

Network Working Group  
Internet Draft  
Expiration Date: June 2005

R. Aggarwal (Juniper)  
D. Papadimitriou (Alcatel)  
S. Yasukawa (NTT)  
Editors

## Extensions to RSVP-TE for Point to Multipoint TE LSPs

[draft-ietf-mpls-rsvp-te-p2mp-00.txt](http://www.ietf.org/drafts/mpls-rsvp-te-p2mp-00.txt)

### Status of this Memo

By submitting this Internet-Draft, we certify that any applicable patent or IPR claims of which we are aware have been disclosed, and any of which we become aware will be disclosed, in accordance with [RFC 3668](http://www.rfc-editor.org/rfc/rfc3668.txt).

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

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.''

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

### Abstract

This document describes extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for the setup of point-to-multipoint (P2MP) Label Switched Paths (LSPs) in Multi-Protocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks. The solution relies on RSVP-TE without requiring a multicast routing protocol in the Service Provider core. Protocol elements and procedures for this solution are described. There can be various applications for P2MP TE LSPs such as IP multicast. Specification of how such applications will use a P2MP TE LSP is outside the scope of this document.

## 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 [RFC-2119](#) [KEYWORDS].

## Authors' Note

Some of the text in the document needs further discussion between authors and feedback from MPLS WG. This has been pointed out when applicable. A change log and reviewed/updated text will be made available online.

## Table of Contents

<a href="#">1</a>	Introduction.....	<a href="#">3</a>
<a href="#">2</a>	Terminology.....	<a href="#">3</a>
<a href="#">3</a>	Mechanisms.....	<a href="#">4</a>
<a href="#">3.1</a>	P2MP Tunnels.....	<a href="#">4</a>
<a href="#">3.2</a>	P2MP LSP Tunnels.....	<a href="#">5</a>
<a href="#">3.3</a>	P2P Sub-LSPs.....	<a href="#">5</a>
<a href="#">3.3.1</a>	Representation of a P2P sub-LSP.....	<a href="#">5</a>
<a href="#">3.3.2</a>	P2P Sub-LSPs and Path Messages.....	<a href="#">5</a>
<a href="#">3.4</a>	Explicit Route Encoding.....	<a href="#">6</a>
<a href="#">4</a>	Sub-Groups.....	<a href="#">8</a>
<a href="#">5</a>	Path Message Format.....	<a href="#">9</a>
<a href="#">6</a>	Path Message Processing.....	<a href="#">10</a>
<a href="#">6.1</a>	Multiple Path Messages.....	<a href="#">10</a>
<a href="#">6.1.1</a>	Identifying Multiple Path Messages.....	<a href="#">11</a>
<a href="#">6.2</a>	Multiple P2P Sub-LSPs in One Path Message.....	<a href="#">12</a>
<a href="#">7</a>	RESV Message Format.....	<a href="#">13</a>
<a href="#">8</a>	RESV Message Processing.....	<a href="#">14</a>
<a href="#">8.1</a>	RRO Processing.....	<a href="#">15</a>
<a href="#">8.2</a>	Resv Message Throttling.....	<a href="#">15</a>
<a href="#">9</a>	Transit Fragmentation.....	<a href="#">16</a>
<a href="#">10</a>	Grafting.....	<a href="#">16</a>
<a href="#">11</a>	Pruning.....	<a href="#">17</a>
<a href="#">11.1</a>	P2MP TE LSP Teardown.....	<a href="#">18</a>
<a href="#">11.2</a>	Path Tear Message Format.....	<a href="#">19</a>
<a href="#">12</a>	Refresh Reduction.....	<a href="#">19</a>
<a href="#">13</a>	State Management.....	<a href="#">19</a>
<a href="#">13.1</a>	Incremental State Update.....	<a href="#">19</a>
<a href="#">13.2</a>	Combining Multiple Path Messages.....	<a href="#">20</a>
<a href="#">14</a>	Error Processing.....	<a href="#">21</a>
<a href="#">14.1</a>	Branch Failure Handling.....	<a href="#">21</a>
<a href="#">15</a>	Notify and ResvConf Messages.....	<a href="#">22</a>



<a href="#">16</a>	Control of Branch Fate Sharing.....	<a href="#">23</a>
<a href="#">17</a>	Admin Status Change.....	<a href="#">23</a>
18	Label Allocation on LANs with Multiple Downstream Nodes.	24
<a href="#">19</a>	Make-Before-Break.....	<a href="#">24</a>
<a href="#">19.1</a>	P2MP Tree re-optimization.....	<a href="#">24</a>
<a href="#">19.2</a>	Re-optimization of a subset of P2P sub-LSPs .....	<a href="#">24</a>
<a href="#">20</a>	Fast Reroute.....	<a href="#">25</a>
<a href="#">20.1</a>	Facility Backup.....	<a href="#">25</a>
<a href="#">20.2</a>	One to One Backup.....	<a href="#">25</a>
<a href="#">21</a>	Support for LSRs that are not P2MP Capable.....	<a href="#">26</a>
22	Reduction in Control Plane Processing with LSP Hierarchy	28
<a href="#">23</a>	P2MP LSP Tunnel Remerging and Cross-Over.....	<a href="#">28</a>
<a href="#">23.1</a>	PathErr Message Format.....	<a href="#">30</a>
<a href="#">24</a>	New and Updated Message Objects.....	<a href="#">31</a>
<a href="#">24.1</a>	P2MP SESSION Object.....	<a href="#">31</a>
<a href="#">24.2</a>	P2MP LSP Tunnel SENDER_TEMPLATE Object.....	<a href="#">32</a>
<a href="#">24.2.1</a>	P2MP LSP Tunnel IPv4 SENDER_TEMPLATE Object.....	<a href="#">33</a>
<a href="#">24.2.2</a>	P2MP LSP Tunnel IPv6 SENDER_TEMPLATE Object.....	<a href="#">33</a>
<a href="#">24.3</a>	P2P SUB-LSP Object.....	<a href="#">34</a>
<a href="#">24.3.1</a>	P2P IPv4 SUB-LSP Object.....	<a href="#">34</a>
<a href="#">24.3.2</a>	P2P IPv6 SUB-LSP Object.....	<a href="#">35</a>
<a href="#">24.4</a>	FILTER_SPEC Object.....	<a href="#">35</a>
<a href="#">24.5</a>	SUB EXPLICIT ROUTE Object (SER0).....	<a href="#">36</a>
<a href="#">24.6</a>	SUB RECORD ROUTE Object (SRRO).....	<a href="#">36</a>
<a href="#">25</a>	IANA Considerations.....	<a href="#">37</a>
<a href="#">26</a>	Security Considerations.....	<a href="#">37</a>
<a href="#">27</a>	Acknowledgements.....	<a href="#">37</a>
<a href="#">28</a>	Appendix.....	<a href="#">37</a>
<a href="#">28.1</a>	Example.....	<a href="#">37</a>
<a href="#">29</a>	References.....	
.....	<a href="#">39</a>	
<a href="#">30</a>	Authors.....	<a href="#">40</a>
<a href="#">31</a>	Intellectual Property.....	<a href="#">43</a>
<a href="#">32</a>	Full Copyright Statement.....	<a href="#">43</a>
<a href="#">33</a>	Acknowledgement.....	<a href="#">44</a>



## **1. Introduction**

[RFC3209] defines a mechanism for setting up P2P TE LSPs in MPLS networks. [RFC3473] defines extensions to [RFC3209] for setting up P2P TE LSPs in GMPLS networks. However these specifications do not provide a mechanism for building P2MP TE LSPs.

This document defines extensions to RSVP-TE [RFC3209] and [RFC3473] protocol to support P2MP TE LSPs satisfying the set of requirements described in [P2MP-REQ].

This document relies on the semantics of RSVP that RSVP-TE inherits for building P2MP LSP Tunnels. A P2MP LSP Tunnel is comprised of multiple P2P sub-LSPs. These P2P sub-LSPs are set up between the ingress and egress LSRs and are appropriately combined by the branch LSRs using RSVP semantics to result in a P2MP TE LSP. One Path message may signal one or multiple P2P sub-LSPs. Hence the P2P sub-LSPs belonging to a P2MP LSP Tunnel can be signaled using one Path message or split across multiple Path messages.

Path computation and P2MP application specific aspects are outside of the scope of this document.

## **2. Terminology**

This document uses terminologies defined in [RFC3031], [RFC2205], [RFC3209], [RFC3473] and [P2MP-REQ]. In addition the following terms are used in this document.

P2P sub-LSP: A P2MP TE LSP is constituted of one or more P2P sub-LSPs. A P2P sub-LSP refers to the portion of the label switched path from the ingress LSR to a particular egress LSR. The egress LSR is the destination of the P2P sub-LSP.

## **3. Mechanism**

This document describes a solution that optimizes data replication by allowing non-ingress nodes in the network to be replication/branch nodes. A branch node is a LSR that is capable of replicating the incoming data on two or more outgoing interfaces. The solution uses RSVP-TE in the core of the network for setting up a P2MP TE LSP.

The P2MP TE LSP is set up by associating multiple P2P TE sub-LSPs and relying on data replication at branch nodes. This is described further in the following sub-sections by describing P2MP tunnels and how they relate to P2P sub-LSPs.



### **3.1. P2MP Tunnels**

The specific aspect related to P2MP TE LSP is the action required at a branch node, where data replication occurs. For instance, in the MPLS case, incoming labeled data is appropriately replicated to several outgoing interfaces with different labels.

A P2MP TE tunnel comprises of one or more P2MP LSPs referred to as P2MP LSP tunnels. A P2MP TE Tunnel is identified by a P2MP SESSION object. This object contains the P2MP ID defined as a destination identifier, a tunnel ID and an extended tunnel ID.

The fields of a P2MP SESSION object are identical to those of the SESSION object defined in [\[RFC3209\]](#) except that the Tunnel Endpoint Address field is replaced by the P2MP Identifier (P2MP ID) field.

This identifier encodes the P2MP ID and identifies the set of destination(s) of the P2MP LSP Tunnel.

### **3.2. P2MP LSP Tunnel**

A P2MP TE tunnel comprises of one or more P2MP LSPs referred to as P2MP LSP Tunnels. A P2MP LSP Tunnel is identified by the combination of the P2MP ID, Tunnel ID, and Extended Tunnel ID that are part of the P2MP SESSION object, and the IPv4 or IPv6 tunnel sender address and LSP ID fields of the P2MP SENDER\_TEMPLATE object. The new P2MP SENDER\_TEMPLATE object is defined in [section 24.2](#)

### **3.3. P2P Sub-LSPs**

A P2MP LSP Tunnel is constituted of one or more P2P sub-LSPs.

#### **3.3.1. Representation of a P2P Sub-LSP**

A P2P sub-LSP exists within the context of a P2MP LSP Tunnel. Thus it is identified by the P2MP ID, Tunnel ID, and Extended Tunnel ID that are part of the P2MP SESSION, the IPv4 or IPv6 tunnel sender address and LSP ID fields of the P2MP SENDER\_TEMPLATE object, and the P2P sub-LSP destination address that is part of the P2P\_SUB\_LSP object. The P2P\_SUB\_LSP object is defined in [section 24.3](#).

Additionally, a sub-LSP ID contained in the P2P\_SUB\_LSP object may be used depending on further discussions about the make-before-break procedures described in [section 19](#).

An EXPLICIT\_ROUTE Object (ERO) or SUB\_EXPLICIT\_ROUTE Object (SERO) is used to optionally specify the explicit route of a P2P sub-LSP. Each ERO or a SERO that is signaled corresponds to a particular





P2P\_SUB\_LSP object. Details of explicit route encoding are specified in [section 3.4](#).

### **3.3.2. P2P Sub-LSPs and Path Messages**

The mechanism in this document allows a P2MP LSP Tunnel to be signaled using one or more Path messages. Each Path message may signal one or more P2P sub-LSPs. Multiple Path messages are desirable as one Path message may not be large enough to fit all the P2P sub-LSPs; and they also allow separate manipulation of sub-trees of the P2MP LSP Tunnel. The reason for allowing a single Path message, to signal multiple P2P sub-LSPs, is to optimize the number of control messages needed to setup a P2MP LSP Tunnel.

### **3.4. Explicit Route Encoding**

When a Path message signals a single P2P sub-LSP (that is, the Path message is only targeting a single leaf in the P2MP tree), the EXPLICIT\_ROUTE object encodes the path from the ingress LSR to the egress LSR. The Path message also includes the P2P\_SUB\_LSP object for the P2P sub-LSP being signaled. The <[EXPLICIT\_ROUTE], <P2P\_SUB\_LSP>> tuple represents the P2P sub-LSP. The absence of the ERO should be interpreted as requiring hop-by-hop routing for the sub-LSP based on the P2P sub-LSP destination address field of the P2P\_SUB\_LSP object.

The absence of the ERO should be interpreted as requiring hop-by-hop routing for the sub-LSP based on the P2P sub-LSP destination address field of the P2P\_SUB\_LSP object.

When a Path message signals multiple P2P sub-LSPs the path of the first P2P sub-LSP, from the ingress LSR to the egress LSR, is encoded in the ERO. The first P2P sub-LSP is the one that corresponds to the first P2P\_SUB\_LSP object in the Path message. The P2P sub-LSPs corresponding to the P2P\_SUB\_LSP objects that follow are termed as subsequent P2P sub-LSPs. The path of each subsequent P2P sub-LSP is encoded in a SUB\_EXPLICIT\_ROUTE object (SERO). The format of the SERO is the same as an ERO (as defined in [\[RFC3209\]](#)). Each subsequent P2P sub-LSP is represented by tuples of the form [<SUB\_EXPLICIT\_ROUTE>] <P2P\_SUB\_LSP>. There is a one to one correspondence between a P2P\_SUB\_LSP object and a SERO. The absence of a SERO should be interpreted as requiring hop-by-hop routing for that sub-LSP. Note that the destination address is carried in the P2P sub-LSP object. The encoding of the SERO and P2P sub-LSP object are described in detail in [section 24](#).

The motivation behind the use of the SERO object is to provide explicit route compression when a Path message signals simultaneously



multiple P2P sub-LSPs. One approach to encode the explicit route of a subsequent P2P sub-LSP is to include the path from the ingress to the egress of the P2P sub-LSP. However this implies potential repetition of hops that can be learned from the ERO or explicit routes of other P2P sub-LSPs. Explicit route compression using SEROs attempts to minimize such repetition. A SERO for a particular P2P sub-LSP includes only the path from a certain branch LSR to the egress LSR if the path to that branch LSR can be derived from the ERO or other SEROs.

Explicit route compression is illustrated using the following figure.

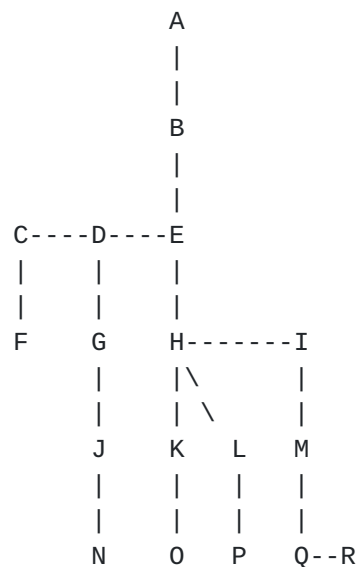


Figure 1. Explicit Route Compression

Figure 1. shows a P2MP LSP Tunnel with LSR A as the ingress LSR and six egress LSRs: (F, N, O, P, Q and R). When all the six P2P sub-LSPs are signaled in one Path message let us assume that the P2P sub-LSP to LSR F is the first P2P sub-LSP and the rest are subsequent P2P sub-LSPs. Following is one way for the ingress LSR A to encode the P2P sub-LSP explicit routes using compression:

P2P sub-LSP-F:	ERO = {B, E, D, C, F},	P2P_SUB_LSP Object-F
P2P sub-LSP-N:	SERO = {D, G, J, N},	P2P_SUB_LSP Object-N
P2P sub-LSP-O:	SERO = {E, H, K, O},	P2P_SUB_LSP Object-O
P2P sub-LSP-P:	SERO = {H, L, P},	P2P_SUB_LSP Object-P,
P2P sub-LSP-Q:	SERO = {H, I, M, Q},	P2P_SUB_LSP Object-Q,
P2P sub-LSP-R:	SERO = {Q, R},	P2P_SUB_LSP Object-R,

After LSR E processes the incoming Path message from LSR B it sends a



Path message to LSR D with the P2P sub-LSP explicit routes encoded as follows:

```
P2P sub-LSP-F:  ERO = {D, C, F}, P2P_SUB_LSP Object-F
P2P sub-LSP-N:  SERO = {D, G, J, N}, P2P_SUB_LSP Object-N
```

LSR E also sends a Path message to LSR H and following is one way to encode the P2P sub-LSP explicit routes using compression:

```
P2P sub-LSP-O:  ERO = {H, K, O}, P2P_SUB_LSP Object-O
P2P sub-LSP-P:  SERO = {H, L, P}, P2P_SUB_LSP Object-P,
P2P sub-LSP-Q:  SERO = {H, I, M, Q}, P2P_SUB_LSP Object-Q,
P2P sub-LSP-R:  SERO = {Q, R}, P2P_SUB_LSP Object-R,
```

After LSR H processes the incoming Path message from E it sends a Path message to LSR K, LSR L and LSR I. The encoding for the Path message to LSR K is as follows:

```
P2P sub-LSP-O:  ERO = {K, O}, P2P_SUB_LSP Object-O
```

The encoding of the Path message sent by LSR H to LSR L is as follows:

```
P2P sub-LSP-P:  ERO = {L, P}, P2P_SUB_LSP Object-P,
```

Following is one way for LSR H to encode the P2P sub-LSP explicit routes in the Path message sent to LSR I:

```
P2P sub-LSP-Q:  ERO = {I, M, Q}, P2P_SUB_LSP Object-Q,
P2P sub-LSP-R:  SERO = {Q, R}, P2P_SUB_LSP Object-R,
```

The explicit route encodings in the Path messages sent by LSRs D and Q are left as an exercise to the reader.

This compression mechanism reduces the Path message size. It also reduces extra processing that can result if explicit routes are encoded from ingress to egress for each P2P sub-LSP. No assumptions are placed on the ordering of the subsequent P2P sub-LSPs and hence on the ordering of the SEROs in the Path message. All LSRs need to process the ERO corresponding to the first P2P sub-LSP. A LSR needs to process a P2P sub-LSP descriptor for a subsequent P2P sub-LSP only if the first hop in the corresponding SERO is a local address of that LSR. The branch LSR that is the first hop of a SERO propagates the corresponding P2P sub-LSP downstream.



#### 4. Sub-Groups

As with all other RSVP controlled LSP Tunnels, P2MP LSP Tunnel state is managed using RSVP messages. While use of RSVP messages is the same, P2MP LSP Tunnel state differs from P2P LSP state in a number of ways. The two most notable differences are that a P2MP LSP Tunnel is targeted at multiple P2P Sub-LSPs and that, as a result of this, it may not be possible to represent full state in a single IP datagram and even more likely that it can't fit into a single IP packet. It must also be possible to efficiently add and remove endpoints to and from P2MP TE LSPs. An additional issue is that P2MP LSP Tunnels must also handle the state "remerge" problem, see [[P2MP-REQ](#)].

These differences in P2MP state are addressed through the addition of a sub-group identifier (Sub-Group ID) and sub-group originator (Sub-Group Originator ID) to the SENDER\_TEMPLATE and FILTER\_SPEC objects. Taken together the Sub-Group ID and Sub-Group Originator ID are referred to as the Sub-Group fields.

The Sub-Group fields, together with rest of the SENDER\_TEMPLATE and SESSION objects, are used to represent a portion of a P2MP LSP Tunnel's state. The portion of P2MP LSP Tunnel state identified by specific subgroup field values is referred to as a signaling sub-tree. It is important to note that the term "signaling sub-tree" refers only to signaling state and not data plane replication or branching. For example, it is possible for a node to "branch" signaling state for a P2MP LSP Tunnel, but to not branch the data associated with the P2MP LSP Tunnel. Typical applications for generation and use of multiple subgroups are adding an egress and semantic fragmentation to ensure that a Path message remains within a single IP packet.

#### 5. Path Message Format

This section describes modifications made to the Path message format as specified in [[RFC3209](#)] and [[RFC3473](#)]. The Path message is enhanced to signal one or more P2P sub-LSPs. This is done by including the P2P sub-LSP descriptor list in the Path message as shown below.

```
<Path Message> ::=      <Common Header> [ <INTEGRITY> ]
                        [ [ <MESSAGE_ID_ACK> | <MESSAGE_ID_NACK> ] ... ]
                        [ <MESSAGE_ID> ]
                        <SESSION> <RSVP_HOP>
                        <TIME_VALUES>
                        [ <EXPLICIT_ROUTE> ]
                        <LABEL_REQUEST>
                        [ <PROTECTION> ]
```





```
[ <LABEL_SET> ... ]
[ <SESSION_ATTRIBUTE> ]
[ <NOTIFY_REQUEST> ]
[ <ADMIN_STATUS> ]
[ <POLICY_DATA> ... ]
<sender descriptor>
[P2P sub-LSP descriptor list]
```

Following is the format of the P2P sub-LSP descriptor list.

```
<P2P sub-LSP descriptor list> ::= <P2P sub-LSP descriptor>
                                   [ <P2P sub-LSP descriptor list> ]
```

```
<P2P sub-LSP descriptor> ::= <P2P_SUB_LSP> [ <SUB_EXPLICIT_ROUTE> ]
```

Each LSR MUST use the common objects in the Path message and the P2P sub-LSP descriptors to process each P2P sub-LSP represented by the P2P sub-LSP object and the SUB-/EXPLICIT\_ROUTE object combination.

The first P2P\_SUB\_LSP object's explicit route is specified by the

ERO. Explicit routes of subsequent P2P sub-LSPs are specified by the corresponding SERO. A SERO corresponds to the following P2P\_SUB\_LSP object.

The RRO in the sender descriptor contains the hops traversed by the Path message and applies to all the P2P sub-LSPs signaled in the Path message.

Path message processing is described in the next section.

## 6. Path Message Processing

The ingress-LSR initiates the set up of a P2P sub-LSP to each egress-LSR that is the destination of the P2MP LSP Tunnel. Each P2P sub-LSP is associated with the same P2MP LSP Tunnel using common P2MP SESSION object and <Source Address, LSP-ID> fields in the SENDER\_TEMPLATE object. Hence it can be combined with other P2P sub-LSPs to form a P2MP LSP Tunnel. Another P2P sub-LSP belonging to the same instance of this P2P sub-LSP (i.e. the same P2MP LSP Tunnel) can share resources with this LSP. The session corresponding to the P2MP TE tunnel is determined based on the P2MP SESSION object. Each P2P sub-LSP is identified using the P2P\_SUB\_LSP object. Explicit routing for the P2P sub-LSPs is achieved using the ERO and SEROs.

As mentioned earlier, it is possible to signal P2P sub-LSPs for a given P2MP LSP Tunnel in one or more Path messages. And a given Path



message can contain one or more P2P sub-LSPs."

### **6.1. Multiple Path messages**

As described in [section 3](#), <EXPLICIT\_ROUTE> <P2P SUB-LSP> or <SUB\_EXPLICIT\_ROUTE> <P2P\_SUB\_LSP> tuple is used to specify a P2P sub-LSP. Multiple Path messages can be used to signal a P2MP LSP Tunnel. Each Path message can signal one or more P2P sub-LSPs. If a Path message contains only one P2P sub-LSP, each LSR along the P2P sub-LSP follows [\[RFC3209\]](#) procedures for processing the Path message besides the P2P SUB-LSP object processing described in this document.

Processing of Path messages containing more than one P2P sub-LSP is described in [Section 6.2](#).

An ingress LSR may use multiple Path messages for signaling a P2MP LSP. This may be because a single Path message may not be large enough to signal the P2MP LSP Tunnel. Or it may be while adding leaves to the P2MP LSP Tunnel the new leaves are signaled in a new Path message. Or an ingress LSR MAY choose to break the P2MP tree into separate manageable P2MP trees. These trees share the same root and may share the trunk and certain branches. The scope of this management decomposition of P2MP trees is bounded by a single tree and multiple trees with a single leaf each. Per [\[P2MP-REQ\]](#), a P2MP LSP Tunnel must have consistent attributes across all portions of a tree. This implies that each Path message that is used to signal a P2MP LSP Tunnel is signaled using the same signaling attributes with the exception of the P2P sub-LSP information.

The resulting sub-LSPs from the different Path messages belonging to the same P2MP LSP Tunnel SHOULD share labels and resources where they share hops to prevent multiple copies of the data being sent.

In certain cases a transit LSR may need to generate multiple Path messages to signal state corresponding to a single received Path message. For instance ERO expansion may result in an overflow of the resultant Path message. There are two cases occurring in such circumstances, either the message can be decomposed into multiple Path messages such that each of the message carries a subset of the incoming P2P sub-LSPs carried by the incoming message or the message can not be decomposed such that each of the outgoing Path message fits its maximum size value."



### **6.1.1. Identifying Multiple Path Messages**

Multiple Path messages generated by a LSR that signal state for the same P2MP LSP are signaled with the same SESSION object and have the same <Source address, LSP-ID> in the SENDER\_TEMPLATE object. In order to disambiguate these Path messages a <Sub-Group Originator ID, sub-Group ID> tuple is introduced (also referred to as the Sub-Group field). Multiple Path messages generated by a LSR to signal state for the same P2MP LSP have the same Sub-Group Originator ID and have a different sub-Group ID. The Sub-Group Originator ID SHOULD be set to the Router ID of the LSR that originates the Path message. This is either the ingress LSR or a LSR which re-originates the Path message with its own Sub-Group Originator ID. Cases when a transit LSR may change the Sub-Group Originator ID of an incoming Path message are described below. The <Sub-Group Originator ID, sub-Group ID> tuple is globally unique. The sub-Group ID space is specific to the Sub-Group Originator ID. Therefore the combination <Sub-Group Originator ID, sub-Group ID> is network-wide unique. Also, a router that changes the Sub-Group originator ID MUST use the same value of the Sub-Group Originator ID for a particular P2MP LSP Tunnel and should not vary it during the life of the P2MP LSP Tunnel.

Note: This version of the document assumes that these additional fields i.e., <Sub-Group Originator ID, sub-Group ID> are part of the SENDER\_TEMPLATE object."

### **6.2. Multiple P2P Sub-LSPs in one Path message**

The P2P sub-LSP descriptor list allows the signaling of one or more P2P sub-LSPs.

in one Path message. It is possible to signal multiple P2P sub-LSP object and ERO/SERO combinations in a single Path message. Note that these two objects are the ones that differentiate a P2P sub-LSP. Each LSR can use the common objects in the Path message and the P2P sub-LSP descriptors to process each P2P sub-LSP.

All LSRs need to process, when it is present, the ERO corresponding to the first P2P sub-LSP. If one or more SEROs are present an ERO must be present. The first P2P sub-LSP is propagated in a Path message by each LSR along the explicit route specified by the ERO. A LSR needs to process a P2P sub-LSP descriptor for a subsequent P2P sub-LSP only if the first hop in the corresponding SERO is a local address of that LSR. If this is not the case the P2P sub-LSP descriptor is included in the Path message sent to LSR that is the next hop to reach the first hop in the SERO. This next hop is determined by using the ERO or other SEROs that encode the path to the SERO's first hop. If this is the case and the LSR is also the



egress the P2P sub-LSP descriptor is not propagated downstream. If this is the case and the LSR is not the egress the P2P sub-LSP descriptor is included in a Path message sent to the next-hop determined from the SERO. Hence a branch LSR only propagates the relevant P2P sub-LSP descriptors on each downstream link. A P2P sub-LSP descriptor that is propagated on a downstream link only contains those P2P sub-LSPs that are routed using that link. This processing may result in a subsequent P2P sub-LSP in an incoming Path message to become the first P2P sub-LSP in an outgoing Path message.

Note that if one or more SEROs contains loose hops, expansion of such loose hops may result in overflowing the Path message size. [Section 9](#) describes how signaling of the set of P2P sub-LSPs can be split in more than one Path message.

The Record Route Object (RRO) contains the hops traversed by the Path message and applies to all the P2P sub-LSPs signaled in the path message. A transit LSR appends its address in an incoming RRO and propagates it downstream. A branch LSR forms a new RRO for each of the outgoing Path messages. Each such updated RRO is formed by appending the branch LSR's address to the incoming RRO.

If a LSR is unable to support a P2P sub-LSP setup, a PathErr message MUST be sent for the impacted P2P sub-LSP, and normal processing of the rest of the P2MP LSP Tunnel SHOULD continue. The default behavior is that the remainder of the LSP is not impacted (that is, all other branches are allowed to set up) and the failed branches are reported in PathErr messages in which the Path\_State\_Reomved flag MUST NOT be set. However, the ingress LSR may set a LSP Integrity flag (see [section 25](#)) to request that if there is a setup failure on any branch the entire LSP should fail to set up.

## **[7. Resv Message Format](#)**

The Resv message follows the [[RFC3209](#)] and [[RFC3473](#)] format:

```
<Resv Message> ::=      <Common Header> [ <INTEGRITY> ]
                        [ [ <MESSAGE_ID_ACK> | <MESSAGE_ID_NACK> ] ... ]
                        [ <MESSAGE_ID> ]
                        <SESSION> <RSVP_HOP>
                        <TIME_VALUES>
                        [ <RESV_CONFIRM> ] [ <SCOPE> ]
                        [ <NOTIFY_REQUEST> ]
                        [ <ADMIN_STATUS> ]
                        [ <POLICY_DATA> ... ]
                        <STYLE> <flow descriptor list>
```





```
<flow descriptor list> ::= <FF flow descriptor list>
                           | <SE flow descriptor>
```

```
<FF flow descriptor list> ::= <FF flow descriptor>
                              | <FF flow descriptor list>
                              <FF flow descriptor>
```

```
<SE flow descriptor> ::= <FLOWSPEC> <SE filter spec list>
```

```
<SE filter spec list> ::= <SE filter spec>
                          | <SE filter spec list> <SE filter spec>
```

The FF flow descriptor and SE filter spec are modified as follows to identify the P2P sub-LSPs that they correspond to:

```
<FF flow descriptor> ::= [ <FLOWSPEC> ] <FILTER_SPEC> <LABEL>
                          [ <RECORD_ROUTE> ] [ <P2P sub-LSP descriptor
list> ]
```

```
<SE filter spec> ::=      <FILTER_SPEC> <LABEL> [ <RECORD_ROUTE> ]
                          [ <P2P sub-LSP descriptor list> ]
```

FILTER\_SPEC is defined in [section 24.4](#).

The P2P sub-LSP descriptor has the same format as in [section 5.1](#) with the difference that a SUB\_RECORD\_ROUTE object is used in place of a SUB\_EXPLICIT\_ROUTE object. The SUB\_RECORD\_ROUTE objects follow the same compression mechanism as the SUB\_EXPLICIT\_ROUTE objects. Note that that a Resv message can signal multiple P2P sub-LSPs that may belong to the same FILTER\_SPEC object or different FILTER\_SPEC objects. The same label is allocated if the FILTER\_SPEC object is the same.

However different upstream labels are allocated if the <Source Address, LSP-ID> of the FILTER\_SPEC object is different as that implies different P2MP LSP Tunnels.



## **8. Resv Message Processing**

The egress LSR follows normal RSVP procedures while originating a Resv message. The Resv message carries the label allocated by the egress LSR.

A subsequent node allocates its own label and passes it in the Resv message upstream. The node may combine multiple flow descriptors, from different Resv messages received from downstream, in one Resv message sent upstream. A Resv message is not sent upstream until at least one Resv message has been received from a downstream neighbor except when the integrity bit is set in the LSP\_ATTRIBUTE object.

Each FF flow descriptor or SE filter spec sent upstream in a Resv message includes a P2P sub-LSP descriptor list. Each such FF flow descriptor or SE filter spec for the same P2MP LSP Tunnel (whether on one or multiple Resv messages) is allocated the same label.

This label is associated by that node with all the labels received from downstream Resv messages for that P2MP LSP Tunnel. Note that a transit node may become a replication point in the future when a branch is attached to it. Hence this results in the setup of a P2MP LSP Tunnel from the ingress-LSR to the egress LSRs.

The ingress LSR may need to understand when all desired egresses have been reached. This is achieved using <P2P\_SUB\_LSP> objects.

Each branch node can potentially send one Resv message upstream for each of the downstream receivers. This may result in overflowing the Resv message, particularly when considering that the number of messages increases the closer the branch node is to the ingress.

Transit nodes MUST replace the Sub-Group ID fields received in the FILTER\_SPEC objects with the value that was received in the Sub-Group ID field of the Path message from the upstream neighbor, when the node set the Sub-Group Originator field in the associated Path message. ResvErr messages generation is unmodified. Nodes propagating a received ResvErr message MUST use the Sub-Group field values carried in the corresponding Resv message.

The solution for this issue is for further discussion.

### **8.1. RRO Processing**

A Resv message contains a record route per P2P sub-LSP that is being signaled by the Resv message if the sender node requests route recording by including a RRO in the Path message. The same rule is used during signaling of P2MP LSP Tunnels i.e. insertion of the RRO



in the Path message used to signal one or more P2P sub-LSP triggers the inclusion of an RRO for each sub-LSP.

The record route of the first P2P sub-LSP is encoded in the RRO. Additional RROs for the subsequent P2P sub-LSPs are referred to as SUB\_RECORD\_ROUTE objects (SRRROs). Their format is specified in [section 24.5](#). The ingress node then receives the RRO and possibly the SRRRO corresponding to each subsequent P2P sub-LSP. Each P2P\_SUB\_LSP object is followed by the RRO/SRRRO. The ingress node can then determine the record route corresponding to a particular P2P sub-LSP. The RRO and SRRROs can be used to construct the end to end Path for each P2P sub-LSP.

## **8.2. Resv Message Throttling**

A branch node may have to send the Resv message being sent upstream whenever there is a change in a Resv message for a P2P sub-LSP received from downstream. This can result in excessive Resv messages sent upstream, particularly when the P2P sub-LSPs are established for the first time. In order to mitigate this situation, branch nodes can limit their transmission of Resv messages. Specifically, in the case where the only change being sent in a Resv message is in one or more SRRRO objects, the branch node SHOULD transmit the Resv message only after a delay time has passed since the transmission of the previous Resv message for the same session. This delayed Resv message SHOULD include SRRROs for all branches. Specific mechanisms for Resv message throttling are implementation dependent and are outside the scope of this document.

## **9. Transit Fragmentation**

In certain cases a transit LSR may need to generate multiple Path messages to signal state corresponding to a single received Path message. For instance ERO expansion may result in an overflow of the resultant Path message. It is desirable not to rely on IP fragmentation in this case. In order to achieve this, the multiple Path messages generated by the transit LSR, are signaled with the Sub-Group Originator ID set to the TE Router ID of the transit LSR and a distinct

Sub-Group ID. Thus each distinct Path message that is generated by the transit LSR for the P2MP LSP Tunnel carries a distinct <Sub-Group Originator ID, Sub-Group ID> tuple.

When multiple Path messages are used by an ingress or transit node, each Path message SHOULD be identical with the exception of the P2P sub-LSP related information (e.g., SERO), message and hop information (e.g., INTEGRITY, MESSAGE\_ID and RSVP\_HOP), and the SENDER\_TEMPLATE



objects. Except when performing a make-before-break operation, the tunnel sender address and LSP ID fields MUST be the same in each message, and for transit nodes, the same as the values in the Path message.

As described above one case in which the Sub-Group Originator ID of a received Path message is changed is that of transit fragmentation. The Sub-Group Originator ID of a received Path message may also be changed in the outgoing Path message and set to that of the LSR originating the Path message based on a local policy. For instance a LSR may decide to always change the Sub-Group Originator ID while performing ERO expansion. The Sub-Group ID MUST not be changed if the Sub-Group Originator ID is not being changed.

## **10. Grafting**

The operation of adding egress LSR(s) to an existing P2MP LSP Tunnel is termed as grafting. This operation allows egress nodes to join a P2MP LSP Tunnel at different points in time.

There are two methods to add P2P sub-LSPs to a P2MP LSP Tunnel. The first is to add new P2P sub-LSPs to the P2MP LSP Tunnel by adding them to an existing Path message and refreshing the entire Path message. Path message processing described in [section 6](#) results in adding these P2P sub-LSPs to the P2MP LSP Tunnel. Note that as a result of adding one or more P2P sub-LSPs to a Path message the ERO compression encoding may have to be recomputed.

The second is to use incremental updates described in [section 13.1](#). The egress LSRs can be added/removed by signaling only the impacted P2P sub-LSPs in a new Path message. Hence other P2P sub-LSPs do not have to be re-signaled.

## **11. Pruning**

The operation of removing egress LSR(s) from an existing P2MP LSP Tunnel is termed as pruning. This operation allows egress nodes to leave a P2MP LSP Tunnel at different points in time.

The P2P sub-LSP(s) being removed from the P2MP LSP Tunnel are signaled in a PathTear message. The PathTear message includes the P2P sub-LSP descriptor list which is included before the sender descriptor. Note that the PathTear message contains only the P2P sub-LSP(s) being removed and rest of the P2MP LSP Tunnel does not have to be re-signaled. This results in removal of the state corresponding to these P2P sub-LSPs. State for rest of the P2P sub-LSPs is not





modified.

This section describes various mechanisms to perform pruning. Further discussion and feedback is needed to finesse these mechanisms. In the first mechanism in order to delete one or more P2P Sub-LSPs, a PathTear message is sent with the list of P2P sub-LSPs being deleted. This is a form of explicit tear down. A single PathTear message can only contain P2P sub-LSPs that were signaled by the ingress using the same <Sub-Group Originator ID, Sub-Group ID> tuple. The PathTear message is signaled with the SESSION and SENDER\_TEMPLATE objects corresponding to the P2MP LSP Tunnel and the <Sub-Group Originator ID, Sub-Group ID> tuple corresponding to the P2P sub-LSPs that are being deleted. A transit LSR that propagates the PathTear message downstream MUST ensure that it sets the <Sub-Group Originator ID, Sub-Group ID> tuple in the PathTear message to the values used to generate the last Path message that corresponds to the P2P sub-LSPs signaled in the PathTear message that it generates. The transit LSR may need to generate multiple PathTear messages for an incoming PathTear message if it had performed transit fragmentation for the corresponding incoming Path message.

The Path messages from which the P2P sub-LSPs were deleted need to be refreshed with the remaining P2P sub-LSPs. Note that as a result of deleting one or more P2P sub-LSPs from a Path message the ERO compression encoding may have to be recomputed.

When the last P2P sub-LSP is to be removed from a Path state, i.e., there are no remaining P2P sub-LSPs to send in a Path message, a PathTear message SHOULD be sent carrying the Sub-Group ID of the Path message that no longer has any P2P sub-LSPs.

The second mechanism to delete P2P sub-LSPs is implicit teardown which uses standard RSVP message processing. Per standard RSVP processing, a P2P sub-LSP may be removed from a P2MP TE LSP by sending a modified message for the Path or Resv message that previously advertised the P2P sub-LSP. This message MUST list all P2P sub-LSPs that are not being removed. When using this approach, a node processing a message that removes a P2P sub-LSP from a P2MP TE LSP MUST ensure that the P2P sub-LSP is not included in any other Path state associated with session before interrupting the data path to that egress. All other message processing remains unchanged.

The third mechanism is an explicit teardown mechanism that defines new syntax and semantics for a PathTear message. This new mechanism minimizes signaling required to remove a subset of P2P sub-LSPs set signaled in a Path message, and thereby reduces associated processing. When using this mechanism each identified P2P sub-LSP is removed from the P2MP LSP Tunnel state, even if the P2P sub-LSP is



advertised in multiple Path message.

When using this approach, a PathTear message is generated. The PathTear message MUST identify each P2P sub-LSP to be removed, via a P2P\_SUB\_LSP object per P2P Sub-LSP, and include a SENDER\_TEMPLATE object corresponding to the Path state being modified. The Sub-Group ID valued contained in the SENDER\_TEMPLATE object message MUST be set to zero (0). Subsequent Path messages associated with the P2MP LSP Tunnel MUST NOT contain the removed P2P sub-LSPs, unless that P2P sub-LSP is being re-added to the P2MP LSP.

To support the third mechanism, the receiver of PathTear message that is associated with a P2MP LSP Tunnel MUST check the value of a received Sub-Group ID fields. When there is no SENDER\_TEMPLATE object present or the value of the Sub-Group ID fields is non-zero, then PathTear processing as defined in the above explicit tear down mechanism must be followed. When the Sub-Group ID field is zero (0), then the processing node MUST remove the identified egresses from all control plane state associated with the P2MP LSP Tunnel and adjust the data path appropriately.

### **11.1. P2MP TE LSP Tear Down**

This operation is accomplished by listing all the P2P sub-LSPs in a PathTear message.

A PathTear message must be generated for each Path message used to signal the P2MP LSP Tunnel.

### **11.2. PathTear message Format**

The format of the PathTear message is as follows:

```
<PathTear Message> ::= <Common Header> [ <INTEGRITY> ]
                        [ [ <MESSAGE_ID_ACK> | <MESSAGE_ID_NACK> ...
                          ]
                          [ <MESSAGE_ID> ]
                          <SESSION> <RSVP_HOP>
                          [ <sender descriptor> ]
                          [ <P2P sub-LSP descriptor list> ]
```

er descriptor> ::= (see earlier definition)

Note: it is assumed that the P2P sub-LSP descriptor will not include the SUB\_EXPLICIT\_ROUTE object associated with each P2P\_SUB\_LSP being deleted



## **12. Refresh Reduction**

The refresh reduction procedures described in [[RFC2961](#)] are equally applicable to P2MP LSP Tunnels described in this document. Refresh reduction applies to individual messages and the state they install/maintain, and that continues to be the case for P2MP LSP Tunnels.

## **13. State Management**

State signaled by a P2MP Path message is managed by a local implementation using the <P2MP ID, Tunnel ID, Extended Tunnel ID> as part of the SESSION object and <Tunnel Sender Address, LSP ID, Sub-Group Originator ID, Sub-Group ID> as part of the SENDER\_TEMPLATE object.

Additional information signaled in the Path message is part of the state created by a local implementation. This mandatorily includes PHOP and SENDER\_TSPEC object.

### **13.1. Incremental State Update**

RSVP as defined in [[RFC2205](#)] and as extended by RSVP-TE [[RFC3209](#)] and GMPLS [[RFC3473](#)] uses the same basic approach to state communication and synchronization, namely full state is sent in each state advertisement message. Per [[RFC2205](#)] Path and Resv messages are idempotent. Also, [[RFC2961](#)] categorizes RSVP messages into two types: trigger and refresh messages and improves RSVP message handling and scaling of state refreshes but does not modify the full state advertisement nature of Path and Resv messages. The full state advertisement nature of Path and Resv messages has many benefits, but also has some drawbacks. One notable drawback is when an incremental modification is being made to a previously advertised state. In this case, there is the message overhead of sending the full state and the cost of processing it. It is desirable to overcome this drawback and add/delete P2P sub-LSPs to a P2MP LSP Tunnel by incrementally updating the existing state.

It is possible to use the procedures described in this document to allow P2P sub-LSPs to be incrementally added or deleted from the P2MP LSP by allowing a Path or a PathTear message to incrementally change the existing P2MP LSP Tunnel Path state.

As described in [section 6.1](#), multiple Path messages can be used to signal a P2MP LSP Tunnel. The Path messages are distinguished by different <Sub-Group Originator ID, sub-Group ID> tuples in the



SENDER\_TEMPLATE object. In order to perform incremental P2P sub-LSP state addition a separate Path message with a new sub-Group ID is used to add the new P2P sub-LSPs, by the ingress LSR. The Sub-Group Originator ID MUST be set to the TE Router ID [[RFC3477](#)] of the node that sets the Sub-Group ID.

This maintains the idempotent nature of RSVP Path messages; avoids keeping track of individual P2P sub-LSP state expiration and provides the ability to perform incremental P2MP LSP Tunnel state updates.

### **13.2. Combining Multiple Path Messages**

There is a tradeoff between the number of Path messages used by the ingress to maintain the P2MP LSP Tunnel and using full state refresh to add P2P sub-LSPs. It is possible to combine P2P sub-LSPs previously advertised in different Path messages in a single Path message in order to reduce the number of Path messages needed to maintain the P2MP LSP. This can also be done by a transit node that performed fragmentation and at a later point is able to combine multiple Path messages that it generated into a single Path message. This may happen when one or more P2P sub-LSPs are pruned from the existing Path states.

The new Path message is signaled by the node that is combining multiple Path messages with all the P2P sub-LSPs that are being combined in a single Path message. This Path message contains a new Sub-Group ID field value. When a new Path and Resv message that is signaled for an existing P2P sub-LSP is received by a transit LSR, state including the new instance of the P2P sub-LSP is created.

The P2P sub-LSP SHOULD continue to be advertised in both the old and new Path messages until a Resv message listing the P2P sub-LSP and corresponding to the new Path message is received by the combining node. Hence until this point state for the P2P sub-LSP SHOULD be maintained as part of the Path state for both the old and the new Path message [[Section 3.1.3](#), 2205]. At that point the P2P sub-LSP SHOULD be deleted from the old Path state using a PathTear message. The P2P sub-LSP should also be removed from the old Path message and the old Path message should be signaled again, if there are other remaining P2P sub-LSPs in the old Path message.

A Path message with a sub-Group\_ID(n+1) may signal a set of P2P sub-LSPs that belong partially or entirely to an already existing Sub-Group\_ID(i),  $i \leq n$ , the SESSION object and <Sender Tunnel Address, LSP-ID, Sub-Group Originator ID> being the same. Or it may signal a strictly non-overlapping new set of P2P sub-LSPs with a strictly higher sub-Group\_ID value.





- 1) If  $\text{sub-Group\_ID}(i) = \text{sub-Group\_ID}(n+1)$ ,  $i \leq n$  then either a full refresh is indicated by the Path message or a P2P Sub-LSP is added to/deleted from the group signaled by  $\text{sub-Group\_ID}(n+1)$
- 2) If  $\text{sub-Group\_ID}(i) \neq \text{sub-Group\_ID}(n+1)$ ,  $i \leq n$  then the Path message is signaling a set of P2P sub-LSPs that belong partially or entirely to an already existing  $\text{sub-Group\_ID}(i)$  or a strictly non-overlapping set of P2P sub-LSPs.

## **14. Error Processing**

PathErr and ResvErr messages are processed as per RSVP-TE procedures. Note that a LSR on receiving a PathErr/ResvErr message for a particular P2P sub-LSP changes the state only for that P2P sub-LSP. Hence other P2P sub-LSPs are not impacted. In case the ingress node requests the maintenance of the 'LSP integrity', any error reported within the P2MP TE LSP must be reported at (least at) any other branching nodes belonging to this LSP. Therefore, reception of an error message for a particular P2P sub-LSP MAY change the state of any other P2P sub-LSP of the same P2MP TE LSP.

### **14.1. Branch Failure Handling**

During setup and during normal operation, PathErr messages may be received at a branch node. In all cases, a received PathErr message is first processed per standard processing rules. That is: the PathErr message is sent hop-by-hop to the ingress/branch LSR for that Path message. Intermediate nodes until this ingress/branch LSR MAY inspect this message but take no action upon it. The behavior of a branch LSR that generates a PathErr message is under the control of the ingress LSR.

The default behavior is that the PathErr does not have the Path\_State\_Removed flag set. However, if the ingress LSR has set the 'LSP integrity' flag on the Path message (see LSP\_ATTRIBUTE object in [section 24](#)) and if the Path\_State\_Removed flag is supported, the LSR generating a PathErr to report the failure of a branch of the P2MP LSP Tunnel SHOULD set the Path\_State\_Removed flag.

A branch LSR that receives a PathErr message with the Path\_State\_Removed flag set MUST act according to the wishes of the ingress LSR. The default behavior is that the branch LSR clears the Path\_State\_Removed flag on the PathErr and sends it further upstream. It does not tear any other branches of the LSP. However, if the LSP

integrity flag is set on the Path message, the branch LSR MUST send PathTear on all downstream branches and send the PathErr message upstream with the Path\_State\_Removed flag set.



A branch LSR that receives a PathErr message with the Path\_State\_Removed flag clear MUST act according to the wishes of the ingress LSR. The default behavior is that the branch LSR forwards the PathErr upstream and takes no further action. However, if the LSP integrity flag is set on the Path message, the branch LSR MUST send PathTear on all downstream branches and send the PathErr upstream with the Path\_State\_Removed flag set (per [\[RFC3473\]](#)).

In all cases, the PathErr message forwarded by a branch LSR MUST contain the P2P sub-LSP identification and explicit routes of all branches that are errored (reported by received PathErr messages) and all branches that are explicitly torn by the branch LSR.

## **[15. Notify and ResvConf Messages](#)**

Notify messages, see [\[RFC3473\]](#), may contain either SENDER\_TEMPLATE or FILTER\_SPEC objects, but are sent in a targeted fashion. This means that the Sub-Group fields cannot be updated in transit and is unlikely to provide any value to the Notify message recipient. Therefore, the receiver of a Notify message MUST identify the sender state referenced in the message based on the Source address and LSP ID contained in the received SENDER\_TEMPLATE or FILTER\_SPEC objects rather than, as is normally done, based on the whole objects.

ResvConf messages may contain FILTER\_SPEC objects and may also be sent in a targeted fashion. As with Notify messages, the receiver of a ResvConf message MUST identify the state referenced in the message based on the address and LSP ID contained in the received FILTER\_SPEC object rather than, as is normally done, based on the whole objects.

## **[16. Control of Branch Fate Sharing](#)**

An ingress LSR can control the behavior of an LSP if there is a failure during LSP setup or after an LSP has been established. The default behavior is that only the branches downstream of the failure are not established, but the ingress may request 'LSP integrity' such that any failure anywhere within the LSP tree causes the entire P2MP LSP Tunnel to fail.

The ingress LSP may request 'LSP integrity' by setting bit [TBD] of the Attributes Flags TLV. The bit is set if LSP integrity is required.

It is RECOMMENDED to use the LSP\_ATTRIBUTES Object for this flag and not the LSP\_REQUIRED\_ATTRIBUTES Object.



A branch LSR that supports the Attributes Flags TLV and recognizes this bit MUST support LSP integrity or reject the LSP setup with a PathErr carrying the error "Routing Error"/"Unsupported LSP Integrity"

## **17. Admin Status Change**

A branch node that receives an ADMIN\_STATUS object processes it normally and also relays the ADMIN\_STATUS object in a Path on every branch. All Path messages may be concurrently sent to the downstream neighbors.

Downstream nodes process the change in the status object per [RFC3473], including generation of Resv messages. When the last received upstream ADMIN\_STATUS object had the R bit set, branch nodes wait for a Resv message with a matching ADMIN\_STATUS object to be received (or a corresponding PathErr or ResvTear message) on all branches before relaying a corresponding Resv message upstream.

## **18. Label Allocation on LANs with Multiple Downstream Nodes**

A sender on a LAN uses a different label for sending traffic to each node on the LAN that belongs to the P2MP LSP Tunnel. Thus the sender performs replication. It may be considered desirable on a LAN to use the same label for sending traffic to multiple nodes belonging to the same P2MP LSP Tunnel, to avoid replication. Procedures for doing this are for further study. Given the relatively small number of receivers on LANs typically deployed in MPLS networks, this is not currently seen as a practical problem. Furthermore avoiding replication at the sender on a LAN requires significant complexity in the control plane. Given the tradeoff we propose the use of replication by the sender on a LAN.

## **19. Make-before-break**

Let's consider the following cases where make-before-break is needed:

### **19.1. P2MP Tree Re-optimization**

In this case all the P2P sub-LSPs are signaled with a different LSP ID by the ingress-LSR and follow make-before-break procedure[RFC3209] Thus a new P2MP LSP Tunnel instance is established. Each P2P sub-LSP is signaled with a different LSP ID, corresponding to the new P2MP TE LSP. The ingress can, after moving traffic to the new instance, tear down the previous P2MP LSP Tunnel instance.



### **19.2. Re-optimization of a subset of P2P sub-LSPs.**

One way to accomplish re-optimization of a subset of P2P sub-LSPs that belong to a P2MP LSP Tunnel is to resignal the entire tree with a new LSP-ID as described in the previous subsection.

(There is NO-CONSENSUS between the authors on rest of the text in this subsection and it needs further discussion.)

It is possible to accomplish re-optimization of one or more P2P sub-LSPs without re-signaling rest of the P2MP LSP. To achieve this a sub-LSP ID is used to identify each P2P sub-LSP. This is encoded in the P2P sub-LSP object. Each re-optimized P2P sub-LSP is signaled with a different sub-LSP ID and hence a new P2P sub-LSP is established. Once the new setup is complete, the old P2P sub-LSP can be torn down. In some cases this may result in transient data duplication.

## **20. Fast Reroute**

[RSVP-FR] extensions can be used to perform fast reroute for the mechanism described in this document.

### **20.1. Facility Backup**

Facility backup as described in [[RSVP-FR](#)] can be used to protect P2MP LSP Tunnels.

If link protection is desired, a bypass tunnel is used to protect the link between the PLR and next-hop. Thus all P2P sub-LSPs that use the link can be protected in the event of link failure. Note that all such P2P sub-LSPs belonging to a particular instance of a P2MP tunnel will share the same outgoing label on the link between the PLR and the next-hop. This is the P2MP LSP label on the link. Label stacking is used to send data for each P2MP LSP in the bypass tunnel. The inner label is the P2MP LSP Tunnel label allocated by the nhop. During failure Path messages for each P2P sub-LSP, that is effected, will be sent to the MP, by the PLR. It is recommended that the PLR use the sender template specific method to identify these Path messages. Hence the PLR will set the source address in the sender template to a local PLR address. The MP will use the LSP-ID to identify the corresponding P2P sub-LSPs.

The MP MUST not use the <sub-group originator ID, sub-group ID> while identifying the corresponding P2P sub-LSPs.

In order to further process a P2P sub-LSP it will determine the





protected P2P sub-LSP using the LSP-id and the P2P sub-LSP object.

If node protection is desired, the bypass tunnel must intersect the path of the protected P2P sub-LSPs somewhere downstream of the PLR. This constrains the set of P2P sub-LSPs being backed-up via that bypass tunnel to those that pass through a common downstream MP. The MP will alloca

te the same label to all such P2P sub-LSPs belonging to a particular instance of a P2MP tunnel. This will be the inner label used during label stacking. This may require the PLR to be branch capable as multiple bypass tunnels may be required to backup the set of P2P sub-LSPs passing through the protected node. Else all the P2P sub-LSPs being backed up must pass through the same MP.

## **20.2. One to One Backup**

One to one backup as described in [[RSVP-FR](#)] can be used to protect a particular P2P sub-LSP against link and next-hop failure. Protection may be used for one or more P2P sub-LSPs between the PLR and the next-hop. All the P2P sub-LSPs corresponding to the same instance of the P2MP tunnel, between the PLR and the next-hop share the same P2MP LSP Tunnel label.

All or some of these P2P sub-LSPs may be protected.

The detour P2P sub-LSPs may or may not share labels, depending on the detour path. Thus the set of outgoing labels and next-hops for a P2MP LSP Tunnel that was using a single next-hop and label between the PLR and next-hop before protection, may change once protection is triggered.

Its is recommended that the path specific method be used to identify a backup P2P sub-LSP. Hence the DETOUR object will be inserted in the backup Path message. A backup P2P sub-LSP MUST be treated as belonging to a different P2MP tunnel instance than the one specified by the LSP-id. Furthermore multiple backup P2P sub-LSPs MUST be treated as part of the same P2MP tunnel instance if they have the same LSP-id and the same DETOUR objects. Note that as specified in [section 3](#) P2P sub-LSPs between different P2MP tunnel instances use different labels.

If there is only P2P sub-LSP in the Path message, the DETOUR object applies to that sub-LSP. If there are multiple P2P sub-LSPs in the Path message the DETOUR applies to all the P2P sub-LSPs.



## **21. Support for LSRs that are not P2MP Capable**

It may be that some LSRs in a network are capable of processing the P2MP extensions described in this document, but do not support P2MP branching in the data plane. If such an LSR is requested to become a branch LSR by a received Path message, it MUST respond with a PathErr message carrying the Error Value "Routing Error" and Error Code "Unable to Branch".

Its also conceivable that some LSRs, in a network deploying P2MP capability, may not support the extensions described in this document. If a Path message for the establishment of a P2MP LSP Tunnel reaches such an LSR it will reject it with a PathErr because it will not recognize the C-Type of the P2MP SESSION object.

LSRs that do not support the P2MP extensions in this document may be included as transit LSRs by the use of LSP-stitching and LSP-hierarchy [[LSP-HIER](#)]. Note that LSRs that are required to play any other role in the network (ingress, branch or egress) MUST support the extensions defined in this document.

The use of LSP-stitching and LSP-hierarchy [[LSP-HIER](#)] allows to build P2MP LSP Tunnels in such an environment. A P2P LSP segment is signaled from the previous P2MP capable hop of a legacy LSR to the next P2MP capable hop. Of course this assumes that intermediate legacy LSRs are transit LSRs and cannot act as P2MP branch points. Transit LSRs along this LSP segment do not process control plane messages associated with a P2MP LSP Tunnel. Furthermore these LSRs also do not need to have P2MP data plane capability as they only need to process data belonging to the P2P LSP segment. Hence these LSRs do not need to support P2MP MPLS. This P2P LSP segment is stitched to the incoming P2MP LSP Tunnel. After the P2P LSP segment is established the P2MP Path message is sent to the next P2MP capable LSR as a directed Path message. The next P2MP capable LSR stitches the P2P LSP segment to the outgoing P2MP LSP Tunnel.

In packet networks, the P2P sub-LSPs may be nested inside the outer P2P LSP Tunnel. Hence label stacking can be used to enable use of the same LSP Tunnel segment for multiple P2MP LSP Tunnels. Stitching and nesting considerations and procedures are described further in [[INT-REG](#)].

It may be an overhead for an operator to configure the P2P LSP segments in advance, when it is desired to support legacy LSRs. It may be desirable to do this dynamically. The ingress can use IGP extensions to determine non P2MP capable LSRs. It can use this information to compute P2P sub-LSP paths such that they avoid these legacy LSRs. The explicit route object of a P2P sub-LSP path may



contain loose hops if there are legacy LSRs along the path. The corresponding explicit route contains a list of objects upto the P2MP capable LSR that is adjacent to a legacy LSR followed by a loose object with the address of the next P2MP capable LSR. The P2MP capable LSR expands the loose hop using its TED. When doing this it determines that the loose hop expansion requires a P2P LSP to tunnel through the legacy LSR. If such a P2P LSP exists, it uses that P2P LSP. Else it establishes the P2P LSP. The P2MP Path message is sent to the next P2MP capable LSR using non-adjacent signaling. The P2MP capable LSR that initiates the non-adjacent signaling message to the next P2MP capable LSR may have to employ a fast detection mechanism such as [\[BFD\]](#) to the next P2MP capable LSR.

This may be needed for the directed Path message Head-End to use node protection FRR when the protected node is the directed Path message tail.

Note that legacy LSRs along a P2P LSP segment cannot perform node protection of the tail of the P2P LSP segment.

## **[22. Reduction in Control Plane Processing with LSP Hierarchy](#)**

It is possible to take advantage of LSP hierarchy [\[LSP-HIER\]](#) while setting up P2MP LSP Tunnels, as described in the previous section, to reduce control plane processing along transit LSRs that are P2MP capable. This is applicable only in environments where LSP hierarchy can be used. Transit LSRs along a P2P LSP segment, being used by a P2MP LSP Tunnel, do not process control plane messages associated with the P2MP LSP Tunnel. Infact they are not aware of these messages as they are tunneled over the P2P LSP segment. This reduces the amount of control plane processing required on these transit LSRs.

Note that the P2P LSP segments can be dynamically setup as described in the previous section or preconfigured. For example in Figure 2, PE1 can setup a P2P LSP to P1 and use that as a LSP segment. The Path messages for PE3 and PE4 can now be tunneled over the LSP segment. Thus P3 is not aware of the P2MP LSP Tunnel and does not process the P2MP control messages.



### **23. P2MP LSP Tunnel Remerging and Cross-Over**

The functional description described so far assumes that multiple Path messages received by a LSR for the same P2MP LSP Tunnel arrive on the same incoming interface. However this may not always be the case. Further discussion is needed for this section.

P2MP tree remerging or cross-over occurs when a transit or egress node receives the signaling state i.e. Path message for the same P2MP TE LSP from more than one previous hop. If the re-merged P2P sub-LSPs are sent out on different interfaces there is no data plane issue. However if the re-merged P2P sub-LSPs are sent out on the same interface it can result in data duplication downstream. In order to describe identification of cross over and remerging by a LSR let us list the various cases when state for a P2P sub-LSP is received by a LSR.

Case1: P2P sub-LSP already exist as part of an existing Path state. The following are the various sub-cases.

- a) The new P2P sub-LSP uses the same PHOP and outgoing interface as the existing P2P sub-LSP. This is either a refresh or can occur when multiple existing Path messages are combined in a new Path message.
- b) The new P2P sub-LSP uses the same PHOP but different outgoing interface as the existing P2P sub-LSP. This is a case of re-routing.
- c) The new P2P sub-LSP uses a different PHOP and same outgoing interface as the existing P2P sub-LSP. This is a case of re-merging.
- d) The new P2P sub-LSP uses a different PHOP and a different outgoing interface as compared to the existing P2P sub-LSP. This is a case of cross-over.

Case2: P2P sub-LSP does not exist as part of an existing Path state. The following are the sub-cases.

- a) The new P2P sub-LSP uses a PHOP and outgoing interface that is same as the PHOP and outgoing interface used by an existing P2P sub-LSP. This is a legal case of signaling a new P2P sub-LSP.
- b) The new P2P sub-LSP uses a PHOP that is same as that used by an existing P2P sub-LSP. However the outgoing interface is different from the outgoing interfaces used by existing P2P sub-LSPs. This is a legal case of signaling a new P2P sub-LSP.
- c) The new P2P sub-LSP uses a different PHOP than that used by any of the existing P2P sub-LSP. However the outgoing interface is same as the outgoing interface used by an existing P2P sub-LSPs. This is a case of remerging.
- d) The new P2P sub-LSP uses a different PHOP than that used by any of the existing P2P sub-LSP. Also the outgoing interface is different from the outgoing interfaces used by existing P2P sub-LSPs. This is a



case of cross-over.

[draft-ietf-mpls-rsvp-te-p2mp-00.txt](#)

[Page 29]

Cases 1(d) and 2(d) above identify cross-over and this is considered legal. Cases 1(c) and 2(c) above identify remerging in the data plane. If the LSR is capable of remerging in the data plane this is considered legal.

The below procedure applies for remerging.

The remerge error case is detected by checking incoming Path messages that represent new P2MP TE LSP state and seeing if they represent both known LSP state and a different P2P sub-LSP list. Specifically, the remerge check MUST be performed when processing Path messages that contain SESSION, SENDER\_TEMPLATE and RSVP\_HOP objects that have not previously been seen on a particular interface. The remerge check consists of attempting to locate state that has the same values in the SESSION object and in the tunnel sender address and LSP ID fields of the SENDER\_TEMPLATE object.

If no matching state is located, then there is no remerge condition.

If matching state is found, then the list of P2P Sub-LSPs associated with the new Path message is compared against the list present in the located state. If any addresses in the lists of P2P sub-LSPs match, then it is the legal LSP rerouting case mentioned here above.

If there are no overlap in the lists, and the LSR is capable of remerging in the data plane, this is considered legal. Else the new Path message MUST be handled according to remerge error processing as described below.

The LSR generates a PathErr message with Error Code "Routing Problem/P2MP Rmerge Detected" towards the upstream node (i.e. the node that sent the Path message) until it reaches the node that caused the remerge condition. Identification of the offending node requires special processing by the nodes upstream of the error. A node that receives a PathErr message that contains a the error "Routing Problem/P2MP Rmerge Detected" MUST check to see if it is the offending node. This check is done by comparing the P2P sub-LSPs listed in the PathErr message with existing LSP state. If any of the egresses are already present in any Path state associated with the P2MP TE LSP other than the one associated with the <SESSION, SENDER\_TEMPLATE> objects signaled in the PathErr message, then the node is the signaling branch node that caused the remerge condition. This node SHOULD then correct the remerge condition by adding all P2P sub-LSPs listed in the offending Path state to the Path state (and Path message) associated to these P2P sub-LSPs. Note that the new Path state may be sent out the same outgoing interface in different Path messages in order to meet IP packet size limitations. If use of a new outgoing interface violates one or more SERO constraint, then a



PathErr message containing the associated egresses and any identified valid egresses SHOULD be generated with the error code "Routing Problem" and error value of "ERO Resulted in Reremerge".

This process may continue hop-by-hop until the ingress is reached. The only case where this process will fail is when all the listed P2P sub-LSPs are deleted prior to the PathErr message propagating to the ingress. In this case, the whole process will be corrected on the next (refresh or trigger) transmission of the offending Path message.

In all cases where a reremerge error is not detected, normal processing continues.

### **23.1. PathErr Message Format**

As described above, in the case where remerging is detected, a PathErr message will include one or more P2P\_SUB\_LSP objects. The resulting modified for a PathErr Message is:

```
<PathErr Message> ::=      <Common Header> [ <INTEGRITY> ]
                             [ [ <MESSAGE_ID_ACK> | <MESSAGE_ID_NACK> ]
... ]
                             [ <MESSAGE_ID> ]
                             <SESSION> <ERROR_SPEC>
                             [ <ACCEPTABLE_LABEL_SET> ... ]
                             [ <POLICY_DATA> ... ]
                             <sender descriptor>
                             [ <P2P sub-LSP descriptor list> ]
```

PathErr messages generation is unmodified, but nodes that set the Sub-Group Originator field and propagate a received PathErr message upstream MUST replace the Sub-Group fields received in the PathErr message with the value that was received in the Sub-Group fields of the Path message from the upstream neighbor. Note the receiver of a PathErr message is able to identify the errored outgoing Path message, and outgoing interface, based on the Sub-Group fields received in the error message.



## 24. New and Updated Message Objects

This section presents the RSVP message related formats as modified by this document.

### 24.1. P2MP LSP Tunnel SESSION Object

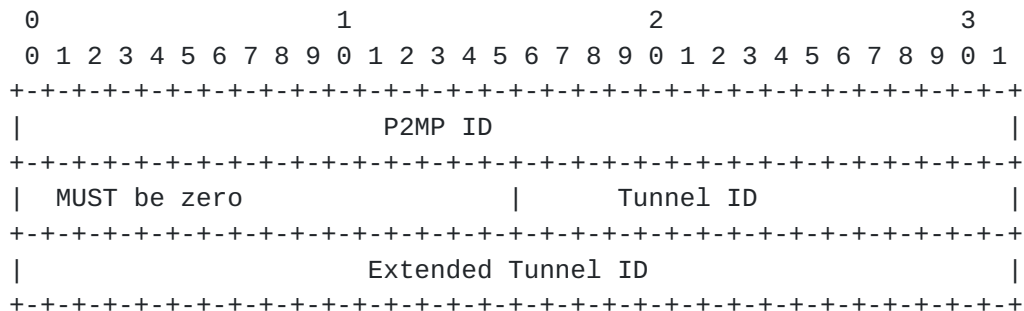
A P2MP LSP Tunnel SESSION object is used. This object uses the existing SESSION C-Num. New C-Types are defined to accommodate a logical P2MP destination identifier of the P2MP Tunnel. This SESSION object has a similar structure as the existing point to point RSVP-TE SESSION object. However the destination address is set to the P2MP ID instead of the unicast Tunnel Endpoint address. All P2P sub-LSPs part of the same P2MP LSP Tunnel share the same SESSION object. This SESSION object identifies the P2MP Tunnel.

The combination of the SESSION object, the SENDER\_TEMPLATE object and the P2P SUB-LSP object, identifies each P2P sub-LSP. This follows the existing P2P R

SVP-TE notion of using the SESSION object for identifying a P2P Tunnel which in turn can contain multiple LSP Tunnels, each distinguished by a unique SENDER\_TEMPLATE object.

#### 24.1.1. P2MP IPv4 LSP SESSION Object

Class = SESSION, P2MP\_LSP\_TUNNEL\_IPv4 C-Type = TBD



#### P2MP ID

A 32-bit identifier used in the SESSION object that remains constant over the life of the P2MP tunnel. It encodes the P2MP ID and identifies the set of destinations of the P2MP LSP Tunnel.

#### Tunnel ID

A 16-bit identifier used in the SESSION object that remains constant over the life of the P2MP tunnel.



## Extended Tunnel ID

A 32-bit identifier used in the SESSION object that remains constant over the life of the P2MP tunnel. Normally set to all zeros. Ingress nodes that wish to narrow the scope of a SESSION to the ingress-PID pair may place their IPv4 address here as a globally unique identifier [RFC3209].

#### 24.1.2. P2MP IPv6 LSP SESSION Object

This is same as the P2MP IPv4 LSP SESSION Object with the difference that the extended tunnel ID may be set to a 16 byte identifier [RFC3209].

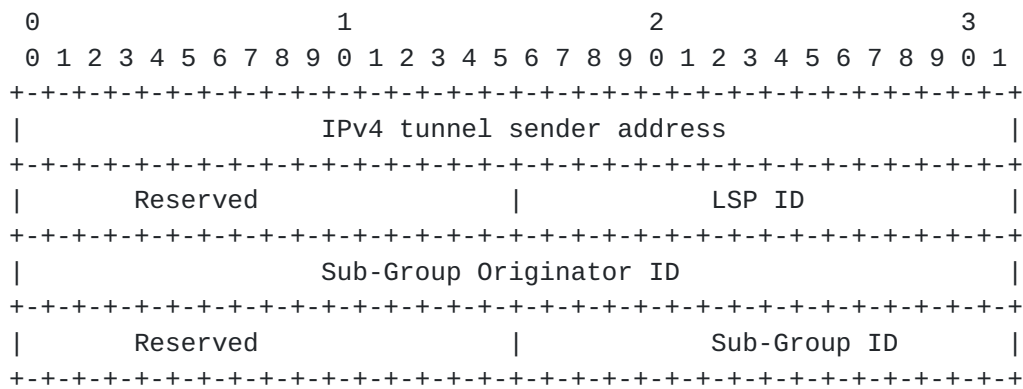
## 24.2. SENDER\_TEMPLATE object

The sender template contains the ingress-LSR source address. LSP ID can be changed to allow a sender to share resources with itself. Thus multiple instances of the P2MP tunnel can be created, each with a different LSP ID. The instances can share resources with each other, but use different labels. The P2P sub-LSPs corresponding to a particular instance use the same LSP ID.

As described in [section 6.1](#) it is necessary to distinguish different Path messages that are used to signal state for the same P2MP LSP Tunnel by using a <Sub-Group ID Originator ID, Sub-Group ID> tuple. There are various methods to encode this information. This document proposes the use of the SENDER\_TEMPLATE object and modifies it to carry this information as shown below. This encoding is subject to review by the MPLS WG.

#### 24.2.1. P2MP IPv4 LSP Tunnel SENDER\_TEMPLATE Object

Class = SENDER\_TEMPLATE, P2MP\_LSP\_TUNNEL\_IPv4 C-Type = TBD







IPv4 tunnel sender address

See [[RFC3209](#)]

Sub-Group Originator ID

The Sub-Group Originator ID is set to the TE Router ID of the LSR that originates the Path message. This is either the ingress LSR or a LSR which re-originates the Path message with its own Sub-Group Originator ID.

Sub-Group ID

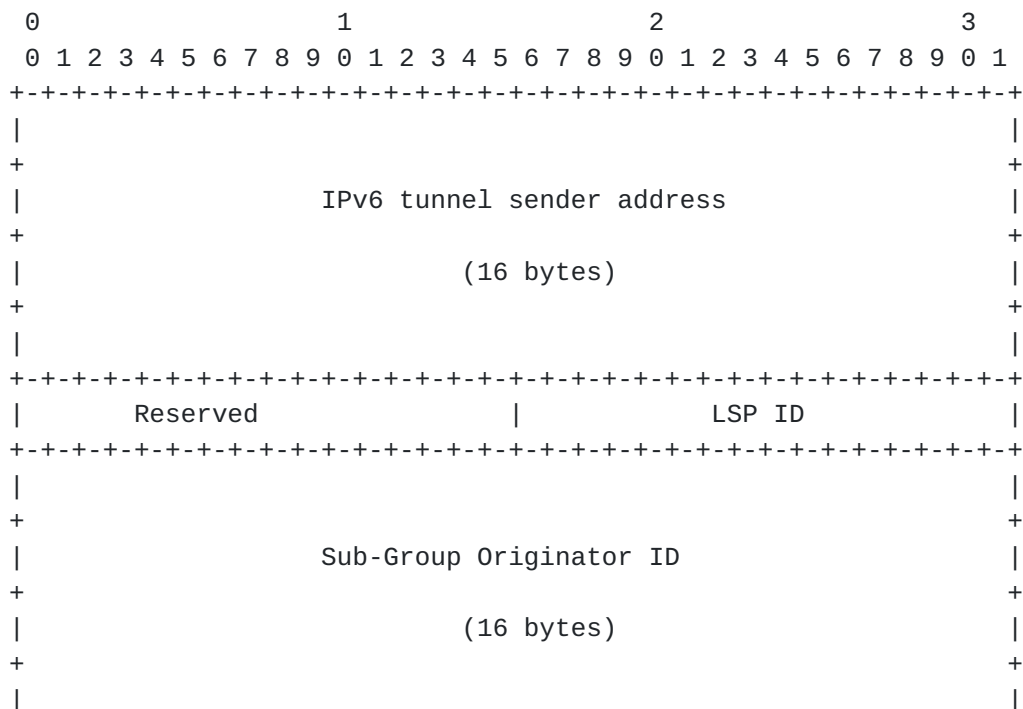
An identifier of a Path message used to differentiate multiple Path messages that signal state for the same P2MP LSP. This may be seen as identifying a group of one or more egress nodes targeted by this Path message. If the third mechanism for pruning is used as described in [section 11](#), the Sub-Group ID value of zero (0) has special meaning and MUST NOT be used with P2MP LSP Tunnels in messages other than PathTear messages. Use of a Sub-Group ID value of zero (0) in PathTear messages is defined below.

LSP ID

See [[RFC3209](#)]

#### 24.2.2.2. P2MP LSP Tunnel IPv6 SENDER\_TEMPLATE Object

Class = SENDER\_TEMPLATE, P2MP\_LSP\_TUNNEL\_IPv6 C-Type = TBD





```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Reserved           |           Sub-Group ID           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

IPv6 tunnel sender address

See [<RFC3209>]

Sub-Group Originator ID

The Sub-Group Originator ID is set to the IPv6 TE Router ID of the LSR that originates the Path message. This is either the ingress LSR or a LSR which re-originates the Path message with its own Sub-Group Originator ID.

Sub-Group ID

As above.

LSP ID

See [<RFC3209>]

### **24.3. P2P SUB-LSP IPv4 Object**

A new P2P Sub-LSP object identifies a particular P2P sub-LSP belonging to the P2MP LSP Tunnel.

#### **24.3.1. P2P SUB-LSP IPv4 Object**

SUB\_LSP Class = TBD, P2P\_SUB\_LSP\_IPv4 C-Type = TBD

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           IPv4 P2P Sub-LSP destination address           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| MUST be zero           |           Sub-LSP               |
ID      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

IPv4 Sub-LSP destination address

IPv4 address of the P2P sub-LSP destination.

(There is NO-CONSENSUS amongst the authors on the sub-LSP ID described below and it needs more discussion)

Sub-LSP ID

A 16-bit identifier that identifies a particular instance



of a P2P sub-LSP. It can be varied for P2P sub-LSP make-before-break. Different P2P sub-LSPs, with the same SESSION object and LSP ID, follow the label merge semantics described in [section 3](#) to form a particular instance of the P2MP tunnel.

#### **[24.3.2.](#) P2P SUB-LSP IPv6 Object**

SUB\_LSP Class = TBD, P2P\_SUB\_LSP\_IPv6 C-Type = TBD

This is same as the P2P IPv4 Sub-LSP object, with the difference that the destination address is a 16 byte IPv6 address.

#### **[24.4.](#) FILTER\_SPEC Object**

The FILTER\_SPEC object is canonical to the P2MP SENDER\_TEMPLATE object.

##### **[24.4.1.](#) P2MP LSP\_TUNNEL\_IPv4 FILTER\_SPEC Object**

Class = FILTER SPEC, P2MP LSP\_TUNNEL\_IPv4 C-Type = TBD

The format of the P2MP LSP\_TUNNEL\_IPv4 FILTER\_SPEC object is identical to the P2MP LSP\_TUNNEL\_IPv4 SENDER\_TEMPLATE object.

##### **[24.4.2.](#) P2MP LSP\_TUNNEL\_IPv6 FILTER\_SPEC Object**

Class = FILTER SPEC, P2MP LSP\_TUNNEL\_IPv6 C\_Type = TBD

The format of the P2MP LSP\_TUNNEL\_IPv6 FILTER\_SPEC object is identical to the P2MP LSP\_TUNNEL\_IPv6 SENDER\_TEMPLATE object.

#### **[24.5.](#) SUB\_EXPLICIT\_ROUTE Object (SER0)**

The SER0 is defined as identical to the ER0. The CNums are TBD and TBD of the form 11bbbbbb.

#### **[24.6.](#) SUB\_RECORD\_ROUTE Object (SRRO)**

The SRRO is defined as identical to the RR0. The CNums are TBD and TBD of the form 11bbbbbb.



## **25. IANA Considerations**

### **25.1. New Message Objects**

IANA considerations for new message objects will be specified after the objects used are decided upon.

### **25.2. New Error Codes**

Two new Error Codes are defined for use with the Error Value "Routing Error". IANA is requested to assign values.

The Error Code "Unable to Branch" indicates that a P2MP branch cannot be formed by the reporting LSR.

The Error Code "Unsupported LSP Integrity" indicates that a P2MP branch does not support the requested LSP integrity function.

### **25.3. LSP Attributes Flags**

IANA has been asked to manage the space of flags in the Attributes Flags TLV carried in the LSP\_ATTRIBUTES Object [LSP-ATTRIB]. This document defines two new flags as follows:

Suggested Bit Number:	3
Meaning:	LSP Integrity Required
Used in Attributes Flags on Path:	Yes
Used in Attributes Flags on Resv:	No
Used in Attributes Flags on RRO:	No
Referenced Section of this Doc:	16

Suggested Bit Number:	4
Meaning:	Branch Reoptimization Allowed
Used in Attributes Flags on Path:	Yes
Used in Attributes Flags on Resv:	No
Used in Attributes Flags on RRO:	No
Referenced Section of this Doc:	17.3





## **26. Security Considerations**

This document does not introduce any new security issues. The security issues identified in [[RFC3209](#)] and [[RFC3473](#)] are still relevant.

## **27. Acknowledgements**

This document is the product of many people. The contributors are listed in [Section 26](#).

Thanks to Yakov Rekhter, Der-Hwa Gan, Arthi Ayyanger and Nischal Sheth for their suggestions and comments. Thanks also to Dino Farninacci for his comments.

## **28. Appendix**

### **28.1. Example**

Following is one example of setting up a P2MP LSP Tunnel using the procedures described in this document.

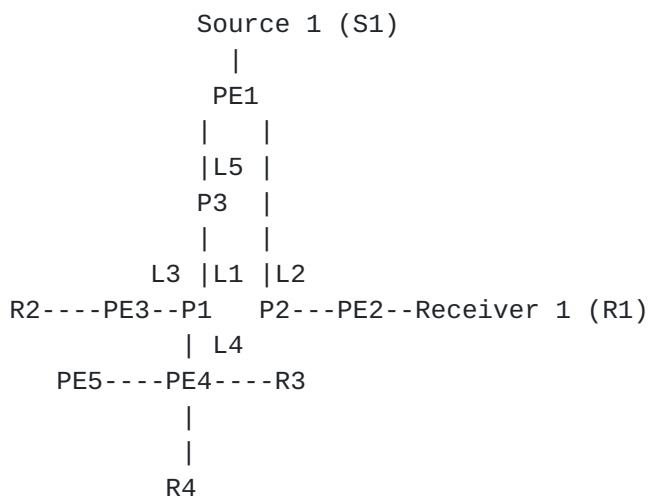


Figure 2.

The mechanism is explained using Figure 2. PE1 is the ingress-LSR. PE2, PE3 and PE4 are Egress-LSRs.

a) PE1 learns that PE2, PE3 and PE4 are interested in joining a P2MP tree with a P2MP ID of P2MP ID1. We assume that PE1 learns of the



egress-LSRs at different points.

b) PE1 computes the P2P path to reach PE2.

c) PE1 establishes the P2P sub-LSP to PE2 along <PE1, P2, PE2>

d) PE1 computes the P2P path to reach PE3 when it discovers PE3. This path is computed to share the same links where possible with the sub-LSP to PE2 as they belong to the same P2MP session.

e) PE1 establishes the P2P sub-LSP to PE3 along <PE1, P3, P1, PE3>

f) PE1 computes the P2P path to reach PE4 when it discovers PE4. This path is computed to share the same links where possible with the sub-LSPs to PE2 and PE3 as they belong to the same P2MP session.

g) PE1 signals the Path message for PE4 sub-LSP along <PE1, P3, P1, PE4>

e) P1 receives a Resv message from PE4 with label L4. It had previously received a Resv message from PE3 with label L3. It had allocated a label L1 for the sub-LSP to PE3. It uses the same label and sends the Resv messages to P3. Note that it may send only one Resv message with multiple flow descriptors in the flow descriptor list. If this is the case and FF style is used, the FF flow descriptor will contain the P2P sub-LSP descriptor list with two entries: one for PE4 and the other for PE3. For SE style, the SE filter spec will contain this P2P sub-LSP descriptor list. P1 also creates a label mapping of (L1 -> {L3, L4}). P3 uses the existing label L5 and sends the Resv message to PE1, with label L5. It reuses the label mapping of {L5 -> L1}.

## **29. References**

### **29.1. Normative References**

[LSP-HIER] K. Kompella, Y. Rekhter, "LSP Hierarchy with Generalized MPLS TE", [draft-ietf-mpls-lsp-hierarchy-08.txt](#).

[LSP-ATTR] A. Farrel, et. al. , "Encoding of Attributes for Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Establishment Using RSVP-TE", [draft-ietf-mpls-rsvpte-attributes-03.txt](#), March 2004, work in progress.

[RFC3209] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan,



G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels",  
[RFC3209](#), December 2001

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2205] Braden, R., Zhang, L., Berson, S., Herzog, S. and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1, Functional Specification", [RFC 2205](#), September 1997.
- [RFC3471] Lou Berger, et al., "Generalized MPLS - Signaling Functional Description", [RFC 3471](#), January 2003
- [RFC3473] L. Berger et.al., "Generalized MPLS Signaling - RSVP-TE Extensions", [RFC 3473](#), January 2003.
- [RFC2961] L. Berger, D. Gan, G. Swallow, P. Pan, F. Tommasi, S. Molendini, "RSVP Refresh Overhead Reduction Extensions", [RFC 2961](#), April 2001.
- [RFC3031] Rosen, E., Viswanathan, A. and R. Callon, "Multiprotocol Label Switching Architecture", [RFC 3031](#), January 2001.
- [RSVP-FRR] P. Pan, G. Swallow, A. Atlas (Editors), "Fast Reroute Extensions  
to RSVP-TE for LSP Tunnels",  
[draft-ietf-mpls-rsvp-lsp-fastreroute-07.txt](#).
- [RFC3477] K. Kompella, Y. Rekther, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)".

## **29.2. Informative References**

- [BFD] D. Katz, D. Ward, "Bidirectional Forwarding Detection",  
[draft-katz-ward-bfd-01.txt](#).
- [BFD-MPLS] R. Aggarwal, K. Kompella, "BFD for MPLS LSPs",  
[draft-raggarwa-mpls-bfd-00.txt](#)
- [IPR-1] Bradner, S., "IETF Rights in Contributions", [BCP 78](#),  
[RFC 3667](#), February 2004.
- [IPR-2] Bradner, S., Ed., "Intellectual Property Rights in IETF Technology", [BCP 79](#), [RFC 3668](#), February 2004.
- [INT-REG] JP Vasseur, A. Ayyangar, "Inter-area and Inter-AS MPLS Traffic Engineering", [draft-vasseur-ccamp-inter-area-as-te-00.txt](#).



[P2MP-REQ] S. Yasukawa, et. al., "Requirements for Point-to-Multipoint capability extension to MPLS",  
[draft-ietf-mpls-p2mp-requirement-04.txt](#).

[RFC2209] R. Braden, L. Zhang, "Resource Reservation Protocol (RSVP) Version 1 Message Processing Rules", [RFC 2209](#).

## **[30.](#) Author Information**

### **[30.1.](#) Editor Information**

Rahul Aggarwal  
Juniper Networks  
1194 North Mathilda Ave.  
Sunnyvale, CA 94089  
Email: [rahul@juniper.net](mailto:rahul@juniper.net)

Seisho Yasukawa  
NTT Corporation  
9-11, Midori-Cho 3-Chome  
Musashino-Shi, Tokyo 180-8585 Japan  
Phone: +81 422 59 4769  
Email: [yasukawa.seisho@lab.ntt.co.jp](mailto:yasukawa.seisho@lab.ntt.co.jp)

Dimitri Papadimitriou  
Alcatel  
Francis Wellesplein 1,  
B-2018 Antwerpen, Belgium  
Phone: +32 3 240-8491  
Email: [Dimitri.Papadimitriou@alcatel.be](mailto:Dimitri.Papadimitriou@alcatel.be)

### **[30.2.](#) Contributor Information**

John Drake  
Calient Networks  
Email: [jdrake@calient.net](mailto:jdrake@calient.net)

Alan Kullberg  
Motorola Computer Group  
120 Turnpike Road 1st Floor  
Southborough, MA 01772  
Email: [alan.kullberg@motorola.com](mailto:alan.kullberg@motorola.com)





Lou Berger  
Movaz Networks, Inc.  
7926 Jones Branch Drive  
Suite 615  
McLean VA, 22102  
Phone: +1 703 847-1801  
EMail: lberger@movaz.com

Liming Wei  
Redback Networks  
350 Holger Way  
San Jose, CA 95134  
Email: lwei@redback.com

George Apostolopoulos  
Redback Networks  
350 Holger Way  
San Jose, CA 95134  
Email: georgeap@redback.com

Kireeti Kompella  
Juniper Networks  
1194 N. Mathilda Ave  
Sunnyvale, CA 94089  
Email: kireeti@juniper.net

George Swallow  
Cisco Systems, Inc.  
300 Beaver Brook Road  
Boxborough , MA - 01719  
USA  
Email: swallow@cisco.com

JP Vasseur  
Cisco Systems, Inc.  
300 Beaver Brook Road  
Boxborough , MA - 01719  
USA  
Email: jpv@cisco.com

Dean Cheng  
Cisco Systems Inc.  
170 W Tasman Dr.  
San Jose, CA 95134  
Phone 408 527 0677  
Email: dcheng@cisco.com

Markus Jork



Avici Systems  
101 Billerica Avenue  
N. Billerica, MA 01862  
Phone: +1 978 964 2142  
EMail: mjork@avici.com

Hisashi Kojima  
NTT Corporation  
9-11, Midori-Cho 3-Chome  
Musashino-Shi, Tokyo 180-8585 Japan  
Phone: +81 422 59 6070  
EMail: kojima.hisashi@lab.ntt.co.jp

Andrew G. Malis  
Tellabs  
2730 Orchard Parkway  
San Jose, CA 95134  
Phone: +1 408 383 7223  
Email: Andy.Malis@tellabs.com

Koji Sugisono  
NTT Corporation  
9-11, Midori-Cho 3-Chome  
Musashino-Shi, Tokyo 180-8585 Japan  
Phone: +81 422 59 2605  
EMail: sugisono.koji@lab.ntt.co.jp

Masanori Uga  
NTT Corporation  
9-11, Midori-Cho 3-Chome  
Musashino-Shi, Tokyo 180-8585 Japan  
Phone: +81 422 59 4804  
EMail: uga.masanori@lab.ntt.co.jp

Igor Bryskin  
Movaz Networks, Inc.  
7926 Jones Branch Drive  
Suite 615  
McLean VA, 22102

Adrian Farrel  
Old Dog Consulting  
Phone: +44 0 1978 860944  
EMail: adrian@olddog.co.uk

Jean-Louis Le Roux  
France Telecom  
2, avenue Pierre-Marzin



22307 Lannion Cedex  
France  
E-mail: [jeanlouis.leroux@francetelecom.com](mailto:jeanlouis.leroux@francetelecom.com)

### **31. Intellectual Property**

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

### **32. Full Copyright Statement**

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#) and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.



### **33. Acknowledgement**

Funding for the RFC Editor function is currently provided by the Internet Society.