TEAS Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 10, 2015

Tarek Saad, Ed.
Rakesh Gandhi, Ed.
Zafar Ali
Cisco Systems, Inc.
Robert H. Venator
Defense Information Systems Agency
Yuji Kamite
NTT Communications Corporation

March 9, 2015

Extensions to Resource Reservation Protocol For Re-optimization of Loosely Routed Point-to-Multipoint Traffic Engineering LSPs
draft-ietf-teas-p2mp-loose-path-reopt-02

Abstract

For a Traffic Engineered (TE) Point-to-Multipoint (P2MP) Label Switched Path (LSP), it is preferable in some cases to re-evaluate and re-optimize the entire P2MP-TE LSP by re-signaling all its Source-to-Leaf (S2L) sub-LSP(s). Existing mechanisms, a mechanism for an ingress Label Switched Router (LSR) to trigger a new path re-evaluation request and a mechanism for a mid-point LSR to notify an availability of a preferred path, operate on an individual or a sub-group of S2L sub-LSP(s) basis only.

This document defines RSVP-TE signaling extensions to allow an ingress node of a P2MP-TE LSP to request the re-evaluation of the entire LSP tree containing one or more S2L sub-LSPs whose paths are loose (or abstract) hop expanded, and for a mid-point LSR to notify to the ingress node that a preferable tree exists for the entire P2MP-TE LSP. For re-optimizing a group of S2L sub-LSP(s) in a tree, an S2L sub-LSP descriptor list can be used to signal one or more S2L sub-LSPs in an RSVP message. This document defines markers to indicate beginning and end of an S2L sub-LSP descriptor list when the RSVP message needs to be fragmented due to large number of S2L sub-LSPs in the message when performing sub-group based re-optimization.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute

working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

Copyright Notice

Table of Contents

1. Introduction

This document defines Resource Reservation Protocol - Traffic Engineering (RSVP-TE) [RFC2205] [RFC3209] signaling extensions for re-optimizing loosely routed Point-to-Multipoint (P2MP) Traffic Engineered (TE) Label Switched Paths (LSPs) [RFC4875] in an Multi-Protocol Label Switching (MPLS) and/or Generalized MPLS (GMPLS) networks.

A P2MP-TE LSP is comprised of one or more source-to-leaf (S2L) sub-LSPs. A loosely routed P2MP-TE S2L sub-LSP is defined as one whose path does not contain the full explicit route identifying each node along the path to the egress node at the time of its signaling by the ingress node. Such an S2L sub-LSP is signaled with no Explicit Route Object (ERO) [RFC3209], or with an ERO that contains at least one loose hop, or with an ERO that contains an abstract node that is not a simple abstract node (that is, an abstract node that identifies more than one node). This is often the case with inter-domain P2MP-TE LSPs where Path Computation Element (PCE) is not used [RFC5440].

As per [RFC4875], an ingress node may re-optimize the entire P2MP-TE LSP by re-signaling all its S2L sub-LSP(s) or may re-optimize individual or group of S2L sub-LSP(s) i.e. individual or group of destination(s).

[RFC4736] defines RSVP signaling extensions for re-optimizing loosely routed Point-to-Point (P2P) TE LSP(s) as follows:

o A mid-point LSR that expands loose next-hop(s) sends a solicited or unsolicited PathErr with the Notify error code (25 as defined in [RFC3209]) with sub-code 6 to indicate "Preferable Path Exists" to the ingress node.

o An ingress node triggers a path re-evaluation request at all mid-point LSR(s) that expands loose next-hop(s) by setting the "Path Re-evaluation Request" flag (0x20) in SESSION_ATTRIBUTES Object in the Path message.

o The ingress node upon receiving this PathErr either solicited or unsolicited initiates re-optimization of the LSP with a different LSP-ID.

Following Sections discuss the issues that may arise when using existing mechanisms defined in [RFC4736] for re-optimizing loosely routed P2MP-TE LSPs.

1.1. Loosely Routed Inter-domain P2MP-TE LSP Tree

An example of a loosely routed inter-domain P2MP-TE LSP tree is shown in Figure 1. In this example, the P2MP-TE LSP tree consists of 3 S2L sub-LSPs, to destinations (i.e. leafs) R10, R11 and R12 from the ingress node (i.e. source) R1. Nodes R2 and R5 are branch nodes and nodes ABR3, ABR4, ABR7, ABR8 and ABR9 are area border routers. For the S2L sub-LSP to destination R10, nodes ABR3, ABR7 and R10 are defined as loose hops. For the S2L sub-LSP to destination R11, nodes ABR3, ABR8 and R11 are defined as loose hops. For the S2L sub-LSP to destination R12, nodes ABR4, ABR9 and R12 are defined as loose hops.

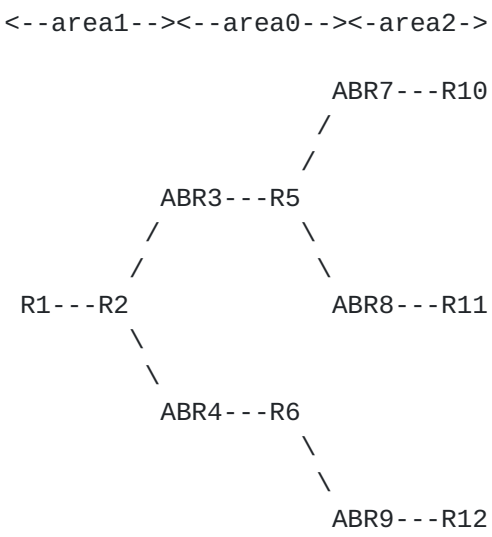


Figure 1: An Example of Loosely Routed Inter-domain P2MP-TE LSP Tree

1.2. Existing Mechanism For Tree-Based P2MP-TE LSP Re-optimization

[RFC4736] does not define signaling extensions specific for re-optimizing entire P2MP-TE LSP tree. Mechanisms defined in [RFC4736] can be used for signaling the re-optimization of individual or group of S2L sub-LSP(s). However, to use [RFC4736] mechanisms for re-optimizing an entire P2MP-TE LSP tree, an ingress node needs to send the path re-evaluation requests on all (typically 100s of) S2L sub-LSPs and the mid-point LSR to notify PathErrs for all S2L sub-LSPs. Such mechanisms may lead to the following issues:

o A mid-point LSR that expands loose next-hop(s) may have to accumulate the received path re-evaluation request(s) for all S2L sub-LSPs (e.g. by using a wait timer) and interpret them as a re-optimization request for the whole P2MP-TE LSP tree. Otherwise, a

mid-point LSR may prematurely notify "Preferable Path Exists" for one or a sub-set of S2L sub-LSPs.

o Similarly, the ingress node may have to heuristically determine when to perform entire P2MP-TE LSP tree re-optimization versus per S2L sub-LSP re-optimization, for example, to delay re-optimization long enough to allow all PathErr(s) to be received. Such procedures may produce undesired results due to timing related issues.

o The ingress node that receives (un)solicited PathErr notification(s) for individual S2L sub-LSP(s), may prematurely start re-optimizing the sub-set of S2L sub-LSPs. However, as mentioned in [RFC4875] Section 14.2, such sub-group based re-optimization procedure may result in data duplication that can be avoided if the entire P2MP-TE LSP tree is re-optimized using a different LSP-ID, especially if the ingress node eventually receives PathErr notifications for all S2L sub-LSPs of the P2MP-TE LSP tree.

In order to address above mentioned issues and to align re-optimization of P2MP-TE LSP with P2P LSP [RFC4736], there is a need for a mechanism to trigger re-optimization of the LSP tree by re-signaling all S2L sub-LSPs with a different LSP-ID. To meet this requirement, this document defines RSVP-TE signaling extensions for the ingress node to trigger the re-evaluation of the P2MP LSP tree on every hop that has a next-hop defined as a loose or abstract hop for one or more S2L sub-LSP path, and a mid-point LSR to signal to the ingress node that a preferable LSP tree exists (compared to the current path) or that the whole P2MP-TE LSP must be re-optimized (because of maintenance required on the TE LSP path).

1.3. Existing Mechanism For Sub-Group-Based P2MP-TE LSP Re-optimization

Based on [RFC4875] (Section 14.2 "Sub-Group-Based Re-Optimization"), an ingress node may trigger path re-evaluation requests using the procedures defined in [RFC4736] for a set of S2L sub-LSPs and combining multiple Path messages using S2L sub-LSP descriptor list. Similarly, a mid-point LSR may send a PathErr message (with Error code 25, sub-code 6) containing a list of S2L sub-LSPs transiting through the LSR using an S2L sub-LSP descriptor list to notify the ingress node. This method can be used for re-optimizing a sub-group of S2L sub-LSPs within an LSP tree using the same LSP-ID. This method can alleviate the scale issue associated with sending RSVP messages for individual S2L sub-LSPs. However, this procedure can lead to the following issues when used to re-optimize the LSP tree:

o Path message that is intended to carry the path re-evaluation

request as defined in [RFC4736] with a full list of S2L sub-LSPs in S2L sub-LSPs descriptor list will be decomposed at branching LSRs, and only a subset of the S2L sub-LSPs that are routed over the same next-hop will be added in the descriptor list of the Path message propagated to downstream mid-point LSRs. Consequently, when a preferable path exists at such mid-point LSRs, the PathErr can only include the sub-set of S2L sub-LSPs traversing the LSR. In this case, at the ingress node there is no way to distinguish which mode of re-optimization to invoke, i.e. sub-group based re-optimization using the same LSP-ID or tree based re-optimization using a different LSP-ID.

o An LSR may fragment a large RSVP message (when a combined message may not be large enough to fit all S2L sub-LSPs). In this case, the ingress node may receive multiple PathErrs with sub-sets of S2L sub-LSPs in each (either due to the combined Path message got fragmented or combined PathErr message got fragmented) and would require additional logic to infer to re-optimize the LSP tree (for example, waiting for some time to aggregate all possible PathErr messages before taking an action).

In order to address the above mentioned issue due to the RSVP message fragmentation, this document defines markers to indicate beginning and end of an S2L sub-LSP descriptor list when combining large number of S2L sub-LSPs in an RSVP message.

2. Conventions Used in This Document

2.1. Key Word Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The reader is assumed to be familiar with the terminology in [RFC4875] and [RFC4736].

2.2. Abbreviations

ABR: Area Border Router.

AS: Autonomous System.

ERO: Explicit Route Object.

LSR: Label Switching Router.

TE LSP: Traffic Engineering Label Switched Path.

TE LSP ingress: Head-end/source of the TE LSP.

TE LSP egress: Tail-end/destination of the TE LSP.

2.3. Nomenclatures

Domain: Routing or administrative domain such as an IGP area and an autonomous system.

Interior Gateway Protocol Area (IGP Area): OSPF area or IS-IS level.

Inter-area TE LSP: A TE LSP whose path transits across at least two different IGP areas.

Inter-AS MPLS TE LSP: A TE LSP whose path transits across at least two different Autonomous Systems (ASes) or sub-ASes (BGP confederations).

S2L sub-LSP: Source-to-leaf sub Label Switched Path.

3. Signaling Procedure For Loosely Routed P2MP-TE LSP Re-optimization

3.1. Tree-Based Re-optimization

To evaluate an entire P2MP-TE LSP tree on mid-point LSRs that expand loose next-hop(s), an ingress node MAY send a Path message with "P2MP-TE Tree Re-evaluation Request" defined in this document. The ingress node SHOULD select one of the S2L sub-LSPs of the P2MP-TE LSP tree transiting a mid-point LSR to trigger the re-evaluation request. The ingress node MAY send a re-evaluation request to each border LSR on the path of the LSP tree.

A mid-point LSR that expands loose next-hop(s) for one or more S2L sub-LSP path(s), and that receives a Path message with the "P2MP-TE Tree Re-evaluation Request" bit set:

o The mid-point LSR SHOULD check for a preferable P2MP-TE LSP tree by re-evaluating all S2L sub-LSP(s) that are expanded paths of the loose next-hops of the P2MP-TE LSP.

o If a preferable P2MP-TE LSP tree is found, the mid-point LSR MAY send an RSVP PathErr to the ingress node with Error code 25 (Notify defined in [RFC3209] and sub-code "Preferable P2MP-TE Tree Exists" defined in this document. The mid-point LSR, in turn, SHOULD NOT propagate the "P2MP-TE Tree Re-evaluation Request" bit in subsequent

RSVP Path messages sent downstream for the re-evaluated P2MP-TE LSP.

o If no preferable tree for P2MP-TE LSP can be found, the recommended mode is that the mid-point LSR that expands loose next-hop(s) for one or more S2L sub-LSP path(s) SHOULD propagate the request downstream by setting the "P2MP-TE Tree Re-evaluation Request" bit in the LSP_ATTRIBUTES Object of RSVP Path message.

A mid-point LSR MAY send an unsolicited PathErr message with "Preferable P2MP-TE Tree Exists" PathErr to the ingress node to notify of a preferred P2MP-TE LSP tree when it determines it exists. In this case, the mid-point LSR that expands loose next-hop(s) for one or more S2L sub-LSP path(s) SHOULD select one of the S2L sub-LSP(s) of the P2MP-TE LSP tree to send this PathErr message to the ingress node.

The sending of an RSVP PathErr Notify message "Preferable P2MP-TE Tree Exists" to the ingress node SHALL notify the ingress node of the existence of a preferable P2MP-TE LSP tree and upon receiving this PathErr, the ingress node MAY trigger re-optimization of the LSP using a different LSP-ID.

3.2. Sub-Group-Based Re-optimization Using Markers

It might be preferable, as per [RFC4875], to re-optimize the entire P2MP-TE LSP by re-signaling all of its S2L sub-LSP(s) (Section 14.1, "Make-before-Break") or to re-optimize individual or group of S2L sub-LSP(s) i.e. individual or group of destination(s) (Section 14.2 "Sub-Group-Based Re-Optimization" in [RFC4875]), both using the same LSP-ID. For loosely routed S2L sub-LSPs, this can be achieved by using the procedures defined in [RFC4736] to re-optimize one or more S2L sub-LSP(s) of the P2MP-TE LSP.

An ingress node may trigger path re-evaluation requests using the procedures defined in [RFC4736] for a set of S2L sub-LSPs by combining multiple Path messages using an S2L sub-LSP descriptor list [RFC4875]. An S2L sub-LSP descriptor list is created using a series of S2L_SUB_LSP Objects as defined in [RFC4875]. Similarly, a mid-point LSR may send a PathErr message (with Error code 25, sub-code 6, Preferable Path Exists) containing a list of S2L sub-LSPs transiting through the LSR using an S2L sub-LSP descriptor list to notify the ingress node of preferable paths available.

As per [RFC4875] (Section 5.2.3, "Transit Fragmentation of Path State Information"), when a Path message is not large enough to fit all S2L sub-LSPs in the descriptor list, an LSR may fragment the message. In this case, the LSR MAY add S2L_SUB_LSP_MARKER_BEGIN and

S2L_SUB_LSP_MARKER_END Objects defined in this document at the beginning and at the end of the S2L sub-LSP descriptor list, respectively.

Both S2L_SUB_LSP_MARKER_BEGIN and S2L_SUB_LSP_MARKER_END Objects defined in this document are optional. However, a node MUST add the S2L_SUB_LSP_MARKER_END Object if it has added S2L_SUB_LSP_MARKER_BEGIN Object in the S2L sub-LSP descriptor list.

A mid-point LSR SHOULD wait to accumulate all S2L sub-LSPs before attempting to re-evaluate preferable path when a Path message for "Path Re-evaluation Request" is received with S2L_SUB_LSP_MARKER_BEGIN Object. An ingress node SHOULD wait to accumulate all S2L sub-LSPs before attempting to trigger re-optimization when a PathErr message with "Preferable Path Exists" is received with S2L_SUB_LSP_MARKER_BEGIN Object.

New objects S2L_SUB_LSP_MARKER_BEGIN and S2L_SUB_LSP_MARKER_END defined in this document have a wider applicability other than the P2MP-TE LSP re-optimization but it is outside the scope of this document.

4. Message and Object Definitions

4.1. P2MP-TE Tree Re-evaluation Request Flag

In order to trigger a tree re-evaluation request, a new flag is defined in Attributes Flags TLV of the LSP_ATTRIBUTES Object [RFC5420] as follows:

Bit Number (to be assigned by IANA): P2MP-TE Tree Re-evaluation Request flag

The "P2MP-TE Tree Re-evaluation Request" flag is meaningful in a Path message of a P2MP-TE S2L sub-LSP and is inserted by the ingress node.

4.2. Preferable P2MP-TE Tree Exists Path Error Sub-code

In order to indicate to an ingress node that a preferable P2MP-TE LSP tree exists, the following new sub-code for PathErr code 25 (Notify Error) [RFC3209] is defined:

Sub-code (to be assigned by IANA): Preferable P2MP-TE Tree Exists sub-code

When a preferable path for P2MP-TE LSP tree exists, the mid-point LSR

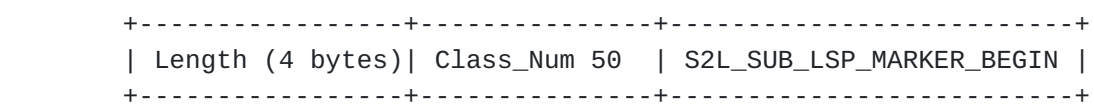
sends a solicited or unsolicited "Preferable P2MP-TE Tree Exists" PathErr notification to the ingress node of the P2MP-TE LSP.

4.3. Markers For S2L sub-LSP Descriptor

An S2L_SUB_LSP Object [RFC4875] identifies a particular S2L sub-LSP belonging to the P2MP-TE LSP. An S2L sub-LSP descriptor list is created using a series of S2L_SUB_LSP Objects as defined in [RFC4875]. In order to indicate the beginning and end of the S2L sub-LSP descriptor list when the RSVP message needs to be fragmented due to large number of S2L sub-LSPs, the following new types are defined for the S2L_SUB_LSP Object [RFC4875].

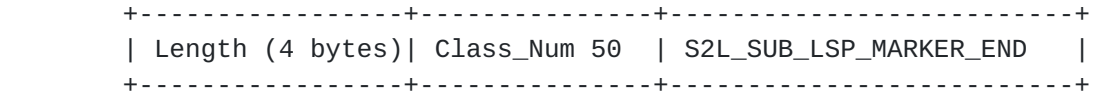
S2L_SUB_LSP_MARKER_BEGIN :

Class-Num 50, C-Type TBA by IANA



S2L_SUB_LSP_MARKER_END :

Class-Num 50, C-Type TBA by IANA



The S2L_SUB_LSP_MARKER_BEGIN Object is added before adding the first S2L_SUB_LSP_IPv4 or S2L_SUB_LSP_IPv6 Object and the S2L_SUB_LSP_MARKER_END Object is added after adding the last S2L_SUB_LSP_IPv4 or S2L_SUB_LSP_IPv6 Object in the S2L sub-LSP descriptor list.

5. Compatibility

The LSP_ATTRIBUTES Object has been defined in [RFC5420] with class numbers in the form 11bbbbbb, which ensures compatibility with non-supporting nodes. Per [RFC2205], nodes not supporting this extension will ignore the new flag defined in this document but forward it without modification.

The S2L_SUB_LSP_MARKER_BEGIN and S2L_SUB_LSP_MARKER_END Objects have been defined with class numbers in the form 11bbbbbb, which ensures

compatibility with non-supporting nodes. Per [RFC2205], nodes not supporting new S2L_SUB_LSP_MARKER_BEGIN and S2L_SUB_LSP_MARKER_END Objects will ignore them but forward it without modification.

6. Security Considerations

This document defines RSVP-TE signaling extensions to allow an ingress node of a P2MP-TE LSP to request the re-evaluation of the entire LSP tree, and for a mid-point LSR to notify the ingress node of the existence of a preferable tree by sending a PathErr. As per [RFC4736], in the case of a P2MP-TE LSP S2L sub-LSP spanning multiple domains, it may be desirable for a mid-point LSR to modify the RSVP PathErr message defined in this document to preserve confidentiality across domains. Furthermore, an ingress node may decide to ignore this PathErr message coming from a mid-point LSR residing in another domain. Similarly, a mid-point LSR may decide to ignore the P2MP-TE tree re-evaluation request originating from another ingress domain.

This document also defines markers to indicate beginning and end of an S2L sub-LSP descriptor list when combining large number of S2L sub-LSPs in an RSVP message and the message needs to be fragmented. The introduction of these markers, by themselves, introduce no additional information to signaling. For a general discussions on MPLS and GMPLS related security issues, see the MPLS/GMPLS security framework [RFC5920].

7. IANA Considerations

IANA is requested to administer assignment of new values for namespace defined in this document and summarized in this section.

7.1. P2MP-TE Tree Re-evaluation Request Flag

IANA maintains a name space for RSVP-TE TE parameters "Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Parameters" (see http://www.iana.org/assignments/rsvp-te-parameters). From the registries in this name space "Attribute Flags", allocation of new flag is requested (Section 4.1).

The following new flag is defined for the Attributes Flags TLV in the LSP_ATTRIBUTES Object [RFC5420]. The numeric value is to be assigned by IANA.

o P2MP-TE Tree Re-evaluation Request Flag:

Bit No	Attribute Flag Name	Carried in Path	Carried in Resv	Carried in RRO	Reference
TBA by IANA	P2MP-TE Tree Re-evaluation	Yes	No	No	This document

7.2. Preferable P2MP-TE Tree Exists Path Error Sub-code

IANA maintains a name space for RSVP protocol parameters "Resource Reservation Protocol (RSVP) Parameters" (see http://www.iana.org/assignments/rsvp-parameters). From the sub-registry "Sub-Codes - 25 Notify Error" in registry "Error Codes and Globally-Defined Error Value Sub-Codes", allocation of a new error code is requested (Section 4.2).

As defined in [RFC3209], the Error Code 25 in the ERROR SPEC Object corresponds to a Notify Error PathErr. This document adds a new sub-code for this PathErr as follows:

o Preferable P2MP-TE Tree Exists sub-code:

Sub-code value	Sub-code Description	PathErr Code	PathErr Name	Reference
TBA by IANA	Preferable P2MP-TE Tree Exists	25	Notify Error	This document

7.3. BEGIN and END Markers For S2L sub-LSP Descriptor

IANA maintains a name space for RSVP protocol parameters "Resource Reservation Protocol (RSVP) Parameters" (see http://www.iana.org/assignments/rsvp-parameters). From the sub-registry "Class Types or C-Types 50 S2L_SUB_LSP" in registry "Class Names, Class Numbers, and Class Types", allocation of new C-Types is requested (Section 4.3).

As defined in [RFC4875], S2L_SUB_LSP Object is defined with Class-Number 50 to identify a particular S2L sub-LSP belonging to the P2MP-TE LSP. This document adds two new object types for this object as follows:

o S2L_SUB_LSP_MARKER_BEGIN and S2L_SUB_LSP_MARKER_END Object types:

C-Type value	Description	Reference
TBA by IANA	S2L_SUB_LSP_MARKER_BEGIN	This document
TBA by IANA	S2L_SUB_LSP_MARKER_END	This document

8. Acknowledgments

The authors would like to thank Loa Andersson, Sriganesh Kini, Curtis Villamizar, Dimitri Papadimitriou and Nobo Akiya for reviewing this document. The authors would also like to thank Ling Zeng for implementing mechanisms defined in this document.

9. References

9.1. Normative References

[RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC4875] Aggarwal, R., Papadimitriou, D., and S. Yasukawa, "Extensions to Resource Reservation Protocol Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, May 2007.

[RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.

9.2. Informative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4736] Vasseur, JP., Ikejiri, Y. and Zhang, R, "Reoptimization of Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Loosely Routed Label Switched Path (LSP)", RFC 4736, November 2006.

[RFC5440] Vasseur, JP., Ed., and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

[RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

Author's Addresses

Tarek Saad (editor)
Cisco Systems

Email: tsaad@cisco.com

Rakesh Gandhi (editor)
Cisco Systems

Email: rgandhi@cisco.com

Zafar Ali
Cisco Systems

Email: zali@cisco.com

Robert H. Venator
Defense Information Systems Agency

Email: robert.h.venator.civ@mail.mil

Yuji Kamite
NTT Communications Corporation

Email: y.kamite@ntt.com