Internet Engineering Task Force Internet-Draft Intended status: Informational Expires: January 11, 2014 C. Jacquenet France Telecom Orange Q. Zhao Huawei Technology B. Zhang Telus Communications E. Metz KPN July 10, 2013

Requirements for Receiver-Driven Traffic Engineered Point-to-Multi-Point (P2MP) MPLS Tree Structures draft-jacquenet-mpls-rd-p2mp-te-requirements-03.txt

Abstract

This document presents a set of requirements for the establishment and maintenance of Receiver-Driven Point-to-Multipoint (RD-P2MP) and Multipoint-to-Multipoint (RD-MP2MP) Traffic-Engineered (TE) Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of $\underline{BCP 78}$ and $\underline{BCP 79}$.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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

This Internet-Draft will expire on January 11, 2014.

Copyright Notice

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

This document is subject to <u>BCP 78</u> and the IETF Trust's Legal Provisions Relating to IETF Documents (<u>http://trustee.ietf.org/license-info</u>) in effect on the date of publication of this document. Please review these documents

Jacquenet, et al. Expires January 11, 2014 [Page 1]

July 2013

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

<u>1</u> . Int	roduction						2
<u>1.1</u> .	Rationale						<u>2</u>
1.2.	Typical Use Case: Interconnecting PIM Islands	Т	Thro	bugh	А		
	MPLS Backbone						<u>3</u>
<u>1.3</u> .	Scope and Assumptions						<u>4</u>
<u>2</u> . Ter	minology						<u>4</u>
<u>3</u> . Req	uirements						<u>5</u>
<u>3.1</u> .	Operation						<u>5</u>
<u>3.2</u> .	Forwarding						<u>6</u>
<u>3.3</u> .	Signaling						<u>7</u>
<u>3.4</u> .	Robustness						7
<u>3.5</u> .	Performance and Scalability						<u>8</u>
<u>3.6</u> .	Management						<u>10</u>
<u>3.7</u> .	Backward Compatibility						<u>10</u>
<u>4</u> . Ack	nowledgements						<u>10</u>
<u>5</u> . IAN	A Considerations						<u>11</u>
<u>6</u> . Sec	urity Considerations						<u>11</u>
<u>7</u> . Ref	erences	•				•	<u>11</u>
<u>7.1</u> .	Normative References						<u>12</u>
<u>7.2</u> .	Informative References						<u>13</u>
Authors	'Addresses						<u>13</u>

1. Introduction

<u>1.1</u>. Rationale

The dramatic growth of multicast-based IP services such as live TV broadcasting raises new challenges for network operators. The sole use of IP multicast becomes challenging, given the QoS-demanding nature of the applications, especially in light of the deployment of High Definition or 3D TV.

The specification of traffic-engineered, MPLS-based, Point-to-Multi-Point (P2MP) tree structures [6] is meant to address both quality and robustness issues, but failed to be massively adopted and deployed by operators so far, mostly because the current standard assumes a priori knowledge of all the routers involved in the establishment and the maintenance of the tree structure, at the cost of extra management complexity. Current technological state-of-the-art shows that most of implementations assume the static configuration of the basic information that will be used by routers to build a traffic-engineered P2MP LSP path a la [6]. Scalability issues may also be raised by the current standard as core routers are likely to maintain a number of RSVP states that will grow with the number of leaf routers.

The receiver-initiated paradigm of IP multicast is also a true incentive for faster convergence in MPLS environments, yielding an improved computation and establishment of traffic-engineered P2MP MPLS tree structures.

In addition, the inability to take into account the network access capabilities of the receivers, so that the computation of the tree structure can dynamically adapt to such access conditions becomes even more challenging with the ever-growing number of multicast channels that need to be delivered to the receivers.

Thus, we believe that the combination of the dynamics of receiverinitiated IP multicast distribution trees with the hard QoS guarantees brought by a MPLS-based traffic engineered tree computation scheme is promising, for the sake of optimized convergence and facilitated (if not automated) hard-guaranteed service design and operation.

Taking the best of breed between PIM-computed [<u>16</u>], receiver-driven multicast distribution trees and MPLS traffic engineering capabilities can indeed significantly improve

<u>1.2</u>. Typical Use Case: Interconnecting PIM Islands Through A MPLS Backbone

Taking the best of breed between PIM-computed, receiver-driven multicast distribution trees and MPLS traffic engineering capabilities can indeed significantly improve the level of quality associated to the delivery of multicast services.

One typical example would be a network design where PIM "islands" (for example, PIM-enabled access infratsructures that connect multicast receivers, as well as the IPTV head end that connects the multicast source) connect to a MPLS backbone infrastructure. Multicast traffic is forwarded along the RD-P2MP LSP paths without soliciting an additional routing protocol such as BGP ([14], [15]) to learn the multicast routes and exchange the PIM multicast states of each PIM island, so that P2MP LSP paths can be established accordingly.

The Receiver-Driven approach should solely rely upon the PIM protocol and the use of a PIM/MPLS inter-working function at the border between the PIM islands and the MPLS backbone infrastructure, so that no PIM state has to be maintained by the MPLS routers of the backbone.

<u>1.3</u>. Scope and Assumptions

This document details the requirements for the dynamic establishment and maintenance of receiver-driven, traffic-engineered Point-to-Multi-Point (RD-P2MP) tree structures. As such, considerations about Label Distribution Protocol (LDP) extensions for the computation of P2MP tree structures [10] are out of the scope of this draft.

Likewise, this document exclusively focuses on traffic-engineered P2MP MPLS tree structures. As such, traffic-engineered Multi-Pointto-Multi-Point (MP2MP) MPLS tree structures are not taken into consideration, mostly because it is believed that the primary applicability of MPLS traffic engineering capabilities for multicastenabled services remains IPTV services that assume a 1:N group communication scheme that is typical of Source-Specific Multicast designs. Nevertheless, the requirements detailed in this document are also valid for receiver-driven, traffic-engineered MP2MP tree structures.

2. Terminology

The following terms are used in this document:

- o Sender: The source of the content, as in [9].
- o Receiver: The Receiver of the content multicast by the source, as in [9].
- o Upstream: The direction of a given traffic from a Receiver towards a Sender, as defined in $[\underline{9}]$.
- o Downstream: The direction of multicast traffic from a Sender towards a Receiver, as defined in $[\underline{9}]$.
- Path-Sender: The sender of a RSVP_PATH message, with NO correlation to the directionality of the corresponding multicast traffic.
- Path-Receiver: The receiver of a RSVP_PATH message, with N0 correlation to the directionality of the corresponding multicast traffic.

- o Path-Initiator: The Path-Sender that originated a RSVP PATH message. A Path-Initiator is different from a Path-Sender in that an intermediate node (a node that is part of the receiver-driven P2MP tree structure) can be a Path-Sender.
- o Path-Terminator: The Path-Receiver that does NOT propagate a RSVP PATH message. A Path-Terminator is different from the Path-Receiver in that an intermediate node (a node that is part of the receiver-driven P2MP tree structure) can be a Path-Receiver.
- o Receiver-Driven P2MP Label Switched Path (RD P2MP LSP): A trafficengineered P2MP MPLS Label Switched Path that is dynamically computed and established at the initiative of the receivers.

3. Requirements

[5] specifies requirements for P2MP traffic-engineered MPLS Label Switched Paths. These requirements are equally applicable to Receiver-Driven traffic-engineered MPLS Label Switched Paths. This section details the additional requirements raised by the dynamic computation and establishment of RD P2MP LSPs.

3.1. Operation

- REQ-1 Grafting and pruning of branches of any given RD P2MP tree structure MUST be dynamic and receiver-initiated.
- REQ-2 Leaves of RD P2MP LSP paths MUST be aware of the corresponding P2MP LSP identifiers for tree computation and maintenance purposes.
- REQ-3 A leaf router MUST initiate RSVP PATH messages that will be sent towards the root router. RSVP PATH messages are triggered by typical signaling subcription procedures originated by receivers, such as IGMP/MLD Report messages that may be directly processed by the leaf routers of a given RD P2MP LSP.
- REQ-4 A node that receives a RSVP_PATH message MUST first decide if this message will make itself a branch Label Switch Router (LSR) or not. In the case that it will become a transit LSR because of this RSVP PATH message, the router will allocate required resources on the interface through which the RSVP PATH message is received, before forwarding it upstream (for the sake of multicast traffic delivery efficiency).
- REQ-5 In case the node that received the RSVP PATH message is already a branch or transit node for the multicast content associated to the said RSVP PATH message, the node MUST allocate

required resources (if available) on the interface through which the RSVP PATH message is received. This node MUST NOT send the RSVP PATH message upstream.

- REQ-6 Upon receipt of a RSVP PATH message, a Path-Terminator MUST send back a RSVP RESV message that MUST be forwarded along the RD P2MP LSP to the Path-Sender.
- REQ-7 A node receiving a RSVP RESV message SHOULD interpret it as a successful resource reservation from the upstream node for the establishment of the RD P2MP LSP.
- REQ-8 The termination of a L2S (Leaf-to-Source) sub-path that belongs to a given RD P2MP LSP MUST be one of the ingress routers where multicast data sent by a source enter the said RD P2MP LSP.
- REQ-9 Label allocation MUST be done prior to sending RSVP PATH messages upstream.
- REQ-10 To facilitate the computation of RD P2MP LSP paths within a network whose LSRs do not all support the same capabilities with respect to RD P2MP LSP signaling and data forwarding, the capability of a given LSR to support the RSVP-TE signaling and forwarding features for RD P2MP LSP path computation purposes MUST be advertised to its neighbor LSRs.

3.2. Forwarding

- REQ-11 Just like typical P2P MPLS [3], any given multicast traffic characterized by a multicast group address is associated with a FEC (Forwarding Equivalence Class). All packets that belong to a particular FEC MUST be forwarded along the corresponding RD P2MP LSP.
- REQ-12 RD P2MPLSPs MAY be deployed over multi-access media such as Ethernet. In these environments, it is possible that the entry point to the network segment is a branch LSR of the RD P2MPLSP. To avoid all replicated data are sent through the same port and carried on the same segment, a mechanism SHOULD be supported by the said branch LSR so as to send a single copy of the multicast data onto the multi-access network to reach several downstream nodes
- REQ-13 The RD P2MP LSP computation scheme SHOULD provide a means for a Branch LSR to send a single copy of the multicast data through an Ethernet LAN interface to reach several downstream nodes.

- REQ-14 As a consequence of the previous requirement, the same label MUST be negotiated with all downstream LSRs for any given RD P2MP LSP.
- REQ-15 When there are several candidate upstream LSRs conected to a given LAN segment, the RD P2MP LSP path computation scheme SHOULD provide a means for all downstream LSRs to select the same upstream LSR, so as to avoid traffic replication.
- REQ-16 The RD P2MP LSP path computation scheme SHOULD allow for an efficient balancing of a set of P2MP LSPs among a set of candidate upstream LSRs connected to a LAN segment.

3.3. Signaling

Certain parameters (such as priority and bandwidth) are associated with an LSP. The parameters are installed by means of the RSVP signalling specific to the establishment and the maintenance of RD P2MP LSPS. As such:

- REQ-17 Downstream or upstream LSRs MUST NOT alter the attributes set and signaled by a leaf router of any given RD P2MP LSP.
- REQ-18 A consistent QoS policy SHOULD be enforced from the root to all leaves of a single P2MP/MP2MP LSP.
- REQ-19 Some leaves of a given tree may yield the enforcement of a different QoS policy, depending on the various access capabilities of the receivers. Still, content will be delivered to these receivers by using the same (core) P2MP/MP2MP tree structure.
- REQ-20 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.

3.4. Robustness

- REQ-21 The detection of a more optimal path (e.g., from a cost standpoint) is an example of a situation where re-routing of the RD P2MP LSP MAY be required. While re-routing is in progress, duplicate bandwidth reservation over the common parts between the old and new RD P2MP LSP MUST be avoided.
- REQ-22 RD P2MP LSP path computation, establishment and maintenance MUST support a Make-Before-Break procedure to ensure that there is minimal traffic disruption during re-routing operations.

- REQ-23 Scope of Make-Before-Break procedures MUST be restricted to the relevant sub-LSP portion of any given RD P2MP LSP, for the sake of resource optimization and overall service performance.
- REQ-24 The support of a Make-Before-Break procedure MUST include reoptimization facilities for any impacted sub-LSP portion of a given RD P2MP LSP.
- REQ-25 Any Make-Before-Break operation MUST not impact the rest of the RD P2MP LSP, from both a signaling and operational standpoints.
- REQ-26 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 to be re-optimized. Sub-LSP reoptimization initiated by a downstream LSR MUST be carried out without dramatically impacting the forwarding of multicast traffic along the corresponding RD P2MP LSP.
- REQ-27 Any downstream node located on the sub-LSP that is being reoptimized SHOULD have control on the re-optimization procedure.
- REQ-28 Overtime, a given optimized path may become sub-optimized. Path re-computation capabilities SHOULD be supported and they could be triggered by updated IGP cost metrics, time interval or a combination thereof.
- REQ-29 Traffic load balancing capabilities SHOULD be supported by the RD P2MP LSP path computation scheme, to enforce least-fillbased load sharing computation strategies.

<u>3.5</u>. Performance and Scalability

- REQ-30 The RD P2MP LSP computation, establishment and maintenance MUST be scalable: The switching performances of the routers that are part of a RD P2MP LSP MUST NOT be affected during RD P2MP LSP path computation, establishment and operation.
- REQ-31 Likewise, scalability of RD P2MP LSP path design and operation MUST NOT be jeopardized by the number of receivers, nor by the number of RD P2MP LSPs maintained at any given time by the routers of the network. in particular, RD P2MP LSP path computation, establishment and operation SHOULD accommodate the need to deliver multicast traffic to several millions of receivers, without any perceived overall service performance degradtion, including the preservation of customer's Quality of Experience.

- REQ-32 Dynamic grafting and pruning operations pertaining to any given RD P2MP LSP MUST NOT affect multicast traffic forwarding efficiency, from a packet loss and one-way transit delay perspectives, among other possible Quality of Service metrics.
- REQ-33 Dynamic grafting and pruning operations pertaining to any given RD P2MP LSP SHOULD NOT infer any other additional processing than the processing specific to the portion of the RD P2MP LSP from the added/removed leaf LSR to the corresponding branch LSR.
- REQ-34 Dynamic grafting and pruning operations pertaining to any given RD P2MP LSP MUST NOT disrupt the forwarding of multicast traffic along the RD P2MP LSP at any given time.
- REQ-35 In order to scale to a large number of branches, RD P2MP LSPs SHOULD be unambiguously identified by means of P2MP ID identifiers that MUST be persistent for the whole RD P2MP LSP, regardless of the number of branches and leaves.
- REQ-36 In order to accommodate a growing number of leaves, the amount of a P2MP LSP states on a given LSR, for one particular RD P2MPLSP SHOULD only depend on the number of adjacent LSRs that belong to the same RD P2MP LSP.
- REQ-37 RD P2MP LSP performance and scalability assessment SHOULD rely upon the following metrics (and a combination thereof):
- o Number of receivers
- o Number of sources
- o Number of leaf routers
- o Number of branch routers
- o Number of branches
- o Number of RD P2MP LSPs to be deployed over the MPLS network
- REQ-38 Scalability of control plane operation (setup, maintenance, modification, and teardown) MUST be considered. In particular:
- The amount of refresh processing associated with maintaining a RD P2MP LSP.
- The amount of protocol state that must be maintained by ingress and transit LSRs along a RD P2MP LSP.

- 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 RD P2MP LSP after failure or to perform make-before-break.
- The amount of protocol information transmitted to manage a RD P2MP LSP (i.e., the amount of management traffic as a function of the global bandwidth resources available).
- o The amount of signaling traffic required for RD P2MP LSP path compution, setup and operation.
- 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.
- The amount of control plane processing required by the ingress, transit, and egress LSRs to add/delete a branch LSP to/from an existing RD P2MP LSP.

3.6. Management

REQ-39 Design and operation of RD P2MP TE LSP paths MUST support management capabilities as per the Specific Management Functional Areas (SMFAs), namely Fault, Configuration, Accounting, Performance and Security management capabilities. In particular, RD P2MP TE LSP-based and Source-to-Leaf (S2L) traffic statistics management MUST be supported.

<u>3.7</u>. Backward Compatibility

- REQ-40 The RD P2MP LSP path computation scheme MUST allow the continued use of existing techniques to establish P2P and legacy P2MP LSPs (Traffic-engineered or not) within the same network, and MUST allow the coexistence of Receiver-Driven P2MP/MP2MP LSPSs with P2P and legacy P2MP LSPs within the same network.
- REQ-41 RD P2MP LSP paths MUST be able to coexist with legacy P2P and P2MP LSPs within the same network.
- REQ-42 Signaling capabilities of the RD P2MP LSP path computation scheme MUST NOT prevent the signaling of "legacy" P2P and P2MP LSP paths.

4. Acknowledgements

We would like to thank authors of [5] and [13] who inspired some of the text of this draft.

5. IANA Considerations

This draft makes no request to IANA.

6. Security Considerations

This document does not define any protocol extensions and does not, therefore, make any changes to any security models. It is a requirement that any RD P2MP LSP design MUST include mechanisms to enable the secure establishment and management. of RD P2MP LSPs. This means in particular that:

- o A receiver MUST be authenticated before it is allowed to trigger the grafting of an additional leaf of a RD P2MP LSP tree structure,
- o Mechanisms to provide some guarantees about the identity of an ingress LSR that belongs to a given RD P2MP LSP SHOULD be supported,
- o Mechanisms to ensure that communicating signaling entities can verify each other's identities SHOULD be supported,
- o Mechanisms to ensure that control plane messages are protected against spoofing and tampering SHOULD be supported,
- o Mechanisms to ensure that unauthorized leaves or branches are not added to any given RD P2MP/ LSP; and mechanisms to protect signaling messages from snooping MUST be supported.

Note that RD P2MP LSP signaling mechanisms built on P2P RSVP-TE signaling and RSVP-TE P2MP signalling are likely to inherit all the security techniques and problems associated with RSVP-TE. These problems may be exacerbated in P2MP situations where security associations may need to be maintained between any given ingress LSR and multiple egress LSRs. Such issues are similar to security issues raised by the IP multicast transmission scheme.

7. References

7.1. Normative References

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", <u>BCP 14</u>, <u>RFC 2119</u>, March 1997.
- [2] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", <u>RFC 2702</u>, September 1999.
- [3] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", <u>RFC 3031</u>, January 2001.
- [4] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", <u>RFC 3209</u>, December 2001.
- [5] Yasukawa, S., "Signaling Requirements for Point-to-Multipoint Traffic-Engineered MPLS Label Switched Paths (LSPs)", <u>RFC 4461</u>, 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)", <u>RFC 4875</u>, 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)", <u>RFC 4420</u>, February 2006.
- [8] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", <u>RFC 4206</u>, October 2005.
- [9] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", <u>RFC 2205</u>, September 1997.
- [10] Wijnands, IJ., Minei, I., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", <u>RFC 6388</u>, November 2011.
- [11] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", <u>RFC 4090</u>, May 2005.

- [12] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", <u>RFC 3471</u>, January 2003.
- [13] Le Roux, JL. and T. Morin, "Requirements for Point-to-Multipoint Extensions to the Label Distribution Protocol", <u>RFC 6348</u>, September 2011.
- [14] Rosen, E. and R. Aggarwal, "Multicast in MPLS/BGP IP VPNs", <u>RFC 6513</u>, February 2012.
- [15] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", <u>RFC 6514</u>, February 2012.
- [16] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", <u>RFC 4601</u>, August 2006.

7.2. Informative References

- [17] Andersson, L. and G. Swallow, "The Multiprotocol Label Switching (MPLS) Working Group decision on MPLS signaling protocols", <u>RFC 3468</u>, February 2003.
- [18] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", <u>RFC 3473</u>, January 2003.
- [19] Le Faucheur, F. and W. Lai, "Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering", <u>RFC 3564</u>, July 2003.
- [20] Berger, L., Takacs, A., Caviglia, D., Fedyk, D., and J. Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label Switched Paths (LSPs)", <u>RFC 5467</u>, March 2009.

Authors' Addresses

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

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

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

Email: qzhao@huawei.com

Boris Zhang Telus Communications 200 Consilium Pl. Floor 15 Toronto, ON M1H 3J3 Canada

Email: Boris.Zhang@telus.com

Eduard Metz KPN Netherlands

Email: eduard.metz@kpn.com