Stream: Internet Engineering Task Force (IETF)
RFC: 9961
Category: Standards Track
Published: April 2026
ISSN: 2070-1721
Authors:

H. Bidgoli, Ed.	Z. Ali	Z. Zhang	A. Budhiraja	D. Voyer
Nokia	*Cisco System*	*Juniper Networks*	*Cisco System*	*Cisco System*

RFC 9961
MPLS Segment Routing Point-to-Multipoint (P2MP) Policy Ping

Abstract

Segment Routing (SR) Point-to-Multipoint (P2MP) Policies are used to define and manage explicit P2MP paths within a network. These policies are typically calculated via a controller-based mechanism and installed via, e.g., a Path Computation Element (PCE). In other cases, these policies can be installed using the Network Configuration Protocol (NETCONF) / YANG or a Command Line Interface (CLI). They are used to steer Multicast traffic along optimized paths from a Root to a set of Leaf routers.

This document defines extensions to ping and traceroute mechanisms for an SR P2MP Policy with MPLS encapsulation to provide Operations, Administration, and Maintenance (OAM) capabilities. The extensions enable operators to verify connectivity, diagnose failures, and troubleshoot forwarding issues within SR P2MP Policy Multicast trees.

By introducing new mechanisms for detecting failures and validating path integrity, this document enhances the operational robustness of P2MP Multicast deployments. Additionally, it ensures that existing MPLS and SR-based OAM tools can be effectively applied to networks utilizing SR P2MP Policies.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9961.

Copyright Notice

Table of Contents

1. Introduction

[RFC9960] explains the concept of the SR P2MP Policy and its candidate paths (CPs). It also explains the concept of how a CP is selected to be the active CP. To enable seamless global optimization, a CP may consist of multiple P2MP tree instances (PTIs), allowing for Make-Before-Break procedures between an active PTI and a newly established, optimized PTI. A PTI is the actual P2MP tunnel set up from the Root to a set of Leaves via transit routers. A PTI is identified on the Root node by the PTI's instance ID.

To ensure reliable network operation, it is essential to verify end-to-end connectivity for both active and backup CPs, as well as all associated PTIs. This document specifies a mechanism for detecting data plane failures within an SR P2MP Policy CP and its associated PTIs, enabling operators to monitor and troubleshoot Multicast path integrity.

This specification applies exclusively to Replication Segments (Replication-SIDs) that use MPLS encapsulation for forwarding and does not cover Segment Routing over IPv6 (SRv6). The mechanisms described herein build upon the concepts established in [RFC6425] for P2MP MPLS OAM. All considerations and limitations described in Section 6 of [RFC6425] apply to this document as well.

1.1. Terminology

The readers of this document should familiarize themselves with the following documents and sections for terminology and details regarding the implementation of the SR P2MP Policy.

[RFC9524], Section 1.1 defines terms specific to an SR Replication segment and also explains the node terminology in a Multicast domain, including the Root node, Leaf node, and Bud node.

[RFC9960], Section 1.1 defines terms and concepts specific to the SR P2MP Policy including the CP and the PTI.

2. Conventions Used in This Document

The key words "**MUST**", "**MUST NOT**", "**REQUIRED**", "**SHALL**", "**SHALL NOT**", "**SHOULD**", "**SHOULD NOT**", "**RECOMMENDED**", "**NOT RECOMMENDED**", "**MAY**", and "**OPTIONAL**" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Motivation

An SR P2MP Policy and its corresponding Replication segments are typically provisioned via a centralized controller or configured using NETCONF/YANG or CLI. The Root and the Leaves are discovered in accordance with [RFC9960], and the Multicast tree is computed from the Root to the Leaves. However, there is no underlay signaling protocol to distribute the SR P2MP Policy from the Root to the Leaf routers. Consequently, when a P2MP tree fails to deliver user traffic, identifying the failure can be challenging without ping and traceroute mechanisms to isolate faults along the tree.

To address this challenge, SR P2MP Policy ping and traceroute can be utilized to detect and localize faults within the P2MP tree and its associated Replication segments, as defined in [RFC9524]. These OAM tools enable periodic ping operations to verify connectivity between the Root and the Leaves. In cases where a ping fails, a traceroute can be initiated to determine the point of failure along the tree. This diagnostic process can be initiated from the node responsible for establishing the SR P2MP Policy, ensuring proactive monitoring and fault detection.

3.1. MPLS SR P2MP Policy Ping and Traceroute

Ping/traceroute packets are forwarded based upon the SR P2MP Policy on a specific CP and its PTI toward the designated Leaf routers. These packets are replicated at the replication point based on the Replication segment forwarding information on the corresponding router.

MPLS packets are processed based on the standard behavior when their Time to Live (TTL) expires or when they reach the egress (Leaf) router. The appropriate response is sent back to the Root node following the procedures outlined in [RFC6425].

3.1.1. Applicability of RFC 6425 to MPLS SR P2MP Policy

The procedures in [RFC6425] define fault detection and isolation mechanisms for P2MP MPLS Label Switched Paths (LSPs) and extend the LSP ping techniques described in [RFC8029] such that they may be applied to P2MP MPLS LSPs, ensuring alignment with existing fault management tools. [RFC6425] emphasizes the reuse of existing LSP ping mechanisms designed for Point-to-Point (P2P) LSPs, adapting them to P2MP MPLS LSPs to facilitate seamless implementation and network operation.

The fault detection procedures specified in [RFC6425] are applicable to all P2MP MPLS protocols, including P2MP RSVP-TE and Multicast LDP and now the SR P2MP Policy. While [RFC6425] highlights specific differences for P2MP RSVP-TE and Multicast LDP, this document introduces considerations unique to SR P2MP Policies, including:

1. Egress Address P2MP Responder sub-TLVs: Multicast LDP, as per Section 3.2.1 of [RFC6425], does not allow for the inclusion of Egress Address P2MP Responder sub-TLVs, as upstream Label Switching Routers (LSRs) lack visibility into downstream Leaf nodes. Similarly, SR P2MP Policies often rely on a PCE for programming transit routers. This is why in the SR P2MP domain, transit routers do not have knowledge of the Leaf nodes. Only the Root node, where the SR P2MP Policy is programmed, has visibility into the Leaf nodes. Consequently, these sub-TLVs **SHOULD NOT** be used when an echo request carries an SR P2MP Policy MPLS CP Forwarding Equivalence Class (FEC). If a node receives the Egress Address P2MP Responder sub-TLVs in an echo request, then it will not respond since it is unaware of whether it lies on the path to the address in the sub-TLV.
2. End of processing for traceroutes: As per Section 4.3.1 of [RFC6425], it is **RECOMMENDED** that traceroute orations provide for a configurable upper limit on TTL values. This is because, for some protocols like Multicast LDP, there may not be an easy way to figure out the end of the

traceroute processing, as the initiating LSR might not always know about all of the Leaf routers. In the case of an SR P2MP Policy, the Root node has visibility of the Leaf nodes; as such, there is a definitive way to estimate the end of processing for a traceroute, and a configurable upper limit on TTL may not be necessary. However, a configurable upper limit on the TTL value is an implementation choice.

3. Identification of the LSP under test: Section 3.1 of [RFC6425] defines distinct identifiers for P2MP RSVP-TE and Multicast LDP when identifying an LSP under test. As each protocol has its own identifier, this document introduces a new Target FEC Stack TLV specific to SR P2MP Policies to uniquely identify their CPs and PTIs. These modifications ensure that SR P2MP Policy OAM mechanisms are properly aligned with existing MPLS ping and traceroute tools while addressing the specific operational characteristics of SR P2MP Policies.

3.1.2. Conformance to Existing Procedures and Additional Considerations

In addition to major differences outlined in the previous section, SR P2MP Policies **SHOULD** follow the common procedures specified in [RFC6425] for P2MP MPLS LSPs. Furthermore, this specification reuses the same destination UDP port as defined in [RFC8029] for consistency with the existing MPLS OAM mechanism.

Implementations **MUST** account for the fact that an SR P2MP Policy may contain multiple CPs, and each CP may consist of multiple PTIs. As such, implementations **SHOULD** support the ability to individually test each CP and its corresponding PTI using ping and traceroute mechanisms. The ping and traceroute packets are forwarded along the specified CP and its PTI, traversing the associated Replication segments. When a downstream node capable of understanding the Replication-SID receives a ping or traceroute packet, it **MUST** process the request and generate a response even if the CP and its PTI are not currently the active path.

3.1.3. Considerations for Interworking with Unicast Paths

As per [RFC9960], there are two ways to build a P2MP tree:

1. P2MP tree with non-adjacent Replication segments
2. P2MP tree with adjacent Replication segments

For the case of adjacent Replication segments, there are no special considerations for the TTL or Hop Limit propagation, and the TTL should be decremented hop by hop as the OAM packet traverses the Replication segments of a P2MP tree.

For the case of non-adjacent Replication segments (for example, two Replication segments that are connected via an SR Policy or similar technology), there are special considerations. In such scenarios, SR P2MP Policy OAM tools should be used to verify the connectivity of the non-adjacent Replication segments that are building the P2MP tree, while the unicast OAM tools should be used to verify the connectivity of the unicast path connecting the two non-adjacent Replication segments. In these scenarios, the Replication-SID should not be exposed or examined in the unicast path. Proper TTL handling to copy the Replication segment TTL to the unicast path can be achieved via the hierarchical MPLS TTL mode being used (e.g., Pipe Mode vs. Uniform Mode) as per [RFC3270]. For the P2MP tree traceroute, the TTL mode **MUST** be set to Pipe Mode on the router that the unicast path starts. This will ensure that the unicast path TTL is set to a large value that allows the traceroute packet to be delivered to the downstream Replication segment. For ping, either the Pipe Mode or the Uniform Mode can be used depending on the implementation. The unicast path failure detection is considered out of scope for this document.

3.2. Packet Format and New TLVs

The packet format used in this specification follows Section 3 of [RFC8029]. However, additional TLVs and sub-TLVs are required to support the new functionality introduced for SR P2MP Policies. These extensions are described in the following sections.

3.2.1. Identifying a P2MP Policy

[RFC8029] defines a standardized mechanism for detecting data plane failures in MPLS LSPs. To correctly identify the Replication segment associated with a given CP and PTI, the echo request message **MUST** include a Target FEC Stack TLV that explicitly specifies the CP and PTI under test.

This document introduces a new sub-TLV, referred to as the SR MPLS P2MP Policy Tree Instance sub-TLV, which is defined as follows:

Sub-Type	Length	Value Field
41	Variable	SR MPLS P2MP Policy Tree Instance

Table 1

Further details regarding the structure and processing of this sub-TLV are provided in subsequent sections.

3.2.1.1. SR MPLS P2MP Policy Tree Instance FEC Stack Sub-TLVs

The SR MPLS P2MP Policy Tree Instance sub-TLV value field follows the format specified in Section 2.3 of [RFC9960]. The structure of this sub-TLV is illustrated in the figure below. It should be noted that this sub-TLV is testing a specific PTI within a specific CP and it is not testing the CP.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Address Family         | Address Length|   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                             Root                              ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Tree-ID                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Instance-ID               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Address Family:

2 octets. IPv4/IPv6 Address Family Numbers as specified in [IANA-AF], indicating the Address Family of the Root. Any other Address Family, except IPv4/IPv6, is not supported by this document.

Address Length: 1 octet. Specifies the length of the Root Address in octets (4 octets for IPv4, and 16 octets for IPv6).

Reserved: **MUST** be set to zero by the sender and should be ignored by the receiver.

Root: Variable length depending on the Address Family field. The Root node of the SR P2MP Policy, as defined in [RFC9960].

Tree-ID: 4 octets. A unique identifier for the P2MP tree, as defined in [RFC9960].

Instance-ID: 2 octets. Identifies the specific Path-Instance, as defined in [RFC9960].

3.3. Limiting the Scope of Response

As specified in Section 3.2 of [RFC6425], four sub-TLVs are used within the P2MP Responder Identifier TLV that is included in the echo request message.

The sub-TLVs for IPv4 and IPv6 egress addresses of the P2MP responder are aligned with Section 3.2.1 of [RFC6425].

The sub-TLVs for IPv4 and IPv6 node addresses of the P2MP responder are aligned with Section 3.2.2 of [RFC6425].

These mechanisms ensure that responses are appropriately scoped to limit unnecessary processing and improve the efficiency of P2MP OAM procedures.

4. IANA Considerations

IANA has assigned a sub-type value for the SR MPLS P2MP Policy Tree Instance sub-TLV in the "Sub-TLVs for TLV Types 1, 16, and 21" registry under the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry group. The sub-type value has been assigned from the Standards Action range of 0-16383 as shown below. Note that the sub-TLV has been assigned from Type 1 (Target FEC Stack) of the "TLVs" registry.

Sub-Type	Sub-TLV Name
41	SR MPLS P2MP Policy Tree Instance

Table 2

5. Security Considerations

Overall, the security needs for SR P2MP Policy ping are the same as in [RFC9960], [RFC6425], and [RFC8029]. SR P2MP Policy ping is susceptible to the same three attack vectors as explained in Section 5 of [RFC8029]. The same procedures and recommendations explained in Section 5 of [RFC8029] should be taken and implemented to mitigate these attack vectors for SR P2MP Policy ping as well.

In addition, the security considerations in Section 8 of [RFC6425] should be followed, specifically the security recommendations from [RFC4379], which recommend the following:

> To avoid potential Denial-of-Service attacks, it is **RECOMMENDED** that implementations regulate the LSP ping traffic going to the control plane. A rate limiter **SHOULD** be applied to the well-known UDP port [allocated for this service].

6. References

6.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3270] Le Faucheur, F., Ed., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, DOI 10.17487/RFC3270, May 2002, <https://www.rfc-editor.org/info/rfc3270>.

[RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, DOI 10.17487/RFC4379, February 2006, <https://www.rfc-editor.org/info/rfc4379>.

[RFC6425] Saxena, S., Ed., Swallow, G., Ali, Z., Farrel, A., Yasukawa, S., and T. Nadeau, "Detecting Data-Plane Failures in Point-to-Multipoint MPLS - Extensions to LSP Ping", RFC 6425, DOI 10.17487/RFC6425, November 2011, <https://www.rfc-editor.org/info/rfc6425>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC9524] Voyer, D., Ed., Filsfils, C., Parekh, R., Bidgoli, H., and Z. Zhang, "Segment Routing Replication for Multipoint Service Delivery", RFC 9524, DOI 10.17487/RFC9524, February 2024, <https://www.rfc-editor.org/info/rfc9524>.

[RFC9960] Parekh, R., Ed., Voyer, D., Ed., Filsfils, C., Bidgoli, H., and Z. Zhang, "Segment Routing Point-to-Multipoint Policy", RFC 9960, DOI 10.17487/RFC9960, April 2026, <https://www.rfc-editor.org/info/rfc9960>.

6.2. Informative References

[IANA-AF] IANA, "Address Family Numbers", <http://www.iana.org/assignments/address-family-numbers>.

Authors' Addresses

Hooman Bidgoli (EDITOR)
Nokia
Ottawa
Canada
Email: hooman.bidgoli@nokia.com

Zafar Ali
Cisco System
San Jose,
United States of America
Email: zali@cisco.com

Zhaohui Zhang
Juniper Networks
Boston,
United States of America
Email: zzhang@juniper.net

Anuj Budhiraja
Cisco System
San Jose,
United States of America
Email: abudhira@cisco.com

Daniel Voyer
Cisco System
Montreal
Canada
Email: davoyer@cisco.com