

MPLS Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: October 14, 2013

Zafar Ali  
Rakesh Gandhi  
Tarek Saad  
Cisco Systems, Inc.  
Robert H. Venator  
Defense Information Systems Agency  
Yuji Kamite  
NTT Communications Corporation  
April 12, 2013

**Signaling RSVP-TE P2MP LSPs in an Inter-domain Environment**  
**draft-ietf-mpls-inter-domain-p2mp-rsvp-te-lsp-01**

Abstract

Point-to-MultiPoint (P2MP) Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering Label Switched Paths (TE LSPs) are established using signaling procedures defined in [RFC4875]. However, [RFC4875] does not address several issues that arise when a P2MP-TE LSP is signaled in inter-domain networks. One such issue is the computation of a loosely routed inter-domain P2MP-TE LSP paths that are re-merge free. Another issue is the reoptimization of the inter-domain P2MP-TE LSP tree vs. an individual destination(s), since the loosely routing domain ingress border node is not aware of the reoptimization scope. This document defines the required protocol extensions needed for establishing and reoptimizing P2MP MPLS and GMPLS TE LSPs in inter-domain networks.

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."

This Internet-Draft will expire on October 14, 2013.

## Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Expires October 14, 2013

[Page 2]

## Table of Contents

|                       |  |                    |
|-----------------------|--|--------------------|
| <a href="#">1.</a>    | <a href="#">Introduction . . . . .</a>                                 | <a href="#">4</a>  |
| <a href="#">1.1.</a>  | <a href="#">Summary of Solutions . . . . .</a>                         | <a href="#">5</a>  |
| <a href="#">1.2.</a>  | <a href="#">Path Computation Techniques . . . . .</a>                  | <a href="#">6</a>  |
| <a href="#">1.3.</a>  | <a href="#">Use cases . . . . .</a>                                    | <a href="#">6</a>  |
| <a href="#">2.</a>    | <a href="#">Conventions used in this document . . . . .</a>            | <a href="#">7</a>  |
| <a href="#">3.</a>    | <a href="#">Control Plane Solution For Re-merge Handling . . . . .</a> | <a href="#">7</a>  |
| <a href="#">3.1.</a>  | <a href="#">Single Border Node For All S2Ls . . . . .</a>              | <a href="#">7</a>  |
| <a href="#">3.2.</a>  | <a href="#">Crankback and PathErr Signaling Procedure . . . . .</a>    | <a href="#">7</a>  |
| <a href="#">4.</a>    | <a href="#">Data Plane Solution For Re-merge Handling . . . . .</a>    | <a href="#">9</a>  |
| <a href="#">4.1.</a>  | <a href="#">P2MP-TE Re-merge Recording Request Flag . . . . .</a>      | <a href="#">9</a>  |
| <a href="#">4.2.</a>  | <a href="#">P2MP-TE Re-merge Present Flag . . . . .</a>                | <a href="#">9</a>  |
| <a href="#">4.3.</a>  | <a href="#">Signaling Procedure . . . . .</a>                          | <a href="#">10</a> |
| <a href="#">5.</a>    | <a href="#">Intra-domain P2MP-TE LSP Re-merge Handling . . . . .</a>   | <a href="#">11</a> |
| <a href="#">6.</a>    | <a href="#">Reoptimization Handling . . . . .</a>                      | <a href="#">12</a> |
| <a href="#">6.1.</a>  | <a href="#">P2MP-TE Tree Re-evaluation Request Flag . . . . .</a>      | <a href="#">12</a> |
| <a href="#">6.2.</a>  | <a href="#">Preferable P2MP-TE Tree Exists sub-code . . . . .</a>      | <a href="#">12</a> |
| <a href="#">6.3.</a>  | <a href="#">Signaling Procedure . . . . .</a>                          | <a href="#">12</a> |
| <a href="#">7.</a>    | <a href="#">Compatibility . . . . .</a>                                | <a href="#">14</a> |
| <a href="#">8.</a>    | <a href="#">Security Considerations . . . . .</a>                      | <a href="#">14</a> |
| <a href="#">9.</a>    | <a href="#">IANA Considerations . . . . .</a>                          | <a href="#">14</a> |
| <a href="#">9.1</a>   | <a href="#">P2MP-TE Re-merge Recording Request Flag . . . . .</a>      | <a href="#">14</a> |
| <a href="#">9.2</a>   | <a href="#">P2MP-TE Tree Re-evaluation Request Flag . . . . .</a>      | <a href="#">15</a> |
| <a href="#">9.3</a>   | <a href="#">P2MP-TE Re-merge Present Flag . . . . .</a>                | <a href="#">15</a> |
| <a href="#">9.4</a>   | <a href="#">Preferable P2MP-TE Tree Exists sub-code . . . . .</a>      | <a href="#">15</a> |
| <a href="#">10.</a>   | <a href="#">Acknowledgments . . . . .</a>                              | <a href="#">15</a> |
| <a href="#">11.</a>   | <a href="#">References . . . . .</a>                                   | <a href="#">16</a> |
| <a href="#">11.1.</a> | <a href="#">Normative References . . . . .</a>                         | <a href="#">16</a> |
| <a href="#">11.2.</a> | <a href="#">Informative References . . . . .</a>                       | <a href="#">16</a> |
|                       | <a href="#">Author's Addresses . . . . .</a>                           | <a href="#">17</a> |

Expires October 14, 2013

[Page 3]

## 1. Introduction

[RFC4875] describes procedures to set up Point-to-Multipoint (P2MP) Traffic Engineering Label Switched Paths (TE LSPs) for use in MultiProtocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

As with Point-to-Point (P2P) TE LSPs, P2MP TE LSP state is managed using RSVP messages. While the use of RSVP messages is mostly similar to their P2P counterpart, P2MP LSP state differs from P2P LSP in a number of ways. In particular, the P2MP LSP must also handle the "re-merge" problem described in [\[RFC4875\] section 18](#).

The term "re-merge" refers to the situation when two source-to-leaf (S2L) sub-LSPs branch at some point in the P2MP tree, and then intersect again at another node further downstream the tree. This may occur due to discrepancies in the routing algorithms used by different nodes, errors in path calculation or manual configuration, or network topology changes during the establishment of the P2MP LSP. Such re-merges are inefficient due to the unnecessary duplication of data and also consume additional network resources. Consequently one of the requirements for signaling P2MP LSPs is to choose a P2MP path that is re-merge free. In some deployments, it may also be required to signal P2MP-TE LSPs that are both re-merge and crossover free [\[RFC4875\]](#).

For the purposes of this document, a domain is considered to be any collection of network elements within a common sphere of address management or path computational responsibility. Examples of such domains include Interior Gateway Protocol (IGP) areas and Autonomous Systems (ASes). A border node is a node between different routing domains.

The re-merge free requirement becomes more acute to address when P2MP LSP spans multiple domains. This is because in an inter-domain environment, the ingress node may not have topological visibility into other domains to be able to compute and signal a re-merge free P2MP LSP. In that case, the border node for a new domain will be given loose next hops for one or more destinations in a P2MP LSP. A border node computes paths in its domain by individually expanding the loose next hops for the destinations when signaled to it. A border node can ensure that it computes the re-merge free paths while performing loose hop ERO expansions by individually grafting destinations. Note that the computed P2MP tree by a border node in this case may not be optimal. When processing a Path message, the border node may not have knowledge of all the destinations of the P2MP LSP; for example, in the case when not all S2L sub-LSPs pass through this border node. In that case, existing protocol mechanisms

Expires October 14, 2013

[Page 4]

do not provide sufficient information for it to be able to expand the loose hop(s) such that the overall P2MP LSP tree is guaranteed to be re-merge free.

[RFC4875] specifies two approaches to handle re-merge conditions. The first method is based on control plane handling the re-merge. In this case the node detecting the re-merge condition, i.e. the re-merge node initiates the removal of the re-merge sub-LSP(s) by sending a PathErr message(s) towards the ingress node. However, this can lead to a deadlock in setting up the P2MP LSP in certain cases; for example, when the first S2L setup causes the re-merge with all subsequent S2Ls in the tree. The second method is based on the data plane handling the re-merge condition. In this case, the re-merge node allows the re-merge condition to persist, but data from all but one incoming interface is dropped at the re-merge node. This ensures that duplicate data is not sent on any outgoing interface. However, network resources (such as bandwidth capacity) are wasted as long as re-merge condition persists which is inefficient.

[RFC4736] defines procedures and signaling extensions for reoptimizing an inter-domain P2P LSP. Specifically, an ingress node sends a "path re-evaluation request" to a border node by setting a flag (0x20) in SESSION\_ATTRIBUTES object in a Path message. A border node sends a PathErr code 25 (notify error defined in [RFC3209]) with sub-code 6 to indicate "preferable path exists" to the ingress node. The ingress node upon receiving this PathErr may initiate reoptimization of the LSP. [RFC4736] however does not define a procedure to reoptimize the entire P2MP LSP as a whole tree.

As per [RFC4875] [Section 14](#), for a P2MP LSP, an ingress node may reoptimize the entire P2MP LSP by resignaling all destinations ([Section 14.1](#), "Make-before-Break") or may reoptimize individual the destinations ([Section 14.2](#) "Sub-Group-Based Re-Optimization"). Generally speaking make-before-break is considered available for "whole" P2MP LSP reoptimization, but it can also be used for reoptimizing physical routes for specific sub-LSP(s). The Sub-Group-Based reoptimization is not always applicable because it can lead to data duplication inside the backbone.

### **[1.1. Summary of Solutions](#)**

This document defines RSVP-TE signaling procedures for P2MP LSP to handle the re-merge condition when using either the control plane or data plane approach. The procedures are applicable to both MPLS TE and GMPLS networks.

The control plane solution for the re-merge problem makes use of the crankback signaling mechanism of the RSVP protocol. [RFC5151]



Expires October 14, 2013

[Page 5]

describes such mechanisms for applying crankback to inter-domain P2P LSPs, but does not cover P2MP LSPs. Also, crankback mechanisms for P2MP LSPs are not addressed by [\[RFC4875\]](#). This document describes how crankback signaling extensions for MPLS and GMPLS RSVP-TE defined in [\[RFC4920\]](#) can be used for setting up P2MP TE LSPs to resolve re-merges.

The data plane solution for the re-merge problem described in [\[RFC4875\]](#) is extended by using a new flag in the LSP\_ATTRIBUTES TLV (in a Path message) and a new flag in RRO Attributes Sub-object (in a Resv message) in RSVP. The LSP\_ATTRIBUTES TLV (in a Path message) and RRO Attributes Sub-object (in a Resv message) have been defined in [\[RFC5420\]](#). This document describes how these new flags can be used to handle P2MP re-merge conditions efficiently.

For P2MP LSP, a border node may have loosely routed entire or part of the P2MP LSP by expanding EROs in Path messages of the destinations. Border node does not know with the signaling procedure defined in [\[RFC4736\]](#) if an ingress node is requesting a reoptimization for an individual destination(s) or reoptimization of the entire P2MP tree. Signaling extension and procedure are defined in this document to handle reoptimization of an individual destination(s) and the reoptimization of the entire P2MP tree. Basically, a new query message is defined in LSP\_ATTRIBUTES TLV to request for a "P2MP-TE Tree Re-evaluation" and a new sub-code is defined for PathErr message to indicate "Preferable P2MP-TE Tree Exists".

## **1.2. Path Computation Techniques**

This document focuses on the case where the ingress node does not have full visibility of the topology of all domains and is therefore not able to compute the complete P2MP tree. Rather, it includes loose hops to traverse the domains for which it does not have full visibility and ingress border nodes(s) of each transit domain is responsible for expanding those loose hops.

The solution presented in this document do not guarantee optimization of the overall P2MP tree across all domains. Path Computation Element (PCE) can be used, instead, to address global optimization of the overall P2MP tree.

## **1.3. Use cases**

Service providers having a network with multiple routing domains are interested to use the network for P2MP-TE LSPs. This allows the service providers to use the network to carry multicast and broadcast traffic (such as video). Service providers can deploy the VPLS and MVPN services in the network using inter-domain P2MP TE LSPs. The use

Expires October 14, 2013

[Page 6]

case is for P2MP TE LSPs across multiple routing domains that belong to a single administrative area. Use case for the Multiple administrative domains (e.g. autonomous systems) is outside the scope of this document.

## **2. Conventions used in this document**

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 [\[RFC2119\]](#).

## **3. Control Plane Solution For Re-merge Handling**

It is RECOMMENDED that boundary re-routing is requested for P2MP LSPs traversing multiple domains. This is because border nodes that are expanding loose hops are typically best placed to correct any re-merge errors that occur within their domain, not the ingress node.

### **3.1. Single Border Node For All S2Ls**

It is RECOMMENDED that the ingress node of a P2MP LSP selects the same ingress border node in the loose hop ERO for all sibling S2L sub-LSPs that transit through a given domain. The reason is that it will increase the possibility of re-merge downstream if two or more border nodes have roles simultaneously to expand loose EROs. An ingress border node that performs the loose ERO expansion for individual sub-LSP(s) has the necessary state information for the destinations transiting through its domain to ensure computed P2MP tree is re-merge free.

### **3.2. Crankback and PathErr Signaling Procedure**

As mentioned in [\[RFC4875\]](#), in order to avoid duplicate traffic, the re-merge node MAY initiate the removal of the re-merge S2L sub-LSPs by sending a PathErr message to the ingress node of the S2L sub-LSP.

Crankback procedures for rerouting around failures for P2P RSVP-TE LSPs are defined in [\[RFC4920\]](#). These techniques can also be applied to P2MP LSPs to handle re-merge conditions, as described in this section.

If an ingress border node on the path of the P2MP LSP is unable to find a route that can supply the required resources or that is re-merge free, it MUST generate a PathErr message for the subset of the S2L sub-LSPs which it is not able to route. For this purpose the ingress border node SHOULD try to find a minimum subset of S2L sub-LSPs for which the PathErr needs to be generated towards the ingress

Expires October 14, 2013

[Page 7]

node. These are the S2L sub-LSPs on an incoming interface that has less number of S2L sub-LSPs compared to the second incoming interface that is causing the re-merge condition.

The RSVP-TE Notify messages do not include S2L\_SUB\_LSP objects and cannot be used to send errors for a subset of the S2L sub-LSPs in a Path message. For that reason, the error generating node SHOULD use a PathErr message rather than a Notify message to communicate the error. In the case of a re-merge error, the node SHOULD use the error code "Routing Problem" and the error value "ERO resulted in re-merge" as specified in [[RFC4875](#)].

A border node receiving a PathErr message for a set of S2L sub-LSPs MAY hold the message and attempt to signal an alternate path that can avoid re-merge through its domain for those S2L sub-LSPs that pass through it. However, in the case of a re-merge error for which some of the re-merging S2L sub-LSPs do not pass through the border node, it SHOULD propagate the PathErr upstream towards the ingress node. If the subsequent attempt by the border node is successful, the border node discards the held PathErr and follows the crankback roles of [[RFC4920](#)] and [[RFC5151](#)]. If repeated subsequent attempts by the border node are unsuccessful, the border node MUST send the held PathErr upstream towards the ingress node.

If the ingress node receives a PathErr message with error code "Routing Problem" and error value "ERO resulted in re-merge", then it SHOULD attempt to signal an alternate path through a different domain or through a different border node for the affected S2L sub-LSPs. The ingress node MAY use the error node information from the PathErr for this purpose.

However, it may be that the ingress node or an ingress border node does not have sufficient topology information to compute an Explicit Route that is guaranteed to avoid the re-merge link or node. In this case, Route Exclusions [[RFC4874](#)] may be particularly helpful. To achieve this, [[RFC4874](#)] allows the re-merge information to be presented as route exclusions to force avoidance of the re-merge link or node.

As discussed in [[RFC4920](#)] [section 3.3](#), border node MAY keep the history of PathErrs. In case of P2MP LSPs, ingress node and border nodes may keep re-merge PathErrs in history table until S2L sub-LSPs have been successfully established or until local timer expires.

Expires October 14, 2013

[Page 8]

#### **4. Data Plane Solution For Re-merge Handling**

As mentioned in [[RFC4875](#)], a node may accept the re-merging S2Ls but only send the data from one of these interfaces to its outgoing interfaces. That is, the node MUST drop data from all but one incoming interface causing the re-merge. This ensures that duplicate data is not sent on any outgoing interface. Note that data plane may be either programmed to drop the incoming traffic for the S2L sub-LSP or not programmed at all.

It is desirable to avoid the persistent re-merge condition associated with data plane based solution in the network in order to optimize bandwidth resources in the network.

The following sections define the RSVP-TE signaling extensions for "P2MP- TE Re-merge Recording Request" and "P2MP-TE Re-merge Present" messages.

##### **4.1. P2MP-TE Re-merge Recording Request Flag**

In order to indicate to traversed nodes that P2MP-TE re-merge recording is desired, a new flag in the Attribute Flags TLV of the LSP\_ATTRIBUTES object defined in [[RFC5420](#)] is defined as follows:

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

The "P2MP-TE Re-merge Recording Request" flag is meaningful in a Path message and is inserted by the ingress node or a border node in the LSP\_ATTRIBUTES object.

If the "P2MP-TE Re-merge Recording Request" Flag is set, it implies that the "P2MP-TE Re-merge Present" flag defined in the next section MUST be used to indicate to the ingress and ingress border nodes of the transit domains that a re-merge condition is present for this S2L sub-LSP but accepted, and that incoming traffic is being dropped for this S2L sub-LSP.

The rules of the processing of the Attribute Flags TLV of the LSP\_ATTRIBUTES object follow [[RFC5420](#)].

##### **4.2. P2MP-TE Re-merge Present Flag**

The "P2MP-TE Re-merge Present" Flag is the counter part of the "P2MP-TE Re-merge Recording Request" flag defined above. Specifically, RSVP signaling extension is defined to indicate to the



Expires October 14, 2013

[Page 9]

upstream node of the re-merge condition and that incoming traffic is being dropped for the given S2L.

When a node accepts a re-merge condition by dropping traffic from an incoming interface for an S2L due to the re-merge condition, and if it understands the "P2MP-TE Re-merge Recording Request" in the Attribute Flags TLV of the LSP\_ATTRIBUTES object of the Path message, the node MUST set the newly defined "P2MP-TE Re-merge Present" flag in the RRO Attributes sub-object defined in [[RFC5420](#)] in RRO.

The following new flag for RRO Attributes Sub-object is defined as follows:

Bit Number (same as bit number assigned for "P2MP-TE Re-merge Recording Request" flag): P2MP-TE Re-merge Present flag

The "P2MP-TE Re-merge Present" flag indicates that the S2L is causing a re-merge. The re-merge has been accepted but the incoming traffic on this S2L is dropped by the reporting node.

The rules of the processing of the RRO Attribute Sub-object in the Resv message follow [[RFC5420](#)].

### **4.3. Signaling Procedure**

When a node that does not support data plane based re-merge handling receives an S2L sub-LSP Path message with LSP Attributes sub-object that has "P2MP-TE Re-merge Recording Request" Flag set, and if the S2L is causing a re-merge condition, the node MUST reject the S2L sub-LSP Path message and send the PathErr with the error code "Routing Problem" and the error value "ERO resulted in re-merge" as specified in [[RFC4875](#)]. If a node is capable of data plane based re-merge handling but operator may have disabled it via a configuration, the the node MUST also reject the re-merge and send this PathErr.

When a Path message is received at a transit node for an S2L sub-LSP and "P2MP-TE Re-merge Recording Request" Flag is set in the LSP Attributes sub-object, the node MAY decide to accept the re-merge S2L sub-LSP based on the local policy and node capability. In this case, before the Resv message is sent to the upstream node for this S2L sub-LSP, the node MUST add the RRO Attributes sub-object in the Resv RRO if not already present and set the "P2MP-TE Re-merge Present" Flag if traffic from the incoming interface of this S2L sub-LSP will be dropped. This same incoming interface can still be used for a different S2L sub-LSP in the P2MP LSP to forward traffic and "P2MP-TE Re-merge Present" flag will not be set for that S2L sub-LSP. Note

Expires October 14, 2013

[Page 10]

that rules for adding or modifying the other RRO sub-objects do not change due to this flag.

When a transit node receives a Resv message for an S2L that is causing a re-merge condition, the node MUST set the "P2MP-TE Re-merge Present" flag in the RRO Attributes sub-object in the Resv message if it decides to drop the incoming traffic of this S2L. The "P2MP-TE Re-merge Present" flag in RRO Attribute sub-object is not set for the S2L(s) whose incoming interface is selected to receive and forward the traffic.

An ingress node MAY immediately start sending traffic on all S2Ls in up state even when re-merge conditions are present on some S2Ls of the P2MP LSP.

The proposed signaling extensions allow an ingress node and an ingress border node to have a complete view of the re-merge conditions on the entire S2L path and on all S2Ls of the P2MP tree. The ingress or ingress border node in this case can take appropriate actions to resolve the re-merge conditions and optimize network bandwidth resources usage. This can be achieved by computing and selecting alternate path(s) for the S2L(s) bypassing the re-merge node(s).

The proposed signaling extensions are equally applicable to single domain scenarios.

A node where re-merge is present, may decide to select a different incoming interface to forward traffic from in the future. In that case, a Resv change message with updated "P2MP-TE Re-merge Present" flag in the RRO is sent upstream for all effected S2Ls. For the new set of S2L sub-LSPs whose traffic from the incoming interface is dropped, "P2MP-TE Re-merge Present" flag will be set.

A border node due to local policy MAY remove the record route object from the Resv message of the S2L sub-LSP and propagate Resv message towards the ingress node. When such a policy is provisioned, the border node may attempt to correct the re-merge condition in its domain. If the border node is not able to resolve the re-merge condition, the border node SHOULD send the PathErr with the error code "Routing Problem" and the error value "ERO resulted in re-merge" as specified in [[RFC4875](#)].

## **5. Intra-domain P2MP-TE LSP Re-merge Handling**

Re-merges between S2Ls in a single domain can occur due to provisioning errors or path computation errors in the environment

Expires October 14, 2013

[Page 11]

where IGP-TE or PCE is used. Re-merges can also happen in the environment where static routing or static path selection policy is applied at ingress (e.g., CSPF calculation is disabled due to some operational reasons), regardless of using loose or static hops. In either case, procedures described in this document are equally applicable to the intra-domain (i.e. single domain) P2MP-TE LSPs.

## **6. Reoptimization Handling**

### **6.1. P2MP-TE Tree Re-evaluation Request Flag**

In order to query border nodes to check if a preferable P2MP tree exists, 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 an S2L sub-LSP and is inserted by the ingress node.

### **6.2. Preferable P2MP-TE Tree Exists sub-code**

In order to indicate to an ingress node that a preferable P2MP-TE tree is available, 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 P2MP-TE tree is found, the border node MUST send "Preferable P2MP-TE Tree Exists" PathErr to the ingress node in order to reoptimize the entire P2MP LSP.

### **6.3. Signaling Procedure**

Using signaling procedure defined in [[RFC4736](#)], an ingress node MUST initiate "path re-evaluation request" query to reoptimize a destination in a P2MP LSP. Note that this message MUST be used to reoptimize a single or a sub-set of the destinations in a P2MP LSP. Ingress node MUST send this query in a Path message for each destination it is reoptimizing.

When a Path message for a destination in a P2MP LSP with "path re-evaluation request" flag [[RFC4736](#)] is received at the border node,

Expires October 14, 2013

[Page 12]

it MUST re-compute the loose-hop ERO to see if a preferable path exists for that destination. A border node MUST send PathErr code 25 (notify error defined in [[RFC3209](#)]) with "preferable path exists" sub-code to indicate that a preferable path exists for the requested destination AND border node is capable of per destination reoptimization. A border node MUST terminate the path query. Alternatively, a border node not capable of per destination reoptimization MAY respond with "Preferable P2MP-TE Tree Exists" PathErr by checking for a preferable P2MP tree instead of a preferable single destination.

It is often desired to reoptimize the entire P2MP LSP. In order to query border nodes to check if a preferable P2MP tree exists, an ingress node MUST send a Path message with "P2MP-TE Tree Re-evaluation Request" defined in this document. An ingress node MAY send this message for all destinations in a P2MP LSP or a sub-set of the destinations.

A border node receiving the "P2MP-TE Tree Re-evaluation Request" MUST check for a preferable P2MP LSP for the destinations it is loosely routing by loose-hop ERO expansions. The border node if a preferable P2MP-TE tree is found, MUST reply with "Preferable P2MP-TE Tree Exists" sub-code defined in this document with PathErr 25 (notify error defined in [[RFC3209](#)] and terminate the path query.

Note that a border node MAY send "Preferable P2MP-TE Tree Exists" with PathErr code 25 to indicate the ingress node in order to reoptimize the entire P2MP LSP message unsolicited or in a response to "path re-evaluation query" for a destination or in a response to "P2MP-TE Tree Re-evaluation Request" message.

If an ingress node initiated a "path re-evaluation request" query for a single destination for per S2L sub-LSP reoptimization and receives "Preferable P2MP-TE Tree Exists" PathErr, the ingress node MAY cancel the per S2L reoptimization and initiate P2MP-TE tree reoptimization. This may happen in case when a border node is not capable of per destination reoptimization.

Note that even if per destination reoptimization, not whole P2MP LSP Tree reoptimization, is sufficient, ingress node often needs to re-signal whole P2MP LSP tree to complete route optimization for that destination. In this case, make-before-break reoptimization scheme is used (see [[RFC4875](#)] [Section 14.1](#)), and all S2L sub-LSPs are re-signaled with a different LSP-ID. That is, the procedure of signaling a re-optimization by an ingress node is separate from the matter if PathErr reply was "Preferable Path Exists" or "Preferable P2MP-TE Tree Exists".



Expires October 14, 2013

[Page 13]

## **7. Compatibility**

The LSP\_ATTRIBUTES TLV and RRO Attributes sub-object have been defined [[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 TLV, sub-object and the new flags defined in this document but forward it, unexamined and unmodified, in all messages resulting from this message.

## **8. Security Considerations**

This document does not introduce any additional security issues above those identified in [[RFC3209](#)], [[RFC4875](#)], [[RFC5151](#)], [[RFC4920](#)] and [[RFC5920](#)].

## **9. IANA Considerations**

IANA maintains a name space for RSVP-TE TE parameters "Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Parameters". From the registries in this name space "Attribute Flags" and "Record Route Object Sub-object Flags" allocation of new flags are requested (sections [9.1](#) - [9.3](#)).

IANA also maintains a name space for RSVP protocol parameters "Resource Reservation Protocol (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 9.4](#)).

### **9.1 P2MP-TE Re-merge Recording Request Flag**

The following new flag is defined for the Attributes Flags TLV in the LSP\_ATTRIBUTES object [[RFC5420](#)]. The numeric values are to be assigned by IANA.

- o P2MP-TE Re-merge Recording Request Flag:
  - Bit Number: To be assigned by IANA.
  - Attribute flag carried in Path message: Yes
  - Attribute flag carried in Resv message: No

Expires October 14, 2013

[Page 14]

## **9.2 P2MP-TE Tree Re-evaluation Request Flag**

The following new flag is defined for the Attributes Flags TLV in the LSP\_ATTRIBUTES object [[RFC5420](#)]. The numeric values are to be assigned by IANA.

- o P2MP-TE Tree Re-evaluation Request Flag:
  - Bit Number: To be assigned by IANA.
  - Attribute flag carried in Path message: Yes
  - Attribute flag carried in Resv message: No

## **9.3 P2MP-TE Re-merge Present Flag**

The following new flag is defined for the RRO Attributes sub-object in the RECORD\_ROUTE object [[RFC5420](#)]. The numeric values are to be assigned by IANA.

- o P2MP-TE Re-merge Present Flag:
  - Bit Number: To be assigned by IANA.
  - Attribute flag carried in Path message: No
  - Attribute flag carried in RRO Attributes sub-object in RRO of the Resv message: Yes

## **9.4 Preferable P2MP-TE Tree Exists sub-code**

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 as follows for this PathErr:

- o Preferable P2MP-TE Tree Exists sub-code:
  - Sub-code for Notify PathErr code 25. To be assigned by IANA.

## **10. Acknowledgments**

The authors would like to thank N. Neate for his contributions on the draft.

Expires October 14, 2013

[Page 15]

## **11. References**

### **11.1. Normative References**

- [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.
- [RFC5151] Farrel, A., Ayyangar, A., and JP. Vasseur, "Inter-Domain MPLS and GMPLS Traffic Engineering -- Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 5151](#), February 2008.
- [RFC4920] Farrel, A., Satyanarayana, A., Iwata, A., Fujita, N., and G. Ash, "Crankback Signaling Extensions for MPLS and GMPLS RSVP-TE", [RFC 4920](#), July 2007.
- [RFC5920] L. Fang, Ed., "Security Framework for MPLS and GMPLS Networks", [RFC 5920](#), July 2010.
- [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.
- [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.
- [RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangar, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", [RFC 5420](#), February 2009.

### **11.2. Informative References**

- [RFC4726] Farrel, A., Vasseur, J., and A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", [RFC 4726](#), November 2006.

Expires October 14, 2013

[Page 16]

Author's Addresses

Zafar Ali  
Cisco Systems

Email: [zali@cisco.com](mailto:zali@cisco.com)

Rakesh Gandhi  
Cisco Systems

Email: [rgandhi@cisco.com](mailto:rgandhi@cisco.com)

Tarek Saad  
Cisco Systems

Email: [tsaad@cisco.com](mailto:tsaad@cisco.com)

Robert H. Venator  
Defense Information Systems Agency

Email: [robert.h.venator.civ@mail.mil](mailto:robert.h.venator.civ@mail.mil)

Yuji Kamite  
NTT Communications Corporation

Email: [y.kamite@ntt.com](mailto:y.kamite@ntt.com)



Expires October 14, 2013

[Page 17]