

Workgroup: Network Working Group
Internet-Draft:
draft-hb-pim-p2mp-policy-ping-03
Published: 13 November 2021
Intended Status: Standards Track
Expires: 17 May 2022
Authors: H. Bidgoli, Ed. V. Voyer P. Parekh
 Nokia Bell Canada Cisco System
 Z. Zhang
 Juniper Networks

P2MP Policy Ping

Abstract

SR P2MP policies are set of policies that enable architecture for P2MP service delivery. A P2MP Policy consists of candidate paths that connects the Root of the Tree to a set of Leaves. The P2MP policy is composed of replication segments. A replication segment is a forwarding instruction for a candidate path which is downloaded to the Root, transit nodes and the leaves.

This document describes a simple and efficient mechanism that can be used to detect data plane failures in P2MP Policy Candidate Paths (CPs) and Path Instances (PIs).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and 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 <https://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 17 May 2022.

Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.

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

- [1. Introduction](#)
- [2. Conventions used in this document](#)
- [3. Motivation](#)
 - [3.1. MPLS P2MP Policy Ping and Traceroute](#)
 - [3.2. Packet format and new TLVs](#)
 - [3.2.1. Identifying a P2MP Policy](#)
 - [3.2.1.1. P2MP Policy CP FEC Stack Sub-TLVs](#)
 - [3.3. Limiting the Scope of Response](#)
- [4. IANA Consideration](#)
- [5. Security Considerations](#)
- [6. Acknowledgments](#)
- [7. References](#)
 - [7.1. Normative References](#)
 - [7.2. Informative References](#)
- [Authors' Addresses](#)

1. Introduction

Each P2MP Policy can have multiple CPs. The CP with highest preference is the active CP while all other CPs are the backup CPs. A CP can have multiple PIs to allow global optimization of the CP via Make Before Break procedures between the active PI and the newly setup and optimized PI.

It is desirable to test the end to end connectivity of a CP, whether it is the active CP or a backup CP and all the CP's PIs. This document describes a mechanism that can be used to detect data plane failures in P2MP Policy Candidate Paths (CP) and its associate Path Instances (PI).

This draft is only for replication SIDs that use MPLS encap for their forwarding and not SRv6. It reuses most of the concepts in [[RFC6425](#)]

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

A P2MP Policy and its corresponding Replication Segments are usually setup via a controller, the root and the leaves are discovered as per [[draft-ietf-pim-sr-p2mp-policy-02](#)]. The tree is calculated from the root to the leaves. There is no underlay protocol to signal the P2MP Policy from the root to the Leaf routers, as such when a P2MP tree fails to deliver user traffic, the failure can be difficult to pin point without a ping/traceroute mechanism to isolate the fault in the P2MP tree. The P2MP Policy ping/traceroute can be utilized to detect faults on the path of the tree and its associated replication segments [[draft-ietf-spring-segment-routing-policy-13](#)]. These tools can be used to periodically ping the leaves to ensure connectivity. If the ping fails, trace route can be initiated to determine where the failure lies. The ping/traceroute can be initiated from the node that initiates the P2MP policy.

3.1. MPLS P2MP Policy Ping and Traceroute

Ping/Traceroute packets are forwarded on the P2MP Policy, on a specific CP or its active PI toward the leaf routers. They are replicated at the replication point based on the replication segment forwarding information on the corresponding node. This draft only addresses the replication segments that use MPLS encap only, future drafts will address the SRv6 forwarding. The packets are processed accordingly when their TTL expires or when the egress router (leaf) is reached. The appropriate response is sent back to the root as per procedures in [[RFC6425](#)]

This draft reuses most procedures for mLDP in RFC [[RFC6425](#)]

This draft also reuses the same destination UDP port as [[RFC4379](#)]

The implementation should take into account that there can be many CPs under the P2MP Policy and the implementation should allow each CP and its corresponding PI to be tested via Ping and Trace route. The Ping and Traceroute packet is forwarded via that specific CP and its corresponding replication segments. On the egress node the corresponding CP and its PI should respond irrespective if it is the active CP or a backup CP.

Two replication segments can be connected via a unicast SR domain. In this scenario the SR tunnel labels need to be programmed with the right TTL depending on the which type of hierarchical MPLS TTL mode is used. As an example pipe vs uniform mode. When in SR domain the P2MP Tree PING and Traceroute will be processed on the two

connecting replication segments based on the replication SID and its TTL. As such the SR domain will act as a single hop on the replication SID and the replication SID TTL is subtracted by one before the unicast SR SIDs are pushed on the replication SID. To detect failure in SR domain is beyond the scope of this draft.

3.2. Packet format and new TLVs

The packet format used is as per [\[RFC4379\]](#) section 3. Some new TLVs and sub-TLVs are required to support the new functionality. They are described in the following sections.

3.2.1. Identifying a P2MP Policy

[\[RFC4379\]](#) defines how an MPLS TE LSP under test may be identified in an echo request. In order to identify the correct replication segment for the CP and its PI, the echo request message MUST carry a Target FEC Stack TLV, and this MUST carry exactly one of one new sub-TLV: a P2MP policy MPLS CP FEC Stack sub-TLV. The new sub-TLV is assigned Sub-Type identifiers as follows, and are described in the following sections.

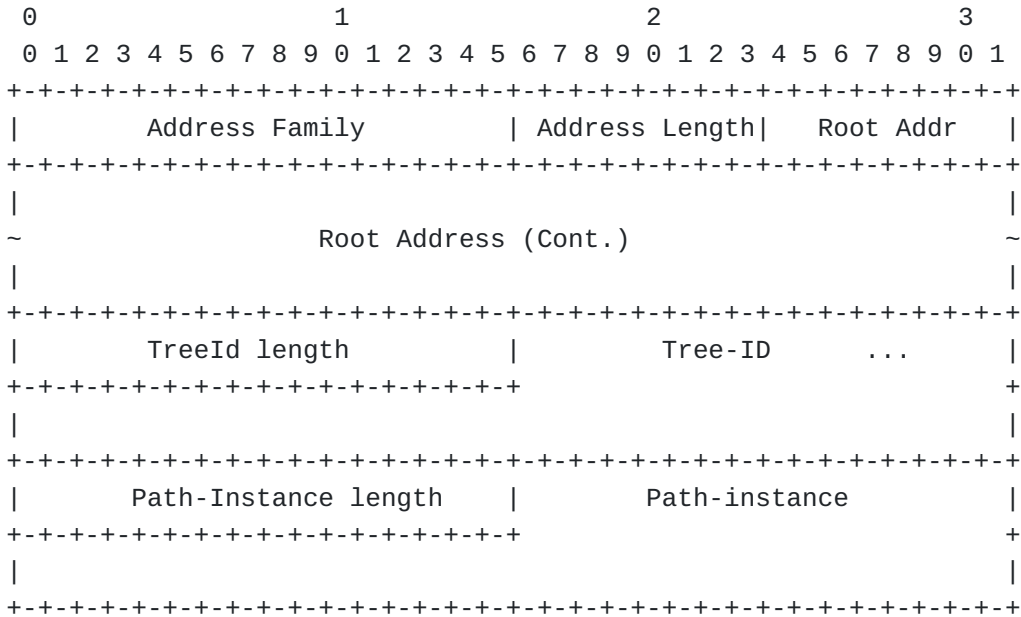
artwork

Sub-Type	Length	Value Field
-----	-----	-----
42	xx	P2MP Policy MPLS CP

3.2.1.1. P2MP Policy CP FEC Stack Sub-TLVs

The format of the P2MP Policy Session sub-TLV value field is specified in the following figure. The value fields are taken from the definitions of the P2MP Policy section 3 of [\[draft-ietf-pim-sr-p2mp-policy-02\]](#)

artwork



*Address Family: Two-octet quantity, containing a value from ADDRESS FAMILY NUMBERS in [IANA-AF] that encodes the address family for the Root Address.

*Address Length: Length of the Root Address in octets.

*Root Address: The address of Root node of P2MP tree instantiated by the SR P2MP Policy

*TreeId Length: Length of the TreeID in octets. This should be set to 4 octets

*Tree-ID: A identifier that is unique in context of the Root. This is an unsigned 32-bit number.

*Path-instance Length: Length of the path instance ID in octet.

*path-instance: path instance ID to be tested [[draft-ietf-spring-segment-routing-policy-13](#)]

3.3. Limiting the Scope of Response

As per [[RFC6425](#)] section 3.2 Four sub-TLVs are used for the inclusion in the P2MP Responder Identifier TLV carried on the echo request message.

The Sub-TLVs for IPv4 and IPv6 egress address P2MP responder is in par with [[RFC6425](#)] section 3.2.1

The Sub-TLVs for IPv4 and IPv6 node address P2MP responder is in par with [[RFC6425](#)] section 3.2.2

4. IANA Consideration

Two new sub-TLV types are defined for inclusion within the LSP ping [RFC4379] Target FEC Stack TLV (TLV type 1).

5. Security Considerations

TBD

6. Acknowledgments

7. References

7.1. Normative References

[RFC2119] "S. Brandner, "Key words for use in RFCs to Indicate Requirement Levels"", March 1997.

[RFC8174] "B. Leiba, "ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words"", May 2017.

7.2. Informative References

[draft-ietf-pim-sr-p2mp-policy-02] "D. Yoyer, C. Filsfils, R.Prekh, H.bidgoli, Z. Zhang, "draft-voyer-pim-sr-p2mp-policy"", October 2019.

[draft-ietf-spring-segment-routing-policy-13] "D. Yoyer, C. Filsfils, R.Prekh, H.bidgoli, Z. Zhang, "draft-voyer-pim-sr-p2mp-policy "draft-voyer-spring-sr-replication-segment"", July 2020.

[RFC4379] "K. Kompella, G. Swallow", February 2006.

[RFC6425] "S. Saxena, G. Swallow, Z. Ali, A. Farrel, S. Yasukawa, T.Nadeau "Detecting Data-Plane Failures in Point-to-Multipoint MPLS"", November 2011.

Authors' Addresses

Hooman Bidgoli (editor)
Nokia
Ottawa
Canada

Email: hooman.bidgoli@nokia.com

Daniel Voyer
Bell Canada
Montreal

Canada

Email: daniel.yover@bell.ca

Rishabh Parekh
Cisco System
San Jose,
United States of America

Email: riparekh@cisco.com

Zhaohui Zhang
Juniper Networks
Boston,
United States of America

Email: zzhang@juniper.com