Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: April 20, 2011
C. Jacquenet
France Telecom
Q. Zhao
Huawei Technology
October 17, 2010

Receiver Driven P2MP TE Requirements
draft-jacquenet-mpls-rd-p2mp-te-requirements-00.txt

Abstract

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

There is no intent to specify solution-specific details or application-specific requirements in this document.

The requirements presented in this document not only apply to packet-switched networks under the control of MPLS protocols, 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

Multiparty multimedia applications are getting greater attention in the telecom world. Such applications are QoS-demanding and can therefore benefit from the activation of MPLS traffic engineering capabilities that lead to the dynamic computation and establishment of MPLS LSPs whose characteristics comply with application-specific QoS requirements. P2MP-RE-REQ[RFC4461] specifies the signalling requirements to setup multipoint LSPs from sender to receiver. P2MP-TE [RFC4875] defines the procedure to setup multipoint LSPs from sender to receiver. This document presents a set of requirements specific for the establishment and maintenance of receiver-driven P2MP-TE LSPs.

1.1. Motivation

IP multicast distribution trees are receiver-initiated and dynamic by nature. IP multicast-enabled applications are also bandwidth savvy, especially in the area of residential IPTV services, where several hundreds of thousands of IPTV receivers need to be served with the appropriate level of quality. Current source-driven P2MP LSP establishment assumes a prior knowledge of receiver(s) location for the sake of the P2MP LSP tree structure's forwarding efficiency. But the receiver's location information is not available a priori for the root MPLS router to compute and establish the relevant P2MP tree structure.

Receiver-driven MPLS P2MP tree structure do not require sender to maintain/discover receiver information a priori, and their design is meant to better accommodate the receiver-specific QoS conditions, such as network access capabilities.

1.2. Terminology

With the receiver-driven concept, we have re-defined the following terms:

- Sender: Sender refers to the Originator (and hence) Sender of the content/payload. As in [RFC2205].

- Receiver: Receiver refers to the Receiver of the content/payload. As in [RFC2205].

- Upstream: The direction of flow from content Receiver toward content Sender. As defined in [RFC2205].

- Downstream: The direction of flow from content Sender toward content Receiver. As defined in [RFC2205].

o Path-Sender: The sender of the RSVP PATH message, with NO correlation to direction of content/payload flow. All other control message flows discussed in this document also assume no correlation with the direction of content flows.

o Path-Receiver: The receiver of the RSVP PATH message, with NO correlation to direction of content/payload flow.

o Path-Initiator: The Path-Sender that originated the RSVP PATH message. This being different from Path-Sender because an intermediate node can be a Path-Sender, but different from the node that created and initiated the RSVP PATH message, the Path-Initiator.

o Path-Terminator: The Path-Receiver that does NOT propagate the Path message. This being different from Path-Receiver because an intermediate node can be a Path-Receiver.

1.3. What is not covered in this document

This document does not specify any requirements for the following functions.

o Discovery of root for a P2MP LSP.

o Algorithms for computing P2MP distribution trees.

o Hierarchical P2MP LSPs.

o OAM for P2MP LSPs.

o Inter-area and inter-AS P2MP TE LSPs.

o Applicability of P2MP MPLS TE LSPs to service scenarios.

o Application-specific requirements.

o Multipoint-to-point LSPs.

o Multipoint-to-multipoint LSPs.

o Routing protocols.

o Construction of the traffic engineering database.

o Distribution of the information used to construct the traffic engineering database.

2. Basic Requirements

[RFC4461] specifies requirements for P2MP traffic engineering over MPLS. It describes sender driven P2MP traffic engineering in detail. Those definitions and requirements which are equally applicable to receiver driven traffic engineering in a point-to-multipoint service environment, they are not repeated here.

Following are key requirements that are specific to the receiver-driven paradigm:

1. In order to scale well with a large number of leaves it is RECOMMENDED to follow a leaf-initiated P2MP LSP setup approach. For that purpose, leaves will have to be aware of the P2MP LSP identifier. The ways a Leaf LSR discovers P2MP LSPs identifiers rely on the applications that will use P2MP LSPs, and are out of the scope of this document.

2. The RD P2MP TE mechanism MUST allow the dynamic addition and removal of leaves to and from a P2MP LSP, without any restriction (provided there is network connectivity). It is RECOMMENDED that these operations be leaf-initiated.

3. These operations for the dynamic addition and removal of leaves to and from a RD P2MP LSP MUST not impact the data transfer (packet loss, duplication, delay) towards other leaves.

4. It is RECOMMENDED that these operations do not cause any additional processing except on the path from the added/removed Leaf LSR to the Branch LSR.

5. Receivers MUST trigger the grafting of a new leaf of a given RD P2MP tree structure. This assumes the processing of an IGMP or MLD Report message by the leaf LSR router directly connected to the receiver, so that it can behave as a Path_Initiator and send the RSVP_PATH message accordingly.

6. The destination of a L2S (Leaf to Source) sub-path SHOULD be the ingress router of the RD P2MP multicast tree and this ingress router is the entry point where the contents transported by the RD P2MP TE LSP enter from

7. RSVP P2MP PATH messages MUST traverse along the path from receiver to senders.

8. RSVP P2MP RESV messages MUST traverse along the path from the sender to the receiver if the leaf is the first leaf of the RD P2MP TE LSP. Otherwise it will be forwarded by the PATH Terminator towards the PATH Initiator, which, in that case, is

an intermediate node of the RD P2MP tree structure.

9. A node receiving a RSVP RESV message SHOULD be interpreted as successful resource reservation from the upstream node.

10. Label allocation on incoming interface MUST be done prior to sending RSVP PATH messages upstream.

11. A node receiving a RSVP PATH message MUST first decide if this RSVP PATH message will make itself a branch LSR or not. In the case that it will become a transit LSR because of this PATH message, then it will allocate required resources on the interface through which the RSVP PATH message is received, before sending the RSVP PATH message upstream. So that the upstream node can send traffic soon after successfully reserving resources on the downstream link, on which the RSVP PATH message SHOULD be received. In the case that the node is a branch or transit node already before it receives the PATH message, then it will allocate required resources on the interface through which the RSVP PATH message is received, and send the RESV message to the node which sends the PATH message without sending the PATH message to the upstream node.

3. Detailed Requirements

3.1. RD P2MP LSP

The RD P2MP TE extensions MUST be applicable to the signaling of LSPs for different switching types. For example, it MUST be possible to signal a P2MP TE LSP in any switching medium, whether it is packet or non-packet based (including frame, cell, TDM, lambda, etc.).

As with P2P MPLS technology [RFC3031], traffic is classified with a FEC in this extension. All packets that belong to a particular FEC and that travel from a particular node MUST follow the same RD 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 P2ID) that is constant for the whole LSP regardless of the number of branches and/or leaves.

3.2. P2MP TE LSP Establishment, Teardown, and Modification Mechanisms

The P2MP TE solution MUST support the dynamic establishment, maintenance, and teardown of RD P2MP TE LSPs in a manner that is at least scalable in a linear way. 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 RD P2MP TE LSP establishment and teardown mechanisms, the solution SHOULD support a partial RD P2MP tree modification mechanism, for the sake of path computation optimization.

For the purpose of adding sub-P2MP TE LSPs to an existing RD 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 cause no additional processing in nodes that are not along the path from the grafting or pruning node to the upstream branch node. Moreover, both grafting and pruning operations MUST NOT disrupt traffic that is being forwarded along the P2MP tree.

There is no assumption that the explicitly routed P2MP LSP remains on an optimal path after several grafts and prunes have occurred. In this context, scalable refers to the signaling process for the RD P2MP TE LSP. The TE nature of the LSP allows that re-optimization may take place from time to time to restore the optimality of the LSP.

3.3. Re-Optimization of RD 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 to avoid double bandwidth reservation (over the common parts between the old and new LSP) through the use of resource sharing.

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

Make-before-break that only applies to a sub-P2MP tree without impacting the data on all the other parts of the P2MP tree MUST be supported.

The solution SHOULD allow for make-before-break re-optimization of any subdivision of the P2MP LSP, it SHOULD do so by not having any signaling affect the stability of the rest of the P2MP LSP, and without affecting the ability of the management plane to manage the LSP.

The solution SHOULD also provide the ability for any downstream LSR to have control over the re-optimization process.

Where sub-LSP 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-LSP that is to be re-optimized. Sub-LSP re-optimization initiated by a downstream LSR MUST be carried out with the same regard to minimizing the impact on active traffic as was described above for other re-optimization.

3.4. Merging of Tree Branches

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

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 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 upstream interfaces. This case represents data flows that cross at the LSR but that 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 downstream interfaces, and where the upstream 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 upstream

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.

3.5. Support for LAN interfaces

RD 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 LSR of the P2MP LSP.

To avoid all replicated data are sent through the same port and carried on the same segment, a solution MUST provide a mechanism for a branch LSR to send a single copy of the data onto a multi-access network to reach multiple (adjacent) downstream nodes.

The RD P2MP TE mechanism SHOULD provide a way for a Branch LSR to send a single copy of the data onto an Ethernet LAN interface and reach multiple adjacent downstream nodes. This requires that the same label be negotiated with all downstream LSRs for the LSP.

When there are several candidate upstream LSRs on a LAN interface, the RD P2MP TE mechanism SHOULD provide a way for all downstream LSRs of a given P2MP LSP to select the same upstream LSR, so as to avoid traffic replication. In addition, the RD P2MP TE mechanism SHOULD allow for an efficient balancing of a set of P2MP LSPs among a set of candidate upstream LSRs on a LAN interface.

3.6. P2MP MPLS Label

A RD P2MP TE solution MUST allow the continued use of existing techniques to establish P2P and legacy P2MP LSPs (TE and otherwise) within the same network, and MUST allow the coexistence of P2P and legacy P2MP LSPs within the same network as RD P2MP TE LSPs.

A RD P2MP TE solution MUST be specified in such a way that it allows legacy P2MP and P2P TE LSPs to be signaled on the same interface.

3.7. Advertisement of RD P2MP Capability

To facilitate the computation of RD P2MP trees using TE constraints within a network that contains LSRs that do not all have the same capability levels with respect to RD P2MP signaling and data forwarding, the capability of an LSR to support the RD signalling and forwarding SHOULD be advertised to its neighbor LSRs.

3.8. 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:

- o the number of recipients
- o the number of egress LSRs
- o the number of branch LSRs
- o the number of branches

Both scalability of control plane operation (setup, maintenance, modification, and teardown) MUST be considered.

Key considerations MUST include:

- o the amount of refresh processing associated with maintaining a P2MP TE LSP.
- o the amount of protocol state that must be maintained by ingress and transit LSRs along a P2MP tree.
- o the number of protocol messages required to set up or tear down a RD P2MP LSP as a function of the number of egress LSRs.
- o the number of protocol messages required to repair a P2MP LSP after failure or to perform make-before-break.
- o the amount of protocol information transmitted to manage a P2MP TE LSP (i.e., the message size).
- o the amount of additional data distributed in potential routing extensions.
- o 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
- o 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.
- o It is expected that the applicability of each solution will be evaluated with regards to the aforementioned scalability criteria.

In order to scale well with an increase in the number of leaves, it is RECOMMENDED that the size of a P2MP LSP state on a LSR, for one particular LSP, depend only on the number of adjacent LSRs on the LSP.

3.9. 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:

- o No attributes set and signaled by the first leaf LSR of a P2MP LSP may be altered by other downstream LSRs or upstream LSRs.

- o There MUST be homogeneous QoS from the root to all leaves of a single RD P2MP LSP.

- o Changing the parameters for the whole tree MAY be supported, but the change MUST apply to the whole tree from ingress LSR to all egress LSRs.

4. Backwards Compatibility

It SHOULD be an aim of any RD RSVP-TE P2MP solution to offer as much backward compatibility as possible.

Also, it is a requirement that RD P2MP TE LSPs MUST be able to coexist with legacy P2MP TE multicast networks.

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

5. Acknowledgements

We would like to thank authors of [RFC4461] and the authors of draft-ietf-mpls-mp-ldp-reqs from which some text of this document has been inspired.

6. IANA Considerations

This memo includes no request to IANA.

7. Security Considerations

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

It is a requirement that any P2MP solution developed to meet some or all of the requirements expressed in this document MUST include mechanisms to enable the secure establishment and management of P2MP MPLS-TE LSPs. This includes, but is not limited to:

- o A receiver MUST be authenticated before it is allowed to establish P2MP LSP with source, in addition to hop-by-hop security issues identified by in RFC 3209 and RFC 4206.

- o Mechanisms to ensure that the ingress LSR of a P2MP LSP is identified;

- o Mechanisms to ensure that communicating signaling entities can verify each other's identities;

- o Mechanisms to ensure that control plane messages are protected against spoofing and tampering;

- o Mechanisms to ensure that unauthorized leaves or branches are not added to the P2MP LSP; and

- o Mechanisms to protect signaling messages from snooping.

- o Note that RD P2MP signaling mechanisms built on P2P RSVP-TE signaling are likely to inherit all the security techniques and problems associated with RSVP-TE. These problems may be exacerbated in P2MP situations where security relationships may need to be maintained between an ingress LSR and multiple egress LSRs. Such issues are similar to security issues for IP multicast.

- o It is a requirement that documents offering solutions for RD P2MP LSPs MUST have detailed security sections.

8. References

8.1. Normative References

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

[2] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J.

McManus, "Requirements for Traffic Engineering Over MPLS", RFC 2702, September 1999.

[3] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

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

[5] Yasukawa, S., "Signaling Requirements for Point-to-Multipoint Traffic-Engineered MPLS Label Switched Paths (LSPs)", RFC 4461, April 2006.

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

[7] Farrel, A., Papadimitriou, D., Vasseur, J., and A. Ayyangar, "Encoding of Attributes for Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Establishment Using Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 4420, February 2006.

[8] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.

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

[10] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

[11] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[12] Aggarwal, R. and J. Roux, "MPLS Upstream Label Assignment for RSVP-TE", draft-ietf-mpls-rsvp-upstream-05 (work in progress), March 2010.

[13] Roux, J., "Requirements for Point-To-Multipoint Extensions to the Label Distribution Protocol", draft-ietf-mpls-mp-ldp-reqs-04 (work in progress), February 2008.

8.2. Informative References

[14] Andersson, L. and G. Swallow, "The Multiprotocol Label Switching (MPLS) Working Group decision on MPLS signaling protocols", RFC 3468, February 2003.

[15] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[16] Le Faucheur, F. and W. Lai, "Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering", RFC 3564, July 2003.

[17] Berger, L., Takacs, A., Caviglia, D., Fedyk, D., and J. Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label Switched Paths (LSPs)", RFC 5467, March 2009.

Authors' Addresses

Christian Jacquenet
France Telecom
4 rue du Clos Courtel
35512 Cesson-Sevigne
France

Phone: +33 2 99 12 43 82
Email: christian.jacquenet@orange-ftgroup.com

Quintin Zhao
Huawei Technology
125 Nagog Park
Acton, MA 01919
US

Phone:
Email: qzhao@huawei.com