-protocol vsi id 2
LSC2(config-if)#exit

LSC2(config)#int xtagatm22//LSR2 port 2.2
```

```

LSC2(config-if)#extended-port atml/0 bpx 2.2
LSC2(config-if)#tag-switching ip
LSC2(config-if)#tag-switching atm cos available 49
LSC2(config-if)#tag-switching atm cos standard 50
LSC2(config-if)#tag-switching atm cos premium 0
LSC2(config-if)#tag-switching atm cos control 1
LSC2(config-if)#ip unnumbered loopback0
LSC2(config-if)#exit

LSC2(config)#int xtagatm23//LSR2 port 2.3
LSC2(config-if)#extended-port atml/0 bpx 2.3
LSC2(config-if)#tag-switching ip
LSC2(config-if)#tag-switching atm cos available 49
LSC1(config-if)#tag-switching atm cos standard 50
LSC1(config-if)#tag-switching atm cos premium 0
LSC1(config-if)#tag-switching atm cos control 1
LSC2(config-if)#ip unnumbered loopback0
LSC2(config-if)#exit

LSC2(config)#int loopback0//configure loopback0 interface
LSC2(config-if)#ip address 200.200.200.2 255.255.255.255
LSC2(config-if)#exit

LSC2(config)#ip routing//enable IP routing
LSC2(config)#ip cef//enable Cisco Express Forwarding Protocol
LSC2(config)#router ospf 10
LSC2(config-router)#network 200.200.200.2 0.0.0.0 area 0
LSC2(config-router)#end

```

## Edge LSR Configurations

### LSR1

```

LSR1LSR1#config t
LSR1(config)#int atml/0//LSR1 interface
LSR1(config-if)#no shut
LSR1(config-if)#exit
LSR1(config)#interface atml/0.1 tag-switching//create tag sub-interface
LSR1(config-subif)#ip unnumbered loopback0
LSR1(config-subif)#tag-switching atm multi-vc//enable multi-vc mode (4 VCs)
LSR1(config-subif)#tag-switching ip

LSR1(config)#int loopback0//configure loopback0 interface
LSR1(config-if)#ip address 200.200.100.1 255.255.255.255

LSR1(config)#ip routing//enable IP routing
LSR1(config)#ip cef//enable Cisco Express Forwarding Protocol
LSR1(config)#router ospf 10
LSR1(config-router)#network 200.200.100.1 0.0.0.0 area 0
LSR1(config-router)#exit

```

In default multiple LVC mode, there are four MPLS Cos LVCs created by cos-map with clp set to off. The four classes of service are available (0/4), standard (1/5), premium (2/6), and control (3/7).

## LSR2

```
LSR2#config t
LSR2LSR2(config)#int atm2/0//LSR2 interface
LSR2(config-if)#no shut
LSR2(config-if)#exit
LSR2(config)#interface atm2/0.1 tag-switching//create tag sub-interface
LSR2(config-if)#ip unnumbered loopback0
LSR2(config-if)#tag-switching ip

LSR2(config)#int loopback0//configure loopback0 interface
LSR2(config-if)#ip address 200.200.100.2 255.255.255.255

LSR2(config)#ip routing//enable IP routing
LSR2(config)#ip cef//enable Cisco Express Forwarding Protocol
LSR2(config)#router ospf 10
LSR2(config-router)#network 200.200.100.2 0.0.0.0 area 0
LSR2(config-router)#end

LSR2(config)#tag-switching cos-map 1//configure Cos-Map
LSR2(config-tag-cos-map)#end//for now use default 4 VCs
LSR2#sho tag-switching cos-map//there should be 4 VCs w/ clp off
LSR2#config t
LSR2(config)#access-list 1 permit 200.200.100.1 0.0.0.0 //create access list for
network 200.200.100.1
LSR2(config)#tag-switching prefix-map 1 access-list 1 cos-map 1//map access-list to
cos-map 1
LSR2(config)#show tag forward 200.200.100.1 32 detail//verify forwarding table
```

Verify that the LSC/LSR is operational and BPXs have clear alarms.

LSR1 should be able to ping to LSR2 successfully.

Check that VSI resources have been allocated that the and controller was added successfully.

BPXs should have clear alarms and no software log and trunk errors.

## BPX1/BPX1

```
dsprtrks//successful with no alarms
dsvpsipartinfo //verify lcns and bandwidth are allocated successfully
dsplns//no alarm
dspctrlrs//controller ID is added successfully
```

Check that LSC/Edge LSR interfaces are operational and TDP bindings are successful.

## LSC1 and LSC2

```
LSC1#sho tag interface //xtagatm interfaces are operational
LSC1#sho xtag cross-connect//verify crosss-connect
LSC1#sho xtag vc//verify tag vc
LSC1#sho control vsi descriptor//verify VSI VPI range and Bw
LSC1#sho control vsi control-interface//verify number of connections for each
cross-connect
LSC1#sho control vsi traffic//verify traffic statistics
LSC1#sho tag atm bind //verify tag atm bindings
LSC1#sho tag atm sum//verify local/remote connections
```

## LSR1 and LSR2

```
LSR1#sho tag interface //xtagatm interfaces are operational
LSC2#sho tag tdp disc//verify tdp session rx/tx
LSC2#sho atm vc//verify atm pvc and tvc
```

**Note**

---

MPLS CoS Multiple LVC mode lets you reconfigure the classes for different traffic configurations. You have the flexibility to modify the four LVCs for any CoS. For example, you have the choice of assigning a “high” weight to a low class (that is, available CoS =60 and control CoS = 20).

---



## MPLS VPNS with BPX 8650

---

This chapter describes MPLS VPNs with the use of the BPX 8650 ATM Label Switch Router (LSR) and provides an example of the configuration of IOS to support VPNs:

- Introduction: MPLS-Enabled VPNs
- MPLS VPNs over IP+ATM Backbones
- Configuration Example
  - Configuring the BPX 8650 ATM LSR
  - Configuring VRFs
  - Configuring BGPs
  - Configuring Import and Export Routes
  - Verifying VPN Operation
- Command List

For an overview of Virtual Private Networks, including their features and benefits, see Chapter 1, “Introduction to MPLS.”

### Related Documents

- *MPLS VPNs Feature Module*
- *Cisco IOS Network Protocols Command Reference, Part 1*

## Introduction: MPLS-Enabled VPNs

Service providers can use MPLS to build an entirely new class of IP VPNs. MPLS-enabled IP VPNs are connectionless networks with the same privacy as VPNs built using Frame Relay or ATM VCs.

Cisco MPLS solutions offer multiple IP service classes to enforce business-based policies. Providers can offer low-cost managed IP services because they can consolidate services over common infrastructure and make provisioning and network operations much more efficient.

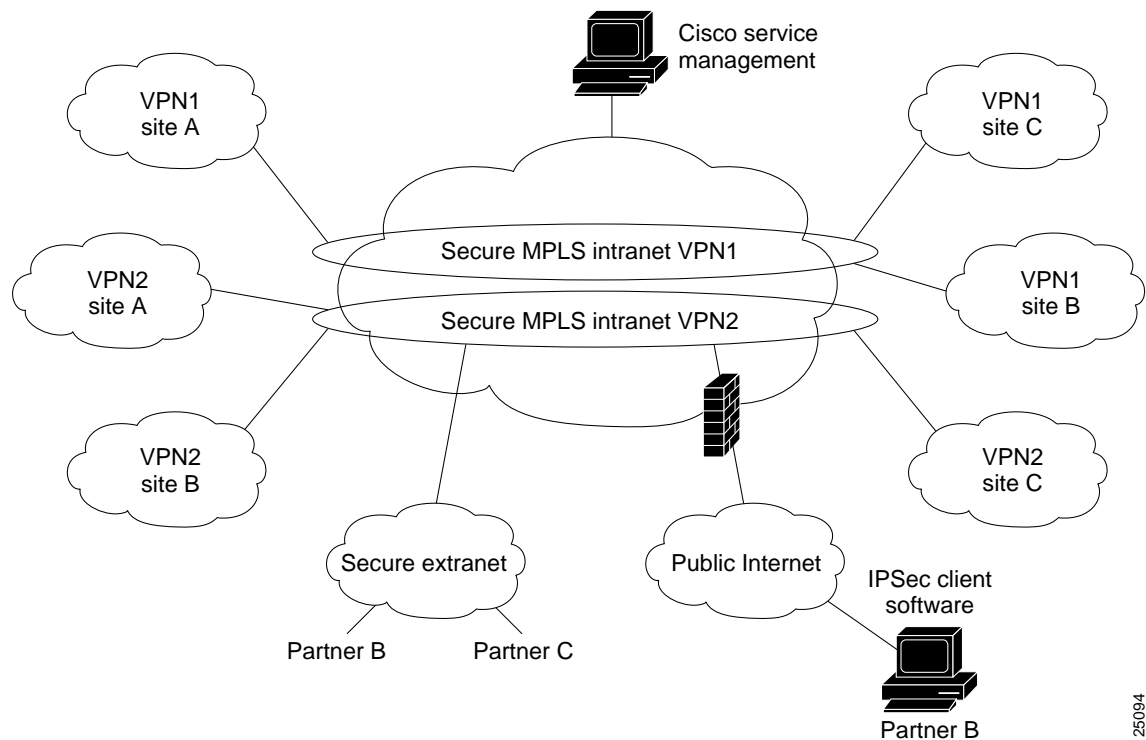
Although Frame Relay and multiservice ATM deliver privacy and Class of Service, IP delivers any-to-any connectivity, and MPLS on Cisco IP+ATM switches, such as the BPX 8650 ATM LSR, enables providers to offer the benefits of business-quality IP services over their ATM infrastructures.

MPLS VPNs, created in Layer 3, are connectionless, and therefore substantially more scalable and easier to build and manage than conventional VPNs.

In addition, value-added services, such as application and data hosting, network commerce, and telephony services, can easily be added to a particular MPLS VPN because the service provider's backbone recognizes each MPLS VPN as a separate, connectionless IP network. MPLS over IP+ATM VPN networks combine the scalability and flexibility of IP networks with the performance and QoS capabilities of ATM.

From a single access point, it is now possible to deploy multiple VPNs, each of which designates a different set of services (Figure 7-1). This flexible way of grouping users and services makes it possible to deliver new services more quickly and at a much lower cost. The ability to associate closed groups of users with specific services is critical to service provider value-added service strategies.

**Figure 7-1 VPN Network**



The VPN network must be able to “see” traffic by application type, such as voice, mission-critical applications, or e-mail, for example. The network should easily separate traffic based on its associated VPN without configuring complex, point-to-point meshes.

The network must be “VPN aware” so that the service provider can easily group users and services into intranets or extranets with the services they need. In such networks, VPNs offer service providers a technology that is highly scalable and allows subscribers to quickly and securely provision extranets to new partners. MPLS brings “VPN awareness” to switched or routed networks. It enables service providers to quickly and cost-effectively deploy secure VPNs of all sizes, all over the same infrastructure.

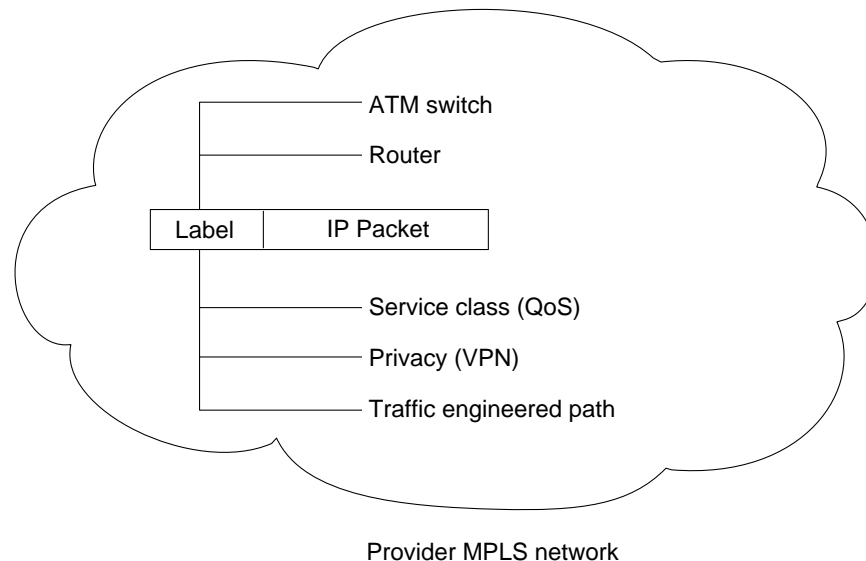
25094



## MPLS Labeling Criteria

For enabling business IP services, the most significant benefit of MPLS is the ability to assign labels that have special meanings. Sets of labels distinguish destination address as well as application type or service class, as discussed in the following sections (see Figure 7-2).

**Figure 7-2 Benefits of MPLS Labels**



The MPLS label is compared to precomputed switching tables in core devices, such as the BPX ATM LSR, allowing each switch to automatically apply the correct IP services to each packet. Tables are precalculated, so there is no need to reprocess packets at every hop. This strategy not only makes it possible to separate types of traffic, such as best-effort traffic from mission-critical traffic, it also makes an MPLS solution highly scalable.

Because MPLS uses different policy mechanisms to assign labels to packets, it decouples packet forwarding from the content of IP headers. Labels have local significance, and they are used many times in large networks. Therefore, it is nearly impossible to run out of labels. This characteristic is essential to implementing advanced IP services such as QoS, large-scale VPNs, and traffic engineering.

## Quality of Service

As part of their VPN services, service providers may wish to offer premium services defined by SLAs to expedite traffic from certain customers or applications. QoS in IP networks gives devices the intelligence to preferentially handle traffic as dictated by network policy.

The QoS mechanisms give network managers the ability to control the mix of bandwidth, delay, jitter, and packet loss in the network. QoS is not a device feature, it is an end-to-end system architecture. A robust QoS solution includes a variety of technologies that interoperate to deliver scalable, media-independent services throughout the network, with system-wide performance monitoring capabilities.

For a detailed discussion of QoS design issues, refer to Chapter 4, “Quality of Service in MPLS Networks.” VPNs are used with the Class of Service (CoS) feature for MPLS. MPLS-enabled IP VPN networks provide the foundation for delivering value-added IP services, such as multimedia application

support, packet voice, and application hosting, all of which require specific service quality and privacy. Because QoS and privacy are an integral part of MPLS, they no longer require separate network engineering.

Cisco's comprehensive set of QoS capabilities enable providers to prioritize service classes, allocate bandwidth, avoid congestion, and link Layer 2 and Layer 3 QoS mechanisms:

- **Committed Access Rate (CAR)**  
CAR classifies packets by application and protocol, and specifies bandwidth allocation.
- **Weighted Fair Queuing (WFQ) and Class-Based Queuing (CBQ)**  
Both of these techniques implement efficient bandwidth usage by always delivering mission-critical application traffic and deferring noncritical application traffic when necessary.
- **Weighted Random Early Detection (WRED)**  
WRED provides congestion avoidance to slow transmission rates before congestion occurs and ensures predictable service for mission-critical applications that require specific delivery guarantees.

MPLS makes it possible to apply scalable QoS across very large routed networks and Layer 3 IP QoS in ATM networks, because providers can designate sets of labels that correspond to service classes. In routed networks, MPLS-enabled QoS substantially reduces processing throughout the core for optimal performance. In ATM networks, MPLS makes end-to-end Layer 3-type services possible.

Traditional ATM and Frame Relay networks implement CoS with point-to-point virtual circuits, but this is not scalable because of high provisioning and management overhead. Placing traffic into service classes at the edge enables providers to engineer and manage classes throughout the network. If service providers manage networks based on service classes, rather than point-to-point connections, they can substantially reduce the amount of detail they must track and increase efficiency without losing functionality.

Compared to per-circuit management, MPLS-enabled CoS in ATM networks provides virtually all the benefits of point-to-point meshes with far less complexity. Using MPLS to establish IP CoS in ATM networks eliminates per-VC configuration. The entire network is easier to provision and engineer.

## Security

Subscribers want assurance that their VPNs are in fact private and that their applications and communications are isolated and secure. Cisco offers many robust security measures to keep information confidential:

- encrypted data
- access restricted to authorized users
- user tracking after they are connected to the network
- real-time intrusion auditing

In intranet and extranet VPNs based on Cisco MPLS, packets are forwarded using a unique route distinguisher (RD). RDs are unknown to end users and uniquely assigned automatically when the VPN is provisioned. To participate in a VPN, a user must be attached to its associated logical port and have the correct RD. The RD is placed in packet headers to isolate traffic to specific VPN communities.

MPLS packets are forwarded using labels attached in front of the IP header. Because the MPLS network does not read IP addresses in the packet header, it allows the same IP address space to be shared among different customers, simplifying IP address management.

Service providers can deliver fully managed MPLS-based VPNs with the same level of security that users are accustomed to in Frame Relay/ATM services, without the complex provisioning associated with manually establishing PVCs and performing per-VPN customer premises equipment (CPE) router configuration.

QoS addresses two fundamental requirements for applications that run on a VPN: predictable performance and policy implementation. Policies are used to assign resources to applications, project groups, or servers in a prioritized way. The increasing volume of network traffic, along with project-based requirements, results in the need for service providers to offer bandwidth control and to align their network policies with business policies in a dynamic, flexible way.

## Manageability

As service providers build VPNs that include WAN switches, routers, firewalls, and Cisco IOS software, they need to seamlessly manage these devices across the network infrastructure and provide service-level agreements to their customers. They also need to enable business customers to personalize their access to network services and applications.

The Cisco Service Management (CSM) System addresses these needs with a suite of service management solutions to enable service providers to effectively plan, provision, operate, and bill VPN services.

## Scalability

VPNs based on Cisco MPLS technology scale to support tens of thousands of business-quality VPNs over the same infrastructure. MPLS-based VPN services solve peer adjacency and scalability issues common to large virtual circuit (VC) and IP tunnel topologies. Complex permanent virtual circuit/switched virtual circuit (PVC/SVC) meshes are no longer needed, and providers can use new, sophisticated traffic engineering methods to select predetermined paths and deliver IP QoS to premium business applications and services.

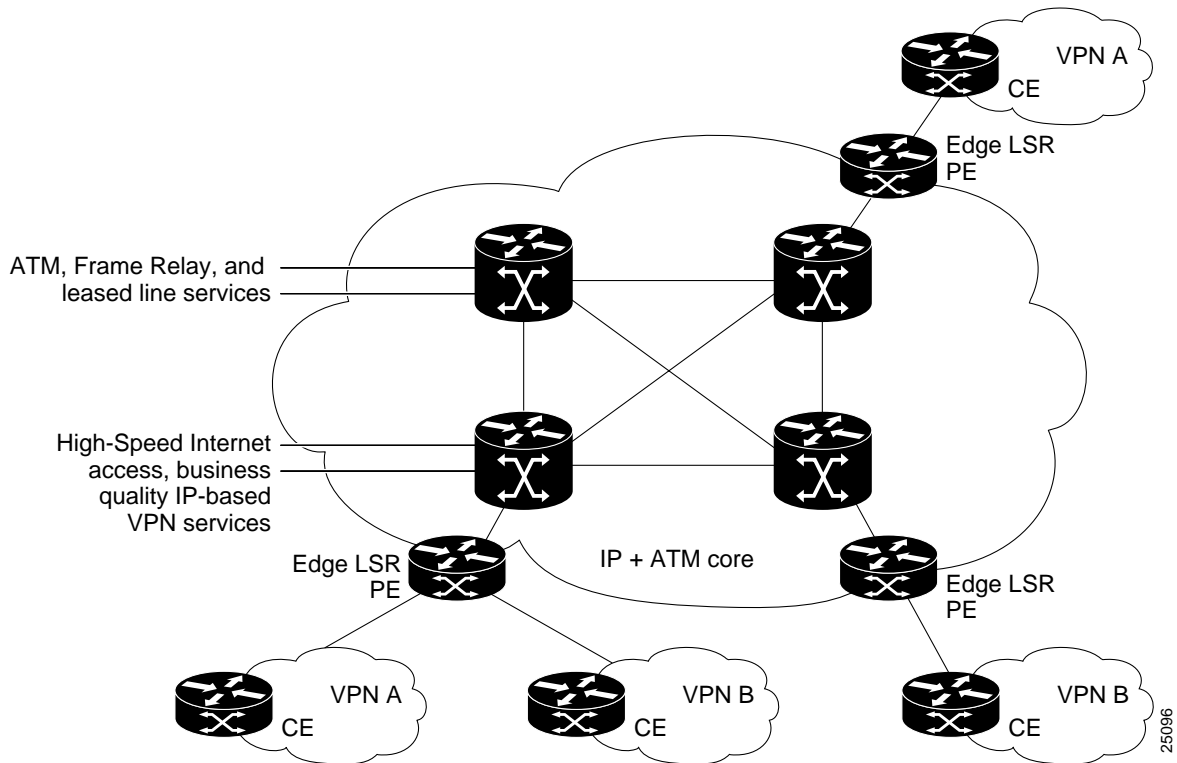
## MPLS VPNS over IP+ATM Backbones

Service providers can use MPLS to build intelligent IP VPNs across their existing ATM networks. Because all routing decisions are precomputed into switching tables, MPLS both expedites IP forwarding in large ATM networks at the provider edge and makes it possible to apply rich Layer 3 services via Cisco IOS technologies in Layer 2 cores.

A service provider with an existing ATM core can deploy MPLS-enabled edge switches or routers (LSRs) to enable the delivery of differentiated business IP services. The service provider needs only a small number of VCs to interconnect provider edge switches or routers to deliver extremely large numbers of secure VPNs.

Cisco IP+ATM solutions give ATM networks the ability to intelligently “see” IP application traffic as distinct from ATM/Frame Relay traffic. By harnessing the attributes of both IP and ATM, service providers can provision intranet or extranet VPNs. Cisco enables IP+ATM solutions with MPLS, uniting the application richness of Cisco IOS software with carrier-class ATM switches, as shown in Figure 7-3.

Figure 7-3 MPLS VPNS in Cisco IP+ATM Network



Without MPLS, IP transport over ATM networks require a complex hierarchy of translation protocols to map IP addresses and routing into ATM addressing and routing.

MPLS eliminates complexity by mapping IP addressing and routing information directly into ATM switching tables. The MPLS label-swapping paradigm is the same mechanism that ATM switches use to forward ATM cells. This solution has the added benefit of allowing service providers to continue to offer their current Frame Relay, leased-line, and ATM services portfolio while enabling them to offer differentiated business-quality IP services.

## Built-In VPN Visibility

To cost-effectively provision feature-rich IP VPNs, providers need features that distinguish between different types of application traffic and apply privacy and QoS—with far less complexity than an overlay IP tunnel, Frame Relay, or ATM “mesh.”

Compared to an overlay solution, an MPLS-enabled network can separate traffic and provide privacy without tunneling or encryption. MPLS-enabled networks provide privacy on a network-by-network basis, much as Frame Relay or ATM provides it on a connection-by-connection basis. The Frame Relay or ATM VPN offers basic transport, whereas an MPLS-enabled network supports scalable VPN services and IP-based value added applications. This approach is part of the shift in service provider business from a transport-oriented model to a service-focused one.

In MPLS-enabled VPNs, whether over an IP switched core or an ATM LSR switch core, the provider assigns each VPN a unique identifier called a route distinguisher (RD) that is different for each intranet or extranet within the provider network. Forwarding tables contain unique addresses, called VPN-IP

addresses (see Figure 7-4), constructed by concatenating the RD with the customer IP address. VPN-IP addresses are unique for each endpoint in the network, and entries are stored in forwarding tables for each node in the VPN.

**Figure 7-4 VPN-IP Address Format**

RD	IP Address/Mask Length	General format
0.1.0.99	130.101.0.0/16	VPN-IPv4 example

RD is a 64-bit route distinguisher

- Never carried on packets, only in Label tables

Each customer network can use:

- Registered IP addresses
- Unregistered addresses

Private addresses (RFC 1918, for example, 10.x.x.x)

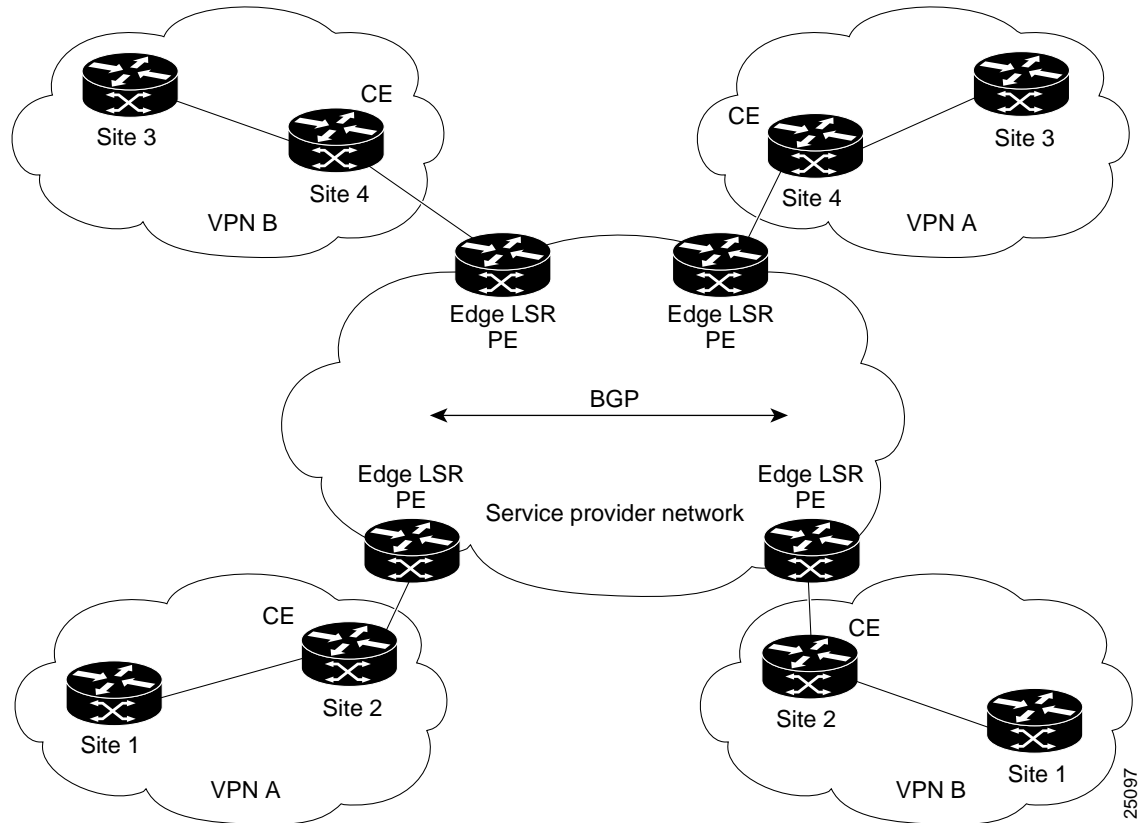
25100

## BGP Protocol

Border Gateway Protocol (BGP) is a routing information distribution protocol that defines who can talk to whom using multiprotocol extensions and community attributes. In an MPLS-enabled VPN, BGP distributes information about VPNs only to members of the same VPN, providing native security through traffic separation. Figure 7-5 shows an example of a service provider network with ATM backbone switches (P), service provider Edge Label Switch Routers (PE), and customer edge routers (CE).

Additional security is assured because all traffic is forwarded using LSPs, which define a specific path through the network that cannot be altered. This label-based paradigm is the same property that assures privacy in Frame Relay and ATM connections.

Figure 7-5 VPN with Service Provider Backbone



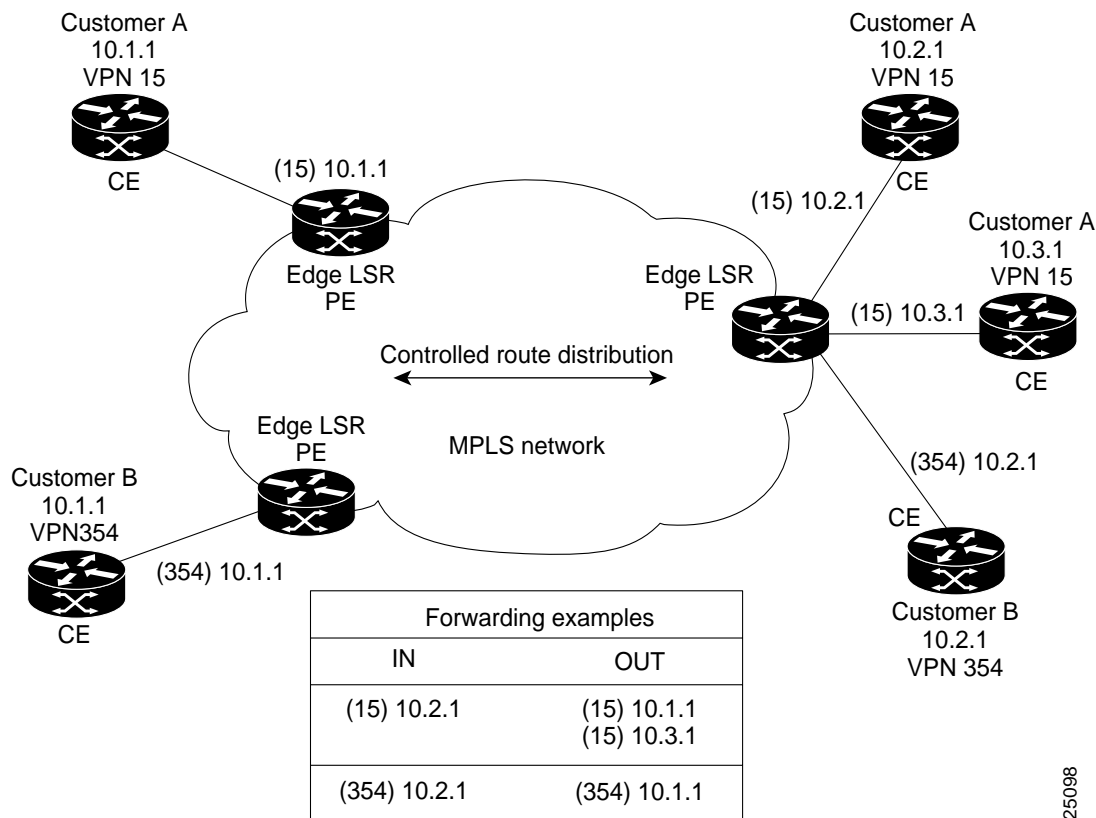
25097

The provider, not the customer, associates a specific VPN with each interface when the VPN is provisioned. Within the provider network, RDs are associated with every packet, so VPNs cannot be penetrated by attempting to “spoof” a flow or packet. Users can participate in an intranet or extranet only if they reside on the correct physical port and have the proper RD. This setup makes Cisco MPLS-enabled VPNs virtually impossible to enter, and provides the same security levels users are accustomed to in a Frame Relay, leased-line, or ATM service.

PN-IP forwarding tables contain labels that correspond to VPN-IP addresses. These labels route traffic to each site in a VPN (see Figure 7-6).

Because labels are used instead of IP addresses, customers can keep their private addressing schemes, within the corporate Internet, without requiring Network Address Translation (NAT) to pass traffic through the provider network. Traffic is separated between VPNs using a logically distinct forwarding table for each VPN. Based on the incoming interface, the switch selects a specific forwarding table, which lists only valid destinations in the VPN, as specified by BGP. To create Extranets, a provider explicitly configures reachability between VPNs. (NAT configurations may be required.)

Figure 7-6 Using MPLS to Build VPNs



25098

One strength of MPLS is that providers can use the same infrastructure to support many VPNs and do not need to build separate networks for each customer. VPNs loosely correspond to “subnets” of the provider network.

This solution builds IP VPN capabilities into the network itself, so providers can configure a single network for all subscribers that delivers private IP network services such as intranets and extranets without complex management, tunnels, or VC meshes. Application-aware QoS makes it possible to apply customer-specific business policies to each VPN. Adding QoS services to MPLS-based VPNs works seamlessly; the provider Edge LSR assigns correct priorities for each application within a VPN.

MPLS-enabled IP VPN networks are easier to integrate with IP-based customer networks. Subscribers can seamlessly interconnect with a provider service without changing their intranet applications, because these networks have application awareness built in, for privacy, QoS, and any-to-any networking. Customers can even transparently use their private IP addresses without NAT.

The same infrastructure can support many VPNs for many customers, removing the burden of separately engineering a new network for each customer, as with overlay VPNs.

It is also much easier to perform adds, moves, and changes. If a company wants to add a new site to a VPN, the service provider only has to tell the CPE router how to reach the network, and configure the LSR to recognize VPN membership of the CPE. BGP updates all VPN members automatically.

This scenario is far easier, faster, and less expensive than building a new point-to-point VC mesh for each new site. Adding a new site to an overlay VPN entails updating the traffic matrix, provisioning point-to-point VCs from the new site to all existing sites, updating OSPF design for every site, and reconfiguring each CPE for the new topology.

## Virtual Routing/Forwarding

Each VPN is associated with one or more VPN routing/forwarding instances (VRFs). A VRF table defines a VPN at a customer site attached to a PE router. A VRF table consists of:

- an IP routing table
- a derived Cisco Express Forwarding (CEF) table
- a set of interfaces that use the forwarding table
- a set of rules and routing protocol variables that determine what goes into the forwarding table

A 1-to-1 relationship does not necessarily exist between customer sites and VPNs. A given site can be a member of multiple VPNs. However, a site may be associated with one (and only one) VRF. A site's VRF contains all the routes available to the site from the VPNs of which it is a member.

Packet forwarding information is stored in the IP routing table and the CEF table for each VRF. (Together, these tables are analogous to the forwarding information base (FIB) used in Label Switching.)

A logically separate set of routing and CEF tables is constructed for each VRF. These tables prevent information from being forwarded outside a VPN and also prevent packets that are outside a VPN from being forwarded to a router within the VPN.

## VPN Route-Target Communities

The distribution of VPN routing information is controlled through the use of VPN route target communities, implemented by BGP extended communities.

Here is how distribution works:

When a VPN route is injected into BGP, it is associated with a list of VPN route target extended communities. Typically the list of VPN communities is set through an export list of extended community-distinguishers associated with the VRF from which the route was learned.

Associated with each VRF is an import list of route-target communities. This list defines the values to be verified by the VRF table before a route is eligible to be imported into the VPN routing instance.

For example, if the import list for a particular VRF includes community-distinguishers of A, B, and C, then any VPN route that carries any of those extended community-distinguishers—A, B, *or* C—will be imported into the VRF.

## IBGP Distribution of VPN Routing Information

A service provider edge (PE) router can learn an IP prefix from a customer edge (CE) router (by static configuration, through a Border Gateway Protocol (BGP) session with the CE router, or through the routing information protocol (RIP) with the CE router).

Once it learns the prefix, the router generates a VPN-IPv4 (vpngv4) prefix based on the IP prefix by linking an 8-byte route distinguisher to the IP prefix. This extended VPN-IPv4 address uniquely identifies hosts within each VPN site, even if the site is using globally nonunique (unregistered private) IP addresses.

The route distinguisher (RD) used to generate the VPN-IPv4 prefix is specified by a configuration command on the PE.



BGP uses VPN-IPv4 addresses to distribute network reachability information for each VPN within the service provider network. BGP distributes routing information between IP domains (known as autonomous systems) using messages to build and maintain routing tables. BGP communication takes place at two levels: within the domain (interior BGP or IBGP) and between domains (external BGP or EBGP).

BGP propagates VPNV4 information using the BGP multiprotocol extensions for handling these extended addresses. (See RFC 2283, *Multiprotocol Extensions for BGP-4*.) BGP propagates reachability information (expressed as VPN-IPv4 addresses) among PE routers; the reachability information for a given VPN is propagated only to other members of that VPN. The BGP multiprotocol extensions identify the valid recipients for VPN routing information. All the members of the VPN learn routes to other members.

## Label Forwarding

Based on the routing information stored in the IP routing table and the CEF table for each VRF, Cisco label switching uses extended VPN-IPv4 addresses to forward packets to their destinations.

An MPLS label is associated with each customer route. The PE router assigns the label that originated the route, and directs the data packets to the correct CE router.

Label forwarding across the provider backbone, is based on either dynamic IP paths or Traffic Engineered paths. A customer data packet has two levels of labels attached when it is forwarded across the backbone:

- the top label directs the packet to the correct PE router
- the second label indicates how that PE router should forward the packet

The PE router associates each CE router with a forwarding table that contains only the set of routes that should be available to that CE router.

## Configuration Example

Before configuring VPN operation, your network must be running the following Cisco IOS services:

- Label Switching connectivity with generic routing encapsulation (GRE) tunnels configured among all provider (PE) routers with VPN service, or label switching in all provider backbone (P) routers
- Label Switching with VPN code in all provider routers with a VPN edge service (PE) routers
- BGP in all routers providing a VPN service
- CEF switching in every label-enable router
- GRE
- CoS enabled on all routers
- 7000 series routers

Perform these tasks to configure and verify VPNs:

1. Configuring the BPX 8650 ATM LSR
2. Configuring VRFs
3. Configuring BGPs
4. Configuring Import and Export Routes
5. Verifying VPN Operation

## Configuring the BPX 8650 ATM LSR

For MPLS VPN operation, you must first configure the BPX 8650 ATM LSR, including its associated 6400, 7200, or 7500 LSC for MPLS or for MPLS QoS.

You configure network VPN operation on the Edge LSRs that act as PE routers.

The BPX 8650, including its LSC, requires no configuration beyond enabling MPLS and QoS.

## Configuring VRFs

To configure a VRF and associated interfaces, perform these steps on the PE router:

	Command	Purpose
Step 1	Router(config)# <b>ip vrf</b> <i>vrf-name</i>	Enter VRF configuration mode and specify the VRF name to which subsequent commands apply.
Step 2	Router(config-vrf)# <b>rd</b> <i>route-distinguisher</i>	Define the instance by assigning a name and an 8-byte route distinguisher.
Step 3	Router(config-if)# <b>ip vrf forwarding</b> <i>vrf-name</i>	Associate interfaces with the VRF.
Step 4	Router(config-router)# <b>address-family ipv4 vrf</b> <i>vrf-name</i>	Configure BGP parameters for the VRF CE session to use BGP between the PE and VRF CE.  The default setting is off for auto-summary and synchronization in the VRF address-family submode.  To ensure that addresses learned through BGJP on a PE router from a CE router are properly treated as VPN IPv4 addresses, you must enter the command <b>no bgp default ipv4-activate</b> before configuring and CE neighbors.
Step 5	Router(config-router)# <b>address-family ipv4 vrf</b> <i>vrf-name</i>	Configure RIP parameters for use between the PE and VRF CEs.
Step 6	Router(config-router-af)# <b>exit-address-family</b>	Exit from address-family configuration mode.
Step 7	Router(config)# <b>ip route</b> [ <i>vrf vrf-name</i> ]	Configure static routes for the VRF.

## Configuring BGPs

To configure a BGP between provider routes for distribution of VPN routing information, perform these steps on the PE router:

	Command	Purpose
Step 1	Router(config-router)# <b>address-family</b> { <i>ipv4 vpn4</i> } [ <i>unicast multicast</i> ]	Configure BGP address families.
Step 2	Router(config-router-af)# <b>neighbor</b> { <i>address peer-group</i> } <b>remote-as</b> <i>as-number</i>	Define a BGP session.
Step 3	Router(config-router)# <b>no bgp default ipv4-activate</b>	Activate a BGP session. Prevents automatic advertisement of address family IPv4 for all neighbors.
Step 4	Router(config-router)# <b>neighbor</b> <i>address</i> <b>remote-as</b> <i>as-number</i>	Configure an IBGP to exchange VPNv4 NLRI.
Step 5	Router(config-router)# <b>neighbor</b> <i>address</i> <b>update-source</b> <i>interface</i>	Define an IBGP session.
Step 6	Router(config-router-af)# <b>neighbor</b> <i>address</i> <b>activate</b>	Activate the advertisement of VPNv4 NLRI.

## Configuring Import and Export Routes

To configure import and export routes to control the distribution of routing information, perform these steps on the PE router:

	Command	Purpose
Step 1	Router(config)# <b>ip vrf</b> <i>vrf-name</i>	Enter VRF configuration mode and specify a VRF.
Step 2	Router(config-vrf)# <b>route-target import</b> <i>community-distinguisher</i>	Import routing information to the specified extended community.
Step 3	Router(config-vrf)# <b>route-target export</b> <i>community-distinguisher</i>	Export routing information to the specified extended community.
Step 4	Router(config-vrf)# <b>import map</b> <i>route-map</i>	Associate the specified route map with the VRF.

## Verifying VPN Operation

To verify VPN operation, perform these steps:

	Command	Purpose
Step 1	Router# <b>show ip vrf</b>	Display the set of defined VRFs and interfaces.
Step 2	Router# <b>show ip vrf detail</b>	Display VRF information including import and export community lists.
Step 3	Router# <b>show ip route vrf vrf-name</b>	Display the IP routing table for a VRF.
Step 4	Router# <b>show ip protocols vrf vrf-name</b>	Display the routing protocol information for a VRF.
Step 5	Router# <b>show ip cef vrf vrf-name</b>	Display the CEF forwarding table associated with a VRF.
Step 6	Router# <b>show ip interface interface-number</b>	Display the VRF table associated with an interface.
Step 7	Router# <b>show ip bgp vpnv4 all [tags]</b>	Display VPNv4 NLRI information.
Step 8	Router# <b>show tag-switching forwarding vrf vrf-name [prefix mask/length][detail]</b>	Display label forwarding entries that correspond to VRF routes advertised by this router.

## Configuration Example

Here is a sample configuration file from a PE router.

```

! CEF switching is a pre-requisite for Tag
ip cef distributed
frame-relay switching
!
! Define two VPN Routing instances, named 'vrf1' and 'vrf2'
ip vrf vrf1 rd 100:1
ip vrf vrf2 rd 100:2
!
! Configure the import and export VPN route-target list for each VRF
ip vrf vrf1 route-target both 100:1
ip vrf vrf2 route-target both 100:2
ip vrf vrf2 route-target import 100:1
! Configure an import route-map for vrf2
ip vrf vrf2 import map vrf2_import
! 'vrf2' should not install PE-CE addresses in the global routing table
no ip vrf vrf2 global-connected-addresses
!
interface lo0
 ip address 10.13.0.13 255.255.255.255
 no shut
! Backbone link to another Provider router
interface atm9/0/0
!
interface atm9/0/0.1 tag-switching
 tag-switching ip
ip unnumbered lo0
!
! Set up an Ethernet interface as a VRF link to a CE router
interface Ethernet5/0/1
 ip vrf forwarding vrf1
 ip address 10.20.0.13 255.255.255.0
!
! Set up a Frame-Relay PVC sub-interface a link to another CE router
interface hssi 10/1/0
 hssi internal-clock
 encaps fr
 frame-relay intf-type dce
 frame-relay lmi-type ansi
!
interface hssi 10/1/0.16 point-to-point
 ip vrf forwarding vrf2
 ip address 10.20.1.13 255.255.255.0
 frame-relay interface-dlci 16
!
! Configure BGP sessions
router bgp 1
! Define an IBGP session with another PE
 no bgp default ipv4-activate
 neighbor 10.15.0.15 remote-as 1
 neighbor 10.15.0.15 update-source lo0
 no synchronization
! Define some VRF (CE) sessions.
 neighbor 10.20.1.11 remote-as 65535
 neighbor 10.20.1.11 update-source h10/1/0.16
! Deactivate the default IPv4 session
 neighbor 10.20.0.60 remote-as 65535
 neighbor 10.20.0.60 update-source e5/0/1
!
! Activate PE peer for exchange of VPNv4 NLRI

```

```

address-family vpnv4 unicast
  neighbor 10.15.0.15 activate
  exit-address-family
!
! If exchange of IPv4 NLRI with 10.15.0.15 is desired, activate it:
address-family ipv4 unicast
  neighbor 10.15.0.15 activate
  exit-address-family
!
! Define BGP parameters for PE - CE sessions
! Activate sessions with peers in VRFs vrf1 and vrf2.
address-family ipv4 unicast vrf vrf1
  neighbor 10.20.0.60 activate
  no auto-summary
  redistribute static
  exit-address-family
!
address-family ipv4 unicast vrf vrf2
  neighbor 10.20.1.11 activate
  no auto-summary
  redistribute static
  exit-address-family
!
! Define a VRF static route
ip route vrf vrf1 12.0.0.0 255.0.0.0 e5/0/1 10.20.0.60

```

## Command List

This section lists new or modified commands. All other commands used with this feature are documented in the Cisco IOS command references, for Cisco IOS commands, and in the *Cisco WAN Switch Command Reference* for BPX 8650 CLI commands. For information on using the following commands, refer to the *Cisco MPLS VPN Feature Guide*.

- **address-family**
- **clear ip route vrf**
- **exit-address-family**
- **ip route vrf**
- **ip vrf forwarding**
- **ip vrf global-connected-addresses**
- **ip vrf**
- **neighbor activate**
- **show ip bgp vpnv4**
- **show ip cef vrf**
- **show ip protocols vrf**
- **show ip route vrf**
- **show ip vrf**
- **show tag-switching forwarding vrf**



## MPLS Redundancy for IP+ATM Networks

---

This chapter describes the Cisco Label Switch Controller (LSC) redundancy architecture that provides IP+ATM networks using MPLS with a level of reliability comparable to the hot-standby redundancy used in router networks but without the difficulty of implementing it. The hot LSC redundancy model provides the fastest reroute recovery time for IP+ATM networks.

This chapter covers these topics:

- What Is LSC Redundancy
- Benefits of LSC Redundancy
- LSC Redundancy Architecture
- LSC Hot Redundancy
- How the LSC, ATM Switch, and VSI Work Together
- Implementing LSC Redundancy
- Sample LSC Redundancy Configuration

### What Is LSC Redundancy

In traditional router IP networks, network managers ensure reliability by creating multiple paths through the network from every source to every destination. If a device or link on one path fails, IP traffic uses an alternate path to reach its destination.

Unlike router networks, circuit switch networks like ATM and Frame Relay transfer data by establishing circuits or virtual circuits. To ensure reliability, network managers incorporate redundant switch components: backup backplanes, power supplies, line cards, trunk cards, and so on.

But unlike router networks, switches take some time to reroute traffic when a failure occurs. Switch connection routing software, such as AutoRoute, PNNI, and MPLS, require calculating routes and reprogramming hardware for each connection. That's why router networks can reroute large aggregates of traffic more quickly than most connection-oriented networks.

Cisco's LSC redundancy recognizes that the LSC is the single point of failure for an IP+ATM network. Whether an LSC is an external router such as the Cisco 7204 router or an internal Routing Processor Module (RPM) in a BPX or MGX switch, an LSC is in the critical path for network reliability. If the LSC fails or if the LSC's port adapter goes down, the control function of the ATM LSR is disabled. The rest of the network can no longer trust that the ATM LSR has the correct MPLS label connections, and therefore will no longer use the links to the ATM LSR to carry MPLS traffic. Connectivity to some destinations in the network might be impossible unless there are alternative routes that avoid the failed ATM LSR.

Because label switch controllers are a critical component of IP+ATM networks, they must be robust and restore service quickly despite equipment or software failures.

Cisco's LSC redundancy is an alternative way to increase reliability in IP networks. This reliability is nearly equivalent to that provided with the use of hot-standby routing processes. But the result is in general terms the same: if the primary controller fails, traffic can be almost instantly routed by a secondary controller. In addition, the Cisco LSC redundancy architecture reroutes traffic much faster than conventional rerouting processes.

LSC redundancy basically consists of:

- Two controllers, such as two MPLS controllers
- The Virtual Switch Interface (VSI)
- Equal-cost multipath IP routing

In essence, two independent MPLS controllers, via VSI, control separate partitions in the IP+ATM switch, creating a set of two identical subnetworks. Multipath IP routing chooses to use both subnetworks equally, leading to identical connections in both subnetworks. If a controller in one subnetwork fails, then multipath IP routing very quickly diverts traffic to the other subnetwork.

LSC redundancy differs from hot-standby redundancy in that the LSCs do not need copies of each other's internal state or database, thus increasing reliability. LSC redundancy is simpler than hot-standby redundancy because it is not necessary to set up new connections when a controller fails. The LSC redundancy architecture requires the same amount of equipment as a network with hot-standby controllers, except that the controllers act independently, rather than in hot-standby mode.

## Benefits of LSC Redundancy

By implementing the LSC redundancy model, you eliminate the single point of failure between the LSC and the ATM switch it controls. If one LSC fails, the other LSC takes over and routes the data on the other path. The other benefits of LSC redundancy are now described in more detail.

## LSC Redundancy Allows Different Software Versions

The LSCs work independently; there is no interaction between the controllers. They do not share the controller's state or database, as other redundancy models require. Therefore, you can run different versions of the IOS software on the LSCs.

The advantage of this is that you can test the features of the latest version of software without risking reliability. You can run the latest version of the IOS software on one LSC and an older version of the IOS software on a different LSC. If the LSC running the new IOS software fails, the LSC running the older software takes over.



### Note

---

Using different IOS software version on different LSCs is *not* recommended except as a temporary measure. Different versions of IOS software in a network could be incompatible, although it is unlikely. For best results, run the same version of IOS software on all devices.

---



## LSC Redundancy Does Not Use Shared States or Databases

In the LSC redundancy model, the LSCs do not share states or databases, which increases reliability. Sometimes, when states and databases are shared, an error in the state or database information can cause both controllers to fail simultaneously.

Also, new software features and enhancements do not affect LSC redundancy. Because the LSCs do not share states or database information, you do not have to worry about ensuring redundancy during every step of the update.

## LSC Redundancy Lets You Use Different Hardware

You can use different models of routers in this LSC redundancy model. For example, one LSC can be a Cisco 7200 series router. The other LSC could be based in a Cisco 6400 edge switch series router. Using different hardware in the redundancy model reduces the chance that a hardware fault might interrupt network traffic.

## LSC Redundancy Provides An Easy Migration from Stand-alone LSCs to Redundant LSCs

You can migrate from a stand-alone LSC to a redundant LSC and back again without affecting network operations. Because the LSCs work independently, you can add a redundant LSC without interrupting the other LSC.

## LSC Redundancy Allows Configuration Changes in a Live Network

The hot LSC redundancy model provides two parallel, independent networks. Therefore, you can disable one LSC without affecting the other LSC. This feature has two main benefits:

- LSC redundancy model facilitates configuration changes and updates. After you finish with configuration changes or image upgrades to the LSC, you can add it back to the network and resume the LSC redundancy model.
- The redundancy model protects the network during partitioning of the ATM switch. You can disable one path and perform partitioning on that path. While you are performing the partitioning, data uses the other path. The network is safe from the effects of the partitioning, which include breaking and establishing LVC connections.

## LSC Redundancy Provides Fast Reroute in IP+ATM Networks

The hot LSC redundancy model offers redundant paths for every destination. Therefore, reroute recovery is very fast. Other rerouting processes in IP+ATM networks require many steps and take more time.

In normal IP+ATM networks, the reroute process consists of the following steps:

- Detect the failure
- Converge the Layer 2 routing protocols
- Complete label distribution for all destinations
- Establish new connections for all destinations

After this reroute process, the new path is ready to transfer data. However, rerouting data by using this process takes time.

The hot LSC redundancy method allows you to quickly reroute data in IP+ATM networks without using the normal reroute process. Hot LSC redundancy creates active parallel paths. Every destination has at least one alternative path. If a device or link along the path fails, the data uses the other path to reach its destination. The hot LSC redundancy model provides the fastest reroute recovery time for IP+ATM networks.

## LSC Redundancy Architecture

The architecture is distinguished by two main features:

1. Multiple controllers share the resources of the same switch, creating two independent IP networks
2. The resulting subnetworks are both linked at the Edge Label Switch Routers (LSR)

Consider a basic IP network of switches with one MPLS controller (or a hot-standby pair of them) and MPLS Edge Label Switch Routers (LSR) feeding the edge of the network.

The LSC redundancy architecture adds to this basic network two independent controllers of the same type (such as MPLS), enabled by the Virtual Switch Interface (VSI) to control two separate partitions on the same IP+ATM switch. The pair of controllers on the switch form two separate MPLS control planes for the network that effectively create two independent parallel IP subnetworks.

Provided that the two independent MPLS controllers on each switch have identical shares of the switch's resources and link capacity, the two subnetworks are identical. The two identical, parallel IP subnetworks exist on virtually the same equipment that would otherwise support only one IP network.



### Note

---

Each control plane and partition might have a redundant pair of controllers, but these are coupled. Note that the two independent controllers must be of the same type. Also, the equipment must have sufficient connection capacity for the doubled-up connections.

---

The LSC hot redundancy solution differs from hot-standby redundancy in that the MPLS controllers need not have copies of each other's internal state.

The second feature of the LSC redundancy architecture is the linkage of the two parallel subnetworks on the same physical ATM LSR at the edge.

This LSC redundancy network might use the Open Shortest Path First (OSPF) protocol with equal-cost multipath or a similar IP routing protocol with multipath capability. Because there are two identical, parallel IP subnetworks, there are at least two equally good paths from every Edge LSR to every other Edge LSR, one in each subnetwork.

OSPF equal-cost multipath chooses to distribute traffic evenly across both sets of paths (and hence both subnetworks). Because of this, MPLS sets up two identical sets of connections for the two MPLS control planes. IP traffic is shared evenly across the two sets of connections, across both control planes.

LSC redundancy works with either of the two interior gateway routing protocols:

- OSPF
- Intermediate System to Intermediate System (IS-IS)

LSC redundancy also works with either of the two label distribution protocols for hop-by-hop routed MPLS:

- Tag Distribution Protocol (TDP)
- Label Distribution Protocol (LDP)

If there were a failure in one MPLS controller in one switch, some paths in one of the subnetworks would no longer work. If there were only one subnetwork, there would be an undesirable interruption in passing data while other switches break connections and reroute them around the failed node.

However, because all connections are mirrored in the secondary subnetwork, there are already alternative paths for the traffic without the need to establish new links. All that is required is for multipath routing to detect the failure of one set of paths and to divert the traffic onto the remaining good paths. Because connections on the other paths have already been set up, the interruption to traffic flow is much smaller than if new connections were required.

## Operational Modes

The LSC redundancy architecture supports these operational modes:

- **Transparent Mode**  
The primary and backup LSCs have identical images and startup configurations. Thus, when one controller fails, the other takes over seamlessly.
- **Upgrade Mode**  
You can upgrade the redundant system and change resources of the switch without rebooting the system. You can use this mode to change the resources between different partitions of the slave ATM switch.

## LSC Hot Redundancy

You can configure two LSC controllers for hot redundancy, which provides the fastest rerouting and equal cost routing.

Each of the two LSCs:

- Use VSI to control two separate partitions of the same IP+ATM switch
- Run in parallel with independent LDPs

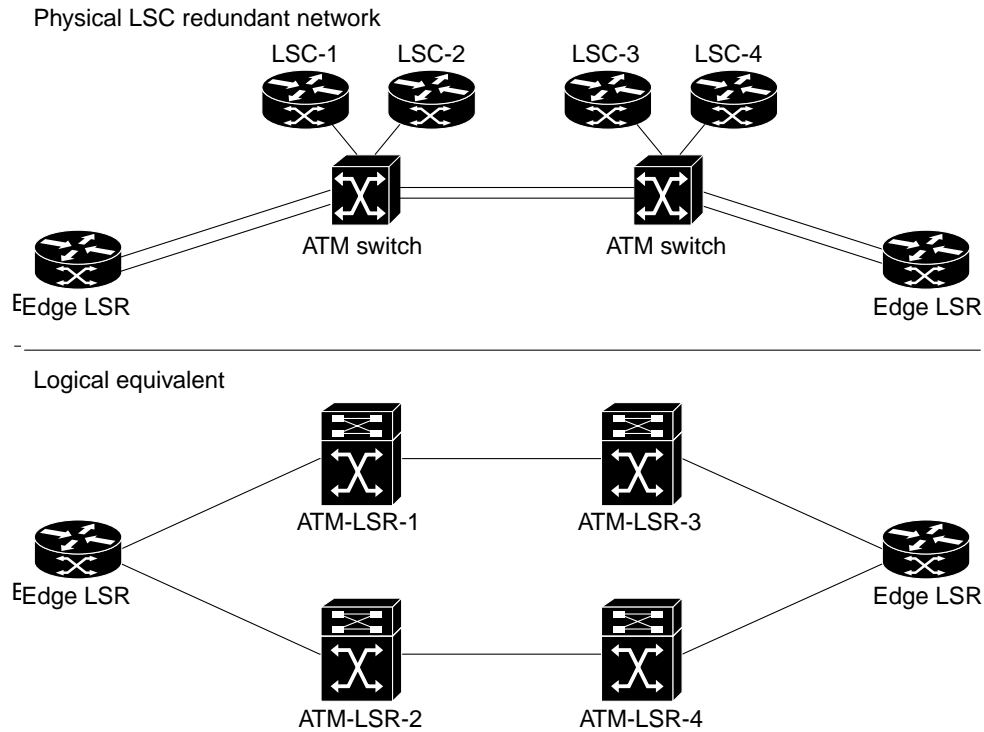
The backup MPLS controller provisions connections in tandem with the primary controller. Both controllers are active or “hot” at all times, giving each destination two independent paths, each path generated by one of the two controllers.

Hot redundancy (Figure 8-1 or Figure 8-2) uses two independent paths to route traffic. You set up both paths to use equal cost multipath routing, so that traffic is load balanced between the two paths. In other words, the two partitions on the switch must be configured with equal bandwidth and cross-connect space.

Also, both LSCs must run the same routing protocols.

The result is that the Edge LSRs have multiple routes to the same destination and request multiple labels. If one controller fails, only one of the two paths fails; the secondary controller already has the labels established and immediately provides an active backup path to handle the traffic with no time lost for rerouting or setting up labels.

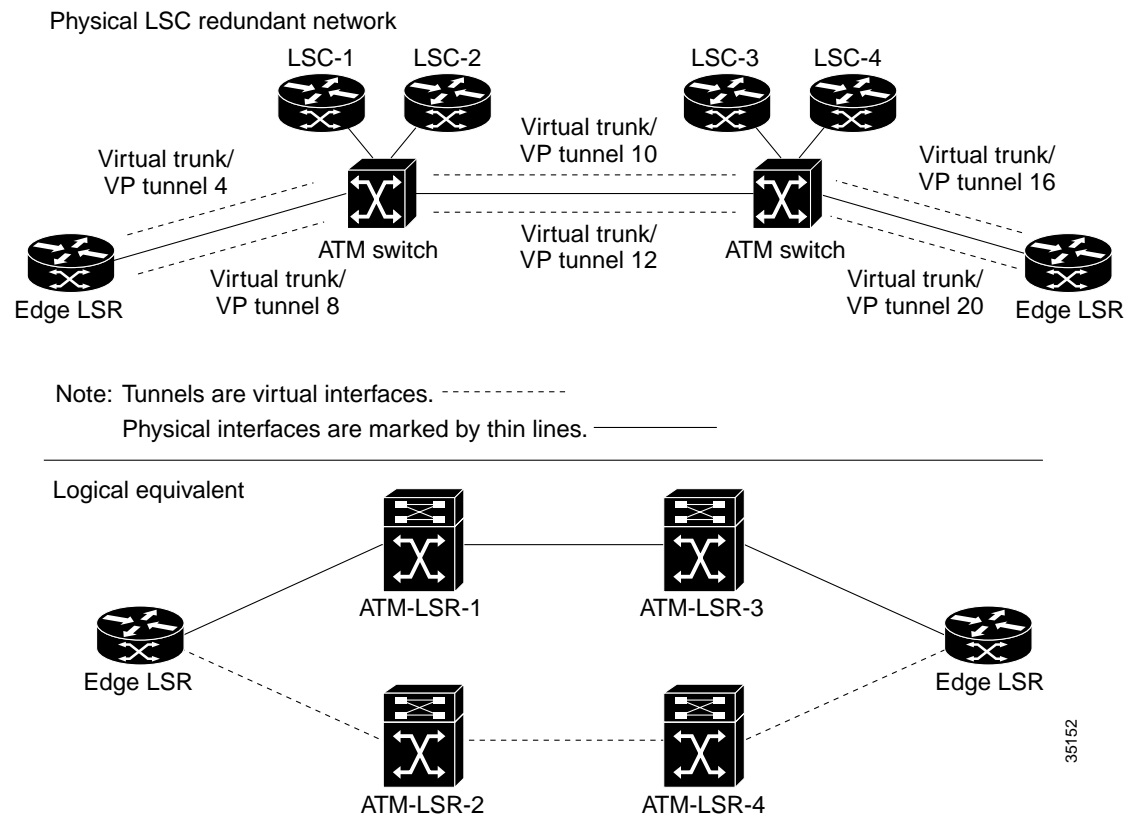
**Figure 8-1** LSC Redundancy with Physically Separate Trunks



**Note**

By placing two LSCs on an ATM switch, they become two logically separate ATM LSRs, which is what the “Logical equivalent” is showing. It’s important to clearly distinguish between an LSC and an ATM LSR: an LSC is not an ATM LSR, it is merely part of one.

Figure 8-2 LSC Redundancy with Shared Trunks



## How the LSC, ATM Switch, and VSI Work Together

The LSC and slave ATM switch have these characteristics:

- The LSC runs all of the control protocols
- The ATM switch forwards the data
- Each physical interface on the slave ATM switch maps to an XtagATM interface on the LSC. Each XtagATM interface has a dedicated LDP session with a corresponding interface on the edge. The XtagATM interfaces are mapped in the routing topology and the ATM switch behaves as a router.
- The LSC can also function as an Edge LSR. The data for the Edge LSR passes through the control interface of the router.

If a component on the LSC fails, the ATM switch's IP switching function is disabled. The stand-alone LSC is the single point of failure.

The VSI implementation includes these characteristics:

- The VSI allows multiple, independent control planes to control a switch. The VSI ensures that the control processes (SS7, MPLS, PNNI, and so on) can act independently of each other by using a VSI slave process to control the resources of the switch and apportion them to the correct control planes.
- In MPLS, each physical interface on the slave ATM switch maps to an XtagATM interface on the LSC through the VSI. In other words, physical interfaces are mapped to their respective logical interfaces.
- The routing protocol on the LSC generates route tables entries. The master sends connection requests and connection release requests to the slave.
- The slave sends the configured bandwidth parameters for the ATM switch interface to the master in the VSI messages. The master includes the bandwidth information in the link state topology. You can override these bandwidth values by manually configuring the bandwidth on the XtagATM interfaces.

## Implementing LSC Redundancy

To make an LSC redundant, you perform these basic steps:

- Partition the resources of the slave ATM switch
- Implement a parallel VSI model
- Assign redundant LSCs to each switch
- Create redundant LSRs

## Partitioning the Resources of the ATM Switch

In the LSC redundancy model, two LSCs control different partitions of the ATM switch. When you partition the ATM switch for LSC redundancy, follow these guidelines:

- Make the MPLS partitions identical. If you create two partitions, make sure both partitions have the same amount of resources. (You can have two MPLS VSI partitions per switch.) Use the **cnfrsrc** command to configure the partitions.
- If the partitions are on the same switch card, perform these steps:
  - Create different control VCs for each partition.  
For example, there can be only one (0, 32) control VC on the XtagATM interface. To map two XtagATM interfaces on the same ATM switch interface, use a different control VC for the second LSC. Use the **tag-switching atm control-vc** command.
  - Create the LVC on the XtagATM interfaces using nonintersecting VPI ranges.  
Use the **tag-switching atm vpi** command.
- Specify the bandwidth information on the XtagATM interfaces. Normally, this information is read from the slave ATM switch. When you specify the bandwidth on the XtagATM interface, the value you enter takes precedence over the switch-configured interface bandwidth.
- Configure the logical channel number (LCN) ranges for each partition according to the expected number of connections.

See the *Cisco BPX 8600 Series* documentation for more information about configuring the slave ATM switch.

## Implementing the Parallel VSI Model

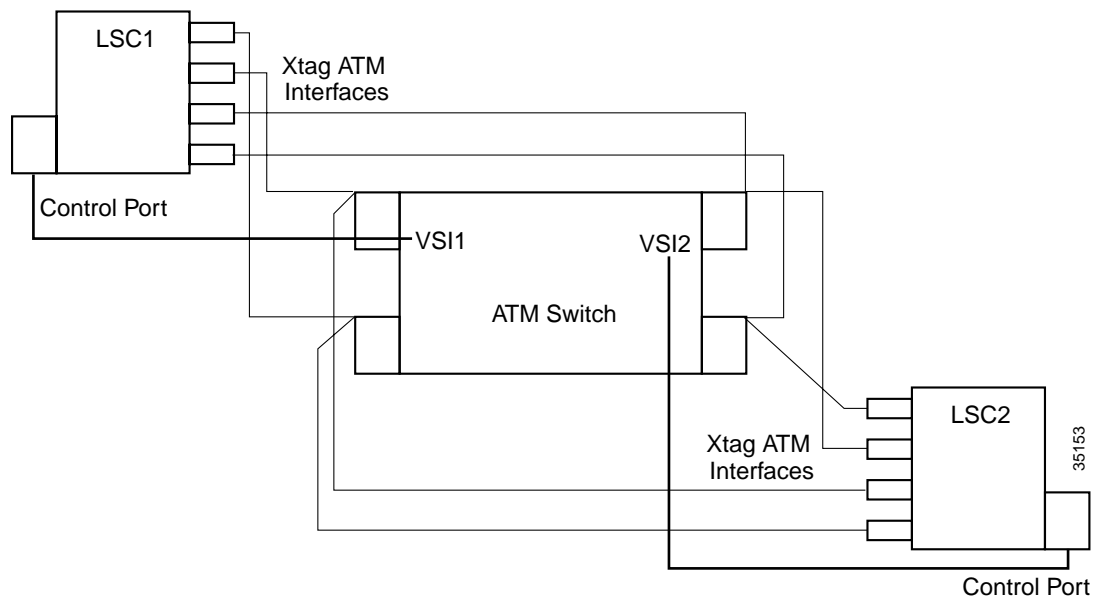
The parallel VSI model means that the physical interfaces on the ATM switch are shared by more than one LSC. For example:

- LSC1 maps VSI slave interfaces 1 to N to the ATM switch's physical interfaces 1 to N.
- LSC2 maps VSI slave interfaces to the ATM switch's physical interfaces 1 to N.
- LSC1 and LSC2 share the same physical interfaces on the ATM switch.

With this mapping, you achieve fully meshed independent masters.

Figure 8-3 shows four ATM physical interfaces mapped as four XtagATM interfaces at LSC1 and LSC2. Each LSC is unaware that the other LSC is mapped to the same interfaces. Both LSCs are active all the time. The ATM switch runs the same VSI protocol on both partitions.

**Figure 8-3** XtagATM Interfaces

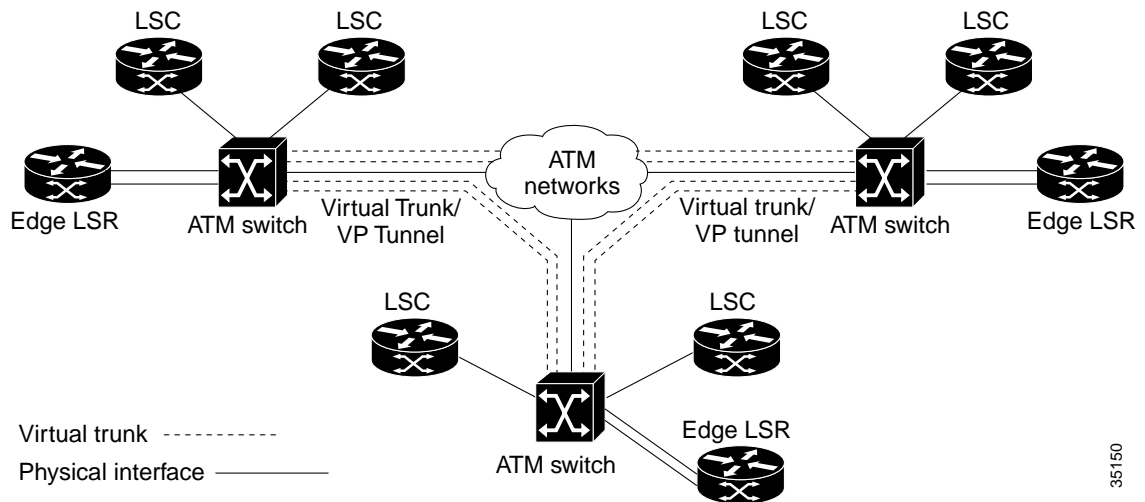


## Adding Interface Redundancy

To ensure reliability throughout the LSC redundant network, you can also implement:

- Redundant interfaces between the Edge LSR and the ATM LSR.  
Most Edge LSRs are co-located with the LSCs. Creating redundant interfaces between the Edge LSRs and the ATM LSRs reduces the chance of a disruption in network traffic by providing parallel paths.
- Redundant virtual trunks and VP tunnels between slave ATM switches.  
To ensure hot redundancy between the ATM switches, you can create redundant virtual trunks and VP tunnels. See Figure 8-4.

Figure 8-4 Interface Redundancy



35150

## Implementing Hot LSC Redundancy

Hot redundancy provides instant failover to the other path when an LSC fails. When you set up hot redundancy, both LSCs are active and have the same routing costs on both paths. To ensure that the routing costs are the same, run the same routing protocols on the redundant LSCs.

In hot redundancy, the LSCs run parallel and independent Label Distribution Protocols (LDPs). At the Edge LSRs, when the LDP has multiple routes for the same destination, it requests multiple labels. It also requests multiple labels when it needs to support Class of Service (CoS). When one LSC fails, the labels distributed by that LSC are removed.

To achieve hot redundancy, you can implement these redundant components:

- Redundant physical interfaces between the Edge LSR and the ATM LSR to ensure reliability in case one physical interface fails.
- Redundant interfaces or redundant VP tunnels between the ATM switches.
- Slave ATM switches, such as the BPX 8650, can have redundant control cards and switch fabrics. If redundant switch fabrics are used and the primary switch fails, the other switch fabric takes over.
- Redundant LSCs.
- The same routing protocol running on both LSCs. (You can have different label distribution protocols.)



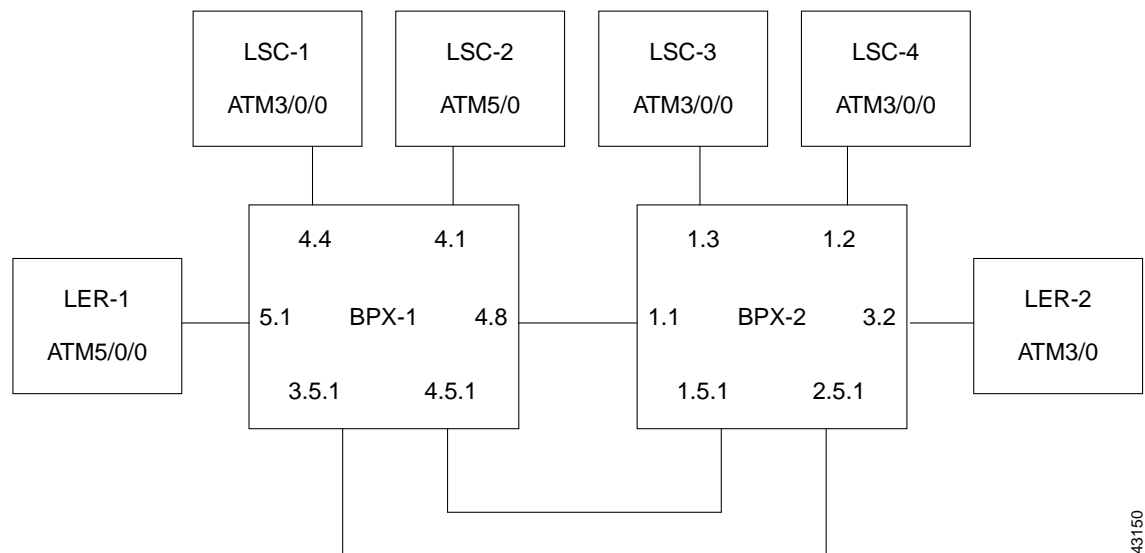
# Sample LSC Redundancy Configuration

The diagram in Figure 8-5 indicates the connections to support two active independent controllers on each BPX switch with two independent paths for each destination; that is, hot redundancy.

The sample configuration settings shown in this section assumes a network topology with two BPX switches (BPX1 and BPX2). Each BPX is connected to its own Edge Label Switch Router (Edge LSR) and each BPX supports two LSCs in separate partitions.

The Edge LSR and LSC must be forced to use different control VCs for the two partitions on the links between the LSC and BPX. In this example, this is done by using "tag-switching atm control-vc" commands, in the LSC and Edge LSR, for the second partition. The commands for the interfaces at both ends of a link must match, that is, specify the same control VC.

**Figure 8-5** Topology for Sample Hot Redundancy Configuration



## Note

Virtual trunks are not necessary and are not recommended in practice (except in very specific circumstances) because they break parts of LSC redundancy. Please disregard the virtual trunk interfaces in this example: 3.5.1, 4.5.1, 1.5.1, 2.5.1

The two LSCs on each BPX control different partitions, 1 and 2. The correct partition ID must be configured for all partitions controlled by each controller, including the partition on the LSC control interface.

43150

## Connections to BPX1

```
LSC1      4.4
  atm port atm3/0
LSC2      4.1
  atm port atm5/0
LER1      5.1
  Trunk port 4.8
  atm port atm5/0/0
```

## Connections to BPX2

```
LSC3      1.3
  atm port atm3/0/0
LSC4      1.2
  atm port atm3/0/0
LER2      3.2
  Trunk Port 1.1
  atm port atm3/0
```

## BPX1 Resource Parameter Settings

Use the **cnfrsrc** command to configure all VSI and AutoRoute resources. The following **dsprsrc** command screens show the recommended settings the BPX1 side of the basic topology. Naturally, depending on your network, you will need to adjust the resource parameters to maximize efficiency.

The configuration for BPX2 and its LER2, LSC3 and LSC4 are almost identical to those of the BPX1 but with different addresses for the BPX, the router ATM port, and loopback.

```
-----
Port/Trunk : 4.4

Maximum PVC LCNS:          256      Maximum PVC Bandwidth:148207
                               (Statistical Reserve: 5000)

Partition 1

Partition State :          Enabled
Minimum VSI LCNS:         0
Maximum VSI LCNS:        4096
Start VSI VPI:           100
End VSI VPI :            200
Minimum VSI Bandwidth :   0         Maximum VSI Bandwidth :    200000
VSI ILMF Config :        0

Last Command: dsprsrc 4.4 1
```

```
-----
Port/Trunk : 4.4
Maximum PVC LCNS:          256      Maximum PVC Bandwidth:148207
                                   (Statistical Reserve: 5000)
```

```
Partition 2
Partition State :          Disable
```

```
Last Command: dsprsrc 4.4 2
```

```
-----
Port/Trunk : 4.1
Maximum PVC LCNS:          256      Maximum PVC Bandwidth:148207
                                   (Statistical Reserve: 5000)
```

```
Partition 1
Partition State :          Disable
```

```
Last Command: dsprsrc 4.1 1
```

```
-----
Port/Trunk : 4.1
Maximum PVC LCNS:          256      Maximum PVC Bandwidth:148207
                                   (Statistical Reserve: 5000)
```

```
Partition 2
Partition State :          Enabled
Minimum VSI LCNS:          0
Maximum VSI LCNS:          4096
Start VSI VPI:             201
End VSI VPI :              300
Minimum VSI Bandwidth :    0      Maximum VSI Bandwidth :      200000
VSI ILMI Config :          0
```

```
-----
Port/Trunk : 4.8
```

```
Maximum PVC LCNS:          256      Maximum PVC Bandwidth:148207
                               (Statistical Reserve: 5000)
```

```
Partition 1
```

```
Partition State :          Enabled
Minimum VSI LCNS:          0
Maximum VSI LCNS:         4096
Start VSI VPI:            100
End VSI VPI :             200
Minimum VSI Bandwidth :    0      Maximum VSI Bandwidth :      200000
VSI ILMF Config :         0
```

```
Last Command: dsprsrc 4.8 1
```

```
-----
Port/Trunk : 4.8
```

```
Maximum PVC LCNS:          256      Maximum PVC Bandwidth:148207
                               (Statistical Reserve: 5000)
```

```
Partition 2
```

```
Partition State :          Enabled
Minimum VSI LCNS:          0
Maximum VSI LCNS:         4096
Start VSI VPI:            201
End VSI VPI :             255
Minimum VSI Bandwidth :    0      Maximum VSI Bandwidth :      100000
VSI ILMF Config :         0
```

```
Last Command: dsprsrc 4.8 2
```

```
-----
Virtual Trunk : 4.5.1
```

```
Maximum PVC LCNS:          256      Maximum PVC Bandwidth:867
                               (Statistical Reserve: 1000)
```

```
Partition 1
```

```
Partition State :          Enabled
Minimum VSI LCNS:          0
Maximum VSI LCNS:         4096
Start VSI VPI:            35
End VSI VPI :             35
Minimum VSI Bandwidth :    0      Maximum VSI Bandwidth :      1000
VSI ILMF Config :         0
```

```
Last Command: dsprsrc 4.5.1 1
```

```

-----
Virtual Trunk : 3.5.1

Maximum PVC LCNS:          256      Maximum PVC Bandwidth:867
                                (Statistical Reserve: 1000)

Partition 1

Partition State :          Enabled
Minimum VSI LCNS:         0
Maximum VSI LCNS:        4096
Start VSI VPI:           36
End VSI VPI :            36
Minimum VSI Bandwidth :   0          Maximum VSI Bandwidth :      1000
VSI ILMI Config :         0

Last Command: dsprsrc 3.5.1 1

-----
Port/Trunk : 5.1

Maximum PVC LCNS:          256      Maximum PVC Bandwidth:30000

Partition 1

Partition State :          Enabled
Minimum VSI LCNS:         0
Maximum VSI LCNS:        4096
Start VSI VPI:           100
End VSI VPI :            200
Minimum VSI Bandwidth :   0          Maximum VSI Bandwidth :      30000
VSI ILMI Config :         0

Last Command: dsprsrc 5.1 1

-----
Port/Trunk : 5.1

Maximum PVC LCNS:          256      Maximum PVC Bandwidth:30000

Partition 2

Partition State :          Enabled
Minimum VSI LCNS:         0
Maximum VSI LCNS:        4096
Start VSI VPI:           201
End VSI VPI :            255
Minimum VSI Bandwidth :   0          Maximum VSI Bandwidth :      30000
VSI ILMI Config :         0

Last Command: dsprsrc 5.1 2

```

You can then enable the two LSCs to control two different partitions by using:

```

addshelf 4.4 v 1 1          [partition id=1, controller id=1]
addshelf 4.1 v 2 2          [partition id=2, controller id=2]

```

This example uses a single physical link between BPX nodes, from interface 4.8 on BPX1 to 1.1 on BPX2. This physical trunk has two VSI partitions, one under the control of each LSC.

An alternative configuration, not shown in the diagram, would be to create two virtual trunks on the link, for example, with the BPX1 endpoints numbered 4.8.1 and 4.8.2.

The virtual trunk 4.8.1 would have only partition 1 enabled, and 4.8.2 would have only partition 2. Such a configuration is not recommended in practice, because it would prevent the sharing of spare bandwidth between the two “virtual networks” under the control of the two sets of LSCs.

The controller ID for the two LSCs must be different. It may equal the partition ID, but it could be a different value. Note that the controller ID must be specified both on the LSC (in the tag-control-protocol command) and the BPX switch (in **addshelf**).

## LER1 Configuration File

```

!
version 12.1
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname 7500-12
!
boot system slot0:rsp-pv-mz.121-1.1.T
enable secret 5 $1$QvGU$NDhlWJM9eYcXN3gJfgZcc1
enable password cisco
ip subnet-zero
ip cef
!
no ip domain-lookup
clns routing
tag-switching tdp router-id Loopback0
!
interface Loopback0
 ip address 12.12.12.12 255.255.255.255
!
interface ATM5/0/0
 no ip address
 no ip route-cache distributed
 atm framing cbitplcp
 no atm ilmi-keepalive
 tag-switching ip
!
interface ATM5/0/0.1 tag-switching
 ip unnumbered Loopback0
 tag-switching atm vpi 100-200
 tag-switching ip
!
interface ATM5/0/0.2 tag-switching
 ip unnumbered Loopback0
 tag-switching atm control-vc 201 40
 tag-switching atm vpi 201-255
 tag-switching ip
!
router ospf 50
 network 12.12.12.12 0.0.0.0 area 5
!
ip classless
ip route 0.0.0.0 0.0.0.0 172.29.113.1
no ip http server
!
!

```

```
tftp-server slot0:rsp-jsv-mz.120-6.5.T4
!
line con 0
  exec-timeout 0 0
  transport input none
line aux 0
line vty 0 4
  exec-timeout 0 0
  password cisco
  login
!
no scheduler max-task-time
end
```

## LSC1 Configuration File

```
!
version 12.1
no service pad
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
service compress-config
!
hostname R7200-12
!
boot system slot0:c7200-p-mz.121-1.1.T
enable password cisco
!
!
!
!
!
ip subnet-zero
ip cef
no ip domain-lookup
!
tag-switching tdp router-id Loopback0
!
!
interface Loopback0
  ip address 112.112.112.112 255.255.255.255
  no ip route-cache
  no ip mroute-cache
!
interface ATM3/0
  no ip address
  no ip mroute-cache
  tag-control-protocol vsi id 1
  ! set controller id to 1 (the default)
  no atm ilmi-keepalive
!
!
interface XTagATM48
  ip unnumbered Loopback0
  no ip route-cache cef
  extended-port ATM3/0 bpx 4.8
  tag-switching ip
!
```

```

interface XTagATM51
 ip unnumbered Loopback0
 no ip route-cache cef
 extended-port ATM3/0 bpx 5.1
 tag-switching ip
!
router ospf 50
 network 112.112.112.112 0.0.0.0 area 5
!
ip classless
ip route 0.0.0.0 0.0.0.0 172.29.113.1
no ip http server
!
!
line con 0
 exec-timeout 0 0
 password cisco
 transport input none
line aux 0
line vty 0 4
 exec-timeout 0 0
 password cisco
 login
!
end

```

## LSC2 Configuration File

```

!
version 12.1
no service pad
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname R7200-13
!
boot system tftp c7200-p-mz.121-1.1.T 172.29.113.87
enable password cisco
!
!
!
!
!
ip subnet-zero
ip cef
no ip domain-lookup
!
tag-switching tdp router-id Loopback0
!
!
!
!
!

```



```
interface Loopback0
 ip address 13.13.13.13 255.255.255.255
interface ATM5/0
 no ip address
 tag-control-protocol vsi id 2
 ! set controller id to 2
 no atm ilmi-keepalive
 tag-switching ip
 !
interface Hssi6/0
 no ip address
 no ip mroute-cache
 shutdown
 fair-queue
 !
interface XTagATM48
 ip unnumbered Loopback0
 no ip route-cache cef
 extended-port ATM5/0 bpx 4.8
 tag-switching atm control-vc 201 40
 tag-switching ip
 !
interface XTagATM51
 ip unnumbered Loopback0
 extended-port ATM5/0 bpx 5.1
 tag-switching atm control-vc 201 40
 tag-switching ip
 !
router ospf 50
 network 13.13.13.13 0.0.0.0 area 5
 !
ip classless
ip route 0.0.0.0 0.0.0.0 172.29.113.1
no ip http server
 !
 !
 !
line con 0
 exec-timeout 0 0
 transport input none
line aux 0
line vty 0 4
 exec-timeout 0 0
 no login
 !
no scheduler max-task-time
end
```





---

## A

**ATM LSR** An ATM label switching router with a number of LC-ATM interfaces. The router forwards the cells among these interfaces using labels carried in the VPI/VCI field.

---

## C

**CAR** Committed Access Rate (packet classification). CAR is the main feature supporting packet classification. CAR uses the type of service (TOS) bits in the IP header to classify packets. CAR classification commands are used to classify and reclassify a packet.

**CEF** Cisco Express Forwarding. An advanced Layer 3 IP switching technology. CEF optimizes network performance and scalability for networks with large and dynamic traffic patterns.

**CE router** Customer edge router. A router that is part of a customer network and that interfaces to a provider edge (PE) router.

**CoS** Class of service. A feature that provides scalable, differentiated types of service across a label switched network.

---

## D

**DWFQ** VIP-Distributed WFQ.

**DWRED** VIP-Distributed WRED.

---

## E

**Edge ATM Edge LSR** A router that is connected to the ATM LSR cloud through LC-ATM interfaces. The edge ATM LSR adds labels to unlabeled packets and removes labels from unlabeled packets.

**Edge ATM LSR** A switch router that is connected to the ATM LSR cloud through LC-ATM interfaces. The edge ATM LSR adds labels to unlabeled packets and strips labels from unlabeled packets.

**Edge Label Switch Router (LSR)** The edge device that performs initial packet processing and classification and applies the first label. This device can be either a router, such as the Cisco 7500, or a switch with built-in routing, such as the Cisco BPX 8650.

---

**G**

**GRE** Generic routing encapsulation. A tunneling protocol developed by Cisco that can encapsulate a wide variety of protocol packet types inside IP tunnels, creating a virtual point-to-point link to Cisco routers at remote points over an IP internetwork. By connecting multiprotocol subnetworks in a single-protocol backbone environment, IP tunneling that uses GRE allows network expansion across a single-protocol backbone environment.

---

**I**

**IGP** Interior Gateway Protocol. An Internet protocol used to exchange routing information within an autonomous system. Examples of common IGPs include IGRP, OSPF, and RIP.

**IP Precedence** A 3-bit value in TOS byte used for assigning Precedence to IP packets.

**IS-IS** Intermediate system-to-intermediate system. OSI link-state hierarchical routing protocol in which ISs (routers) exchange routing information based on a single metric in order to determine network topology.

---

**L**

**Label** A label is a header used by an LSR to forward packets. The header format depends upon network characteristics. In router networks, the label is a separate, 32-bit header. In ATM networks, the label is placed into the virtual channel identifier/virtual path identifier (VCI/VPI) cell header. In the core, LSRs read only the label, not the packet header. One key to the scalability of MPLS is that labels have only local significance between two devices that are communicating.

**Label-Controlled ATM Interface (LC-ATM interface)** An interface on a router or switch that uses label distribution procedures to negotiate label VCs.

**Label Distribution Protocol (LDP)** Provides communication between edge and core devices. It assigns labels in edge and core devices to establish Label Switched Paths (LSPs) in conjunction with routing protocols such as OSPF, IS-IS, Enhanced Interior Gateway Routing Protocol (EIGRP), or BGP.

**Label Imposition** The act of putting the first label on a packet.

**LSA** Link-state advertisement. A broadcast packet used by link-state protocols. The LSA contains information about neighbors and path costs and is used by the receiving router to maintain a routing table.

**Label-Switched Path (LSP)** A sequence of hops (R0...Rn) in which a packet travels from R0 to Rn through MPLS Switching mechanisms. A label-switched path can be established dynamically, based on normal routing mechanisms, or through configuration.

**Label-Switched Path (LSP) Tunnel** A configured connection between two routers, in which label Switching is used to carry the packet.

---

**L**

- Label Switch Router (LSR)** The core device that switches labeled packets according to precomputed switching tables. It can also be a switch or a router
- Label VC (LVC)** An ATM virtual circuit that is set up through ATM LSR label distribution procedures.

---

**M**

- MPLS** Multiprotocol Label Switching. Networks using MPLS, transport IP packets over ATM using label switching, thereby realizing the flexibility and scalability of TCP/IP along with the switching speed and reliability of ATM.

---

**N**

- NLRI** Network layer reachability information. BGP sends routing update messages containing NLRI to describe a route and how to get there. In this context, an NLRI is a prefix. A BGP update message carries one or more NLRI prefixes and the attributes of a route for the NLRI prefixes; the route attributes include a BGP next hop gateway address, community values, and other information.

---

**P**

- PE Router** Provider edge router. A router that is part of a service provider's network and that is connected to a customer edge (CE) router. The PE router function is a combination of an MLS edge label switch router (LSR) function with some additional functions to support VPNs.

---

**Q**

- QoS** Quality of service. Measure of performance for a transmission system that reflects its transmission quality and service availability.

---

**R**

- RED** Random early detection. Congestion avoidance algorithm in which a small percentage of packets are dropped when congestion is detected and before the queue in question overflows completely.
- RD** Route distinguisher. An 8-byte value that is concatenated with an IPv4 prefix to create a unique VPN IPv4 prefix.
- RIP** Routing Information Protocol. Used to exchange routing information within an autonomous system, RIP uses hop count as a routing metric.

---

**T**

- Traffic Engineering** The techniques and processes used to cause routed traffic to travel through the network on a path other than the one that would have been chosen if standard routing methods had been used.
- Traffic Engineering Tunnel** A label-switched path tunnel that is used for engineering traffic. It is set up through means other than normal Layer 3 routing and is used to direct traffic over a path different from the one that Layer 3 routing would cause it to take.
- Tunneling** Architecture providing the services necessary to implement any standard point-to-point data encapsulation scheme.
- TOS** Type of Service. A byte in the IPv4 header.

---

**V**

- VPN** Virtual private network. A secure network that shares resources with one or more physical networks. A VPN can contain one or more geographically dispersed sites that can communicate securely over a shared backbone.
- VPNv4** Used as a keyword in commands to indicate VPN-IPv4 prefixes. These prefixes are customer VPN addresses, each of which has been made unique by the addition of an 8-byte route distinguisher.
- VRF** VPN routing/forwarding instance. A VRF consists of an IP routing table, a derived forwarding table, a set of interfaces that use the forwarding table, and a set of rules and routing protocols that determine what goes into the forwarding table. In general, a VRF includes the routing information that defines a customer VPN site that is attached to a PE router.

---

**W**

- WEPD** Weighted Early Packet Discard.
- WFQ** Weighted Fair Queueing. A queue management algorithm that provides a certain fraction of link bandwidth to each of several queues, based on a relative bandwidth applied to each of the queues.
- WRED** Weighted RED. A variant of RED in which the probability of a packet being dropped depends on either, its IP Precedence, CAR marking, or Label Switching CoS (as well as the other factors in the RED algorithm).

---

**X**

- xBGP** Border Gateway Protocol. Interdomain routing protocol that exchanges reachability information with other BGP systems. It is defined in RFC 1163.



---

## A

### ATM LSR

adding redundancy 8-9

### ATM switch

adding redundancy 8-9

configuring for LSC redundancy 8-8

partitioning 8-8

### ATM-TSR

*See* ATM label switch router; terminology

---

## B

### bandwidth

on XtagATM interfaces 8-8

specifying 8-8

### boot

sequence 5-23

boot sequence 5-23

---

## C

### Cisco IOS software

basics 5-33

getting help 5-35

modes of operation 5-33

saving the configuration 5-35

### commands

config terminal 5-32

copy running-config startup-config 5-33

enable 5-32

config terminal command 5-32

### configuration

displaying 5-35

manual 5-24

saving changes to 5-35

conventions, documentation **xxi**

copy running-config startup-config command 5-33

customer premises equipment. *See* CPE 1-6

---

## D

### documentation

conventions **xxi**

---

## E

### Edge LSR

adding redundancy 8-9

enable command 5-32

---

## G

global configuration mode 5-34

---

## H

hot LSC redundancy 8-10

---

## I

### installation

changing configuration register settings 5-33

## L

## Label Distribution Protocol

compared to Tag Distribution Protocol **xix**

## LCN

configuring **8-8**

## LDP

configuring for hot redundancy **8-10**

logical channel number. *See* LCN. **8-8**

## LSC redundancy

benefits **8-2**

configuring hot redundancy **8-10**

configuring the ATM switch **8-8**

configuring the VSI **8-9**

migrating from a standalone LSC **8-3**

switching from hot to warm **8-3**

## M

## Multiprotocol Label Switching

terminology **xix**

Multiprotocol Label Switching. *See* MPLS.

## P

privileged EXEC mode **5-34**

## R

ROM monitor mode **5-34**

## routing protocols

for hot redundancy **8-10**

## S

## setup

command facility **5-24**

manual configuration **5-32**

## software

initial setup program **5-24**

System Configuration Dialog, using **5-24**

## T

tag, TER, TFIB, TSR, TSC, TSP, TVC, *any name containing tag*

*See* terminology **xix**

Tag Distribution Protocol **xix**

*See also* Label Distribution Protocol; terminology

Tag Switching **xix**

*See also* Multiprotocol Label Switching; terminology

## TDP

*See* Tag Distribution Protocol

terminology **xix**

traffic engineering **Glossary-4**

traffic engineering tunnel **Glossary-4**

## U

user EXEC mode **5-34**

## V

## virtual circuit

creating different control VCs **8-8**

## virtual trunks

adding redundancy **8-9**

## VSI

configuring for LSC redundancy **8-9**

## X

## XtagATM

specifying bandwidth **8-8**