MFA Forum Technical Committee Contribution

Subject: Straw Ballot Text for MPLS InterCarrier Interconnect Technical Specification

---

Contacts: Nabil Bitar
Verizon
40 Sylvan Rd
Waltham, MA 02451
Phone: +1 781 466 2161
Email: nabil.n.bitar@verizon.com

Contributors:
Matthew Bocci, Alcatel-Lucent
Rao Cherkuri, Juniper Networks
Ross Callon, Juniper Networks
Anne Exter, Verizon
Reda Haddad, Ericsson
Martin Halstead, Nexagent
John Kenney, Tellabs
Andrew Malis, Verizon
David McDysan, Verizon
Donnald O'Connor, Fujitsu
Nikhil Shah, Lucent Technologies
David Sinicrope, Ericsson

Date: September, 2007

Location:

Distribution: Participants in the MFA Forum Technical Committee A&D Working Group

---

Abstract: This document is the result of the straw ballot resolution in San Jose, California And conference calls on 8/6/2007 and 8/13/2007. Not all notes in the meeting minutes from these two calls have been addressed in this version.

Declaration of IPR:
The contributor has read and supports the MFA Forum's IPR Policy Statement. To the best of the contributor's knowledge, the contributing member: does not have any IPR that would be infringed by the Technical Specification proposed by any member. The objective of this Technical Specification is to enable a set of MPLS-based applications across administrative domains operated by different carriers.

For additional information contact:
The MFA Forum, 39355 California Street,
Suite 307, Fremont, CA 94538
510-608-3997 phone
info@mplsforum.org

Contribution History

Version	Change	Date
00	Merged mpls2007.00.01 that contains baseline text for the MPLS-ICI with the following contributions: (1) mpls2207.016.00 on MPLS OAM,, (2) mpls2207.004.00 on MIBS, (3) mpls2007.006.00 on Load balancing, MTU handling, TTL processing, and (4) mpls2007.014.01 on IPVPN. In addition, Annex F was removed and it is to be substituted by appendix B that illustrates via examples how EXP maoping is performed at ASBRs between two providers to effectively map a CoS in one provider network to s CoS in the other provider network.	01/29/2007
01	Embedded commented demarked by NB from the Richardson Interim meeting	02/12/2007
02	Addressed the comments from the Richardson Interim meeting. The comments from the Richardson meeting are left in place for tracking. All new modification addressing these comments are demarked by NB with date 0/0/2007 for easy tracing The following modifications were made: : (1) Updated the rference list and reformatted, (2) Updated the reference diagrams in Section 6, (3) Included text for LSP ping and RSVP-TE for path MTU discovery in Sections 11.2.2 and 11.2.3, (4) Re-orgamized the OAM section 12.1 and 12.2 as indicated in the respective section, and (5) added graceful restart in Annex B, C and D. Inserted a section for MS-PW OAM in Annex C to be included from other contributions. In addition, updated the numbering style for appendices.	03/03/2007
03	- Split references between informational and normative	03/05/2007
04	Comments and online edits from the Chicago meeting 3-6 March 2007	
05	Merged in contribution mpls2007.21.02, mpls20007.025.01, mpls2007.027.02 and edits that came from comments during the Richardson and Chicago meetings, including edits contributed by Matthew Bocci on PW OAM contribution mpls2007.044.00 and the figures by Don O'Connor and Davie McDysan on the protection models in Section 9.1.1 along with suggested edits in some places.	04/17/2007
06	Checked in edits from version 05	06/04/2007
07	Incorporated comments from contribution mpls2007.022.01, with some editorials, in Section 9.1 and in Annex C. Also included a missing sentence in Section 11.1 that was agreed on in Chicago. In addition, changed the word on fast reroute to fast switchover to address some of Dan O'Connor comment in section 9. Also addressed Matthew Bocci's comment about graceful restart and that routers with hot control plane will not need to support graceful restart in restart mode.	06/04/2007
08	Accepted changes from mpls2007.017.07	06/22/2007
09	Made the following additional modification with some minor edits to be consistent with the edit prposed in mpls2007..022.01: - edited section 9.1 and C1.1 the following sentence: Section 1 of RFC 4447 states that " packets that are transmitted from one end of the pseudowire to the other are MPLS packets which must be transmitted through an MPLS tunnel. However, if the pseudowire endpoints are immediately adjacent and penultimate hop popping behavior is in use, the MPLS tunnel may not be	06/22/2007

	necessary." - Eliminated a paragraph that talks about why when a PSN tunnel is needed or not in Annx C	
10	Include edits from the San Jose MFA Metting	07/18/2007
11	Addrssed all the comments from the SJ meeting the 8/6 call except as noted and as it is related to Fujitsu comments on section 9. There are comments embedded in this document from the 8/13 call in Annex C that were documented as result of the call discussion. Fujitsu has commited to addressing their comments by proposing text.	08/20/2007
12	Incorporated changes from the 08/20/2007 and 08/21/2007 calls. Next revision will update Annex C per the call on 8/27/2007 and figures in Section 9 that pertain to the protection model.	09/03/2007
13	Addressed the comments on the 8/27 call and incorporated the text that ties Annex C protection models to section 9 with minor modifications that can be seen by comparing this version to the previous one	09/06/2007
14	Straw ballot text. Updated refrence for IP-BFD and accpeted the changes from version 13	09/06/2007
15	Straw ballot text. Fixed some figure formatting problems with Figures 11 and 12.	09/06/2007

MFA Forum

MPLS Inter-Carrier Interconnect (MPLS-ICI)

Technical Specification

MFA Forum Technical Committee

September 2007

Note: The user's attention is called to the possibility that implementation of the MFA Forum Technical Specification contained herein may require the use of inventions covered by patent rights held by third parties. By publication of this MFA Forum Technical Specification the MFA Forum makes no representation that the implementation of the specification will not infringe on any third party rights. The MFA Forum take no position with respect to any claim that has been or may be asserted by any third party, the validity of any patent rights related to any such claims, or the extent to which a license to use any such rights may not be available.

Editors:

Nabil Bitar
Verizon

Contributors:

Matthew Bocci, Alcatel-Lucent
Rao Cherkuri, Juniper Networks
Ross Callon, Juniper Networks
Anne Exter, Verizon
Reda Haddad, Ericsson
Martin Halstead, Nexagent
John Kenney, Tellabs
Andrew Malis, Verizon
David McDysan, Verizon
Donnald O'Connor, Fujitsu
Nikhil Shah, Lucent Technologies
David Sinicrope, Ericsson

For more information contact:

The MFA Forum

48377 Fremont Blvd., Suite 117
Fremont, CA 94538
Phone: +1-510-492-4056
Fax: +1-510-492-4001
E-mail: info@mfaforum.org
WWW: http://www.mfaforum.org/

Full Notice

Table of Contents

1. Introduction

This document presents the MPLS InterCarrier Interconnect (MPLS-ICI) Technical Specification. The objective of this Technical Specification is to enable a set of MPLS services, enlisted in Section 2, across administrative domains operated by different carriers.

The MPLS-ICI is a bi-directional IP-MPLS logical link between an Autonomous System Border Router (ASBR) in one service provider network and an ASBR in another service provider network. The logical link appears internally to each ASBR as a single logical interface. There may be more than one MPLS-ICI between a pair of ASBRs. An ASBR, as specified in this document, is not required to support IP forwarding of user datagrams when the service does not require it

One or more Label Switched Paths (LSPs) may cross the MPLS-ICI. A given ASBR will terminate an LSP that crosses the MPLS-ICI, or it will switch the LSP toward the destination.

This Technical Specification focuses on mechanisms that enable the establishment of inter-carrier LSPs, including pairs of LSPs that form Pseudowires. It covers configuration, routing, signaling, and forwarding as well as management of the label switched paths.

This document is organized into a base document and services annexes. The base document describes the base functional elements of the MPLS-ICI. It also describes the protocols, information elements and procedures required to provide the base functional elements. This information applies to all MPLS services that utilize the MPLS-ICI. The scope section enlists the MPLS services covered in this document.

Each optional MPLS-service annex describes the protocol information content, and procedures for an MPLS service that uses the MPLS-ICI. An MPLS-service annex describes the service, the use of the base MPLS-ICI elements, any additional attributes conveyed in the MPLS-ICI messages, if needed, and procedural processing of the attributes specific to the service.

This document also specifies the Management Information Base (MIB) modules as standardized by the IETF to configure and manage the various aspects of an MPLS-ICI service in the appropriate sections. Compliance groups from the SNMP MIB's are referenced using the descriptive name and specific Management Information Base (MIB) in this document. An ASBR should support the management information model consistent with these groups. An informative appendix for each major MIB provides a summary description of what each MIB group or notification performs.

Although the document uses MIBs to describe the management information model, an ASBR may implement various management interfaces for read-only or read/write support such as CLI, XML, enterprise MIBs, or configuration files besides SNMP.

2. Scope

The purspose of this document is to specify capabilities that enable inter-carrier MPLS services. This document is targeted to equipment vendors to specify the necessary ASBR requirements and to service providers to provide guidance on how to use these capabilities. This document uses protocols developed in appropriate standard bodies (e.g., IETF, ITU) where applicable.

Following are the MPLS-ICI services covered in this document:

- BGP/MPLS IP Virtual Private Networks Multi-AS Backbone Option A
- BGP/MPLS IP Virtual Private Networks Multi-AS Backbone Option B
- Inter-domain tunnel establishment using BGP for IPv4
- Pseudowires and multi-segment Pseudowires (statically signaled and routed, statically routed and stiched with signaled segments)
- Interdomain Traffic Engineered tunneled (statically signaled and routed, dynamically signaled and routed)

For the services addressed in this document, following are within the scope of this Technical Specification:

- MPLS forwarding
- IP version 4 (IPv4) processing when processing of IP header is required as a matter of forwarding MPLS-labeled packets
- IP verion 4 (IPv4) processing for signaling and routing control
- MPLS Pseudo wires with static configuration
- Option ‘A’ of the ‘Multi-AS Backbones’ section of [RFC4364]. This is a case where the IP traffic forwarded on the MPLS-ICI is unlabeled IP but the MPLS service is a BGP/MPLS IPVPN service.
- TE Tunnels and RSVP-TE signaling
- LDP signaling for Pseudo Wires.
- MPLS IP-VPNs
- BGP for inter-domain routing and establishing label switched paths
- MPLS OAM
- Connection Admission Control
- Resiliency
- Classes of Service
- Policies
- Security
- MIBs: where no previous standards work exists, then this document describes a management information model in the form of an Enterprise MIB that a r would need to implement in order to support the MPLS-ICI base function or application.

Following are out of the scope of this Technical Specification:

- Inter-Carrier MPLS-based services that do not involve MPLS label processing at the interconnect for forwarding purposes. Therefore, usage of the MPLS-ICI by non-MPLS services (e.g., IP) is outside the scope of the MPLS-ICI. The only exception is RFC4364 Multi-AS Backbone Option A.
- IPv6: will be the subject of future consideration.

3. Definitions

must, shall or mandatory — the item is an absolute requirement of this Technical Specification.

should — the item is desirable.

may or optional — the item is not compulsory, and may be followed or ignored according to the needs of the implementer.

4. Acronyms

ATM	Asynchronous Transfer Mode
AS	Autonomous System
ASBR	Autonomous System Border Router
BGP	Border Gateway Protocol
MP-BGP	Multi-Protocol BGP
CE	Customer Edge
CoS	Class of Service
DoS	Denial of Service
ERO	Explicit Route Object
IETF	Internet Engineering Task Force
IP	Internet Protocol
ITU-T	InternationalTelecommunication Union - Telecom
LDP	Label Distribution Protocol
LSP	Label Switched Path
LSR	Label Switching Router
MIB	Management Information Base
MPLS	Multi Protocol Label Switching
PE	Provider Edge
PW	Pseudowire
QoS	Quality of Service
RSVP-TE	Resource Reservation Protocol with Traffic Engineering Extensions
SP	Service Provider
S-PE	Switching PE
TE	Traffic Engineering
T-PE	Terminating PE
VPN	Virtual Private Network

5. References

5.1. Normative References

[RFC4364] E. Rosen, Y. Rekhter, "BGP/MPLS IP Virtual Private Networks," IETF, RFC 4364, February 2006

[RFC3209] D. Awduche, et al., "RSVP-TE: Extensions to RSVP for LSP Tunnels," IETF, RFC 3209, December 2001.

[RFC3107] Y. Rekhter and E. Rosen, "Carrying Label Information in BGP-4," IETF, RFC 3107, May 2001.

[inter-AS-RSVP-TE] A. Ayyangar, J.P. Vasseur, "Inter domain GMPLS Traffic Engineering - RSVP-TE extensions," IETF work in progress.

[RFC4447] L. Martini, et al., "Pseudowire Setup and Maintenance using the Label Distribution Protocol (LDP)," IETF, RFC 4447, April 2006.,

[RFC3270] F. Le Faucheur, et al, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services," IETF, RFC3270, May 2002.

[RFC4124] F. Le Faucheur et al., "Protocol extensions for support of Differentiated-Service-aware MPLS Traffic Engineering," IETF, RFC 4124, June 2005

[RFC3478] Y. Rekhter, et al., "Graceful Restart Mechanism for Label Distribution Protocol," IETF, RFC 3478, February 2003.

[RFC4724] S. Sangli, et al., "Graceful Restart Mechanism for BGP," IETF, RFC 4724, January 2007.

[RFC4781] Y. Rekhter and R. Aggarwal, "Graceful Restart Mechanism for BGP," IETF, RFC 4781, January 2007.

[RFC3471] L. Berger, et al., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description," IETF, RFC 3471, January 2003.*

[RFC4379] K. Kompella and G. Swallow, "Detecting MPLS Data Plane Failures," RFC 4379, February 2006 .

[BFD-Base] D. Katz and D. Ward, "Bidirectional Forwarding Detection," draft-ietf-bfd-base-06.txt, IETF work in progress.

[IP-BFD] D. Katz and D. Ward, "BF for IPv4 and IPv6 (Single Hop)," draft-ietf-bfd-v4v6-1hop-06.txt, IETF work in progress.

[vccv] T. Nadeau, C. Pignataro, R. Aggarwal, et al.. "Pseudo Wire Virtual Circuit Connectivity Verification (VCCV)," draft-ietf-pwe3-vccv-13.txt, IETF work in progress.

[MPLS-BFD] R. Aggarwal, et al., "BFD For MPLS LSPs," draft-ietf-mpls-bfd-04.txt, IETF work in progress.

[RFC1191] J. Mogul and S. Deering, "Path MTU Discovery," IETF, RFC 1191, November 1990.

[interdomain-PDPC] J.P. Vasseur, et al., " A Per-domain path computation method for establishing Inter-domain Traffic Engineering (TE) Label Switched Paths (LSPs)," draft-ietf-ccamp-inter-domain-pd-path-comp-05.txt, IETF work in progress.

[RFC3032] E. Rosen, et al., "MPLS Label Stack Encoding," IETF, RFC 3032, January 2001.

[RFC3443] P. Agarwal and B. Akyol, "Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks," IETF, RFC 3443, January 2003.

[BGP-ORF] Chen, E., and Rekhter, Y., "Outbound Route Filtering Capability for BGP-4", draft-ietf-idr-route-filter-16.txt, IETF work in progress.

[MPLS-CNI] Packet based GMPLS Client to Network Interconnect (CNI) MFAF X.0.0

[RFC4623] A. Malis and M. Townsley, "Pseudowire Emulation Edge-to-Edge (PWE3) Fragmentation and Reassembly", IETF RFC 4623, August 2003.

[RFC4684] P. Marques et al, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", IETF RFC 4684, November 2006.

5.2. Informative References

[RFC2475] S. Blake, et al., "An Architecture for Differentiated Service," IETF, RFC 2475, December 1998.

[MS-PW_Requirements] L. Martini, N. Bitar, and M. Bocci., "Requirements for multi-Segment Pseudo-Wires," IETF work in progress.

[RFC2697] J. Heinanen and R. Guerin, "A Single Rate Three Color Marker," IETF, RFC 2697, September, 1999.

[RFC2698] J. Heinanen and R. Guerin, "A Two Rate Three Color Marker," IETF, RFC 2698, September 1999.

[RFC4385] S. Bryant et al., "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", IETF RFC 4385, February 2006.

[RFC4111] L. Fang, et al.,"Security Framework for Provider Provisioned Virtual Private Networks," IETF, RFC 4111, July 2005.

[RFC4778] M. Kaeo, "Current Operational Security Practices in Internet Service Provider Environments," IETF, RFC 4778, January 2007.

[RFC3871] G. Jones, et al, "Operational Security Requirements for Large Internet Service Provider (ISP) IP Network Infrastructure," IETF, RFC3871, September 2004.

[SecEfforts] C. Lonvick and D. Spak, "Security Best Practices Efforts and Documents," IETF work in progress.

[RFC4273] J. Haas and S. Hares, "Definitions of Managed Objects for BGP-4," IETF, RFC 4273, January, 2006,

[RFC2206] F. Baker, J. Krawczyk, and A. Sastry, "RSVP Management Information Base using SMIv2," IETF, RFC2206, September 1997.

[RFC3812] C. Srinivasan, et al., "MPLS Traffic Engineering (TE) MIB," IETF, RFC 3812, June 2004.

[RFC3813] C. Srinivasan, et al., "MPLS Label Switching Router (LSR) MIB," IETF, RFC 3813, June 2004.

[RFC3815] J. Cucchiara, et al., "Definitions of Managed Objects for the MPLS Label Distribution Protocol (LDP)," IETF, RFC 3815, June 2004.

[RFC3289] F. Baker, K. Chan, and A. Smith, "Management Information Base for the Differentiated Services Architecture," IETF, RFC 3289, May 2002.

[RFC3290] Y. Bernet, et al., "An Informal Management Model for Diffserv Routers," IETF, RFC 3290, May 2002.

[RFC4265] B. Schliesser and T. Nadeau, "Definition of Textual Conventions for Virtual Private Network (VPN) Management," IETF, RFC 4265, November 2005.

[RFC4382] T. Nadeau, et al., "MPLS/BGP Layer 3 Virtual Private Network (VPN) Management Information Base," IETF, RFC 4382, Febuary 2006.

[PWMIB] T. Nadeau , et al., "Pseudo Wire (PW) Management Information Base," IETF work in progress.

[PWMPLSMIB] D. Zelig and T. Nadeau, "Pseudo Wire (PW) over MPLS PSN MIB," IETF work in progress.

[BFDMIB] T. Nadeau and Z. Ali, "Bidirectional Forwarding Detection Management Information Base," IETF work in progress.

[Y.1541] ITU-T Recommendation Y.1541, "Network Performance Objectives for IP-Based Services," 2002/2006

[MIT_WP] "Inter-provider Quality of Service", MIT Communications Futures Program, November 2006.

[RFC3386] W. Lai and D. McDysan, "Network Hierarchy and Multilayer Survivability," IETF RFC 3386, November 2002.

[Y.1221] ITU-T, Recommendation Y.1221, "Traffic control and congestion control in IP based networks," 2003/2002.

[RFC 3086] K. Nichols and B. Carpenter,"Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification," RFC3086, April 2001

6. Reference Architecture

Following are reference model diagrams that describe the architecture of the MPLS inter-carrier interconnection (MPLS-ICI), illustrating the control protocols and service extensions across an inter-carrier MPLS interface, and the protocol stack. These reference models are applicable to all MPLS services discussed in this document.

Figure 1 illustrates how BGP Routing and MPLS Signaling (e.g., BGP, LDP or RSVP-TE) are used to establish an LSP or LSP segment between two Autonomous System Border Routers (ASBRs) across the MPLS-ICI. It depicts the establishment of tunnel LSP segments across each provider MPLS network and across the MPLS-ICI. The concatenation of these LSP segments forms an edge-to-edge tunnel LSP as shown in Figure 2. Figure 1 also depicts the establishment of service LSP segments (e.g., Pseudo Wire segment or BGP/MPLS IPVPN LSP segment) across each provider MPLS network and across the MPLS-ICI. The concatenation of these service LSP segments forms an edge-edge service LSP as shown in Figure 2. These reference diagrams do not prohibit an edge-to-edge service LSP to be established between PE1 and PE2 while being transparent to ASBR1 and ASBR2. That type of service LSP will be tunneled in the tunnel LSP. However, this case is out of the scope of this Technical Specification as it does not involve any direct action on the MPLS ICI. In addition, these reference models do not mandate that a tunnel LSP segment be established across the MPLS ICI to establish a service LSP segment across the MPLS ICI. These models describe the more generic case. The tunnel LSP segment across the MPLS-ICI can be NULL. In addition, the tunnel LSP itself can be the service LSP.

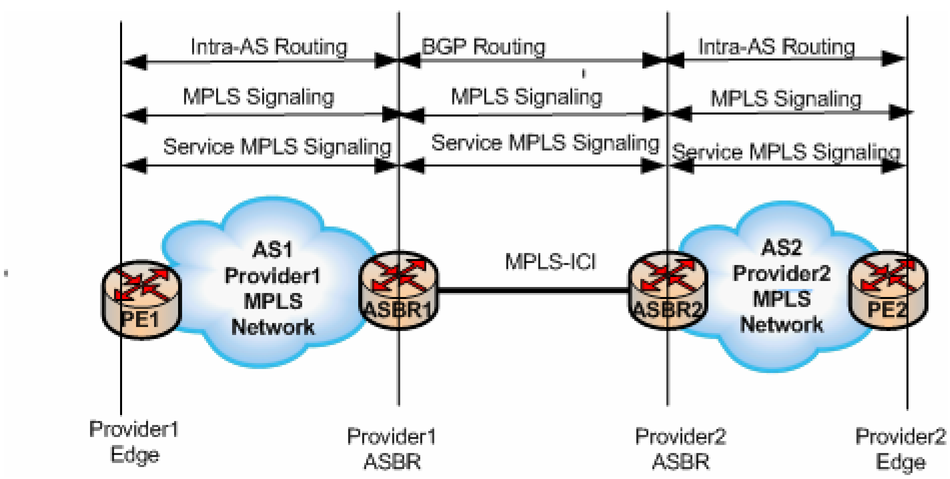


Figure 1: Control Protocols involved in extending MPLS services across an MPLS-ICI interconnecting two provider networks

Figure 2 illustrates the edge-to-edge MPLS services and LSPs used to transport these services between two providers. Such services may include (but are NOT limited to) BGP/MPLS IP-VPN, Layer2 Pseudo Wires (PWs) (e.g., Ethernet PWs), and data trunk tunneling through another Service Provider’s network

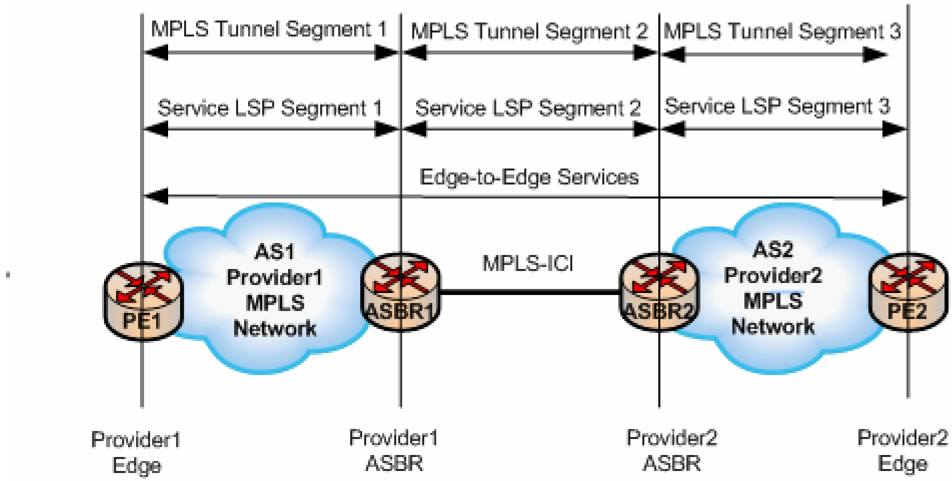


Figure 2: Services and LSPs extended across an MPLS-ICI interconnecting two provider networks

Figure 3 illustrates the protocol stack associated with the MPLS-service transport between two provider networks. Some of the layers in the stack may be transparent to ASBRs, or may be NULL if not used for a given application.

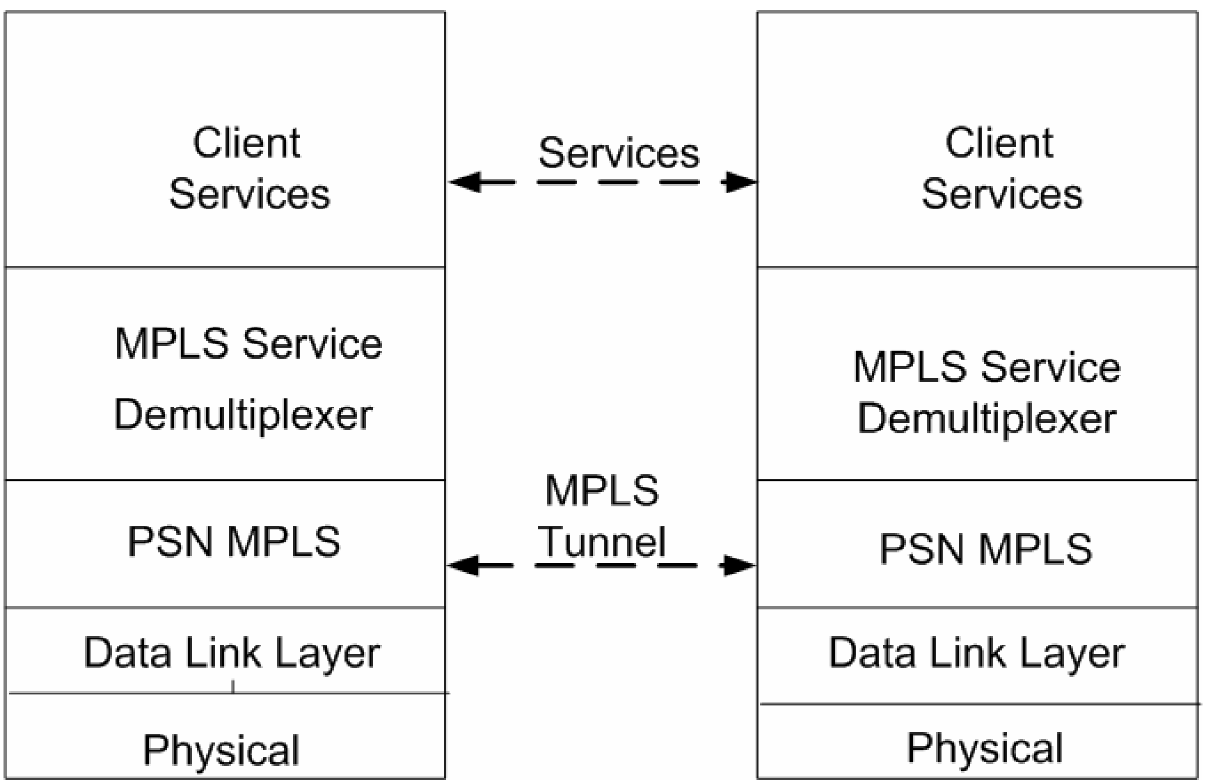


Figure 3: Protocol Stack

7. Physical layer (Layer 1)

MPLS operates over a variety of physical interfaces including TDM/SONET, Ethernet and optical wavelengths, operating at various speeds ranging from T1 to OC192 and beyond. This Technical Specification applies over any physical layer that supports the transport of IP and MPLS packets. An ASBR compliant with this Technical Specification must enable management of Layer 1 interfaces and protocols using the semantics of standard MIBs.

8. Data Link Layer (Layer 2)

MPLS operates over a variety of layer 2 technologies including HDLC, Packet over SONET, 802.1q, Ethernet, ATM and Frame Relay. This Technical Specification does not mandate or prohibit any specific layer 2 technologies that enable the transport of IP and MPLS packets. An ASBR compliant with this Technical Specification must enable management of Layer 2 interfaces and protocols using the semantics of standard MIBs.

9. Mechanisms for LSP establishment

This section identifies three mechanisms of LSP establishment across a provider domain boundary: all-static configuration, statically configured and signaled establishment, and dynamic establishment. It specifically identifies mechanisms for InterCarrier TE tunnel establishment, Multi-Segment Pseudo Wire establishment, and labeled BGP path establishment for IP-VPN and IP routes. These MPLS-services-specific mechanisms, along with associated signaling, routing and management, are discussed in the associated annexes.

9.1. All-Static configuration

An ASBR must be able to support static LSPs to establish Pseudo Wires (PWs) (Single-Segment PWs and Multi-Segment PWs) and MPLS tunnels across the domain boundary without the use of dynamic signaling. As stated in sections 5.2 and 5.3 of RFC 4447, "a (bidirectional) pseudowire consists of a pair of unidirectional LSPs, one in each direction". Section 1 of RFC 4447 states that "packets that are transmitted from one end of the pseudowire to the other are MPLS packets which must be transmitted through an MPLS tunnel. However, if the pseudowire endpoints are immediately adjacent and penultimate hop popping behavior is in use, the MPLS tunnel may not be necessary." Therefore, the procedures for setting up the LSP's and associated resiliency in the case of all-static configuration are covered in this section in a uniform model.

All-Static (manual) configuration refers to administratively or manually configuring an LSP or LSP segment that spans the interconnection between two domains' Autonomous System Border Routers (ASBRs). There is no MPLS signaling protocol between these ASBRs, and MPLS labels are manually assigned. All-Static configuration is labor intensive and introduces challenges in providing for resiliency or reroute around failures, but may be required to satisfy a provider's security and operational procedures. These security requirements may prohibit signaling between domains or require hiding the reachability of a PE that is the tail-end of the LSP in one domain to the other domain.

All-Static configuration can be used to establish an LSP that spans the interconnection between two ASBRs and terminates at these ASBRs, or to establish an LSP segment that gets stitched to one or more LSPs in the service provider domains on either side of the MPLS-ICI. All static configuration here refers to the establishment of the LSP or LSP segment between two ASBRs.

9.1.1. Resiliency

Resiliency of statically-configured LSPs or LSP segments across an MPLS-ICI requires protection against link failure between the same ASBR pair interconnected by the MPLS-ICI or end-end LSP protection.

Upon the failure of an MPLS-ICI carrying a statically configured LSP or LSP segment between two ASBRs, an ASBR should be able to redirect the LSP traffic to redundant MPLS-ICI. An implementation may support admission control of the LSP over the redundant MPLS-ICI prior to rerouting. Protection against MPLS-ICI failure must be configured on an LSP basis. In certain cases, there is no parallel MPLS-ICI between the two ASBRs that can satisfy the constraints of the LSP (e.g., bandwidth). In that case, the LSP must be rerouted through other intermediate nodes between the same ASBR pair. Such a redundant path should be configured at the time of the LSP/LSP segment configuration.

9.1.1.1. Protection Models

This Technical Specification describes several protection models that apply to statically configured LSPs, including tunnel LSPs and LSPs that form MPLS Pseudowises. These models do not cover layer 1 or layer 2 protection switching. If such protection switching is present, consideration should be given to applying appropriate holdoff timers to faciltate multi-layer protection per RFC 3386 [RFC3386] Section 3.5. ASBRs compliant with this Technical Specification must support the One-hop MPLS-ICI protection model below:

- One-hop MPLS-ICI protection as illustrated in Figure 4: In this case, two ASBRs are interconnected by a pair of MPLS-ICIs. Ideally, these MPLS-ICIs should be diversely routed over layer1 and layer2 networks to reduce fate sharing. The LSP segment is provisioned over a primary MPLS-ICI and a redundant MPLS-ICI. The LSP segment headend shall route the LSP segment over the primary or redundant MPLS-ICI depending on the state of the MPLS-ICI path or the state of the LSP segment. MPLS-ICI failure detection can be based on lower layer mechanisms or IP BFD [IP-BFD]. If the LSP segment has any constraints (e.g., bandwidth), the LSP segment should be admitted over both the primary and redundant MPLS-ICIs (from LSP segment aspect). The LSP segment over the redundant MPLS-ICI should be setup prior to the failure and activated upon failure detection.

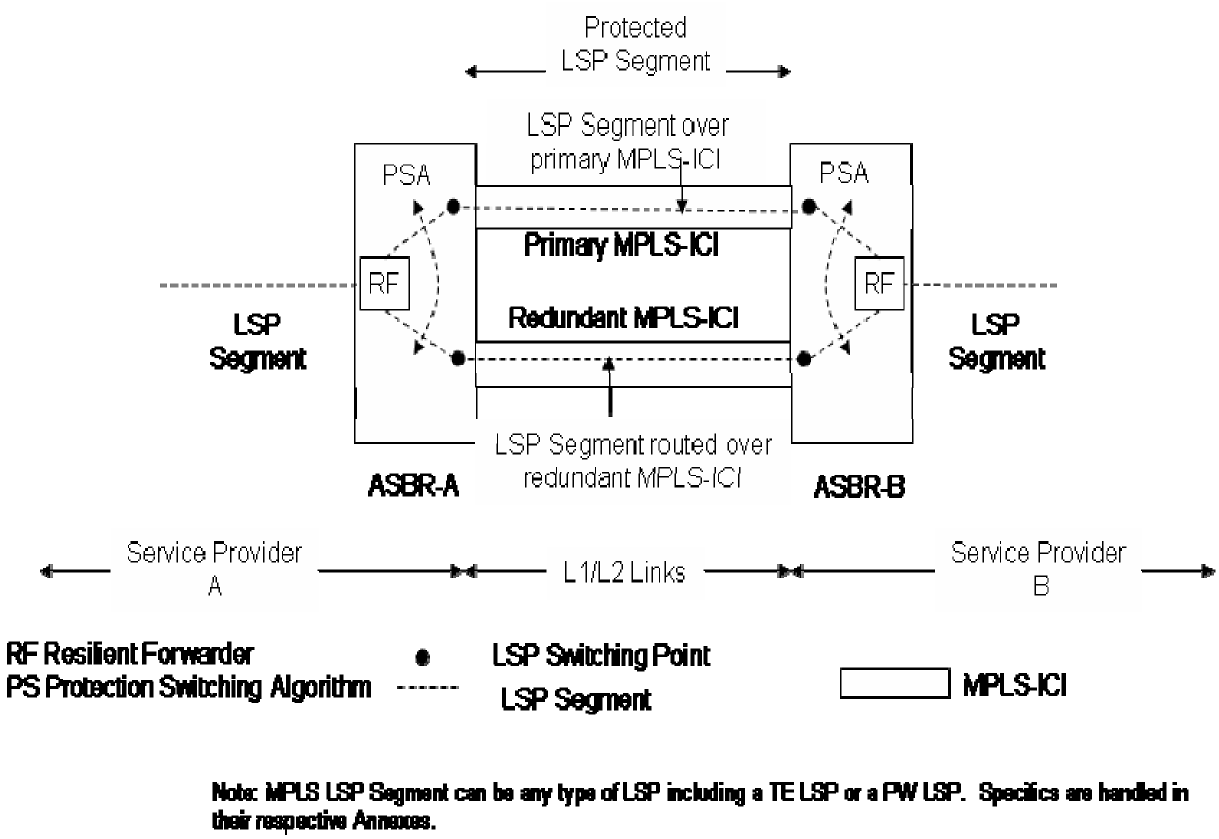


Figure 4: One-hop MPLS-ICI Protection

ASBRs compliant with this Technical Specification should support the following protection models:

- Multi-hop LSP protection over an MPLS-ICI as illustrated in Figure 5: In this case, an LSP segment is setup between two ASBRs over a direct primary MPLS-ICI interconnecting the two ASBRs. The LSP segment is also setup between the same ASBR pair across multiple hops by configuration, with a switching point connecting the segments (e.g., as shown in Figure 5). The multiple segment hops are ASBRs. ASBR-A2 and ASBR-B2 must be configured so that ASBR-A1 and ASBR-B1 use the same LSP label for both the primary MPLS ICI and the Multi-hop redundant path. Multi-hop protection may be required when two ASBRs do not have more than one direct MPLS-ICI interconnecting them, or when none of the existing parallel direct MPLS-ICIs can satisfy the LSP constraints, if any. Upon failure detection and notification, the LSP packets are rerouted over the Multi-hop redundant path.

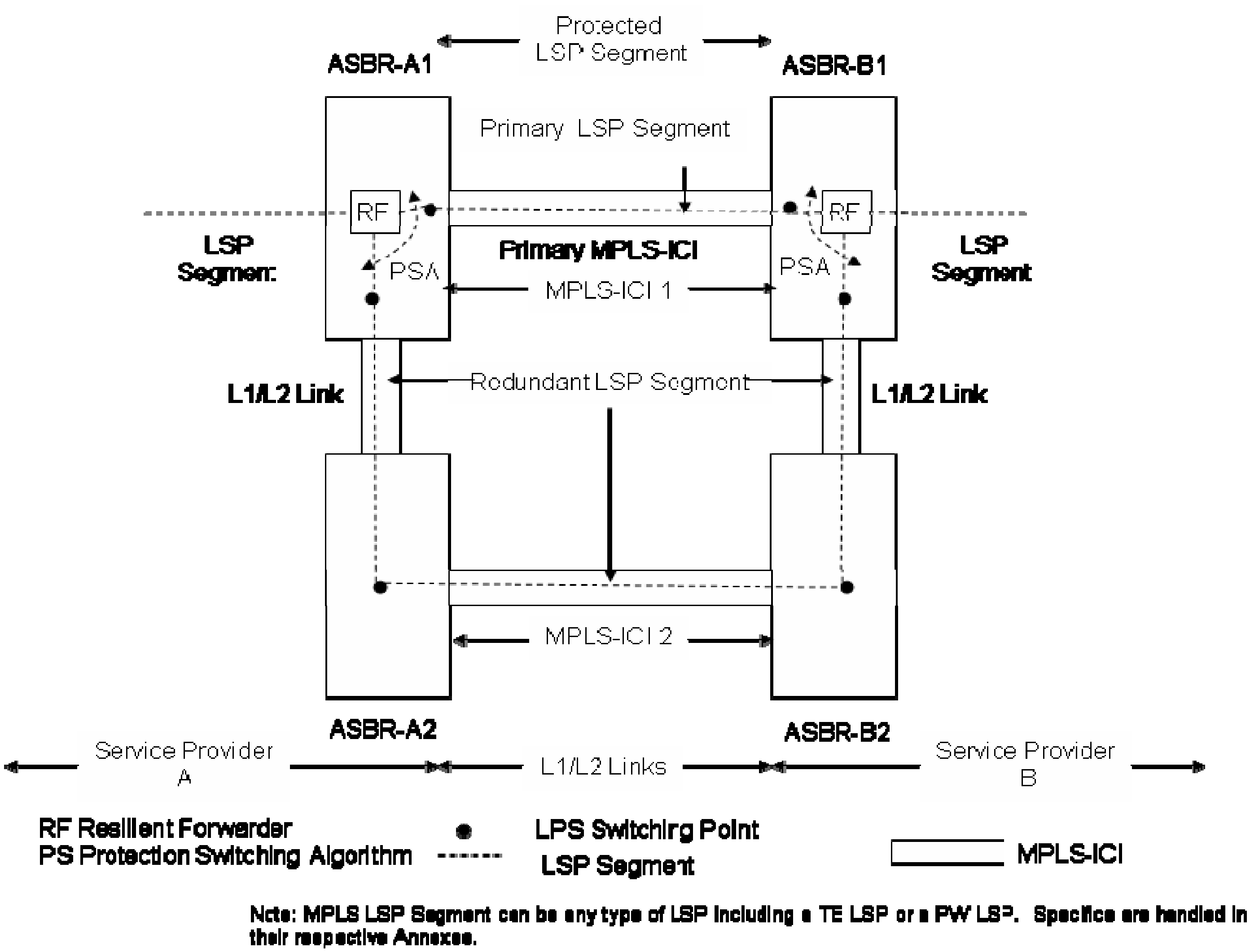


Figure 5: Multi-Hop LSP protection over an MPLS-ICI

- Multi-hop tunnel bypass protection for the primary MPLS-ICI as illustrated in Figure 6: An LSP bypass tunnel is configured between two ASBRs to protect one or more MPLS-ICIs interconnecting the same ASBR pair. The LSP bypass tunnel is routed over a L1/L2 path which is diverse from the primary MPLS-ICI and may span one or more other ASBRs but still provide connectivity between the same ASBR pair. The two ASBRs that initiate and terminate the LSP segment are configured to associate that LSP segment with the bypass tunnel for the purpose of protection. When the headend of the LSP segment receives an MPLS-ICI failure notification message, it will redirect the LSP packets to the bypass tunnel. The bypass tunnel LSP can carry the protection traffic of one or more LSP segments. The intermediate ASBR nodes only see the bypass tunnel and not the individual LSP segments tunneled through it, as illustrated by the tunnel in Figure 6. These intermediate ASBRs are not required to support individual LSP switching points along the path. In addition, the same MPLS label used on the primary MPLS-ICI for a given LSP segment is used for that LSP segment when it is tunneled over the bypass tunnel since the two ASBRs terminating that segment are still directly connected by the bypass tunnel. This capability should be configurable on a per LSP segment basis (i.e., not all LSP segments established over a protected MPLS-ICI are protected). When associating an LSP segment with a bypass tunnel, that LSP segment should be admitted to the bypass tunnel based on tunnel available resources and attributes and the LSP segment constraints. In addition, the bypass tunnel must be established before protection is triggered. One MPLS bypass tunnel may protect many LSP segments, but the bypass tunnel need only be established once before LSP segments can be rerouted over it.

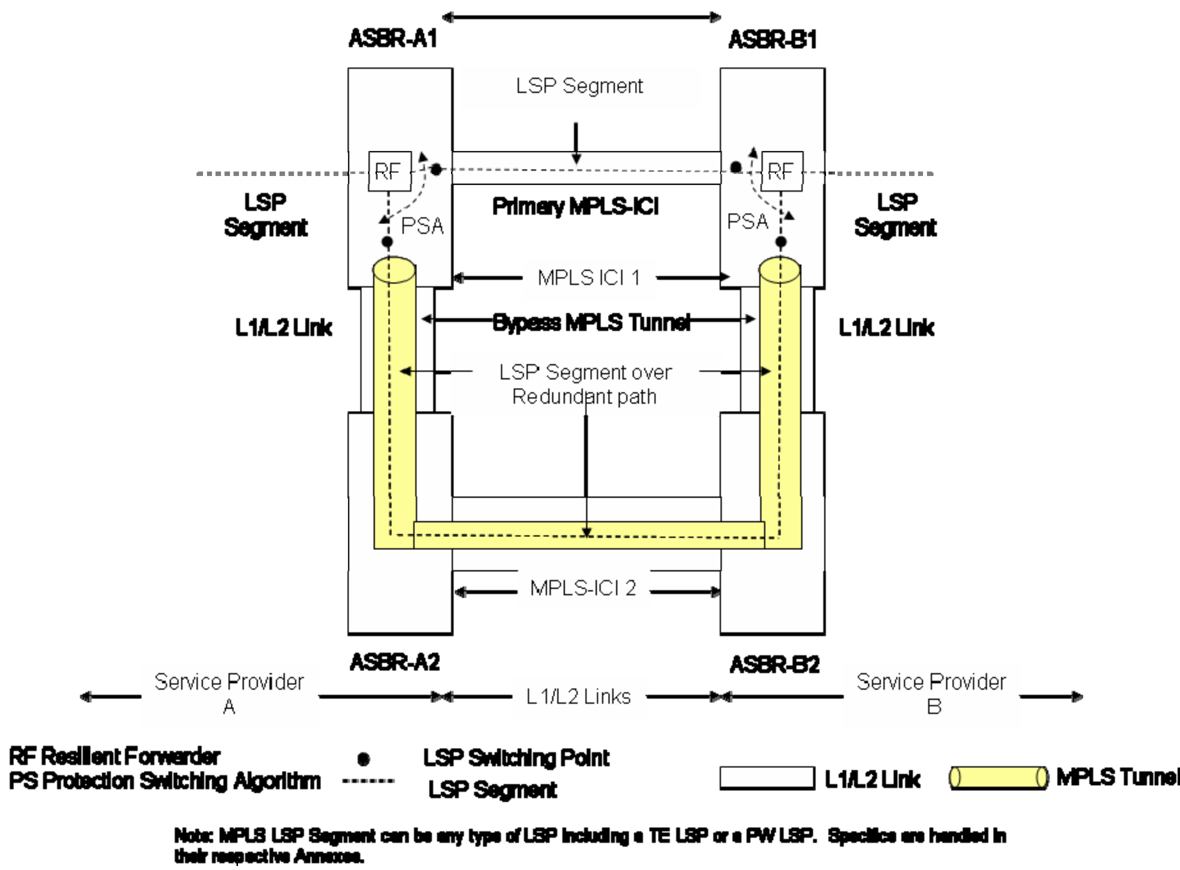


Figure 6: Multi-Hop bypass tunnel (Many:One) protection for the MPLS-ICI

- End-to-end LSP path protection: There may also be instances of end-to-end LSP protection across the MPLS-ICI where the protection switching functions does not reside in the ASBR, In this case, the only role that ASBRs can play in activating protection is failure detection and notification or failure notification relay. End-to end protection may be used simultaneously with other LSP segment protection mechanisms described earlier to obtain a more optimum end-end path while effecting fast reroute of the traffic at the point of failure. However it may be necessary to implement appropriate holdoff timers or similar measures for the end-to-end protection instances when operating in conjunction with other forms of protection configured on intermediate segments.

9.1.1.2. Failure Detection and Notification

For all statically configured LSPs, failure detection across an MPLS-ICI can be accomplished by detecting the loss of the physical layer signal, Layer 1 or Layer 2 OAM, or by utilizing BFD [IP-BFD]. ASBR's must support one-hop IP-BFD at minimum as a mechanism for unidirectional and birectional MPLS-ICI failure detection. When a failure is detected, local protection may be triggered, if configured, as discussed in the previous section. For the One-hop and multi-hop MPLS ICI protection measures, described in the previous section, and for the case of pseudowires, an ASBR may send a notification message to the preceding segment of an LSP using either BFD or LDP Status messages. More details for the pseudowire case are provided in Annex C. Alternatively, end-end OAM continuity check or loopback tests may be continuously run to detect failures on the LSP path that could not be locally corrected, but this is out of the out of the scope of this document. MPLS OAM mechanisms for LSPs, other than those composing MPLS PWs, are described in Section 12 while PW OAM mechanisms are decribed in Annex C.

9.1.1.3. Switchover

The LSP segment MPLS-ICI protection models described in Section 9.1.1.1 enploy unidirectional protection switching for the MPLS-ICI path to a standby LSP-segment path upon failure detection and notification for the MPLS-ICI as described in Section 9.1.1.2. Referring to Figure 4-Figure 6, upon reception of defect indication for the LSP segment on which traffic is currently being sent (defined in Section 9.1.1.2), the Resilient Forwarding Function in ASBR at the headend of the LSP-segment will switch the packet flow from the currently active LSP Segment path to the standby LSP segment path. An ASBR compliant with this Technical Specification shall be able to achieve switchover times within 100 msecs, after receipt of the failure notification, using the protection models described in this section.

9.1.2. MIBs

The ASBR should support the management information model consistent with mplsLsrModuleFullCompliance module of the LSR MIB as specified in [RFC3813] for static MPLS LSPs, which specifies read-write access. In addition, Read-write should be supported for the following groups:

- mplsInterfaceGroup
- mplsInSegmentGroup
- mplsOutSegmentGroup
- mplsXCGroup, mplsPerfGroup

Read-write should be supported for the following:

- mplsLabelStackGroup
- mplsLsrNotificationsGroup

The ASBR should support the management information model consistent with the read-only module compliance of the LSR MIB as specified in [RFC3813] for dynamic MPLS LSPs. An ASBR should provide Read-only support for the following groups: mplsInterfaceGroup, mplsInSegmentGroup, mplsOutSegmentGroup, mplsXCGroup, mplsPerfGroup. The ASBR should also support the management information model consistent with mplsLsrNotificationsGroup.

An informative description of these groups and their constituent objects is contained in Appendix IAppendix I.1.

9.2. Statically Configured and Signaled LSPs: PWs and TE-Tunnels

A statically-configured and signaled LSP is an LSP that is configured manually at both ASBRs but that is established via dynamic signaling between the ASBRs. As with the All-Static LSP, the statically-configured and signaled LSP, spanning the interconnect between two ASBRs, can be administratively or dynamically stitched to other LSPs in either domain. The advantage of the Static Signaled LSP is that ASBRs, rather than a human administrator, manage label assignments. The choice of the signaling protocol will depend on the service/application and/or the provider policy. Specifically, there are two service-dependent mechanisms that fit in this category: (1) targeted LDP for multi-segment Pseudo Wire (MS-PW) signaling discussed in Annex C, and (2) TE-tunnels, discussed in Annex D. Resiliency for this LSP establishment mechanism is covered in Annex C and Annex D.

9.3. Dynamically Established LSPs: Inter-AS BGP/MPLS IPVPN Option B, BGP labeled IPv4 paths, and TE-Tunnels

Dynamic LSPs across a domain boundary are those established without any static (manual) configuration at any intermediate point (e.g., ASBR) of the LSP. Implementations that support this Technical Specification shall provide for the dynamic establishment of LSPs via multi-protocol-BGP (MP-BGP) and TE-tunnels using RSVP-TE. The establishment of label-switched paths using MP-BGP for IPVPN and IPv4 routes is discussed in Annex A and Annex B, respectively. TE-tunnel establishment is discussed in Annex D. Resiliency for this LSP establishment mechanism is covered in Annex A, Annex B, and Annex D.

9.4. MIBS

The MIBs and management information for dynamically established LSPs in support of Pseudowires are described in Annex C. The MIBs and management information for dynamically established TE LSPs are described in Annex D.

10.Connection Admission Control

An ASBR compliant with this Technical Specification must be able to exercise admission control during the setup of an LSP. Connection admission control applies to all mechanisms of LSP establishment and must be based on policies that apply to a traffic profile per CoS and/or to other administrative policies. Admission control based on bandwidth constraints must be administratively enabled/disabled at the interface level. Other policies that affect admission control must be configurable and will vary per MPLS service. Admission control specific to MPLS services within the scope of this document is discussed within the respective annexes. ASBRs compliant with this Technical Specification must support the following generic capability for statically configured LSPs (all-static and statically-configured but signaled) with bandwidth and CoS constraints:

When an LSP with bandwidth and CoS requirements is configured between two ASBRs compliant with this Technical Specification, the ASBR must be capable of applying admission control to the LSP over one or more TE-tunnels or MPLS-ICI links between the ASBRs to accommodate the bandwidth requirement of the LSP per CoS. As a result, if the LSP is admitted, it is bound to the chosen link or TE tunnel. If the LSP cannot be admitted because of insufficient resources, the ASBR should keep the LSP down and send a trap to a network management element notifying it of the failure. If the LSP being setup is a segment of an end-end LSP that extends into either domain, the end-end LSP setup must fail when the LSP segment setup fails.

11.Forwarding

11.1. Traffic management

A class of service (CoS) is an aggregate for all traffic that shares similar QoS objectives and similar characteristics. We use the term CoS in this document interchangeably with service class, in the same manner that has been used in [RFC 4594]. The Diffserv architecture [RFC2475], MPLS Diffserv extensions [RFC3270] and Diffserv-aware MPLS traffic engineering extensions [RFC4124] defined by the IETF provide the architecture for supporting the transport and signaling of multiple classes of service [RFC4124][RFC3270]. However, they have not yet addressed most of the issues that arise when establishing inter-domain inter-provider MPLS connections.

In the forwarding plane, a CoS is associated with a Per Hop Behavior (PHB). PHBs as defined in [RFC2475] "are implemented in nodes by means of some buffer management and packet scheduling mechanisms". In MPLS networks that support Diffserv, the PHB that a packet should receive at a node can be inferred from the MPLS label and exp bits (L-LSPs) or from the EXP field value (E-LSPs) [RFC3270]. This document addresses E-LSPs only. For IP packets, as in multi-AS Backbone Option A, the PHB that a packet should receive is identified by the Diffserv Code Point (DSCP) [RFC2475].

11.1.1. Classes of Service and EXP Mappings in the Forwarding Plane

Providers may use different EXP-field values to correspond to similar classes of service. Providers may also implement different classes of service in their respective networks. In the data path, the CoS (for E-LSPs) and drop priority identifiers are encoded in the EXP field of the MPLS header of an MPLS packet. Thus, in the data path, the EXP field value used in one provider network may need to be mapped to another EXP field value used in the other provider network for packets to receive the same treatment across both networks or to perform CoS and drop priority mappings. When providers implement different classes of service, they may decide to also map one CoS in one network to another CoS in the other network. This CoS mapping is often done as an agreement between providers. In the forwarding plane, the agreed-on CoS mapping translates to a mapping between corresponding EXP field values whereby each EXP field value identifies a PHB. The desired externally observable behavior is that packets with exp field value are directed to a partuclar queue and drop priority and they leave the ASBR with a particular exp field value that may possibly be different from the one received at the ASBR.

In order to support interprovider operations as described earlier, ASBRs that support this Technical Specification must enable the configuration of EXP field value mapping between incoming EXP values and outgoing EXP values. A provider must be able to apply an EXP map at the ASBR ingress for traffic received over an MPLS-ICI and at the ASBR egress for traffic sent over an MPLS-ICI. If the mapping is applied to the egress of an interface, an ASBR performs the mapping and (re)writes the EXP field value in the outer label header and possibly the inner header (based on configuration) with the outgoing EXP field value prior to transmitting the packet to the next ASBR over the MPLS-ICI. If the mapping is applied at ingress, an ASBR performs the mapping when it receives the packet and (re)writes the EXP field value in the outer label header, and possibly inner label header, with the outgoing EXP field value when transmitting the packet within its network. Appendix II illustrates via examples how EXP mapping could be configured at ASBRs interconnected by an MPLS-ICI. ASBRs should be capable of also programming the exp to PHB map dynamically based on signaling a Diffserv-MPLS TE tunnel as defined in [RFC3270] section 5.2.1. When performing label popping or imposition, equipment must support the pipe model, short-pipe model and uniform models as defined in [RFC3270]. The model applied must be configurable per LSP and per interface.

An ASBR compliant with this Technical Specification must be able to associate an EXP field value with a queue and a buffer management profile associated with that queue before or after the EXP field mapping. A queue is served by a scheduling structure. A buffer management profile applied to a queue is used to decide whether to enqueue the packet or drop it based on the state of the queue, the length of the packet and the color of the packet [RFC2475].

11.1.2. Traffic Policing

A downstream provider may choose to enforce the traffic contract of an LSP, or an aggregate of LSPs by policing traffic forwarded by the upstream provider. ASBRs compliant with this Technical Specification must be able to apply a policer to traffic that belongs to an LSP, a group of LSPs, and/or a CoS identified by an EXP field value on an interface. The traffic contract itself can be derived from signaling and/or administrative configuration.

For policers, ASBRs must implement the one-rate-three-color-marker [RFC2697] and the two-rate-three-color-marker [RFC2698] metering algorithms in both color-blind and color-aware mode. The actions taken by the policers can be to pass the packet, drop the packet, or to increase its drop priority. Marking drop priority is done in the EXP field value of the MPLS header.

11.1.3. Traffic Shaping

An ASBR compliant with this section of the Technical Specification must support a single-rate shaper per EXP per interface or per IPVPN attachment circuit and/or IP Precedence/DSCP (as it applies to multi-AS Backbone Option A). Optinally should support a single-rate shaper per EXP per LSP or group of LSPs.

If a downstream node polices traffic per LSP, group of LSPs, and/or EXP, or per Attachment circuit and/or DSCP/Precedence, an upstream node that sends the traffic over the MPLS-ICI should be able to shape traffic to the traffic profile used in the metering and policing action. An ASBR compliant with this section of the Technical Specification must support a single-rate shaper per EXP and per interface or per IPVPN attachment circuit and/or IP Precedence/DSCP (as it applies to multi-AS Backbone Option A). An ASBR should support a single-rate shaper per EXP per LSP or group of LSPs. ASBRs should support dual-rate shapers. The ASBRs must enable/disable shaping on queues by configuration.

11.2. Path MTU Handling and Fragmentation

Maximum Transmission Unit (MTU) is the largest packet size that an interface can transmit without fragmentation. MTU is measured in Bytes (B) and it is media-dependent. It represents the largest payload that can be carried in a Layer-2 frame, and it does not include Layer-2 headers or trailers.

While MTU is defined with respect to interfaces, the end-to-end packet transport depends on the Path MTU (PMTU), which is defined as the smallest MTU along the packet path from the source to the destination. The packet sender is usually concerned with the PMTU, rather than the MTU of the attached link, because it is the PMTU that determines whether the packet will be delivered to the receiver and whether it will be fragmented in the process.

The PMTU is determined by the Path MTU Discovery (PMTUD) process, which is defined in [RFC 1191] for IPv4 (PMTDUv4). PMTUDv4 [RFC 1191] operates as follows:

- IP sender sets the "Don't Fragment" (DF) flag in the IP header.
- IP sender sends a datagram corresponding to the MTU of the first hop.
- When an ASBR is encountered with the next-hop link MTU too small to send the packet without fragmentation, it discards the packet (because DF is set) and sends to the IP sender an Internet Control Message Protocol (ICMP) message “Destination Unreachable” with a code meaning "Fragmentation Needed and DF Set."
- IP sender receives the ICMP message and sends a smaller packet.
- The process is repeated until a packet is small enough to get through without generating an ICMP message.
- This smallest packet becomes the PMTU used by the IP sender.

.

It should be noted that PMTUD could potentially be used to create a DOS attack on the control plane of the ASBRs that generate ICMP messages in response to packets that exceed the outgoing interface MTU and that cannot be fragmented.

11.2.1. Label stack depth

Considering that MPLS operates between Layer 2 and Layer 3, as far as Layer 2 is concerned, the MPLS label stack represents a part of the payload. Therefore, the MPLS label stack may change the status of the Layer 2 payload from “not exceeding the MTU size” to “exceeding the MTU size.” Thus, MTU scenarios for MPLS-ICI may further depend on whether a new header is provided for the MPLS-ICI tunnel, and whether this header represents an additional layer in the label stack:

1. contiguous tunnel – no new header for the ICI, the stack depth does not change; same MTU determination for the ICI as for the connected SP domains
2. stitched tunnel – new header for the ICI, but the stack depth does not change; same MTU determination for the ICI as for the connected SP domains
3. nested tunnel – new header for the ICI, the stack depth increases; labeled packets may exceed the MTU of the ICI.

11.2.2. Fitting packet sizes to PMTU

Different domains may implement different approaches to fitting packet sizes to the PMTUs, including the following:

1. sending packets that correspond to the minimum MTU (576 Bytes for IPv4 networks)

2. fragmenting packets (in any node along the path for IPv4)
3. using the Path MTU Discovery (PMTUD) process to determine the PMTU
4. using the RSVP-TE procedures defined in [RFC3209] Section 2.6 (Path MTU) when signaling an RSVP-TE LSP
5. using LSP trace [RFC4379] and the downstream mapping object included in echo reply

Approach-1 is wasteful of bandwidth if the PMTU is larger than 576 Bytes. Approach-2 is wasteful of network processing resources and creates delay that may be unacceptable for real time applications that send large packets, e.g., Video Conferencing. Approach-3 is most efficient compared to approach-1 and approach-2 from bandwidth and processing viewpoint. The only drawback of option-3 is that it could result in attacks on a router control plane if that router is not properly protected. Approach-4 is best suited when signaling an RSVP-TE LSP as this this approach does not require additional processing on routers processing the RSVP-TE messages. Approch-5 could be useful when using LSP ping trace mode for path trace. In addition, LSP ping trace-mode could be specifically used for this processing. However, as stated the OAM section, LSP trace could be construed as a potential DoS attack on routers.

11.2.3. MTU Requirements for MPLS-ICI

- An ASBR should support PMTU Discovery for IPv4.
- The PMTU Discovery process for MPLS tunnels must follow the process described in [RFC3032], Section 3.6.
- An ASBR compliant with this Technical Specification must be able to create, receive and pass “Destination Unreachable” ICMP messages with a code meaning "Fragmentation Needed and DF Set".
- Providers must be able to configure an option for silently dropping (i.e., without generating ICMP messages) packets whose lengths exceed the outgoing interface MTU.
- An ASBR compliant with this Technical Specification must be able to detect whether the payload of an MPLS packet is an IP packet or not in order to perform the right fragmentation procedure if fragmentation is permitted.
- An ASBR compliant with this Technical Specification must resist Denial of Service (DoS) attacks that use ICMP messages. In particular, an ASBR must be able to rate limit the number of generated ICMP messages when the packet forwader at the ASBR sends packets to an interface with a size that exceeds the interface MTU.
- An ASBR compliant with this Technical Specification should support a configuration parameter called “Maximum Initially Labeled IP Datagram Size) as described in RFC 3032, Section 3.2.
- An ASBR compliant with this Technical Specification must implement the procedures defined in RFC 3032, Section 3.4, for processing IPv4 datagrams that are too big.
- An ASBR compliant with this Technical Specification must support the MTU processing as required by Section 2.6 of RFC 3209 [RFC3209]
- An ASBR compliant with this Technical Specification must support MTU processing as required for LSP ping trace-mode [RFC4379]. In particular, the MTU returned by an ASBR in the echo reply must be the minimum MTU of the path to the next hop (and associated label map) returned by the ASBR.

[RFC4623] defines generic PW encapsulation and procedures that allow a PW packet fragmentation before insertion of the PW into the PSN at the ingress PE and the reassembly of the PW packet fragments at the egress PE. Similar procedures can be applied at an ASBR if the ASBR is a switching point (S-PE) of the PW (c.f., see [MS-PW_Requirements] for S-PE). An ASBR compliant with this specification should support [RFC4623]. In this case, an ASBR may segment and/or reassemble a PW packet. The choice of the outgoing MTU may be discovered or configured. In the discovery model, an ASBR may be able to use the path MTU discovery mechanisms to discover the path MTU of a PW segment that extends between two ASBRs for instance. If the PW is an IP PW, procedures described in [RFC3032] and discussed earlier in this section are applicable. A ASBR that complies with this specification must have the capability of turning off PW fragmentation if that capability exists and should provide the capability for allowing fragmentation/reassembly selectively for certain pseudowires but not others.

11.3. Load Balancing

A ASBR compliant with this Technical Specification must be able to load balance MPLS traffic across equal cost providers' paths, including interconnects. The ASBR must enable the operator to define the label stack depth used in load-balancing MPLS labeled packets. It must be capable of enabling/disabling load balancing and defining load balancing criteria when enabled. Implementations must ensure that packets from the same microflow do not get load-balanced across different paths when load balancing is enabled. For instance, packets from the same PW in a given direction must follow the same path. If load balancing involves MPLS labels only, this can be ensured. If the load balancing criterion includes IP information, the forwarding path must be able to detect whether the payload of an MPLS packet is IP [RFC4385] and apply load balancing only when applicable.

It should be noted that Equal Cost Multi Path (ECMP) at intermediate nodes as a matter of packet forwarding applies to non-reserved traffic, that is, traffic whose path is not reserved. ECMP also introduces challenges in tracing data paths and conducting performance measurements across all possible paths. These challenges are operational in nature.

11.4. TTL Processing

[RFC3443] supports multiple models for TTL processing in hierarchical MPLS networks: (1) uniform, (2) short pipe with and without Penultimate Hop Popping (PHP), and (3) pipe. An ASBR in conformance to this Technocal Specificaton must support all these models. Specifically, the pipe model and short-pipe model may be most useful for providers who like to hide the length of the path across their network when tunneling LSPs and/or IP packets through MPLS tunnels in their networks.

11.5. MIBs

The configuration and management of Diffserv-related traffic management components are specified in [RFC3289]. An informative description of how these components can work together is given in [RFC3290]. The Diffserv MIB defines a means to define a sequence of components on either the ingress or egress direction of an interface. The MIB defines some of these components as well as provides a generic row pointer mechanism to point to other MIBs that define the configuration and management of other components (e.g., an MPLS label and EXP bit classifier Table). The generic components defined in [RFC3289] are: classification, metering, action, and queuing. The ASBR should support the management information model consistent with the following groups with read-write access as specified in [RFC3289]:

- diffServMIBDataPathGroup
- diffServMIBClfrGroup
- diffServMIBClfrElementGroup
- diffServMIBMultiFieldClfrGroup
- diffServMIBActionGroup
- diffServMIBAlgDropGroup
- diffServMIBQGroup
- diffServMIBSchedulerGroup
- diffServMIBMaxRateGroup
- diffServMIBMinRateGroup
- diffServMIBCounterGroup

The ASBR should also support the management information model consistent with the following groups with read-write access as specified in [RFC3289]:

- diffServMIBMeterGroup
- diffServMIBTBParamGroup
- diffServMIBDscpMarkActGroup
- diffServMIBRandomDropGroup

An informative description of these objects, their indexing, content, suggested usage and context is contained in Appendix IAppendix I.2.

Since diffServMultiFieldClfrTable in [RFC3289] does not cover MPLS labels and EXP bits, the ASBR should support an enterprise MIB or equivalent information element that contains classification parameters that match certain patterns of the EXP bits and LSP IDs for an MPLS packet in the ingress and egress directions of an interface as specified in section 11.1.1. The parameters shall support matching on a set of MPLS labels and patterns of EXP bits. The DiffServClfrEntry associated with these MPLS labels and EXP parameters would then specify the next Diffserv component (e.g., a meter).

Since diffServDscpMarkActTable only (re)marks the DSCP value in the IP packet header, there is a need for an Enterprise MIB or equivalent information model that contains the value of the EXP bits markings to be applied in the MPLS packet header.

12.OAM

The MPLS OAM functionality can be split into two types: (1) ‘always-on’ defect detection and handling and (2) ‘on-demand’ diagnostics. OAM is vital to network operations and impacts provider Operational Support Systems (OSS). Due to its importance, ASBRs compliant with this Technical Specification must support both MPLS OAM functions: defect detection and handling and Diagnostics. The defect detection and handling function needs to be as simple as possible to minimize the processing cost. The diagnostics tools must include functions such as echo request/reply and Path Trace.

This section discusses OAM capabilities at an MPLS-ICI related to all of MPLS services described in the Annexes. OAM processing requirements in this section only apply to OAM packets that are destined to ASBRs interconnected by an MPLS-ICI. The focus of this section is OAM capabilities that apply to LSP segments established across the MPLS-ICI and to other segments of the same LSP that extend beyond the MPLS-ICI. End-End OAM is out of the scope of this document. The LSP segment established across the MPLS-ICI is a segment of an LSP that extends from PE to PE across ASes. This LSP segment is dynamically or statically established and stitched.

12.1. Connection Verification: "Always On" Defect Detection and handling

ASBRs compliant with this Technical Specification must support Bidirecional Forwarding Detection (BFD) [MPLS-BFD] in asynchronous mode. They must support the following:

- Timer Parameters:
    - time interval between successively transmitted protocol messages per LSP
    - minimum receive interval for protocol messages per LSP
    - failure detection criterion in terms of the number of successive messages that must be lost in case of BFD to trigger the declaration of an LSP down. This must be configurable per LSP.
- Failure notification
    - An ASBR should be able to send an SNMP trap upon LSP failure detection.
    - An ASBR must notify all client protocols that depend on the liveliness of the LSP being monitored when that LSP fails

An ASBR acting as an endpoint for an LSP must be able to avert OAM-related DOS attacks by:

- Droping the protocol messages received on an LSP if the protocol is not enabled for that LSP.
- Policing protocol messages per LSP basis to enforce the OAM message rate configured for that LSP
- Policing the aggregate of OAM protocol messages for all LSPs to a traffic profile. The traffic profile must be configurable by the Operator or dynamically generated and installed as sessions are enabled/disabled. The traffic profile should be set so that compliant messages are not inadvertently dropped.
- Supporting the configuration of per-LSP policing, aggregate policing or both for BFD protocol messages

It should be noted that policing BFD messages used for liveliness check may result in a false failure detection. Thus, the policing parameters must be set so that "legitimate" messages used for liveliness check are not impacted by policing unless they exceed their allocated rate.

ASBRs compliant with this Technical Specification must allow enabling an MPLS BFD session per LSP basis. When the ASBRs are transit points for an LSP, BFD running between the LSP endpoints is outside of the scope of this document. An LSP extended between ASBRs can be stitched to other LSP segments to form an end-end LSP. In that case, a BFD session that runs end-end between the LSP endpoints is transparent to the ASBRs. However, ASBRs must support the co-existenence of an end-end BFD session and a BFD session between the ASBRs for an LSP segment. MPLS packets carrying the BFD messages corresponding to the ASBR BFD sessions must have TTL set to 1 to force these messages to be processed at the ASBRs rather than being switched across.

ASBRs should support the authentication option for multihop and single-hop BFD sessions between two ASBRs as discussed in [BFD-Base]. Authentication must be enabled by configuration and the same key should be shared for all sessions between the same ASBR pair.

ASBR's compliant with this Technical Specification must support the bootstrapping method via MPLS ping for exchanging Your Discriminator and My Discriminator values used in BFD control messages. An ASBR compliant with this Technical Specification should also support the configuration of the local Discrimantor for a BFD session (My Discriminator value in the BFD messages the ASBR sends to the peer ASBR at the other end, and Your Discrminator value in the BFD messages it receives for the session), and the configuration of the peer ASBR Discrimantor for a BFD session (My Discriminator value in the BFD messages the ASBR receives from the peer ASBR at the other end, and Your Discrminator value in the BFD messages it sends for the session),

12.2. On Demand Diagnostics

ASBR's compliant with this Technical Specification must support LSP-PING [RFC4379] in both ping and trace modes to verify unidirectional connectivity and perform path tracing of MPLS label switched paths on inter-carrier interconnect segments, respectively. Equipment compliant with this Technical Specification must also support BFD in echo mode [MPLS-BFD] for performing loopback tests.

ASBR's compliant with this Technical Specification must support LSP-ping [RFC4379] in ping mode for checking the liveliness of the LSP and the data plane state against the control plane state for that LSP. This applies to LSPs with endpoints on the ASBRs at either end of an MPLS-ICI. An ASBR that is compliant with this Technical Specification and supports a service described in the annexes must support the associated LSP-ping FEC sub-TLV and their stackings as indicated in the following table:

Sub-type	Length	Field value	Application
1	5	LDP IPv4 prefix	LDP-PSN Tunnel
3	20	RSVP IPv4 LSP	TE-Tunnels
6	13	VPN IPv4 prefix	IPVPN
8	14	L2 VPN endpoint	Pseudo-wires
10	14	"FEC 128" Pseudowire	Pseudo wires
11	16	"FEC 129" Pseudowire	Pseudo Wires (Note)
12	5	BGP labeled IPv4 prefix	Labeled IPv4 routes

Note: If Annex C is supported.

LSP-ping must support the following reply modes specified in [RFC4379]:

Value	Meaning
1	Do not reply
2	Reply via an IPv4 UDP packet
4	Reply via application level control channel

LSP-ping messages for LSPs that do not terminate on an MPLS-ICI ASBR transit the MPLS-ICI. Corresponding LSP-ping replies could include replies with the router alert option. Such replies will cause every router on the path of the LSP to process the MPLS-ping messages. In order to prevent ping replies originated in one provider domain to be processed on every router in another provider domain on the path to the destination, the reply mode should not include the router alert option. An ASBR must be configurable to drop or rate-limit received echo reply packets with the router alert option to avert overloading or attacking the ASBR control plane and that of other routers within the ASBR AS. Another way for a network to avoid DoS attacks is to transparently pass the packets with the router alert option albeit this will also prevent processing other packets with the router alert option.

The following must also be supported for LSP ping in ping mode:

- Timer parameters:
    - time interval between sucessive echo requests per LSP or globally when initiating an LSP ping test
    - failure detection criterion in terms of the number of successively missed echo request replies that triggers the declaration of an LSP down. This must be settable per LSP.
- Failure notification
    - An ASBR should be able to send an SNMP trap upon LSP failure detection.
    - An ASBR should notify all client protocols that depend on the liveliness of the LSP being monitored when that LSP fails

An ASBR acting as an endpoint for an LSP must be able to avert OAM-related DoS attacks by:

- Droping the LSP ping messages received on an LSP if the protocol is not enabled for that LSP.
- Policing LSP ping echo requests to enforce the message rate configured for all LSP. Per LSP policing is optinal.

Trace mode capability of LSP-PING can be used for fault isolation. LSP path tracing enables the identification of the path(s) traversed by an LSP and hop-by-hop fault localization. ASBRs compliant with this Technical Specification must provide the capability to rate limit or drop LSP tracing messages arriving at an ASBR from another provider to be processed at the ASBR.When an ASBR drops an LSP ping message it will disrupt the end-end path trace. An ASBR should also support the option to respond at the domain boundary without including a downstream label map..

When responding to path trace messages from another provider ASBR, an ASBR must be configurable to respond with either of the following two options: (1) respond without a downstream label map to the next hop, (2) respond with a downstream label map to the next hop. Protection of confidential information when the path trace message is targeted to a router behind the ASBR within an AS is still an issue to be addressed. The detection of the LSP ping messages at an ASBR may require deep packet inspection and may not be feasible in certain cases. If a router inside an AS has knowledge that the LSP being traced is a cross-AS LSP it may drop the LSP ping echo request or respond to the LSP ping ping echo request without the downstream label map.

12.3. MIBS

The ASBR should support a management information model that provides for the configuration and management of BFD in a way that is consistent with [BFDMIB] compliance for the following groups:

- bfdSessionGroup,
- bfdSessionPerfGroup,
- bfdSessionPerfHCGroup,
- bfdNotificationGroup

An informative description of these objects, their indexing, content, suggested usage and context is contained in Appendix IAppendix I.3.

The information model in the BFD MIB contains objects that can be configured to cause a ASBR to generate a significant number of packets, potentially create a DOS attack on other routers or networks. Therefore, when the MIB is used the ASBR shall support only SNMPv3 for configuration of the BFD MIB. When only read operations are performed SNMPv2 is sufficient. When other network management protocols are used, a mechanism with a comparable level of security should be used.

13.Security and Confidentiality

This section discusses security capabilities that are important at the MPLS-ICI, and at devices (including ASBRs) which support the ICI. Security threats, capabilities, and practices that occur or may be deployed elsewhere in the network are outside of the scope of this section. More network-wide or general discussions of threats, capabilities, and practices can be found in [RFC4111] [RFC3871] [RFC4778]. The security capabilities stated in this section should be considered as complementary to security considerations addressed in the individual protocol specifications and/or security frameworks. This section discusses threats and capabilities at the InterCarrier-Interconnect related to all of the use cases described in the Annexes below.

Security vulnerabilities and exposures may be propagated across multiple networks because of security vulnerabilities arising in one peer’s network. Threats to security originate from accidental, administrative and intentional sources. Intentional threats include events such as spoofing and Denial of Service (DoS) attacks.

The level and nature of threats, as well as security and availability requirements, may vary over time and from network to network. This section therefore discusses capabilities that need to be available in equipment deployed for support of the MPLS-ICI. Whether any particular capability is used in any one specific instance of the ICI is up to the service providers managing the ASBRs offering/using the ICI services.

13.1. Control Plane Protection

This section discusses capabilities for control plane protection, including protection of routing, signaling, and OAM capabilities.

13.1.1. Authentication of Signaling Sessions

ASBRs compliant with this Technical Specification must support MD5 authentication for all TCP-based protocols within the scope of the MPLS-ICI (i.e., LDP signaling, and BGP routing) and MD5 authentication for the RSVP-TE Integrity Object to interoperate with current practices.

An ASBR compliant with this Technical Specification should be able to support exchange of all signaling and routing (LDP, RSVP-TE, and BGP) protocol messages over a single IPSec tunnel in tunnel or transport mode with authentication but with NULL encryption, between the peering ASBRs. IPSec, if supported, must be supported with HMAC-MD-5 and optionally SHA-1. It is expected that authentication algorithms will evolve over time and support can be updated as needed. An ASBR must support key configuration for an IPsec session.

OAM Operations across the MPLS-ICI could also be the source of security threats on the provider infrastructure as well as the service offered over the MPLS-ICI. A large volume of OAM messages could overwhelm the processing capabilities of an ASBR if the ASBR is not probably protected. Maliciously-generated OAM messages could also be used to bring down an otherwise healthy service (e.g., MPLS Pseudo Wire), and therefore effecting service security. MPLS-ping [RFC4379] does not support authentication today and that support should be subject for future considerations. Bidirectional Forwarding Detection (BFD) [BFD-Base] however, does have support for carrying an authentication object. It also supports Time-To-Live (TTL) processing as anti-replay measure. Implementations conformant to this MPLS-ICI should support BFD authentication using MD-5 and must support the procedures for TTL processing.

13.1.2. Protection against DoS attacks in the Control Plane

A ASBR compliant with this Technical Specification must provide the ability to filter signaling, routing, and OAM packets destined for itself, and must provide the ability to rate limit such packets. Packet filters should be capable of being separately applied per interface, and should have minimal or no performance impact. For example, this allows an operator to filter or rate-limit signaling, routing, and OAM messages that can be sent by a peer provider and limit such traffic to a traffic profile.

In the presence of a control plane DoS attack against an ASBR, the ASBR should guarantee sufficient resources to allow network operators to execute network management commands to take corrective action, such as turning on additional filters or disconnecting an interface which is under attack. DoS attacks on the control plane should not adversely affect data plane performance.

ASBRs which support BGP must support the ability to limit the number of BGP routes received from any particular peer. Furthermore, in the case of IPVPN, a ASBR must be able to limit the number of routes learned from a BGP peer per IPVPN. In the case that an ASBR has multiple BGP peers, it should be possible for the limit to vary between peers.

13.1.3. Protection against Malformed Packets

ASBRs compliant with this Technical Specification should be robust in the presence of malformed protocol packets. For example, malformed routing, signaling, and OAM packets should be treated in accordance to the relevant protocol specification.

13.1.4. Ability to Enable/Disable Specific Protocols

ASBRs compliant with this Technical Specification must allow an administrator to enable or disable a protocol (by default a protocol is disabled unless administratively enabled) on an interface basis. ASBRs must be able to drop any signaling or routing protocol messages when these messages are to be processed by the ASBR but the corresponding protocol is not enabled on the interface on which the messages were received. This dropping should not adversely affect data plane or control plane performance.

13.1.5. Protection against Incorrect Cross Connection

ASBRs compliant with this Technical Specification must support MPLS LSP Ping [RFC4379]. This may be used to verify end to end connectivity for the LSP (e.g., PW, TE Tunnel, VPN LSP, etc), and to verify PE to PE connectivity for L3 VPN services.

Routers (ASBRs and Route Reflectors) which support BGP operation must allow a means to restrict which Route Target attributes are sent to and accepted from a BGP peer across an ICI. Equipment (ASBRs, RRs) should also be able to inform the peer regarding which Route Target attributes it will accept from the peer [BGP-ORF]. This is due to the fact that a peer which sends an incorrect Route Target can result in incorrect cross-connection of VPNs. Also, sending inappropriate route targets to a peer may disclose confidential information. Further Security Consideration for inter-provider BGP/MPLS IPVPN operations are discussed in the IPVPN Annex.

13.1.6. Protection Against Spoofed Updates and Route Advertisements

ASBRs compliant with this Technical Specification must support signaling and routing authentication per section 13.1.1. They must also support route filtering of routes received via a BGP peering session by applying policies that include one or more the following: AS path, BGP next hop, standard community and/or extended community.

13.1.7. Protection of Confidential Information

ASBRs compliant with this Technical Specification should provide the ability to identify and prohibit messages that can reveal confidential information about network operation (e.g., performance OAM messages, LSP Traceroute messages). Service Providers must have the flexibility of handling these messages at the ASBR. For example, equipment supporting LSP Traceroute may limit which addresses replies can be sent to.

Note: This capability should be used with care. For example, if a service provider chooses to prohibit the exchange of LSP PING messages at the MPLS-ICI, it may make it more difficult to debug incorrect cross-connection of LSPs or other problems.

A provider may decide to progress these messages if they are incoming from a trusted provider and are targeted to specific agreed-on addresses. Another provider may decide to traffic police, reject or apply policies to these messages. Solutions must enable providers to control the information that is relayed to another provider about the path that an LSP takes. For example, in RSVP-TE record route object or MPLS-ping trace, a provider must be able to control the information contained in corresponding messages when sent to another provider.

13.2. Data Plane Protection

13.2.1. Protection against DoS in the Data Plane

This is provided via traffic policing as described in Section 11.1.2.

13.2.2. Protection against Label Spoofing

ASBRs compliant with this Technical Specification must be able to verify that a label received across an MPLS-ICI was actually assigned to an LSP arriving from the provider across that MPLS-ICI. If the label was not assigned to an LSP which arrives at this ASBR from the correct neighboring provider, the packet must be dropped. This verification can be applied to the top label only. The top label is the received top label and every label that is exposed by label popping to be used for forwarding decisions.

ASBRs compliant with this Technical Specification must provide the capability of dropping MPLS-labeled packets if all labels in the MPLS label stack are not processed at the ASBR. The presence of non-processed MPLS headers is detected if the the S-bit in the last MPLS header processed is set to 0. This behavior must be configurable on per interface basis. Enabling this behavior on an interface prevents some applications across that interface. However, when enabled, it provides a carrier the capability of guaranteeing that every label that enters its domain from another carrier was actually assigned to that carrier and avoids the possibility of potential security attack on a service within its domain.

Annex A IP VPN

This annex covers the IP-VPN MPLS service based on BGP/MPLS IP VPNs as defined in [RFC4364]. In particular, this annex discusses option `A' and option 'B' of the 'Multi-AS Backbones' section of [RFC4364].

Service providers can typically meet most enterprise requirements for IP-VPN services on their own IP-MPLS network or networks. However, in many cases, a single IP-VPN provider may not have the footprint necessary to serve all sites of an enterprise IP-VPN customer. Thus, providers find themselves forming partnerships with other providers, the nature of which depends for example on mutual business benefit or by an enterprise demanding that a certain service provider be included in a specific solution.

Multi-AS Option A, 'VRF-to-VRF connections at the AS (Autonomous System) border routers' is described in section 10 item (a) of [RFC4364]. Multi-AS option A requires that unlabeled IPv4 unicast addresses be exchanged between the ASBRs and that forwarding be based on information in the IP header of the received IP datagram. Customer DSCP/IP precedence preservation is especially important when the two interconnected providers's networks use different DSCP values to indicate the same Per-Hop-Behaviors (PHBs) or use different sets of PHBs.

Multi-AS Option B, 'EBGP redistribution of labelled VPN-IPv4 routes from AS to neighboring AS' is described in section 10 item (b) of RFC [RFC4364]. In the forwarding plane for Multi-AS Option B, an ASBR forwards packets based on the MPLS label-value associated with each VPN-IPv4 route.

The remainder of this Annex specifies policy features for addressing operational requirements of the MPLS-ICI without changing the procedures of [RFC4364].

A.1 Routing

BGP is specified in [RFC 4364] as the routing protocol used across service providers boundaries, e.g. MPLS-ICI This section discusses specific additional policies and mechanisms that must be supported. across the MPLS-ICI.

A.1.1 Route Target Allocations – Multi-AS Option B

An ASBR, in compliance with this Technical Specification must provide an SP with the ability to define import/export policies between two (MP)-BGP peers that apply to the Route Target (RT) community attribute(s) of labeled VPN-IPv4 routes being exchanged over the corresponding (MP)-BGP session. In addition to supporting the policies defined in [RFC4382], the following policy must be supported:

An SP must be able to export labeled VPN-IPv4 routes that match specified RT values, but with possibly different RT(s) specified as part of the policy. This RT mapping function should be capable of being performed on a one-to-one, one-to-many, many-to-one, or many to many basis. An example of a 'many to many' RT mapping policy could be where a single prefix has three RT values assigned. The ASBR will match all three RT values and replace them with three new RT values..

Labeled VPN-IPv4 routes are identified by amongst others, the Route Target (RT) extended community attribute. In [RFC4364] Multi-AS option ‘B’, labeled VPN-IPv4 routes along with associated RTs are exchanged between service provider operated domains. Service providers however usually have difficulty in provisioning RT values that are assigned by external parties on their PE network elements. This difficulty is due to the provider having automated or manual IP-VPN service provisioning systems and operational processes that apply to single provider domains only. In order to simplify operation, an ASBR, in compliance with this Technical Specification must be capable of mapping an incoming RT value to an outgoing RT value as specified earlier based on import/export policies.

An ASBR may receive a labeled VPN-IPv4 route with associated RT(s) from both internal and external peers and should be capable of re-advertising the route with one or more mapped RT value(s). This mapping function should be supported with labeled VPN-IPv4 routes received from both an MPLS-ICI and an internal peer.

A.1.2 Route Distinguisher Allocations – Multi-AS Option B

An ASBR in compliance with this Technical Specification must provide an SP the capability to configure the following policies that apply to the RD part of a labeled VPN-IPv4 route:

- Reject routes whose RD matches specific values.
- Reject routes whose RDs match a specified type, administrator subfield, or assigned number subfield.

In addition, an ASBR ahould provide an SP the capability to configure an RD rewrite policy as follows:

- Upon exporting a route, replace the RD part of the route with an RD value specified in the policy. The policy should allow replacing the full 8-byte RD value, the administrator subfield and/or the assigned number subfield.

In [RFC4364] Multi-AS Option `B’, labeled VPN-IPv4 routes are exchanged among ASs via (MP)-BGP [RFC4364]. A VPN-IPv4 address is composed of an 8-byte Route Distinguisher (RD) pre-pended to a 4-byte IPv4 address. An RD pre-pended to an IPv4 address must uniquely define the VPN-IPv4 address. When labeled VPN-IPv4 routes are exchanged across provider domain boundaries, VPN-IPv4 addresses must not overlap. In order to ensure that RD values that transit the MPLS-ICI are unique and do not collide with those used inside the AS of the receiving ASBR, RD values must in principle contain a registered AS Number (ASN) [RFC4364]. To enforce that, an ASBR in compliance with this Technical Specification that receives VPN-IPv4 routes from another provider-AS must be capable of rejecting routes whose RD or part of their RD matches a provider-specified value. For instance, an SP may want to reject routes that it receives from a peering SP if their administrator subfield value matches its own ASN.

An ASBR must in addition be capable of re-writing the RD value within a VPN-IPv4 address based on SP specified policies applicable to both received and advertised routes. When a public ASN is used as part of the encoding of type 0 or type 2 RD values [RFC4364], RD rewrite should only involve changes to the 2 or 4 byte administrator subfield, leaving the assigned number subfield unchanged. This assignment policy should be carefully coordinated amongst service providers for VPNs that will transit AS borders.

RD rewrite may be invoked to provide for: (1) security and confidentiality, (2) simplified management of the Inter-provider IPVPN, (3) diversified routes through alternative ASBRs and potentially load-balancing when Equal Cost Multipaths exist, or (4) enabling traffic routing across different ASBRs to the destination based on geographic location. This rewrite function ensures that for security purposes, VPN-IPv4 addresses are unique to each service provider domain and that they do not reveal any numbering scheme used within the provider network. In addition, RD rewrite provides for PEs within an SP domain to learn reachability to the same IP VPN destination in the peering provider domain or beyond through different ASBRs that have (MP)-BGP peering sessions with that peering provider. A PE may select an ASBR as the exit point to a peering provider based on the cost to reach the ASBR, or may load-balance traffic across multiple ASBRs when the paths through these ASBRs have equal costs.

A.1.3 Route Target Advertisement – Multi-AS Option B

The ASBR should be capable of restricting RT values to only be advertised to and received from interested parties across interconnects. This is particularly relevant where the ASBR is peered with more than one other ASBR. [RFC4684] describes this function across ASes. In order to hide the intra-domain IP-VPN topology from an external service provider, RTs advertised across the MPLS-ICI should consist of different values to the RT's provisioned within a service provider's domain.

A.1.4 VPN-IPv4 Route Summarization

An ASBR must be capable of summarising VPN-IPv4 routes learned from or advertised across the MPLS-ICI boundary on a per VPN basis [RFC4364]. Per VPN route summarisation allows existing operational processes for reducing numbers of customer routes from being advertised across a single provider's network to apply at the AS border. In Multi-AS option 'A' of [RFC4364], route summarization must be configurable at an ASBR on a per BGP session basis in the context of a VRF. For Multi-AS option 'B' of [RFC4364], route summarization must be configurable at an ASBR on a per VPN basis. In the case of Option B, summarization should be practiced with caution as an advertised label across the MPLS-ICI must map to a label associated with the PE that advertised the summarized VPN-IPv4 routes within the advertising ASBR AS.

A.2 Label Value Acceptance – Multi-AS Option B

For Inter-provider IP-VPN based on [RFC4364] Multi-AS option 'B', (MP)-BGP is used to exchange labels for VPN-IPv4 routes. The LSP segment established via BGP across the MPLS-ICI is a segment of an LSP that extends from PE to PE across ASes. As per paragraph 2 of Section 10 (b) of RFC 4364, "An ASBR should never accept a labeled packet from an EBGP peer unless it has actually distributed the top label to that peer". However, the VPN-IPv4 label may not be the top label on the received labeled packet as the packet may be tunneled by another LSP (e.g., TE-LSP) and the VPN label is the bottom label in the stack. An ASBR must never accept a labeled packet from an EBGP peer unless it has actually distributed all labels exposed at the ASBR to the peer. These include the top arriving label and every label exposed after popping a label. An ASBR should log the number of packets discarded as a result of this processing for diagnostic purposes.

A.3 DSCP/Precedence-EXP Mapping

An ASBR, that supports Multi-AS Backbone Option A as specified in [RFC4364] must enable the following:

- Type 1 map: Configuration of a map between incoming IP Diffserv Code Point values (DSCP) and outgoing EXP field values, and the association of that map with an MPLS-ICI or a logical channel over the MPLS-ICI (i.e., an IPVPN attachement circuit [RFC4364]). Upon receiving an IP packet over an MPLS-ICI, an ASBR must be able to classify the IP packet based on DSCP and determine the corresponding EXP value according to the configured map. The ASBR must be able to write the outgoing EXP value in the outer MPLS header and possibly inner header (based on configuration) when sending the resulting MPLS labeled packet on one of its core interfaces

- Type 2 map: Configuration of a map between incoming IP Precedence values and outgoing EXP field values, and the association of that map with an MPLS-ICI or a logical channel over the MPLS-ICI (i.e., an IPVPN attachement circuit [RFC4364]). Upon receiving an IP packet over an MPLS-ICI, an ASBR must be able to classify the IP packet based on Precedence bits and determine the corresponding outgoing EXP value according to the configured map. The ASBR must be able to write that EXP value in the outer MPLS header and possibly inner header (based on configuration) when sending the resulting MPLS labeled packet on one of its core interfaces.

- Type 3 map (applies to Ethernet 802.1q interfaces): Configuration of a map between incoming EXP field value to an ASBR and priority bits in the 802.1q VLAN Tag and the association of that map with an MPLS-ICI or a logical channel over that MPLS-ICI (e.g., the VLAN ID that acts as an attachment circuit identifier to a VRF [RFC4364]). Such map may be simply to copy the EXP value to the 3 priority bits in the VLAN tag. An ASBR must be able to apply that map and write the priority bits in the VLAN tag upon encapsulating the IP packet in an 802.1q Ethernet frame prior to transmission to the next ASBR across the MPLS-ICI.

- Type 4 map (applies to Ethernet 802.1q interfaces): Configuration of a map between incoming priority bits carried in 802.1q attachement circuit VLAN tag and outgoing EXP value and the association of that map with an MPLS-ICI or a logical channel (attachment circuit) over that MPLS-ICI. Upon receiving an IP packet encapsulated in a 802.1 Ethernet frame, the ASBR must be able to apply that map, and accordingly write the EXP value in the MPLS label header(s) it imposes on the IP packet prior to transmission towards its AS-bound core interfaces.

An ASBR, that supports Multi-AS Backbone Option A as specified in [RFC4364] should enable one or more of the following:

- Type 5 map: Configuration of a map between incoming EXP field values and outgoing DSCP values, and the association of that map with an MPLS-ICI or a logical channel over the MPLS-ICI (i.e., an IPVPN attachement circuit [RFC4364]). Upon receiving an IP packet over an interface, an ASBR must be able to classify the received MPLS packet based on outer label EXP field value and determine the corresponding outgoing DSCP value according to the configured map. The ASBR must be able to re-write that DSCP value in the IP header when sending the packet over the MPLS-ICI. This map type does not maintain transparency to customer IP DSCP marking, often an undesirable behavior.

- Type 6 map: Configuration of a map between incoming EXP field values and outgoing precednce values, and the association of that map with an MPLS-ICI. Upon receiving an IP packet over an interface, an ASBR must be able to classify the received MPLS packet based on outer label EXP field value and determine the corresponding outgoing precedence value according to the configured map. The ASBR must be able to re-write that precedence value in the IP header when sending the IP packet over the MPLS-ICI. This map type does not maintain transparency to customer IP precedence marking.

- Type 7 map: Configuration of a map between incoming EXP field values plus DSCP value and outgoing DSCP values, and the association of that map with an MPLS-ICI or a logical channel over the MPLS-ICI (i.e., an IPVPN attachement circuit [RFC4364]). Upon receiving an IP packet over a core interface (An AS-internal link), an ASBR must be able to classify the received MPLS packet based on outer label EXP field value + DSCP of the encapsulated IP packet and determine the corresponding outgoing DSCP value according to the configured map. The ASBR must be able to re-write that DSCP value in the IP header when sending the IP packet over the MPLS-ICI. This map type does not maintain transparency to customer IP DSCP marking.

- Type 8 map: Configuration of a map between incoming EXP field values + IP Precedence value and outgoing precedence values, and the association of that map with an MPLS-ICI a logical channel over the MPLS-ICI (i.e., an IPVPN attachement circuit [RFC4364]). Upon receiving an IP packet over a core interface (An AS-internal link), an ASBR must be able to classify the received MPLS packet based on outer label EXP field value + IP Precedence of the encapsulated IP packet and determine the corresponding outgoing precedence value according to the configured map. The ASBR must be able to re-write that precedence value in the IP header when sending the IP packet over the MPLS-ICI. This map type does not maintain transparency to customer IP precedence marking.

- Type 9 map: Configuration of a map between incoming EXP field values + IP Precedence value and outgoing DSCP values, and the association of that map with an MPLS-ICI or a logical channel over the MPLS-ICI (i.e., an IPVPN attachement circuit [RFC4364]). Upon receiving an IP packet over a core interface (An AS-internal link), an ASBR must be able to classify the received MPLS packet based on outer label EXP field value + IP Precedence of the encapsulated IP packet and determine the corresponding outgoing DSCP value according to the configured map. The ASBR must be able to re-write the DSCP value in the IP header when sending the IP packet over the MPLS-ICI. Specifically, when writing the DSCP value, the lower 3 bits, corresponding to IP precedence must remain untouched, This map type maintains transparency to customer IP precedence marking if the ASBR on the other end implements Type 8 map.

- Type 10 map: Configuration of a map between incoming DSCP value and outgoing precedence value and EXP, and the association of that map with an MPLS-ICI or a logical channel over the MPLS-ICI (i.e., an IPVPN attachement circuit [RFC4364]). Upon receiving an IP packet over an MPLS-ICI, an ASBR must be able to classify the received IP packet based on DSCP and determine the corresponding outgoing IP precedence value according to the configured map. The ASBR must be able to re-write the DSCP value in the IP header with the leftmost 3 bits preserved in the DSCP field as received and the rightmost three bits re-written with 0. This map type, cojointly with Type 7 map apply at the other end of the MPLS-ICI, maintains transparency to customer IP precedence marking.

A.4 Resiliency

For both multi-AS Option A and Option B, an ASBR should support the configuration of a multi-hop BGP session with another ASBR between corresponding loopback addresses. When an ASBR interconnects to one or more ASBRs via MPLS-ICIs, failure of one MPLS-ICI should trigger fast traffic re-direction to the surviving interfaces through which the destination is reachable. Implementations should target reroute times on the order of 50-100 mseconds or better.

When an ASBR peers with two or more ASBRs and an IPv4 route (Multi-AS Option A) or a VPN-IPv4 route (Multi-AS Option B) is learned via two or more of its ASBR neighbors, the failure of one ASBR should result in fast traffic re-direction to the surviving ASBR(s). For Multi-AS Option B, ASBRs must allow the tunneling of VPN-IPv4 labels via MPLS-TE tunnels protected by MPLS fast reroute for link protection.

For Multi-AS Option B, an ASBR compliant with this Technical Specification must support the rewrite of the RD part of a VPN-IPv4 route received across the MPLS-ICI as described earlier. This allows reachability to an IPv4 destination in a VPN to be learned through two different ASBRs providing for faster convergence. In addition, an ASBR compliant with this Technical Specification must support MP-BGP graceful restart [RFC4724] [RFC4781] for the VPN-IPv4 AFI/SAFI and IPv4 in the case of multi-AS Option A for the data plane to continue forwarding in the presence of control plane failure. It should be noted that the validity of these routes is dependent on the routes being reachable through the respective AS of the ASBR. Proper operation requires that the ASBR performing graceful restart for eBGP also performs graceful restart for the IGP on the AS-bound interfaces as well as graceful restart for the MPLS protocols (e.g., LDP and/or RSVP-TE) that provide for tunneling the IP VPN packets to the PE connected to the customer site or the far-end ASBR connected to another AS. If the ASBR supports a non-stop control plane (i.e., maintain routing state and signaling state on an active control card and a standby control card, and resume signaling and routing after control card switchover without interruption), control plane switchover will not be noticed by the rest of the network. In that case, the ASBR need only implement graceful restart in helper mode. Otherwise, the ASBR must implement graceful restart in both helper and restart modes.

A.5 Admission Control Policy: VPN-IPv4 Route Learning Restriction Requirement

An ASBR receiving IP-VPN traffic must be capable of restricting the number of VPN-IPv4 routes learned on a per VPN basis when supporting [RFC4364] Multi-AS option 'A' and option 'B'. These VPN-IPv4 routes could be learned from the service provider's own network or from other service provider domains. Per VPN route capping allows existing operational processes for restricting the number of customer routes from being advertised across a single provider network to apply at the AS border. In addition, an ASBR must provide the capability to limit or cap the number of unique routes learnt per (MP)-BGP session or a set of BGP sessions per VPN.

For Multi-AS Option A, an ASBR must provide the capability of limiting the number of IPv4 routes per VRF on the ASBR. In addition, an ASBR must provide the capability to limit the number of unique routes learnt per BGP session and a peering group that applies to multiple BGP sessions potentially with the same peer SP on the ABR.

For Multi-AS Option B, an ASBR can identify the routes belonging to a VPN based on RT(s). An ASBR must support policies that cap the number of unique VPN-IPv4 routes learnt per VPN on a BGP session basis and peering group basis.

The configuration of capabilities discussed in this section must be left to the discretion of the SP operating the ASBR. Specifically, an SP may choose to configure route caps per VPN, BGP session, or peering group or impose no caps.

A.6 MIBs

The ASBR should support the management information model consistent with the BGP-4 MIB objects defined in [RFC4273],

The ASBR should support a management information model consistent with the full compliance of the following MIB groups as specified in [RFC4382]:

- mplsL3VpnScalarGroup
- mplsL3VpnVrfGroup
- mplsL3VpnIfGroup
- mplsL3VpnPerfGroup
- mplsL3VpnVrfRteGroup
- mplsL3VpnVrfRTGroup
- mplsL3VpnSecGroup
- mplsL3VpnNotificationGroup

An informative description of these objects, their indexing, content, suggested usage and context is contained in Appendix I.4.

The following BGP MPLS VPN configuration and management functions are not specified in [RFC4382]. The ASBR should support an informantion element in the form of enterprise MIB or equivalent to implement the following functions:

The ASBR shall be configurable on a per peer basis with the maximum number of routes that can be received from a peer. The configuration shall allow the user to specify the action taken in the event this threshold is crossed as either 1) sending an alarm message, 2) dropping routes greater than the configured value, or both of these actions.

A.7 BGP/MPLS IPv4 VPN – Multi-AS Option A (non-normative)

(This section does not form an integral part of this specification)

An Inter-Provider IP-VPN service is depicted in Figure 8 using [RFC4364] Multi-AS OptionA. In Option A, described as ‘VRF-to-VRF connections at the AS (Autonomous System) border’, one ASBR in one AS attaches to an ASBR in the other AS via one or more attachment circuits. There is at least one attachment circuit for each IPVPN whose routes are exchanged via these two ASBRs. Each such circuit is an IP sub-interface to an IPVPN VRF configured on the corresponding ASBR IPVPN Forwarding decisions at the ASBRs are based on IP destination address lookup within the corresponding IPVPN VRF forwarding table. In Multi-AS option A, non-labeled IP packets are exchanged between the ABRs across an MPLS-ICI.

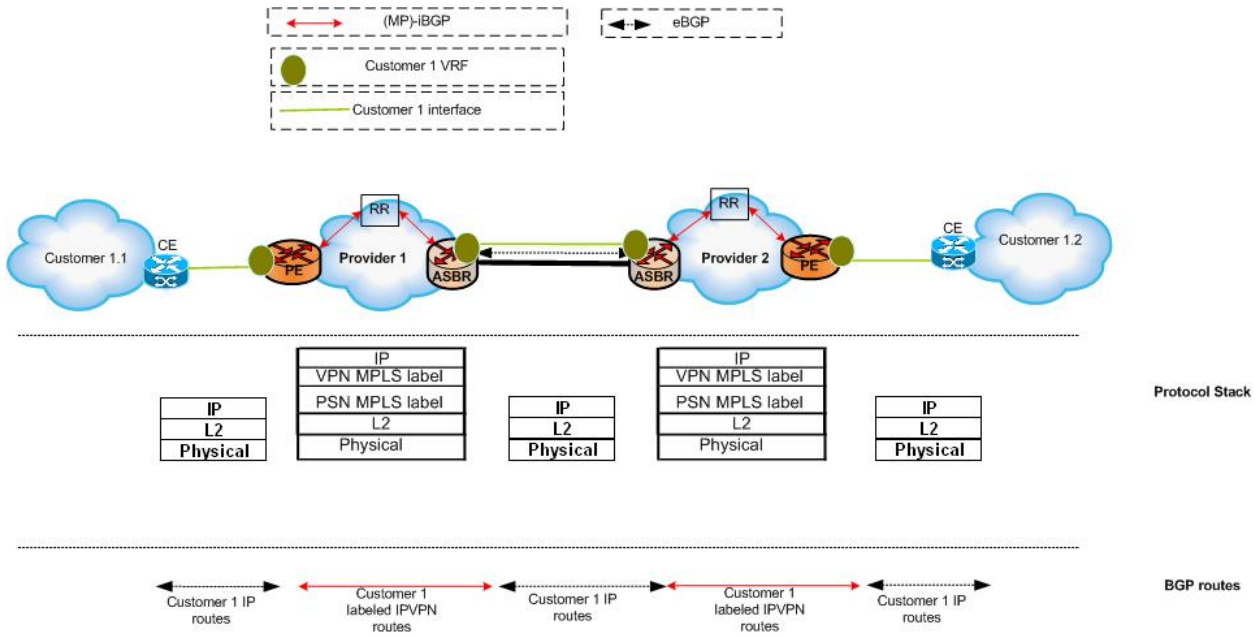


Figure 7: Inter-provider IPVPN service using RFC 4364 Multi-AS Option A

A.8 BGP/MPLS IPv4 VPN – Multi-AS Option B (non-normative)

(This section does not form an integral part of this specification)

Service providers can typically meet most enterprise requirements for IP-VPN services on their own IP-MPLS network or networks. However, in many cases, a single IP-VPN provider may not have the footprint necessary to serve all sites of an enterprise IP-VPN customer. Thus, providers find themselves forming partnerships with other providers, the nature of which depends either on mutual business benefit or by an enterprise demanding that a certain service provider be included in a specific solution.

An Inter-Provider IP-VPN service is depicted in Figure 8 using [RFC4364] Multi-AS Option B. In Multi-AS Option B, described as 'EBGP redistribution of labelled VPN-IPv4 routes from AS to neighboring AS', each ASBR receives labeled VPN-IPv4 routes from within its domain and redistributes these labeled routes to connected ASBRs with which it maintains e-BGP peering sessions supporting labeled VPN-IPv4 routes. . When redistributing VPN-IPv4 routes, the ASBR assigns MPLS label values to the routes from its own label space(s). In addition, the ASBR may modify other attributes associated with these routes. In the forwarding plane, an ASBR forwards packets based on the MPLS label-value associated with each VPN-IPv4 route.

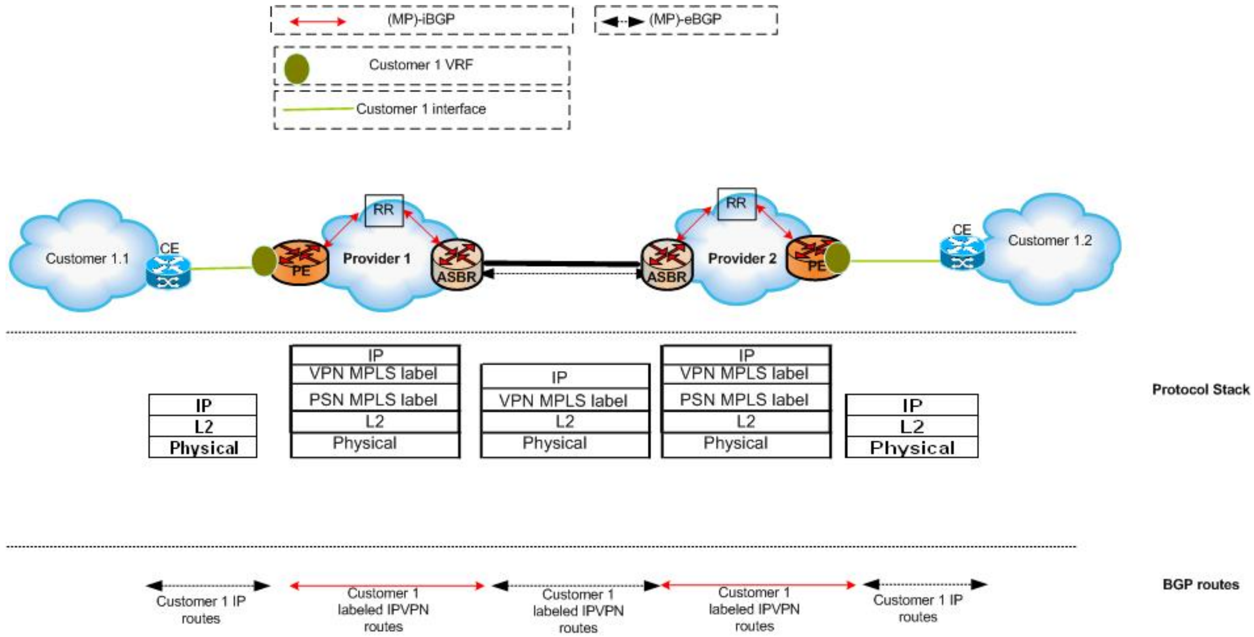


Figure 8: Inter-provider IP-VPN service using RFC 4364 Multi-AS Option B.

A.9 Interconnect Performance Measurement (non-normative)

(This section does not form an integral part of this specification)

Service providers need to measure performance metrics such as round-trip delay, one-way delay, delay variation and packet loss in order to provide for SLA guarantees across an MPLS-ICI. In the case of [RFC4364] Multi-AS options 'A' and 'B', these performance metrics can be to either designated host addresses within customer VPNs or to an ASBR address for ASBR to ASBR MPLS-ICI data plane performance measurements. Service providers need to have the ability to host measurement probes, as well as carry and report performance measurements across an MPLS-ICI. Guidelines and methodologies for performing QoS measurements and budgeting for impairments across multiple provider domains can be found in [Y.1541] amd [MIT_WP].

Annex B Labeled IPv4 Routes

Implementations that conform to this specification must support the distribution of Multiprotocol Label Switching (MPLS) label for IPv4 routes across an MPLS-ICI using Multi-Protocol (MP) Border Gateway routing Protocol (BGP) running between the ASBRs at either side of the ICI. An ASBR may in turn distribute a label for that route internal to its Autonomous System (AS) using an Internal BGP session (iBGP) or other label distribution protocol (e.g., LDP). Doing so allows the establishment of a label switched path to a destination route across an MPLS-ICI. The benefit of distributing a label with an IPv4 route advertised by BGP is that it allows the establishment of LSPs to IPv4 destinations as well as distributes reachability information for IPv4 routes and other routes via a single protocol, namely BGP.

B.1 MP-BGP – Label Distribution

[RFC3107] defines the MP-BGP extensions necessary to distribute one or more label with an IPv4 route advertised by a BGP update message and to withdraw an advertised labeled IPv4 route. Implementations that support this specification must support [RFC3107].

A system must allow enabling/disabling RFC3107 capability on a session basis. When RFC3107 is enabled on a session, route update processing must be controllable via BGP route policies. At a BGP speaker, a route may be learned with multiple labels and the same BGP next hop. Similarly, the same route may be learned with multiple labels and different BGP next hops. In either case, the receiving ASBR should be able to keep all labels and load balance traffic across equal cost labeled rotes. The minimum number of equal cost multipaths that can be supported should be 4.

For security purposes, MP-BGP sessions with labeled route advertisements should be run on direct point-to-point links and the BGP next hop for labeled routes must be constrained to be the advertising speaker. All other labeled routes must be rejected at the receiving speaker. In addition, the label stack (there could be one or more label per stack) must be processed at the BGP next hop. Finally, implementations must support per-interface label spaces. When a labeled packet arrives on an interface and the outer label is outside of that interface space, the packet must be dropped.

B.2 Routing

MP-BGP with extension defined in [RFC3107] is the routing protocol used to distribute an IPv4 route and the label associated with that route in addition to other routing information across an MPLS-ICI.

B.3 Resiliency

The ability to reroute around failure in this case is dependent on BGP convergence. When ECMP entries are available, a system should be capable of removing a failed route from the set of ECMP entries, keeping the other surviving ECMP entries, prior to full BGP route re-selection so that traffic is directed away from the failed path fast enough. A system should also enable tunneling a BGP session over a TE tunnel between the BGP speakers with MPLS fast Reroute enabled on the tunnel for next hop and/or link protection. In that case, to maintain security, labels associated with BGP routes should be allocated from a label space restricted to the TE tunnel, treated as an interface, rather than other types of interfaces. This will allow link/node protection to work while maintaining a level of confidence in the received label authenticity.

An ASBR compliant with this Technical Specification must support BGP graceful restart [RFC4724] and [RFC4781] for the labeled IPv4 route AFI/SAFI in order to continue forwarding for external BGP routes upon control plane switchover. It should be noted that the validity of these routes is dependent on the routes being reachable through the respective AS of the ASBR. Proper operation requires that the ASBR performing graceful restart for eBGP also performs graceful restart for the IGP on the AS-bound interfaces. If the ASBR supports a non-stop control plane (i.e., maintain routing state and signaling state on an active control card and a standby control card, and resume signaling and routing after control card switchover without interruption), control plane switchover will not be noticed by the rest of the network. In that case, the ASBR need only implement graceful restart in helper mode. Otherwise, the ASBR must implement graceful restart in both helper and restart modes.

B.4 Policies and Admission Control

An implementation in compliance with this specification must support the definition of MP-BGP import and export policies that control the processing of labeled routes. In addition, capability must provided to control the number of labeled routes that could be kept or accepted per BGP session basis from a BGP speaker or set of BGP speakers that happen to belong to same peering AS. In addition, the number of ECMP entries per labeled route must be configurable.

B.5 MIBs

The ASBR should support the management information model consistent with the BGP-4 MIB objects defined in [RFC4273]. There is no standard MIB for BGP status updates or statistics. Therefore, a router enterprise MIB or equivalent information model is necessary.

The ASBR shall be configurable to enable/disable BGP state change notifications.

If BGP is enabled, the ASBR shall generate a status change notification for at least the following events.

- BGP peer transitions from up state to down state
- BGP peer transitions from down state to up state
- Number of BGP routes crosses pre-configured threshold
- Number of BGP route changes over a pre-configured time interval crosses a pre-configured threshold

The ASBR shall be configurable to generate an alarm message if a number of routes greater than the threshold value is received. This shall be configurable on a per BGP session basis. The action taken upon shall be configurable as either accept or drop number of routes received greater than the threshold.

Annex C Pseudo Wires

C.1 Statically-configured and dynamically signaled MPLS Pseudo Wires

C.1.1 Signaling

Figure 9 shows the reference architecture in support of the MPLS-ICI multi-segment Pseudo Wire service adapted from [MS-PW_Requirements] to the MPLS-ICI configuration and terminology. The the Single-Segment PW reference architecture, [RFC3985] calls out a PSN tunnel that interconnects the provider edges' (PEs) at the terminating ends of the PW segment and tunnels that PW segment. The objective of the PSN tunnel is to hide the MPLS PW label exchanged between the PEs from the PSN network, specifically the intermideiate nodes on the path between these PEs. When the PEs terminating the PW segment are directly connected, as in the case of the PW segment extending between two ASBRs (PEs) directly connected by the MPLS-ICI, there is no need to have a PSN tunnel that tunnels the PW over the direct link. As stated in Section 1 of RFC 4447, "packets that are transmitted from one end of the pseudowire to the other are MPLS packets which must be transmitted through an MPLS tunnel. However, if the pseudowire endpoints are immediately adjacent and penultimate hop popping behavior is in use, the MPLS tunnel may not be necessary."

Figure 10 shows the reference architecture in support of the MPLS-ICI multi-segment Pseudo Wire service, adapted from [MS-PW_Requirements] to the MPLS-ICI configuration and terminology, where the PW segment (e..g., PWI1) over the MPLS-ICI is carried over the MPLS-ICI directly witout a PSN tunnel.

Implementation in support of this Technical Specification shall support [RFC4447] for single segment Pseudo Wire (SS-PW) signaling (e.g., SS-PW-A1, SS-PW-I1 and SS-PW-B1). For the establishment of Multi-segment static switching of PW segments at ASBRs (i.e., Switching PEs PE-A2 and PE-B2 in Figure 9 and Figure 10) that interconnect service provider domains (e.g., Service Provider A and Service Provider B in Figure 9 and Figure 10) . Each PW segment, specifically the PW segment over the MPLS-ICI, is established using PW control LDP signaling [RFC4447].

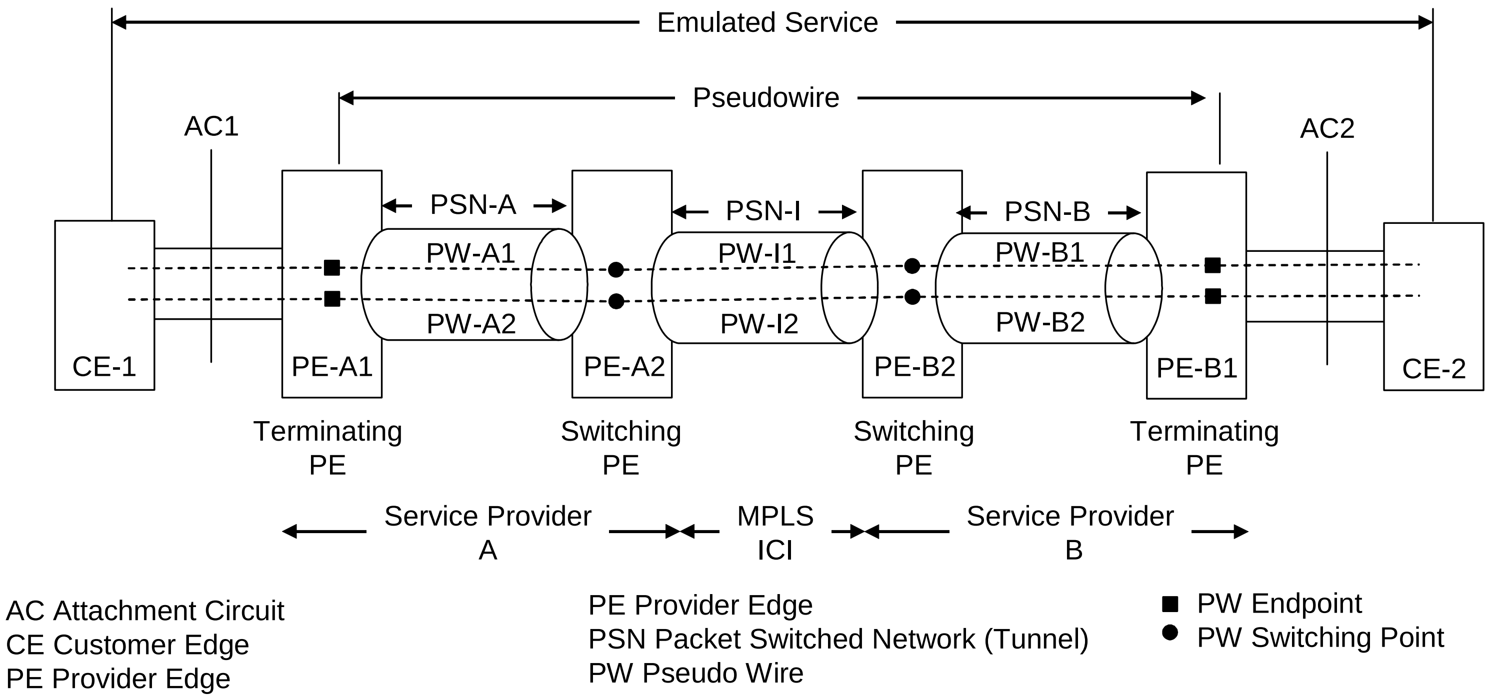


Figure 9: Multi-Segment PW reference model

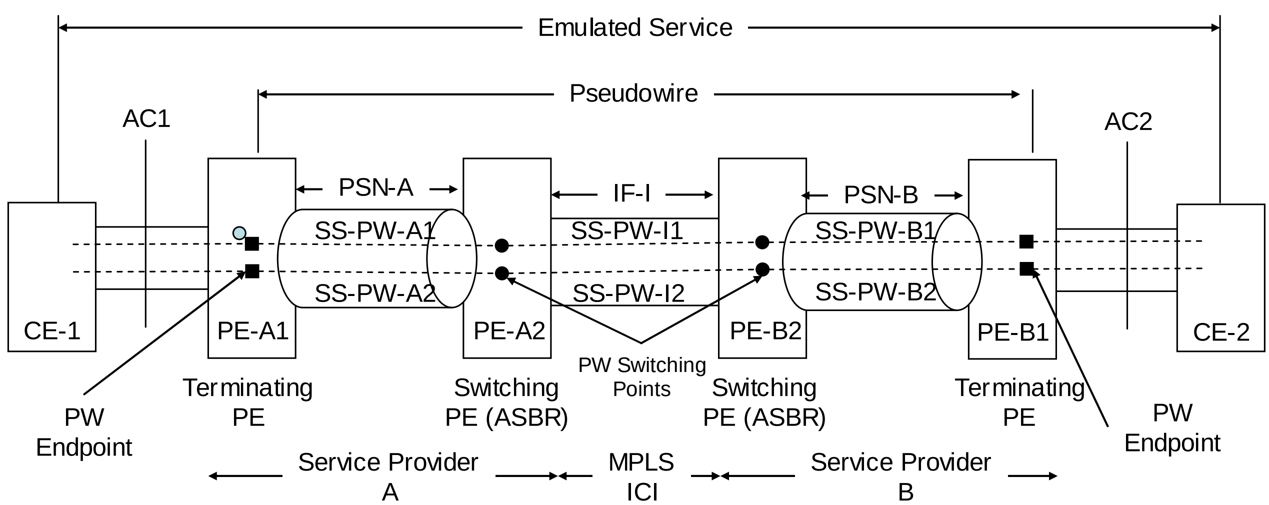


Figure 10: Multi-Segment PW Reference Model without an MPLS Tunnel on the MPLS-ICI tunneling the PW.

C.1.2 Connection Admission Control

The connection admission control capability based on traffic parameters described in Section 10 for statically configured LSPs (each direction of PW is an LSP) is the only requirement for ASBRs compliant with this Technical Specification. In this case, the traffic parameters and CoS of the PW segment over the MPLS-ICI must be configurable at the segment endpoints (ASBR at both ends of the MPLS-ICI) since they are not carried in signaling messages.

C.1.3 Routing

One ASBR at the head of a PW segment that spans an MPLS-ICI needs to know how to reach (i.e., over what interface) the ASBR IP address at the other end of the PW segment over the MPLS-ICI. This IP reachability can be learned via BGP or static routing. An implementation that supports this Technical Specification may support either method.

C.1.4 Resiliency

An ASBR compliant with this Technical Specification must support LDP graceful restart [RFC3478] in helper mode for PW ID FEC 128 [RFC 4447]. In addition, the ASBR should support LDP graceful restart [RFC3478] in helper mode for Generalized ID FEC 129 [RFC4447]. In addition, the ASBR must support graceful restart in helper mode for eBGP running with the neighbor ASBR if eBGP is used to determine reachability to that ASBR, as well as graceful restart for the MPLS tunneling protocol used to tunnel the PW if the LDP PW segment is tiunneled over another MPLS tunnel. An ASBR compliant with this Technical Specification must also support graceful restart in restart mode for the same LDP FECs, eBGP and MPLS tunneling protocol if the failure of its control plane disrupts the LDP session(s) with its neighbors(s), eBGP sessions or the PSN tunnels (e.g., RSVP-TE tunnels). If the ASBR supports a non-stop control plane (i.e., maintain routing state and signaling state on an active control card and a standby control card, and resume signaling and routing after control card switchover without interruption), control plane switchover will not be noticed by the rest of the network. In that case, the ASBR need only implement graceful restart in helper mode. Otherwise, the ASBR must implement graceful restart in both helper and restart modes.

C.1.4.1 PW Protection Models

This section covers the following protection models:

- End-End MS-PW Protection
- MPLS-ICI Protection for a PW segment over the MPLS-ICI

C.1.4.1.1 Support for End-to-End Protection Model

C.1.4.1.1.1 ASBR Role

In the End-to-End protection model depicted in Figure 11, an ASBR (S-PE) with a PW segment setup over an MPLS-CI is responsible for detecting failures on the MPLS-ICI,, and for subsequently sending failure notification to the ASBR on the other end of the MPLS-ICI and to the PW segment within its own AS towards the T-PE. The ASBR (S-PE) is also responsible for detecting the failure of the PW segment within its own AS, terminated on the

ASBR, and for sending failure notification to the other end of the segment and/or the ASBR on the other end of the MPLS-ICI. The ASBR is also responsible for relaying failure notifications it receives from its own AS or from the ASBR on the other end of the MPLS-ICI toward the T-PE in the direction of the received failure notification. In this model, T-PE's should receive the MS-PW failure notification and take action to switchover traffic to an alternative path. However, T-PE behavior is out of the scope of this Technical Specification. Failure detection and notification procedures performed by the ASBR are defined in the PW OAM section.

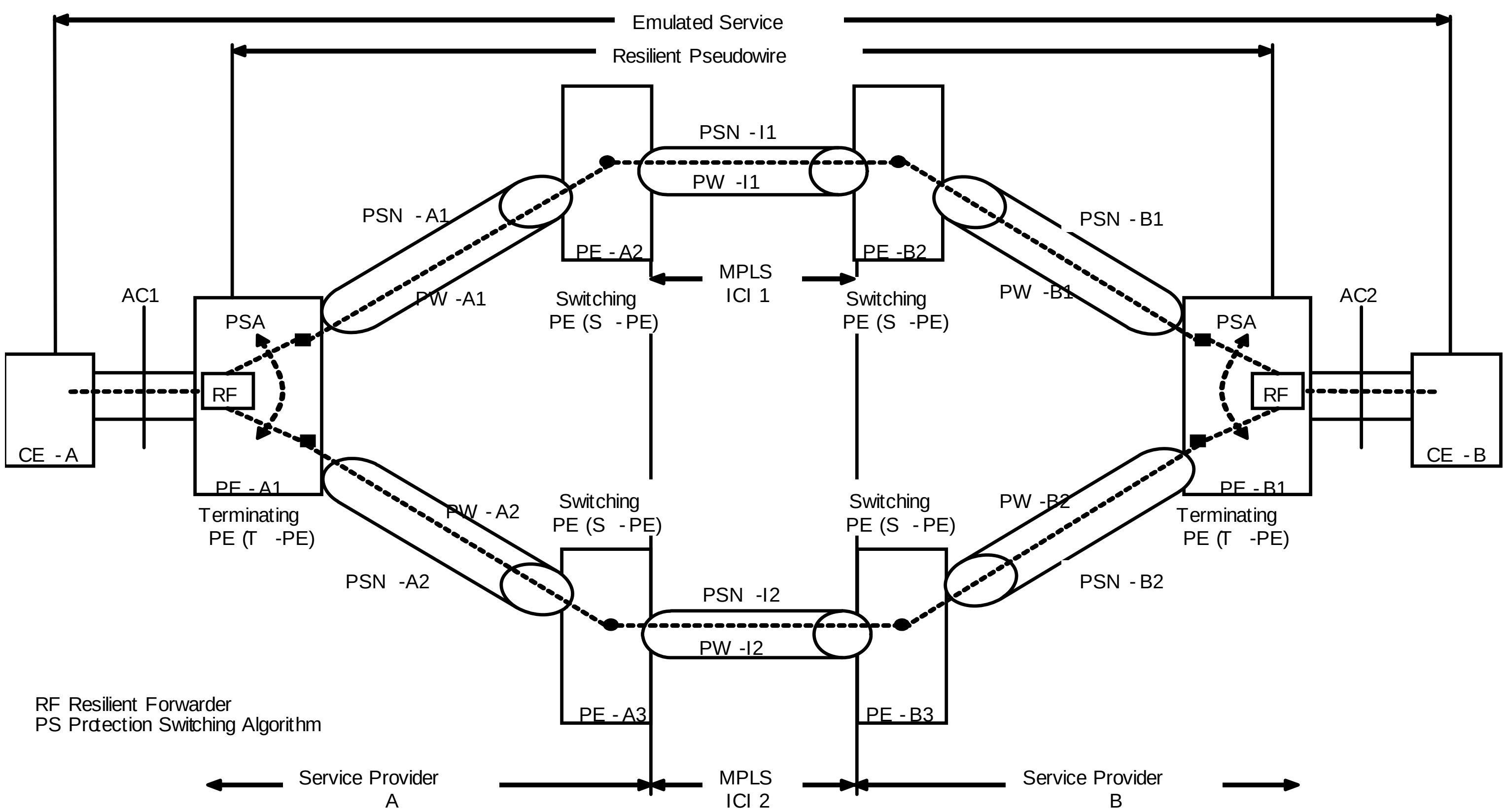


Figure 11: ASBR Fault Detection/Notification at the MPLS-ICI usage in End-End Protection (non-normative)

C.1.4.1.1.2 T-PE Role (non-normative)

(This section does not form an integral part of this Technical Specification)

Figure 11 depicts the end-end protection model whereby a MS-PW has a primary path through one MPLS-ICI and a backup path through another MPLS-ICI. In both cases, a PW segment is setup between the ASBRs interconnected by these MPLS-ICIs. MPLS-ICI failure detection and failure notification to the T-PEs is the responsibility of the ASBRs, ,while traffic switchover between primary and redundant paths is the responsibility of the T-PEs.

The backup path can be a hot standby or warm standby. The paths need only be specified in terms of T-PEs and S-PEs. The primary path has at least one PW segment setup between two ASBRs (S-PEs) interconnected by an MPLS-ICI while the backup path has at least one PW segment setup over correspondingly diverse MPLS-ICI(s). It should be noted that in the context of this Technical Specification, an MS-PW is a concatenation of SS-PWs whereby at least one PW segment is setup over an MPLS-ICI.

The selection between the primary and backup path is the responsibility of each T-PE with an endpoint of the MS-PW. Since both directions of the MS-PW must follow the same path, each T-PE must be consistently configured with which path is primary and which path is backup and must consistently elect to send traffic on one of the two paths from either direction. When there is no failure, T-PEs must select the primary path for sending traffic. Signaling extensions that allows each T-PE to indicate the primary path and switchover provide operational enhancements but are not yet defined and such definition is out of scope of this Technical Specification. In order to support switchover from a primary path to a backup path, an ASBR that detects the failure of a PW segment must be able to send a failure notification towards the T-PEs to trigger switchover from the primary path to the backup path or vice-versa. Since there is no informational element defined for signaling to indicate which path is being used by a T-PE or which path has a fault, In-band OAM via vccv must be used to inform T-PEs with the identification of the faulty path. Alternatively, two different PWs (different FEC128 PWIs, or (SAII, TAII)) can be used for primary and backup paths albeit they must be associated with the same Attachment Circuit at the T-PEs. In that case, only one PW can be up at a time. When a T-PE signals to the remote T-PE a PW different from the one currently used, it should be taken by that T-PE as an indication to switchover. In this latter case, LDP procedures for failure notification can be used. During the transition from one path to the other, T-PEs may have elected different paths due to differential delays in receiving and processing OAM messages. LDP failure notification may use status notification messages [RFC4447] or label withdrawal messages if status message is not supported. This specification recommends that equipment implement the status notification message.

C.1.4.1.2 MPLS-ICI Protection for a PW segment over the MPLS-ICI

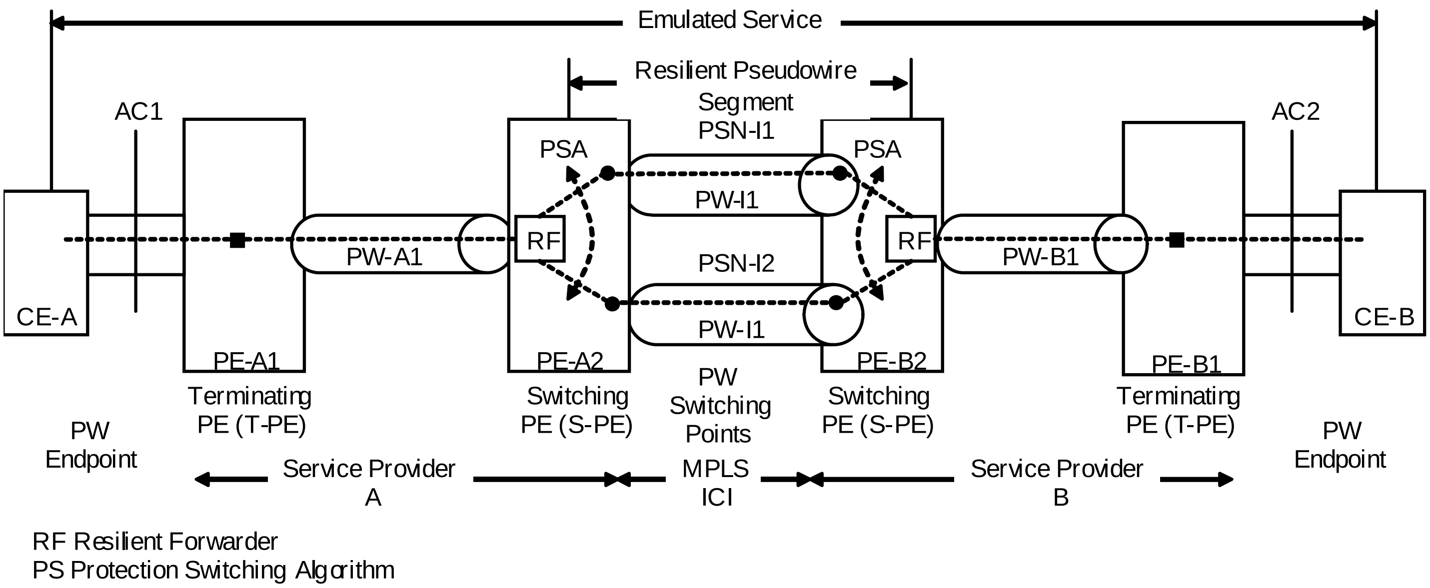


Figure 12: this figure depicts MS-PW segment protection over an MPLS-ICI when the two ASBRs at both ends of the segment are interconnected by parallel MPLS-ICIs.

An ASBR across an ICI, in conformance with this specification should support local PW segment protection. In local PW segment protection, there will be a primary path and a backup path for a MS-PW segment between the same two ASBRs. The path may be a direct MPLS-ICI or a tunnel between the same ASBR pairs traversing one or more MPLS-ICIs. Figure 12 depicts the case when the the MS-PW segment primary and redundant paths are setup over parallel MPLS-ICIs between the same ASBR pair where a PSN tunnel is present. When two ASBRs are interconnected by parallel MPLS-ICIs, and all is needed is link protection, the backup path should not have to be configured. That is, an ASBR that detects the failure of the PW segment over an MPLS-ICI should attempt to automatically select an alternative MPLS-ICI that satisfies the PW requirement for sending the PW traffic on. If the backup path is pre-setup (i.e., the alternative interface is pre-selected, the ASBR should be able to switch traffic to the alternative interface as soon it detects the failure (i.e., it does not have to make an admission control decision). If the PW segment between two ASBRs is setup over a tunnel, tunnel reroute will cause seamless reroute of the carried PW segment. It should be noted that if the ASBRs are not interconnected by parallel MPLS-ICI or if an ASBR may need to reroute a PW over a multi-hop path to the ASBR at the other end of the PW segment, the two ASBRs must be interconnected by a tunnel over the new path that tunnels the rerouted PW segment between the same ASBR pair. Segment protection between the same pair of ASBRs may be preferred over end-to-end path protection for efficiency reasons and/or the speed in seamless local reroute around a link failure.

For MPLS ICI PW segment protection, there are four protection models:

1) PW segment without a tunnel – PW protection is provided by a redundant MPLS-ICI (redundant L1/L2 links)

The redundant MPLS-ICI path is provided by a single hop L1/L2 path. The applicable protection model is the one-hop model defined in Section 9.1.1.1 and Figure 4. The MPLS-ICI PW segment can be either statically configured per Section 9.1 or signaled by LDP per section C.1.1.

ASBR shall support this ICI PW segment protection model

2) Multi-hop PW segment protection

A multi-hop L1/L2 path can provide the redundant path. The applicable protection model is the multi-hop protection defined in Section 9.1.1 and Figure 5.

ASBR shall support this ICI PW segment protection model

(3) Multi-hop PW segment protection over an MPLS tunnel (one in each direction)

A multi-hop L1/L2 path can provide the redundant path. The applicable protection model is the protection defined in Section 9.1.1 and Figure 6.

ASBR shall support this ICI PW segment protection model

4) PW segment over MPLS PSN TE-tunnel (one in each direction) – TE-tunnel fast reroute [inter-AS-RSVP-TE] as discussed in Annex D.

The TE-tunnel in this section is effectively the “MPLS LSP” that is protected by the bypass tunnel in Section 9.1.1.1. The “bypass tunnel” is the facility bypass tunnel defined in [inter-AS-RSVP-TE]. This facility bypass tunnel protects the TE tunnels and as a result protects the PW’s over it.

ASBR shall support this ICI PW segment protection model

C.1.4.1.3 Revertive path configuration

Equipment should enable the configuration of a primary and backup path in revertive or non-revertive mode. Path reversion may cause traffic in transit to arrive at the destination out of order. On the other hand, under certain circumstances, the primary path may be more optimum than the backup path. In that case, an operator may want to revert the PW path to the primary path when it becomes active again. The detection of primary path availability at the T-PE is out of scope of this document. However, an ASBR with an MPLS-ICI that protects its own segment of a MS-PW should be able to revert back to a primary path, if enabled by configuration, or to reroute the PW segment over a more optimum path if it becomes available. Optimality here alludes to a path between two interconnected ASBRs and often related to a logical link between the two ASBRs. A more optimum path can be one with more available bandwidth after rerouting the PW segment over it, or one that traverses less number of hops if it is riding on a multi-hop tunnel.

C.1.4.2 Switchover

Upon detecting the failure of a PW segment, a PE (T-PE or S-PE), with pre-configured and established local protection, must be able to switchover the PW or PW segment to the alternative path in 50-100 msecs. This specification is concerned with segments setup over an MPLS-ICI and the ASBRs at either end of that MPLS-ICI, be it S-PEs or T-PEs. Protection switchover of an MS-PW crossing an MPLS-ICI at a T-PE should also be supported.

C.1.5 MIBs

The ASBR should support the management information model consistent with the following MIB groups as specified in [RFC3815] in read-write mode:

- mplsLdpGeneralGroup
- mplsLdpNotificationsGroup

An informative description of these objects, their indexing, content, suggested usage and context is contained in Appendix I.5..

The ASBR should support the management information model consistent with the following MIB groups as specified in [PWMIB] in read-write mode:

- pwBasicGroup
- pwNotificationsGroup
- pwIDGroup
- pwGeneralizedFECGroup
- pwPwStatusGroup
- pwAttachmentGroup
- pwPerformanceIntervalGroup and pwHCPerformanceIntervalGroup
- pwSignalingGroup
- pwNotificationControlGroup

An informative description of these objects, their indexing, content, suggested usage and context is contained in Appendix I.6..

The router PW fault detection times should be configurable via an enterprise MIB for use in mapping to emulated service status indication protocols (e.g., TDM OAM, FR LMI, ATM OAM, Ethernet OAM).

The ASBR should support the management information model consistent with the following MIB groups as specified in [PWMPLSMIB] in read-write mode:

- pwMplsGroup
- pwMplsOutboundMainGroup and pwOutboundTeGroup
- pwMplsInboundGroup
- pwMplsMappingGroup

An informative description of these objects, their indexing, content, suggested usage and context is contained in Appendix I.7..

An enterprise MIB or a configuration model is necessary to implement the following:

Configuration of whether a particular PW segment is active or standby

Configuration of whether the T-PE local protection switching action should revert to the active PW segment if it is determined to be available via either LDP status notification and/or VCCV continuity checking

Configuration of whether the Control Channel (CC) [vccv] for a PW segment is to be processed by this ASBR

C.2 PW OAM

[MS_PW_Requirements] describes requirements for MS-PW OAM mechanisms. These requirements are applicable to MS-PWs setup across an MPLS-ICI with ASBRs as Switching PEs (S-PEs). ASBRs interconnected by an MPLS-ICI can also be Terminating PEs (T-PEs) with native attachment circuits, as defined in [MS-PW_arch].This section covers PW OAM mechanisms and capabilities that are required for operations across an MPLS-ICI. It addresses both SS-PWs and MS-PWs.

The OAM mechanisms described in this section capitalize on mechanisms described in [vccv]. They require some configuration capabilities on ASBRs interconnected by an MPLS-ICI and some procedures to provide for simultaneous end-to-end and segment-based OAM operations for MS-PWs. This Technical Specification is limited in scope to segment based OAM as it applies to an MPLS-ICI as a segment. The objectives of the OAM mechanisms described in this section are to provide for:

- PW segment failure detection and notification to other segments of a MS-PW
- Transparent MS-PW OAM support end-to-end across the network. S-PEs (ASBRs) must be able to pass T-PE to T-PE PW OAM messages transparently
- support for switch over between 1:1 protected LSPs end-to-end.
- Support for switchover between 1:1 protected PW segments across an MPLS-ICI

The MPLS-ICI may represent an un-trusted boundary between providers. In such cases, it may be undesirable for an -ASBR in one provider network to respond to certain PW OAM messages originating in another provider's network and received on an MPLS-ICI. Therefore, ASBRs when actng as S-PEs for PWs, should support the ability to discard PW OAM messages received on an MPLS-ICI. Procedures described in this section provide for such capability.

C.2.1 PW OAM Mechanisms

C.2.1.1 PW Connection Verification: "always on" failure detection and notification

As specified in this Technical Specification, The PW segment established across the MPLS-ICI is statically configured but could be statically established or dynamically established using LDP PWE3 control [PWE3_control][MS-PW-switching]. In addition, end-end OAM may or may not be configured for the MS-PW.

In cases where end-end OAM is not enabled for the MS-PW and the segment across the MPLS-ICI is statically established, vccv [vccv] becomes the only available mechanism to check the liveliness of the MS-PW segment across the MPLS-ICI and relay the status of the PW to other segments by interworking with LDP.

Equipment compliant with this Technical Specification must support the following capabilities:

- In-band VCCV [vccv] based on the PW associated channel for the vccv control channel traffic as specified in [vccv]
- BFD with IPv4/UDP header encapsulation (ed note: we may want to state that IPv4/UDP encapsulation is optional) for fault detection and status notification as the continuity verification type (CV). The CV type is carried in the control word [vccv]. It is either configured on both ends of a PW or negotiated via targeted LDP when targeted LDP is enabled on the MPLS-ICI for PW setup.
- Interworking between BFD status message and LDP status notification message as discussed in the faiure detection and notification section.
- Asynchronous BFD for continuity check. The following must be configurable for a BFD session at the ASBRs on either end of the PW segment:
    - IP address used as the source IP address when sending BFD messages and as the destination IP address when receiving BFD message
    - The transmission interval for generating the BFD messages must be configurable. In addition, the failure detection time must be configurable as a multiple of the transmission interval from the peer.
    - Min RX interval
    - Exp bits assigned to PW labeled packets that contain OAM messages. Continuity check messages should be marked with exp bits that result in the most assured and lowest latency forwarding behavior available so that false failure detection does not occur.
- The ASBR may allow the configuration of :
    - Your Discriminator [BFD_base]
    - My Descrimiator [BFD_base]

- If a BFD session fails, the ASBR that detects the failure must send a failure notification to the remote end of the session via a BFD status message. It must also map that failure to the preceding segment via LDP status notification as discussed in the failure detection and notification section
- If the ASBR is a T-PE for a PW, the ASBR should map the failure notification to an OAM mechanism native to the attachment circuit [OAM_MAP]

When LDP and vccv are used to indicate status of a PW segment, the resolution of the status of a PW segment must follow [vccv] Section 4.1. In cases where the ASBR-ASBR PW segment is statically setup, there is no LDP to indicate failure of a PW across an MPLS-ICI.

Equipment compliant with this Technical Specification must allow for end-to-end OAM and segment OAM simultaneously. In particular, within the scope of this specification is the MPLS-ICI segment. Equipment compliant with this Technical Specification must enable testing the liveliness of a PW segment across the MPLS-ICI and failure notification or status up messages across such a segment. In order to provide for these capabilities, the following must be provided on the ASBR:

- The ability to manually configure a BFD session per PW/PW segment on each ASBR at either end of that segment across the MPLS-ICI
- The ability to set the PW label TTL value for each BFD message generated by an ASBR to 1 and the TTL value in the IP header of the encapsulated BFD message to 255. An ASBR that receives a PW MPLS PDU with TTL 1 will have TTL expire and intercept the PDU for further processing. The ASBR should perform the following checks in processing the PDU. The more of these checks are done in the forwarding plane without impacting forwarding engine performance, the more resilient the receiving ASBR will be to potential DOS attacks. These checks may be implemented in any order on the local system. The order of the checks as expressed in this section represents an example.
    - The ASBR processes the packet if the PW label is a valid label in accordance with section 13 (i.e., the label was assigned by the ASBR to the neighbor). Otherwise, the packet is dropped and a corresponding counter is incremented.
    - The ASBR checks if OAM is enabled for that segment. If it is, it proceeds to the next step. Otherwise, the ASBR discards the PDU and increases a counter for TTL expired PW PDUs.
    - If the PDU is not dropped, the ASBR checks if there is a control word. If there is, processing proceeds with the next step. Otherwise, the packet is dropped.
    - If the CV type matches the configured type for that PW, processing continues. Otherwise, the packet is dropped.
    - If the CV type is that of BFD with IP encapsulation and fault detection, the TTL value in the IP header is checked. If TTL is less than 255, the packet may be assumed to have been compromised and the packet is dropped. Otherwise, processing of the packet continues.
    - If the PDU contains an authentication header, the message is authenticated. If authentication succeeds, processing of the packet continues. Otherwise, the packet is dropped and a counter is incremented for dropped OAM packets that failed authentication. This counter should be maintained per PW.
    - If the PDU is not dropped, the ASBR checks the IP address of the source. If the source IP addresses matches that of the peer, processing of the packet proceeds with the next step. Otherwise, the packet is dropped.

- o If the PDU is not dropped, the ASBR checks that the destination IP address corresponds to itself and is what was used for that BFD session. If this check passes, processing proceeds. Otherwise, the PDU is dropped.
- o If the PDU is not dropped, the ASBR checks that the "Your Discriminator" field in the message matches the locally configured "My Discriminator" field and that "My Discriminator" field in the messages matches the locally configured "Your Discriminator". If this check passes, processing proceeds. Otherwise, the PDU is dropped.

In order to allow for simultaneous operation of end-to-end OAM and segment based OAM as described in this section, end-end OAM messages must have TTL set to 255. This avoids expiry of TTL at an S-PE.

C.2.1.2 Diagnostics

Diagnostics allows on-demand path verification and checking of control plane against data plane state and for path trace. An ASBR compliant with this Technical Specification must support the following:

- vccv with CV type being LSP-ping and CC being the PW associated control channel.
- LSP-ping in ping mode as specified in [vccv] Section 5.4.1
- The ability to initiate LSP ping test. An ASBR must be able to test the PW segment extending to the immediate neighbor across the MPLS-ICI by setting TTL to 1 in the PW label header
- The ability of an LSP-ping test within the provider network. Details of the operations of this test are outside of the scope of this document.

A MS-PW may have multiple switching points within one provider network. A provider may not want to reveal the number of segments traversed by an MS-PW within its domain to another provider. One simple approach to deal with this problem is that an ASBR that receives an LSP-ping echo request across an MPLS-ICI does not reply back with the downstream label map. This benavior must be configurable. However, this behavior will prevent the trace initiator from initiating an LSP-ping test in ping mode to a point beyond the ASBR as it will not be able to include the correct label stack.

C.2.1.3 Failure Detection and Notification Procedures

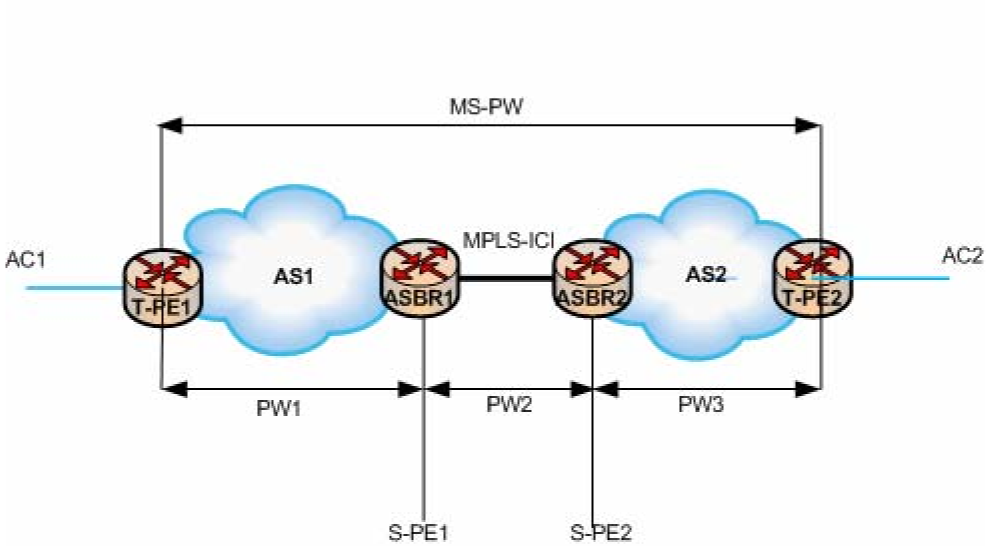


Figure 13: Reference model for a MS-PW across an MPLS-ICI

PW status signaling via BFD is done using diagnostic codes. [BFD_base] defines the following diagnostic codes

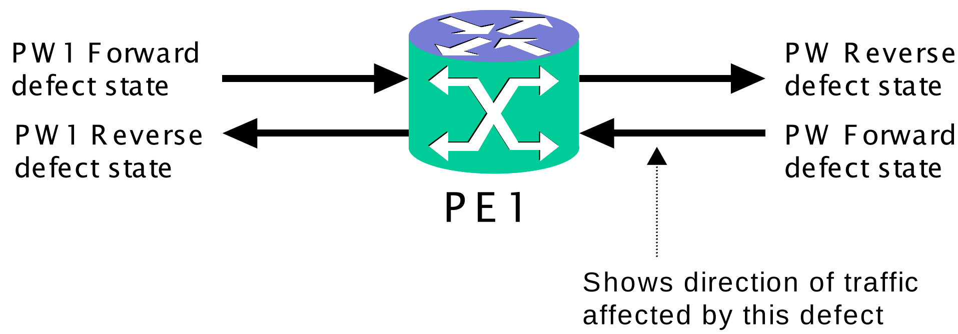
0 -- No Diagnostic
1 -- Control Detection Time Expired
2 -- Echo Function Failed
3 -- Neighbor Signaled Session Down
4 -- Forwarding Plane Reset
5 -- Path Down
6 -- Concatenated Path Down
7 -- Administratively Down
8 -- Reverse Concatenated Path Down
9-31 -- Reserved for future use

[RFC4447] defines status notification messages for SS-PWs using status notification LDP messages. [RFC4446] defines the following PW status codes used in status notification messages for PWs:

0x00000000 - Pseudowire forwarding (clear all failures)
0x00000001 - Pseudowire Not Forwarding
0x00000002 - Local Attachment Circuit (ingress) Receive Fault
0x00000004 - Local Attachment Circuit (egress) Transmit Fault
0x00000008 - Local PSN-facing PW (ingress) Receive Fault
0x00000010 - Local PSN-facing PW (egress) Transmit Fault

Figure 13 depicts the reference model used for defining failure detection and notification procedures for MS-PWs in this document.

The PW fault directions are defined as follows:



The context of the PW fault directions is equivalent to that for attachement circuits and PWs as specified in MFA

Code points for forward defect and reverse defect notifications have not been specified for PW Status signaling. In this version of this Specification, we refer to "forward defect" and "reverse defect" indications. The following mapping may be performed between code points defined in RFC 4446 for PW status signaling:

Forward defect - corresponds to::
[Local Attachment Circuit (ingress) Receive Fault
and logical OR
Local PSN-facing PW (egress) Transmit Fault]

Reverse defect - corresponds to:
[Local Attachment Circuit (egress) Transmit Fault
and logical OR
Local PSN-facing PW (egress) Transmit Fault]

For BFD status notification messages, the following mapping can be performed:

Forward defect - corresponds to:
[Control Detection Time Expired
and logical OR
Reverse concatenated path down]

Reverse defect - corresponds to:
[Control Detection Time Expired
and logical OR
Concatenated path down]

The merging of the received PW status and the local status for the PW segments at an S-PE at the ICI can be summarized as follows:

1. When the local status for both PW segments is UP, the S-PE passes any received AC or PW status bits unchanged, i.e., the status notification TLV is unchanged but the VCid in the case of a FEC 128 TLV is set to value of the PW segment to the next hop. Where LDP is used on one side of the S-PE and BFD is used on the other side, the mapping of the equivalent state if performed as described above.

2. When a local fault is detected by the S-PE, a PW status message (either LDP or BFD) is sent in both directions along the PW. Since there are no attachment circuits on an S-PE, only the following status messages are relevant:

0x00000008 - Local PSN-facing PW (ingress) Receive Fault
0x00000010 - Local PSN-facing PW (egress) Transmit Fault

The following failure notification procedure refer to Figure 13 above and must be supported when BFD is used for failure detection and notification on the MPLS-ICI between S-PE1 and S-PE2 for PW2, and LDP is used on segments PW1 and PW3 with status notification capability:

1. S-PE1/SPE2 receives an LDP status notification message from the session associated with PW1/PW2.
    a. If status code received is 0x00000002 or 0x00000010, S-PE1/S-PE2 sends a BFD control message to S-PE2/S-PE1 with Diagnostics code of 6 (Concatenated Path Down). Upon receiving a BFD message with Diagnostic code 6, S-PE1/S-PE2 sends a BFD message with Diagnostic code 8 to S-PE2/S-PE1 and LDP status notification message with status code 0x0000002 for PW3/PW1, respectively. In this case, the PW is considered as failed.
    b. If the status code received is 0x00000004 or 0x00000008, S-PE1/S-PE2 sends a BFD control message to S-PE2/S-PE1 with Diagnostic code of 7 (Reverse Concatenated Path Down). Upon receiving a BFD message with Diagnostic code 7, S-PE1/S-PE2 sends a BFD message with Diagnostic code 6 to S-PE2/S-PE1 and LDP status notification with code 0x00000004 for PW3/PW1, respectively. The PW is considered as failed.
    c. If the status code recived is 0x00000001 (PW not forwarding), S-PE1/S-PE2 sends a BFD control message to S-PE2/S-PE1 with Diagnostic code 5 (Path Down). Upon receiving a BFD message with code 5, S-PE2/S-PE1 sends an LDP status notification message with status code 0x00000001 to T-PE2/T-PE1. The PW is considered as failed.
2. If failure is the result of an LDP session timing out on S-PE1/SP2 and there are no graceful retsart procedures in progress for the corresponfing LDP session, S-PE1/S-PE2 will declare all the PWs associated with that LDP session down. SPE1/S-PE2 sends a BFD message with status code 0x00000004 to S-PE2/S-PE1. Upon reciving this BFD message, S-PE2/S-PE1 sends a status notification message with status code 0x00000002 to T-PE2/T-PE1.

3. If a failure is detected locally on the MPLS-ICI at S-PE1/S-PE2, there are the following cases
    a. If PW2 is carried over the MPLS-ICI directly (i.e., without encapsulation in an MPLS tunnel) and the BFD session session did not time out
        i. If there is a tunnel setup between S-PE1 and S-PE2 and another between S-PE2 and S-PE1 and both tunnels are up, and and PW2 can be admitted over each tunnel and rerouted over the tunnel in time (before failure is detected based on BFD control session timeout), PW2 will be rerouted transparently. There are no PW failures.
        ii. Otherwise, if both S-PE1 S-PE2 cannot reroute PW2, S-PE1/S-PE2 must send an LDP status notification message with status code 0x00000004 for PW1/PW2. If for some reason S-PE1/S-PE2 could reroute its direction of

PW2 but S-PE2/S-PE1 could not, S-PE1/S-SPE2 will detect failure of the other direction based on BFD connectivity verification procedures. That is S-PE1/S-PE2 will not receive BFD conituity verification messages for the period determined by the dtection multiplier and BFD message transmission interval. As a result, S-PE1 and S-PE2 will declare PW2 down and S-PE1/S-PE2 will send an LDP status notification message with status code 0x00000002 to T-PE1/T-PE2 and S-PE2/S-PE1 sends an LDP status notification message with status code 0x00000004 to T-PE2/T-PE1.

b. If PW2 is carried over an MPLS tunnel between S-PE1 and SPE-2 and another between S-PE2 and S-PE1, and the tunnels are rerouted in time before BFD session time out, PW2 will be rerouted transparently and no failures will be detected. Otherwise, if either MPLS tuneel cannot be rerouted in time before detection of failure, the associated direction of the PW will fail and the procedures described in previous case (a) will follow.
c. If failure is the result of BFD session timing out at S-PE1/S-PE2, S-PE2/S-PE1 sends a BFD control message with code 1 to S-PE1/S-PE2 and LDP status notification message with status code 0x00000002 to T-PE2/T-PE1.

The following failure notification procedure must be supported when LDP is used for failure notification between S-PE1 and S-PE2, and LDP is used on segments PW1 and PW3 with status notification capability <MB> Aren't you mixing up failure notifications with protection switching?

1. If S-PE1/S-PE2 receives an LDP status notification message from the session associated with PW1/PW3 declaring failure, this status notification must be relayed via LDP. The local PW segment must be declared down. If the failure notification recived contains a PW switching TLV [PW-segmeneted] and S-PE1/SP-E2 can determine that the failure point that initiated the LDP status notification message is within its own domain (may not be a simple task as the PW-switching sub-TOVs does not include an AS number), S-PE1/S-PE2 may replace the information in the PW switching TLV with its own based on policy or remove the optional PW switching TLV based on a policy, for confientialty reasons.
2. If S-PE1/S-PE2 detects PW failure on the MPLS-ICI, there are the following cases
    a. If PW2 is carried over the MPLS-ICI directly (i.e., without encapsulation in an MPLS tunnel) and the BFD session session over the MPLS-ICI, if one exists, did not time out
        i. If there is a tunnel setup between S-PE1 and S-PE2 and another between S-PE2 and S-PE1 and both tunnels are up, and and PW2 can be admitted over each tunnel and rerouted over the tunnel in time (before failure is detected based on BFD control session timeout), PW2 will be rerouted transparently. There are no PW failures.
        ii. Otherwise, if both S-PE1 S-PE2 cannot reroute PW2, S-PE1/S-PE2 must send an LDP status notification message with status code 0x00000004 for PW1/PW2. If for some reason S-PE1/S-PE2 could reroute its direction of PW2 but S-PE2/S-PE1 could not, S-PE1/S-SPE2 could detect failure of the other direction based on BFD connectivity verification procedures if a BFD session is configured for that PW. That is S-PE1/S-PE2 will not receive BFD conituity verification messages for the period determined by the dtection multiplier and BFD message transmission interval. As a result, S-PE1 and S-PE2 will declare PW2 down and S-PE1/S-PE2 will send an LDP status notification message with status code 0x00000002 to T-PE1/T-PE2 and S-PE2/S-PE1 sends an LDP status notification message with status code 0x00000004 to T-PE2/T-PE1. It should be noted that if there is a routing

path between S-PE1 and S-PE2, the LDP session may still be alive even if the PW is not.

b. If PW2 is carried over an MPLS tunnel between S-PE1 and SPE-2 and another between S-PE2 and S-PE1, and the tunnels are rerouted in time before BFD session time out if out exists, PW2 will be rerouted transparently and no failures will be detected. Otherwise, if either MPLS tuneel cannot be rerouted in time before detection of failure, the associated direction of the PW will fail and the procedures described in previous case (a) will follow.
c. If failure is the result of BFD session timing out at S-PE1/S-PE2, S-PE2/S-PE1 sends a BFD control message with code 1 to S-PE1/S-PE2 and LDP status notification message with status code 0x00000002 to T-PE2/T-PE1.

C.2.2 Security Considerations

ASBR must also be able to protect confidentiality and security of its own AS as described in the security Section 13. In addition, procedures described in this section are intended to provide additional security and confidentiality measures that apply to PWs.

Mechanism must exist to prevent control processor overload due to expiry of large numbers of VCCV packets at an ASBR. As indicated in Section 13 (security section), in order to prevent DOS attacks , an ASBR must provide capability for policing OAM messages to a configured rate in the forwarding plane. The ASBR must provide the capability to configure a policer per PW, aggregate for all PWs, or both. The policing parameters should be configurable. For manual setting, the rate should be configured to match an expected rate. Alternatively, the system may have the capability to compute a rate for continuity check messages from session transmission rates and program or update a corresponding policer that applies to this type of messages. In addition, when vccv is not enabled for a PW segment or if messages with the wrong CV types are received, vccv messages received on that PW segment to be locally processed must be dropped without impact on the ASBR control plane. A corresponding counter should be incremented.

Annex D Traffic Engineering (TE)-Tunnels

D.1 Statically-configured and signaled TE-tunnels

D.1.1 Signaling

Equipment must provide the capability of configuring a TE-tunnel LSP between two provider border nodes (ASBRs) interconnected by a direct interface. The LSP itself can be the only LSP segment (i.e., the segment starts at one ASBR and terminates at another ASBR at the other end of an interface that interconnects the two), or it can be a segment of an end-end LSP initiated in one provider domain and terminated in another, transiting the border nodes.

Equipment must have the capability of establishing a configured TE-tunnel between the two ASBRs via RSVP-TE signaling. Equipment must provide the capability of stitching, by administrative action, that LSP to another LSP in each provider domain to form an LSP that crosses the domain boundaries.

Equipment must enable the definition of a label space assigned to a group of one or more interfaces. If an MPLS packet arrives on an interface with a label that is out of the space associated with that interface, the packet must be dropped. This must also apply to any label being processed on that interface. On an ASBR, there could be a label space assigned to each peer and the associated connections to it. At the same time there will be a label space assigned to all AS-internal interfaces. Thus, AS-internal labels can have global-space from the point of view of all AS-internal interfaces.

Equipment must also provide the capability of configuring a traffic profile for the LSP at the ingress of the traffic. The traffic profile should include the classes of service supported over the LSP in case of E-LSP and the bandwidth profile for each CoS traffic carried over that LSP. Section 10 addresses the requirements for traffic management in more details.

When statically stitching LSP segments, an LSP must be declared as up end to end if all segments of the LSP are up. When dynamically establishing LSP segments across the boundaries, it is preferred that any intermediate segment establishment be triggered by RSVP-TE signaling, starting at the head end of the end-end LSP.

RSVP-TE signaling must support signaling of multiple classes of service for each LSP based on the extension defined in [MPLS-CNI].

D.1.4 MIBs

The ASBR should support the management information model consistent with the following read-write MIB groups as specified in [RFC3812]:

- mplsTunnelGroup
- mplsTunnelScalarGroup
- mplsTunnelManualGroup
- mplsTeNotificationGroup

An informative description of the objects in these groups, their indexing, content, suggested usage and context is contained in Appendix I.8.

D.2 Dynamically Established TE-Tunnels

D.2.1 Signaling

This section states the required and recommended capability for establishing InterCarrier traffic engineering (TE) tunnels that carry traffic for one or more service classes. TE tunnels are label switched paths with constraints (e.g., bandwidth, preemption priority, etc.). In particular, this section defines the capabilities required for dynamic setup of TE tunnels across service provider domain boundaries.

An ASBR must be capable of supporting the setup of a TE-LSP from one router in one service provider domain X to another router in service provider domain Y. Currently, the simplest way to achieve this is by extending existing intra-Area and intra-domain mechanisms for RSVP-TE, including capability of signaling DS-TE LSPs [RFC4124].

ASBRs compliant with this Technical Specification and this section must be capable of establishing dynamic inter-domain TE-LSPs using interdomain RSVP-TE extensions as defined in [inter-AS-RSVP-TE]. Specifically, a ASBR must support three types of inter-AS TE tunnels: (1) contiguous, (2) stitched and (3) nested. [inter-AS-RSVP-TE] is still in IETF working group draft status. In addition, ASBRs compliant with this implementation must support TE-path computation as defined in [interdomain-PDPC]. [interdomain-PDPC] is also still in IETF draft stage. In this section, stitching is referred to as dynamic stitching. Dynamic stitching, in contrast to the static stitching discussed in section 6.1.2, refers to stitching individually signaled RSVP sessions together in the data plane and control plane.

D.2.2 Routing

This specification requires support for per domain path computation procedures at an ASBR to compute the path segment to the first hop in the Explicit Route Object (ERO) or to the TE tunnel destination. At an ASBR ingress to an MPLS-ICI, the ASBR must be able to select the neighbor ASBR and associated link that satisfies the TE constraints and provide for reachability to the first hop in the ERO and the TE tunnel tail-end. At an ASBR on the egress from an MPLS-ICI, the ASBR must compute a path (generally a partial path) that satisfies the TE constraints and can reach the first hop in the ERO and TE tunnel tail-end.

The only routing inforation required in this specification to be exchanged over the MPLS-ICI dynamically is that of reachability relayed via BGP update messages. That is, this specification does not require any change to normal BGP operations between two ASBRs.

D.2.3 Resiliency

Equipment in compliance with this Technical Specification must support MPLS fast Reroute (MPLS FRR), providing for reroute around link failure and node failure to Next to Next hop. This specification requires support for bypass tunnel for MPLS fast reroute. Equipment in compliance with specification should support one-one detours for MPLS Fast Reroute. It should be noted that provding for node protection requires revealing the node ID next to the ASBR in the forwarding direction as well as the label assigned by that node. This may be considered by some carriers as reveling confidential information and may not be allowed.

In order to provide for continuing forwarding when the ASBR control plane fails, an ASBR compliant with this Technical Specification must support RSVP-TE graceful restart [RFC3471] as well as graceful restart for the routing protocols on which RSVP-TE path computation depends (i.e., eBGP for IPv4 routes [RFC4724] on the MPLS-ICI and IGP on intra-AS core interfaces). If the ASBR supports a non-stop control plane (i.e., maintain routing state and signaling state on an active control card and a standby control card, and resume signaling and routing after control card switchover without interruption), control plane switchover will not be noticed by the rest of the network. In that case, the ASBR need only implement graceful restart in helper mode. Otherwise, the ASBR must implement graceful restart in both helper and restart modes.

D.2.4 MIBs

The ASBR should support the management information model consistent with [RFC 2206] for provisioning an RSVP session per destination, including at least the following: Sender IP address (rsvpSenderAddr) , Destination IP address (rsvpSessionDestAddr) , RSVP message refresh interval (rsvpSenderInterval).

The ASBR should also support the following additional RSVP session parameters using an Enterprise MIB and CLI: Enable/disable RSVP refresh reduction, RSVP session status, LSP and Tunnel ID's provisioned by a particular session.

The ASBR shall support injection of a TE-LSP into the local Routing Information Base (RIB) using a configured metric via an enterprise MIB or equivalent configuration.

The ASBR shall support an enterprise MIB or equivalent that supports configuration of the MPLS PSN tunnel fault detection times for each of the OAM protocols configured to detect faults (e.g., L1, L2).

The ASBR should support the management information model consistent with the following read-write MIB groups as specified in [RFC3812]:

- mplsTunnelGroup
- mplsTunnelScalarGroup
- mplsTunnelSignaledGroup
- mplsTeNotificationGroup

An informative description of the objects in these groups, their indexing, content, suggested usage and context is contained in Appendix I.8.

D.3 Admission Control

A ASBR that supports admission control in compliance with this Technical Specification must provide the following admission control capabilities:

- When administratively stitching TE-LSPs at a domain boundary, one LSP segment may be signaled up to a preconfigured endpoint where it is administratively stitched to an LSP segment. If the arriving traffic profile in the RSVP-TE path message differs from that configured at the boundary, the ASBR must be capable of rejecting the LSP setup by sending a path error message to the source of the path message.

- When dynamically stitching TE-LSPs, a ASBR must be able to apply admission control on the arriving path message. If the traffic profile of the arriving LSP per CT exceeds that of the LSP with which stitching is required, the LSP setup is rejected via a path error message. If the arriving path message has bandwidth requirements that are below those of the configured LSP, the LSP is considered to pass the bandwidth admission control. Once admitted, no other LSP can be stitched to the same InterCarrier LSP segment.

- When establishing a contiguous LSP, the egress ASBR at a domain boundary must have the capability of applying admission control on the arriving path message. If the requested bandwidth per CT exceeds the available bandwidth over the interface to the other carrier, the ASBR must reject the LSP setup

- When establishing a contiguous LSP, the egress ASBR at a domain boundary must have the capability of applying admission control on the arriving path message. If the requested bandwidth per CT exceeds the available bandwidth allocated for the carrier across the ICI at that boundary, the ASBR must reject the LSP setup.

- When establishing a nested TE-LSP, the ASBR at a domain boundary must have the capability of applying admission control on the arriving path message. If the requested bandwidth per CT exceeds the available bandwidth on the LSP that will tunnel the LSP being requested, the LSP is rejected. Otherwise, the LSP is admitted and the tunnel LSP bandwidth is appropriately adjusted. ASBRs should provide options for allowing the expansion of the tunnel LSP bandwidth to accommodate the LSP being tunneled. If the expansion is successful, the LSP will be successfully admitted into the tunnel.

- In addition to making admission based on banadwidth per CT, an ASBR must be able to make admission control decisions based on other types of policies. An ASBR at one of an MPLS-ICI that receives TE-LSP setup requests must enable the configuration of policies that apply to admission control. Specifically, a ASBR must be able to make admission control decisions based on: (1) requested setup priority, (2) requested pre-emption priority, (3) request for route recording, (4) destination address of the LSP, (5) source address of the LSP, (6) neighbor, and (7) affinity of the LSP. A ASBR should be able to admit or reject the setup of an TE-LSP, or modify any of the attributes in the setup message. If an attribute is modified based on a policy, the source should be notified of the modification. Notification procedures will need to be further defined and standardized and are requitred in this specification.

  Policies may be described by a set of one or more MATCH criteria, plus for each MATCH criteria a list of one or more associated ACTIONs.

The MATCH criteria are applied to RSVP Path messages (ie, call requests). If an LSP meets the MATCH criteria, then the associated ACTIONs are applied to that LSP. MATCH may be based, for example, on any combination of:

1. Interface
2. Neighbor (for point to point links this is implied by the interface, but for multipoint links there may be multiple neighbors per interface).
3. Direction (LSP is incoming or outgoing on the associated interface)
4. Source Address (or prefix)
5. Destination Address (or prefix)
6. setup and/or holding Priority
7. affinity

D.4 Protection of Confidential Information

Implementations supporting RSVP-TE signaling at the ICI must allow operators to configure policies that control the handling of ERO and RRO objects for confidentiality reasons.

Note1: Some explanation may be appropriate regarding the "per-interface" nature of Policies. In general, equipment that supports the ICI may have some interfaces which support ICI with "untrusted" neighbors (such as ICI interfaces to other service providers). Equipment that supports the ICI may have some interfaces which support ICI with "trusted" neighbors (such as interfaces to other ASs operated by the same service provider, or to other service providers that are known to operate their networks in a relatively secure fashion). The amount of trust may also vary based on whether the interface is authenticated, or other factors. Similarly it may be necessary for security and availability reasons to restrict LSPs from neighbors which fail to implement refresh reduction. Equipment that supports the ICI may have other interfaces which are internal to the service provider (ie, that don't represent an ICI interface). In many cases Policies may vary between neighbors based on factors such as these.

Note2: The following description outlines one way that Policies may be realized. This is not intended to require any particular implementation, and any implementation which provides equivalent functionality is permitted.

Implementations supporting RSVP-TE signaling at the ICI should allow operators to set policies which limit which EROs will be accepted on incoming RSVP Path messages. Policies should be configurable on a per-interface basis. For example, policies may allow operators to prohibit EROs, to restrict which addresses may be included in EROs (such as by prohibiting addresses of P routers internal to the service provider, while allowing PE routers or external addresses).

Implementations supporting RSVP-TE signaling at the ICI should allow operators to set policies restricting the transmission of RROs. Policies should be configurable on a per-interface basis. For example, policies may prohibit transmission of RROs (implying that RROs would be stripped out prior to transmission of PATH messages), or may prohibit the transmission of the addresses of internal P routers (so that the internal addresses would be stripped out of the RRO prior to transmission across an ICI).

Implementations supporting RSVP-TE signaling at the ICI should allow operators to set policies restricting the transmission of PathErr Messages across the ICI. Policies should be configurable on a per-interface basis.For example, PathErr messages received from inside the service provider which are being propagated across the ICI may have the error code and/or sub-code changed to avoid exposing internal topology information. Tthus for example the PathErr message may be propagated across the ICI with error code=2 ("Policy Control failure") and subcode-tbd ("Inter-Domain Policy failure), rather than using a more specific error code and subcode. See [RFC3209] [[inter-AS-RSVP-TE]. Similarly PathErr messages originated by nodes within the service provider may have the source addresses changed (eg, to the PE router supporting the ICI) in order to preserve the confidentiality of the nodes internal to the SP.

D.5 Class Type Mappings

When there is a mismatch in the implementation of CoS indicated by Class type (CT) and or TE classes (CT plus priority) between two providers that want to establish an MPLS interconnection, mappings must be administratively defined at each provider's boundary at the time the interconnection is configured. Using these mappings, translation must be performed at the ingress ASBR while signaling the LSP across a provider domain boundary. ASBRs in compliance with this Technical Specification must be able to perform this translation. Specifically, ASBRs must enable the configuration of an incoming map that defines the mapping of incoming CT/TE-class and PHBs from one domain to outgoing CT/TE-class and PHBs that apply in the next domain whose border node performs the mapping.

Annex E Voice Over IP

The voice over IP application uses the TE-tunnel transport services discussed in the Annex D and does not add any additional requirements on the MPLS-ICI.

Service providers need to have the ability to host measurement probes, as well as carry and report performance measurements across an MPLS-ICI in support of intercarrier VoIP peering. Probes and performance measurements are beyon the scope of the MPLS-ICI. Guidelines and methodologies for performing QoS measurements and budgeting for impairments across multiple provider domains can be found in [Y.1541] amd [MIT_WP].

Appendix I – Informative Description of MIBs

The subsections of this Appendix contain an informative description of the objects contained in MIB groups specified in normative requirements in the body of this specification. It includes a summary of their naming, indexing, content, suggested usage and context. For further detail and information, the implementer should consult the cited references.

Appendix I.1 MPLS LSR MIB [RFC3813]

The following MPLS MIB Tables apply to the configuration of static MPLS LSPs or for use in quieries of MPLS cross connect s that occur as a result of dynamic MPLS signaling: These tables can be used to provision and manage point-to-point, point-to-multipoint, or multipoint-point MPLS LSP segments.

The ASBR uses the mplsInterfaceGroupTable in the mplsInterfaceGroup as specified in [RFC3813] as a sequence of MplsInterfaceEntry objects indexed by {mplsInterfaceIndex} for the purpose of defining and querying the status of MPLS-enabled interfaces. This table includes information regarding the label range that can be received or transmitted on an interface level basis, total bandwidth of the interface, and whether per platform or per interface label space is used.

The ASBR uses the mplsInterfacePerfGroupTable in the mplsPerfGroup as specified in [RFC3813] as a sequence of MplsInterfacePerfEntry as an extension to the mplsInterfaceEntry table for the purpose of tracking statistics regarding label usage, lookup failures, and fragmented packets on MPLS-enabled interfaces.

The ASBR uses the mplsTunnelPerfTable in the mplsInterfaceGroup as an augment to the MPLSTunnelTable as specified in the MPLS TE MIB of RFC3812 for measuring tunnel performance. This table includes .32- and 64-bit counters of the number of packets forwarded by the tunnel. 32- and 64-bit counters of the bytes forwarded by the tunnel, and a 32-bit counter of the packets dropped because of errors or other reasons.

The ASBR uses the mplsInSegmentTable in the mplsInSegmentGroup as specified in [RFC3813] as a sequence of mplsInSegmentEntry as indexed by {mplsInSegmentIndex}. This table includes the IP address of the previous hop, incoming interface, incoming label, whether label popping should occur, the creator of the cross connect (e.g., manual, LDP, or RSVP-TE), an index to the traffic parameters, and an index to a cross-connect table.

The ASBR uses the mplsInSegmentPerfTable in the mplsInSegmentGroup as specified in [RFC3813] as a sequence of mplsInSegmentPerfEntry as an augment to mplsInSegmentEntry for monitoring the statistics of an incoming MPLS LSP segment. This table includes 32-bit counters for packets, octets, errors and discards in the receive direction of this segment and a 64-bit counter of octets received.

The ASBR uses the mplsOutSegmentTable in the mplsOutSegmentGroup as specified in [RFC3813] as a sequence of mplsOutSegmentEntry indexed by {mplsOutSegmentIndex}. This table includes identification of the IP address of the next hop, the outgoing interface, outgoing label, whether a label push(es) should occur and the a pointer to a label stack table entry, the creator of the cross connect (e.g., manual, LDP, or RSVP-TE), an index to the traffic parameters, and an index to a cross-connect table.

The ASBR uses the mplsOutSegmentPerfTable in the mplsInterfaceGroup as specified in [RFC3813] as a sequence of mplsOutSegmentPerfEntry as an augment to mplsOutSegmentEntry for monitoring the statistics of an outgoing MPLS LSP segment. This table includes 32-bit counters for packets, octets, errors and discards related to packets destined for this segment and a 64-bit counter of octets sent.

The ASBR uses the mplsXCTable mplsXCGroup as a sequence of mplsXCEntry objects indexed by {mplsXCIndex, mplsXCInSegmentIndex, mplsXCOutSegmentIndex}. This table includes contains the indeces to the incoming segment(s) and outgoing segment(s), an LSP ID, the creator of the cross connect (e.g., manual, LDP, or RSVP-TE), the administrative and operational status of the cross-connect.

The ASBR uses the mplsLabelStackTable in the mplsLabelStackGroup as a sequence of mplsLabelStackEntry objects indexed by {mplsLabelStackIndex, mplsLabelStackLabelIndex} as specified in [RFC3813]. This table includes a list of label values to be pushed beneath the top label.

The ASBR uses the mplsLsrNotificationsGroup as specified in [RFC3813] that indicates whether an mpls cross-connect (XC) is up or down.

Appendix I.2 Differentiated Services MIB [RFC3289]

The ASBR uses the diffServDataPathTable of the diffServMIBDataPathGroup as specified in [RFC3289] as a sequence of diffServDataPathEntry indexed by {ifIndex, diffServDataPathIfDirection} for the purposes of specifying the interface index, direction (i.e., ingress or egress) and a pointer to the first Diffserv component of the components configured on this router and interface.

The ASBR uses the diffServClfrElementTable of the diffServMIBClfrGroup as specified in [RFC3289] as a sequence of DiffServClfrEntry for the purposes of identifying the classifier employed, the parameters of the classifier, the precedence order in which it is applied, and a pointer to the next Diffserv component if that classifier applied to the packet is matched.

[RFC3289] specifies L3/L4 classification parameters in the diffServMultiFieldClfrTable of the diffServMIBClfrGroup as a sequence of diffServMultiFieldClfrEntry indexed by a row pointer that contains Source/Destination IPv4/IPv6 address prefixes, DSCP, and source/destination transport layer port numbers. The ASBR may support this MIB to allow configuration of ACLs that result in only admitting certain types of packets to the ASBR. This MIB can be used to support the security requirements specified in section 13.

The ASBR uses the diffServMeterTable of the diffServMIBMeterGroup as specified in [RFC3289] as a sequence of diffServMeterEntry indexed by a row pointer that contains row pointers to the parameters of the meter, as well as row pointers to the next Diffserv component to be invoked if the meter succeeds or fails the conformance test.

The ASBR uses the diffServTBParamTable of the diffServMIBTBParamGroup as specified in [RFC3289] as a sequence of diffServTBParamEntry indexed by a row pointer that defines the Token Bucket (TB) parameters for a simple Token bucket, SrTCM both color-blind and color-aware, TrTCM both color-blind and color-aware. Additional values may be specified in other MIBs.

The ASBR uses the diffServActionTable of the diffServMIBActionGroup as specified in [RFC3289] as a sequence of diffServActionEntry indexed by a row pointer that contains row pointers to the parameters of the action, as well as a row pointer to the next Diffserv component to be invoked.

[RFC3289] defines tables for specific Diffserv-related actions for counting, DSCP (re) marking, algorithmic dropping and random dropping as follows.

The ASBR uses the diffServCountActTable of the diffServMIBCounterGroup as specified in [RFC3289] as a sequence of DiffServCountActEntry indexed by row pointer that contains 64-bit counters of packets and octets.

The ASBR uses the diffServDscpMarkActTable of the diffServMIBDscpMarkActGroup as specified in [RFC3289] as a sequence of diffServDscpMarkActEntry indexed by a row pointer which replaces the DSCP value in the packet with the value stored in the table.

The ASBR uses the diffServAlgDropTable of the diffServMIBAlgDropGroup as specified in [RFC3289] as a sequence of diffServAlgDropEntry indexed by row pointer. This table contains the type (e.g., tailDrop, headDrop, randomDrop, alwaysDrop), a row pointer to the next Diffserv component, identification of the queue, the threshold applied to that queue, a pointer to a table of parameters specific to the drop algorithm, and 64-bit counters for the number of packets and octets dropped for the algorithmic and random cases.

The ASBR uses the diffServRandomDropTable of the diffServMIBRandomDropGroup as specified in [RFC3289] as a sequence of diffServRandomDropEntry indexed by row pointer. This table contains parameters specific to the random drop (i.e., WRED) algorithm for the minimum threshold packets and bytes, maximum threshold packets and bytes, maximum drop probability, and the weighting for past queue history and the sampling interval for the WRED algorithm.

The ASBR uses the diffServQTable of the diffServMIBSchedulerGroup as specified in [RFC3289] as a sequence of DiffServQEntry indexed by a row pointer. This table contains a pointer to an entry in the diffServSchedulerTable, and row pointers to the minimum and maximum rate tables that contain the values that the scheduler should use.

The ASBR uses the diffServSchedulerTable of the diffServMIBSchedulerGroup as specified in [RFC3289] as a sequence of diffServSchedulerEntry indexed by a row pointer. This table contains a pointer to the next Diffserv component, which if it is another scheduler is a way to specify hierarchical scheduling. This table contains the scheduler method (e.g., priority, WRR, WFQ), and pointers into the MinRate and MaxRate tables.

The ASBR uses the diffServMinRateTable of the diffServMIBSchedulerGroup as specified in [RFC3289] as a sequence of diffServMinRateEntry indexed by a row pointer. This table contains parameters relevant to the scheduler method: a priority value (larger numeric value has higher priority), minimum absolute rate, minimum relative rate as a fraction of the interface speed.

The ASBR uses the diffServMaxRateTable of the diffServMIBSchedulerGroup as specified in [RFC3289] as a sequence of diffServMaxRateEntry indexed by row pointer and diffServMaxRateLevel. This table contains an absolute maximum absolute and relative rate for a non-work-conserving scheduler (i.e., shaper).

Appendix I.3 Bi-directional Forwarding Detection (BFD) MIB [BFD MIB]

The ASBR uses the bfdSessTable of the bfdSessionGroup as specified in [BFDMIB] as a sequence of BfdSessEntry indexed by {bfdSessIndex} for the purposes of configuring and managing BFD sessions. This table includes the application ID (e.g., MPLS VPN), local and remote discriminator values, UDP port number, session state (admin down, down, initializing, up), a Boolean indication regarding receipt of packets from the remote, operational mode (asynchronous or demand with or without echo), Boolean indications of the local preference for demand and echo modes, Boolean indication of operation through a control plane disruption, session address type (IPv4 or IPv6) and address, the desired minimum transmit interval, minimum receive and echo intervals, the detection time multiplier, and a Boolean indication of local desire to use Authentication.

The ASBR uses the bfdSessPerfTable of the bfdSessionPerfGroup as specified in [BFDMIB] as a sequence of BfdSessPerfEntry as an augment to BfdSessEntry for the purposes of collecting statistics on each BFD session. This table includes 32-bit counters for the number of BFD messages received and sent, the last time and diagnostic code when communication was lost with the neighbor, the number of times this session has entered the up state since the last ASBR reboot, the last discontinuity time, and 64-bit high-capacity counters for for the number of BFD messages received and sent.

The ASBR uses the bfdSessMapTable of the bfdSessionGroup as specified in [BFDMIB] as a sequence of bfdSessMapEntry indexed by {bfdSessApplicationId, bfdSessDiscriminator, bfdSessAddrType, bfdSessAddr} contains the value of bfdSessMapBfdIndex, which is the index into bfdSessTable.

The ASBR uses the bfdSessNotificationsEnable of the bfdNotificationGroup Boolean scalar to enable/disable the BFD session up and down notifications. A single BFD session up and down notification is issued for a contiguous range of entries in bfdSessTable to minimize the emission of a large number of notifications.

The ASBR uses BFD notifications when a new detection time or detection multiplier is negotiated.

The ASBR uses BFD notifications when a change occurs in the BFD session operating mode.

Appendix I.4 BGP MPLS L3VPN MIB [RFC4382]

State context of 4384 Multi-AS option a and b.

The ASBR uses the following scalar objects from the mplsL3VpnScalarGroup of [RFC4382] : number of configured VRFs, number of active VRFs, number of connected VPN interfaces, enabling/disabling all notifications, maximum total number of routers across all VRFs, the minimum interval for issuing the maximum route threshold notification, and a threshold for the number of illegally received labels above which a notification is issued.

The ASBR uses the mplsL3VpnIfConfTable of the mplsL3VpnIfGroup as a sequence of MplsL3VpnIfConfEntry as specified in [RFC4382] and indexed by {mplsL3VpnVrfName, mplsL3VpnIfConfIndex} for the purposes of configuring the VPN attributes of individual interfaces. This includes, IP VPN classification (enterprise, carrier's carrier, or inter-AS), and the route distribution protocol used (e.g., BGP, RIP, OSPF).

The ASBR uses the mplsL3VpnVrfConfTable of the mplsL3VpnVrfGroup as specified in [RFC4382] and indexed by {mplsL3VpnVrfName} for the purpose of configuring and managing a Virtual Routing and Forwarding (VRF) instance. This table includes, the human-readable name (mplsL3VpnVrfName), the VPN ID from [RFC4265] which can indicate a specific VPN or all VPNs, operational and administrative status, number of associated interfaces, number of active interfaces, mid- and high-level water marks for the number of routes, and the maximum number of routes allowed.

The ASBR uses the mplsL3VpnVrfPerfTable of the mplsL3VpnPerfGroup as specified in [RFC4382], which augments mplsL3VpnVrfTable to provide performance counters for each VRF. This table includes the number of routes added and the number of routes removed since the last discontinuity, the current number of routes, the number of routes dropped due to exceeding the maximum threshold, and the time of the last discontinuity.

The ASBR uses the mplsL3VpnVrfRouteTable of the mplsL3VpnVrfRteGroup as specified in [RFC4382] and indexed by {mplsL3VpnVrfName,mplsL3VpnVrfRteInetCidrDestType, mplsL3VpnVrfRteInetCidrDest, mplsL3VpnVrfRteInetCidrPfxLen, mplsL3VpnVrfRteInetCidrPolicy, mplsL3VpnVrfRteInetCidrNHopType, mplsL3VpnVrfRteInetCidrNextHop} for the purpose of configuring and managing individual route entries in the VRF. This table includes the destination IP address prefix and length, the next hop IP address prefix and length, the next hop type (e.g., local, remote, black hole, reject), mechanism by which the route was learned, AS number of the next hop, primary and alternate metrics, and an index into mplsXCTable [RFC3813] to manually configure the label stack and MPLS cross-connect to be used for this route.

The ASBR uses the MplsVpnVrfRTTable of the mplsL3VpnVrfRTGroup as specified in [RFC4382] and indexed by {mplsL3VpnVrfName, mplsL3VpnVrfRTIndex, mplsL3VpnVrfRTType} for the purpose of configuring and managing individual Route Target (RT) entries in the specified VRF. This table includes the actual RT and its type (import ,export or both) .

The ASBR uses the following notifications of the mplsL3VpnNotificationGroup as specified in [RFC4382] : one interface associated with the VRF is up, one interface associated with the VRF is down, and indication when the mid- or high-level number of threshold is crossed, the number of routes in this VRF has fallen below the high-level threshold, the number of illegal VRF labels exceeds the threshold.

Appendix I.5 MPLS LDP MIB [RFC3815]

The following apply to statically configured and signaled via LDP LSPs.

The ASBR uses the mplsInSegmentLdpLspTable of the mplsLdpLspGroup as specified in [RFC3815] as indexed by {mplsLdpEntityLdpId, mplsLdpEntityIndex, mplsLdpPeerLdpId, mplsInSegmentLdpLspIndex} for the purposes of associated an LDP signaled LSP with an MPLS cross connect (XC) as defined in the MPLS LSR MIB of [RFC3813]. This table includes the value of mplsInSegmentIndex in mplsInSegmentTable of [RFC3813], the label type (generic, FR, ATM), and LSP type (terminating, originating, or cross-connecting).

The ASBR uses the mplsOutSegmentLdpLspTable of the mplsLdpLspGroup as specified in [RFC3815] as indexed by {mplsLdpEntityLdpId, mplsLdpEntityIndex, mplsLdpPeerLdpId, mplsOutSegmentLdpLspIndex} for the purposes of associated an LDP signaled LSP with an MPLS cross connect (XC) as defined in the MPLS LSR MIB of [RFC3813]. This table includes the value of mplsOutSegmentIndex in mplsOutSegmentTable of [RFC3813], the label type (generic, FR, ATM), and LSP type (terminating, originating, or cross-connecting).

The following requirements apply to targeted LDP-established LSPs.

The ASBR uses the SNMPv3 MIB mplsLdpEntityGenericLRTable of the mplsLdpGeneralGroup as indexed by {mplsLdpEntityLdpId mplsLdpEntityIndex, mplsLdpEntityGenericLRMin, mplsLdpEntityGenericLRMax} from [RFC3815] for assigning a label range to an interface or a platform for the indexed LDP entity.

The ASBR uses the mplsLdpEntityTable of the mplsLdpGeneralGroup as indexed by {mplsLdpEntityLdpId, mplsLdpEntityIndex } of the MPLS LDP MIB [RFC3815] as indexed by (mplsLdpEntityLdpId, mplsLdpEntityIndex) to configure/set-up potential LDP sessions on a specific LSR/LER/PE.

The ASBR populates the mplsLdpPeerTable of the mplsLdpGeneralGroup as a sequence of mplsLdpPeerEntry objects indexed by (mplsLdpEntityLdpId, mplsLdpEntityIndex, mplsLdpPeerLdpId} of the MPLS LDP MIB [RFC3815] as determined by LDP through initialization or discovery with information about LDP Peers known to LDP Entities through LDP signaling.

The ASBR populates the mplsLdpSessionTable of the mplsLdpGeneralGroup as a sequence of mplsLdpSessionEntry objects indexed by {mplsLdpEntityLdpId, mplsLdpEntityIndex, mplsLdpPeerLdpId, mplsLdpSessionPeerAddrIndex} of the MPLS LDP MIB [RFC3815] with information learned about a peer via LDP, such as; state, role, protocol version, keep-alive timer and discontinuity time.

The ASBR populates the mplsLdpEntityStatsTable of the mplsLdpGeneralGroup as an augment to mplsLdpEntityTable of the MPLS LDP MIB [RFC3815] to keep statistical information about the LDP Entities on the ASBR, such as; session-level errors, LDP message content errors, and timer expirations.

The ASBR populates the mplsLdpHelloAdjacencyTable of the mplsLdpGeneralGroup as a sequence of mplsLdpHelloAdjacencyEnry objects indexed by {mplsLdpEntityLdpId, mplsLdpEntityIndex, mplsLdpPeerLdpId, mplsLdpHelloAdjacencyIndex} of the MPLS LDP MIB [RFC3815] with statistical information learned about a peer, such as the hold timer and adjacency type.

The ASBR uses LDP Notifications of the mplsLdpNotificationsGroup as defined in the MPLS LDP MIB [RFC3815]. The intent of this requirement is to autonomous notification of the management plane of LDP neighbor adjacency changes.

The ASBR uses the mplsFecTable of the mplsLdpGeneralGroup as specified in [RFC3815] as a sequence of mplsFecEntry as indexed by {mplsFecIndex} for the purposes of configuring Forwarding Equivalence Class (FEC) information. This table contains the FEC type (prefix or host), the address type (IPv4, IPv6), the address (or prefix), and the prefix length.

Appendix I.6 Pseudowire MIB [PWMIB]

The ASBR creates an entry in pwTable in the pwBasicGroup as specified in [PWMIB] for all locally configured PW types (e.g., Ethernet, TDM, FR, ATM, etc.) to hold generic parameters related to PW creation and monitoring. The ASBR creates corresponding entries in the [PWMPLSMIB] and the specific emulated service MIB as associated by the common pwType.

The ASBR supports the pwTable in the pwBasicGroup as specified in [PWMIB] as a sequence of pwEntry objects indexed by {pwIndex} for the purpose of generic configuration and status monitoring specific pseudowire. This table includes information to specify the PW type, PSN tunnel type, setup/holding priority, address type, local stitching information, attachment IDs, configuration of control word option, MTU size, group ID, inbound and outbound PW label, local and remote operational status and other management information.

The ASBR supports the pwIndexMappingTable in the pwPwIdGroup as specified in [PWMIB] as a sequence of pwIndexMappingEntry objects indexed by {pwIndexMappingPwType, pwIndexMappingPWID, pwIndexMappingPeerAddrType, pwIndexMappingPeerAddr} for the purpose of performing a reverse mapping only if the basic PW FEC ID is used.

The ASBR supports the pwPeerMappingTable in the pwMappingTablesGroup as specified in [PWMIB] as a sequence of pwPeerMappingEntry objects indexed by { pwPeerMappingPeerAddrType, pwPeerMappingPeerAddr, pwPeerMappingPwType, pwPeerMappingPwID} for the purpose of querying PW's based upon peer, type and pwID.

The ASBR uses configuration control of notifications as specified in the pwNotificationsGroup of [PWMIB] controlling whether up/down and pw deleted notifications are sent, and the maximum overall rate of notifications that the ASBR can send.

The ASBR uses the pwPerfCurrentTable in the pwPerformanceIntervalGroup and pwHCPerformanceIntervalGroup as a sequence of pwPerfCurrentEntry objects indexed by {pwIndex} as specified in [PWMIB] for the purpose of reporting per-PW performance information for the current interval. .This table includes 32-and 64-bit counters of the number of packets received from the PSN, number of bytes received from the PSN, number of packets sent to the PSN, and number of bytes sent to the PSN.

The ASBR uses the pwPerfIntervalTable in the pwPerformance1DayIntervalGroup as indexed by {pwIndex, pwPerfIntervalNumber} as specified in [PWMIB] for the purpose of per-PW performance information for historical intervals. This table contains up to 96 entries (e.g., 15 minute intervals for a 24-hour period) of the statistics contained in pwPerfCurrentTable, the duration of the interval and an indication of data validity.

The ASBR uses the pwPerfTotalTable in the pwPerformanceGeneralGroup indexed by {pwIndex} as specified in [PWMIB] for the purpose of per-PW performance information for the time since PW establishment or the latest management application reset. This table contains the statistics in pwPerfCurrentTable for the total time since the most recent time that any row counter suffered a discontinuity.

Appendix I.7 Psuedowire to MPLS PSN MIB [PWMPLSMIB]

The ASBR uses the pwMPLSTable of the pwMPLSGroup as specified in [PWMPLSMIB] as a sequence of pwMPLSEntry objects indexed by {pwIndex} for the purpose of associating an MPLS

PSN tunnel with a specific pseudowire. This MIB table includes information to counfigure the PSN tunnel type (TE or non-TE), the EXP bits used in the PW label, the PW label TTL value, and the peer LDP ID.

The ASBR uses the pwMPLSOutboundTable of the pwOutboundMainGroup and pwOutboundTeGroup as specified in [PWMPLSMIB] as a sequence of pwMPLSOutboundEntry objects which augments an pwMPLSEntry for the purpose of statically mapping a specific PW to a PSN tunnel in the outbound direction.

The ASBR uses the pwMPLSInboundTable of the pwMplsInboundGroup as specified in [PWMPLSMIB] as a sequence of pwMPLSInboundEntry objects indexed by {pwIndex} for the purpose of indicating the inbound PSN tunnel when LDP signaling is used. The single value, pwMPLSInboundXcIndex indicates the XC index representing this PW in the inbound direction.

The ASBR uses the pwMPLSNonTeMappingTable of the pwMplsMappingGroup as specified in [PWMPLSMIB as a sequence of pwMPLSNonTeMappingEntry objects indexed by {pwMPLSNonTeMappingDirection, pwMPLSNonTeMappingXcIndex, pwMPLSNonTeMappingIfIndex, pwMPLSNonTeMappingVcIndex } for indicating the PSN tunnel under a variety of non-TE conditions. An application can use this table to quickly retrieve the PW carried over specific non-TE MPLS outer tunnel or physical interface for use in test access.

Appendix I.8 MPLS Traffic Engineering (TE) MIB [RFC3812]

The ASBR uses the mplsTunnelTable from [RFC3812] of the mplsTunnelGroup as a sequence of mplsTunnelEntry objects indexed by{mplsTunnelIndex, mplsTunnelInstance, mplsTunnelIngressLSRId, mplsTunnelEgressLSRId} for establishing MPLS-TE tunnels and tracking their state. This table identifies the Ingress and Egress LSR IDs, whether the tunnel corresponds to an interface or the node, setup/holding priority, session attributes (e.g., FRR, pinning), FRR local protection mode (facility, detour), up time, operational status, alarm state and other attributes.

The ASBR uses the mplsTunnelResourceTable of the mplsTunnelGroup from [RFC3812] which is a sequence of mplsTunnelResourceEntry objects as indexed by {mplsTunnelResourceIndex} for setting up the tunnel resources. This table contains the mean bit rate and maximum burst size for the LSP.

The ASBR uses the mplsTunnelHopTable of the mplsTunnelGroup from [RFC3812] as a sequence of MplsTunnelHopEntry objects indexed by {mplsTunnelHopListIndex, mplsTunnelHopPathOptionIndex, mplsTunnelHopIndex} for specifying strict or loose source routed MPLS tunnel hops.

The ASBR uses the mplsTunnelARHopTable of the mplsTunnelGroup from [RFC3812] as a sequence of MplsTunnelARHopEntry objects indexed by {mplsTunnelARHopListIndex, mplsTunnelARHopIndex} to indicate the individual hops for an MPLS tunnel defined in mplsTunnelTable if it is reported by the MPLS signalling protocol (i.e., the record route option has been configured for a particular tunnel).

The ASBR uses the mplsTunnelCHopTable of the mplsTunnelGroup from [RFC3812] as a sequence of MplsTunnelCHopEntry objects indexed by {mplsTunnelCHopListIndex, mplsTunnelCHopIndex} to indicate the hops for an MPLS tunnel defined in mplsTunnelTable as computed by a constraint-based routing protocol, based on the mplsTunnelHopTable for the outgoing direction of the tunnel.

The ASBR uses the following notifications of the mplsTeNotificationGroup from [RFC3812]: Tunnel Up, Tunnel down , Tunnel rerouted, Tunnel reoptimized. The Tunnel Up/Down traps shall be capable of being enabled/disabled.

The ASBR uses the following scalar objects of the mplsTunnelScalarGroup from [RFC3812]: number of tunnels configured, label distribution protocols configured maximum hops that can be specified for a tunnel, and the maximum rate at which tunnel notifications can be sent.

Appendix II MPLS-ICI Forwarding Behavior and EXP bit Mapping Configuration Example

AppendixII.1 Overview

Forwarding behaviors are defined by several standards bodies. The IETF has defined a set of Differentiated Services (Diffserv) Per Hop Behavior (PHB) that is used in definition of Per Domain Behavior (PDB) that define service objectives [RFC 3086]. Similarly, ITU-T Y.1541 has defined a set of six recommended QoS classes that define objectives that are mapped to PHBs as summarized in section X.4. The example in this Appendix assumes that the ASBR defines a mapping for MPLS EXP bits to local traffic management forwarding behaviors as a local implementation decision.

In the examples of this Appendix, service providers connecting across an ICI agree to a set of forwarding behaviors that they will support. These need not be the same, may be a different number, and may not be defined by the same standards body, if any. A set of these forwarding behaviors would be identified by EXP bits for each MPLS-ICI logical interface. As defined in [RFC3270], there are only three bits and hence at maximum 8 PHBs assignable per LSP. Associating the EXP mapping with an incoming label map [RFC3270] would effectively extend the number of bits available. An EXP<->PHB mapping function as specified in section 11.1.1 is required to map the forwarding behaviors supported in one network to those supported in the network on the other side of the ICI. An EXP-bit remarking function is needed if these mappings are not identical on the ingress and egress interfaces.

Service providers may agree to EXP-bit and forwarding behavior mappings as described in the examples. These mappings should not be interpreted as a hard requirement, but rather represent examples of the generic mapping and remarking functions specified in section 11.1.1 and section 11.4.

AppendixII.2 Example of Bi-Lateral Agreement Forwarding Behavior and EXP-bit Mappings

The following tables show one example of the bi-lateral agreement forwarding behavior and EXP-bit mapping and remarking for two Service Providers (SP), SP-1 using the currently defined ITU-T Y.1541 QoS Classes (i.e., 0 through 5) and a descriptive phrase taken from the guidance for these QoS classes from section 5.3.5 of Y.1541 and SP-2 using IETF Diffserv PHBs. Note that per RFC 3270, EXP marking determines the PHB for E-LSPs. In the interest of brevity drop precedence was not included in this example. For an example of how drop precedence could be encoded in the EXP bits for E-LSPs, see [aggr]. This appendix does not set any requirements on how these forwarding behaviors would be specified or signaled.

In addition to the forwarding behavior and EXP bit mapping and remarking, the service providers may also specify policing, scheduling and/or shaping across the MPLS-ICI as specified in sections 11.1.2 and 11.1.3 using the Diffserv MIB information model described in section 11.4.

Table 1: Example Usage of EXP↔PHB Mappings at SP-1's ASBR

SP-1 EXP↔PHB Mapping Incoming from SP-1 Domain		Generic PHB "Name"	SP-2 EXP↔PHB Mapping Outgoing to MPLS-ICI		
EXP	**Y.1541 QoS Class, Guidance "PHB"**		**Queue**	**Diffserv PHB**	**EXP**
0	0, Real-time, jitter sensitive	Alpha	Q2.1	EF	5
2	2, Signaling	Alpha	Q2.1	EF	5
3	3, Transaction Data, interactive	Gamma	Q2.2	AF41	4
4	4, Low Loss Only	Delta	Q2.3	AF31	3
5	5, Default IP	Epsilon	Q2.4	BE	0

SP-2 EXP↔PHB Mapping Incoming from MPLS-ICI		Generic PHB "Name"	SP-1 EXP↔PHB Mapping Outgoing to SP-1 Domain		
EXP	**Diffserv PHB**		**Queue**	**Y.1541 QoS Class, Guidance "PHB"**	**EXP**
5	EF	Alpha	Q1.1	0, Real-time, jitter sensitive	0
		Beta	Q1.2	2, Signaling	2
4	AF41	Gamma	Q1.3	3, Transaction Data, interactive	3

3	AF31	Delta	Q1.4	4, Low Loss Only	4
0	BE	Epsilon	Q1.5	5, Default IP	5

Table 2: Example Usage of EXP↔PHB Mappings at SP-2's ASBR

SP-2 EXP↔PHB Mapping Incoming from SP-2 Domain		**Generic PHB "Name"**	**SP-1 EXP↔PHB Mapping Outgoing to MPLS-ICI**		
EXP	**Diffserv PHB**		**Queue**	**Y.1541 QoS Class, Guidance "PHB"**	**EXP**
5	EF	EXP5	Q2.1	0, Real-time, jitter sensitive	0
4	AF41	EXP4	Q2.2	3, Transaction Data, interactive	3
3	AF31	EXP3	Q2.3	4, Low Loss Only	4
2	AF21	EXP2	Q2.3	4, Low Loss Only	4
0	BE	EXP0	Q2.4	5, Default IP	5

SP-1 EXP↔PHB Mapping Incoming at MPLS-ICI		**Generic PHB "Name"**	**SP-2 EXP↔PHB Mapping Outgoing to SP-2 Domain**		
EXP	**Y.1541 QoS Class, Guidance "PHB"**		**Queue**	**Diffserv PHB**	**EXP**
0	0, Real-time, jitter sensitive/ 2, Signaling	EXP5	Q2.1	EF	5
3	3, Transaction Data, interactive	EXP4	Q2.2	AF41	4
4	4, Low Loss Only	EXP3	Q2.3	AF31	3
		EXP2	Q2.3	AF21	2
5	5, Default IP	EXP0	Q2.4	BE	0

In this example, the service providers SP-1 and SP-2 agreed to an equivalence mapping of the ITU-T QoS classes and IETF PHBs, which is indicated in the column titled PHB Name in the middle of the pairs of Tables X-1 and X-2. The generic PHB names in SP-1's ASBR refer to some internal tag used within the ASBR. EXP bit remarking would occur on the egress interface in this ASBR. The generic PHB names in SP-2's ASBR refer to EXP bit markings. In SP-2's ASBR, the EXP bits would be remarked on the ingress interface before being sent to the egress interface. All that needs to be agreed is an equivalence mapping to the generic PHB name that meets the QoS objectives agreed to between the providers. For each ASBR there is a table for each direction, the first in the direction incoming from the SP domain outgoing to the MPLS-ICI, and the second in the direction incoming from the MPLS-ICI and outgoing to the SP domain. Note that not all forwarding behaviors supported by one provider may have a corresponding forwarding behavior in another provider network. Each table shows the EXP to PHB mapping and queue used in the incoming direction in the leftmost three columns, the generic PHB Name in the fourth column and the corresponding queue used, the PHB, and the EXP bit marking for the outgoing direction in the rightmost three columns.

In this example, SP-1 has five forwarding behaviors based on [Y.1541] while SP-2 also has five forwarding behaviors based on Diffserv on [RFC2475]. In this example, SP-2 uses five queues, while SP-1 uses only four to illustrate the possibility that many PHBs may map to a single queue, for example, in the first table of X-2. Note that neither SP followed each of these standards exactly in this example, which does occur in the current state of the industry. The mapping between Y.1541 QoS Classes and IETF PHBs also has more entries than that contained in Appendix VI of Y.1541 described in section X.4 below. In most cases in this example, the mapping was one-one. Additionally, the cases of many-one and a null mapping input are also illustrated in this example. The first table in X-1 shows an example of a many-one mapping from SP-1 to SP-2 where two forwarding behaviors from SP-1 are mapped to the single forwarding behavior in SP-2's ASBR, specifically; SP-1 has distinct signaling and real-time QoS classes, while SP-1 only has an Expedited Forwarding (EF) PHB. A similar many-one case exists in SP-2's ASBR in the first table in X-2 for the mapping of IETF PHBs for AF3 and AF2 to only ITU QoS Class 4 in SP-1. The second table in X-1 shows an example of null mapping, since SP-2 has only the EF PHB, there is no mapping entry since SP-2 has no support for ITU QoS Class 2. A similar case exists in the second table in X-2 since SP-1 has no corresponding PHB to AF21 in SP-2.

AppendixII.3 Another Example of Bi-Lateral Forwarding Behavior Mapping

The following tables show another example of the bi-lateral agreement forwarding behavior and EXP-bit mapping and remarking for two Service Providers (SP), SP-A using the currently defined (i.e., 0 through 5) as well as provisional (i.e., 6 and 7) ITU-T Y.1541 QoS Classes.The descriptive phrase taken from the guidance of section 5.3.5 of Y.1541 is defined for QoS classes 0-5, but not for 6 and 7. SP-1 has also used section 3.2 of RFC 4594 to define the "network control" generic PHB. In this example, SP-B uses IETF Diffserv PHBs, with the application guidance defined in the pargraph number from [RFC 4594] cited to the right of the IETF defined Diffserv PHB acronym in the SP-specific PHB columns of the table..

The format and meaning of these tables is the same as described in the previous example. Also as described above, SP-A and SP-B have agreed on an equivalence to some of the generic PHB names (and hence likely similar queuing behaviors as well), as indicated in the fourth column of each of the tables. These generic PHB terms are an extension of those described in [aggr].

Table 3: Example Usage of EXP↔PHB Mappings at SP-A's ASBR

SP-A EXP↔PHB Mapping Incoming from SP-A Domain		Generic PHB "Name"	SP-B EXP↔PHB Mapping Outgoing to MPLS-ICI		
EXP	**Y.1541 QoS Class, Guidance "PHB"**		**Queue**	**RFC 4594 PHB**	**EXP**
6	2, Signaling, RFC 4594 3.2	Network Control	QB.1	EF, 4.1	5
5	6	Real-Time	QB.1	EF, 4.1	5
3	7	Near Real-Time	QB.2	AF41, 4.5	4
2	3, Transaction Data, interactive	Non Real Time	QB.3	AF31, 4.7	3
1	4, Low Loss Only	Best Effort	QB.4	BE, 4.9	0

SP-B EXP↔PHB Mapping Incoming from MPLS-ICI		Generic PHB "Name"	SP-A EXP↔PHB Mapping Outgoing to SP-A Domain		
EXP	**RFC 4594 PHB**		**Queue**	**Y.1541 QoS Class, Guidance "PHB"**	**EXP**
		Network Control	QA.1	2, Signaling, RFC 4594 3.2	6
5	EF, 4.1	Real-Time	QA.2	6	5
4	AF41, 4.5	Near Real-Time	QA.3	7	4
3	AF31/AF21, 4.7/4.8	Non Real-Time	QA.4	3, Transaction Data, interactive	3
0	BE, 4.9	Best Effort	QA.5	4, Low Loss Only	0

Table 4: Example Usage of EXP↔PHB Mappings at SP-B's ASBR

SP-B EXP↔PHB Mapping Incoming from SP-B Domain		**Generic PHB "Name"**	**SP-A EXP↔PHB Mapping Outgoing to MPLS-ICI**		
EXP	**RFC 4594 PHB**		**Queue**	**Y.1541 QoS Class, Guidance "PHB"**	**EXP**
5	EF, 3.2	Real-Time	QA.1	6	5
4	AF41, 4.5	Near Real-Time	QA.2	7	4
3	AF31, 4.7	Non Real-Time 1	QA.3	3, Transaction Data, interactive	3
2	AF21, 4.8	Non Real-Time 2	QA.3	3, Transaction Data, interactive	3
0	BE, 4.9	Best Effort	QA.4	4, Low Loss Only	0

SP-A EXP↔PHB Mapping Incoming from MPLS-ICI		**Generic PHB "Name"**	**SP-B EXP↔PHB Mapping Outgoing to SP-B Domain**		
EXP	**Y.1541 QoS Class, Guidance "PHB"**		**Queue**	**RFC 4594 PHB**	**EXP**
5	2, Signaling/ 6	Real Time	QB.1	EF, 4.1	5
3	7	Near Real-Time	QB.2	AF41, 4.5	4
2	3, Transaction Data, interactive	Non Real Time 1	QB.3	AF31, 4.7	3
		Non Real Time 2	QB.4	AF21, 4.8	2
1	4, Low Loss Only	Best Effort	QB.5	BE, 4.9	0

In this example, the service providers are more closely aligned in terms of the generic PHB names and the EXP-bit assignments than in the previous example. However, there are some differences and the EXP↔Mapping functions are still needed. Both SP-A and SP-B have five PHBs. There is a many to one mapping in SP-A's ASBR where packets in ITU QoS classes 2 and 6 destined for SP-B are mapped to IETF PHB EF, and a null mapping for network control since SP-B has no such generic PHB. There is a many to one mapping in SP-B's ASBR where AF31 and AF21 are both mapped to

ITU QoS class 3 in SP-A, and a null mapping since SP-A has support for only one non real time generic PHB while SP-B supports two non real time generic PHBs.

AppendixII.4 Example of Canoncial Forwarding Behavior Mapping

Appendix VI of ITU-T Recommendation Y.1541 [Y.1541] provides the following mapping between the two IP transfer capabilities defined in ITU-T Recommendation Y.1221 and IETF Diffserv Per Hop Behavior (PHB) specifications [RFC2475] as shown below.

Table VI.1/Y.1541 – Association of Y.1541 QoS classes with Y.1221 transfer capabilities and differentiated services PHBs

Y.1221 transfer capabilities	Associated DiffServ PHBs	IP QoS class	Remarks
Best-effort (BE)	Default	Unspecified QoS class 5	A legacy IP service, when operated on a lightly loaded network may achieve a good level of IP QoS.
Delay-sensitive Statistical Bandwidth (DSBW)	AF	QoS classes 2, 3, 4	The IPLR objective only applies to the IP packets in the higher priority levels of each AF class. The IPTD applies to all packets.
Dedicated Bandwidth (DBW)	EF	QoS classes 0 and 1	

Service providers connected via an MPLS-ICI could adopt this common, canonical definition of a one-one mapping of three forwarding behaviors. However, they could still use different EXP bits to represent them and a remarking would be required on the MPLS-ICI if they did so. A mapping of EXP bits to a local forwarding behavior is still also required in this case.

END OF DOCUMENT