Abstract

This document presents a set of requirements for the establishment and maintenance of Point-to-Multipoint (P2MP) Traffic Engineered (TE) Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs).

There is no intent to specify solution specific details nor application specific requirements in this document.

It is intended that the requirements presented in this document are not limited to the requirements of packet switched networks, but also encompass the requirements of Layer two Switching (L2SC), Time Division Multiplexing (TDM), lambda and port switching networks managed by Generalized MPLS (GMPLS) protocols. Protocol solutions developed to meet the requirements set out in this document must attempt to be equally applicable to MPLS and GMPLS.
Table of Contents

1. Introduction .................................................. 03
   1.1 Non-Objectives ........................................... 05
2. Definitions ................................................... 06
   2.1 Acronyms ................................................ 06
   2.2 Terminology ............................................. 06
       2.2.1 Terminology for Partial LSPs .................... 07
   2.3 Conventions ............................................. 08
3. Problem Statement ............................................ 08
   3.1 Motivation ............................................... 08
   3.2 Requirements Overview ................................... 09
4. Detailed requirements for P2MP TE extensions .................. 11
   4.1 P2MP LSP tunnels ........................................ 11
   4.2 P2MP explicit routing ................................... 11
   4.3 Explicit Path Loose Hops and Widely Scoped Abstract Nodes . 12
   4.4 P2MP TE LSP establishment, teardown, and modification
       mechanisms .................................................. 13
   4.5 Fragmentation ............................................ 14
   4.6 Failure Reporting and Error Recovery ...................... 14
   4.7 Record route of P2MP TE LSP tunnels ..................... 15
   4.8 Call Admission Control (CAC) and QoS Control mechanism
       of P2MP TE LSPs ........................................... 16
   4.9 Variation of LSP Parameters ................................ 16
   4.10 Re-optimization of P2MP TE LSPs ........................ 16
   4.11 Tree Remerge ............................................ 17
   4.12 Data Duplication ........................................ 18
   4.13 IPv4/IPv6 support ....................................... 19
   4.14 P2MP MPLS Label ........................................ 19
   4.15 Routing advertisement of P2MP capability ................. 19
   4.16 Multi-Area/AS LSP ........................................ 19
   4.17 Multi-access LANs ....................................... 20
   4.18 P2MP MPLS OAM .......................................... 20
   4.19 Scalability ............................................. 20
       4.19.1 Absolute Limits ................................... 21
   4.20 Backwards Compatibility ................................ 23
   4.21 GMPLS .................................................. 23
   4.22 Requirements for Hierarchical P2MP TE LSPs ............. 24
   4.23 P2MP Crankback routing ................................ 24
5. Security Considerations ....................................... 24
6. IANA Considerations ........................................... 25
7. Acknowledgements ............................................. 25
8. References ..................................................... 25
   8.1 Normative References .................................... 25
   8.2 Informational References ................................ 26
9. Editor’s Address ............................................... 27
10. Authors’ Addresses ........................................... 27
11. Intellectual Property Consideration .......................... 28
12. Full Copyright Statement ..................................... 29
1. Introduction

Existing MPLS Traffic Engineering (MPLS-TE) allows for strict QoS guarantees, resources optimization, and fast failure recovery, but is limited to point-to-point (P2P) applications. Requirements have been expressed for the provision of services over point-to-multipoint (P2MP) traffic engineered tunnels and this clearly motivates enhancements of the base MPLS-TE tool box in order to support P2MP MPLS-TE LSPs.

[RFC2702] specifies requirements for traffic engineering over MPLS. It describes traffic engineering in some detail, and those definitions and objectives are equally applicable to traffic engineering in a point-to-multipoint service environment. They are not repeated here, but it is assumed that the reader is fully familiar with them.

[RFC2702] also explains how MPLS is particularly suited to traffic engineering, and presents the following eight reason.

1. Explicit label switched paths which are not constrained by the destination based forwarding paradigm can be easily created through manual administrative action or through automated action by the underlying protocols.
2. LSPs can potentially be efficiently maintained.
3. Traffic trunks can be instantiated and mapped onto LSPs.
4. A set of attributes can be associated with traffic trunks which modulate their behavioral characteristics.
5. A set of attributes can be associated with resources which constrain the placement of LSPs and traffic trunks across them.
6. MPLS allows for both traffic aggregation and disaggregation whereas classical destination only based IP forwarding permits only aggregation.
7. It is relatively easy to integrate a "constraint-based routing" framework with MPLS.
8. A good implementation of MPLS can offer significantly lower overhead than competing alternatives for Traffic Engineering.

These points are equally applicable to point-to-multipoint traffic engineering. Points 1. and 7. are particularly important.

That is, the traffic flow for a point-to-multipoint LSP is not constrained to the path or paths that it would follow during multicast routing or shortest path destination-based routing, but can be explicitly controlled through manual or automated action.

Further, the explicit paths that are used may be computed using
algorithms based on a variety of constraints to produce all manner of tree shapes. For example, an explicit path may be cost-based [STEINER], shortest path, QoS-based, or may use some fair-cost QoS algorithm.

[RFC2702] also describes the functional capabilities required to fully support Traffic Engineering over MPLS in large networks.

1. A set of attributes associated with traffic trunks which collectively specify their behavioral characteristics.

2. A set of attributes associated with resources which constrain the placement of traffic trunks through them. These can also be viewed as topology attribute constraints.

3. A "constraint-based routing" framework which is used to select paths for traffic trunks subject to constraints imposed by items 1) and 2) above. The constraint-based routing framework does not have to be part of MPLS. However, the two need to be tightly integrated together.

These basic requirements should also be supported by P2MP traffic engineering.

This document presents a set of requirements for Point-to-Multipoint(P2MP) Traffic Engineering (TE) extensions to Multiprotocol Label Switching (MPLS). It specifies functional requirements for solutions to deliver P2MP TE LSPs.

It is intended that solutions that specify procedures for P2MP TE LSP setup satisfy these requirements. There is no intent to specify solution specific details nor application specific requirements in this document.

It is intended that the requirements presented in this document are not limited to the requirements of packet switched networks, but also encompass the requirements of TDM, lambda and port switching networks managed by Generalized MPLS (GMPLS) protocols. Protocol solutions developed to meet the requirements set out in this document must attempt to be equally applicable to MPLS and GMPLS.

Existing MPLS TE mechanisms such as [RFC3209] do not support P2MP TE LSPs so new mechanisms must be developed. This should be achieved with maximum re-use of existing MPLS protocols. Note that there is a separation between routing and signaling in MPLS TE. In particular, the path of the MPLS TE LSP is determined by performing a constraint-based computation (such as CSPF) on a
traffic engineering database (TED). The contents of the TED may be collected through a variety of mechanism - extensions to the IGPs are a popular mechanism for P2P MPLS TE.

This document focuses on requirements for establishing and maintaining P2MP MPLS TE LSPs through signaling protocols; and routing protocols are out of scope. No assumptions are made about how the TED used as the basis for path computations for P2MP LSPs is formed.

A P2MP TE LSP will be set up with TE constraints and will allow efficient packet or data replication at various branching points in the network. Note that the notion of "efficient" packet replication is relative and may have different meanings depending on the objectives (see section 4.2).

P2MP TE LSP setup mechanisms MUST include the ability to add/remove receivers to/from an existing P2MP TE LSP.

1.1 Non-Objectives

For clarity, this section lists some items that are out of scope of this document.

It is assumed that some information elements describing the P2MP TE LSP are known to the ingress LSR prior to LSP establishment. For example, the ingress LSRs knows the IP addresses that identify the egress LSRs of the P2MP TE LSP. The mechanisms by which the ingress LSR obtains this information is outside the scope of P2MP TE signaling and so is not included in this document. Other documents may complete the description of this function by providing automated, protocol-based ways of passing this information to the ingress LSR.

The following are non-objectives of this document.

- Non-TE LSPs (such as per-hop, routing-based LSPs).
- Discovery of egress leaves for a P2MP LSP
- Hierarchical P2MP LSPs
- OAM for P2MP LSPs
- Inter-area and inter-AS P2MP TE LSPs
- Applicability of P2MP MPLS TE LSPs to service scenarios
- Specific application or application requirements
- Algorithms for computing P2MP distribution trees
- Multipoint-to-point LSPs
- Multipoint-to-multipoint LSPs
- Routing protocols
- Construction of the traffic engineering database
- Distribution of the information used to construct the traffic engineering database

2. Definitions

2.1 Acronyms

P2P:
Point-to-point

P2MP:
Point-to-multipoint

2.2 Terminology

The reader is assumed to be familiar with the terminology in [RFC3031] and [RFC3209].

P2MP TE LSP:
A traffic engineered label switched path that has one unique ingress LSR (also referred to as the root) and one or more egress LSRs (also referred to as the leaf).

P2MP tree:
The ordered set of LSRs and links that comprise the path of a P2MP TE LSP from its ingress LSR to all of its egress LSRs.

ingress LSR:
The LSR that is responsible for initiating the signaling messages that set up the P2MP TE LSP.

egress LSR:
One of potentially many destinations of the P2MP TE LSP. Egress LSRs may also be referred to as leaf nodes or leaves.
bud LSR:

An LSR that is an egress, but also has one or more directly connected downstream LSRs.

branch LSR:

An LSR that has more than one directly connected downstream LSR.

graft LSR:

An LSR that is already a member of the P2MP tree and is in process of signaling a new sub-P2MP tree.

prune LSR:

An LSR that is a member of the P2MP tree and is in process of tearing down an existing sub-P2MP tree.

P2MP-ID (Pid):

A unique identifier of a P2MP TE LSP, that is constant for the whole LSP regardless of the number of branches and/or leaves.

2.2.1 Terminology for Partial LSPs

It is convenient to sub-divide P2MP trees for functional and representational reasons. A tree may be divided in two dimensions:

- A division may be made along the length of the tree. For example, the tree may be split into two components each running from the ingress LSR to a discrete set of egress LSRs
- A tree may be divided at a branch LSR (or any transit LSR) to produce a component of the tree that runs from the branch (or transit) LSR to all downstream egress LSRs.

These two methods of splitting the P2MP tree can be combined, so it is useful to introduce some terminology to allow the partitioned trees to be clearly described.

Use the following designations:
- Source (ingress) LSR - S
- Leaf (egress) LSR - L
- Branch LSR - B
- Transit LSR - X
Define three terms:

Sub-LSP
A component of the P2MP LSP that runs from one LSR to another without (or ignoring) any branches.

Sub-tree
A component of the P2MP LSP that runs from one LSR to more than one other LSR by branching.

Tree
A component of the P2MP LSP that runs from one LSR to all downstream LSRs.

Using these new concepts we can define any combination or split of the P2MP tree. For example:

S2L sub-LSP
The path from the source to one specific leaf.

S2L sub-tree
The path from the source to a set of leaves.

B2L tree
The path from a branch LSR to all downstream leaves.

X2X sub-LSP
A component of the P2MP LSP that is a simple path with no branches.

2.3 Conventions

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. Problem Statement

3.1 Motivation

As described in section 1, Traffic Engineering and Constraint Based Routing, including Call Admission Control(CAC), explicit source routing and bandwidth reservation, are required to enable efficient resource usage and strict QoS guarantees. Such mechanisms also make it possible to provide services across a congested network where conventional "shortest path first" forwarding paradigms would fail.

Existing MPLS TE mechanisms [RFC3209] and GMPLS TE mechanisms...
[RFC3473] only provide support for P2P TE LSPs. While it is possible to provide P2MP TE services using P2P TE LSPs, any such approach is potentially suboptimal since it may result in data replication at the ingress LSR, or in duplicate data traffic within the network.

Hence, to provide P2MP MPLS TE services in a fully efficient manner it is necessary to specify specific requirements. These requirements can then be used to define mechanisms for the use of existing protocols and/or extensions to existing protocols and/or new protocols.

3.2. Requirements Overview

This document states basic requirements for the setup of P2MP TE LSPs. The requirements apply to the signaling techniques only, and no assumptions are made about which routing protocols are run within the network, nor about how the information that is used to construct the Traffic Engineering Database (TED) is distributed. These factors are out of the scope of this document.

A P2MP TE LSP path will be computed taking into account various constraints such as bandwidth, affinities, required level of protection and so on. The solution MUST allow for the computation of P2MP TE LSP paths satisfying constraints with the objective of supporting various optimization criteria such as delays, bandwidth consumption in the network, or any other combinations. This is likely to require the presence of a TED, as well as the ability to signal the explicit path of an LSP.

A desired requirement is also to maximize the re-use of existing MPLS TE techniques and protocols where doing so does not adversely impact the function, simplicity or scalability of the solution.

This document does not restrict the choice of signaling protocol used to set up a P2MP TE LSP, but it should be noted that [RFC3468] states

... the consensus reached by the Multiprotocol Label Switching (MPLS) Working Group within the IETF to focus its efforts on "Resource Reservation Protocol (RSVP)-TE: Extensions to RSVP for Label-Switched Paths (LSP) Tunnels" (RFC 3209) as the MPLS signaling protocol for traffic engineering applications...

The P2MP TE LSP setup mechanism MUST include the ability to add/remove egress LSRs to/from an existing P2MP TE LSP and MUST allow for the support of all the TE LSP management procedures already defined for P2P TE LSP such as the non disruptive rerouting (the so called "Make before break" procedure).
The computation of P2MP TE trees is implementation dependent and is beyond the scope of the solutions that are built with this document as a guideline.

Consider the following figure.

![Diagram of P2MP TE tree](image)

Figure 1

Figure 1 shows a single ingress (I-LSR1), and four egresses (E-LSR2, E-LSR3, E-LSR4 and E-LSR5). I-LSR1 is attached to a traffic source that is generating traffic for a P2MP application. Receivers R1, R2, R3 and R4 are attached to E-LSR2, E-LSR3 and E-LSR4.

The following are the objectives of P2MP LSP establishment and use.

a) A P2MP TE LSP tree which satisfies various constraints is pre-determined and supplied to ingress I-LSR1.

Note that no assumption is made on whether the tree is provided to I-LSR1 or computed by I-LSR1. Note that the solution SHOULD also allow for the support of partial path by means of loose routing.

Typical constraints are bandwidth requirements, resource class affinities, fast rerouting, preemption, to mention a few of them. There should not be any restriction on the possibility to support the set of constraints already defined for point to point TE LSPs. A new constraint may specify which LSRs should be used as branch points for the P2MP LSR in order to take into account some LSR capabilities or network constraints.

b) A P2MP TE LSP is set up from I-LSR1 to E-LSR2, E-LSR3 and E-LSR4 using the tree information.
c) In this case, the branch LSR1 should replicate incoming packets or data and send them to E-LSR3 and E-LSR4.

d) If a new receiver (R5) expresses an interest in receiving traffic, a new tree is determined and a sub-P2MP tree from LSR2 to E-LSR5 is grafted onto the P2MP tree. LSR2 becomes a branch LSR.

4. Detailed requirements for P2MP TE extensions

4.1 P2MP LSP tunnels

The P2MP TE extensions MUST be applicable to the signaling of LSPs of different traffic types. For example, it MUST be possible to signal a P2MP TE LSP to carry any kind of payload being packet or non-packet based (including frame, cell, TDM un/structured, etc.)

As with P2P MPLS technology [RFC3031], traffic is classified with a FEC in this extension. All packets which belong to a particular FEC and which travel from a particular node MUST follow the same P2MP tree.

In order to scale to a large number of branches, P2MP TE LSPs SHOULD be identified by a unique identifier (the P2MP ID or Pid) that is constant for the whole LSP regardless of the number of branches and/or leaves. Therefore, the identification of the P2MP session by its destination addresses is not adequate.

4.2 P2MP explicit routing

Various optimizations in P2MP tree formation need to be applied to meet various QoS requirements and operational constraints.

Some P2MP applications may request a bandwidth guaranteed P2MP tree which satisfies end-to-end delay requirements. And some operators may want to set up a cost minimum P2MP tree by specifying branch LSRs explicitly.

The P2MP TE solution therefore MUST provide a means of establishing arbitrary P2MP trees under the control of an external tree computation process or path configuration process or dynamic tree computation process located on the ingress LSR. Figure 4 shows two typical examples.
One example is the Steiner P2MP tree (Cost minimum P2MP tree) [STEINER]. This P2MP tree is suitable for constructing a cost minimum P2MP tree so as to minimize the bandwidth consumption in the core. To realize this P2MP tree, several intermediate LSRs must be both MPLS data terminating LSRs and transit LSRs (LSRs E, F, G, H, I, J, and K in the figure 4). This means that the LSRs must perform both label swapping and popping at the same time. Therefore, the P2MP TE solution MUST support a mechanism that can setup this kind of bud LSR between an ingress LSR and egress LSRs. Note that this includes constrained Steiner trees that allow for the computation of a minimal cost trees with some other constraints such as a bounded delay between the source and every receiver.

Another example is a CSPF (Constraint Shortest Path First) P2MP tree. By some metric (which can be set upon any specific criteria like the delay, bandwidth, a combination of those), one can calculate a shortest path P2MP tree. This P2MP tree is suitable for carrying real time traffic.

The solution MUST allow the operator to make use of any tree computation technique. In the former case an efficient/optimal tree is defined as a minimal cost tree (Steiner tree) whereas in the later case it is defined as the tree that provides shortest path between the source and any receiver.

To support explicit setup of any reasonable P2MP tree shape, a P2MP TE solution MUST support some form of explicit source-based control of the P2MP tree which can explicitly include particular LSRs as branch nodes. This can be used by the ingress LSR to setup the P2MP TE LSP. For instance, a P2MP TE LSP can be simply represented as a whole tree or by its individual branches.

4.3 Explicit Path Loose Hops and Widely Scoped Abstract Nodes

A P2MP tree is completely specified if all of the required branches and hops between a sender and leaf LSR are indicated.
A P2MP tree is partially specified if only a subset of intermediate branches and hops are indicated. This may be achieved using loose hops in the explicit path, or using widely scoped abstract nodes such as IPv4 prefixes shorter than 32 bits, or AS numbers. A partially specified P2MP tree might be particularly useful in inter-area and inter-AS situations although P2MP requirements for inter-area and inter-AS are beyond the scope of this document.

Protocol solutions SHOULD include a way to specify loose hops and widely scoped abstract nodes in the explicit source-based control of the P2MP tree as defined in the previous section. Where this support is provided, protocol solutions MUST allow downstream LSRs to apply further explicit control to the P2MP tree to resolve a partially specified tree into a (more) completely specified tree.

Protocol solutions MUST allow the P2MP tree to be completely specified at the ingress where sufficient information exists to allow the full tree to be computed.

In all cases, the egress nodes of the P2MP TE LSP must be fully specified.

In case of a tree being computed by some downstream LSRs (e.g. the case of hops specified as loose hops), the solution MUST provide the ability for the ingress LSR of the P2MP TE LSP to learn the full P2MP tree. Note that this requirement MAY be relaxed in some environments (e.g. Inter-AS) where confidentiality must be preserved.

4.4 P2MP TE LSP establishment, teardown, and modification mechanisms

The P2MP TE solution MUST support establishment, maintenance and teardown of P2MP TE LSPs in a scalable manner. This MUST include both the existence of very many LSPs at once, and the existence of very many destinations for a single P2MP LSP.

In addition to P2MP TE LSP establishment and teardown mechanism, it SHOULD implement partial P2MP tree modification mechanism.

For the purpose of adding sub-P2MP TE LSPs to an existing P2MP TE LSP, the extensions SHOULD support a grafting mechanism. For the purpose of deleting a sub-P2MP TE LSPs from an existing P2MP TE LSP, the extensions SHOULD support a pruning mechanism.

It is RECOMMENDED that these grafting and pruning operations do not cause any additional processing in nodes except along the path to the grafting and pruning node and its downstream nodes. Moreover, both grafting and pruning operations MUST not be traffic disruptive for the traffic currently forwarded along the P2MP tree.
4.5 Fragmentation

The P2MP TE solution MUST handle the situation where a single protocol message cannot contain all of the information necessary to signal the establishment of the P2MP LSP. It MUST be possible to establish the LSP in these circumstances.

This situation may arise in either of the following circumstances.
   a. The ingress LSR cannot signal the whole tree in a single message.

   b. The information in a message expands to be too large (or is discovered to be too large) at some transit node. This may occur because of some increase in the information that needs to be signaled or because of a reduction in the size of signaling message that is supported.

The solution to these problems SHOULD NOT rely on IP fragmentation, it is RECOMMENDED to rely on some protocol procedures specific to the signaling solution.

It is NOT RECOMMENDED that fragmented protocol messages are re-combined at any downstream LSR.

4.6 Failure Reporting and Error Recovery

Failure events may cause egress nodes or sub-P2MP LSPs to become detached from the P2MP TE LSP. These events MUST be reported upstream as for a P2P LSP.

The solution SHOULD provide recovery techniques such as protection and restoration allowing recovery of any impacted sub-P2MP TE LSPs. In particular, a solution MUST provide fast protection mechanisms applicable to P2MP TE LSP similar to the solutions specified in [FRR] for P2P TE LSPs. Note also that no assumption is made on whether backup paths for P2MP TE LSPs should or should not be shared with P2P TE LSPs backup paths.

Note that the functions specified in [FRR] are currently specific to packet environments and do not apply to non-packet environments. Thus, while solutions MUST provide fast protection mechanisms similar to those specified in [FRR], this requirement is limited to the subset of the solution space that applies to packet switched networks only.

Note that application-specific requirement documents may introduce even more stringent requirement, such as no packet loss, as a trade-off for the relaxation of other requirements, such as increased...
bandwidth consumption.

The solution SHOULD also support the ability to meet other network recovery requirements such as bandwidth protection and bounded propagation delay increase along the backup path during failure.

A P2MP TE solution MUST support P2MP fast protection mechanism to handle P2MP applications sensitive to traffic disruption.

The report of the failure of delivery to fewer than all of the egress nodes SHOULD NOT cause automatic teardown of the P2MP TE LSP. That is, while some egress nodes remain connected to the P2MP tree it should be a matter of local policy at the ingress whether the P2MP LSP is retained.

When all egress nodes downstream of a branch node have become disconnected from the P2MP tree, and the some branch node is unable to restore connectivity to any of them by means of some recovery or protection mechanisms, the branch node MAY remove itself from the P2MP tree provided that it is not also an egress LSR. Since the faults that severed the various downstream egress nodes from the P2MP tree may be disparate, the branch node MUST report all such errors to its upstream neighbor. The ingress node can then decide to re-compute the path to those particular egress nodes, around the failure point.

Solutions MAY include the facility for transit LSRs and particularly branch nodes to recompute sub-P2MP trees to restore them after failures. In the event of successful repair, error notifications SHOULD NOT be reported to upstream nodes, but the new paths are reported if route recording is in use. Crankback requirements are discussed in Section 4.23.

4.7 Record route of P2MP TE LSP tunnels

Being able to identify the established topology of P2MP TE LSP is very important for various purposes such as management and operation of some local recovery mechanisms like Fast Reroute [FRR]. A network operator uses this information to manage P2MP TE LSPs. Therefore, topology information MUST be collected and updated after P2MP TE LSP establishment and modification process.

The P2MP TE solution MUST support a mechanism which can collect and update P2MP tree topology information after P2MP LSP establishment and modification process. For example, the P2P MPLS TE mechanism of route recording could be extended and used if RSVP-TE was used as the P2MP signaling protocol.
It is RECOMMENDED that the information is collected in a data format by which the sender node can recognize the P2MP tree topology without involving some complicated data calculation process.

The solution MUST support the recording of both outgoing interfaces and node-id.

4.8 Call Admission Control (CAC) and QoS Control mechanism of P2MP TE LSPs

P2MP TE LSPs may share network resource with P2P TE LSPs. Therefore it is important to use CAC and QoS in the same way as P2P TE LSPs for easy and scalable operation.

P2MP TE solutions MUST support both resource sharing and exclusive resource utilization to facilitate co-existence with other LSPs to the same destination(s).

P2MP TE solution MUST be applicable to DiffServ-enabled networks that can provide consistent QoS control in P2MP LSP traffic.

Any solution SHOULD also satisfy the DS-TE requirements [RFC3564] and interoperate smoothly with current P2P DS-TE protocol specifications.

Note that this requirement document does not make any assumption on the type of bandwidth pool used for P2MP TE LSPs which can either be shared with P2P TE LSP or be dedicated for P2MP use.

4.9 Variation of LSP Parameters

Certain parameters (such as priority and bandwidth) are associated with an LSP. The parameters are installed by the signaling exchanges associated with establishing and maintaining the LSP.

Any solution MUST NOT allow for variance of these parameters within a single P2MP LSP. That is:
- No attributes set and signaled by the ingress of a P2MP LSP may be varied by downstream LSRs.
- There MUST be homogeneous QoS from the root to all leaves of a single P2MP LSP.

Variation of parameters may be allowed so long as it applies to the whole LSP from ingress to all egresses.

4.10 Re-optimization of P2MP TE LSPs

The detection of a more optimal path (for example, one with a lower
overall cost) is an example of a situation where P2MP TE LSP re-routing may be required. While re-routing is in progress, an important requirement is avoiding double bandwidth reservation (over the common parts between the old and new LSP) thorough the use of resource sharing.

Make-before-break MUST be supported for a P2MP TE LSP to ensure that there is minimal traffic disruption when the P2MP TE LSP is re-routed.

It is possible to achieve make-before-break that only applies to a sub-P2MP tree without impacting the data on all of the other parts of the P2MP tree.

The solution SHOULD allow for make-before-break re-optimization of any subdivision of the P2MP LSP (S2L sub-tree, S2X sub-LSP, S2L sub-LSP, X2L sub-tree, B2L sub-tree, X2L tree, or B2L tree) with no impact on the rest of the P2MP LSP (no label reallocation, no change in identifiers, etc.).

The solution SHOULD also provide the ability for the ingress LSR to have a strict control on the re-optimization process. The ingress LSR SHOULD be able to limit all re-optimization to be source-initiated.

Where sub-tree re-optimization is allowed by the ingress LSR, such re-optimization MAY be initiated by a downstream LSR that is the root of the sub-tree that is to be re-optimized. Sub-tree re-optimization initiated by a downstream LSR MUST be carried out with the same regard to minimizing the hit on active traffic as was described above for other re-optimization.

4.11 Tree Remerge

It is possible for a single transit LSR to receive multiple signaling messages for the same P2MP LSP but for different sets of destinations. These messages may be received from the same or different upstream nodes and may need to be passed on to the same or different downstream nodes.

This situation may arise as the result of the signaling solution definition or implementation options within the signaling solution. Further, it may happen during make-before-break reoptimization (section 4.10), or as a result of signaling message fragmentation (section 4.5).

It is even possible that it is necessary to construct distinct upstream branches in order to achieve the correct label choices
in certain switching technologies managed by GMPLS (for example, photonic cross-connects where the selection of a particular lambda for the downstream branches is only available on different upstream switches).

The solution MUST support the case where of multiple signaling messages for the same P2MP LSP are received at a single transit LSR and refer to the same upstream interface. In this case the result of the protocol procedures SHOULD be a single data flow on the upstream interface.

The solution SHOULD support the case where multiple signaling messages for the same P2MP LSP are received at a single transit LSR and refer to different upstream interfaces, and where each signaling message results in the use of different downstream interfaces. This case represents data flows that cross at the LSR but which do not merge.

The solution MAY support the case where multiple signaling messages for the same P2MP LSP are received at a single transit LSR and refer to different upstream interfaces, and where the downstream interfaces are shared across the received signaling messages. This case represents the merging of data flows. A solution that supports this case MUST ensure that data is not replicated on the downstream interfaces.

An alternative to supporting this last case is for the signaling protocol to indicate an error such that the merge may be resolved by the upstream LSRs.

4.12 Data Duplication

Data duplication refers to the receipt by any recipient of duplicate instances of the data. In a packet environment this means the receipt of duplicate packets - although this should be a benign (if inefficient) situation, it may be catastrophic in certain existing and deployed applications. In a non-packet environment this means the duplication in time of some part of the signal that may lead to the replication of data or to the scrambling of data.

Data duplication may legitimately arise in various scenarios including re-optimization of active LSPs as described in the previous section, and protection of LSPs. Thus, it is impractical to regulate against data duplication in this document.

Instead, the solution:
- SHOULD limit to transitory conditions only the cases where network bandwidth is wasted by the existence of duplicate
delivery paths.
- MUST limit the cases of delivery of duplicate data to an application to error cases or rare transitory conditions.

4.13 IPv4/IPv6 support

Any P2MP TE solution MUST be equally applicable to IPv4 and IPv6.

4.14 P2MP MPLS Label

A P2MP TE solution MUST support establishment of both P2P and P2MP TE LSPs and MUST NOT impede the operation of P2P TE LSPs within the same network. A P2MP TE solution MUST be specified in such a way that it allows P2MP and P2P TE LSPs to be signaled on the same interface. Labels for P2MP TE LSPs and P2P TE LSPs MAY be assigned from shared or dedicated label space(s). Label space shareability is implementation specific.

4.15 Routing advertisement of P2MP capability

Several high-level requirements have been identified to determine the capabilities of LSRs within a P2MP network. The aim of such information is to facilitate the computation of P2MP trees using TE constraints within a network that contains LSRs that do not all have the same capabilities levels with respect to P2MP signaling and data forwarding.

These capabilities include, but are not limited to:

- the ability of an LSR to support branching.
- the ability of an LSR to act as an egress and a branch for the same LSP.
- the ability of an LSR to support P2MP MPLS-TE signaling.

It is expected that it may be appropriate to gather this information through extensions to TE IGPs (see [RFC3630] and [IS-IS-TE]), but the precise requirements and mechanisms are out of the scope of this document. It is expected that a separate document will cover this requirement.

4.16 Multi-Area/AS LSP

P2MP TE solutions SHOULD support multi-area/AS P2MP TE LSPs.

The precise requirements in support of multi-area/AS P2MP TE LSPs is out of the scope of this document. It is expected that a separate document will cover this requirement.
4.17 Multi-access LANs

P2MP MPLS TE may be used to traverse network segments that are provided by multi-access media such as Ethernet. In these cases, it is also possible that the entry point to the network segment is a branch point of the P2MP LSP.

Two options clearly exist:

- the branch point replicates the data and transmits multiple copies onto the segment
- the branch point sends a single copy of the data to the segment and relies on the exit points to discriminate the reception of the data.

The first option has a significant scaling issue since all replicated data must be sent through the same port and carried on the same segment. Thus, a solution SHOULD provide a mechanism for a branch node to send a single copy of the data onto a multi-access network and reach multiple (adjacent) downstream nodes.

4.18 P2MP MPLS OAM

Management of P2MP LSPs is as important as the management of P2P LSPs.

The MPLS and GMPLS MIB modules will be enhanced to provide P2MP TE LSP management in line with whatever signaling solutions are developed.

In order to facilitate correct management, P2MP TE LSPs MUST have unique identifiers since otherwise it is impossible to determine which LSP is being managed.

OAM facilities will have special demands in P2MP environments especially within the context of tracing the paths and connectivity of P2MP TE LSPs. Further and precise requirements and mechanisms for OAM are out of the scope of this document. It is expected that separate documents will cover these requirements and mechanisms.

4.19 Scalability

Scalability is a key requirement in P2MP MPLS systems. Solutions MUST be designed to scale well with an increase in the number of any of the following:

- the number of recipients
- the number of branch points
Both scalability of performance and operation MUST be considered. Key considerations MUST include:
- the amount of refresh processing associated with maintaining a P2MP TE LSP.
- the amount of protocol state that must be maintained by ingress and transit LSRs along a P2MP tree.
- the number of protocol messages required to set up or tear down a P2MP LSP as a function of the number of egress LSRs.
- the number of protocol messages required to repair a P2MP LSP after failure or perform make-before-break.
- the amount of protocol information transmitted to manage a P2MP TE LSP (i.e. the message size).
- the amount of potential routing extensions.
- the amount of additional control plane processing required in the network to detect whether an add/delete of a new branch is required, and in particular, the amount of processing in steady state when no add/delete is requested.
- the amount of control plane processing required by the ingress, transit and egress LSRs to add/delete a branch LSP to/from an existing P2MP LSP.

It is expected that the applicability of each solution will be evaluated with regards to the aforementioned scalability criteria.

4.19.1 Absolute Limits

In order to achieve the best solution for the problem space it is helpful to clarify the boundaries for P2MP TE LSPs.

- Number of recipients.
  A P2MP TE LSP MUST reduce to similar scaling properties as a P2P LSP when the number of recipients reduces to one.

  It is important to classify the problem as a Traffic Engineering problem. It is anticipated that the initial deployments of P2MP TE LSPs will be limited to a maximum of around a hundred recipients, but that medium term deployments may increase this to several hundred, and that future deployments may require significantly larger numbers.

  An acceptable solution, therefore, is one that scales linearly with the number of recipients.

  Solutions that scale worse than linear (that is, exponential or polynomial) are not acceptable whatever the number of recipients they could support.
- Number of branch points.
  Solutions MUST support all possibilities from one extreme of a single branch point that forks to all leaves on a separate branch, to the greatest number of branch points which is \((n-1)\) for \(n\) recipients. Assumptions MUST NOT be made in the solution regarding which topology is more common, and the solution MUST be designed to ensure scalability in all topologies.

- Dynamics of P2MP tree.
  Recall that the mechanisms for determining which recipients should be added to an LSP, and for adding and removing recipients from that group are out of the scope of this document. Nevertheless, it is useful to understand the expected rates of arrival and departure of recipients since this can impact the selection of solution techniques.
  Again, it must be recalled that this document is limited to Traffic Engineering, and in this model the rate of change of recipients may be expected to be lower than in an IP multicast group.
  Although the absolute number of recipients coming and going is the important element for determining the scalability of a solution, it may be noted that a percentage may be a more comprehensible measure but that this is not as significant for LSPs with a small number of recipients.
  A working figure for an established P2MP TE LSP is less than 10% churn per day. That is, a relatively slow rate of churn.
  We could say that a P2MP LSP would be shared by multiple multicast groups and dynamics of P2MP LSP would be relatively small.
  Considering applicability that P2MP LSP to use L2 multi-access path technology, we can consider stable P2MP L2 path even when we transfer IP multicast traffic over the path.

  Solutions MUST optimize around such relatively low rates of change and are NOT REQUIRED to optimize for significantly higher rates of change.

- Rate of change within the network.
  It is also important to understand the scaling with regard to changes within the network. That is, one of the features of a P2MP TE LSP is that it can be robust or protected against network failures, and can be re-optimized to take advantage of newly available network resources.

  It is more important that a solution be optimized for scaling with respect to recovery and re-optimization of the LSP, than for change in the recipients, because P2MP is used as a TE tool.
  The solution MUST follow this distinction.
4.20 Backwards Compatibility

It SHOULD be an aim of any P2MP solution to offer as much backward compatibility as possible. An ideal which is probably impossible to achieve would be to offer P2MF services across legacy MPLS networks without any change to any LSR in the network.

If this ideal cannot be achieved, the aim SHOULD be to use legacy nodes as both transit non-branch LSRs and egress LSRs.

It is a further requirement for the solution that any LSR that implements the solution SHALL NOT be prohibited by that act from supporting P2P TE LSPs using existing signaling mechanisms. That is, unless administratively prohibited, P2P TE LSPs MUST be supported through a P2MP network.

Also, it is a requirement that P2MP TE LSPs MUST be able to co-exist with IP unicast and IP multicast networks.

4.21 GMPLS

The requirement for P2MP services for non-packet switch interfaces is similar to that for PSC interfaces. Therefore, it is a requirement that reasonable attempts must be made to make all the features/mechanisms (and protocol extensions) that will be defined to provide MPLS P2MP TE LSPs equally applicable to P2MP PSC and non-PSC TE-LSPs. If the requirements of non-PSC networks over-complicate the PSC solution a decision may be taken to separate the solutions. This decision must be taken in full consultation with the MPLS and CCAMP working groups.

Solutions for MPLS P2MP TE-LSPs when applied to GMPLS P2MP PSC or non-PSC TE-LSPs MUST be backward and forward compatible with the other features of GMPLS including:

- control and data plane separation (IF_ID RSVP_HOP and IF_ID ERROR_SPEC),
- full support of numbered and unnumbered TE links (see [RFC 3477] and [GMPLS-ROUTE]),
- use of the GENERALIZED_LABEL_REQUEST, the GENERALIZED_LABEL (C-Type 2 and 3), the SUGGESTED_LABEL and the RECOVERY_LABEL, in conjunction with the LABEL_SET and the ACCEPTABLE_LABEL_SET object,
- processing of the ADMIN_STATUS object,
- processing of the PROTECTION object,
- support of Explicit Label Control,
- processing of the Path_State_Removed Flag,
- handling of Graceful Deletion procedures,
- E2E and Segment Recovery procedures.
- support of Graceful Restart.
In addition, since non-PSC TE-LSPs may have to be processed in environments where the "P2MP capability" could be limited, specific constraints may also apply during the P2MP TE Path computation. Being technology specific, these constraints are outside the scope of this document. However, technology independent constraints (i.e. constraints that are applicable independently of the LSP class) SHOULD be allowed during P2MP TE LSP message processing. It has to be emphasized that path computation and management techniques shall be as close as possible to those being used for PSC P2P TE LSPs and P2MP TE LSPs.

4.22 Requirements for Hierarchical P2MP TE LSPs

[LSP-HIER] defines concepts and procedures for P2P LSP hierarchy.

The P2MP MPLS-TE solution SHOULD support the concept of region and region hierarchy (PSC1<PSC2<PSC3<PSC4<L2SC<TDM<LSC<FSC).

Particularly it SHOULD allow a Region i P2MP TE LSP to be nested into a region j P2MP TE LSP or multiple region j P2P TE LSPs, providing that i<j.

The precise requirements and mechanisms for this function are out of the scope of this document. It is expected that a separate document will cover these requirements.

4.23 P2MP Crankback routing

P2MP solutions SHOULD support crankback requirements as defined in [CRANKBACK]. In particular, they SHOULD provide sufficient information to a branch LSR from downstream LSRs to allow the branch LSR to re-route a sub-tree around any failures or problems in the network.

5. Security Considerations

This requirements document does not define any protocol extensions and does not, therefore, make any changes to any security models.

It should be noted that P2MP signaling mechanisms built on P2P RSVP-TE signaling are likely to inherit all of the security techniques and problems associated with RSVP-TE. These problems may be exacerbated in P2MP situations where security relationships may need to maintained between an ingress and multiple egresses. Such issues are similar to security issues for IP multicast.

It is a requirement that documents offering solutions for P2MP LSPs MUST have detailed security sections.
6. IANA Considerations

This informational draft does not introduce any new encodings or code points. It requires no action from IANA.

7. Acknowledgements

The authors would like to thank George Swallow, Ichiro Inoue, Dean Cheng, Lou Berger and Eric Rosen for their review and suggestions.

Thanks to Loa Andersson for his help resolving the final issues in this document.

8. References

8.1 Normative References


8.2 Informational References


9. Editor’s Address

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

10. Authors’ Addresses

Dimitri Papadimitriou  
Alcatel  
Francis Wellensplein 1,  
B-2018 Antwerpen, Belgium  
Phone: +32 3 240 8491  
Email: dimitri.papadimitriou@alcatel.be

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

Yuji Kamite  
NTT Communications Corporation  
Tokyo Opera City Tower  
3-20-2 Nishi Shinjuku, Shinjuku-ku, Tokyo 163-1421, Japan  
Email: y.kamite@ntt.com

Rahul Aggarwal  
Juniper Networks  
1194 North Mathilda Ave.  
Sunnyvale, CA 94089  
Email: rahul@juniper.net
11. Intellectual Property Consideration

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.

12. Full Copyright Statement

Copyright (C) The Internet Society (2005). 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.